UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE ED221 AUGUSTIN COURNOT

[UMR 7522]

THÈSE présentée par :

Stefano Bianchini

soutenue le : **23 juin 2015**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Sciences économiques

Three essays on firm growth, innovation, and persistent performance

THÈSE dirigée par :

M. LLERENA Patrick	Professeur, Université de Strasbourg
M. BOTTAZZI Giulio	Professeur, Scuola Superiore Sant'Anna, Pise (Italie)

RAPPORTEURS :

M. FOLTA Timothy	Professeur, University of Connecticut, Storrs (US)
M. VIVARELLI Marco	Professeur, Università Cattolica del Sacro Cuore, Milan (Italie)

AUTRES MEMBRES DU JURY :

M. DOSI Giovanni	Professeur, Scuola Superiore Sant'Anna, Pise (Italie)
M. LISSONI Francesco	Professeur, Université de Bordeaux

To Angela and Lorenzo

Acknowledgement

I would like to express my profound and sincere gratitude to my supervisor Patrick Llerena who guided me, encouraged and supported along each stage of this doctoral program.

I would like to acknowledge the importance of Professors Giulio Bottazzi and Federico Tamagni, whose help has been fundamental for the realization of this thesis. Through their unfailing guidance I have learned a lot.

I am especially indebted to Professor Francesco Lissoni. Perhaps I would never considered to start a PhD if it was not for his inspiration. His stimulus are always of crucial importance.

My thanks to Professors Giovanni Dosi, Timothy Folta and Marco Vivarelli. I am honoured that they accepted to be members of the jury.

Last but not least, a special thanks to my beloved family.

Contents

1 Introduction **3**

2 Economic and financial determinants of persistent high-growth performance **9**

2.1 Introduction . . . 9
2.2 Background literature and motivation . . . 11
2.3 Data and identification of persistent high-growth firms . . . 13
2.3.1 Firm characteristics . . . 17
2.4 Distributional analysis . . . 19
2.5 Regression analysis . . . 22
2.6 The role of innovation . . . 27
2.7 Size and age . . . 28
2.8 Alternative regression models . . . 32
2.9 Conclusion . . . 35
2.10 Appendix . . . 36

3 Innovation strategies and firm growth **39**

3.1 Introduction . . . 39
3.2 Background framework . . . 40
3.3 Data . . . 44
3.3.1 Data and sample . . . 44
3.3.2 Main variables . . . 45
3.4 Growth and innovation: descriptive evidence . . . 47
3.5 Growth and innovation: main results . . . 48
3.5.1 Panel estimates . . . 50
3.5.2 Fixed-Effects quantile regressions . . . 55
3.6 Testing complementarity of innovation activities . . . 59
3.7 Conclusions . . . 63
3.8 Appendix . . . 65

4 Does persistence of innovation spur persistence of growth? **71**

4.1 Introduction . . . 71
4.2 Background literature and research objectives . . . 73
4.2.1 On the persistence of innovation . . . 73
4.2.2 On the persistence of growth . . . 74
4.2.3 Connecting persistence of innovation to persistence of growth . 75
4.3 Data and empirical setting . . . 75
4.3.1 Source and variables . . . 75

4.3.2 Measuring persistence of growth and innovation 76
4.3.3 Empirical methodology 78
4.4 Results 79
4.4.1 Descriptive evidence 79
4.4.2 Main evidence 80
4.4.3 Robustness checks 86
4.5 Conclusions 86

5 French summary **91**

Introduction

Firm growth is a topic that has for long intrigued many constituencies. Economists, in primis, have always looked at firm growth as a tangible proof of the process of market selection, to understand whether markets effectively deliver rewards and punishments in terms of relative sizes or shares according to differential efficiencies (Bartelsman et al., 2005; Lotti et al., 2009; Bottazzi et al., 2010; Dosi et al., 2013). Policy-makers are attracted by growing firms because of their potential in terms of new jobs creation and fostering macroeconomic growth (see, for instance, the recent discussion in Haltiwanger et al., 2013). Last, scholars from strategic management tradition, managers and consultants are interested in understanding the best-practices which are responsible for superior firm performance, thus to replicate them within their own business or the business of their clients (Teece et al., 1997; Katkalo et al., 2010). Among the many companies that populate our economies a small group of firms with extraordinary growth performance, which the literature commonly referred to as high-growth firms or “gazelles”, have been recently the target of academic inquiry and policy initiatives; the latter often regard taxation, regulation, immigration, access to capital, and academic commercialization (Stangler, 2010; Schimke and Mitusch, 2011).

Accordingly, the increasing attention to the topic of firm growth, and lately of high-growth, is reflected in the abundant theoretical and empirical literature developed in the last decades. At the theoretical level contributions come from alternative schools of thought, and despite differences in the underlying assumptions, the process of growth (and high-growth) is typically regarded as the result of the interplay between three main dimensions of the firm, namely productivity (or efficiency), profitability, and financial status (see, for instance, Nelson and Winter, 1982; Jovanovic, 1982; Ericson and Pakes, 1995). Concisely, these models predict that an idiosyncratic shock affecting firm-specific unobserved factors (i.e. technological and organizational traits, capabilities, strategic and managerial practices) leads to heterogeneous efficiency across firms, and firms with higher relative efficiency gain market shares at the expenses of less efficient units. Asymmetries in profitability and financial conditions grant to more productive firms the access to the resources needed to invest and pursue further growth.

But despite the theoretical background have informed the empirical investigations that the candidate key drivers of firm growth must be searched for in terms of efficiency, profitability, and financial status, the existent evidence does not point to a straightforward relationship among these dimensions. Indeed, behind the simplified linear pattern that the theoretical models describe, play both the strong complexity and the idiosyncrasy that are well known to characterize the process of firm growth (Geroski, 2002; Delmar et al., 2003). Thus, while some quite robust stylized facts emerge from the industrial dynamics studies (for instance, think of the negative relationship between age and size with firm growth), we still face a lot of puzzling evidences. For example, against the expectations, some scholars find that more efficient and productive firms have weak behavioural inclination to growth more, and this is also the case for firms with sounder financial conditions (Bottazzi et al., 2002, 2008; Coad, 2007b); other scholars find a significant convergence toward Gibrat-like behaviour in the long-run (Lotti et al., 2009).

The ambiguous empirical regularities might be due to the fact that other factors beyond mere productivity, profitability, and financial variables are playing an important role in shaping the process of firm growth. Not surprisingly some external factors such as macroeconomic and industry specific features (i.e. business cycles, regional peculiarities, industry lifecycle) have been proved to affect firm growth, as well as for other firm-level components, innovativeness among the top (see Coad, 2009 for an exhaustive survey). Concerning the latter, recent contributions show, in addition, that the process leading from innovative input to innovative output may display different effects according to the different positioning of a firm in the growth rates distribution, with high-growing firms deriving more benefits from innovation activities (Coad and Rao, 2008; Hölzl, 2009).

The scenario becomes even more intricate when the concept of persistence in growth patterns is introduced. The existence of persistent profitability differentials hints indeed at an equally persistent growth dynamics, as more efficient firms, exploiting their increased availability of internal resources, progressively erode market shares from competitors. In line with the theoretical background, we should observe some degree of persistence in the process of growth since the "good firms" tend to expand at first rapidly ("success breeds success"-type of dynamics), and then experiencing a progressive slow down. Persistent growth performance are of particular concern for economists and policy-makers since, at least in principle, they should be connected with the presence of exceptional capabilities inside the firm or structural advantages around it (Teece, 2007). Put differently, persistence of firm growth should run again the notion of randomness for which outperforming firms are simply "more lucky" than their competitors (Barney, 1997).

But once again empirics is not reconciled with the theoretical expectations, and these "success breeds success" dynamics are very difficult to detect in the data, so that only few companies are able to sustain their higher growth performance over long periods of time. Empirical studies failed indeed to find a robust degree of persistence in the process of growth; rather, the evidence goes from strong growth autocorrelation (Bottazzi et al., 2001) to no or negative autocorrelation, which supports a more erratic and unpredictable nature of the growth profile (Coad and Hölzl, 2009; Bottazzi et al., 2011; Hölzl, 2014). Recent contributions (Delmar et al., 2003; Capasso et al., 2013; Daunfeldt and Halvarsson, 2013) focus on persistence of high-growth patterns, emphasizing the existence of a group of persistent outperforming units but simply delineating their mere demographic profile (size, age, and industry affiliation).

Complementing the mixed evidence on the persistent nature of firm growth, there is instead agreement about the persistent nature of innovation, at least when input proxies are considered (i.e. R&D expenditures). At the theoretical level persistence in innovation finds several explanations, starting from the Schumpeterian idea that relates market power and innovation, as monopolists have more to lose by not innovating than potential new entrants do, to more recent conjectures which acknowledge the importance of information asymmetry between the innovator and the lender, so that innovative firms tend to rely on retained earnings rather than external funds (Bhattacharya and Ritter, 1983). In line with this view, past innovation success yields profits that can be used to finance current innovation activities, thus inducing persistence in innovation behaviour. Some scholars argue that the high degree of persistence in R&D is very likely to be induced by specific industry barriers such as start-ups sunk costs (Sutton, 1991); others have put forward interpretations related to the process of learning by innovating (Dosi and Marengo, 1994). At the empirical level most of the recent findings (among the many see Cefis, 2003; Peters, 2009; Raymond et al., 2010) do point to a high degree of persistence in R&D activities, whereas patent-based studies find on average little evidence of persistence in innovation (see, for instance, Geroski et al., 1997).

However, contributions that analyze the relationship between innovation persistence and firm performance are still very scarce, and linking more explicitly the evidence on the patterns of innovation with what is known about firms growth, both at the empirical and at the theoretical level, is a hard but urgent challenge for future research (Dosi, 2007).

This dissertation aims to fill some gaps of the extant literature. The three essays focus, by and large, on the process of firm growth, its persistence, and on the role of innovation in affecting firm performance.

In the first essay we concentrate on persistence of high-growth and investigate whether this peculiar growth pattern is actually associated with better operating capabilities. As already stressed, the conspicuous existent literature is primarily focused on the identification of the causes and conditions that led a company to outperform its competitors in a specific, relatively short, period of time. In our work we offer a new perspective as we address if more structural, economic or financial, factors are distinguishing features of companies that are able to exhibit high-growth performance repeatedly over time, above and beyond mere demographic characteristics such as size, age and industry affiliation.

We build our conjectures drawing from the models of firm-industry evolution. Market competition should favor more efficient and profitable firms, and sounder financial conditions should help accessing the external resources needed to finance investment and growth. Therefore, we should expect high-growth firms to be more productive and more profitable than firms displaying less abnormal growth. More interesting is to understand whether the same characteristics also display any association with persistence in high-growth, and whether persistent high-growth firms differ, in terms of these characteristics, from other firms and, in particular, from firms that display “spurts” of high-growth but are not able to consistently sustain this pace over longer periods of time.

We carry out an investigation on a panel data of Italian, French, Spanish and UK incumbents, identifying high-growth and persistent high-growth performers. We analyze, both in a non-parametric and parametric setting, how initial years productivity, profitability and financial factors relate with subsequent growth dynamics. In the exploratory non-parametric analysis we explore if a set of key variables, taken to proxy the operational performance and financial status, display distributional differences across high growers, persistently high growers and other firms. Secondly, we estimate several specifications of a multinomial probit model to identify which variables are more effective in discriminating persistent high-growth firms from "simple" high-growth and other firms. All these exercises are extended in several ways in order to test the role of institutional or other more macro-level factors, sectoral peculiarities and patterns of innovation. In line with the theoretical background we do confirm that economic determinants, and productivity in particular, is significantly associated with high-growth. However, we do not find systematic evidence of any statistically significant difference between high-growth and persistently high-growth firms in term of operating efficiency; none of the considered dimensions therefore seems to work in sustaining high-growth performance repeatedly over time. Although there is a number of other potential factors that may work in sustaining high-growth over time, our results are intriguing as, at least in principle, we cannot exclude that persistent high-growth might occur at random.

The second essay aims to explore the relationship between growth and innovation, taking into account the multidimensional nature of the innovation process. Our contribution is in line with the recent call by Audretsch et al. (2014): "The complexity of R&D activities, together with the diversity of innovation strategies and the multiplicity of growth modes, requires a multidimensional approach to examine the contribution of innovations on firm growth.".

We draw upon information from a CIS-type of dataset on Spanish firms over the period 2004-2011, very peculiar given its longitudinal nature. The contributions we deliver in this study are manifold. First, we provide a broad picture of the relationship between growth and innovation, by looking at a wide set of innovation variables that capture the different sources, modes and types of innovative activity undertaken within firms. More precisely, the set of innovation indicators encompasses internal vs. external R&D, process innovation, types of product innovation, and embodied vs. disembodied technological acquisition. Secondly, we account for the interactions among the above mentioned innovation modes, increasing the complexity level and investigating whether growth is likely to be fostered by specific combinations of innovation activities. In performing this step we test whether complementarities among innovation modes are at stake.

We start to separately investigate the relationship between sales growth and each innovation activity. A common limitation to extant studies exploiting CIS-like data to analyze growth-innovation interplay is that such surveys are run in waves every 3-4 years, hence forcing the econometric setting to consist of a single cross section, and in turn failing to carefully control for unobserved heterogeneity. But the ample degrees of complexity, uncertainty and idiosyncrasy that characterize the innovation process, as well as the huge asymmetries in production efficiencies, would suggest that a systemic control of unobserved heterogeneity is more than a mere technicality. Thus we exploit the longitudinal dimension of our data, detecting how each innovation variable correlates with sales growth applying both a conventional regression framework and a up-to-date quantile regression technique designed to account for firm-level fixed effects.

Our first set of results point to a good deal of heterogeneity in the capacity of different innovation activities to contribute to expanding sales. Indeed, among the innovation indicators we account for, R&D (especially if carried out internally) and embodied technical change represent the primary sources of competitive advantage, in particular for a subset of high-growing companies. While process innovation seems to not directly affect sales growth, it emerges that high-growth firms receive a higher premium when they introduce innovative product into the market.

Being aware that the relationship between growth and innovation may be different for firms which are active simultaneously in all layers, with respect to firms that only performs one or two innovation activities at a time, we account for the interactions among the different innovation modes building what we refer to as "innovation strategies". Higher complexity probably rises more costs and challenging coordination issues, but at the same time can also offer stronger ability to capture and create growth opportunities. We perform a direct test of complementarity (see, e.g., Mohnen and Roller, 2005; Catozzella and Vivarelli, 2014) between a set of innovation modes, which points indeed to some complementarity effects: product innovation is likely to have a stronger effect on growth if internal R&D activities are simultaneously undertaken, and we do find evidence of complementarity also between process and product innovation.

In the third essay we examine the role of persistence of innovation on persistence of growth performance, assessing whether a systematic, rather than sporadic, engagement in innovation activities induce more structure in the process of firm growth. Understing whether persistence of innovation spurs persistence of growth is relevant to verify to what extent, and how quickly, market competition erodes the competitive advantages that firms build upon their innovation success, as well as the longlasting consequences that different innovation behaviours may induce.

Industrial ecomomics and strategic management literature provides some guidelines for our conjectures: as innovation success is likely to be the result of a systematic engagement into innovation activities, we expect firms who are able to persistently translate their innovative efforts into valuable outcomes (for instance patentable products) to overcome their competitors continuously for longer periods, put differently, to show a higher degree of persistence in their profitability and growth paths (Dosi and Nelson, 2010). Since the innovation process involves uncertainty and risk taking, investing for innovation is a necessary but not sufficient condition to succeed, thus we expect that persistence in R&D investments does not necessarily boost structure into growth process.

We test these conjectures exploiting a rich longitudinal data set of Spanish firms (Encuesta Sobre Estrategias Empresariales-ESEE) comprising information about firms and market characteristics, spanning a period of twenty years (from 1990 to 2009). Our empirical strategy is based on the characterization of the innovation status of each firm, that we derive by looking at the frequency with which the company innovates. To this end we consider both input and output proxies for innovation, namely R&D expenditure and number of patent applications. We identify and distinguish three categories of firms: non-innovators, occasional innovators, and persistent innovators. We firstly conduct an explanatory analysis to assess whether a set of key firm-level variables display differences across the different groups. Afterwards, we exploit duration model techniques to investigate the determinants of the length of a period of continuous positive growth (what we refer to as "growth spell"). Put it differently, we take the spell of time in which a firm grows as unit of analysis, and model the probability that the spell will end at any particular time.

Our results show that persistent innovators differ from occasional innovators (and from non-innovators too) in terms of demographic features, being generally larger, older, and active in high-tech sectors. Moreover, in line with the theoretical conjectures, we do find support that persistence of innovation affect persistence of growth, but it does so only when firms build their durable competitive advantages upon systematic innovation success.

In summary, although with some obvious limitations, this thesis contributes to the empirical literature of firm-industry dynamics, and provides new insights for a quite disparate audience ranging from industrial economists, practitioners, strategic management scholars, and eventually policy makers. All in all, this document delivers a set of contributions that concern principally the following points:

1. Micro-level determinants of high-growth patterns;
2. Economic and financial characterization of persistently high-growing units;
3. Relationship between a wide set of innovation indicators and sales growth;
4. Existence of complementarities between innovation modes in fostering growth;
5. Effects of a systematic engagement in innovation activities on persistence of growth performance.

Economic and financial determinants of persistent high-growth performance

2.1 Introduction

Among the many private companies that populate developed economies it is typically possible to identify, within a given time window, a small group of firms with extraordinary growth performance, which are commonly referred to as high-growth firms or "gazelles" (among others, see Schreyer, 2000; Delmar et al., 2003; Acs and Mueller, 2008; Parker et al., 2010). This kind of companies attracts the attention not only of academic scholars, but also of managers, practitioners and policy makers (see for instance the discussion in Schimke and Mitusch, 2011; Stangler, 2010). On the one hand, managers and consultants are interested in understanding the "best-practices" which are responsible for superior firm performance and seek to replicate them within their own business or the business of their clients. On the other hand, policy-makers are particularly interested in the early identification of high-growth firms because of their extraordinary potential in terms of new jobs creation and fostering of macroeconomic growth.

There is a vast empirical literature on high-growth companies, that links the occurrence of high-growth events to macro-economic or institutional factors, external to the firm, and to micro-economic characteristics specific to a given firm. The latter often include demographic variables such as age and size, together with more economic determinants such as the degree of firm innovativeness. This literature focuses on the identification of the causes and conditions that lead a company to outperform its competitors in a specific, relatively short, period of time.

In this essay we offer a different perspective. Instead of searching for the determinants of high-growth at a given point in time, we want to identify the factors that make a firm a *persistent* high-growing firm. The motivation for this shift of focus rests in the consideration that high-growth performance have a more relevant economic impact, and turn more interesting to practitioners and promising to policy makers, if they are long lasting and persistent. This is indeed the kind of growth behavior that is likely connected with the presence of structural comparative advantages and exceptional capabilities inside the firm. As a matter of fact, the dynamics underlying a fast expansion can vary, even in substantial form, from company to company (Delmar et al., 2003): some firms sporadically respond to market shocks, other companies display a more erratic and unpredictable pattern, and only few are able to exhibit a persistent, continuing year after year, fast expansion.

While empirical research has for long concentrated on the persistence of firm growth rates, with mixed results, the study of the persistence in high-growth patterns is only of very recent development. Existing studies limit the attention to the exploration of demographic characteristics of firms, such as size, age or sector of activity. We instead want to address whether persistent high-growth is related to better operating capabilities. Do persistent high-growth firms differ in terms of productivity or profitability with respect to firms that display "spurts" of high-growth, but are not able to consistently replicate high rates of growth over a longer period of time? Answering this question has relevant policy implications, since one wants to understand whether firms able to sustain high-growth over time are also those which can increase the overall efficiency and competitiveness of sectors and countries. We do not know of previous studies making such an attempt.

Theories of firm-industry dynamics with heterogeneous firms, from different traditions (see, e.g., Nelson and Winter, 1982; Jovanovic, 1982; Silverberg et al., 1988; Dosi et al., 1995; Metcalfe, 1998; Ericson and Pakes, 1995; Cooley and Quadrini, 2001; Luttmer, 2007) provide the theoretical background of our analysis. Although none of the models specifically addresses the issue of the relative abundance of high-growth firms and their behavior over time, they all relate growth rates differentials across firms to the presence of competitive advantages due to structural factors, which influence firm performance over a relatively long period of time. The existence of persistent profitability differentials hints at an equally persistent growth dynamics, as more efficient firms, exploiting their increased availability of internal resources, progressively erode market shares from competitors. In practice, however, firm expansion, especially if it is fast and relatively large, must often rely upon external finance, despite the fact that growth events can be considered as a risky enterprise by potential lenders. Thus, both the degree of dependence from external credit and the cost of credit can influence the occurrence of high-growth. In turn, these financial factors are affected by the actual or expected operating capability of the firm and, hence, they have to be included in the analysis in order to avoid a potentially relevant source of endogeneity. In any case, drawing definite theoretical conclusions about the role of external finance in the whole picture is difficult. Indeed, different models assume different market structure for the incumbent firms, which imply very different potential profitability patterns and risk levels.

Our analysis proceeds as follows. Exploiting panel data on Italian, French, Spanish and UK incumbents, we identify high-growth companies, and within this group, those displaying persistent high-growth. It turns out that only a small proportion of firms is able to sustain superior growth performance over time. We then analyze how initial years productivity, profitability and financial factors relate with subsequent growth dynamics. We perform both non-parametric and parametric analyses. First, we investigate whether a set of key variables, taken to proxy both the operational performance and the financial status of firms, display distributional differences across high-growers, persistently high-growers and other firms. Second, we estimate several specifications of a Multinomial Probit model to identify which variables are more effective in discriminating persistent high-growth firms from "simple" high-growth and from other firms.

Our findings are challenging for both academic scholars and policy makers. Indeed, we do confirm that some structural characteristics, and productivity in particular, are significantly associated with high-growth. However, we do not find evidence of any systematic difference between high-growth and persistent high-growth firms, nor in terms of operating efficiency, neither in terms of the other considered dimensions. None of them seems to work in sustaining high-growth performance repeatedly over time. The same pattern is invariant across manufacturing and services, suggesting minor role of sectoral specificities, and it is also stable across countries, suggesting a minor role for institutional or other more macro-level factors. Further, the picture is robust to a number of extensions, including controls for firm-level innovation, disaggregated analysis by size and age, and alternative estimation methodologies.

2.2 Background literature and motivation

Our study is directly related to the empirical literature on the identification and characterization of high-growth companies. The basic "stylized facts" emerge from the seminal study by Schreyer (2000). Based on firm-level data from five OECD countries (Germany, Italy, Netherlands, Spain and Sweden) as well as from Quebec (Canada), high-growth firms are found to be (i) present in all industries and in all regions of the examined countries; (ii) more R&D intensive than "normally growing" firms or than the average incumbent; (iii) younger and smaller than the average firm. Consistent results have been confirmed by subsequent studies.

Concerning the determinants of observed high-growth performance, a stream of literature focuses on the role of factors external to the firm, such as institutions, geography, sectoral or broadly speaking macro-level variables. Among others, Davidsson and Henrekson (2002) investigate the importance of a number of institutions and policy measures such as taxation of entrepreneurial income, incentives for wealth accumulation, wage-setting and labor market regulations. The evidence, from a panel of Swedish firms, shows that the little support to dynamic firms by policy makers can hinder nascent entrepreneurship and the net employment contribution by high-growth firms. Acs and Mueller (2008) stress the role of local knowledge spillover as a driver of firm's birth rate and high-growth, concluding that metropolitan areas offer fertile ground for fast growing firms, whereas small cities facilitate new entry, but not the expansion of rapidly growing units.

More recently, scholars have started to look at more micro-level determinants of high-growth, in particular focusing on innovation-related drivers. Coad and Rao (2008) link innovation to sales growth of incumbent firms in high-tech sectors, finding that innovation is of crucial importance only for a handful of high-growth firms. Hölzl (2009) explores the relationship between R&D and superior growth performance using CIS III data for 16 countries. The findings reveal that R&D is more important to high-growth firms in countries that are closer to the technological frontier, suggesting that high-growth firms derive much of their drive from the exploitation of comparative advantages rather than from other firm-level determinants. Segarra and Teruel (2014) show, on Spanish data, that R&D investment positively affects the probability to be a high-growth firm, but internal and external R&D have asymmetric effects on the firm growth rates distribution, with internal R&D being the only type of investment having a positive impact among high-growth firms. Finally, Colombelli et al. (2014) investigate the innovation strategies of a set of European publicly traded companies by building specific indicators of the structure of knowledge (i.e. variety, coherence, and similarity). The evidence supports the conjecture that high-growth firms tend to adopt exploration rather than exploitation strategies, therefore stimulating the creation of new technological knowledge.

As the influential contributions by Delmar et al. (2003) have highlighted, however, high-growth firms do not all grow in the same way, and results can be sensitive to alternative size-growth proxies as well as to alternative criteria to identify high-growth (Delmar, 2006). As a matter of fact, there exist several types of high-growth patterns, which in turn differ by demographic characteristics such as size, industry affiliation or firm age, and by type of governance. Differences are sharp, ranging from "super absolute growers", which are typically small- and medium-sized firms operating in knowledge intensive manufacturing industries, to the "erratic one-shot growers", which are more common among small-sized firms in low-technology service sectors. It is then plausible to expect that the investigation of the determinants of high-growth can lead to different results, according to the definition of high-growth which is adopted. This consideration motivates us to adopt a multidimensional measurement criterion and to embark into a series of robustness checks with respect to possibly alternative criteria. Moreover, with respect to the studies cited above, which are essentially focused on the explanatory factors of short-run and sporadic high-growth events, we want to include the persistence of such high-growth dynamics into the picture. In this respect it is useful to take a step back and refer more closely to what existing theories suggest us to look at in the search for the drivers of high-growth.

We draw our theoretical background from models of firm-industry evolution with heterogeneous firms, originally developed within the evolutionary disequilibrium approach with no anticipating or strategic agents (see, e.g., Nelson and Winter, 1982; Silverberg et al., 1988; Dosi et al., 1995; Metcalfe, 1998), and revisited within a more standard partial equilibrium frameworks with (possibly bounded) rational agents and strategic interaction (such as in Jovanovic, 1982; Hopenhayn, 1992; Ericson and Pakes, 1995; Cooley and Quadrini, 2001; Melitz, 2003; Asplund and Nocke, 2006; Luttmer, 2007). Despite differences in the core assumptions from alternative schools of thought, these models share a common mechanism of firm selection and growth, which is made explicit in disequilibrium models, while it is implicitly described as the convergence to the equilibrium path in equilibrium models. The predicted pattern starts typically with an idiosyncratic shock to incumbent firms, or with an idiosyncratic initial endowment of entrants, as the first driver. The shock regards firm-specific unobserved factors, such as technological and organizational traits, capabilities, strategic and managerial practices, and it gets reflected into heterogeneous efficiency across firms. Next, firms with higher relative efficiency grow and gain market shares at the expenses of less efficient units, either directly via lower prices, or indirectly via increasing profits which, in combination with sounder financial conditions, grant to more productive firms the access to the resources needed to invest and pursue further growth, possibly with some time lag.

Although these models do not directly discuss high-growth performance, their implications concerning the characterization of high-growth companies are relevant to our study. The common framework, in fact, predicts that the candidate key drivers of high-growth must be searched for in terms of efficiency and profitability, so that we should expect high-growth firms to be more productive and more profitable than firms displaying moderate growth patterns.

More difficult is to draw definite predictions about the role of firms' financial conditions. Availability of more cash flow can ease the need to rely upon external finance, but it can also imply a higher capability of servicing the debt and, consequently, the possibility to sustain a larger debt exposure, *ceteris paribus*. Which effect eventually prevails is uncertain a priori, and it also depends from the capability of the credit market to correctly select the appropriate growth prospects. And, even more relevant for us, the models are uninformative about whether the kind of financial situation that eventually fosters growth can also be considered as drivers of persistent high-growth. Some scholars have even advanced the hypothesis that randomness (or "mere luck") is the most appropriate account of firms' persistent success (Barney, 1986, 1997).

The empirical literature on persistence of firm growth, on the other hand, does not provide any evidence about the determinants of persistence of high-growth performance. Traditionally, this literature looks at persistence in terms of the autocorrelation structure in the growth process, mostly with the aim to test Gibrat's Law. The results are mixed, ranging from the view that growth is indeed a random walk advanced in Geroski (2002), to the evidence of strong autocorrelation (up to the 7^{th} lag) found in Bottazzi et al. (2001). In between, positive serial autocorrelation is found by Geroski et al. (1997) on a panel of UK quoted firms, Wagner (1992) for German manufacturing companies, Weiss (1998) for the Austrian farm sector, and Bottazzi and Secchi (2003) for US manufacturing firms, while negative serial correlation is found, for instance, by Goddard et al. (2002) on Japanese quoted firms, and by Bottazzi et al. (2007) and Bottazzi et al. (2011) for Italian and French manufacturing, respectively. Findings on service firms provide a similarly mixed picture, as in Vennet (2001) on banking companies across OECD countries and Goddard et al. (2004) on US financial services.

It is only recently that empirical research considers the degree of persistence in the entire distribution of the growth rates, adopting tools like quantile autoregression or transition probability matrices. Coad (2007a) and Coad and Hölzl (2009) do observe some persistence, with small high-growth firms displaying negative autocorrelation whereas large and established companies achieving smoother dynamics. Conversely, Capasso et al. (2013) conclude that persistent outperformers are more often present among micro firms. Yet, other studies cast doubt on the very existence of persistent high-growers. Daunfeldt and Halvarsson (2013) claim that high-growth firms are basically "one-hit wonders", and document that firms experiencing strong job losses in one period are most likely to become high-growth units in the next period. The findings in Hölzl (2014) confirm that most of high-growth firms do not replicate their high-growth performance over time, and show that the degree of persistence might also depend upon the criterion adopted for the identification of such companies. All of these studies, however, do not address if more structural, economic or financial, factors are distinguishing features of persistent high-growth companies, above and beyond mere demographic characteristics such as size, age and industry affiliation.

2.3 Data and identification of persistent high-growth firms

In this Section, we present the data, our definitions of high-growth and persistent high-growth firms, and the proxies we use to measure firm characteristics. A key point is that the identification of persistence in high-growth performance requires a reasonably long period of time over which firm growth is evaluated. Our strategy is to divide the time span available in the data into two periods, and exploit the first period to measure "initial" firm characteristics, which we next seek to map into high-growth, persistent high-growth or other growth dynamics over the second period.

Sources and sample

We draw upon firm-level information from the AMADEUS dataset, a well known and widely used commercial database provided by Bureau van Dijk. AMADEUS contains detailed balance sheet and income statement information for firms in all sector of activity, covering all European countries. We have access to data on Italy, Spain, France and the UK. The edition at our disposal (2012) covers a time span of 9 years, from 2004 to 2012. However, to have a time interval with a good coverage of the variables of interest in all countries, our analysis spans the period 2004-2011. In line with previous studies (among the many, see Schreyer, 2000; Delmar et al., 2003; Bottazzi et al., 2011), our attention is on continuing incumbent firms: firms that entered midway after 2004 or exited midway before 2011 have been removed, yielding a balanced panel over the sample time window. Further, our main concern is about internal growth, and we therefore exclude those firms who experience any kind of modification of structure, such as mergers or acquisitions. The survival bias that this selection procedure might possibly introduce is minimal in this case as we will run a comparative analysis across different groups of surviving firms.[1] All the firms are classified according to their sector of principal activity, disaggregation up to 2-digits of NACE 2008 classification. The present study considers both manufacturing and services, covering a final sample of 55,454 firms.[2]

Spain has the higher number of observations followed by Italy, France and the UK. The number of small-medium enterprises (with less than 250 employees), covers approximately 95% of the entire sample. More than 60% of the sample is represented by firms belonging to Services. A screenshot of the data by countries and sectors is in the Appendix.

Definition of high-growth and persistent high-growth

The obvious preliminary step in the analysis is to choose a definition of high-growth firms and to design a strategy to identify persistent high-growth performance. There are no commonly accepted identification criteria in the literature, due to the quite disparate approaches followed in previous studies. In fact, studies on high-growth companies consider a long list of alternative size-growth indicators such as assets,

[1] In the empirical literature on firms dynamics the survival bias is often referred to as *attrition bias*. To be precise, we should not say that we compare high-growth firms with "other firms", but rather high-growth-and-surviving firms with other-and-surviving firms. In fact, it could be the case that this distinction does matter in some instances. Due to the nature of our database, however, we are not in the position to test this hypothesis. We omit any further reference to this issue in what follows.

[2] The sector of principal activity in the AMADEUS dataset is time-invariant, measured in the last available year. Manufacturing includes section C, while Services include sections G, H, I, J, K, L, M, N, R, S.

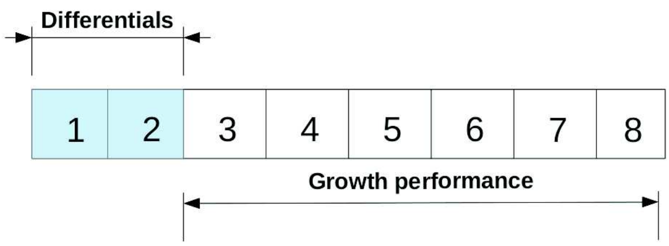

Figure 2.1: Partitioning of the sample time-period. Differences in firm attributes are measured in the first two years (2004-2005), while growth patterns are evaluated over the subsequent six years (2006-2011).

employment, market share, physical output, profits or sales. Moreover, there is a variety of possible criteria to classify a firm as high-growth, once a given size proxy is chosen. At the same time, studies looking at growth rates autocorrelation (on the average or within quantiles) do not provide a criterion to identify persistent high-growth enterprises, beyond sharing the basic intuition that these firms must experience high-growth performance – however defined – consecutively for some years.

Against this background, we implement the following choices. First, we measure annual growth of firm i in year t, in terms of the log difference

$$g_{it} = s_{it} - s_{i,t-1} \quad , \tag{2.1}$$

where

$$s_{it} = log(S_{it}) - \frac{1}{N}\sum_{i} log(S_{it}) \tag{2.2}$$

and firm size S_{it} is measured as either sales or number of employees, and the sum is computed over the N firms populating the same (2-digit) sector. In this way firm size and, thus, the growth rates are normalized by their annual sectoral average. The normalization implicitly removes common trends, such as inflation and business cycles effects in sectoral demand.

In the data, both employment and sales growth rates distributions display the usual fat-tails shape already found in previous studies. In case of employment growth, maximum likelihood estimates of the shape parameter b of a Power Exponential distribution (see Bottazzi and Secchi, 2006) range indeed from 0.45 for French firms to 0.87 for UK firms. The distributions of sales growth rates have b very close to 1 in all countries, thus close to a Laplace distribution.[3] The results are stable over the years of the sample period.

Given a sample period of 8 years, we reserve the first two years (2004-2005) to measure the firm characteristics that we want to map into growth performance, while we exploit the last six years (2006-2011) to identify firms displaying different "growth status" (see Figure 2.1). To identify high-growth (HG) firms, we compute the total growth rates experienced by each firm in terms of both sales and employment over

[3]The Subbotin family of densities possesses the following functional form: $f_s(x) = e^{-\frac{1}{b}|\frac{x-\mu}{a}|^b}/(2ab^{1/b}\Gamma(1/b+1))$, where $\Gamma(x)$ is the Gamma function. The distribution has three parameters, the mean μ, the dispersion parameter a and the shape parameter b. When $b = 2$ the distribution is a Gaussian, while it has fat tails for $b < 2$ and in particular it is a Laplace distribution if $b = 1$.

the six years spanning the second part of the sample period, and then define as HG firms those companies falling into the top 10% of the total growth rates distribution, in terms of at least one of the two growth measures. To define persistent high-growth (PHG) firms, we examine the annual growth rates of each firm, again over the last six years of the sample period, and define as persistent high-growth companies those firms falling for at least four (out of five) years into the top 10% of the yearly cross-sectional distribution of either sales or employment growth (or both). We then assign all firms not passing the two criteria to a residual category of "other firms".

With our definition we expect to have from 10% to 19% of firms classified as HG firms. The lower bound corresponds to the case of perfect cross-correlation between employment growth and sales growth, whereas the upper bound corresponds to the case in which the two growth rates measures are uncorrelated. At the same time, under the hypothesis of serially uncorrelated growth rates, we expect the fraction of PHG firms to be in between 0.045% (for perfect cross-correlation between sales growth and employment growth) and 0.65% (for no cross-correlation). Of course, if there is perfect serial correlation in growth rates, then all HG firms are also PHG firms (and the above upper and lower bounds apply).

The choice to consider both sales and employment growth in the definition of HG and PHG firms allows for a multidimensional description of the growth process, responding to the idea advanced in the literature that no single "best" indicator of size exists, with each alternative proxy measuring different aspects of the firm growth process. Indeed, sales is more a proxy of success on the market, while employment is more related to establishing capacity.[4] Definitions based on a single size indicator can considerably reduce the number of PHG firms, undermining the reliability of the empirical analysis. We have however verified that our main empirical findings do not change if we identify HG and PHG firms based exclusively on employment or exclusively on sales.

The strategy to identify HG firms through annualized average growth over some years is in line with the literature, and reflects the consideration that growth is quite unstable over time, so that one single big growth shock in one year is not enough to capture true high-growers. Depending on the study, the number of years considered may vary from 3 to 6 years, but the main idea is common to the vast majority of previous works. There is instead less consensus on whether the threshold employed to distinguish high-growth from "normal" growth needs to be in absolute value (for instance, defining as HG a firm that hires at least 100 employees) or in relative terms, that is looking at percentage growth over time relative to other firms. Our definition implicitly follows the latter approach. Absolute growth implies a bias towards larger firms, whereas the percentage measure allows smaller firms to enter the HG group. Also questionable is our choice to consider the top 10% of annualized average growth. We have however experimented with other definitions appearing in the literature (top 15%), and the main conclusions from the empirical analysis do not change.

Given the lack of a precise definition of PHG firms in previous studies, our iden-

[4]Sales and employment are indeed the most frequently chosen size proxies in the literature. They are relatively easily accessible, they can be compared within and between industries (for instance physical output do not benefit of the same property), and they are not too much related to the capital intensity of the industry (as opposed to total assets). Also notice that the inter-sectoral comparability is improved by the use of the normalized shares defined in Equation (4.2).

Table 2.1: # of High-growth and persistent high-growth firms, Manufacturing

	Italy			Spain			France			UK		
NACE	**Total**	**HG**	**PHG**	**Total**	**HG**	**PHG**	**Total**	**HG**	**PHG**	**Total**	**HG**	**PHG**
10	724	117	14	927	129	8	415	65	4	140	19	1
11	143	21	2	173	16	1	58	10	2	38	4	0
12	2	1	0	2	0	0	0	0	0	2	0	0
13	507	67	6	310	38	1	68	10	0	28	3	0
14	280	63	8	179	30	0	41	11	1	15	3	0
15	268	47	4	206	35	2	31	4	0	1	0	0
16	176	24	4	386	65	6	186	28	1	23	2	0
17	249	23	1	135	12	1	57	7	1	46	5	0
18	145	21	0	506	67	2	187	22	2	64	8	2
19	38	7	1	7	1	0	5	0	0	8	1	0
20	447	64	5	265	26	3	112	18	1	116	17	1
21	114	20	1	29	1	1	22	4	0	34	4	0
22	553	78	2	350	50	3	196	25	3	70	6	1
23	459	72	5	516	109	6	169	25	1	47	6	1
24	363	50	3	194	27	0	37	3	1	34	3	0
25	1422	178	24	1511	263	10	615	78	8	166	20	3
26	279	44	6	92	14	2	111	20	2	88	17	2
27	404	66	8	160	22	5	69	8	1	55	11	1
28	1231	178	22	442	49	6	202	22	9	139	18	4
29	173	28	1	162	30	2	69	11	2	44	2	2
30	88	16	5	28	9	0	27	4	0	28	6	0
31	310	42	7	425	73	3	80	13	0	31	4	2
32	197	29	7	169	26	4	94	13	0	184	22	3
33	115	19	2	363	50	6	290	48	3	56	9	1
Total	8687	1275	138	7537	1142	72	3141	449	42	1457	190	24

tification criterion balances between the aim to capture firms that outperform other firms continuously over a reasonably long number of years, and the time constraint imposed by the available data. We have anyhow checked that the results presented in the following empirical analysis do not change if we apply a less stringent criterion where persistent high-growth firms are defined as firms passing the 10% threshold for just 3 out of 5 years. On the other hand, being more restrictive and imposing that PHG firms pass the threshold in all years, substantially reduces the number of PHG firms, making the statistical analysis unfeasible.

Tables 2.1 and 2.2 show the number of HG and PHG firms by country and sectors, as resulting from the identification criteria we adopted to identify growth status over the period 2006-2011. The incidence of HG firms is comparable across countries, varying between 12.9% and 15.1% of the total sample. These numbers are compatible with non-zero cross-correlation between sales and employment growth. The number of PHG companies is small, ranging from 0.9% to 2% of the total sample in the different countries. This is in line with previous studies, even when adopting different identification criteria, and suggests some degree, although not perfect, of serial correlation in employment and sales growth rates. We also observe a relatively higher incidence of PHG firms within services than in manufacturing.[5]

2.3.1 Firm characteristics

We map growth status into a set of indicators of structural performance including productivity, profitability and financial condition, together with the more traditionally investigated characteristics in terms of size and age.

We compute Total Factor Productivity (TFP) as the residual of production function estimation performed through the IV-GMM modified Levinsohn-Petrin estima-

[5] The number of PHG firms increases, but it never exceeds the 5% of the total population if we consider 3 out of 5 years as the identification criterion for the definition of PHG firms.

Table 2.2: # of High-growth and persistent high-growth firms, Services

	Italy			Spain			France			UK		
NACE	**Total**	**HG**	**PHG**	**Total**	**HG**	**PHG**	**Total**	**HG**	**PHG**	**Total**	**HG**	**PHG**
45	773	82	10	1596	178	7	1115	137	8	337	27	0
46	2949	417	54	5092	776	99	2122	267	40	555	56	7
47	782	106	13	3627	530	42	1753	245	26	202	22	7
49	320	40	7	992	126	11	466	66	3	147	13	1
50	22	4	1	32	5	0	6	0	0	15	2	0
51	11	1	1	5	0	0	1	0	0	24	4	0
52	292	43	4	252	39	2	94	11	4	74	11	1
53	4	1	0	22	4	0	3	1	0	5	1	0
55	162	22	0	443	34	2	312	22	3	112	8	0
56	105	10	2	1171	151	7	456	65	5	73	8	0
58	84	10	4	137	22	2	83	18	1	61	9	0
59	16	3	0	43	9	2	31	3	0	21	3	0
60	22	2	0	29	7	0	6	0	0	7	2	0
61	18	5	0	68	11	1	16	4	0	42	10	1
62	184	33	1	237	40	6	119	21	4	135	30	3
63	72	12	2	15	1	0	20	3	0	15	2	0
64	41	10	3	33	6	0	71	22	5	157	19	4
66	17	1	1	40	6	1	8	1	0	29	6	2
68	160	27	6	218	84	13	75	42	6	61	10	2
69	70	6	1	298	21	2	57	4	0	11	2	0
70	155	32	4	125	25	1	89	17	4	282	41	5
71	99	14	9	271	45	11	150	30	3	46	6	1
72	23	4	1	20	2	0	16	2	0	15	2	1
73	85	17	1	202	43	5	68	11	1	39	6	1
74	51	10	4	188	49	6	34	2	3	44	6	2
75	0	0	0	29	5	0	1	0	0	1	0	0
77	43	7	0	174	41	6	82	12	2	81	14	1
78	10	5	0	8	1	0	7	0	1	55	14	1
79	64	12	2	117	23	3	10	1	0	32	6	0
80	37	3	0	49	8	0	15	1	1	10	1	0
81	82	17	3	234	42	2	204	26	5	26	2	1
82	91	19	3	86	14	3	78	15	2	199	37	5
90	15	3	2	40	7	1	24	8	1	9	1	0
91	6	0	0	6	4	0	11	1	0	1	0	0
92	6	1	0	87	14	1	39	1	0	10	2	0
93	74	13	5	176	31	0	52	7	1	40	4	0
94	0	0	0	6	0	0	0	0	0	8	1	0
95	28	5	0	103	21	1	35	4	1	3	0	0
96	52	6	0	239	26	4	282	30	3	102	12	2
Total	7025	1003	144	16510	2451	241	8011	1100	133	3086	400	48

tor proposed in Wooldridge (2009).[6] As our profitability proxy, we consider an index of Return on Assets (ROA), defined as operating margins over total assets. Financial conditions are taken into account by looking at two indicators capturing different dimensions of financial status: a flow measure of the capacity to meet financial obligations, computed as the ratio between interest expenses and total sales (IE/S) in a given year, and a standard measure of leverage (LEV), computed as the ratio between total debt and total assets. Age is computed exploiting information on the year of foundation. Lastly, we proxy for size through annual sales in distributional and regression analysis, and we use employment to define the size-classes in the analysis of Section 2.7.

Table 2.3 provides basic descriptive statistics in three reference years. The broad picture reflects well known differences across the countries. TFP displays indeed similar rankings across countries, with UK, France and Italian companies displaying higher average efficiency than Spanish counterpart. Notice that UK service firms are characterized by the highest average value of TFP. Concerning profitability, the average ROA is higher in the UK and France, in all years, while similar across

[6]The estimates are performed pooling firms within the same 2-digit level sector, taking number of employees and fixed tangible assets as measures of labour and capital inputs, respectively, and value added as the proxy for output, while we use the cost of material inputs as instrument to control for endogeneity of labour inputs. Alternatively, we have also considered a standard labour productivity index computed as the ratio between value added and number of employees. Results are in line with those presented along this essay.

Table 2.3: Descriptive statistics

	MANUFACTURING						SERVICES					
	2004		**2007**		**2010**		**2004**		**2007**		**2010**	
Variable	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**	**Mean**	**Std**
Italy												
log(TFP)	4.62	0.56	4.76	0.55	4.69	0.58	4.5346	0.7425	4.6905	0.7444	4.6353	0.7503
ROA	0.0235	0.0542	0.0302	0.0564	0.0190	0.0555	0.0207	0.0678	0.0255	0.0646	0.0184	0.0646
IE/S	0.0136	0.0201	0.0153	0.0222	0.0107	0.0140	0.0136	0.0326	0.0156	0.0326	0.0114	0.0396
LEV	0.6097	0.1991	0.6193	0.2000	0.5589	0.2070	0.6746	0.2129	0.6760	0.2084	0.6235	0.2250
Age	22.92	14.79	25.92	14.79	28.92	14.79	19.98	14.82	21.98	14.82	23.98	14.82
Size (sales)	24057.56	125718.60	30555.89	153283.90	28630.10	121112.50	27633.70	125823.70	33886.27	140855.40	35523.14	152103.80
Size (no. employees)	85.73	256.60	91.82	288.71	89.76	292.12	106.49	1555.98	100.35	518.16	109.49	563.37
Spain												
log(TFP)	3.64	0.61	3.83	0.58	3.74	0.62	3.5769	0.6899	3.7907	0.6893	3.7051	0.7065
ROA	0.0335	0.0879	0.0393	0.0741	-0.0096	0.1336	0.0365	0.0934	0.0402	0.1118	-0.0004	0.1266
IE/S	0.0145	0.0216	0.0171	0.0190	0.0189	0.0293	0.0127	0.0297	0.0157	0.0358	0.0167	0.0448
LEV	0.6493	0.3031	0.5328	0.2596	0.5863	0.3309	0.6857	0.3029	0.5730	0.3195	0.6177	0.3828
Age	14.24	11.20	17.24	11.20	20.24	11.20	11.96	9.07	13.96	9.07	15.96	9.07
Size (sales)	13144.34	236536.90	16900.90	333398.00	15330.09	322342.90	13398.45	433976.10	19305.39	640646.50	20100.02	710875.00
Size (no. employees)	51.52	816.53	57.24	1161.73	52.91	1115.56	55.47	1561.01	71.55	2177.18	77.15	2430.92
France												
log(TFP)	4.18	0.55	4.30	0.54	4.29	0.55	4.2131	0.6195	4.3182	0.6224	4.3485	0.6402
ROA	0.0498	0.0978	0.0594	0.1007	0.0396	0.1128	0.0583	0.1182	0.0632	0.1067	0.0492	0.1220
IE/S	0.0074	0.0089	0.0073	0.0096	0.0062	0.0201	0.0079	0.0164	0.0071	0.0126	0.0060	0.0135
LEV	0.5346	0.2122	0.5705	0.2250	0.5361	0.2670	0.5911	0.3069	0.6216	0.2768	0.5918	0.3476
Age	22.39	19.14	25.39	19.14	28.39	19.14	17.61	14.98	19.61	14.98	21.61	14.98
Size (sales)	18866.41	196887.50	22486.79	238035.60	21827.05	263363.70	30415.01	551751.80	39785.44	730391.10	43108.86	820173.70
Size (no. employees)	86.62	715.23	88.99	784.36	87.65	835.33	163.06	3132.35	227.27	4484.88	225.69	4671.12
UK												
log(TFP)	3.93	1.95	4.03	1.31	4.02	1.34	4.9161	1.1644	5.0666	1.1777	4.9549	1.1892
ROA	0.0470	0.0926	0.0537	0.1024	0.0557	0.1018	0.0490	0.1216	0.0570	0.1166	0.0428	0.3575
IE/S	0.0109	0.0154	0.0136	0.0186	0.0115	0.0216	0.0185	0.0421	0.0242	0.1362	0.0170	0.0438
LEV	0.5945	0.2449	0.5686	0.2661	0.5290	0.2601	0.6673	0.3492	0.6485	0.4025	0.6140	0.3703
Age	30.69	25.99	33.69	25.99	36.69	25.99	24.16	23.53	26.16	23.53	28.16	23.53
Size (sales)	179168.80	1106028.00	210478.00	1317031.00	212923.40	1462093.00	287814.70	1865298.00	323652.50	2034431.00	338985.20	2415635.00
Size (no. employees)	781.15	4306.20	869.86	5197.60	850.39	5342.37	1632.55	11239.11	1750.51	11613.20	1836.24	12900.66

Notes: Annual mean and standard deviation (Std) of the main firm characteristics in 3 reference years. Total Factor Productivity (TFP) via the modified Levinshon-Petrin estimator in Wooldridge (2009). Return on Assets (ROA) as operating margins-to-assets ratio. Coverage ratio as interest expenses over sales (IE/S). Leverage (LEV) as total debt over total assets. Age is computed from year of foundation. Sales are in thousands of Euros, and number of employees in units.

the other two countries. The pattern is robust across manufacturing and services. Productivity and profitability measures also reveal the fingerprints of the current financial crisis in a sharp decrease in the last reported year, common to all countries even if more modest in the UK and particularly marked in Spain. The financial ratios display interesting patterns across sectors and countries. In manufacturing, French and UK firms are relatively more solid on average along both the proxies, followed by Italian firms and with Spanish firms coming last as the most vulnerable, especially in the last year, again possibly connecting with the current crisis. Similar patterns appear in Services, but here average leverage is higher than in manufacturing, in all countries, suggesting larger debt exposure of service firms. Average firm size in terms of sales is definitively larger in the UK, similar across Italy and France, while Spanish firms are smaller on average. UK firms are also bigger in terms of employment, again with the average Spanish firms being smaller than the average French and Italian companies in the sample. This may also be part of the explanation of the comparatively lower average productivity observed for Spanish firms. Finally, notice the differences in age, with Spanish firms on average younger, reflecting the typical structure of the economy. The average age of the firms (above 10 years old in all countries) is obviously influenced by the choice to only look at incumbent firms along the considered time window.

2.4 Distributional analysis

We start by assessing statistical differences in the empirical distributions of firm characteristics across the three groups of HG, PHG and “other” firms.

Table 2.4: Distributional comparisons - HG vs. "other" firms

	Country	#Other firms	#HG	ROA	IE/S	LEV	log(TFP)	AGE	log(SIZE)
Manufacturing									
	Pooled	17490	3056	2.330	3.127*	11.251**	-0.796	-18.210**	-8.585**
	IT	7274	1275	2.564*	2.207	9.035**	-2.941*	-12.886**	-8.068**
	ES	6323	1142	0.252	2.453*	6.442**	0.749	-10.979**	-4.789**
	FR	2650	449	1.619	0.225	3.689**	0.083	-8.494**	-3.779**
	UK	1243	190	0.902	-0.079	0.316	1.500	-1.701	-2.957*
Services									
	Pooled	29112	4954	2.032	1.998	8.400**	0.917	-19.426**	-9.098**
	IT	5878	1003	0.666	0.223	4.374**	-0.659	-11.546**	-7.355
	ES	13818	2451	1.139	0.817	4.915**	3.562**	-11.743**	-6.777**
	FR	6778	1100	2.032	2.877*	4.083**	-0.664	-8.059**	-4.078**
	UK	2638	400	1.798	-0.309	2.946*	-0.851	-7.112**	-2.514

Notes: Fligner-Policello (FP) test of stochastic equality. HG firms as benchmark: a positive and significant FP statistic means that HG firms dominates. Asterisks denote significance levels: * $p < 0.01$, ** $p < 0.001$.

To reduce the impact of possible outliers, we compute the average of the variables (ROA, financial indicators, TFP, age and size in terms of sales) over the two initial years which are not used to identify HG and PHG patterns. On this set of variables, we make pairwise comparisons across the three groups of firms, by applying the Fligner and Policello (1981) procedure (hereafter, FP) to test stochastic equality between two empirical distributions. While usual tests try to assess differences up to a shift of location (in mean or in median), the FP test looks at the stochastic dominance between two compared samples, asking which of the two compared distributions is statistically more likely to have more probability mass in the right part of the support. Because of its very mild assumptions, the test is particularly suitable in case of uneven samples, it does not require equality of variances, and it allows for asymmetries. All these features are found in our data.[7]

We take the HG firms as the reference category, so that a statistically significant and positive (negative) FP statistic indicates that HG firms have larger (smaller) probability to display a larger value of the considered variable, with respect to the compared group of "other" or PHG firms. The analysis is run separately by manufacturing and services, within each country or pooling across countries. Since univariate analysis is likely to be polluted by unobserved heterogeneity, we only discuss highly significant (more than 1%) differences.

In Table 2.4 we compare HG firms versus "other firms". Our first noticeable finding is that demographic characteristics are confirmed to distinguish outstanding growing firms. Indeed, in agreement with the literature, HG firms are smaller and younger in distribution. The result is generally valid in our data, across countries and sectors. The exception is the UK, where HG and other firms are statistically

[7] A drawback of the FP test, common to other non-parametric tests, is the need to have more data points to achieve the same power of basic tests for differences in mean. We have verified that our conclusions from the FP test, however, remain valid if we use a standard two-sample Student's t test for equality of the mean across samples with unequal variances, and also if we employ a Wilcoxon-Mann-Whitney test for equality of medians.

Table 2.5: Distributional comparisons - HG vs. PHG firms

	Country	#HG	#PHG	ROA	IE/S	LEV	log(TFP)	AGE	log(SIZE)
Manufacturing									
	Pooled	3056	276	1.323	-1.594	-4.365**	1.978	4.794**	3.784**
	IT	1275	138	2.029	-1.967	-3.407**	4.870**	3.861**	7.732**
	ES	1142	72	0.156	-0.944	-1.914	2.284	4.679**	3.742**
	FR	449	42	-0.251	0.323	-2.002	0.899	1.523	1.731
	UK	190	24	-0.555	-0.832	-3.012*	-0.557	2.040	0.375
Services									
	Pooled	4954	566	0.773	-0.849	-3.458**	0.148	5.129**	1.933
	IT	1003	144	1.579	-1.177	-2.427	2.532	2.366	3.330**
	ES	2451	241	0.012	-1.469	-1.857	1.437	4.572**	2.646*
	FR	1100	133	0.324	1.177	-1.631	1.133	1.731	1.415
	UK	400	48	-1.261	0.126	1.892	-2.412	3.318**	1.484

Notes: Fligner-Policello (FP) test of stochastic equality. HG firms as benchmark: a positive and significant FP statistic means that HG firms dominates. Asterisks denote significance levels: * $p < 0.01$, ** $p < 0.001$.

equal in terms of age in manufacturing, and in terms of size in services.

The picture is more nuanced when we turn to more structural characteristics. First, we find a very weak association of high-growth performance with profitability. Indeed, the null of equality is rejected only for Italian manufacturing firms, with a positive FP statistic in agreement with the prediction that HG firms are more profitable. No differences against other firms are detected in all other countries, irrespective of the sector. Second, we find similarly lacking evidence of significant differences in terms of efficiency. The exceptions in this case are in Italian manufacturing, where HG firms have lower TFP in distribution, and in Spanish services, where HG firms appear as more efficient than the group of other firms. The negative sign for Italian manufacturing firms is somewhat at odds with theoretical predictions, but it could be driven by the strong correlation between capital intensity and size. The multivariate regression analysis in the following sections will shed light on this conjecture. Third, moving to financial factors, the estimates on the IE/S ratio provide mixed results. HG firms active in Spanish manufacturing and in French services have a larger share of sales "absorbed" by annual debt servicing, while we do not observe statistically significant differences with respect to the group of other firms in all the other country-sector combinations. Conversely, we find robust evidence that HG firms differ in terms of leverage. In all cases, with the only exception of UK services, we obtain strongly significant and positive statistics, implying that HG firms feature an heavier reliance on debt as compared to own assets. Since both leverage and IE/S ratio are here measured ex-ante, in the years before the actual HG status is realized, the implication is that will-be HG firms do have access to external finance, so they are not completely credit rationed, but have to pay more for it. Overall, we find signals that structural characteristics matter, but the evidence is not that conclusive as one could expect from theory. Beyond age and size, only a relatively high degree of ex-ante indebtedness is clearly standing out as a distinguishing feature of high-growth firms. Variation of this picture across sectors and countries is minor.

Even more striking, structural characteristics play an even weaker role in the comparison between HG and PHG firms, reported in Table 2.5. Profitability and the IE/S ratio never display statistically significant differences across the two groups of firms, in all countries and across both manufacturing and services. Second, PHG firms tend to be less productive than HG firms in Italian manufacturing, while in all other country-sector combinations PHG and HG firms have statistically identical TFP distribution. Third, we can say that leverage displays some, albeit very limited, discriminatory power. PHG firms are more indebted than HG firms, in proportion to their total assets, in Italy and in the UK in manufacturing, and if we pool the data altogether in services. Finally, the results cast also doubts on the role of size and age. In manufacturing, PHG firms tend to be younger and smaller in Italy and Spain, while we cannot reject the equality of age and sales distributions for France and the UK. In services, age plays a role in Spain and the UK only, again with the expected sign, while size only matters in Italy and Spain. Overall, a fair reading of the distributional analysis is that persistent high-growth firms do not seem to differ in any systematic way from high-growth firms.

2.5 Regression analysis

We next turn to a more standard multivariate regression analysis, investigating the role of firm characteristics in predicting the probability that a firm belongs to the three groups of HG, PHG and "other firms". The dependent variable is a discrete indicator

$$y_i = \begin{cases} 0 & \text{if firm } i \text{ is "other firm"}, \\ 1 & \text{if firm } i \text{ is HG firm}, \\ 2 & \text{if firm } i \text{ is PHG firm}, \end{cases} \tag{2.3}$$

according to our classification of growth status observed in the second part of the sample period.

The probability to belong to each category is then modeled as a function of a vector $\mathbf{v}_i$ of explanatory variables

$$P_j := \text{Prob}\left[y_i = j \mid \mathbf{v}_i\right] = F(\beta_j \, \mathbf{v}_i) \ , \tag{2.4}$$

with $\beta_j, (j = 0, 1, 2)$ the coefficient to be estimated. The vector of explanatory variables includes all the dimensions of firm characteristics and performance: ROA, TFP, IE/S, leverage, age and size (as sales). As we did in the above distributional analysis, all the regressors enter as their average across 2004-2005. Regressors are z-scored, allowing to compare coefficient magnitudes across variables and also across specifications. The lag between growth status (measured in the second time span) and initial firm characteristics (measured in the first time span) reduces potential simultaneity bias.

We estimate a Multinomial Probit model, via full maximum likelihood. This estimation method is a natural choice, since the growth status is unordered (we might have inverted the assignments without any effect) and, by construction of the three groups, we cannot hold the independence from irrelevant alternatives assumption required by Logit-type estimators.[8] Moreover, we do not apply ordered probit or logit models, since these models assume that the observed PHG or HG status of a

[8] See Section 2.8 for a discussion of alternative methods.

Table 2.6: Multinomial Probit - Main estimates

	Pooled (1)	Italy (2)	Spain (3)	France (4)	UK (5)
Other firms:					
ROA	-0.0260*	-0.0388	-0.0076	-0.0359	-0.0378
	(0.0110)	(0.0222)	(0.0162)	(0.0250)	(0.0346)
IE/S	-0.0289**	0.0008	-0.0427**	-0.0596*	0.0425
	(0.0110)	(0.0239)	(0.0163)	(0.0240)	(0.0324)
LEV	-0.1042***	-0.1803***	-0.0724***	-0.1056***	-0.1023**
	(0.0110)	(0.0209)	(0.0137)	(0.0207)	(0.0363)
log(TFP)	-0.1263***	-0.1551***	-0.1699***	-0.0495	-0.0423
	(0.0150)	(0.0255)	(0.0297)	(0.0290)	(0.0388)
AGE	0.1740***	0.2227***	0.1746***	0.1399***	0.1692***
	(0.0117)	(0.0228)	(0.0193)	(0.0294)	(0.0377)
log(SIZE)	0.1473***	0.2914***	0.1146***	0.0796**	0.0919*
	(0.0146)	(0.0246)	(0.0246)	(0.0282)	(0.0365)
Service dummy	0.0141	0.0146	0.0101	0.0263	0.0105
	(0.0184)	(0.0339)	(0.0347)	(0.0512)	(0.0744)
Country dummies	yes	-	-	-	-
PHG firms:					
ROA	0.0081	0.0097	0.0102	-0.0214	0.0631
	(0.0184)	(0.0385)	(0.0342)	(0.0443)	(0.0466)
IE/S	0.0195	0.0078	0.0350	-0.1260*	0.1186*
	(0.0110)	(0.0329)	(0.0204)	(0.0626)	(0.0514)
LEV	0.0252	0.1121*	-0.0023	0.0074	0.0506
	(0.0151)	(0.0498)	(0.0285)	(0.0262)	(0.0520)
log(TFP)	-0.0260	-0.0395	-0.0387	-0.0391	0.1735*
	(0.0238)	(0.0548)	(0.0404)	(0.0667)	(0.0718)
AGE	-0.0986**	-0.0576	-0.1261	-0.0453	-0.2456
	(0.0354)	(0.0608)	(0.0659)	(0.0765)	(0.1638)
log(SIZE)	-0.1342***	-0.2083***	-0.0831	-0.0792	-0.1104
	(0.0270)	(0.0571)	(0.0584)	(0.0633)	(0.0691)
Service	0.1724***	0.1954**	0.2342***	0.1447	-0.0377
	(0.0449)	(0.0674)	(0.0602)	(0.0971)	(0.1400)
Country dummies	yes	-	-	-	-
Observation	55,454	15,712	24,047	11,152	4,543
Log Pseudo-likelihood	-26,544.17	-7,523.93	-11,549.00	-5,296.98	-2,066.29
Chi-2	1,199.867	646.898	272.015	144.793	113.719

Notes: Coefficient estimates of Multinomial Probit regression from different specifications of model (2.4), taking High-Growth firms as the baseline category. Explanatory variables in z-scores. Bootstrapped standard errors (100 runs) in parenthesis below the coefficients. Asterisks denote significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

firm is just the result of a differential reaction to different values of the independent variables, while the underlying mechanism connecting firm characteristics to different growth patterns is the same across HG and PHG firms. Instead, we want to test, and not assume, whether PHG firms can be considered the results of a "stronger" treatment effect.

Despite some computational burden related to the underlying specification of a multivariate Normal distribution, the estimation outcome is simple to interpret as the multiple choice version of a usual binary choice probit, once a baseline category is chosen. In presenting the results, we select the HG firms as the baseline, so that a positive (negative) estimated coefficient capture if the corresponding regressor increases (decreases) the odds of belonging to the group of "other firms" or to the group of PHG firms, with respect to be in the HG group. We report estimated

Table 2.7: Multinomial Probit - Manufacturing

	Pooled (1)	Italy (2)	Spain (3)	France (4)	UK (5)
Other firms:					
ROA	-0.0540**	-0.1094***	-0.0086	-0.0184	-0.0251
	(0.0194)	(0.0283)	(0.0281)	(0.0468)	(0.0711)
IE/S	-0.0005	0.0302	-0.0208	-0.0265	0.0400
	(0.0186)	(0.0344)	(0.0344)	(0.0450)	(0.0681)
LEV	-0.1666***	-0.2490***	-0.1206***	-0.1588***	-0.0351
	(0.0166)	(0.0287)	(0.0291)	(0.0419)	(0.0648)
log(TFP)	-0.1737***	-0.1550***	-0.2129***	-0.1700**	-0.1323
	(0.0242)	(0.0362)	(0.0475)	(0.0583)	(0.0766)
AGE	0.1967***	0.2134***	0.2252***	0.2229***	0.0492
	(0.0223)	(0.0324)	(0.0339)	(0.0645)	(0.0641)
log(SIZE)	0.2224***	0.3111***	0.1426**	0.1963***	0.1977**
	(0.0244)	(0.0327)	(0.0438)	(0.0527)	(0.0672)
Country dummies	yes	-	-	-	-
PHG firms:					
ROA	0.0475	0.0246	0.0496	0.0285	0.1132
	(0.0329)	(0.0627)	(0.0587)	(0.1056)	(0.0952)
IE/S	0.0202	0.0323	0.0159	-0.1333	0.0642
	(0.0278)	(0.0674)	(0.0567)	(0.0957)	(0.1432)
LEV	0.0516*	0.1229	-0.0380	0.0726	0.1624
	(0.0256)	(0.0642)	(0.0641)	(0.0665)	(0.1004)
log(TFP)	-0.0573	-0.0206	-0.0452	-0.1266	0.0875
	(0.0545)	(0.0864)	(0.0956)	(0.1595)	(0.1110)
AGE	-0.1199*	-0.0307	-0.3510*	-0.0427	-0.2112
	(0.0593)	(0.0725)	(0.1441)	(0.1514)	(0.2204)
log(SIZE)	-0.2274***	-0.3232***	-0.2001	-0.0587	-0.0381
	(0.0585)	(0.0827)	(0.1233)	(0.1293)	(0.1350)
Country dummies	yes	-	-	-	-
Observations	20,822	8,687	7,537	3,141	1,457
Log Pseudo-likelihood	-9,754.81	-4,069.50	-3,506.94	-1,460.48	-669.87
Chi-2	587.089	401.984	154.803	80.845	25.995

Notes: Coefficient estimates of Multinomial Probit regression from different specifications of model (2.4), taking High-Growth firms as the baseline category. Explanatory variables in z-scores. Bootstrapped standard errors (100 runs) in parenthesis below the coefficients. Asterisks denote significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

coefficients together with robust standard errors computed via bootstrap. Given the relatively large number of regressors, we avoid to comment 10% significance levels, as they are likely to be spurious.[9]

Table 2.6 shows our main estimates, where we pool together manufacturing and services, and thus regressors also include a dummy for service firms. The top panel report the estimates obtained for the odds of being in the “other firms” category

[9]Since the variables are in z-scores, the marginal effects at the sample mean of the covariates are proportional to the corresponding coefficients. Standard errors are obtained out of 100 bootstrap runs, which were enough to obtain convergence. We have also applied the usual sandwich-White type of robust standard errors, obtaining the same patterns of statistical significance. The same conclusion holds with respect to all the results presented in the rest of the paper.

against being an HG firm, while results in the bottom panel show how firm characteristics associate with the odds of being a PHG firm rather than a HG firm.

In Column 1, we pool all the data across countries. The signs and significance of the coefficient obtained for the "other firms" imply that HG firms are more profitable, pay higher interests per unit of sales, have a disproportionately larger debt-to-asset ratio, and are more efficient. The results complement the univariate distributional analysis, confirming the relevance of leverage, but they also match with the theoretical expectation that profitability and productivity performance do have a discriminatory power. This was not the case in the above distributional comparisons, where we were not controlling for other firm characteristics. Moreover, we still observe a significant role of both age and size, with "other firms" being older and larger than HG firms. Age, in particular, displays the stronger association (coefficient is 0.174), followed by size (0.147). TFP and Leverage have a weaker and similar coefficient (about 0.11, considering the standard errors), while ROA and IE/S play a secondary role, with much smaller coefficients (about 0.02) and weaker statistical significance.

The picture changes completely when we look at the estimated association of regressors with the probability to fall into the PHG category. In this model, indeed, none of the structural firm attributes displays a statistically significant coefficient. The estimates for age and size are negative and significant, matching previous evidence that PHG firms are more likely to be younger and smaller companies than HG firms. Also notice that the service dummy has a positive and significant coefficient, reflecting the fact that the proportion of persistent high-growth firms is larger within services than in manufacturing.

Pooling the data helps increasing the number of observations available for the estimation, especially given the relatively small number of PHG firms. In columns 2-5, we re-estimate the same specification separately for each country. This provides some more flexibility than country dummies in evaluating whether results are invariant to institutional and other country-specific factors. Results fully agree with the picture from the pooled analysis. First, looking at the HG vs "other firms" results, we confirm that leverage and productivity play the major role, together with age and size, in distinguishing HG firms from "other firms". Italy is the country where coefficients are larger for all variables, with size and age having a strong relevance (point estimate about 0.29 and 0.22). Age is the factor with stronger association with being HG in Spain, France and the UK. IE/S is barely significant in Spain and France, while profitability is never significant.

Second, the estimates for the HG/PHG odds confirm the lack of any systematic association between persistence of high-growth performance and all the considered firm characteristics. There are few exceptions, which are however barely significant: IE/S in France and the UK, leverage in Italy, and TFP in the UK. Moreover, age turns out as never statistically significant, while size has a relatively large and significant coefficient only in the case of Italian firms.

In Table 2.7 and 2.8 we present a disaggregated analysis distinguishing by manufacturing and services. Results confirm the core evidence. First, in manufacturing, efficiency and financial leverage, together with size and age emerge as the key characteristics distinguishing HG from "other firms", with HG firms generally more efficient, more indebted relatively to own assets, and also smaller and younger. Profitability has a role only in Italy. Yet, PHG firms do not differ systematically from

Table 2.8: Multinomial Probit - Services

	Pooled (1)	Italy (2)	Spain (3)	France (4)	UK (5)
Other firms:					
ROA	-0.0068	0.0387	-0.0076	-0.0310	-0.0452
	(0.0140)	(0.0344)	(0.0205)	(0.0281)	(0.0478)
IE/S	-0.0450***	-0.0160	-0.0566***	-0.0708**	0.0317
	(0.0125)	(0.0334)	(0.0170)	(0.0229)	(0.0415)
LEV	-0.0707***	-0.1076***	-0.0511*	-0.0871***	-0.1340**
	(0.0127)	(0.0257)	(0.0199)	(0.0253)	(0.0483)
log(TFP)	-0.1108***	-0.1502***	-0.1583***	-0.0230	0.0059
	(0.0176)	(0.0445)	(0.0304)	(0.0272)	(0.0481)
AGE	0.1614***	0.2325***	0.1552***	0.1131**	0.2411***
	(0.0168)	(0.0388)	(0.0250)	(0.0360)	(0.0596)
log(SIZE)	0.1109***	0.2587***	0.1053***	0.0498	0.0336
	(0.0167)	(0.0421)	(0.0260)	(0.0305)	(0.0428)
Country dummies	yes	-	-	-	-
PHG firms:					
ROA	-0.0016	0.0058	-0.0045	-0.0325	0.0424
	(0.0229)	(0.0473)	(0.0396)	(0.0524)	(0.0527)
IE/S	0.0235	-0.0043	0.0509*	-0.1216	0.1201
	(0.0141)	(0.0458)	(0.0213)	(0.0852)	(0.0736)
LEV	0.0207	0.1030	0.0053	-0.0078	0.0250
	(0.0165)	(0.0618)	(0.0322)	(0.0304)	(0.0741)
log(TFP)	-0.0146	-0.0387	-0.0375	-0.0164	0.2164*
	(0.0294)	(0.0821)	(0.0379)	(0.0776)	(0.0895)
AGE	-0.0901	-0.0713	-0.0978	-0.0495	-0.3023
	(0.0471)	(0.0719)	(0.0762)	(0.0801)	(0.2647)
log(SIZE)	-0.0894*	-0.1224	-0.0500	-0.0780	-0.1503
	(0.0360)	(0.0717)	(0.0494)	(0.0734)	(0.0803)
Country dummies	yes	-	-	-	-
Observations	34,632	7,025	16,510	8,011	3,086
Log Pseudo-likelihood	-16,741.13	-3,436.38	-8,027.73	-3,826.12	-1,388.81
Chi-2	310.385	187.897	168.073	53.093	43.664

Notes: Coefficient estimates of Multinomial Probit regression from different specifications of model (2.4), taking High-Growth firms as the baseline category. Explanatory variables in z-scores. Bootstrapped standard errors (100 runs) in parenthesis below the coefficients. Asterisks denote significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

HG firms along any of the included dimensions, with the only exception of size in Italy. Second, the picture is quite similar when we look at services. The main differences with manufacturing are that in services the cost of debt servicing (IE/S) is significantly higher for high-growth firms in most countries (not in the UK), while profitability is never significant in this sector. But we fully confirm that PHG firms do not differ from HG firms under any of the firm attributes. The result is even stronger than in manufacturing, since here even size does not display any statistically significant coefficient. Our general conclusion is that the drivers of growth predicted by the theory, productivity and leverage in particular, play some role in shaping high-growth patterns, whereas they do not discriminate persistent from sporadic high-growth. Notice that this absence of statistical correlation also downplays the concerns with endogeneity and omitted variables, which, if any, would bias our estimates upward.

2.6 The role of innovation

The recent empirical literature on high-growth firms suggests that innovativeness might represent a distinguishing feature of this type of firms. High-growth firms tend to be more concentrated in high-tech sectors or in sectors closer to the technological frontier, and they also tend to be more involved than other firms into R&D and patenting activity. There is no direct evidence, however, about the innovation patters of persistent high-growth firms. In this section we replicate our regression analysis including among the regressors the value of intangible assets (INTASS) as a firm-level proxy for innovativeness.[10]

In Table 2.9 we show the results of the specification pooling data across manufacturing and services. Point estimates and patterns of statistical significance for the economic and financial variables are substantially identical to the main results presented in the previous Section. The picture is basically unchanged also concerning size and age, although adding intangibles affect the estimated coefficient of these two latter variables. Intangibles present themselves a negative and significant coefficient in the odds of being "other" vs HG firms, at least in some cases (pooled analysis, and in Italy and the UK), in accordance with the evidence that high-growth firms tend to be more innovative. However, intangibles do not have any statistically significant discriminatory power in distinguishing persistent high-growth performers from "simple" high growers.

We find consistent results also when distinguishing manufacturing and services, in Table 2.10 and Table 2.11. Once again, structural and demographic variables can discriminate between HG and other firms, but they have limited role in distinguishing persistent high-growth firms. The only noticeable difference with respect to the aggregate analysis is in services, where intangible assets are found to increase the probability to be in the HG group (in Italy and in the UK) or in the PHG category (in France), but at very low levels of statistical significance.[11]

Overall, we confirm our key finding that firms who display a subsequent pattern of persistent high-growth performance are neither more productive, nor more profitable, nor characterized by peculiar financial conditions in the initial years.

[10] The Amadeus data are known to lack information about R&D expenditures, while information on patenting activity is available only for very few firms. Intangible assets have instead a good coverage and represent a suitable alternative proxy, repeatedly adopted in innovation studies, since, e.g., Hall (1999).

[11] As an alternative way to explore the role of innovation, we have also estimated our baseline Multinomial Probit augmented with dummy indicators identifying groups of sectors by their innovative characteristics. For manufacturing, we have experimented with dummies for Low vs. High-Tech industries (EUROSTAT classification) and distinguishing the four classical Pavitt (1984) taxonomy classes. For services, we distinguished KIS vs. non-KIS sectors (EUROSTAT taxonomy). The results about the main structural and demographic characteristics replicate our main conclusions. Moreover, sectoral dummies turn out as statistically significant only in some specific cases, and thus provide a weak contribution to predict persistence of high-growth status.

Table 2.9: Intangible assets - Multinomial Probit, main estimates

	Pooled (1)	Italy (2)	Spain (3)	France (4)	UK (5)
Other firms:					
ROA	-0.0291**	-0.0492*	-0.0081	-0.0379	-0.0388
	(0.0105)	(0.0218)	(0.0170)	(0.0229)	(0.0378)
IE/S	-0.0235*	0.0102	-0.0412*	-0.0533**	0.0584
	(0.0096)	(0.0236)	(0.0163)	(0.0205)	(0.0419)
LEV	-0.1056***	-0.1836***	-0.0719***	-0.1120***	-0.1073***
	(0.0103)	(0.0225)	(0.0159)	(0.0235)	(0.0320)
log(TFP)	-0.1257***	-0.1534***	-0.1694***	-0.0519	-0.0443
	(0.0151)	(0.0271)	(0.0274)	(0.0291)	(0.0390)
AGE	0.1700***	0.2093***	0.1743***	0.1376***	0.1551***
	(0.0122)	(0.0245)	(0.0248)	(0.0308)	(0.0390)
log(SIZE)	0.1699***	0.3371***	0.1203***	0.1003**	0.1500***
	(0.0151)	(0.0290)	(0.0255)	(0.0321)	(0.0414)
log(INTASS)	-0.0454***	-0.0857***	-0.0120	-0.0434	-0.1135**
	(0.0121)	(0.0210)	(0.0170)	(0.0238)	(0.0382)
Service dummy	0.0140	0.0137	0.0101	0.0255	0.0096
	(0.0195)	(0.0384)	(0.0281)	(0.0432)	(0.0675)
Country dummies	yes	-	-	-	-
PHG firms:					
ROA	0.0053	0.0120	0.0065	-0.0244	0.0623
	(0.0170)	(0.0419)	(0.0300)	(0.0463)	(0.0466)
IE/S	0.0235*	0.0079	0.0388*	-0.1061	0.1252*
	(0.0111)	(0.0357)	(0.0198)	(0.0621)	(0.0615)
LEV	0.0238	0.1141*	-0.0007	-0.0041	0.0493
	(0.0134)	(0.0454)	(0.0292)	(0.0266)	(0.0550)
log(TFP)	-0.0256	-0.0376	-0.0353	-0.0434	0.1744**
	(0.0223)	(0.0563)	(0.0361)	(0.0539)	(0.0637)
AGE	-0.1026**	-0.0534	-0.1313	-0.0479	-0.2424
	(0.0322)	(0.0529)	(0.0775)	(0.0894)	(0.1590)
log(SIZE)	0.1147***	-0.2148***	-0.0580	-0.0407	-0.0960
	(0.0251)	(0.0482)	(0.0512)	(0.0576)	(0.0788)
log(INTASS)	-0.0409	0.0173	-0.0624	-0.0865	-0.0243
	(0.0237)	(0.0430)	(0.0386)	(0.0501)	(0.0736)
Service dummy	0.1722***	0.1957**	0.2356**	0.1402	-0.0340
	(0.0441)	(0.0712)	(0.0769)	(0.0904)	(0.1527)
Country dummies	yes	-	-	-	-
Observations	55,454	15,712	24,047	11,152	4,543
Log Pseudo-likelihood	-26,534.74	-7,514.24	-11,547.47	-5,294.18	-2,061.79
Chi-2	1,121.293	762.728	264.793	114.849	90.821

Notes: Coefficient estimates of Multinomial Probit regression from different specifications of model (2.4), taking High-Growth firms as the baseline category. Explanatory variables in z-scores. Bootstrapped standard errors (100 runs) in parenthesis below the coefficients. Asterisks denote significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

2.7 Size and age

The statistical exercises presented so far provide mixed results on the role of age and size, and in particular concerning the discriminatory power of such demographic characteristics across HG and PHG firms. Despite there is some variation across sectors and countries, distributional comparisons suggest that PHG firms tend to be smaller and younger than HG firms, especially in manufacturing. On the contrary, in the Multinomial Probit regressions we do not find systematic evidence that persistent high-growth firms differ from high-growth firms in terms of age and size.

Table 2.10: Intangible assets - Multinomial Probit, Manufacturing

	Pooled (1)	Italy (2)	Spain (3)	France (4)	UK (5)
Other firms:					
ROA	-0.0569**	-0.1149***	-0.0087	-0.0207	-0.0263
	(0.0192)	(0.0336)	(0.0329)	(0.0474)	(0.0596)
IE/S	0.0039	0.0381	-0.0205	-0.0162	0.0590
	(0.0193)	(0.0354)	(0.0409)	(0.0394)	(0.0768)
LEV	-0.1664***	-0.2487***	-0.1204***	-0.1673***	-0.0406
	(0.0165)	(0.0295)	(0.0331)	(0.0474)	(0.0635)
log(TFP)	-0.1699***	-0.1489***	-0.2127***	-0.1701**	-0.1264
	(0.0244)	(0.0389)	(0.0424)	(0.0655)	(0.0813)
AGE	0.1929***	0.2054***	0.2251***	0.2175**	0.0394
	(0.0224)	(0.0287)	(0.0460)	(0.0691)	(0.0615)
log(SIZE)	0.2400***	0.3350***	0.1438**	0.2289***	0.2349***
	(0.0259)	(0.0354)	(0.0458)	(0.0689)	(0.0674)
log(INTASS)	-0.0372	-0.0497	-0.0025	-0.0693	-0.0836
	(0.0218)	(0.0328)	(0.0322)	(0.0457)	(0.0672)
Country dummies	yes	-	-	-	-
PHG firms:					
ROA	0.0492	0.0272	0.0446	0.0313	0.1128
	(0.0325)	(0.0561)	(0.0661)	(0.0892)	(0.0997)
IE/S	0.0211	0.0360	0.0188	-0.1744	0.0547
	(0.0315)	(0.0693)	(0.0584)	(0.1083)	(0.1587)
LEV	0.0523*	0.1220	-0.0348	0.0917	0.1594
	(0.0250)	(0.0711)	(0.0584)	(0.0732)	(0.0984)
log(TFP)	-0.0582	-0.0224	-0.0388	-0.1282	0.0805
	(0.0552)	(0.0842)	(0.1005)	(0.1373)	(0.1132)
AGE	-0.1167	-0.0278	-0.3561*	-0.0351	-0.2049
	(0.0600)	(0.0687)	(0.1568)	(0.1679)	(0.2212)
log(SIZE)	-0.2342***	-0.3304***	-0.1735	-0.1249	-0.0551
	(0.0640)	(0.0863)	(0.1185)	(0.1431)	(0.1323)
log(INTASS)	0.0185	0.0221	-0.0641	0.1476	0.0447
	(0.0497)	(0.0648)	(0.0697)	(0.1088)	(0.1387)
Country dummies	yes	-	-	-	-
Observations	20,822	8,687	7,537	3,141	1,457
Log Pseudo-likelihood	-9,752.11	-4,067.63	-3,506.58	-1,457.33	-668.84
Chi-2	666.365	451.100	143.440	104.552	25.179

Notes: Coefficient estimates of Multinomial Probit regression from different specifications of model (2.4), taking High-Growth firms as the baseline category. Explanatory variables in z-scores. Bootstrapped standard errors (100 runs) in parenthesis below the coefficients. Asterisks denote significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 2.11: Intangible assets - Multinomial Probit, Services

	Pooled (1)	Italy (2)	Spain (3)	France (4)	UK (5)
Other firms:					
ROA	-0.0099	0.0224	-0.0083	-0.0327	-0.0455
	(0.0158)	(0.0288)	(0.0214)	(0.0269)	(0.0398)
IE/S	-0.0388**	-0.0009	-0.0547**	-0.0662**	0.0444
	(0.0119)	(0.0360)	(0.0177)	(0.0214)	(0.0443)
LEV	-0.0730***	-0.1194***	-0.0505*	-0.0919***	-0.1396*
	(0.0138)	(0.0278)	(0.0206)	(0.0274)	(0.0543)
log(TFP)	-0.1123***	-0.1605***	-0.1580***	-0.0254	-0.0015
	(0.0184)	(0.0437)	(0.0348)	(0.0331)	(0.0437)
AGE	0.1576***	0.2133***	0.1546***	0.1118***	0.2242***
	(0.0165)	(0.0352)	(0.0281)	(0.0302)	(0.0529)
log(SIZE)	0.1345***	0.3291***	0.1123***	0.0646	0.0993*
	(0.0209)	(0.0467)	(0.0286)	(0.0340)	(0.0489)
log(INTASS)	-0.0471***	-0.1195***	-0.0151	-0.0308	-0.1201*
	(0.0137)	(0.0317)	(0.0214)	(0.0268)	(0.0477)
Country dummies	yes	-	-	-	-
PHG firms:					
ROA	-0.0062	0.0076	-0.0079	-0.0391	0.0421
	(0.0234)	(0.0509)	(0.0414)	(0.0520)	(0.0704)
IE/S	0.0299	-0.0040	0.0553*	-0.0815	0.1273
	(0.0162)	(0.0497)	(0.0232)	(0.0740)	(0.0793)
LEV	0.0172	0.1052	0.0067	-0.0307	0.0241
	(0.0187)	(0.0695)	(0.0347)	(0.0346)	(0.0792)
log(TFP)	-0.0164	-0.0375	-0.0354	-0.0274	0.2147*
	(0.0315)	(0.0825)	(0.0436)	(0.0674)	(0.0894)
AGE	-0.0968*	-0.0664	-0.1030	-0.0517	-0.3012
	(0.0459)	(0.0819)	(0.0735)	(0.0937)	(0.2920)
log(SIZE)	-0.0584	-0.1230	-0.0256	-0.0105	-0.1238
	(0.0313)	(0.0864)	(0.0534)	(0.0678)	(0.0970)
log(INTASS)	-0.0667*	0.0069	-0.0603	-0.1579**	-0.0453
	(0.0267)	(0.0639)	(0.0447)	(0.0611)	(0.1064)
Country dummies	yes	-	-	-	-
Observations	34,632	7,025	16,510	8,011	3,086
Log Pseudo-likelihood	-16,733.67	-3,428.25	-8,026.59	-3,821.87	-1,385.56
Chi-2	437.888	200.566	176.130	70.426	57.515

Notes: Coefficient estimates of Multinomial Probit regression from different specifications of model (2.4), taking High-Growth firms as the baseline category. Explanatory variables in z-scores. Bootstrapped standard errors (100 runs) in parenthesis below the coefficients. Asterisks denote significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 2.12: Distributional comparison by Age and Size - Manufacturing

	Country	#HG	#PHG	ROA	IE/S	LEV	log(TFP)
Young	Pooled	626	87	0.948	0.613	-1.381	0.713
	IT	223	39	1.031	0.796	-1.277	3.235*
	ES	288	33	0.074	-0.423	-0.113	0.070
	FR	91	12	-0.401	1.917	-2.166	0.438
	UK	24	3	0.520	-0.233	-0.605	0.442
Middle/Old	Pooled	2430	189	-2.059	-3.324**	0.857	1.726
	IT	1052	99	1.726	-2.532	-2.661*	3.205*
	ES	854	39	0.121	-0.693	-1.168	2.380
	FR	358	30	0.128	-0.602	-1.068	0.538
	UK	166	21	-1.015	-0.832	-2.856*	-0.699
Micro-Small	Pooled	2355	239	1.815	-1.391	-3.295**	0.019
	IT	906	126	2.316	-1.558	-2.454	3.377**
	ES	1042	71	-0.057	-0.967	-1.715	1.454
	FR	364	36	-0.148	0.220	-2.095	0.752
	UK	43	6	-1.928	0.395	-0.673	-0.995
Medium-Large	Pooled	701	37	-1.173	-0.320	-2.105	1.602
	IT	369	12	-1.416	-0.572	-1.182	0.688
	ES	100	1	-	-	-	-
	FR	85	6	-0.384	0.075	-0.232	-0.138
	UK	147	18	0.478	-1.247	-3.025*	-0.109

Notes: Fligner-Policello (FP) test of stochastic equality. Young firms are ≤ 5 years old in 2004. Micro-Small firms defined as firms with < 50 employees in 2004. HG firms as benchmark within each class: a positive and significant FP statistic means that HG firms dominates. Asterisks denote significance levels: * $p < 0.01$, ** $p < 0.001$.

Motivated by the emphasis given to age and size in the literature, we propose a further look at the role of these firm attributes. We want to explore whether age and size interplay with our negative result about the lacking association between persistence of high-growth and economic or financial attributes. That is, it may be the case that although efficiency, profitability and financial indicators *on average* cannot discriminate between HG and PHG firms, the association of the same variables with PHG status vary across firms of different size and age. We therefore propose a further exercise dividing the firms into age and size classes, and then, within each size and age class, repeat the Fligner-Policello test of stochastic dominance to compare the empirical distribution of productivity, profitability and financial indicators across HG and PHG firms.[12] To have a reasonable number of observations in each class, we build two size classes based on employment exploiting the standard EUROSTAT distinction between Micro-Small ($<$ 50 employees) and Medium-Large ($\geq$ 50 employees) firms, and we define Young firms as those with age ≤ 5, to be compared against Medium-Old-aged firms with age ≥ 6 years. The assignment to the different classes is defined according to age and size in the first year of the sample.

We focus on results breaking down by countries and sector of activity. The general finding is that the null of equality of distributions can be rejected only in few particular cases, and generally at low levels of significance. In manufacturing (see Table 2.12), we find that young HG firms outperform young PHG firms in terms

[12]Regression analysis within each class is prevented by the small number of PHG firms falling into each class, especially when breaking down the analysis by countries and sectors.

Table 2.13: Distributional comparison by Age and Size - Services

	Country	#HG	#PHG	ROA	IE/S	LEV	log(TFP)
Young	Pooled	1333	218	-1.971	0.795	0.600	-2.589*
	IT	225	47	-0.263	-0.702	-0.977	1.120
	ES	743	105	-2.049	0.891	2.040	-1.314
	FR	268	43	-0.907	2.368	-1.335	-1.241
	UK	97	23	-0.535	-1.362	-0.308	-0.331
Middle/Old	Pooled	3621	348	2.778*	-1.501	-3.113*	1.065
	IT	778	97	2.121	-1.008	-1.687	1.893
	ES	1708	136	2.132	-2.865*	-2.509	2.224
	FR	832	90	1.262	-0.034	-0.808	2.182
	UK	303	25	-1.455	2.433	-1.511	-3.649**
Micro-Small	Pooled	4210	505	0.882	-0.495	-3.148*	-0.535
	IT	811	122	1.064	-0.537	-2.339	2.157
	ES	2298	235	-0.304	-1.424	-1.438	0.891
	FR	966	126	0.731	1.293	-2.098	0.606
	UK	135	22	0.341	-0.346	-1.918	-1.974
Medium-Large	Pooled	744	61	-0.073	-1.274	-1.049	-0.633
	IT	192	22	1.632	-1.922	-0.333	1.052
	ES	153	6	2.823*	-0.208	-4.805**	0.965
	FR	134	7	-0.750	-0.273	1.285	1.518
	UK	265	26	-2.152	0.159	-0.659	-1.805

Notes: Fligner-Policello (FP) test of stochastic equality. Young firms are ≤ 5 years old in 2004. Micro-Small firms defined as firms with < 50 employees in 2004. HG firms as benchmark within each class: a positive and significant FP statistic means HG firms dominates. Asterisks denote significance levels: * $p < 0.01$, ** $p < 0.001$.

of TFP in Italy, whereas mid-old HG firms are more productive, but less leveraged than mid-old PHG firms in the same country. Higher leverage also characterize mid-old PHG firms in the UK. A similar ranking in productivity in Italy also holds if we look at size, with micro-small HG firms more productive than micro-small PHG firms. And leverage also plays a role in the UK, where we see that medium-large PHG firms are more indebted than HG firms in the same age class.

Concerning services (in Table 2.13), the evidence is of an even weaker statistical difference across PHG and HG firms. Within young firms the null of distributional equality is basically never rejected, whereas within mid-old firms only TFP seems to play some more strongly statistically significant role, but only in the UK, with PHG more productive than HG firms. Disaggregating by size only adds that medium-large PHG firms are significantly more leveraged than medium-large HG firms in Spain.

Once again, we corroborate our main conclusion that the set of economic and financial characteristics does not provide any robust discriminatory power in distinguishing firms experiencing persistent high-growth performance.

2.8 Alternative regression models

Our identification strategy of HG and PHG firms makes standard panel regression models not viable. In principle, with the data at hand, one could implement a panel regression approach with the HG annual growth rate as the dependent variable. This

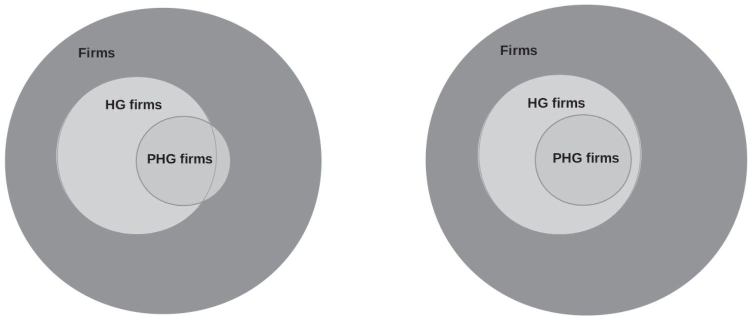

Figure 2.2: Definition of HG, PHG and other firms. Left panel: PHG is not a subset of HG. Right panel: PHG is a subset of HG

would basically correspond to a modification of the usual augmented Gibrat model, limiting the scope of the analysis to a very specific notion of persistence based on the simple autocorrelation structure of the growth rates. We thus decided not to pursue that approach (and a large literature already did, as discussed in Section 2).

Further notice that the two categories of HG and PHG firms, as we define them, are not nested. In fact, a company can be a PHG firm without falling at the same time in the HG group (see left plot in Figure 2.2). This is the case, for instance, of a firm displaying a powerful growth record in four out of the final five years of the sample, but next having such a poor performance in the remaining year that the firm is outside the top decile of the annualized average growth rate distribution. The Multinomial Probit is theoretically superior in this situation. However, since only a quite small number of PHG firms are not HG firms in the data (76 firms in total), we provide a robustness check estimating an alternative econometric specification where we impose that PHG firms are a subset of the HG category (c.f. the right panel in Figure 2.2), as if the decision to be PHG is nested into or dependent from the decision to be HG.

Such structure, implying that a firm can be PHG only conditional upon being HG, naturally leads to a two-step conditional probit. In practice, we modify the definition of growth status by assigning to the "other firms" group all the firms which are PHG, but not HG. Next, we estimate a first probit model for the probability to be selected in the HG set (which now includes the PHG set)

$$P^1 := \text{Prob}\left[y_i \in \{1,2\} \,|\, \mathbf{v}_i\right] = F(\beta^1\, \mathbf{v}_i) \ , \tag{2.5}$$

and then a second-step probit on the probability to be selected in the PHG set, conditional on being in the HG set

$$P^2 := \text{Prob}\left[y_i = 2 \,|\, y_i \in \{1,2\}, \mathbf{v}_i\right] = F(\beta^2\, \mathbf{v}_i) \ . \tag{2.6}$$

This two-step model assumes that the idiosyncratic term in the second conditional regression is independent from the error term in the first step regression. In this sense it represents the restriction of the multinomial probit to a degenerate error variance-covariance matrix.[13]

[13]As in the Multinomial Probit specifications, the regressors include the averages of the firm attributes over computed over 2004-2005, in z-scores, and we report standard errors computed over 100 bootstrap runs.

Table 2.14: Conditional Probit

	Pooled (1)	Italy (2)	Spain (3)	France (4)	UK (5)
First step probit: dep. variable is Prob(HG=1)					
ROA	0.0198**	0.0305*	0.0059	0.0233	0.0376
	(0.0075)	(0.0145)	(0.0115)	(0.0174)	(0.0238)
IE/S	0.0243***	0.0006	0.0363**	0.0378*	-0.0035
	(0.0062)	(0.0142)	(0.0120)	(0.0157)	(0.0256)
LEV	0.0786***	0.1395***	0.0530***	0.0783***	0.0793**
	(0.0074)	(0.0159)	(0.0111)	(0.0161)	(0.0249)
log(TFP)	0.0899***	0.1104***	0.1202***	0.0327	0.0437
	(0.0103)	(0.0207)	(0.0189)	(0.0220)	(0.0245)
AGE	-0.1322***	-0.1678***	-0.1335***	-0.1035***	-0.1371***
	(0.0094)	(0.0167)	(0.0142)	(0.0207)	(0.0299)
log(SIZE)	-0.1160***	-0.2315***	-0.0852***	-0.0639**	-0.0764**
	(0.0111)	(0.0202)	(0.0169)	(0.0213)	(0.0246)
Service dummy	-0.0005	-0.0015	0.0035	-0.0043	-0.0121
	(0.0147)	(0.0211)	(0.0202)	(0.0356)	(0.0496)
Country dummies	yes	-	-	-	-
Observations	55,454	15,712	24,047	11,152	4,543
Log Pseudo-likelihood	-23,731.62	-6,637.67	-10,438.27	-4,718.40	-1,850.92
Chi-2	632.493	445.340	290.036	76.821	51.610
Second step probit: dep. variable is Prob(PHG=1)					
ROA	0.0208	0.0346	0.0149	-0.0028	0.0899
	(0.0173)	(0.0369)	(0.0308)	(0.0462)	(0.0577)
IE/S	0.0346**	0.0161	0.0556**	-0.0929	0.1289
	(0.0126)	(0.0316)	(0.0179)	(0.0624)	(0.0701)
LEV	0.0528**	0.1315**	0.0283	0.0392	0.0615
	(0.0198)	(0.0474)	(0.0266)	(0.0372)	(0.0946)
log(TFP)	-0.0115	-0.0218	-0.0229	-0.0376	0.1725**
	(0.0239)	(0.0406)	(0.0361)	(0.0503)	(0.0643)
AGE	-0.1093***	-0.0785*	-0.1509**	-0.0385	-0.2306
	(0.0286)	(0.0392)	(0.0543)	(0.0528)	(0.1304)
log(SIZE)	-0.1215***	-0.2356***	-0.0486	-0.0819	-0.1099
	(0.0274)	(0.0408)	(0.0370)	(0.0530)	(0.0817)
Service dummy	0.1546***	0.1678*	0.2013**	0.1816	-0.0727
	(0.0392)	(0.0731)	(0.0654)	(0.1071)	(0.1711)
Country dummies	yes	-	-	-	-
Observations	8,776	2,530	3,874	1,711	661
Log Pseudo-likelihood	-2,526.40	-777.98	-983.43	-528.37	-212.60
Chi-2	175.799	92.236	47.739	10.963	21.515

Notes: Explanatory variables in z-scores. Bootstrapped standard errors (100 runs) in parenthesis below the coefficients. Asterisks denote significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

In Table 2.14 we show results for the specifications pooling the data across manufacturing and services. The patterns of statistical significance exactly match with the estimates from the corresponding Multinomial Probit models reported in Table 2.6 above. The point estimates are also quite similar, providing a similar conclusion about the relatively weak power of economic and financial factors in predicting persistence in high-growth performance. We therefore conclude that our findings are robust to the alternative estimation method.[14]

[14] In unreported estimates we have repeated the two-step conditional probit for all the other specifications presented in the previous sections. Results are in accordance with the Multinomial Probit analysis.

2.9 Conclusion

Persistent high-growth performance is a topic of great interest for its potential implications for both academic scholars and policy makers, but we are still missing a deep understanding of this phenomenon. From models of firm-industry dynamics we might expect to find a significant association between efficiency, profitability and financial conditions, on the one hand, and the ability of firms to succeed in achieving high-growth records, but the literature does not provide a theoretical framework explicitly targeting persistent high-growth as an emergent property.

In this essay, exploiting cross-country data on Italian, French, Spanish and UK firms, we have addressed empirically the question whether there is a relationship between that set of key firm characteristics and persistent high-growth. To the best of our knowledge, this is the first study posing this question. Previous studies have indeed so far revealed that outstanding persistent growth performers appear as rare exceptions, more common among small and young firms, but we lack of attempts to investigate the more structural economic and financial determinants of persistent high-growth.

We do find some support that economic and financial characteristics (efficiency and TFP in particular) are associated with high-growth. However, none of the supposedly key drivers of growth systematically stand out as significant predictors of persistently high-growth performance. The result is robust across countries, it does not change across manufacturing and services, and it also holds within groups of firms of different age and size. Moreover, we also find that firm innovativeness (as proxied by intangible assets) is not able to discriminate persistent high-growth from simple high-growth, and that firm size and age do not play a systematic role, although persistently high-growers are younger and smaller in some countries.

Of course, there is a number of other potential factors that may help sustaining high-growth over time and that we have not directly explored in this study. An interesting extension of the analysis would be to include factors of more direct derivation from management research, for which we do not have data, e.g. looking deeper into organizational characteristics, or exploring the role of differences in the underlying firm strategies and managerial or entrepreneurial characteristics. And one cannot rule out, at least in principle, that persistent high-growth primarily occurs at random, guided by "mere luck", so that it would be interesting to test the explanatory power of null models providing random assignment of growth performance.

The research agenda has just begun and many avenues for further research are open. Yet, with all their limitations, our findings represent a challenge for the theory and also raise concerns about the longer run effectiveness of existing policies targeting high-growth companies. The lacking association between efficiency and persistence in high-growth performance, in particular, suggests that supporting high-growth firms could have no impact on the overall competitiveness of sectors and countries.

2.10 Appendix

Table 2.15: Number of firms by country and sector - Manufacturing

NACE	IT	ES	FR	UK
10	724 (688)	927 (906)	415 (399)	140 (77)
11	143 (140)	173 (163)	58 (57)	38 (18)
12	2 (0)	2 (0)	0 (0)	2 (1)
13	507 (485)	310 (306)	68 (63)	28 (15)
14	280 (265)	179 (177)	41 (39)	15 (10)
15	268 (258)	206 (204)	31 (30)	1 (1)
16	176 (169)	386 (385)	186 (182)	23 (15)
17	249 (233)	135 (129)	57 (52)	46 (30)
18	145 (140)	506 (506)	187 (185)	64 (49)
19	38 (35)	7 (6)	5 (5)	8 (7)
20	447 (421)	265 (257)	112 (95)	116 (70)
21	114 (89)	29 (17)	22 (14)	34 (17)
22	553 (532)	350 (344)	196 (183)	70 (43)
23	459 (440)	516 (505)	169 (159)	47 (30)
24	363 (337)	194 (187)	37 (34)	34 (24)
25	1422 (1386)	1511 (1504)	615 (595)	166 (127)
26	279 (261)	92 (84)	111 (98)	88 (64)
27	404 (381)	160 (154)	69 (57)	55 (32)
28	1231 (1178)	442 (436)	202 (191)	139 (93)
29	173 (149)	162 (142)	69 (65)	44 (20)
30	88 (81)	28 (27)	27 (23)	28 (11)
31	310 (306)	425 (423)	80 (79)	31 (19)
32	197 (193)	169 (167)	94 (92)	184 (136)
33	115 (111)	363 (363)	290 (284)	56 (41)
Total	8687 (8278)	7537 (7392)	3141 (2981)	1457 (950)

Note: Number of firms with less than 250 employees in parenthesis.

Table 2.16: Number of firms by country and sector - Service

NACE	IT	ES	FR	UK
45	773 (770)	1596 (1592)	1115 (1110)	337 (234)
46	2949 (2887)	5092 (5048)	2122 (2074)	555 (429)
47	782 (721)	3627 (3604)	1753 (1732)	202 (100)
49	320 (293)	992 (978)	466 (448)	147 (75)
50	22 (21)	32 (32)	6 (6)	15 (8)
51	11 (10)	5 (2)	1 (1)	24 (12)
52	292 (265)	252 (247)	94 (81)	74 (42)
53	4 (3)	22 (22)	3 (3)	5 (2)
55	162 (156)	443 (436)	312 (311)	112 (80)
56	105 (92)	1171 (1162)	456 (447)	73 (29)
58	84 (75)	137 (130)	83 (75)	61 (28)
59	16 (15)	43 (43)	31 (30)	21 (14)
60	22 (22)	29 (27)	6 (4)	7 (3)
61	18 (17)	68 (61)	16 (15)	42 (26)
62	184 (172)	237 (230)	119 (103)	135 (111)
63	72 (68)	15 (15)	20 (18)	15 (13)
64	41 (26)	33 (12)	71 (42)	157 (101)
66	17 (15)	40 (39)	8 (6)	29 (25)
68	160 (148)	218 (217)	75 (75)	61 (44)
69	70 (66)	298 (294)	57 (57)	11 (9)
70	155 (127)	125 (114)	89 (39)	282 (106)
71	99 (91)	271 (262)	150 (139)	46 (28)
72	23 (22)	20 (18)	16 (14)	15 (8)
73	85 (83)	202 (202)	68 (66)	39 (31)
74	51 (50)	188 (187)	34 (34)	44 (33)
75	0 (0)	29 (29)	1 (1)	1 (1)
77	43 (41)	174 (171)	82 (77)	81 (60)
78	10 (8)	8 (6)	7 (4)	55 (35)
79	64 (62)	117 (112)	10 (10)	32 (22)
80	37 (31)	49 (45)	15 (14)	10 (5)
81	82 (58)	234 (215)	204 (191)	26 (9)
82	91 (86)	86 (82)	78 (75)	199 (128)
90	15 (13)	40 (40)	24 (24)	9 (5)
91	6 (2)	6 (6)	11 (11)	1 (1)
92	6 (5)	87 (84)	39 (38)	10 (2)
93	74 (73)	176 (175)	52 (51)	40 (30)
94	0 (0)	6 (6)	0 (0)	8 (7)
95	28 (27)	103 (103)	35 (34)	3 (1)
96	52 (48)	239 (236)	282 (281)	102 (68)
Total	7025 (6669)	16510 (16284)	8011 (7741)	3086 (1965)

Note: Number of firms with less than 250 employees in parenthesis.

Innovation strategies and firm growth

3.1 Introduction

The relationship between innovation and firm performance has for long interested economists. Obviously the general intuition is that innovation is among the key determinants of comparative advantages of firms over competitors, thus contributing to the ability of firms to grow and gain market shares. Against this simplistic prediction, however, play both the ample degrees of complexity, uncertainty and idiosyncrasy that are well known to characterize the innovation process. Innovation is indeed the search for, and the discovery, development, improvement, adoption and commercialization of, new processes, new products and new organizational structures and procedures. It involves indeed uncertainty, risk taking, probing and re-probing, experimenting and testing. Thus the process of innovation itself, and its ensuing effects on various aspects of firm performance, can be extremely heterogeneous and difficult to predict (Dosi, 1988).

Within the vast literature, this essay contributes to the studies that seek to identify the links between innovation and firm growth, focusing in particular on the linkages between innovative activities and success on the market in terms of sales growth. In spite of the increasing availability of firm level data over the last 10-15 years, especially following the attempt undertaken by the EU to provide regular surveys of innovation across members states (the CIS-*Community Innovation Survey* exercise), this literature is still underdeveloped under several respects, in turn motivating the contributions that we want to pursue in this study.

First, our major contribution is to provide a broad picture of the relationship between growth and innovation, by looking at a wide set of innovation indicators that capture different sources, modes and output of the innovative efforts undertaken by firms. Extant empirical studies on growth and innovation mostly focus on traditional proxies such as R&D and patents. On the contrary, exploiting a rich dataset on Spanish firms, we look at different measures of innovative input (distinguishing between internal vs. external R&D, investment in innovative machinery and equipment, purchase of licenses or know-how from other firms), at different output of innovation (process vs. product innovation), also distinguishing different types of product innovation (new-to-the-firm or new-to-the-market). In this respect our paper is closely related to the recent work by Hölzl (2009) focusing on high-growth firms. The cross-sectional nature of that study, however, represents a limitation we want to improve upon.

Indeed, our second contribution stems from the possibility to work with a panel of firms observed over several years. A common limitation to studies exploiting

CIS-like data is that such surveys are run in waves every 3-4 years, often on rotating samples of firms. Thus, previous studies can typically exploit a single cross section, or they can follow just a few firms overt time, in turn failing to control for unobserved heterogeneity. This point is not merely a technical econometric drawback, given the inherently idiosyncratic nature of the process and outcomes of innovation. The dataset of Spanish firms available to us is a CIS-type dataset in terms of the rich and detailed information about innovative activity, but it is longitudinal in nature, since a consistent data collection methodology ensures to have information on the same set of firms over time.

Third, and relatedly, we also contribute to the recent literature (Coad and Rao, 2008; Falk, 2012; Segarra and Teruel, 2014) that adopts quantile regressions to show that while innovation can have mixed or nil effect on the average growth rate in a cross section of firms, innovation is instead more beneficial for fast, or high-growing, firms. Besides suffering from the above-mentioned limitation of focusing only on patents or R&D, these studies apply basic quantile regression techniques. Exploiting the longitudinal dimension of our data, we can instead apply up-to-date quantile regression techniques designed to account for firm fixed effects. To the best of our knowledge, this is the first attempt in this direction within the growth-innovation literature.

Finally, we provide an empirical assessment of the complementarities existing between the different innovation activities in fostering sales growth. Recent studies exploit the notion of modularity of the innovation function to investigate the complementarity of innovation inputs or knowledge sources in successful generation of innovation outputs. We apply the same conceptual and methodological apparatus to ask whether different combinations of basic innovation activities (R&D, process and product innovation, embodied and disembodied technical change) help improving growth performance, above and beyond the contribution of each single activity alone.

Our results point to a good deal of heterogeneity in the way different innovation activities contribute to expanding sales. Indeed, among the innovation indicators we account for, internal R&D turns out as the main driver of sales growth, on average. Other innovation activities, with exception of acquisition of disembodied knowledge and process innovation, have a positive association with growth only in the top quartile of the growth distribution, that is for high-growth performance. We also document a complementarity effect between internal R&D and product innovation, and between product and process innovation. This evidence emphasizes the complexity underlying the growth-innovation relationship and provides a potential explanation for the inconclusive results of previous studies which adopted a more standard unidimensional approach.

3.2 Background framework

In this Section we discuss some of the most relevant contributions in the field, in turn motivating the gaps in the literature that we tackle in the present paper. We devote more attention to studies investigating sales growth, which are more directly related to our analysis.[1]

Whilst theoretical models acknowledge the importance of innovation as a major driver of firm growth and success on the market (see Aghion and Howitt 1992; Aghion et al. 2005, among others), the empirical literature does not fully support the theoretical expectations. A long tradition of studies do find some positive effect of innovative activity, especially of R&D, on growth. The early papers documenting this fact go back to classical studies of the 1960s. Mansfield (1962) carries out a detailed assessment of the steel and petroleum sectors by using a long time series and finds that successful innovators grow faster. Similar results are also found in Scherer (1965), analyzing the patenting activity of the 365 largest US companies, and in Mowery (1983), looking at the effect of R&D employment on the growth of US manufacturing industries over a 25-years period. In their influential paper, Geroski and Machin (1992) concentrate on 539 quoted UK firms that introduced at least a major innovation, observed over more than ten years. They find that innovating firms are more profitable and grow faster, but the increase in sales is transitory, lasting only until the firm loses proprietary control over the new knowledge employed. Storey (1994) corroborates this finding and underlines the important magnifying role played by the initial size, with smaller firms achieving a more rapid growth after having been successful in innovating. Stam and Wennberg (2009) explicitly target new start-ups and show that the effects of R&D on new products development and hence on growth is present only in high-tech sectors.

By contrast, however, there is also a considerable stream of research that does not find any significant effect of innovation on sales growth, like in Geroski and Mazzucato (2002). The contribution in Bottazzi et al. (2001) is particularly relevant for the unusually detailed level of analysis (product-level). They target the top-150 world pharmaceutical firms, and conclude that the innovative position of a firm (measured either by the discovery of new chemical entities or by the share of patented products) is not associated with growth of sales. More recently, Demirel and Mazzucato (2012), also targeting a sample of US quoted companies in pharmaceutical, conclude that R&D investments affect average sales growth of large companies, whereas for smaller units such an effect is significant only if there is a certain degree of persistence in the innovation activities undertaken by the firm.

The mixed empirical support for the existence of a strong link between innovation and sales growth might be related to the extreme complexity of the firms' innovative process. In turn, a robust stylised fact emerging from industrial economics is that the firm growth rates distribution is characterised by wide heterogeneity and a tent shape (see Stanley et al., 1996; Bottazzi and Secchi, 2006), whatever the level of sectoral aggregation considered (Dosi, 2007). In this respect, due to its inherent nature, the process leading from innovative input to innovative output may show different effects according to the different positioning of a firm in the growth rates distribution, whereas more traditional regression studies are only informative about the "average firm". Freel (2000), analyzing a sample of 228 small UK manufacturing disaggregated by differential innovativeness, shows that, although in-

[1] There also exists a huge literature on the effects of innovation on growth of employment, where the main focus is on the labour-saving vs. labour augmenting role of innovation, and topics related to skill-bias technical change (see Vivarelli, 2014, for an exhaustive survey on the topic). We do not discuss this literature here, as we are more interested in a measure of growth capturing success on the market.

novation does not necessarily determine firm growth, it may be relevant in boosting high-growth. More recent literature provides similar conclusions, applying quantile regression techniques to disentangle the effect of innovation proxies along the spectrum of the distribution of growth rates. The approach is also popular within studies looking at those companies labeled as high-growth firms or 'gazelles'. For instance, Coad and Rao (2008) work with firms active in four US sectors with fast changing technologies, and find that innovation, measured in terms of R&D and patents, has an asymmetric impact over the sales growth distribution, with high-growth firms deriving the greatest benefits from their innovative efforts. Hölzl (2009) analyzes CIS-III data for 16 countries, and shows that R&D is much more important for high-growth SMEs in countries that are closer to the technological frontier, arguing that such firms derive much of their drive from the exploitation of comparative advantages. Falk (2012), exploiting a sample of Austrian firms active both in manufacturing and service, confirms the same asymmetric effect of R&D expenditure on growth. He claims, in addition, that the return of R&D investments seems to decrease over time, reaching the minimum values in the period just before the financial crisis. Colombelli et al. (2013) compare the effect of product and process innovation on sales growth, along three waves of the French CIS. They show that innovative companies, regardless the type of innovation, have stronger tendency to expand their sales and that the marginal effect of innovation is heightened for firms located in the upper quantiles of the growth rates distribution.

While the application of quantile regressions has allowed to, at least partially, reconcile the evidence with the theoretical expectation of a strong influence of innovation on firm growth, the literature still suffers from several limitations. In particular, studies on the subject tend to focus on traditional proxies of innovative activity such as R&D and patents. As recently emphasized in the extensive literature review in Audretsch et al. (2014), however, the great variety of innovation strategies undertaken by firms calls for a multidimensional approach to assess the actual contribution of innovation on corporate growth.

In this essay, we respond to that call by enlarging the picture on the role of different innovation activities or strategies in shaping firm growth. The set of innovation indicators that we use are intended to capture different aspects of the innovative process. They cover the usual dichotomy between innovative inputs vs. innovative outputs, but they also allow to investigate the role of internal vs. external sourcing of knowledge. The existing literature does not provide conclusive evidence on their effect on sales growth.

First, concerning the output side of innovation, many studies highlight the merits of innovation surveys in providing direct proxies for product and process innovations (see Griffith et al. 2006; Parisi et al. 2006; Hall et al. 2008, 2009), beyond traditional focus on patents as the only measured outcome of the innovation process. Only few works however consider the relationship between sales growth and proxies of innovative output alternative to patents. On the one hand, there is practically no evidence about the direct impact of process innovation on sales growth, as indeed most studies focus on the relationship between new processes and productivity (see Griffith et al. 2006; Hall et al. 2009). A notable exception is in Goedhuys and Veugelers (2012), where it is shown that process innovation has no effect on sales growth, for a sample of Brazilian manufacturing firms. The suggested interpretation is that of a mediating role of productivity, such that process innovation has direct effect on cost efficient production, while it may show its beneficial effects on sales in later stages, after an initial period of process restructuring. On the other hand, concerning product innovation, theory would predict a positive link between the introduction of new products and sales growth, as indeed efforts directed to creation and commercialization of new products represent the primer strategy for expansion and growth (Hay and Kamshad, 1994). But the evidence is mixed. Cucculelli and Ermini (2012) find that the mere introduction of products (dummy for product innovation) does not affect sales growth if one does not control for unobserved product characteristics, which in the study is proxied through the tenure from last product introduction. Other empirical investigations confirm that the product characteristics matter beyond the simple introduction of new products, by looking at the two measures of product innovation that we also use, that is distinguishing between products-new-to-the-firm vs. products-new-to-the market. Hölzl (2009) shows that the share in total sales due to products new-to-the-market is of great importance for high-growth firms, in particular for those located in countries closer to the international technological frontier. This evidence lends support to the intuition that products-new-to-the-market, capturing more original and complex innovation, are those that really matter for competing and gaining market shares. Conversely, however, Corsino and Gabriele (2011) find that sales growth is positively affected also by more incremental product innovations introduced in the recent past, although these are often considered as related to less valuable innovation or imitative efforts.

Second, moving to the evidence about the relationship between sales growth and innovation inputs, also in this case we observe a sort of resilience to abandon traditional measures such as expenses in in-house formal R&D. An exception is in the recent Segarra and Teruel (2014), making a distinction between R&D carried out internally and purchases of outsourced R&D. The results show that while internal R&D has a significant positive impact in the upper quantiles of the sales growth distribution, that is for high-growth, external R&D appears to be important only up to the median.

However, and third, we lack further attempts to exploit the distinction between internal and external sources of knowledge and innovation. Innovation surveys provide information not only about activities like outsourced R&D, but also about acquisition of innovative technology, both embodied (investment in new machinery and equipment) and disembodied (acquisition of patents, know-how, licenses, etc.), but their respective impact on sales growth has not been explored yet. This is quite unfortunate, given the central interest devoted to these activities in innovation studies. Theoretically, indeed, the acquisition of new knowledge or new techniques from outside the boundaries of the firm has uncertain effects. On the one hand, external sourcing can help improving the knowledge base and, thus, the overall innovative capabilities of firms. But, on the other hand, exploitation of external sourcing is also subject to constraints due to absorptive capacities, to the complexity of and coordination within the user-producer interactions, and to the differential ability to adapt the outsourced innovative inputs to the specific characteristics, competences and needs of each firm. Empirical studies tend to support a positive impact of external sources on the outputs of innovation. Santamaria et al. (2009) show that non-R&D activities are crucial for both product and process innovation, while Pellegrino et al. (2012) and Conte and Vivarelli (2014) provide evidence that embodied technical change fosters the share of sales due to new or significantly improved products, especially in low-tech industries and across small and young firms. However, we lack systematic evidence about the effect of external sourcing on sales growth.

All in all, from reading the literature, and to our knowledge at least, the already mentioned Goedhuys and Veugelers (2012) represents the only attempt to reconstruct the relationship between sales growth and at least a subset of the various innovation strategies that firms have at their disposal. The paper exploits the logic of a standard augmented CDM model (Crepon et al., 1998) to recursively assess the relevance of internal vs. external R&D for product and process innovation, and then to estimate the ensuing impact of successful new processes or products on stimulating sales growth.

We provide a different contribution. We look at a broader set of innovative indicators, encompassing internal vs. external R&D, process innovation, different types of product innovation, and embodied vs. disembodied technological acquisition, and explore both their direct effects and the complementarities among them in fostering sales growth.

3.3 Data

3.3.1 Data and sample

We exploit a firm-level dataset drawn from the Spanish Technological Innovation Panel (henceforth PITEC), jointly developed by the Spanish National Statistic Institute (INE), the Spanish Foundation for Science and Technology (FECYT), and the Foundation for Technical Innovation (COTEC). The data are collected following the Oslo Manual guidelines (OECD, 1997) and, as such, they can be considered a Community Innovation Survey (CIS)-type dataset. Thus, PITEC includes a rich set of variables that measure firms' engagement in innovation activity, economic and non-economic measures of the effects of innovation, self-reported evaluations of factors hampering or fostering innovation, participation in cooperative innovation activities, access to public funding, use of patents and other means of appropriability, and some complementary innovation activities such as organizational innovation and marketing. The main limitation, common to other CIS-type surveys, lies in the small set of variables about more structural and industrial characteristics of firms, which essentially cover only annual turnover and total employment, industry affiliation, founding year, export status, industrial group, and few others.

The key feature that distinguishes PITEC from the majority of European CIS-type datasets is its longitudinal nature. Indeed, since 2003 systematic data collection ensures a consistent representativeness of the population of Spanish manufacturing and service firms over time, allowing to follow the same firms over a considerable number of years. This allows to control for unobserved factors that could have an impact on the relationship between innovation variables and patterns of sales growth. Another advantage of the data is that there is no need to care about the sample-selection issues commonly arising in CIS-type surveys.

We select our working sample from an initial dataset of 100,016 firm-year observations over the period 2004-2011. We focus on manufacturing firms, and we look at "organic growth", hence discarding all firms involved in M&A transactions. The resulting sample is an unbalanced panel of 26,386 firm-year observations for which the variables used in our empirical exercise are non-missing. Table 1 shows that the

Table 3.1: Composition of the panel

Time obs	# firms	%	%Cum	Obs
3	140	2.76	2.76	140
4	230	4.54	7.31	460
5	250	4.94	12.24	750
6	328	6.48	18.72	1,312
7	972	19.19	37.91	4,860
8	3,144	62.09	100	18,864
Total	5,064	100		26,386

Note: Time obs. indicate the minimum number of years over which firms are observed: T=3 refers to firms that are observed for at least three periods: T=4 corresponds to firms that are observed for at least four periods, and so on.

large majority of firms (62.09 %) is observed over the entire sample period, whereas another 19.19% persists in the data for 7 years, and only a negligible percentage (7,31%) for less than 5 years.

3.3.2 Main variables

Our dependent variable is firm growth measured in terms of sales. This is defined as the log-difference:

$$G_{it} = s_{it} - s_{i,t-1} \quad , \tag{3.1}$$

where

$$s_{it} = log(S_{it}) - \frac{1}{N}\sum_i log(S_{it}) \quad , \tag{3.2}$$

and S_{it} is sales (annual turnover) of firm i in year t, and the sum is computed over the N firms populating the same (2-digit) sector. In this way size and, thus, the growth rates are normalized by their annual sectoral average. The normalization implicitly removes common trends, such as inflation and business cycles effects in sectoral demand.

In our attempt to provide a multidimensional view about innovation activity of firms, we employ the following innovation indicators, available for each firm in each year:

1. *Internal R&D* (intensity): Intramural R&D expenditures, normalized by total turnover.
2. *External R&D* (intensity): Extramural R&D expenditures, normalized by total turnover.
3. *Prod New-to-the-firm*: Share of firm's total sales due to sale of new or significantly improved products, which were new only for the firm.
4. *Prod New-to-the-market*: Share in firm's total sales due to sales of new or significantly improved products, which were new to both the firm and the market.

Table 3.2: Innovation variables - Descriptives

	Mean	Std.Dev.	Median	Min	Max
Internal R&D	0.031	0.161	0.004	0	7.986
External R&D	0.006	0.055	0	0	3.353
Prod. New-to-firm	0.248	0.352	0.056	0	1
Prod. New-to-MKT	0.099	0.225	0	0	1
Proc. Innov	0.633	0.482	1	0	1
Emb.Tech.Change	0.006	0.047	0	0	3.441
Disemb.Tech.Change	0.000	0.005	0	0	0.555

Notes: Table reports basic descriptive statistics on the different innovation variables. Figures computed pooling over the working sample - 26,386 observations.

5. *Process Innov*: Binary indicator equal to 1 if the firm introduces new or significantly improved processes.

6. *Embodied technological change* (intensity): Investment in innovative machinery and equipment, normalized by total turnover.

7. *Disembodied technological change* (intensity): Acquisition of external knowledge (patents, know-how, and other types of knowledge from other enterprises or organizations), normalized by total turnover.

The definitions of these proxies from PITEC are equivalent to their counterpart in innovation surveys from other countries. The interpretation is in most cases well accepted. R&D indicators just measure expenditures in different R&D activities, and we also follow the usual approach to take the ratio to total turnover instead of absolute figures. Concerning product innovation, the introduction of products perceived as new-to-the-market connects with the ability to perform “more important” innovation, resulting in more valuable products, while products new-to-the-firm are usually considered as a proxy of more “incremental” and less valuable innovation. The dummy for process innovation has the standard interpretation as capturing reorganization of production or implementation of new processes, and we also follow the common practice to interpret acquisition of new machineries and of external knowledge as proxies for, respectively, acquisition of embodied and disembodied technical change.

In Table 3.2 we report descriptive statistics for the innovation indicators. Notice, first, that all the indicators display highly skewed distributions, suggesting considerable heterogeneity in the innovative behaviour. Second, firms in our sample appear more prone to undertake internal generation of knowledge rather than searching for external sources. Indeed, on average, intramural formalized R&D amounts to 3.1% of annual sales, while we observe an average 0.6% share in sales for both extramural R&D and for acquisition of innovative machineries and equipment, and such share is close to zero in the case of acquisition of disembodied knowledge. Further, from the indicators of innovative output, we see that a relatively large fraction of the sample performs process innovation (approximately 63% of the observations). On

Table 3.3: Descriptive statistics for sales growth by innovation status

		Mean	Median	Min	Max	#Obs
Internal R&D	NO	-0.040	-0.016	-4.813	3.853	11,225
	YES	0.009	0.006	-3.821	4.674	15,161
External R&D	NO	-0.025	-0.008	-4.813	3.853	18,999
	YES	0.022	0.012	-3.821	4.674	7,387
Prod.New-to-firm	NO	-0.021	-0.007	-4.813	4.674	17,200
	YES	0.005	0.006	-3.603	3.57	9,186
Prod.New-to-MKT	NO	-0.027	-0.011	-4.813	4.674	10,237
	YES	-0.002	0.002	-3.958	3.57	16,149
Proc. Innov.	NO	-0.032	-0.016	-4.813	4.674	10,290
	YES	0.001	0.006	-3.958	3.57	16,096
Embod.Tech.Change	NO	-0.018	-0.006	-4.813	4.674	21,780
	YES	0.018	0.011	-2.839	3.253	4,606
Dis.Tech.Change	NO	-0.013	-0.003	-4.813	4.674	25,826
	YES	0.016	0.001	-2.759	2.615	560

Notes: descriptive statistics of G_t by “Innovators” vs. “Non-innovators” defined as firms that do (YES) or do not (NO) engage in innovation, according to the different innovation variables. Figures computed pooling over the working sample - 26,386 observations.

the other hand, concerning product innovation, the share in total sales due to products new-to-the-market is on average smaller than the share of sales from products new-to-the-firms (9.9% vs. 24.8%). This hints that “truly” innovative products are more difficult to achieve and more rare than incremental innovation, and thus may contribute less to sales.

3.4 Growth and innovation: descriptive evidence

As a first assessment of the relationship between sales growth and innovation, we compare the growth rates across “innovators” and “non-innovators”, that is splitting the sample between firms that do or do not undertake each specific innovative activity. Of course, non-innovators according to one variable may still be innovative firms, in the sense that they may be engaged in other types of innovative activity. This issue becomes relevant when we build a set of “innovation strategies” (see Section 3.6).

Table 3.3 shows basic descriptives of sales growth across the different subgroups. We see that “Innovators” tend to display larger mean and median growth rates than “non-innovators”, regardless the innovation variable. The median, in particular, is positive for “innovators” and negative for “non-innovators” for all the proxies.

We next look at the unconditional distribution of sales growth rates, again across “innovators” and “non-innovators”. Kernel densities (on log-scale) are reported in Figure 3.1. The estimates reveal differences between the two groups, with "non-innovators" generally more concentrated in the left part of the support. These asymmetries in the left tail are particularly pronounced for the two R&D indicators. The differences in the right tails are less clear-cut, with the two distributions substantially overlapping, irrespective of the innovation variable considered. This implies that "non-innovators" are nevertheless able to enjoy extreme positive growth events. The visual inspection is confirmed by a Fligner and Policello (1981) test of distributional equality (henceforth FP), allowing to assess which of the two distributions stochastically dominates the other along each innovation variable considered. The null hypothesis of stochastic equality is always rejected (except for technological acquisition) and the positive FP statistics imply that "innovators" present a larger probability to experience superior growth performance than "non-innovators".

Overall, the observed distributional asymmetries suggest that the larger average growth observed within innovators can be due to innovators being more able to avoid below-average growth rates, rather than to stably reach a positive and high-growth performance. All these findings, however, just provide an unconditional picture.

3.5 Growth and innovation: main results

In this Section we present our main analyses. The empirical strategy is to separately investigate the relationship between sales growth and each innovation activity, conditional on a set of controls. We first look at the effect of innovation variables on average growth, through standard panel techniques, and then exploit fixed-effects quantile regressions to estimate asymmetries in the innovation-growth relationship across growing and shrinking firms.

The baseline empirical model is a panel regression equation

$$G_{i,t} = \alpha \; INNOV_{i,t-1} + \boldsymbol{\beta} \times \mathbf{Z_{i,t-1}} + u_i + \epsilon_{i,t} \quad , \tag{3.3}$$

where $INNOV$ stands alternatively for one of the different innovation variables, $\mathbf{Z}$ is a set of firm-level control variables, u_i is a firm fixed-effect, and $\epsilon_{i,t}$ a standard error term.

Both $INNOV$ and the controls enter with a 1-year lag, at least partially controlling for potential simultaneity.[2] The set of controls includes the lagged dependent variable (G_{t-1}), a proxy for size in terms of number of employees (in log, $\ln Empl$), firm age computed by year of foundation (in log, $\ln Age$) and three dummy variables, respectively taking value 1 if firm i is exporting ($Export$), or receiving public financial support to innovation ($PubFund$), or belonging to an industrial group ($Group$) in year $t-1$, and zero otherwise.[3] Table 3.4 reports the corresponding descriptive statistics. All the specifications also include a full set of industry (2-digit) and year dummies.

[2]Since one might argue that it takes time for innovation to be "translated" into sales growth, we also checked models including a full lag structure for the innovation variables. The baseline model with 1-year lag distance between $INNOV$ and growth was chosen through sequential rejection of the statistical significance of more distant lags.

[3]The $PubFund$ dummy records any kind of public financial support for innovation activities from Spanish local or government authorities and from the EU bodies, including tax credits or deductions, grants, subsidized loans, and loan guarantees. It excludes research and other innovation activities entirely conducted for the public sector under a specific contract.

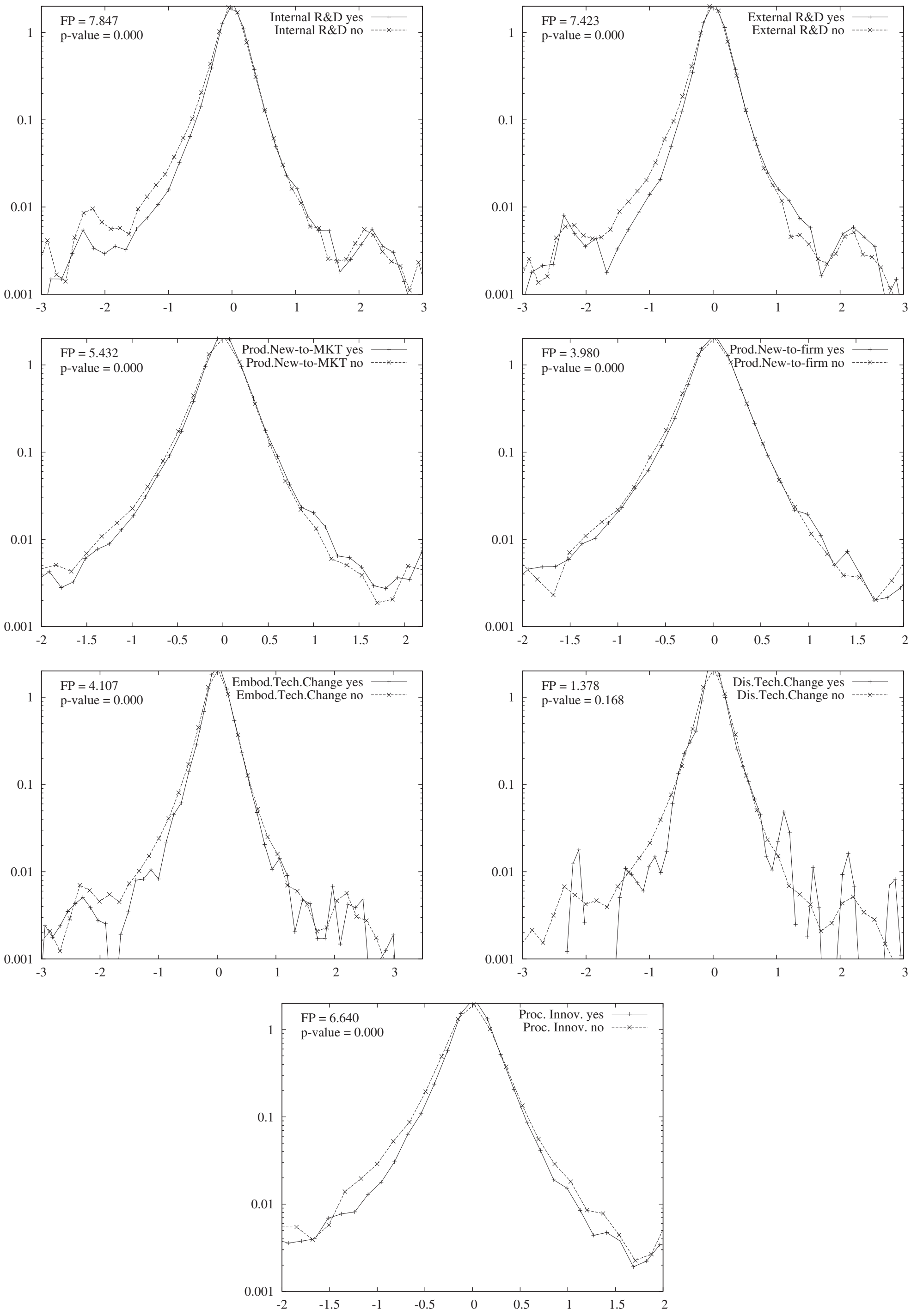

Figure 3.1: Kernel estimates (Epanenchnikov kernel) of sales growth rates densities for "innovators" vs. "non-innovators", defined as firms that do (YES) or do not (NO) engage in each innovation activity. Innovation proxies are Internal or External R&D (first row), Products new-to-the-firm or new-to-the-market (second row), Embodied vs. Disembodied technical change (third row), and Process Innovation (bottom row). Figures also report a Fligner and Policello (1981) test of stochastic dominance: a positive and significant FP statistic indicates that innovators dominate non-innovators along the innovation proxy considered. Results obtained pooling over the working sample - 26,386 observations.

Table 3.4: Descriptive statistics for the control variables

	Mean	Std.Dev.	Median	Min	Max
G_{t-1}	0.026	0.376	0.027	-4.813	4.739
$\ln Empl_{t-1}$	4.088	1.309	3.932	0	9.234
$\ln Age_{t-1}$	3.223	0.598	3.258	0	5.088
$Export_{t-1}$	0.796	0.403	1	0	1
$PubFund_{t-1}$	0.354	0.478	0	0	1
$Group_{t-1}$	0.378	0.485	0	0	1

Notes: Figures computed pooling over the working sample - 26,386 observations.

The coefficient of primer interest is of course α, capturing the effect of each specific innovation activity on sales growth. Inclusion of firm fixed-effects implies that the main parameter is identified through within-firm changes of the $INNOV$ proxies over time. This helps mitigating standard omitted variable bias, which in our case can provide a relatively severe source of incorrect estimation, due to the limited number of firm-level controls available in PITEC (as common also to other innovation surveys). In particular, we do not have data to compute a reliable measure of productivity, which is theoretically a crucial determinant of both growth and innovation, especially for its mediating role between input and output of innovation suggested by innovation studies. Firm fixed-effects absorb at least the time-invariant component of efficiency, while the time varying component remains unobserved and thus it is possibly interacting with other controls like age, size and export status. A similar reasoning applies for other unmeasured factors jointly influencing growth and innovation, such as financial constraints, managerial and organizational characteristics, or input quality. Such an endogeneity issue is controlled for in standard panel-GMM estimators. Conversely, quantile regression approaches jointly controlling for fixed effects and endogenous covariates are still under development.[4]

3.5.1 Panel estimates

We start presenting standard panel analysis of Equation (3.3). As a reference, we first show the results obtained with the Fixed Effects-Within (FE) estimator, although this might be severely biased due to endogeneity and the presence of the lagged dependent variable. Secondly, we apply the GMM-DIFF estimator (Arellano and Bond, 1991), that mitigates endogeneity via exploiting lags of the regressors as instruments after differencing the estimation equation.[5] The instruments included in the GMM procedure vary depending on the estimated equation. We always use $\ln Age$, $Group$ and year dummies as exogenous variables, while different lags of

[4] Harding and Lamarche (2009) and Harding and Lamarche (2014) are, to our knowledge, the only works tackling both issues. However there are difficulties in implementing the methods since one does not have an equivalent to panel-GMM allowing for internal instruments.

[5] We prefer this estimator over the alternative GMM-SYS estimator (Blundell and Bond, 1998) since firm growth is known to display weak persistence over time, and thus time-differences of growth are poor instruments for growth levels. This is also confirmed by our results below.

Table 3.5: Panel estimates - R&D intensity

	Innovation Proxy			
	Internal R&D		External R&D	
	FE	GMM	FE	GMM
Dep.Var. is G_t	(1)	(2)	(3)	(4)
INNOV_{t-1}	0.2156***	0.3837***	0.4912*	0.5049
	(0.078)	(0.063)	(0.289)	(0.511)
G_{t-1}	-0.3087***	-0.2225	-0.3122***	-0.0534
	(0.013)	(0.178)	(0.012)	(0.155)
$\ln Empl_{t-1}$	-0.1605***	-0.1752	-0.1615***	-0.2010
	(0.022)	(0.229)	(0.022)	(0.199)
$\ln Age_t$	-0.1718***	-0.1888**	-0.1952***	-0.1933**
	(0.053)	(0.038)	(0.055)	(0.097)
$Export_{t-1}$	0.0037	-0.0801**	0.0034	-0.0779**
	(0.015)	(0.038)	(0.015)	(0.038)
$PubFund_{t-1}$	0.0014	-0.0076	0.0032	-0.0085
	(0.007)	(0.021)	(0.007)	(0.021)
$Group_{t-1}$	-0.0205	-0.0229	-0.0201	-0.0285
	(0.020)	(0.030)	(0.020)	(0.034)
Obs	26,386	21,291	26,386	21,291
AR(1)		0.016		0.001
AR(2)		0.600		0.518
Sargan		0.118		0.371
Hansen		0.333		0.370

Notes: Fixed Effects-Within (FE) and GMM-DIFF estimates of Equation (3.3). Regressions include a full set of year and sector dummies. Robust standard errors in parenthesis, clustered at firm-level: ***, ** and * indicate significance at 1%, 5% and 10% level, respectively. We also report p-values of Arellano-Bond test for first and second order serial correlation, AR(1) and AR(2), together with p-values of usual Sargan and Hansen tests for overidentifying restrictions.

G, $INNOV$, $\ln Empl$, $Export$ and $PubFund$ are included, based on the standard Arellano-Bond tests for serial correlation and on Sargan/Hansen tests for overidentifying restrictions. We mainly comment on the GMM results, since these are in principle more reliable.

In Table 3.5 we explore the estimates of the models including the two measures of R&D intensity as innovation proxy. The FE results reveal a positive and strongly significant relationship between sales growth and internal R&D intensity, whereas a barely significant (10% level only) association is detected with extra-mural R&D activity. The GMM estimates corroborate the results, and external R&D in this case looses statistical significance. The point estimates across the two estimation methods differ in magnitude, but cannot be considered as statistically different within 1-standard error confidence band. These findings confirm the central role of R&D as a driver of corporate growth and success on the market. At the same time, however, they suggest that it is internally developed research that pays off. Outsourced R&D does not support sales growth, possibly due to problems related to absorptive capacity or to coordination failures with the external provider of R&D services. Another explanation can be that firms tend to outsource more marginal R&D projects, i.e. those more loosely linked to the core activity or less relevant to product development, and thus less likely to impact on sales and market shares.[6]

Concerning the control variables, the estimated coefficients display robust patterns, irrespective of the innovation proxy considered. We comment on GMM results which tackles the good deal of endogeneity potentially affecting the analysis. First, we do not find any significant autocorrelation of sales growth over time. This is in line with the the vast literature on size-growth relationships and Gibrat's law, where attempts to quantify growth rates autocorrelation provided quite mixed results, supporting the notion that growth follows a quite erratic and difficult to predict pattern. Second, and confirming one of the implications of Gibrat's Law, the coefficient on lagged size (in terms of employment) is not statistically different from zero. Third, age is always negatively correlated with firm growth, at strong significance level, confirming the intuition that younger firms are typically growing more rapidly than more mature firms. Fourth, export status has a negative and significant coefficient. This may be unexpected, since the literature on micro-empirics of exports suggest that exporters typically reach superior performance than non-exporters. Recall however that here the coefficient captures the effect of over time, within-firm changes of export status, so that the result says that becoming exporters is associated to a reduction in sales growth. Finally, we observe a common pattern for the dummy variables identifying public support to innovation and group membership: both do not exert any statistically significant relationship with sales growth.

Next, in Table 3.6 we present the estimates obtained with the indicators of product innovation, looking at shares of sales of products new-to-the firm and of products new-to-the-market. Both variables turn out as not significant. The result is striking at first, since one expects the simple selling of new products should spur growth. But, what we measure here is whether the effect of an increase in the share of sales due to new products "translates" into an increase of overall sales. The result may suggest that this share is overall small and that only few new products have a deep impact on sales, so that in the end the contribution of product innovation vanishes, on average.

The results on the control variables (once again focusing on GMM estimates) are generally in agreement with the patterns emerged above in the models including internal and external R&D. The main difference is that in the specification with products new-to-the-firm, we find a negative autocorrelation of sales growth, and a negative effect of lagged size on subsequent sales growth. Both regressors loose their statistical significance in the model with products new-to-the-market. For all the other controls, point estimates and patterns of significance are similar across the two specifications . In line with the models including R&D variables, we confirm a negative and significant effect of age and export status, while the dummy variables indicating public support and group membership are confirmed to lack any statistically significant relationship with sales growth.

Table 3.7 presents the estimates concerning the other innovation proxies. In columns 1-2 we exploit the binary indicator for process innovation. Both FE and

[6]Our data, unfortunately, do not allow us to test such conjectures.

Table 3.6: Panel estimates - Product Innovation

	Innovation Proxy			
	Prod.New-to-firm		Prod.New-to-MKT	
	FE	GMM	FE	GMM
Dep.Var. is G_t	(1)	(2)	(3)	(4)
INNOV_{t-1}	-0.0046	0.0771	0.0148	0.0464
	(0.009)	(0.048)	(0.014)	(0.030)
G_{t-1}	-0.3143***	-0.3129**	-0.3144***	-0.1112
	(0.012)	(0.156)	(0.012)	(0.157)
$\ln Empl_{t-1}$	-0.1620***	-0.4146**	-0.1620***	-0.2557
	(0.022)	(0.211)	(0.022)	(0.199)
$\ln Age_t$	-0.2079***	-0.3170***	-0.2083***	-0.2794***
	(0.057)	(0.082)	(0.057)	(0.074)
$Export_{t-1}$	0.0040	-0.1058***	0.0040	-0.0909**
	(0.015)	(0.039)	(0.015)	(0.038)
$PubFund_{t-1}$	0.0049	-0.0043	0.0045	0.0007
	(0.007)	(0.019)	(0.007)	(0.019)
$Group_{t-1}$	-0.0201	-0.0240	-0.0202	-0.0274
	(0.020)	(0.029)	(0.020)	(0.032)
Obs	26,386	21,291	26,386	21,291
AR(1)		0.021		0.002
AR(2)		0.377		0.761
Sargan		0.086		0.317
Hansen		0.336		0.261

Notes: Fixed Effects-Within (FE) and GMM-DIFF estimates of Equation (3.3). Regressions include a full set of year and sector dummies. Robust standard errors in parenthesis, clustered at firm-level: ***, ** and * indicate significance at 1%, 5% and 10% level, respectively. We also report p-values of Arellano-Bond test for first and second order serial correlation, AR(1) and AR(2), together with p-values of usual Sargan and Hansen tests for overidentifying restrictions.

GMM results reveal that process innovation does not affect growth. The estimated coefficient are small and not significant. One explanation, already suggested above, is that the role of process innovation on firm growth is mediated by productivity. Activities intended to change production processes or to eventually restructure the organization of production, tend to enhance firm efficiency, rather than directly affecting sales growth. We thus observe here the result of a lacking relationship between productivity and growth, recently suggested in several studies documenting that markets do not work as efficient selectors in redistributing market shares in favour of the more efficient firms (Bottazzi et al., 2008, 2010; Dosi et al., 2013). [7]

A similar reasoning can also apply to explaining the estimated effect of embodied

[7] We tested the correlation of process innovation with a rough proxy of labour productivity, namely the ratio bewteen total turnover and number of employees. The estimated coefficient was positive, very high in magnitude, and strongly significant.

Table 3.7: Panel estimates - Process Innov. and Embodied vs. Disembodied Tech. Change

	Innovation Proxy					
	Proc. Innov.		Emb.Tech.Change		Dis.Tech.Change	
	FE	GMM	FE	GMM	FE	GMM
Dep.Var. is G_t	(1)	(2)	(3)	(4)	(5)	(6)
INNOV_{t-1}	-0.0001	0.0058	0.3499***	-0.0004	0.9572	0.4413
	(0.009)	(0.162)	(0.125)	(0.002)	(0.730)	(1.139)
G_{t-1}	-0.3143***	-0.0710	-0.3134***	-0.0633	-0.3144***	0.0497
	(0.012)	(0.199)	(0.012)	(0.050)	(0.012)	(0.092)
$\ln Empl_{t-1}$	-0.1621***	-0.1189	-0.1610***	-0.3686*	-0.1620***	-0.4371
	(0.022)	(0.235)	(0.022)	(0.204)	(0.022)	(0.295)
$\ln Age_t$	-0.2077***	-0.2782***	-0.2031***	-0.2528***	-0.2045***	-0.1979**
	(0.057)	(0.085)	(0.056)	(0.068)	(0.056)	(0.078)
$Export_{t-1}$	0.0040	-0.0814**	0.0036	-0.0968**	0.0040	-0.2175**
	(0.015)	(0.038)	(0.015)	(0.038)	(0.015)	(0.100)
$PubFund_{t-1}$	0.0048	0.0095	0.0032	-0.0091	0.0049	-0.0163
	(0.007)	(0.038)	(0.007)	(0.019)	(0.007)	(0.059)
$Group_{t-1}$	-0.0201	-0.0288	-0.0203	-0.0272	-0.0199	-0.0273
	(0.020)	(0.033)	(0.020)	(0.033)	(0.020)	(0.034)
Obs	26,386	21,291	26,386	21,291	26,386	21,291
AR(1)		0.006		0.000		0.000
AR(2)		0.678		0.115		0.048
Sargan		0.257		0.119		0.061
Hansen		0.164		0.271		0.353

Notes: Fixed Effects-Within (FE) and GMM-DIFF estimates of Equation (3.3). Regressions include a full set of year and sector dummies. Robust standard errors in parenthesis, clustered at firm-level: ***, ** and * indicate significance at 1%, 5% and 10% level, respectively. We also report p-values of Arellano-Bond test for first and second order serial correlation, AR(1) and AR(2), together with p-values of usual Sargan and Hansen tests for overidentifying restrictions.

technical change (columns 3-4) through acquisition of new machineries. Indeed one can think that this specific activity has direct effects on productive efficiency or capacity utilization, while it only indirectly impacts on sales growth. And also in this case, we find no evidence of statistically significant effect in the GMM estimate.

In columns 5-6 we next find that also the proxy of disembodied technical change does not have a significant effect on subsequent growth. In line with the interpretation put forward above about the effect of external R&D, an explanation for the result calls for difficulties in managing the integration and the exploitation of knowledge sources (know how, patents, licenses) acquired outside the boundaries of the firm. Or, again making a parallel between external R&D and acquisition of external knowledge, it may also be that firms tend to source from outside only marginal “ingredients” of their overall innovation process, such that the effect on sales growth is at best indirect and in the end nil.[8]

To sum up, the (causal) effect of innovation is, in general, quite modest. Once controlling for endogeneity, only investing in R&D carried out internally stands out as a robust driver of subsequent sales growth. Of course, this conclusion only applies to the effect on the average of the conditional distribution. In this sense, our findings are not surprising, since they just extend to a large set of proxies of innovation the existing evidence that the "true" contribution of innovative activity is to spur extreme growth events, not an effect on the average growth rate. The next Section explores exactly this issue via Fixed-Effects quantile regressions.

Another interpretation is that, since we exploit within-firm variation, the contribution to sales growth coming from innovation is related to the sticky components of innovation activities, washed away with firm fixed-effects. Consider, for instance, the lacking effect we find for sales due to new products. Identification trough within-firm changes over time leaves open the possibility that we do not see a significant relationship with growth because much of the impact on growth goes through the persistent component of sales of new products. In other words, our negative results would be explained by the fact that product innovators keep a relatively persistent share of sales due to new products, while non-innovators hardly can manage to become innovators over time. And a similar reasoning, mutatis mutandis, can be extended to the other innovation variables for which we do not find significant results. This explanation, however, can have some relevance only in the case of the dummy indicator of process innovation. That variable is indeed fairly persistent, since "innovators" and "non-innovators" tend to remain like that over the sample period. The other innovation proxies are instead continuous variables that change over time: for all of them, although there is some persistence, we have verified that there is also considerable variation within firms.[9] Recall, finally, that we tested longer lag structures, so that the lacking effect estimated for most innovation variables cannot simply be explained by arguing that it takes more than one year for innovation to affect growth.

3.5.2 Fixed-Effects quantile regressions

The distributional analysis provided in Section 3.4 recalls one of the major stylized fact of industrial dynamics, stating that firm growth rates are characterized by a fat-tail distribution. This implies that standard regression analysis, capturing the effect on the expected value of the dependent, can only deliver a partial picture. Quantile regressions have become popular in recent years in the literature on firm growth and innovation (see review in Section 3.2), exactly because one can uncover the asymmetries characterizing the innovation-growth relationship along the spectrum of the growth rates distribution. Existing studies, however, beyond focusing only

[8] Also recall that only few firms engage in this activity, see Table 3.3 above.

[9] As a further check that the results are not driven by too little within-firm variation of the innovation proxies, we also performed a Correlated Random Effects estimation, adding the within-firm time series averages of both innovation variables and controls as further regressors. The coefficient estimates on the lagged innovation regressors are by definition equivalent to the FE estimates reported above. The coefficient on the average components, capturing the time invariant part of innovation activities, is positive and significant for all the innovation proxies but for external R&D and disembodied technical change. However, Correlated Random Effects do not tackle endogeneity.

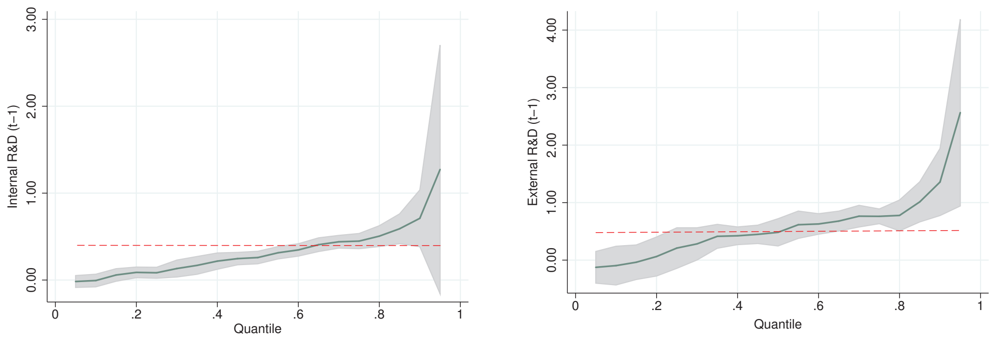

Figure 3.2: Fixed-Effects quantile regression estimates of coefficient α from Equation (3.3). Innovation proxies are Internal (left) and External (right) R&D intensity. The shaded areas represent 99% confidence band via bootstrapped standard errors. The horizontal line depicts FE estimates of α as benchmark.

on R&D and patents, apply basic quantile regression methods, which are easy to implement, but come at the cost of not controlling for unobserved firm-specific factors.

In this Section we exploit the Fixed-Effects quantile regression estimator developed in Canay (2011), explicitly allowing for firm-specific unobserved heterogeneity. Essentially, the method consists of a transformation of the response variable that allows to "wash out" the firm fixed effect. First rewrite our baseline Equation (3.3) as

$$Y_{i,t} = X'_{i,t}\beta + u_i + \epsilon_{i,t} \quad \text{, with} \quad E(\epsilon_{i,t}|X_i, u_i) = 0 \tag{3.4}$$

where the dependent $Y_{i,t}$ is sales growth G as defined above, $X_{i,t}$ contains the set of explanatory variables (each innovation indicator $INNOV$, alternatively, plus the controls), while u_i and $\epsilon_{i,t}$ are the firm fixed effect and the standard disturbance term.

Next, the Canay (2011) estimator proceeds in two steps: (i) obtain an estimate of the individual fixed effect through $\hat{u}_i = E_T[Y_{i,t} - X'_{i,t}\hat{\beta}]$, where $E_T(.) = T^{-1}\sum_{t=1}^{T}(.)$ and $\hat{\beta}$ is the standard Fixed-Effects (Within) estimator of β; (ii) build a transformed response variable $\hat{Y}_{i,t} = Y_{i,t} - \hat{u}_i$ and then obtain quantile regression coefficients through

$$\hat{\beta}(\tau) = \underset{\beta \in B}{\operatorname{argmin}} E_{nT}\left[\rho_\tau\left(\hat{Y}_{i,t} - X'_{i,t}\beta\right)\right] \quad , \tag{3.5}$$

which is just a quantile regression as in Koenker and Bassett (1978) on the transformed dependent variable.[10]

As we did with standard panel regression, we estimate our baseline Equation (3.3) separately for each innovation variable. In Figure 3.2, 3.3 and 3.4 we provide a

[10]The key assumption in the Canay estimator is that the fixed-effects are location shifters, meaning they affect all quantiles in the same way. An alternative FE quantile regression method is in Koenker (2004). The drawback of this solution rests in the large number of parameters to estimate, increasing the computation burden and the risk of non-convergence. In addition, a key assumption is that the longitudinal dimension is long enough to reduce the incidental parameter problem. Instead, Canay's procedure can be implemented on short longitudinal data.

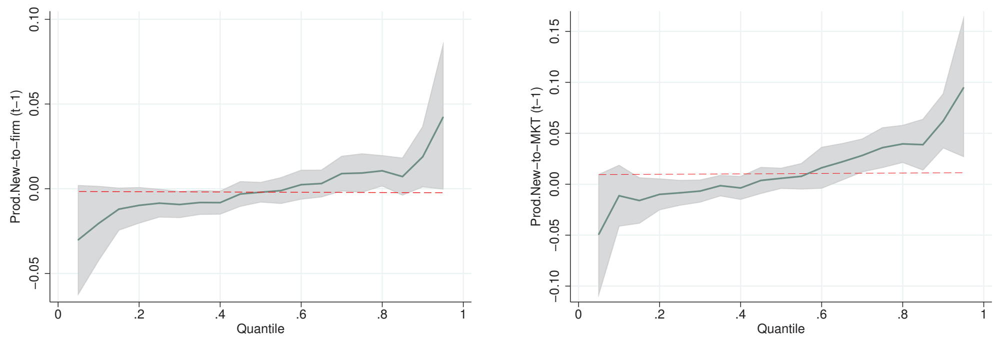

Figure 3.3: Fixed-Effects quantile regression estimates of coefficient α from Equation (3.3). Innovation proxies are % of sales due to products new-to-the-firm (left) and % of sales due to products new-to-the-market (right). Shaded areas represent 99% confidence band via bootstrapped standard errors. The horizontal line depicts FE estimates of α as benchmark.

graphical representation of the results, plotting the coefficient associated to the different innovation variables across the quantiles of the growth rates distribution.[11] To evaluate statistical significance, we also show a 99% confidence band, obtained from bootstrapped standard errors, as recommended in Koenker (2004) and Canay (2011). We also report an horizontal line indicating the FE coefficients estimated in the standard regression analysis.

Figure 3.2 shows the results for the two measures of internal and external R&D. The quantile regression curves reveal clear heterogeneity in the effect of each indicator across the growth rates distribution. Two results are worth noticing here, common across the two proxies. First, for shrinking firms, that is in the bottom quartile, R&D expenditures have a weak or not statistically significant association with growth. Second, the coefficient estimates increase and become positive and significant in the central part (till the third quartile), and then are larger than the FE estimates in the top quartile. These asymmetries reveal that R&D provides a strong contribution to growth performance of high-growth firms. The estimated coefficient on external R&D is twice as larger, but so is the standard error. The nil effect of R&D for firms belonging to the left tail is open to several interpretations. On the one hand, it may be that uncertainty of innovation often leads to unsuccessful outcomes, thereby making R&D efforts no more than a waste of resources. Our finding would imply that shrinking and non-growing firms are more often engaged in such unsuccessful dynamics. On the other hand, we can admit that R&D produces successful outcomes even for these firms, but the impact on sales is not strong enough to counteract a loss of market shares due to, for instance, a generally weak competitiveness of a firm in the market.

We comment on product innovation variables in Figure 3.3. For both variables the estimates resemble the patterns emerged for R&D. Indeed, there is no statistically significant association with growth until the top 15-20% of the growth rate distribution, and the estimates turn positive and significant in these top quantiles. This means that product innovation is relevant for high-growth. Noteworthy is the different magnitude of the estimated coefficients in the top quantiles: consistently

[11] See the tables in Appendix for full set of coefficient estimates (innovation variables and controls).

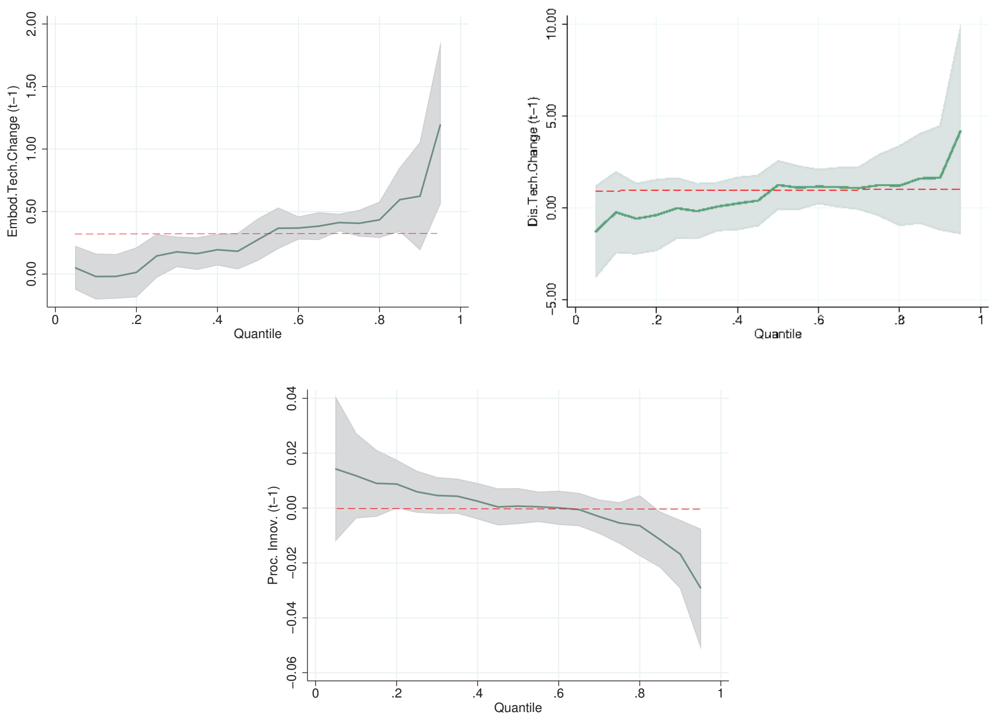

Figure 3.4: Fixed-Effects quantile regression estimates of coefficient α from Equation (3.3). Innovation proxies are Embodied (top-left) vs. Disembodied (top-right) technical change, and Process Innovation (bottom). Shaded areas represent 99% confidence band via bootstrapped standard errors. The horizontal line depicts FE estimates of α as benchmark.

with expectations, sales due to products new-to-the-market display a stronger association with total sales growth than sales due to products new-to-the-firm.

There is instead a peculiar behaviour in the left side of the support for the share of sales from products new to the firm. For shrinking firms, indeed, the estimated coefficient is negative. A temptative interpretation is that shrinking firms try to survive to market selection by imitating competitors and readjusting their product range, but competitive pressure is however too strong and hampers a recovery.

Next, in the top plots of Figure 3.4, we report the findings about embodied and disembodied technical change. Results for embodied technical change mimic what we observe for the R&D variables. The estimates tend to be small or not even significant in the first quartile, and then become positive and significant starting from the median and through the upper quartile. Conversely, disembodied technical change does not show any significant coefficient across the entire spectrum of the growth rates distribution.

The same result applies in the bottom plot of Figure 3.4, where we see that process innovation does not provide direct benefits in terms of sales growth. If anything, there might even be a mild negative effect among top-growing firms.

Overall, quantile regressions allow for two major qualifications of the standard panel analysis. First, the positive effect of internal R&D on growth is confirmed, but we discover that it actually originates for the most part from growing and fast-growing firms. Second, we find that some of the innovation variables which do not affect average growth do have, instead, a positive and significant effect on sales growth of high-growth firms in the top quantiles. This is the case for external

R&D, for product innovation (new-to-the-market, in particular) and also for technical changed embodied in the acquisition of new machineries. We instead fully confirm the minor role of process innovation and disembodied technical change.

3.6 Testing complementarity of innovation activities

In this Section we explore if sales growth originates from combinations of different innovation activities, rather than from each single one. Indeed, firms in reality often pursue different innovation strategies, undertaking different innovation activities at the same time. The outcome of innovation in terms of sales growth can be different depending on the complexity of the strategy pursued, in terms of the number and the type of activities performed at the same time. Each different combination may entail specific costs and challenging coordination issues, while also increasing the ability to create and capture growth opportunities.

The key question is whether different innovation activities are complements in their effect on growth. We explore this issue through the concept of super-modularity. In general terms, consider a function $f(\mathbf{X})$, where $\mathbf{X}$ is a vector of binary arguments, $\mathbf{X}=\{X_1, X_2, \ldots, X_n\}$, with $X_j = \{0, 1\}$ depending whether a certain action j is undertaken or not. Action X_j and X_i are complements if f is supermodular in X_j and X_i, that is

$$f(X_j \vee X_i) + f(X_j \wedge X_i) \geq f(X_j^c) + f(X_i^c) \quad , \tag{3.6}$$

where X^c stands for "non-X".

The idea is simply that the effect of choosing X_j on the objective function f is larger if also X_i is chosen at the same time, as compared to other possible combinations where X_j appears but X_i is not chosen. This approach to complementarity is adopted by a number of studies exploring complementarity of different innovation inputs, or of obstacles to innovation, in generating innovation outputs (see, e.g., Mohnen and Roller, 2005; Cassiman and Veugelers, 2006; Catozzella and Vivarelli, 2014). We apply the same framework to explore super-modularity of the growth function with respect to innovation activities.

We proceed as follows. Firstly, we group our original seven innovation activities into four categories, capturing the different types of innovation output (product vs. process) and the different innovation inputs (R&D vs. other inputs), also distinguishing between internal vs. external sources. Accordingly, we define the following dummy variables:

- Internal innovation (INT) = 1 if the firm performs intra-mural R&D, 0 otherwise.
- External innovation (EXT) = 1 if the firm performs extra-mural R&D or acquires embodied or disembodied knowledge, 0 otherwise.
- Product innovation (NEWP) = 1 if the firms introduces new products.
- Process innovation (PROC) = 1 if the firms introduces new or significantly improved processes.

Table 3.8: Innovation strategies

Strategy	INT	EXT	NEWP	PROC	Combination
STR_0	0	0	0	0	No inno
STR_1	0	0	0	1	PROC
STR_2	0	0	1	0	NEWP
STR_3	0	0	1	1	NEWP&PROC
STR_4	0	1	0	0	EXT
STR_5	0	1	0	1	EXT&PROC
STR_6	0	1	1	0	EXT&NEWP
STR_7	0	1	1	1	EXT&NEWP&PROC
STR_8	1	0	0	0	INT
STR_9	1	0	0	1	INT&PROC
STR_{10}	1	0	1	0	INT&NEWP
STR_{11}	1	0	1	1	INT&NEWP&PROC
STR_{12}	1	1	0	0	INT&EXT
STR_{13}	1	1	0	1	INT&EXT&PROC
STR_{14}	1	1	1	0	INT&EXT&NEWP
STR_{15}	1	1	1	1	INT&EXT&NEWP&PROC

Notice that these are mutually exclusive categories. Of course, firms may engage in none, just one or more of these activities at the same time. We build all the possible combinations among these four categories, ending up with a total of $2^4 = 16$ possible "innovation strategies". These are listed in Table 3.8. So, for instance, STR_0 is a dummy that takes value 1 if a firm does not engage in any of the four basic activities. This is also conventionally indicated as S_{0000}. STR_1 is a strategy where a firm only engage in process innovation (S_{0001}), and so on.

Next, we specify the growth function as a regression of sales growth against the set of alternative strategies

$$G(S, \mathbf{Z}) = f(S_{0001}, S_{0010}, ..., S_{1111}, \mathbf{Z}), \tag{3.7}$$

where G is sales growth, $\mathbf{Z}$ is the usual set of lagged controls as in the main Equation (3.3), and we normalize S_{0000} to zero. Notice that the strategy dummies are measured in $t-1$ and change over time.

The definition of super-modularity of G with respect to the lattice S means

$$G(S' \vee S'', \mathbf{Z}) + G(S' \wedge S'', \mathbf{Z}) \geq G(S', \mathbf{Z}) + G(S'', \mathbf{Z}) \quad . \tag{3.8}$$

The number of non trivial inequalities implied by the definition is $2^{(K-2)} \sum_{i=1}^{K-1} i$, where K is the number of basic categories for which one wants to assess pairwise complementarity (Topkis, 1998). In our case, $K = 4$ and we thus have a total of 24 nontrivial inequality constraints, 4 for each pairwise combination of basic innovation activities. Labeling as b_j the coefficient on the dummy STR_j estimated from Equation (3.7), the constraints can be compactly written as:

- Complementarity INT-EXT: $b_{8+s} + b_{4+s} \leq b_{0+s} + b_{12+s}$ with $s = 0, 1, 2, 3$

- Complementarity INT-NEWP: $b_{8+s} + b_{2+s} \leq b_{0+s} + b_{10+s}$ with $s = 0, 1, 4, 5$

- Complementarity INT-PROC: $b_{8+s} + b_{1+s} \leq b_{0+s} + b_{9+s}$ with $s = 0, 2, 4, 6$

- Complementarity EXT-NEWP: $b_{4+s} + b_{2+s} \leq b_{0+s} + b_{6+s}$ with $s = 0, 1, 8, 9$

- Complementarity EXT-PROC: $b_{4+s} + b_{1+s} \leq b_{0+s} + b_{5+s}$ with $s = 0, 2, 8, 10$

- Complementarity NEWP-PROC: $b_{2+s} + b_{1+s} \leq b_{0+s} + b_{3+s}$ with $s = 0, 4, 8, 12$

For each pair, the constraints must hold jointly. To implement the test, we exploit the Wald-type statistic and the procedure derived in Kodde and Palm (1986). Let $\gamma = (b_{0001}, b_{0010}, ..., b_{1111})'$ the coefficients to be estimated from the growth function in (3.7). Then, the test statistic is given as

$$D = (C\tilde{\gamma} - C\hat{\gamma})'(C'cov(\hat{\gamma})C)^{-1}(C\tilde{\gamma} - C\hat{\gamma}) \tag{3.9}$$

with

$$\tilde{\gamma} = \underset{\gamma}{\operatorname{argmin}}(C\gamma - C\hat{\gamma})'(C'cov(\hat{\gamma})C)^{-1}(C\gamma - C\hat{\gamma}) \quad s.t. \quad C\gamma \leq 0 \tag{3.10}$$

where $\hat{\gamma}$ is the estimate of γ from the growth function in (3.7) and $cov(\hat{\gamma})$ the associated covariance matrix, while C is a matrix that maps the coefficients into the inequality constraints stated above. The set of coefficient $\tilde{\gamma}$ is obtained as the closest value to the estimates of γ under the restrictions imposed by the matrix C, and it can be computed via quadratic minimization under inequality constraints. The D statistic does not have an exact distribution, but Kodde and Palm (1986) provide lower and upper bounds for different levels of significance. The null of complementarity is accepted for values of D below the lower bound and it is rejected for values above the upper bound, whereas the test is inconclusive if the estimated D falls between the two bounds.

The main requirement for the procedure to work is that $\hat{\gamma}$ is a consistent estimate of γ. We estimate the growth function via the GMM-DIFF estimator. This allow, once again, to control for firm fixed-effects and endogeneity of innovation strategies and controls.

Results are presented in Table 3.9. In the left panel we show the estimates of the growth function. The set of instruments includes lags of growth and controls, as well as lag-2 of the innovation strategies in the set S. The coefficients on the strategies are all positive, but most of them are not significant, except for STR_4 (i.e., EXT alone), STR_8 (INT alone), STR_{10} (combination of INT and NEWP), and STR_{13} (INT+EXT+PROC).

The coefficients as such convey little information, as they do not provide a formal test of complementarity. The super-modularity tests are presented in the right panel. We report the estimated D statistic for the different pairwise combinations of the basic innovation activities. Cases where the null of complementarity cannot be rejected are in bold (at the 10% level, which seems standard in previous studies).

Results support complementarity only in two cases. First, we find that there is complementarity between INT and NEWP, meaning that these two activities are more important for growth when done together, than when they are carried out separately. We therefore confirm the crucial role of internal R&D, but we can

Table 3.9: Estimation results & complementarity test

Dep.Var. is G_t	Estimation (1)
$\mathrm{STR}_{1,t-1}$	0.0293
	(0.063)
$\mathrm{STR}_{2,t-1}$	0.0769
	(0.147)
$\mathrm{STR}_{3,t-1}$	0.1986
	(0.155)
$\mathrm{STR}_{4,t-1}$	0.4623**
	(0.208)
$\mathrm{STR}_{5,t-1}$	0.0309
	(0.080)
$\mathrm{STR}_{6,t-1}$	-0.1448
	(0.454)
$\mathrm{STR}_{7,t-1}$	0.1330
	(0.140)
$\mathrm{STR}_{8,t-1}$	0.1798**
	(0.090)
$\mathrm{STR}_{9,t-1}$	0.0300
	(0.102)
$\mathrm{STR}_{10,t-1}$	0.1849*
	(0.105)
$\mathrm{STR}_{11,t-1}$	0.1464
	(0.114)
$\mathrm{STR}_{12,t-1}$	0.0886
	(0.123)
$\mathrm{STR}_{13,t-1}$	0.1888**
	(0.095)
$\mathrm{STR}_{14,t-1}$	0.2091
	(0.129)
$\mathrm{STR}_{15,t-1}$	0.1160
	(0.104)
G_{t-1}	-0.3042***
	(0.092)
$\ln Empl_{t-1}$	-0.1279
	(0.180)
$\ln Age_t$	-0.3182***
	(0.074)
$Export_{t-1}$	-0.2848**
	(0.120)
$PubFund_{t-1}$	0.0324
	(0.069)
$Group_{t-1}$	-0.0323
	(0.031)
Obs	21,291
AR(1)	0.000
AR(2)	0.133
Sargan	0.120
Hansen	0.131

Complementarity test: Pair	Wald statistic
INT-EXT	5.3215
INT-NEWP	**1.6045**
INT-PRO	4.8413
EXT-NEWP	3.0288
EXT-PRO	6.0155
NEWP-PRO	**1.6156**

Notes: GMM-DIFF estimates of Equation (3.7). Regression includes a full set of year dummies. Robust standard errors in parenthesis, clustered at firm-level: ***, ** and * indicate significance at 1%, 5% and 10% level, respectively. We also report p-values of Arellano-Bond test for first and second order serial correlation, AR(1) and AR(2), together with p-values of usual Sargan and Hansen tests for overidentifying restrictions.

Complementarity test: bold values indicate acceptance of complementarity at 10% significance level (lower bound = 1.642, upper bound = 7.094).

add that internal R&D pays even more in terms of growth when it is carried out together with product innovation. At the same time, we recover here a role for product innovation, suggesting that introduction of new products is more likely to impact on growth when formal R&D activities are carried out internally.

Second, there is complementarity between process and product innovation. This result, on the one hand, further highlights that product innovation is more beneficial when coupled with other activities, as we just saw for its combination with R&D. On the other hand, we recover here a role for process innovation. While in the panel and quantile analysis we concluded that process innovation alone does not directly affect growth, we now find that it has an effect in combination with the capacity to introduce new products. The result indeed indicates that restructuring of production processes is effective if combined with a simultaneous change in the share of new products (new-to-the-market). Conversely, process innovation is not complement with the other innovation activities, confirming the overall weak role of such variable in fostering sales growth.

Finally, we do not detect any complementarity of external sourcing of knowledge with none of the other innovation activities. This finding once again calls for an already emerged difficulty in integrating knowledge and technologies produced outside the firm, due, e.g., to complex coordination with external "providers" or to weak absorptive capacity. In this respect, complementarity analysis confirm the conclusion emerging from the analysis of the separate role of external R&D and of embodied and disembodied technical change.

3.7 Conclusions

The relationship between innovation and firm growth is a classical, yet still puzzling topic. While theory tends to predict a strong positive link, the empirical literature provides mixed results. Moreover, most studies tend to focus on the effect of innovation on productivity and employment growth, perhaps given the important implications for economic growth, job creation and destruction.

This essay, by taking advantage of a rich panel on innovation activity of Spanish manufacturing firms, provides new evidence on the relationships between success on the market, in terms of sales growth, and a richer set of innovation dimensions, capturing innovation inputs and outputs as well as different modes of sourcing new knowledge.

The overall picture emerging from the analysis suggests a good deal of heterogeneity in the capacity of different innovation activities to support expansion of sales and market shares.

First, from standard panel regression analysis, controlling for firm fixed-effects and endogeneity, we find that internal R&D is the only innovation indicator significantly (and positively) related with growth. Conversely, we are not able to find any significant effect of external R&D, process innovation, increases in the sales due to new products, as well as of acquisition of embodied or disembodied new technologies. This negative result is striking, at first, as one generally thinks that innovation spurs growth. But we contribute to provide explanations for some of these findings. The lacking correlation between activities that involve external sourcing of new knowledge or new technologies (external R&D, disembodied and embodied technical change) supports the view that valuable knowledge is inherently firm-specific. Firms may face difficulties in establishing effective collaboration with external providers, or may lack of specific absorptive capacities in integrating external knowledge and technologies. The equally lacking effect of process innovation can be interpreted as a signal that new processes are primarily designed to improve efficiency or to change production modes, and may affect sales growth only indirectly. And, the weak role of product innovation may just reflect that the share of sales due to new products is on average small. Overall, our findings conform with previous studies that highlight how the effect of innovation activities on average growth may be difficult to detect.

We recover a positive effect for most of the innovation variables when we look at their association with growth along the entire spectrum of the growth rates distribution. Estimates of quantile regression controlling for firm fixed-effects show that all variables, with the exception of process innovation and disembodied technical change, have a positive and significant coefficient in the upper quartile, that is for high-growth episodes. In this respect, we support that many innovation activities are beneficial, but only for fast growing units. Notice that this result adds to the emerging literature underlying the peculiarities of high-growth firms, which has so far explored a more limited set of innovation indicators (R&D and patents) as drivers of growth.

The analysis of the complementarities between innovation activities adds further insights. We confirm the importance of internal R&D as a driver of sales growth, but we also recover a role for both product and process innovation. Indeed, we find that the beneficial effect of internal R&D is stronger when coupled with product innovation, and that process and product innovation have a stronger association with growth if carried out together.

The research agenda is of course open to further developments, in particular to extend the analysis on the interactions among different innovation activities we consider here. We foresee many possible extensions, perhaps requiring richer datasets. A first strategy could be to further exploit the distinction between innovative inputs (internal and external R&D, embodied and disembodied technical change) and innovative outputs (sales due to new products, and process innovation). One could exploit a standard Crepon et al. (1998) type of framework, possibly modified to also explore complementarities and modularity, to reconstruct how different inputs contribute to different outputs of innovation, and the ensuing effect of different innovative configurations on growth on the market. Or, perhaps with richer datasets longer in time, one might identify the effect of sequential adoption of basic innovation strategies, exploring in more details whether, e.g., acquisition of new machineries turns out to have a positive impact on growth only after a subsequent process innovation related to that acquisition is implemented. Second, although our analysis of complementarities already incorporates the idea that firms engage in a different number and in different types of basic innovation activities, one can imagine to deepen the analysis of the relationship between growth and the "complexity level" of firms' innovation strategies. For instance, one may think of taxonomies seeking to characterize complexity in terms of some measure of the coherence among the different innovation activities performed within each firm, and assess whether this translates into differential patterns of growth. Our results, so far, suggest that a combination of internal R&D, process and product innovation is the key candidate to provide the more effective mix of growth-enhancing strategies, especially in view of their observed strong relationship with high-growth episodes.

3.8 Appendix

For completeness, we present here tables reporting all the coefficient estimates from fixed-effects quantile regressions applied to our baseline model in Equation (3.3). Graphical analysis of the results obtained for each innovation variable, and related comments, are presented in the main text.

We remark on the estimates obtained for the set of controls. Firstly, across all the specifications, that is irrespective of the innovation proxy considered, we observe a negative growth autocorrelation coefficient across all the quantiles. This result suggests that all firms, either growing or shrinking in one year, are unlikely to repeat the same growth performance in the following year. Second, and again robustly across different innovation indicators, we observe a negative correlation of size and age with sales growth. In both cases, moreover, the estimated coefficient is increasing (in absolute value) when moving from the left to the right tails of the growth rate distribution. The evidence connects to the well known finding that smaller and younger firms tend to grow faster, although the quantile profile here allows to add that the "detrimental effect" of age and size seems stronger for big positive jumps. Finally, across all the innovation dimensions, we observe some variability across quantiles in the coefficient estimates of the three control dummies on export status, public financial support and group membership. The export dummy plays a positive and significant association at lower quantiles, while the association becomes negative and significant for high-growth firms. This evidence recalls results in Hölzl (2009) who finds a negative relationship between export and growth performance in countries of Southern Europe (Italy, Portugal, Greece, Spain). Conversely, being part of industrial group is negatively related with sales growth across almost all quantiles, while it has a positive coefficient on the very top tail of the growth distribution. Public financial support to innovation does not have any significant relationship with sales growth, a result that might cast doubts on the effectiveness of such supporting schemes.

Table 3.10: Quantile regressions – Internal R&D

	Quantile (%)				
	10	25	50	75	90
Internal R&D$_{t-1}$	-0.006	0.083**	0.259***	0.448***	0.710***
	(0.047)	(0.037)	(0.038)	(0.053)	(0.127)
G$_{t-1}$	-0.225***	-0.209***	-0.211***	-0.223***	-0.236***
	(0.019)	(0.009)	(0.009)	(0.010)	(0.014)
$\ln Empl_{t-1}$	-0.136***	-0.151***	-0.161***	-0.171***	-0.186***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
$\ln Age_t$	-0.138***	-0.157***	-0.171***	-0.181***	-0.201***
	(0.005)	(0.003)	(0.002)	(0.003)	(0.005)
$Export_{t-1}$	0.026***	0.011***	0.000	-0.010*	-0.032***
	(0.010)	(0.004)	(0.003)	(0.005)	(0.010)
$PubFund_{t-1}$	0.006	0.000	-0.006**	-0.004	-0.015**
	(0.006)	(0.003)	(0.003)	(0.003)	(0.006)
$Group_{t-1}$	-0.054***	-0.030***	-0.022***	-0.014***	0.014**
	(0.008)	(0.003)	(0.003)	(0.004)	(0.007)
Observations	26,386	26,386	26,386	26,386	26,386
Industry dummies	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes

Notes: Fixed-effects quantile regression estimates of Equation (3.3). Bootstrapped standard errors in parenthesis. ***, ** and * indicate significance on a 1%, 5% and 10% level, respectively.

Table 3.11: Quantile regressions – External R&D

	Quantile (%)				
	10	25	50	75	90
External R&D$_{t-1}$	-0.096	0.209	0.483***	0.761***	1.358***
	(0.165)	(0.161)	(0.122)	(0.085)	(0.495)
G$_{t-1}$	-0.225***	-0.212***	-0.214***	-0.225***	-0.244***
	(0.019)	(0.009)	(0.009)	(0.011)	(0.015)
$\ln Empl_{t-1}$	-0.136***	-0.151***	-0.162***	-0.173***	-0.190***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
$\ln Age_t$	-0.159***	-0.180***	-0.194***	-0.208***	-0.230***
	(0.005)	(0.003)	(0.002)	(0.003)	(0.005)
$Export_{t-1}$	0.025**	0.010**	0.001	-0.010**	-0.029***
	(0.010)	(0.004)	(0.003)	(0.005)	(0.010)
$PubFund_{t-1}$	0.004	-0.001	-0.004	0.001	-0.005
	(0.006)	(0.004)	(0.002)	(0.003)	(0.006)
$Group_{t-1}$	-0.054***	-0.029***	-0.022***	-0.013***	0.015**
	(0.008)	(0.004)	(0.003)	(0.004)	(0.007)
Observations	26,386	26,386	26,386	26,386	26,386
Industry dummies	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes

Notes: Fixed-effects quantile regression estimates of Equation (3.3). Bootstrapped standard errors in parenthesis. ***, ** and * indicate significance on a 1%, 5% and 10% level, respectively.

Table 3.12: Quantile regressions – Prod.New-to-firm

	Quantile (%)				
	10	25	50	75	90
Prod.New-to-firm$_{t-1}$	-0.020**	-0.009**	-0.002	0.009*	0.019**
	(0.009)	(0.004)	(0.004)	(0.006)	(0.009)
G$_{t-1}$	-0.224***	-0.211***	-0.217***	-0.226***	-0.244***
	(0.019)	(0.009)	(0.009)	(0.010)	(0.015)
$\ln Empl_{t-1}$	-0.135***	-0.151***	-0.162***	-0.174***	-0.192***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
$\ln Age_t$	-0.169***	-0.192***	-0.206***	-0.220***	-0.241***
	(0.005)	(0.003)	(0.002)	(0.003)	(0.005)
$Export_{t-1}$	0.028***	0.011***	0.001	-0.012**	-0.036***
	(0.010)	(0.004)	(0.003)	(0.005)	(0.010)
$PubFund_{t-1}$	0.001	-0.001	-0.003	0.004	0.004
	(0.006)	(0.003)	(0.003)	(0.003)	(0.006)
$Group_{t-1}$	-0.055***	-0.031***	-0.022***	-0.012***	0.017**
	(0.008)	(0.004)	(0.003)	(0.004)	(0.007)
Observations	26,386	26,386	26,386	26,386	26,386
Industry dummies	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes

Notes: Fixed-effects quantile regression estimates of Equation (3.3). Bootstrapped standard errors in parenthesis. ***, ** and * indicate significance on a 1%, 5% and 10% level, respectively.

Table 3.13: Quantile regressions – Prod.New-to-MKT

	Quantile (%)				
	10	25	50	75	90
Prod.New-to-MKT$_{t-1}$	-0.011	-0.008	0.006	0.036***	0.062***
	(0.017)	(0.006)	(0.006)	(0.010)	(0.013)
G$_{t-1}$	-0.223***	-0.212***	-0.216***	-0.228***	-0.244***
	(0.019)	(0.009)	(0.009)	(0.010)	(0.015)
$\ln Empl_{t-1}$	-0.135***	-0.151***	-0.162***	-0.174***	-0.192***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
$\ln Age_t$	-0.170***	-0.192***	-0.207***	-0.221***	-0.242***
	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)
$Export_{t-1}$	0.027***	0.011***	0.001	-0.012**	-0.036***
	(0.010)	(0.004)	(0.003)	(0.005)	(0.010)
$PubFund_{t-1}$	0.003	-0.001	-0.002	0.003	0.004
	(0.006)	(0.003)	(0.002)	(0.003)	(0.006)
$Group_{t-1}$	-0.056***	-0.031***	-0.022***	-0.012***	0.017**
	(0.008)	(0.004)	(0.003)	(0.004)	(0.007)
Observations	26,386	26,386	26,386	26,386	26,386
Industry dummies	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes

Notes: Fixed-effects quantile regression estimates of Equation (3.3). Bootstrapped standard errors in parenthesis. ***, ** and * indicate significance on a 1%, 5% and 10% level, respectively.

Table 3.14: Quantile regressions – Process Innovation dummy

	Quantile (%)				
	10	25	50	75	90
Proc. Innov_{t-1}	0.012	0.006*	0.001	-0.005	-0.017***
	(0.007)	(0.004)	(0.003)	(0.004)	(0.006)
G_{t-1}	-0.224***	-0.215***	-0.216***	-0.228***	-0.241***
	(0.020)	(0.009)	(0.009)	(0.010)	(0.014)
$\ln Empl_{t-1}$	-0.136***	-0.152***	-0.162***	-0.175***	-0.191***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
$\ln Age_t$	-0.168***	-0.191***	-0.206***	-0.220***	-0.243***
	(0.005)	(0.003)	(0.002)	(0.003)	(0.005)
$Export_{t-1}$	0.024**	0.010**	0.001	-0.010*	-0.032***
	(0.010)	(0.004)	(0.003)	(0.005)	(0.010)
$PubFund_{t-1}$	-0.003	-0.003	-0.003	0.005	0.007
	(0.006)	(0.003)	(0.003)	(0.003)	(0.006)
$Group_{t-1}$	-0.053***	-0.031***	-0.022***	-0.012***	0.015**
	(0.008)	(0.004)	(0.003)	(0.004)	(0.007)
Observations	26,386	26,386	26,386	26,386	26,386
Industry dummies	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes

Notes: Fixed-effects quantile regression estimates of Equation (3.3). Bootstrapped standard errors in parenthesis. ***, ** and * indicate significance on a 1%, 5% and 10% level, respectively.

Table 3.15: Quantile regressions – Embod.Tech.Change

	Quantile (%)				
	10	25	50	75	90
$\text{Emb.Tech.Change}_{t-1}$	-0.019	0.145	0.275***	0.407***	0.623***
	(0.077)	(0.095)	(0.106)	(0.059)	(0.163)
G_{t-1}	-0.227***	-0.212***	-0.217***	-0.232***	-0.249***
	(0.019)	(0.009)	(0.009)	(0.011)	(0.014)
$\ln Empl_{t-1}$	-0.135***	-0.151***	-0.161***	-0.174***	-0.190***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
$\ln Age_t$	-0.165***	-0.187***	-0.202***	-0.215***	-0.237***
	(0.005)	(0.003)	(0.002)	(0.003)	(0.005)
$Export_{t-1}$	0.026**	0.011***	0.001	-0.008*	-0.032***
	(0.010)	(0.004)	(0.003)	(0.005)	(0.010)
$PubFund_{t-1}$	0.001	-0.002	-0.003	0.002	0.001
	(0.006)	(0.003)	(0.003)	(0.003)	(0.006)
$Group_{t-1}$	-0.053***	-0.031***	-0.022***	-0.012***	0.018**
	(0.008)	(0.004)	(0.003)	(0.004)	(0.007)
Observations	26,386	26,386	26,386	26,386	26,386
Industry dummies	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes

Notes: Fixed-effects quantile regression estimates of Equation (3.3). Bootstrapped standard errors in parenthesis. ***, ** and * indicate significance on a 1%, 5% and 10% level, respectively.

Table 3.16: Quantile regressions – Disemb.Tech.Change

	Quantile (%)				
	10	25	50	75	90
Dis.Tech.Change$_{t-1}$	-0.249	-0.024	1.238*	1.230*	1.625
	(1.077)	(0.856)	(0.737)	(0.742)	(1.376)
G$_{t-1}$	-0.225***	-0.213***	-0.216***	-0.226***	-0.242***
	(0.020)	(0.009)	(0.009)	(0.010)	(0.015)
$\ln Empl_{t-1}$	-0.135***	-0.151***	-0.162***	-0.175***	-0.192***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
$\ln Age_t$	-0.166***	-0.189***	-0.203***	-0.217***	-0.239***
	(0.005)	(0.003)	(0.002)	(0.003)	(0.005)
$Export_{t-1}$	0.028***	0.010**	0.001	-0.011**	-0.033***
	(0.010)	(0.004)	(0.003)	(0.005)	(0.010)
$PubFund_{t-1}$	-0.000	-0.001	-0.002	0.004	0.004
	(0.006)	(0.003)	(0.003)	(0.003)	(0.006)
$Group_{t-1}$	-0.055***	-0.030***	-0.021***	-0.012***	0.017**
	(0.008)	(0.004)	(0.003)	(0.004)	(0.007)
Observations	26,386	26,386	26,386	26,386	26,386
Industry dummies	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes

Notes: Fixed-effects quantile regression estimates of Equation (3.3). Bootstrapped standard errors in parenthesis. ***, ** and * indicate significance on a 1%, 5% and 10% level, respectively.

Does persistence of innovation spur persistence of growth?

4.1 Introduction

This essay examines the role of persistence of innovation on persistence of growth at the firm-level. The main issue we address in our study is to empirically assess whether a systematic engagement in innovation activities induces a less erratic structure in the process of firm growth. Or, put more simply, whether persistent innovators exhibit persistently superior growth performance.

To the best of our knowledge this is the first attempt linking explicitly the evidence on the patterns of innovation with what is known about persistent growth.[1] The lack of empirical regularities on the subject is, however, somewhat surprising since in the last decades two abundant streams of literature, one focused on persistence of growth, the other on persistence of innovation, have been developed in parallel but never reconciled one another.

Concerning the former, contributions from economics and management help in providing theoretical conjectures on why some firms persistently outperform in our economy. The rationale can be simplified as follows: efficiency (or productivity) is the central channel through which companies achieve growth and gain market shares at the expenses of less efficient units, either directly because higher efficency gets reflected into lower prices, or indirectly because higher efficiency implies increasing profits which, in combination with sounder financial conditions, grant the access to additional resources needed to further investments (among the many contributions see the primordial discussion in Penrose, 1959, or the theoretical models of firm-industry evolution with heterogeneous agents as in Nelson and Winter, 1982; Jovanovic, 1982; Cooley and Quadrini, 2001). Higher efficiency is assumed to be strongly related to more attractive firm-specific unobserved factors, technological and organizational traits in primis. These dimensions create value for consumers, are unique (or at least not identical than that possessed by rivals), durable, not (in principle) inimitable, and generate returns which are appropriable (Teece et al., 1997; Teece, 2007). Therefore, the degree of persistence in the process of growth that one should observe is connected to the structural competitive advantages (comparative and differential) that the firm exploits over time, at least until the fingerprint of market competition erodes such advantages (Geroski, 2002).

Persistence of innovation also finds several theoretical explanations. The "suc-

[1] As recently stressed in Dosi (2007) theoretical and empirical contributions in this direction are an hard but urgent challenge for future research on industry dynamics.

cess breeds success" hypothesis is perhaps among the most accredited: succesful innovation broadens both a firm's technological opportunities and its market power which, in turn, positively affect the conditions for subsequent innovations (Flaig and Stadler, 1994). In addition to that, the cumulative and additionality nature of technological knowledge is such that firms are likely to experience some dynamic increasing returns in their innovation activities (Stiglitz, 1987). Some scholars have put forward interpretations related to a process of learning by innovating which enhance knowledge stocks and technological capabilities, at last inducing state dependence in innovation behaviour (Dosi et al., 2008). Finally, innovation activities are typically characterized by high start-up sunk costs which are necessary to set up the entire R&D apparatus (research infrastructures and the staff), thus once the decision has been taken, the opportunity costs to stop are typically too high (Sutton, 1991).

There exists a vast literature linking growth events to innovation potential and, certainly, the latter is considered by economists and scholars from strategic management tradition as one among the key factors of competitiveness. However, the degree of persistence in innovation which is necessary to build durable competitive advantages, that in turn might get reflected into sustained sales expansion, is somewhat under investigated and some questions remain still unanswered. Are firms that systematically engaged into innovation activities those firms experiencing more persistent growth patterns? Or, on the opposite, can a sporadic innovative behaviour be sufficient to overcome competitors over long periods of time? Answering these questions is relevant to understand to what extent, and how quickly, market competition erodes the competitive advantages that firms build upon their innovation success, as well as the longlasting consequences that different innovation behaviours may induce.

Against this background the research question that we pose in this essay is the following: does persistence of innovation spur persistence of growth?

To answer this question we exploit a rich panel of Spanish firms comprising micro-level information about firms and market characteristics and spanning a period of twenty years, from 1990 to 2009. We empirically investigate the determinants of persistent sales growth by mean of duration model techniques: more precisely, we relate the probability and the length of a period of continuous positive growth (what we will define as "growth spell") to the "innovation status" defined in a time window prior the persistent expansion. We build the innovation status of each firm on the basis of the frequency with which it undertakes R&D activities or applies for patents, in turn distinguishing three distinct categories of firms: non-innovators, occasional, and persistent innovators.

Although we do observe striking differences among groups, being persistent innovators generally larger, older, more productive and active in high-tech sectors, we find weak support to the conjecture that a systematic engagement in R&D activities is beneficial to persistent sales expansion. However, when we focus on innovation outcome, proxied by patent applications, we do find a positive and significant association between persistence of innovation and persistent of growth performance. Put it differently, we observe that firms which are persistently able to produce innovate outputs display subsequent longer periods of sustained positive growth. These findings are robust to a number of extensions and alternative definitions of the innovation status. All in all our evidence suggests that a firm's competitive advantages have a higher propensity to be sustainable only if the firm is able to persistently, and not just occasionally, broaden its technological capabilities.

4.2 Background literature and research objectives

4.2.1 On the persistence of innovation

Although, as discussed in the previous Section, the theoretical literature provides several explanations supporting the path-dependent and persistent nature of innovation, the empirical evidence is far from conclusive. The latter is indeed very dependent on the proxy of innovation considered (i.e. input vs. output measures) as well as on data specificities (i.e. country, industry).

The vast majority of the empirical literature on the subject has focused on the output side of the firm's innovation activity, relying on different measures such as patents, share of sales stemming from innovative products, and realization of process and/or product innovations.

Patent-based studies, on average, have largely found little evidence of persistence. Malerba and Orsenigo (1999), using European Patent office (EPO) data for 6 different countries, find that only a negligible proportion of firms persist in its patenting activity. Interestingly, the number of patents granted to these firms increases exponentially over time, suggesting that persistent innovators, although few in number, are responsible for a very large share of total patents. Along the same line, Geroski et al. (1997), drawing upon a representative sample of UK innovative firms observed for the period 1969-1988, come to similar conclusions: very few firms innovate persistently, and they do so only after reaching an initial threshold of five patents. Cefis and Orsenigo (2001) and Cefis (2003), applying transition probabilities approach to EPO data, further corroborate this evidence. Their results show also that persistence varies greatly across industries and size dimensions.

As well known, patent is a peculiar proxy of innovation activity since only a very small percentage of inventions are actually patented. This might, to some extent, explain the low level of persistence of innovation detected by previous studies. At the same time, persistence measured using patent data has the advantage to implicitly indicate persistence of innovative leadership (Duguet and Monjon, 2004) and, as we will point out later in this Section, this issue is particularly relevant in our context.

Other scholars have focused on different indicators of innovative success, such as the realization of products and process innovations and the share of innovative sales. König et al. (1994) and Flaig and Stadler (1994), using a panel of German manufacturing firms find support of high persistence for both process and product innovations. Raymond et al. (2010), by applying a dynamic two-type tobit model on Dutch CIS panel data, provide evidence of state dependence only in the high-tech industries. A similar conclusion is reached by Antonelli et al. (2012) which, accounting for complementarity effects between the product and process innovation, find robust evidence of persistence for Italian companies.

Finally, some contributions have considered the input side of the innovation activity, analysing the degree of persistence of R&D activities. In this respect, regardless the econometric methodology and the data used, a high level of persistence in innovation activity is detected (see Peters 2009; Mañez et al. 2009; Arqué-Castells 2013; García-Quevedo et al. 2014; Triguero et al. 2014). Unlike output proxies, measures of innovation input denote the firm's willingness to realize any type of innovation, rather than exclusively success in achieving such innovation. In this context, as previously discussed, the high degree of persistence in R&D is very likely to be induced by specific industry barriers such as start-ups sunk costs (Sutton, 1991; Mañez et al., 2009).

4.2.2 On the persistence of growth

Persistence of firms' differential performance is a widely investigated topic in the industrial economic literature. Perhaps one of the major stylized fact emerging from the vast body of empirical studies is that firms persistently differ, even within the same lines of activities, in their productivity and profitability. Much more controversial is the evidence on persistence of corporate growth.

At the theoretical level, as discussed in Section 4.1, one should expect some structure in the growth process. The conjecture is relatively straightforward: competitive advantages are at the basis of growth and the presence of such advantages relies upon the possession of finest resources, routines, technological and organizational capabilities. All these core competencies create value on the market, are typically unique (or at least better than that possessed by rivals), durable, generate returns which are appropriable, and they are (or should be) inimitable. The uniqueness of such traits leads to heterogeneity of firms (i.e. asymmetries in the production efficiencies), and the different levels of performance are the result of the different value created for consumers. Since competitive advantages are typically durable and competencies difficult to imitate, superior performance tend to persist over long periods of time (Teece et al., 1997; Geroski, 2002).

Notwithstanding this, empirics do not fully reconcile with the expectations. Many studies have analyzed the structure in the process of growth, conventionally by looking at the autocorrelation of growth rates over time (under different autoregressive lags). Even though ideally the time series should be long enough to estimate unbiased parameters, positive values imply that growth process goes beyond the randomic case and displays memory.[2] The evidence is quite mixed and highly dependent on the sample (e.g. country and sector peculiarities), the measure of size, and the statistical methodolgy. Results ranges from the presence of strong autocorrelation in the growth path as in Bottazzi et al. (2001) to the claim that firm growth can be approximated by a random walk process (Geroski, 2002).

On the other hand, scholars agree on the fact that our economy is populated by some persistently growing firms (and a handful of persistently high-growing ones) but the factors underlying such sustained growth still remains largely a black box. Some authors have even suggested that the presence of persistent superior performance, when referring to disproportionately higher profits, might be the simple result of fortune (Barney, 1986). The industrial economic literature is indeed underdeveloped and the few existing studies, mainly focused on high-growth units, aim to sketch merely a demographic profile of persistently growing companies (Coad, 2007a; Coad and Hölzl, 2009; Capasso et al., 2013). Typically, larger and older firms are those who achieve more persistent and smoother growth dynamics, while small firms tend

[2] More recently scholars have started exploiting quantile autoregression techniques to take into consideration the entire distribution of the growth rates. This allows to verify whether some degree of persistence is detected in the full spectrum of growth events, rather than on the "average" firm.

to be characterized by more erratic patterns.

4.2.3 Connecting persistence of innovation to persistence of growth

As already stressed along this dissertation (see Introduction and Chapter 2) we do not know of previous attempts linking other economic or financial factors to persistent growth behaviour. To search for the suitable candidates we need to do a step back and refer more closely to the literature investigating the determinants of firm growth. Among the possible explanations for asymmetries in firm performance, the ability to innovate and the production efficiencies are certainly the core drivers. In particular, if markets operate as efficient selectors, firms featured by good innovative capabilities and/or high production efficiency should grow and gain market shares at the expenses of less innovative and efficient competitors. Likewise, since such innovative capabilities have been proved to be persistent over time, persistently higher capabilities should translate into persistently higher firm performance, profitability and growth in primis (Dosi and Nelson, 2010).

This background helps us to formulate some conjectures on the effects that persistence of innovation may play on persistence of corporate growth. Indeed, as innovation success is likely to be the result of a systematic engagement into innovation activities, we expect firms who are able to persistently translate their innovative efforts into valuable outcomes to overcome their competitors continuously for longer periods. Put differently, to show a higher degree of persistence in their growth paths. Since the innovation process involves uncertainty and risk taking, it is characterize by degrees of complexity, uncertainty and idiosyncrasy, investing for innovation is however a necessary but not sufficient condition to succeed. Thus we expect that persistence in R&D investments does not necessarily spur persistent growth.

4.3 Data and empirical setting

In this Section we present our data, define the main variables and introduce the empirical methodology that we exploit to test whether persistence of innovation spurs persistence of growth.

4.3.1 Source and variables

We draw upon firm-level data from the Survey on Business Strategies (Encuesta Sobre Estrategias Empresariales, henceforth ESEE) which has been conducted yearly since 1990 by the SEPI foundation, on behalf of the Spanish Ministry of Industry. This annual survey gathers extensive information on around 2000 manufacturing companies operating in Spain and employing at least ten workers. The sampling procedure ensures representativeness for each two-digit NACE-CLIO[3] manufacturing sector, following both exhaustive and random sampling criteria. More in detail,

[3]NACE is the usual industrial classification of economic activities within the European Union while CLIO is the nomenclature used by the Spanish input–output tables. The Spanish Accounting Economic System (National Institute of Statistics:http://www.ine.es/) officially recognises both classifications.

in the first year of the survey all Spanish manufacturing firms employing more than 200 workers were required to participate (715 in 1990) and a sample of firms employing between 10 and 200 workers were selected by stratified sampling (stratification across 20 manufacturing sectors and four size intervals) with a random start (1473 firms in 1990). In order to guarantee a high level of representativeness, all newly created companies with more than 200 employees (rate of response around 60%) together with a random sample of firms with fewer than 200 workers and more than 10 (rate of response around 4%) have been incorporated in the survey every year.

In this paper, we refer to data that were obtained between 1990 and 2009, restricting our analysis on a sample of 331 continuing incumbent firms observed for 20 years (6,620 observations). The survival bias that the balancing procedure might possibly introduce is minimal in this case as we will run a comparative analysis across different groups of surviving firms.

We measure firm growth in terms of total sales (henceforth GRS). This is defined as the log-difference:

$$GRS_{it} = s_{it} - s_{i,t-1} \quad , \tag{4.1}$$

where

$$s_{it} = log(S_{it}) - \frac{1}{N}\sum_{i} log(S_{it}) \quad , \tag{4.2}$$

and S_{it} is sales (annual turnover) of firm i in year t, and the sum is computed over the N firms populating the same (2-digit) sector. In this way size and, thus, the growth rates are normalized by their annual sectoral average. This normalization implicitly removes common trends, such as inflation and business cycles effects in sectoral demand.

Innovation performance is measured by using two conventional and widely adopted input and output indicators, namely R&D-to-sales intensity (total R&D expenditure over total turnover, henceforth RDI) and the number of patent applications (PAT).[4] Beside demographic characteristics such as age, size (proxied by number of employees) and industry affiliation (2-digit)[5] we consider and introduce in our analysis a set of control variables which might influence the pattern of growth: a standard labour productivity index computed as the ratio between value added and number of employees (LP), the export intensity computed as the ratio between sales to foreign markets and total sales (EXPI), and a measure of financial leverage computed as percentage which represent the stockholders' equity on total liabilities (EQ/DEBT).

4.3.2 Measuring persistence of growth and innovation

Our measure of persistence of growth performance is a "growth spell" measured as the number of successive years in which a firm shows positive growth rates (as defined above in Equation 4.1). Thus a growth spell is considered as starting in year t if the firm did not growth in $t-1$, and analogously the spell will end in year T

[4] The data available to us allow to distinguish between the number of patent applications filed only in Spain and abroad. In this study we consider the total number of patent applications, as the sum of those filed in Spain and abroad.

[5] We will consider a broad distinction between high-tech (HT) and low-tech (LT) industries according to the Eurostat classification.

when a firm stops growing, after one or more consecutive years of positive growth performance.

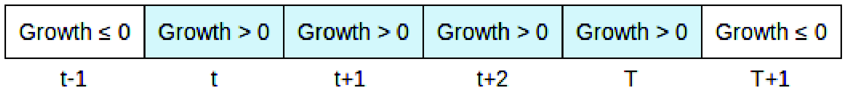

Figure 4.1: A growth spell of four years

As already stressed, we want to uncover what factors are positively related to the probability and the length of growth spells, in other words what factors contribute to support sales growth over time. We pointed out in Section 4.2 that our main concern is to verify if, and to what extent, a systematic engagement in innovation activities, as well as persistent innovation success, is connetected to longer periods of sustainable expansion. Notice that a problem of endogeneity might arise if one considered the "innovation status" precisely during the growth spell, as the former is likely to be pre-determined by the latter and reverse causality issues may be relevant. A very similar problem was noticed in the influencial paper by Geroski et al. (1997), studying how the persistence of innovation varies according to the intensity of patenting. Interestingly, they solved the issue looking for the relationship between the number of innovations produced by a firm just prior to an innovation spell and the length of the spell that follows. We borrow their strategy and we rearrange it in line with our framework.

Therefore we define the innovation status at the beginning of each growth spell on the basis of the frequency with which a firm has shown positive R&D expenditure or applied for patents in a time window of five years prior the growth spell (see Figure 4.2).

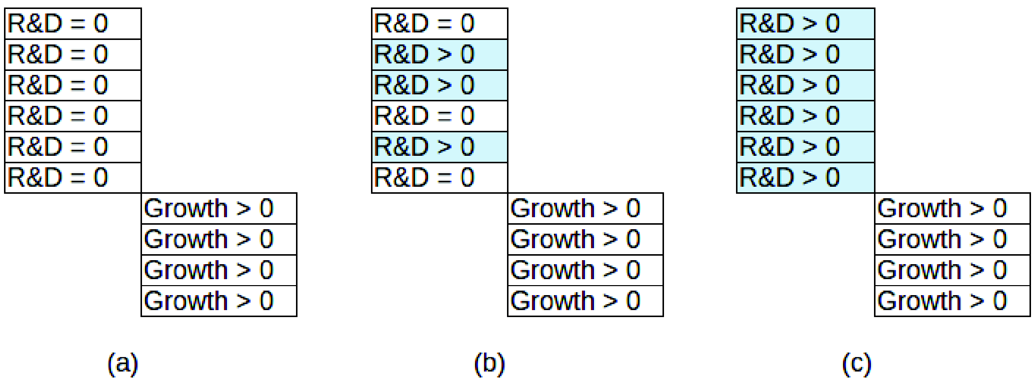

Figure 4.2: Definition of the innovation status based on R&D expenditures (definition based on patent applications is analogous). These figures represent an example of growth spell of four years and the R&D status defined in the time window of five years pre-spell. Panel(a): firms who never invested in R&D and classified as NOI. Panel(b): firms who occasionally invested in R&D and classified as OI. Panel(c): firms who persistently invested in R&D and classified as PI.

We define three mutually exclusive categories, namely: (i) non-innovators (NOI), as those firms that have never invested in R&D, or alternatively never applied for a patent in the time window considered, (ii) occasional innovators (OI), as those firms that exhibit a discontinuos and irregular innovation behaviour, and (iii) persistent innovators (PI), as those firms that have continuously, year after year, spent for R&D or applied for patents.

The choice of building a time window of five years, as well as to build an innovation status collapsing the information into three categories only, balances between the aim at capturing a reasonably long number of years and a reasonable time lag through which innovation may affect growth. Anyhow, we check that our findings are not driven by the specific definitions that we impose to our data (see Section 4.4.3 for details).

4.3.3 Empirical methodology

The empirical setting we choose to model the degree of duration dependence in the growth path is based on survival methods. In a nutshell, we take the spell of time in which a firm grows as unit of analysis, and model the probability that the spell will end at any particular time.

The primary goal of our multivariate analysis is to assess the effect of the innovation status (i.e. non-innovator, occasional innovator, persistent innovator) on the hazard of growth spell ending, while controlling for a set of additional variables which take into account other micro and macro-level features that might also influence a firm's growth pattern. To this end we adopt a discrete time proportional hazard model with gamma mixture distribution to summarize unobserved individual heterogeneity (also said "frailty"), as proposed by Meyer (1990).[6]

The methodology consists of estimating a discrete time representation of the underlying continuous time proportional hazard function, parameterized as follows:

$$\lambda_i(t) = \theta_i \lambda_0(t) \exp\{z_i(t)'\beta\} \tag{4.3}$$

where $\lambda_i(t)$ is the hazard function, $\lambda_0(t)$ is the baseline hazard at time t (an arbitrary non-negative function of t), $z_i(t)$ is a vector of fixed or time-varying explanatory variables (covariates) for firm i, and β is a vector of parameters. Finally, θ_i is a random variable that is assumed to be independent of $z_i(t)$, and represents the frailty which is incorporated multiplicatively.[7]

Following the approach introduced by Prentice and Gloeckler (1978) one can assume that the discrete hazard is given by a complementary log logistic (*cloglog* function), hence obtaining a discrete time counterpart of the underlying countinuous time proportional hazard of equation 4.3, that is:

$$\lambda_i(t) = 1 - \exp\{-\exp(z_i(t)'\beta + \lambda_0(t)) + \theta_i\} \tag{4.4}$$

[6]The choice of implementing a discrete rather than a continuous time approach is motivated by the nature of our data. Indeed, even if growth transitions occured during the year, we record data only on annual basis.

[7]In principle, any continuous distribution with positive support, mean one and finite variance, can be employed to represent the frailty distribution. However, the gamma distribution gives a closed form expression for the likelihood function, avoiding numerical integration and problems of convergence (Meyer, 1990).

To estimate the model of equation 4.4 the unit of analysis must be the growth spell, therefore the dataset must be re-organised so that, for each firm, there are as many data rows as there are time intervals at risk of the event occurring. In practical terms our dependent variable will be a simple binary indicator taking value zero for the survival period (positive growth rate), and one when the growth spell ends (what we previously denoted as period T, see Figure 4.1). We choose the natural logarithm of time ($\ln(t)$) as baseline hazard function to account for the risk of spell ending exclusively attributed to the passage of survival time.[8] It might be the case that some growth spells are in operation before the first year used for the analysis or in the last year of observation, a problem known respectively as left or right censuring. To account for this issue, we carefully include in our model two control variables for left-censored and right-censored spells.

The model is estimated separately to study the effect of persistence in R&D engagement and persistence of patenting. For both exercises we produce several specifications with the aim of assessing the reliability of the estimates: we include step-by-step a set of control variables based on the theoretical background put forward in Section 4.2. The first specification – Model(1) – looks only at the effect of the innovation status, based on R&D or patent applications, on growth spell. The second specification – Model(2) – mimics the first but it allows for dummies to accurately control for censoring. In the third – Model(3) – we include basic demographic features, in particular the age and the size of the firm along with its industry affiliation.[9] The last specification – Model(4) – contains the full set of covariates so that the efficiency, export propensity, ongoing investments in R&D activities and financial status are also taken into account.

4.4 Results

4.4.1 Descriptive evidence

Before moving to the econometric analysis, in this Section, we provide descriptive statistics on our working sample and some preliminary univariate evidence regarding the role of persistence of innovation on persistence of growth performance.

Table 4.1 presents the distribution of growth spell lengths and maximum spell lengths. As expected, the number of spells is much greater than the number of firms, denoting that some firms have more than one spell. However, as can be seen, this dissimilarity is mostly accounted for by spell lengths of 1 years, which represent the 50% of the total numbers of spells. On the other hand, for spell lengths of 7 years or more, the total number of spells is exactly equal to the number of firms that have maximum duration spells of that length. Finally, it is worth nothing that, although we track firms over a period of 20 years, the maximum length of the spells is ten years. This evidence reinforces the idea that persistent growth is a rare phenomenon.

[8]To check whether our findings were influenced by the choice of the baseline hazard, we re-estimated the model of equation 4.4 using a flexible high-order polynomial in survival time. We tested both quadratic and cubic polynomial specifications; results were in line with those presented in this document.

[9]We broadly distinguish between high-tech and low-tech industries according to the Eurostat taxonomy. Indeed the computational burden of the estimation did not allow us to include a full set of industry dummies as the converenge of the likelihood maximization was impossible to achieve.

Table 4.1: Distribution of sales growth spell lengths and maximum spell lengths by firms

Lenght years	All spells		Maximum spells	
	No.	%	No.	%
1	825	50.77	4	1.21
2	458	28.18	83	25.08
3	182	11.20	113	34.14
4	76	4.68	56	16.92
5	53	3.26	47	14.20
6	21	1.29	18	5.44
7	4	0.25	4	1.21
8	3	0.18	3	0.91
9	2	0.12	2	0.60
10	1	0.06	1	0.30
Total	1,625	100.00	331	100.00

Tables 4.2 and 4.3 show basic statistics of the main variables broken down by each innovation status. A first comment on the size of each subsample is in order: as expected, the largest groups are non-innovators, both in terms of R&D and patent applications; the group of occasional innovators based on R&D expenditure is similar in number to the one of persistent innovators, whereas we do observe that only a handful of firms persistently engage in patenting activities.[10] We can see some sharp differences among firms belonging to different innovation statuses: indeed, PI are older and larger than their counterparts, they are more R&D intensive, more efficient, export oriented and predominantly operating in high-tech industries. On the other hand we do not observe any profound difference in terms of growth potential or financial status. To notice that the same characterization is valid for OI with respect to NOI.

As preliminary evidence, we estimate Kaplan-Meier survival functions for the sample of persistent innovator vs. all the other firms (non and occasional innovators) and present them in Figure 4.3. This univariate analysis suggests that persistent innovators enjoy longer growth spells than "sporadic" innovators (indeed the survival function of former is always above that one of the latter). In addition, the difference is particular sizeable when innovation is measured in terms patenting activity.

4.4.2 Main evidence

We turn to the estimation of the discrete time proportional hazard model with gamma frailty and discuss the main findings on the determinants of hazard rates. Table 4.4 and 4.5 report results of four econometric specifications to test, respectively, the effect of persistence of R&D engagement and persistence of patenting on the duration of the growth spell. Since our criterion to define the innovation status leads to three mutually exclusive categories, we select the non-innovative firms

[10] Although this evidence was expected, to check whether our results were influenced by the too small subsample size we replicate the analyses by using a continuous variable (instead of collapsing the information into three dummies) that ranges from 0 to the lenght of the time window pre-spell, depending on the frequency with which a firm undertake R&D activities or apply for patents. We elaborate more in depth this point and provide robustness checks in Section 4.4.3.

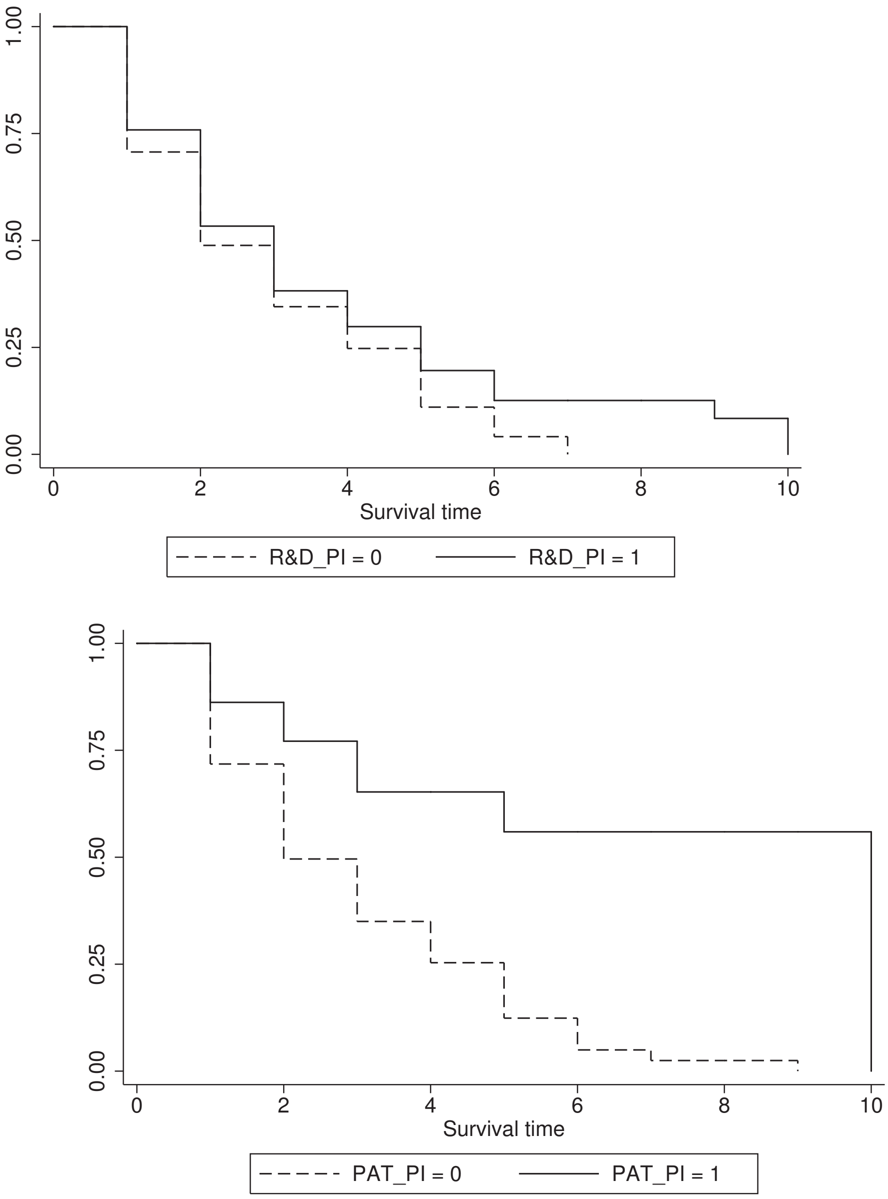

Figure 4.3: Kaplan-Mayer survival function (PI based on 5 years pre-growth spell). Upper panel: R&D measure. Bottom panel: Patent applications measure

Table 4.2: Descriptive statistics broken down by innovation status (R&D)

	Mean	Std.Dev.	Median	Min	Max
NOI (1,241 obs):					
GRS	0.16	0.18	0.12	0.00	1.93
ln(AGE)	3.28	0.45	3.26	2.08	4.71
ln(SIZE)	3.44	1.00	3.22	1.61	7.34
RDI	0.00	0.00	0.00	0.00	0.09
ln(LP)	3.48	0.55	3.46	1.03	5.57
EQ/DEBT	0.05	0.02	0.05	0.00	0.10
EXPI	0.08	0.17	0.00	0.00	1.00
HT	0.12	0.32	0.00	0.00	1.00
OI (533 obs):					
GRS	0.19	0.22	0.13	0.00	1.91
ln(AGE)	3.33	0.55	3.30	1.79	4.71
ln(SIZE)	4.30	1.25	3.91	2.20	8.99
RDI	0.01	0.02	0.00	0.00	0.27
ln(LP)	3.79	0.68	3.69	0.59	6.62
EQ/DEBT	0.05	0.02	0.05	0.00	0.09
EXPI	0.25	0.30	0.10	0.00	1.00
HT	0.14	0.34	0.00	0.00	1.00
PI (588 obs):					
GRS	0.17	0.21	0.11	0.00	2.37
ln(AGE)	3.62	0.54	3.64	2.08	4.76
ln(SIZE)	5.43	1.29	5.51	2.30	8.95
RDI	0.02	0.03	0.01	0.00	0.26
ln(LP)	3.90	0.52	3.86	0.81	6.31
EQ/DEBT	0.05	0.02	0.05	0.00	0.09
EXPI	0.39	0.30	0.35	0.00	1.00
HT	0.31	0.46	0.00	0.00	1.00

(R&D_NOI or PAT_NOI) as the benchmark so that estimated coefficients for the other two groups represent the effect of being occasional or persistent innovator relative to the reference.

Estimated coefficients portray the effect of covariates on the hazard of ending a growth spell. Thus a negative (positive) coefficient is interpreted as a decrease (increase) in the hazard rate, or an increase (decrease) in the expected duration of the growth spell.

From Table 4.4, we can conclude that persistence of R&D investments plays only a very weak role on persistence of sales growth. In the baseline model – Column(1) – we obtain a negative and statistical significant coefficient for R&D_PI, pointing to a degree of association between the persistence of R&D activities and the length of subsequent growth spells. The lack of significance for R&D_OI indicates that firms who occasionally undertake R&D do not differ in terms of persistence of growth from firms that do not invest in R&D at all.[11] The inclusion of censoring dummies in Column(2) slightly changes the magnitudes of the coefficients but do not significantly alter their significance. However, once basic demographic controls are added to the model, the pattern previously identified vanishes and the conclusion we derive is that the duration of the growth spell is fully independent from the innovation status

[11] A simple t-test allows to assess whether the coefficients for occasional and persistent innovators statistically differ. They actually do, with p-value<0.001.

Table 4.3: Descriptive statistics broken down by innovation status (patent applications)

	Mean	Std.Dev.	Median	Min	Max
NOI (1,980 obs):					
GRS	0.17	0.19	0.12	0.00	1.93
ln(AGE)	3.35	0.50	3.33	1.79	4.75
ln(SIZE)	3.98	1.31	3.61	1.61	8.95
RDI	0.01	0.02	0.00	0.00	0.27
ln(LP)	3.62	0.60	3.60	0.59	6.62
EQ/DEBT	0.05	0.02	0.05	0.00	0.10
EXPI	0.18	0.26	0.03	0.00	1.00
HT	0.16	0.36	0.00	0.00	1.00
OI (353 obs):					
GRS	0.17	0.21	0.12	0.00	2.37
ln(AGE)	3.45	0.54	3.47	2.08	4.76
ln(SIZE)	4.81	1.64	4.74	1.61	8.99
RDI	0.01	0.03	0.00	0.00	0.26
ln(LP)	3.77	0.59	3.78	1.78	6.31
EQ/DEBT	0.05	0.02	0.04	0.00	0.09
EXPI	0.32	0.33	0.19	0.00	0.98
HT	0.20	0.40	0.00	0.00	1.00
PI (29 obs):					
GRS	0.09	0.06	0.07	0.00	0.24
ln(AGE)	4.19	0.37	4.28	3.26	4.65
ln(SIZE)	5.94	0.71	5.97	4.34	7.40
RDI	0.06	0.05	0.03	0.00	0.14
ln(LP)	4.24	0.48	4.33	3.24	4.93
EQ/DEBT	0.06	0.01	0.05	0.04	0.08
EXPI	0.22	0.21	0.15	0.02	0.80
HT	0.52	0.51	1.00	0.00	1.00

defined at the commence of the spell. Specification in Column(4) confirms this evidence, but we recover an association (although significant only at 10% level) between size and growth spell duration. The negative coefficient for size suggests that larger firms have a higher tendency to experience longer periods of persistent sales expansion. Interestingly, none of the other potential determinats plays a role so that factors such as higher efficency as well as sounder financial conditions seem not to favour sustained growth patterns.

Results on innovation status proxied by patent applications are instead very stable (see Table 4.5). In line with our conjectures, we find a positive association between persistence of patenting and persistence of sales growth. The baseline specification – Column(1) – indicates that occasional innovators enjoy longer growth spells compared to non-innovators but persistent innovators are those actors who experience the longest period of sustained positive sales growth. The difference in magnitude of the estimated coefficients, so as the statistical signicance, for PAT_OI and PAT_PI is indeed vary sharp. Again, a simple t-test tells us that the coefficients for PI is statistically larger than that for OI (p-value<0.001). The scenario doesn not change when we examine the estimates for the alternative econometric specifications, and a common feature of the regressions shown in Column(3) and (4) is a size effect similar to the one discussed above (here the statistical significance is definitely stronger). Still, all the other explanatory variables do not exert any effect.

All these evidences point to the following interpretation: *while engagement in* *innovation activities may lead to longer spells of growth, it does so only when firms (i) translate their innovative efforts into valuable outcomes and (ii) produce such outcomes persistently, year after year, over time.* Our results suggest that market competition erodes very quickly the competitive advantages that firms build upon their innovation success, at least if one assume that such advantages get reflected into gain of market shares. Therefore, structural competitive advantages come to the cost of being able to persistently overcome competitors in terms of higher innovation differentials. Not coincidentally it emerges a weak association between occasional innovation success and duration of growth spell.

Table 4.4: Estimates of the Proportional Hazard Model - growth spell and R&D persistence

	(1)	(2)	(3)	(4)
R&D_NOI	ref	ref	ref	ref
R&D_OI	0.0345	-0.1116	0.0174	0.0727
	(0.190)	(0.192)	(0.196)	(0.213)
R&D_PI	-0.5860***	-0.4633**	-0.2151	-0.1055
	(0.210)	(0.206)	(0.240)	(0.278)
ln(AGE)			-0.0005	0.0628
			(0.153)	(0.157)
ln(SIZE)			-0.1121	-0.1379*
			(0.071)	(0.082)
RDI				1.3181
				(3.964)
ln(LP)				-0.0951
				(0.146)
EXPI				0.0838
				(0.358)
EQ/DEBT				1.7779
				(3.664)
HT			0.0884	-0.0563
			(0.227)	(0.245)
ln(t)	3.4823***	3.5424***	3.2798***	3.1504***
	(0.319)	(0.337)	(0.319)	(0.340)
Censoring dummies	no	yes	yes	yes
Year dummies	yes	yes	yes	yes
Testing frailty:				
LR test on Gamma var.	1688.92***	1580.03***	1551.75***	1526.84***
Log likelihood	-653.01	-650.31	-660.93	-641.04
Obs	2362	2362	2359	2171

Notes: Coefficient estimates of regression from different specifications of model (4.4), taking non-innovators as the baseline category. The improvement in log-likelihood relative to the no-frailty model is deteced with a standard likelihood-ratio test (LR). Standard errors in parenthesis: ***, ** and * indicate significance at 1%, 5% and 10% level, respectively.

Table 4.5: Estimates of the Proportional Hazard Model - growth spell and patent applications

	(1)	(2)	(3)	(4)
PAT_NOI	ref	ref	ref	ref
PAT_OI	-0.3671*	-0.3627*	-0.3995*	-0.3920*
	(0.208)	(0.206)	(0.218)	(0.221)
PAT_PI	-3.2918***	-3.2484***	-3.0772***	-3.2300***
	(1.037)	(1.028)	(1.077)	(1.232)
ln(AGE)			0.0349	0.0628
			(0.152)	(0.149)
ln(SIZE)			-0.1381**	-0.1600**
			(0.062)	(0.072)
RDI				2.9219
				(3.575)
ln(LP)				-0.1466
				(0.135)
EXPI				0.2693
				(0.327)
EQ/DEBT				1.1568
				(3.405)
HT			0.0019	-0.0380
			(0.218)	(0.227)
ln(t)	3.2227***	3.1869***	3.2423***	2.9341***
	(0.303)	(0.311)	(0.330)	(0.338)
Censoring dummies	no	yes	yes	yes
Year dummies	yes	yes	yes	yes
Testing frailty:				
LR test on Gamma var.	1674.63***	1557.43***	1563.23***	1512.87***
Log likelihood	-656.25	-658.82	-652.48	-645.79
Obs	2362	2362	2359	2171

Notes: Coefficient estimates of regression from different specifications of model (4.4), taking non-innovators as the baseline category. The improvement in log-likelihood relative to the no-frailty model is deteced with a standard likelihood-ratio test (LR). Standard errors in parenthesis: ***, ** and * indicate significance at 1%, 5% and 10% level, respectively.

4.4.3 Robustness checks

In this Section we report a series of additional exercises we carried out to assess whether our findings were invariant to alternative definitions of innovation status.

A first natural extension we propose consists of changing the length of the time window pre-spell along which the three categories (NOI, OI, PI) are defined. In this respect we choose a shorter window of four years and a longer window of six years, and for both cases, the three groups are created according to the original criterion (see Figure 4.2). Thus persistent innovators are still those firms who continuously, year after year, invest in R&D or apply for patents over the new time windows.

A second extension regards the nature of the variable defining the innovation status. Indeed one might argue that collapsing all the information into three dummies might not be the best choice, as for example firms who innovate one out of five years pre-spell of growth have been treated as firms that innovate for four out of five periods pre-spell. Hence we define two new continuous variables (R&D_PERS and PAT_PERS) whose ranges go from 0 to the lenght of the time window pre-spell, depending on the frequency with which a firm undertake R&D activities or apply for patents, respectively.

Tables 4.6 and 4.7 show the estimates for the model with the full set of covariates. The story we deliver is in line with to the one proposed in Section 4.4: we confirm the lack of association between persistence of R&D expenditure and persistence of sales growth – see Column(1) – whereas we still can claim that persistent innovators, when innovation outcome is proxied by patent applications, display the highest propensity to experience longer spells of growth – see Columns(2).[12] As shown in Column(3) and (4) this result holds also when we estimate the model including the continuous variables which alternatively define the innovation status of firms.

A final extension aims to take into account two characteristics of technological knowledge, namely cumulability and non-exhaustibility (David, 1992). Both features could have implications in our framework, as the generation of new knowledge is typically conditioned upon the existing stock that can be used as an input because of its non-exhaustibility. Thus, instead of defining the innovation status of each firm in a finite time window pre-growth spell, we want to trace the innovation behaviour of the firm from the first year of observation up the beginning of the spell, put it differently we want to account for the full innovation history pre-spell. This task can be accomplished by removing the threshold of the window prior the spell where the innovation status was defined, and simply defining two continuous variables (R&D_PERS_FULLH and PAT_PERS_FULLH) that represent the number of years in which a firm has engaged in R&D activities or applied for patents. Results are shown in Table 4.8 and, on the whole, we can still confirm our main findings.

4.5 Conclusions

In this work we have explored the link between the degree of persistence of innovation activities and the persistence of growth performance, using both input and output

[12] Although in the main results shown in Table 4.5 the coefficient for PAT_OI was barely significant, now we loose such significance. In a way we can state that firms who sporadically patent their innovations do not differ in terms of persistence of growth performance from firms who do not patent at all.

Table 4.6: Estimates of the Proportional Hazard Model - innovation status defined on a time window of 4 years pre-growth spell

	(1)	(2)	(3)	(4)
R&D_NOI	ref			
R&D_OI	-0.2478			
	(0.242)			
R&D_PI	-0.0212			
	(0.292)			
PAT_NOI		ref		
PAT_OI		0.0229		
		(0.251)		
PAT_PI		-3.5172**		
		(1.368)		
R&D_PERS			-0.0383	
			(0.069)	
PAT_PERS				-0.3655***
				(0.135)
ln(AGE)	0.1944	-0.0187	0.0055	0.0093
	(0.174)	(0.162)	(0.160)	(0.156)
ln(SIZE)	-0.1574*	-0.1573*	-0.1230	-0.1630**
	(0.089)	(0.081)	(0.086)	(0.078)
RDI	1.6887	3.0875	0.1829	2.6069
	(4.391)	(3.886)	(4.281)	(3.832)
ln(LP)	-0.0171	-0.1892	-0.1102	-0.1038
	(0.161)	(0.153)	(0.153)	(0.149)
EXPI	0.0106	0.1804	0.2604	0.3164
	(0.390)	(0.367)	(0.371)	(0.351)
EQ/DEBT	5.5726	6.6153*	5.4036	4.4876
	(3.959)	(3.860)	(3.821)	(3.682)
HT	-0.2630	-0.0987	-0.0903	-0.1398
	(0.267)	(0.256)	(0.253)	(0.245)
ln(t)	3.6364***	3.5854***	3.4157***	3.3663***
	(0.373)	(0.359)	(0.341)	(0.334)
Censoring dummies	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
Testing frailty:				
LR test on Gamma var.	1683.17***	1685.19***	1664.60***	1658.03***
Log likelihood	-629.08	-626.33	-638.75	-640.18
Obs	2269	2269	2269	2269

Notes: Standard errors in parenthesis: ***, ** and * indicate significance at 1%, 5% and 10% level, respectively. Columns (1) and (2) report estimates of the innovation status (NOI, OI, PI) defined on a time window of 4 years pre-growth spell. Columns (3) and (4) report estimates of the continuous version of the innovation status, still defined on a time window of 4 years pre-growth spell. The improvement in log-likelihood relative to the no-frailty model is deteced with a standard likelihood-ratio test (LR).

Table 4.7: Estimates of the Proportional Hazard Model - innovation status defined on a time window of 6 years pre-growth spell

	(1)	(2)	(3)	(4)
R&D_NOI	ref			
R&D_OI	-0.0407			
	(0.201)			
R&D_PI	-0.0188			
	(0.274)			
PAT_NOI		ref		
PAT_OI		-0.1810		
		(0.210)		
PAT_PI		-2.8346**		
		(1.282)		
R&D_PERS			-0.0050	
			(0.043)	
PAT_PERS				-0.2507***
				(0.092)
ln(AGE)	0.2068	0.2108	0.1883	0.1102
	(0.162)	(0.156)	(0.158)	(0.156)
ln(SIZE)	-0.1763**	-0.1415**	-0.1851**	-0.1181
	(0.080)	(0.070)	(0.077)	(0.072)
RDI	0.8007	2.2462	1.2139	2.5166
	(3.858)	(3.496)	(3.733)	(3.591)
ln(LP)	-0.1634	-0.1002	-0.0582	-0.0519
	(0.138)	(0.133)	(0.137)	(0.139)
EXPI	0.4606	0.3195	0.3420	0.2651
	(0.340)	(0.320)	(0.336)	(0.334)
EQ/DEBT	1.4203	-0.5772	-2.1584	0.5456
	(3.590)	(3.432)	(3.483)	(3.541)
HT	-0.0355	-0.0355	-0.0374	-0.1449
	(0.239)	(0.229)	(0.232)	(0.234)
ln(t)	2.8591***	2.6459***	2.6936***	2.7851***
	(0.330)	(0.347)	(0.322)	(0.352)
Censoring dummies	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
Testing frailty:				
LR test on Gamma var.	1279.08***	1278.99***	1256.61***	1283.46***
Log likelihood	-646.42	-645.61	-658.37	-643.08
Obs	1977	1977	1977	1977

Notes: Standard errors in parenthesis: ***, ** and * indicate significance at 1%, 5% and 10% level, respectively. Columns (1) and (2) report estimates of the innovation status (NOI, OI, PI) defined on a time window of 6 years pre-growth spell. Columns (3) and (4) report estimates of the continuous version of the innovation status, still defined on a time window of 6 years pre-growth spell. The improvement in log-likelihood relative to the no-frailty model is deteced with a standard likelihood-ratio test (LR).

Table 4.8: Estimates of the Proportional Hazard Model - definitions based on full history

	(1)	(2)	(3)	(4)
R&D_PERS_FULLH	-0.0377*	-0.0049		
	(0.022)	(0.025)		
PAT_PERS_FULLH			-0.1253**	-0.1368**
			(0.050)	(0.057)
ln(AGE)	0.0355	0.1466	0.0704	0.1452
	(0.161)	(0.167)	(0.155)	(0.163)
ln(SIZE)	-0.0266	-0.1395	-0.0828	-0.1248
	(0.086)	(0.094)	(0.076)	(0.088)
RDI		-0.4433		-1.2409
		(4.728)		(4.540)
ln(LP)		-0.1965		-0.1019
		(0.167)		(0.166)
EXPI		0.1477		0.2287
		(0.411)		(0.391)
EQ/DEBT		-2.5127		3.8987
		(4.033)		(4.018)
HT	0.2161	0.1306	0.1585	0.2619
	(0.271)	(0.284)	(0.268)	(0.278)
ln(t)	4.6466***	4.3874***	4.5542***	4.1873***
	(0.430)	(0.415)	(0.430)	(0.397)
Censoring dummies	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
Testing frailty:				
LR test on Gamma var.	2447.88***	2349.45***	2442.79***	2361.42***
Log likelihood	-678.96	-667.69	-677.77	-658.71
Obs	3059	2831	3059	2831

Notes: Standard errors in parenthesis: ***, ** and * indicate significance at 1%, 5% and 10% level, respectively.

proxies to characterize the innovation status of firms.

Although there are many studies linking firm growth to micro and macro-level variables, the literature targeting the determinants of persistent growth is only of recent development and the underlying factors shaping and limiting this growth behaviour are still largely unknown. This is, however, a topic that should deserve much more attention and contributions. Indeed, if one rules out that persistent superior performance occur at random (and this might be per se an interesting research question), this kind of growth behaviour may be considered as a tangible proof of competetive advantages at work: firms who possess better operating capabilities and persistenlty higher innovation differentials, defeat competitors and gain market shares repeatedly over time.

Previous contributions provide a mere demographic characterization of persistently growing units, and little consensus exists on the strategies that companies can adopt to maintain desired growth rates. Motivated by the theoretical background we focus on innovation as core driver of growth and, more precisely, we test whether persistence of innovation spurs, to some extent, persistence of growth.

We carried out an empirical investigation on a long and rich panel of Spanish firms comprising micro-level information and frame our analyses considering different indicators of innovation in order to appreciate differentiated patterns. Our results suggest, in a rather robust fashion, that periods of sustained sales expansion are, on average, anticipated by persistent innovation success. While this systematic success allows some firms to build durable competitive advantages, it is not the case for firms who occasionally translate their innovative efforts into innovative outcomes. Moreover, we do find that persistence in R&D investments is not enough to inject structure in the growth process.

Although we believe that our study provides novel insights to the industrial economic literature, it obviously suffers of some limitations that might be object of further investigations. A first extension might be to carefully account for entry and exit phenomena. Although some difficulties may arise in defining when firms do actually fail and exit the market, the growth dynamics of nascent enterprises who are typically the recepee of subsidies for innovation might be of particular concern for policy making.

Secondly, we have to admit that we are fully silent with respect to factors of more direct derivation from management research; we do recognize that such factors are likely to influence the growth behaviour of firms. Among the possible extensions one may foreseee, we think that it would particularly relevant to look deeper into organizational traits, or exploring the role of differences in the underlying firm strategies and managerial characteristics. Are some growth strategies more effective than others? Or, are some organizational and governance structures more helpful to convert persistent innovation success into sustained growth than others? Answering these questions not only would complement our preliminary evidence but would also increase the overall understanding of some intricate firm-industry dynamics.

French summary

La croissance de l'entreprise est un sujet qui a pendant longtemps intrigué les économistes. Ces derniers l'ont toujours considéré comme une preuve tangible de la procédure de sélection de marché. Le but étant de comprendre si les marchés réussissent à récompenser et sanctionner les entreprises en terme de leur taille relative ou leur part de marché selon le principe de l'efficacité différentielle (Bartelsman et al., 2005; Lotti et al., 2009; Bottazzi et al., 2010; Dosi et al., 2013). Les responsables politiques s'intéressent aux entreprises en raison de leur potentiel de croissance, qui se traduit par la création de nouveaux emplois et favorise ainsi la croissance macroéconomique (voir par exemple la discussion récente Haltiwanger et al., 2013. Enfin, les chercheurs de la tradition de la gestion stratégique, les gestionnaires et les consultants tentent de déterminer les meilleures pratiques qui engendrent la performance optimale de l'entreprise. Ce qui leur permet par la suite de l'appliquer à leur propre entreprise, ou à celle de leurs clients (Teece et al., 1997; Katkalo et al., 2010). En outre, parmi les nombreuses entreprises présentes dans nos économies un petit groupe se détache grâce à sa performance de croissance exceptionnelle. Ces entreprises communément appelées entreprises à forte croissance ou gazelles, ont été récemment la cible de recherches universitaires ainsi que d'initiatives politiques.

Cet intérêt particulier pour la croissance des entreprises, et plus récemment, des entreprises à forte croissance, se reflète dans l'abondance des études théoriques et empiriques écrites ces dernières décennies. Sur le plan théorique, bien que les écoles de pensées divergent quant aux hypothèses-sous-jacentes, le processus de croissance (et éventuellement de forte croissance) est considéré comme le résultat de l'interaction entre les trois dimensions principales de l'entreprise: à savoir la productivité (ou l'efficacité), la rentabilité et la situation financière (voir, Nelson and Winter, 1982; Jovanovic, 1982; Ericson and Pakes, 1995). Généralement, ces modèles prédisent qu'un choc idiosyncrasique conduit à une efficacité hétérogène entre les entreprises. Cette différence s'explique par l'influence du choc sur des facteurs non-observables spécifiques à l'entreprise (tels que l'aspect technologique et organisationnel, les capacités de l'entreprise ainsi que ses pratiques stratégiques et de gestion). Ainsi, les entreprises avec une efficacité relative plus élevée acquièrent des parts de marché aux dépens d'entreprises moins efficaces. Ces asymétries en termes de rentabilité et de situation financière permettent aux entreprises les plus productives d'accéder plus facilement aux ressources nécessaires à leur développement.

Bien que le contexte théorique ait incité les études empiriques, il n'existe pas de relation directe entre ces dimensions (efficacité, profitabilité et situation financière) et la croissance de l'entreprise. En effet, derrière le modèle linéaire simplifié décrit par les études théoriques résident la forte complexité et l'idiosyncrasie, bien connus pour caractériser le processus de croissance de l'entreprise (Geroski, 2002; Delmar et al., 2003).

Ainsi, alors que certains faits stylisés assez robustes émergent des études dynamiques industrielles (pensez à la relation négative entre l'âge et la taille avec la croissance des entreprises par exemple), nous sommes toujours confrontés à de nombreuses preuves énigmatiques concernant la croissance de l'entreprise. Par exemple, contre toute attente, certains chercheurs constatent que les entreprises les plus efficaces et plus productives ont une faible inclinaison comportementale à la croissance, ce qui est également le cas pour les entreprises ayant de meilleures conditions financières (Bottazzi et al., 2002, 2008; Coad, 2007b).

La preuve ambiguë pourrait être due au fait que d'autres facteurs, indépendants de la simple productivité, la rentabilité et des variables financières, jouent un rôle important dans le processus de croissance des entreprises. Sans grande surprise, il a été prouvé que certains facteurs externes tels que les caractéristiques macroéconomiques et les caractéristiques spécifiques aux industries (par exemple les cycles économiques, les caractéristiques régionales, le cycle de vie de l'industrie) affectent la croissance des entreprises, ainsi que d'autres composants de niveau de l'entreprise tel que l'innovation parmi les meilleures entreprises (voir Coad, 2009 pour une étude approfondie). Par rapport à ce dernier élément, des études récentes démontrent que le processus entrée/sortie de l'innovation peut revêtir des effets différents selon le positionnement d'une entreprise dans la distribution des taux de croissance. Ainsi, les entreprises à forte croissance bénéficient de plus d'avantages provenant des activités innovatrices (Coad and Rao, 2008; Holzl, 2009).

Le scénario se complique lorsque le concept de persistance est introduit dans les modèles de croissance. En effet, conformément aux études théoriques, nous devrions observer un certain degré de persistance dans le processus de croissance des entreprises. Les «bonnes» entreprises ont tendance à se développer d'abord rapidement (une dynamique de type "la réussite engendre la réussite"), puis rencontrent par la suite un ralentissement progressif. Les économistes et responsables politiques sont particulièrement préoccupés par les performances de croissance persistante. Ces dernières doivent être en principe associées à la présence de capacités exceptionnelles au sein de l'entreprise ou bien d'avantages structurels extérieurs (Teece, 2007). Autrement dit, la persistance dans la croissance des entreprises va à l'encontre la notion de hasard qui stipule que seule la chance peut expliquer la surperformance de certaines entreprises (Barney, 1997).

Mais une fois de plus les études empiriques ne s'accordent pas aux théories selon lesquelles la « réussite engendre la réussite ». En réalité ces dynamiques sont très difficiles à déceler dans les données statistiques. Les études empiriques ne permettent pas de trouver un degré robuste de persistance dans le processus de croissance. Les études supposent davantage une croissance d'auto corrélation ou une baisse d'auto corrélation (Bottazzi et al., 2001). Cela implique une nature plus erratique et imprévisible de la croissance (Coad and Holzl, 2009; Bottazzi et al., 2011; Holzl, 2014). Des études récentes (Delmar et al., 2003; Capasso et al., 2013; Daunfeldt and Halvarsson, 2013) portent sur la persistance de modèles à forte croissance. Ces dernières l'existence d'un groupe d'entreprises surperformantes persistantes en simplement différenciant leur profil démographique (c'est-à-dire leur taille, leur maturité et leur affiliation).

Les preuves incompatibles sur la nature persistante de la croissance sont complétées par un accord sur la nature persistante de l'innovation. Sur le plan théorique, la persistance de l'innovation est expliquée par plusieurs phénomènes. En liant pouvoir de marché et innovation, Schumpeter met en avant l'idée selon laquelle les entreprises en situation de monopole ont d'avantage intérêt à innover par rapport aux nouveaux entrants potentiels. Puis, les études récentes ont reconnu l'importance de l'asymétrie d'information entre l'innovateur et le prêteur, de sorte que les entreprises innovantes puisent s'appuyer sur les résultats non distribués plutôt que des fonds externes (Bhattacharya et Ritter, 1983). A cet égard, le succès de l'innovation passé offre des bénéfices qui peuvent être utilisés pour financer des activités innovatrices actuelles, induisant ainsi la persistance des comportements d'innovation. Enfin, d'autres chercheurs ont élaboré des interprétations liées au processus d'apprentissage par l'innovation (Dosi et Marengo, 1994). Sur le plan empirique, la plupart des études récentes (parmi lesquelles Cefis, 2003; Peters, 2009; Raymond et al, 2010; García-Quevedo et al, 2014) prouvent l'existence d'un fort degré de persistance dans les activités innovatrices, en particulier pour les entreprises qui déposent un grand nombre de brevets ainsi que les entreprises opérant dans les secteurs de haute technologie.

Toutefois, les études analysant la relation entre la persistance de l'innovation et la performance des entreprises sont encore très rares. Expliciter le lien entre les motifs de l'innovation avec ce que l'on sait sur la croissance des entreprises, tant au empirique et sur le plan théorique, est un défi aussi important qu'urgent pour la recherche future (Dosi, 2007).

Cette thèse vise à combler certaines lacunes de la littérature existante. Les trois études traitent, en général, du processus de croissance des entreprises, sa persistance, ainsi que du rôle de l'innovation dans la performance des entreprises.

Dans la première étude, nous nous concentrons sur la persistance de la forte croissance des entreprises et examinons si ce modèle de croissance spécifique est associé à de meilleures capacités d'exploitation. La littérature existante, comme déjà souligné, est principalement axée sur l'identification des causes et des conditions qui ont conduit une entreprise à surpasser ses concurrents durant une période de temps relativement courte. Notre travail vise à mettre en avant une nouvelle perspective. Nous tentons de déterminer si des facteurs plus structurels, économiques ou financiers, peuvent être représentatifs des entreprises présentant des performances en forte croissance à plusieurs reprises au fil du temps (au-delà de simples caractéristiques démographiques tels que la taille, l'âge et distinctif affiliation de l'industrie).

Nos hypothèses reposent sur des modèles d'évolution entreprise-industrie. Ainsi, la concurrence du marché devrait favoriser les entreprises les plus rentables et efficaces ; de même que de bonnes conditions financières devraient permettre d'accéder aux ressources extérieures nécessaires pour financer l'investissement et la croissance. Par conséquent, nous devrions nous attendre à ce que les entreprises à forte croissance soient plus productives et plus rentables que les entreprises illustrant une croissance "moins anormale". Il est intéressant de déterminer si ces mêmes caractéristiques sont associées avec une forte croissance persistante. Par ailleurs, nous tentons de comprendre si les entreprises à forte croissance persistante diffèrent, en fonction de ces caractéristiques, des autres entreprises et, en particulier, des entreprises qui ont des poussées d'accélération sur de courtes périodes.

Nous réalisons une enquête sur un panel d'entreprises italiennes, françaises, espagnoles et britanniques déjà en exercice. Nous tentons d'identifier celles qui relèvent d'une forte croissance et celles qui relèvent une forte croissance persistante. Nous analysons, à la fois dans un cadre non-paramétrique et paramétrique, comment les premières années de la productivité, la rentabilité et les facteurs financiers sont liés à la dynamique de croissance ultérieures. Dans l'analyse non-paramétrique nous examinons si un ensemble de variables clés, pris pour représenter la performance opérationnelle et la situation financière, montre les différences de répartition à travers et des entreprises à forte croissance, à forte croissance persistante et des autres. Tous ces exercices sont conduits de plusieurs façons afin de tester le rôle de différents facteurs institutionnels ou autre indicateurs macroéconomique sur les particularités sectorielles et les tendances d'innovation.

Conformément aux études théoriques, nous confirmons que les déterminants économiques, et la productivité en particulier, sont significativement associés à une forte croissance. Cependant, nous ne trouvons pas de preuve systématique de toute différence statistiquement significative entre les entreprises à forte croissance et celles à forte croissance persistantes en termes d'efficacité d'exploitation. Aucune des dimensions considérées ne semble donc fonctionner dans le maintien du statut à forte croissance dans le temps. Malgré le fait qu'il existe un grand nombre d'autres facteurs pouvant expliquer une croissance soutenue dans le temps, nos résultats sont particuliers. Il semblerait que la forte croissance persistante soit le fruit du hasard, telle que le résultat d'une ≪ simple chance ≫.

La deuxième étude vise à explorer la relation entre la croissance et l'innovation, en tenant compte de la nature multidimensionnelle du processus d'innovation. Notre contribution fait référence au travail d'Audretsch et al. (2014): ≪ la complexité des activités de R&D, couplée avec la diversité des stratégies d'innovation et la multiplicité des modes de croissance, nécessite une approche multidimensionnelle pour examiner la contribution des innovations sur la croissance des entreprises ≫.

Nous nous appuyons sur un ensemble de données d'informations de type ECI sur les entreprises espagnoles. L'étude est basée sur la période 2004-2011, qui est très particulière compte tenu de sa nature longitudinale. Nous souhaitons contribuer à cette étude par de multiples façons. Dans un premier temps, nous dressons un portrait de la relation entre la croissance et l'innovation. Pour cela, nous observons un large ensemble de variables d'innovation qui saisissent des sources différentes, des modes et types d'activités innovantes mises en œuvre au sein des entreprises.

Plus précisément, l'ensemble des indicateurs de l'innovation comprend : la R&D interne vs externe, l'innovation de procédé, les types d'innovation de produit ainsi que l'acquisition incorporée et désincorporé. Dans un second temps, nous tenons compte des interactions entre les modes d'innovation mentionnés ci-dessus (ce que nous appelons "stratégie d'innovation"), et augmentons le niveau de complexité afin de déterminer si la croissance est susceptible d'être entraînée par une combinaison spécifique des activités d'innovation. Ce faisant, nous vérifions si les complémentarités entre les différents modes d'innovation sont en jeu.

Nous commençons par enquêter séparément sur la relation entre la croissance des ventes et chaque activité innovatrice. Les études actuelles exploitant des données similaires à l'ECI qui analysent l'interaction entre la croissance et l'innovation ont toutes une limitation commune : l'écart de 3-4 ans entre la réalisation de chaque enquête. Par conséquent, cela limite les paramétrages économétriques à une seule coupe transversale, ils échouent donc à leur tour dans le contrôle précis de l'hétérogénéité non-observée. Pourtant, l'ampleur de la complexité, l'incertitude et l'idiosyncrasie qui caractérisent le processus d'innovation, ainsi que les asymétries dans l'efficacité de production, indiquent qu'un contrôle systématique de l'hétérogénéité non observée n'est plus qu'une simple formalité. Ainsi, nous exploitons la dimension longitudinale de nos données, afin de déterminer la corrélation de chaque variable de l'innovation avec la croissance des ventes. Nous appliquons à la fois un cadre de régression classique et une technique actualisée de régression quantile conçue pour tenir compte des effets fixes au niveau de l'entreprise (Canay, 2011).

Nos résultats préliminaires soulignent une contribution hétérogène des différentes activités d'innovation à l'expansion des ventes. En effet, parmi les indicateurs de l'innovation que nous avons pris en compte, la R&D (en particulier celle réalisée en interne) et le changement technique de l'entreprise représentent les principales sources d'avantage concurrentiel des entreprises, et plus particulièrement pour certaines en pleine croissance. Alors que l'innovation de processus ne semble pas influencer directement la croissance des ventes, il en ressort que les entreprises à forte croissance ont un plus grand avantage quand elles introduisent un produit innovant sur le marché.

Nous tenons compte des interactions entre les différents modes d'innovation créant ainsi ce que nous appelons des «stratégies d'innovation ». Nous notons par ailleurs que la relation entre la croissance et l'innovation peut être différente pour les entreprises actives simultanément, par rapport aux entreprises qui n'innovent qu'une ou deux fois. Un niveau plus élevé de complications implique certainement davantage de coûts et des problèmes de coordination, mais d'un autre côté il peut induire et créer des opportunités de croissance. Nous effectuons donc un test direct de complémentarité entre différents modèles d'innovation (voir, par exemple, Mohnen et Roller, 2005; Catozzella et Vivarelli, 2014). Ce test souligne certains effets complémentaires : l'innovation des produits est susceptible d'avoir un effet plus important sur la croissance si les activités internes de R&D sont menées simultanément. Nous trouvons également des preuves de complémentarité entre le processus et aussi l'innovation de produit.

Dans la troisième étude, nous analysons l'effet de l'innovation persistante sur la croissance des ventes et sa persistance. Comme nous l'avons déjà souligné, la politique ainsi que la théorie économique trouvent un intérêt particulier dans l'étude minutieuse de la nature de cette relation. L'innovation en général, et la R&D en particulier, est considéré comme un facteur clé de la performance des entreprises. C'est pourquoi l'innovation persistante peut être vue à la fois comme un moteur important de l'avantage concurrentiel durable et comme un changement durable dans la performance des entreprises (Teece, 2007). En effet, conformément à la théorie, un marché compétitif favorise les entreprises rentables et efficaces, deux dimensions qui sont à leur tour fortement influencées par le succès de l'innovation. Ce succès constitue un avantage concurrentiel pour l'entreprise, de sorte que l'on s'attende à ce que les innovateurs persistants se démultiplient. En outre, la persistance dans l'innovation est susceptible de subir des avantages concurrentiels ainsi que de soutenir une position de leadership sur le marché au fil du temps. Ainsi, le modèle de croissance des innovateurs continus devrait être plus régulier (ou persistant).

Nous testons ces hypothèses via un riche ensemble de données longitudinales des entreprises espagnoles (Encuesta Sobre Estrategias Empresariales-ESEE). Ces données couvrent une période de vingt ans et comprennent les informations spécifiques aux entreprises ainsi qu'aux caractéristiques du marché.

Notre stratégie empirique est basée sur la représentation de l'état de l'innovation de chaque entreprise, que nous tirons en regardant la fréquence d'innovation. Cette classification mène à trois groupes d'entreprises distincts: les entreprises non-innovatrices, les innovatrices occasionnelles et les innovatrices persistantes. Nous menons une analyse plus approfondie afin de déterminer si les différences de répartition à travers différents groupes d'entreprises peuvent s'expliquer par un ensemble de variables clés. Ensuite, nous utilisons les techniques de modélisation de duration pour étudier les déterminants de la durée d'une période continue et positive de croissance des ventes. Puis nous associons ces résultats au comportement innovant persistant.

Nos résultats montrent que les innovateurs persistants diffèrent des innovateurs occasionnels (et des non-innovateurs également) en termes de caractéristiques démographiques, étant généralement plus grandes, plus matures, et actives dans les secteurs de haute technologie. En outre, conformément à nos hypothèses, nous prouvons que la persistance dans l'innovation affecte la croissance, de même qu'un investissement systématique dans les activités en matière de brevets est bénéfique à l'expansion des ventes. Par ailleurs, nous établissons également des preuves selon lesquelles la persistance dans l'innovation conduit à une plus grande structure dans le processus de croissance. Autrement dit, les innovateurs persistants connaissent de plus grandes périodes de croissance basées sur une performance continuelle et croissante.

Bibliography

Acs, Z. J. and P. Mueller (2008): "Employment effects of business dynamics: Mice, gazelles and elephants," *Small Business Economics*, 30, 85–100.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005): "Competition and innovation: An inverted-U relationship," *The Quarterly Journal of Economics*, 120, 701–728.

Aghion, P. and P. Howitt (1992): "A Model of growth through creative destruction," *Econometrica*, 60, 323–51.

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

Arellano, M. and S. Bond (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The Review of Economic Studies*, 58, 277–297.

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

Asplund, M. and V. Nocke (2006): "Firm turnover in imperfectly competitive markets," *Review of Economic Studies*, 73, 295–327.

Audretsch, D. B., A. Coad, and A. Segarra (2014): "Firm growth and innovation," *Small Business Economics*, 43, 743–749.

Barney, J. B. (1986): "Strategic factor markets: Expectations, luck, and business strategy," *Management science*, 32, 1231–1241.

——— (1997): *On flipping coins and making technology choices: Luck as an explanation of technological foresight and oversight*, Cambridge University Press: Cambridge, UK.

Bartelsman, E., S. Scarpetta, and F. Schivardi (2005): "Comparative analysis of firm demographics and survival: evidence from micro-level sources in OECD countries," *Industrial and Corporate Change*, 14, 365–391.

Blundell, R. and S. Bond (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of econometrics*, 87, 115–143.

Bottazzi, G., E. Cefis, and G. Dosi (2002): "Corporate growth and industrial structure. Some evidence from the Italian manufacturing industry," *Industrial and Corporate Change*, 11, 705–723.

Bottazzi, G., E. Cefis, G. Dosi, and A. Secchi (2007): "Invariances and diversities in the evolution of Italian manufacturing industry," *Small Business Economics*, 29, 137–159.

Bottazzi, G., A. Coad, N. Jacoby, and A. Secchi (2011): "Corporate growth and industrial dynamics: Evidence from French manufacturing," *Applied Economics*, 43, 103–116.

Bottazzi, G., G. Dosi, N. Jacoby, A. Secchi, and F. Tamagni (2010): "Corporate performances and market selection. Some comparative evidence," *Industrial and Corporate Change*, 19, 1953–1996.

Bottazzi, G., G. Dosi, M. Lippi, F. Pammolli, and M. Riccaboni (2001): "Innovation and corporate growth in the evolution of the drug industry," *International Journal of Industrial Organization*, 19, 1161–1187.

Bottazzi, G. and A. Secchi (2003): "Common properties and sectoral specificities in the dynamics of US manufacturing companies," *Review of Industrial Organization*, 23, 217–232.

——— (2006): "Explaining the distribution of firms growth rates," *The RAND Journal of Economics*, 37, 235–256.

Bottazzi, G., A. Secchi, and F. Tamagni (2008): "Productivity, profitability and financial performance," *Industrial and Corporate Change*, 17, 711–751.

Canay, I. A. (2011): "A simple approach to quantile regression for panel data," *The Econometrics Journal*, 14, 368–386.

Capasso, M., E. Cefis, and K. Frenken (2013): "On the existence of persistently outperforming firms," *Industrial and Corporate Change*, forthcoming.

Cassiman, B. and R. Veugelers (2006): "In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition," *Management science*, 52, 68–82.

Catozzella, A. and M. Vivarelli (2014): "The catalysing role of in-house R&D in fostering complementarity among innovative inputs," *Industry and Innovation*, 21, 179–196.

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

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

Coad, A. (2007a): "A closer look at serial growth rate correlation," *Review of Industrial Organization*, 31, 69–82.

——— (2007b): "Testing the principle of growth of the fitter: the relationship between profits and firm growth," *Structural Change and economic dynamics*, 18, 370–386.

——— (2009): *The growth of firms: A survey of theories and empirical evidence*, Edward Elgar, Cheltenham, UK.

Coad, A. and W. Hölzl (2009): "On the autocorrelation of growth rates," *Journal of Industry, Competition and Trade*, 9, 139–166.

Coad, A. and R. Rao (2008): "Innovation and firm growth in high-tech sectors: A quantile regression approach," *Research Policy*, 37, 633–648.

Colombelli, A., N. Haned, and C. Le Bas (2013): "On firm growth and innovation: Some new empirical perspectives using French CIS (1992–2004)," *Structural Change and Economic Dynamics*, 26, 14–26.

Colombelli, A., J. Krafft, and F. Quatraro (2014): "High-growth firms and technological knowledge: do gazelles follow exploration or exploitation strategies?" *Industrial and Corporate Change*, 23, 261–291.

Conte, A. and M. Vivarelli (2014): "Succeeding in innovation: key insights on the role of R&D and technological acquisition drawn from company data," *Empirical Economics*, 47, 1317–1340.

Cooley, T. F. and V. Quadrini (2001): "Financial markets and firm dynamics," *American Economic Review*, 91, 1286–1310.

Corsino, M. and R. Gabriele (2011): "Product innovation and firm growth: evidence from the integrated circuit industry," *Industrial and Corporate Change*, 20, 29–56.

Crepon, B., E. Duguet, and J. Mairessec (1998): "Research, innovation And productivity: An econometric analysis at the firm level," *Economics of Innovation and New Technology*, 7, 115–158.

Cucculelli, M. and B. Ermini (2012): "New product introduction and product tenure: What effects on firm growth?" *Research Policy*, 41, 808–821.

Daunfeldt, S.-O. and D. Halvarsson (2013): "Are high-growth firms one-hit wonders? Evidence from Sweden," *Small Business Economics*, 1–23.

David, P. A. (1992): "Knowledge property and the system dynamics of technological change," *World Bank Economic Review*, 215–248.

Davidsson, P. and M. Henrekson (2002): "Determinants of the prevalance of start-ups and high-growth firms," *Small Business Economics*, 19, 81–104.

Delmar, F. (2006): "Measuring growth: methodological considerations and empirical results," in *Entrepreneurship and the Growth of Firms*, ed. by P. Davidsson, F. Delmar, and J. Wiklund, University of Michigan Press, Ann Arbor.

Delmar, F., P. Davidsson, and W. B. Gartner (2003): "Arriving at the high-growth firm," *Journal of business venturing*, 18, 189–216.

Demirel, P. and M. Mazzucato (2012): "Innovation and firm growth: Is R&D worth it?" *Industry and Innovation*, 19, 45–62.

Dosi, G. (1988): "Sources, procedures, and microeconomic effects of innovation," *Journal of economic literature*, 1120–1171.

——— (2007): "Statistical regularities in the evolution of industries. A guide through some evidence and challenges for the theory," in *Perspectives on Innovation*, ed. by F. Malerba and S. Brusoni, Cambridge University Press.

Dosi, G., M. Faillo, and L. Marengo (2008): "Organizational capabilities, patterns of knowledge accumulation and governance structures in business firms: an introduction," *Organization Studies*, 29, 1165–1185.

Dosi, G., O. Marsili, L. Orsenigo, and R. Salvatore (1995): "Learning, market selection and evolution of industrial structures," *Small Business Economics*, 7, 411–36.

Dosi, G., D. Moschella, E. Pugliese, and F. Tamagni (2013): "Productivity, market selection and corporate growth: comparative evidence across US and Europe," *LEM Working Paper Series*.

Dosi, G. and R. R. Nelson (2010): "Technical change and industrial dynamics as evolutionary processes," *Handbook of the Economics of Innovation*, 1, 51–127.

Duguet, E. and S. Monjon (2004): "Is innovation persistent at the firm level? An econometric examination comparing the propensity score and regression methods," *An Econometric Examination Comparing the Propensity Score and Regression Methods (July 2004). University of Paris I-Cahiers de la MSE Working Paper*, 75.

Ericson, R. and A. Pakes (1995): "Markov-perfect industry dynamics: A framework for empirical work," *Review of Economic Studies*, 62, 53–82.

Falk, M. (2012): "Quantile estimates of the impact of R&D intensity on firm performance," *Small Business Economics*, 39, 19–37.

Flaig, G. and M. Stadler (1994): "Success breeds success. The dynamics of the innovation process," *Empirical Economics*, 19, 55–68.

Fligner, M. A. and G. E. Policello (1981): "Robust rank procedures for the Behrens-Fisher problem," *Journal of the American Statistical Association*, 76, 141–206.

Freel, M. S. (2000): "Do small innovating firms outperform non-innovators?" *Small Business Economics*, 14, 195–210.

García-Quevedo, J., G. Pellegrino, and M. Vivarelli (2014): "R&D drivers and age: Are young firms different?" *Research Policy*, 43, 1544–1556.

Geroski, P. and S. Machin (1992): "Do innovating firms outperform non-innovators?" *Business Strategy Review*, 3, 79–90.

Geroski, P. and M. Mazzucato (2002): "Learning and the sources of corporate growth," *Industrial and Corporate Change*, 11, 623–644.

Geroski, P. A. (2002): "The growth of firms in theory and in practice," in *Competence, Governance, and Entrepreneurship - Advances in Economic Strategy Research*, Oxford University Press: Oxford and New York.

Geroski, P. A., J. Van Reenen, and C. F. Walters (1997): "How persistently do firms innovate?" *Research Policy*, 26, 33–48.

Goddard, J., W. J., and B. P. (2002): "Panel tests of Gibrat's law for Japanese manufacturing," *International Journal of Industrial Organization*, 20, 415–433.

Goddard, J., P. Molyneux, and J. Wilson (2004): "Dynamics of growth and profitability in banking," *Journal of Money, Credit, and Banking*, 36, 1069–1090.

Goedhuys, M. and R. Veugelers (2012): "Innovation strategies, process and product innovations and growth: Firm-level evidence from Brazil," *Structural Change and Economic Dynamics*, 23, 516–529.

Griffith, R., E. Huergo, J. Mairesse, and B. Peters (2006): "Innovation and Productivity across Four European Countries," *Oxford Review of Economic Policy*, 22, 483–498.

Hall, B. H. (1999): "Innovation and market value," NBER Working Papers 6984, National Bureau of Economic Research, Inc.

Hall, B. H., F. Lotti, and J. Mairesse (2008): "Employment, innovation, and productivity: evidence from Italian microdata," *Industrial and Corporate Change*, 17, 813–839.

——— (2009): "Innovation and productivity in SMEs: Empirical evidence for Italy," *Small Business Economics*, 33, 13–33.

Haltiwanger, J., R. S. Jarmin, and J. Miranda (2013): "Who creates jobs? Small versus large versus young," *Review of Economics and Statistics*, 95, 347–361.

Harding, M. and C. Lamarche (2009): "A quantile regression approach for estimating panel data models using instrumental variables," *Economics Letters*, 104, 133–135.

——— (2014): "Estimating and testing a quantile regression model with interactive effects," *Journal of Econometrics*, 178, 101–113.

Hay, M. and K. Kamshad (1994): "Small firm growth: Intentions, implementation and impediments," *Business Strategy Review*, 5, 49–68.

Hölzl, W. (2009): "Is the R&D behaviour of fast-growing SMEs different? Evidence from CIS III data for 16 countries," *Small Business Economics*, 33, 59–75.

——— (2014): "Persistence, survival, and growth: a closer look at 20 years of fast-growing firms in Austria," *Industrial and Corporate Change*, 23, 199–231.

HOPENHAYN, H. A. (1992): "Entry, exit and firm dynamics in long run equilibrium," *Econometrica*, 60, 1127–1150.

JOVANOVIC, B. (1982): "Selection and the evolution of industry," *Econometrica*, 50, 649–70.

KATKALO, V. S., C. N. PITELIS, AND D. J. TEECE (2010): "Introduction: On the nature and scope of dynamic capabilities," *Industrial and Corporate Change*, dtq026.

KODDE, D. A. AND F. C. PALM (1986): "Wald criteria for jointly testing equality and inequality restrictions," *Econometrica: journal of the Econometric Society*, 1243–1248.

KOENKER, R. (2004): "Quantile regression for longitudinal data," *Journal of Multivariate Analysis*, 91, 74–89.

KOENKER, R. AND J. BASSETT, GILBERT (1978): "Regression quantiles," *Econometrica*, 46, 33–50.

KÖNIG, H., F. LAISNEY, M. LECHNER, AND W. POHLMEIER (1994): "On the dynamics of process innovative activity," in *The explanatory power of business cycle surveys*, ed. by G. P. K. Oppenlander, Avebury, 243–262.

LOTTI, F., E. SANTARELLI, AND M. VIVARELLI (2009): "Defending Gibrat's law as a long-run regularity," *Small Business Economics*, 32, 31–44.

LUTTMER, E. G. J. (2007): "Selection, growth and the size distribution of firms," *The Quarterly Journal of Economics*, 122, 1103–1144.

MALERBA, F. AND L. ORSENIGO (1999): "Technological entry, exit and survival: An empirical analysis of patent data," *Research Policy*, 28, 643–660.

MAÑEZ, J. A., M. E. ROCHINA-BARRACHINA, A. SANCHIS, AND J. A. SANCHIS (2009): "The role of sunk costs in the decision to invest in R&D," *The Journal of Industrial Economics*, 57, 712–735.

MANSFIELD, E. (1962): "Entry, Gibrat's Law, innovation, and the growth of firms," *The American Economic Review*, 52, 1023–1051.

MELITZ, M. J. (2003): "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 71, 1695–1725.

METCALFE, S. J. (1998): *Evolutionary economics and creative destruction*, London, UK: Routledge.

MEYER, B. D. (1990): "Unemployment insurance and unemployment spells," *Econometrica*, 58, 757–782.

MOHNEN, P. AND L.-H. ROLLER (2005): "Complementarities in innovation policy," *European Economic Review*, 49, 1431–1450.

MOWERY, D. C. (1983): "Economic theory and government technology policy," *Policy Sciences*, 16, 27–43.

NELSON, R. R. AND S. G. WINTER (1982): *An evolutionary theory of economic change*, The Belknap Press of Harvard University Press: Cambridge, MA.

PARISI, M. L., F. SCHIANTARELLI, AND A. SEMBENELLI (2006): "Productivity, innovation and R&D: Micro evidence for Italy," *European Economic Review*, 50, 2037–2061.

PARKER, S. C., D. J. STOREY, AND A. VAN WITTELOOSTUIJN (2010): "What happens to gazelles? The importance of dynamic management strategy," *Small Business Economics*, 35, 203–226.

PAVITT, K. (1984): "Sectoral pattern of technical change: Towards a taxonomy and a theory," *Research Policy*, 13, 343–373.

PELLEGRINO, G., M. PIVA, AND M. VIVARELLI (2012): "Young firms and innovation: A microeconometric analysis," *Structural Change and Economic Dynamics*, 23, 329–340.

PENROSE, E. T. (1959): *The theory of the growth of the firm*, Oxford university press.

PETERS, B. (2009): "Persistence of innovation: stylised facts and panel data evidence," *The Journal of Technology Transfer*, 34, 226–243.

PRENTICE, R. L. AND L. A. GLOECKLER (1978): "Regression analysis of grouped survival data with application to breast cancer data," *Biometrics*, 57–67.

RAYMOND, W., P. MOHNEN, F. PALM, AND S. S. VAN DER LOEFF (2010): "Persistence of innovation in Dutch manufacturing: Is it spurious?" *The Review of Economics and Statistics*, 92, 495–504.

SANTAMARIA, L., M. J. NIETO, AND A. BARGE-GIL (2009): "Beyond formal R&D: Taking advantage of other sources of innovation in low- and medium-technology industries," *Research Policy*, 38, 507–517.

SCHERER, F. M. (1965): "Corporate inventive output, profits, and growth," *Journal of Political Economy*, 73, 290–297.

SCHIMKE, A. AND K. MITUSCH (2011): "Gazelles-High-growth companies," EUROPE INNOVA - Sectoral Innovation Watch Report Final report, European Commission.

SCHREYER, P. (2000): "High-growth firms and employment," Science, Technology and Industry Working Papers 2000/03, OECD Publishing.

SEGARRA, A. AND M. TERUEL (2014): "High-growth firms and innovation: an empirical analysis for Spanish firms," *Small Business Economics*, 43, 805–821.

SILVERBERG, G., G. DOSI, AND L. ORSENIGO (1988): "Innovation, diversity and diffusion: A Self-organisation Model," *Economic Journal*, 98, 1032–1054.

STAM, E. AND K. WENNBERG (2009): "The roles of R&D in new firm growth," *Small Business Economics*, 33, 77–89.

Stangler, D. (2010): “High-growth firms and the future of the American economy,” *Available at SSRN 1568246*.

Stanley, M. H. R., L. A. N. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger, and H. E. Stanley (1996): “Scaling behaviour in the growth of companies,” *Nature*, 379, 804–806.

Stiglitz, J. E. (1987): “Learning to learn, localized learning and technological progress,” *Economic policy and technological performance*, 125–153.

Storey, D. J. (1994): *Understanding the small business sector*, London: Cengage Learning EMEA.

Sutton, J. (1991): *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration*, MIT press.

Teece, D. J. (2007): “Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance,” *Strategic management journal*, 28, 1319–1350.

Teece, D. J., G. Pisano, and A. Shuen (1997): “Dynamic capabilities and strategic management,” *Strategic Management Journal*, 18, 509–533.

Topkis, D. M. (1998): *Supermodularity and complementarity*, Princeton university press.

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

Vennet, R. V. (2001): “The law of proportionate effect and OECD bank sectors,” *Applied Economics*, 33, 539–546.

Vivarelli, M. (2014): “Innovation, employment and skills in advanced and developing countries: A survey of economic literature,” *Journal of Economic Issues*, 48, 123–154.

Wagner, J. (1992): “Firm size, firm growth, and persistence of chance: Testing Gibrat’s law with establishment data from Lower Saxony, 1978-1989,” *Small Business Economics*, 4, 125–131.

Weiss, C. R. (1998): “Size, growth, and survival in the upper Austrian farm sector,” *Small Business Economics*, 10, 305–312.

Wooldridge, J. M. (2009): “On estimating firm-level production functions using proxy variables to control for unobservables,” *Economics Letters*, 104, 112–114.

Stefano BIANCHINI

Three essays on firm growth, innovation, and persistent performance

Résumé

Les trois études traitent du processus de croissance des entreprises, sa persistance, ainsi que du rôle de l'innovation dans la performance des entreprises. Dans la première étude, nous nous concentrons sur la persistance de la forte croissance des entreprises et examinons si ce modèle de croissance spécifique est associé à de meilleures capacités d'exploitation. La deuxième étude vise à explorer la relation entre la croissance et l'innovation, en tenant compte de la nature multidimensionnelle du processus d'innovation. Nous observons un large ensemble de variables d'innovation qui saisissent des sources différentes, des modes et types d'activités innovantes mises en œuvre au sein des entreprises. Dans le troisième essai, nous examinons le rôle de la persistance de l'innovation sur la persistance des performances en terme de croissance.

Mots-clés: Croissance des entreprises; Innovation; Persistance

Résumé en anglais

The three essays focus on the process of firm growth, its persistence, and on the role of innovation in affecting firm performance. In the first essay we concentrate on persistence of high-growth and investigate whether this peculiar growth pattern is associated with better operating capabilities. The second essay aims to explore the relationship between growth and innovation, taking into account the multidimensional nature of the innovation process. We provide a broad picture of the relationship between growth and innovation, by looking at a wide set of innovation variables that capture the different sources, modes and types of innovative activity undertaken within firms. In the third essay we examine the role of persistence of innovation on persistence of growth performance, assessing whether a systematic, rather than sporadic, engagement in innovation activities induces more structure in the process of firm growth.

Mots-clés: Firm growth ; Innovation ; Persistence