Innovation and patterns in extreme firm growth and decline

Matin Khadem & Paul Kattuman
University of Cambridge

Abstract

A noteworthy stylized fact that has emerged recently in the literature on firm growth is that both among firms on the upper tail and firms on the lower tail of the short run firm growth rate distribution, from one year to the next, some persist, while significant proportions bounce to the opposite tail. We find that innovation plays an important role in generating this pattern in firm growth. Innovation, in terms of its component stages, increase both the probability of sustaining or rebounding to extreme positive growth, and the probability of sustaining or bouncing down to extreme negative decline, with the former dominating the latter. Potential rewards of innovation are highest for firms that are growing most rapidly. These findings have important implications for innovation policy.

1. Introduction

It is universally acknowledged that innovation is crucially important for the long-term fortunes of countries. The Schumpeterian theory of creative destruction, as also other industrial organisation theories, emphasise the central role of firm level innovation in national economic fortunes (Aghion and Howitt, 1992). The causal effectiveness of innovation policy continues to attract a large amount of empirical research, not least at the behest of policy makers who seek to support high-growth firms (for example see European Commission, 2011, p. 8). However, empirical research has struggled to unearth a compellingly robust relationship between innovation and firm growth.

In the quest to bridge the gap between theory and empirics, a recent branch in the investigation has focused on whether the growth effect of innovation is conditioned on the rate at which the innovating firm has been growing. The robust finding is that innovation does have significantly larger effects on growth at the upper quantiles of the growth rate distribution (Capasso et al., 2015; Coad and Rao, 2008; Falk, 2012; Segarra and Teruel, 2014). At the same time, another stylized fact that has emerged from work by Capasso et al. (2014) is the coexistence of persisting and 'bouncing' (or more accurately, 'recoiling') firms. Striking patterns of extreme 'growth' over the short-term have been identified using quantile regressions and transition probabilities based trajectory analyses. The authors highlight the uncertainty intrinsic to the innovation process as a potential source of the extreme patterns in growth. This is in line with the notion of a Schumpeterian Mark I [regime] (Capasso et al., 2014, p. 1023): innovation effort may generate commercially viable knowledge, but may also fail to do so. However, the authors do not attempt to identify the determinants that place firms on the very different extreme growth paths.

In this paper we study the role of the innovation process in generating the striking heterogeneity found among firms at the extremes of the short-run firm growth rate distribution, comprising both persistence and recoil in growth rates. To do this we disaggregate the innovation process in terms of its stages: R&D expenditure, the resulting output in the form of patents, and the value of this output in terms of citations achieved; and evaluate the importance of each of these stages in the innovation process in contributing to the above described firm growth patterns.

The next section presents a review of the literature related to innovation and extreme growth of firms. Building on this, section 3 presents hypotheses on the relationship between firm growth and the various stages of innovation. Section 4 describes the data we analyse. Section 5 outlines the econometric approach. Section 6 presents results and Section 7 concludes.

2. Literature Review

In the empirical literature, R&D is consistently found to impact the growth of economies positively (Zachariadis, 2003); but a robust link between R&D and firm level growth has been elusive, even when conditioning the effect of R&D on firm characteristics such as age, size and industry (Coad et al., 2016). Recent work has gone further, examining whether the impact of innovation is conditional on firm growth rate itself (Coad and Rao, 2008).

One motivation for our study is the empirical finding that innovation tends to have a larger positive impact on growth as a firm moves up the conditional growth distribution (Capasso et al., 2015; Coad and Rao, 2008; Coad et al., 2016; Falk, 2012). The effect is not linear; it pertains only to the fastest growers, for whom it constitutes an important driver for their high-growth status. The effect of innovation on firms at lower quantiles of the growth rate distribution is generally insignificant (Falk, 2012) or even negative (Coad et al., 2016). One explanation for this finding draws on an argument from Freel (2000, p. 208) that R&D investment is inherently uncertain, and at some level it has the nature of a 'bet', in that it is not guaranteed to produce valuable knowledge. In this broad sense, three sub-classes of innovation can be conceived: 'tried and succeeded', 'tried and failed' and 'not tried'. For the second among these categories, innovation constitutes a waste of resources and hence it is likely that higher innovation actually worsens the growth performance for such firms, explaining the negative coefficients on the lowest quantiles found by Coad and Rao (2008) and Coad et al. (2016).

A related motivation for our study flows from growth patterns observed among fast growing firms and declining firms by Capasso et al. (2014)[1]. They highlight the two sub-populations that coexist, at both ends of the growth rate distribution: recoiling firms, whose growth rates jump from one extreme to the other over the short run; and persistent firms, who maintain their extreme growth performance from one year to the next, whether growing fast or declining. The above mentioned study does not offer an explanation for these strikingly heterogeneous growth paths. Indeed, the few papers that directly explore the relationship between persistently outperforming firms and innovation have generally reported no discernible link. Bianchini et al. (2017) attempt to identify the characteristics that distinguish persistently high-growing firms (over at least four of five years), from one-off high-growth firms and find that the usual drivers of growth including innovation are not relevant to persistent outperformance, supporting a 'mere luck' conjecture. Similarly, Guarascio and Tamagni (2016) find that persistent innovators do not exhibit higher growth persistence. These papers focus on the long-term association whereas the concept of persistence that

[1]See also Coad and Hölzl (2009); Hölzl (2014).

we consider in this paper is akin to Capasso et al. (2014) and Coad and Hölzl (2009), and pertain to the short-run.

In the management literature, a discussion of how R&D may vary conditioned on firm performance draws on the behavioral theory of the firm to offer an explanation for why R&D expenditure may be driven up relatively more at either end of the performance range (Greve, 2003). At the poor performance end, falling short of aspiration can stimulate R&D aimed at the perceived causes of poor performance. Managers are likely to consider any performance below their aspirational level as a loss situation, leading them to greater tolerance for risk in the quest for improvement (Kahneman and Tversky, 1979). In studies of organizational change and risk taking, greater risk taking has been found to accompany such problemistic search (Bolton, 1993; Bromiley, 1991; Greve, 1998; Grinyer and McKiernan, 1990; Miller and Leiblein, 1996).

At the outperformance end, insofar as good performance generates slack in resources, "innovations that would not be approved in the face of scarcity but have strong subunit [e.g., R&D department] support" are more readily accommodated (Cyert and March, 1963, p. 208). Such firms will have greater latitude for experimentation and for organizational change, which, when combined with less strict performance monitoring, affords both the resources and managerial patience that enable risky innovation effort.

These arguments offer theoretical bases for why firms at the extremes of the growth rate distribution may differ from firms in the middle range of growth in the quantum of R&D effort they choose to undertake. Combined with the inherent uncertainty underlying R&D outcomes, a plausible explanation for the co-existence of the recoiling firms and persistent firms at both ends of the growth distribution emerges.

3. Hypotheses

On the basis of the above discussion, we formulate a set of hypotheses concerning the effects of innovation on firm growth. In Section 5, we set out to test each hypothesis empirically.

3.1. Baseline

We begin by examining whether the effect of innovation on growth is positive for the representative firm, in line with past findings such as Coad and Rao (2008) and Falk (2012):

Hypothesis 1a: On average, greater input into the innovation process (R&D expenditure) translates into a higher growth rate for the firm, controlling for the productivity of such innovation input.

Turning from the effect of input to that of outcome, we next examine whether innovation output and its quality contribute positively to firm growth, controlling for the input into the innovation process. The rationale is the obvious one that growth value of innovation input should be less uncertain once a tangible outcome of innovation in terms of quantity and quality are observed:

Hypothesis 1b: On average, the greater the magnitude of innovation output, higher the growth rate, controlling for the level of innovation input.

Hypothesis 1c: On average, the greater the quality of innovation output, higher the growth rate, controlling for the level of innovation input.

3.2. Dynamics at the extremes

Before turning to the role of innovation in generating extreme growth patterns at the extremes, we explore whether the patterns of persistence and recoil found by Capasso et al. (2014) are evident in our sample:

Hypothesis 2 - Persistent growth: The probabilities that firms with growth rates at either extreme of the range of growth rates continue in their respective growth categories in the following period are higher than the probability that a firm in the middle range of growth rates experience an extreme growth event in the following period.

Hypothesis 3 - Recoiling growth: The probability that a firm at the underperforming extreme of the range of growth rates bounces up to the outperforming extreme in the following period; and the probability that a firm at the outperforming extreme bounces down to the underperforming extreme in the following period, are both higher than the probability that a firm in the middle range of growth rates experience an extreme growth event in the following period.

3.3. Heterogeneity in innovation outcome

We now turn to the characterisation of innovation as a "bet", to see if the pattern of higher input into innovation being associated with *both* a higher probability of ending up in the right *and* left tails of the growth distribution can be explained in the following terms:

Hypothesis 4a: On average, greater input into the innovation process increases *both* the probability of underperformance and outperformance in the following period.

It is reasonable to expect that, controlling for innovation input, greater innovation output and higher quality signals higher R&D productivity, and reduces the uncertainty of growth payoff from R&D investment. Greater innovation output and higher quality of such output must improve growth prospects of the firm at all points of the growth rate distribution:

Hypothesis 4b: On average, greater innovation output reduces the probability of underperformance and raises the probability of outperformance in the subsequent period, controlling for the level of input.

Hypothesis 4c: On average, more valuable innovation output reduces the probability of underperformance and raises the probability of outperformance in the subsequent period, controlling for the level of input.

Hypothesis 4a relates to the heterogeneity in growth outcomes that result even when innovation input is the same; to similar firms experiencing divergent growth paths due to the uncertainty inherent in the innovation process. Heterogeneity may also be across subpopulations of firms: R&D and patenting may be riskier or more rewarding for different types of firms. For example, if previously outperforming firms have a stronger R&D culture on average, then it is plausible that the increase in the probability of outperformance in the next period resulting from an increase in R&D may be higher for the average top tier firm compared to a middle tier firm. This possibility is explored in Section 5.

4. Data

Our sample is constructed by matching firm level data from Compustat with the NBER patent database[2] which has information on all USPTO patents granted between 1970 and 2006. The sample is restricted to North American Compustat firms with positive sales and R&D expenditures.[3] Firm-years in which the firm has been engaged in a merger or an acquisition have been excluded, as also the entry and exit years for the firm. For comparability

[2] The version used can be found at: `https://sites.google.com/site/patentdataproject/Home`

[3] Excluding observations with zero R&D mitigates the issue of misreporting (Capasso et al., 2015).

with existing literature, we limit the analysis to manufacturing firms (SIC codes 2000-3999).

We follow the matching process outlined in Hall et al. (2001) to extract the number of patents and forward citations received each year for each firm from the NBER database, resulting in a final sample which contains both patenters and non-patenters. The NBER database suffers from two main truncation issues. The first pertains to patent counts and arises from the fact that the average lag between patent application and grant is two years; hence the final few years of data only capture a fraction of patents that were eventually granted. Secondly, citation counts for patents in later years suffer in a similar way since forward citations typically pick up over lengthy durations. To mitigate these problems we limit the analysis, insofar as innovation output is concerned, to patents that were applied for in years up until and including 2000. In terms of the number of patents per year, Coad and Rao (2008, p. 639) find evidence of a clear structural break for certain sectors at the beginning of the 1980s which they suggest may be a result of reforms to patent regulations. This leads us to limit our analysis to the period after 1980. The final sample consists of an unbalanced panel of 3636 firms that meet the above restrictions and have data for at least two consecutive years between 1981 and 2000. To investigate how the effect of innovation has changed over time, and to check the robustness of our results we also split the overall duration into four 5-year sub-periods and replicate the analysis.

4.1. Descriptive Statistics

In what follows, we adopt the terminology used by Capasso et al. (2014) and use 'underperformer' to refer to firms in the bottom 10% of the growth rate distribution, and 'outperformer' or 'high-growth' to refer to firms in the top 10% of the growth rate distribution in the given year. Table 1 gives the summary statistics for the key variables categorised by their location in the sales growth rate distribution (bottom 10%, middle 80% and top 10% of firms in ascending order of growth rate) for the full sample of 27,137 firm years over the period 1981-2000.

Firms at both extremes of the growth distribution have significantly higher R&D to sales ratios on average compared to firms in the middle 80%; this is seen in Figures 1 and 2 which present the kernel densities of the logarithm of the ratio of R&D to sales for each growth category, and average ratios of R&D to sales by growth quantile. Kolmogorov-Smirnov tests for equality of distributions between each extreme sub-sample and the middle 80% sub-sample reject equality at the 1% level in both cases. Insofar as R&D effort poses both risk as well as reward with in terms of growth, these descriptive findings offer a preliminary indication that differences in inputs into innovation may have a bearing on the polarised

Table 1: Descriptive statistics by growth outcome (1981-2000)

	Mean	Median	p25	p75	S.D.	# Obs
Bottom 10%						
Public Age	14.62	11	7	19	10.59	2,725
Total Sales	793.7	27.69	7.235	143.4	4,888	2,725
Citations-per-Patent*	11.81	8	4	14.25	14.03	1,007
R&D to Sales	1.650	0.0785	0.0223	0.226	7.607	2,725
Patents to R&D	0.264	0	0	0.199	0.749	2,725
Middle 80%						
Public Age	19.29	16	9	28	12.62	21,687
Total Sales	2,704	202.7	44.42	1,105	11,692	21,687
Citations-per-Patent*	11.63	8.592	4.667	14.08	12.60	11,341
R&D to Sales	0.125	0.0340	0.0132	0.0803	0.888	21,687
Patents to R&D	0.324	0.0393	0	0.339	0.760	21,687
Top 10%						
Public Age	9.568	7	5	11	7.489	2,725
Total Sales	567.6	43.44	10.37	193.4	3,401	2,725
Citations-per-Patent*	17.69	10.50	4.750	22	22.55	1,157
R&D to Sales	0.613	0.0866	0.0307	0.237	2.262	2,725
Patents to R&D	0.318	0	0	0.210	0.893	2,725

Note: Total Sales and R&D expenditure in deflated ($million)

* Of patenting firms

growth patterns of firms at the extremes of the growth rate distribution. It is also notable that high-growth firms have, on average, significantly higher citations-per-patent than others.

5. Method

We borrow the broad analytical framework of the structural CDM model outlined by Crepon et al. (1998), which has been a workhorse in empirical research on firm level innovation over the past 20 years (Lööf et al., 2016). In its original application, the model was used to understand how research investment translates into innovation output, which in turn translates into increased productivity. Our focus being on firm performance, it is important to distinguish between alternative dimensions of firm performance. Improved productivity from innovation is likely to translate into higher sales growth on average; but the effect on employment growth could be negative if innovation takes the form of labour-saving process innovation. While employment growth is also of policy interest, sales growth has been

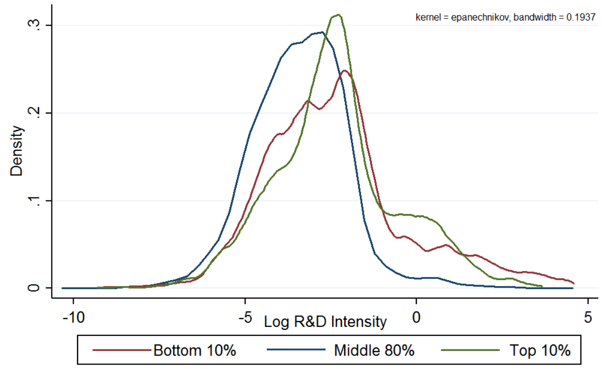

Fig. 1. Kernel density of Log R&D to Sales by growth outcome (1981-2000)

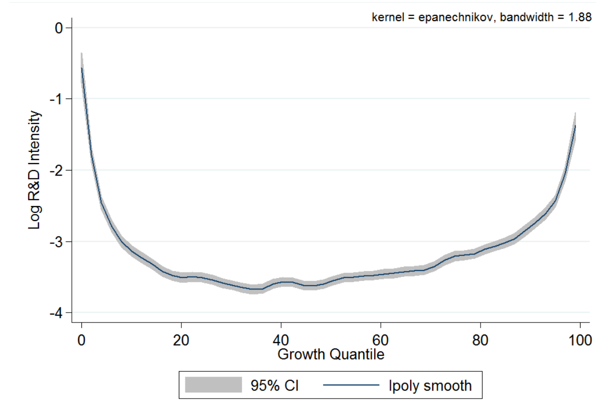

Fig. 2. Kernel smoothing of Log R&D to Sales by growth quantile (1981-2000)

highlighted as the primary channel through which innovation improves firm profits, and as a particularly meaningful indicator of post-innovation performance (Geroski and Machin, 1992; Scherer, 1965; Coad and Rao, 2008). We follow the latter in measuring firm size in terms of sales and firm growth rate as the difference of log sales.

In the most general form, the data-generating process may be written as:

$$Growth_{it} = \alpha + \beta_{1it-1} InnovationFlow'_{it-1} + \boldsymbol{X}'_{it-1}\boldsymbol{\beta}_{2it-1} + \mu_i + \theta_t + \varepsilon_{it} \tag{1}$$

Where $InnovationFlow$ is a measure of the addition to the knowledge stock generated in year $t-1$ by firm i, and $\boldsymbol{X}$ is a vector of other explanatory variables including age, size and industry. This unrestricted form, with parameter values allowed to vary between firms and time periods, cannot be estimated with the available data, but it serves as a guide for introducing heterogeneous effects.

5.1. Choice of innovation measures

To bring the innovation process to data, we use an empirical specification which adheres to the conceptual framework outlined in Hall et al. (2005). In this study, the knowledge creation process begins with R&D expenditure as an input, which translates with uncertainty into codifiable knowledge in the form of patents, which may be considered a quantity measure of the output of the innovation process. Though widely used as a proxy for innovation output, patent counts have two distinct drawbacks. First, they do not capture all new knowledge generated from research efforts. This is because there are firms who choose not to patent innovations, opting instead for secrecy to protect their intellectual property. The second drawback is that patents can differ hugely in economic value. To overcome this second issue, we use the number of forward citations for any given patent as a measure of its 'importance', and thus a proxy for the value of the innovation output (Akcigit and Kerr, 2016; Akcigit et al., 2016).[4]

However, including citations directly in the model would censor the sample to only firms that have patented successfully. We solve this problem by conducting two parallel analyses: one which includes R&D and an unweighted patent count measure, and another which includes R&D and a measure where patents are weighted by the number of forward citations received. We follow Trajtenberg (1990) in using a linear weighting scheme in which a value of one is assigned to each patent and each citation. The citation counts are summed to give the 'weighted patent count' for each firm for each year:

$$WPC_{it} = \sum_{k=0}^{K} (C_k + 1) \tag{2}$$

Where K is the total number of patents received by firm i at time t and C_k gives the total number of citations subsequently received by patent k.

While the weighted patent count does not have a rigorous theoretical justification,[5] it nevertheless has the useful feature of being able to distinguish between three kinds of firms: those that successfully patented and in doing so generated commercially valuable knowledge (large WPC), those that successfully patented yet produced patents of little or no value (non-zero but small WPC), and firms that were unsuccessful in translating R&D into patents (zero WPC). In contrast, the unweighted patent count (UPC) fails to discriminate between the first two. Due to high collinearity between UPC and WPC, we do not include both simultaneously in the same regression, but run separate regressions. For brevity, we do not present separate tables for each, and instead report both sets of estimates in the same table sequentially. The reported estimates for the R&D intensity are from the regression that includes WPC.[6]

Table 2 describes the relevant variables. Given the focus on short-term growth patterns, we follow the extant literature[7] in using the logarithm of deflated R&D expenditure divided by deflated sales in the previous period as the R&D intensity measure, whose first lag enters the model. As explained by Hall et al. (2005, p. 17), a specification that already includes R&D intensity implies that the additional value of patents is captured by the number of

[4] A number of studies have directly investigated the relationship between citation counts and patent value and found a robust link (Harhoff et al., 1999; Trajtenberg, 1990).

[5] See Shane and Klock (1997, p. 137) for an exposition and alternative weighting scheme.

[6] Estimates for R&D from regressions including UPC are very similar.

[7] For example, see Capasso et al. (2015); Falk (2012); Segarra and Teruel (2014) for similar specifications.

patents per unit ($million) of R&D intensity. As such, our patent measures enter into the regression models as the first lags of the respective patent counts divided by deflated R&D expenditure, with the names weighted and unweighted patent yields (WPY and UPY hereafter). Patents are counted in the year in which they were applied for, not the year in which they were granted; this provides a rationale for including them contemporaneously with R&D. While there is some gestation lag between the outlay of R&D expenditure and the application for patents, it is significantly shorter than the average lag between the application for the patent and its grant, with Trajtenberg (1990, p.183) commenting that 'there is a strong statistical association between patents and R&D expenditure; this relationship appears to be mostly contemporaneous... supporting previous findings of short gestation lags'.

Graphing sales growth rate against patent counts suggests a positive yet diminishing marginal effect of patents on growth; based on this, we accommodate potential non-linearity in the effect of patent yields through a squared term.[8]

Table 2: Variable definitions

Variable	Definition
GrSales	Annual growth rate calculated by taking the difference of logs of firm size as measured by deflated sales in $million
RDIntensity	Logarithm of the ratio of deflated R&D expenditure in $million to lagged deflated sales in $million
UPY	Ratio of unweighted patent count to deflated R&D expenditure
WPY	Ratio of weighted patent count to deflated R&D expenditure
Logsize	Logarithm of deflated sales in $million
Logage	Logarithm of age in Compustat database

5.2. Linear Model

We specify the empirical model as:

$$\begin{aligned} GrSales_{it} = \beta_1\ RDIntensity_{it-1} + \beta_2\ PATY_{it-1} + \beta_3\ PATY^2_{it-1} \\ + \beta_4\ GrSales_{it-1} + \boldsymbol{Controls}'_{it-1}\boldsymbol{\beta}_5 + \mu_i + \theta_t + \varepsilon_{it} \end{aligned} \tag{3}$$

The above model is estimated twice; for $PATY = UPY$ and $PATY = WPY$.

[8]In further exercises, omission of the squared term left the pattern of results unchanged.

Insofar as R&D intensity measures input into the innovation process, and patents and citations measure the corresponding output and its quality, β_1 captures the growth effect of increased input into the innovation process, controlling for the productivity of R&D. This accommodates the fact that firms may vary considerably in their R&D productivity, and the omission of innovation output from the regressor set can be a source of bias for the R&D input coefficient – if more productive firms tend to do more R&D, higher productivity would be falsely attributed to higher R&D, biasing β_1 upwards. In the model that includes R&D intensity, β_2 and β_3 measure the growth effects of higher innovation output and its quality, for given input. These interpretations are in line with extant research.

We begin with hypothesis 1, and estimate the growth effect of innovation for the representative firm. To control for short run autocorrelation in growth, we include lagged growth rate as a regressor (Coad and Rao, 2008; Demirel and Mazzucato, 2012). All regressions include lagged age and size as controls, as well as year fixed effects (θ_t) to control for common macroeconomic shocks. OLS and GMM regressions also include industry dummies. Full sample regressions are complemented with analyses of four 5-year sub-periods – 1981-1985, 1986-1990, 1991-1995, 1996-2000 – to ensure that the results are not affected by structural breaks. All reported standard errors are robust to heteroscedasticity and clustered at the firm level.

System GMM estimation is the preferred specification as it is well suited to the 'small T, large N' setting. It also directly accounts for the dynamic nature of growth and can deal with endogeneity arising both from the 'feedback' effect of growth on R&D intensity and the presence of firm fixed effects which influence both R&D decisions and growth outcomes, as captured by μ_i. We also present pooled OLS, and fixed effects estimates but note that with fixed T, the within estimator is inconsistent if the data generating process is dynamic (Nickell, 1981).

We use a two-step system GMM procedure[9] using first-differencing and Windmeijer (2005) corrected cluster-robust standard errors. Coad et al. (2016) highlight the possibility of endogeneity arising from feedback effects between growth and innovation, although it has been suggested by Capasso et al. (2015, p. 49) that the influence of this channel is more pertinent to long-term analyses. Nevertheless, we follow Demirel and Mazzucato (2012) in treating all innovation measures as endogenous, instrumenting each with their lagged values in the GMM estimation.

[9] Near random walk nature of growth rates can lead to weak instruments and hence inconsistency when using difference GMM, especially as T is small (Blundell and Bond, 1998). Stationarity, which is required to satisfy the necessary moment conditions was confirmed through a series of Harris-Tzavalis (1999) panel unit root tests on sales growth rates for each sub-period, which rejected the presence of a unit root at the 1% level.

5.3. *Multinomial Logit*

Turning to the role of innovation in the observed persistence and recoil behaviour reported in Capasso et al. (2014), we estimate a multinomial 'choice' model with three growth outcomes:

$$CatGrSales_{it} = \begin{cases} 0, & \text{if the growth rate of firm } i \text{ is in the bottom 10\% at time } t \\ 1, & \text{if the growth rate of firm } i \text{ is in the middle 80\% at time } t \\ 2, & \text{if the growth rate of firm } i \text{ is in the top 10\% at time } t \end{cases} \tag{4}$$

The parsimony in the number of outcome categories is in order to focus on the drivers of growth at the extremes. Discrete choice models are suggested as an alternative to conditional quantile regressions by Koenker and Hallock (2001, p. 148). With the multinomial specification, partial effects have a natural interpretation as changes in the probability of ending up in different parts of the conditional growth distribution, and also permits comparisons with the transition probability matrices in Capasso et al. (2014).

We estimate the multinomial model with the same set of regressors as in the linear model, and address hypotheses 2 and 3. To test hypothesis 4 we determine the average partial effects of innovation – the average increase in the probability that a firm experiences underperformance or outperformance in the following period associated with a 1% increase in R&D intensity. Finally, to elucidate polarisation in growth behaviour at the extremes, we consider how average partial effects differ based on past growth performance. Inference is based on bootstrapped standard errors (100 replications) with clustering at the firm level. In all results, the middle 80% is treated as the base outcome.

The choice between the multinomial logit (MNL) and probit (MNP, in which errors may be correlated and heteroscedastic) comes down to the distribution of the error term: type-1 extreme value, i.i.d. vs. multivariate normal. The nature of the data precludes estimation of an unrestricted MNP – for any given firm-year we observe only one growth outcome, and no 'alternative-specific' variables, in the data. The (unrestricted) MNP relaxes the independence of irrelevant alternatives (IIA) assumption, and allows for temporal correlation of the errors. In our case, the Small-Hsiao tests consistently failed to reject IIA at the 10% level across all sub-periods. However, independence also requires unobserved variation in the growth rate to be independent over time. In general terms this may not be unreasonable, as growth rates are known to be best approximated by a random walk, in the spirit of Gibrat's Law (Coad, 2009; Geroski, 1999).[10] As in the linear case, we include lagged growth rate as a regressor

[10]We re-ran all specifications using the independent MNP model (Kropko, 2007, p. 11) – these returned near identical partial effects as the MNL model.

in order to control for short run autocorrelation.

Notwithstanding the natural ordering of the dependent variable, we preferred the multinomial specification to the ordered specification, as the latter would constrain the effect of innovation to be equal across all growth outcomes: this would defeat the purpose of this study by obviating one of the central hypotheses to be tested. We therefore estimated the multinomial logit model via full maximum likelihood.

The probability that firm i experiences growth outcome j at time t, $Pr(CatGrSales_{it} = j)$, is given by equation 5 below:

$$\frac{e^{\alpha_j+\beta_{j,1}RDIntensity_{it-1}+\beta_{j,2}PATY_{it-1}+\beta_{j,3}PATY^2_{it-1}+\beta_{j,4}GrSales_{it-1}+\boldsymbol{Controls}'_{it-1}\boldsymbol{\beta}_{j,5}}}{\sum_{k=0}^{2} e^{\alpha_k+\beta_{k,1}RDIntensity_{it-1}+\beta_{k,2}PATY_{it-1}+\beta_{k,3}PATY^2_{it-1}+\beta_{k,4}GrSales_{it-1}+\boldsymbol{Controls}'_{it-1}\boldsymbol{\beta}_{k,5}}} \tag{5}$$

Controls include lagged age and size, as well as year and industry dummies. As with the linear model, we estimate the multinomial model separately: for $PATY = UPY$ and $PATY = WPY$.

The partial effect, the change in the probability of ending up in the top 10% at time t for firm i, of a 1% increase in that firm's R&D intensity at $t-1$, holding constant all other regressors is:

$$\frac{\partial Pr(CatGrSales_{it} = 2)}{\partial RDIntensity_{it-1}} = Pr(CatGrSales_{it} = 2)\left[\beta_{2,1} - \sum_{k=0}^{2} Pr(CatGrSales_{it} = k)\beta_{k,1}\right] \tag{6}$$

Because the slope parameters, β_ks, are alternative-specific, the above partial effects are not constrained to be the same across all outcomes. This accommodates the possibility that an increase in R&D intensity simultaneously raises the probability of ending up in the bottom 10% and in the top 10%. The average partial effect is found by evaluating equation 6 for each firm at their respective covariate values and taking the mean. As eluded to at the end of section 3, it is also possible that there are heterogeneous effects across different kinds of firms. The above outlined multinomial framework can be used to analyse this hypothesis by evaluating the average partial effect for subsets of the population to identify, for example, whether R&D is riskier and/or more rewarding conditional on past growth or on firm size.

We do not run a fixed effects multinomial logit regression due to the incidental parameters problem which, due to fixed T, leads to the estimates not being consistent with respect to the number of firms. Valid inference cannot be performed on any quantities that are a function of the fixed effects and thus the partial effects of innovation cannot be reliably estimated.

This is problematic because for the non-linear model it is precisely these partial effects which are of interest and not the coefficient estimates. That said, the closeness of the OLS, FE and GMM estimates are encouraging.

6. Results

6.1. Linear Model

Results for the linear model are presented in Table 3. The consistently positive and significant coefficient on R&D intensity across all specifications offers support for hypothesis 1a. This is in line with results in Demirel and Mazzucato (2012): R&D intensity retains a significantly positive impact on sales growth even after accounting for innovation output. There is some evidence in favour of hypotheses 1b and 1c, with a positive linear term and negative quadratic term suggesting positive but diminishing marginal effects of higher quantity and quality of patents on sales growth.

The similarity of the FE and GMM results reassures us that feedback effects are not substantial in the data, since the latter additionally instruments for endogeneity of the measures of innovation. Moreover the OLS coefficient on $RDintensity_{t-1}$ is a little over half of the magnitude of the FE estimate which is in line with findings by Coad and Rao (2008, p. 644). GMM sub-period results[11] to be found in Table A1 in the appendix suggest that these magnitudes, although consistently positive, have not been stable over time; the estimates suggest increasing average returns to R&D over the 20-year period. In their totality, these results offer strong evidence for innovation having a discernible positive effect on the growth of the representative firm; and for the quantity and quality of innovation output positively affecting firm growth over and above the effects of R&D expenditure.

6.2. Multinomial Logit

Evidence of persistence and recoil patterns (hypotheses 2 and 3) can be found in Table 4; if the probability of ending up in any particular decile in the next period was uniformly distributed for any given growth rate in the current period, then we would expect the values not to be statistically different from 0.1. Instead what we see is that firms who previously experienced extreme growth events have predicted probabilities that are significantly greater than 0.1 of having either an extreme positive or negative growth event at time t, and also larger predicted probabilities compared to firms previously in the middle 80%. In particular,

[11] FE and OLS sub-period results are available on request.

Table 3: Linear Model Results (1981-2000)

	OLS	FE	GMM
$GrSales_{t-1}$	0.0228	-0.0371**	-0.00544
	(0.0160)	(0.0171)	(0.0232)
$RDIntensity_{t-1}$	0.0396***	0.0731***	0.0623***
	(0.00288)	(0.00830)	(0.0160)
WPY_{t-1}	0.000935***	0.000528**	0.000814***
	(0.000153)	(0.000257)	(0.000240)
WPY^2_{t-1}	-6.36e-07***	-7.79e-07***	-6.18e-07***
	(1.21e-07)	(1.64e-07)	(1.67e-07)
UPY_{t-1}	0.0244***	0.0211***	0.0181**
	(0.00521)	(0.00653)	(0.00896)
UPY^2_{t-1}	-0.00175**	-0.00169*	-0.000687
	(0.000886)	(0.000926)	(0.00119)
Observations	27,137	27,137	27,137
R-squared	0.068	0.162	
# Instruments			74
Hansen OverID test			0.139
Arellano-Bond AR(2) test			0.682

All regressions include lagged age, size and year dummies.
OLS and GMM include industry dummies. Robust standard errors clustered at the firm level in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

persistence among outperformers stands out, with almost a quarter of outperforming firms coming from the group of firms whose growth rates were in the top 10% in the previous period. These patterns can be seen visually in Table A2 in the appendix which contains a heat map generated from sub-period results. Such results corroborate the TPM results in Capasso et al. (2014, p. 1029), which led them to conclude that: "if a firm, at a given year, experiences an extreme growth event, it is safe to say that the same firm is unlikely to be stable in the following year and can be expected to experience another extreme event".

We now turn to the extent to which innovation can explain these extreme growth patterns. To investigate hypothesis 4a, we present the average partial effect of $RDIntensity_{t-1}$ for all firms in Table 5, represented graphically in Figure 3. Sub-period results can be found in the appendix in Table A3. The pattern in Figure 3 points to the inherent uncertainty of R&D, with the average partial effects being positive and significant for extreme outcomes at both ends at time t. On average, higher R&D intensity is associated with increases in both the probability of extreme positive growth and the probability of extreme negative growth - although the probability of the latter increases by less. Combining this finding with the descriptive results presented in Section 2 that show that firms experiencing extreme growth events tend to have a higher R&D intensity, a possible explanation for the observed patterns

Table 4: Multinomial Logit Predicted Probabilities (1981-2000)

Growth Pattern	
Bot 10% Persist	0.148***
	(0.00531)
Bot $10\%_{t-1}$ to Top $10\%_t$	0.128***
	(0.00418)
Mid $80\%_{t-1}$ to Bot $10\%_t$	0.0926
	(0.00219)
Mid $80\%_{t-1}$ to Top $10\%_t$	0.0745
	(0.00191)
Top $10\%_{t-1}$ to Bot $10\%_t$	0.115***
	(0.00533)
Top 10% Persist	0.245***
	(0.00691)
Observations	27,137

All regressions include lagged age, size and growth rate as well as year and industry dummies
Bootstrapped standard errors (100 reps) clustered at the firm. level in parentheses.
*** p<0.01, ** p<0.05, * p<0.1 (Null of H0: Coeff $\leq$ 0.1)

of persistence and recoil emerges: firms at the extremes of the growth distribution engage in more R&D than other firms, and this carries both risk and reward. This finding explicates Bianchini et al. (2017, p. 653) who suggest that "innovativeness... is not able to discriminate persistent high-growth". Moreover, on the basis of the sub-period results in Table A3, R&D appears to have become riskier over time, which may explain the results in Table A2 which show that the probability of underperformance by previously outperforming firms has been increasing over time.

Average partial effects of patent yields can also be found in Table 5. These offer some support for hypothesis 4b insofar as it points towards a higher quantity of innovation output being associated with improved growth outcomes at the lower end of the growth distribution; increases in UPY are associated with a reduced probability of underperformance in the following period. However unweighted patent yields appear not to have a discernible effect on the probability of outperformance. In contrast, there is strong support that more valuable innovation improves growth outcomes along the entire growth distribution (hypothesis 4c), with a higher WPY being associated with a significant reduction in the probability of underperformance and increase in the probability of outperformance.

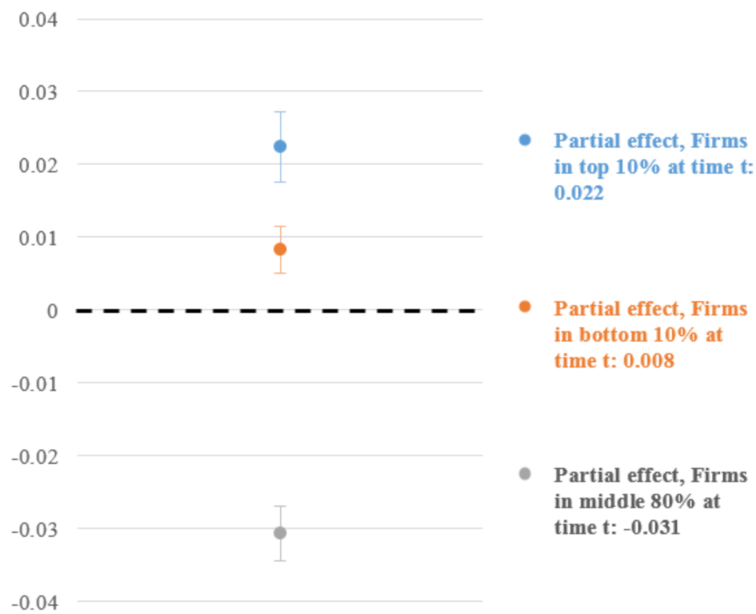

Fig. 3. Average Partial Effects of $RDIntensity_{t-1}$ (1981-2000)

Table 5: Multinomial Logit Innovation Average Partial Effects (1981-2000)

		1981-2000
RDIntensity$_{t-1}$	$CatSales_t = 0$	0.00823***
		(0.00163)
	$CatSales_t = 2$	0.0224***
		(0.00188)
UPY$_{t-1}$	$CatSales_t = 0$	-0.0228***
		(0.00532)
	$CatSales_t = 2$	0.00563
		(0.00463)
WPY$_{t-1}$	$CatSales_t = 0$	-0.000711***
		(0.000221)
	$CatSales_t = 2$	0.000586***
		(0.000131)
Observations		27,137

All regressions include lagged age, size and growth rate as well as year and industry dummies

Bootstrapped standard errors (100 reps) clustered at the firm. level in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

We turn now to a distinct channel in which innovation may play a role in generating observed growth patterns; through heterogeneous effects across different kinds of firms. To investigate this possibility, we present the partial effects of innovation by lagged growth outcome in Table 6.[12] These are illustrated in Figures 4 and 5. R&D tends to be more rewarding for firms that previously experienced extreme growth compared to firms previously in the middle 80% of growth rates. In Figure 4, this is captured by the 'V-shaped' blue line. Such a finding suggests that the comparatively higher R&D intensities of outperformers and underperformers may be justified. In contrast, it is not clear whether R&D is riskier for any particular growth based subpopulation, as captured by the flat orange line. It is however possible that the true risk of previously underperforming firms persisting as underperformers into the subsequent period is actually higher due to a left truncation arising from the exclusion of the exit year of non-surviving firms. This would imply that the reported increase in the probability of persisting as an underperformer associated with higher R&D represents a lower bound. Similarly, as can be seen in Figure 5, the pattern of partial effects for WPY across lagged growth subpopulations supports the above findings that innovation tends to be more rewarding for firms that previously experienced extreme growth: higher weighted patent yields are associated with both significantly greater reductions in the probability of underperformance at time t and significantly greater increases in the probability of outperformance at time t for firms who previously experienced an extreme growth event.

A possible explanation for why higher R&D intensity and patent yields are more rewarding for firms that previously experienced extreme growth is that such firms are, on average, younger and smaller than firms in the middle 80%, as seen in Table 1. Balasubramanian and Lee (2008) find that firm age and size have a significant and negative association with the technical quality of innovation, thereby attenuating the positive impact of R&D intensity. We find evidence supporting this conjecture. Figure 6 gives the partial effects of $RDIntensity$ on the probability of outperformance and underperformance evaluated at each decile of the log size distribution: the increase in the probability of outperformance associated with higher $RDIntensity$ is decreasing in firm size. Interestingly however, this same inverse relationship appears to hold true for the effect of $RDIntensity$ on the probability of underperformance. This finding offers further support for the second explanation which links innovation to patterns of recoil and persistence, namely that R&D is both riskier and more rewarding for different firm subpopulations.

[12] Sub-period results available on request.

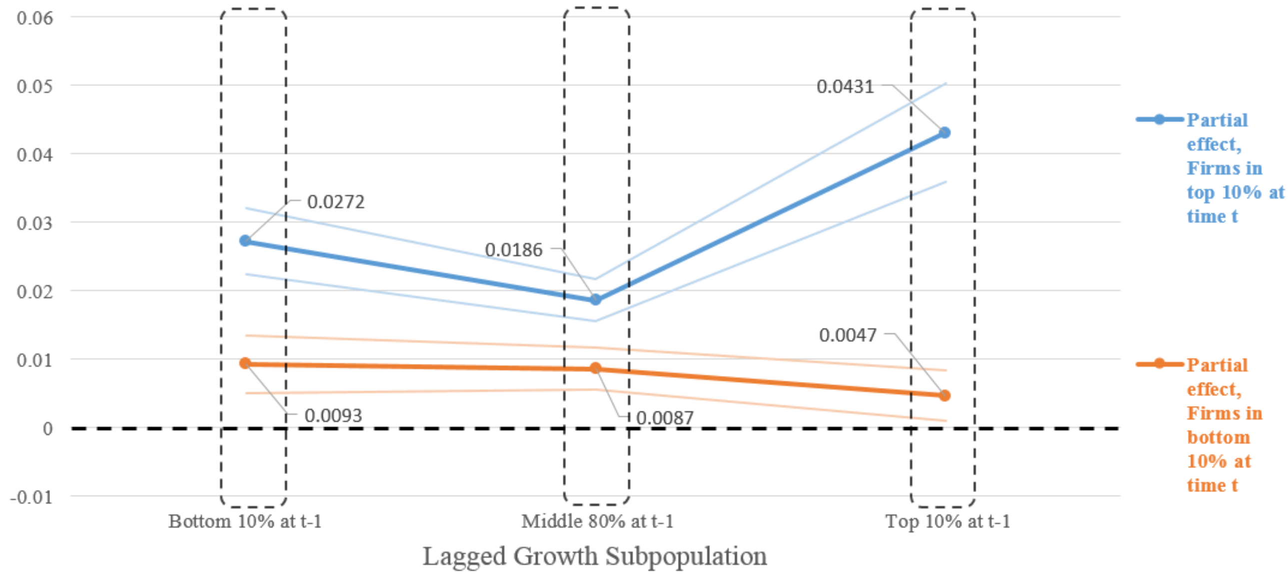

Fig. 4. Partial Effects of $RDIntensity_{t-1}$ by lagged growth outcome (1981-2000)

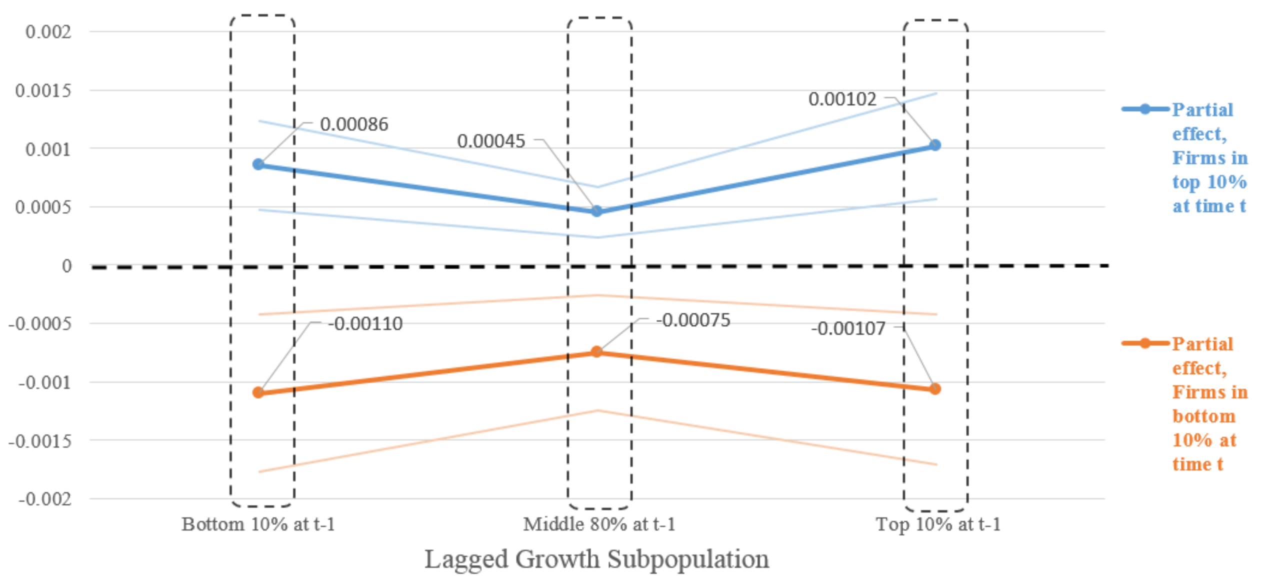

Fig. 5. Partial Effects of WPY_{t-1} by lagged growth outcome (1981-2000)

Table 6: Multinomial Logit Innovation Partial Effects by Lagged Growth Outcome (1981-2000)

		Growth outcomes at time $t-1$		
		$CatSales_{t-1} = 0$	$CatSales_{t-1} = 1$	$CatSales_{t-1} = 2$
RDIntensity$_{t-1}$	$CatSales_t = 0$	0.00929***	0.00866***	0.00467**
		(0.00219)	(0.00154)	(0.00188)
	$CatSales_t = 2$	0.0272***	0.0186***	0.0431***
		(0.00245)	(0.00156)	(0.00368)
UPY$_{t-1}$	$CatSales_t = 0$	-0.0316***	-0.0212***	-0.0212***
		(0.00730)	(0.00499)	(0.00617)
	$CatSales_t = 2$	0.00912	0.00423	0.0120
		(0.00584)	(0.00381)	(0.00903)
WPY$_{t-1}$	$CatSales_t = 0$	-0.00102***	-0.000644***	-0.000908***
		(0.000306)	(0.000207)	(0.000252)
	$CatSales_t = 2$	0.000796***	0.000470***	0.00117***
		(0.000108)	(0.000108)	(0.000255)
Observations		27,137	27,137	27,137

All regressions include lagged age, size and growth rate as well as year and industry dummies.
Bootstrapped standard errors (100 reps) clustered at the firm level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

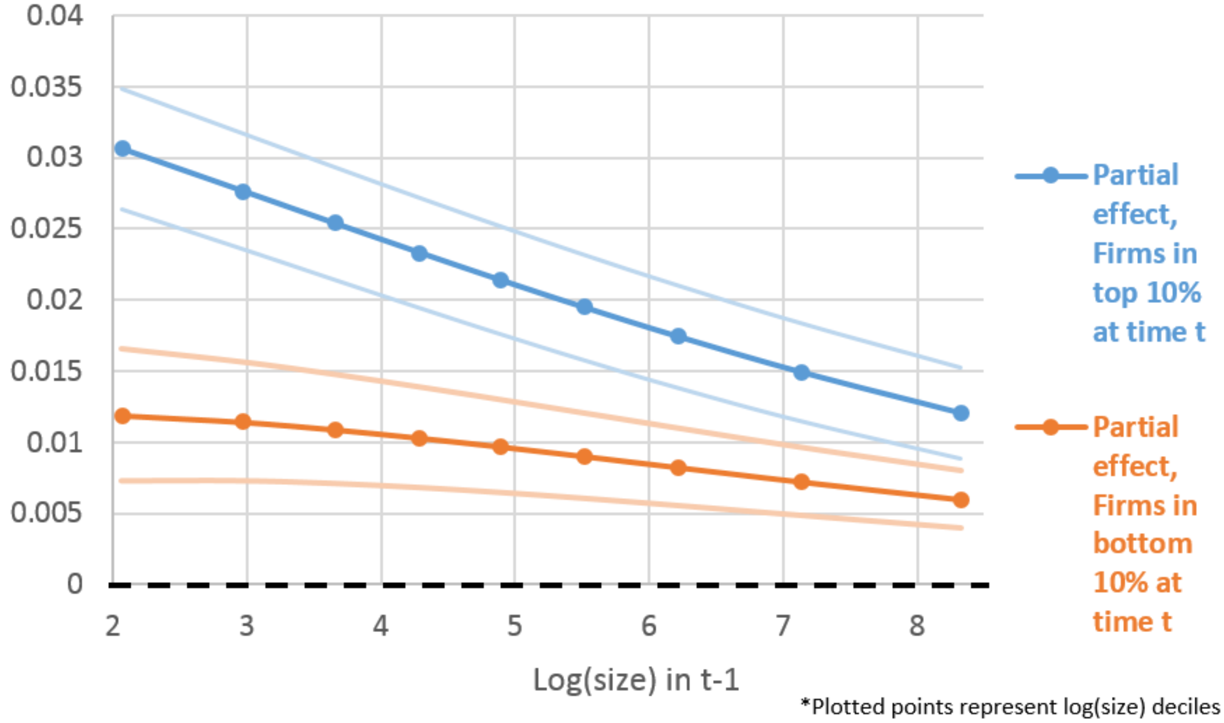

Fig. 6. Partial Effect of $RDIntensity_{t-1}$ on the probability of outperformance by logsize decile (1981-2000)

7. Conclusions

We investigated the role of innovation in generating the observed patterns in firm growth. The aim was to reconcile the observed coexistence of persistent and recoiling firms at both extremes of the growth rate distribution, with the pattern in the growth effects of innovation which has been found to depend on firm growth itself. The primary results can be summarised as follows: while the growth effect of R&D intensity on the representative manufacturing firm is positive, the short term growth effects of innovation on firms at the two extreme ends of the conditional growth rate distribution are decidedly polarised, with significantly higher probabilities of both extreme positive and of extreme negative growth rates in the next period. Insofar as a firm is able to generate valuable patents, the downside risk is reduced. R&D may have become riskier over time, but is not intrinsically riskier for firms growing at extreme rates.

These results suggest that the inherently uncertain nature of innovation can help explain the striking patterns in extreme firm growth. In terms of policy, support for innovation is justified by the finding that innovation has a positive growth effect on average, and further, by the finding that for firms at growth rate extremes, innovation increases the probability of sustaining or rebounding to positive extreme growth more than it increases the probability of sustaining or bouncing down to negative extreme growth. Potential rewards of innovation are highest for firms experiencing extreme growth.

The multinomial model we estimated does not exploit the panel structure of the data. But results (available on request) of separate binary fixed-effects logistic regressions,[13] one for the bottom growth decile and another for the top growth decile, both relative to the mid-range growers, produced very similar results that reinforce our main finding: innovation increases the probability of both underperformance and outperformance. Further work employing methods that allow for heterogeneity of outcomes in a panel setting, for example panel quantile regression (Coad et al., 2016), is the next step. Selection effects due to including only innovating firms (Segarra and Teruel, 2014), failing to control for firm survival (Capasso et al., 2015), and not controlling for sectoral heterogeneity within manufacturing (Coad and Rao, 2008) are also to be addressed in further work.

The focus on short-term growth has led us to consider only innovation in the immediately preceding year. This masks some of the complexity in the dynamics of the innovation process and its effect on firm growth over time (Coad and Rao, 2010).[14] It would be worthwhile to extend the analysis to consider more longer term growth patterns (for example, in the vein of Bianchini et al., 2017; Capasso et al., 2015; Stam and Wennberg, 2009). Alternative innovation measures are also worth considering: for example, self-citation ratio that characterise patents as 'external' or 'internal', as in Akcigit and Kerr (2016, p. 9)[15], and measures of technological spillover from the innovation network that links patent fields together (Acemoglu et al., 2016).

[13] Koenker and Hallock (2001) highlight the validity of estimating 'a family of binary response models for the probability that the response variable exceeded some prespecified cutoff values'.

[14] In defence of a short-term focus, see Geroski and Machin (1992, p. 81).

[15] See also Segarra and Teruel (2014).

References

Acemoglu, D., Akcigit, U., Kerr, W. R., 2016. Innovation network. Proceedings of the National Academy of Sciences 113, 11483–11488.

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

Akcigit, U., Hanley, D., Stantcheva, S., 2016. Optimal Taxation and R&D Policies. Working Paper 22908, National Bureau of Economic Research.

Akcigit, U., Kerr, W. R., 2016. Growth Through Heterogeneous Innovations. Working Paper 11-044., Harvard Business School Entrepreneurial Management.

Balasubramanian, N., Lee, J., 2008. Firm age and innovation. Industrial and Corporate Change 17, 1019.

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

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

Bolton, M. K., 1993. Organizational Innovation and Substandard Performance: When is Necessity the Mother of Innovation? Organization Science 4, 57–75.

Bromiley, P., 1991. Testing a Causal Model of Corporate Risk Taking and Performance. The Academy of Management Journal 34.

Capasso, M., Cefis, E., Frenken, K., 2014. On the existence of persistently outperforming firms. Industrial and Corporate Change 23, 997.

Capasso, M., Treibich, T., Verspagen, B., 2015. The medium-term effect of R&D on firm growth. Small Business Economics 45, 39–62.

Coad, A., 2009. The Growth of Firms: a Survey of Theories and Empirical Evidence. Edward Elgar, Cheltenham, UK .

Coad, A., Hölzl, W., 2009. On the Autocorrelation of Growth Rates. Journal of Industry, Competition and Trade 9, 139–166.

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

Coad, A., Rao, R., 2010. Firm growth and R&D expenditure. Economics of Innovation and New Technology 19, 127–145.

Coad, A., Segarra, A., Teruel, M., 2016. Innovation and firm growth: Does firm age play a role? Research Policy 45, 387–400.

Crepon, B., Duguet, E., Mairesse, J., 1998. Research, Innovation, and Productivity: An Econometric Analysis at the Firm Level. Working Paper 6696, National Bureau of Economic Research.

Cyert, R. M., March, J. G., 1963. A Behavioral Theory of the Firm. SERBIULA (sistema Librum 2.0) .

Demirel, P., Mazzucato, M., 2012. Innovation and Firm Growth: Is R&D Worth It? Industry and Innovation 19, 45–62.

European Commission, 2011. Policies in support of high-growth innovative SMEs. Bonn: An INNO-Grips Policy Brief by empirica Communication and Technology Research .

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

Freel, M. S., 2000. Do Small Innovating Firms Outperform Non-Innovators? Small Business Economics 14, 195–210.

Geroski, P., Machin, S., 1992. Do Innovating Firms Outperform Non-Innovators? Business Strategy Review 3, 79–90.

Geroski, P. A., 1999. The Growth of Firms in Theory and in Practice. CEPR Discussion Papers 2092, C.E.P.R. Discussion Papers.

Greve, H., 1998. Performance, Aspirations, and Risky Organizational Change. Administrative Science Quarterly 43, 58.

Greve, H. R., 2003. A Behavioral Theory of R&D Expenditures and Innovations: Evidence from Shipbuilding. The Academy of Management Journal 46, 685–702.

Grinyer, P., McKiernan, P., 1990. Generating Major Change in Stagnating Companies. Strategic Management Journal 11, 131–146.

Guarascio, D., Tamagni, F., 2016. Persistence of innovation and patterns of firm growth. LEM Papers Series 2016/31, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

Hall, B. H., Jaffe, A., Trajtenberg, M., 2005. Market value and patent citations. The RAND Journal of Economics 36, 16–38.

Hall, B. H., Jaffe, A. B., Trajtenberg, M., 2001. The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. Working Paper 8498, National Bureau of Economic Research.

Harhoff, D., Narin, F., Scherer, F. M., Vopel, K., 1999. Citation Frequency And The Value Of Patented Inventions. The Review of Economics and Statistics 81, 511–515.

Harris, R., Tzavalis, E., 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. Journal of Econometrics 91, 201–226.

Hölzl, W., 2014. Persistence, survival, and growth: a closer look at 20 years of fast-growing firms in Austria. Industrial and Corporate Change 23, 199.

Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. Econometrica 47, 263–91.

Koenker, R., Hallock, K. F., 2001. Quantile Regression. Journal of Economic Perspectives 15, 143–156.

Kropko, J., 2007. Choosing between multinomial logit and multinomial probit models for analysis of unordered choice data. ProQuest Dissertations and Theses .

Lööf, H., Mairesse, J., Mohnen, P., 2016. CDM 20 Years After. Working Paper Series in Economics and Institutions of Innovation 442, Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies.

Miller, K. D., Leiblein, M. J., 1996. Corporate risk-return relations: Returns variability versus downside risk. Academy of Management Journal 39, 91–122.

Nickell, S., 1981. Biases in Dynamic Models with Fixed Effects. Econometrica 49, 1417–1426.

Scherer, F. M., 1965. Corporate Inventive Output, Profits, and Growth. Journal of Political Economy 73, 290–297.

Segarra, A., Teruel, M., 2014. High-growth firms and innovation: an empirical analysis for Spanish firms. Small Business Economics 43, 805–821.

Shane, H., Klock, M., 1997. The Relation Between Patent Citations and Tobin's Q in the Semiconductor Industry. Review of Quantitative Finance and Accounting 9, 131–146.

Stam, E., Wennberg, K., 2009. The roles of R&D in new firm growth. Small Business Economics 33, 77–89.

Trajtenberg, M., 1990. A Penny for Your Quotes: Patent Citations and the Value of Innovations. The RAND Journal of Economics 21, 172–187.

Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of Econometrics 126, 25–51.

Zachariadis, M., 2003. R&D, innovation, and technological progress: a test of the Schumpeterian framework without scale effects. Canadian Journal of Economics/Revue canadienne d'conomique 36, 566–586.

Appendix A.

Table A1: GMM Results by Sub-period

	1981-1985	1986-1990	1991-1995	1996-2000
$GrSales_{t-1}$	0.201***	0.0112	-0.0222	-0.0862**
	(0.0565)	(0.0582)	(0.0509)	(0.0396)
$RDIntensity_{t-1}$	0.0289	0.0311	0.0587**	0.0837***
	(0.0354)	(0.0485)	(0.0269)	(0.0319)
WPY_{t-1}	-0.000129	0.000274	0.00107	0.000417
	(0.000413)	(0.000517)	(0.000695)	(0.00132)
WPY^2_{t-1}	-3.68e-08	9.49e-08	-7.21e-07	-2.46e-06
	(2.56e-07)	(8.03e-07)	(9.71e-07)	(5.69e-06)
UPY_{t-1}	-0.00656	0.0175	0.0131	0.00398
	(0.0187)	(0.0161)	(0.0164)	(0.0245)
UPY^2_{t-1}	0.00300	-0.00117	0.00139	-0.00117
	(0.00234)	(0.00225)	(0.00241)	(0.00370)
Observations	6,161	6,465	6,796	7,715
Firm Count	1,696	1,853	1,883	2,213
# Instruments	49	59	59	59
Hansen OverID test	0.596	0.239	0.209	0.160
Arellano-Bond AR(2) test	0.0421	0.985	0.536	0.328

All regressions include lagged age, size as well as industry and year dummies.
Robust standard errors clustered at the firm level in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A2: Multinomial Logit Predicted Probabilities by Sub-period

Growth Pattern	1981-1985	1986-1990	1991-1995	1996-2000
Bot 10% Persist	0.159	0.143	0.158	0.154
Bot $10\%_{t-1}$ to Top $10\%_t$	0.0880	0.112	0.144	0.154
Mid $80\%_{t-1}$ to Bot $10\%_t$	0.0945	0.0937	0.0911	0.0897
Mid $80\%_{t-1}$ to Top $10\%_t$	0.0744	0.0737	0.0745	0.0732
Top $10\%_{t-1}$ to Bot $10\%_t$	0.0941	0.112	0.117	0.129
Top 10% Persist	0.273	0.264	0.235	0.233

Warmer colours represent higher probability

Table A3: Multinomial Logit Innovation Average Partial Effects by Sub-period

		1981-1985	1986-1990	1991-1995	1996-2000
RDIntensity_{t-1}	$CatSales_t = 0$	-0.00472	0.00833**	0.00654**	0.0127***
		(0.00372)	(0.00378)	(0.00279)	(0.00300)
	$CatSales_t = 2$	0.0191***	0.0247***	0.0189***	0.0263***
		(0.00386)	(0.00371)	(0.00350)	(0.00317)
UPY_{t-1}	$CatSales_t = 0$	-0.0269**	-0.0149	-0.0230**	-0.0289***
		(0.0133)	(0.0126)	(0.0105)	(0.0104)
	$CatSales_t = 2$	0.000658	0.00608	0.00409	0.0159**
		(0.00916)	(0.00882)	(0.0101)	(0.00734)
WPY_{t-1}	$CatSales_t = 0$	-0.000825	-0.000768	-0.000698*	-0.00193**
		(0.000643)	(0.000471)	(0.000385)	(0.000818)
	$CatSales_t = 2$	0.000317	0.000931***	0.000523	0.00139**
		(0.000198)	(0.000323)	(0.000324)	(0.000555)
Observations		6,161	6,465	6,796	7,715

All regressions include lagged age, size and growth rate as well as year and industry dummies.
Bootstrapped standard errors (100 reps) clustered at the firm level in parentheses
*** p<0.01, ** p<0.05, * p<0.1