Persistence of Innovation: Stylised Facts and Panel Data Evidence — Source link

Discussion Paper No. 05-81

Persistence of Innovation: Stylised Facts and Panel Data Evidence

Bettina Peters

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 05-81

Persistence of Innovation: Stylised Facts and Panel Data Evidence

Bettina Peters

Non–technical Summary

Innovation is widely considered to be a key long–term driving force for economic growth. In 1993, the Community Innovation Surveys (CIS) were initiated in the European countries to investigate firms' innovation activities. However, there is still very little empirical evidence on the dynamics in firms' innovation behaviour. Looking at innovation indicators at the aggregate or industry level, we can identify a high and quite stable share of innovators in Germany over the last ten years. One interesting question, however, cannot be answered by such macroeconomic numbers: Is it the same group of firms that always set themselves at the cutting edge by introducing new products and processes or is there a steady entry into and exit from innovation activities at the firm level, with the aggregate level remaining more or less stable over time?

This paper analyses the dynamics in firms' innovation behaviour. In particular, it focuses on the following two questions: First of all, is innovation persistent at the firm–level? And secondly, if persistence is prevalent, what drives this phenomenon? In principle, there are various potential sources for persistent behaviour: First, it might be caused by true state dependence. This means that a causal effect exists in the sense that the decision to innovate in one period itself enhances the probability of innovating in the subsequent period. The theoretical literature delivers several potential explanations for state dependent behaviour: success breeds success, dynamic increasing returns (learning effects) or sunk costs in R&D investments. On the other hand, firms may possess certain characteristics which make them more likely to innovate. To the extent that these characteristics themselves show persistence over time, they will induce persistence in innovation behaviour. Such firm–specific attributes can be classified into observable characteristics, like firm size, competitive environment, skills or financial resources, and unobservable ones. For instance, technological opportunities, managerial abilities or risk attitudes are important for the firms' decision to innovate, but are typically not observed (unobserved heterogeneity).

The answers to both questions are important from a theoretical as well as a policy point of view. From a theoretical point of view they are interesting because endogenous growth models differ in their underlying assumptions about the innovation frequency of firms (creative destruction versus creative accumulation). From a policy point of view the distinction between permanent innovation activities due to firm–inherent factors and true state dependence has some important implications. If innovation is state dependent, innovation–stimulating policy measures such as government support programmes are supposed to have a more profound effect because

they do not only affect the current innovation activities but are also likely to induce a permanent change in favour of innovation. If, on the other hand, individual heterogeneity induces persistent behaviour, support programmes are unlikely to have long–lasting effects and economic policy should concentrate more on measures which have the potential to improve innovation–relevant firm–specific factors.

The paper presents some stylised facts regarding the permanence of German manufacturing and service firms' innovation behaviour in the period 1994–2002. In a second step, the sources for persistent behaviour are analysed and identified by means of a dynamic random effects binary choice model using the estimator recently proposed by Wooldridge (2005). This panel data approach allows to control for individual heterogeneity, a potential source of bias which was not taken into account in most of the previous empirical studies due to data restrictions.

A first main finding is that innovation behaviour is permanent at the firm level to a very large extent. Year–to–year transition rates indicate that in manufacturing nearly nine out of ten innovating firms in one period persisted in innovation activities in the subsequent period and about 84 per cent of non–innovators remained inactive in the following period. Yet innovation is not a once and for all phenomenon. 45 per cent of manufacturing and 55 per cent of service firms experienced at least one change in their innovation behaviour. In general, persistence is less pronounced in the service sector and exhibits a higher variance across time.

The econometric results confirm the hypothesis of true state dependence. Depending on the calculation method, about one third to one half of the difference in the propensity to innovate between previous innovators and non–innovators in manufacturing can be traced back to true state dependence. In the service sector, persistence is generally less prevalent and state dependence effects are less pronounced, yet still highly significant.

The results further confirm and highlight the role of unobserved heterogeneity as well as innovative capabilities on the dynamics in firms' innovation behaviour. That is, in addition to past innovation experience, knowledge provided by skilled employees was found to be important in generating innovations over time.

Persistence of Innovation: Stylised Facts and Panel Data Evidence

Bettina Peters*

November 2005

Abstract: This paper investigates whether firms innovate persistently or discontinuously over time using an innovation panel data set on German manufacturing and service firms for the period 1994–2002. We find that innovation behaviour is permanent at the firm–level to a very large extent. Using a dynamic random effects discrete choice model and a new estimator recently proposed by Wooldrigde (2005), we further shed some light on the driving forces for this phenomenon. The econometric results confirm the hypothesis of true state dependence for manufacturing as well as for service sector firms. In addition to past innovation experience, the results further highlight the important role of knowledge provided by skilled employees and unobserved individual heterogeneity in explaining the persistence of innovation.

Keywords: innovation, persistence, state dependence, unobserved heterogeneity, dynamic random effects panel probit model

JEL Classification: O31, C23, C25, L20

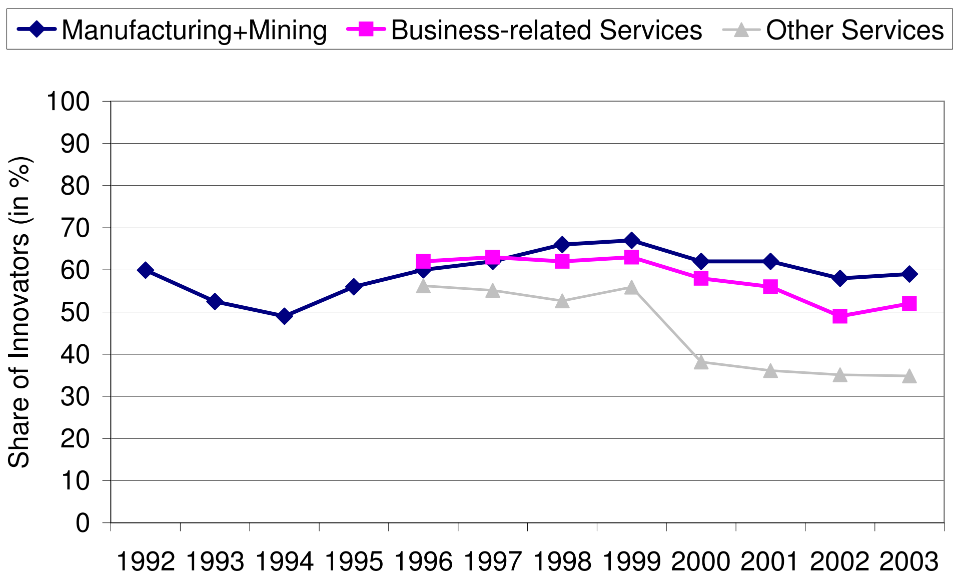
* Centre for European Economic Research (ZEW), Department of Industrial Economics and International Management, P.O. Box 10 34 43, D-68034 Mannheim, Germany, phone: +49-621-1235-174, fax: +49-621-1235-170, b.peters@zew.de.

The author gratefully acknowledge the comments of Martin Biewen, Francois Laisney, Uli Kaiser, Hans Christian Kongsted, Jacques Mairesse, Pierre Mohnen, Christian Rammer and Jeffrey M. Wooldridge. I would also like to thank all participants in the Annual Conference of the Danish Econometric Society (Sandbjerg, May 15, 2004), the 2^{nd} ZEW conference on Innovation and Patenting (Mannheim, September 19–20, 2005), the 7^{th} Conference on Innovations and Intellectual Property Values of the Applied Econometrics Association (AEA, Paris, October 20–21, 2005) and in the ZEW Brown Bag Seminar (Mannheim, March 22, 2005) for valuable comments on previous versions. Any errors remain those of the author.

1 Introduction

Innovation is widely considered to be a key long–term driving force for economic growth. In 1993, the Community Innovation Surveys (CIS) were initiated in the European countries to investigate firms' innovation activities. These rich and internationally harmonised data sets have served as the starting–point for many empirical studies which have analysed various aspects of innovation activities. However, there is still very little evidence on the dynamics in firms' innovation behaviour. Looking for example at innovation performance indicators at the aggregate or industry level, we can identify a high and quite stable share of innovators in the manufacturing as well as in the service sector in Germany over the last ten years (see Figure 1). One interesting question, however, cannot be answered by such macroeconomic numbers: Is it the same group of firms that always set themselves at the cutting edge by introducing new products and processes or is there a steady entry into and exit from innovation activities at the firm level, with the aggregate level remaining more or less stable over time?

Figure 1: Share of Innovators 1992–2003



Notes:
Business–related services include telecommunication, financial intermediation, data processing, technical services, consultancies and other business-related services. Wholesale, retail, transport/storage, post, real estate and renting are summarised as other services. All figures are expanded to the target population of German firms with 5 or more employees.
Comparability of figures for other services before and after 2000 is reduced due to slight changes in the definition of innovation.
Source: Rammer et al. (2005).

This paper analyses the dynamics in firms' innovation behaviour. In particular, it focuses upon the following two research questions: First of all, is innovation persistent at the firm–level? Persistence occurs when a firm which has innovated in one period innovates once again in the subsequent period. And secondly, if persistence is prevalent, what drives this phenomenon?

In principle, there are various potential sources for persistent behaviour (see Heckman 1981a,b): First, it might be caused by true state dependence. This means that a causal behavioural effect exists, in the sense that the decision to innovate in one period in itself enhances the probability to innovate in the subsequent period. The theoretical literature delivers several potential explanations for state dependent behaviour. The most prominent ones relate to (i) the hypothesis of success breeds success (Mansfield 1968), (ii) the hypothesis that innovations involve dynamic increasing returns (see, e.g., Nelson and Winter 1982 and Malerba and Orsenigo 1993), and (iii) sunk costs in R&D investments (Sutton 1991). Secondly, firms may possess certain characteristics which make them particularly "innovation-prone", i.e., more likely to innovate. To the extent that these characteristics themselves show persistence over time, they will induce persistence in innovation behaviour. Such firm–specific attributes can be classified into observable characteristics[1], like firm size, competitive environment or financial resources, and unobservable ones. For instance, technological opportunities, managerial abilities or risk attitudes are important for the firms' decision to innovate, but are typically not observed. If these unobserved determinants are correlated over time, but are not appropriately controlled for in estimation, past innovation may appear to affect future innovation merely because it picks up the effect of the persistent unobservable characteristics. In contrast to true state dependence this phenomenon is therefore called spurious state dependence. And thirdly, serial correlation in exogenous shocks to the innovation decision can cause permanent behaviour over time.

The answers to both research questions are important for several reasons. First, they are interesting from a theoretical point of view. Endogenous growth models for example differ in their underlying assumptions about the innovation frequency of firms. While Romer (1990) assumes that innovation behaviour is persistent at the firm level to a very large extent, the process of creative destruction leads to a perpetual renewal of innovators in the model of Aghion and Howitt (1992). Thus, empirical knowledge about the dynamics in firms' innovation behaviour is a tool to assess different endogenous growth models (Cefis 2003). Furthermore, it might help to improve current theories of industrial dynamics, where some forms of dynamic

[1] Observable characteristics means known to the econometrician.

increasing returns play a major role in determining degrees of concentration, the evolution of market shares and their stability over time (Geroski 1995). Second, from a managerial point of view permanent innovation activities are seen as a crucial factor for strengthening competitiveness. And last but not least, the distinction between permanent innovation activities due to firm–inherent factors as opposed to true state dependence has important implications for technology and innovation policy. If innovation performance shows true state dependence, innovation–stimulating policy measures such as government support programmes are supposed to have a more profound effect because they do not only affect the current innovation activities but are likely to induce a permanent change in favour of innovation. If, on the contrary, individual heterogeneity induces persistent behaviour, support programmes are unlikely to have long–lasting effects and economic policy should concentrate more on measures which have the potential to improve innovation–relevant firm–specific factors and circumstances.

To answer the first question, the paper presents some stylised facts of how permanently German manufacturing and service firms innovated in the period 1994–2002. While in most of the other European countries innovation surveys take place every 4 years, the German innovation survey is conducted annually. This provides us with rather long panel data which are appropriate to study whether the innovation behaviour is persistent at the firm-level. In a broader sense, this part ties in with the literature about the existence of innovation persistence effects using patents (see Geroski et al. 1997, Malerba and Orsenigo 1999 and Cefis 2003) and R&D indicators (see Manez Castillejo et al. 2004).

In a second step, the sources for persistent behaviour are analysed and identified by means of a dynamic random effects binary choice model. This panel data approach allows us to control for individual heterogeneity, a potential source of bias which was not taken into account in most of the previous empirical studies due to data restrictions.

The paper contributes to the existing literature in that it is one of the first which investigates firm–level persistence using innovation data (see section 3) and that it is able to exploit data from a unique long panel, which are nonetheless internationally comparable. Furthermore, a new estimation method recently proposed by Wooldridge (2005) is applied, and the paper is the first to provide empirical evidence on innovation persistence in service firms. Investigating the dynamics in the innovation behaviour of service firms is interesting not only because the service sector has experienced a rapid development over the last two decades, but also from a theoretical point of view. Looking at the potential theoretical explanations for true state dependence listed above, the third one in particular is strongly related to R&D, which is less important and less common in the service sector. Thus, one hypothesis that will be investigated is that innovation activities are less permanent in this sector compared to manufacturing.

The outline of the paper is as follows. Section 2 sketches some theoretical arguments in favour of and against state dependence in innovation behaviour at the firm level. Section 3 summarises the main empirical firm–level results so far. The panel data set underlying this study is explored in section 4 and section 5 briefly comments on some measurement issues. The following section 6 depicts some stylised facts about the entry into and exit from innovation activities at the firm level during the period 1994–2002. Section 7 presents the econometric model and its empirical implementation. It further explores the estimation methods used and sets forth the econometric results. Section 8 draws some conclusions on the persistence of firm–level innovation activities and discusses the main findings.

2 Theoretical Explanations

Economic theory provides at least three potential explanations of why innovation behaviour might demonstrate state dependence over time.

The first one is the well–known hypothesis of "success breeds success". However, this view is based on different arguments in the literature. Phillips (1971), for instance, argued that successful innovations positively affect the conditions for subsequent innovations via an increasing permanent market power of prosperous innovators.[2] Mansfield (1968) and Stoneman (1983), however, emphasised that a firm's innovation success broadens its technological opportunities which make subsequent innovation success more likely. Based on this idea of dynamic intra–firm spill–overs, Flaig and Stadler (1994) developed a stochastic optimisation model in which firms maximise their expected present value of profits over an infinite time horizon by simultaneously choosing optimal sequences of both product and process innovations. Both were shown to be dynamically interrelated in this model. Another line of reasoning is the existence of financial constraints. Usually, information asymmetries about the risk and the failure probability of an innovation project exist between the innovator and external financial investors. This leads to adverse

[2] In contrast to Schumpeter, who assumed that the increasing market power is a temporary phenomenon and is eroded by the entry of imitators or innovators, Phillips argued that success favours growing barriers to entry that eventually allow a few increasingly successful firms to permanently dominate an industry.

selection and moral hazard problems which usually force firms to finance innovation projects by means of internal funds (see, e.g., Stiglitz and Weiss 1981). Successful innovations provide firms with increased internal funding and hence can be used to finance further innovations (Nelson and Winter 1982). Common to all these various "success breeds success" theories is the notion that a firm can gain some kind of locked–in advantage over other firms due to successful innovations (Simons 1995).

The second hypothesis is based on the idea that knowledge accumulates over time as represented by the changes in an organisations repertories of operating and dynamic routines (Nelson and Winter 1982). Evolutionary theory states that technological capabilities are a decisive factor in explaining innovation. Firms' innovative capabilities in turn are primarily determined by human capital, i.e., by the knowledge, skills and creativity of their employees. Experience in innovation is associated with dynamic increasing returns in the form of learning–by–doing and learning–to–learn effects which enhance knowledge stocks and, therefore, the probability of future innovations. Since a firm's absorptive capacity, i.e. its ability to recognise the value of new, external information, to assimilate and apply it to commercial ends, is likewise a function of the level of knowledge, learning in one period will furthermore permit a more efficient accumulation of external knowledge in subsequent periods (Cohen and Levinthal 1990). The cumulative nature of knowledge should therefore induce state dependence in innovation behaviour (see, e.g., Nelson and Winter 1982 and Malerba and Orsenigo 1993).[3]

The hypothesis of sunk costs in R&D investments is a third argument in favour of state dependence (see Sutton 1991 or Manez Castillejo et al. 2004). It is stressed that R&D decisions are subject to a long time horizon, and if a firm decides to take up R&D activities, it has to incur start–up costs in building up an R&D department or hiring and training R&D staff. These fixed outlays, once made, are usually not recoverable and can therefore be considered as sunk costs.[4] With respect to persistence, sunk costs represent a barrier to both entry into and exit from R&D activities. Sunk costs may prevent non–R&D performers from taking up such activities because, unlike established R&D performers, potential entrants have to take these costs into account in determining their prices. Conversely, sunk costs may represent a barrier to exit for established R&D performers because they are

[3] Theories which focus on how firms accumulate technological capabilities may also be considered as "success breeds success" theories since technological capabilities might substantiate sustained competitive advantages (Teece and Pisano 1994). However, learning can also occur as a result of unsuccessful innovations.

[4] In contrast to most other kinds of sunk costs, firms can decide strategically upon the amount of R&D expenditure. Costs incurred in this manner are therefore referred to as endogenous sunk costs (Sutton 1991).

not recovered in the case that the firm stops R&D and the firm has to incur them again if it decides to re–enter in future periods.

However, even if firms experience sunk costs or knowledge accumulation due to innovations, there are several theoretical explanations of why they may exit from innovation activities in future periods with the consequence that persistence does not emerge. The first two arguments are related to the demand–pull theory which emphasises that innovations are stimulated by demand (see Schmookler 1966). If there is, at least in the firm's perception of consumer demand, no need for further innovations due to its own previous introduction of new products or processes, the firm will at least temporarily cease to innovate. This is particularly true if a firm only offers one or a few products and typical product life cycles are several years. Closely related is the second argument that states that unfavourable market conditions in general (i.e. expected decrease in demand) might prevent firms from carrying on with innovation, especially with respect to the timing of the market introduction of new products. This is one argument in the literature on innovation and business cycles and will be explored in more detail in section 6. Finally, an incumbent innovator might fear that the introduction of further new products or processes will cannibalise his rents from previous innovations and thus stop innovating (Schumpeter 1942). Patent race models, for instance, predict that an incumbent invests less in R&D than challengers because it would erode current monopoly and profits (see, e.g., Reinganum 1983).

3 What do we know so far? Previous Empirical Findings

Though economic theory emphasises that innovation is an inherently dynamic process between heterogenous firms (see Blundell et al. 1995), firm–level empirical evidence on persistence in innovation activities is scarce. We can broadly classify the existing literature into three categories according to how the authors measure innovation: patent–based, R&D–based and innovation–based studies.

Patent–based studies have mainly focused on the question whether innovation persistence exists, irrespective of its origin. Malerba and Orsenigo (1999) examined this question using data of manufacturing firms from six countries (France, Germany, Italy, Japan, USA and the UK) which had requested at least one patent at the European Patent Office (EPO) between 1978 and 1991. Their results corroborated substantial entry into and exit from patent activities implying that the population of innovators changed remarkably over time. In terms of employment, entrants and exiters showed nearly the same size as incumbent innovators, but, in terms of the number of patents, both were much smaller. The high entry and exit rates were associated with a large proportion of enterprises that innovated only once and than ceased to innovate further. Only a small fraction of entrants were able to persist in patent activities as time went on. However, these firms became rather large innovators (in terms of patents) over time, resulting in the fact that persistent innovators, although small in absolute numbers, accounted for an important part of all patents. The same result, that patent activities among patenting firms exhibited only a little degree of persistence, was confirmed by Geroski et al. (1997) who concentrated on patents as well, but used data of UK manufacturing firms which had at least one patent granted in the US between 1969 and 1988.

Cefis (1999, 2003) used a UK sub–sample of the data set of Malerba and Orsenigo (1999) and applied a non–parametric approach based on transition probabilities matrices. Cefis and Orsenigo (2001) extended this kind of analysis to a firm–level cross–country comparison over time for the original six countries. In their studies they distinguished four states in each year: occasional (zero patents)[5], small (one patent), medium (two to five patents) and great innovators (at least 6 patents). They corroborated previous evidence that in general only a low degree of state persistence in patenting was prevalent in all countries which furthermore declined as time went by. Only occasional and great innovators had a high probability of remaining in their state while persistence was much lower in the intermediate classes for which a strong tendency towards the non-innovator state was ascertained. Moreover, persistence was found to differ across industries, but inter–sectoral differences were by and large consistent across countries suggesting that persistence is at least partly technology–specific. However, cross–country differences showed up in the relationship between persistence and firm size. While a strictly positive impact was found in Italy, France, USA and the UK, this was not observed in Japan and Germany.

In contrast to the other studies, Geroski et al. (1997) also examined potential sources of persistence. To test the hypothesis of dynamic economies of scale, they focused on patent spells, which measured the number of successive years in which a firm produced a patent. In this setting, dynamic economies of scale would imply that the probability of the spell ending at any particular time $t + \Delta t$, given it has lasted until t, decreases with the initial level of patents and with the length of time a firm has already spent in that spell. While the first relationship was confirmed by their data, the second one was rejected. All in all, their results suggested that dynamic economies might have led to more persistent patent spells, but only when

[5] Firms with zero patents in a given period are nevertheless referred to as occasional innovators since they had at least one patent in the whole period under consideration.

the threshold of initial patent activities was high enough to overcome the reversed within–spell effects. And only a few firms ever reached this threshold.

One explanation of why patent–based studies revealed only a small degree of persistence might be the well–known fact that patents measure only some aspects of innovative activity, see Griliches (1990). However, in the context of persistence analysis, patents have an additional drawback, because in this kind of winner–takes–all contest, to be classified as permanent innovators firms have to win the patent race continuously, see Kamien and Schwartz (1975). This means that patent data measure the persistence of innovative leadership rather than the persistence of innovation, as was stressed by Duguet and Monjon (2004).

Instead of patents, another strand of literature uses R&D activities. Mairesse et al. (1999) and Mulkay et al. (2001) for instance estimated dynamic equations for physical as well as R&D investment rates. Based on samples of large French and US manufacturing firms they found evidence that R&D investment rates are highly correlated over time, even more highly correlated than physical capital investments. This reflects the inter–temporal nature of R&D and the fact that about half of the R&D expenditure consists of labour costs for R&D staff. Using a sample of small and large Spanish manufacturing firms between 1990 and 2000 and a dynamic discrete choice model, Manez Castillejo et al. (2004) asserted that past R&D experience had significantly affected the current decision to engage in R&D and interpreted this as an indication for sunk costs in building up R&D. Their results further indicated a rapid depreciation of R&D experience in that there was no significant difference between the re–entry costs of a firm that last performed R&D activities two or three years ago and a firm that had never previously conducted R&D.

Though R&D is an important input to innovation, it does not capture all aspects pertinent to innovation. Innovation activities close to the market, for instance, are not captured by the concept of R&D. Such activities of small and medium–sized manufacturing as well as service sector firms are heavily underestimated by patents as well as R&D indicators.

Hence, another strand of literature uses the broader concept implied by innovation data. So far, only a few studies have attempted to estimate the dynamics in the innovation process at the firm–level and empirical results are inconclusive. König et al. (1994) and Flaig and Stadler (1994, 1998) were the first to examine dynamic effects using innovation data from a panel of manufacturing firms in West Germany in the eighties. Applying a dynamic panel probit model, empirical evidence of state dependence in process innovation activities was supported by the first study. This result was corroborated for process as well as product innovations by the second authors. Duguet and Monjon (2004) for French firms and Rogers (2004) for Australian firms also reported persistence effects. However, due to data limitations both studies did not carry out a dynamic panel data analysis and thus did not control for unobserved individual heterogeneity, which leads to biased estimates if heterogeneity is present.[6]

Conversely, Geroski et al. (1997) and Raymond et al. (2005) could not ascertain persistence effects in the occurrence of innovations for UK and Dutch manufacturing firms. But Raymond et al. (2005) pointed out that among continuous innovators the innovation success, measured in terms of sales due to new products, had a positive impact on future success.

Among the other things highlighted, this review makes clear that previous studies focused solely on manufacturing. One aim is therefore to extend this kind of analysis to a comparison between the manufacturing and service sector.

4 Data Set

The research makes use of firm level data from the Mannheim Innovation Panel (MIP) in the German manufacturing (NACE 10–45) and service sector (NACE 50–52, 60–74, 90).[7] The MIP is based on innovation surveys carried out by the Centre for European Economic Research (ZEW) on behalf of the German Federal Ministry of Education and Research. The target population covers all legally independent firms with 5 or more employees and the surveys are drawn as stratified random samples (stratified by firm size, branches of industries and East/West region). The survey methodology and definitions of innovation indicators comply with the recommendations manifested in the OSLO-Manual, see OECD and EUROSTAT (1997), thereby yielding internationally comparable data on innovation activities of German firms. In 1993 (CIS1), 1997 (CIS2) and 2001 (CIS3) the surveys were the German contribution to the European–wide harmonised CIS.

While in most other European countries innovation surveys take place every 4 years, they are conducted annually in Germany. In manufacturing, we refer in our analysis to the surveys 1995 to 2003, in the service sector the first usable wave was that of 1997.[8] Thus, 9 waves in manufacturing and 7 in services are available. The

[6] Both studies applied a cross–sectional probit approach, including a dummy variable for whether the firm was an innovator in the previous period as an explanatory variable.

[7] For a detailed definition, see Table 14 in the appendix.

[8] In manufacturing, the survey started in 1993. However, due to a major refreshment and

data of each survey refers to the previous year, hence we focus on the period 1994–2002 in manufacturing and 1996–2002 in the service sector. This relatively long period ensures that we can observe firms' innovation behaviour over different phases of the business cycle, and the observation period is also longer than the average product life cycle in industry.

The samples are constructed as panels and about 10,000 firms in manufacturing and 12,000 service firms are questioned each year. Since participation is voluntary, response rates vary between 20 to 25 %,[9] and although the survey is designed as a panel study, we have to assert that the main part of the firms participated only once or twice.[10] Furthermore, for analysing the dynamics in firms' innovation behaviour with econometric methods, only those firms which have answered consecutively can be taken into account. Therefore, in the following we distinguish two panel data sets: Panel U is an unbalanced panel comprising all firms for which at least 4 successive observations are available and Panel B is the balanced sub–sample. The latter is needed for estimation purposes (see section 7.2).

Table 1: Characteristics of the Unbalanced and Balanced Panel

	Manufacturing	Services
Panel U: Unbalanced Panel		
Number of observations	13558	7901
Number of firms	2256	1528
Minimum number of consecutive obs. per firm	4	4
Average number of consecutive obs. per firm	6.0	5.2
Panel B: Balanced Panel		
Number of observations	3933	1974
Number of firms	437	282
Number of consecutive obs. per firm	9	7
Time Period	1994–2002	1996–2002

Source: Own calculations.

Table 1 summarises the main characteristics of both samples. Given our interest in analysing the persistence of innovation behaviour and the need to estimate a

enlargement of the initial sample in 1995 and the need to construct a balanced panel for estimation purposes, I decided to discard the first two waves. In the service sector, the first survey took place in 1995, with a break in 1996.

[9] The low response rates are in line with those of comparable voluntary surveys of German firms. In order to control for a response bias in the net sample, non-response analyses are carried out each year, covering a similar number of firms of the net sample and collecting information on product and process innovations by the means of telephone interviews. They come up with the result that the share of innovators is only slightly underestimated in the net sample.

[10] Table 15 in the appendix sheds some light on the individual participation behaviour of the sampled firms.

dynamic specification with a lagged endogenous variable, I have chosen to maximise the time dimension of the panel. As a result, in manufacturing as well as in the service sector this choice leads to a marked reduction in the number of observations and the resulting panel data sets might not be representative for the total sample. To check representativeness, Tables 16 and 17 in the appendix compare the distribution of firms by industry, size class, region and innovation status in the total sample of all observations, the unbalanced panel and the balanced sub–sample. It turns out that in manufacturing large firms with 100 or more employees are slightly over–represented in the unbalanced and balanced panel compared to the total sample, while the opposite applies to the service sector. Moreover, the share of East German firms is slightly higher in both panels in manufacturing as well as in the service sector. The tables further demonstrate that the share of innovators is lower in both panels used. But while the difference for instance between the balanced panel and the total sample is rather small in manufacturing, it amounts to 8.5 percentage points in the service sector. That is, the service firms in our sample are less likely to engage in innovation activities. Based on these comparisons, we argue that by and large the panels still reflect total–sample distributional characteristics quite well in manufacturing and don't give any obvious cause for selectivity concerns. Admittedly, in the service sector selectivity might be a more severe problem in the resulting panels since innovators are less represented.

5 Measurement Issues

One problem in studying state dependence in innovation behaviour with CIS data is the fact that the indicator whether a firm has introduced an innovation is related to a 3–year–reference period, that is, using this indicator for yearly waves would induce an artificial high persistence due to overlapping time periods and double counting.[11] Both studies of Duguet and Monjon (2004) or Raymond et al. (2005) suffer from this overlapping of time periods problem in their dependent variable. However, information on innovation expenditure is available on a yearly basis. Innovation expenditure include outlays for intramural and extramural R&D, acquisition of external knowledge, machines and equipment, training, market introduction, design and other preparations for product and/or process innovation activities in a given year.[12] Therefore, and in contrast to the previously mentioned studies, we define an

[11] As an example, in the 2001 survey a firm is defined as an innovator if it has introduced an innovation in the period 1998–2000, in the 2002 survey this indicator is related to 1999–2001.

[12] R&D expenditure accounted for 50–55 % of innovation expenditure in the period under consideration, see Gottschalk et al. (2002).

innovator as a firm which exhibits positive innovation expenditure in a given year, i.e., which decides to engage in innovation activities. This implies that we analyse the persistence in innovation input rather than in innovation outcome behaviour.

From a theoretical point of view it is not unambiguous whether state dependence in innovation behaviour should be tested in terms of an input or an output measure. The literature on sunk costs usually models the decision to invest in R&D by a rational profit–maximising firm, so that an input measure seems advisable. In contrast, the "success breeds success" hypothesis is clearly outcome–oriented. By stressing the accumulative nature of innovation and the importance of learning effects in the innovation process, the evolutionary theory is likewise rather outcome–oriented since the process of learning involves successful implementation rather than just dedicating some resources to innovation projects, see Blundell et al. (1993). Econometric evidence shows that, on average, innovation output is significantly determined by innovation input (see, e.g., Crepon et al. 1998, Lööf and Heshmati 2001, Love and Roper 2001 or Janz et al. 2004), implying that input persistence should to a certain degree be converted into output persistence. However, it is possible that more than one period is needed to translate innovation effort into new products or processes and furthermore firms can not necessarily control their innovation outcome because serendipity might play an important role in the innovation process, see Kamien and Schwartz (1982) or Flaig and Stadler (1998).[13]

6 Stylised Facts

In what follows we want to give an answer to the first research question of "How persistently do firms innovate?". To investigate this question, transition probabilities are an appropriate method. Tables 2 and 3 show corresponding figures for the whole period and differentiated by years. First of all, it turns out that there are hardly any differences between our much larger unbalanced panel and the smaller balanced panel which has to be used for estimation purposes. Table 2 clearly indicates that innovation behaviour is permanent at the firm-level to a very large extent. In the period 1994–2002, nearly 89 % of innovating firms in manufacturing in one period persisted in innovation activities in the subsequent period while 11 % stopped their engagement. Similarly, about 84 % of non–innovators maintained this status in the following period while 16 % entered into innovation activities. That also means that the probability of being innovative in period $t + 1$ was about 72 percentage points

[13] I checked the robustness of my results by applying the output-oriented 3–period innovation indicator and taking only every third survey into account, see section 7.5.

higher for innovators than for non-innovators in t which can be interpreted as a measure of state dependence. Against the background of the sunk costs hypothesis, it is interesting that using the narrower concept of R&D expenditure, Manez Castillejo et al. (2004) found slightly higher exit rates in Spanish manufacturing for the period 1990–2000, while not surprisingly the entry into R&D activities is much less frequent than for innovation activities.[14]

Table 2: Transition Probabilities, Whole Period[a)]

	Innovation status in $t+1$					
	Unbalanced Panel			Balanced Panel		
Innovation status in t	Non–Inno	Inno	Total	Non–Inno	Inno	Total
Manufacturing						
Non–Inno	83.6	16.4	100.0	85.3	14.7	100.0
Inno	11.2	88.8	100.0	11.2	88.8	100.0
Total	41.9	58.1	100.0	44.5	55.5	100.0
Services						
Non–Inno	82.9	17.1	100.0	83.9	16.1	100.0
Inno	29.2	70.8	100.0	30.2	69.8	100.0
Total	62.6	37.4	100.0	64.0	36.0	100.0

Notes:
[a)] Manufacturing: 1994–2002, service sector: 1996–2002.
Source: Own calculations.

In services, persistence effects are also clearly observable, though less prevalent than in manufacturing. Non–innovative service firms had pretty much the same propensity to enter into innovation activities as manufacturing firms. However, in any given year the probability of an innovative service firm remaining in innovation activities in the subsequent year was significantly lower (70 %) than for a manufacturing firm. This implies that the state dependence effect in the service sector was clearly lower with approximately 54 percentage points. Several arguments could explain this finding, one being the fact the sunk cost hypothesis is strongly related to R&D investments. However, R&D is less important and less common in most of the service sectors compared to manufacturing. This result might also occur because, on average, the time needed to develop an innovation is shorter in services and hence covers two calendar years less often. Alternatively, individual or industry heterogeneity, for example in the technological opportunities or in the demand for new innovations, might explain this difference.

[14] Manez Castillejo et al. (2004) reported only transition rates for small and large firms. Using a weighted average, one would get an exit rate of about 17 % and an entry probability of 8 %.

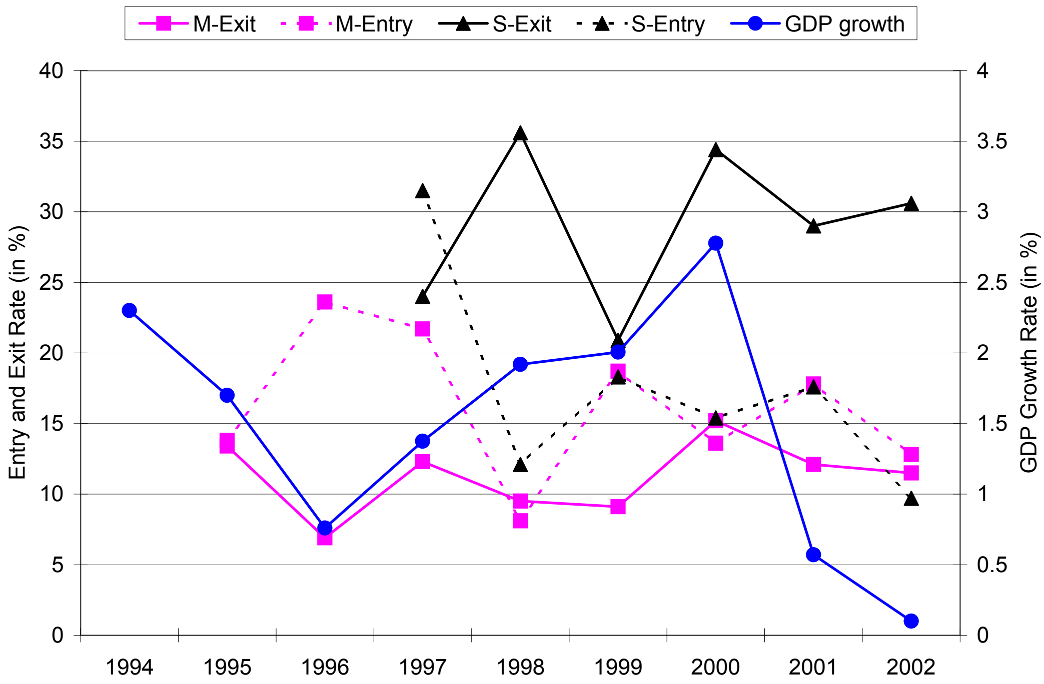
Table 3: Transition Probabilities by Year

Innovation Status		Years							
Year t	Year $t+1$	94–95	95–96	96–97	97–98	98–99	99–00	00–01	01–02
Manufacturing									
Non–Inno	Non–Inno	86.2	76.4	78.3	91.9	81.3	86.4	82.2	87.2
	Inno	13.8	23.6	21.7	8.1	18.7	13.6	17.8	12.8
Inno	Non–Inno	13.4	6.9	12.3	9.5	9.1	15.2	12.1	11.5
	Inno	86.6	93.1	87.7	90.5	90.9	84.8	87.9	88.5
Services									
Non–Inno	Non–Inno	–	–	68.5	87.9	81.7	84.6	82.4	90.3
	Inno	–	–	31.5	12.1	18.3	15.4	17.6	9.7
Inno	Non–Inno	–	–	24.0	35.6	20.9	34.4	29.0	30.6
	Inno	–	–	76.0	64.4	79.1	65.6	71.0	69.7

Notes:
Sample: Unbalanced Panel.
Source: Own calculations.

There is a related strand of literature investigating the interrelationship between business cycles and innovation activity. According to the technology-push argument, science and technology are a major driver for innovation activities and consequently the business cycle, see Schumpeter (1939) or Kleinknecht (1990) for an empirical assessment. In contrast, the demand-pull hypothesis states that innovation behaviour depends on demand conditions and thus on the level of economic activity, see Schmookler (1966). Within this body of literature, arguments for both pro- as well as counter–cyclical relationships can be found. Pro–cyclical effects are expected to occur because cash–flow as an important source of finance innovations is positively correlated with economic activity, see Himmelberg and Petersen (1994). Furthermore, Judd (1985) argued that markets have a limited capacity for absorbing new products and thus firms are more likely to introduce new products in prosperous market conditions. Aghion and Saint-Paul (1998) showed that firms tend to invest more in productivity growth (i.e. process innovations) during recessions, since the opportunity cost in terms of forgone profits of investing capital in technological improvements is lower during recessions. During the period 1994–2002 the German economy underwent different business cycles. 1993 was characterised by a deep recession, followed by an upswing in 1994-1995 which came to a near halt in 1996. Since 1997 economic growth steadily increased again, reaching its peak in 2000. Since 2001 the German economy has again been fighting a significant cyclical slump. Table 3 shows that despite different business cycles, the propensity to remain innovative and correspondingly the exit rates were quite stable over time in manufacturing, with one remarkable exception in the peak period 2000 where the flow out of innovating sharply increased.[15] At the same time, the entry rate was more volatile across the periods in manufacturing. In the service sector, the propensity to remain innovative was not only lower but also exhibited a higher variance across time.[16] However, contrasting both exit and entry rates with the annual GDP growth rate, no clear pro– or counter–cyclical link to the level of economic activity can be found.

Figure 2: Innovation Entry and Exit Rates and Business Cycles



Notes:
The innovation exit rate in any given year t is defined as the share of innovators in year $t-1$ which flow out of innovation activities in year t. Similarly, the innovation entry rate in t is the share of non–innovators in year $t-1$ which start innovation activities in year t. GDP growth denotes the annual percentage change of real GDP (in constant prices of 1995). M and S denote manufacturing and services, respectively. Sample: Unbalanced Panel.
Source: GDP growth rates: Sachverständigenrat (2004). Own calculation.

Table 4 and Figure 3 provide some information on innovation persistence by size class and industry. As expected, innovation behaviour was more stable in larger firms, though also relatively permanent in small firms. This result holds for manu-

[15] This result coincides with the decline in the share of innovators at the aggregate level, see Figure 1. A main cause for this somewhat astonishing development was a severe shortage in high–qualified personnel in 2000, hampering a large number of SMEs in their innovative efforts (Janz et al. 2002).

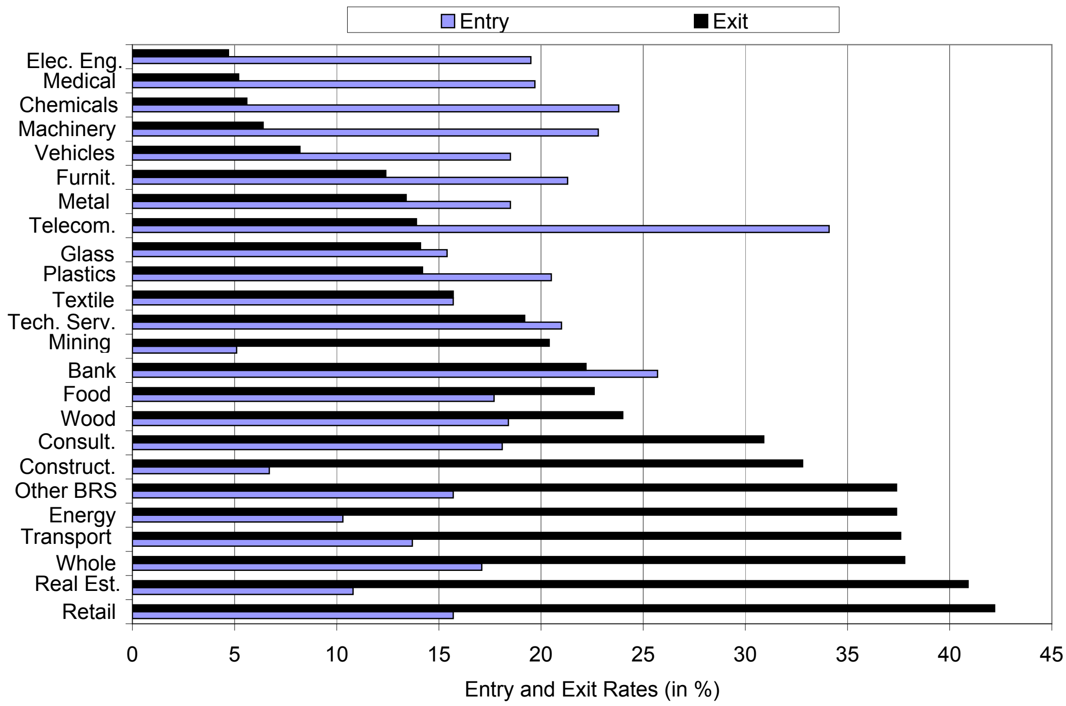
[16] The standard deviation of exit and entry rates is 2.6 and 5.1 in manufacturing and 5.8 and 7.6 in the service sector.

Table 4: Transition Probabilities by Size Class

Innovation Status		Years					
Year t	Year $t+1$	< 10	10–19	20–49	50–99	100–499	>=500
Manufacturing							
Non–Inno	Non–Inno	91.3	87.4	83.9	81.4	78.0	79.0
	Inno	8.7	12.6	16.2	18.6	22.0	21.0
Inno	Non–Inno	32.7	20.3	17.7	12.9	10.7	7.2
	Inno	67.3	79.7	82.4	87.1	89.3	92.8
Services							
Non–Inno	Non–Inno	87.1	84.6	85.3	79.6	76.0	77.1
	Inno	12.9	15.5	14.7	20.4	24.0	22.9
Inno	Non–Inno	40.5	40.4	30.7	21.4	28.8	12.8
	Inno	59.5	59.6	69.3	78.6	71.2	87.2

Notes:
Manufacturing: 1994–2002, service sector: 1996–2002. Sample: Unbalanced Panel.
Source: Own calculations.

Figure 3: Entry into and Exit from Innovation Activities by Industry



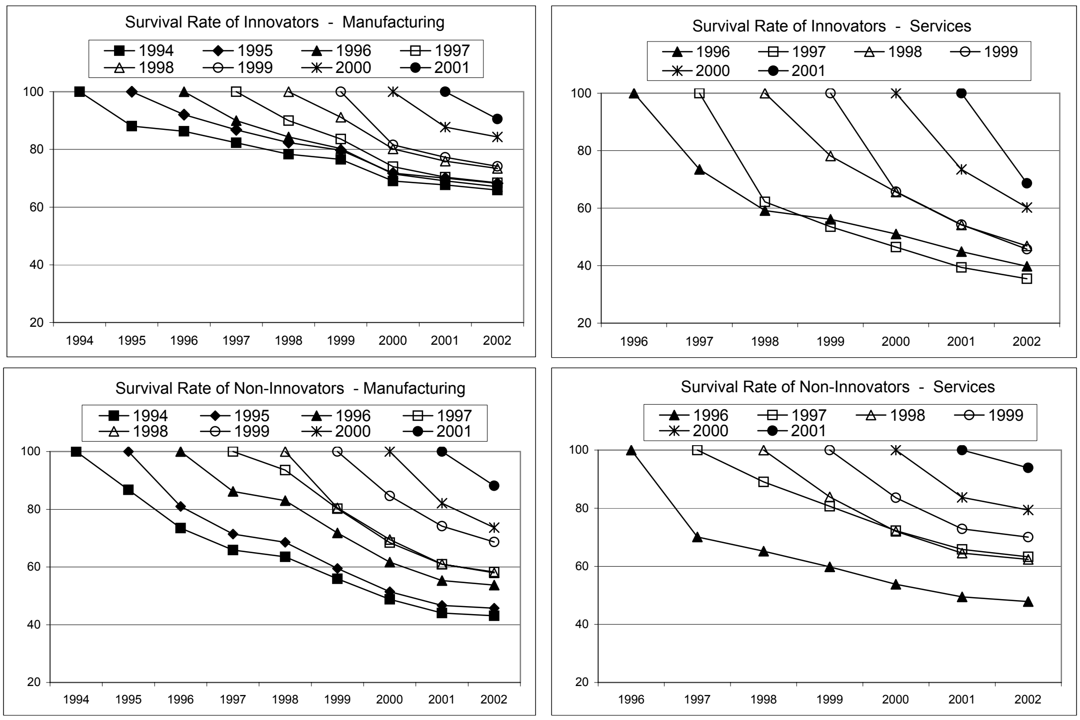
Notes:
See Figure 2.
Source: Own calculation.

facturing and, by and large, for service firms as well. The propensity to remain innovative increased with firm size, while at the same time the propensity for non–

innovators to take up such activities rose as well. Nevertheless, the (unconditional) state dependence effect measured as the difference between the probabilities of being innovative in period $t+1$ for innovators and for non–innovators in t was more pronounced in large manufacturing firms (approximately 72 percentage points for firms with more than 500 employees) than in small ones (59 percentage points for firms with less than 10 employees). The same picture emerges in services with a difference of 64 and 47 percentage points.

Figure 3 further demonstrates that innovation activities at the firm level are found to be more persistent in high–technology industries, though also quite permanent in some low–technology manufacturing and business–related service industries. For instance, the lowest exit rates can be found in R&D intensive industries like chemicals, vehicles, electrical engineering, medical instruments or machinery while exiting innovation activities is much more likely in the wood/paper, energy/water or construction industry or in most service industries.

Figure 4: Survival Rates of Innovator and Non–Innovator Cohorts by Years (in %)



Notes:
Sample: Balanced Panel.
Source: Own calculation.

Finally, we look at the innovative history of firms. Figure 4 depicts the survival rates of different innovator as well as non–innovator "cohorts" by years and Table 5 reports the number of (re–)entry into and (re–)exit from innovation. The survival rate for instance for the innovator cohort 1994 is the proportion of innovators in year

$t = 1994$ that was still innovating in year $t + s$, for $s = 1, 2 \ldots$. In manufacturing, the 3–year survival rates were quite similar for different cohorts, amounting to 78 % on average (based on cohorts 1994 to 1999). After 5 years, on average 71 % of the innovators were still innovating and 66 % of initially innovative firms (i.e., cohort 1994) were continuously engaged in innovation throughout the whole period. In services, the survival rates are smaller and exhibit higher variances. On average only 51 % of the innovators were still involved in innovation after three years, and the share of incessant innovators (40 %) is also much lower (even though the period for services is shorter). Survival rates of non–innovator cohorts in manufacturing turned out to be lower in general than for innovators, for instance 67 % on average after three years. About 43 % of the initial non–innovators kept out of innovation activities throughout the whole period. In the service sector these last two figures were very similar with 67 % and 48 %.

Table 5 further indicates that concerning those firms which experienced at least one change in their innovation behaviour (45 % in manufacturing and 55 % in services), we find a stronger tendency to return to the initial innovation status. This also implies that re–entry into innovation occurs to a non–negligible extent.

Table 5: Innovation History of Firms: Number of Entries into and Exits from Innovation Activities

	Manufacturing			Services		
Number of changes	Total	Non–Inno in $t = 0$	Inno in $t = 0$	Total	Non–Inno in $t = 0$	Inno in $t = 0$
0	54.9	43.1	65.9	45.0	47.8	39.8
1	11.2	13.7	8.9	13.1	6.5	25.5
2	19.0	24.2	14.2	22.7	28.3	12.2
3	8.5	10.4	6.6	10.3	7.6	15.3
4	4.8	6.6	3.1	6.4	8.2	3.1
5	1.1	1.4	0.9	1.8	0.5	4.1
6	0.5	0.5	0.5	0.7	1.1	0.0
Total	100	100	100	100	100	100

Notes:
Figures are calculated as share of total firms, initial non–innovators and innovators, respectively.
Sample: Balanced Panel.
Source: Own calculations.

7 Econometric Analysis

7.1 Econometric Model

Though interesting, transition rates only depict the degree of persistence, but don't offer a clue to the causes of this phenomenon since we do not control for observed or unobserved individual characteristics. In the following we therefore investigate whether and to which extent the observed persistence is due to underlying differences in individual characteristics and / or due to a genuine causal effect of past on future innovations, using a dynamic random effects probit model. The same model was applied for studying state dependence effects in poverty state (Biewen 2004) or export behaviour (Kaiser and Kongstedt 2004). This panel data approach allows us to distinguish between the sources of the persistence over time observed in the data and to control for individual heterogeneity. If individual heterogeneity is present but not controlled for, the coefficients of the observed characteristics are likewise biased if both are correlated.

We start on the assumption that a firm i will invest in innovation in period t if the expected present value of profits accruing to the innovation investment y_{it}^* is positive. The expected profit depends on the previous (realised) innovation experience $y_{i,t-1}$, on some observable explanatory variables summarised in the k–dimensional row vector x_{it} and on unobservable firm–specific attributes which are assumed to be constant over time and captured by μ_i. The structural model is thus given by:

$$y_{it}^* = \gamma\, y_{i,t-1} + x_{it}\,\beta + \mu_i + \varepsilon_{it} \qquad i = 1, \ldots, N, \qquad t = 1, \ldots, T \tag{1}$$

The effect of other time–varying unobservable determinants is summarised in the idiosyncratic error ε_{it}. It is assumed that $\varepsilon_{it}|y_{i0}, \ldots, y_{i,t-1}, x_i$ is $i.i.d.$ as $N(0,1)$ and that $\varepsilon_{it} \perp (y_{i0}, x_i, \mu_i)$ where $x_i = (x_{i1}, \ldots, x_{iT})$. N is the number of firms and the index t runs from 1 to 8 in manufacturing and 1 to 6 in services. If y_{it}^* is larger than a constant threshold (without any loss of generality we assume zero) we observe that firm i engages in innovation where I denotes the indicator function:

$$y_{it} = I\,[y_{it}^* > 0] \tag{2}$$

7.2 Estimation Method

For estimation purposes we have to solve two important theoretical and practical problems. First, the treatment of the unobserved heterogeneity μ_i, and secondly the treatment of the initial value y_{i0}. A random effects (RE) model rests on assumptions about the distribution of μ_i given the observables, whereas a fixed effects (FE) model assumes that μ_i is random but without specification of the distribution making it in fact preferable. However, there is no general solution in the literature how to estimate dynamic FE binary choice panel models because no general transformation is known how to elimate unobserved effects, i.e., unlike in linear models a first difference or within transformation does not eliminate μ_i in non–linear models. Honoré and Kyriazidou (2000) proposed a semiparametric estimator for the FE logit model, but their estimator is extremely data demanding and cannot be used here. Carro (2003) suggested a modified maximum likelihood estimator for the dynamic probit model, but the estimator is only consistent when T goes to infinity.[17] Therefore, I decide to apply a RE model.

Concerning the second problem, there are in general three different ways of handling the initial condition y_{i0} in parametric dynamic non–linear models. The first one is to assume that y_{i0} is a non–random constant, which is usually not a realistic assumption. The second solution is to allow for randomness of y_{i0} and to attempt to find the joint density for y_{i0} and all outcomes y_{it} conditional on strictly exogenous variables x_i. This approach starts on the joint distribution $(y_{i0}, \ldots, y_{iT})|\mu_i, x_i$ and it requires us to specify the distributions of $y_{i0}|\mu_i, x_i$ and that of $\mu_i|x_i$ to integrate out the unobserved effect. However, the joint distribution can only be found in very special cases. Heckman (1981b) thus suggested a method to approximate the conditional distribution. Another possibility is to assume that y_{i0} is likewise random and to specify the distribution of μ_i conditional on y_{i0} and x_i, which leads to the joint density of $(y_{i1}, \ldots, y_{iT})|y_{i0}, x_i$. This was first suggested by Chamberlain (1980) for a linear AR(1) model without covariates and Wooldridge (2005) used the same assumption to develop an estimator for dynamic nonlinear RE models, for instance dynamic RE probit, logit or tobit models. Following this latter estimation strategy, I further assume that the individual heterogeneity depends on the initial condition and the strict exogenous variables in the following way:

$$\mu_i = \alpha_0 + \alpha_1 y_{i0} + \bar{x}_i \alpha_2 + a_i, \tag{3}$$

where $\bar{x}_i = T^{-1} \sum_{t=1}^{T} x_{it}$ denotes the time–averages of x_{it}. Adding the means of the explanatory variables as a set of controls for unobserved heterogeneity is intuitive

[17] But Monte Carlo studies have shown that this estimator performs quite well for 8 or more time periods. The estimator is based on the idea of getting a reparametrisation in such a way that the incidental parameters are information orthogonal to the other parameters which reduces the order of the bias of the ML without increasing its asymptotic variance (see Cox and Reid 1987).

in the sense that we are estimating the effect of changing x_{it} but holding the time average fixed.[18] For the error term a_i we assume:

$$a_i \;\sim\; i.i.d. \;\; N(0,\sigma_a^2) \quad \text{and} \quad a_i \perp (y_{i0},\bar{x}_i) \tag{4}$$

and thus $\mu_i|y_{i0},\bar{x}_i$ follows a $N(\alpha_0+\alpha_1 y_{i0}+\bar{x}_i\,\alpha_2,\sigma_a^2)$ distribution. Having specified the distribution of the individual heterogeneity in this way, Wooldridge (2005) showed that the probability of being an innovator is given by:

$$P\left(y_{it}=1|y_{i0},\ldots,y_{i,t-1},x_i,\bar{x}_i,a_i\right)=\Phi\left(\gamma\, y_{i,t-1}+x_{it}\,\beta+\alpha_0+\alpha_1\, y_{i0}+\bar{x}_i\,\alpha_2+a_i\right) \tag{5}$$

Integrating out a_i in (5) yields a likelihood function that has the same structure as in the standard RE probit model, except that the explanatory variables are enriched by the initial condition and the time averages of the strict exogenous variables:

$$z_{it}=(1,x_{it},y_{i,t-1},y_{i0},\bar{x}_i) \tag{6}$$

Identification of the parameters requires that the exogenous variables vary across time and industry. If the structural model contains time–invariant regressors like industry dummies, one can include them in the regression to increase explanatory power. However, it is not possible to separate out the direct effect and the indirect effect via the heterogeneity equation unless it is assumed a priori that μ_i is partially uncorrelated with the industry dummies. Time dummies which are the same for all i are excluded from $\bar{x}_i$.

The first advantage of the proposed estimator is that it is computationally attractive. The approach further allows selection and panel attrition to depend on the initial condition (innovation state). The third advantage is that partial effects are identified and can be estimated. This is not possible in semiparametric approaches since they don't specify the distribution of individual heterogeneity on which partial effects depend. This allows us not only to determine whether true state dependence exists by referring to the significance level of the coefficient of the lagged dependent variable, but also on the importance of this phenomenon. One problem in estimating partial effects is the fact that firm heterogeneity is unobservable. Two alternative calculation methods have been proposed to deal with this shortcoming. The first way is to estimate the partial effect as in the standard probit

[18] Instead of $\bar{x}_i$ the original estimator used $x_i=(x_{i1},\ldots,x_{iT})$ in equation (3), but time–averages are allowed to reduce the number of explanatory variables (see Wooldridge 2005).

model and to assume that the individual heterogeneity μ_i takes its average value (PEA). $E(\mu_i) = \alpha_0 + \alpha_1 E(y_{i0}) + E(\bar{x}_i)\,\alpha_2$ and can be consistently estimated by $\widehat{E(\mu_i)} = \widehat{\alpha}_0 + \widehat{\alpha}_1\,\bar{y}_0 + \bar{x}\,\widehat{\alpha}_2$, where $\bar{y}_0 = \sum_{i=1}^{N} y_{i0}$ and $\bar{x} = \sum_{i=1}^{N} \bar{x}_i$. For the binary lagged dependent variable we can therefore calculate the marginal effect as the discrete change in the probability as the dummy variables changes from 0 to 1:

$$\widehat{PEA} = \Phi\left[\widehat{\gamma} + x_i\,\widehat{\beta} + \widehat{\alpha}_0 + \widehat{\alpha}_1\,\bar{y}_0 + \bar{x}\,\widehat{\alpha}_2\right] - \Phi\left[x_i\,\widehat{\beta} + \widehat{\alpha}_0 + \widehat{\alpha}_1\,\bar{y}_0 + \bar{x}\,\widehat{\alpha}_2\right] \quad (7)$$

This estimate suffers from the fact that usually the average value only represents a small fraction of firms. Alternatively, one can estimate partial effect after averaging the unobserved heterogeneity across firms. The average partial effect (APE) for the lagged dependent variable is estimated by:

$$\begin{aligned}\widehat{APE} &= \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\Phi\left[\widehat{\gamma}_a + x^o\,\widehat{\beta}_a + \widehat{\alpha}_{0a} + \widehat{\alpha}_{1a}\,y_{i0} + \bar{x}_i\,\widehat{\alpha}_{2a}\right] \\ &\quad -\frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\Phi\left[x^o\,\widehat{\beta}_a + \widehat{\alpha}_{0a} + \widehat{\alpha}_{1a}\,y_{i0} + \bar{x}_i\,\widehat{\alpha}_{2a}\right] \end{aligned} \quad (8)$$

where the subscript a denotes the original parameter estimates multiplied by $(1 + \widehat{\sigma}_a^2)^{-0.5}$ and x^o and y^o are fixed values that have to be chosen (here sample means averaged across i and t are used).[19]

One limitation of the estimator is that it was derived for balanced panels which evidently reduces the number of observations included. But using the sub–sample of balanced data still leads to consistent ML estimators under certain assumptions. More critical is the fact that, as in alternative estimation methods for dynamic discrete choice panel models (e.g., Heckman 1981a,b or Honoré and Kyriazidou 2000), the consistency hinges on the strict exogeneity assumption of the regressors

[19] The APE of one of the explanatory variables measures the change of the expected probability of innovating at time t due to an infinitesimal increase in that variable, where the expectation is over the distribution of the individual heterogeneity μ_i. Wooldridge (2005) calculated the APE for a specific point in time. However, since the panel data set used in this study has a rather long time dimension, we would like to calculate an average APE over the whole period. In the case of the lagged dependent variable, for instance, we are interested in:

$$\begin{aligned} APE &= E\left[P(y = 1 | y_{i,t-1} = y^o_{-1} = 1, x_{it} = x^o, y_{i0}, \bar{x}_i)\right] \\ &\quad -E\left[P(y = 1 | y_{i,t-1} = y^o_{-1} = 0, x_{it} = x^o, y_{i0}, \bar{x}_i)\right]. \end{aligned} \quad (9)$$

In that case a consistent estimator can be yielded by equation (8). I would like to thank J. Wooldridge and F. Laisney for helpful discussions on this point. Any errors remain those of the author.

and the estimator leads to inconsistent results if the distributional assumptions are not valid. Blindum (2003) and Biewen (2004) both extended the estimator to allow for endogenous dummy variables, but not for a continuous variable that fails strict exogeneity which seems to be more critical in our analysis. Honoré and Lewbel (2002) and Lewbel (2005) recently proposed a semiparametric approach which does not require the strict exogeneity assumption. However, their estimator is based on the existence of one "very exogenous" regressor, and there seems to be no variable at hand that satisfies this assumption in our case.[20]

7.3 Empirical Model Specification

Theoretical and empirical studies have identified a whole array of innovation determinants; firm size and market structure are the oldest and most prominent ones (see Schumpeter 1942). Firm size is measured by the log number of employees in the previous period (SIZE) and the market structure is captured by the Herschmann–Herfindahl index (HHI) from the previous year measured on a three–digit level, see Table 6 for more detailed variable definitions.

The modern innovation literature stresses that there are additional firm–level determinants other than firm size and market structure. Cohen (1995) distinguished between *firm* and *industry or market* characteristics. Widely–considered firm characteristics explaining innovation activities are product diversification (Nelson 1959), the degree of internationalisation, the availability of financial resources (e.g., Müller 1967, Bond et al. 1999 or Kukuk and Stadler 2001) and technological capabilities. As the data set does not contain information on product diversification for all years, we cannot take this hypothesis into account. The degree of international competition is measured by the export intensity (EXPORT) and the availability of financial resources is proxied by an index of creditworthiness (RATING). While a positive impact of EXPORT is expected, the hypothesis is that RATING negatively affects the propensity to innovate since the index ranges from 1 (best rating) to 6 (worst rating) and thus a higher value of RATING implies that less external funding is

[20] The key assumption is that of conditional independence. This means that when the values of the other covariates x_{it} are known, the conditional distribution of $\mu_i + \varepsilon_{it}$ is not altered by additional knowledge of the "very exogenous" continuous regressor v_{it}, i.e., $f(\mu_i + \varepsilon_{it}|x_{it}) = f(\mu_i + \varepsilon_{it}|x_{it}, v_{it})$. In my case the idiosyncratic errors and fixed effects capture, for instance, risk attitudes, innovation preferences, management abilities or technological opportunities. The assumption will hold if a continuous explanatory variable exists, that is assigned to firms independently of these unobserved attributes. However, there seems to be no variable at hand that satisfies this assumption. In labour supply models, government benefits income might fulfill this requirement.

available and that it is more costly due to higher interest rates, making fewer innovation projects profitable. The data set does not contain a measure for internal financial resources, like profit or cash–flow. On the other side, both enter the index of creditworthiness, and thus RATING also reflects internal financial capabilities. In addition to innovation experience, technological capabilities are mainly determined by the skills of employees. Hence, I operationalise this construct by means of three variables: the share of employees with a university degree (HIGH), a dummy variable equaling 1 if a firm has not invested in training its employees in the previous period (NOTRAIN) and the amount of training expenditure per employee (TRAINEXP).

One aim of government support programmes is to promote innovation activities. To test whether public funding induces a permanent change in favour of innovation, I further include a dummy variable equaling 1 if the enterprise has received any public financial support for innovation activities in the previous period (PUBLIC). Since all firms which receive financial support are innovators by definition, PUBLIC is an interaction term and measures the additional effect of supported compared to non–supported innovators.

The estimation also controls for ownership structure by distinguishing between public limited companies (PLC), private limited liability companies (LTD) and private partnerships (PRIVPART). One hypothesis stressed by the principal agency theory is that managers prefer to carry out less risky investment and innovation projects than owners because managers are more closely related to the company and they will be threatened with the loss of their job if the investment fails while owners can spread their risk by diversification strategies (Jensen and Meckling 1976).

In addition, firm–specific variables reflecting firm age (AGE), location (EAST), whether the firm is part of an enterprise group (GROUP) and whether the group's headquarter is located abroad (FOREIGN) are included. On the one side, enterprises which are part of a conglomerate may have easier access to external capital in a world of capital market imperfections and we would therefore hypothesise a positive relationship. But clearly, GROUP may also capture other effects of the companies' organisational structure on innovative activities. On the other side, some authors have stressed that foreign owned firms are less engaged in innovation activities. One argument in favour of a negative link is that R&D plays a crucial role in the long term strategic planning of a company and managers wish to maintain direct control over such activities, therefore R&D activities usually take place at or in close proximity to the companies' headquarters (see Howell 1984 or Bishop and Wiseman 1999).[21] The observed period is characterised by the catching–up process of the

[21] Kleinknecht and Poot (1992) linked this argument into a product life cycle approach. They

Eastern German economy after reunification and the share of innovators had been found higher at the aggregate level in East than in West Germany until the end of the ninetees (see Rammer et al. 2005). Therefore, we expect a higher propensity to innovate for East German firms.[22]

As mentioned above, market or industry characteristics – alone or in combination with firm–specific features – may be important for innovation activities. In this context technological opportunities are expected to play a significant role. The concept of technological opportunities can be summarised by the fact that the prevailing technological dynamics (basic inventions, spillover potentials of new technologies) in some industries spur innovation stronger than in other industries. Nelson (1988) showed in a theoretical model that improved technological opportunities increase the incentive to invest in R&D. Technological opportunities are measured by the product life cycle of a firm's main product (LCYCLE) and industry dummies.

Table 6: Variable Definition

Variable	Type[a)]	Definition
Alternative endogenous variables		
INNO	0/1	1 if a firm i has positive innovation expenditure in year t. Innovation expenditure includes expenditure for intramural and extramural R&D, acquisition of external knowledge, machines and equipment, training, market introduction, design and other preparations for product and/or process innovations.
INNO_RD	0/1	1 if a firm i has positive expenditure for intramural and extramural R&D in year t.
INNO_NRD	0/1	1 if a firm i has positive innovation expenditure in year t, but no R&D activities.
Explanatory variables:		
Variables varying across individuals and time		
SIZE	c	Number of employees of firm i in year $t-1$, in logarithm.
LCYCLE	c	Length of product life cycle (in years) of firm's i main product in year $t-1$, in logarithm.

Continued on next page.

argue that early stages of a cycle are associated with considerable R&D activities which are therefore carried out close to the headquarters, while less R&D activities are necessary in later stages for incremental product or process modifications and can therefore be decentralised.

[22] Note, that the catching–up process in East Germany was patronised by special government support programmes.

Table 6 – *continued from previous page*

Variable	Type[a)]	Definition
RATING	c	Credit rating index for firm i in year $t-1$, originally ranging between 100 (highest creditworthiness) and 600 (worst creditworthiness), divided by 100 to get appropriately scaled coefficients.
AGE	c	Age of firm i at the beginning of year t, in logarithm.
GROUP	0/1	1 if firm i belongs to a group in year t.
PUBLIC	0/1	1 if firm i received public funding for innovation projects in year $t-1$.
NOTRAIN	0/1	1 if firm i has no training expenditure in year $t-1$.
TRAINEXP	c	Training expenditure per employee (in logarithm) of firm i in year $t-1$ if NOTRAIN=0, otherwise 0.
HIGH	c	Share of employees with a university or college degree in firm i in year $t-1$, divided by 100.
EXPORT	c	Export intensity of firm i in year $t-1$ defined as exports/sales.
EXPORT2	c	Squared export intensity.
Variables varying across industries and time		
HHI	c	Hirschman–Herfindahl Index in year $t-1$, on the 3–digit industry NACE level, divided by 100 to get appropriately scaled coefficients. Only available for manufacturing.
Time–constant individual–specific variables		
FOREIGN	0/1	1 if firm i is a subsidiary of a foreign company.
EAST	0/1	1 if firm i is located in Eastern Germany.
PLC	0/1	1 if firm i is a public limited company (*AG*).
LTD	0/1	1 if firm i is a private limited liability company (*GmbH, GmbH & Co. KG*).
PRIVPART	0/1	1 if firm i is a private partnership (*Personengesellschaft, OHG, KG*).
IND	0/1	System of 15 and 9 dummies grouping industries and services respectively, see Table 14.
Time–varying individual–constant variables		
TIME	0/1	System of time dummies for each year.

Notes:

[a)] c: continuous variable.

Table 7 reports the descriptive statistics of the variables used in the estimations. It turned out that for almost all variables the variation across firms (between variation) is much higher compared to that within a firm over time. The variables FOREIGN, EAST, PLC, LTD and PRIVPART can vary across i and t. However, due to the fact that hardly any within variation showed up, we treated them as time–constant firm–specific variables in the estimation and include them only in equation (3).

Table 7: Descriptive Statistics[a)]

		Manufacturing						Services					
	Unit	Mean	Std.dev. Overall	Std.dev. Between	Std.dev. Within	Min	Max	Mean	Std.dev. Overall	Std.dev. Between	Std.dev. Within	Min	Max
INNO	[0/1]	0.555	0.497	0.419	0.268	0	1	0.360	0.480	0.372	0.304	0	1
INNO_RD	[0/1]	0.465	0.499	0.442	0.231	0	1	0.158	0.365	0.308	0.195	0	1
INNO_NRD	[0/1]	0.090	0.287	0.163	0.236	0	1	0.202	0.402	0.254	0.312	0	1
SIZE[b)]	no. empl.	2018.7	14121.3	13566.9	3964.5	1	243638	1782.0	18143.6	18107.3	1512.0	1	271078
LCYCLE[b)]	years	15.4	21.4	21.0	4.3	0.3	200	16.2	22.6	22.0	5.2	1	100
RATING	[Index: 1–6]	2.088	0.600	0.548	0.244	1	6	2.194	0.440	0.407	0.167	1	6
AGE[b)]	years	21.8	23.0	22.5	4.9	0	142	22.3	21.0	20.9	2.4	1	141
GROUP	[0/1]	0.360	0.480	0.409	0.253	0	1	0.223	0.416	0.349	0.227	0	1
PUBLIC	[0/1]	0.243	0.429	0.351	0.246	0	1	0.096	0.295	0.248	0.161	0	1
NOTRAIN	[0/1]	0.176	0.381	0.315	0.215	0	1	0.255	0.436	0.377	0.220	0	1
TRAINEXP[b)]	€	663.2	1135.8	872.0	728.9	0	7702	1223.1	3164.0	2264.1	2213.5	0	25791
HIGH	[0–1]	0.110	0.136	0.117	0.069	0	1	0.200	0.260	0.236	0.110	0	1
EXPORT	[0–1]	0.196	0.246	0.232	0.083	0	1	0.025	0.096	0.084	0.046	0	1
HHI	[Index: 0–100]	4.7	6.1	5.6	2.4	0.1	43.2	—	—	—	—	—	—
FOREIGN	[0/1]	0.059	0.236	0.196	0.131	0	1	0.018	0.134	0.118	0.064	0	1
EAST	[0/1]	0.344	0.475	0.469	0.075	0	1	0.420	0.494	0.491	0.054	0	1
PLC	[0/1]	0.078	0.268	0.268	0.000	0	1	0.053	0.225	0.221	0.042	0	1
LTD	[0/1]	0.830	0.376	0.375	0.028	0	1	0.692	0.462	0.457	0.072	0	1
PRIVPART	[0/1]	0.085	0.279	0.278	0.028	0	1	0.220	0.414	0.410	0.063	0	1
Obs				3496						1692			

Notes:
[a)] For the period 1995–2002 (manufacturing) and 1997–2002 (services). In case of lagged explanatory variables, periods used are 1994–2001 and 1996–2001.
[b)] Variable values shown are not log–transformed. For estimation purposes, however, a log–transformation of these variables is used to take the skewness of the distribution into account.

7.4 Econometric Results

Table 8 reports the estimation results of the dynamic RE probit model, including the Schumpeter determinants (size and market structure), product life cycle, and industry and time dummies as exogenous variables, and compares the results with the static pooled model and static RE model. In all tables M_ denotes the individual time–average of the corresponding variable. Note that marginal effects are reported. In the dynamic RE model they are calculated at the average value of the firm–specific error.[23] Furthermore, in the case of the static pooled model, the standard errors have been adjusted to account for the fact that observations are not necessarily independent within firms.

The first main result is that including the lagged dependent variable is an important part of the model specification. That is, even after accounting for individual unobserved heterogeneity, the variable turns out to be highly significant in both manufacturing and services, confirming therefore the hypotheses of true state dependence. The results further show that some of the variables which are significant in the static estimation lose this property in the dynamic specification; for instance, firm size is no longer significant in services. One interpretation of this result is that firm size, which is likewise highly time–persistent, simply picks up the impact of the lagged dependent variable in the static case.

As mentioned above, one problem of the dynamic RE panel probit model is the fact that strict exogeneity of the exogenous variables is assumed. This implies that no feedback effects from the innovation variable on future values of the explanatory variables are allowed, which seems to be contestable for some of the variables usually explaining innovation behaviour, e.g. firm size, market structure or export behaviour. To assess the impact of including variables which potentially fail the strict exogeneity assumption on the estimated state dependence effect, I apply a stepwise procedure. That is, I start estimating an extremely parsimonious specification (1) including only LCYCLE and industry and time dummies as exogenous variables. Specification (2) then adds the Schumpeter determinants (which underlie the comparison) and (3) incorporates some firm characteristics for which strict exogeneity seems to be statisfied.[24] Specifications (4) and (5) further include some variables that are presumably not strictly exogenous. The estimation results are summarised in Tables 9 and 10 for manufacturing and services, respectively.

[23] However, the calculation of the marginal effect of a variable k neglects that an infinitesimal increase in x_k also changes the mean value $\bar{x}_{i,k}$.

[24] I used the procedure proposed by Wooldridge (2002), i.e. I added the lead of the corresponding variable and tested on the significance of the coefficient.

Table 8: Comparison: Marginal Effects in Static Pooled, Static Random Effects and Dynamic Random Effects Probit Model

	Manufacturing			Services		
	Pooled Probit	Static RE Probit	Dynamic RE Probit	Pooled Probit	Static RE Probit	Dynamic RE Probit
$INNO_{-1}$	—	—	0.358***	—	—	0.127***
			(0.035)			(0.044)
LCYCLE	-0.055**	-0.089***	-0.052	0.017	-0.003	0.003
	(0.021)	(0.016)	(0.044)	(0.047)	(0.061)	(0.109)
SIZE	0.141***	0.216***	0.129**	0.086***	0.134***	0.019
	(0.014)	(0.016)	(0.062)	(0.014)	(0.021)	(0.064)
HERFIN	0.018	0.034	0.050	—	—	—
	(0.041)	(0.034)	(0.056)			
$INNO_0$	—	—	0.535***	—	—	0.457***
			(0.045)			(0.063)
M_LCYCLE	—	—	0.018	—	—	-0.041
			(0.050)			(0.099)
M_SIZE	—	—	-0.035	—	—	0.042
			(0.063)			(0.066)
M_HERFIN	—	—	-0.038	—	—	—
			(0.070)			
σ_a	—	1.861	0.801	—	1.367	0.928
		(0.103)	(0.082)		(0.119)	(0.107)
ρ	—	0.776	0.391	—	0.651	0.463
		(0.019)	(0.049)		(0.040)	(0.058)
LR_ρ	—	0.000	0.000	—	0.000	0.000
W_{TIME}	0.005	0.008	0.010	0.000	0.000	0.000
W_{IND}	0.000	0.000	0.000	0.000	0.000	0.000
$\ln L$	-1820.1	-1249.7	-1107.2	-935.2	-760.4	-722.5
$\ln L_{Cons}$	-2402.1	-1403.1	-1403.1	-1105.5	-828.9	-828.9
R^2_{MF}	0.242	0.109	0.211	0.154	0.083	0.128
R^2_{MZ}	0.476	—	—	0.303	—	—
Obs Prob	55.5	55.5	55.5	36.0	36.0	36.0
Pred Prob	57.7	71.8	64.7	34.8	26.6	28.5
Corr Pred	71.5	69.9	85.4	72.4	72.5	77.0
Corr Pred 1	77.8	80.0	86.0	42.2	43.5	59.4
Corr Pred 0	63.7	57.4	84.7	89.4	88.8	86.8
Obs	3496	3496	3496	1692	1692	1692

Notes:
$***$, $**$ and $*$ indicate significance on a 1%, 5% and 10% level, respectively. Standard errors in pooled probit model adjusted for clustering on firms. A constant (significant at the 1% level in each regression) as well as time and industry dummies are included in each regression, but not reported.

It emerges from this exercise that the marginal effect of the lagged dependent variable is nearly unaltered in the different estimations. That is, even after accounting for individual unobserved heterogeneity, past innovation has a behavioural effect: Conditional on observed and unobserved firm characteristics, an innovator in $t-1$ has a probability of innovating which is approximately 36 percentage points higher than that of a non–innovator in manufacturing. For service companies the marginal effect amounts to roughly 13 percentage points.

The results further show that the initial condition is also highly significant in both samples. This implies a substantial correlation between firms' initial innovation status and the unobserved heterogeneity.

A third important finding is that in addition to past innovation experience, knowledge provided by skilled employees has a crucial influence on generating innovations over time. In both industries the variables NOTRAIN and TRAIN, and in manufacturing also HIGH, turn out to be significant in the equation explaining individual heterogeneity across firms. That is, firms which do not invest in further training of their employees have a significantly lower propensity to innovate, while for those firms which do invest, an increase in training expenditure per employee of 1 per cent raises the probability of innovating by about 5.5 percentage points in both industries. All in all, these results confirm and highlight the role of innovative capabilities in the dynamics of firms' innovation behaviour.

Fourth, the results provide evidence that unobserved heterogeneity is a key factor for innovation persistence. The importance of the unobserved heterogeneity in explaining the total variance can be gauged from $\rho = \sigma_a^2/(1+\sigma_a^2)$.[25] Table 8 has already shown that introducing the lagged dependent variable leads to a distinct reduction of the importance of the unobserved heterogeneity. Unobserved heterogeneity still explains between 30 and 43 % of the variation in the dependent variable in manufacturing depending on the specification of μ_i. In the service sector this effect is in a similar range with 37 to 48 %.

[25] Note that $\varepsilon_{it}|y_{i0},\ldots,y_{i,t-1},X_i \sim N(0,1)$ and $\mu_i|y_{i0},\bar{X}_i \sim N(0,\sigma_a^2)$.

Table 9: Robustness of Dynamic RE Probit Estimations in Manufacturing

Regression	(1)	(2)	(3)	(4)	(5)
Structural Equation					
$INNO_{-1}$	0.364***	0.358***	0.359***	0.358***	0.333***
	(0.034)	(0.035)	(0.035)	(0.035)	(0.036)
LCYCLE	-0.049	-0.052	-0.057	-0.043	-0.053
	(0.044)	(0.044)	(0.045)	(0.031)	(0.044)
SIZE	—	0.129**	0.122**	0.111*	0.100*
		(0.062)	(0.062)	(0.062)	(0.061)
HERFIN	—	0.050	0.051	0.048	0.055
		(0.056)	(0.056)	(0.057)	(0.056)
RATING	—	—	-0.059	-0.066	-0.068
			(0.044)	(0.044)	(0.044)
AGE	—	—	-0.075*	-0.071*	-0.067*
			(0.038)	(0.038)	(0.037)
GROUP	—	—	0.053	0.052	0.062
			(0.050)	(0.050)	(0.050)
NOTRAIN	—	—	—	-0.123	-0.116
				(0.162)	(0.160)
TRAINEXP	—	—	—	0.014	0.014
				(0.017)	(0.017)
HIGH	—	—	—	-0.100	-0.103
				(0.214)	(0.216)
EXPORT	—	—	—	0.459***	0.473***
				(0.136)	(0.130)
PUBLIC	—	—	—	—	0.174***
					(0.045)
TIME	yes	yes	yes	yes	yes
Individual Heterogeneity					
$INNO_0$	0.625***	0.535***	0.538***	0.460***	0.341***
	(0.042)	(0.045)	(0.045)	(0.045)	(0.047)
M_LCYCLE	0.023	0.018	0.021	0.030	0.017
	(0.051)	(0.050)	(0.050)	(0.050)	(0.049)
M_SIZE	—	-0.035	-0.035	-0.047	-0.056
		(0.063)	(0.064)	(0.064)	(0.063)
M_HERFIN	—	-0.038	-0.040	-0.038	-0.044
		(0.070)	(0.071)	(0.069)	(0.067)
M_RATING	—	—	0.030	0.026	0.032
			(0.062)	(0.061)	(0.059)
M_AGE	—	—	0.119**	0.116**	0.100**
			(0.051)	(0.050)	(0.047)
M_GROUP	—	—	0.024	-0.020	-0.026
			(0.085)	(0.082)	(0.078)
FOREIGN	—	—	-0.128	-0.162**	-0.125
			(0.084)	(0.083)	(0.079)

Continued on next page.

Table 9 – *continued from previous page*

Regression	(1)	(2)	(3)	(4)	(5)
EAST	—	—	0.016	0.047	-0.051
			(0.051)	(0.051)	(0.051)
PLC	—	—	-0.209*	-0.201**	-0.168*
			(0.110)	(0.102)	(0.097)
PRIVPART	—	—	0.025	0.038	0.025
			(0.069)	(0.064)	(0.060)
M_NOTRAIN	—	—	—	-0.638***	-0.651***
				(0.247)	(0.236)
M_TRAINEXP	—	—	—	0.053*	0.054**
				(0.029)	(0.027)
M_HIGH	—	—	—	0.646**	0.157
				(0.316)	(0.312)
M_EXPORT	—	—	—	0.347*	0.289
				(0.198)	(0.194)
M_PUBLIC	—	—	—	—	0.370***
					(0.091)
IND	yes	yes	yes	yes	yes
σ_a	0.876	0.801	0.792	0.709	0.623
	(0.083)	(0.082)	(0.082)	(0.077)	(0.077)
ρ	0.434	0.391	0.386	0.334	0.280
	(0.047)	(0.049)	(0.049)	(0.049)	(0.050)
LR_ρ	0.000	0.000	0.000	0.000	0.000
W_{TIME}	0.013	0.010	0.006	0.007	0.009
W_{IND}	0.000	0.000	0.000	0.010	0.030
ln L	-1132.2	-1107.3	-1099.8	-1077.5	-1047.1
ln L_{Cons}	-1403.1	-1403.1	-1403.1	-1403.1	-1403.1
R^2_{MF}	0.193	0.211	0.216	0.232	0.254
Obs Prob	55.5	55.5	55.5	55.5	55.5
Pred Prob	63.8	64.7	64.7	64.6	65.7
Corr Pred	83.6	85.4	85.6	86.1	87.4
Corr Pred 1	84.1	86.0	86.4	86.4	87.2
Corr Pred 0	83.0	84.7	84.7	85.7	87.7
Obs	3496	3496	3496	3496	3496

Notes:
$***$, $**$ and $*$ indicate significance on a 1%, 5% and 10% level, respectively. Marginal effects are reported, calculated at the average value of the individual–specific error. Columns (4) and (5) report the marginal effect of EXPORT, corrected for the fact that the original regressions also contain the quadratic term. Standard errors were calculated using the delta method. Original coefficient estimates in (4) and (5): EXPORT: 1.604 (0.784) and 1.762 (0.770), EXPORT2: -2.659 (0.906) and -2.710 (0.882). W_{IND} and W_{TIME} test for the null hypothesis that the industry and time dummies are jointly equal to zero. Estimations are based on Gauss–Hermite quadrature approximations using 8 quadrature points. The accuracy of the results have been checked using the STATA command quadchk. Most coefficients change by less than 0.01% and none change by more than 1%, so that the model can be reliably fitted using the quadrature approach.

Table 10: Robustness of Dynamic RE Probit Estimations in Services

Regression	(1)	(2)	(3)	(4)	(5)
Structural Equation					
INNO_{-1}	0.126***	0.127***	0.128***	0.128***	0.103**
	(0.044)	(0.044)	(0.045)	(0.045)	(0.047)
LCYCLE	-0.009	0.003	-0.002	-0.005	-0.039
	(0.109)	(0.109)	(0.109)	(0.109)	(0.112)
SIZE	—	0.019	0.016	0.011	0.006
		(0.064)	(0.064)	(0.066)	(0.069)
RATING	—	—	-0.210**	-0.209**	-0.206**
			(0.099)	(0.099)	(0.103)
AGE	—	—	0.053	0.050	0.057
			(0.060)	(0.059)	(0.062)
GROUP	—	—	0.006	0.009	0.010
			(0.063)	(0.063)	(0.066)
NOTRAIN	—	—	—	-0.060	-0.068
				(0.155)	(0.161)
TRAINEXP	—	—	—	0.003	0.007
				(0.020)	(0.021)
HIGH	—	—	—	-0.027	-0.016
				(0.127)	(0.133)
EXPORT	—	—	—	0.109	0.084
				(0.311)	(0.320)
PUBLIC	—	—	—	—	0.294***
					(0.102)
TIME	yes	yes	yes	yes	yes
Individual Heterogeneity					
INNO_0	0.532***	0.457***	0.434***	0.370***	0.335***
	(0.059)	(0.063)	(0.064)	(0.065)	(0.064)
M_LCYCLE	-0.047	-0.041	-0.046	-0.044	-0.021
	(0.097)	(0.099)	(0.098)	(0.099)	(0.102)
M_SIZE	—	0.042	0.021	0.022	0.021
		(0.066)	(0.068)	(0.070)	(0.073)
M_RATING	—	—	0.084	0.122	0.176
			(0.123)	(0.122)	(0.125)
M_AGE	—	—	-0.149**	-0.127*	-0.118
			(0.075)	(0.075)	(0.076)
M_GROUP	—	—	0.071	0.070	0.057
			(0.106)	(0.104)	(0.105)
FOREIGN	—	—	0.270	0.214	0.278
			(0.203)	(0.202)	(0.193)
EAST	—	—	0.040	0.022	-0.025
			(0.062)	(0.063)	(0.063)
PLC	—	—	0.216	0.211	0.281*
			(0.166)	(0.162)	(0.158)

Continued on next page.

Table 10 – *continued from previous page*

Regression	(1)	(2)	(3)	(4)	(5)
PRIVPART	—	—	-0.064	-0.049	-0.015
			(0.059)	(0.058)	(0.060)
M_NOTRAIN	—	—	—	-0.594**	-0.649**
				(0.270)	(0.273)
M_TRAINEXP	—	—	—	0.055*	0.056*
				(0.034)	(0.034)
M_HIGH	—	—	—	0.151	0.010
				(0.205)	(0.209)
M_EXPORT	—	—	—	0.201	0.006
				(0.428)	(0.460)
M_PUBLIC	—	—	—	—	0.528***
					(0.159)
IND	yes	yes	yes	yes	yes
σ_μ	0.966	0.928	0.886	0.850	0.777
	(0.109)	(0.107)	(0.105)	(0.104)	(0.102)
ρ	0.482	0.463	0.440	0.420	0.376
	(0.056)	(0.058)	(0.059)	(0.060)	(0.062)
LR_ρ	0.000	0.000	0.000	0.000	0.000
W_{TIME}	0.000	0.000	0.000	0.000	0.000
W_{IND}	0.000	0.000	0.000	0.145	0.138
ln L	-729.7	-722.5	-712.1	-703.9	-680.4
ln L_{Cons}	-828.9	-828.9	-828.9	-828.9	-828.9
R^2_{MF}	0.120	0.128	0.141	0.150	0.179
Obs Prob	36.0	36.0	36.0	36.0	36.0
Pred Prob	28.0	28.8	28.8	28.4	30.5
Corr Pred	76.7	77.0	78.7	79.1	80.1
Corr Pred 1	63.4	59.4	62.7	63.6	63.6
Corr Pred 0	84.3	86.8	87.7	87.9	89.4
Obs	1692	1692	1692	1692	1692

Notes:
$***$, $**$ and $*$ indicate significance on a 1%, 5% and 10% level, respectively. Marginal effects are reported, calculated at the average value of the individual–specific error. The Wald test statistics W_{IND} and W_{TIME} test for the null hypothesis that the industry and time dummies are jointly equal to zero, respectively. As in manufacturing, the accuracy of the results have been proved using the STATA command quadchk.

In addition to prior innovation experience, skills and unobserved heterogeneity, some observed firm characteristics are also found to be crucial factors in explaining innovation. These results are by and large in line with the literature and with what we expected. Firms that are more financially constrained are less likely to engage in innovation. This effect is highly significant in services and slightly significant in manufacturing (p–value: 0.128 in the preferred specification (4)). Moreover, firms which receive public funding in the previous period exhibit a higher propensity to innovate in the subsequent period than innovators without financial support in both industries. In contrast, firm size is only important in manufacturing, not in the service sector. This is likewise the case for the degree of internationalisation, a result which is maybe not that surprising because exporting is less prevalent in services.[26] Firms which are more active on international markets have a higher propensity to innovate in manufacturing. However, we find an inverse U–shaped relationship for the export intensity with an estimated point of inflexion at 33 % in specification (4). It is also only in manufacturing that ownership matters. That is, public limited companies, in which conflicts of interests between managers and shareholders might arise, have a significantly lower conditional probability of being innovative. However, regarding the second Schumpeterian determinant, we do not find any significant impact of market concentration on innovation. But admittedly, this may be due to the fact that HHI is a bad proxy of market structure.

All in all, our model seems to fit the data quite well. The McFadden's pseudo R2 varies between 20 and 25 % in manufacturing and based on the preferred specification (4) the model correctly predicts the innovation behaviour for 86 % of the observations. This number is much higher than in the static model. Correct predictions in the service sector are likewise high with 79 %. However, the model clearly performs worse in predicting the occurrence of innovation for service firms.

As mentioned above, partial effects at average value (PEA) suffer from the fact that usually the average value only represents a small fraction of firms. To amplify what has been said so far on the importance of state dependence effects, Table 11 contrasts the PEA with the estimated average partial effect (APE). It is quite plain that averaging the unobserved heterogeneity across firms reduces the estimates of the state dependence effects. Section 6 has shown that the propensity to innovate in period $t + 1$ was approximately 74 percentage points higher for innovators than for non–innovators in period t in panel B. Controlling for differences in observed and unobserved characteristics, this differences reduces to 36 percentage points using PEA and 23 percentage points using APE. This implies that depending on the calculation method between nearly one third (APE) to one half (PEA) of the innovation persistence in manufacturing can be traced back to true state dependence, while the rest was due to observed and unobserved characteristics. In the service sector state dependence accounts for about 15 (APE) to 25 (PEA) % of the observed persistence.

[26] I also experimented with dummy variables for the export status or export classes, but in no case does export exhibit a significant impact on innovation in services.

Table 11: Importance of State Dependence Effects in Manufacturing and Services

	OSD	PEA[a)]				APE[b)]			
		$\widehat{P(1\|1)}$	$\widehat{P(1\|0)}$	$\widehat{PEA}$		$\widehat{P(1\|1)}$	$\widehat{P(1\|0)}$	$\widehat{APE}$	
				abs.	rel.			abs.	rel.
Manufacturing	74.1	79.3	43.5	35.8	*48.3*	68.9	45.9	23.0	*31.0*
Services	53.7	36.9	24.0	12.9	*25.0*	41.1	32.9	8.2	*15.3*

Notes:
OSD: Observed state dependence effect calculated .
a) $\widehat{P(1|1)}$ and $\widehat{P(1|0)}$ denote estimates of the probabilities $P(y_{it} = 1|y_{i,t-1} = 1, x_i, \mu_i)$ and $P(y_{it} = 1|y_{i,t-1} = 0, x_i, \mu_i)$ at the average value of μ_i.
b) $\widehat{P(1|1)}$ and $\widehat{P(1|0)}$ are estimates of the expected probabilities of $P(y_{it} = 1|y^o_{i,t-1} = 1, x^o_i, \mu_i)$ and $P(y_{it} = 1|y^o_{i,t-1} = 0, x^o_i, \mu_i)$ where the expectation is over the distribution of μ_i.
All estimates are based on specification (4) in Tables 9 and 10.

7.5 Sensitivity Analysis

In this section, some further sensitivity analyses are carried out to check on the robustness of the results. Firstly, using each value $x_i = (x_{i1}, \ldots, x_{iT})$ in equation (3) instead of individual time–averages as proposed by Wooldridge (2005) leaves the results nearly unaltered. They are therefore not reported here, but are available upon request.

Secondly, Table 12 differentiates betweeen R&D–performing and non–R&D–performing innovators to examine whether persistence is mainly driven by R&D activities and whether this can explain the difference found between manufacturing and services. The results suggest that significant state dependence effects exist for both kinds of innovation activities in both samples. But as expected, persistence effects are much higher for R&D–performing than for non–R&D–performing innovators. Furthermore, the marginal effect of past R&D experience is nearly three times higher in manufacturing with 50 percentage points than in the service sector with 16 percentage points. On the other hand, in case of innovators without R&D activities the impact of past innovation experience on the propensity to remain innovative is very much the same in manufacturing with 7 and in services with 9 percentage points. By and large, the main conclusions drawn in the previous section still hold in the separate estimations.

Table 12: Persistence of Non-R&D– and R&D–Performing Innovators

	Manufacturing		Services	
Dep. Var.	INNO_NRD	INNO_RD	INNO_NRD	INNO_RD
Structural Equation				
INNO_NRD_{-1}	0.070***	—	0.093***	—
	(0.022)		(0.034)	
INNO_RD_{-1}	—	0.500***	—	0.159**
		(0.037)		(0.077)
LCYCLE	-0.014	-0.025	-0.093	-0.062**
	(0.010)	(0.058)	(0.061)	(0.031)
SIZE	-0.007	0.158**	0.011	-0.011
	(0.014)	(0.076)	(0.037)	(0.016)
HERFIN	-0.003	0.073	—	—
	(0.014)	(0.063)		
RATING	-0.006	-0.059	-0.070	-0.010
	(0.012)	(0.052)	(0.056)	(0.028)
AGE	-0.005	-0.083	0.033	-0.005
	(0.008)	(0.052)	(0.033)	(0.013)
GROUP	-0.004	0.073	0.017	-0.004
	(0.011)	(0.060)	(0.038)	(0.014)
NOTRAIN	0.020	-0.090	-0.000	-0.022
	(0.051)	(0.192)	(0.095)	(0.027)
TRAINEXP	0.004	-0.005	-0.003	0.003
	(0.004)	(0.021)	(0.012)	(0.005)
HIGH	-0.051	-0.036	-0.100	0.029
	(0.052)	(0.242)	(0.079)	(0.030)
EXPORT	-0.017	0.637***	-0.106	0.063
	(0.027)	(0.157)	(0.196)	(0.078)
Individual Heterogeneity				
INNO_NRD_0	0.059**	—	0.172***	—
	(0.026)		(0.049)	
INNO_RD_0	—	0.472***	—	0.166***
		(0.061)		(0.059)
M_LCYCLE	0.012	-0.008	0.039	-0.050*
	(0.011)	(0.066)	(0.055)	(0.028)
M_SIZE	-0.009	-0.029	-0.016	0.018
	(0.015)	(0.079)	(0.039)	(0.017)
M_HERFIN	-0.004	-0.054	—	—
	(0.017)	(0.080)		
M_RATING	0.022	-0.004	0.062	-0.010
	(0.014)	(0.073)	(0.067)	(0.033)
M_AGE	0.004	0.117*	-0.040	-0.008
	(0.010)	(0.066)	(0.040)	(0.017)
M_GROUP	0.025	-0.088	-0.009	0.003
	(0.017)	(0.100)	(0.057)	(0.023)

Continued on next page.

Table 12 – *continued from previous page*

	Manufacturing		Services	
Dep. Var.	INNO_NRD	INNO_RD	INNO_NRD	INNO_RD
FOREIGN	-0.008	-0.125	0.004	0.057
	(0.015)	(0.084	(0.088)	(0.078)
EAST	-0.016*	0.139**	0.033	-0.013
	(0.009)	(0.67)	(0.034)	(0.013)
PLC	0.018	-0.175*	-0.060	0.078
	(0.027)	(0.104)	(0.045)	(0.079)
PRIVPART	0.012	-0.063	0.001	-0.018*
	(0.016)	(0.089)	(0.032)	(0.010)
M_NOTRAIN	-0.034**	-1.088***	-0.247*	-0.061
	(0.017)	(0.345)	(0.142)	(0.056)
M_TRAINEXP	0.010*	0.106***	0.016	0.003
	(0.006)	(0.037)	(0.018)	(0.007)
M_HIGH	0.115*	0.899**	0.089	0.030
	(0.076)	(0.374)	(0.116)	(0.043)
M_EXPORT	0.004	0.286	-0.156	0.029
	(0.042)	(0.229)	(0.250)	(0.096)
σ_a	0.590	0.828	0.689	0.713
	(0.078)	(0.105)	(0.095)	(0.178)
ρ	0.258	0.407	0.322	0.337
	(0.049)	(0.061)	(0.060)	(0.111)
LR_ρ	0.000	0.000	0.000	0.001
W_{TIME}	0.041	0.004	0.000	0.504
W_{IND}	0.504	0.017	0.308	0.126
ln L	-858.6	-820.6	-694.5	-298.5
R^2_{MF}	0.084	0.320	0.088	0.330
APE: INNO_NRD	0.088	—	0.088	—
APE: INNO_RD	—	0.292	—	0.170
Obs	3496	3496	1692	1692

Notes:
$***$, $**$ and $*$ indicate significance on a 1%, 5% and 10% level, respectively. Marginal effects are reported, calculated at the average value of the individual–specific error. Time and industry dummies are included in each regression. The Wald test statistics W_{IND} and W_{TIME} test for the null hypothesis that the industry and time dummies are jointly equal to zero, respectively.

Moreover, as was set out in section 5, the results so far measured the persistence in innovation input. For manufacturing, the picture can be completed by examining the output persistence for the same set of firms. I use a dummy variable indicating whether the firm has introduced a new product or process within a 3–year period (INOUT) and take only every third survey into account to avoid overlapping, i.e. I used the periods 1994–1996, 1997–1999 and 2000–2002. This strategy leads to a larger reduction of the number of observations. It turns out that the lagged dependent variable is highly significant again and the partial effects are very similar in magnitude, as can be seen from Table 13. That is, the results corroborate true state dependence in innovation output as well. Furthermore, the other main findings asserted for the innovation input are confirmed for the innovation output indicator.

Table 13: Innovation Input and Output Persistence in Manufacturing

	Innovation Input	Innovation Output
Dependent Variable	INNO	INOUT
PEA	35.8	34.2
APE	23.0	21.5
Obs	3496	874

Notes:
Estimates are based on the same specification as in column (4) in Table 9.

8 Conclusion

In this paper I analysed the persistence of innovation behaviour of firms based on data for German manufacturing and services during the period 1994–2002. Using the estimator recently proposed by Wooldridge (2005) for dynamic binary choice panel data models, I have analysed whether innovation behaviour shows persistence at the firm level and whether state dependence drives this phenomenon.

A first main finding is that innovation behaviour is permanent at the firm level to a very large extent. Year–to–year transition rates indicate that in manufacturing nearly nine out of ten innovating firms in one period persisted in innovating in the subsequent period and about 84 per cent of non–innovators maintained their state in the following period. Yet innovation is not a once and for all phenomenon. 45 per cent of manufacturing and 55 per cent of service firms experienced at least one change in their innovation behaviour. In general, persistence is less pronounced in the service sector and exhibits a higher variance across time. Less surprisingly, persistence turns out to be higher in larger firms and in high–technology industries, but is nevertheless relatively high in small firms.

The econometric results confirm the hypothesis of true state dependence. Partial effects were calculated highlighting the importance of this phenomenon. Depending on the calculation method, about one third to one half of the difference in the propensity to innovate between previous innovators and non–innovators in manufacturing can be traced back to true state dependence. In the service sector, persistence is generally less prevalent and state dependence effects are less pronounced, yet still highly significant. The fact that innovation performance exhibits true state dependence implies that innovation–stimulating policy measures such as governmental supporting programmes have the potential of long–lasting effects because they do not only affect the current innovation activities but are likely to induce a permanent change in favour of innovation.

The results confirm and highlight the role of innovative capabilities on the dynamics in firms' innovation behaviour. In addition to past innovation experience, knowledge provided by skilled employees has found to be important in generating innovations over time.

The results further emphasise the important role of unobserved heterogeneity in explaining the persistence of innovation. Leaving out this source of persistence in the empirical analysis can lead to highly misleading results. Some observed firm characteristics like size or export behaviour (determinants which themselves show high persistence) also make some firms also more innovation–prone than others.

One topic on the agenda of future research is to test for dynamic completeness, that is, to extend the estimator to allow for more complex lag structures of the lagged endogenous variable. So far I have assumed that dynamics are correctly specified by a first order process.

References

Aghion, P. and Howitt, P. (1992). A Model of Growth Through Creative Destruction, *Econometrica* 60, 323–351.

Aghion, P. and Saint-Paul, G. (1998). Virtues of Bad Times: Interaction Between Productivity Growth and Economic Fluctuations, *Macroeconomic Dynamics* 2, 322–344.

Biewen, M. (2004). Measuring State Dependence in Individual Poverty Status: Are There Feedback Effects to Employment Decisions and Household Composition?, *IZA Discussion Paper 1138*, Bonn.

Bishop, P. and Wiseman, N. (1999). External Ownership and Innovation in the United Kingdom, *Applied Economics* 31(4), 443–450.

Blindum, S. (2003). Relaxing the Strict Exogeneity Assumption in a Dynamic Random Probit Model, *CAM Working Papers 2003-04*, Copenhagen.

Blundell, R., Griffith, R. and Van Reenen, J. (1993). Knowledge Stocks, Persistent Innovation and Market Dominance: Evidence from a Panel of British Manufacturing Firms, *IFS Working Paper W93/19*, London.

Blundell, R., Griffith, R. and Van Reenen, J. (1995). Dynamic Count Data Models of Technological Innovation, *The Economic Journal* 105(429), 333–344.

Bond, S., Harhoff, D. and Van Reenen, J. (1999). Investment R&D and Financial Constraints in Britain and Germany, *IFS Working Paper W99/5*, London.

Carro, J. M. (2003). Estimating Dynamic Panel Data Discrete Choice Models with Fixed Effects, *CEMFI Working Paper 0304*, Madrid.

Cefis, E. (1999). *Persistence in Innovative Activities. An Empirical Analysis*, PhD thesis, European University Institute, Florence.

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

Cefis, E. and Orsenigo, L. (2001). The Persistence of Innovative Activities. A Cross–Country and Cross–Sectors Comparative Analysis, *Research Policy* 30(7), 1139–1158.

Chamberlain (1980). Analysis of Covariance with Qualitative Data, *Review of Economic Studies* 47, 225–238.

Cohen, W. (1995). Empirical Studies of Innovative Activity, *in:* Stoneman, P. (ed.), *Handbook of the Economics of Innovation and Technological Change*, Oxford, Cambridge, 182–264.

Cohen, W. M. and Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation, *Administrative Science Quarterly* 35, 128–158.

Cox, D. and Reid, N. (1987). Parameter Orthogonality and Approximate Conditional Inference, *Journal of the Royal Statistical Society* 49, 1–39.

Crépon, B., Duguet, E. and Mairesse, J. (1998). Research Innovation and Productivity: An Econometric Analysis at the Firm Level, *Economics of Innovation and New Technology* 7, 115–158.

Duguet, E. and Monjon, S. (2004). Is Innovation Persistent at the Firm Level? An Econometric Examination Comparing the Propensity Score and Regression Methods, *Cahiers de la Maison des Sciences Economiques v04075*, Paris.

EUROSTAT (1992). *Nace Rev. 1*, Paris.

Flaig, G. and Stadler, M. (1994). Success breeds Success. The Dynamics of the Innovation Process, *Empirical Economics* 19, 55–68.

Flaig, G. and Stadler, M. (1998). On the Dynamics of Product and Process Innovations. A Bivariate Random Effects Probit Model, *Jahrbücher für Nationalökonomie und Statistik* 217, 401–417.

Geroski, P. (1995). What do we Know about Entry?, *International Journal of Industrial Organization* 13, 421–440.

Geroski, P. A., Van Reenen, J. and Walters, C. F. (1997). How Persistently Do Firms Innovate?, *Research Policy* 26(1), 33–48.

Gottschalk, S., Janz, N., Peters, B., Rammer, C. and Schmidt, T. (2002). Innovationsverhalten der deutschen Wirtschaft: Hintergrundbericht zur Innovationserhebung 2001, *ZEW Dokumentation 02–03*, Mannheim.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey, *Journal of Economic Literature* 28, 1661–1707.

Heckman, J. J. (1981a). Heterogeneity and State Dependence, *in:* Rosen, S. (ed.), *Studies in Labour Markets*, Chicago, 91–139.

Heckman, J. J. (1981b). Statistical Models for Discrete Panel Data, *in:* Manski, C. and McFadden, D. (eds), *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge, Mass., 114–178.

Himmelberg, C. P. and Petersen, B. C. (1994). R&D and internal Finance: A Panel Study of Small Firms in High-Tech Industries, *Review of Economics and Statistics* 38–51.

Honoré, B. and Kyriazidou, E. (2000). Panel Data Discrete Choice Models with Lagged Dependent Variables, *Econometrica* 68, 839–874.

Honoré, B. and Lewbel, A. (2002). Semiparametric Binary Choice Panel Data Models Without Strictly Exogenous Regressors, *Econometrica* 70(5), 2053–2063.

Howell, J. (1984). The Location of Research and Development: Some Observations and Evidence from Britain, *Regional Studies* 18, 13–29.

Janz, N., Ebling, G., Gottschalk, S., Peters, B. and Schmidt, T. (2002). *Innovation Activities in the German Economy – Report on Indicators from the Innovation Survey 2001*, Mannheim.

Janz, N., Lööf, H. and Peters, B. (2004). Firm Level Innovation and Productivity - Is There a Common Story Across Countries?, *Problems and Perspectives in Management* 2, 184–204.

Jensen, M. C. and Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure, *Journal of Financial Economics* 3, 305–360.

Judd, K. (1985). On the Performance of Patents, *Econometrica* 53, 567–585.

Kaiser, U. and Kongsted, H. C. (2004). True Versus Spurious State Dependence in Firm Performance: The Case of West German Exports, *CEBR Working Paper 2004–07*, CEBR.

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

Kamien, M. and Schwartz, N. (1982). *Market Structure and Innovation*, Cambridge, Mass.

Kleinknecht, A. (1990). Are There Schumpeterian Waves of Innovations?, *Cambridge Journal of Economics* 14, 81–92.

Kleinknecht, A. and Poot, T. (1992). Do Regions Matter for R&D?, *Regional Studies* 26(3), 221–232.

König, H., Laisney, F., Lechner, M. and Pohlmeier, W. (1994). On the Dynamics of Process Innovative Activity: An Empirical Investigation Using Panel Data, *in:* Oppenländer, K.-H. and Poser, G. (eds), *The Explanatory Power of Business Cycles Surveys*, Avebury, 243–262.

Kukuk, M. and Stadler, M. (2001). Financing Constraints and the Timing of Innovations in the German Services Sector, *Empirica* 28, 277–292.

Lewbel, A. (2005). *Simple Endogenous Binary Choice and Selection Panel Model Estimators*. mimeo.

Lööf, H. and Heshmati, A. (2001). Knowledge Capital and Performance Heterogeneity: A Firm Level Innovation Study, *International Journal of Production Economics* 0, 1–25.

Love, J. H. and Roper, S. (2001). Location and Network Effects on Innovation Success: Evidence for UK, German and Irish Manufacturing Plants, *Research Policy* 30, 643–661.

Mairesse, J., Hall, B. and Mulkay, B. (1999). Firm–Investment in France and the United States: An Exploration of What We Have Learned in Twenty Years, *Annales d'Economie et de Statistiques* 55–56, 27–67.

Malerba, F. and Orsenigo, L. (1993). Technological Regimes and Firm Behaviour, *Industrial Corporate Change* 2(1), 45–71.

Malerba, F. and Orsenigo, L. (1999). Technological Entry, Exit and Survival: An Empirical Analysis of Patent Data, *Research Policy* 28, 643–660.

Manez Castillejo, J. A., Rochina Barrachina, M. E., Sanchis Llopis, A. and Sanchis Llopis, J. (2004). *A Dynamic approach to the Decision to Invest in R&D: The Role of Sunk Costs.* mimeo.

Mansfield, E. (1968). *Industrial Research and Technological Innovation: An Econometric Analysis*, New York.

Mulkay, B., Hall, B. and Mairesse, J. (2001). Investment and R&D in France and in the United States, *in:* Deutsche Bundesbank (ed.), *Investing Today for the World of Tomorrow: Studies on the Investment Process in Europe*, New York.

Müller, D. C. (1967). The Firm‘s Decision Process: An Econometric Investigation, *Quarterly Journal of Economics* 81, 58–87.

Nelson, R. (1988). Modelling the Connections in the Cross-Section between Technical Progress and R&D Intensity, *RAND Journal of Economics* 19, 478–485.

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

Nelson, R. and Winter, S. (1982). *An Evolutionary Theory of Economic Change*, Cambridge.

OECD and EUROSTAT (1997). *Oslo Manual, Proposed Guidelines for Collecting and Interpreting Technological Innovation Data*, Paris.

Phillips, A. (1971). *Technology and Market Structure: A Study of the Aircraft Industry*, Lexington.

Rammer, C., Aschhoff, B., Doherr, T., Peters, B. and Schmidt, T. (2005). *Innovation in Germany – Results of the German Innovation Survey 2004*, Mannheim.

Raymond, W., Mohnen, P., Palm, F. and Schim van der Loeff, S. (2005). *The Dynamics of the Innovation Process in Dutch Manufacturing: A Panel Data Analysis.* mimeo.

Reinganum, J. F. (1983). Uncertain Innovation and the Persistence of Monopoly, *American Economic Review* 73(4), 741–748.

Rogers, M. (2004). Networks, Firm Size and Innovation, *Small Business Economics* 22(2), 141–153.

Romer, P. (1990). Endogenous Technological Change, *Journal of Political Economy* 98(5), 71–102.

Sachverständigenrat (2004). *Jahresgutachten 2004/05: Erfolge im Ausland - Herausforderungen im Inland*, Wiesbaden.

Schmookler, J. (1966). *Invention and Economic Growth*, Cambridge.

Schumpeter, J. (1939). *Business Cycles*, 2 edn, New York.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*, New York.

Simons, K. (1995). *Shakeouts: Firm Survival and Technological Change in New Manufacturing Industries*, PhD thesis, Carnegie Mellon University.

Stiglitz, J. and Weiss, A. (1981). Credit Rationing in Markets with Imperfect Information, *American Economic Review* 71, 393—410.

Stoneman, P. (1983). *The Economic Analysis of Technological Change*, New York.

Sutton, J. (1991). *Sunk Costs and Market Structure*, Cambridge.

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

Wooldridge, J. (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.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*, Cambridge.

Appendix: Tables

Table 14: Branches of Industry Covered by the MIP

Industry Sector		Service Sector	
Branches of Industry	NACE[a)]	Branches of Industry	NACE[a)]
Mining	$10-14$	Distributive services	
Manufacturing		Wholesale	51
Food	$15-16$	Retail/repairing	$50, 52$
Textile	$17-19$	Transport/storage/post	$60-63, 64.1$
Wood/paper/printing	$20-22$	Real estate/renting	$70-71$
Chemicals	$23-24$	Business related services	
Plastic/rubber	25	Banks/insurances	$65-67$
Glass/ceramics	26	Computer/telecomm.	$72, 64.2$
Metals	$27-28$	Technical services	$73, 74.2-74.3$
Machinery	29	Consultancies	$74.1, 74.4$
Electrical engineering	$30-32$	Other BRS[b)]	$74.5-74.8, 90,$
MPO[c)] instruments	33		$92.1-92.2$
Vehicles	$34-35$		
Furniture/recycling	$36-37$		
Energy/water	$40-41$		
Construction	45		

Notes:
[a)] The industry definition is based on the classification system NACE Rev.1 (Nomenclature gnrale des activits conomiques dans les les Communauts Europennes) as published by EUROSTAT (1992), using 2–digit or 3–digit levels.
[b)] Business related services.
[c)] Medical, precision and optical instruments.

Table 15: Individual Participation Pattern

	Total			Manufacturing		Services	
No. of	firms		obs	*firms*[b)]	obs	*firms*[a)]	obs
Participation[a)]	#	%	#	#	#	#	#
1	5949	43.3	5949	*2803*	2803	*3146*	3146
2	2499	18.2	4998	*1223*	2446	*1276*	2552
3	1769	12.9	5307	*876*	2629	*893*	2678
4	1109	8.1	4436	*575*	2298	*535*	2138
5	803	5.8	4015	*464*	2320	*339*	1695
6	590	4.3	3540	*323*	1936	*267*	1604
7	560	4.1	3920	*337*	2360	*223*	1560
8	253	1.8	2024	*253*	2024	–	–
9	220	1.6	1980	*220*	1980	–	–
Total	13752	100	36169	*7074*	20796	*6678*	15373

Notes:
[a]) The number of utilisable observations is higher than that which would arise from the participation pattern. This can be explained by the fact that since 1998 the survey is sent only to a sub-sample of firms in even years due to cost reasons. However, to maintain the panel structure with yearly waves, the most relevant variables are asked retrospectively for the preceding year in odd years.
[b]) Some firms have changed their main business activity which defines their industry assignment and have switched between manufacturing and services during the considered period. The number of firms is the average number of firms, calculated as the number of observations divided by the number of participation.
Source: ZEW, own calculations.

Table 16: Distribution of the Unbalanced and Balanced Panel in Manufacturing

Distribution by:	Panel[a)]			Difference		Distribution by:	Panel[a)]			Difference	
	T	U	B	B-T	B-U		T	U	B	B-T	B-U
Industry[b)]						**Size[b)]**					
Mining	2.0	2.1	1.7	−0.3	−0.4	0-4	2.7	1.8	1.6	−1.2	−0.3
Food	6.3	6.0	5.5	−0.8	−0.5	5-9	6.9	6.5	5.5	−1.3	−1.0
Textile	5.2	4.9	4.9	−0.3	−0.0	10-19	12.1	11.6	10.2	−1.8	−1.4
Wood/printing	6.7	6.5	6.4	−0.3	−0.0	20-49	17.8	18.2	19.7	+1.9	+1.5
Chemicals	6.6	6.8	8.7	+2.1	+1.9	50-99	15.2	15.7	14.3	−0.8	−1.3
Plastic/rubber	6.8	7.7	8.4	+1.6	+0.8	100-199	13.0	13.7	13.8	+0.8	+0.2
Glass/ceramics	4.7	5.0	5.5	+0.8	+0.6	200-499	15.5	16.4	17.5	+2.0	+1.1
Metals	13.2	13.4	11.5	−1.6	−1.8	500-999	7.6	8.0	8.3	+0.7	+0.3
Machinery	14.3	14.5	13.0	−1.3	−1.5	1000+	8.9	8.2	9.1	+0.3	+1.0
Electrical engineering	8.0	7.8	7.8	−0.2	+0.0						
Medical instr.	6.5	6.8	7.8	+1.3	+1.1	**Region[b)]**					
Vehicles	4.6	4.5	4.4	−0.2	−0.1	West	68.2	66.8	65.7	−2.6	−1.1
Furniture/recycling	4.2	3.6	3.8	−0.4	+0.2	East	31.8	33.2	34.3	+2.6	+1.1
Energy/water	4.4	4.8	5.9	+1.5	+1.1						
Construction	6.6	5.9	4.6	−2.0	−1.3	**Innovators[b)]**	59.3	57.8	55.1	−4.2	−2.7
Total Obs	27116	13558	3933				27116	13558	3933		

Notes:
[a)] T: Unbalanced panel of all firms within the period 1994–2002. U: Unbalanced panel of firms with at least 4 consecutive observations within 1994–2002. B: Balanced panel of firms within 1994–2002.
[b)] Calculated as share of total number of observations (in %).
Source: Own calculations.

Table 17: Distribution of the Unbalanced and Balanced Panel in the Service Sector

Distribution by:	Panel[a)]			Difference		Distribution by:	Panel[a)]			Difference	
	T	U	B	B-T	B-U		T	U	B	B-T	B-U
Industry[b)]						**Size[b)]**					
Wholesale	11.4	12.0	10.7	−0.7	−1.2	0-4	7.3	7.2	9.4	+2.1	+2.1
Retail	10.4	12.8	11.9	+1.5	−0.8	5-9	13.9	15.4	14.2	+0.3	+1.1
Transport	15.4	18.8	18.8	+3.4	+0.0	10-19	17.7	19.5	19.1	+1.4	+0.4
Bank/insurance	11.1	10.0	9.2	−1.8	−0.8	20-49	19.5	22.2	20.0	+0.4	−2.2
Computer	8.3	6.8	7.1	−1.1	+0.3	50-99	11.3	12.1	12.9	+1.6	+0.8
Technical serv.	14.4	13.5	11.5	−2.9	−2.0	100-199	9.6	9.8	11.0	+1.4	+1.2
Consultancies	7.8	6.7	8.2	+0.4	+1.5	200-499	8.0	7.0	6.5	−1.5	−0.5
Other BRS	13.8	12.0	12.8	−1.0	+0.8	500-999	4.5	2.8	1.8	−2.7	−0.9
Real estate/renting	6.7	7.5	9.7	+3.0	+2.2	1000+	7.9	4.1	5.2	−2.7	+1.1
						Region[b)]					
						West	62.5	57.4	57.9	−4.6	+0.5
						East	37.5	42.6	42.1	+4.6	−0.5
						Innovators[b)]	44.5	37.6	35.8	−8.6	−1.8
Total Obs	20493	7901	1974				20493	7901	1974		

Notes:
[a)] T: Unbalanced panel of all firms within the period 1996–2002. U: Unbalanced panel of firms with at least 4 consecutive observations within 1996–2002. B: Balanced panel of firms within 1996–2002.
[b)] Calculated as share of total number of observations (in %).
Source: Own calculations.