OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 0305-9049
doi: 10.1111/obes.12482

On the Estimation of True State Dependence in the Persistence of Innovation*

MARTA F. ARROYABE† and MARTIN SCHUMANN‡

†*Essex Business School, University of Essex, Colchester CO4 3SQ, UK (e-mail: mf17255@essex.ac.uk)*
‡*Department of Quantitative Economics, School of Business and Economics, Maastricht University, Maastricht, Limburg, Netherlands (e-mail: m.schumann@maastrichtuniversity.nl)*

Abstract

This paper explores the persistence often found in firms' innovation and advances current research by investigating its actual nature. Previous studies have aimed at disentangling true state dependence from spurious state dependence by using a random effects (RE) dynamic panel probit approach, thereby imposing strong conditions on the underlying structure of the unobserved heterogeneity. Building on recent advances in the econometric literature, which allows for true fixed effects estimation of dynamic nonlinear panel data models, we demonstrate that relaxing the assumptions on the unobserved heterogeneity can have a considerable effect on the estimates of true state dependence. While we confirm the existence of a strong persistence of innovation in firms, we however find that true state dependence only explains about half of the persistent behaviour displayed by firms; this is in contrast to the popular RE methodology that attributes 70% to 100% of persistence to true state dependence. Our results suggest that policy programs aimed at encouraging initial innovations alone are useful but may not possess a long-term stimulating effect on innovation activity.

I. Introduction

Innovation is considered one of the key determinants of firms' growth (Aghion and Howitt, 1992; Aghion, Harris and Vickers, 1997). Innovation activities lead to the creation of new products that satisfy consumers' needs or to the development of new processes that lower the production costs, thus increasing market share, sales and profits (Fagerberg, Mowery and Nelson, 2005; Pavitt, 1990). The capacity of firms to innovate over long periods of time provides firms with the possibility of obtaining a sustained competitive advantage, which helps explaining differences in firm performance in the long run (Conner, 1991; Patel and Pavitt, 1991; Dosi *et al*., 1995;

JEL Classification numbers: C23, D22, L20, O31.
*We thank the Editor Brian Bell and two anonymous referees for comments that greatly improved this paper. We also thank Katrin Hussinger, Pierre Mohnen, Gautam Tripathi and seminar participants at the Academy of Management 2019 conference.

Hitt, Hoskisson and Kim, 1997; Teece, Pisano and Shuen, 1997). Given the importance of understanding the sources of long-term differences in firm performance, the capacity of firms to perform sustained innovation has received wide attention in the scholarly community (e.g. Grant, 1991; Lengnick-Hall, 1992; Cefis, 2003; Guarascio and Tamagni, 2019).

Persistence of innovation is defined as the level of continuity of innovation activities and outcomes over time (Heckman, 1981; Flaig and Stadler, 1994; Ganter and Hecker, 2013). From a theoretical point of view, the existence of persistent behaviour in innovation substantiates endogenous growth models (Aghion and Howitt, 1992) and provides an explanation for the often observed differences in long-term performance across firms (Flaig and Stadler, 1994; Geroski, Van Reenen and Walters, 1997). From a policy point of view, the existence of innovation persistence gives grounds for the implementation of policies aiming to spur innovation, as these are expected to have long-term effects.

Persistence of innovation can be traced to either true or spurious state dependence (Raymond *et al*., 2010; Antonelli, Crespi and Scellato, 2013; Tavassoli and Karlsson, 2015). If the persistence is due to true state dependence, conducting innovation in the past has a structural impact on the probability of conducting innovation in the future. Spurious state dependence occurs if the persistence is falsely attributed to past innovation experience (e.g. when serially correlated firm characteristics are insufficiently accounted for).

Empirically, previous studies have aimed at (1) showing the existence of persistence of innovation, and (2) analysing the nature of this persistence (i.e. differentiating between true or spurious state dependence). While obtaining a set of mixed results, the majority of these studies have based their analysis on the same empirical strategy for the identification of spurious and true state dependence, that is, they use a dynamic panel data specification with random effects (RE). In particular, a popular approach among recent papers is to implement a RE dynamic probit model using Wooldridge's maximum likelihood estimator (MLE) (Wooldridge, 2005). While computationally convenient, RE estimators impose strong assumptions on the conditional distribution of the unobserved heterogeneity and its correlation with the observable characteristics (Honoré, 2002; Carro, 2007).

In this paper, we make use of recent econometric advances in order to disentangle true from spurious state dependence by adopting the true fixed effects (FE) approach for dynamic nonlinear panel data of Fernández-Val and Weidner (2016). We are thus able to correctly measure the degree of true state dependence in the persistence of innovation by isolating the effect of observed and unobserved firm characteristics. As compared to the RE methodology, FE approaches do neither require specifying the distribution of the unobserved heterogeneity (which often leads to a misspecification bias) nor do they impose restrictions on the dependence structure between observed characteristics and the unobserved heterogeneity (which makes FE estimators robust to arbitrary correlations between observed and unobserved characteristics).

Our results suggest that RE approaches may overestimate the extent to which persistence of innovation can be regarded as a consequence of true state dependence. This means that after accounting for unrestricted unobserved individual heterogeneity, the estimated magnitude of the coefficient measuring the effect of experiencing innovation in the previous period is much smaller as compared to the coefficient estimate based on the RE approach. Specifically, while the RE approach attributes between 70% and 100% of persistence of innovation to true state dependence, the FE estimates indicate that this proportion is closer to 50%. While we do not suggest that earlier studies based on the RE methodology are necessarily invalid, our results do indicate the need for a critical evaluation of distributional assumptions used in the measurement of true state dependence.

Our findings have important implications for policy makers, as they partially question the long-term effectiveness of programs that aim at fostering initial innovations. While we do not discourage these policies (as we do find true state dependence in the persistence of innovation), our results indicate that programs targeting sustained innovative activity may also need to focus on stimulating firm characteristics that drive innovation.

II. Theory and literature review

Theoretical explanations for persistence

Persistence of innovation refers to the degree of intertemporal continuity in innovative behaviour and describes the influence of past innovation activities on current and future innovation activities and success (Flaig and Stadler, 1994; Ganter and Hecker, 2013; Guarascio and Tamagni, 2019). At the macroeconomic level, the idea that technology develops in an evolutionary fashion has been a critical issue in the literature of innovation and competition (e.g. Nelson and Winter, 1982). The existence of innovation persistence validates endogenous growth models (e.g. Aghion and Howitt, 1992), which attribute long-run growth to the continuous accumulation of new and valuable knowledge, and highlights the role of incumbent firms (creative accumulation) as driver of industry dynamics (see Malerba and Orsenigo, 1996). At the microeconomic level, innovation persistence provides an explanation for firms' competitive advantage and for the sustained differences in performance across firms (Flaig and Stadler, 1994; Geroski *et al*., 1997).

Because of its importance both at the macro- and micro-levels, previous studies have aimed at understanding the emergence of persistence of innovation. These studies have often modelled innovation activity as being state dependent, where, besides other firm characteristics, firms' current innovator status depends on the past innovator status. As explained by Heckman (1981), this state dependence can be attributed to true and spurious state dependence.

True state dependence occurs when past experience of an event (e.g. of innovation) has a structural effect on the probability of experiencing that event in the future, regardless of other individual characteristics (Heckman, 1981). For firms, this means that conducting innovation has a behavioural effect on the decision to innovate in the future, such that otherwise identical firms that did not conduct innovation behave differently in future periods (Heckman, 1981; Hecker and Ganter, 2014). On the other hand, the propensity to innovate may depend on other firm specific characteristics. If these characteristics are not controlled for while being correlated over time, previous innovation activities may incorrectly appear to be determinants of future innovative activities (as past activities act as a proxy for the unobservable/uncontrolled characteristics) (Heckman, 1981; Hecker and Ganter, 2014). This is referred to as spurious state dependence.

Previous literature provides three major theoretical explanations for the occurrence of true state dependence. The first theory, *success-breeds-success*, argues that innovation activities are capital intensive, risky and require large amounts of resources that firms often obtain from external sources (Hall, 2002; Piga and Vivarelli, 2004; Brown, Fazzari and Petersen, 2009). Because of the difficulties for external financers to evaluate these innovation activities, firms often face financial constrains (Czarnitzki and Hottenrot, 2010), which are mitigated when firms show previous success on innovation (Flaig and Stadler, 1994). Moreover, previous innovation success allows firms to reinvest the profits into R&D, increasing the probability of success (Schumpeter, 1934; Nelson and Winter, 1982; Latham and Le Bas, 2006).

The second theory, *learning by doing*, is rooted in the evolutionary theory and argues that R&D shows dynamic increasing returns and that knowledge is cumulative (Arrow, 1962; Cohen and Levinthal, 1989; Klevorick *et al*., 1995; Cefis and Orsenigo, 2001). The generation of new knowledge, which fuels innovation, is a process in which firms recombine prior knowledge and external knowledge (dependent on the absorptive capacity, which in turn is a function of previous knowledge) to generate new ideas (Rosenberg, 1976; Nelson and Winter, 1982; Weitzman, 1998). In this framework, as knowledge and experience accumulate, unique competences allow firms to maintain innovative performance along the technological trajectories (Nelson and Winter, 1982; Kogut and Zander, 1992; Dosi and Marengo, 1994).

The third theory, *sunk costs* of R&D investment, argues that R&D activities require large start-up costs (e.g. R&D facilities, equipment, hiring and training of scientific and specialized staff), which are largely unrecoverable (Sutton, 1991; Cohen and Klepper, 1996). Hence, these costs represent a barrier of exit for innovating firms and a barrier of entry for non-innovating firms.

While less explored, spurious state dependence has been mainly attributed to a variety of firm-specific factors (Cefis and Orsenigo, 2001). Spurious state dependence is rooted in the resource based view (RBV) theory of the firm (Wernerfelt, 1984; Barney, 1991), in which innovation persistence is explained by the initial allocation of the firm's innovation capabilities. These characteristics are heterogeneously distributed among firms, stable and hard to change, as they possess a high level of inertia (Helfat, 1994; Stuart and Podolny, 1996; Clausen *et al*., 2011). Previous studies have pointed out that strategic positioning (Clausen *et al*., 2011), corporate culture (Khazanchi, Lewis and Boyer, 2007), research abilities (Zucker, Darby and Brewer, 1998; Baumol, Schilling and Wolff, 2009), managerial talent (Bloom and Van Reenen, 2007, 2010; Custódio, Ferreira and Matos,), organizational routines (Cyert and March, 1963; Nelson and Winter, 1982; Levitt and March, 1988) and firms' organizational (Nelson and Winter, 1982; Hannan and Freeman, 1984) and dynamic capabilities (Teece, 2007; Barreto, 2010) are key internal factors in firms' initiation and continuous adoption of innovative practices and activities that ensure a lasting process of innovation (Cefis and Orsenigo, 2001).

These theories not only support a pattern of innovation persistence across different types of innovation activities, but they also provide arguments explaining differences in the degree of persistence among these activities (Le Bas and Scellato, 2014; Tavassoli and Karlsson, 2015; Guarascio and Tamagni, 2019). For instance, technological innovations refer to those innovation activities aimed at improving the performance of a product or service in terms of its quality, cost, number of features or speed of delivery and comprise product and process innovation (Schumpeter, 1934; Cohen and Levinthal, 1990). While technological innovations have been found to display high levels of persistence, the sunk-cost and success-breeds-success theories suggest that product innovation will show higher levels of persistence (Tavassoli and Karlsson, 2015; Guarascio and Tamagni, 2019). The sunk-cost theory might not be relevant in many industries as firms do not conduct R&D to develop new processes themselves, but rather buy machinery and process equipment from specialized firms in the machinery industry which are the ones that conduct the R&D. Similarly, the success-breeds-success arguments cannot be directly applicable to process innovation as successful process innovations do not directly translate into higher market power, which firms can leverage for obtaining better finance for subsequent innovation or to exploit economies of scale.

Previous empirical evidence

While early studies exploring persistence of innovation made use of patent data, more recent studies are based on various country-level innovation surveys (in its majority the Community Innovation Survey (CIS)).

Patent-based studies employed descriptive statistics, transition probability matrices (TPM) and duration models (Weibull or Cox models) to investigate the existence of persistence. In general, these studies found little evidence of persistence, with strong persistence only present in the case of high patenting firms (e.g. Crépon and Duguet, 1997; Geroski *et al*., 1997; Malerba and Orsenigo, 1999; Cefis and Orsenigo, 2001; Cefis, 2003). These results have been attributed to the use of patent data, which implicitly requires firms to have applied for a patent, so that the analysis might effectively reflect the continuous winners of patent races or the persistence of innovative leadership (Duguet and Monjon, 2004; Peters, 2009; Raymond *et al*., 2010). Another shortcoming of these models is that the methodology and database employed are not adequate for differentiating between true and spurious state dependence (Cefis and Orsenigo, 2001).

More recent studies have taken advantage of the CIS to analyse innovation persistence at the firm level, exploring both innovation input (e.g. R&D expenditures) and output (e.g. product and process innovation). These studies have aimed at analysing the existence and the nature (true state dependence vs. spurious state dependence) of persistence of innovation activity, making use of dynamic RE probit models. Within this set of studies, a particularly popular approach has been the dynamic RE probit model proposed by Wooldridge (Wooldridge, 2005) which circumvents the initial condition problem (Arulampalam and Stewart, 2009). As shown in Table 1, since Peters (2009) first applied Wooldridge's approach in the context of persistence of innovation, numerous studies have made use of this approach to explore different persistence-related questions.

TABLE 1

Literature review

Authors	*Sample*	*Measure for innovation*	*Main findings*	*Evidence of persistence*	*Point estimates [Marginal effects]*
Peters (2009)	German CIS: manufacturing (1994–2002) and services (1996–2002)	Dummy for positive innovation expenditure	True state dependence:	High	Manufacturing: 0.333 [23pp]
		Dummy for product/process innovation introduction	Higher for R&D performers		Services: 0.103 [8.2pp]
			Stronger in manufacturing		
Martinez-Ros and Labeaga (2009)	Spanish ESEE: manufacturing (1990–9)	Dummy for the introduction of a new product	Persistence is present in both product and process innovations:	High	Product: 0.948 [N/A]
		Dummy for the introduction of a new process	It increases the probability to develop more innovations		Process: 0.789 [N/A]
Raymond *et al*. (2010)	Dutch CIS: manufacturing (1994–2002)	Dummy for the introduction of new/improved product or process (TPP)	True persistence in the probability of innovating in high tech (spurious for low tech)	High & Low	TPP (high tech): 0.273 [N/A]
		Shares of sales coming from new or improved products	Small effect of past innovation output in current innovation output (only in high tech)		TPP (low tech): 0.199 [N/A]
					Sales (high tech): 0.132 [N/A]
					Sales (high tech): 0.077 [N/A]

(*Continued*)

TABLE 1

(*Continued*)

Authors	*Sample*	*Measure for innovation*	*Main findings*	*Evidence of persistence*	*Point estimates [Marginal effects]*
Clausen *et al*. (2011)	Norwegian CIS: industry (1996–2004)	Dummy for the introduction of new product	Spurious product innovation in the high-tech sector	Medium	Product: 0.436 [0.166]
		Dummy for the introduction of a new process	Persistence of process innovation in the low-tech sector		Process: 0.323 [0.115]
			Firms' initial strategies (firms' heterogeneity) is highly explanatory of innovation		
Huergo and Moreno (2011)	Spanish ESEE: manufacturing (1990–2005)	Dummy for engaging in R&D activities	Existence of true state dependence both in the decision of R&D investment and in the production of innovations	High	Engagement: N/A [0.586]
		Dummies for achieving product and/or process innovations			
		Total expenditures in R&D over total employment			
Antonelli, Crespi and Scellato (2012)	Italian MCC: manufacturing (1998–2006)	Dummies for performing product and/or process innovation	Differentiated patterns of persistence:	High	Product: 0.419 [N/A]
		Dummy for positive R&D expenditures	Higher level of persistence is found for the R&D based innovation activities		Process: 0.218 [N/A]
			Higher innovation persistence in the group of R&D performers		R&D: 0.238 [N/A]
			Higher persistence for product than for process innovation		

(*Continued*)

TABLE 1
(*Continued*)

Authors	*Sample*	*Measure for innovation*	*Main findings*	*Evidence of persistence*	*Point estimates [Marginal effects]*
Arqué-Castells (2013)	Spanish ESEE: manufacturing (1998–2009)	Dummy variable for positive R&D expenditures	Confirm existence of true state dependence	High	R&D: 1.57 [0.36]
			Subsidies contribute to persistence		
Antonelli *et al.* (2013)	Bureau Van Dick: Italian manufacturing firms (1996–2005)	Dummy for positive total factor productivity on the year $t-1$	Persistence is path-dependent	High	TFP: 0.668 [N/A]
			Persistence is affected by the accessibility to external and internal knowledge		
Ganter and Hecker (2013)	German CIS: all industries (2002–8)	Dummies for adoption of organizational, technological (product and/or process), product (new to the firm), product (new to the market), process innovation	True persistence of product innovation (new to the market)	Low	Product (new to the firm): 0.06 [0.6pp]
			No persistence for product innovation (new to the firm), process innovation, or organizational innovation		Product (new to the market): 1.35 [17.7 pp]
					Process: 0.04 [0.8pp]
Triguero and Corcoles (2013)	Spanish ESEE: manufacturing (1990–2008)	Dummy for conducting or contracting R&D activities	Innovation input and output are highly persistent at the firm level	High	R&D activities: N/A [0.509]
		Dummy for the introduction of product and/or process innovation	Persistence is higher in R&D activities		Innovation: N/A [0.306]

(*Continued*)

TABLE 1

(*Continued*)

Authors	*Sample*	*Measure for innovation*	*Main findings*	*Evidence of persistence*	*Point estimates [Marginal effects]*
Haned, Mothe and Nguyen-Thi (2014*)	French CIS: manufacturing (2002–08)	Dummy for the introduction of product innovation (only)	Organizational innovation explains persistence in technological (product and/or process) innovation	High	Product: 0.233 [N/A]
		Dummy for the introduction of process innovation (only)			Process: 0.164 [N/A]
		Dummy for the introduction of product and process innovation			
Hecker and Ganter (2014*)	German CIS: manufacturing and services (2002–8)	Dummies for adoption of organizational, product and process innovation	True state dependence for product innovation	High & Low	Product: 0.59 [N/A]
			State dependence of process and organizational innovation is driven by time-invariant and firms' unobserved characteristics		Process: 0.04 [N/A]
					Organizational: −0.28 [N/A]

(*Continued*)

TABLE 1
(*Continued*)

Authors	*Sample*	*Measure for innovation*	*Main findings*	*Evidence of persistence*	*Point estimates [Marginal effects]*
Lhuillery (2014*)	French CIS: manufacturing (1998–2008)	Share of innovative sales for innovative products	Innovation success is persistent	High	Total sales: 0.093 [N/A]
		Share of innovative sales for innovative products not new to the firm (incremental)	Innovation marketing does not positively influence persistent innovation success in low-tech industries		Incremental: 0.085 [N/A]
		Share of innovative sales for innovative products new to the market (radical)	Innovation marketing positively influences persistent innovation success for incremental innovation but negatively influences it for radical innovation		Radical: 0.094 [N/A]
Woerter (2014*)	Swiss KOF: all industries (1996–2008)	Sales share of R&D expenditures	Persistence is related to market competition	High & Low	R&D expenditures: N/A [0.501]
			Persistence is observed in markets with few competitors		
Diana Suarez (2014)	Argentinian INDEC: manufacturing (1998–2002)	Dummy variable for achieving at least one type of innovation (product, process, organization and/or commercialization)	Instability of the environment nullifies innovation persistence	Low	Innovation: −0.114 [N/A]

(*Continued*)

TABLE 1

(*Continued*)

Authors	*Sample*	*Measure for innovation*	*Main findings*	*Evidence of persistence*	*Point estimates [Marginal effects]*
Le Bas and Poussing (2014)	Luxembourgish CIS: manufacturing and services(2002–8)	Dummy for the introduction of product or process innovation (single)	Firms conducting both product and process innovation are more likely to remain persistent than single innovators	High	Product or process: 0.858 [N/A]
		Dummy for the introduction of product and process innovation (complex)			Product and process: 0.995 [N/A]
Tavassoli and Karlsson (2015)	Swedish CIS: manufacturing and services (2004–12)	Dummies for the introduction of process, product, organizational or marketing innovation	Product innovation displays the strongest persistence	High & Low	Product: 0.354 [15.7pp]
			Product, process and organizational present true state dependency		Process: 0.199 [12pp]
			Marketing innovation's persistence is spurious		Organizational: 0.328 [12pp] Marketing: 0.200 [6pp]
Cefis and Marsili (2015)	Dutch Central Bureau of Statistics (CIS and BR): manufacturing (1994–2002)	Dummy for the introduction of new/improved product/service/process or for positive R&D expenditures	Previous M&A activity increases persistence	High	Innovative activity: 3.19 [N/A]
		Dummy for positive turnover of technologically new or improved products/service	Larger firms (as compared to medium sized or small firms) are the greatest persistent innovators and benefit the most from M&As		

(*Continued*)

TABLE 1
(*Continued*)

Authors	*Sample*	*Measure for innovation*	*Main findings*	*Evidence of persistence*	*Point estimates [Marginal effects]*
Badillo and Moreno (2016)	Spanish CIS: manufacturing and services (2002–10)	Dummy for cooperating in innovation activities either horizontally, vertically or with institutions	Innovation collaboration displays persistence	High	Cooperation: N/A [0.329]
			Persistence is higher in the case of vertical collaboration		
Muinelo-Gallo and Suanes (2018)	Uruguay's National Agency for Research and Innovation: manufacturing (2001–9)	Dummy for the introduction of a new product	Persistent effects only for product innovations	High & Low	Product: N/A [0.562]
		Dummy for the introduction of a new process			Process: N/A [0.215]
Ayllón and Radicic (2019)	Spanish ESEE: manufacturing (2001–14)	Dummy for the introduction of a new product	Persistence is present in both product and process	High	Product: 1.160 [0.226]
		Dummy for the introduction of a new process	Particularly strong in process innovation		Process: 1.032 [0.293]
Nam and Tram (2019)	Vietnam's biannual surveys of SMEs: 2007–15	Dummy for the introduction of a new product	No evidence of persistence in any of the types of innovation	Low	New products: 0.035 [N/A]
		Dummy for the improvement of a existing product			Improving existing products: 0.047 [N/A]
		Dummy for the introduction of a new process			New processes: −0.040 [N/A]

(*Continued*)

TABLE 1
(*Continued*)

Authors	*Sample*	*Measure for innovation*	*Main findings*	*Evidence of persistence*	*Point estimates [Marginal effects]*
Antonioli *et al.* (2019)	Italian MET survey: manufacturing (2005–13)	Dummy for the introduction of a new product	In times of crisis, persistence is lower and moderated by public support	Low	Product: 0.146 [N/A]
		Dummy for the introduction of a new process	Persistence is lower in SMEs		Process: 0.441 [N/A]
		Dummy for patent filing			Patent: 0.340 [N/A]

Notes: *Paper included in a special issue on persistence of innovation in *Economics of Innovation and New Technology*, edited by Le Bas and Scellato (2014) (2014, volume 23, issue 5–6).

About half of the studies find clear evidence of strong persistence. The rest of the studies find a mix of weak evidence or no evidence for the existence of persistence. Our research aims to shed light on this debate by exploring this question using a recently developed bias corrected fixed effects (BCFE) approach developed by Fernández-Val and Weidner (2016). In particular, our paper explores the following research questions: *Is innovation persistent?* And if so, *what is the source of this persistence (i.e. true state dependence or spurious state dependence)?*

III. Methodology

In this section, we review the specification of a model of persistence of innovation together with the main econometric challenges. We then proceed by discussing the well-known FE and RE approaches together with the recently BCFE estimators while summarizing their respective advantages and shortcomings in the measurement of true state dependence in the persistence of innovation. From now on, we let n denote the number of individuals and T the panel length.

Measuring true state dependence

Following Peters (2009), innovation activity is conducted whenever the value of expected profits from innovation denoted by Y_{it}^* is positive. While Y_{it}^* is typically latent, we observe $Y_{it} = 1\{Y_{it}^* > 0\}$, where $1\{\cdot\}$ denotes the indicator function. As explained in section 'Theoretical explanations for persistence', past innovation experience may have a structural impact on the probability of conducting innovation in future periods. We refer to this case as 'true state dependence'. However, many firm-specific characteristics that are of great importance in determining firms' innovation activities are stable over time and heterogeneously distributed across firms. Moreover, many of these determinants are not available in commonly used data sets or may even generally be unobservable. The presence of unobserved firm-specific differences leads to serial correlation in innovation activity, so that past innovation experience may appear to have a structural effect on the probability of future innovation activity whereas in fact it does not, thus creating 'spurious state dependence'. The preceding discussion suggests that a dynamic econometric model which allows for the presence of unobserved effects should be used to disentangle true from spurious state dependence. Thus, $Y_{it} = 1\{\rho Y_{it-1} + X_{it}'\theta + c_i + \varepsilon_{it} > 0\}$, where X_{it} denotes a vector of strictly exogenous covariates. Typically, distributional assumptions are imposed on the error term ε_{it} in order to account for the nonlinearity of this model. A popular choice is to assume that $\varepsilon_{it}|Y_{i0},\ \ldots,\ Y_{it-1},\ X_i,\ c_i \sim \mathrm{NIID}(0,\ 1)$, where $X_i = (X_{i1},\ \ldots,\ X_{iT})$ denotes the time series of explanatory variables and Y_{i0} is the initial condition. We thus arrive at a dynamic probit model with

$$P(Y_{it}|Y_{i0},\ldots,\ Y_{it-1},\ X_i,\ c_i) = \Phi(\rho Y_{it-1} + X_{it}'\theta + c_i). \tag{1}$$

The main econometric challenges in the estimation of the parameters in model (1) are to account for the dependence on the initial condition (for which assuming independence of the unobserved heterogeneity c_i is often unreasonable) and to disentangle the estimation of $\rho\rho$ (which measures the structural impact of past innovation experience on future innovation activity) from the effect of c_i. While estimates of the coefficients provide information about the sign and the relative magnitude of the effect of an explanatory variable, they do not provide information on the absolute magnitude of the effect. Therefore, in many situations, the ultimate object of interest is the average partial effect (APE) rather than the coefficient of an explanatory variable. Unfortunately, the APE also depends on the initial condition and the unobserved heterogeneity, so that similar econometric issues arise.

Fixed effects (FE) estimation

Ideally, one would account for unobserved heterogeneity by treating c_i as a fixed effect, thus avoiding any restrictions on the distributional features of the unobserved effects (Heckman, 1987). In this case, one would model the loglikelihood of individual i (scaled by the factor T^{-1}) as

$$\ell_{iT}(\rho,\ \theta,\ c_i) = \frac{1}{T}\sum_{t=1}^{T}[Y_{it}\Phi(\rho Y_{it-1} + X'_{it}\theta + c_i) + (1 - Y_{it})(1 - \Phi(\rho Y_{it-1} + X'_{it}\theta + c_i))],$$

where we condition on the initial condition and the unobserved effect, so that c_i acts as an additional parameter in the likelihood. As c_i is unobserved, this likelihood can, however, not be used for estimation and inference, which therefore has to be based on a 'pseudo-likelihood'. A popular choice is the 'concentrated' or 'profile' likelihood $\ell_{iT}(\rho,\ \theta,\ \hat{c}_i(\rho,\ \theta))$ where $\hat{c}_i(\rho,\ \theta) = \arg\max_c \ell_{iT}(\rho,\ \theta,\ c)$. The MLE of $(\rho,\ \theta')'$ is then obtained as

$$(\hat{\rho}_{MLE},\ \hat{\theta}'_{MLE})' = \arg\max_{(\rho,\ \theta')'} \sum_{i=1}^{n} \ell_{iT}(\rho,\ \theta,\ \hat{c}_i(\rho,\ \theta)), \tag{2}$$

whereas the FE APE of Y_{it-1} is calculated as

$$\widehat{APE}_{FE} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\Phi(\hat{\rho}_{MLE} + X'_{it}\hat{\theta}_{MLE} + \hat{c}_i(\hat{\rho}_{MLE},\ \hat{\theta}_{MLE})) - \Phi(X'_{it}\hat{\theta}_{MLE} + \hat{c}_i(\hat{\rho}_{MLE},\ \hat{\theta}_{MLE})). \tag{3}$$

While the FE MLE can easily be implemented in standard statistical software (for instance by including individual dummy variables), it is unfortunately often unreliable in practice due to the incidental parameters problem (IPP) first noted in Neyman and Scott (1948). As can be seen from the definition of the FE MLE in equation (2), the dependence of $\hat{c}_i(\rho,\ \theta)$ on the parameters of interest leads to a contamination of the estimators for $(\rho,\ \theta')'$. It can then theoretically be shown that the FE MLE suffers from a bias of order $O(T^{-1})$ while its asymptotic distribution is only correctly centred around the true value if $n/T \to 0$, that is, when the panel length is much larger than the number of individuals (see, for instance, Hahn and Newey, 2004 or Arellano and Hahn, 2007). The latter condition is surely unreasonable in most microeconomic panel data sets in which typically $n >> T$. In practice this means that the FE MLE and the FE APE are severely biased in short panels and are thus of limited use (see, e.g.

Heckman, 1987; Greene, 2004 or Czarnowske and Stammann, 2019 for simulation evidence on the performance of the FE MLE).

Random effects (RE) estimation

One approach that circumvents the IPP and the initial condition problem is the correlated RE estimator of Wooldridge (2005). Here, instead of treating the unobserved heterogeneity as a parameter to be estimated, the distribution of c_i is explicitly modelled conditional on observable covariates and the initial condition Y_{i0}, which allows for certain forms of dependence between the aforementioned variables. In the following, we discuss a ‘parsimonious’ version of Wooldridge’s RE estimator, which has been particularly popular in earlier literature on persistence in innovation. This means that instead of modelling the dependence of the unobserved effect on the observed characteristics using the full time series $X_{i1}, \ldots, X_{iT}$ of the explanatory variables, the mean of the unobserved effect is modelled as a function of the mean over time of the explanatory variables $T^{-1}\sum_{t=1}^{T} X_{it}$.[1] Hence, resembling (Wooldridge 2005 equation 15), the unobserved effect c_i is modelled as $c_i = \alpha_0 + \alpha_1 Y_{i0} + \bar{X}_i' \alpha_2 + a_i$, where $a_i | Y_{i0}, \bar{X}_i \sim N(0, \sigma_a^2)$. This leads to the model

$$P(Y_{it}|X_{it}, Y_{it-1}, \bar{X}_i, Y_{i0}, a_i) = \Phi(\rho Y_{it-1} + X_{it}'\theta + \alpha_0 + \alpha_1 Y_{i0} + \bar{X}_i'\alpha_2 + a_i), \qquad (4)$$

with respect to its conditional density (see Wooldridge, 2005, equation 21). The APE of previous innovation activity on current innovation activity is then computed as

$$APE_{RE} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}[\Phi(\hat{\rho}_a + X_{it}'\hat{\theta}_a + \hat{\alpha}_{0,a} + \hat{\alpha}_{1,a} Y_{i0} + \bar{X}_i'\hat{\alpha}_{2,a}) - \Phi(X_{it}'\hat{\theta}_a + \hat{\alpha}_{0,a} + \hat{\alpha}_{1,a} Y_{i0} + \bar{X}_i'\hat{\alpha}_{2,a})], \qquad (5)$$

where the ‘a’ in the index indicates that parameter estimates have been scaled by $(1 + \sigma_a^2)^{-1/2}$.

If the distribution of the unobserved effect c_i is correctly specified, integrating out a_i from equation (4) leads to a genuine likelihood function whose maximization in turn yields estimators that are consistent and have an asymptotically correctly centred distribution as $n \to \infty$, irrespective of the panel length. Thus, under correct specification, the correlated RE approach is an attractive tool for the measurement of true state dependence in the persistence of innovation. However, whether or not the RE framework is suitable in a particular situation needs to be critically assessed: as Wooldridge notes, a fully parametric specification of the conditional distribution of the unobserved heterogeneity naturally incorporates the risk of misspecification, which leads to inconsistent estimates for the parameters of interest. For instance, the RE approach imposes that the unobserved effect c_i depends on observed characteristics and the initial condition only via its mean. The latter is, however, very restrictive since, for

[1] As noted in Rabe-Hesketh and Skrondal (2013), including the explanatory variables in the initial period, that is, using $\bar{X}_i^* = \frac{1}{T+1}\sum_{t=0}^{T} X_{it}$ instead of $\bar{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it}$ can lead to severely biased estimates.

instance, it imposes constant variance of the unobserved heterogeneity. In reality, this assumption may not be reasonable, as the variance of the unobserved heterogeneity may for instance depend on firm size. For example, the spread in unobserved managerial talent may be wider for small firms run by managers with no formal education, whereas large firms are usually run by professional managers with more homogenous managerial skills. Moreover, as mentioned in (Wooldridge 2005 section 3), the normality assumption can be regarded as a choice mainly driven by convenience, leading to simple estimators of the parameters of interest, which are however inconsistent if the distribution of the unobserved effect is non-normal. For example, it is well-known that the innovation performance distribution of inventors is highly skewed (Lotka, 1926; Price, 1965; Narin and Breitzman, 1995), which may also indicate skewness of the distribution of research abilities across firms. Similarly, if unobserved managerial capability was quantifiable, one would also expect a highly skewed and therefore non-normal distribution, which violates the assumptions imposed by Wooldridge's approach.

Bias corrected fixed effects

Since distributional assumptions on the unobserved effects imposed by RE estimators are often hard to justify (and lead to misspecification bias when violated) while FE estimators suffer from considerable bias due to the IPP, BCFE estimation has been developed as an alternative estimation approach in more recent econometric literature. Rather than as an inconsistency problem as $n \to \infty$ while T remains fixed, the IPP is regarded as a asymptotic bias problem as both n, $T \to \infty$, which allows for the use of various bias correction techniques. It is for instance possible to correct the MLE directly (e.g. Hahn and Newey, 2004 or Fernández-Val, 2009) or to use jackknife (e.g. Dhaene and Jochmans, 2015) or implicit pseudo-likelihood (e.g. Arellano and Bonhomme, 2009 or Schumann, Severini and Tripathi, 2021) bias correction methods. All aforementioned approaches have in common that they reduce the stochastic order of the incidental parameter bias from $O(T^{-1})$ to $O(T^{-2})$. Moreover, the asymptotic distribution of BCFE estimators is correctly centred around the true value if $n/T^3 \to 0$ as n, $T \to \infty$, that is, when n can be regarded as being much larger than T.

In this paper, we apply the bias correction outlined in Fernández-Val and Weidner (2016). This approach permits the estimation of dynamic nonlinear panel data models with individual and time FE while allowing for arbitrary dependence of the initial condition on the FE. Fernández-Val and Weidner use stochastic expansions to derive additive correction terms for the FE MLEs of the coefficients and the APE.[2] Simulation results (for instance in Fernández-Val, 2009, tables 5–8) suggest that the correction term can to large extent eliminate the incidental parameters bias in moderately long (i.e. $T \geq 8$) panels. To compute the estimator, we make use of a

[2]The correction terms consist of sample averages of combinations of various likelihood derivatives (see section 4 of Fernández-Val and Weidner, 2016). The precise form of the correction terms is complex and requires specific notation. It is thus omitted here.

recently implemented software package (see Cruz-Gonzalez, Fernández-Val and Weidner, 2017a).[3]

Random vs. fixed effects in practice

In applications, the different assumptions on the unobserved effects can lead to a very different behaviour of RE and FE estimators. For instance, it is well-known that in FE models it is not possible to estimate the effect of observed covariates that to not vary over time (Wooldridge, 2010, section 10.5). In fact, the effect of time-invariant regressors cannot be identified, since it cannot be distinguished from the effect of the unobserved effect if the distribution of the latter is unrestricted. To see this, let γ_i denote a vector of time-invariant regressors while X_{it} and α_i denote a vector of time-varying regressors and the fixed effect, respectively, and note that the model $Y_{it} = X_{it}'\theta + \gamma_i'\beta + \alpha_i + U_{it}$ cannot be distinguished from the model $Y_{it} = X_{it}'\theta + \alpha_i^* + U_{it}$, where $\alpha_i^* = \alpha_i + \gamma_i'\beta$ is simply another unobserved effect.

Imposing a distribution on the unobserved effect (leading to a RE model) allows researchers to also include time-invariant regressors. However, as noted in (Wooldridge 2005 section 3), even in correlated RE models one cannot separately identify the partial effect of a time-invariant regressor from its partial correlation with the unobserved effect. Thus, while estimates of parameters of time-invariant variables that cannot be identified in a FE specification are reported for the RE specification, they cannot be interpreted in isolation from the effect of the unobserved heterogeneity.

Another consequence of leaving the distribution of the unobserved effect unrestricted is that individuals who do not change status (i.e. whose outcome variable does not change over time) cannot be used for estimation of the parameter of interest. The intuitive reason for this is that a time-constant outcome can be perfectly explained by a time-constant unrestricted unobserved effect, so that observed covariates do not provide further information that may be useful for identifying the parameters. The same logic does, however, not apply to RE models, where a specific distribution is imposed on the unobserved effects. As a consequence, individuals with time-constant outcome variables become informative for the parameter of interest, whereby the informative value of these individuals depends on the distribution imposed on the unobserved effects. In order to assess the latter effect, we also include the RE estimator based on only those individuals that change innovator status over time in our analysis.

Small sample properties of random and fixed effects estimators

As mentioned in section 'Random effects (RE) estimation', if the distribution of the unobserved effects is correctly specified, the RE estimator is consistent as $n \to \infty$ for any panel length while the FE estimator requires the panel length T to grow faster than n for valid inference. In practice, however, both RE and FE estimators of the coefficients in

[3]In **R**, the 'bife'-package offers a fast implementation of the BCFE estimator of Fernández-Val (2009), which diminishes potential disadvantages of BCFE relative to RE estimators. Notice further that this estimator coincides with the one used here in models without time FE (see Cruz-Gonzalez *et al*., 2017a, section 2.5).

dynamic probit models can suffer from a considerable bias in small samples. While the source of the FE bias is the incidental parameters problem (which can be corrected), the RE estimator usually suffers from a misspecification bias. Interestingly, as shown in Arellano and Bonhomme (2009) and Arellano and Bonhomme (2011), both the incidental parameters bias and the misspecification bias become negligible at the rate T^{-1} as the panel length increases. Therefore, in long panels, FE and RE estimators of the parameter of interest are likely to yield similar results. However, if the distributional assumptions required for RE estimation cannot be sufficiently justified, it is preferable in short or moderately long panels to use BCFE estimators, as they neither suffer from misspecification bias nor from a large incidental parameters bias.

While estimates of the coefficients provide information about the sign and the relative magnitude of the effect of an explanatory variable, they do not provide information on the absolute magnitude of the effect. Therefore, in many situations, the ultimate object of interest is the APE rather than the coefficient of an explanatory variable.

As in the coefficient estimation, FE estimators of APEs suffer from an incidental parameters problem which becomes negligible at rate T^{-1}. Although simulation results suggest that the magnitude of the bias is less alarming than in the estimation of the coefficients (e.g. see Fernández-Val, 2009), we use the small sample correction outlined in Fernández-Val and Weidner (2016) which again reduces the order of the bias to $O(T^{-2})$. As before, RE estimators of APEs are biased when the distribution of the unobserved is misspecified. Unlike in RE estimation of the coefficients, the misspecification bias does, however, not vanish as T increases, as is again shown in Arellano and Bonhomme (2009) and Arellano and Bonhomme (2011). Therefore, under misspecification of the RE distribution, one should expect different results for RE and FE estimates of APEs, irrespective of the length of the panel.

In summary, RE estimators should be preferred if the distributional assumption is well-justified. If the latter is, however, at least questionable, one should use BCFE estimators, particularly when the panel length is moderate (i.e. $T \geq 8$) and the main interest lies in the APE.

IV. Data and variables

Data set

Our study focuses on Spain, a moderate and slow-growing innovator. The Spanish economy is composed of traditional industries, with an important development of more technologically advanced industries in the recent years (Ministerio de Industria, C. y. T., 2020). Spain has an above average share of non-innovators with potential to innovate, which makes it an interesting case for study in the context of innovation persistence (European Commission, 2021). In 2019, the average expenditure on R&D was 1.14% of the GDP, as compared to the 2.19% of the EU average (European Commission, 2021). Similarly, Spain's overall innovation performance is below that of other EU27 countries scoring 85% in terms of the Summary Innovation Index, an aggregate innovation-performance index reported in the EU innovation scoreboard (European Commission, 2021).

We employ firm-level data from the Spanish Technological Innovation Panel (PITEC), collected on a yearly basis by the Spanish National Institute of Statistics (INE) with the support of the Spanish Foundation for Science and Technology (FECYT). To ensure international comparability, the methodology of the survey and the definition of innovation follow the guidelines of the Oslo Manual (OECD and Eurostat, 1997), which allows the data to serve as input for the CIS. The CIS is considered a reliable tool for the understanding of innovation and is one of the most used data sets in the area of innovation (e.g. Tether, 2002; Miotti and Sachwald, 2003; Belderbos, Carree and Lokshin, 2004; Laursen and Salter, 2006, or more recently, Bianchini *et al*., 2018).

This database has compiled information for a representative sample of over 10,000 firms a year since 2003. The population framework of PITEC is the Central Directory of Companies (DIRCE), which includes Spanish firms located in the national territory. A census is used for the population of firms with more than 200 employees and a stratified sample for firms with less than 200 employees (with internal R&D being the stratum variable). PITEC has a sectoral coverage of agricultural, industrial, construction and service firms, following the NACE-2009 classification. The data are collected via mail, telephone and personal interviews, and covers the whole of the national territory (INE, 2018).[4]

Our analysis includes 10 years of PITEC (corresponding to CIS 2006, CIS 2008, CIS 2010, CIS 2012 and CIS 2014), covering the period of 2005–14.[5] Our sample is restricted to manufacturing firms[6] with an average of at least 10 employees.[7] This yields two panel data sets: an unbalanced data set comprising all firms with at least six successive observations (37,458 firm-year observations corresponding to 4,424 firms), and a balanced sub-sample of the unbalanced data set with firms observed in all time periods included (28,098 firm-year observations corresponding to 3,122 firms).

Variables

Following previous studies on persistence of innovation, we analyse the persistence of innovation in the innovation outcome (e.g. Ganter and Hecker, 2013; Suarez, 2014; Cefis and Marsili, 2015; Tavassoli and Karlsson, 2015). Using the design of PITEC's questionnaire, we distinguish two binary variables, each of them intended to reflect whether the firm has introduced at least one product/process innovation in the period t to $t-2$ (see Table 2 for specific definitions). Following previous studies (e.g. Raymond *et al*., 2010; Tavassoli and Karlsson, 2015), we explain a firm's probability of being an innovator with the lagged product/process innovation experience and a set of

[4]PITEC sampling errors: Coefficient of variation of expenditure on innovation: 0.35%. Coefficient of variation in the number of innovative firms: 1.38%. Coefficient of variation in the number of innovative technology firms: 1.76%. Coefficient of variation in the number of innovative non-technological firms: 1.57%. PITEC non-response rate: 6.83%.

[5]Following previous studies, the first observed sample period is used as the initial condition.

[6]Industry classification codes (NACE Rev. 2): 05 to 43.

[7]Moreover, we also exclude those observations for which incidents in the recording of data are noted (e.g. confidentiality issues, mergers, closures or employment incidents).

TABLE 2

Definition of variables

Variable	*Type*	*Description*
Dependent variables		
Product	B	1 if firm i has introduced a product innovation into the market between year t and year $t-2$, and 0 otherwise. A product innovation is the market introduction of a new or significantly improved good or service with respect to its capabilities, user friendliness, components or sub-systems. It must be new to the firm but not necessarily new to the market.
Process	B	1 if firm i has introduced a process innovation into the market between year t and year $t-2$, and 0 otherwise. A process innovation is the implementation of a new or significantly improved production process, distribution method or supporting activity (excludes organizational innovations). It must be new to the firm but not necessarily new to the market.
Independent variables		
Size	C	Number of employees in firm i in year $t-1$ (in log).
Innovation input	C	Total expenditures in R&D (i.e. expenditures in internal R&D, purchase of external R&D, purchase of machinery, equipment and external knowledge, expenditures on employee training, expenditures on market introduction, design and production of innovations) in firm i in year $t-1$ (in log).
Cooperation	B	1 if firm i cooperated in any innovation activity with either customers, competitors, suppliers or external institutions over the past two years, and 0 otherwise.
Continuous R&D	B	1 if firm i performed internal R&D activities on a continuous basis in year $t-1$, and 0 otherwise.
Exports	C	Amount of exports per employee of firm i in year $t-1$ (in log).
Physical capital	C	Investment in buildings and machineries of firm i in year $t-1$ (in log).
Human capital	C	Percentage of R&D employees with higher education of firm i in year $t-1$.
Demand	C	Sales (in euros) of firm i in year $t-1$ (in log).
Domestic	B	1 if firm i belongs to a group and is a domestic multinational firm in year t, and 0 otherwise.
Foreign	B	1 if firm i belongs to a group and is a foreign multinational firm in year t, and 0 otherwise.
Uninational	B	1 if firm i belongs to a group and is a uninational firm in year t, and 0 otherwise.

Note: B denotes binary variables and C denotes continuous variables.

observable firm characteristics such as innovation input, continuity of R&D activities, physical and human capital, demand-side factors, level of internationalization and cooperation or ownership structure (see Table 2 for exact definitions). We also include time dummies to control for time-specific effects that might affect individual firms' propensity to innovate.[8]

[8]While we do not report industry dummies, including them does not affect our results. However, given that they are approximately constant in our sample, including industry dummies increases the computational burden on our estimators without providing additional information.

TABLE 3

Descriptive statistics (balanced sample)

Variable	*Mean*	*Std. Dev.*			*Min*	*Max*	*Obs.*		
		Overall	*Between*	*Within*			*N*	*n*	*T*
Product	0.649	0.477	0.351	0.323	0	1	28,098	3,122	9
Process	0.633	0.482	0.328	0.353	0	1	28,098	3,122	9
Size	4.456	1.222	1.201	0.228	0	9.234	28,098	3,122	9
Innov. input	9.833	5.442	4.242	3.409	0	19.442	28,098	3,122	9
Cooperation	0.342	0.474	0.354	0.315	0	1	28,098	3,122	9
Continuous R&D	0.553	0.497	0.404	0.290	0	1	28,098	3,122	9
Exports	9.676	6.833	5.655	3.836	0	19.458	28,098	3,122	9
Physical capital	5.881	9.070	6.474	6.354	0	27.392	28,098	3,122	9
Human capital	32.375	33.630	27.456	19.426	0	100	28,098	3,122	9
Demand	16.595	1.591	1.554	0.341	7.928	22.567	28,098	3,122	9
Domestic	0.053	0.225	0.091	0.205	0	1	28,098	3,122	9
Foreign	0.395	0.498	0.410	0.266	0	1	28,098	3,122	9
Uninational	0.008	0.090	0.040	0.081	0	1	28,098	3,122	9

V. Empirical results

In order to allow for a comparison with previous studies, our estimations are primarily focused on the balanced panel data set.

Descriptive statistics

Tables 3 and 4 present the descriptive statistics of the variables used in the estimations.[9] About two thirds of our firm-year observations correspond to firms introducing either product or process innovations. The firms in our sample have on average 207 employees and a volume of sales of 68.5 million Euros; approximately one third of those employees working in the R&D department possess higher education. Firms spend about 1.5 million Euros in R&D, invest 82,300 Euros in physical capital, and their exports amount to 8.9 million Euros per year. About 55% declare to perform innovation on a continuous basis while 34% have cooperated in innovation activities. Finally, around 40% were multinational firms belonging to a group, and about 5% and 1% were domestic multinational and uninational firms belonging to a group respectively.

Table 5 shows the transition probabilities from period $t-1$ to t for both product and process innovation. Generally, we find evidence of strong persistence: about 85% of the non-innovating firms in period $t-1$ remain in that state in period t; similarly, about 90% of firms conducting either product or process innovation in period $t-1$ also innovate in period t. The last column of Table 5 reports the unconditional state dependence (USD), which shows how much of the probability of conducting

[9]Note that for those variables for which we have taken the *log* (e.g. size or exports), the comments included in this subsection refer to the mean of the variable before being transformed.

TABLE 4

Descriptive statistics for the means and the initial conditions (balanced sample)

Variable	*Mean*	*Std. Dev.*	*Min.*	*Max.*	*Obs.*
M_Size$_i$	4.456	1.201	2.035	9.137	28,098
M_Innov. input$_i$	9.833	4.241	0	19.147	28,098
M_Cooperation$_i$	0.342	0.354	0	1	28,098
M_Continuous R&D$_i$	0.553	0.404	0	1	28,098
M_Exports$_i$	9.676	5.655	0	18.689	28,098
M_Physical Cap.$_i$	5.881	6.473	0	26.078	28,098
M_Human Cap.$_i$	32.375	27.452	0	100.067	28,098
M_Demand$_i$	16.595	1.554	12.160	22.371	28,098
M_Domestic$_i$	0.053	0.091	0	0.222	28,098
M_Foreign$_i$	0.395	0.410	0	1	28,098
M_Uninational$_i$	0.008	0.040	0	0.222	28,098
Product$_{i0}$	0.678	0.467	0	1	28,098
Process$_{i0}$	0.668	0.471	0	1	28,098

TABLE 5

Transition probabilities (balanced sample)

		Innovator in year t + *1*		
Innovator in year t		*NO*	*YES*	*USD*
Product	NO	85.84%	14.16%	
	YES	9.22%	90.78%	76.62
Process	NO	84.40%	15.60%	
	YES	11.31%	88.69%	73.09

Note: USD, Unconditional State Dependence. Obs.: 28,098.

innovation in year t can be explained by the difference between being innovator or non-innovator in year $t-1$.[10] This can be expressed as:

$$USD = P(Y_{it} = 1 | Y_{it-1} = 1) - P(Y_{it} = 1 | Y_{it-1} = 0).$$

Table 5 shows that the probability of conducting product innovation in year t is about 76pp higher for firms that also reported product innovation in year $t-1$ as compared to those that did not report any product innovation activities in year $t-1$. Similarly, for process innovation, the probability of being innovative in year t is 73pp higher for firms that also conducted process innovation in year $t-1$ as compared to firms that did not conduct process innovation in year $t-1$.

While overall Table 5 indicates a pattern of strong persistence in our sample, it does not provide any information on the nature of this persistence. The next section aims at distinguishing true from spurious state dependence, considering a set of models that allow us to control for observed and unobserved firm characteristics.

[10]Note that unconditional or observed state dependence does not condition on any observed or unobserved characteristics of the firm.

Estimation results

Tables 6 and 7 report the estimation results for product and process innovation respectively. For both tables, odd-numbered columns present a simple specification in which only the lag of innovation and time dummies are included as explanatory variables, while even-numbered columns include the full set of covariates. Columns (1) and (2) show the results for the pooled probit model with the full balanced sample (i.e. including all firms that do or do not change innovator status over time) and columns (3) and (4) present the RE specification over the full balanced sample. For the sub-sample of firms with time-varying innovator status, columns (5) and (6) report the FE specification, columns (7) and (8) the pooled probit and columns (9) and (10) the RE specification.

For all the specifications in Table 6, the coefficient for the lag variable of product innovation is positive and significant at a 1% level, meaning that firms introducing product innovations in year $t-1$ are more likely to introduce product innovations in year t. As expected, controlling for firms' observable characteristics reduces the magnitude of the lag coefficient.

In addition to previous product innovation activities, we also find that the different specifications confirm some observable characteristics as explanatory factors of firms' product innovation behaviour. Firms' investment in R&D and continuous R&D activities in year $t-1$ increase firms' probability of introducing product innovation in year t. Moreover, firms belonging to domestic multinational groups have a higher probability of introducing product innovations in year t. These results are significant across the different specifications. Similarly, R&D input is found to be a significant determinant of process innovation, as shown in Table 7.

In all RE specifications, the initial conditions and the means of the explanatory variables employed are jointly significant,[11] indicating that there exists a correlation between the unobserved heterogeneity and the independent variables. This underlines the importance of accounting for the presence of unobserved heterogeneity, since failure to do so may lead to biased parameters estimates.

Table 8 shows the APEs for lagged product and process innovation, which correspond to the average level of true state dependence in persistence of product and process innovation. Each of the columns corresponds to the respective specifications in Tables 6 and 7. From the policy point of view, the APEs are often the ultimate object of interest. Unlike the parameter estimates, the APEs provide the magnitude of the average effect of changes in the regressors on the response probabilities. While the estimated APEs in the pooled probit and the RE specifications are similar, there are substantial differences to the APE estimates based on the FE specification.[12]

[11]For product innovation (Table 6, columns 4 and 10), the χ^2 statistics are 2250.50 and 40.74, and for process innovation (Table 7, columns 4 and 10) 132.97 and 47.25, all of them corresponding to P-values of 0.000.

[12]The similarities between the pooled probit and the RE estimates are not surprising, given that we find low levels of intra-class correlation (ICC), which is the proportion of the total variance contributed by the panel-level variance component. A low level of ICC is sometimes regarded as proof against the presence of unobserved heterogeneity. However, this is not correct for at least two reasons: first, the level of ICC depends on the variance that is imposed on the model error in order to identify the parameters. In probit, the latter is typically set to one. Second, the ICC is calculated based on the RE specification, which by design does not allow for certain types of variation (e.g. heteroscedasticity is ruled out). This may lead to low levels of the panel-level variance by design.

TABLE 6

Balanced product innovation

Variables	*(1) Pooled*	*(2) Pooled*	*(3) RE*	*(4) RE*	*(5) BCFE*	*(6) BCFE*	*(7) Pooled*	*(8) Pooled*	*(9) RE*	*(10) RE*
Product_{it-1}	2.456***	2.229***	2.235***	2.108***	1.702***	1.611***	1.768***	1.667***	1.797***	1.701***
	(0.022)	(0.023)	(0.031)	(0.031)	(0.029)	(0.030)	(0.025)	(0.026)	(0.026)	(0.026)
Product_{i0}			0.389***	0.162***					−0.100***	−0.132***
			(0.036)	(0.030)					(0.026)	(0.027)
Size_{it-1}		−0.019		0.100*		0.098		−0.029		0.116*
		(0.018)		(0.056)		(0.067)		(0.021)		(0.061)
Innov. input_{it-1}		0.046***		0.030***		0.098		0.040***		0.039***
		(0.003)		(0.003)		(0.067)		(0.003)		(0.004)
$\text{Cooperation}_{it-1}$		0.043*		−0.046		0.035		−0.006		−0.018
		(0.024)		(0.034)		(0.040)		(0.027)		(0.037)
Continuous $\text{R\&D}_{it-1}$		0.195***		0.061		0.107**		0.082***		0.085**
		(0.028)		(0.039)		(0.046)		(0.032)		(0.042)
Exports_{it-1}		0.005***		0.004		0.005*		0.002		0.006*
		(0.002)		(0.003)		(0.003)		(0.002)		(0.003)
Physical Cap._{it-1}		0.003**		0.002		0.003*		0.003*		0.003
		(0.001)		(0.002)		(0.002)		(0.002)		(0.002)
Human Cap._{it-1}		0.001		0.000		0.000		0.000		0.000
		(0.000)		(0.001)		(0.001)		(0.000)		(0.001)
Demand_{it-1}		0.006		−0.022		−0.028		0.005		−0.033
		(0.015)		(0.037)		(0.043)		(0.016)		(0.040)
Domestic_{it}		0.131**		0.170**		0.256***		0.167***		0.256***
		(0.055)		(0.077)		(0.092)		(0.064)		(0.087)
Foreign_{it}		−0.026		−0.018		0.035		−0.037		0.024
		(0.026)		(0.057)		(0.067)		(0.030)		(0.062)
Uninational_{it}		−0.115		−0.209		−0.225		−0.163		−0.216
		(0.109)		(0.135)		(0.153)		(0.118)		(0.147)
M_Size_{i}				−0.129**						−0.167**
				(0.060)						(0.065)

(*Continued*)

TABLE 6
(*Continued*)

Variables	*(1) Pooled*	*(2) Pooled*	*(3) RE*	*(4) RE*	*(5) BCFE*	*(6) BCFE*	*(7) Pooled*	*(8) Pooled*	*(9) RE*	*(10) RE*
M_Innov. input$_i$				0.051***						0.003
				(0.006)						(0.007)
M_Cooperation$_i$				0.130***						0.039
				(0.050)						(0.056)
M_Continuous R&D$_i$				0.141**						0.004
				(0.061)						(0.067)
M_Exports$_i$				−0.001						−0.005
				(0.004)						(0.004)
M_Physical Cap.$_i$				−0.002						−0.001
				(0.003)						(0.003)
M_Human Cap.$_i$				−0.000						−0.000
				(0.001)						(0.001)
M_Demand$_i$				0.025						0.049
				(0.040)						(0.044)
M_Domestic$_i$				−0.272**						−0.275**
				(0.121)						(0.135)
M_Foreign$_i$				0.019						−0.047
				(0.070)						(0.076)
M_Uninational$_i$				0.668***						0.308
				(0.254)						(0.271)
Observations	28,098	28,098	28,098	28,098	15,957	15,957	15,957	15,957	15,957	15,957
Number of firms	3,122	3,122	3,122	3,122	1,773	1,773	1,773	1,773	1,773	1,773
Firm FE	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO
Firm RE	NO	NO	YES	YES	NO	NO	NO	NO	YES	YES
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Log likelihood	−9,304	−8,765	−9,226	−8,634	−6,551	−6,361	−7,302	−7,088	−7,295	−7,067
Chi-square	13,323	12,311	7,909	9,091	8,329	8,709	5,658	5,611	6,124	6,579

Notes: Robust standard errors in parentheses. ***P<0.01, **P<0.05, *P<0.1

TABLE 7

Balanced process innovation

Variables	*(1) Pooled*	*(2) Pooled*	*(3) RE*	*(4) RE*	*(5) BCFE*	*(6) BCFE*	*(7) Pooled*	*(8) Pooled*	*(9) RE*	*(10) RE*
Process_{it-1}	2.273***	2.135***	2.173***	2.064***	1.698***	1.592***	1.754***	1.670***	1.789***	1.700***
	(0.021)	(0.021)	(0.028)	(0.028)	(0.027)	(0.028)	(0.023)	(0.024)	(0.024)	(0.024)
Process_{i0}			0.171***	0.103***					−0.148***	−0.146***
			(0.028)	(0.027)					(0.024)	(0.024)
Size_{it-1}		0.025		0.036		0.032		0.008		0.054
		(0.017)		(0.054)		(0.064)		(0.020)		(0.059)
Innov. input_{it-1}		0.052***		0.040***		0.062***		0.051***		0.051***
		(0.003)		(0.003)		(0.004)		(0.003)		(0.004)
$\text{Cooperation}_{it-1}$		0.111***		−0.051		0.062		0.043*		−0.004
		(0.022)		(0.032)		(0.038)		(0.025)		(0.035)
Continuous $\text{R\&D}_{it-1}$		0.015		0.026		0.003		−0.075**		0.009
		(0.028)		(0.038)		(0.044)		(0.031)		(0.041)
Exports_{it-1}		−0.001		0.002		0.002		−0.004**		0.002
		(0.002)		(0.003)		(0.003)		(0.002)		(0.003)
Physical Cap._{it-1}		0.002		0.003		0.003		0.000		0.003*
		(0.001)		(0.002)		(0.002)		(0.001)		(0.002)
Human Cap._{it-1}		−0.002***		−0.001		−0.001		−0.002***		−0.001
		(0.000)		(0.001)		(0.001)		(0.000)		(0.001)
Demand_{it-1}		0.016		0.046		0.090**		0.012		0.066*
		(0.014)		(0.034)		(0.041)		(0.016)		(0.038)
Domestic_{it}		0.059		0.133*		0.147*		0.042		0.142*
		(0.054)		(0.072)		(0.084)		(0.059)		(0.078)
Foreign_{it}		0.026		0.090*		0.108*		0.024		0.114**
		(0.025)		(0.053)		(0.063)		(0.028)		(0.057)
Uninational_{it}		−0.122		−0.010		0.113		−0.018		0.044
		(0.116)		(0.126)		(0.142)		(0.120)		(0.134)
M_Size_{i}				−0.002						−0.046
				(0.057)						(0.063)

(*Continued*)

TABLE 7

(*Continued*)

Variables	*(1) Pooled*	*(2) Pooled*	*(3) RE*	*(4) RE*	*(5) BCFE*	*(6) BCFE*	*(7) Pooled*	*(8) Pooled*	*(9) RE*	*(10) RE*
M_Innov. input$_i$				0.036***						0.001
				(0.006)						(0.006)
M_Cooperation$_i$				0.333***						0.132**
				(0.048)						(0.052)
M_Continuous R&D$_i$				−0.099*						−0.134**
				(0.060)						(0.064)
M_Exports$_i$				−0.007**						−0.008**
				(0.003)						(0.004)
M_Physical Cap.$_i$				−0.004						−0.006**
				(0.003)						(0.003)
M_Human Cap.$_i$				−0.003***						−0.001
				(0.001)						(0.001)
M_Demand$_i$				−0.041						−0.056
				(0.038)						(0.042)
M_Domestic$_i$				−0.180						−0.169
				(0.116)						(0.125)
M_Foreign$_i$				−0.080						−0.088
				(0.065)						(0.070)
M_Uninational$_i$				−0.391						−0.313
				(0.240)						(0.244)
Observations	28,098	28,098	28,098	28,098	18,522	18,522	18,522	18,522	18,522	18,522
Number of firms	3,122	3,122	3,122	3,122	2,058	2,058	2,058	2,058	2,058	2,058
Firm FE	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO
Firm RE	NO	NO	YES	YES	NO	NO	NO	NO	YES	YES
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Log likelihood	−10,288	−9,867	−10,267	−9,794	−7,560	−7,297	−8,455	−8,223	−8,436	−8,180
Chi-square	12,536	11,910	8,910	9,856	9,876	10,402	6,643	6,567	7,466	7,978

Notes: Robust standard errors in parentheses. ***P<0.01, **P<0.05, *P<0.1

TABLE 8

Average partial effect (APE) (Balanced sample)

	(1) Pooled	(2) Pooled	(3) RE	(4) RE	(5) BCFE	(6) BCFE	(7) Pooled	(8) Pooled	(9) RE	(10) RE
Product	0.762	0.684	0.686	0.636	0.342	0.319	0.598	0.555	0.605	0.564
Process	0.723	0.669	0.686	0.626	0.392	0.360	0.590	0.553	0.599	0.559

For the full balanced sample (columns 1 to 4), the APEs of the pooled probit and RE models are between 0.63 and 0.76. This means that the probability of introducing product or process innovation in t is between 63PP and 76PP higher for firms that are innovators in $t-1$ as compared to firms that are non-innovators in $t-1$. Moreover, between 82% (column 4) and 100% (column 1) of persistence of innovation is attributed to true state dependence, while observed and unobserved characteristics explain less than 18% of the persistence of both product and process innovation.[13] For the sub-sample of firms with time-varying innovator status (columns 7 to 10), the estimated APEs are reduced to 0.55–0.60. In this case, the probability of introducing product or process innovations in t is between 55PP and 60PP higher for firms that were innovators in $t-1$ than for non-innovators in $t-1$, after controlling for observed and unobserved firm characteristics. These APE estimates also suggest that about 70% to 78% of persistence correspond to true state dependence.

The estimated APEs in the FE specification range between 0.31 and 0.36. As compared to the pooled probit and the RE specification, this translates into a smaller difference in the probability of introducing product or process innovation in t between innovators and non-innovators in $t-1$. The probability of introducing product or process innovations in year t, controlling for firms' observed and unobserved characteristics, is around 31PP–36PP higher for firms that innovate in $t-1$ as compared to non-innovating firms. Moreover, the APEs of the FE allocate only between 41% and 53% of innovation persistence to true state dependence.

Robustness checks and further analysis

In this section we address potential concerns about the validity of our results.

First of all, our main results, following previous studies, are based on a balanced panel. This can raise doubts regarding the representativeness of our sample, since only firms that have been active for 10 consecutive years are represented. In the appendix, we report the results using an unbalanced panel data set (see Tables A1–A5). As compared to the balanced data set, the unbalanced data set contains more observations and is thus more representative of the population of firms, as it is less subject to survivorship bias (Raymond *et al*., 2010). The results based on the unbalanced data set

[13]The proportion of persistence attributed to true state dependence is calculated as the ratio of the APE over the USD (as computed in Table 5).

are not substantially different from the main results using the balanced data set (see Tables A3 and A4).

For product innovation, we find that the APE estimates of the pooled probit and the RE model are between 0.55 and 0.78. Moreover, estimates from these models attribute between 71% and 100% of the persistence to true state dependence. In comparison, estimates from the FE model allocate a much lower weight to true state dependence, with APE estimates ranging between 0.31 and 0.34 (see Table A5). For process innovation, we find similar results, with pooled probit and RE models estimating the probability of conducting innovation in year t to be between 55PP and 74PP higher for firms that conducted process innovation in $t-1$ as compared to those that did not; this difference in probabilities, however, is between 34PP and 38PP for the FE specification. As with product innovation, the proportion of persistence attributed to true state dependence and the difference in the probability of introducing product or process innovation in t between innovators and non-innovators in $t-1$ is substantially smaller for the FE estimates as compared to the pooled or RE estimations.

Furthermore, we re-estimated the FE specifications distinguishing industries according to the level of technology intensity (see Tables A6, A7 and A8). We follow the OECD (2011) classification and distinguish between low-tech, medium-tech and high-tech industries. We find that the APE estimates are higher the lower the technological regime. In particular, the probability of conducting innovation in year t is between 33PP and 39PP higher for low-tech firms that conducted process innovation in $t-1$ as compared to those that did not, while for high-tech firms this difference in probabilities is between 25PP and 31PP.

While our independent variables are measured on a yearly basis, the dependent variable is assessed on blocks of 3 years. This is an unfortunate feature of the CIS: while measuring items such as firms' size or investment in R&D on a yearly basis, it asks firms to declare innovation activities of the past 2 years. For example, the CIS of 2009 will cover the innovation activities of 2009, 2008 and 2007; the CIS of 2008 will cover the innovation activities of 2008, 2007 and 2006. Thus, by including consecutive waves of the CIS in our sample, we are generating an overlap in the reporting of firms' innovative activities that might induce an upward bias in the measures of persistence. This is true independent of the estimation method employed. Hence, it might be that our APEs are overestimated, so that the actual true state dependence is lower than the one suggested by our results. In order to assess the magnitude of this bias, we re-estimate our results using non-overlapping waves of the CIS (CIS in 2014, 2011, 2008 and 2005; and CIS in 2013, 2010, 2007 and 2004). We restrict our robustness check to the pooled probit, as this specification is not as data demanding as the RE (which requires an extra year of data for the initial condition) and the FE (which does not perform well when T is small due to the incidental parameters bias). The APE estimates based on the non-overlapping samples are around 0.50 (about 0.28 lower than estimates based on the corresponding overlapping sample), which suggests the existence of an upwards bias in the APE estimates in our main results. This, if anything, further reinforces our main finding that true state dependence in persistence of innovation may be overestimated.

VI. Discussion and conclusion

Our paper explores the existence and the nature of persistence of innovation at the firm level, making use of the Spanish CIS for the period 2005–14. The aim of the paper is to investigate whether innovation activity is persistent over time and, if so, to which extent this persistence can be traced back to spurious or true state dependence. To determine the nature of persistence of innovation, our paper employs a FE approach. This constitutes a novel methodology in the persistence literature that complements previous studies using RE, as it accounts for arbitrary time-invariant unobserved individual heterogeneity across firms.

Our results show the existence of a high level of persistence of innovation, confirming to a large extent findings of earlier studies (e.g. Martínez-Ros and Labeaga, 2009; Peters, 2009; Huergo and Moreno, 2011; Triguero and Córcoles, 2013; Cefis and Marsili, 2015). They also indicate that process and product innovation display similar levels of persistence; this is in contrast to previous literature that confirms innovation persistence in product innovation but finds weaker evidence of persistence in process innovation (e.g. Ganter and Hecker, 2013; Hecker and Ganter, 2014; Tavassoli and Karlsson, 2015). Regarding the nature of this persistence, our approach allows us to distinguish between true state dependence (i.e. when past experience in innovation has a structural impact on the probability of conducting innovation in the future) and spurious state dependence (i.e. when firms' observed and unobserved characteristics determine the probability to innovate, yet the effect is falsely attributed to past experience in innovation) without imposing distributional assumptions on the unobserved heterogeneity. As compared to the RE approach used in previous literature, our study attributes a much smaller fraction of the persistence to true state dependence, suggesting a more modest structural effect of experiencing innovation on future innovation activity. Our results thus suggest that distributional assumptions on the unobserved effects may have a substantial effect on the resulting estimates. Consequently, we recommend to critically assess the plausibility of distributional assumptions.

Our findings suggest that besides relevant theories such as success-breeds-success, learning by doing or sunk costs, additional literature should also be considered in the context of persistence of innovation. For instance, firm characteristics such as managerial talent, research ability, organizational culture and organizational routines shape firms' technological and organizational capabilities (Cefis and Orsenigo, 2001). The latter, in turn, determine firms' long-term strategies, which affect decisions such as the establishment of R&D labs or the level of innovative activities (Nelson and Winter, 1982; Teece *et al*., 1997; Clausen *et al*., 2011).

From a policy point of view, high levels of true state dependence are convenient, as they imply that programs aiming at fostering initial innovations have a long-term effect on firms' innovation activity. Our results however suggest that unobserved firm-specific characteristics may play a substantial role, thus making optimal policies more complex. While we do not discourage programs that mainly aim at establishing initial innovations (as we do find considerable levels of true state dependence), policy makers may need to consider targeting firm characteristics such as organizational culture or routines, which is a more demanding objective as these characteristics are heavily embedded in the organization and are very difficult to change (Helfat, 1994; Stuart and Podolny, 1996; Clausen *et al*., 2011).

Finally, from the empirical point of view, our paper highlights the importance of taking unobserved individual heterogeneity into consideration when assessing the nature of persistence. While Wooldridge's approach yields an estimator of convenient simplicity that accounts for *some* unobserved heterogeneity, recent advances in the treatment of nonlinear dynamic models allow researchers to take into account *any* type of unobserved heterogeneity with arbitrary dependence on the observed characteristics. BCFE approaches are implemented in popular statistical software packages, applied researchers now have additional tools at their disposal that allow for more robustness without sacrificing computational simplicity.

Limitations and further research

As any, our paper is not free of limitations. Although the CIS has proven to be a very good source of information on firms' innovation activity, there are several limitations that arise from the design of the questionnaire. First, our study measures innovation persistence in terms of the frequency of innovation activities. Thus, future research may explore the depth of these innovations (e.g. how many types of product and process innovations have been introduced) to enrich the analysis of persistence.

Second, our study focuses on firms that have at least 10 employees. While in the context of the PITEC sample, the number of firms discarded is very small (less than 5% of the sample), the so-called 'micro-firms' (between one and nine employees) represent about 40% of all companies in Spain (Ministerio de Industria, C. y. T., 2020). In this respect, the results from our study need to be interpreted with care, as there is a large proportion of Spanish firms which, while conducting little or no innovation (Mulet-Melia, 2020), is not represented in our sample. Thus, further research may be needed to investigate the phenomenon of innovation persistence in this subset of small firms.

Third, our analysis is restricted to the Spanish context. As explained above, Spain is currently considered a moderate and slow-growing innovator; future studies should aim at also re-investigating the issue of innovation persistence in other low and high innovating countries to have a more complete picture. In fact, future research could follow some existing studies in the field of innovation (e.g. Bianchini, Bottazzi and Tamagni, 2017; Cirillo, Sostero and Tamagni, 2017) and provide a cross-country comparison.

Fourth, while simulation evidence suggests that BCFE estimators yield reliable estimates in moderately long panels as considered here, future studies may exploit the availability of longer panel data sets to gain additional robustness to the incidental parameters problem.

Finally, our study focuses on the measurement of innovation persistence, focusing on the outcomes (product and process innovation). Future studies could also explore the question of innovation persistence in other outcomes such as organizational or marketing innovation. Moreover, future research might want to also study how our results impact research on the role of innovation persistence in some key indicators such as firm growth (Guarascio and Tamagni, 2019; Bianchini *et al*., 2018), employment creation (Triguero, Córcoles and Cuerva, 2014) or firms' performance (Cefis and Ciccarelli, 2005).

Appendix: A. On the estimation of true state dependence in the persistence of innovation

TABLE A1

Descriptive statistics (unbalanced sample)

	Mean	*Std. Dev.*			*Min*	*Max*	*Obs.*				
Variable		*Overall*	*Between*	*Within*			*N*	*n*	*T avg.*	*T min.*	*T max.*
Product	0.584	0.493	0.375	0.323	0	1	37,458	4,424	8.47	5	9
Process	0.577	0.494	0.352	0.350	0	1	37,458	4,424	8.47	5	9
Size	4.214	1.255	1.226	0.263	0	9.234	37,458	4,424	8.47	5	9
Innov. input	8.485	6.008	4.904	3.531	0	19.442	37,458	4,424	8.47	5	9
Cooperation	0.291	0.454	0.342	0.297	0	1	37,458	4,424	8.47	5	9
Continuous R&D	0.466	0.499	0.413	0.280	0	1	37,458	4,424	8.47	5	9
Exports	8.759	7.037	5.967	3.780	0	19.458	37,458	4,424	8.47	5	9
Physical capital	4.913	8.562	6.118	5.930	0	27.392	37,458	4,424	8.47	5	9
Human capital	27.417	33.026	27.102	18.723	0	100	37,458	4,424	8.47	5	9
Demand	16.259	1.653	1.610	0.378	0.693	22.567	37,458	4,424	8.47	5	9
Domestic	0.048	0.214	0.093	0.193	0	1	37,458	4,424	8.47	5	9
Foreign	0.337	0.473	0.399	0.250	0	1	37,458	4,424	8.47	5	9
Uninational	0.009	0.092	0.044	0.082	0	1	37,458	4,424	8.47	5	9

TABLE A2

Transition probabilities (unbalanced sample)

		Innovator in year t+1		
Innovator in year t		*NO*	*YES*	*USD*
Product	NO	88.96%	11.04%	
	YES	10.87%	89.13%	78.09
Process	NO	87.42%	12.58%	
	YES	12.78%	87.22%	74.64

Note: USD, Unconditional State Dependence. Obs.: 37,458.

TABLE A3

Unbalanced product innovation

Variables	*(1) Pooled*	*(2) Pooled*	*(3) RE*	*(4) RE*	*(5) BCFE*	*(6) BCFE*	*(7) Pooled*	*(8) Pooled*	*(9) RE*	*(10) RE*
Product_{it-1}	2.527***	2.244***	2.280***	2.130***	1.734***	1.622***	1.792***	1.677***	1.842***	1.724***
	(0.019)	(0.020)	(0.028)	(0.027)	(0.026)	(0.027)	(0.022)	(0.023)	(0.023)	(0.024)
Product_{i0}			0.406***	0.115***					−0.184***	−0.190***
			(0.034)	(0.027)					(0.023)	(0.024)
Size_{it-1}		−0.001		0.103**		0.148**		−0.022		0.157***
		(0.016)		(0.047)		(0.058)		(0.019)		(0.053)
Innov. input_{it-1}		0.058***		0.036***		0.057***		0.051***		0.048***
		(0.002)		(0.003)		(0.003)		(0.003)		(0.003)
$\text{Cooperation}_{it-1}$		0.038*		−0.050		0.041		−0.000		−0.016
		(0.022)		(0.031)		(0.037)		(0.025)		(0.034)
Continuous $\text{R\&D}_{it-1}$		0.198***		0.063*		0.095**		0.082***		0.080**
		(0.026)		(0.035)		(0.042)		(0.029)		(0.038)
Exports_{it-1}		0.007***		0.004		0.005*		0.002		0.005*
		(0.001)		(0.002)		(0.003)		(0.002)		(0.003)
Physical Cap._{it-1}		0.003**		0.003**		0.004**		0.002		0.004**
		(0.001)		(0.002)		(0.002)		(0.001)		(0.002)
Human Cap._{it-1}		0.001		0.000		−0.000		−0.000		0.000
		(0.000)		(0.001)		(0.000)		(0.000)		(0.001)
Demand_{it-1}		0.023*		0.042		0.048		0.020		0.040
		(0.013)		(0.031)		(0.037)		(0.015)		(0.034)
Domestic_{it}		0.087*		0.123*		0.216**		0.094		0.210***
		(0.051)		(0.070)		(0.085)		(0.059)		(0.079)
Foreign_{it}		−0.032		−0.006		0.068		−0.035		0.047
		(0.024)		(0.052)		(0.062)		(0.028)		(0.057)
Uninational_{it}		−0.145		−0.277**		−0.306**		−0.218**		−0.282**
		(0.100)		(0.115)		(0.138)		(0.111)		(0.131)

(*Continued*)

TABLE A3

(*Continued*)

Variables	*(1) Pooled*	*(2) Pooled*	*(3) RE*	*(4) RE*	*(5) BCFE*	*(6) BCFE*	*(7) Pooled*	*(8) Pooled*	*(9) RE*	*(10) RE*
M_Size$_i$				−0.115**						−0.197***
				(0.051)						(0.057)
M_Innov. input$_i$				0.066***						0.005
				(0.005)						(0.006)
M_Cooperation$_i$				0.121***						0.054
				(0.046)						(0.051)
M_Continuous R&D$_i$				0.117**						0.012
				(0.056)						(0.062)
M_Exports$_i$				0.001						−0.003
				(0.003)						(0.003)
M_Physical Cap.$_i$				−0.005**						−0.004
				(0.003)						(0.003)
M_Human Cap.$_i$				−0.000						−0.000
				(0.001)						(0.001)
M_Demand$_i$				−0.038						−0.021
				(0.035)						(0.038)
M_Domestic$_i$				−0.316***						−0.327***
				(0.110)						(0.123)
M_Foreign$_i$				0.007						−0.062
				(0.064)						(0.071)
M_Uninational$_i$				0.724***						0.380
				(0.216)						(0.238)
Observations	37,458	37,458	37,458	37,458	20,990	20,990	20,990	20,990	20,990	20,990
Firm FE	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO
Firm RE	NO	NO	YES	YES	NO	NO	NO	NO	YES	YES
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Log likelihood	−12,442	−11,317	−12,355	−11,111	−8,327	−7,951	−9,497	−9,043	−9,464	−8,991
Chi-square	18,560	16,702	11,289	13,777	11,832	12,585	7,470	7,430	8,657	9,602

Notes: Robust standard errors in parentheses. ***P<0.01, **P<0.05, *P<0.1

TABLE A4

Unbalanced process innovation

Variables	*(1) Pooled*	*(2) Pooled*	*(3) RE*	*(4) RE*	*(5) BCFE*	*(6) BCFE*	*(7) Pooled*	*(8) Pooled*	*(9) RE*	*(10) RE*
Process_{it-1}	2.352***	2.160***	2.221***	2.091***	1.719***	1.588***	1.767***	1.668***	1.817***	1.712***
	(0.018)	(0.019)	(0.025)	(0.025)	(0.024)	(0.025)	(0.021)	(0.021)	(0.021)	(0.022)
Process_{i0}			0.199***	0.076***					−0.207***	−0.198***
			(0.027)	(0.024)					(0.021)	(0.022)
Size_{it-1}		0.047***		0.060		0.080		0.018		0.085*
		(0.015)		(0.045)		(0.055)		(0.018)		(0.050)
Innov. input_{it-1}		0.060***		0.044***		0.069***		0.058***		0.057***
		(0.002)		(0.003)		(0.003)		(0.003)		(0.003)
$\text{Cooperation}_{it-1}$		0.099***		−0.070**		0.060*		0.033		−0.018
		(0.021)		(0.029)		(0.035)		(0.023)		(0.032)
Continuous $\text{R\&D}_{it-1}$		0.004		0.022		−0.000		−0.077***		0.013
		(0.025)		(0.034)		(0.040)		(0.028)		(0.037)
Exports_{it-1}		−0.001		0.002		0.002		−0.004**		0.001
		(0.001)		(0.002)		(0.003)		(0.002)		(0.003)
Physical Cap._{it-1}		0.002		0.002		0.002		−0.000		0.002
		(0.001)		(0.002)		(0.002)		(0.001)		(0.002)
Human Cap._{it-1}		−0.002***		−0.001		−0.001**		−0.002***		−0.001
		(0.000)		(0.000)		(0.001)		(0.000)		(0.001)
Demand_{it-1}		0.024**		0.069**		0.120***		0.017		0.097***
		(0.012)		(0.030)		(0.036)		(0.014)		(0.033)
Domestic_{it}		0.041		0.107		0.090		−0.021		0.097
		(0.049)		(0.065)		(0.077)		(0.054)		(0.071)
Foreign_{it}		0.031		0.114**		0.152***		0.036		0.149***
		(0.023)		(0.049)		(0.058)		(0.026)		(0.053)
Uninational_{it}		−0.119		−0.036		0.026		−0.023		0.005
		(0.101)		(0.109)		(0.129)		(0.111)		(0.120)

(*Continued*)

TABLE A4

(*Continued*)

Variables	*(1) Pooled*	*(2) Pooled*	*(3) RE*	*(4) RE*	*(5) BCFE*	*(6) BCFE*	*(7) Pooled*	*(8) Pooled*	*(9) RE*	*(10) RE*
M_Size$_i$				−0.005						−0.066
				(0.049)						(0.054)
M_Innov. input$_i$				0.052***						0.002
				(0.005)						(0.005)
M_Cooperation$_i$				0.347***						0.140***
				(0.044)						(0.048)
M_Continuous R&D$_i$				−0.162***						−0.154***
				(0.055)						(0.059)
M_Exports$_i$				−0.007**						−0.006*
				(0.003)						(0.003)
M_Physical Cap.$_i$				−0.004*						−0.005**
				(0.002)						(0.003)
M_Human Cap.$_i$				−0.003***						−0.001
				(0.001)						(0.001)
M_Demand$_i$				−0.060*						−0.085**
				(0.033)						(0.037)
M_Domestic$_i$				−0.164						−0.150
				(0.105)						(0.113)
M_Foreign$_i$				−0.111*						−0.119*
				(0.060)						(0.064)
M_Uninational$_i$				−0.294						−0.104
				(0.203)						(0.216)
Observations	37,458	37,458	37,458	37,458	24,043	24,043	24,043	24,043	24,043	24,043
Firm FE	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO
Firm RE	NO	NO	YES	YES	NO	NO	NO	NO	YES	YES
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Log likelihood	−13,614	−12,708	−13,583	−12,584	−9,544	−9,083	−10,869	−10,421	−10,821	−10,344
Chi-square	17,458	16,323	12,481	14,478	13,572	14,495	8,621	8,575	10,225	11,179

Notes: Robust standard errors in parentheses. ***P<0.01, **P<0.05, *P<0.1.

TABLE A5

Average partial effect (APE) (Unbalanced sample)

	(1) Pooled	(2) Pooled	(3) RE	(4) RE	(5) BCFE	(6) BCFE	(7) Pooled	(8) Pooled	(9) RE	(10) RE
Product	0.778	0.659	0.686	0.594	0.335	0.306	0.599	0.543	0.611	0.555
Process	0.740	0.652	0.690	0.595	0.377	0.339	0.587	0.539	0.599	0.549

TABLE A6

Product Innovation: industry splits (fixed effects regression)

Variables	(1) High tech	(2) High tech	(3) Medium tech	(4) Medium tech	(5) Low tech	(6) Low tech
Product_{it-1}	1.720***	1.642***	1.682***	1.589***	1.711***	1.618***
	(0.123)	(0.127)	(0.053)	(0.054)	(0.037)	(0.038)
Size_{it-1}		0.352		−0.029		0.146*
		(0.284)		(0.133)		(0.083)
Innov. input_{it-1}		0.006		0.051***		0.047***
		(0.023)		(0.008)		(0.005)
$\text{Cooperation}_{it-1}$		0.295**		0.091		−0.024
		(0.155)		(0.074)		(0.051)
Continuous $\text{R\&D}_{it-1}$		−0.033		0.136		0.104**
		(0.215)		(0.084)		(0.057)
Exports_{it-1}		−0.006		0.007		0.005
		(0.014)		(0.006)		(0.004)
Physical Cap._{it-1}		0.007		0.002		0.004
		(0.008)		(0.004)		(0.003)
Human Cap._{it-1}		0.005		0.000		0.000
		(0.003)		(0.001)		(0.001)
Demand_{it-1}		0.083		0.051		−0.078
		(0.172)		(0.089)		(0.053)
Domestic_{it}		−0.099		0.334**		0.268**
		(0.317)		(0.172)		(0.117)
Foreign_{it}		−0.081		−0.050		0.088
		(0.228)		(0.120)		(0.087)
Uninational_{it}		−4.930		0.190		−0.230
		(8.978)		(0.348)		(0.180)
Observations	1,026	1,026	5,157	5,157	9,774	9,774
Firm FE	YES	YES	YES	YES	YES	YES
Year dummies	YES	YES	YES	YES	YES	YES
Log likelihood	−408	−392	−2,071	−2,007	−4,061	−3,934
Chi-square	539	570	2,542	2,670	5,195	5,448

Notes: Robust standard errors in parentheses. ***P<0.01, **P<0.05, *P<0.1

TABLE A7

Process innovation: industry splits (fixed effects regression)

Variables	(1) High tech	(2) High tech	(3) Medium tech	(4) Medium tech	(5) Low tech	(6) Low tech
Process$_{it-1}$	1.641***	1.591***	1.731***	1.662***	1.690***	1.556***
	(0.107)	(0.110)	(0.047)	(0.048)	(0.035)	(0.036)
Size$_{it-1}$		−0.053		0.235**		−0.065
		(0.190)		(0.117)		(0.083)
Innov. input$_{it-1}$		0.102***		0.052***		0.064***
		(0.025)		(0.008)		(0.005)
Cooperation$_{it-1}$		0.061		0.026		0.086*
		(0.144)		(0.066)		(0.050)
Continuous R&D$_{it-1}$		−0.428**		0.111		−0.021
		(0.208)		(0.077)		(0.056)
Exports$_{it-1}$		−0.010		−0.005		0.008*
		(0.013)		(0.005)		(0.004)
Physical Cap.$_{it-1}$		−0.009		0.003		0.005
		(0.007)		(0.003)		(0.003)
Human Cap.$_{it-1}$		0.000		−0.001		−0.001
		(0.003)		(0.001)		(0.001)
Demand$_{it-1}$		0.104		0.001		0.128**
		(0.131)		(0.080)		(0.052)
Domestic$_{it}$		0.151		0.206		0.119
		(0.287)		(0.151)		(0.110)
Foreign$_{it}$		0.161		0.202*		0.048
		(0.208)		(0.111)		(0.083)
Uninational$_{it}$		−0.028		0.205		0.108
		(0.527)		(0.324)		(0.167)
Observations	1,233	1,233	6,246	6,246	11,043	11,043
Firm FE	YES	YES	YES	YES	YES	YES
Year dummies	YES	YES	YES	YES	YES	YES
Log likelihood	−500	−484	−2,504	−2,438	−4,541	−4,343
Chi-square	653	684	3,502	3,634	5,736	6,132

Notes: Robust standard errors in parentheses. ***P<0.01, **P<0.05, *P<0.1

TABLE A8

Average partial effect (APE): industry splits (fixed effects regression)

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	High tech	High tech	Medium tech	Medium tech	Low tech	Low tech
Product	0.278	0.257	0.323	0.299	0.361	0.337
Process	0.313	0.297	0.400	0.376	0.399	0.358

Final Manuscript Received: January 2020

References

Aghion, P., Harris, C. and Vickers, J. (1997). ‘Competition and growth with step-by-step innovation: an example’, *European Economic Review*, Vol. 41, pp. 771–782.

Aghion, P. and Howitt, P. (1992). ‘A model of growth through creative destruction’, *Econometrica*, Vol. 60, pp. 323–351.

Antonelli, C., Crespi, F. and Scellato, G. (2012). ‘Inside innovation persistence: new evidence from Italian micro-data’, *Structural Change and Economic Dynamics*, Vol. 23, pp. 341–353.

Antonelli, C., Crespi, F. and Scellato, G. (2013). ‘Internal and external factors in innovation persistence’, *Economics of Innovation and New Technology*, Vol. 22, pp. 256–280.

Antonioli, D. and Montresor, S. (2019). ‘Innovation persistence in times of crisis: an analysis of Italian firms’, *Small Business Economics*, pp. 1–26.

Arellano, M. and Bonhomme, S. (2009). ‘Robust priors in nonlinear panel data models’, *Econometrica*, Vol. 77, pp. 489–536.

Arellano, M. and Bonhomme, S. (2011). ‘Nonlinear panel data analysis’, *Annual Review of Economics*, Vol. 3, pp. 395–424.

Arellano, M. and Hahn, J. (2007). ‘Understanding bias in nonlinear panel models: some recent developments’, *Econometric Society Monographs*, Vol. 43, p. 381.

Arqué-Castells, P. (2013). ‘Persistence in R&D performance and its implications for the granting of subsidies’, *Review of Industrial Organization*, Vol. 43, pp. 193–220.

Arrow, K. J. (1962). ‘The economic implications of learning by doing’, *The Review of Economic Studies*, Vol. 29, pp. 155–173.

Arulampalam, W. and Stewart, M. B. (2009). ‘Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators’, *Oxford Bulletin of Economics and Statistics*, Vol. 71, pp. 659–681.

Badillo, E. R. and Moreno, R. (2016). ‘Are collaborative agreements in innovation activities persistent at the firm level? Empirical evidence for the Spanish case’, *Review of Industrial Organization*, Vol. 49, pp. 71–101.

Barney, J. (1991). ‘Firm resources and sustained competitive advantage’, *Journal of Management*, Vol. 17, pp. 99–120.

Barreto, I. (2010). ‘Dynamic capabilities: a review of past research and an agenda for the future’, *Journal of Management*, Vol. 36, pp. 256–280.

Baumol, W. J., Schilling, M. A. and Wolff, E. N. (2009). ‘The superstar inventors and entrepreneurs: how were they educated?’, *Journal of Economics & Management Strategy*, Vol. 18, pp. 711–728.

Belderbos, R., Carree, M. and Lokshin, B. (2004). ‘Cooperative R&D and firm performance’, *Research Policy*, Vol. 33, pp. 1477–1492.

Bianchini, S., Bottazzi, G. and Tamagni, F. (2017). ‘What does (not) characterize persistent corporate high-growth?’, *Small Business Economics*, Vol. 48, pp. 633–656.

Bianchini, S., Pellegrino, G. and Tamagni, F. (2018). ‘Innovation complementarities and firm growth’, *Industrial and Corporate Change*, Vol. 27, pp. 657–676.

Bloom, N. and Van Reenen, J. (2007). ‘Measuring and explaining management practices across firms and countries’, *The Quarterly Journal of Economics*, Vol. 122, pp. 1351–1408.

Bloom, N. and Van Reenen, J. (2010). ‘Why do management practices differ across firms and countries?’, *Journal of Economic Perspectives*, Vol. 24, pp. 203–224.

Brown, J. R., Fazzari, S. M. and Petersen, B. C. (2009). ‘Financing innovation and growth: cash flow, external equity, and the 1990s R&D boom’, *The Journal of Finance*, Vol. 64, pp. 151–185.

Carro, J. M. (2007). ‘Estimating dynamic panel data discrete choice models with fixed effects’, *Journal of Econometrics*, Vol. 140, pp. 503–528.

Cefis, E. (2003). ‘Is there persistence in innovative activities?’, *International Journal of Industrial Organization*, Vol. 21, pp. 489–515.

Cefis, E. and Ciccarelli, M. (2005). ‘Profit differentials and innovation’, *Economics of Innovation and New Technology*, Vol. 14, pp. 43–61.

Cefis, E. and Marsili, O. (2015). 'Crossing the innovation threshold through mergers and acquisitions', *Research Policy*, Vol. 44, pp. 698–710.
Cefis, E. and Orsenigo, L. (2001). 'The persistence of innovative activities: a cross-countries and crosssectors comparative analysis', *Research Policy*, Vol. 30, pp. 1139–1158.
Cirillo, V., Sostero, M. and Tamagni, F. (2017). 'Innovation and within-firm wage inequalities: empirical evidence from major European countries', *Industry and Innovation*, Vol. 24, pp. 468–491.
Clausen, T., Pohjola, M., Sapprasert, K. and Verspagen, B. (2011). 'Innovation strategies as a source of persistent innovation', *Industrial and Corporate Change*, Vol. 21, pp. 553–585.
Cohen, W. M. and Klepper, S. (1996). 'A reprise of size and R&D', *The Economic Journal*, Vol. 106, pp. 925–951.
Cohen, W. M. and Levinthal, D. A. (1989). 'Innovation and learning: the two faces of R&D', *The Eeconomic Journal*, Vol. 99, pp. 569–596.
Cohen, W. M. and Levinthal, D. A. (1990). 'Absorptive capacity: a new perspective on learning and innovation', *Administrative Science Quarterly*, pp. 128–152.
Conner, K. R. (1991). 'A historical comparison of resource-based theory and five schools of thought within industrial organization economics: do we have a new theory of the firm?', *Journal of Management*, Vol. 17, pp. 121–154.
Crépon, B. and Duguet, E. (1997). 'Estimating the innovation function from patent numbers: GMM on count panel data', *Journal of Applied Econometrics*, Vol. 12, pp. 243–263.
Cruz-Gonzalez, M., Fernández-Val, I. and Weidner, M. (2017a). 'Bias corrections for probit and logit models with two-way fixed effects', *The Stata Journal*, Vol. 17, pp. 517–545.
Cyert, R. M. and March, J. G. (1963). *A Behavioral Theory of the Firm*, Englewood Cliffs, NJ.
Czarnitzki, D. and Hottenrot, H. (2010). 'Financing constraints for industrial innovation: what do we know?', *Review of Business and Economic Literature*, Vol. 55, pp. 346–363.
Czarnowske, D. and Stammann, A. (2019). Binary Choice Models with High-Dimensional Individual and Time Fixed Effects, *arXiv preprint arXiv:1904.04217*.
Dhaene, G. and Jochmans, K. (2015). 'Split-panel Jackknife estimation of fixed-effect models', *The Review of Economic Studies*, Vol. 82, pp. 991–1030.
Dosi, G. and Marengo, L. (1994). 'Towards a theory of organisational competencies', *Evolutionary Concepts in Contemporary Economics*, Ann Arbor, Michigan University Press, Michigan.
Dosi, G., Marsili, O., Orsenigo, L. and Salvatore, R. (1995). 'Learning, market selection and the evolution of industrial structures', *Small Business Economics*, Vol. 7, pp. 411–436.
Duguet, E. and Monjon, S. (2004). 'Is Innovation Persistent at the Firm Level? An Econometric Examination Comparing the Propensity Score and Regression Methods', *University of Paris I - Cahiers de la MSE Working Paper*, 2004(75).
European Commission (2021). 'European Innovation Scoreboard', data retrieved from https://ec.europa.eu/growth/industry/policy/innovation/scoreboards_en
Fagerberg, J., Mowery, D. C. and Nelson, R. R. (2005). *The Oxford Handbook of Innovation*, Oxford University Press, Oxford.
Fernández-Val, I. (2009). 'Fixed effects estimation of structural parameters and marginal effects in panel probit models', *Journal of Econometrics*, Vol. 150, pp. 71–85.
Fernández-Val, I. and Weidner, M. (2016). 'Individual and time effects in nonlinear panel models with large N, T', *Journal of Econometrics*, Vol. 192, pp. 291–312.
Flaig, G. and Stadler, M. (1994). 'Success breeds success. The dynamics of the innovation process', *Empirical Economics*, Vol. 19, pp. 55–68.
Ganter, A. and Hecker, A. (2013). 'Persistence of innovation: discriminating between types of innovation and sources of state dependence', *Research Policy*, Vol. 42, pp. 1431–1445.
Geroski, P. A., Van Reenen, J. and Walters, C. F. (1997). 'How persistently do firms innovate?', *Research Policy*, Vol. 26, pp. 33–48.
Grant, R. M. (1991). 'The resource-based theory of competitive advantage: implications for strategy formulation', *California Management Review*, Vol. 33, pp. 114–135.
Greene, W. (2004). 'The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects', *The Econometrics Journal*, Vol. 7, pp. 98–119.

Guarascio, D. and Tamagni, F. (2019). 'Persistence of innovation and patterns of firm growth', *Research Policy*, Vol. 48, pp. 1493–1512.
Hahn, J. and Newey, W. (2004). 'Jackknife and analytical bias reduction for nonlinear panel models', *Econometrica*, Vol. 72, pp. 1295–1319.
Hall, B. H. (2002). 'The financing of research and development', *Oxford Review of Economic Policy*, Vol. 18, pp. 35–51.
Haned, N., Mothe, C. and Nguyen-Thi, T. U. (2014). 'Firm persistence in technological innovation: the relevance of organizational innovation', *Economics of Innovation and New Technology*, Vol. 23, pp. 490–516.
Hannan, M. T. and Freeman, J. (1984). 'Structural inertia and organizational change', *American Sociological Review*, pp. 149–164.
Hecker, A. and Ganter, A. (2014). 'Path and past dependence of firm innovation', *Economics of Innovation and New Technology*, Vol. 23, pp. 563–583.
Heckman, J. J. (1981). 'Statistical models for discrete panel data', *Structural Analysis of Discrete Data with Econometric Applications*, pp. 114–178.
Heckman, J. J. (1987). The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process and Some Monte Carlo Evidence. University of Chicago Center for Mathematical studies in Business and Economics.
Helfat, C. E. (1994). 'Evolutionary trajectories in petroleum firm R&D', *Management Science*, Vol. 40, pp. 1720–1747.
Hitt, M. A., Hoskisson, R. E. and Kim, H. (1997). 'International diversification: effects on innovation and firm performance in product-diversified firms', *Academy of Management Journal*, Vol. 40, pp. 767–798.
Honoré, B. E. (2002). 'Nonlinear models with panel data', *Portuguese Economic Journal*, Vol. 1, pp. 163–179.
Huergo, E. and Moreno, L. (2011). 'Does history matter for the relationship between R&D, innovation, and productivity?', *Industrial and Corporate Change*, Vol. 20, pp. 1335–1368.
INE (2018). *Panel de innovación tecnológica (PITEC)*. Instituto Nacional de Estadistica, Madrid.
Khazanchi, S., Lewis, M. W. and Boyer, K. K. (2007). 'Innovation-supportive culture: the impact of organizational values on process innovation', *Journal of Operations Management*, Vol. 25, pp. 871–884.
Klevorick, A. K., Levin, R. C., Nelson, R. R. and Winter, S. G. (1995). 'On the sources and significance of interindustry differences in technological opportunities', *Research Policy*, Vol. 24, pp. 185–205.
Kogut, B. and Zander, U. (1992). 'Knowledge of the firm, combinative capabilities, and the replication of technology', *Organization Science*, Vol. 3, pp. 383–397.
Latham, W. R. and Le Bas, C. (2006). *The Economics of Persistent Innovation: An Evolutionary View*. Springer, Berlin.
Laursen, K. and Salter, A. (2006). 'Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms', *Strategic Management Journal*, Vol. 27, pp. 131–150.
Le Bas, C. and Poussing, N. (2014). 'Are complex innovators more persistent than single innovators? An empirical analysis of innovation persistence drivers', *International Journal of Innovation Management*, Vol. 18, p. 1450008.
Le Bas, C. and Scellato, G. (2014). 'Firm innovation persistence: a fresh look at the frameworks of analysis', *Economics of Innovation and New Technology*, Vol. 23, pp. 423–446.
Lengnick-Hall, C. A. (1992). 'Innovation and competitive advantage: what we know and what we need to learn', *Journal of Management*, Vol. 18, pp. 399–429.
Levitt, B. and March, J. G. (1988). 'Organizational learning', *Annual Review of Sociology*, Vol. 14, pp. 319–338.
Lhuillery, S. (2014). 'Marketing and persistent innovation success', *Economics of Innovation and New Technology*, Vol. 23, pp. 517–543.
Lotka, A. J. (1926). 'The frequency distribution of scientific productivity', *Journal of the Washington Academy of Sciences*, Vol. 16, pp. 317–323.
Malerba, F. and Orsenigo, L. (1996). 'Schumpeterian patterns of innovation are technology-specific', *Research policy*, Vol. 25, pp. 451–478.

Malerba, F. and Orsenigo, L. (1999). 'Technological entry, exit and survival: an empirical analysis of patent data', *Research Policy*, Vol. 28, pp. 643–660.
Martínez-Ros, E. and Labeaga, J. M. (2009). 'Product and process innovation: persistence and complementarities', *European Management Review*, Vol. 6, pp. 64–75.
Ministerio de Industria, C. y. T. (2020). 'Estructura y dinamica empresarial en Espana (Datos a 1 de enero de 2019)', Publicaciones de la Administracion General de Estado.
Miotti, L. and Sachwald, F. (2003). 'Co-operative R&D: why and with whom?: An integrated framework of analysis', *Research Policy*, Vol. 32, pp. 1481–1499.
Mulet-Melia, J. (2020). 'La Innovacion y la I+D espanolas en 2018 y su comparacion internacional', *Estudios sobre la Economia Espanola*, 05.
Nam, V. H. and Bao Tram, H. (2019). 'Business environment and innovation persistence: the case of small-and medium-sized enterprises in Vietnam', *Economics of Innovation and New Technology*, pp. 1–23.
Narin, F. and Breitzman, A. (1995). 'Inventive productivity', *Research Policy*, Vol. 24, pp. 507–519.
Nelson, R. and Winter, S. (1982). *An Evolutionary Theory of Economic Change*, Harvard Business School Press, Cambridge.
Neyman, J. and Scott, E. L. (1948). 'Consistent estimates based on partially consistent observations', *Econometrica: Journal of the Econometric Society*, pp. 1–32.
OECD (2011). Technology Intensity Definition. Classification of Manufacturing Industries into Categories Based on R&D Intensities (ISIC REV. 3). OECD, Paris.
OECD and Eurostat (1997). *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data*. OECD, Paris.
Patel, P. and Pavitt, K. (1991). 'Europe's technological performance', *Technology and the Future Of Europe: Global Competition and the Environment in the 1990s*.
Pavitt, K. (1990). 'What we know about the strategic management of technology', *California Management Review*, Vol. 32, pp. 17–26.
Peters, B. (2009). 'Persistence of innovation: stylised facts and panel data evidence', *The Journal of Technology Transfer*, Vol. 34, pp. 226–243.
Piga, C. A. and Vivarelli, M. (2004). 'Internal and external R&D: a sample selection approach', *Oxford Bulletin of Economics and Statistics*, Vol. 66, pp. 457–482.
Price, D. J. D. S. (1965). 'Networks of scientific papers', *Science*, pp. 510–515.
Rabe-Hesketh, S. and Skrondal, A. (2013). 'Avoiding biased versions of Wooldridge's simple solution to the initial conditions problem', *Economics Letters*, Vol. 120, pp. 346–349.
Raymond, W., Mohnen, P., Palm, F. and Van Der Loeff, S. S. (2010). 'Persistence of innovation in Dutch manufacturing: is it spurious?', *The Review of Economics and Statistics*, Vol. 92, pp. 495–504.
Rosenberg, N. (1976). *Perspectives on Technology*, Cambridge University Press, Cambridge.
Schumann, M., Severini, T. A. and Tripathi, G. (2021). 'Integrated likelihood based inference for nonlinear panel data models with unobserved effects', *Journal of Econometrics*, Vol. 223, pp. 73–95.
Schumpeter, J. A. (1934). *The Theory of Economic Development*, Harvard University, Cambridge, MA.
Stuart, T. E. and Podolny, J. M. (1996). 'Local search and the evolution of technological capabilities', *Strategic Management Journal*, Vol. 17, pp. 21–38.
Suarez, D. (2014). 'Persistence of innovation in unstable environments: continuity and change in the firm's innovative behavior', *Research Policy*, Vol. 43, pp. 726–736.
Sutton, J. (1991). *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*, MIT Press, Cambridge, Massachusetts.
Tavassoli, S. and Karlsson, C. (2015). 'Persistence of various types of innovation analyzed and explained', *Research Policy*, Vol. 44, pp. 1887–1901.
Teece, D. J. (2007). 'Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance', *Strategic Management Journal*, Vol. 28, pp. 1319–1350.
Teece, D. J., Pisano, G. and Shuen, A. (1997). 'Dynamic capabilities and strategic management', *Strategic Management Journal*, Vol. 18, pp. 509–533.
Tether, B. S. (2002). 'Who co-operates for innovation, and why: an empirical analysis', *Research Policy*, Vol. 31, pp. 947–967.

Triguero, A. and Córcoles, D. (2013). 'Understanding innovation: an analysis of persistence for Spanish manufacturing firms', *Research Policy*, Vol. 42, pp. 340–352.

Triguero, A., Córcoles, D. and Cuerva, M. (2014). 'Measuring the persistence in innovation in Spanish manufacturing firms: empirical evidence using discrete-time duration models', *Economics of Innovation and New Technology*, Vol. 23, pp. 447–468.

Weitzman, M. L. (1998). 'Recombinant growth', *The Quarterly Journal of Economics*, Vol. 113, pp. 331–360.

Wernerfelt, B. (1984). 'A resource-based view of the firm', *Strategic Management Journal*, Vol. 5, pp. 171–180.

Woerter, M. (2014). 'Competition and persistence of R&D', *Economics of Innovation and New Technology*, Vol. 23, pp. 469–489.

Wooldridge, J. M. (2005). 'Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity', *Journal of Applied Econometrics*, Vol. 20, pp. 39–54.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Massachusetts.

Zucker, L. G., Darby, M. R. and Brewer, M. B. (1998). 'Intellectual human capital and the birth of US biotechnology enterprises', *The American Economic Review*, Vol. 88, p. 290.