High growth firms, innovation and competition: the case of the US pharmaceutical industry

Mariana Mazzucato and Stuart Parris

University of Sussex

Guidelines for authors

Papers shall be submitted in pdf or Word format. They should contain a title, an abstract, and keywords. Papers should be submitted to one of the Editors, who will process them and send them to the appropriate Associate Editor. Two members of SPRU will be asked to provide a short written review within three weeks. The revised versions of the paper, together with a reply to the reviewers, should be sent to the Associate Editor, who will propose to the Editors its publication on the series. When submitting the authors should indicate if the paper has already undergone peer-reviewing, in which case the Associate Editors may decide to skip internal review process.

Website

SWPS: www.sussex.ac.uk/spru/research/sewps

IDEAS: ideas.repec.org/s/sru/ssewps.html

High growth firms, innovation and competition:

the case of the US pharmaceutical industry

MARIANA MAZZUCATO
University of Sussex,
Brighton, BN1 9RH, United Kingdom.
Tel: 01273 877943.
Email: m.mazzucato@sussex.ac.uk

STUART PARRIS
The Open University,
Walton Hall, Milton Keynes, MK7 6AA,
United Kingdom.

ABSTRACT

Innovation is key to economic growth. But firms, across sectors and regions, are highly skewed in their ability to engage with innovation, and even more skewed in their ability to translate investments in innovation into higher growth. While there was initially much attention on 'small' firms (SMEs), due to the assumption that they are more entrepreneurial and innovative, recent evidence that small firms contribute less to innovation and employment than commonly believed, has caused attention to move towards 'high growth innovative' firms (HGF). There is however the risk that this newly emphasized category of firms is also being 'hyped up' given how short a time period 'high growth' lasts, and how 'high growth' appears to only be important when combined with other firm specific conditions. Our paper is dedicated to exploring under what conditions high growth firms matter, in a dynamic setting over the history of the US pharmaceutical industry from 1963-2002. Following Coad and Rao (2008), we use quantile regression techniques to study the R&D-growth relationship in high growth firms compared to low growth firms. We find that the relationship is influenced by a mix of firm level characteristics including R&D intensity, R&D scale and venture capital funding. But more importantly we find that this relationship is sensitive to the changing competitive environment over the industry's history.

Key words: R&D, growth, venture capital, quantile regression, pharmaceutical industry.

JEL Classifications: L1 *Market Structure, Firm Strategy, and Market Performance,* ***O3*** *Technological Change*

1. Introduction

Innovation has long been recognized a key driver of economic growth, at both the company level and at the macro-economic level (Freeman 1995; Grossman and Helpman 1991; Verspagen, 2005). This has caused economists and policy makers to spend much time thinking about what kind of firms will be the most 'innovative' and how to support them, both directly and indirectly (EC, 2010; OECD, 2010). Indeed, across sectors and regions, firms are highly skewed in their ability to engage with innovation, and even more skewed in their ability to translate investments in innovation into higher growth. Thus understanding what types of firms are most important for generating both innovation and economic growth has become a focus of both innovation policy and growth policy. Critically however, this debate has developed without consideration to *how* changes in the external environment influence the performance of firms. Our contribution is to investigate when different factors can be used to understand high growth, specifically by investigating the effect of competition on the relationship between innovation and growth. In so doing we ask 'when' and 'why' differences between firms matter for high growth performance (Nelson, 1991).

For many years the emphasis has been on the role that small firms play in the growth process, due to the assumption that they are more innovative and 'entrepreneurial' (Porter 1985; Acs and Yeung 1999). This emphasis builds on Schumpeter's early work (1934) where he focused on the role of small entrepreneurial firms. Indeed many national and transnational policies have focused on filling the 'funding gap' faced by small firms[1], precisely due to the assumption that filling this gap will lead to more innovation and employment over time (from the UK's Macmillan Report in 1931 to the European Commission's Horizon 2020 program).

Yet recently the emphasis on small firms has been questioned given the tiny percentage of small firms that contribute to new jobs, sales and innovation. Shane (2008) has shown that in the US only 0.05% of new firms contributed 11% of sales and 6% of jobs. Kirchhoff (1994) found that 10% of fastest-growing firms contributed to three quarters of new jobs during an eight-year observation period within a cohort of firms founded in the US in 1978. In the UK, Storey (1994) shows that only 4% of new firms born in any given year accounted for 50% of all the jobs created by surviving firms within that cohort after ten years. Likewise, Bloom and Van Reenan (2006) have focused on low productivity, and low innovation rates of most small companies who suffer from managerial difficulties and are less prone to making the investments needed to increase productivity.

Given this core result that only a small group of firms generate innovation, productivity increases and jobs, different studies have been increasingly emphasizing not firm size but 'high growth' firms (Bartelsman et al, 2005; Veugelers and Cincera, 2010; Mason and Brown, 2013). While potential high growth firms are often 'young' firms (Haltiwanger et al, 2013), in fact high growth firms are found of all sizes (NESTA, 2009). Shane (2009) has argued that policy should focus on creating high quality enterprises, those with potential for high growth, rather than subsidizing a high volume of small firms with a low likelihood of success, whether success is measured in terms of innovation, economic growth, or jobs. Using quantile regression techniques that differentiate high growth firms (HGF) from low growth firms (LGF), Coad and

---

[1] The 'funding gap' happens when demand for financial resources from small-firms is not met by private agents, e.g. commercial banks.

Rao (2008) find further evidence for the importance of high growth firms in one of the most studied relationships, the relationship between spending on innovation inputs (e.g. R&D) and growth. They find that R&D and patenting, when measured as a composite variable, was only a significant predictor of growth for HGF; for other firms investing in R&D and innovation appeared to be only a cost. In the face of such evidence, policy makers are increasingly trying to target high growth innovative companies (EC 2020; OECD 2012).

Given this recent attention to 'high growth' firms, the paper focuses on better understanding what we actually know about them and whether there is the danger that they too (like SMEs in the past) are becoming a fad of policy makers, with little evidence to justify the attention. Indeed, as Delmar et al. (2003) note, given that high growth firms are a 'temporally unstable population' (p.195), it is fundamental to understand the degree to which the HGF category is capturing something structural, that can indeed be targeted by policy makers, versus a transient state, dependent on idiosyncratic features, such as the simple fact that beginning from a lower position will enable higher growth.

We focus on this problem by looking at high growth and low growth firms in one particular sector—the pharmaceutical industry— and ask whether a key relationship which industrial economists worry about, the R&D-growth relationship, can tell us anything about the relevance of high growth firms. Like Coad and Rao (2008) we use quantile regression techniques to study the R&D-growth relationship for high growth and low growth firms, but by focusing on only one industry over a long time frame (1963-2002), we are able to introduce a further dimension to the study: changing levels of competition.

Indeed, innovation studies have found that classic questions concerning the relationship between variables like size and innovation, and industry structure, depend on the phase of the industry's history. For example, it has been found that small firms tend to have greater advantage in periods in which there are lower scale economies and more market share instability (Klepper, 1997). Thus abstracting from dynamic changes in an industry's history implies abstracting from competitive forces. In particular, we are interested in whether changing market structure, and the degree of inter-firm competition, affects the degree to which 'high growth' firms matter. Given the well-known inverted U-shape relationship between innovation and market structure (Scherer, 1987; Aghion et al. 2005), and the understanding we have of how R&D intensity changes over an industry's life-cycle (Gort and Klepper 1982), the question is whether these relationships also affect the importance of high growth firms, and what insights this can provide to the recent interest that such firms have received. And given that firms differ so much in their growth rates, it also helps us to answer the fundamental question posed by Nelson (1991): *when and why do firm differences matter*—a subject we return to in the final conclusion.

The pharmaceutical industry is a particularly interesting industry to look at for several reasons. First, it is an innovation intensive industry, with higher than average R&D spending. It is also an industry in which the ability of R&D to generate growth has been questioned, given the low productivity of this spending when looked at in relation to new medicines coming out of the sector. Second, it is an industry in which the characteristics of innovation have evolved significantly over the course of its history. The Bayh-Dole act in 1980 (which allowed publicly funded research to be patented) increased the number of entrants in the industry, and the search regime itself went through big changes (Demirel and Mazzucato, 2012). The latter has

been described as having gone from a period of 'random search' to one of more 'guided search', due to the advances in IT and genomics (Gambardella, 1995). These changes were accompanied by changes in the level of concentration, but also in the level of market share instability in the industry. While we do not focus on causal issues (e.g. why a lower level of concentration should impact the R&D-growth relationship), the changing competitive landscape (number of firms, concentration, instability) is enough to allow us to see whether changing industry structure affects the degree to which 'high growth firms' differ (in their R&D-growth relationship) from their low growth counterparts.

The paper is structured as follows. Section 2 reviews the recent literature on three key areas related to the study. Section 3 introduces the data used in the study, why the pharmaceutical industry was chosen, and presents some descriptive statistics and figures which raise some initial questions about the usefulness of the *high growth firm* category. Section 4 discusses the characteristics of the competitive environment in the pharma industry that are used to determine two different competition regimes. Section 5 discusses quantile regression techniques used to study the relationship between R&D and performance. Section 6 reports the quantile estimates and Section 7 concludes with discussion of our key result: high growth are indeed the best performing (in terms of the R&D-growth relationship) but this only holds in the period in which competition is fiercest, and variety is the highest.

2. Innovation, Firm Size, Growth, and Competition
As our study looks at the R&D-growth relationship, for high/low growth firms, over the course of industry history, we begin by reviewing three strands of literature that relate to this objective. The first is the literature which focuses on the need for innovation policy to target SMEs and/or 'high growth' firms. The second is the literature that has highlighted the problematic aspects of the R&D-growth relationship. And the third is the literature that has focused on the changing relationship between size, growth and innovation in different competitive environments and how those environments change over the industry life-cycle.

2.1. R&D, Innovation and Growth
We focus on the R&D-growth relationship due to the importance of R&D in national policies aimed at achieving 'smart' growth (European Commission 2010), and due to the great heterogeneity in the ability of firms to achieve growth from their R&D spending (Demirel and Mazzucato 2012). In fact, it has been observed that returns to R&D are generally at a 45 year low (Lang, 2009) with R&D investment in fast growing sectors growth outpacing the resultant firm growth (Ejermo et al. 2011).

While endogenous growth theory (Grossman and Helpman 1991) assumes a direct relationship between R&D and growth at the macroeconomic level, at the microeconomic level the relationship has been harder to find (Verspagen, 2005). While some studies find a *positive* relationship between firm spending on R&D and firm growth (Del Monte and Papagni 2003), of course others have found either *no significant impact* (Almus and Nerlinger 1999) or sometimes even a *negative* impact with regards to employment growth (Brouwer et al. 1993). This creates confusion for policy makers that have been arguing for target levels of R&D spending at the national level (e.g. see EC 2020, European Commission 2010).

The core issue of course is that outcomes of R&D spending vary dramatically between firms. Even among firms with a track record of innovative behavior there are clear differences in the

nature of internal and external processes, relationships, capabilities and general innovative persistence (Demirel and Mazzucato 2012). One variable that has been found to be particularly illuminating is the degree to which firms differ in their 'persistence' of innovation. Geroski et al. (1997) have shown that persistence of innovation is quite rare; firms struggle to be innovative for sustained periods, yet those that can achieve a level of persistence are often responsible for the majority of innovations. Cefis and Orsenigo (2001) find that persistence is more common in Schumpeter type 1 sectors that have higher barriers to entry (as opposed to Schumpeter 2 sectors with lower barriers, and more instability). Demirel and Mazzucato (2012) find that the R&D growth relationship in pharmaceuticals (a Schumpeter 1 industry, science based with high barriers) is indeed strongest for firms that have patented persistently.

Studies have also found that access to venture capital is important in determining the degree to which R&D spending by firms leads to better economic performance (Gompers and Lerner 2001). This is not surprising given that the high uncertainty associated with innovation means that access to risk capital is fundamental. But how do VC investors decide which firms to provide risk capital to? Recent literature has pointed to their attention to patenting and a lookout for high growth 'gazelle' type characteristics (Baum and Silverman 2004). It follows that high growth firms are expected to have better access to venture capital (Gompers and Lerner 2001). VC investors are also expected to have expertise in the identification of firms with disruptive potential, as they also profit from the success of funded firms, and are able to offer expertise to help firm succeed and manage the difficulties of high growth development (Baum and Silverman 2004). High growth innovative firms should potentially grow faster with venture capital support, as it reduces financial and managerial constraints. For these reasons, below we use a VC related variable as a control.

Our work builds most closely on the literature which acknowledges that the impact of R&D and innovation on growth is indeed different for different types of firms, i.e. slow growing/fast growing; above-average R&D intensity/below-average R&D intensity, low tech/high tech (Coad and Rao 2008; Brouwer et al. 1993, and Del Monte and Papagni 2003). To help contribute to the debate regarding the relationship between high growth, SME and R&D, we also focus on the importance of R&D scale, in addition to R&D intensity.

2.2. Innovation, size and high growth firms
Our study contributes to the understanding of high growth firms by looking at a range of firm specific and industry specific conditions that must be in place for the high growth category to be meaningful. We interpret its 'meaningfulness' with respect to a specific relationship that is very important to policy makers: the relationship between R&D and growth discussed above.

Innovation policy has long targeted small firms, due to the assumption that smaller firms are more willing to challenge the status quo, and more 'entrepreneurial' and innovative (Porter 1985; Acs and Yeung 1999). This assumption has its roots in Schumpeter's early work (Schumpeter, 1934) in which he argued that it was the entrepreneur who was the key force of capitalism. Of course the SME category is not homogenous, including hi-tech and low-tech firms (Nunes et al. 2012); firms with different attitudes to risk taking (Qian and Lee 2003), and firms that are not particularly innovative, yet often require higher profit margins. Indeed, the Neo-Schumpeterian literature has taken a more subtle approach, asking 'under what conditions' are large or small firms more innovative. Acs and Audretsch (1990) found that small firms are most innovative in industries that are less capital intensive and in industries where innovation is

mainly product not process oriented. Others still have focused on the degree to which innovation small/large firms are better at competence destroying/enhancing innovations (Christensen, 1997). Regardless of these studies which qualify the importance of small firms, based on firm-specific, technology specific and industry specific variables, many countries continue to indistinctively target small firms in the hope that this will stimulate innovation and growth. Indeed, Hughes finds that the UK government spends £8 billion/year on small firms, more than on the police force, and close to what it spends on universities (Hughes, 2008).

More recently, different studies have argued that the focus on small firms is misplaced given their low productivity (Bloom and Van Reenan, 2006), low innovation rates (NESTA, 2009), and low job creation (Storey, 1994). Kirchhoff (1994) finds that only the 10% of fastest-growing firms contributed to three quarters of new jobs during an eight-year observation period within a cohort of firms started in the US in 1978. Haltiwanger et al. (2013) find that there is indeed no systematic relationship between firm size and growth. Most of the effect is from age: young firms (and business start-ups) contribute substantially to both gross and net job creation.

These critiques have caused the attention to move from small firms, to young firms and 'high growth' firms. Some have focused on the role of 'gazelles' understood to be young innovative firms (Veugelers and Cincera, 2010). Yet Acs et al. (2008) find that it is older firms, on average 25 years old, that demonstrate both high employment and high sales growth. Indeed, more recent focus has been on a category of firms called 'high growth' firms. This is because there is increasing evidence that is a small percentage of firms that achieve high growth, but also that these firms are likely to be mature and relatively large (Audretsch, 2012). In the UK, NESTA, a think tank, has published a series of influential reports, arguing that between 2002 and 2008, most of the employment generation in the UK was due to high growth firms across a wide variety of sectors (high and low tech) (NESTA 2009) and that in the post-crisis period high growth businesses accounted for a disproportionate amount of job growth (NESTA 2011). These reports have basis in the literature, and relate to studies including Coad and Rao (2008) and Hölzl (2009) which find that often only the fastest growing firm reap any benefit from their innovation (proxied by an indicator that combines both R&D and patents) in terms of sales or employment growth. Relatedly, Colombelli et al. (2012) find evidence that HGF are key actors in the creation of new knowledge, indicating some support for the relationship between innovation and HGF.

However, some recent work casts doubt on the relationship between innovation and HGFs. For example, Henrekson and Johansson (2010) find that technology firms are under-represented in the population of HGF; they find HGF in all industries, but particularly in service based industries. Other studies have found innovative high growth firms lack significant levels of investment in R&D (Mason and Brown, 2013), or find evidence of a negative relationship between innovation and growth (Freel and Robson, 2004). However, Stam and Wennberg (2009) in their study of new firms offer some support that R&D does matter to firm growth, including HGFs, but only applies to a select group of new and hi-tech firms. They find no relationship between R&D and growth for the majority of new firms in their study.

Some of the complexities of understanding growth are expected to lie in both the use and measurement of the term growth. Delmar et al. (2003) recognize 'high growth' as a

heterogeneous phenomenon, multidimensional in nature, such that different measures of high growth are not necessarily related. They propose that studies using single measures of growth must be careful to ensure their measure reflects the theoretical perspectives applied, and propose that other more complex multidimensional or composite measures of growth may capture the complexities of the growth phenomenon.

A feature common to studies of R&D and growth is a failure to account for changes in industry market structure over time. If policy is to successfully stimulate the innovation-growth relationship, we argue policy must be sensitive to the wider dynamics of industry, and specifically the competitive environment. This is the focus of the final section of our literature review. Evaluating how the determinants of growth vary with industry structure during different time periods helps us answer an important question: does variety always matter (Nelson, 1991)?

2.3 Competition and Innovation over an industry's history
A corollary of Schumpeter's concern with the relationship between size and innovation was his concern for the relationship between industry structure and concentration. In situations in which small firms are more innovative, industry structure will tend to be less concentrated. Studies which have tested this proposition have found that there is an inverted U shape relationship between innovation and concentration (Scherer and Ross, 1990): modest degrees of concentration are the ones which foster the most spending on R&D. The reason is that when there is too much competition, there is low ability to appropriate innovation hence little incentive to spend on it. And when there is too much concentration firms become more inertial and can also protect their market power without feeling the pressure to innovate (Aghion et al. 2005). The latter effect is strongly related to the role of patents in allowing firms to appropriate the returns from new knowledge (Arrow, 1962).

A further important area of study in this vein is how industry structure, and its relationship with innovation, changes over time, as an industry matures. The 'evolutionary' strand of the Schumpeterian literature has indeed focused much on the role of different types of 'selection' pressures over time, which winnow in on the variety that is created in the process of inter-firm competition (Nelson and Winter, 1982). Industry life-cycle studies, within the 'evolutionary' tradition, have shown that the relationship between R&D spending and firm growth changes over time (Klepper, 1997). In the early stage, when R&D spending tends to focus more on 'product' innovation, entry rates are high as is also market share instability. Only once a certain product 'standard' is established, does R&D tend to focus more on process innovation, which is also more susceptible to dynamic increasing returns to scale that favor large firms. This process of dynamic increasing returns tends to lead to more concentrated markets and less market share instability. Mazzucato (2002) has shown how in reality this evolution is more complex, with some industries following this pattern in a classic form (e.g. autos), while others experience more instability in the later part of their life-cycle (e.g. personal computers) depending on the 'power' of new entrants to challenge the status quo. This was the case for the PC industry when the Intel-Platform replaced the IBM platform in 1990, allowing those firms that had already entered to finally innovate outside the IBM-compatible boundaries.

In our study we focus on the dynamics, rather than on causal relationships. That is, we do not ask what the effect of concentration is on the difference between the performance of high growth and low growth firms; we simply ask whether it is different. We of course do

hypothesize about this in the discussion, but it is not something we test for. We note the major changes that occurred in the industry (e.g. the Bayh-Dole act which allowed new firms to enter), and 'measure' competition using different indicators, and see whether the importance of high growth firms differs in different competitive environments as measured by those indicators— regardless of the reasons which we cannot study causally in this particular study. We choose both classic measures of competition, e.g. industry concentration, as well as more dynamic indicators, e.g. market share instability (Hymer and Pashigian, 1962). Indeed this latter proxy we believe is especially important because even with a constant level of concentration, increasing competitive pressures will force firms to differentiate themselves, even if this does not affect the number of firms it may affect their market shares. And it is market share that is the key concern of competing firms.

3. Data and descriptive analysis

We focus on one industry so to be able to study the effect of R&D on growth, as innovation dynamics and competition unfold. In particular we select pharma as an opportunity to study the long term dynamics of an industry where R&D and innovation have historically played a critical role in determining the success of firms, as a result of translating pioneering R&D in drug discovery and first mover advantages into significant financial reward (Grabowski and Vernon 1987; Dosi and Mazzucato 2006). However, since the late 1990s the productivity of R&D has fallen (Paul et al. 2010), and firms have struggled to produce blockbuster products, resulting in a rising cost for each new molecular entity (NME) approved (Scannell et al. 2012). The evolution of the pharmaceutical industry has received significant scholarly attention, documenting the changing nature of competition as a result of new innovation, and regulatory changes (Gambardella 1995; Orsenigo et al. 2001), and so makes an ideal industry for our analysis that seeks understand the relationship between R&D and growth under different competitive regimes. This allows us to when high growth matters more. We begin by looking at high growth firms using descriptive statistics which later help us to interpret the regression results with respect to the literature identified in Section 2.

3.1 Data

Our dataset includes observations on pharmaceutical firms (GICS 352020) tracked by COMPUSTAT related to the financial years 1950-2007. The COMPUSTAT database covers the entry and exit of pharma firms trading on North American stock exchanges since 1950, covering a total of 370 firms. From this database we collect information on annual investment in R&D, sales, assets and acquisition values. Firms with observations relating to a broken time series are cleaned, so that only the longest spell is considered, resulting in observations on 303 firms. All financial data is deflated to 1982-84 level.

Our dataset includes the major pharmaceutical players with worldwide sales, plus US and major non-US firms. COMPUSTAT's listing of pharma firms recorded total sales of $492bn in 2006, accounting for approximately 77% of world pharmaceutical sales activity[2]. COMPUSTAT data provides a good representation of the global pharmaceutical industry, including firms with headquarters in US and Europe as well in more recently firms in developing countries such as India and China.

---

[2] Based on estimates of global pharmaceutical sales from IMS (2007)

Our data reveals a striking feature of the pharmaceutical industry, in contrast to its popular image as formed entirely of large MNC, which is a significant presence of SMEs. The composition of SME (defined as less than 500 employees) in the sample increases overtime, from zero in 1963, to 37% in 1980, rising to a maximum of 63% in 1996. The presence of such number of SMEs in the latter stages of the pharmaceutical industry lifecycle is both a distinctive feature and relevant given our focus on high growth literature as described in Section 2.

Merger and acquisition activity is also a feature of the pharmaceutical industry, and COMPUSTAT records firm annual expenditure on acquisition activity. However COMPUSTAT does not report accounting information to determine the long term performance of acquired firms or subsidiaries, so it is not possible to report statistics on long term organic vs. acquired growth.

In cases when a firm recorded in the COMPUSTAT database is acquired, the historical annual observations of the target firm are recorded, but target firm activity post-acquisition is recorded with that of the parent firm. In a minority of cases where it is deemed that at least two firms have merged to create a new entity, such as in the case of Hoechst Marion Roussel and Rhone Poulenc Rorer to create Aventis, the historical records of each merging firm are maintained in COMPUSTAT to the point of merger, and a new company entry records post-merger activity.

From extensive searches of Who Owns Whom, LexisNexis International, Hoover's Company Profiles, Bloomberg Business Week profiles, and Edgar SEC filings, it was possible to identify a total of 83 firms exiting the dataset as a result of M&A activity during the period 1950 to 2008 although all occurred post 1980. Of the total acquisition targets recorded in our sample, 61 could be traced to a parent firm also listed in our sample data. As a result we can record firm-years where reported sales, and sales growth and other firm measures, will likely to have been directly influenced by acquisition activity of competing pharma. This does not account for all M&A activity, as pharma in recent decades has relied increasingly on acquisition of small firms to support their pipeline which may have long term performance implications. However, it does include major acquisitions likely to have a transforming short-term influence on big pharma accounts[3].

Using the list of COMPUSTAT pharma firms as a base population we merge the NBER patent database (Hall et al. 2001) and Thomson Private Equity database. The patent data from the NBER project and subsequent updates for 2002 and 2006 were used to provide information on patent counts for each COMPUSTAT listed firm, allocating patents according to ownership at the point of application (Hall et al. 2001). The coverage of the NBER dataset is from 1962 until 2008, thus we do not have reliable patent data for the period 1950 to 1962. Secondly as we count firms with granted patent status, because of the lag between patent application and grant, applications made in periods from 2003 onwards show a lower grant rate in the database, as when the dataset was compiled there were still being reviewed. Therefore we take a conservative approach and use patents applied for between 1962 and 2002.

The names of pharma firms in COMPUSTAT were cross referenced with the Thomson Private Equity database recording the amount and year of investment for those firms that received

---

[3] We make use of both the COMPUSTAT reported acquisition data, and our own analysis of M&A activity to perform robustness checks on our model estimates reported in Section 6

private equity and/or venture capital funding. Searches included the historical names of firms listed in COMPUSTAT to capture their full investment history. As the vast majority of firms receiving investment from private equity funds had previously received venture capital, for the remainder of the paper we refer to these firms as venture capital backed.

3.2 Preliminary descriptive analysis of high growth

One reason why the high growth category might be problematic is that a high growth event might be a very short period in a firm's history, and hence not a good proxy for those features that policy can target. Likewise, from a policy perspective it is unclear how targeting high growth per se is helpful. Instead we suggest it's more useful to understand the underlying characteristics of high growth firms. For instance if high growth arises because very small firms are encouraged to start up at an inefficient size, and so must grow rapidly in order to survive, under these circumstances high growth would only represent a positive outcome if the firm sustained high growth over multiple periods (Nightingale and Coad 2011). In fact our data (Figure 1) indicates that of the 130 firms achieving at least one year of high growth (logged differenced sales growth rate above the 90th percentile), over half (81) achieve this performance for two years or less. At most a firm achieves growth rates above the 90th percentile for a total of 6 years; a rare occurrence.

Fig. 1 Duration of top firm growth performance showing number of years top performing firm spend above 90th percentile sales growth
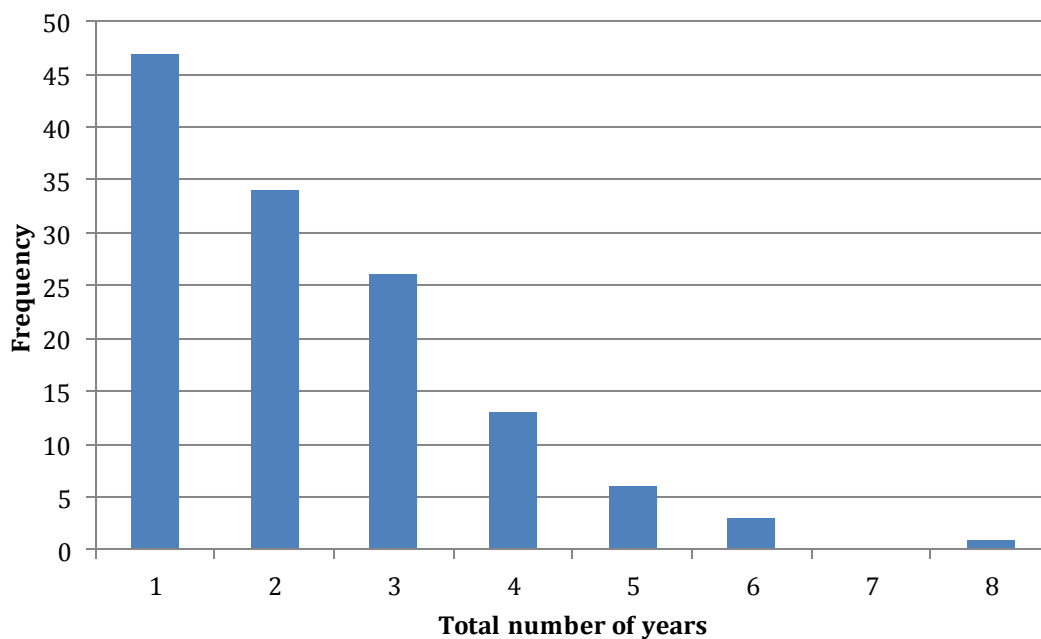
Fig. 2a Epanechnikov kernel density plot of sales growth (1951-2007), darker shades indicate higher density (bandwidth = 0.1)
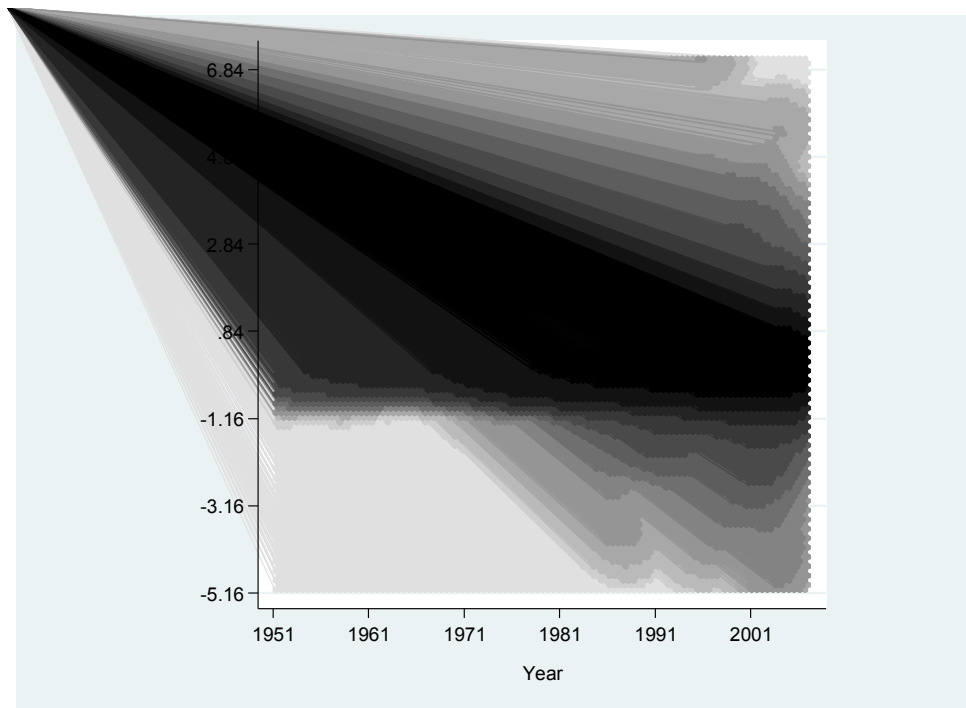

Fig. 2b Density plot of sales growth (1963-2002) vs. normal distribution
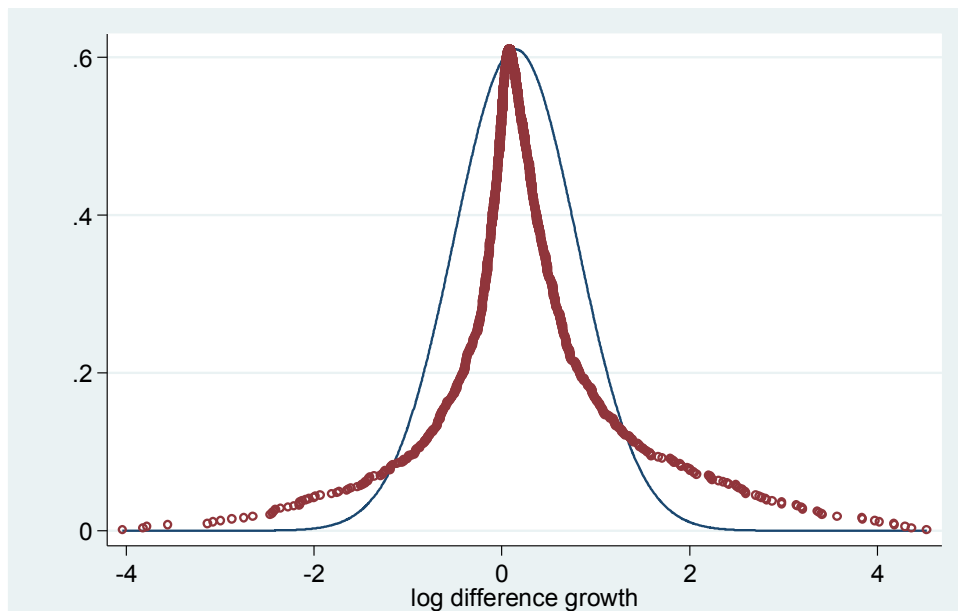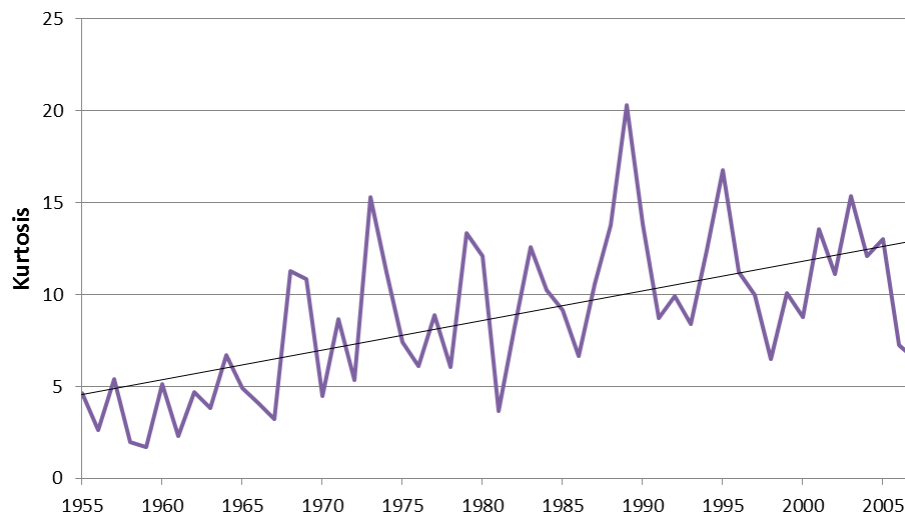
Fig. 3 Kurtosis plot for sales and employment growth

Figures 2 and 3 also provide an important context for the study showing the evolution of the pattern of growth over time: the kernel density plot of the evolution of logged differenced sales growth rates over time (Figure 2a) with an overall probability density plot of sales growth for the period 1963-2002 (Figure 2b) and the changing level of kurtosis of growth rates (Figure 3). Taken together, these figures help motivate our study since they show the importance of not assuming a normal distribution. Figure 2a shows clearly how the range of growth rates represents increasing variety of firm performance over time. We notice that from the 1980s onwards both the range of growth rates increases dramatically, and the density of the distribution close to the mean increases. However, the probability of firms attaining extreme high (and low) growth, as measured via the kurtosis of growth (Figure 3) also increases over time. Over time high growth event observations increase in size, but also occur with greater frequency, such that the distribution of growth becomes increasingly characterized as fat tailed.

Part of the evolution of the nature of high growth may be explained by the presence of SME in our data. We previously noted the increasing composition of SME firms in our data over time. Investigating the growth rates of HGF in our data show SME firms achieving growth rates above either the 80% or 90% quantile were approximately 65% and 78% of firms, respectively. A similar pattern is found for the worst performing firms, where 80% and 55% of observations below the 10% and 20% quantiles, respectively, are also due to SME performance. In the mid-range quantiles of growth, the reverse is true; where at least 69% of growth observations in the mid-range quantiles belong to large firms. Thus extreme performance appears to have a strong connection to SME firms, whereas moderate performance is more likely to result from the traditional large pharma. But what is happening during this time in terms of the underlying patterns of individual high growth firms?

Next in Table 1a we examine a list of fastest annual growth observations by firm (log differenced sales growth), including details of firm characteristics, the year of the high growth event and market share for that year. We include both patenting persistence (those firms that have patented 5 years in a row) and whether the firm had venture capital backing as well. Evaluating the names of companies in this list indicates a low impact group, in terms of market

share. In fact, firms with high growth observations are also likely to experience low growth too (as indicated by firm names in italics).

Table 1a: Top 20 high growth firms 1951-2007 (italics indicate firm also features in bottom 10% growth performance list in at least one year)

Name	R&D intensity*	VC#	Persistent patenting#	Log difference growth	Sales ($m)+	Year	Public age (years)	Mkt share %
NOVADEL PHARMA INC	56%	1	0	5.43	0.17	2004	7	0.0001%
EMISPHERE TECHNOLOGIES INC	85%	0	1	4.52	1.49	1996	9	0.0010%
CHANTAL PHARMACEUTICAL CORP	10%	0	0	4.37	3.53	1995	12	0.0027%
TEKMIRA PHARMACEUTICALS CORP	40%	0	0	4.30	2.22	2001	7	0.0015%
MEDEVA PLC	157%	0	0	4.18	62.10	1990	2	0.0826%
QUIGLEY CORP	5%	0	1	4.18	32.59	1997	2	0.0230%
SPECTRUM PHARMACEUTICALS INC	96%	0	0	4.02	0.92	2002	8	0.0006%
NATUREWELL INC	2%	0	0	3.97	0.45	2002	2	0.0003%
CARDIOME PHARMA CORP	46%	1	1	3.84	0.93	2000	3	0.0006%
GUILFORD PHARMACEUTICAL INC	67%	1	1	3.84	13.32	1996	5	0.0091%
DURECT CORP	72%	1	1	3.57	1.33	2000	3	0.0009%
ADOLOR CORP	26%	1	1	3.41	0.56	2001	8	0.0004%
MGI PHARMA INC	15%	1	0	3.41	2.40	1990	12	0.0032%
ANORMED INC	15%	0	1	3.38	2.11	2001	10	0.0014%
MARSAM PHARMACEUTICALS INC	7%	0	0	3.36	3.90	1989	4	0.0055%
NEXSTAR PHARMACEUTICALS	116%	1	1	3.21	23.93	1994	5	0.0204%
ABLE LABORATORIES INC	27%	0	0	3.20	6.51	1997	9	0.0046%
VEREX LABORATORIES INC	9%	0	0	3.19	0.28	1988	7	0.0004%
CORTEX PHARMACEUTICALS INC	113%	1	0	3.15	1.37	1999	12	0.0010%
ANORMED INC	13%	0	1	3.14	3.76	1998	7	0.0025%

*R&D intensity is calculated as a percentage of lagged asset value. Excludes firms with fewer than six years of recorded data. #Measured as a dummy variable; (e.g. For VC: 1 = venture capital backed 0 = no venture capital). +Rebased to 1982-84

Table 1b: Top 20 firms (1951-2007) by number of years of growth in the 90th percentile

Name	Mean sales ($m)	Mean R&D intensity	VC#	Persistent patenting#	Freq	% time above 90th growth percentile
MDRNA INC	2.22	53%	0	0	8	38%
LESCARDEN INC	0.18	36%	0	0	6	35%
MACROCHEM CORP/DE	0.17	32%	0	0	6	35%
MGI PHARMA INC	16.79	28%	1	0	6	24%
CHANTAL PHARMACEUTICAL CORP	1.71	57%	0	0	5	38%
COLUMBIA LABORATORIES INC	5.88	45%	1	1	5	26%
ARONEX PHARMACEUTICALS INC	1.29	70%	1	1	5	56%
CARACO PHARMACEUTICAL LABS	17.89	42%	0	0	5	36%
DEPOMED INC	2.87	102%	1	1	5	45%
IMMTECH PHARMACEUTICALS INC	1.22	328%	1	0	5	50%
ALCIDE CORP	4.05	17%	1	1	4	21%
CYPRESS BIOSCIENCE INC	1.22	45%	1	0	4	21%
NOVA PHARMACEUTICAL CORP	10.87	30%	1	1	4	57%
PHARMACEUTICAL FORMULATIONS	19.46	11%	0	0	4	18%
VEREX LABORATORIES INC	0.24	41%	0	0	4	31%
EMISPHERE TECHNOLOGIES INC	1.58	45%	0	1	4	21%
REPROS THERAPEUTICS INC	1.61	47%	1	0	4	36%
IMPAX LABORATORIES INC	6.71	28%	1	0	4	67%
MELDEX GROUP PLC	4.51	3%	0	0	4	57%
ANGIOTECH PHARMACEUTICALS	38.14	14%	0	1	4	44%

*R&D intensity is calculated as a percentage of lagged asset value. Excludes firms with fewer than six years of recorded data. #Measured as a dummy variable; (e.g. For VC: 1 = venture capital backed 0 = no venture capital).

Table 1a suggests considerable variation within the high growth category, indicating potential difficulties in trying to target high-growth firms without an understanding of the characteristics which may drive their growth. Our measures of firm characteristics in Table 1a show firms with a wide range of R&D intensities; firms with and without venture capital investment; as well as persistent and non-persistent patenting firms. In addition other characteristics, such as size (sales) or years on the stock market (proxy for age) also show variation. Even those firms repeatedly achieving growth performance above the 90th percentile, as shown in Table 1b, highlight variability in firm R&D, VC, and patenting persistence characteristics. Although, the

names of the largest pharmaceutical firms are notably absent from Table 1a and 1b, the full list of firms achieving growth above the 90th percentile includes firms with annual sales in excess of $1billion. Frequently it is expected that high growth is directly linked with innovative, young SME firms; our descriptive analysis presents a far more complex picture. Next, we examine the pattern of competition and innovation in pharma.

4. Dynamics and the competitive environment

Because in Section 5 we will look at the different behavior of high growth and low growth firms over industry's life-cycle, it is important to get a sense of how the industry has changed over time. Gambardella (1995) argues that before 1980 the 'search regime' in pharma was characterized by a 'random search' process in which large quantities of molecules needed to be screened. This phase did not favor large firms as much as the following period, which he calls the 'guided search regime' in which advances in genomics and IT allowed the search process to become much more guided and driven by scale. Another important change that occurred in early 1980s was the advent of the Bayh-Dole act (1980) which allowed publicly funded research, very much relied on in pharma, to be patented, raising the number of patents significantly, as can be seen in Figure 4. Weighting these patents by citations however reveals that much of the increase was not due to more important innovations but simply the legal changes (Mowery and Ziedonis 2002). The act caused many scientists to set up their own firms, causing the number of firms to increase (as many entered the industry through a patent), as revealed in Figure 5 (entry/exit figure). Orsenigo et al. (2001) have looked at how these changing regimes have affected the relationship between innovation and market structure, concluding that during the 1980s and 1990s the knowledge base of the pharmaceutical industry grew and become more complex in structure as new technologies allowed greater exploration and specialization. During this post 1980 period the number of entrants in pharma increased, typically SMEs, as new technological opportunities spurred entry, with early movers gaining significant advantage. Existing large R&D intensive firms were also able to absorb new knowledge by interacting with entrants, although large pharma firms maintained a hierarchical relationship over new entrants.

Fig. 4 Patent statistics for COMPUSTAT pharma firms
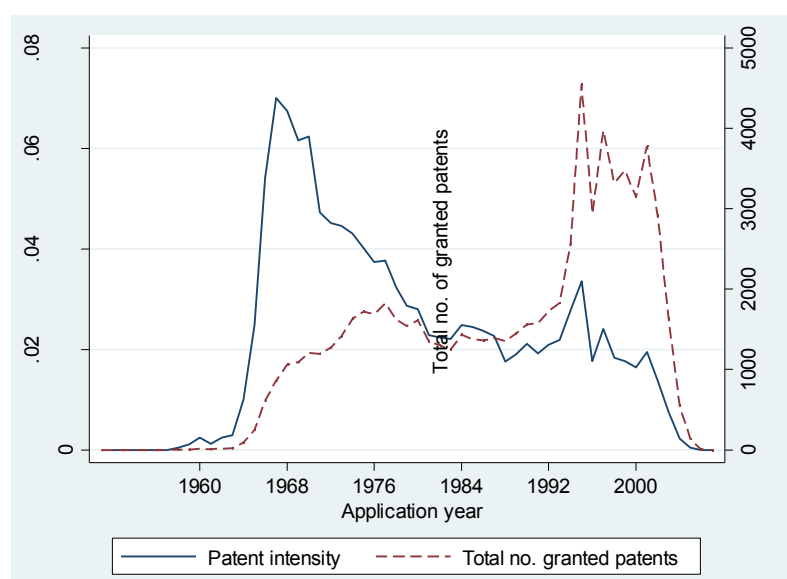
Fig. 5 Entry and Exit of pharma firms listed in COMPUSTAT
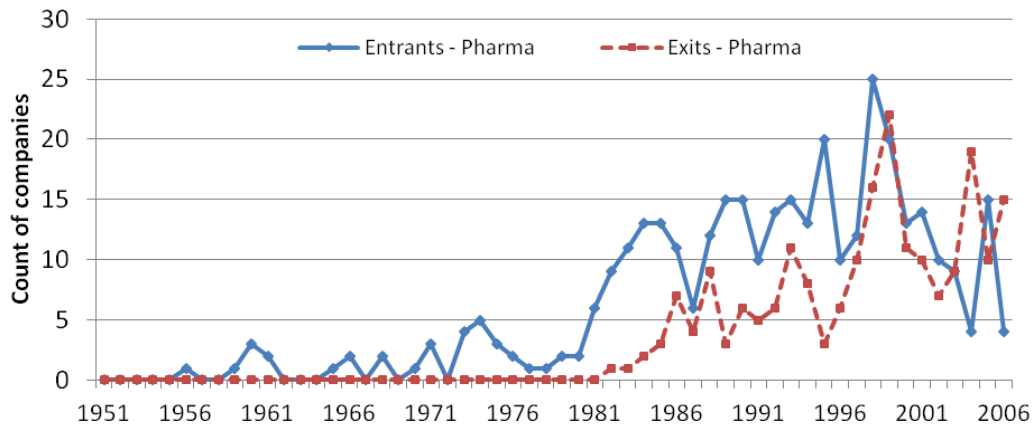
To consider the different 'competitive environments' in pharma, we evaluate the evolution of key environment measures using our pharma dataset, evaluating industry concentration, market share instability and industry level R&D. Based on our assessment of the data we conclude there are two distinctive competition periods which we approximate to pre and post 1980.

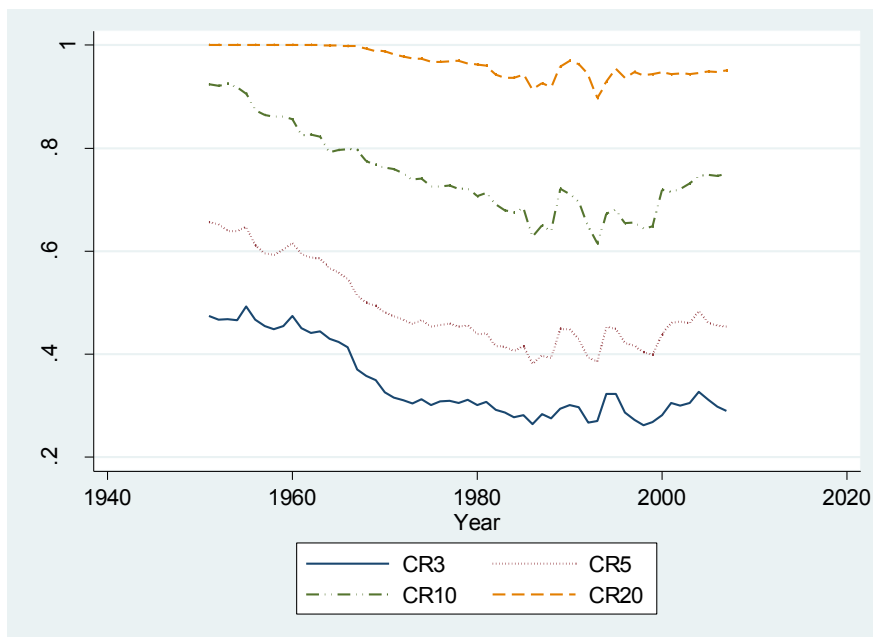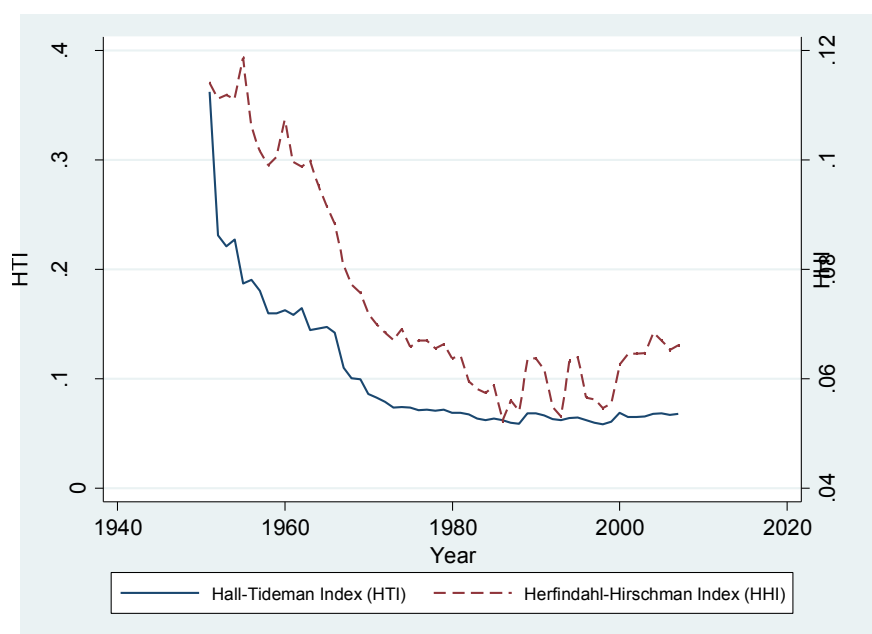**Fig. 6** Pharma concentration ratios

Fig. 7 Pharma HHI and HTI concentration indices

First we look at the degree of competition in pharma, looking at the overall concentration of the industry[4]. As discussed above, different studies have found that competitive pressures affect how much money is spent on R&D: too much competition threatens appropriation; too much concentration reduces innovation incentives and capabilities (Scherer 1967; Aghion et al. 2005). We look at both traditional indices of competition (concentration) and more dynamic indices that have been used by those emphasizing the importance of 'churn' factors.

As regards concentration, our data show from 1950 to 2007 an industry structure formed of a consistent core of large firms, as evidenced by the CR20 (Figure 6). In fact over the full data period, only 43 different firms feature in the CR20, of which 28 appear for more than 10 years. However, despite a general persistence of concentration in market shares within the top 20 firms, the concentration ratios and particularly the two concentration indices Herfindahl – Hirschman (HHI) and Hall-Tidemann (HTI)[5] (Figure 7) reflect a trend of declining market concentration until the 1980s. Declining concentration occurred as the number of entrants rapidly increased, whilst exits remained low. In the post 1980 period concentration remained low, whilst entry and exit were high.

Next we consider the intensity of competition in terms of instability, churn and turbulence, the 'market share instability index', *I* is used to capture the intensity of inter-firm competition, regardless of the degree of concentration (Hymer and Pashigian 1962). It measures the summation of the absolute change in market shares in each year. Hymer and Pashigian argued that this was a superior measure of competition because it could be very high even in industries with few firms. And conversely, if it was low in industries with many firms, this was a sign that

---

[4] Note here market is defined as the total annual sales of COMPUSTAT firms.
[5] The Hall-Tideman Index (Hall and Tideman 1967), an alternative to the Herfindahl-Hirshman index (HHI) (HHI is calculated as a sum of the squares of market share for each Pharma firm), includes information on the number of firms in the concentration index and takes into account entry.

those firms were not feeling the pressure to differentiate from competitors (seeking higher market share).

$$I = \sum_{i=1}^{n}\left[\left|s_{ij} - s_{i,j-1}\right|\right]$$

where s_{ij} = the market share of firm i at time j.

Fig. 8 Pharma instability index and turbulence
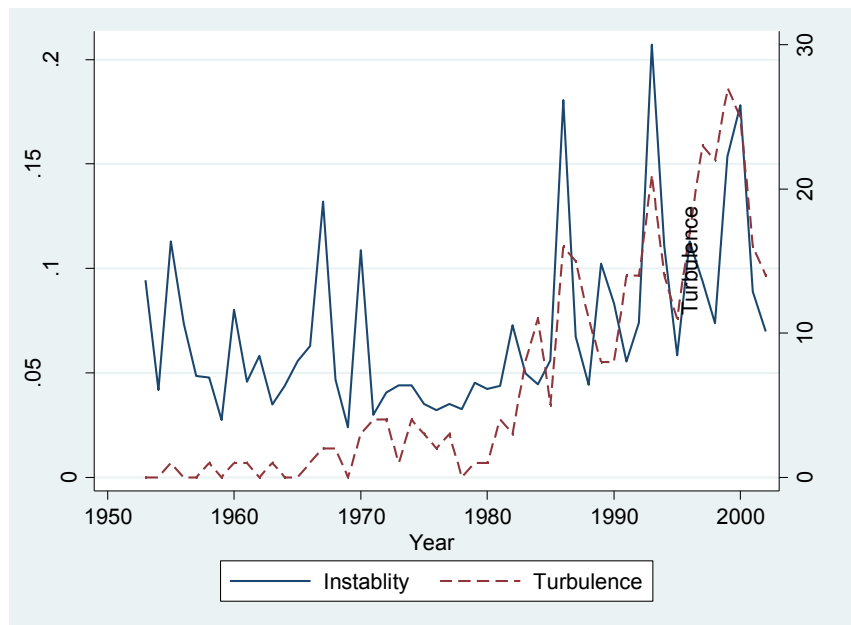
Figure 8 shows a pattern of instability highlighting a divide between pre and post 1980 activity in pharma. We also include a simple measure of turbulence based on the sum of entry and exit, shown as the dashed line. Figure 8 suggests a trend of rising instability over time, particularly in period's post-1980 with increasingly higher peaks of instability. Instability index values peak at 0.21 and 0.19 in the post-1980 period, compared to peak index values of 0.13 and 0.11 in the pre-1980 period. Whilst the pre-1980 period reflected instability as a result of changing market share among a relatively stable and small group of firms, post 1980 the increasing entry and exit contributed to greater changes in market shares of individual firms. Mazzucato (2002) argues that the instability index is particularly useful for capturing periods of creative destruction. In many industries the instability index is highest precisely when innovation is the most radical, not necessarily the period in which there is just high entry (Mazzucato 2002).

Examining the relative level of instability, concentration, entry and exit can be used to summarize the evolution of the competitive environment broadly into two periods, pre and post 1980, as shown in Table 2[6]. During the first period relatively few new firms entered the pharmaceutical industry and still fewer gained significant market share upon entry. Post 1980,

---

[6] Here we are not arguing that there should be a specific date or breakpoint in the competitive regime, but that the pre/post 1980 distinction is a useful method to group the data to capture differences in the competitive environment faced by pharma over time.

the industry restructured operations, moving from 'random screening' to 'guided search' (Gambardella 1995), taking advantage of changes in the continued development of new research tools (Orsenigo et al. 2001), whilst, simultaneously, the introduction of Bayh-Dole (1980) caused major changes to the "legal conditions for the appropriation of new knowledge" (Dosi and Mazzucato 2006, p.4). Thus in the post 1980 period, technical, operational and legal changes in the industry are reflected in increasing instability, high entry and exit and lower levels of concentration.

Table 2: Summary industrial dynamics of Pharma industry

Industry characteristic	1962 – 1980	1981-2002
Instability	Moderate	High
Concentration	High (to moderate)	Low
Entry	Low	High
Exit	Low	High
Industry investment in R&D	Stable over time	Rising over time

Finally, as the pharmaceutical industry is currently questioning the value of R&D given falling R&D productivity, Figure 9 reflects rising R&D intensity (R&D/lagged Sales) particularly in the post 1980 period. In the pre-1980 period total industry sales grew whilst R&D intensity remained fairly constant during the 1960s and 1970s. In the post-1980 period total sales continued to grow, but accompanied with a similar rising trend in R&D intensity. Figure 9 gives further weight to evaluating the influence of R&D and innovation on firm growth during the pre/post 1980 time periods.

Fig. 9 Pharma sales and R&D activity
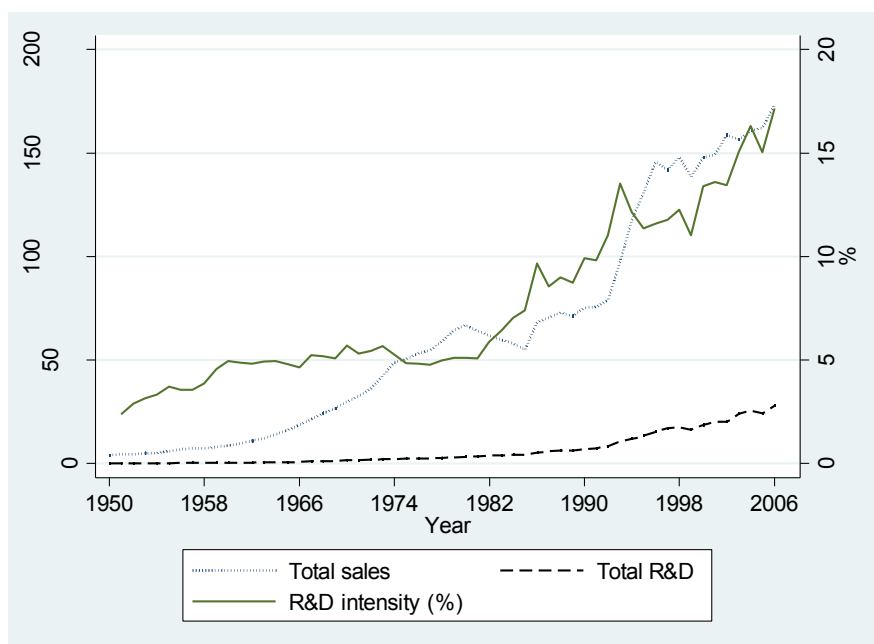
In summary, the competitive environment in pharma can be broadly viewed as formed of two different periods. In the pre-1980 period our measures, summarized in Table 2, signal a weaker

competitive environment; competition occurred within a concentrated firm group at a time when industry revenue was rising, whilst R&D intensity was relatively stable. In the post 1980 period competition was fierce, reflected by rising instability of market shares (a proxy for strong competition) increasing variety in growth performance, and changes in the underlying technology together with a sharp rise in R&D spending.

We use these two periods to investigate how the competitive environment in pharma influences the relationship between firm characteristics and growth, paying special attention to the innovation growth relationship. We describe our method of analysis in Section 5.

5. HIGH GROWTH FIRMS: A QUANTILE REGRESSION APPROACH

Quantile regression methods allow us to look at high growth firms more systematically. We test for the relationship between R&D/innovation and growth, conditional on growth, controlling for other firm characteristics, looking for differences in these relationships pre and post 1980. Previous studies of the R&D growth relationship have used quantile regression, but our approach is to use penalized fixed effect quantile regression, which is able to take account of both the heterogeneity of growth rates in pharma (as shown in Figure 2) as well as to control for firm specific fixed effects.

5.1. Quantile regression approach

Analysis of logged difference sales growth data in Figures 2 and 3 suggest a fat-tailed distribution which prevents using Gaussian assumptions (see Bottazzi and Secchi 2006 for an extended discussion). In such cases it is appropriate to use quantile regression techniques to produce estimates robust to outliers in the dependent variable (Koenker and Bassett 1978). In quantile regression, observations of extreme growth (either positive or negative) are not disregarded as outliers and can be used to estimate the influence of firm characteristics on sales growth at different quantiles of the conditional distribution in order to understand the determinants of growth for top performers.

We estimate the coefficients of each independent variable (i.e. the partial derivative of the conditional quantile of growth (gr) with respect to a particular independent variable) at 10%, 25%, 50%, 75%, 80%, 90% quantiles using Koenker (2004) penalized fixed effect quantile regression to control for unobserved firm fixed effects, applying bootstrapped standard errors[7].

For comparison we also report regression results for OLS and least squares (within) fixed effect estimators with robust errors using Huber-White sandwich estimators. We do not report the standard Koenker and Basset (1978) quantile regression results as they yield very similar results to those reported for the penalized fixed quantile estimates.

5.2 Variables

Our dependent variable is growth, measured as an annual growth rate using the sales revenue of firms, taking logged differences of annual sales observations. As previously described, we find that high growth is a short term feature of firms and we measure annual growth rates to account for the changing dynamics of growth over time. As some pharma firms, particularly in the post 1980 period enter COMPUSTAT for relatively short periods, using annual growth rates

---

[7] See supplementary online material for details.

provides a series of observations to evaluate changes in performance[8]. Although a variety of different growth measures are proposed in the literature, we use sales growth because this is the most commonly used measure of growth in the literature, thus permitting a comparison of our work, and also represents a key performance metric upon which firms are judged (Delmar et al. 2003).

Independent variables calculated from our data sample include, *size*, *R&D intensity*, venture capital (*VC*), innovation persistence (*persist), age, R&D scale,* and a set of variables related to M&A.

To measure firm size we use sales data recorded by COMPUSTAT, in agreement with previous studies, such as Coad and Rao (2008). We also investigate using employment data as an alternative measure of size, and find very similar results, and a strong correlation between sales and employment. However, as employment has missing observations we prefer to use sales as our main measure of size. Given our data includes both large and small surviving firms, based on previous literature we expect size to have a negative relationship with growth (Audretsch 1995).

To measure R&D intensity we combine six different R&D measures into one single composite R&D intensity 'index' variable via the use of principle component analysis. The choice to combine different R&D measures helps to avoid potential bias from focusing on any individual R&D measures. Our R&D variable is calculated from R&D intensity and R&D stock intensity measures. R&D intensity in a period is given by the annual R&D expenditure scaled by a one period lagged measure of firm size. R&D stock intensity is obtained by first calculating an annual measure of R&D stock, using the perpetual inventory method outlined by Hall and Mairesse (1995). R&D stock is then depreciated using an annual discount rate and scaled by a one period lagged measure of firm size to give R&D stock intensity. Although sales is frequently used as a choice of scaling factor in the literature, Brown et al. (2009) suggest asset value to be a more stable scaling factor than revenue. We use both sales and revenue as a size scaling factor to give two sets of R&D intensity measures, one set scaled by lagged sales, the other scaled by assets. In each set we include a simple measure of R&D intensity, and two measures of R&D stock intensity; using a depreciation discount of 15% and then 30% (Hall and Oriani 2006).). The result is that using three different calculations, with two different scaling factors, we produces six measures of R&D intensity as inputs into the principle component analysis.

A principle component analysis was applied to the six different measures to capture as much as possible of the original variance in the different R&D indicators. The first component was retained explaining 52% of the total variance analyzed. We use the first principle component as a composite index measure of '*R&D intensity*' in our analysis[9]. However we stress that the broad pattern of results described in Section 6 are similar if we use individual measures of R&D intensity discussed.

---

[8] We check our model (reported in Section 6) using 3 year logged differenced growth rates and find similar results compared to using the one period measure, however because of the lag structure of model used in our estimations it does significantly reduce the sample size, excluding firms with less than eight years of data. If our analysis is to help reveal the characteristics of all high growth firms and inform policy, it is important to avoid excluding firms, as policy makers cannot know ex-ante which firms will fail.

[9] The six individual measures of R&D have positive loadings of roughly equal size (between 0.28 and 0.5) in the first component.

Our R&D intensity measure is used to determine whether innovation/R&D activity has any influence on growth experienced by firms. Recent quantile based results indicate we should expect the impact of R&D activity tends to increase as we move up the conditional distribution of sales growth (Coad and Rao 2008).

To control for the growth effect of firms receiving venture capital, we construct a time varying dummy variable equal to one for any firm after receiving venture capital (zero for those that have never received VC). Once a firm has received VC, this reflects a permanent change in the dummy variable. One potential concern would be that VC is responsive to the signaling effect of high growth potential firms. In the signaling role VC firms are only responsive to selecting firms with strong potential, but not necessarily acting as a 'coach' to drive high growth in the firm. However it is important to note that in this application the majority of firms receive VC at an early stage of development, well in advance of the time period covered by this dataset and this reduces the possibility of bias in our results from endogeneity of the VC measure. In many cases the direct influence of VC diminishes post IPO, as investors sell their shareholding to obtain return, so the measure is used to assess whether in general receiving VC influences growth performance.

Our measure of VC tests the lasting impact of VC involvement on the growth performance of firms. It cannot distinguish between the coaching and selection performance of VCs as described by Baum and Silverman (2004), but does act as a control for the effects of VC. We would expect given the selection effect and the contribution of resources and coaching, should both have a positive influence on growth, so we expect to find a positive and lasting 'imprint' of VC activity on performance.

To control for the relationship between innovative output and performance we calculate a measure of innovative output persistence using the method proposed by Demirel and Mazzucato (2012). In pharma an output would be considered a pharmaceutical product. In the absence of detailed historical data on pharma outputs, patents can be used as a proxy for innovative outputs. However, as a pharmaceutical innovation usually requires more than a single patent, successful pharma are expected to be persistent innovators, able to produce a series of patents. Thus we use the NBER patent database to generate a measure of innovation persistence, *persist*, measured using a dummy variable equal to one for any firm that has patented consecutively in the previous four years. We note that we already use a measure of R&D intensity which includes components related to R&D stock, and so may reflect R&D 'input' persistence. We therefore use patenting persistence as a measure of innovation 'output' persistence. We would expect innovation persistence to have a positive influence on growth, particularly for HGF.

To control for the effects of age on growth, we determine the public age of firms in the dataset. Following Brown et al. (2009) we include the log of firm age to control for the life-cycle of each firm using the number of years in the COMPUSTAT dataset as a proxy for age[10]. Davidsson et al., (2002) find a negative relationship between age and growth; likewise our expectation is that younger firms will grow faster.

---

[10] For firms present in the first year of COMPUSTAT we manually adjust their age to account for their date of flotation.

In addition to our R&D intensity composite measure, we also measure the scale of R&D operations using total R&D expenditure of each firm. However, rather than measure absolute R&D spending, we express R&D scale as a relative measure to capture the effect of investing greater amounts of R&D than competitors. This reflects the expectation that pharma gains from 'economies of scale in experimentation' (Nightingale 2000), such that achieving significant R&D scale, allows firms to gain greater experimental throughput, and so a greater likelihood of innovative success and ultimately revenue growth. We calculate a variable, *R&D scale* to measure above average annual R&D expenditure relative to all the firms in the dataset each year. This is expressed as a dummy variable, equal to zero when a firm's R&D expenditure is either below or equal to the annual average, otherwise it is set to one. As we already include a measure of firm size based on sales revenue, our R&D scale variable is constructed to control for the effects of being above average in the amount of R&D investment after controlling for R&D intensity, and provides a different perspective on the R&D growth relationship[11]. We expect, after controlling for other characteristics, those firms that increase R&D scale above the industry mean will gain growth advantage.

To control for the effect of acquisitions on growth we investigate two different measures of acquisition activity. First we use the COMPUSTAT recorded annual acquisition expense variable *(Acq value)* as a general measure to record the amount invested in all acquisition activity in a given firm-year. However, as firms may undertake acquisitions for a variety of long-term strategic reasons, not simply to acquire a short term increase of sales, we also use our own collected data on M&A to calculate a second measure using a dummy variable (*Acq DV*) to capture acquisitions of other significant pharma which are expected to directly influence short-term sales growth. If a firm makes an acquisition of a COMPUSTAT list pharma in a particular fiscal year, then we set the dummy variable, *Acq DV*, for that firm-year equal to one. The acquisition of another pharma firm can be expected to dramatically increase the sales growth rate of the parent pharma. Neither variable accounts for changes in the long term performance of a firm as a result of acquisition activity.

5.2. Modeling the R&D-growth relationship
Our methodology follows a trajectory established by studies (Gibrat 1931; Mansfield 1962; Hall 1987; Hart and Oulton 1996) that seek to determine the existence of regularities of firm growth in industrial dynamics.

In this section we outline the model used to evaluate how sales growth (gr) is related to R&D and innovation. Our core contribution is to firstly understand how changing industry level dynamics influence the R&D-growth relationship and secondly investigate additional innovation variables that help explain growth performance. Our selected model extends the conventional Gibrat type sales growth equation where firm growth rates are usually regressed on size, here measured as log of sales and one period lag of growth ($gr_{i,t-1}$) to control for autocorrelation of growth (Evans 1987). Our model includes a composite index measure of R&D intensity (*R&D intensity*) to test for the relationship between R&D and growth, a dummy variable for venture capital (VC), patent persistence (persist), and R&D scale (R&Dscale), all of which enter with one

---

[11] R&D scale is strongly correlated with annual R&D expenditure (corr = 0.78), however rather than investigate the effect of small changes in R&D expenditure on growth, we investigate the extent to which above average annual R&D expenditure has on growth – therefore giving a relative R&D performance measure. R&D scale also has a weak, but negative correlation with R&D intensity.

period lags. We also include the log of public age (*age*), and to control for common macroeconomic influences we also include time dummies for each year (∂_j).

To account for the possibility of patenting activity being correlated with the errors in the model, potentially as a result of growth and patenting being stimulated by an unobserved variable, endogeneity tests were run using instrumental variables for patenting activity including lagged measures of patenting persistence and also a patenting dummy variable recording years in firms successfully applied for a patent. The null hypothesis of exogeneity in Durbin and Wu–Hausman tests were not rejected.

The basic model is defined as:
$$gr_{it} = \alpha + \beta_1 \log(Sales_{i,t-1}) + \beta_2 gr_{i,t-1} + \beta_3 R\&Dins_{i,t-1} + \beta_4 VC_{i,t-1} + \beta_5 Persist_{i,t-1} + \beta_6 \log(Age_{it}) + \beta_7 R\&Dscale_{i,t-1} + \delta_t + u_{it}$$

and
$$gr_{it} = \log(Sales_{it}) - \log(Sales_{i,t-1})$$

As a robustness test we extend the model to test for the effects of M&A and growth persistence and report results for the 90% growth quantile[12]. To control for acquisitions we include the log of acquisition value (*Acq value*) to capture the current effect on growth, with a one (*Acq value1*) and two (*Acq value2*) period lag to try and capture any effect of acquisition on future growth[13]. We repeat and re-estimate the model including our alternative dummy variable measure of acquisition activity (*Acq DV, Acq DV1, Acq DV2*), using the same lag structure. Finally, via stepwise regressions we report the effects of additional two, and three period growth lags[14].

Although the construction of the basic regression requires a minimum of 3 consecutive years of observations for each firm, because as a robustness check we test for two additional lags in the model, we use a consistent data set which excludes firms with greater than five consecutive years of history[15]. The sample is restricted to 1963-2002 because of the unreliability of patent data in the 1950-1962 and 2003-2007 periods, as noted in Section 3.1[16].

In Table 3 we report summary annual statistics for our data sample data which only includes firms with greater than 5 years of data history. In Table 4 we report the correlation coefficient for the sample. Next we discuss the estimates of the model.

---

[12] Results for all quantiles were inspected, and are available. However, as our main interest is HGF, and the main results are robust to including M&A activity, we have only reported the results for the top performing firms.

[13] Similar to Danzon et al. (2007) we look for the effect of M&A on subsequent firm performance, as well as in the acquisition year.

[14] Note, lags of R&D intensity, R&D scale, Sales, VC, and Persist were tested, but were either highly collinear or statistically insignificant and could be dropped without implication to the results reported.

[15] In fact adjusting this restriction to allow for firms with between a minimum of three (without additional lags in the model) and ten years of history yields very similar estimation results to those reported in Section 6 for R&D intensity and VC.

[16] Regressions run on the 1951-2007 period without patenting persistence give a similar pattern of results for the remaining coefficients to those described in Section 6.

Table 3 Summary descriptive statistics (firms with greater than five years of consecutive observations)

	Obs	Mean	Std.Dev	Variance	Skewness	Kurtosis	Median	Percentile		Min	Max
								10th	90th		
Year: 1963-2002	n=196, T = 12.04										
gr	2360	0.146	0.654	0.427	0.997	15.941	0.083	-0.231	0.627	-4.040	4.524
Sales	2360	4.027	3.253	10.583	-0.222	2.236	3.782	-0.085	8.281	-6.981	9.913
R&D intensity (index)	2360	-0.069	1.418	2.010	11.927	198.263	-0.362	-0.471	0.364	-0.521	33.371
VC	2360	0.206	0.404	0.163	1.458	3.125	0	0	1	0	1
Persist	2360	0.371	0.483	0.234	0.533	1.284	0	0	1	0	1
Age	2360	2.574	0.820	0.673	-0.136	2.442	2.565	1.382	3.689	0.693	4.263
R&D scale	2360	0.267	0.442	0.196	1.054	2.110	0	0	1	0	1
Year: 1963-1980	n=43, T =11.26										
gr	484	0.110	0.199	0.039	1.466	12.596	0.095	-0.036	0.282	-0.747	1.443
Sales	484	5.846	2.254	5.079	-0.831	2.568	6.507	2.203	8.196	-0.254	9.102
R&D intensity (index)	484	-0.402	0.089	0.008	3.560	24.429	-0.422	-0.485	-0.322	-0.521	0.289
VC	484	0.039	0.194	0.038	4.745	23.515	0	0	0	0	1
Persist	484	0.545	0.498	0.248	-0.183	1.033	1	0	1	0	1
Age	484	2.575	0.734	0.539	-0.784	3.170	2.890	1.609	3.611	0.693	3.931
R&D scale	484	0.419	0.494	0.244	0.327	1.107	0	0	1	0	1
Year: 1981-2002	n=196; T = 9.57										
gr	1876	0.155	0.726	0.527	0.873	13.123	0.077	-0.331	0.695	-4.040	4.524
Sales	1876	3.558	3.306	10.932	-0.010	2.275	3.180	-0.574	8.304	-6.981	9.913
R&D intensity (index)	1876	0.017	1.578	2.490	10.719	160.115	-0.339	-0.468	0.565	-0.521	33.371
VC	1876	0.248	0.432	0.187	1.165	2.356	0	0	1	0	1
Persist	1876	0.326	0.469	0.220	0.741	1.550	0	0	1	0	1
Age	1876	2.528	0.835	0.697	0.011	2.421	2.485	1.368	3.738	0.693	4.263
R&D scale	1876	0.228	0.419	0.176	1.299	2.688	0	0	1	0	1

n = number of firms; T = average duration of firm; Sales and Age are reported in natural logs.

Table 4 Correlation coefficients

Variable (n=2360)	1	2	3	4	5	6	7	8	9	10
1 gr	1.000									
2 Sales	-0.013	1.000								
3 R&D intensity	0.209	-0.350	1.000							
4 VC	0.094	-0.261	0.164	1.000						
5 Persist	-0.052	0.549	-0.052	-0.057	1.000					
6 log age	-0.149	0.625	-0.130	-0.133	0.541	1.000				
7 R&D scale	-0.088	0.741	-0.126	-0.274	0.571	0.619	1.000			
8 R&D invest	-0.049	0.810	-0.090	-0.103	0.644	0.619	0.779	1.000		
9 Acq DV	0.021	0.132	-0.026	0.013	0.060	0.093	0.098	0.149	1.000	
10 Acq Value	0.002	0.486	-0.114	-0.09	0.319	0.276	0.363	0.449	0.167	1.000

Note *R&D invest* is the log of annual R&D expenditure, and is used to generate R&D scale. We include *R&D invest* in the correlation matrix for reference only.

Table 5 OLS, fixed effect (FE) and penalized quantile regression with fixed effect estimates 1963-2002:

	OLS	FE	Quantile: 10%	20%	25%	50%	75%	80%	90%
Sales	-0.0239**	-0.367***	0.06637***	-0.12929**	0.01166*	-0.01488***	-0.0508***	-0.05364***	-0.11159***
	(0.0115)	(0.0187)	(0.01572)	(0.05584)	(0.00636)	(0.00478)	(0.00648)	(0.00843)	(0.01237)
$gr_{i,t-1}$	-0.0898*	-0.0158	0.00482	0.02594***	0.04949	0.04571	-0.00359	-0.02695	-0.11496***
	(0.0458)	(0.0212)	(0.07125)	(0.00751)	(0.03893)	(0.02936)	(0.03143)	(0.04144)	(0.0384)
R&D intensity	0.0645***	0.0459***	-0.16053	0.0328	-0.01547	0.05966***	0.10087***	0.11385**	0.16023**
	(0.0189)	(0.0131)	(0.12276)	(0.04428)	(0.04376)	(0.01555)	(0.03609)	(0.04582)	(0.07182)
VC	0.0639	-0.234**	0.0318	-0.01598	-0.00835	0.02921	0.03355	0.09397**	0.16368***
	(0.0446)	(0.0987)	(0.06823)	(0.05253)	(0.03227)	(0.02314)	(0.03565)	(0.04272)	(0.06124)
Persist	0.0558	0.0827*	-0.00893	0.00849	0.00558	0.0034	0.01118	-0.00139	0.02477
	(0.0346)	(0.0494)	(0.03044)	(0.0391)	(0.01654)	(0.01105)	(0.01474)	(0.01828)	(0.02344)
Age	-0.0675***	0.0702	0.01908	-0.003	0.01991	0.0103	0.0016	-0.01238	0.0099
	(0.0257)	(0.0562)	(0.02136)	(0.01999)	(0.01227)	(0.01069)	(0.01503)	(0.01792)	(0.02151)
R&D scale	0.0969***	0.105	-0.09753**	0.02115	-0.02676	-0.0016	0.04295**	0.05083***	0.13353***
	(0.0391)	(0.0687)	(0.04125)	(0.01484)	(0.01926)	(0.0135)	(0.01969)	(0.02458)	(0.0355)
α	0.470***	1.078***	-0.42342***	-0.12929**	-0.04123	0.18952	0.49825***	0.56322**	0.89305***
	(0.0708)	(0.173)	(0.10036)	(0.05584)	(0.04795)	(0.01227)	(0.04729)	(0.05849)	(0.07214)
Obs	2175	2175	2175	2175	2175	2175	2175	2175	2175

Note: *** 1% ** 5% * 10% significance level, standard error in brackets.

Table 6 OLS, fixed effect (FE) and penalized quantile regression with fixed effect estimates 1963-1980

	OLS	FE	Quantile:						
			10%	20%	25%	50%	75%	80%	90%
Sales	-0.0174*	-0.237***	0.01504	-0.10916	0.00517	-0.01945**	-0.02882***	-0.03411***	-0.0642***
	(0.00981)	(0.0270)	(0.00915)	(0.09625)	(0.00972)	(0.00912)	(0.00992)	(0.01223)	(0.01277)
$gr_{i,t-1}$	0.366***	0.209***	0.20059*	0.013	0.16209*	0.21667**	0.29817**	0.32015**	0.36843***
	(0.0808)	(0.0434)	(0.10407)	(0.00947)	(0.08726)	(0.08884)	(0.13612)	(0.13224)	(0.0771)
R&D intensity	-0.188	-0.113	0.19432	0.20981**	0.06148	-0.02282	-0.16913	-0.31105*	-0.56129***
	(0.125)	(0.246)	(0.17379)	(0.08998)	(0.14084)	(0.10716)	(0.12995)	(0.161)	(0.18392)
VC	0.0374	0.124	-0.07082	0.00401	-0.00132	0.00544	0.07428	0.08267	0.18454***
	(0.0582)	(0.0816)	(0.11355)	(0.15154)	(0.07386)	(0.05631)	(0.0511)	(0.0588)	(0.07059)
Persist	0.00447	0.0424	-0.00009	0.01569	-0.00829	0.00184	-0.00447	-0.01584	-0.01143
	(0.0159)	(0.0318)	(0.01979)	(0.08141)	(0.01257)	(0.01036)	(0.01519)	(0.01731)	(0.02099)
Age	0.0159	-0.0955*	0.05143	-0.00646	0.02783	0.03734*	0.02218	0.01444	0.02826
	(0.0236)	(0.0499)	(0.03245)	(0.01448)	(0.02472)	(0.02261)	(0.0318)	(0.034)	(0.03314)
R&D scale	0.0373**	0.00258	0.01004	0.0238	0.00711	0.01801	0.01913	0.02683	0.03737*
	(0.0181)	(0.0267)	(0.0236)	(0.02638)	(0.01352)	(0.01293)	(0.01661)	(0.01852)	(0.01993)
α	0.0668	1.362***	-0.13975	-0.10916	-0.03576	0.05748	0.12811	0.1286	0.19195*
	(0.0713)	(0.171)	(0.11218)	(0.09625)	(0.09477)	(0.02472)	(0.08685)	(0.09217)	(0.10124)
Obs	452	452	452	452	452	452	452	452	452

Note: *** 1% ** 5% * 10% significance level, standard error in brackets.

28

Table 7 OLS, fixed effect (FE) and penalized quantile regression with fixed effect estimates 1981-2002:

	OLS	FE	Quantile: 10%	20%	25%	50%	75%	80%	90%
Sales	-0.0251*	-0.424***	0.094***	-0.37999***	0.01719*	-0.01312**	-0.05503***	-0.06072***	-0.13183***
	(0.0131)	(0.0224)	(0.01962)	(0.07713)	(0.00882)	(0.00637)	(0.00802)	(0.01101)	(0.01427)
$gr_{i,t-1}$	-0.0972**	-2.72e-05	-0.06049	0.03708***	0.02134	0.02957	-0.02851	-0.05584	-0.10271**
	(0.0465)	(0.0239)	(0.07256)	(0.01109)	(0.04276)	(0.03228)	(0.0341)	(0.04811)	(0.04036)
R&D intensity	0.0643***	0.0402***	-0.06226	0.02098	-0.00997	0.06245***	0.10211***	0.1083**	0.14354**
	(0.0192)	(0.0146)	(0.11866)	(0.04964)	(0.04581)	(0.01553)	(0.03669)	(0.04581)	(0.06763)
VC	0.0638	-0.343***	0.0058	-0.01028	-0.01655	0.02596	0.03292	0.08144*	0.1203**
	(0.0476)	(0.131)	(0.07366)	(0.0551)	(0.03635)	(0.02621)	(0.03834)	(0.04674)	(0.05757)
Persist	0.0654	0.0579	0.00336	-0.00153	0.02947	0.0106	0.02703	0.00479	0.03337
	(0.0441)	(0.0619)	(0.04778)	(0.04218)	(0.02783)	(0.01919)	(0.0222)	(0.02912)	(0.03207)
Age	-0.0789***	0.132	0.03579	0.00141	0.02071	0.00476	-0.01598	-0.02854	-0.00735
	(0.0288)	(0.0808)	(0.02676)	(0.03322)	(0.01608)	(0.01401)	(0.02012)	(0.02518)	(0.02785)
R&D scale	0.113*	0.142	-0.2514***	0.02436	-0.08037**	-0.02681	0.06224*	0.08232*	0.25391***
	(0.0593)	(0.134)	(0.07312)	(0.01971)	(0.03931)	(0.02839)	(0.03763)	(0.04419)	(0.05725)
α	0.320***	1.154***	-0.86522***	-0.37999***	-0.25713***	0.06225	0.4156***	0.53054***	0.91491***
	(0.0740)	(0.181)	(0.13687)	(0.07713)	(0.06401)	(0.01608)	(0.05707)	(0.08404)	(0.10686)
Obs	1723	1723	1723	1723	1723	1723	1723	1723	1723

Note: *** 1% ** 5% * 10% significance level, standard error in brackets.

6. Quantile regression results

In Tables 5, 6 and 7 we present the results for the estimates of the coefficients of the basic regression model described in Section 5. We briefly discuss the results for the estimates of the full data period (1963-2002) in Table 5. The results in Table 6 and 7 provide our main focus showing how the coefficient of R&D intensity changes between the two time periods. Table 6 shows the results for the period 1963-1980, Table 7 shows the results for the period 1981-2002. In each table we report the estimates of model coefficients for the pooled OLS, fixed effect estimator as well as the penalized fixed effect (FE) quantile regression. The results of the robustness tests for the 90% quantile, including measures of acquisition and additional growth lags, are reported in Tables 8 and 9.

The quantile regression results in Table 5 for the full period sample (1963-2002) using the basic model presented in Section 5.2 are broadly in agreement with previous studies. Our measure of firm size, sales, is negative for high growth firms, indicating that smaller firms grow faster. Lagged growth also has a negative influence on firm growth performance in the 90% quantile of growth, indicating an absence of high growth persistence.

Our main measure of interest, R&D intensity is positive and statistically significant at the 50% quantile and above. For firms in low growth quantiles, increasing R&D intensity has no statistically significant effect, investing in R&D would only seem a good strategy for high growth firms. In agreement with previous studies (e.g. Coad and Rao, 2008) the importance of R&D on growth, rises over performance quantiles. In fact the size of the coefficient of R&D intensity at the 90% quantile is close to three times greater than at the median level. Interestingly, the benchmark of pooled OLS and FE panel regression results show a statistically significant and positive influence of R&D on growth, reporting coefficients similar in strength to those reported at the 50% quantile, but severely underestimate the strength of this relationship at the higher growth quantiles.

In agreement with our approach to understanding the characteristics of firms, Table 5 shows our measure of VC is found to only influence the growth performance at the higher quantiles. For instance at the 90% conditional quantile, holding all other independent variables constant, a firm that has received VC experiences an increase in the growth measure of 0.16. Considering the log difference growth rate of the sample at the 90% quantile is 0.63, the coefficient of VC at 90% conditional quantile represents a significant additional growth advantage for those VC backed firms. Notably the pooled OLS regression does not show any influence of VC on growth, corresponding with our observation that the VC – growth relationship is only present at high growth quantiles.

Our measure of innovation, *Persist* is not found to be statistically significant at any growth quantile (or in the sub period analysis presented in Table 6 and 7). This result may not be surprising as many persistent patenting firms are large, and so the influence of patenting persistence may be controlled for by a combination of R&D scale, as well as VC which targets smaller firms with a history of patenting[17]. Also, by construction, a measure of patenting persistence may not show considerable variation over time – being a firm strategy which by definition can only be achieved with significant and prolonged investment, so may be part of the

---

[17] If we exclude R&D scale and VC we find evidence that Persist is statistically significant at growth quantiles above 50%.

estimated fixed effect. In quantile regression models without fixed effects, patenting persistence is found to have a statistically significant and positive influence on growth at high growth quantiles.

We note that age is not found to be a statistically significant predictor of growth. We note that age is strongly and positively correlated with size (sales), patenting persistence (persist) and R&D scale; it is likely that some of the effects of age reported in previous studies are controlled for by additional factors included in our model.

Finally we note that including R&D scale has a variable effect on growth, conditional on the growth quantile. High growth firms gain additional benefit from increasing their R&D investment above industry average, even after controlling for R&D intensity. In contrast for those firms in lower growth quantiles, the reverse is true, investing in R&D to achieve above average R&D expenditure reduces their growth performance. This supports our position that high growth is not an exclusive characteristic of small firms.

Whilst the results in Table 5 capture the effects of variables in our model on growth over the full period, our main interest is to evaluate the impact of the changing competitive environment on the R&D-growth relationship. Figure 10 shows how the coefficient of R&D intensity changes over time, using the same model reported in Table 5, however applying a rolling quantile regression over 9 year periods[18]. It shows that for HGFs the coefficient of R&D intensity at the 90% is always positive post 1980, but for significant parts of the pre-1980 period the coefficient of R&D intensity is negative. The reverse is true for the poorest performing firms in each nine year window, that appear to gain from increases to R&D intensity in the pre-1980 period, but lose from increasing R&D intensity in the post 1980 period. In order to further investigate the influence of the competitive environment on the R&D-growth relationship we re-estimate our model results for two periods 1962-1980 and 1981-2002.

---

[18] Here we use quantile regression, but without fixed effect as the rolling regression is produced using only a 9 year window. We also exclude patenting persistence for periods prior to 1962. Data plotted for mid-point of each window.

Fig. 10 Rolling regression (9 year window) for the coefficient of R&D intensity, over time (10%, 50% and 90% quantiles of growth only)
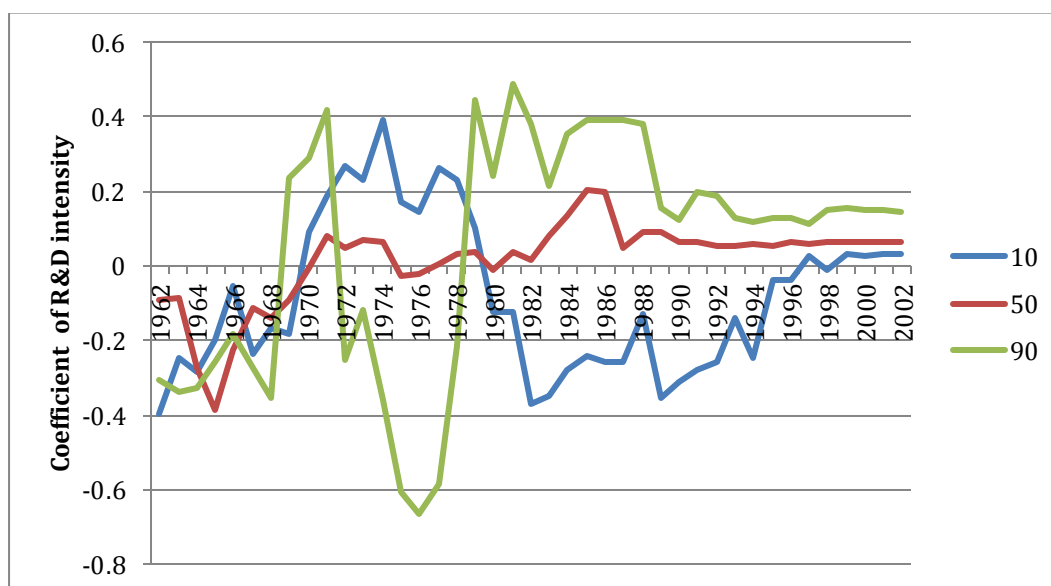

Table 6 shows estimates of our model for the first sub period 1962-1980. We note that firm size, innovation persistence, and age are similar to those reported in Table 5. We also note that this more stable competitive environment, as might be expected, is broadly characterized with positive growth persistence, with previous growth a significant and positive predictor of growth. VC remains a significant predictor of growth, but for growth in the 90% quantile only.

Our main variable of interest is coefficient of R&D intensity and R&D scale. R&D intensity is statistically significant at the 80% and 90% quantiles, but has a negative influence on growth for the fastest growing firms. In fact we only find evidence of a positive influence of R&D intensity on growth at the 20% growth quantile. At remaining quantiles we find no influence of R&D on growth. In contrast the influence of R&D scale, although positive, is only statistically significant for top growth performers. In this pre 1980 period increasing R&D intensity appears to act as a drag for the fastest performers and irrelevant to the growth performance of the majority of firms, but achieving above average R&D scale counts for HGFs.

Table 7 shows the results of the model estimated for the period 1981-2002. The post 1980 period indicates similar outcomes to results reported for the full period (Table 5). In this more turbulent period, small firms appear to grow faster, as shown by the coefficient of sales, but there is an absence of growth persistence – with past growth inversely related to current growth at the 90% quantile. The results for this more competitive period, show that only the fastest growing firms benefit from R&D, whether from increases to R&D intensity, or achieving R&D scale. As before VC remains important for the fastest performing firms having an influence of similar strength to that reported in Table 5.

6.1 Robustness checks

Table 8 and 9 reports results for the 90% quantile only, i.e. they show estimates conditional on the high growth tail of the growth distribution. These tables report a series of robustness checks on our main results, extending the model described in Section 5.2 with additional control variables. Table 8 shows the results for the pre-1980 period, Table 9 for the post 1980 period.

Table 8 Extending basic model 1963-80- penalized quantile regression with fixed effect estimates for 90% quantile only

	1 Coefficient	1 Std error	2 Coefficient	2 Std error	3 Coefficient	3 Std error	4 Coefficient	4 Std error	5 Coefficient	5 Std error
Sales	-0.05763***	0.01297	-0.0642***	(0.01277)	-0.05786***	0.01634	-0.04489***	0.01533	-0.04407***	0.01595
$gr_{i,t-1}$	0.3918***	0.08233	0.36843***	(0.0771)	0.29387***	0.09331	0.18454***	0.06733	0.19954***	0.06942
R&D intensity	-0.58892***	0.20572	-0.56129***	(0.18392)	-0.54451**	0.24155	-0.01381	0.24519	-0.03362	0.23971
VC	0.1814**	0.07797	0.18454***	(0.07059)	0.11231*	0.06483	0.01667	0.06754	0.0028	0.0796
Persist	-0.00487	0.02197	-0.01143	(0.02099)	-0.01897	0.0271	-0.01349	0.02459	-0.01909	0.02815
Age	0.03195	0.03558	0.02826	(0.03314)	0.0353	0.04038	0.05771	0.03901	0.04964	0.03962
R&D scale			0.03737*	(0.01993)	0.04549*	0.02673	0.00613	0.02546	-0.00596	0.02703
$gr_{i,t-2}$					0.2216**	0.086	0.12365**	0.05597	0.14327***	0.05424
$gr_{i,t-3}$							0.20374**	0.09191	0.18535*	0.0973
Acq value									0.00423	0.00348
Acq value$_{t-1}$									-0.00138	0.00259
Acq value$_{t-2}$									-0.00348	0.00263
α	0.14715	0.11332	0.19195*	(0.18392)	0.12946	0.11474	0.20683	0.12764	0.2342*	0.12402
Obs	452		452		452		452		452	

Note: *** 1% ** 5% * 10% significance level, standard error in brackets. Column 2 is repeated from Table 6, as Acq DV has no recorded values prior to 1980.

Table 9 Extending basic model 1981-2002 - penalized quantile regression with fixed effect estimates for 90% quantile only

	1		2		3		4		5	
	Coefficient	Std error	Coefficient	Std error	Coefficient	Std error	Coefficient	Std error	Coefficient	Std error
Sales	-0.09524***	0.01150	-0.13208***	0.01393	-0.11731***	0.01522	-0.11024***	0.01696	-0.11089***	0.01678
$gf_{i,t-1}$	-0.14257***	0.04030	-0.10747***	0.04073	-0.12599***	0.04114	-0.09366*	0.05130	-0.10961**	0.05285
R&D intensity	0.18385**	0.08230	0.14544**	0.06859	0.15986***	0.06385	0.2531**	0.10311	0.24995**	0.10198
VC	0.13284*	0.07119	0.11513*	0.06483	0.10975**	0.07594	0.10536	0.07446	0.11082	0.06861
Persist	0.04307	0.03484	0.02248	0.03395	0.0168	0.03782	0.01738	0.03752	0.00965	0.04003
Age	0.00911	0.02492	-0.00734	0.02529	-0.04897*	0.02847	-0.01112	0.03196	-0.00032	0.03456
R&D scale			0.24069***	0.05595	0.2112***	0.05666	0.14867**	0.06191	0.13421**	0.06343
$gf_{i,t-2}$					-0.07237	0.06001	-0.13225**	0.06287	-0.10418	0.06355
$gf_{i,t-3}$							0.02069	0.0466	0.03529	0.04932
Acq DV			0.39783***	0.0819	0.36128***	0.08948	0.37732***	0.09198		
Acq DV1			0.09596	0.07026	0.11469	0.07163	0.11067	0.07498		
Acq DV2			0.00721	0.06749	0.03689	0.06641	0.0424	0.07404		
Acq value									0.00803**	0.00414
Acq value $_{t-1}$									0.00668	0.00466
Acq value $_{t-2}$									-0.00287	0.00324
α	0.7939***	0.09955	0.94201***	0.11334	1.01018***	0.1305	0.90869***	0.14685	0.93151***	0.14286
Obs	1723		1723		1723		1723		1723	

Note: *** 1% ** 5% * 10% significance level, standard error in brackets.

In each table, the first column shows the results without including R&D scale in the model, in case including this variable changes the estimate of R&D intensity. The second column shows the effect of including the dummy control variables for acquisition (Acq DV); this measure applies to the post -1980 period only[19]. The third and fourth column shows the influence of controlling for additional growth lags. Finally, the fifth column repeats the acquisition analysis, this time using the logs of acquisition value (Acq value) reported in COMPUSTAT for both periods, instead of the acquisition dummy variable (Acq DV).

We note that our results for R&D intensity and R&D scale reported for the post 1980 period (Table 9) are robust to including additional control variables and consistent with those reported in Table 7. We note that including additional lags of growth appears to control for the effects of VC on growth which may be a result of VC targeting individual firms with an outstanding history of growth.

However, in the pre-1980 period when we add a third lagged growth variable, the results in Table 8 show that R&D intensity and R&D scale become insignificant and the size of the coefficients reduce dramatically. The effects of including additional lags lend further evidence to confirm that pre-1980 growth is more predictable and, at best, unrelated to investment in R&D. However, although post-1980 growth is generally more erratic and less persistent, top performing firms can gain advantage from investing to increase R&D intensity or by achieving significant R&D scale.

7. Discussion and conclusion

The results provide insights on the degree to which 'high growth' is a meaningful category in innovation policy. We find that it is a very transient process of particular firms—50% of the time lasting for 2 years or less, and that it is not homogeneous, but instead includes different types of firms, both small and large, and with both high and low R&D intensities. We find that fast growing firms gain increased performance from investing more intensely in R&D, but can also gain additional growth from achieving a scale of R&D expenditure above the industry average. In combination these firm characteristics offer an important insight into the drivers of high growth, and of outlier behavior. The growth advantage pharma firms gain from investing in R&D is specific to the type of R&D investment (achieving higher intensity vs. achieving above average scale) and also determined by the competitive environment facing the firm.

Unlike other studies, we control for the competitive environment and find that it matters. Differences in the relationship between R&D and growth (between quantiles) matters only when proxies for the competitive environment are controlled for. High growth firms that focus on increasing R&D intensity fail to realize any growth advantage when the competition regime is relatively weak (high concentration, market share stability and relatively stable industry investment in R&D). In contrast, when the competition regime is relatively fierce (less concentration and more market share instability, rising industry R&D investment), increasing R&D intensity benefits those firms with growth performance at the median level or above, with the strongest effect found for high growth firms. Furthermore, during periods of fierce competition high growth firms that invest to achieve above average investment in R&D also appear to gain growth advantage. In this sense we seem to confirm Nelson's (1991) untested

---

[19] The absence of acquisitions pre-1980 is discussed in Section 3.1.

hypothesis that variety matters more in periods in which there is higher intensity of innovation and dynamic change. In times of strong competition firms must grow or die, and it is through innovation that growth occurs.

We believe this suggests that industrial policy must be clearly linked to competition policy but where competition is understood dynamically. It is futile to target 'high growth firms' without also targeting the competitive conditions in which they operate.

Acknowledgements

References

Acs, Z.J., & Audretsch, D.B. (1990). The determinants of small-firm growth in US manufacturing. *Applied Economics* 22, 143-153

Acs, Z.J., & Yeung, B. (1999). Conclusion. In Acs, Z., & Yeung, B. (Eds.), *Small and Medium-Sized Enterprises in the Global Economy* (pp.164-173)*.* Ann Arbor: University of Michigan Press.

Acs, Z.J., Parsons, W., & Spencer, T. (2008). *High impact firms: Gazelles revisited*. Technical report, Office of Advocacy Working Paper, U.S. Small Business Administration.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and Innovation: An Inverted-U Relationship. *Quarterly Journal of Economics*, 120(2), 701-728.

Almus, M., & Nerlinger, E.A. (1999). Growth of new technology-based firms: Which factors matter? *Small Business Economics* 13 (2),141-154.

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

Audretsch, D. B. (1995). *Innovation and Industry Evolution*, Cambridge: MIT Press

Audretsch, D.B. (2012). Determinants of High-Growth Entrepreneurship, OECD/DBA Report. Available from: http://www.oecd.org/cfe/leed/Audretsch_determinants%20of%20high-growth%20firms.pdf. (Accessed 1 November 2013).

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

Baum, J.A.C., & Silverman. B.S. (2004). 'Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology start-ups'. *Journal of Business Venturing* 19, 411–436.

Bloom, N. and J. van Reenen. 2006. *Measuring and Explaining Management Practices across Firms and Countries.* London: Centre for Economic Performance.

Bottazzi, G. & Secchi, A. (2006). Explaining the distribution of firm growth rates. *The RAND Journal of Economics* 37 (2), 235-256.

Brouwer, E., Kleinknecht, A., & Reijnen, J. O. N. (1993). Employment growth and innovation at the firm level. *Evolutionary Economics* 3, 153-159.

Brown, J. R., Fazzari, S. M., & Petersen, B. C. (2009). Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *The Journal of Finance* 64 (1), 151-185.

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

Christensen, C. M. 1997. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Boston: Harvard Business Press.

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

Colombelli, A., Krafft, J., & Quatraro, F. (2012). *High Growth Firms and Technological Knowledge: Do gazelles follow exploration or exploitation strategies?* Available from: http://halshs.archives-ouvertes.fr/hal-00666707/. Accessed 1 November 2013

Danzon, P.M., Epstein, A., & Nicholson, S. (2007). Mergers and acquisitions in the pharmaceutical and biotech industries. *Managerial and Decision Economics* 28(4-5), 307-328.Davidsson, P., & Delmar, F. (2006). High-growth firms and their contribution to employment: The case of Sweden. In Davidsson, P., Delmar, F., and Wiklund, J. (Eds.), *Entrepreneurship and the Growth of Firms* (pp. 156-178), Cheltenham, UK: Edward Elgar.

Delmar, F., Davidsson, P., and Gartner, W. (2003). Arriving at the high-growth firm. *Journal of Business Venturing*, 18(2), 189-216.

Del Monte, A., & Papagni, E. (2003). R&D and the growth of firms: empirical analysis of a panel of Italian firms. Research Policy 32 (6), 1003-1014.

Demirel, P., & Mazzucato, M. (2012). Innovation and firm growth: Is R&D worth it? *Industry and Innovation*, 19(1), 45-62.

Dosi, G., & Mazzucato, M. (2006). *Knowledge Accumulation and Industry Evolution. The Case of Pharma-Biotech*. New York: Cambridge University Press.

Ejermo, O., Kander, A., & Henning, M.S. (2011). The R&D-growth paradox arises in fast-growing sectors. *Research Policy* 40, 664–672.

Evans, D.S. (1987). Tests of alternative theories of firm growth. *The Journal of Political Economy* 95 (4), 657-674.

European Commission (2010). *Europe 2020: A strategy for smart, sustainable and inclusive growth*. Available from: http://ec.europa.eu/eu2020/pdf/COMPLET%20EN%20BARROSO%20%20%20007%20-%20Europe%202020%20-%20EN%20version.pdf. Accessed 5 May 2013.

Freel, M., & Robson, P. (2004). Small firm innovation, growth and performance. International Small Business Journal, 22(6), 561–575.

Freeman, C. 1995. 'The "National System of Innovation" in Historical Perspective'. *Cambridge Journal of Economics* 19, no. 1: 5–24.

Grabowski, H., & Vernon, J., (1987). Pioneers, imitators, and generics: a simulation model of Schumpeterian competition. *Quarterly Journal of Economics* 102(3), 491-525.

Gambardella, A. (1995). *Science and Innovation in the US Pharmaceutical Industry*. Cambridge: Cambridge University Press.

Geroski, P.A., Van Reenen, J., & Walters, C.F. (1997). How persistently do firms innovate? *Research Policy* 26(1), pp.33-48.

Gibrat, R. (1931). *Les Inégalités Economiques*. Librairie du Recueil Sirey, Paris.

Gompers, P.A., & Lerner, J. (2001). The venture capital revolution. *Journal of Economic Perspectives* 15(2), 145-168.

Gort, M. and Klepper, S. (1982). Time paths in the diffusion of product innovations. *Economic Journal,* Royal Economic Society, 92(367): : 630-53.

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

Hall, B.H. (1987). The relationship between firm size and firm growth in the US manufacturing sector. *Journal of Industrial Economics* 35(4), 583-600.

Hall, B.H., Jaffe, A.B., & Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools. NBER Working Papers 8498, National Bureau of Economic Research.

Hall, B.H., & Mairesse, J. (1995). Exploring the relationship between R&D and productivity in French manufacturing firms. *Journal of Econometrics* 65, 263-293.

Hall, M., & Tideman, N. (1967). Measures of concentration. *Journal of the American Statistical Association* 62 (317), 162-168.

Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics*, 95(2), 347-361.Hart, P.E., & Oulton, N. (1996). The growth and size of firms. *The Economic Journal* 106(3), 1242-1252.

Henrekson, M., & Johansson, D. (2010). Gazelles as job creators: a survey and interpretation of the evidence. *Small Business Economics* 35, 227–244

Hughes, A. 2008. 'Entrepreneurship and Innovation Policy: Retrospect and Prospect'. *Political Quarterly* 79, issue supplement s1 (September): 133–52.

Hymer, S., & Pashigian, P. (1962). Firm size and rate of growth. *Journal of Political Economy* 70(6), 556-569.

IMS. (2007). *IMS health reports global pharmaceutical market Grew 7.0 percent in 2006, to $643 Billion*, Press Release, March 20th. http://www.imshealth.com/media. Accessed 10 April 2010.

Kirchhoff, B. A. (1994). Entrepreneurship and Dynamic Capitalism the Economics of Business Firm Formation and Growth, Westport, CT: Greenwood Publishing.

Koenker, R.W. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*. 91(1), 74–89.

Klepper, S. (1997). Industry Life Cycles. Industrial and Corporate Change, 6(1): 145-182

Koenker, R.W., & Basset, G.J.(1978). Regression quantiles. *Econometrica* 46(1), 33–50.

Lang, G. (2009). Measuring the returns of R&D – An empirical study of the German manufacturing sector over 45 years. *Research Policy* 38, 1438–1445.

Mansfield, E. (1962). Entry, Gibrat's law, innovation, and the growth of firms. *The American Economic Review* 52(5), 1023-1051.

Mowery, D.C., & Ziedonis, A.A. (2002). Academic patent quality and quantity before and after the Bayh–Dole Act in the United States. *Research Policy* 31(3), 399-418.

Mason, C. and Brown, R. (2013). Creating good public policy to support high-growth firms, *Small Business Economics*, 40, 211-225.

Mazzucato, M. (2002). The PC industry: New economy or early life-cycle. *Review of Economic Dynamics* 5 (2), 318-345.

Mazzucato, M. (2013). *The Entrepreneurial State*. London: Anthem. ISBN 978-0-85728-252-1.

Nelson, R. (1991). Why do firms differ, and how does it matter? *Strategic Management Journal* 12, 61-74.

Nelson, R.R., & Winter, S.G. (1982). *An Evolutionary Theory of Economic Change*. Cambridge, Mass :Bellknap.

NESTA. (2009). *The vital 6 per cent, how high growth innovative businesses generate prosperity and jobs, Research Summary*. October 2009. Available from: http://www.nesta.org.uk/publications/reports/assets/features/the_vital_6_per_cent. Accessed 17 January 2012.

NESTA (2011). *Vital Growth,* March 2011. Available from: http://www.nesta.org.uk/publications/reports/assets/features/vital_growth. Accessed 17 January 2012.

Nightingale, P. (2000). Economies of scale in experimentation: knowledge and technology in pharmaceutical R&D. *Industrial and Corporate Change*, 9 (2), 315-359.

Nightingale, P., & Coad, A. (2011) *Muppets and Gazelles: Ideological and methodological biases in entrepreneurship research*. FINNOV Discussion paper. http://www.finnov-fp7.eu/sites/default/files/FINNOV_DP8.5.pdf. Accessed 17 January 2012.

Nunes, P.M., Serrasqueiro, Z., & Leitão, J. (2012). Is there a linear relationship between R&D intensity and growth? Empirical evidence of non-high-tech vs. high-tech SMEs. *Research Policy* 41, 36–53.

OECD (2010), "The OECD Innovation Strategy: Getting a Head Start on Tomorrow", ISBN 978-92-64-08470-4, http://www.oecd.org/sti/theoecdinnovationstrategygettingaheadstartontomorrow.htm.. Accessed 20 January 2012.

Orsenigo, L., Pammolli, F., & Riccaboni, M. (2001). Technological change and network dynamics. *Research Policy* 30 (3), 485-508.

Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R., & Schacht, A. L. (2010). How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery* 9 (3), 203-214.

Porter, M.E. (1985). *Competitive Advantage*. New York: Free Press.

Qian, G., & Lee, L. (2003). Profitability of small and medium sized enterprises in high-tech industries: the case of the biotechnology industry. *Strategic Management Journal* 24, 881–887.

Scherer, F. M. (1987). *Market Structure.* Amsterdam: The New Palgrave.

Scherer, F. M., & Ross, D. (1990). *Industrial Market Structure and Economic Performance*. Boston: Houghton Mifflin Company.

Scannell, J.W., Blanckley, A., Boldon, H., & Warrington, B. (2012). Diagnosing the decline in pharmaceutical R&D efficiency, *Nature Review of Drug Discovery* 11, 191–200.

Schumpeter, J. (1934). Capitalism, socialism, and democracy. New York: Harper & Row.

Shane, S., 2008. *The Illusions of Entrepreneurship: the Costly Myths That Entrepreneurs, Investors, and Policy Makers Live By*. New Haven: Yale University Press.

Shane, S. (2009). Why encouraging more people to become entrepreneurs is bad public policy, *Small Business Economics*, 33(2), 141-149.Singh, A., &Whittington, G. (1975). The size and growth of firms. *Review of Economic Studies* 42(1), 15-26.

Stam, E., and Wennberg, K. (2009). The roles of R&D in new firm growth. *Small Business Economics*, 33, 77-89.

Storey, D.J. (1994). *Understanding the Small Business Sector*. London: Routledge.

Verspagen, B. 2005. "Innovation and Economic Growth". In: J. Fagerberg; D.C. Mowery; and R.R. Nelson, 2005, *The Oxford Handbook of Innovation*, Oxford: Oxford University Press.

Veugelers, R., & Cincera, M. (2010). *Europe's missing Yollies*. Bruegel Discussion Paper, 2010/06. http://www.bruegel.org/publications/publication-detail/publication/430-europes-missing-yollies/. Accessed 22 January 2013.

Online only Appendix: **Electronic Supplementary Material**

Penalized fixed effect quantile regression

Koenker (2004) penalized fixed effect quantile regression is implemented by estimating the coefficients, $\beta(\tau_k)$ of each independent variable (i.e. the partial derivative of the conditional quantile of growth (gr) with respect to a particular independent variable), at quantiles τ_k, where k indexes the quantiles at 10%, 25%, 50%, 75%, 80%, 90%. Estimation is implemented using Koenker (2004) penalised fixed effect quantile regression, minimising Expression 1:

Expression 1: Penalised fixed effect quantile regression estimator

$$\min_{\alpha,\beta} \sum_{k=1}^{q} \sum_{i=1}^{n} \sum_{t=1}^{t_i} w_k \rho_{\tau_k} \big(gr_{it} - \alpha_i - X'_{it}\beta(\tau_k) \big) + \lambda \sum_{i=1}^{n} |\alpha_i|$$

Where i is an index for the n firms, and j the observations for each firm and $\rho_\tau(u) = u(\tau - I(u < 0))$, is the piecewise linear quantile loss function used for each quantile regression as described by Koenker and Bassett (1978).

The penalized fixed effect estimator produces slope coefficient estimates for each selected quantile, but the fixed effect for each firm α_i are independent of quantiles, thereby avoiding estimating fixed effects for each different quantile regression. Estimation is performed by minimising the expression selecting w_k weights for each of the q quantiles which control for the influence of quantiles on the estimation of α_i and applying a penalty term to penalise the vectors of fixed effect coefficients, shrinking these coefficients toward zero. The importance of the penalty term is controlled by selecting a value of λ (Koenker, 2004). We set $\lambda = 1$ and select an even weighting of w_k across quantiles. Varying the choice of weights made only minor changes to the reported results. For a detailed discussion of this method see Koenker (2004). Standard errors are constructed via a generalized bootstrap with unit exponential weights sampled for each firm.

For comparison we also report regression results for OLS and least squares (within) fixed effect estimators with robust errors using Huber-White sandwich estimators. We also run standard quantile regression estimation using Koenker and Basset (1978) method, but preferring bootstrapped standard errors with 1000 repetitions due to the presence of heteroskadasticity.

Recent papers in the SPRU Working Paper Series:

SWPS 2013-05. Fulvio Castellacci. September 2013. "Institutional voids or entry barriers? Business groups, innovation and market development in emerging economies".

SWPS 2013-06. Tommaso Ciarli and Michele Di Maio. October 2013. "Theoretical arguments for industrialisation-driven growth and economic development".

SWPS 2013-07. William Lazonick, Mariana Mazzucato, and Öner Tulum. October 2013. "Apple's changing business model: what should the world's richest company do with all those profits?".

SWPS 2013-08. Phil Johnstone. October 2013. "Planning reform, rescaling, and the construction of the post-political: The case of The Planning Act 2008 and nuclear power consultation in the UK".

SWPS 2013-09. Mariana Mazzucato and Alan Shipman. October 2013. "Accounting for productive investment and value creation".

SWPS 2013-10. Anders Bornhäll, Sven-Olov Daunfeldt, and Niklas Rudholm. October 2013. "Sleeping Gazelles: High profits but no growth".

SWPS 2013-11. Alex Coad and Christina Guenther. November 2013. "Processes of firm growth and diversification: theory and evidence".

SWPS 2013-12. Sally Brooks. November 2013. "Investing in Food Security? Philanthrocapitalism, Biotechnology and Development".

SWPS 2013-13. Vicente Salas-Fumás, J. Javier Sanchez-Asin, and David Storey. November 2013. "Entrepreneurs and Output: Theory and Empirical Evidence with Spanish Data".

SWPS 2013-14. Diego Chavarro, Puay Tang, and Ismael Rafols. November 2013. "Interdisciplinarity and research on local issues: evidence from a developing country".

SWPS 2013-15. Paul Nightingale. November 2013. "Schumpeter's Theological Roots? Harnack and the origins of creative destruction".