SPRU
Science and Technology Policy Research

SPRU Electronic Working Paper Number 202

New Venture Survival and Growth: When does the fog lift?

Alex Coad, Julian S Frankish, Richard G Roberts, David J Storey

9 November 2012

University of Sussex
SPRU – Science & Technology Policy Research

SPRU – Science and Technology Policy Research

University of Sussex
Falmer, Brighton, BN1 9SL
http://www.sussex.ac.uk/spru

New Venture Survival and Growth: When does the fog lift?

Alex Coad[+], Julian S Frankish*, Richard G Roberts*, David J Storey[+]
+University of Sussex
*Barclays Bank

ABSTRACT:

Does our ability to predict the performance of new ventures improve in the years after start-up? We investigate the growth and survival of 6247 new ventures that are tracked through the customer records at Barclays Bank. We develop the concept of Gamblers Ruin as a valid approach to understanding new venture growth and survival. The Nagelkerke R^2 obtained from growth rate regressions decreases significantly in the years after start-up, which suggests that the fog gets thicker with respect to growth. However, the Nagelkerke R^2 of survival regressions increases in the years after start-up. Interestingly, a blip in year 5 suggests that macro-economic factors may have a strong effect on the amount of 'fog.' Our results offer little support for Strategic Entrepreneurship theory.

Received: 9 November 2012

JEL CODES: L26, L25

KEYWORDS: Entrepreneurship, firm growth, survival analysis, coefficient of determination, selection environment

1. INTRODUCTION

New ventures are characterised by a high likelihood of closure and by considerable variation between firms and over short periods of time within firms. Dahl and Sorenson (2012) for Denmark, show 52% of new ventures with a least one employee do not survive for 4 years. They also show that, although mean profitability is 127,000 Dkr, this varies across the sample from profits of approximately 8 million Dkr to losses of 100 million DKr.

Given this variance, the challenge facing scholars is to offer insights into the factors influencing that performance. Are there explanations for why some prosper but most do not? Even more challenging is whether these explanations can then be used to forecast the survival and/or growth of individual, or groups of, new ventures.

This paper addresses this challenge. It sets out two contrasting theoretical approaches and then tests their predictions on a sample of 6247 new ventures, all of which began to trade in the UK in the same quarter of 2004. Our chosen approach is Gambler's Ruin (GR) which we use to seek to explain the performance of new ventures – where performance includes both growth/ decline/stability[1] and survival/ non-survival.[2]

Gamblers Ruin assumes that new venture growth/decline/stability is a random walk, whereas new venture survival is strongly influenced by access to financial resources. This contrasts it with Strategic Entrepreneurship (SE) which implies both survival and growth are strongly related to the human capital of the founder(s), most notably their ability to identify and exploit opportunities [Ireland et al 2003].

The key test of these theoretical approaches is whether, as the new venture matures, its performance becomes more predictable and more closely linked to the human capital of the founder(s). SE implies that performance – growth and survival - is linked to human capital and becomes more predictable with time. In terms of the title of this paper – the fog lifts. In contrast, the key assumption in Gamblers Ruin is that growth is random and hence the fog never lifts. However Gamblers Ruin assumes that survival does become more predictable with time, since the longer-established survivors acquire financial resources which are not available to (most) new ventures and which can be used to "even out" the inevitable fluctuations that characterise the revenue streams in a new venture [Storey 2011]. The survival "fog" does lift under Gamblers Ruin.

[1] Prior empirical work has measured growth in terms of numerous metrics such as employment, sales, profits, business valuation. In the empirical section this paper follows the recommendation of Zimmerman and Zeltz (2002) and uses a measure of sales – credit turnover - as its chosen metric. Zimmerman and Zeltz (2002, p417) write that "growth in sales is especially important for new ventures since their economies of scale are too important for them to continue without increasing their scale of operations."

[2] Although conceptually simple, defining survival can be problematic (Storey and Greene, 2010). In the theory section of this paper we take survival to be the continuation of the same business between two points in time for which we observe turnover activity.

The key empirical novelty of the paper is, by taking the R^2 statistic as a measure of the density of the fog, to investigate how the fog changes as the new venture matures. We show the evolution of the R^2 statistic obtained from growth regressions and survival regressions in the 6 years after entry, and observe that while the ability to predict growth deteriorates in the years after entry, the ability to predict survival improves somewhat. This approach is only possible because the large dataset available to us monitors every financial transaction by a new venture from the day it begins trading. These trends in the R^2 coefficient are not linear or monotonic, however, and we suspect that macroeconomic phenomena (such as the start of a major recession) play a role in determining the amount of 'fog' that obfuscates the performance landscape.

The remainder of the paper is set out as follows: Section 2 discusses the predictions emerging from theory, which leads up to the development of hypotheses in Section 3. Section presents our methodology and our bootstrap algorithm. Section 5 presents the dataset, and we test our hypotheses in Section 6. Section 7 concludes.

2. THEORY

Gambler's Ruin (GR): Wilcox (1971) provided an early formulation of Gamblers Ruin (GR). Its central tenet is that firm growth is a random walk, as proposed by Gibrat (1931) as a statistical explanation for the observed log-normal firm size distribution. Here the probability of a proportionate change in size over time is the same for all firms in a given industry, irrespective of their size at the beginning of the period (Mansfield, 1962). Gibrat's Law therefore implies the firm lacks control over the environment in which it either buys inputs or sells its output – reflecting the textbook model of perfect competition. Gibrat's Law models growth in terms of the following random process:

$$x_t = x_{t-1} + \varepsilon_t, \qquad (1)$$

where x_t is the logarithm of firm size at time t, and ε is a random shock (additive in logs, but multiplicative on a linear scale). Despite having some empirical support,[3] Gibrat's Law – and hence the random walk hypothesis – is frequently dismissed on two grounds. First it is deemed to have failed key empirical tests. Second, and more intuitively, it appears at odds with the reported views of those operating enterprises.

The first empirical test failed by Gibrat's Law is that the smallest firms – if they survive – grow faster than either medium or larger firms (Hart and Oulton, 1996). The second test is that there are

[3] Storey (2011) shows the percentage of new firms having sales growth or decline in each six-month period approximates closely to a coin-tossing (Bernoulli) trial, where p = 0.5, with 34.9 percent having two periods of growth and 35.5 percent having three over four periods of time. Parker et al (2011) show fast growth in one period of time is uncorrelated with that in a later period. Coad (2007) shows that, for smaller firms, serial correlation is negative.

identifiable factors associated with either firm survival or growth – which is incompatible with a random walk. These factors include the skills and talents of the owner(s) reflected in the managerial decisions they make. For example, in their review of the factors influencing growth, Wiklund et al (2009) identify five groups of factors: entrepreneurial orientation; the external environment; strategic fit; the role of resources; and attitudes; all of which have been shown to influence firm performance.

The second reason why Gibrat's Law is frequently rejected is an intuitive unease with firm performance being seen as a pure random walk. This appears to be clearly at variance with the explanations given by business owners for the success of their enterprise, implying that the qualities of their owners/ managers or the markets in which they operate are unrelated to how a firm performs. These qualities might be the ability to recognise and exploit opportunities (Wiklund et al 2009, Shane, 2003) or the ability to manage other workers (Casson and Wadeson, 2007).

These limitations notwithstanding, Gibrat's Law can be considered to be a reasonable approximation for an examination of new venture growth (Coad et al, 2012). Gamblers Ruin (GR) takes the stochastic growth mechanism from Gibrat's Law, and uses the analogy of a gambler playing a game of chance, with access to a given stock of resources – gambling chips. The gambler at a table in a casino is constantly facing a choice of whether or not to stay at the table (continue in business) or to leave the table (quit). It is assumed that the owner derives positive utility from being at the table (Hamilton 2000[4]) but quits for one of three reasons: the first is when their supply of financial resources – chips – becomes exhausted. The second is following a very large win which eliminates any further need to gamble. The third is to 'quit while ahead' (perhaps returning to the table at some point in the future when they are more confident of success). Gambler's Ruin therefore generalizes Gibrat's growth process to make predictions concerning survival (see also Levinthal, 1991).

Each of these gambling situations has clear parallels in the new venture context. The case where the gambler leaves the table because of an exhausted supply of resources is comparable to the business owner who has no finance with which to continue trading. The firm's survival, S, therefore depends on whether it is above a minimum threshold size x*:

$S = 1$ if $x > x^*$; otherwise $S = 0$ (2)

Modifications and extensions of this survival condition can also be mentioned. For example, there is the "special case" of the individual who is a major winner and who is able to sell the business without ever having to work again.[5] Much more typical[6]is the case where the business owner quits without

[4] Hamilton (2000) finds that "most entrepreneurs enter and persist in business despite the fact that they have lower initial earnings and lower earnings growth than in paid employment implying a median earnings differential of 35% for individuals in business for 10 years"

[5] This is referred to as the "special case" because of its rarity. About 500,000 businesses are started in the UK annually of which two thirds cease to trade in six years. Of these, less than 2% have sales of £1m in six years. As for the ultimate "win" of an IPO launch, Ritter (2011) shows that, even in the US, there were only 7617 between 1980 and 2011, or less than 250 per year. So, if the UK had the same rate of conversion to IPOs as the US and none of the IPOs were acquisitions, then it would imply a chance of approximately 1 in 10,000 new businesses becoming an IPO.

having exhausted either their own sources of finance or any credit that might be supplied by others. Their choice to quit is assumed to be influenced by the current availability of more attractive options – such as waged employment or retirement – but also by how they expect the business to perform in the future (Gimeno et al, 1997).[7]

Gamblers Ruin (GR) therefore retains the random walk component of Gibrat's Law but supplements it with the key concept of the stock of financial resources. These are assumed to enhance the survival of the firm, meaning that firms with access to them are more likely to survive, and weather the inevitable series of adverse shocks, than otherwise similar firms without such resources. In the GR model these resources are exclusively financial. They do not include the wider concept of resources captured within the Resource-Based View of the firm (Barney 2001).

In summary, Gambler's Ruin assumes financial resources are acquired in three ways. The first is ***direct acquisition from "wins"*** in a game of chance. Applying this in a new venture context, a "win" might be a substantial order from a customer generating an income stream. The income from such wins can then be used to purchase new inputs or may be used to fund periods of time when there are no wins, and without such wins the firm would have closed. The second way of generating financial resources is their ***acquisition from external parties*** as a result of "wins." Because the market for external finance is opaque, external providers of finance have to rely heavily on signals of quality from those seeking funding (Berger and Udell, 2006). One key signal is obtaining a "win", which provides the financier with confidence. Wins provide a signal in this opaque marketplace that is taken to point to an increased likelihood that more sales will follow. In gambling terms it may be taken to imply that the punter has valuable knowledge/ insight/ information. The third way is ***having access to them in a way which is unrelated to the performance of the firm***.

Apart from family, finance for current ventures may also have been accumulated from prior successful business ventures – visits to the table. GR attributes this to prior luck, but it may also reflect entrepreneurial talent reflected in those with prior experience performing better than novices (Plehn-Dujowich 2010). To reach the latter conclusion however requires being able to distinguish between the successful and the less successful, whilst avoiding the use of self-report data (Bertrand and Mullainathan 2001) and to distinguish between the finance and experience. Evidence on this is scarce, but Metzger (2006) found that those whose previous business resulted in bankruptcy were more likely to quit their current business than those who closed through choice. Although not definitive, it implies that access to finance is a factor influencing current performance.

Prior wealth is therefore assumed to have two influences on the pay-off matrix. The first is to lower the downside risk of new venture creation because the individual knows that the consequences for them of exit are small since, even if it happens, their personal financial resources or the finance they

[6]Headd (2003) shows for the US that only one third of new businesses closed under circumstances that their owners considered unsuccessful. Interestingly in the current context Headd also shows that "having the resources to be or get larger leads to survival." p56. Watson (2010), using Australian cases in retailing, makes similar points.
[7]They however note: "we found survival and growth to be governed by similar stochastic processes." P.387

can access from others will cover any shortfall. The second influence is that access to wealth enables the individual to continue playing the lottery of business ownership when those with lower wealth have to exit, so enabling them to continue to be eligible for "the big win".

In summary, GR assumes business growth/decline/stability is a random walk but that access to financial resources influences the survival of the business and so, indirectly, its performance. GR predicts that two finance-related factors positively influence new firm survival. The first is that those that begin large, because they have access to prior wealth, are able to "ride-out" the inevitable vicissitudes of trade faced by a new venture. The second is the role of "early wins". These generate not only internal resources but also external resources from providers in an opaque marketplace that interpret these "wins" as a signal of quality or talent.

Strategic entrepreneurship (SE): SE is the antithesis of GR (and *a fortiori* Gibrat's Law) because it assumes there are specific and identifiable characteristics of the business or its owner(s) that enable the new venture to shape its economic environment. The definition used by Ireland et al (2003) is that:

"Strategic entrepreneurship (SE) involves simultaneous opportunity-seeking and advantage-seeking behaviours and results in superior firm performance."

These "opportunity-seeking and advantage-seeking behaviours" reflect the psychological make-up of the owner – referred to as the entrepreneurial mind-set, most clearly reflected in the ability to recognise opportunities. However the duality of SE is its combination of both recognising, but also exploiting, opportunities. The latter is captured by the ability to manage strategically three forms of entrepreneurial capital-financial, human and social. At the heart of strategic entrepreneurship is knowledge or talent which may be in the form of the human capital of the owners and co-workers but also crucially, an ability to learn from past trading experience. Ireland et al (2003) say:

"Collected over time and events, knowledge, especially tacit knowledge, represents much of what the firm knows—about how to compete in its industry, to innovate, and to identify and exploit entrepreneurial opportunities" p.975

However what is distinctive about SE is that it explicitly links these identifiable characteristics of individuals, or combinations of individuals, to "superior firm performance". It therefore excludes, or implies only a modest role for, chance. This places it in direct contrast with GR.

We now take the two models (GR and SE) and identify seven dimensions in which they differ. These are shown as the seven columns of Table 1. The first column shows the extent to which GR and SE are expected to explain new enterprise performance. The measure of explanatory power we use is R^2. Table 1 shows that there is expected to be a striking difference between the explanatory power of the SE and GR models. The latter assumes financial resources affect survival but not growth, whereas SE assumes there are clearly identifiable factors that explain both components of performance. For these reasons, higher values of model fit (i.e., R^2) - in combination with significant effects for variables such as business experience, education and sources of advice - –reflecting

human and social capital - constitute support for the SE model. The expected outcome from GR is lower values of R^2, with these variables being insignificant.

Table 1 shows the interpretation of the role of **initial size** at start-up on survival also differs in the two models. In GR the assumption is that initial size captures access to financial resources which then can be used to "ride out" the short term fluctuations in performance that characterise new ventures[8], and therefore provide a buffer of resources that safeguards it against imminent exit. In contrast, SE assumes initial size reflects an informed judgement on the part of the business owner, and those providing (any) external funding, of the entrepreneurial mind-set, culture and leadership of the owner(s) with more able owners being able to access greater funding. Both models, however, point to initial size being positively correlated with survival and growth.

Human Capital is assumed to play a key role in the SE Model. This may be in the form of knowledge (Acs et al 2009), skills (Stuetzer et al 2012), education (Bates 1985), learning (Harrison and Leitch, 2005, Muehlfeld et al 2012), and motivation (Baum and Locke, 2004, Ireland et al 2003). Each of these is assumed to exert a positive influence on firm performance. This is in contrast with the GR model which assumes that, whatever skills the business owner has, do not influence firm performance since this is a lottery. GR therefore expects no observable relationship between these characteristics and firm performance. Prior empirical work on human capital and new venture performance has tended to find positive, yet not consistent links. For example the link between founders educational attainment and new venture survival was found to be positive by Dahl and Sorenson (2012), yet absent by Taylor (1999). A fuller comparison of prior work is found in Storey and Greene (2010).

Column 4 of Table 1 shows the expected role of **non-financial resources**. These play a crucial role in SE most notably as a source of social capital in the form of information drawn from networks. The new venture founder(s) draw upon these information sources to shape and influence the direction of the enterprise (Lechner and Dowling, 2003). However empirical work on new ventures is less clear about the role of such networks. For example Cooper et al (1994) and Lee et al (2001) find weak relationships whereas Chrisman and McMullan 2004; Pons Rotger et al, 2012 find some support for the use of publicly funded networks enhancing new venture performance. In SE, such non-financial resources play a vital role in enhancing both measures of performance. GR does not imply that information is not accessed by new venture founders, but merely that it does not affect either survival or growth.

A fifth column of Table 1 shows how the two models differ in terms of the role of the **age of the business**. In the SE model, a key element is that business owners learn (Harrison and Leitch, 2005). It assumes "the entrepreneur learns from the deals that he does and from the deals that fall through" (Casson 1982 p.387). On those grounds it is to be expected that those businesses that have more trading experience are likely to have accumulated more knowledge/ expertise than younger businesses and this will enable them to perform better. Age of business, under SE will therefore have

[8] See Footnote 3 earlier for evidence of such temporal fluctuations.

a positive and significant effect on both performance measures[9]. In contrast there is no role for learning in GR since, by definition, it is impossible to learn to play the lottery. The more intuitive explanation for an absence of learning amongst new venture owners is that the variety of circumstances that they face are so diverse that learning is (close to) impossible (Frankish et al, 2012). The role of firm age in GR is therefore more nuanced. If growth is a random walk then business age will be unrelated to growth. However, survivor bias may be present, implying that older firms are more likely to have had either prior wins or more resources at start-up. Firm age may therefore be capturing the contribution of accumulated financial resources which can be used to lower the likelihood of closure or to facilitate growth. In turn, survival under GR implies that the longer the firm stays in the game the bigger the chance it has of a big win.

Column 6 of Table 1 shows that **current size** has a positive effect on survival in both the GR and SE Models, although again for different reasons. GR assumes that currently larger firms have higher survival rates because they either had more financial resources initially, or acquired them as a result of "wins". These financial resources then enable the enterprise to more successfully ride out those variations in performance which occur than smaller enterprises without access to such resources. In GR current size also promotes growth since it provides a signal taken by lenders of the credibility of the business – making it more likely they will provide the resources necessary for investment to achieve growth. A very different interpretation of the positive link between current size and performance is provided by the SE model. This assumes that larger new firms are more likely to exhibit the twin qualities required for enhanced performance – "opportunity-seeking and advantage-seeking behaviours."

The final column of Table 1 captures the key concepts of this paper. It shows the GR model implies that several SE variables will provide weak explanations of the growth/decline/stability of a new venture when it begins to trade. It also implies that these explanations will not improve with time. The GR survival model, where the key element is access to financial resources, implies that survivors are those with such resources. The fog over survival does lift somewhat as the new venture matures.

In contrast, the SE model assumes there are identifiable mindset, cultural, leadership and managerial qualities that enhance new venture survival and growth. A model able to capture these characteristics would therefore have good explanatory power. Since learning is a key component of SE, it is to be expected that the explanatory power of such a model would increase with time. In terms of the title of this paper, SE implies the fog is less dense than implied under GR, and lifts over time.

[9] Of course, as recognised earlier, this knowledge may also have been accumulated from experience with enterprises other than their current enterprises.

Table 1: Comparing Gamblers Ruin and Strategic Entrepreneurship

	R^2	Initial size	Human capital	Non-Financial Resources	Age of business	Current size	Change in R^2 over time
GROWTH							
Gambler's Ruin **GR**	Low	No impact	No impact	No impact	No impact	No impact	None
Strategic Entrepreneurship **SE**	High	Positive	Positive	Positive	positive (learning)	Positive	High
SURVIVAL							
Gambler's Ruin **GR**	Low	Positive	No impact except via initial size	No impact	No direct impact	Positive	Low
Strategic Entrepreneurship **SE**	High	Positive	Positive	Positive	positive (learning)	Positive	High

3. DERIVING THE HYPOTHESES

Drawing upon the comparisons presented in Table 1 we now derive some hypotheses to be tested in Section 5. We are primarily interested in investigating the goodness-of-fit of models that seek to explain the survival and growth of new ventures. Hypotheses 1a and 1b relate to the R^2 statistics obtained from growth rate regressions, while Hypothesis 2 relates to the R^2 statistics obtained from survival regressions.

Regarding growth, Gambler's Ruin follows Gibrat's Law in approximating firm growth as a random walk. Growth remains random in the years after start-up. Therefore, Gambler's Ruin would predict that the R^2 from growth regressions is low and remains low over time.

In contrast, SE suggests that firm growth is driven by firm-specific capabilities, routines, and resources that are accumulated by firms and lead to persistent performance differentials. In the years after start-up, firms will accumulate these capabilities and resources to derive increasing competitive advantage. We therefore expect that, in the years after start-up, the explanatory power of growth rate regressions will increase:

Hypothesis 1: in line with Strategic Entrepreneurship, the R^2 from growth regressions is relatively high and increases in the years after start-up

According to Strategic Entrepreneurship, the R^2 from growth regressions can increase for two reasons. First, selection effects weed out the poor performing firms, which brings up the average 'viability' of survivors (the *selection* argument) and thus make the business landscape more predictable as chaotic, short-lived firms are removed. Second, surviving firms gain experience and improve in productivity/viability over time (the *learning* argument) which may make their performance outcomes less vulnerable to random shocks. In contrast, Gambler's Ruin considers growth to be a random process, and expected growth is not higher among firms that have grown in the past (i.e. no selection effect) nor is it higher among firms that have gained experience (i.e. no learning effect).

With regards to survival, the simplest form of GR holds that survival depends on the stock of accumulated financial resources. Gamblers leave the table when they run out of gambling chips. At every point in time, some gamblers can be expected to leave the table (even if they previously had a large stock of resources) if they experience a run of bad luck and deplete their resources. However, as time goes by, players will have had the chance to accumulate resources and protect themselves from the threat of sudden death. Players that remain small in the years after start-up will have higher exit rates, while more fortunate players can be expected to survive. As time goes by, firms that survive the 'liability of newness' enter into a more stable and predictable business environment.

SE makes similar predictions for the evolution of the R^2 statistic in the years after entry. In addition to the 'buffer' effect whereby resources can provide some protection from exit, SE suggests that other factors, such as entrepreneurial skill, business acumen, accumulated valuable resources and superior capabilities can provide further information on a firm's survival chances. If entrepreneurs are able to apply their human capital to learn from their business experience since start-up, then this will provide further explanatory power in terms of predicting which firms will survive.

With regards to the R^2 obtained from survival regressions, we therefore hypothesize:

Hypothesis 2: Both Gambler's Ruin and Strategic Entrepreneurship predict that the R^2 from survival regressions increases in the years after start-up

4 .METHODOLOGY

Of crucial interest to our paper is the measure of what we call 'fog' – the coefficient of determination, or R^2 statistic, which is the proportion of variance explained in contrast to that which remains unexplained. However, there are other indicators of 'goodness of fit' or explanatory power that are preferable to the usual R^2 statistic because they have more desirable properties. For example, the "adjusted-R^2" is often preferred to the standard R^2 because it takes into account the number of explanatory variables (that is, unlike R^2 the adjusted- R^2 only increases if the additional explanatory variables improve the model fit over and above what would be expected by mere chance).[10] More generally, Nagelkerke (1991) sets out a list of 7 desirable properties for an ideal R^2 statistic, and to satisfy these requirements he proposes what has become known as the "Nagelkerke R^2." Our analysis of the goodness of fit of our regression models is based primarily on analysis of the Nagelkerke R^2, although we show that our results are not sensitive to this choice of R^2 statistic.

As a starting point, we run regressions for each year, where the dependent variable is either growth rate or survival probability. For each year we obtain an R^2 statistic. These regressions are shown in Tables 4 and 5.

We begin by plotting the evolution of the Nagelkerke R^2 over time using line charts – one chart for growth, one for survival. This gives a simple overview of the change over time – whether the trend is generally increasing or decreasing over the years. However, to assess whether the differences are statistically significant we will need to apply some tests that involve associating a standard error to the Nagelkerke R^2 statistic.

How can we assess the significance of the changes in the R^2 statistic? In some regressions, a significant change in R^2 is investigated by comparing the R^2 before and after adding some extra variables. In this situation, the same sample of observations is used in both cases, and the same baseline variables are included, but the R^2 changes as new variables are added (e.g. Tanriverdi and Lee, 2008, p390). However, what we have in mind is different – we want to compare the R^2 over different years (i.e. with different cross-sectional subsamples for the different years), although we use the same number of observations and the same set of explanatory variables, to see if the same explanatory variables provide a better fit as firms get older. We therefore follow the following procedure:

[10] In our regressions, however, we keep the same number of explanatory variables in each year, to maintain comparability.

BOOTSTRAP ALGORITHM:

1) Randomly draw a sample of size n=1000 observations for a given year (bootstrapping without replacement)

2) Perform a survival or growth regression on these n=1000 observations, then repeat for r=1000 replications (with other randomly-selected samples of n=1000) to obtain a distribution of r=1000 estimates of the sample R^2 statistic

3) Obtain the mean and standard error for the distribution of R^2 statistics for a given year

4) Repeat for all available years

5) Apply two-sample t-tests to see if the R^2 statistics are significantly different.

5. THE DATASET: BARCLAYS BANK CUSTOMER ACCCOUNTS

To adequately test the above theories we require a dataset that permits the analysis of survival and growth of new businesses. In the UK the bulk[11] of prior empirical work on this topic has used data sources that are restricted to limited companies (Foreman-Peck, 1985) or to firms appearing in official public records (Van Stel and Storey 2004), or to selected survey respondents (Westhead et al 2005) or on the self-employed (Burke et al, 2008, Dawson and Henley, 2013). Compared with the data source used in this paper – new business customers of Barclays Bank – these data sets suffer from up to six serious problems. First their coverage is not only incomplete but subject to the bias noted by Yang and Aldrich (2012) – particularly the inadequate coverage of the smallest and newest enterprises. The second is that of incomplete performance data. For example the self-employment data only identify whether or not an individual is self-employed and not on the performance of the enterprise. The third is that, even when performance data are available, it is only on a discrete basis –at best annually. The limitation of this is that during that 12 month period there are considerable monthly fluctuations in sales and up to 15% of new enterprises cease. The data used in this paper are continuous and collated on a biannual basis. The fourth limitation of many studies is a reliance on self-report data – without the full significance of such individuals being recognised as highly optimistic (Storey 2011, Cassar2010). Our data avoids the use of self-report information, other than that provided by the individual(s) prior to start-up. A fifth problem is that new venture studies can be restricted to a single sector, thus making generalisations difficult (Fan (2010)). The data set used here covers all sectors of the UK economy with the exception of financial services. Finally the size of the data set is considerable; with all financial transaction examined amongst 6247 businesses that began to trade for the first time in the March-May 2004.

[11] The notable exception is Cressy (1996) which used data from National Westminster Bank – now Royal Bank of Scotland (RBS).

We now set out the key definitions employed in our data base.

Start-up: definition

To be eligible for inclusion a customer of Barclays Bank is defined as a start-up enterprise if they provide a positive response to two questions. First, is this the first business current account for the customer with Barclays? If so, is it the first business current account for the business at any bank? If the answer to these two questions is 'yes' then the customer is a start-up. It therefore includes all businesses covered in other official data-sets such as those on self-employment, limited companies or those registered for tax purposes. The only exception would be a totally cash-based business. For reasons described more fully below, our dataset was drawn from non-financial firms identified as start-ups who entered Barclays customer base between March and May 2004. During this three month period there were 23,344 such firms.

There are two important caveats to note regarding this relatively straightforward description of a start-up. One is that a minority of firms will have been trading prior to opening a business current account, since companies and partnerships must trade through such accounts. However, this is not the case for sole traders where use can be made of personal accounts[12]. It is not clear how many start-ups entering Barclays are in this position, although their activity levels will, in most cases, have been on a tiny scale[13]. The other is that a new business does not necessarily start trading immediately upon opening an account. Indeed, for Barclays customers, approximately five percent of start-ups show no activity through their account in the subsequent 12 months. We addressed this by only including firms that showed activity in the month following entry to the customer base.[14] [15]

We therefore focus on a cohort of firms that have the same start date. We consider this to be important, because firms starting in different years may not be readily comparable (especially if the macroeconomic conditions at start-up have persistent effects on firm development in subsequent years. Focusing on a single cohort means that firms face the same macroeconomic conditions at each year of their development, and can therefore be meaningfully compared to each other.

Start-up: data

We now set out the data required from a Barclays customer. Basic owner(s) data was collected on gender and age but, at start-up, respondents were also asked about the following.

To capture one element of human capital identified in Table 1 customers were asked about the highest level of educational qualification obtained by the owner(s). As with other studies (van der Sluis et al 2005) this was intended to act as a proxy for the human capital available to the new firm. The second sought information about prior business experience (Westhead et al 2005). This included previous ownership and/or ownership among immediate family members, so capturing both work-

[12]For sole traders with low turnover an important consideration is that personal current accounts operate on a 'free in credit' basis, whereas equivalent business accounts have explicit charges after an initial period.

[13] On the grounds that the Bank would have identified business-related transactions in the account.

[14]We also included a small proportion of firms who did not show activity in their first full month, but in either May or June 2004. In these cases the start month of the firm was recorded as the month prior to activity.

[15] The UK, unlike many countries in continental Europe, is not characterised by multiple banking (Ongenga and Smith (2000)). The account at a single ban is therefore likely to capture the full trading activities of the new venture.

related capital and family support. Finally, to capture access to non-financial resources, owners were asked about the sources of advice and support they used of prior to start-up.

As the questions were voluntary, only a minority of the start-up customers in the three month period answered them fully. This left us with a sample base of 6247 start-ups, 27% of the total inflow, but more than large enough for our purposes[16].

We then supplement this data with information collected by the bank as part of its general account opening process. These are the age of the owner(s), the legal form of the business, the activity type (sector/branch/market) and its location (standard region) within the UK. Table 2 sets out the data in full.

[16]Response rates of 27%, even from samples where the population is not identified is somewhat above the"mid-range" for comparable articles. For example the reported rates are 14.6% in Ucbasaran et al (2009); 10.3% in Schulze et al (2003), 19.3 % in Hmieleski and Corbett (2008). Our response rateis lower than the 51% obtained byWiklund and Shepherd (2005) but, in none of the above cases, did the number of respondents exceed 500 and in most cases it was closer to 200.

TABLE 2: Variable Names and definitions.

Dependent variables	Description
Open = 1;	Enterprise continues to trade at end of period (Open = 0 if the enterprise exited)
Growth Rate	Growth is measures as Credit turnover – value of payments into a current account - excluding payments from a related account (deposit account). The growth rate is the log-differences of turnover [log(turnover(t)) - log(turnover(t-1)].
Independent variables	
Legal form	Legal form of business: 1 = Company, 2 = Partnership, 3 = Sole Trader
Industry dummies	Business activity: 1 = Agriculture, 2 = Manufacturing, 3 = Construction, 4 = Motor Trades, 5 = Wholesale, 6 = Retail, 7 = Hotels & Catering, 8 = Transport, 9 = Property Services, 10 = Business Services, 11 = Health, Education & Social Work, 12 = Other Services
GOR Region dummies	Region: 1 = East of England, 2 = East Midlands, 3 = London, 4 = North East, 5 = North West, 6 = South East, 7 = South West, 8 = West Midlands, 9 = Yorkshire, 10 = Scotland, 11 = Wales, 12 = Northern Ireland
No. owners	Number of owners
Excess owners	Owners in excess of minimum number for legal form: 0 = No, 1 = Yes
Male owners	Number of male owners
Male owner involved	At least one male owner: 0 = No, 1 = Yes
Age	Mean age of owner(s) at start-up
Age squared	Square of age
Education	Highest level of educational attainment by owner(s): 1 = <NVQ2, 2 = NVQ2, 3 = NVQ3, 4 = NVQ4
No business experience	Previous business experience, None: 0 = No, 1 = Yes
Family business experience	Previous business experience, Family: 0 = No, 1 = Yes
Individual business experience	Previous business experience, Owner: 0 = No, 1 = Yes
Sources of advice:	
EABL	Advice/support (prior to start-up), Enterprise Agency/Businesslink: 0 = No, 1 = Yes
Accountant	Advice/support, Accountant: 0 = No, 1 = Yes
Solicitor	Advice/support, Solicitor: 0 = No, 1 = Yes
College	Advice/support, College: 0 = No, 1 = Yes
SR seminar	Advice/support, (Barclays) Start Right Seminar: 0 = No, 1 = Yes

PYBT	**Advice/support, Princes Trust: 0 = No, 1 = Yes**
Family	**Advice/support, Family/friends: 0 = No, 1 = Yes**
Other	**Advice/support, Other source(s): 0 = No, 1 = Yes**
Turnover	**Turnover**
Volatility	**Volatility of turnover**
Authorized overdraft use	**Use of approved overdraft limit during the period: 0 = No, 1 = Yes**
extend of auth. OD use	**Average proportion of approved overdraft limit during the period**
Overdraft excess	**Excess use of overdraft during the period: 0 = No, 1 = Yes**
OD XS duration	**Proportion of time spent in overdraft excess during year x period y**

Ongoing data

As well as start-up and structural characteristics, any analysis of survival and growth requires data on the size of firms over time. In constructing this dataset we used credit turnover – the value of payments into a current account[17] – as our measure of size. This serves as a very close approximation to sales revenue inclusive of taxes. The much greater granularity of turnover compared with using measures of employee numbers is a particular strength. Drawing such data from bank records also provides other advantages. One is the direct observation of data without the need to survey businesses[18]. Another is the greater frequency with which the data can be observed. Every financial transaction is documented and credit turnover is available monthly rather than being limited to 12 month periods. For our dataset turnover was aggregated across six month periods from the date of start-up.

The frequency of turnover data permits the creation of variables that would not otherwise be possible, but which are vital in the testing of GR. It will be recalled this places considerable emphasis on sales not falling below a given 'reserve' level, determined by the resources available to the owner, and thus inducing exit. The variance of sales, particularly amongst very young firms is therefore likely to be a strong influence on the stay/quit decision. The unique monthly data available to us means it is possible to create a measure of turnover 'volatility' that captures this concept. We define volatility as the standard deviation of turnover for each firm across a six month period divided by total turnover[19].

In addition to the level, growth and variability of turnover, bank records also provide the opportunity to formulate measures of the quality of financial management. For this dataset the key

[17]Excluding payments from related accounts, e.g. deposit accounts held by the business.
[18]This can also be obtained from business accounts, although it can be time-consuming to access these for small firms and, in the case of the UK, a large proportion of the corporate population is not required to supply this information.
[19]This scales the measure to the size of the business.

variable was unauthorised overdraft use. The overdraft is an important financial product in the UK and has traditionally been used as the first source of working capital for small firms. With prior agreement from the relevant bank, the overdraft permits a bank customer to make payments, when the balance in their current account falls below zero. Provided that their balance remains above a given amount then the customer only pays interest on the amount of overdraft used[20]. However, customers can, in most cases, exceed their overdraft limit[21]. If this occurs the bank usually applies both a flat charge and a considerably higher interest rate to the entire balance. The financial costs to exceeding an overdraft are therefore high. While, in extreme circumstances, firms may judge that it is worth incurring these costs, persistent unauthorised use points to poor financial management on the part of the owner(s).

Our dataset includes two variables relating to unauthorised overdraft use. The first is a simple binary variable about whether the business was in unauthorised overdraft at any point during a six month period. The second records the proportion of that period spent in this position. To ensure these measures do not simply reflect more general overdraft use, i.e. that excess use provides additional information, we also include two further variables. One shows whether the firm used their overdraft at all and the other shows the mean proportion of the limit used over the period.

A full listing of the on-going variables is provided in the lower section of Table 2.

Exit and closure

As noted earlier, establishing when a business has closed is perhaps the most challenging aspect of any study of survival and growth. Even for datasets taken from near comprehensive official sources, the date at which exit occurs may be some time after actual closure[22].

When using bank records, there are two main issues to resolve. The first is to distinguish between those businesses that have closed and those that have switched to another bank. For our dataset we used Barclays closure reason codes that record why any given account has been closed. We identified 6.36% of our initial sample as having switched over the six years covered by the dataset, i.e. they had closed their account with Barclays, but continued to trade. This could be an understatement of the true number as imperfections in the coding process meant that not all switchers were recorded[23].

The second issue is judging when a given business has actually closed. While the majority of Barclays customers ceasing to trade clearly close at a specific time when no more transactions take place, an important minority become dormant, i.e. their account remains open, but with no

[20]Although there maybe a periodic charge to maintain it and the bank is able to change or withdraw it at short notice.
[21]Including where the business has no agreed limit.
[22] For example Storey et al (1987, p45), in a study of the closure of 177 Limited Companies that "failed" , identified seven decision-rules that were required to identify the year in which the enterprise ceased to trade.
[23]However, other work suggests we have found most (if not all). Fraser (2005, p90) reports that the annual rate of bank switching is just over 2% over all types of (SME) businesses.

activity.[24] For the firms in our sample we used a simple rule – if the business had shown no turnover in consecutive six month periods then it was deemed to have closed in the first of these periods.[25]

It is important to note that this process identifies closures. It is not limited to business 'failures'. By the latter we mean those firms that cease to trade with some external financial liability. Of course, as noted earlier, a closing firm may, or may not, have met the objectives of its owner(s), although closure may equally reflect that a better opportunity has presented itself.

Summary statistics

Given all the above assumptions Table 3 provides an overview of UK new ventures that began in what is now seen as the prosperous first three months of 2004.

The top half of the table shows the credit turnover (sales) at the end of each year from Q1 2004 to Q1 2010. The top line of the table shows that in Q1 2005 the new venture had arithmetic mean sales

[24]Indeed, some of these may have switched rather than closed.

[25]Some Barclays customers can show little or no activity for a number of months before seeing turnover return to non-negligible levels. This may reflect the nature of many 'micro' businesses.

of £116,724, but that this figure was heavily skewed with the median new venture sales being only £39, 276. It is also important to note that these figures are based on 5192 enterprises, compared with the 6247 that began one year earlier.

The skewness of the sales distribution is reflected in the Table. So, by the end of Year 1, there were 10% of firms with sales of less than £5800, and 10% of firms has sales of more than £260,000.

The rows in the top half of the Table show how sales change among surviving firms over the six years. The mean sales of a new venture that survives for six years is £195,173, but, as we show later, less than 40% survive to this point. Perhaps more pertinently, the sales of the median new venture peak in Year 4 and then decline, perhaps in part due to the difficult macro-economic situation in 2008[26].

The changes in firm sales over the six years are most clearly seen in the lower half of the Table.

Table 3: Summary statistics for size and growth rates

	Mean	SD	10%	25%	Median	75%	90%	Obs
	sales							
year 1	116724	529336	5734	15108	39276	105339	261042	5192
year 2	151939	591640	5750	17199	46260	129972	330283	3878
year 3	177054	693858	5967	17832	49627	143316	388179	3092
year 4	193319	623200	5880	19019	53962	163688	445426	2575
year 5	195632	574910	6194	18610	52443	156450	463580	2184
year 6	195173	713013	5530	17550	48775	150118	461199	1867
	sales gr.							
year 1	-	-	-	-	-	-	-	-
year 2	-0.035	0.912	-0.926	-0.257	0.060	0.361	0.762	3878
year 3	-0.103	0.883	-0.914	-0.289	0.026	0.244	0.568	3092
year 4	-0.094	0.847	-0.813	-0.270	0.016	0.227	0.503	2575
year 5	-0.182	0.891	-0.973	-0.372	-0.065	0.137	0.423	2184
year 6	-0.214	0.800	-0.851	-0.368	-0.080	0.086	0.359	1867

The summary statistics in Table 3 highlight that many of the firms in our sample are remarkably small, especially in comparison to conventional datasets. For example, the median turnover in year 1 is GBP 39,276 which is far smaller than the VAT threshold of GBP73,000 (above which firms start to appear in national administrative datasets). To investigate how our analysis is affected by our rich coverage of micro firms, we complement our baseline results with those obtained from restricting our sample to larger firms (that is, with above-median start-up size), hence making our sample more similar to other work on new ventures that over-samples larger firms (Yang and Aldrich, 2012).

[26] The reservation about whether or not this is a macro-economic effect is that other studies have pointed towards peak size being reached around year 4, irrespective of macro-economic circumstances (Box, 2005; Storey, 1985)

6. TESTING THE HYPOTHESES

We begin with some plots of the evolution of the R^2 statistic over time (Section 5.1), before assessing statistical significance using our Bootstrap Algorithm (Section 5.2).

6.1 Plotting the R^2 statistics

Growth

Table 4 examines the factors influencing the growth rates of new firms in the sample. The first five columns show the results for Years 2 to 6 inclusive. The remaining five columns show the results for firms with above-median start-up size only.

An examination of columns 1-5 shows the explanatory power – defined as the Nagelkerke R^2- falls over time from 0.224 in Year 2 to 0.170 in Year 6. This decline implies that the fog appears to thicken slightly rather than lift.

At the level of the individual variables, in all five equations, lagged size negatively affects the growth rate implying smaller firms have faster growth rates. However, the general absence of significant human capital variables in the equations,[27] and the low values of R^2, provides little evidence in support of SE. For example, key elements of SE such as accessing advice, or any form of prior business experience do not appear to positively enhance growth.[28] Other variables appear as significant in the equations for some years, but not consistently throughout all years, suggesting that different factors matter in different years. Examples of these include Age and Age^2 (Year 2 only); Education (Years 2, 5 and 6); Family business experience (Year 5 only); male owners (Year 2 only); multiple owners (Years 2 and 3 only).

However, two other variables or groups of variables are consistently significant. The first is legal form – showing that enterprises choosing limited company status at start-up were, in all years, more likely to show growth than those choosing to be either a partnership or sole traders.[29] Second, a range of trading variables also strongly influence growth: these are that firms with highly volatile sales on a monthly basis were less likely to grow, as were those with an overdraft facility which they exceeded.

[27] We undertook tests of joint significance of the human capital variables (that is, business experience, education and advice variables) and in most cases they were jointly insignificant, although they were jointly significant at the 5% level in year 5 and year 6.
[28] The one exception is that in Year 6 Start-Right Service does have a significantly positive sign.
[29] Choice of legal form has been shown in prior empirical work to be associated with faster growth amongst new/young firms (Storey 1994).

It will be recalled that a key element of GR was the role played by "wins", especially for survival. We capture this concept as lagged growth –implying that growth in the previous year both provides the resources for future growth directly but also indirectly through signalling the viability of the enterprise to external funders. Previous work has found that lagged growth has a significant effect on future growth as well as survival (Coad, 2009). For this reason Figures 1a and 1b takes the same time period and uses the same set of variables as in Table 4 but with the inclusion of lagged growth (at the cost of losing one year of observations). This lagged-growth variable is only marginally significant (if at all) in explaining growth, and there is no big improvement in the R^2 values in the equation.

Our second test is the examination of whether the explanatory power of our models changes over time – whether indeed the fog lifts.

Figure 1a plots the evolution of the R^2 statistics from Table 4, showing the R^2 values for four subsamples. The R^2 for the baseline sample is shown as a solid line. This we argue captures the SE Model by including a wide range of human capital variables. Other lines in Figure 1a correspond to a sample including lagged growth (at the expense of losing one year of observations),[30] as well as the subsample of firms with above-median start-up size (with or without lagged growth).

[30] We include the R^2 obtained from regressions controlling for lagged growth in the line charts, because lagged growth is a key variable for Gambler's Ruin. However, for the sake of space they are not included in the regression tables.

Figure 1a: OLS Growth regression Nagelkerke R^2 statistics for the first 6 years, for 4 different growth rate regressions.

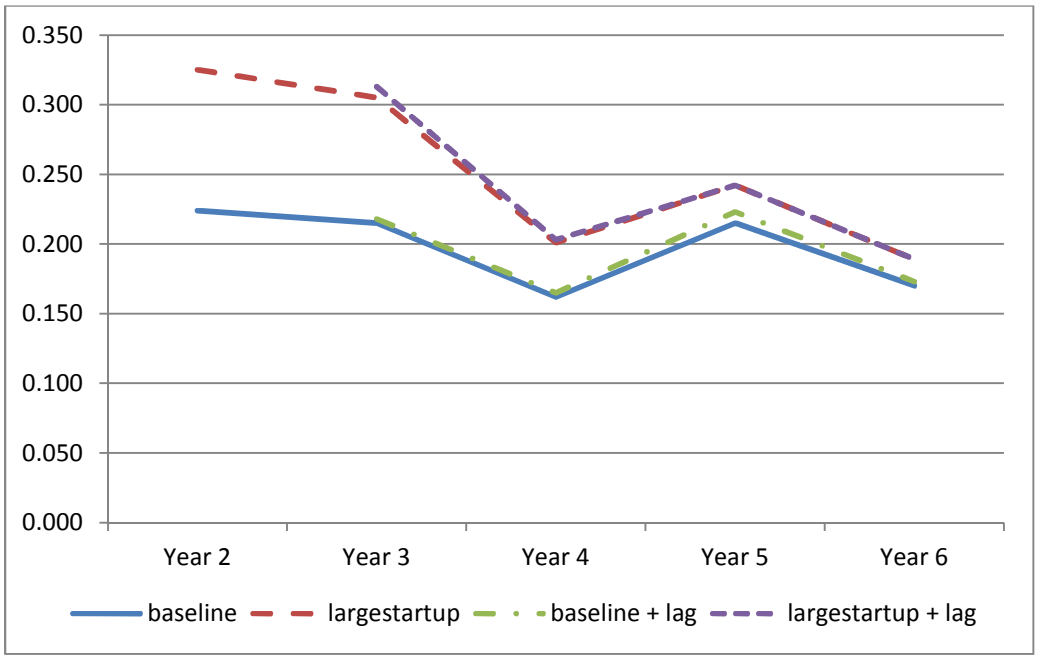
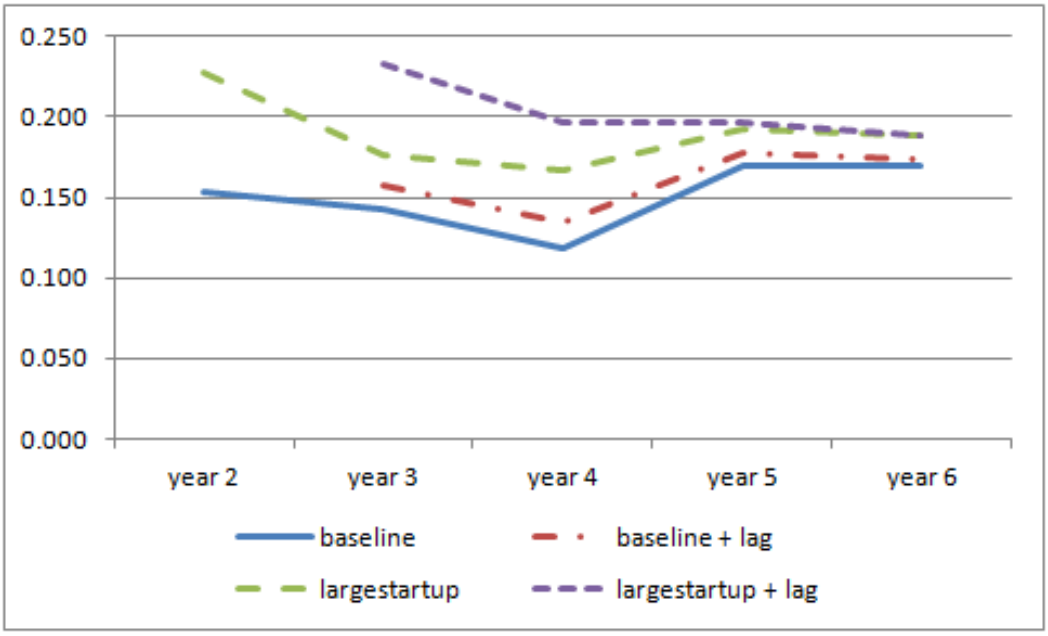


Figure 1b: OLS Growth regression Nagelkerke R^2 statistics for the first 6 years, for 4 different growth rate regressions (focus on subsample of firms that survive until the end of year 6).



KEY: Baseline: full sample. Baseline + lag: full sample controlling for lagged growth. Largestartup: above-median start-up size only. Largestartup + lag: above median start-up size subsample, controlling for lagged growth.

Figure 1a shows that the R^2 measure is generally low, ranging from 0.16 to 0.23 for the baseline estimates. Second, for all years, the Nagelkerke R^2 values for the equations that only include the

larger enterprises are slightly higher than for the baseline sample. This is in keeping with the Yang and Aldrich (2012) observation that much recent research has inflated our apparent ability to correctly explain outcomes by examining only those new enterprises sufficiently large and well-established to be included in publicly available data bases.

Figure 1b focuses on a subsample of firms that are *ex post* observed to have survived until the end of our sample period. Is it easier to predict the performance of firms that we know will survive for longer? Focusing on these *ex post* survivors allows us to leave aside the selection effect and focus only on the learning effect. Figure 1b shows non-monotonic fluctuations in the R^2 coefficient. In each case, the R^2 dips down slightly in the intermediate years before rising in the later years. In some cases ("baseline", and "baseline + lag") we observe that the final R^2 is slightly higher than the initial R^2, while in the other cases ("largestartup" and "largestartup + lag") the final R^2 is slightly lower than the initial value. All in all, however, we do not find evidence of a smoothly-increasing R^2 that would be expected from learning effects. In fact, the R^2 from regressions focusing on ex post survivors ranges from 0.12-0.23 and is lower than the R^2 obtained from focusing on the full sample (which includes short-lived firms).

Survival

To test our Hypothesis 2 – relating to survival – we run year-by-year regressions in Table 5, and plot the evolution of the R^2 statistic in Figure 2. A first observation, in line with previous work,[31] is that lagged size (that is, lagged log turnover) is a significant determinant of survival – and indeed, it is one of the few significant determinants. The only consistent human capital determinant of new firm survival is owner age. Amongst the other human capital variables, most never appear as significant in any equation – prior experience, most advice sources. Furthermore, in some years we observe somewhat unexpected signs. For example education has a negative sign in year 4 and males a positive sign in year 1.[32]

An examination of the survival equations shows, as with the growth equations, that choice of legal form – limited company status – clearly enhances survival. There are also some sectoral and regional influences, but the explanatory power of the baseline model in Table 5 is low. We therefore repeat the analysis with a subsample of above-median start-up size firms (last five columns of Table 5).

This makes a substantial difference. Apart from age, the age-related human capital variables become largely non-significant, although legal form and some regional and sectoral variables continue to be significant. Instead, it is the Trading variables that exert a powerful influence on survival in the directions expected. Survival rates are much lower in enterprises having high volatility in sales income and by two measures of overdraft excess. It also shows the powerful role of (log of) current size, with larger firms having higher survival rates.

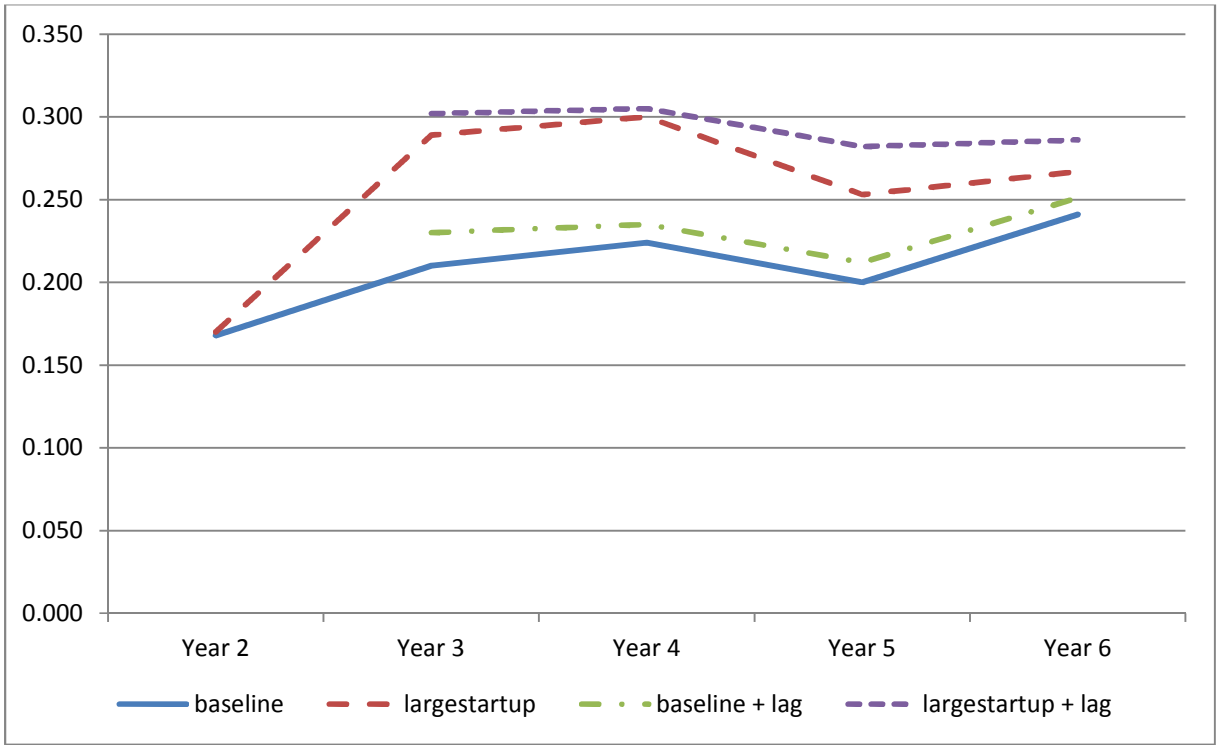
[31] See, among others, Mata and Portugal (2002).

[32] Of the two the negative education sign on survival could perhaps being explained by more educated individuals, having tried business ownership for a sufficient period to accurately assess its returns, being able more easily switch into an alternative form of employment.

In Figure 2 we show the Nagelkerke R^2 statistics for these subsamples. It has several similarities to Figure 1 but one major difference. The first similarity is that the R^2 values are broadly comparable – ranging from around 0.16 to 0.25 for the baseline sample (but rising to 0.30 for firms with above-median start-up size). The second similarity is that the inclusion of a lagged growth variable – as implied by Gamblers Ruin – improves the explanatory power of the model. The third similarity is that excluding the smallest firms also improves the values of the pseudo R^2.

The key difference is that, for three out of the four equations, values rise over the years – whereas they fell in Figure 1. It appears that, when seeking to explain new firm survival, "the fog clears" in the years after entry, but the fog remains dense when the task is to explain growth.

Figure 2: Probit survival regression: Nagelkerke R^2 statistics for years 1-6, for 4 different survival regressions



KEY: Baseline: full sample. Lag: full sample controlling for lagged growth. Large susize: above-median start-up size only. Large susize_lag: above median start-up size subsample, controlling for lagged growth.

6.2 Statistical significance

To address issues of statistical significance, we apply the Bootstrap Algorithm developed in Section 4. Figures 3 and 4 below show the distribution of bootstrapped R^2 statistics obtained for each year.

For growth, there is no monotonic trend (see Figure 3). The R^2 drops from year 2 to year 3, drops further in year 4, but picks up again in year 5 (and then drops in year 6). Year 5 (corresponding to

2008-09) stands out as a 'blip' that seems to interrupt a decreasing trend. The unexpected R^2 statistic for year 5 could be related to the severe macroeconomic conditions in 2008-09 which corresponded to the onset of a major global recession.

Figure 4 shows the corresponding plot for survival, The trend here seems to be that the goodness-of-fit for survival regressions increases in the years since start-up, although – once more – there is a 'blip' in year 5 that breaks an otherwise monotonic trend.

While Figures 3 and 4 provide a first impression, this 'eyeballing' of the changes in the distribution of R^2 needs to be complemented with more formal statistical tests of significance. Tables 6 and 7 below present the outcome of two-sample t-tests with unequal variance to see if the R^2 statistic is significantly different from the value in the first observable year (year 2), and also if it differs significantly from the value in the previous year. Table 6 presents the results for the growth regressions, while Table 7 refers to the survival regressions.

Regarding the growth rate regressions, our results show that the year-on-year change in R^2 is significantly negative – apart from year 5, which jumps upwards. When we compare each year's R^2 with the R^2 observed for the first observed period (that is, year 2) we see that the R^2 in each case is significantly lower. Taken together, the evidence suggests that the R^2 of growth rate regressions decreases over time.

Regarding the survival regressions, the R^2 generally increases over time, but it dips in year 5 to take an unexpectedly low value. Therefore the increase is not monotonic over the years. However, in each year after year 2, the R^2 is significantly higher than the year 2 value. This suggests that the 'fog' affecting survival does lift.

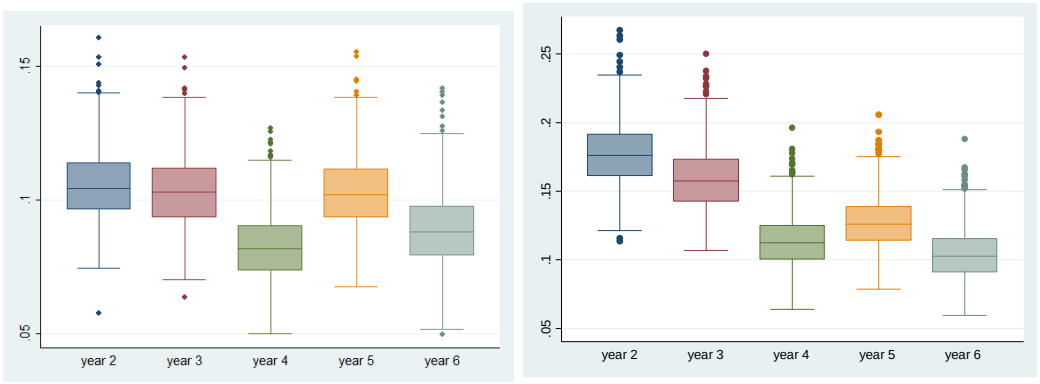


Figure 3: Evolution of bootstrapped Nagelkerke R^2 distribution for growth regressions (full model). Bootstrapped sample size n=1000, replications r=1000. Box plots refer to baseline (left) and above-median start-up size (right).

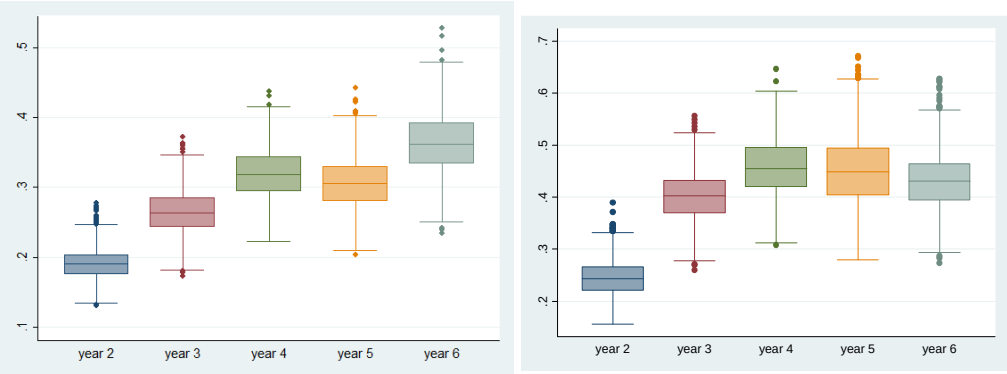


Figure 4: Distribution of bootstrapped Nagelkerke R^2 statistics obtained from probit survival regressions, n=1000, replications r=1000. Box plots refer to baseline (left) and above-median start-up size (right) samples.

7. CONCLUSION

This paper has examined a topic of key importance to individuals, governments and financial institutions – the new venture – the central characteristic of which is its diversity of performance. For each new venture that becomes a major economic player, many thousands of others have a life which is short and sometimes nasty and brutish[33].

We began by identifying two measures of new firm performance – survival and growth – and then compared the ability of two very different theories to offer insight into each performance measure. These theories we called Gamblers Ruin (GR) and Strategic Entrepreneurship (SE).

GR assumes that new venture growth/decline/stabiity is a random walk, whereas new venture survival depends on a combination of chance and access to financial resources. In contrast, SE assumes the skills/ talent/ knowledge of the owner(s) of the new venture are the key influences on both dimensions of its performance and that these qualities can be identified and used to both explain and predict both survival and growth.

On these grounds SE implies the uncertainty – the fog – over both measures of new venture performance is little more than light mist because of strong links between the qualities of the owners and venture performance. Such light mist would also be expected to clear as the owner(s) improve their knowledge through learning.

[33] Thomas Hobbes, 1588-1679

In contrast, GR implies that, since growth is a random walk, it would not be well explained by SE variables. The fog surrounding growth/decline/survival would be thick at start-up, and be expected to remain so with time. GR however expects the fog over survival to be thick initially but become less thick as surviving new ventures acquire resources that enable them to "ride out" the inevitable vicissitudes of trade that characterise the new venture in its early months and years.

Tested on a large sample of a cohort of UK new ventures over the years 2004-2010, our results show that the goodness-of-fit of growth rate regressions generally decrease in the years since start-up. However the goodness-of-fit of survival regressions increase in years since start-up. Taken together, our results suggest that there is some kind of 'uncertainty principle' at work – over time it becomes easier to predict survival but more difficult to predict growth. It is not possible to accurately predict both survival and growth – implying stronger support for GR than SE.

One unexpected, but highly interesting, result is that our results show a 'blip' for year 5 – the growth rate regression R^2 is higher than expected, and the survival R^2 is lower than expected. This blip seems to interrupt what is otherwise a monotonic trend, perhaps due to the financial crisis.

Another contribution of this paper, although it is not our main contribution, is that different factors matter for survival and growth in different years. Furthermore, we observe that many variables predicted to have a significant effect on the growth and survival of new firms (such as education, sources of advice and prior business experience) had little effect on growth or survival.

Our view is that the novelty of these results is strongly, but not exclusively, linked to the quality of the data we have available to us. Cohort data has the advantage that all new ventures face the same macroeconomic phenomena at the same stage of development, the only downside being the insurmountable identification problems of distinguishing between macroeconomic factors and developmental factors in a single cohort of firms.

What remains clear however is that raising the R^2 and, by implication, giving the impression that the fog is less thick, can likely be achieved by having samples of new firms that are unrepresentative of the population. These include having firms that are not really new; that are comparatively large; that are survivors; that have employees; that provide data themselves that cannot be verified; that are restricted to one industry or sector etc. Unfortunately, as we show, almost all other data sets on new venture performance exhibit at least one of these characteristics. The fact that our data set suffers from none of these limitations goes some way to explaining why the fog appears to be thicker amongst our firms than those examined in much other work.

As we show in the paper the fog appears to be considerably less thick once firms become larger and more established giving, what we believe to be, a misleading impression that the talents and skills of the founder(s) dominate the role of chance. Instead our view is that new venture founders can do all the "right things" and fail to survive, but many of the "wrong things" and prosper at least for some time.

The key limitation and area requiring further work is that the SE variables we use in the paper do not reflect the full diversity of Strategic Entrepreneurship. Since we do not use any self-report data we exclude key components of SE such as mind-set and aspirations. We make the case that inclusion of other key SE variables capturing human and social capital appear to play an insignificant role in our performance equations points to a greater role for chance than is implied under SE. However it clearly is an area for further work to seek to improve the coverage of SE concepts whilst linking this to the large-scale and high quality performance data available to us.

REFERENCES

Acs, Z.J. Braunerhjelm, P. Audretsch, D.B. and Carlsson, B (2009),The knowledge spillover theory of entrepreneurship, *Small Business Economics*, 32(1), 15-30

Barney, J. B. (2001), Is the Resource-Based View a useful perspective for Strategic Management research? Yes., *Academy of Management Review,* 26(1), 41-56

Bates, T. (1985), Entrepreneur Human Capital endowments and Minority Business Viability, *Journal of Human Resources*, 20,(4), 540-554

Baum, J. R. and Locke, E. A. (2004),The Relationship of Entrepreneurial Traits, Skill and Motivation to Subsequent Venture Growth. *Journal of Applied Psychology*, 89(4), 587-598

Berger,A.N. and Udell, G.F. (2006), A more complete conceptual model of SME finance, *Journal of Banking and Finance*, 30(11), 2945-2966

Bertrand, M and Mullainathan, S (2001), Do People Mean What They Say? Implications for Subjective Survey Data, *The American Economic Review*, 91(2), Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association (May, 2001) 67-72

Box, M (2005), *New Venture, Survival, Growth: Continuance, termination and Growth of Business Firms and Business Populations in Sweden during the 20th Century*, Stockholm Studies in Economic History, No 48, Almqvist and Wiksell, Stockholm

Burke, A.E., FitzRoy F.E and Nolan, M.A (2008) What makes a die-hard entrepreneur? Beyond the 'employee or entrepreneur' dichotomy, *Small Business Economics* 31(2), 93-115

Cassar, G (2010), Are individuals entering self-employment overly optimistic? An empirical test of plans and projections on nascent entrepreneur expectations, *Strategic Management Journal,* 31(8), 822-840

Casson, M.C. (1982) *The Entrepreneur: An Economic Theory*, Martin Robertson, Oxford.

Casson, M.C. and Wadeson, N (2007), The Discovery of Opportunities: Extending the Economic Theory of the Entrepreneur, *Small Business Economics*, 28(4), 285-300

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

Coad, A, Frankish, J, Roberts, R.G. and Storey, D.J. (2012) "Growth Paths and Survival chances", *Journal of Business Venturing*, forthcoming. DOI: 10.1016/j.jbusvent.2012.06.002

Cooper, A.C., Gimeno-Gascon, F.J. and Woo, C.Y. (1994) "Initial human and financial capital as predictors of new venture performance" *Journal of Business Venturing,* 9 (5), 371-395.

Cressy, R.C. (1996) Are business start-ups debt rationed?, *Economic Journal*, Vol. 106, No. 438, 1253-1270

Chrisman, J.J. and McMullan, W.E. (2004), Outsider Assistance as a Knowledge Resource for New Venture Survival, *Journal of Small Business Management*, 42(3), 229–244

Dahl, M.S. and Sorenson, O. (2012), Home Sweet Home: Entrepreneurs' Location Choices and the Performance of Their Ventures, *Management Science*, 58 (6), 1059-1071

Dawson, C and Henley, A. (2013 forthcoming), Over-optimism and entry and exit from self-employment, *International Small Business Journal,*

Fan, T P C (2010), De Novo Venture Strategy: Arch Incumbency at Inaugural Entry, *Strategic Management Journal,* 31, 19-38

Foreman-Peck, J.S (1985). "Seedcorn or Chaff? New Firm Formation and the Performance of the Interwar Economy ".*The Economic History Review*, New Series, 38(3), 402-422

Frankish, J.S, Roberts, R.G., Coad, A., Spears, T.C. and Storey, D.J.(2012), Do Entrepreneurs really learn? Or do they just tell us that they do? *Industrial and Corporate Change*, forthcoming

Fraser, S (2005), *Finance for Small and medium-sized enterprises*, CSME, University of Warwick, UK

Geroski,P.A., Mata,J and Portugal, P (2010) Founding conditions and the survival of new firms, *Strategic Management Journal,*31, 510-529. "

Gimeno, J, Folta, T.B., Cooper, A.C. and Woo, C.Y. (1997) Survival of the Fittest? Entrepreneurial Human Capital and the persistence of underperforming firms, *Administrative Science Quarterly*, 42(4), 750-783

Harrison, R. T and Leitch, C.M., 2005; Entrepreneurial Learning: Researching the Interface between Learning and the Entrepreneurial Context, *Entrepreneurship Theory and Practice*, 29(4), 351-371

Hart, P.E. and Oulton, N (1996) Growth and Size of Firms, *Economic Journal*, Vol. 106, No. 438, 1242-1252

Headd, B (2003), Redefining Business Success: Distinguishing between closure and failure", *Small Business Economics*, 21(1), 51-61

Hmieleski, K.M. and Baron, R.A. (2009), Entrepreneurs' Optimism and New Venture Performance: A Social Cognitive Perspective, *Academy of Management*, 52 (3), 473-488

Hmieleski, K.M and Corbett, A.C. (2008),The contrasting interaction effects of improvisational behavior with entrepreneurial self-efficacy on new venture performance and entrepreneur work satisfaction, *Journal of Business Venturing,*23(4), 482–496

Hurst, E and Lusardi, A (2004), Liquidity Constraints, Household Wealth, and Entrepreneurship, *Journal of Political Economy* 112 (2), 319-347

Ireland R.D., Hitt, M.A. and Sirmon D.G. (2003), "A Model of Strategic Entrepreneurship: The construct and its Dimensions", *Journal of Management*, 29 (6), 963-989

Lechner, C. and Dowling, M. (2003), Firm Networks: external relationships as sources for the growth and competitiveness of entrepreneurial firms, *Entrepreneurship and Regional Development*, 15(1), 1-26

Lee, C., Lee, K. and Pennings, J.M. (2001), "Internal Capabilities, External Networks and Performance: A Study of Technology-based ventures, *Strategic Management Journal*, 22, 615-640

Levinthal, D (1991). "Random walks and organizational mortality." Administrative Science Quarterly, 397-420.

Lindh, T. and Ohlsson, H, (1996), Self-Employment and Windfall Gains: Evidence from the Swedish Lottery, *Economic Journal* 106, No. 439, 1515-1526

Lotti, F., Santarelli, E.,Vivarelli, M (2009), Defending Gibrat's Law as a long-run regularity, *Small Business Economics*, 32, 31-44.

Mansfield E. (1962), Entry, Gibrats Law, Innovation and the Growth of Firms, *American Economic Review*, 52 (5), 1023-1051

Mata J., Portugal P. (2002). The survival of new domestic and foreign-owned firms, *Strategic Management Journal* 23, 323-343.

Muehlfeld K., Sahib P.R.and van Witteloostuijn, A. (2012), A contextual theory of organizational learning from failures and successes: A study of acquisition completion in the global newspaper industry, 1981-2008, *Strategic Management Journal*, 33: 938–964.

Nagelkerke, N.J.D. (1991), A note on a general definition of the coefficient of determination, *Biometrika,* 78 (3) 691-2.

Ongenga, S. and Smith, D.C. (2000), What determines the number of bank relationships? Cross- Country Evidence, *Journal of Financial Intermediation*, 9(1), 26-56

Pons Rotger, G., Gørtz, M. and Storey, D.J. (2012), Assessing the effectiveness of guided preparation for new venture creation and performance: Theory and Practice, *Journal of Business Venturing*, 27(4), 506–521

Plehn-Dujowich, J. (2010), A theory of serial entrepreneurship, *Small Business Economics,* 35(4), 377-398

Ritter, J. (2011), Initial Public Offerings: VC-backed IPO Statistics through 2011, University of Florida

Shane, S. (2003), *A general theory of entrepreneurship – the individual-opportunity nexus*, Edward Elgar, Cheltenham

Storey, D.J. (1985), Manufacturing Employment Change in Northern England 1965-1978, in Storey, D.J. (ed), *Small Firms in Regional Economic Development*, Cambridge University Press, London

Storey, D.J. (1994), The role of legal status in influencing bank financing and new firm growth, *Applied Economics,* 26(2), 129-136

Storey, D.J. (2011), "Optimism and chance: The elephants in the entrepreneurship room", *International Small Business Journal* 29(4), 303-32

Storey, D.J. and Greene, F.J.. (2010), *Small Business and Entrepreneurship*, Pearson, London

Stuetzer, M., Obschonka, M. andSchmitt-Rodermund, E. (2012), Balanced skills among nascent entrepreneurs. *Small Business Economics (forthcoming)*

Tanriverdi, H and Lee, C-H (2008). Within-Industry diversification and firm performance in the presence of network externalities: evidence from the software industry, *Academy of Management Journal*, 51(2), 381-397

Taylor, M (1999), Survival of the Fittest: An analysis of self-employment duration in Britain, *Economic Journal*, 109(454), C140-155

Ucbasaran, D.,Westhead, P. and Wright, M. (2009),The extent and nature of opportunity identification by experienced entrepreneurs, *Journal of Business Venturing*,24(2), 99–115

Van der Sluis, J., van Praag, M. and Vijverberg, W. (2005), Entrepreneurship Selection and Performance: A Meta-Analysis of the Impact of Education in Developing Economies, *World Bank Economic Review*, 19 (2), 225-261

Van Stel, A.J. and Storey, D.J. (2004),The Link between Firm Births and Job Creation: Is there a Upas Tree Effect? *Regional Studies* 38(4), 893-909

Watson J. (2010), *SME Performance: Separating myth from Reality*", Edward Elgar, Cheltenham, UK

Westhead, P, Ucbasaran, D. and Wright, M (2005), Decisions, Actions, and Performance: Do Novice, Serial, and Portfolio Entrepreneurs Differ? *Journal of Small Business Management***,** 393–417

Wiklund, J. Patzelt, H. and Shepherd, D.A. (2009), Building an Integrative Model of Small Business Growth*, Small Business Economics*, 32(4), 351-374

Wiklund, J. and Shepherd, D. (2005), Entrepreneurial orientation and small business performance: a configurational approach, *Journal of Business Venturing,* 20 (1), 71–91

Wilcox, J. (1971), A gambler's ruin prediction of business failure using accounting data. *Sloan Management Review,* 12(3),1-10.

Yang, T and Aldrich H.E. (2012), Out of sight but not out of mind: Why failure to account for left truncation biases research on failure rates, *Journal of Business Venturing*, 27(4), 477-492

Zimmerman, M.A. and Zeltz, G.Z. (2002), Beyond Survival: Achieving New Venture Growth by achieving Legitimacy, *Academy of Management Review*, 27(3), 414-431

Table 4: Growth Rate regressions for each year. Columns (1) to (5) contain the baseline regressors, while columns (6) to (10) focus on firms with above-median start-up size.

	Baseline					Above-median start-up size				
	year 2	year 3	year 4	year 5	year 6	year 2	year 3	year 4	year 5	year 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log turnover (lagged)	-0.173***	-0.162***	-0.146***	-0.210***	-0.120***	-0.0929***	-0.144***	-0.151***	-0.159***	-0.119***
	(0.0191)	(0.0160)	(0.0205)	(0.0357)	(0.0189)	(0.0210)	(0.0274)	(0.0304)	(0.0291)	(0.0282)
age	-0.00693**	-0.000995	0.00103	-0.00220	-0.000117	-0.00485**	0.000612	0.00161	0.00185	0.00209
	(0.00182)	(0.00181)	(0.00192)	(0.00201)	(0.00220)	(0.00194)	(0.00246)	(0.00226)	(0.00256)	(0.00316)
age squared	0.000206*	-0.000135	-0.000127	-3.14e-05	-0.000137	4.21e-05	-0.000240	-0.000240*	-0.000176	-0.000123
	(0.000116)	(0.000116)	(0.000122)	(0.000139)	(0.000133)	(0.000120)	(0.000151)	(0.000140)	(0.000169)	(0.000171)
educ_dummy_2	-0.00903	-0.122***	-0.0160	0.0549	-0.00523	-0.0292	-0.125***	-0.0488	0.0130	-0.0292
	(0.0340)	(0.0383)	(0.0415)	(0.0475)	(0.0451)	(0.0386)	(0.0458)	(0.0518)	(0.0594)	(0.0607)
educ_dummy_3	-0.0435	-0.0359	0.0158	0.0807	-0.134**	-0.0425	-0.0571	0.0187	-0.00423	-0.146*
	(0.0427)	(0.0440)	(0.0499)	(0.0599)	(0.0578)	(0.0485)	(0.0515)	(0.0541)	(0.0739)	(0.0786)
educ_dummy_4	-0.0150	-0.0639	-0.0225	0.175***	0.0198	0.0382	-0.126**	-0.0125	0.166***	-0.00932
	(0.0418)	(0.0431)	(0.0504)	(0.0540)	(0.0544)	(0.0453)	(0.0516)	(0.0579)	(0.0612)	(0.0648)
Business experience										
family	0.0106	-0.00458	-0.0551	0.0877**	0.0581	0.0211	0.0218	-0.0635	0.0997**	0.0460
	(0.0281)	(0.0303)	(0.0336)	(0.0369)	(0.0384)	(0.0313)	(0.0387)	(0.0436)	(0.0480)	(0.0522)
self	-0.0155	0.0141	-0.0199	0.0339	0.0311	-0.116***	-0.0392	0.00184	0.0492	-0.0348
	(0.0337)	(0.0345)	(0.0377)	(0.0460)	(0.0393)	(0.0366)	(0.0443)	(0.0516)	(0.0647)	(0.0606)
Sources of advice										
EABL	-0.0843	0.0591	-0.0160	0.0344	0.0152	-0.0910	-0.000983	-0.0239	0.00297	0.0646
	(0.0513)	(0.0546)	(0.0689)	(0.0615)	(0.0608)	(0.0633)	(0.0706)	(0.0672)	(0.0640)	(0.0792)
accountant	0.00892	0.00834	-0.0595*	0.0384	-0.0335	0.0236	-0.0122	-0.0881**	-0.0176	-0.0336
	(0.0293)	(0.0314)	(0.0353)	(0.0404)	(0.0388)	(0.0333)	(0.0382)	(0.0419)	(0.0437)	(0.0498)
sollicitor	0.0246	-0.0568	0.125	-0.0834	0.0863	-0.00523	-0.0973	0.205*	0.0179	0.243**
	(0.0775)	(0.0765)	(0.0793)	(0.0981)	(0.0989)	(0.0832)	(0.0976)	(0.112)	(0.119)	(0.114)
college	0.0292	0.0115	0.0641	0.0386	0.0517	-0.0498	0.0385	0.0679	-0.000410	0.130
	(0.0651)	(0.0679)	(0.0556)	(0.0739)	(0.0724)	(0.0616)	(0.0790)	(0.0743)	(0.0921)	(0.0861)
SR seminar	-0.130	0.0228	-0.0872	-0.111	0.314**	-0.0902	-0.140	-0.222	0.0925	0.340***
	(0.219)	(0.199)	(0.277)	(0.123)	(0.143)	(0.0698)	(0.254)	(0.170)	(0.173)	(0.132)
PYBT	0.0994	-0.0619	0.0756	-0.514	0.285	0.120	-0.133	-0.602	-0.940***	0
	(0.125)	(0.172)	(0.154)	(0.321)	(0.196)	(0.170)	(0.217)	(0.378)	(0.216)	(0)
family	0.0447	0.0682**	0.00866	-0.0812*	-0.0573	0.0115	0.0688	-0.0296	-0.0518	-0.0883
	(0.0312)	(0.0348)	(0.0360)	(0.0420)	(0.0423)	(0.0348)	(0.0443)	(0.0447)	(0.0511)	(0.0541)
other	0.0542	-0.0660	-0.102	0.122	-0.126	0.0852	0.104	-0.00648	0.225**	-0.179
	(0.0575)	(0.0674)	(0.0671)	(0.0831)	(0.0892)	(0.0672)	(0.0655)	(0.0817)	(0.0922)	(0.149)
volatility	-0.380***	-0.343***	-0.287***	-0.358***	-0.275***	-0.429***	-0.394***	-0.331***	-0.412***	-0.329***
	(0.0230)	(0.0280)	(0.0319)	(0.0375)	(0.0367)	(0.0332)	(0.0401)	(0.0440)	(0.0582)	(0.0506)
overdraft excess	0.0304	0.0184	0.0807**	0.0274	-0.0371	0.0102	0.0257	0.0391	0.0108	-0.0995
	(0.0322)	(0.0368)	(0.0406)	(0.0465)	(0.0474)	(0.0375)	(0.0478)	(0.0525)	(0.0521)	(0.0668)
OD XS duration	-0.502***	-0.642***	-0.659***	-0.512***	-0.571***	-0.530***	-0.897***	-0.513***	-0.520***	-0.488**
	(0.0903)	(0.0961)	(0.118)	(0.117)	(0.136)	(0.110)	(0.147)	(0.137)	(0.164)	(0.203)
Authorised OD use	0.323***	0.160***	0.0938**	0.0933*	0.0759	0.253***	0.154***	0.0874**	0.0651	0.0431
	(0.0343)	(0.0366)	(0.0446)	(0.0510)	(0.0601)	(0.0347)	(0.0412)	(0.0417)	(0.0542)	(0.0765)
Extent of auth. OC use	-0.135***	-0.124***	-0.134***	-0.0681	-0.0613	-0.0955**	-0.136***	-0.127**	-0.0387	-0.0237
	(0.0438)	(0.0430)	(0.0439)	(0.0525)	(0.0621)	(0.0404)	(0.0526)	(0.0517)	(0.0615)	(0.0859)
No. owners	0.111**	0.0868*	0.0199	-0.0272	0.0106	0.135***	0.0493	0.0316	-0.0512	0.0190
	(0.0450)	(0.0462)	(0.0487)	(0.0564)	(0.0495)	(0.0420)	(0.0511)	(0.0476)	(0.0587)	(0.0535)
Male owner(s)	0.0756**	0.0123	0.0464	0.0458	-0.0465	0.0259	0.0261	-0.0405	-0.0158	-0.0682
	(0.0371)	(0.0429)	(0.0497)	(0.0524)	(0.0473)	(0.0475)	(0.0612)	(0.0616)	(0.0638)	(0.0712)
Legal form: (omitted=company)										
partnership	-0.0807*	-0.121**	-0.101**	-0.206***	-0.0973	0.00800	-0.151***	-0.112*	-0.154**	-0.123
	(0.0425)	(0.0478)	(0.0504)	(0.0664)	(0.0650)	(0.0482)	(0.0575)	(0.0597)	(0.0708)	(0.0926)
sole trader	-0.127***	-0.140***	-0.198***	-0.181***	-0.0483	-0.0383	-0.0721	-0.246***	-0.0682	-0.0485
	(0.0386)	(0.0388)	(0.0467)	(0.0609)	(0.0515)	(0.0390)	(0.0461)	(0.0603)	(0.0588)	(0.0727)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	2.415***	2.203***	1.905***	2.685***	1.532***	1.420***	2.009***	2.234***	2.145***	1.674***
	(0.260)	(0.226)	(0.275)	(0.425)	(0.300)	(0.332)	(0.357)	(0.404)	(0.429)	(0.462)
Observations	3,625	2,898	2,426	2,066	1,763	1,965	1,620	1,378	1,209	1,037
R-squared	0.208	0.198	0.149	0.199	0.155	0.295	0.280	0.181	0.221	0.172
ll	-4382	-3404	-2829	-2492	-1960	-1989	-1759	-1459	-1325	-1138
ll_0	-4806	-3724	-3025	-2722	-2109	-2332	-2025	-1596	-1476	-1236
Cox-Snell R2	0.208	0.198	0.149	0.199	0.155	0.295	0.280	0.181	0.221	0.172
Nagelkerke R2	0.224	0.215	0.162	0.215	0.170	0.325	0.305	0.201	0.242	0.189
Robust standard errors in parentheses										
*** p<0.01, ** p<0.05, * p<0.1										

Table 5: Survival regressions for each year. Columns (1) to (5) contain the baseline regressors, while columns (6) to (10) focus on firms with above-median start-up size.

	Baseline					Above-median start-up size				
	year 2	year 3	year 4	year 5	year 6	year 2	year 3	year 4	year 5	year 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log turnover (lagged)	0.0777***	0.0634***	0.0846***	0.0945***	0.133***	0.0425	0.0417	0.0638*	0.0892**	0.182***
	(0.0179)	(0.0208)	(0.0238)	(0.0246)	(0.0278)	(0.0418)	(0.0394)	(0.0380)	(0.0403)	(0.0411)
age	0.00796**	0.00594**	0.00600*	0.0157***	0.00424	0.0128***	0.00496	0.00905*	0.0121**	0.00588
	(0.00232)	(0.00287)	(0.00351)	(0.00381)	(0.00455)	(0.00362)	(0.00431)	(0.00527)	(0.00574)	(0.00662)
age squared	-0.000286[‡]	-0.000138	-0.000241	-0.000510[‡]	-0.000456	-0.000574[‡]	-4.87e-05	-0.000400	-0.000554	-0.000666
	(0.000163)	(0.000194)	(0.000246)	(0.000258)	(0.000311)	(0.000259)	(0.000296)	(0.000353)	(0.000368)	(0.000421)
educ_dummy_2	0.0270	-0.00410	-0.0655	0.117	-0.00269	0.0583	0.0640	-0.143	0.0837	-0.0171
	(0.0571)	(0.0704)	(0.0864)	(0.0916)	(0.103)	(0.0868)	(0.101)	(0.132)	(0.134)	(0.142)
educ_dummy_3	0.0810	-0.0325	-0.0484	0.0463	-0.0554	0.265**	0.154	-0.247*	-0.0358	-0.213
	(0.0682)	(0.0811)	(0.102)	(0.108)	(0.123)	(0.107)	(0.120)	(0.143)	(0.155)	(0.163)
educ_dummy_4	-0.117*	0.0808	-0.164*	0.0711	-0.0435	-0.124	0.143	-0.172	0.00140	-0.103
	(0.0620)	(0.0790)	(0.0931)	(0.100)	(0.115)	(0.0926)	(0.113)	(0.137)	(0.143)	(0.151)
Business experience										
family	-0.0103	-0.0645	0.0488	-0.118	0.0329	0.0629	-0.00211	0.0176	-0.0417	0.0101
	(0.0441)	(0.0551)	(0.0654)	(0.0720)	(0.0800)	(0.0672)	(0.0820)	(0.0962)	(0.108)	(0.110)
self	-0.0317	0.0369	-0.0298	-0.0642	-0.0226	-0.0121	0.0582	0.0563	-0.147	0.0220
	(0.0492)	(0.0614)	(0.0749)	(0.0811)	(0.0925)	(0.0811)	(0.0962)	(0.118)	(0.137)	(0.136)
Sources of advice										
EABL	-0.0408	0.0356	-0.0743	0.229*	0.134	-0.196*	-0.220	-0.285*	0.107	0.193
	(0.0684)	(0.0893)	(0.102)	(0.122)	(0.136)	(0.114)	(0.145)	(0.157)	(0.196)	(0.219)
accountant	-0.0948**	0.0154	0.0959	-0.0382	0.0525	-0.131**	-0.0255	0.127	-0.0773	-0.0275
	(0.0455)	(0.0560)	(0.0676)	(0.0733)	(0.0809)	(0.0656)	(0.0795)	(0.0954)	(0.104)	(0.108)
sollicitor	0.159	-0.149	-0.0344	-0.131	0.0910	0.155	-0.152	-0.109	-0.0732	0.109
	(0.102)	(0.114)	(0.138)	(0.161)	(0.176)	(0.149)	(0.158)	(0.189)	(0.206)	(0.241)
college	-0.00591	0.349**	-0.0432	-0.0839	0.269	0.153	0.419*	0.00195	-0.493**	0.0977
	(0.100)	(0.145)	(0.140)	(0.170)	(0.198)	(0.157)	(0.233)	(0.208)	(0.225)	(0.267)
SR seminar	-0.0615	-0.175	0.348	-0.135	-0.671	-0.402	-0.125	0.0591	0.727	-0.405
	(0.241)	(0.331)	(0.475)	(0.419)	(0.481)	(0.433)	(0.595)	(0.622)	(0.551)	(0.903)
PYBT	-0.166	0.305	-0.608**	0.459	-0.104	-0.407	0.300	-0.998*	0.497	
	(0.194)	(0.335)	(0.259)	(0.489)	(0.401)	(0.391)	(0.555)	(0.585)	(0.606)	
family	0.0117	0.0900	-0.0979	0.128	-0.143*	0.0664	0.0614	-0.0386	0.0980	-0.0200
	(0.0479)	(0.0603)	(0.0697)	(0.0811)	(0.0865)	(0.0738)	(0.0881)	(0.103)	(0.117)	(0.117)
other	-0.0192	-0.00589	-0.237**	-0.0864	-0.0541	-0.0418	0.453**	-0.156	0.0336	-0.414**
	(0.0833)	(0.105)	(0.116)	(0.147)	(0.163)	(0.140)	(0.206)	(0.184)	(0.226)	(0.207)
volatility	-0.239***	-0.305***	-0.280***	-0.304***	-0.283***	-0.254***	-0.410***	-0.375***	-0.447***	-0.282***
	(0.0268)	(0.0314)	(0.0366)	(0.0411)	(0.0458)	(0.0455)	(0.0501)	(0.0543)	(0.0616)	(0.0668)
overdraft excess	-0.0889*	-0.127**	-0.0586	-0.0823	-0.158*	-0.126*	-0.175*	0.0270	-0.220*	-0.255**
	(0.0500)	(0.0618)	(0.0784)	(0.0850)	(0.0937)	(0.0745)	(0.0900)	(0.110)	(0.122)	(0.123)
OD XS duration	-1.181***	-0.979***	-0.852***	-0.550***	-0.867***	-1.574***	-1.450***	-1.199***	-0.461**	-0.708***
	(0.128)	(0.114)	(0.123)	(0.141)	(0.158)	(0.241)	(0.189)	(0.185)	(0.212)	(0.223)
Authorised OD use	0.126*	0.172**	-0.0148	0.0697	-0.0599	0.150*	0.395***	-0.0171	-0.0825	0.0137
	(0.0684)	(0.0811)	(0.0926)	(0.108)	(0.116)	(0.0863)	(0.109)	(0.123)	(0.137)	(0.147)
Extent of auth. OC use	-0.275**	-0.189**	-0.209**	-0.107	-0.164	-0.151	-0.135	-0.247*	-0.0438	-0.180
	(0.135)	(0.0947)	(0.102)	(0.120)	(0.120)	(0.174)	(0.125)	(0.131)	(0.145)	(0.144)
No. owners	0.0875	-0.00225	0.0797	0.330***	-0.173*	0.127	0.0269	0.253**	0.322**	-0.220*
	(0.0674)	(0.0771)	(0.0952)	(0.108)	(0.102)	(0.0882)	(0.0976)	(0.115)	(0.126)	(0.120)
Male owner(s)	0.121**	0.00766	-0.0458	0.0904	0.112	-0.0684	0.0171	-0.293**	-0.0423	0.0666
	(0.0562)	(0.0730)	(0.0861)	(0.0934)	(0.109)	(0.0994)	(0.124)	(0.142)	(0.149)	(0.161)
Legal form: (omitted=company)										
partnership	-0.284***	-0.314***	-0.368***	-0.0837	0.0183	-0.356***	-0.491***	-0.501***	-0.0338	-0.160
	(0.0705)	(0.0846)	(0.101)	(0.117)	(0.136)	(0.0965)	(0.111)	(0.130)	(0.156)	(0.173)
sole trader	-0.191***	-0.0954	-0.0766	0.0431	0.254**	-0.402***	-0.270***	-0.0421	0.0388	0.139
	(0.0566)	(0.0680)	(0.0834)	(0.0880)	(0.102)	(0.0808)	(0.0975)	(0.118)	(0.129)	(0.144)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	0.554*	0.993**	1.014**	0.220	0.0825	5.914	1.229*	1.109*	1.093	5.070
	(0.322)	(0.394)	(0.421)	(0.490)	(0.511)	(0)	(0.641)	(0.640)	(0.879)	(0)
Observations	4,858	3,625	2,898	2,426	2,066	2,438	1,965	1,620	1,378	1,208
R-squared										
ll	-2458	-1559	-1083	-872.4	-709.9	-1062	-723.7	-531.2	-414.8	-395.3
ll_0	-2752	-1817	-1288	-1019	-861.3	-1199	-913.0	-683.1	-512.8	-492.6
Cox-Snell R2	0.114	0.133	0.132	0.114	0.136	0.106	0.175	0.171	0.133	0.149
Nagelkerke R2	0.168	0.210	0.224	0.200	0.241	0.170	0.289	0.300	0.253	0.267
Robust standard errors in parentheses										
*** p<0.01, ** p<0.05, * p<0.1										

Table 6: tests of significance of changes in R^2 over the years, for growth rate regressions

baseline	Year 2	Year 3	Year 4	Year 5	Year 6
Nagelkerke R2	0.110	0.109	0.088	0.108	0.095
SE	0.016	0.016	0.015	0.017	0.017
t-stats					
comparing with prev. year		2.236	29.429	-27.987	17.329
p-value		0.025	0.000	0.000	0.000
comparing with year 2		2.236	32.112	3.616	21.108
p-value		0.025	0.000	0.000	0.000
above-median start-up size	Year 2	Year 3	Year 4	Year 5	Year 6
Nagelkerke R2	0.176	0.159	0.113	0.127	0.104
SE	0.023	0.023	0.019	0.019	0.019
t-stats					
comparing with prev. year		17.108	48.680	-16.755	27.341
p-value		0.000	0.000	0.000	0.000
comparing with year 2		17.108	68.043	52.806	77.232
p-value		0.000	0.000	0.000	0.000

Table 7: tests of significance of changes in R^2 over the years, for survival regressions

baseline	Year 2	Year 3	Year 4	Year 5	Year 6
Nagelkerke R2	0.207	0.285	0.342	0.335	0.389
SE	0.029	0.038	0.045	0.047	0.055
t-stats					
comparing with prev. year		-50.842	-30.560	3.295	-23.799
p-value		0.000	0.000	0.001	0.000
comparing with year 2		-50.842	-78.863	-72.854	-92.608
p-value		0.000	0.000	0.000	0.000
above-median susize	Year 2	Year 3	Year 4	Year 5	Year 6
Nagelkerke R2	0.245	0.402	0.458	0.452	0.432
SE	0.035	0.049	0.054	0.066	0.055
t-stats					
comparing with prev. year		-83.005	-24.327	2.366	7.023
p-value		0.000	0.000	0.018	0.000
comparing with year 2		-83.005	-104.963	-87.522	-90.967
p-value		0.000	0.000	0.000	0.000

APPENDIX

This appendix contains further robustness analysis, providing results on statistical significance of changes in the R^2 from alternative specifications.

Table A1 Growth Rate regressions for each year. Columns (1) to (5) contain the baseline regressors, while columns (6) to (9) include lagged growth as a further control variable.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Year 2	Year 3	Year 4	Year 5	Year 6	Year 3	Year 4	Year 5	Year 6
log turnover (lagged)	-0.173***	-0.162***	-0.146***	-0.210***	-0.120***	-0.147***	-0.132***	-0.191***	-0.112***
	(0.0191)	(0.0160)	(0.0205)	(0.0357)	(0.0189)	(0.0157)	(0.0200)	(0.0287)	(0.0192)
turnover growth (lagged)						-0.0701*	-0.0644	-0.118*	-0.0633
						(0.0386)	(0.0551)	(0.0683)	(0.0481)
age	-0.00693**	-0.000995	0.00103	-0.00220	-0.000117	-0.00166	0.000869	-0.00215	-0.000120
	(0.00182)	(0.00181)	(0.00192)	(0.00201)	(0.00220)	(0.00181)	(0.00190)	(0.00201)	(0.00221)
age squared	0.000206*	-0.000135	-0.000127	-3.14e-05	-0.000137	-0.000105	-0.000125	-3.16e-05	-0.000138
	(0.000116)	(0.000116)	(0.000122)	(0.000139)	(0.000133)	(0.000116)	(0.000122)	(0.000139)	(0.000133)
educ_dummy_2	-0.00903	-0.122***	-0.0160	0.0549	-0.00523	-0.125***	-0.0211	0.0516	-0.00457
	(0.0340)	(0.0383)	(0.0415)	(0.0475)	(0.0451)	(0.0384)	(0.0411)	(0.0472)	(0.0448)
educ_dummy_3	-0.0435	-0.0359	0.0158	0.0807	-0.134**	-0.0400	0.0132	0.0767	-0.129**
	(0.0427)	(0.0440)	(0.0499)	(0.0599)	(0.0578)	(0.0436)	(0.0501)	(0.0600)	(0.0577)
educ_dummy_4	-0.0150	-0.0639	-0.0225	0.175***	0.0198	-0.0676	-0.0248	0.166***	0.0248
	(0.0418)	(0.0431)	(0.0504)	(0.0540)	(0.0544)	(0.0429)	(0.0503)	(0.0537)	(0.0540)
Business experience									
family	0.0106	-0.00458	-0.0551	0.0877**	0.0581	-0.00600	-0.0571*	0.0803**	0.0591
	(0.0281)	(0.0303)	(0.0336)	(0.0369)	(0.0384)	(0.0303)	(0.0335)	(0.0369)	(0.0384)
self	-0.0155	0.0141	-0.0199	0.0339	0.0311	0.00539	-0.0272	0.0232	0.0254
	(0.0337)	(0.0345)	(0.0377)	(0.0460)	(0.0393)	(0.0349)	(0.0368)	(0.0448)	(0.0395)
Sources of advice									
EABL	-0.0843	0.0591	-0.0160	0.0344	0.0152	0.0613	-0.00580	0.0487	0.0215
	(0.0513)	(0.0546)	(0.0689)	(0.0615)	(0.0608)	(0.0542)	(0.0671)	(0.0609)	(0.0611)
accountant	0.00892	0.00834	-0.0595*	0.0384	-0.0335	0.00748	-0.0607*	0.0331	-0.0333
	(0.0293)	(0.0314)	(0.0353)	(0.0404)	(0.0388)	(0.0314)	(0.0355)	(0.0396)	(0.0388)
sollicitor	0.0246	-0.0568	0.125	-0.0834	0.0863	-0.0606	0.116	-0.0673	0.0877
	(0.0775)	(0.0765)	(0.0793)	(0.0981)	(0.0989)	(0.0759)	(0.0780)	(0.0982)	(0.0983)
college	0.0292	0.0115	0.0641	0.0386	0.0517	0.0165	0.0726	0.0403	0.0500
	(0.0651)	(0.0679)	(0.0556)	(0.0739)	(0.0724)	(0.0673)	(0.0556)	(0.0720)	(0.0723)
SR seminar	-0.130	0.0228	-0.0872	-0.111	0.314**	0.0299	-0.0860	-0.136	0.297**
	(0.219)	(0.199)	(0.277)	(0.123)	(0.143)	(0.202)	(0.275)	(0.126)	(0.136)
PYBT	0.0994	-0.0619	0.0756	-0.514	0.285	-0.0513	0.0769	-0.498	0.268
	(0.125)	(0.172)	(0.154)	(0.321)	(0.196)	(0.175)	(0.152)	(0.321)	(0.191)
family	0.0447	0.0682**	0.00866	-0.0812*	-0.0573	0.0705**	0.0105	-0.0844**	-0.0597
	(0.0312)	(0.0348)	(0.0360)	(0.0420)	(0.0423)	(0.0347)	(0.0360)	(0.0421)	(0.0421)
other	0.0542	-0.0660	-0.102	0.122	-0.126	-0.0608	-0.106	0.110	-0.119
	(0.0575)	(0.0674)	(0.0671)	(0.0831)	(0.0892)	(0.0668)	(0.0676)	(0.0819)	(0.0897)
volatility	-0.380***	-0.343***	-0.287***	-0.358***	-0.275***	-0.348***	-0.289***	-0.365***	-0.281***
	(0.0230)	(0.0280)	(0.0319)	(0.0375)	(0.0367)	(0.0282)	(0.0321)	(0.0374)	(0.0360)
overdraft excess	0.0304	0.0184	0.0807**	0.0274	-0.0371	0.0171	0.0749*	0.0304	-0.0376
	(0.0322)	(0.0368)	(0.0406)	(0.0465)	(0.0474)	(0.0368)	(0.0396)	(0.0462)	(0.0475)
OD XS duration	-0.502***	-0.642***	-0.659***	-0.512***	-0.571***	-0.643***	-0.667***	-0.526***	-0.578***
	(0.0903)	(0.0961)	(0.118)	(0.117)	(0.136)	(0.0967)	(0.118)	(0.119)	(0.136)
Authorised OD use	0.323***	0.160***	0.0938**	0.0933*	0.0759	0.163***	0.0903**	0.0842*	0.0689
	(0.0343)	(0.0366)	(0.0446)	(0.0510)	(0.0601)	(0.0367)	(0.0446)	(0.0498)	(0.0598)
Extent of auth. OC use	-0.135***	-0.124***	-0.134***	-0.0681	-0.0613	-0.124***	-0.138***	-0.0787	-0.0646
	(0.0438)	(0.0430)	(0.0439)	(0.0525)	(0.0621)	(0.0429)	(0.0444)	(0.0536)	(0.0624)
No. owners	0.111**	0.0868*	0.0199	-0.0272	0.0106	0.0867*	0.0149	-0.0434	0.00233
	(0.0450)	(0.0462)	(0.0487)	(0.0564)	(0.0495)	(0.0461)	(0.0489)	(0.0548)	(0.0494)
Male owner(s)	0.0756**	0.0123	0.0464	0.0458	-0.0465	0.0108	0.0427	0.0434	-0.0485
	(0.0371)	(0.0429)	(0.0497)	(0.0524)	(0.0473)	(0.0427)	(0.0494)	(0.0523)	(0.0469)
Legal form: (omitted=company)									
partnership	-0.0807*	-0.121**	-0.101**	-0.206***	-0.0973	-0.114**	-0.0945*	-0.203***	-0.0938
	(0.0425)	(0.0478)	(0.0504)	(0.0664)	(0.0650)	(0.0479)	(0.0500)	(0.0652)	(0.0655)
sole trader	-0.127***	-0.140***	-0.198***	-0.181***	-0.0483	-0.127***	-0.186***	-0.170***	-0.0389
	(0.0386)	(0.0388)	(0.0467)	(0.0609)	(0.0515)	(0.0387)	(0.0473)	(0.0572)	(0.0520)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	2.415***	2.203***	1.905***	2.685***	1.532***	2.066***	1.776***	2.486***	1.452***
	(0.260)	(0.226)	(0.275)	(0.425)	(0.300)	(0.225)	(0.273)	(0.361)	(0.304)
Observations	3,625	2,898	2,426	2,066	1,763	2,898	2,426	2,066	1,763
R2	0.208	0.198	0.149	0.199	0.155	0.201	0.152	0.207	0.157
ll	-4382	-3404	-2829	-2492	-1960	-3398	-2826	-2482	-1958
ll_0	-4806	-3724	-3025	-2722	-2109	-3724	-3025	-2722	-2109
Cox-Snell R2	0.208	0.198	0.149	0.199	0.155	0.201	0.152	0.207	0.157
Nagelkerke R2	0.224	0.215	0.162	0.215	0.170	0.218	0.165	0.223	0.173
Robust standard errors in parentheses									
*** p<0.01, ** p<0.05, * p<0.1									

Table A2: Growth Rate regressions for each year, for firms with above-median start-up size. Columns (1) to (5) contain the baseline regressors, while columns (6) to (9) control for lagged growth.

	year 2	year 3	year 4	year 5	year 6	year 3	year 4	year 5	year 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log turnover (lagged)	-0.0929***	-0.144***	-0.151***	-0.159***	-0.119***	-0.0929***	-0.132***	-0.163***	-0.120***
	(0.0210)	(0.0274)	(0.0304)	(0.0291)	(0.0282)	(0.0228)	(0.0286)	(0.0297)	(0.0303)
turnover growth (lagged)						-0.143**	-0.0604	0.0202	0.00727
						(0.0636)	(0.0768)	(0.0646)	(0.0668)
age	-0.00485**	0.000612	0.00161	0.00185	0.00209	-0.000521	0.00136	0.00185	0.00208
	(0.00194)	(0.00246)	(0.00226)	(0.00256)	(0.00316)	(0.00243)	(0.00229)	(0.00256)	(0.00315)
age squared	4.21e-05	-0.000240	-0.000240*	-0.000176	-0.000123	-0.000212	-0.000233*	-0.000176	-0.000122
	(0.000120)	(0.000151)	(0.000140)	(0.000169)	(0.000171)	(0.000151)	(0.000140)	(0.000169)	(0.000170)
educ_dummy_2	-0.0292	-0.125***	-0.0488	0.0130	-0.0292	-0.130***	-0.0541	0.0137	-0.0292
	(0.0386)	(0.0458)	(0.0518)	(0.0594)	(0.0607)	(0.0458)	(0.0517)	(0.0596)	(0.0608)
educ_dummy_3	-0.0425	-0.0571	0.0187	-0.00423	-0.146*	-0.0610	0.0155	-0.00412	-0.146*
	(0.0485)	(0.0515)	(0.0541)	(0.0739)	(0.0786)	(0.0515)	(0.0538)	(0.0740)	(0.0784)
educ_dummy_4	0.0382	-0.126**	-0.0125	0.166***	-0.00932	-0.127**	-0.0168	0.167***	-0.00976
	(0.0453)	(0.0516)	(0.0579)	(0.0612)	(0.0648)	(0.0517)	(0.0576)	(0.0612)	(0.0649)
Business experience									
family	0.0211	0.0218	-0.0635	0.0997**	0.0460	0.0174	-0.0652	0.101**	0.0460
	(0.0313)	(0.0387)	(0.0436)	(0.0480)	(0.0522)	(0.0383)	(0.0435)	(0.0481)	(0.0523)
self	-0.116***	-0.0392	0.00184	0.0492	-0.0348	-0.0555	-0.00325	0.0491	-0.0350
	(0.0366)	(0.0443)	(0.0516)	(0.0647)	(0.0606)	(0.0448)	(0.0507)	(0.0648)	(0.0604)
Sources of advice									
EABL	-0.0910	-0.000983	-0.0239	0.00297	0.0646	0.00897	-0.0178	0.000821	0.0642
	(0.0633)	(0.0706)	(0.0672)	(0.0640)	(0.0792)	(0.0692)	(0.0661)	(0.0640)	(0.0792)
accountant	0.0236	-0.0122	-0.0881**	-0.0176	-0.0336	-0.00844	-0.0912**	-0.0163	-0.0334
	(0.0333)	(0.0382)	(0.0419)	(0.0437)	(0.0498)	(0.0380)	(0.0419)	(0.0438)	(0.0497)
sollicitor	-0.00523	-0.0973	0.205*	0.0179	0.243**	-0.111	0.193*	0.0144	0.243**
	(0.0832)	(0.0976)	(0.112)	(0.119)	(0.114)	(0.0961)	(0.111)	(0.121)	(0.114)
college	-0.0498	0.0385	0.0679	-0.000410	0.130	0.0368	0.0744	-0.00132	0.131
	(0.0616)	(0.0790)	(0.0743)	(0.0921)	(0.0861)	(0.0799)	(0.0738)	(0.0919)	(0.0864)
SR seminar	-0.0902	-0.140	-0.222	0.0925	0.340***	-0.150	-0.239	0.0938	0.336**
	(0.0698)	(0.254)	(0.170)	(0.173)	(0.132)	(0.262)	(0.189)	(0.173)	(0.133)
PYBT	0.120	-0.133	-0.602	-0.940***	0	-0.116	-0.608	-0.940***	0
	(0.170)	(0.217)	(0.378)	(0.216)	(0)	(0.218)	(0.377)	(0.218)	(0)
family	0.0115	0.0688	-0.0296	-0.0518	-0.0883	0.0707	-0.0282	-0.0518	-0.0881
	(0.0348)	(0.0443)	(0.0447)	(0.0511)	(0.0541)	(0.0438)	(0.0448)	(0.0512)	(0.0538)
other	0.0852	0.104	-0.00648	0.225**	-0.179	0.113*	-0.00656	0.229**	-0.180
	(0.0672)	(0.0655)	(0.0817)	(0.0922)	(0.149)	(0.0661)	(0.0819)	(0.0934)	(0.150)
volatility	-0.429***	-0.394***	-0.331***	-0.412***	-0.329***	-0.407***	-0.334***	-0.411***	-0.329***
	(0.0332)	(0.0401)	(0.0440)	(0.0582)	(0.0506)	(0.0405)	(0.0440)	(0.0580)	(0.0505)
overdraft excess	0.0102	0.0257	0.0391	0.0108	-0.0995	0.0200	0.0353	0.0102	-0.1000
	(0.0375)	(0.0478)	(0.0525)	(0.0521)	(0.0668)	(0.0475)	(0.0514)	(0.0523)	(0.0665)
OD XS duration	-0.530***	-0.897***	-0.513***	-0.520***	-0.488**	-0.883***	-0.533***	-0.520***	-0.486**
	(0.110)	(0.147)	(0.137)	(0.164)	(0.203)	(0.147)	(0.143)	(0.164)	(0.202)
Authorised OD use	0.253***	0.154***	0.0874**	0.0651	0.0431	0.161***	0.0872**	0.0669	0.0434
	(0.0347)	(0.0412)	(0.0417)	(0.0542)	(0.0765)	(0.0417)	(0.0417)	(0.0545)	(0.0767)
Extent of auth. OC use	-0.0955**	-0.136***	-0.127**	-0.0387	-0.0237	-0.129**	-0.129**	-0.0379	-0.0234
	(0.0404)	(0.0526)	(0.0517)	(0.0615)	(0.0859)	(0.0523)	(0.0523)	(0.0620)	(0.0860)
No. owners	0.135***	0.0493	0.0316	-0.0512	0.0190	0.0414	0.0234	-0.0480	0.0198
	(0.0420)	(0.0511)	(0.0476)	(0.0587)	(0.0535)	(0.0502)	(0.0474)	(0.0593)	(0.0544)
Male owner(s)	0.0259	0.0261	-0.0405	-0.0158	-0.0682	0.0251	-0.0428	-0.0149	-0.0679
	(0.0475)	(0.0612)	(0.0616)	(0.0638)	(0.0712)	(0.0613)	(0.0610)	(0.0634)	(0.0710)
Legal form: (omitted=company)									
partnership	0.00800	-0.151***	-0.112*	-0.154**	-0.123	-0.137**	-0.110*	-0.153**	-0.123
	(0.0482)	(0.0575)	(0.0597)	(0.0708)	(0.0926)	(0.0577)	(0.0595)	(0.0707)	(0.0929)
sole trader	-0.0383	-0.0721	-0.246***	-0.0682	-0.0485	-0.0593	-0.240***	-0.0671	-0.0494
	(0.0390)	(0.0461)	(0.0603)	(0.0588)	(0.0727)	(0.0454)	(0.0612)	(0.0585)	(0.0730)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	1.420***	2.009***	2.234***	2.145***	1.674***	1.473***	2.041***	2.187***	1.690***
	(0.332)	(0.357)	(0.404)	(0.429)	(0.462)	(0.323)	(0.403)	(0.436)	(0.482)
Observations	1,965	1,620	1,378	1,209	1,037	1,620	1,378	1,209	1,037
R-squared	0.295	0.280	0.181	0.221	0.172	0.287	0.183	0.221	0.172
ll	-1989	-1759	-1459	-1325	-1138	-1751	-1457	-1325	-1138
ll_0	-2332	-2025	-1596	-1476	-1236	-2025	-1596	-1476	-1236
Cox-Snell R2	0.295	0.280	0.181	0.221	0.172	0.287	0.183	0.221	0.172
Nagelkerke R2	0.325	0.305	0.201	0.242	0.189	0.313	0.203	0.242	0.189
Robust standard errors in parentheses									
*** p<0.01, ** p<0.05, * p<0.1									

Table A3: Survival regressions for each year. Columns (1) to (5) contain the baseline regressors, while columns (6) to (9) include lagged growth as a further control variable.

	year 2	year 3	year 4	year 5	year 6	year 3	year 4	year 5	year 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log turnover (lagged)	0.0777***	0.0634***	0.0846***	0.0945***	0.133***	-0.0156	0.0345	0.0460*	0.0969***
	(0.0179)	(0.0208)	(0.0238)	(0.0246)	(0.0278)	(0.0233)	(0.0270)	(0.0275)	(0.0307)
turnover growth (lagged)						0.243***	0.165***	0.175***	0.145***
						(0.0345)	(0.0398)	(0.0460)	(0.0539)
age	0.00796**	0.00594**	0.00600*	0.0157***	0.00424	0.00778**	0.00611*	0.0153***	0.00446
	(0.00232)	(0.00287)	(0.00351)	(0.00381)	(0.00455)	(0.00291)	(0.00351)	(0.00381)	(0.00454)
age squared	-0.000286*	-0.000138	-0.000241	-0.000510*	-0.000456	-0.000251	-0.000224	-0.000533*	-0.000493
	(0.000163)	(0.000194)	(0.000246)	(0.000258)	(0.000311)	(0.000197)	(0.000249)	(0.000259)	(0.000310)
educ_dummy_2	0.0270	-0.00410	-0.0655	0.117	-0.00269	-0.00574	-0.0382	0.122	-0.0130
	(0.0571)	(0.0704)	(0.0864)	(0.0916)	(0.103)	(0.0706)	(0.0866)	(0.0915)	(0.103)
educ_dummy_3	0.0810	-0.0325	-0.0484	0.0463	-0.0554	-0.0124	-0.0262	0.0533	-0.0640
	(0.0682)	(0.0811)	(0.102)	(0.108)	(0.123)	(0.0819)	(0.102)	(0.108)	(0.123)
educ_dummy_4	-0.117*	0.0808	-0.164*	0.0711	-0.0435	0.0986	-0.141	0.0838	-0.0714
	(0.0620)	(0.0790)	(0.0931)	(0.100)	(0.115)	(0.0798)	(0.0934)	(0.100)	(0.116)
Business experience									
family	-0.0103	-0.0645	0.0488	-0.118	0.0329	-0.0572	0.0609	-0.102	0.0202
	(0.0441)	(0.0551)	(0.0654)	(0.0720)	(0.0800)	(0.0555)	(0.0655)	(0.0724)	(0.0801)
self	-0.0317	0.0369	-0.0298	-0.0642	-0.0226	0.0792	-0.0174	-0.0398	-0.00962
	(0.0492)	(0.0614)	(0.0749)	(0.0811)	(0.0925)	(0.0619)	(0.0751)	(0.0815)	(0.0927)
Sources of advice									
EABL	-0.0408	0.0356	-0.0743	0.229*	0.134	0.0103	-0.113	0.210*	0.105
	(0.0684)	(0.0893)	(0.102)	(0.122)	(0.136)	(0.0910)	(0.102)	(0.124)	(0.138)
accountant	-0.0948**	0.0154	0.0959	-0.0382	0.0525	0.0292	0.0989	-0.0191	0.0447
	(0.0455)	(0.0560)	(0.0676)	(0.0733)	(0.0809)	(0.0564)	(0.0680)	(0.0735)	(0.0812)
sollicitor	0.159	-0.149	-0.0344	-0.131	0.0910	-0.136	-0.0173	-0.167	0.125
	(0.102)	(0.114)	(0.138)	(0.161)	(0.176)	(0.115)	(0.141)	(0.164)	(0.179)
college	-0.00591	0.349**	-0.0432	-0.0839	0.269	0.322**	-0.0528	-0.0836	0.266
	(0.100)	(0.145)	(0.140)	(0.170)	(0.198)	(0.146)	(0.140)	(0.170)	(0.202)
SR seminar	-0.0615	-0.175	0.348	-0.135	-0.671	-0.175	0.358	-0.125	-0.663
	(0.241)	(0.331)	(0.475)	(0.419)	(0.481)	(0.325)	(0.466)	(0.450)	(0.473)
PYBT	-0.166	0.305	-0.608**	0.459	-0.104	0.251	-0.618**	0.468	-0.0533
	(0.194)	(0.335)	(0.259)	(0.489)	(0.401)	(0.332)	(0.263)	(0.512)	(0.400)
family	0.0117	0.0900	-0.0979	0.128	-0.143*	0.0738	-0.113	0.130	-0.125
	(0.0479)	(0.0603)	(0.0697)	(0.0811)	(0.0865)	(0.0604)	(0.0699)	(0.0814)	(0.0868)
other	-0.0192	-0.00589	-0.237**	-0.0864	-0.0541	-0.0349	-0.227*	-0.0714	-0.0824
	(0.0833)	(0.105)	(0.116)	(0.147)	(0.163)	(0.107)	(0.117)	(0.148)	(0.162)
volatility	-0.239***	-0.305***	-0.280***	-0.304***	-0.283***	-0.284***	-0.267***	-0.295***	-0.277***
	(0.0268)	(0.0314)	(0.0366)	(0.0411)	(0.0458)	(0.0324)	(0.0373)	(0.0424)	(0.0477)
overdraft excess	-0.0889*	-0.127**	-0.0586	-0.0823	-0.158*	-0.135**	-0.0620	-0.0899	-0.165*
	(0.0500)	(0.0618)	(0.0784)	(0.0850)	(0.0937)	(0.0622)	(0.0786)	(0.0852)	(0.0939)
OD XS duration	-1.181***	-0.979***	-0.852***	-0.550***	-0.867***	-0.932***	-0.803***	-0.512***	-0.841***
	(0.128)	(0.114)	(0.123)	(0.141)	(0.158)	(0.115)	(0.124)	(0.143)	(0.157)
Authorised OD use	0.126*	0.172**	-0.0148	0.0697	-0.0599	0.169**	0.0283	0.128	-0.0147
	(0.0684)	(0.0811)	(0.0926)	(0.108)	(0.116)	(0.0809)	(0.0926)	(0.107)	(0.116)
Extent of auth. OC use	-0.275**	-0.189**	-0.209**	-0.107	-0.164	-0.163*	-0.202**	-0.104	-0.171
	(0.135)	(0.0947)	(0.102)	(0.120)	(0.120)	(0.0949)	(0.102)	(0.120)	(0.120)
No. owners	0.0875	-0.00225	0.0797	0.330***	-0.173*	0.00550	0.0923	0.362***	-0.143
	(0.0674)	(0.0771)	(0.0952)	(0.108)	(0.102)	(0.0779)	(0.0958)	(0.109)	(0.103)
Male owner(s)	0.121**	0.00766	-0.0458	0.0904	0.112	0.0108	-0.0230	0.0929	0.110
	(0.0562)	(0.0730)	(0.0861)	(0.0934)	(0.109)	(0.0736)	(0.0862)	(0.0934)	(0.109)
Legal form: (omitted=company)									
partnership	-0.284***	-0.314***	-0.368***	-0.0837	0.0183	-0.344***	-0.381***	-0.0948	0.0247
	(0.0705)	(0.0846)	(0.101)	(0.117)	(0.136)	(0.0844)	(0.100)	(0.117)	(0.137)
sole trader	-0.191***	-0.0954	-0.0766	0.0431	0.254**	-0.168**	-0.122	0.00946	0.211**
	(0.0566)	(0.0680)	(0.0834)	(0.0880)	(0.102)	(0.0691)	(0.0846)	(0.0880)	(0.104)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	0.554*	0.993**	1.014**	0.220	0.0825	1.775***	1.498***	0.723	0.488
	(0.322)	(0.394)	(0.421)	(0.490)	(0.511)	(0.416)	(0.441)	(0.510)	(0.526)
Observations	4,858	3,625	2,898	2,426	2,066	3,625	2,898	2,426	2,066
R-squared									
ll	-2458	-1559	-1083	-872.4	-709.9	-1531	-1072	-862.6	-703.0
ll_0	-2752	-1817	-1288	-1019	-861.3	-1817	-1288	-1019	-861.3
Cox-Snell R2	0.114	0.133	0.132	0.114	0.136	0.146	0.138	0.121	0.142
Nagelkerke R2	0.168	0.210	0.224	0.200	0.241	0.230	0.235	0.212	0.251
Robust standard errors in parentheses									
*** p<0.01, ** p<0.05, * p<0.1									

Table A4: Survival regressions for each year, for firms with above-median start-up size. Columns (1) to (5) contain the baseline regressors, while columns (6) to (9) control for lagged growth.

	year 2	year 3	year 4	year 5	year 6	year 3	year 4	year 5	year 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log turnover (lagged)	0.0425	0.0417	0.0638*	0.0892**	0.182***	-0.0910*	0.00406	-0.0176	0.115**
	(0.0418)	(0.0394)	(0.0380)	(0.0403)	(0.0411)	(0.0468)	(0.0483)	(0.0464)	(0.0460)
turnover growth (lagged)						0.281***	0.140**	0.322***	0.263***
						(0.0649)	(0.0654)	(0.0717)	(0.0726)
age	0.0128***	0.00496	0.00905*	0.0121**	0.00588	0.00658	0.00913*	0.0111*	0.00561
	(0.00362)	(0.00431)	(0.00527)	(0.00574)	(0.00662)	(0.00433)	(0.00524)	(0.00568)	(0.00664)
age squared	-0.000574*	-4.87e-05	-0.000400	-0.000554	-0.000666	-7.62e-05	-0.000355	-0.000537	-0.000709*
	(0.000259)	(0.000296)	(0.000353)	(0.000368)	(0.000421)	(0.000299)	(0.000353)	(0.000366)	(0.000421)
educ_dummy_2	0.0583	0.0640	-0.143	0.0837	-0.0171	0.0623	-0.120	0.0988	-0.0340
	(0.0868)	(0.101)	(0.132)	(0.134)	(0.142)	(0.102)	(0.131)	(0.135)	(0.143)
educ_dummy_3	0.265**	0.154	-0.247*	-0.0358	-0.213	0.173	-0.227	-0.0221	-0.216
	(0.107)	(0.120)	(0.143)	(0.155)	(0.163)	(0.121)	(0.143)	(0.155)	(0.165)
educ_dummy_4	-0.124	0.143	-0.172	0.00140	-0.103	0.153	-0.138	0.0106	-0.148
	(0.0926)	(0.113)	(0.137)	(0.143)	(0.151)	(0.114)	(0.137)	(0.143)	(0.151)
Business experience									
family	0.0629	-0.00211	0.0176	-0.0417	0.0101	-0.000263	0.0217	0.00353	-0.00934
	(0.0672)	(0.0820)	(0.0962)	(0.108)	(0.110)	(0.0824)	(0.0961)	(0.109)	(0.110)
self	-0.0121	0.0582	0.0563	-0.147	0.0220	0.0972	0.0535	-0.146	0.0262
	(0.0811)	(0.0962)	(0.118)	(0.137)	(0.136)	(0.0971)	(0.117)	(0.140)	(0.139)
Sources of advice									
EABL	-0.196*	-0.220	-0.285*	0.107	0.193	-0.239	-0.296*	0.0804	0.162
	(0.114)	(0.145)	(0.157)	(0.196)	(0.219)	(0.146)	(0.158)	(0.198)	(0.219)
accountant	-0.131**	-0.0255	0.127	-0.0773	-0.0275	-0.0286	0.123	-0.0585	-0.0260
	(0.0656)	(0.0795)	(0.0954)	(0.104)	(0.108)	(0.0801)	(0.0952)	(0.105)	(0.109)
sollicitor	0.155	-0.152	-0.109	-0.0732	0.109	-0.123	-0.0891	-0.145	0.134
	(0.149)	(0.158)	(0.189)	(0.206)	(0.241)	(0.160)	(0.189)	(0.214)	(0.246)
college	0.153	0.419*	0.00195	-0.493**	0.0977	0.404*	-0.00337	-0.508**	0.104
	(0.157)	(0.233)	(0.208)	(0.225)	(0.267)	(0.236)	(0.208)	(0.228)	(0.280)
SR seminar	-0.402	-0.125	0.0591	0.727	-0.405	-0.126	0.0697	0.786	-0.431
	(0.433)	(0.595)	(0.622)	(0.551)	(0.903)	(0.591)	(0.651)	(0.520)	(0.877)
PYBT	-0.407	0.300	-0.998*	0.497		0.254	-0.996*	0.727	
	(0.391)	(0.555)	(0.585)	(0.606)		(0.550)	(0.602)	(0.555)	
family	0.0664	0.0614	-0.0386	0.0980	-0.0200	0.0561	-0.0532	0.111	-0.0157
	(0.0738)	(0.0881)	(0.103)	(0.117)	(0.117)	(0.0884)	(0.102)	(0.117)	(0.118)
other	-0.0418	0.453**	-0.156	0.0336	-0.414**	0.454**	-0.150	0.0642	-0.483**
	(0.140)	(0.206)	(0.184)	(0.226)	(0.207)	(0.203)	(0.183)	(0.230)	(0.207)
volatility	-0.254***	-0.410***	-0.375***	-0.447***	-0.282***	-0.365***	-0.359***	-0.424***	-0.247***
	(0.0455)	(0.0501)	(0.0543)	(0.0616)	(0.0668)	(0.0523)	(0.0554)	(0.0653)	(0.0700)
overdraft excess	-0.126*	-0.175*	0.0270	-0.220*	-0.255**	-0.183**	0.0281	-0.211*	-0.250**
	(0.0745)	(0.0900)	(0.110)	(0.122)	(0.123)	(0.0903)	(0.110)	(0.124)	(0.124)
OD XS duration	-1.574***	-1.450***	-1.199***	-0.461**	-0.708***	-1.383***	-1.153***	-0.510**	-0.680***
	(0.241)	(0.189)	(0.185)	(0.212)	(0.223)	(0.189)	(0.186)	(0.213)	(0.222)
Authorised OD use	0.150*	0.395***	-0.0171	-0.0825	0.0137	0.385***	0.0158	-0.00409	0.0617
	(0.0863)	(0.109)	(0.123)	(0.137)	(0.147)	(0.109)	(0.121)	(0.138)	(0.147)
Extent of auth. OC use	-0.151	-0.135	-0.247*	-0.0438	-0.180	-0.142	-0.254**	-0.0559	-0.201
	(0.174)	(0.125)	(0.131)	(0.145)	(0.144)	(0.124)	(0.130)	(0.146)	(0.144)
No. owners	0.127	0.0269	0.253**	0.322**	-0.220*	0.0411	0.276**	0.366***	-0.179
	(0.0882)	(0.0976)	(0.115)	(0.126)	(0.120)	(0.0978)	(0.117)	(0.129)	(0.123)
Male owner(s)	-0.0684	0.0171	-0.293**	-0.0423	0.0666	0.0278	-0.275*	-0.00949	0.0813
	(0.0994)	(0.124)	(0.142)	(0.149)	(0.161)	(0.125)	(0.141)	(0.150)	(0.163)
Legal form: (omitted=company)									
partnership	-0.356***	-0.491***	-0.501***	-0.0338	-0.160	-0.518***	-0.495***	-0.0354	-0.145
	(0.0965)	(0.111)	(0.130)	(0.156)	(0.173)	(0.110)	(0.129)	(0.155)	(0.176)
sole trader	-0.402***	-0.270***	-0.0421	0.0388	0.139	-0.313***	-0.0611	0.0372	0.0723
	(0.0808)	(0.0975)	(0.118)	(0.129)	(0.144)	(0.0984)	(0.118)	(0.131)	(0.146)
Industry dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	5.914	1.229*	1.109*	1.093	5.070	2.742***	1.731**	2.292**	5.972
	(0)	(0.641)	(0.640)	(0.879)	(0)	(0.699)	(0.710)	(0.912)	(0)
Observations	2,438	1,965	1,620	1,378	1,208	1,965	1,620	1,378	1,208
R-squared									
ll	-1062	-723.7	-531.2	-414.8	-395.3	-714.6	-528.7	-402.6	-387.7
ll_0	-1199	-913.0	-683.1	-512.8	-492.6	-913.0	-683.1	-512.8	-492.6
Cox-Snell R2	0.106	0.175	0.171	0.133	0.149	0.183	0.174	0.148	0.159
Nagelkerke R2	0.170	0.289	0.300	0.253	0.267	0.302	0.305	0.282	0.286
Robust standard errors in parentheses									
*** p<0.01, ** p<0.05, * p<0.1									

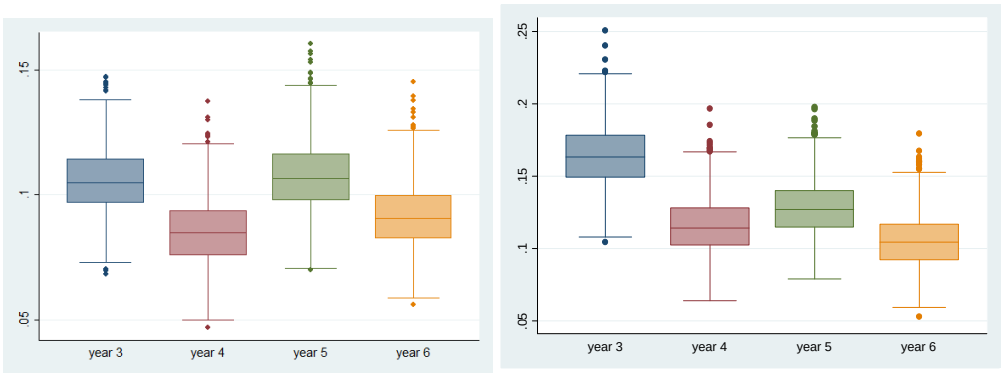


Figure A1: Evolution of bootstrapped Nagelkerke R^2 distribution for growth regressions (full model). Bootstrapped sample size n=1000, replications r=1000. Box plots refer to baseline plus lagged growth (left) and above-median start-up size with lagged growth (right).

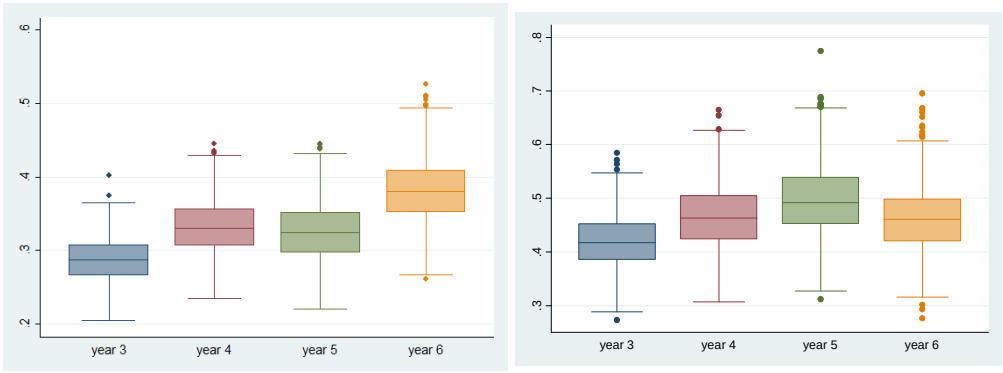


Figure A2: Distribution of bootstrapped NagelkerkeR^2 statistics obtained from probit survival regressions, n=1000, replications r=1000. Box plots refer to baseline plus lagged growth (left) and above-median start-up size with lagged growth (right).

Table A5: Bootstrapped distributions: R^2 statistics and their associated standard errors, obtained from growth rate regressions

baseline		year 2		year 3		year 4		year 5		year 6	
	R2	0.247	0.030	0.236	0.031	0.189	0.029	0.239	0.033	0.192	0.031
	adj R2	0.212	0.031	0.200	0.033	0.151	0.031	0.203	0.035	0.154	0.032
	McFadden R2	0.108	0.016	0.105	0.016	0.085	0.014	0.105	0.016	0.090	0.016
	Cox-Snell R2	0.103	0.015	0.100	0.015	0.081	0.014	0.100	0.015	0.086	0.015
	Nagelkerke R2	0.110	0.016	0.109	0.016	0.088	0.015	0.108	0.017	0.095	0.017
large_susize											
	R2	0.326	0.036	0.314	0.034	0.218	0.032	0.255	0.033	0.208	0.033
	adj R2	0.295	0.038	0.281	0.035	0.181	0.033	0.220	0.035	0.172	0.034
	McFadden R2	0.168	0.021	0.153	0.021	0.107	0.017	0.122	0.018	0.099	0.018
	Cox-Snell R2	0.160	0.020	0.145	0.020	0.102	0.016	0.116	0.017	0.094	0.017
	Nagelkerke R2	0.176	0.023	0.159	0.023	0.113	0.019	0.127	0.019	0.104	0.019
lag											
	R2			0.241	0.030	0.197	0.032	0.254	0.038	0.197	0.031
	adj R2			0.205	0.032	0.158	0.033	0.218	0.040	0.159	0.033
	McFadden R2			0.108	0.016	0.089	0.016	0.112	0.017	0.093	0.016
	Cox-Snell R2			0.103	0.015	0.084	0.015	0.106	0.017	0.089	0.015
	Nagelkerke R2			0.112	0.017	0.092	0.017	0.115	0.018	0.098	0.017
large_susize_lag											
	R2			0.323	0.034	0.222	0.033	0.257	0.034	0.210	0.033
	adj R2			0.290	0.036	0.185	0.034	0.221	0.035	0.173	0.035
	McFadden R2			0.158	0.021	0.110	0.018	0.123	0.018	0.100	0.018
	Cox-Snell R2			0.150	0.020	0.104	0.017	0.117	0.017	0.095	0.017
	Nagelkerke R2			0.164	0.022	0.116	0.020	0.128	0.019	0.105	0.019

Table A6: Bootstrapped distributions: R^2 statistics and their associated standard errors, obtained from survival regressions

baseline		year 2		year 3		year 4		year 5		year 6	
	pseudo-R2	0.148	0.021	0.189	0.025	0.211	0.027	0.200	0.028	0.231	0.032
	McFadden R2	0.148	0.021	0.189	0.025	0.211	0.027	0.200	0.028	0.231	0.032
	Cox-Snell R2	0.141	0.020	0.180	0.024	0.201	0.026	0.190	0.026	0.220	0.030
	Nagelkerke R2	0.207	0.029	0.285	0.038	0.342	0.045	0.335	0.047	0.389	0.055
susize											
	pseudo-R2	0.161	0.022	0.255	0.030	0.273	0.032	0.248	0.035	0.252	0.031
	McFadden R2	0.161	0.022	0.255	0.030	0.273	0.032	0.248	0.035	0.252	0.031
	Cox-Snell R2	0.153	0.021	0.244	0.029	0.261	0.030	0.237	0.033	0.241	0.029
	Nagelkerke R2	0.245	0.035	0.402	0.049	0.458	0.054	0.452	0.066	0.432	0.055
lag											
	pseudo-R2			0.205	0.026	0.219	0.028	0.210	0.029	0.242	0.032
	McFadden R2			0.205	0.026	0.219	0.028	0.210	0.029	0.242	0.032
	Cox-Snell R2			0.195	0.025	0.209	0.026	0.200	0.027	0.230	0.030
	Nagelkerke R2			0.309	0.040	0.354	0.046	0.352	0.049	0.407	0.054
large_susize_lag											
	pseudo-R2			0.266	0.030	0.278	0.033	0.274	0.035	0.270	0.033
	McFadden R2			0.266	0.030	0.278	0.033	0.274	0.035	0.270	0.033
	Cox-Snell R2			0.254	0.029	0.265	0.031	0.261	0.033	0.257	0.032
	Nagelkerke R2			0.418	0.049	0.465	0.057	0.498	0.066	0.462	0.059

Table A7: tests of significance of changes in R^2 over the years, for growth rate regressions

	baseline				
growth	Year 2	Year 3	Year 4	Year 5	Year 6
Nagelkerke R2	0.110	0.109	0.088	0.108	0.095
SE	0.016	0.016	0.015	0.017	0.017
t-stats					
comparing with prev. year		2.236	29.429	-27.987	17.329
p-value		0.025	0.000	0.000	0.000
comparing with year 2		2.236	32.112	3.616	21.108
p-value		0.025	0.000	0.000	0.000
		lags			
		Year 3	Year 4	Year 5	Year 6
Nagelkerke R2		0.112	0.092	0.115	0.098
SE		0.017	0.017	0.018	0.017
t-stats					
comparing with prev. year			26.792	-29.554	21.683
p-value			0.000	0.000	0.000
comparing with year 2			26.792	-3.701	18.676
p-value			0.000	0.000	0.000
	above-median susize				
	Year 2	Year 3	Year 4	Year 5	Year 6
Nagelkerke R2	0.176	0.159	0.113	0.127	0.104
SE	0.023	0.023	0.019	0.019	0.019
t-stats					
comparing with prev. year		17.108	48.680	-16.755	27.341
p-value		0.000	0.000	0.000	0.000
comparing with year 2		17.108	68.043	52.806	77.232
p-value		0.000	0.000	0.000	0.000
		above-median susize with lag			
		Year 3	Year 4	Year 5	Year 6
Nagelkerke R2		0.164	0.116	0.128	0.105
SE		0.022	0.020	0.019	0.019
t-stats					
comparing with prev. year			51.122	-13.913	26.774
p-value			0.000	0.000	0.000
comparing with year 2			51.122	38.726	63.702
p-value			0.000	0.000	0.000

Table A8: tests of significance of changes in R^2 over the years, for survival regressions

	baseline				
survival	Year 2	Year 3	Year 4	Year 5	Year 6
Nagelkerke R2	0.207	0.285	0.342	0.335	0.389
SE	0.029	0.038	0.045	0.047	0.055
t-stats					
comparing with prev. year		-50.842	-30.560	3.295	-23.799
p-value		0.000	0.000	0.001	0.000
comparing with year 2		-50.842	-78.863	-72.854	-92.608
p-value		0.000	0.000	0.000	0.000
		lags			
		Year 3	Year 4	Year 5	Year 6
Nagelkerke R2		0.309	0.354	0.352	0.407
SE		0.040	0.046	0.049	0.054
t-stats					
comparing with prev. year			-23.956	1.032	-23.642
p-value			0.000	0.302	0.000
comparing with year 2			-23.956	-21.867	-46.283
p-value			0.000	0.000	3.71e-310
	above-median susize				
	Year 2	Year 3	Year 4	Year 5	Year 6
Nagelkerke R2	0.245	0.402	0.458	0.452	0.432
SE	0.035	0.049	0.054	0.066	0.055
t-stats					
comparing with prev. year		-83.005	-24.327	2.366	7.023
p-value		0.000	0.000	0.018	0.000
comparing with year 2		-83.005	-104.963	-87.522	-90.967
p-value		0.000	0.000	0.000	0.000
		above-median susize with lag			
		Year 3	Year 4	Year 5	Year 6
Nagelkerke R2		0.418	0.465	0.498	0.462
SE		0.049	0.057	0.066	0.059
t-stats					
comparing with prev. year			-19.768	-11.832	12.794
p-value			0.000	0.000	0.000
comparing with year 2			-19.768	-30.537	-17.997
p-value			0.000	0.000	0.000