INTERNAL AND EXTERNAL FACTORS IN INNOVATION PERSISTENCE

Cristiano Antonelli
Dipartimento di Economia, Università di Torino
BRICK (Bureau of Research on Innovation, Complexity and Knowledge), Collegio Carlo Alberto.

Francesco Crespi
Corresponding author: crespi@uniroma3.it
Dipartimento di Economia, Università Roma Tre
BRICK (Bureau of Research on Innovation, Complexity and Knowledge), Collegio Carlo Alberto.

Giuseppe Scellato
Dipartimento di Ingegneria Gestionale e della Produzione, Politecnico di Torino
BRICK (Bureau of Research on Innovation, Complexity and Knowledge), Collegio Carlo Alberto.

ABSTRACT
This paper contributes the analysis of the persistence of innovation activities, as measured by total factor productivity (TFP) and explores its internal and external determinants stressing its path dependent characteristics. The external conditions, namely the quality of local knowledge pools and the strength of the Schumpeterian rivalry, together with the internal conditions, that is the actual levels of dynamic capabilities, as proxied by the levels of wages and the size of firms, exert a specific and localized effect upon the persistent introduction of innovations. A Multiple Transition Probability Matrixes (MTPMs) approach has been implemented to grasp the contingent effects of external effects on the long-term innovation persistence. The empirical analysis of the dynamics of firm level TFP for a sample of about 7000 Italian manufacturing companies observed during the years1996-2005 is based on both the comparison of different transition probability matrixes and on dynamic discrete choice panel data models. The evidence provided by the test of MTPMs in sub-periods suggests that innovation persistence is path dependent, as opposed to past dependent.

JEL CLASSIFICATION: O31, C23, C25, L20

KEY-WORDS: KNOWLEDGE CUMULABILITY; KNOWLEDGE EXTERNALITIES; INNOVATION; PERSISTENCE; SEQUENTIAL MARKOV CHAINS; PATH DEPENDENCE; TFP.

1.INTRODUCTION

According to the conventional economic wisdom, innovation is an exogenous random shock that like manna falls from heaven. The economics of innovation impinges upon the view that innovation is the deliberate and intentional result of the capability of firms to generate new knowledge and to apply it to new products, new processes, new organizational methods, new combinations of inputs and new markets (Nelson and Winter, 1982; Dosi et al. 1988; Fagerberg et al., 2005).

This approach leads to two quite distinct explanations of innovation persistence. The first, consistent with the resource based theory of the firm, elaborates view that innovation persistence is the result of given, intrinsic characteristics of the firm. Innovation capabilities are like time invariant endowments that display their effects. Innovation persistence is fully driven by the initial allocation of the innovation capabilities: firms are able to learn to learn (Penrose, 1959; Stiglitz, 1987; Teece and Pisano, 1994: Langlois and Foss, 1999).

The second elaborates the view that innovation persistence is a path dependent process where the probability to introduce an innovation at time t is indeed influenced by the introduction of an innovation at time t-1. However, the transition probability might change along time because of the effects of contingent events and specifically by the changing levels of knowledge externalities. The generation of new knowledge and the introduction of innovations are the conditional result of a creative and localized reaction that takes place when firms face unexpected events in both factor and product markets. A number of contextual and ever changing conditions, however, are necessary to make the reaction actually creative so as to lead to the introduction of an innovation as opposed to adaptive. In the latter case the lack of the contextual characteristics would enable firms to change techniques, in a given technical space, but would not lead to the introduction of a productivity enhancing novelty (Schumpeter, 1947).

In order to face un-expected events in both factor and product markets and the consequent out-of-equilibrium conditions events, firms try and mobilize the internal stocks of knowledge accumulated by means of learning processes. The chances that the reaction of the firm actually leads to the successful introduction of innovations, relies upon the access to the knowledge available in the surroundings. In other words, the reaction of firms to unexpected events becomes actually creative when both the competence accumulated by means of internal learning processes and a number of external conditions in terms of access conditions to the local pools of knowledge apply (Antonelli, 2008 and 2011).

According to this view, location, spatial proximity, the actual conditions of knowledge governance and the knowledge communication channels that link firms might enhance the processes of knowledge generation, favoring interactions among agents with diverse knowledge bases. Indeed, firms cluster together mainly for these specific reasons (Swann et al., 1998; Baptista and Swann, 1999). Long-distance coordination among agents and knowledge interactions can also be realized through organized proximity (Rallet and Torre, 2005). In this context, knowledge governance mechanisms and the characteristics of knowledge structure are particularly relevant (Quatraro, 2012).

Moving from the seminal contribution by Griliches (1979), a rich and detailed array of empirical studies confirm the pervasive role of technological spillovers in favoring the economic performances of clustered firms in terms of output, employment, labor productivity and total factor productivity. The following literature has interpreted these empirical findings as a reliable clue to assessing the positive effects of knowledge externalities upon the rate of introduction of technological changes by firms able to use external knowledge as an input in their own innovation process (Acs et al., 2002; Fritsch, 2002 and 2004; Fritsch and Franke, 2004).

Building upon this literature we put forward the hypothesis that innovation persistence is *path dependent*, as opposed to *past dependent*, because it is the result not only of the internal characteristics of firms, as the resource theory of the firm claims, but also of the –changing- characteristics of the context into which firms are localized. Knowledge externalities are strictly necessary for the reaction of firms to become actually creative. At the same time knowledge externalities are indeed external to each firms, but clearly internal to the system. When and if the characteristics of the context change, the results of the innovative effort also change. Hence, innovation persistence cannot be longer regarded as the result of a given, intrinsic capability of the firm that qualifies its action like an endowment given once for ever, but rather as the conditional result of a systemic and interactive process that keeps changing along time (Antonelli and Scellato, 2012).

Persistence is path dependent, as opposed to past dependent, as it is a dynamic process where hysteresis matters but is shaped by a number of complementary and contingent factors that shape locally the dynamics of the process.

The present paper builds upon the recognition that external technological knowledge represents an augmenting and facilitating factor in the introduction of technological innovations and makes a step further arguing that external knowledge is key factor in determining a path dependent innovation persistence characterized by contextual and conditional recursive feedbacks. The paper elaborates the hypothesis that the introduction of innovations is the persistent, emerging property of an economic system characterized by knowledge cumulability and complementarity both internal and external to firms. Indeed, the introduction of new technologies and new organizations methods affects the systems on two counts as it engenders further waves of unexpected events and Schumpeterian rivalry, and, at the same time, makes available new knowledge spillovers that add to the existing stock of external knowledge.

Knowledge cumulability consists in the inter-temporal, diachronic indivisibility of knowledge. As it is well-known the arrovian economics of knowledge assumes that knowledge is characterized by indivisibility and non-exhaustibility. Knowledge vintages add on and build up a stock of knowledge that does not wear out because of repeated use.

Indivisibility articulates in cumulability and complementarity among the different bits of knowledge. Next to the bits of knowledge, internally possessed by each agent, external bits of knowledge, possessed by other agents also play a central role. The generation of new knowledge is indeed possible only 'standing on the shoulders of giants', that is by means of access and use of the existing stock of knowledge. The existing stock of knowledge, however, is both internal to each firm and external, embedded in the other agents that belong to the same system (Colombelli and Von Tunzelmann, 2011).

As Peter Swann has convincingly shown, the evolving structure of the system changes endogenously as a consequence of the changing modes of interaction among firms, their entry and exit, their growth. The introduction of innovations is itself a major factor of change of the architecture of the system. The external conditions into which firms are embedded are at the same time a consequence and a cause of the recursive feedbacks that support the persistence of innovation activities (Swann et al., 1998; Baptista and Swann, 1998 and 1999; Beaudry and Swann, 2009).

Internal and external knowledge cumulability affects the dynamics of economic processes as the stock of knowledge that each firm can access and use internally and externally shapes the chances of generation of new knowledge. Such effects can change through time, as the rates of accumulation and the access conditions are not fixed. Inventions and scientific breakthrough can make some portions of the stock of knowledge obsolete. Changes in the structure of interactions and transactions can modify the access conditions to external knowledge. As such the effects of internal and external knowledge cumulability are typically path dependent, as opposed to past dependent. In the former case, the effects of hysteresis are qualified and shaped by the contingent changes that take place along the process. In the latter case, the process is shaped by the initial conditions only. The dynamics of the process is indeed influenced by a weak irreversibility that allows changes along the process to alter both the rate and the direction of the dynamics with typical path dependent effects (David, 1997 and 2007).

With this approach in the background, the aim of this work is threefold. First, we contribute the literature on the persistence of innovation by providing an empirical analysis based on total factor productivity measures. Second, we qualify the characteristics of the persistence and explore its external determinants, by specifically looking at the role of regional context and the characteristics of the product markets in shaping this process. Third we discuss in detail the methodological and theoretical implications of the use of TPMs with reference to Markov chains theory.

The rest of the paper is structured as it follows. Section 2 reviews the literature on the matter. Section 3 outlines the hypotheses and the research design of this study. Section 4 presents the econometric evidence. The conclusions summarize the main results.

2. PRIOR RESEARCH ON INNOVATION PERSISTENCE

In the special issue of the *International Journal of Industrial Organization* dedicated to the economics of path dependence, Malerba, Orsenigo and Peretto (1997) pave the way to the analysis of the persistence of innovation activities now explored by a growing literature synthesized in Table 1.

[Table 1 about here]

Earlier studies can be grouped into a subset of studies that build upon the analysis of large samples of patents and a subset of empirical studies that make use of data from innovation surveys. The persistence of innovation has been addressed by looking at various specific aspects, such as technological specialization (Malerba et al. 1997), comparative studies of cross-country and cross-sector evolution (Cefis and Orsenigo 2001; Raymond et al., 2010; Clausen et al., 2011), the empirical properties of the distribution of persistence (Cefis 2003) and the diverse typologies of innovative activities (Roper and Hewitt-Dundas, 2008; Peters, 2009; Martínez-Ros and Labeaga, 2009; Le Bas et al., 2011; Antonelli et al., 2012).

Though emerging from different contexts, some convergent conclusions appear to have been reached by previous studies. In particular, both innovators and non-innovators showed a strong tendency to remain within their states. The evidence shows that the degree of persistence varies according to the innovation indicator adopted (Duguet and Monjon, 2004). While the works that have used patents as indicator suggest that the persistence is weak and exhibits strong values only in the case of heavy patentees, empirical analyses based on survey data found stronger evidence of innovation persistence.

A number of factors have been associated with the presence of persistence in innovative activities. Among these the size of firms, profitability and the intensity of R&D activities emerged as crucial, confirming the hypothesis that the accumulation of knowledge over time tends to induce state dependence in innovative behavior and that the availability of internal funds enhance the possibility to continuously engage in innovation (Cefis and Ciccarelli, 2004; Latham and Le Bas, 2006; Peters, 2009).

The evidence suggesting that R&D based innovation activities tend to be associated with higher persistence appears to be of particular evidence as it helps explaining two important results emerging from previous literature. First, several contribution highlighted that innovation persistence is stronger in high-tech, science based industries where R&D activities are mainly concentrated (Raymond et al., 2010; Clausen et al., 2011). Second, when different innovation output indicators have been considered, it turned out that product innovation, very often linked to R&D investments (Crespi and Pianta, 2007), tends to be characterized by higher state dependence with respect to process innovation (Martínez-Ros and Labeaga, 2009; Antonelli et al., 2012). In this respect the complementarities between different types of innovation activities emerged as crucial in shaping different patterns of persistence (Clausen et al., 2011 and Antonelli et al., 2012), including the case of organizational innovation.

In the reviewed studies attention has been paid primarily to internal factors that qualify the persistence as the result of the characteristics of the firms, while the role of external knowledge and local context in shaping innovation persistence is almost neglected. In this respect our paper adds to previous literature as it provides a first attempt to take into account external factors in determining innovation persistence. Building on previous analyses that showed how successful innovative activity is more likely to happen within strong industrial regions (Baptista and Swann, 1999; Swann, Prevezer and Stout, 1998), we claim that the access condition to the stock of knowledge of the other agents in the system are likely to play a major role in assessing innovation persistence. The persistence of innovation is then determined by the twin effects of knowledge cumulability internal to firms and external to firms but internal to their localized context of action. The access to the stock of knowledge external of each other firm and the actual amount of knowledge externalities that qualify the regional and industrial context of action of each firm are a necessary condition for the actual introduction of technological innovations. At the same time, however, knowledge externalities provided by the localized context of action, keep changing over time, albeit with a slow pace. The architecture of interactions and transactions that are the carriers of knowledge externalities change gradually over time as a result of the growth performances of firms, their entry, decline and exit and ultimately the introduction of innovations. Externalities and specifically knowledge externalities are indeed external to each firm, but internal and endogenous to each system (Antonelli and Scellato, 2012).

Furthermore, as evidence on persistence has shown to be in part dependent on the specific innovation activity scrutinized, in order to provide additional empirical indication on innovation persistence we will use total factor productivity growth in order to obtain a general measure of the extent to which innovation is persistent at the firm level. The empirical tests will develop and elaborate the Transition Probability Matrix (TPM) methodology implemented by many authors such as Cefis and Orsenigo (2003), Cefis (2005), Peters (2009), David and Rullani (2008) and Antonelli et al. (2012). In particular, we propose an approach that consists in observing different TPMs for specific sub-time periods, within a longer time interval. This type of analysis enables to identify the changes of the transition probabilities and to interpret them as clues of the effects of the small –external- events that affect the persistence.

3. HYPOTHESES AND RESEARCH DESIGN

The generation of technological knowledge is an activity characterized by significant indivisibility and learning. Knowledge indivisibility and learning to learn exerts strong cumulative effects. The generation of new knowledge and the introduction of innovations is the result of the creation, within corporations, of new functional routines and of research and development laboratories and of the structure of the communication networks that qualify the access to the external knowledge. The generation of new knowledge and the related introduction of innovation are shaped by the twin and joint effect of internal cumulative forces, and external positive feedbacks exerted by the system into which firms are embedded.

As a consequence we retain the hypothesis that innovation is a path dependent -as distinct from a past dependent - process determined by a number of internal and external factors. External factors are characterized by high levels of contingency, as such their changes affect the dynamics of the persistence. Following the resource based theory of the firms we suppose that the following factors matter:

A) The size of firms. The generation of technological knowledge is characterized by substantial sunk costs. Corporations that have innovated once are more likely to keep innovating simply because the incremental costs of the internal facilities designed to generate new technological knowledge and introduce innovations are low (Penrose, 1959; Arrow, 1974; Conner and Prahalad, 1996).

B) The level of wages. The well-known dynamics of the Matthew effect is likely to apply not only to scientists but also to firms for at least two classes of reasons. First, it seems plausible that innovating firms are able to pay higher wages and hence to attract more creative and talented employees. Second, innovating firms are likely to interact with innovative suppliers and innovative customers and, hence, to feed more fertile and productive user-producers interactions. The repeated interaction between the accumulation of knowledge, the creation of routines to valorize and exploit it eventually leads to the creation of dynamic capabilities that favor the systematic reliance upon innovation as a competitive tool (Stiglitz, 1987; Teece and Pisano, 1994; Langlois and Foss, 1999).

C) Price-cost margins. The effects of price cost margins on the persistence of innovation are twofold. On the one hand large price-cost margins should provide access to internal funds and favor the innovative efforts of firms. Hence the effect should be positive. On the other hand, however, large price-cost margins are a clear indicator of barriers to entry and market power. Firms that enjoy market power have lower incentives to persist in funding innovation activities. Hence the effects should be negative, especially when the levels of price-cost margins are very high (Aghion et al., 2005; Antonelli and Scellato, 2011).

D) The investment in intangible capital. The intangible assets intensity captures firms' effort for building innovative competences. R&D expenditures are the traditional indicator used to measure the amount of internal efforts to generate new technological knowledge. However, R&D statistics measure only a partial amount of the overall effort that firms make to introduce new technologies. Accountancy rules provide suitable evidence upon stocks of intangible capital that include capitalised research expenditures as well as purchasing costs for patents and licenses and the costs incurred to build and implement brand and know how (Teece, Pisano, Shuen, 1997).

Next to the internal factors to which the literature on innovation persistence has paid much attention, we argue that external factors play a crucial role. External factors are also contingent as the structure of the system into which external knowledge and rivalry take place keep changing as a consequence of the introduction of innovations. At each point in time the networks of interactions and the types of transactions on factor and product markets change. Yet at each point in time the architecture of the system and the market forms exert a strong effect on the capability of firms to access and use external knowledge and to rely on it for the introduction of further innovations as a competitive tool. As we expect that innovation is a persistent process that takes place when knowledge externalities and external-local feedbacks play a positive role, we introduce, next to the internal factors considered so far, two external factors:

E) The access to the local pools of knowledge stocks generated by the spillover of the innovative activity of other firms co-localized in the proximity of each firm provides a key contribution to the persistence of innovative activities. Such effects are typically inter-industrial: knowledge generated in an industry may be useful in other activities (Jacobs, 1969). Hence, we expect that the levels of total factor productivity of firms co-localized in the same region, irrespective of the industrial sector, favor the persistence of innovation. The higher the levels of total factor productivity of all the firms that are co-localized and the higher we expect to be the innovation persistence.

F) The levels of the innovative activity of firms within the same industry and hence active in the same product markets, measure the extent to which the typical Schumpeterian rivalry based upon the introduction of innovation is at work. The higher are the levels of total factor productivity of rival firms and the stronger is the competitive pressure. The Schumpeterian rivalry pushes firms to innovate in order to survive. Hence we expect that the higher is the efficiency of the rivals within the same industry and the larger is the likelihood that each firms relies upon the introduction of innovation as a competitive tool and hence the stronger is the persistence of innovation (Aghion et al., 2005). These hypotheses are consistent with the model elaborated by Gruber (1992) about the role of sequential product innovations in maintaining the leadership in markets characterized by vertical differentiation.

External factors add to internal ones and shape the context into which the persistence of innovation takes place. The external conditions, namely the quality of local pools of knowledge stocks and the strength of the Schumpeterian rivalry, together with the internal conditions, that is the actual levels of dynamic capabilities, as proxied by the levels of wages and the size of firms, exert a specific and localized effect upon the persistent introduction of innovations. Because externalities are internal to the local system into which firms are embedded the changing conditions exert a path dependent effect upon the sequence of innovations.

In order to study the persistence of innovation we rely upon a classic indicator such as the total factor productivity. We assume in fact that innovation has much a broader scope than indicators such as the generation and introduction of science-based new technologies that patent statistics tend to emphasize, or the specific introduction of new products and processes, measured by innovation counts.

Innovation consists, more generally, in the systematic capability to generate new knowledge and to apply it to the broad array of activities that firms carry on. So far our notion of innovation is much broader and retains a strong Schumpeterian flavor as it includes the introduction of new products and new processes as well as the introduction of changes in the organization, in the mix of inputs and in the product and factor markets into which firms operate. Hence we assume that total factor productivity is better able to capture the general increase in the efficiency of the firm that is engendered by the command of technological, organizational and commercial knowledge.

Clearly our hypothesis here is that the probability to introduce an innovation at time t+1 is conditional both to the introduction of an innovation at time t and to the effects of contingent forces that exert their effect locally so as to affect the sequence of state dependency.

Our two hypotheses lead to a two-step research design that can be summarized as it follows. In a first step we focus the analysis upon the identification of the persistence of the innovative activity as measured by TPM (Transition Probability Matrix) computed using variations in the levels of total factor productivity. Within the period of time considered we explore the possibility that relevant external factors may affect the transition probabilities. We introduce here the Multiple Transition Probability Matrixes (MTPMs) approach that consists in the analysis of sub-periods TPMs in order to test whether transition probabilities change within the stretch of time considered.
The MTPMs involves the computation of a single Markov chain for the full period of time and the comparison of its results with the computation of different Markov chains in relevant sub-periods, identified by significant contingent events that are expected to affect the transition probabilities between the innovative and non-innovative status of the analyzed

companies. We suggest that this approach based upon the comparison of the parameters of the Markov chains in different sub-periods should allow a better identification of the path dependent character of the innovation process. In particular, the observation of different parameters for the Markov chains in different sub-periods might be an indication of the fact that the extent of innovation persistence is indeed affected by contingent events and, hence, innovation can be qualified as a path-dependent process.

In the second step we concentrate the analysis upon the determinants of innovation persistence as we want to qualify the type of persistence at work, as well as the role of non-observable heterogeneity. Our main argument here is that a number of contingent and localized conditions, both internal and external to each firm, exert a significant effect upon the persistence. The persistence of the innovative activity, hence, is path dependent and it is not past dependent. The letter takes place when the dynamics of the process, both in terms of rates and direction, is determined at the onset. Path dependence, instead, takes place when contingent factors that arise through time exert a clear and persistent effect on the direction and the rate of the process (David, 1997 and 2007).

4. THE EMPIRICAL EVIDENCE

4.1. THE DATA

Our analysis is based on an original dataset containing accounting data for a sample of Italian manufacturing firms. The dataset includes financial accounting data for a large sample of manufacturing companies, observed along years 1996-2005. The data have been extracted from the AIDA database provided by Bureau Van Dick, which reports accounting information for public and private Italian firms with a turnover larger than 0.5 millions of Euros. The companies included in the analysis have been founded before year 1995, they are registered in a manufacturing sector according to the Italian ATECO classification, and they are still active by the end of year 2005. The introduction of the latter condition implies that we do not consider market exit/entry.

We have included all the companies with at least 15 employees at the end of fiscal year 1995. In order to drop outliers due to possible errors in the data source, we computed a set of financial ratios and yearly growth rates of employees, sales and fixed capital stock. The final panel is composed of 7020 companies. All financial data have been deflated according to a sectoral two-digit deflator using year 2000 basic prices. In annex 1 we report the sectoral composition of the dataset.

4.2 TOTAL FACTOR PRODUCTIVITY AS A MEASURE OF INNOVATIVITY

We investigate the persistence in innovation activity, as measured by firm level total factor productivity TFP. The rates of increase of TFP are good measures of the degree of innovativeness of the firms. This is especially true with respect to the Italian system where, although the levels of formalized R&D activities and patenting are low, much innovation is based upon informal research activities, tacit knowledge and learning. Hence, we assume that the bottom line increase of efficiency at the firm level is the ultimate indicator of the wide array of interrelated effects of the introduction of changes in products, processes, markets, organization and inputs.

In order to compute firm-level TFP we have firstly estimated a set of Cobb-Douglas production functions with constant returns to scale for each industry included in the sample, so to obtain the correct levels of output elasticity of labor and capital. After the assignment of each firm to an industry we have computed TFP for company i in year t according to the following expression:

$$TFP_{i,t} = \frac{Q_{i,t}}{L_{i,t}^{\beta} K_{i,t}^{1-\beta}} \qquad (1)$$

Where:

$Q_{i,t}$:deflated value added

$L_{i,t}$:number of employees

$K_{i,t}$:fixed capital stock.

Fixed capital stock has been computed using a perpetual inventory technique according to which the first year accounting data, i.e. year 1996 in our case, are used as actual replacement values. The subsequent yearly values of fixed capital are computed using a depreciation parameter δ, assumed equal to 6.5%, and adding deflated yearly investments. The level of yearly depreciation of physical capital has been chosen following the approach applied in previous studies that have applied perpetual inventory techniques to estimate yearly fixed capital levels adopting depreciation parameters in the range 5%-10% for physical capital. On this issue see Olley and Pakes (1996) and Parisi et al. (2006) for the Italian economy. Since the adopted depreciation parameter is constant across industries we should not expected changes in the significance of estimate coefficients for slight changes in δ. The investment parameter ($I_{i,t,}$) has been computed as the yearly variation in net fixed capital in companies' balance sheets plus yearly amortizations. Hence, the time series of fixed capital is defined as follows:

$$K_{i,t} = (1 - \delta)K_{i,t-1} + I_{i,t} / p_t \tag{2}$$

In order to identify the parameter β at industry level to compute equation 2, we have estimated for each industry the following equation:

$$Log\left(\frac{Q_{i,t}}{K_{i,t}}\right) = \beta \times Log\,\frac{L_{i,t}}{K_{i,t}} + \alpha_i + \alpha_t + \varepsilon_{i,t} \tag{3}$$

We have used a fixed effect estimator, where α_i is a firm specific effect and α_t is a time specific effect[1].

The following Table 2 provides summary statistics about the variables that will be used in our analyses.

[INSERT TABLE 2]

4.3 THE RESERCH STRATEGY TO TEST PATH DEPENDENT INNOVATION PERSISTENCE

Consistently with the theoretical discussion, in our modeling framework we follow two complementary approaches. In the first part of the analysis, we investigate the presence of firm-level persistence by means of different transition probability matrixes (MTPMs). In the second part, we explore firm-level innovation persistence by means of discrete dynamic panel data models. Below we discuss the methodological details of such complementary approaches.

4.3.1 The analysis of Multiple Transition Probability Matrixes

Our analysis aims at identifying the path dependent property that characterizes the innovation process. Following an established literature we rely on transition matrices that have been frequently used to test the hypothesis that history matters in innovation processes (Peters, 2008).

However, previous contributions relied on regular transition matrixes[2],

[1] For a discussion of the properties of different estimation approaches see Blundell and Bond (2000) and Olley and Pakes (1996).

which imply that the processes under analysis are ergodic. In contrast, in order to deal with the non ergodic character of innovation persistence we will refer to non-homogeneous Markov chains, that allow us to model time dependent transition probabilities.
Below we try and clarify this point and we discuss how it affects our empirical approach.

The parameters of transition matrixes can be interpreted as the empirical estimation of an underlying Markov process. More specifically, Markov chains are dynamic stochastic processes characterized by the presence of discrete values of the states and, more importantly, by the fact that the conditional probability of a state at time t depends exclusively on the state at time t-1. This implies that the process has no memory and only the last state influences the subsequent state. Technically this amounts to the following definition of state probability along time:

for $t_{k+1} > t_k > t_{k-1} > ...$

$$\Pr\left[X(t_{k+1}) = x_{t_{k+1}} \mid X(t_k) = x_{t_k}, X(t_{k-1}) = x_{t_{k-1}}, ...\right] =$$
$$= \Pr\left[X(t_{k+1}) = x_{t_{k+1}} \mid X(t_k) = x_{t_k}\right] \qquad (1)$$

Given the above property, all the statistical features of a stationary Markov process can be determined from the conditional densities between two subsequent periods t_k and t_{k+1}.

$$f(x_{t_{k+1}}, x_{t_k}) = f(x_{t_{k+1}} \mid x_{t_k}) * f(x_{t_k}) \qquad (2)$$

A Markov process is homogeneous if the conditional density (what we estimate with the TPM) is time invariant while the first order density can vary in time.
If we observe a process that can be described by a TPM and such matrix is regular then in the context of Markov chains we are assuming an underlying stationary ergodic process. In fact, if the underlying process is non-ergodic, it cannot be properly captured by an homogeneous transition matrix.

Suppose that we observe a process in which each company can be at each time in one of just two alternative states: innovative and non innovative. We then compute - by pooling observations along time - the probability of being in state i at time t conditional on being in state j at time t-1, i.e. we compute a 2X2 TPM. According to the previous considerations the observation of such TPM cannot not directly provide evidence on the path

[2] A regular TPM is an irreducible matrix with at least one of the diagonal elements different from zero.

dependent properties of the system based on our definition of path dependency.

The use of regular Markov chains allows us just to state that prior conditions affect future events. Hence, we can say that "history matters" because the innovative status at time t is not randomly distributed in the population of firms. However, in this setting all the past information is incorporated just in the state at time t-1. Moreover, the innovation persistence that stems from a time invariant conditional probability of states is intrinsically ergodic and it is fully consistent with the hypothesis that innovation persistence is the result of a special quality, a talent, embodied in the firm that qualifies it as a part of its intrinsic endowment. Such an innovation persistence is fully consistent with the resource based theory of the firm where learning internal processes display long-term and stable effects.

Results of previous contributions that have made an implicit use of homogeneous Markov chains (through the computation of a single period TPM) confirm that the innovation process is a persistent process characterized by a twin character: 1) the (non) introduction of an innovation at time t affects the probability of (non) introducing an innovation at time t+1; 2) yet the estimated structure of probabilities is time-invariant. According to our hypotheses the innovation process, instead, is characterized by path dependence. Path dependence takes place when contingent events bear a dynamic effect where the past affects the future with changing transition probabilities. Path dependent innovation persistence takes place when future events are affected by present ones with changing weights as the contingent events that take place at time t, do change the transition probability distribution, as well. This implies that what we observe by a single period TPM calculated over a sufficiently long time window is indeed the averaged result of different dynamics in sub-periods.

In the context of economic analysis we might stretch this consideration to state that for a given change in conditions of the system, external to each firm, but internal to the system in sub-period 1 (e.g. a contraction in credit supply, an increase in aggregated demand, the emergence of new technological opportunities, a change in the provision of knowledge externalities) the reaction of companies would be captured by the TPM in sub-period 1 and would also affect the TPMs in following sub-periods. A single period TPM could not grasp these changes. Our approach instead enables to identify and appreciate the differences between the results of sub-period Markov chains. The observation of different processes in properly defined sub-periods can be interpreted not just as the presence of "transitory" phases towards the long run stationary process, but as an evidence of path dependence.

4.3.2 The analysis of persistence through panel data

While the TPM approach is expected to provide only summary evidence on the path dependent persistence of firm level TFP levels along time, and a clue about the effects exerted by the changing characteristics of the system, the panel data analysis aims at identifying the impact of contingent factors on the persistence of innovation.

In order to analyze the persistence of innovation along time we have constructed a time varying dummy variable ($INNO_t$) that equals one if a company has experienced a positive TFP growth rate over a two year period, between year t-2 and year t. We then apply different dynamic discrete choice models in which such variable is regressed against its past realization and a set of appropriate controls. In particular, we test the relationship between the innovation dummy and both internal and external factors. The former group includes a variable of firm size measured as the log of firms' total asset (SIZE), an indicator of the level of human capital as captured by the average wage (WAGE), the price-cost-margin as indicator of firms' profitability (PCM) and an indicator of the incidence of intangible assets (INTANG) defined as the ratio of intangible to tangible assets in a specific year.

The second group of regressors accounts for changes along time in sectoral technological opportunities and for regional conditions. As previously highlighted, we claim that firms' capability to introduce technological innovations can be affected by the specific conditions of the local economic environment. For this reason, as controls for external conditions we include in the model specification a variable (REG_TFP) that for each company i equals the yearly average level of the TFP of all the other companies (included in our sample) and located in the same region of company i. This regressor is expected to capture general regional conditions potentially affecting productivity levels through time, such as the presence of knowledge intensive infrastructure, the local development of financial institutions or specific characteristics in the input markets.

Clearly, changes along time in firm-level TFP are likely to be affected also by non-geographically defined external factors. In order to account for sectoral dynamics of TFP we include in the model the variable SECT_FTP that for each company i equals the yearly average level of the TFP of all the other companies (included in our sample) belonging to the same 2-digit ATECO classification of company i. Since the innovation dummy variable is defined over a two-year period, we have inserted the above mentioned controls with a lag.

As previously highlighted, observed persistence may be due to true state dependence or permanent unobserved heterogeneity across the analysed companies. By a theoretical perspective, if the source of persistence is due to permanent unobserved heterogeneity, individuals show higher propensity to take a decision, but there is no effect of previous choices on current utility and past experience has no behavioural effect (Heckman, 1981).

In our specific context, we can assume that expected drivers of true state persistence include the existence of dynamic increasing return to innovation effort, determined by the sunk R&D costs previously incurred by a company, and the internal cumulativity of the innovation process. On the other side, the source of unobserved serially correlated characteristics that make firms more or less likely to innovate relate to risk attitude of entrepreneurs and other idiosyncratic features. By controlling for a set of observable firm specific dimensions we expect to obtain a clearer view of the contribution of the different potential sources of the observed innovation persistence.

The baseline specification for a dynamic discrete response model is the following, where y_{it} is our innovation indicator:

$$y_{it}^{*} = \gamma y_{it-1} + \beta x_{it} + u_i + \varepsilon_{it} \tag{3}$$

The estimation of the above model requires an important assumption on the initial observations y_{i0} and their relationship with u_i, the unobserved individual effects. In fact, if the start of the analysed process does not coincide with the start of the available observations, y_{i0} cannot be treated as exogenous and its correlation with the error term would give raise to biased estimates of the autoregressive parameter g, which represents our measure of persistence. Two different approaches can be adopted for handling such initial condition problem: Heckman (1981) suggests to specifying the distribution of y_{i0} conditional on u_i and x_i; alternatively, Wooldridge (2005) proposes to specify the distribution of u_i conditional on y_{i0} and x_i.

For sake of robustness in our analysis we have then applied both the methodologies.

The approach by Heckman (1981) adopts a linearized approximation of the reduced form equation for the initial value (t=0) of the latent variable as follows:

$$y_{i0}^{*} = z_{i0}\pi + \eta_i \quad (4)$$

where z_{i0} is a vector of exogenous instruments and includes xi0. The underlying assumption of such specification is that η_i is correlated with ui (see eq. 5) but uncorrelated with ε_{it} for any t>0. Given a $\vartheta > 0$ we can then write the following relation:

$$\eta_i = \vartheta u_i + \varepsilon_{i0} \quad (5)$$

and

$$y_{i0}^{*} = z_{i0}\pi + \vartheta u_i + \varepsilon_{i0} \quad (6)$$

Given the specification of the initial observation (eq. 4), it is then possible to use the joint probability of the observed binary sequence (t=0,…t=T) with maximum likelihood for the estimation of the dynamic model. Stewart (2007) provides an application of this estimator[3]. In our case we have adopted as instruments in equation (6) firm-level pre-sample variables.

Concerning the Wooldridge modelling approach, we follow the methodology applied by Peters (2009) that offers a simplification of the Wooldridge method, by using the first realisation of the innovation indicators (y_{i0}) and the time-averaged covariates as predictors of the individual effect, according to the following relationship:

$$u_i = \alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + c_i \quad (7)$$

Where:

$$\bar{x}_i = T^{-1} \sum_{t=1}^{T} x_{it} \quad (8)$$

Under the assumption that the error term c_i is distributed as $N(0,\sigma_c^2)$ and that $c_i \perp (y_{i0}, \bar{x}_i)$ we obtain that:

$$u_i \mid y_{i0}, \bar{x}_i \approx N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i \, , \sigma_c^2] \quad (9)$$

Hence the dynamic probit model can be rewritten according to the following specification:

$$P(y_{it} = 1 \mid y_{i0}, \ldots, y_{it-1}, x_i, \bar{x}_i, c_i) = \phi(\gamma y_{it-1} + \beta x_{it} + \alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + c_i) \quad (12)$$

[3] The model has been estimated with the STATA routine redprob, developed by Stewart (2007). For more details see http://www2.warwick.ac.uk/fac/soc/economics/staff/faculty/stewart/stata

This second methodology in principles has the advantage of being less restrictive on exogeneity assumptions with respect to the Heckman's one. By a technical point of view the Wooldridge (2005) method amounts to estimating a dynamic random effect probit model in which regressors include a dummy representing the initial realisation of the dependent variable (variable INITIAL in our models) and the time average of those covariates that are expected to be correlated to the individual effect (in our model AVGSIZE, AVGWAGE, AVGPCM, AVGINTANG).

4.4 RESULTS

4.4.1 Evidence from the MTPM analyses

The following Table 3 provides the results for TPMs obtained on the full sample of companies observed in the all period (1998-2005) and in the two sub-periods before and after the year 2001. This year has been chosen as it identifies a major contingent event such as the turning point in the economic cycle into which firms are observed. Moreover it has the advantage to be in the middle of the panel so that we avoid problems of comparability related to strong differences in samples' dimension. Note that the balanced nature of our firm- level dataset avoids possible drawbacks of the TPM analysis.

For each element of transition probability matrixes we have also computed standard errors, adopting the following approach. Let P_{ij} and $\hat{P}_{ij}$ denote the population and sample probabilities of a transition of a company from the status i to the status j. This transition process can also be seen as the outcome of a binomial distribution. Hence, standard errors of the estimated transition probabilities can be calculated as a binomial standard deviation: $\sqrt{P_{ij} * (1-P_{ij})/N}$ where N equals the number of companies in status i. As N increases $\hat{P}_{ij}$ tends to P_{ij} . In the matrixes that will be presented in our analysis clearly the binomial process has just two possible outcomes. Hence the estimated standard error is the same for the elements of each row in the 2X2 matrix.

Our calculations show the presence of strong innovation persistence as both the main diagonal elements of the transition matrix referring to the whole period are greater than 0.5. However, persistence patterns are found to be different in the two sub-periods. Data show that in the second interval the percentage of persistent innovators increases from 45.53 to 66.95. The transition probability from a negative to a positive status rises as well (from 27.42 to 35.6). The analysis of the MTPMs unfolds interesting evidence. There is a remarkable difference among the results of the three TPMs. This difference confirms that contingent events modify the distribution of transition probabilities and yet each of them is statistically significant.

This evidence is quite relevant from both a methodological and a historic viewpoint. From the methodological viewpoint the high levels of statistical significance of all the matrices confirm that the in the innovation process events at time t bear a strong effect on the events at time t+1. The introduction of an innovation at time t does affect the introduction of an innovation at time t+1. At the same time, however, the contingent effects that have been taking place along the period have significantly changed the 'weight' with which the (non) introduction of an innovation at time t has affected the (non) introduction of an innovation at time t+1.

From the historic viewpoint, this evidence can be easily framed in a Schumpeterian perspective. In the final part of the upturn phase of the economic cycle which started in Italy in the second half of 1990s, increasing demand allowed also non innovative and less efficient firms to survive in the market and the incentives to innovate tend to reduce; on the contrary after the turning point competitive pressure increases, profit margins are lower and the incentives to innovate tend to increase in order for firms to survive. In parallel, a more specific interpretation of the evidence can be linked to the diffusion of the ICTs and its economic consequences in Italy, where the rate of penetration of these technologies registered a lagged dynamics with respect to USA and other advanced EU countries. This triggered a process of restructuring of production in the early 2000s that might have affected the pace of innovation as measured by positive increments of TFP. Moreover, as evidenced in previous literature (Quatraro, 2009; 2012), the transformation process related to the diffusion of ICT has been uneven in the Italian territory. For this reason we have split further the sample in 4 macro-regions so as to identify potential differences in the patterns of innovation persistence across regions and across time. Results indicate that the hierarchy in terms of percentage of persistent innovators is completely inverted in the two sub periods with the north-west part of Italy showing the greatest presence of persistent innovators after 2001 and the south part of Italy the lowest. This evidence suggests the presence of a divide. While in the north part of Italy and in particular in north-west the penetration of ICT activated a virtuous process of transformation of the economy that led the majority of firms to rely on the continuous introduction of innovation as a competitive strategy, in the south-part there is no evidence of change in the patterns of persistence as indicating that the transformation process occurred in other parts of Italy has been (at least) less relevant in this case.

This result is important mainly for three reasons. First, it provides original evidence supporting the idea that in the period of observation an uneven process of transformation of the Italian economy has occurred. Second it gives support to the hypothesis that the local contexts can relevantly contribute to shape the patterns of innovation persistence. Third, the evidence suggests that the path and past dependent characters of innovation persistence can be of different relevance across time and location. In the examined case, in the North-western part of Italy, the path dependent character of innovation persistence appears to be dominant while the contrary is true in the Southern regions.

[INSERT TABLE 3]

The data seem to provide initial evidence of significant persistence in innovation, as captured by positive growth rates of TFP. However, we claim that it is important to stress how the above results, although suggesting the presence of some form of inter-temporal stability in innovation effort, do not provide, yet, a sound answer to two key question: how much of the observed persistence can be labeled as true persistence driven only by previous innovation? Moreover, when internal factors are included in the analysis to what extent the observed persistence is still influenced by external factors? In the next section we introduce an econometric analysis specifically devoted to assessing these two points.

4.4.2 Results from dynamic panel data analyses

In the following Table 4 we report our results for different specifications of the persistence model estimated with the Wooldridge dynamic probit approach. The results stress that, even after controlling for a number of internal and external factors, the probability of observing an innovation at time t is positively and significantly affected by the previous realization of the INNO variable.

[INSERT TABLE 4]

It is worth clarifying that the result of the econometric estimates tests the role of a number of controlling factors upon the chances of observing a positive growth rate of TFP, rather than upon the probability that innovators keep introducing innovations along time. In this sense, we obtain that the fact of being located in a region characterized by higher levels of TFP of surrounding firms is positively associated with the probability of introducing some form of innovation. In the following Table 4 we report the results obtained for the model specification based on the Heckman (1981) approach. Also in this case we find a positive and significant correlation along time in the realizations of the innovation variable. The significance of the other variables is most important as it confirms the path dependent character of the process. Among the internal factors the levels of human capital, as measured by average unit wage, significantly enhance the probability of subsequent innovation outcomes. The effects of size enters the model specification through two covariates (Table 4): AVGSIZE and SIZE. The former is time-invariant. The latter is the yearly measure. Our results suggest that the AVGSIZE, i.e. the dimensional class to which each firms belongs has a negative effect. This result is perfectly aligned with the expectations based upon the Gibrat law. The results suggest, instead, that SIZE, i.e. the time varying dimension of the firm, has a positive effect. As expected, the intensity of intangible capital, which is a proxy for the investment along time in research and development and innovation activities, exerts a positive significant impact.

[INSERT TABLE 5]

In both models (Table 4 and Table 5) the local context exerts a strong and positive role upon the persistence of innovation as measured by the levels of TFP of firms co-localized in the proximity within the same region. As expected, the access to the local pools of knowledge and the pecuniary knowledge externalities generated by the regional agglomeration of innovative firms favor the persistence of innovative activities. The intensity of innovation of the firms active in the same industries also favors the persistence of innovation. The stronger is the typical Schumpeterian rivalry among firms that rely upon the introduction of innovations as a competitive tool and the stronger is the persistence of innovation.

Our results confirm the persistence of total factor productivity growth and suggest that such persistence is affected by contingent factors that are both internal and external to each firm. The results can be interpreted as a test of the claim that the persistence is path rather than past dependent. Contingent factors, such as human capital, market rivalry and geographic location would not be significant when the persistence is past dependent because the original conditions would play an exhaustive causal role.

5. CONCLUSIONS

Knowledge cumulability stemming from knowledge indivisibility and knowledge non-exhaustibility plays a central role in path dependent innovation persistence. The introduction of further innovations is easier for firms that can command a larger stock of internal knowledge and have access to larger pools of knowledge stocks of co-localized firms. Much attention has been given to the exploration of internal factors that are at the origin of innovation persistence. This paper provides empirical evidence upon the central role of external factors in determining the path dependent persistence of innovation activities, as measured by total factor productivity levels (TFP).

In particular, the paper makes three contributions to the economics of innovation persistence. First, it provides an interpretative framework based upon the economics of knowledge that privileges the role of knowledge externalities. Second it distinguishes between types of innovation persistence. Past dependent innovation persistence is the result of a given allocation of a specific innovative capability or talent that keeps exerting its effects along time with no changes. Past dependent innovation persistence is consistent with the predictions of the resource based theory of the firm. Path dependent innovation persistence, instead, is the result of systemic interaction. Firms caught in out-of-equilibrium conditions in factor and product markets try and react by means of the introduction of innovations. Their reaction is actually successful, so as to lead to the introduction of productivity enhancing innovations, only when a set of external conditions are met. Such conditions keep changing over time and affect the likelihood that the introduction of an innovation at time t affects the likelihood that an innovation at time t+1 is also introduced.

Third, the paper discusses the methodological implications of the use of a MTPM approach based on the analysis of sub-periods to assess innovation persistence, with specific reference to the Markov chains theory. In particular, we suggest that the comparison of the parameters of different Markov chains across a given stretch of time enables to assess empirically whether contingent events have exerted significant effects on persistency patterns. In this case path dependent innovation persistence applies because the relationship between past and future is altered by the events that take place at time t. Finally, building upon these results, we have investigated the firm-level innovation persistence patterns using dynamic panel methods. The econometric results confirm that the persistence of innovation is affected by contingent and localized events, among which the access conditions to the stock of knowledge of the co-localized agents play a central role. At each point in time the probability of introduction of further innovations is indeed affected by the sequence of innovations introduced in the past but it is also conditional to the actual levels of internal dynamic capabilities of each firm to accumulate and exploit technological knowledge and human capital, the amount of external knowledge that is available in the regional proximities, and the competitive pressure of innovative rivals active in the same product markets.

Innovation persistence exhibits the characters of path dependence because of the effects of contingent factors that emerge through the process and yet are able to alter its dynamics. Contingent and endogenous changes concern typically the provision and the access conditions to knowledge externalities that exhibit changing effects through time. Knowledge externalities are possible only if and when effective communication channels based upon networks of interactions and transactions are available. The architecture of such networks however changes over time because of the conduct of firms and the introduction of innovations. Externalities are external to each firm, but internal to the system.

In terms of policy implications, it seems important to stress that the localized path dependent character of innovation persistence calls for a systematic and systemic approach to technology policy. In the case of 'true' state dependence we would assume that once a firm has been induced to innovate the likelihood that it will keep innovating is enhanced. Hence policy interventions would be redundant. The identification of the central role of external factors in assessing the path dependence of innovation, on the opposite, confirms the need of a national innovation policy to reinforce the internal cumulability of technological knowledge within firms and stresses the scope of action of the design and implementation at the regional level of public interventions devoted to upgrade the architectures of the networks of interactions and transactions that can implement the localized provision of knowledge externalities. The implementation of a twin innovation policy articulated in a national level aimed at firms, especially in technologies where knowledge cumulability is high, and a regional level aimed at strengthening the provision of knowledge externalities is crucial in sustaining the continuous introduction of innovation at the system level.

REFERENCES

Acs, Z.J., Anselin, L. and Varga, A. 2002. Patents and innovation counts as measures of regional production of new knowledge, *Research Policy* 31, 1069-1085.

Aghion, P., Bloom, N., Blundell, R., Griffith, R. and Howitt, P. 2005. Competition and innovation An inverted U relationship, *Quarterly Journal of Economics* 120, 701-28.

Alfranca, O., Rama, R., von Tunzelmann, N. 2002. A patent analysis of global food and beverage firms: The persistence of innovation, *Agribusiness* 18, 349 – 368.

Antonelli, C. 2008. *Localized technological change: Towards the economics of complexity*, Routledge, London.

Antonelli C. (Ed.) 2011. *Handbook on the economic complexity of technological change*, Edward Elgar, Cheltenham.

Antonelli, Scellato, G. 2011. Out of equilibrium, profits and innovation, *Economics of Innovation and New Technology* 20, 405-421.

Antonelli, C., Scellato, G. 2012. Complexity and innovation: Knowledge interactions and firm level total factor productivity, *Journal of Evolutionary Economics*, forthcoming

Antonelli C., Crespi F., Scellato G. 2012. Inside innovation persistence: New evidence from Italian micro-data, *Structural Change and Economic Dynamics* (in press DOI: 10.1016/j.strueco.2012.03.002).

Arellano, M., Bond, S. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies* 58, 277–97.

Arrow, K.J. 1974. *The limits of organization*, W.W. Norton, New York.

Baptista, R. Swann, G.M.P. 1998. Do firms in cluster innovate more?, *Research Policy* 27 (5), 527-542.

Baptista, R. Swann, G.M.P. 1999. The dynamics of firm growth and entry in industrial clusters: A comparison of the US and UK computer industries, *Journal of Evolutionary Economics*, 9 (3), 373-399.

Beaudry, C., Swann, G.M.P. 2009. Firm growth in industrial clusters of the United Kingdom, *Small Business Economics*, 32(4), 409-424.

Blundell R., Bond, S. 1998. Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics* 87, 115-143.

Cefis, E. 2003. Is there persistence in innovative activities? *International Journal of Industrial Organization* 21, 489-515.

Cefis, E., Orsenigo, L. 2001. The persistence of innovative activities. A cross-countries and cross-sectors comparative analysis, *Research Policy* 30, 1139-1158.

Cefis, E., Ciccarelli, M. 2005. Profit differentials and innovation, *Economics of Innovation and New Technology* 14, 43-61.

Clausen T., Pohjola M., Sapprasert K., Verspagen B. 2011. Innovation strategies as a source of persistent innovation *Industrial and Corporate Change,* (in press doi:10.1093/icc/dtr051).

Colombelli, A., von Tunzelmann N. 2011. *Persistence of innovation and path dependence*. In C. Antonelli (ed.) *Handbook on the economic complexity of technological change*. Edward Elgar, Cheltenham.

Conner, K. R., Prahalad, C.K., 1996. A resource based theory of the firm: Knowledge versus opportunism, *Organization Science* 7, 477-501.

Crespi F., Pianta M. 2007. Demand and innovation in European industries, *Economia Politica –Journal of Analytical and Institutional Economics-* n.1 Aprile, pp. 79-112.

David, P. A. 1985. Clio and the economics of QWERTY, *American Economic Review* 75, 332-37.

David, P.A. 1994. Positive feedbacks and research productivity in science: Reopening another black box, in Granstrand, O. (ed.), *The economics of technology*, Elsevier North Holland, Amsterdam.

David, P.A. 1997. Path dependence and the quest for historical economics: One more chorus of ballad of QWERTY, *Oxford University Economic and Social History Series 020, Economics Group*, Nuffield College, University of Oxford.

David, P. A. 2007. Path dependence: A foundational concept for historical social science, *Cliometrica: Journal of Historical Economics and Econometric History* 1, 91-114

David, P. A., Rullani, F. 2008. Dynamics of innovation in an "open source" collaboration environment: lurking, laboring, and launching FLOSS projects on SourceForge. Industrial and Corporate Change, 17, 647-710.

Davison, A. C.; Hinkley, D. 2006. *Bootstrap methods and their applications*, 8th, Cambridge Series in Statistical and Probabilistic Mathematics.

Dosi G. - Freeman C. - Nelson R. - Silverberg G. - Soete L. 1988, *Technical change and economic theory*, London, Pinter.

Duguet, E., 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, Université Panthéon-Sorbonne.

Fagerberg J., Mowery D., Nelson R. (Eds.) 2005. *The Oxford handbook of innovation,* Oxford, Oxford University Press.

Foster, L., Haltiwanger, J. and C.J. Krizan 2002. The link between aggregate and micro productivity growth: Evidence from retail trade, *NBER Working Paper*.

Fritsch, M. 2002. Measuring the quality of regional innovation systems: A knowledge production function approach, *International Review of Regional Science*, 25, 86-101.

Fritsch, M. 2004. Cooperation and the efficiency of regional R&D activities, *Cambridge Journal of Economics*, 28, 829-846.

Fritsch, M. and Franke, G. 2004. Innovation, regional knowledge spillovers and R&D cooperation, *Research Policy*, 33, 245-255.

Geroski, P. 1994. *Market structure corporate performance and innovative activity*, Oxford University Press, Oxford.

Griliches Z. 1979. Issues in assessing the contribution of research and development to productivity growth, *The Bell Journal of Economics*, 10, 92-116.

Gruber, H. 1992. Persistence of leadership in product innovation, *Journal of Industrial Economics* 40, 359-375.

Heckman, J.J. 1981. The incidental parameters problem and the problem of initial conditions in estimating a discrete time - discrete data stochastic process, in C.F. Manski and D. McFadden (eds*.). Structural analysis of discrete data with econometric applications*, MIT Press, Cambridge.

Huang, C.H. 2008. A note on the persistence of firms' innovation behavior: A dynamic random effect probit model approach, *Economics Bulletin*, 15(5), 1-9.

Jacobs, J. 1969. The economy of cities, Random House, New York.

Jang S., Chen J.H., 2011. What determines how long an innovative spell will last? *Scientometrics* 86, 65–76

Langlois, L.N., Foss, N.J. 1999. Capabilities and governance: The rebirth of production in the theory of organization 52(2), 201-18.

Latham, W.R., Le Bas, C. (eds.) 2006. *The economics of persistent innovation: An evolutionary view*, Springer, Berlin.

Malerba, F., Orsenigo, L., Petretto, P. 1997. Persistence of innovative activities sectoral patters of innovation and international technological specialization, *International Journal of Industrial Organization* 15, 801-826.

Martínez-Ros, E. and Labeaga, J. M. 2009. Product and process innovation: Persistence and complementarities, *European Management Review* 6, 64–75.
Nelson R., Winter S. 1982. *An evolutionary theory of economic change*, Cambridge (MA), The Belknap Press of Harvard University Press.
Olley S., Pakes, A. 1996. The dynamics of productivity in the telecommunications equipment industry, *Econometrica* 64, 1263–1297.
Parisi, M.L., Schiantarelli, F., Sembenelli, A. 2006. Productivity innovation and R&D: Microevidence for Italy, *European Economic Review* 50, 2037-2061.
Penrose, E., 1959. *The theory of the growth of the firm*, Oxford University Press, Oxford.
Peters, B. 2009. Persistence of innovation: Stylized facts and panel data evidence, *Journal of Technology Transfer* 36, 226-243.
Quatraro, F. 2009. Innovation structural change and productivity growth. Evidence from Italian regions 1980-2003, *Cambridge Journal of Economics*, 2009, 33, pp. 1001-1022.
Quatraro, F. 2012. *The economics of structural change in knowledge*, Routledge, London.
Raymond, W., Mohnen, P., Palm, F., Schim van der Loeff, S., 2010. Persistence of innovation in Dutch manufacturing: Is it spurious? *Review of Economics and Statistics* 92, 495–504.
Roodman, D. 2006. An introduction to “difference” and “system” GMM in Stata, Working Paper n. 103, Center for Global Development. www.cgdev.org
Roper, S., Hewitt-Dundas, N. 2008. Innovation persistence: Survey and case-study evidence, *Research Policy* 37, 149-162.
Scherer, F.M., Harhoff, D. 2000. Technology policy for a world of skew-distribution outcomes, *Research Policy* 29, 559–566.
Schumpeter, J. A. 1947. The creative response in economic history, *Journal of Economic History* 7,149-159.
Stewart, M.B. 2007. The inter-related dynamics of unemployment and low-wage employment, *Journal of Applied Econometrics*, 22, 511 - 531
Stiglitz, J.E. 1987. Learning to learn localized learning and technological progress, in Dasgupta, P. and Stoneman, P. (eds.), *Economic policy and technological performance*, Cambridge University Press, Cambridge.
Swann, G.M.P., Prevezer, M., Stout, D.K. 1998. *The dynamics of industrial clustering: International comparisons in computing and biotechnology*, Oxford University Press, Oxford.
Teece, D.J., Pisano, G. 1994. The dynamic capabilities of firms: An introduction, *Industrial and Corporate Change* 3, 537-555.
Teece, D.J., Pisano, G., Shuen, A. 1997. Dynamic capabilities and strategic management, *Strategic Management Journal* 18, 509-533.
Wooldridge, J. 2005. Simple solutions to the initial conditions. Problem in dynamic nonlinear panel data models with unobserved heterogeneity, *Journal of Applied Econometrics* 20, 39–54.

TABLES

TABLE 1 Summary of main contributions in the field of innovation persistence

Authors	**Data**	**Methodology**	**Results**
	PATENT DATA ANALYSES		
Malerba, Orsenigo and Peretto (1997)	Patent data from OTAF-SPRU data base for five EU countries (1969-1986)	Dynamic panel data model	The econometric evidence shows that the innovative activity is persistent.
Geroski, Van Reenen and Walters (1997)	Patent records and 'major' innovations of a sample of UK firms (1969-1988)	Proportional hazard function	Only a minority of firms (major innovators) is found to be persistently innovative.
Cefis and Orsenigo (2001)	Patent data on a sample of 1400 manufacturing firms (1978-1993) in Germany, Italy, Japan, US and France	Transition probability matrix	Evidence of weak persistency with both low-innovators and great-innovators generally remain in their classes
Cefis (2003)	Data on 577 UK patenting firms (1978-1991)	Transition probability matrix	Evidence o little persistence characterized by a strong threshold effect. Only great innovators have a stronger probability to keep innovating.
Cefis and Ciccarelli (2005)	Data on 267 UK patenting firms (1988-1992)	Bayesian econometric models	They show that current innovative activity can be positively influenced by past innovation via the greater availability of financial resources.
Alfranca, Rama and von Tunzelmann (2002)	Information on 16,698 patents granted in the United States from 1977 to 1994 to 103 global firms in the food and beverage industry.	Time series analysis	The evidence confirms that global firms in this industry exhibit a stable pattern of technological accumulation in which "success breeds success".
Latham and Le Bas (2006)	Patent data for 3347 French firms (1969-1985)	Duration econometric model	The persistence of innovation is stronger among individuals than among firms.
Huang (2008)	Patent and R&D data on 246 electronics firms listed on the Taiwan Stock Exchange (1998-2003)	Dynamic random effect probit model	Evidence supporting the existence of persistent innovation after controlling for firm heterogeneity.
Jang and Chen (2011)	Patent data on 125 publicly-listed IT firms in Taiwan (1990–2001)	Survival analysis	Evidence of state dependence but transient nature of the competitive advantage attributable to innovative persistence.
	SURVEY DATA ANALYSES		
Duguet and Monjon (2004)	Innovation and census data on 621 French firms operating in manufacturing sectors (1986-1996)	Propensity score matching models	Strong evidence of innovation persistence associated with size and formal R&D activities.

Roper and Hewitt-Dundas (2008)	Data on 3604 plants covered by the Irish Innovative Panel (1991-2002)	Transition probability matrix	Both product and process innovations are found to be strongly persistent.
Peters (2009)	Community Innovation Survey (CIS) data on German manufacturing and service firms (1994-2002)	Transition probability matrix and dynamic probit models	High levels of persistence in undertaking innovation activities.
Martínez-Ros and Labeaga (2009)	ESEE survey on Spanish manufacturing firms (1990-1999)	Random effect probit models	Evidence of persistence with relevant complementarities between product and process innovation.
Raymond et al. (2010)	Unbalanced panel of 2,764 enterprises from the Dutch Community Innovation Surveys (1994-2000).	Maximum likelihood dynamic tobit models	They find true persistence in the probability of innovating in high-tech industries and spurious persistence in the low-tech category.
Clausen et al. (2011)	Panel database constructed from R&D and Community Innovation Surveys in Norway	Dynamic random effects probit models	R&D intensive and science based companies are found to be more likely to be persistent innovators.
Le Bas et al. (2011)	Panel data on 287 firms from Luxembourg (CIS2006, 2008)	Multinomial probit models	Organizational innovation is shown to be a determinant factor for innovation persistence.
Antonelli, Crespi and Scellato (2012)	Data on 451 Italian manufacturing companies observed during the years 1998-2006	Transition probability matrix and dynamic probit model	Clearer evidence of persistence in the case of product innovation with respect to process innovation when complementarity effects are taken into account.

Table 2 – Definition and summary statistics of variables. All reported variables are time varying. Financial variables are deflated using year 2000 basic prices.

Variable	Definition	Mean	Median	Std err.	1^{st} perc	99^{th} perc
SIZE	Log(Total Assets) computed with perpetual inventory method	14.351	14.390	1.387	11.011	17.741
WAGE	Log (Labour costs/number of employees)	10.307	10.232	0.248	9.744	11.015
PCM	Price-cost-margin	0.285	0.279	0.256	0.056	0.671
INNO	Dummy = 1 in year t if TFP_t-TFP_{t-2}>0	0.401	0	0.490	0	1
INTANG	Ratio of intangible to tangible assets	0.158	0.080	0.194	0	0.858
REG_TFP	Average of the log of TFP of all companies in the same region of firm i excluding the contribution of firm i	8.327	8.350	0.162	7.857	8.623
SECT_TFP	Average of the log of TFP of all companies in the same sector of firm i, excluding the contribution of firm i	8.146	8.367	0.712	5.764	9.204

Table3 – Transition probability matrixes for different sub samples and time periods. Standard errors in parentheses.

All period		
	INNOt	NOT INNOt
INNOt-1	57.95%	42.05%
	(0.004)	(0.004)
NOT INNOt-1	32.04%	67.96%
	(0.003)	(0.003)

Before 2001		
	INNOt	NOT INNOt
INNOt-1	45.53%	54.67%
	(0.006)	(0.006)
NOT INNOt-1	27.42%	72.58%
	(0.004)	(0.004)

After 2001		
	INNOt	NOT INNOt
INNOt-1	66.95%	33.05%
	(0.005)	(0.005)
NOT INNOt-1	35.60%	64.40%
	(0.004)	(0.004)

Companies located in North-west		
Before 2001		
	INNOt	NOT INNOt
INNOt-1	43.06%	56.94%
	(0.008)	(0.008)
NOT INNOt-1	25.50%	74.50%
	(0.005)	(0.005)
After 2001		
	INNOt	NOT INNOt
INNOt-1	70.17%	29.83%
	(0.006)	(0.006)
NOT INNOt-1	36.15%	63.85%
	(0.005)	(0.005)

Companies located in North-est		
Before 2001		
	INNOt	NOT INNOt
INNOt-1	45.00%	55.00%
	(0.010)	(0.010)
NOT INNOt-1	27.45%	72.55%
	(0.006)	(0.006)
After 2001		
	INNOt	NOT INNOt
INNOt-1	65.98%	34.02%
	(0.008)	(0.008)
NOT INNOt-1	35.52%	64.48%
	(0.006)	(0.006)

Companies located in central regions		
Before 2001		
	INNOt	NOT INNOt
INNOt-1	48.55%	51.45%
	(0.015)	(0.015)
NOT INNOt-1	31.29%	68.71%
	(0.010)	(0.010)
After 2001		
	INNOt	NOT INNOt
INNOt-1	61.97%	38.03%
	(0.012)	(0.012)
NOT INNOt-1	33.74%	66.26%
	(0.009)	(0.009)

Companies located in South		
Before 2001		
	INNOt	NOT INNOt
INNOt-1	58.14%	48.86%
	(0.025)	(0.025)
NOT INNOt-1	38.49%	61.51%
	(0.021)	(0.021)
After 2001		
	INNOt	NOT INNOt
INNOt-1	58.63%	41.37%
	(0.012)	(0.012)
NOT INNOt-1	36.23%	63.77%
	(0.009)	(0.009)

Table 4 Dynamic random effect probit model with the Wooldridge specification. Dependent variable $INNO_t$.

	MODEL I	MODEL II	MODEL III	MODEL IV
L.INNO	0.668***	0.668***	0.650***	0.650***
	(0.014)	(0.014)	(0.014)	(0.014)
L.SIZE	0.048***	0.048***	0.047***	0.047***
	(0.015)	(0.015)	(0.015)	(0.015)
L.PCM	0.002	0.001	0.002	0.002
	(0.003)	(0.003)	(0.003)	(0.003)
L.WAGE	0.548***	0.552***	0.619***	0.621***
	(0.043)	(0.043)	(0.043)	(0.043)
L.INTANG	0.150**	0.149**	0.155***	0.155***
	(0.059)	(0.059)	(0.059)	(0.059)
REG_TFP		0.203***		0.127**
		(0.062)		(0.062)
SECT_TFP			0.335***	0.334***
			(0.013)	(0.013)
AVGWAGE	0.489***	0.467***	0.562***	0.547***
	(0.055)	(0.055)	(0.055)	(0.056)
AVGSIZE	-0.036**	-0.035**	-0.036**	-0.035**
	(0.016)	(0.016)	(0.016)	(0.016)
AVGPCM	-0.008	-0.007	-0.008	-0.007
	(0.009)	(0.009)	(0.009)	(0.009)
AVGINTANG	0.047	0.049	0.051	0.053
	(0.074)	(0.074)	(0.074)	(0.075)
INITIAL	0.077***	0.075***	0.073***	0.071***
	(0.013)	(0.013)	(0.014)	(0.014)
Industry dummy	yes	yes	yes	yes
Year dummy	yes	yes	yes	yes
Constant	-11.066***	-11.964**	-10.512***	-11.870**
	(3.619)	(5.704)	(2.841)	(5.628)
Observations	49140	49140	49140	49140
Wald Chi-sq	8661.2***	8665.7***	9073.4***	9087.4***
Log likelihood	-28540.4	-28535.5	-28225.3	-28223.5

*Robust standard errors in parentheses,*** p<0.01, ** p<0.05, * p<0.10*

Table 5 - Dynamic random effect probit model with the Heckman approach. Dependent variable: INNOt. Model estimated with the redprob routine by Stewart (2007). Instruments for reduced form: pre-sample levels of firm-level variables.

	MODEL I	MODEL II	MODEL III	MODEL IV
L.INNO	0.624***	0.626***	0.604***	0.605***
	(0.013)	(0.013)	(0.013)	(0.013)
L.SIZE	0.021***	0.023***	0.022***	0.023***
	(0.005)	(0.005)	(0.005)	(0.005)
L.PCM	-0.278***	-0.272***	-0.279***	-0.274***
	(0.023)	(0.023)	(0.023)	(0.023)
L.WAGE	0.284***	0.302***	0.315***	0.329***
	(0.031)	(0.032)	(0.031)	(0.032)
L.INTANG	0.206***	0.205***	0.220***	0.219***
	(0.033)	(0.033)	(0.033)	(0.033)
REG_TFP		0.247***		0.185***
		(0.060)		(0.060)
SECT_TFP			0.327***	0.326***
			(0.013)	(0.013)
Industry dummy	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Constant	2.983***	1.098*	-16.329***	-17.614***
	(0.323)	(0.561)	(0.798)	(0.903)
Observations	49140	49140	49140	49140
Wald Chi-sq	8141.9***	8155.7***	8759.9***	8766.9***
Log likelihood	-29247	-29240	-28938	-28934

Robust standard errors in parentheses,*** p<0.01, ** p<0.05, * p<0.10

ANNEX A – Sectoral distribution of analysed firms

Table A1- Sectoral distribution of companies included in the sample

Industry Classification	Number of companies	Percentage
Food and beverages	561	8.0%
Textile	607	8.6%
Textile product industry	212	3.0%
Leather and leather products manufacturing	249	3.5%
Wood and wood products manufacturing	155	2.2%
Pulp, paper and paper products manufacturing	174	2.5%
Printing	193	2.7%
Chemical industry	401	5.7%
Plastics and rubber manufacturing	421	6.0%
Non-metallic mineral product manufacturing	390	5.6%
Metallurgy	275	3.9%
Metal products manufacturing	983	14.0%
Mechanical machinery and equipment manufacturing	1,078	15.4%
Computer and electronic manufacturing	24	0.3%
Electrical machinery and equipment manufacturing	287	4.1%
Telecommunication machinery and equipment	91	1.3%
Medical, optical and precision equipment	143	2.0%
Transportation equipment manufacturing	122	1.7%
Other transport equipment manufacturing	61	0.9%
Furniture	487	6.9%
Software	106	1.5%
Total	7,020	100.0%

ANNEX B - Robustness control

In the following table we report the results for the dynamic probit model suing a alternative model specific with a three year lag for the computation of the dummy dependent variable. The main results presented in the paper with the two years time lag are confirmed. As expected, we estimate an overall lower level of persistence from the autoregressive covariate due to the fact that we using a longer time window.

Table B1 – Robustness control. Dynamic probit model using a three years time lag for the computation of the dependent variable (INNO).

	MODEL I	MODEL II	MODEL III	MODEL IV
L.INNO	0.440***	0.439***	0.433***	0.441***
	(0.017)	(0.017)	(0.017)	(0.017)
L.SIZE	0.143***	0.143***	0.140***	0.139***
	(0.020)	(0.020)	(0.020)	(0.020)
L.PCM	-0.238***	-0.236***	-0.240***	-0.237***
	(0.054)	(0.054)	(0.054)	(0.054)
L.WAGE	2.155***	2.150***	1.912***	1.888***
	(0.058)	(0.058)	(0.059)	(0.060)
L.INTANG	0.339***	0.342***	0.335***	0.329***
	(0.093)	(0.093)	(0.093)	(0.093)
REG_TFP		0.157**		0.155**
		(0.078)		(0.079)
SECT_TFP			0.304***	0.302***
			(0.161)	(0.161)
AVGWAGE	1.863***	1.844***	1.687***	1.658***
	(0.070)	(0.071)	(0.071)	(0.072)
AVGSIZE	-0.110***	-0.109***	-0.110***	-0.109***
	(0.021)	(0.021)	(0.021)	(0.021)
AVGPCM	-0.008	-0.008	-0.007	-0.007
	(0.012)	(0.012)	(0.011)	(0.011)
AVGINTANG	-0.063	-0.064	-0.052	-0.052
	(0.075)	(0.075)	(0.076)	(0.076)
INITIAL	-0.024	-0.025	-0.022	-0.022
	(0.016)	(0.016)	(0.016)	(0.016)
Industry dummy	2.673***	1.516*	-19.035***	-19.421***
Year dummy	(0.517)	(0.777)	(1.259)	(1.369)
Constant	-6.655***	-6.711***	-6.615***	-6.195***
	(0.586)	(0.607)	(0.595)	(0.476)
Observations	42120	42120	42120	42120
Wald Chi-sq	8061.1***	8074.8***	8277.3***	8213.4***
Log likelihood	-19168.4	-19167.9	-18991.0	-18991.2