SPRU
Science and Technology Policy Research

SPRU Electronic Working Paper Number 195

Growth Paths and Survival Chances

Alex Coad (SPRU)
Julian Frankish (Barclay's Bank, UK)
Richard G. Roberts (Barclay's Bank, UK)
David J Storey (University of Sussex

The Freeman Centre (Sussex University Campus), Falmer, Brighton, BN1 9QE
United Kingdom Tel.: +44(0)1273 877943 Fax: + 44(0)1273 877977

July 2, 2011

University of Sussex
SPRU – Science & Technology Policy Research

SPRU – Science and Technology Policy Research
The Freeman Centre
University of Sussex
Falmer, Brighton, BN1 9QE
http://www.sussex.ac.uk/spru

Growth Paths and Survival Chances *

Alex Coad [a] Julian Frankish [b] Richard G. Roberts [b]
David J. Storey [a]

July 2, 2011

a University of Sussex, Falmer, Brighton, UK
b Barclays Bank, UK

Abstract

We investigate the growth and survival of nascent businesses by analyzing their bank records. We do not find strong evidence in favour of a taxonomy of growth paths, because we observe that every possible growth path seems to occur with roughly equal probability. However, we observe that survival depends on the business' growth path. Controlling for lagged size, we observe that longer lags of growth, and even start-up size, have significant effects on survival.

JEL codes: L25

Keywords: Growth paths; firm growth, firm survival, gambler's ruin, start-up size.

*We are very much indebted to Paul Nightingale and Erik Stam for helpful discussions. We gratefully acknowledge financial support from the ESRC, UK. The usual disclaimer applies.

1 Introduction

The search for a satisfactory explanation of why and when some firms grow has been extensive, but meta reviews of the findings of such research show progress has been modest. Even basic tests of explanatory power, where they are provided, point to low or very low values of R^2 in conventional regressions. Unsurprisingly, there is an absence of papers prepared to even attempt to forecast future growth rates of individual enterprises. Pulling much of that literature together, McKelvie and Wiklund (2010, p. 262) write that "[r]eviews of the growth literature tend to result in a relatively negative account of the state of affairs."

Following on from research into the determinants of growth rates, a number of scholars have sought regularities in the growth paths of firms, rather than focusing on explaining their growth in any particular year. In this vein, Delmar et al. (2003, p. 191) found that "[t]here is no such thing as a typical growth firm. Rather, there are many different types of growth firms with different growth patterns."

This paper contributes to a literature that seeks regularities in the growth paths of firms. In undertaking such a task it is vital, as McKelvie and Wiklund (2010) acknowledge, to make valid comparisons across firms. Using a unique and, close-to ideal, data set, this paper tracks the performance of a cohort of over 5'000 firms that are genuinely new and all of which began to trade in the first half of 2004. Their sales are then tracked on a monthly basis for six complete years until 2010. Most importantly, if a firm exits, there is a complete record of its performance until the exit. If it survives there is also a complete record. As far as we are aware, no other database has all of these qualities.

We investigate the possible existence of structured growth paths with respect to a randomized coin-tossing benchmark. We do not find strong evidence of any structure in growth paths, because each growth path occurs with roughly equal chances. However, a firm's growth path is seen to have a significant and long-term effect on survival. Controlling for lagged size, we observe that long lags of growth improve a firm's chances of survival.

Section 2 surveys the literature on firm growth and growth paths. Section 3 draws upon this review in presenting our theory which is heavily based on Gibrat

(1931), by contrasting it with more modern growth-path theories. This contrast is the basis of the hypotheses set out later on in the section. Section 4 presents the database used to test the hypotheses. The analysis is in Section 5, and Section 6 concludes.

2 The current evidence base

2.1 The literature on firm growth

There have been a number of comprehensive reviews of studies of firm growth generally, and of new and small firms in particular. [Henrekson and Johansson (2010) on gazelles – exceptionally fast growing (small) firms, Leitch et al. (2010) and Wiklund et al. (2009).]

The reviews emphasise the diversity of approaches that have been used to examine this topic and from this we formulate Table 1 which takes eight dimensions of this diversity and provides some brief statements to emphasise this point. When describing our own data and our approach later in Table 3 we use the same eight dimensions.

The first row of Table 1 emphasises that the sample frame from which individual firms are drawn is diverse. Some studies focus on specific types of firms – often those expected to be high performers, or those that actually were high performers, whereas others are closer to random samples of enterprises. The second dimension, and one of only two where there is almost no diversity, is that the explanatory powers of the models are low, very low, not specified, or reach reasonable levels only by including difficult to interpret interaction terms. A third source of variation is the ages of firms analysed. Some are relatively young but many are well established. Of relevance for our work, is the rarity of studies of wholly new firms that are tracked since inception. All the review studies comment on the diversity of growth metrics used, the frequent low corrleation between them when more than one metric is used, and the difficulty of justifying a single 'best' metric. There is more consensus over the importance of seeking to measure only organic growth and so eliminating sales or employment growth that comes about

Table 1: Eight dimensions of studies on firm growth

Growth Characteristic	**Illustrative studies**
Sample derivation: only HGFs?	Some studies seek to capture firms which are expected to perform well such as born globals or those in the high tech sectors (Madsen and Servais, 1997; Coad and Rao, 2008) Others select convenient samples of firms such as the study by Baum et al. (2001) of the architectural woodworking sector
Explanatory power	The explanatory power of conventional growth rate regressions is rather low (see for example Majumdar, 2004). Even including a wide range of interaction terms, Hmieleski and Baron (2009) obtain an adjusted R^2 value of 0.23. Coad (2009) finds that R^2 values above 0.15 are unusual. Even when taking into account the founder's growth motivation, only a minority of the variance can be explained (Wiklund and Shepherd, 2005).
Type of firm: new firms; large/small; sectoral/regional composition	Hmieleski and Baron (2009) examine new firms – average age 5.74 years. Hansen and Hamilton (2011) have a sample of firms in Christchurch New Zealand, in which the youngest firm was 18 years old. Acs et al. (2008) report the average high impact firm is 25 years old. More generally, Bamford et al. (2004, Table 1) provide an excellent survey on the ages of firms in previous work that focuses on nascent businesses. Many of the firms in these (small) samples are already rather old.
Measure of growth: financial, employment, self-assessed	Acs et al. (2008) define high impact to be both a doubling of sales over the last four years and a complex employment growth metric. As well as financial or employment metrics Henrekson and Johansson (2010) and Wiklund et al. (2009) also point to self-assesed measures being used.
With and without acquisition	Delmar et al. (2003) found 10% of their firms grew through acquisition and that it was growth amongst the acquired firms that was the prime contributor to growth. It is therefore important to distinguish growth through acquisition from organic growth.
Duration of growth: period of time over which it is measured	Henrekson and Johansson (2010) report the number of years over which growth was examined. Of the 19 studies specifying a date, 9 were for up to 5 years; 5 were for 6-10 years and 5 were for >10 years.
Variability of growth: year to year variability	Garnsey et al. (2006) and Garnsey and Heffernan (2005) find a wide variety of growth patterns over time. Delmar et al. (2003) also find distinct categories of high-growth firms, but the single most frequent category are what they call erratic one-shots which constitute 16.7% of all fast growers.
Persistence of growth	Coad (2007) finds small firms have negative autocorrelation.

only through acquisition. Unfortunately many of the key studies reviewed are unable to distinguish between these forms of growth. Further diversity characterises the period over which growth is measured, with the studies reviewed by Henrekson and Johansson varing from 3 to 18 years. While some studies measure growth over a year, or a number of years, others focus even on monthly growth (Davila et al., 2003). There is some consistency amongst those studies that have examined the temporal diversity of growth; virtually all emphasising that it is very rarely linear. However there is much less clarity about the patterns of diversity. Even more serious is that most studies use a binary notion of growth which eliminates any opportunity for observing temporal variation. Finally, it is important to acknowledge the possible presence of autocorrelation. In the context of the current paper, the earlier finding by Coad (2007) that smaller firms appear to exhibit negative autocorrelation, points to the need to examine carefully whether or not growth in one period is correlated with that in a later period.

2.2 Renewed impetus: searching for growth paths

The search for determinants of growth rates of firms has been huge, but progress has been slow. There remains a low explanatory power (low R^2) of conventional regressions. This has led researchers to shift their focus from identifying the determinants of growth rates to studying the mode of growth – from 'how much?' to 'how?' (McKelvie and Wiklund, 2010). Our approach is slightly different. Instead of trying to predict growth in any particular year (which is the standard approach in firm growth regressions), we look at a firm's longer-term growth history – it's growth path – in order to better understand the growth process.

The origins of growth path analysis can be found in a firm's growth rate autocorrelation profile – which in its simplest form can be seen as a growth path over a two-year period. Firms have often been observed to experience mild autocorrelation in their growth rates. Larger firms display positive autocorrelation, while smaller firms have negative autocorrelation (Coad, 2007). High-growth firms appear to be unlikely to repeat their high growth in subsequent years (Parker et al., 2010; Coad, 2007). In most cases, though, lagged growth is a poor signal of future growth.

Building on these investigations into growth rate autocorrelation, others have investigated the possibility of a more complex structure in growth rates, or any regularity in growth paths, when more than two periods of growth are considered. These studies have often attempted to classify firms into neat taxonomies of groups of firms that are arranged according to common growth paths. Table 2 contains a literature review of some previous attempts along these lines.

However, earlier research encountered the methodological problem of finding concise representations of multidimensional phenomena, when growth paths are mapped over a number of years. Often there is a high degree of arbitrariness in deciding how 'similar' firms are grouped together in contradistinction to 'dissimilar' ones. We contribute to the nascent stream of literature on firm growth paths by developing a new methodology for analyzing growth paths.

3 Theoretical background: 'Gambler's ruin' theory

This section sets out two contrasting theoretical approaches to understanding new and small firm growth. The first derives from a combination of Gibrat's Law and optimism. It argues that (new) firm growth is best seen as a random walk undertaken by optimistic entrepreneurs. The second asserts this is is incompatible with an evidence-base showing clear growth patterns linked to certain capabilities and resources leading to sustained superior performance.

The theoretical model underpinning our analysis of firm growth and survival is essentially a random process. Gibrat (1931) first suggested that firm growth could be modelled as a random walk. Although it may be, for some, hard to stomach that firm growth behaves as if it is a random process, Gibrat's Law seems to pass what we might call the 'Sherlock Holmes' test: "When you have eliminated the impossible, whatever remains, however improbable, must be the truth."[1]

[1] Sherlock Holmes in "The Sign of Four," 1890.

Table 2: Literature review of previous attempts to sort firms into a taxonomy based on their growth paths

Study	No. categories	Names of categories, and proportion in each category	Notes
McMahon (2001)	3	Low growth (70%); moderate growth (25%), high growth (5%)	Cluster analysis based on age, size and growth rate
Delmar et al. (2003)	7	Super absolute growers (13.5%); Steady sales growers (12.8%); Acquisition growers (10%); Super relative growers (16.3%); Erratic one-shot growers (16.7%); Employment growers (16.0%); Steady overall growers (14.8%)	High-growth firms only
Garnsey and Heffernan (2005)	8	Categories not named	Employment growth
Garnsey et al. (2006)	4	Continuous growth (6%), Growth setback (37%), Early growth and/or plateau (24%), Delayed take-off and growth (14%).	Employment growth
Acs et al. (2008)	6	Constant grower; Mixed grower; Non-changer; Volatile non-changer; Mixed decliner; and Constant decliner.	Employment growth. Analysis of growth before and after high-impact growth events.
Cowling and Liu (2011)	4	No growth (64%); New growth (18%); Contained growth (11%); Sustained growth (7%)	Past employment growth and future growth ambitions.

Geroski (2000) writes that: "The most elementary 'fact' about corporate growth thrown up by econometric work on both large and small firms is that firm size follows a random walk." (p. 169). As a result, we consider it to be a useful benchmark for our empirical investigations into growth paths. Previous research providing a taxonomy of heterogeneous growth paths is not in itself sufficient evidence to falsify the random growth model. A random growth process is capable of generating heterogeneous growth paths, including a minority of consecutive high-growth events. As such, our search for structured growth paths should be compared with the expected taxonomy of growth paths predicted by a random growth model. We therefore generalize Gibrat's random growth model to make predictions concerning both firm growth and also firm survival (following Levinthal, 1991). We refer to this resulting model as the 'Gambler's Ruin' model (Wilcox, 1971).

Gambler's ruin theory can be illustrated by considering a ring of gamblers gathered around a gambling table – a game of chance – each with a stock of resources (such as 'poker chips'). Although the players are involved in a game of chance, they are confident and overoptimistic about their chances of winning (Storey, 2011). Opportunities for learning are precluded, because in this model knowledge of past outcomes cannot be meaningfully applied to novel situations. A 'win' corresponds to an increase in the stock of gambling chips, while a loss decreases this resource base. Wins and losses evolve according to a random process. The player leaves the table either when their stock of resources is zero, or when they are no longer willing to risk losing more, or when they have accumulated sufficient resources to satisfy their requirements. The influences on survival – duration at the table – is some combination of accumulated resources, chance and future expectations.

Gambler's ruin theory predicts an inverted-U-shaped pattern of exit rates over a firm's first few years in business. This is because exit rates are relatively low in the first period because of the 'honeymoon of start-up capital' effect. Thus, those firms that experience a string of negative shocks at startup draw upon their initial stock of resources in order to stay in the game for a little while, with a view to estimating their likelihood of success. However, at some point in time – which varies depending on chance, resources and expectations – the resources become exhausted for the unsuccesful gambler/ enterprise and there is no choice but to quit. This inverted-U-shaped pattern of exit rates has been observed in the prior literature (Frank, 1988), and this has been explained in terms of random processes governing the fortunes of nascent businesses (Levinthal, 1991).

At first sight the Gambler's Ruin model is difficult to reconcile with learning since, by definition, the individual cannot learn to play a game of pure chance. However, although entrepreneurial learning is argued to be widespread (Politis, 2005), the Gambler's ruin model has validity in the entrepreneurial context for four reasons. The first is that, even though the entrepreneur may be alert to learning opportunities, circumstances rarely repeat themselves in an identical format from which unambiguous lessons may be drawn. This implies either a skill or luck in correctly interpreting opaque signals. Second, the skill-sets required to develop a new business may vary at inception – when the requirement may be to make a sale; somewhat later perhaps when ensuring a payment has to be made; later still perhaps when an employee has to found. All these are different skill-sets, and the opportunities for learning from experience may be modest. Thirdly, most individuals start only one or two businesses – so the opportunities to learn from earlier businesses are small. Finally, learning the 'correct' lessons may also prove particularly difficult for optimistic individuals who are very likely to attribute lack of success in their business to third parties. For all these reasons entrepreneurial learning is open to question and therefore not a basis for rejecting the lottery element in Gambler's Ruin (Frankish et al., 2010).

Denrell (2004, p. 923) acknowledges the "underestimation of the role of chance" in the business environment and warns that "[s]purious theories may be developed to account for essentially random phenomena." Mlodinow (2008) gives a fascinating account of the pervasiveness of random processes in everyday situations – and our inability to recognize such phenomena as essentially random.

We now briefly review the alternative and better known case. It is that it is reasonable to expect that firms have certain capabilities and resources that allow them to enjoy superior performance. Leading firms will have sustained superior growth, while backward firms will repeatedly have poor performance. Capabilities and Resources are captured within the concepts of Strategy and Fit. Those taking the economic approach emphasise that the ability to adjust is crucial to survival whereas the ecological approach sees inertia as the key to survival (Geroski et al., 2010).

However, there is a big difference between a firm applying existing capabilities within their existing market base, and expanding into new markets by replicating and adapting these capabilities. Growth from inception may entail novelty, the entrepreneurial pursuit of new opportunities, and this necessary novelty of growth means that learning from past experience offers no advantage for future challenges. In this situation, past growth will not help future growth.

Although scholars have not been able to predict growth in any particular year, nevertheless there might be growth paths in the longer term. A nascent body of literature has begun exploring this hypothesis. This implies:

Hypothesis 1 *Firm growth paths are not entirely random and can be sorted into a taxonomy of growth paths.*

If no support for Hypothesis 1 can be found, then we would prefer the null hypothesis of a randomized benchmark model of random growth. Gambler's Ruin theory predicts no support for Hypothesis 1.

As noted before, Gambler's Ruin theory makes predictions regarding both the survival and the growth of new businesses. In the Gambler's ruin model, firms exit when their stock of resources hits zero. As a result, larger firms have higher chances of survival because their larger stock of accumulated resources provides a buffer that protects them from imminent failure. Even if large firms experience a sequence of unfortunate events, their resource stock will not be quickly depleted.

Hypothesis 2a *Initial resources have a positive effect on survival.*

Previous work into exit rates of new businesses has observed an inverted-U-shaped relationship between exit rates and time since start-up, because of the 'honeymoon of start-up capital' effect. Firms with a smaller start-up size can be expected to have higher exit rates, and furthermore the duration of the initial honeymoon period can be expected to be shorter, because their smaller initial size constitutes a smaller buffer protecting them from exit. Firms with a larger start-up size, however, have a larger buffer stock of resources protecting them from imminent exit. As a result, the Gambler's Ruin model predicts that exit rates of firms with a large start-up size will reach their peak later than exit rates of smaller start-ups.

In the Gambler's ruin model, all that matters for survival is the stock of accumulated resources. As a result, growth since start-up is hypothesized to have a positive impact on survival, because it makes a positive contribution to the resource stock.

Hypothesis 2b *Growth paths have an effect on survival.*

Phillips and Kirchhoff (1989) showed that new enterprises with 1-4 employees at start-up, but with no employment growth, had a six-year survival rate of 26%. This rose to 65% if that firm employed even one extra worker, but the survival 'returns' were small for each subsequent marginal worker. Therefore, we expect Hypothesis 2b to be supported.

Gambler's Ruin theory is a simple model in which a firm's growth history has no effect on survival – apart from its effect on a firm's accumulated stock of resources. However, there are reasons to expect that past growth may influence survival, even controlling for (lagged) size. For example, the owner of a firm that has experienced prolonged growth may consider their firm to have a good future, whereas a 'twin' firm, of comparable size but with a less impressive recent growth history, might be disappointed with its recent performance and so be more likely to exit.

Most notably, Gimeno et al. (1997) argue that exit is strongly affected by factors other than firm size and growth. In particular, they discuss the role of aspiration levels and outside options. A business owner with attractive reservation options will not have the patience to endure low performance – she will prefer to exit. A business owner with no other outside options will be forced to keep working in their business even if they only obtain low returns from this business.

Hypothesis 2c *Growth paths have an effect on survival, even controlling for lagged size.*

Taking this line of reasoning further, we could speculate that there may even be a significant role of start-up size on survival, even survival into the distant future. A business with a small start-up size might correspond to a cash-strapped founder, while a business with a large start-up size would correspond to a well-connected founder with many other business possibilities. Start-up size would be expected to have a negative effect on survival (controlling for lagged size) because, conditional on firm size at time $t-1$, the smaller the start-up size (at time $t-s$) the higher the growth since start-up. Start-up size might therefore act as a firm-specific reference point for success, as a rough proxy for the exit threshold, even several years after start-up.[2]

The differences between Hypotheses 2b and 2c can be illustrated by looking at Figure 1. Figure 1 shows the size (growth) record of two fictitious firms, that are the same size at time t but had different growth experiences for the period $t-2:t$. Gambler's ruin theory (associated with Hypothesis 2b) would predict that both these businesses have equal survival chances, because all that matters for survival is the stock of accumulated resources. Alternative theories of firm survival, however, such as the aspiration level model in Gimeno et al. (1997) which is associated with our Hypothesis 2c, would predict that past growth has a long term effect on survival, even controlling for lagged size. Empirical support for this is provided by Geroski et al. (2010) who show that prior conditions influence current firm survival rates.

4 Database

To investigate our hypotheses relating to growth paths and survival chances, we will exploit what we claim to be an ideal database for studying the growth and survival of nascent businesses. Our database consists of customer records at Barclays Bank, covering new enterprises in England and Wales.

Conventional datasets usually have a limited coverage of new businesses. Ad-

[2]The idea that start-up size acts as a proxy for exit threshold does not take into account the possibility of firms that start small but that have high aspirations, however.

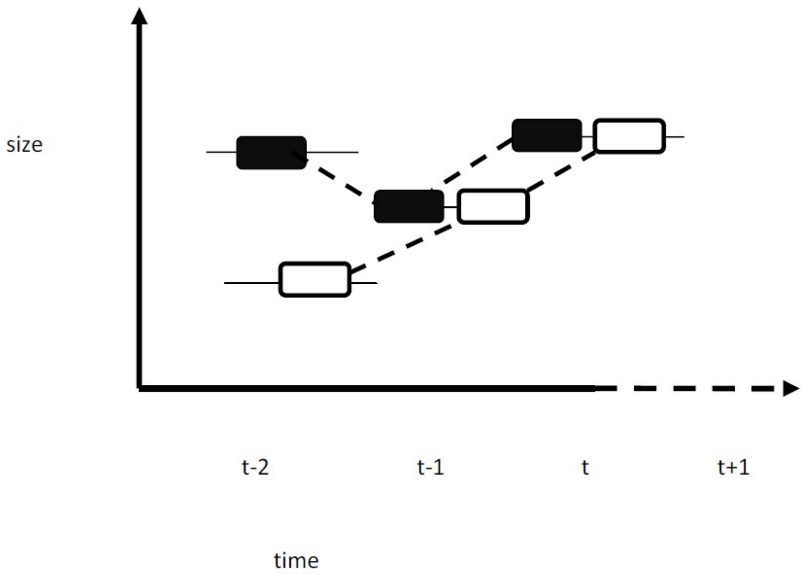

Figure 1: Hypothesizing the relationship between growth, size and survival: do these two businesses have equal survival chances for the period t:t+1?

ministrative datasets, produced by governments either from sample employment censuses or from other administrative or tax related records, have poor coverage of new businesses because they aim to reduce the bureaucratic burden facing small firms. There may also be size thresholds below which registration is not required.[3] In addition, small firms are often associated with accidental misreporting and deliberate tax evasion, which might be a source of measurement error in some datasets. There certainly is no incentive on the part of the enterprise to ensure the information is either correct or timely, so less than annual surveys are very unusual.

Survey datasets (from sources such as telephone or mail surveys), despite being widely used, are not entirely reliable and risk being non-random, and prone to bias with those exiting being less likely to be included – particularly in the run-up to the exit.

[3]For example, in the UK, the threshold for Value-Added Tax (VAT) registration is at about £70,000 per annum.

Finally, for new firms, there remains a doubt about when such firms should be included as 'new'. For example the Global Economic Monitor (GEM) regards a business as new when it becomes profitable; governments view them as new when they register with the authorities; self-employed individuals are new when they did not previously classify themselves as self-employed. Clearly these definitional differences considerably influence the number of new firms.[4]

Our data suffers none of these limitations. It exploits information from business current account records at Barclays Bank. Barclays provides the primary current account facility for just over 20% of all businesses in England and Wales with sales of less than £1 million. Their active customer base in this market is in excess of 500,000 firms, with the Bank and the enterprises having strong commercial incentives to ensure that the data is accurate and timely.

It is important to note that the opening of a business current account is **not** conditional on the provision of any other banking service such as a deposit account, overdraft facility, or term loan.[5]

Our dataset comprises a randomly-selected sample of all new business bank accounts opened with Barclays between March and May 2004, and tracks them for the six years till 2010. As part of the provision of account facilities new business owners were required to complete a questionnaire relating to their prior employment and educational attainment, together with some personal details such as age and gender, as well as information on the sources of advice or support approached prior to start-up. These data are later used to control for the influence of human capital, previous experience, and personal characteristics of the founder. In addition the bank has information on the legal form of the business, the sector in which it operates and its location. These data are described in more detail in Table 7.

While administrative datasets (such as the Companies House data in the UK) have difficulty distinguishing between date of company registration and commencement of trading activities, the bank is able to identify the date of commencement

[4]To illustrate: in the UK Barclays Bank estimated there were almost 500,000 new firms in 2003/4; new self-employed were nearly 400,000; GEM estimated approximately 260,000; Official government figures were about 180,000 (Storey and Greene, 2010, p. 115).

[5]It is also important to note that in England and Wales it is virtually impossible for a business to trade without access to a bank account.

of activities. Only firms showing trading activity over the period April-June 2004 are included, so dormant businesses are excluded. We believe the sample to be as close as possible to being a representative sample of new firm starts in England and Wales in that period. Only the financial services sector is excluded.

Our main variable, used to measure business size (and growth), is 'credit turnover' and it corresponds to sales revenue. We therefore follow the suggestions in Hamilton (2011, p. 9) and Delmar et al. (2003, p. 194), to base our analysis of business growth paths by considering sales growth rather than employment growth. Sales growth is calculated in the usual way by taking log-differences of annual sales.

Uniquely we are also able to observe a number of time-varying business-specific variables relating to the individual bank accounts, such as volatility of monthly turnover, use of overdraft, and also unauthorized overdraft excess. Although we are able to observe data at biannual intervals, our analysis focuses on annual growth rates to avoid complications due to seasonality, and to maintain comparability with previous work.

Table 3 provides an overview of our database and it is structured so as to make a direct comparison with Table 1 to show explicitly how we address the key areas of concern derived from earlier studies covered in Table 1. More details on the variables can be found in Appendix A.

4.1 Summary statistics

To familiarize the reader with our dataset, the evolution of the cohort's first six years is tracked in Table 4. The upper half of the table shows firm sales, the lower half shows sales growth. The main characteristic of these new firms is their small size. When they begin, median sales are £39,276. This is virtually half that required to register for Value-Added Tax (VAT) in the UK and hence appear in official statistics. As time goes by, the average size increases slightly, as does the standard deviation. Five years later, even after two-thirds of firms are no longer trading, median sales are only £48,775 – emphasizing both the small scale of such enterprises and the absence of growth. Most of the change in the firm

Table 3: Overview of our database

Growth Characteristic	Our database
Sample derivation: only HGFs?	Representative sample of all new business starts with Barclays Bank, which has a 20% market share in England and Wales.
Explanatory power:	Pseudo R^2 ranges from 0.101 to 0.163
Type of firm: new firms; large/small; sectoral/regional composition	All sectors apart from financial services. All regions in England and Wales are included. New enterprises are identified directly they begin to trade as a business
Measure of growth: financial, employment, self-assessed	Sales are measured on a monthly basis
With and without acquisition	Since they are new there is almost no acquisition
Duration of growth: period of time over which it is measured	Sales measured on monthly basis but then aggregated to quarterly, six-monthly and then annual basis for up to six years for firms that survive for that time.
Variability of growth: year to year variability	Volatility of sales on a six-monthly and annual basis are calculated
Persistence of growth	Estimated from the data

Table 4: Evolution of firm size and growth rate over the cohort's first six years.

	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 growth								
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

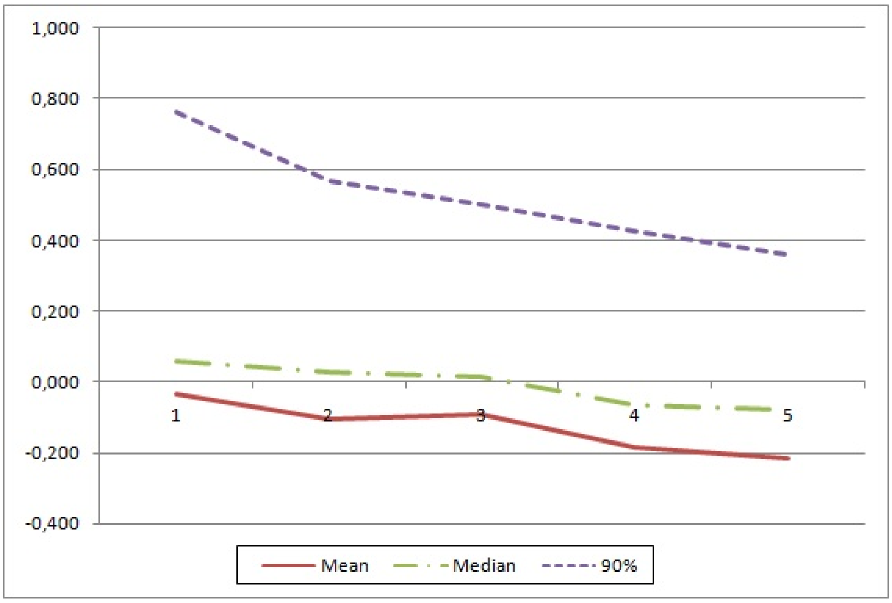

Figure 2: Evolution of the growth rate distribution

size distribution is visible at the 90% decile, where firms almost double in size. In contrast to the firm size distribution, the standard deviation of the growth rates distribution decreases over time, which is consistent with the hypothesis that growth patterns become more orderly over time. Even in the sixth year, however, there is a large disparity in growth rates. By that point, it appears that 'high-growth' events (at the 90% decile) are less spectacular than in previous years, while 'rapid-decline' events (at the 10% decile) remain similar to what was observed in previous years.

A striking feature of the growth rate distribution is that the mean growth rate is negative in each of the growth years in our sample. Figure 2 plots the evolution of the mean, median and 90% percentile of the growth rates distribution. The mean growth rate is negative in each year, and tends to become more negative over time. The median growth rate is positive in the first few years but becomes negative towards the end of our period of analysis. The 90% decile of growth rates also decrease over time.

Further information on the characteristics of our sample can be found in Appendix B, which shows how the composition of the dataset changes over the first six years in terms of regions and sectors.

Our analysis essentially focuses on firms that survive their first five years of business (i.e. their first four growth periods), and tracks their growth performance over these five years. With this database, we consider ourselves to be in a position to address concerns that research into business ventures should focus on a wide range of industries (Bamford et al., 2004, p. 916), and analyze larger sample sizes (Pompe and Bilderbeek, 2005, p. 865).

4.2 Concepts of business exit

Recent contributions to the literature on (small/new) business exit have shown that exit can be seen as both a success and a failure (Headd, 2003; Bates, 2005; DeTienne et al., 2008; Parker et al., 2010). In some cases, the survival of low-performance firms may be due to the obstinate persistence of their owners rather than the existence of attractive business opportunities (DeTienne et al., 2008). Wennberg et al. (2010) explore a range of heterogeneous exit routes – sale of the business (harvest sale, distress sale) and liquidation (harvest liquidation and distress liquidation) – and show that exit cannot always be narrowly equated with business failure.

In our context, however, it seems reasonable to take exit as equivalent to enterprise failure. Since we track businesses since their founding, we do not expect many business exits to be 'success stories' so soon after startup. In our data there are no Initial Public Offerings (IPOs), and acquisitions are negligible.[6] Furthermore, the vast majority of the businesses in our data are too small to be likely to be the objects of business sales.

Our assertion is that closure within six years was not the prime intention of the vast majority of those when starting the business, and so in this sense it can be classified as a forced closure. It is therefore very different from those small

[6]In a separate analysis, we observe that the determinants of survival to not vary greatly across years for the six years coverage of our data, which is consistent with the view that exit can be equated with business failure in our data.

business closures where the owner 'harvests' their wealth by selling or when an ageing owner has a deliberate plan to retire from the business.[7]

Nevertheless, to investigate a possible misspecification of our exit variable with regard to successful exits, we repeat our analysis on a restricted sample in which firms that are suspected of being most likely to experience successful exit are dropped. Previous work has shown that harvest sales are more likely to occur in the cases where firms are large, but that there is no clear relationship between harvest sales and education (Wennberg et al., 2010). We also suspect that successful exit is more likely to affect businesses that have experienced rapid growth in their recent history. As a result, we drop firms with more than £300,000 in sales in the year of their exit, and that experienced positive sales growth in the previous two periods. Repeating our multivariate regressions from Section 5.2 on this subsample does not change our main results.

One type of exit that is more relevant is switching to a rival bank, but this is not a common event. Fraser (2005) finds only 2% of all UK SMEs switched banks in the previous three years. We find that 1.40% of the firm closures in the sixth year of business (i.e. the year we focus on in our survival analysis) are 'switchers'. All switching businesses are identified and dropped from our analysis.

5 Testing the hypotheses

5.1 A coin-flipping benchmark

In this section, we introduce a new methodology that is tailor-made for our research question. The main problem we face is to devise a concise representation of growth paths, when the number of possible growth paths increases exponentially with the number of years considered. We organize our analysis by referring to the median rate of sales growth. In any given year, half of the businesses will be characterized as 'growth' firms (because they experience above-median growth), while the other

[7]For example, Taylor (1999) finds that retirement was the reason for exiting from self-employment for 35% of those in business for 10 or more years. For those in business for 1 year, this was 0%. Similar striking differences are found by Harada (2007)

half will be said to experience 'decline'.[8]

5.1.1 Growth paths

Figure 3 contains our basic findings relating to the structure of growth paths (Hypothesis 1). It plots the number of firms in each growth category with the shaded boxes being years in which sales fell (relative to the median) and white/unshaded areas showing increased sales. So, for example, at the top left hand side the four shaded boxes show the proportion of the sample – 6.50% – that experienced declines in every year since startup. Conversely, at the bottom left hand side are the 7.22% of firms that increased their sales in each year since startup.

There are two key results. The first has already been noted. It is that the probability of a new firm growing its sales in four consecutive years is only 7% – emphasising the rarity of consistent and linear growth.

Our second key result is that every possible growth path is more or less equipopulated – all growth paths occur with roughly equal probability. The most populated category is 'growth-growth-growth-decline' (occuring in 7.58% of cases) while the least populated category is 'growth-decline-decline-growth' (occuring in 5.24% of cases). In some cases, we observe growth in each of the four consecutive periods we analyze. This does not contradict a random model, however, because even when flipping a coin we would expect four consecutive heads in 1/16 (6.25%) of the cases.

For example, the probability of observing 'decline-decline-decline-growth' is 7.49%, while the probability of observing 'growth-decline-decline-growth' is 5.24%. On this basis, if we only observe the growth of a nascent business over growth periods 2, 3 and 4 but do not observe how the firm grew in the first growth period, then it is in fact marginally more likely to have experienced growth in the first growth period also.

In order to assess the statistical significance of the results in Figure 3, we

[8] We note that this kind of distinction between growth and decline (with the median as the threshold) is not possible when employment growth is the metric, because indivisibilities (integer restrictions) in terms of employees meant that many firms in the central mass of the growth rate distribution had employment growth of exactly 0.00%.

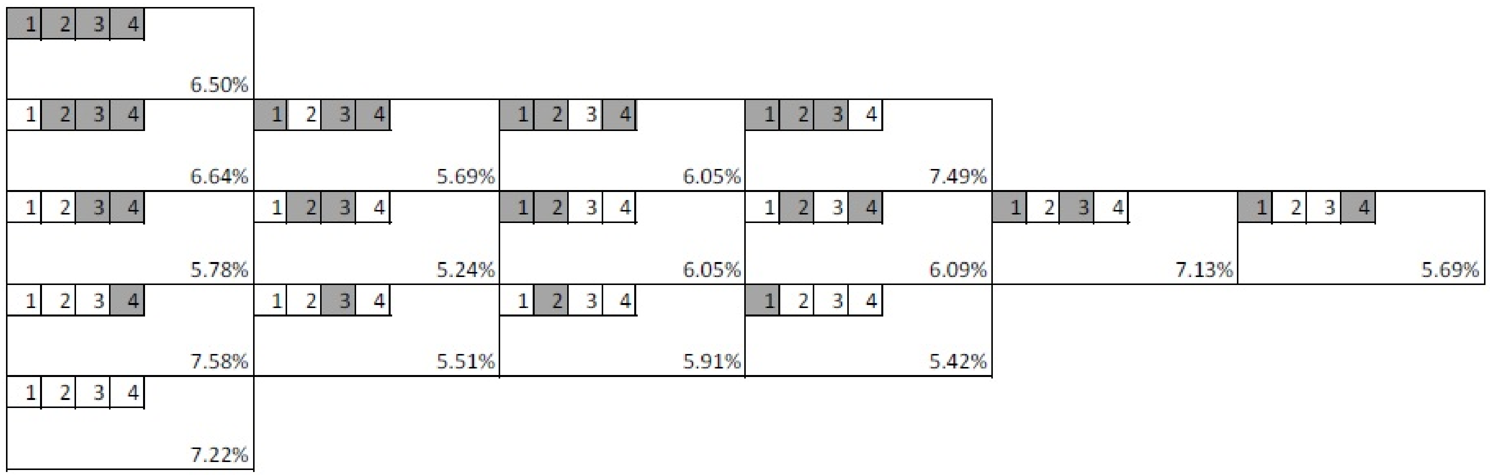

Figure 3: Growth paths of firms surviving until the end of the fifth year. N=2184. **Legend**: White squares correspond to growth years (above-median growth) while dark squares correspond to decline years (below-median growth), for the first four growth years of each business. Frequencies are shown in each box. If there is no structure in growth rates (i.e. a purely random process), each of the 16 growth paths should occur with probability 1/16 = 6.25%.

proceed as follows. Given that the variance of a Bernouilli process is $n \cdot p \cdot$ (1-p), and we take n=2184 and p=$0.5^4 = 0.0625$, assuming each p to be an independent draw,[9] the standard deviation for the proportion of the population of each of the 16 growth path configurations is $\sqrt{((2184 \times 0.0625 \times 0.9375)/2184)} = 0.07197$. Hence the 95% confidence interval around the expected value is 6.25% $\pm$ (1.96 $\times$ 0.07197%), which ranges from 6.109 to 6.391. In the light of the estimated confidence intervals, we see that, in fact, *all* of the growth path configurations are statistically different from the expected value of 6.25%. While a simple inspection of the numbers corresponding to each growth path would suggest that every possible growth path occurs with roughly equal probability, in statistical terms these differences are significant. This appears to be an artifact of our large sample size – differences that appear small in practical terms are in fact significant in statistical terms. Thus far, we cannot reject Hypothesis 1 that growth paths are not random. In the light of further robustness analysis in Section 5.1.3, however, we remain cautious about the possible existence of regularities in growth paths.

5.1.2 Growth paths and survival

We begin our analysis of factors affecting survival by investigating Hypothesis 2a, which predicted that start-up size has a positive influence on survival. Figure 4 plots the exit rates for the firms in our sample. Here we see that exit rates generally decrease over the 12 six-month periods in our sample. Looking first at the exit rates for the full sample, we observe the familiar inverted-U-shape of exit rates, which signals that firms enjoy a honeymoon effect immediately after entry. This honeymoon effect is consistent with the Gambler's Ruin model. Figure 4 also plots the exit rates for two groups of firms, sorted according to the median start-up size. Firms with a larger start-up size have generally lower exit rates, and their exit rates peak about 4 six-month periods after start-up. For firms with a smaller start-up size, however, exit rates are higher and they peak much earlier. In fact,

[9] Strictly speaking, each p is not exactly an independent draw, because of restrictions stemming from our choice of median as threshold (i.e. relative growth rather than absolute growth), and the implications of this for growth paths over four growth periods. However, we consider independence of p to be an appropriate rough approximation for our present purposes.

we cannot detect a inverted-U-shape of exit rates for those firms with the smallest start-up size, presumably because they are so small that their initial resource stock is so small that it is quickly depleted in the event of adverse business events. In the light of this evidence, we consider that Hypothesis 2a is supported.

We begin our investigation of Hypothesis 2b by considering the evidence in Figure 5. In contrast to our previous analysis of growth paths, Figure 5 shows that exit rates vary considerably across growth paths. Firms that experience four consistent periods of decline have an exit rate of 26.39%, while firms with four consistent periods of growth have an exit rate of 5.63%. Firms experiencing a pattern of growth-growth-decline-decline have an exit rate of 26.56%, which is about three times higher than firms experiencing decline-decline-growth-growth (8.96%).

Interestingly enough, Figure 5 implies that early growth of new businesses influences survival prospects some years into the future. Take, for example, a new firm that declines in the second and third periods, but grows in the fourth period. Figure 5 shows that growth performance in the first growth period (that is, growth from year 1 to year 2) can make a considerable difference to likelihood of survival in Year 6. If the firm grew in the first period then its exit rate would be 8.62. whereas if it declined in the first period its exit rate was 15.06. We take this as evidence that growth paths have an effect on survival, and hence that Hypothesis 2b is supported.

5.1.3 Robustness

We now investigate the robustness of our analysis by changing the threshold according to which growth and decline are defined.[10] One might reasonably suppose

[10]In further robustness analysis, we attempted to 'pre-process' the growth rates to clean them of any possible influence of control variables (in particular, all the control variables listed in Table 6 except for lagged growth and start-up size), by running regressions on the growth rates, and sorting firm-year observations by the residuals rather than the raw growth rates. We wanted to do this, because the literature suggests that growth rates show some degree of dependence on certain variables such as lagged size (for a survey, see Coad, 2009). This analysis did not yield interesting results however – firms were very heavily over-represented in the "0000" and "1111" categories, suggesting that by controlling for other influences (which are essentially time-invariant) we introduced an artificial time-invariant characteristic to the (pre-processed) growth

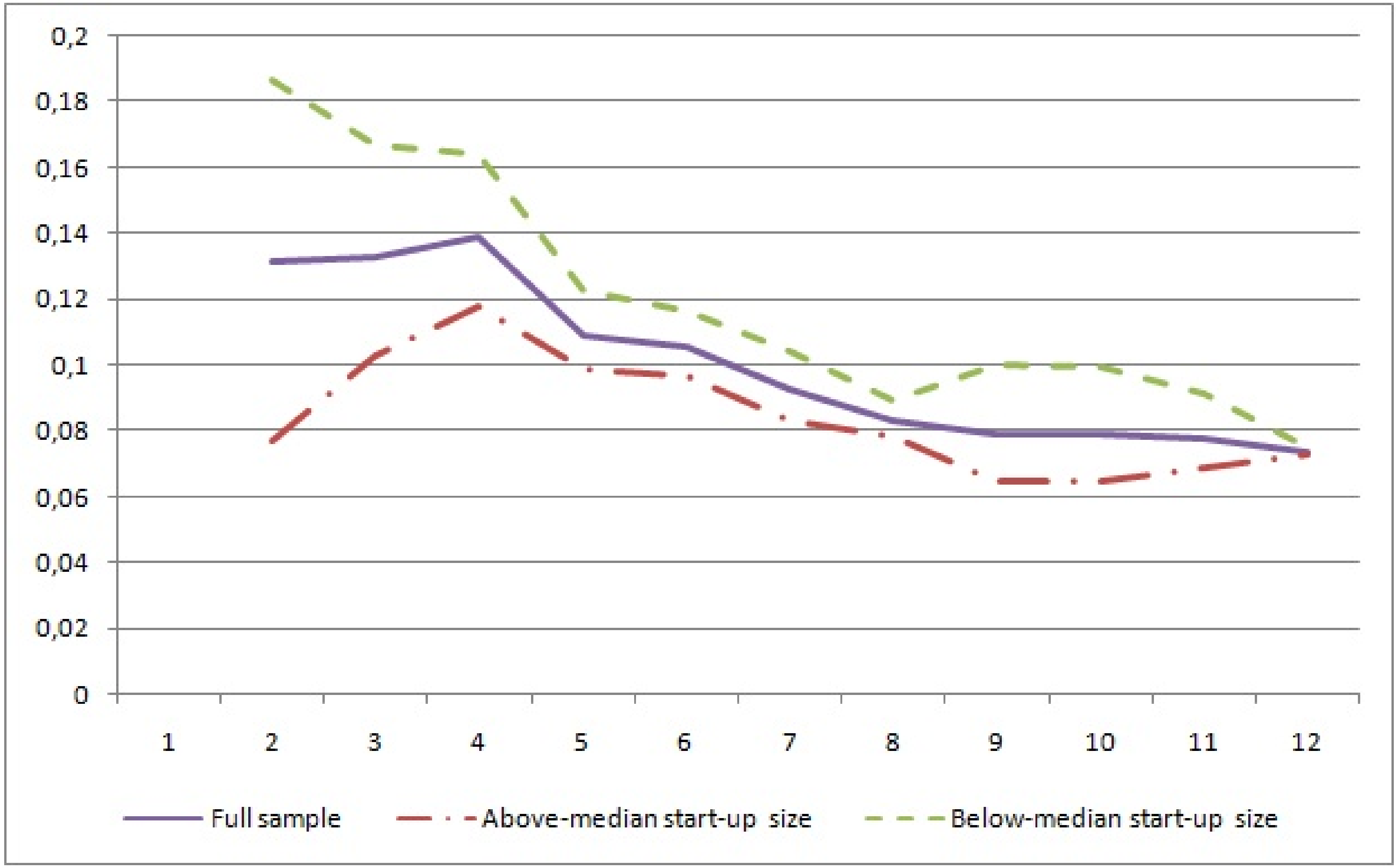

Figure 4: Exit rates in our sample, for firms with below-median start-up size, above-median start-up size, and also for the full sample.

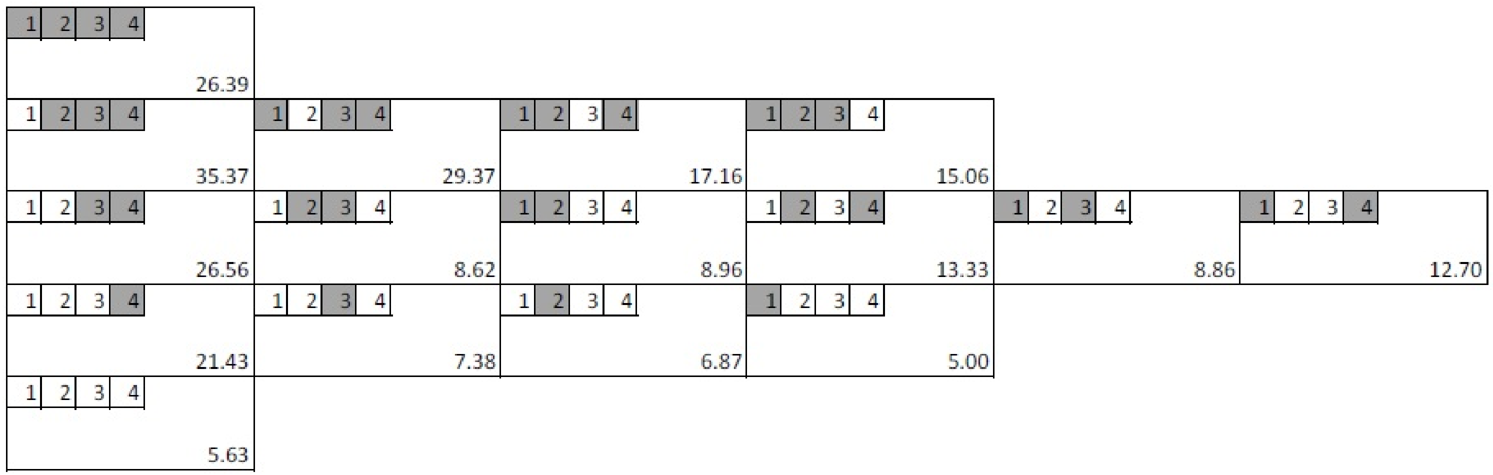

Figure 5: Exit rates rates (during the sixth year) associated with growth paths until the fifth year. N=2184. White squares correspond to growth years (above-median growth) while dark squares correspond to decline years (below-median growth), for the first four growth years of each business.

that the difference between growth (coded here as 1) and decline (0) might simply be due to slender fluctuations around our threshold. To investigate the sensitivity of our results to the median threshold, we now split the sample into the top 45% growth firms, versus the bottom 45% firms, effectively chopping out the central 10% of firms in each year. The proportion of businesses dropped at this stage is bounded from below at 10% (if the central 10% of businesses are the same in each year) and bounded from above at 40% (if in each of the four growth years a different 10% of firms are found in the 'grey zone'). In our data, we observe that removing the central 10% of observations in each of the four years leads us to drop 34.02% of our observations[11] which strongly suggests that it is not the same firms that occupy the central 10% of the growth rate distribution in each year. This can be taken as further evidence of the random nature of the firm growth process.

Table 5 shows the ranking of the different growth path categories, according to number of businesses in each growth path category, and exit rates associated with these firms in each category. We see that the ordering of the populations of growth categories are not robust (growth is random; growth paths are more or less equipopulated, and deviations from the equipopulated benchmark are not robust). On the other hand, the observed dependence of survival on growth path is more robust.

This robustness analysis leads us to be skeptical of Hypothesis 1, although it lends support to Hypothesis 2b. To investigate Hypothesis 2c, we require the inclusion of control variables, and so we now turn to multivariate regressions.

5.2 Multivariate regressions

In this section, we investigate Hypothesis 2c, which posits that growth paths have an effect on survival even after controlling for lagged size. We pursue our analysis of the effect of lagged growth on survival chances by applying a logit duration model (Jenkins, 1995).

Our main interest lies in the dependence of survival on lagged growth and

rate series. Therefore we did not pursue this line of investigation any further.

[11]That is, 1441 out of 2184 observations remain at this stage.

Table 5: Ranking of growth path categories, according to number of businesses in each growth path category, and exit rates associated with these firms in each category.

	Growth path				**Exit rate**			
	Full sample		Central 10% removed		Full sample		Central 10% removed	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Growth path	Freq.	Rank	Freq.	Rank	%	Rank	%	Rank
0000	6.36	6	5.76	1	23.02	13	30.12	13
0001	7.60	1	6.94	8	13.86	10	16.00	10
0010	6.00	9	5.55	16	14.50	11	20.00	11
0011	6.04	8	6.04	13	9.09	7	11.49	7
0100	5.68	12	5.90	12	27.42	15	34.12	15
0101	7.23	3	7.36	4	8.86	6	8.49	5
0110	5.68	13	6.66	7	12.10	8	12.50	8
0111	5.40	14	5.55	6	5.08	2	5.00	3
1000	6.55	5	6.66	15	32.87	16	38.54	16
1001	5.31	16	6.52	5	7.76	5	9.57	6
1010	6.32	7	6.73	2	12.32	9	14.43	9
1011	5.82	11	5.62	10	6.30	4	2.47	1
1100	5.86	10	5.62	9	25.78	14	30.86	14
1101	5.40	15	4.86	14	4.24	1	4.29	2
1110	7.55	2	6.52	3	20.61	12	21.28	12
1111	7.19	4	7.70	11	5.73	3	8.11	4
No. Obs.	2184		1441		2184		1441	
	Corr((1),(3)) = 0.6918				Corr((5),(7)) = 0.9846			
	Corr((2),(4)) = 0.1941				Corr((6),(8)) = 0.9794			

Table 6: Duration models: survival in the sixth year. Logit estimation with robust standard errors in parentheses. Survival is regressed on lagged (log) size, lagged growth, and a list of independent variables. For details on the independent variables, see Table 7. Key to significance levels: $***p < 0.01, **p < 0.05, *p < 0.1$

Dep. Var.: survival	(1)	(2)	(3)	(4)	(5)	(6)
L.log_sales	0.394***	0.270***	0.215***	0.135	0.0777	0.391***
	(0.0581)	(0.0724)	(0.0832)	(0.0842)	(0.0850)	(0.0820)
L.gr_sales			0.191	0.299***	0.365***	
			(0.131)	(0.111)	(0.112)	
L2.gr_sales				0.424***	0.496***	
				(0.147)	(0.142)	
L3.gr_sales					0.190	
					(0.137)	
L4.gr_sales					0.121	
					(0.142)	
start-up size						-0.302***
						(0.0877)
age	0.00813	0.00469	0.00509	0.00404	0.00581	0.00840
	(0.00978)	(0.0107)	(0.0107)	(0.0107)	(0.0107)	(0.0108)
age^2	-0.00106	-0.00108	-0.00111	-0.00113	-0.00119	-0.00125
	(0.000719)	(0.000769)	(0.000773)	(0.000773)	(0.000769)	(0.000766)
education	-0.0290	-0.00277	-0.0150	-0.00197	0.000706	-0.00265
	(0.0947)	(0.100)	(0.0998)	(0.101)	(0.102)	(0.101)
bexp_none	0.148	-0.0244	-0.0460	-0.0840	-0.105	-0.0991
	(0.266)	(0.272)	(0.273)	(0.282)	(0.285)	(0.278)
adv_entbl	0.108	0.0776	0.0393	-0.0209	-0.0500	-0.0321
	(0.342)	(0.349)	(0.354)	(0.355)	(0.358)	(0.357)
adv_acc	0.111	0.155	0.158	0.195	0.200	0.189
	(0.203)	(0.207)	(0.210)	(0.210)	(0.212)	(0.211)
adv_sol	-0.193	-0.214	-0.198	-0.231	-0.216	-0.186
	(0.371)	(0.396)	(0.400)	(0.408)	(0.412)	(0.406)
adv_coll	1.531*	1.400	1.439	1.404	1.334	1.308
	(0.883)	(0.859)	(0.883)	(0.898)	(0.868)	(0.846)
adv_srs	-1.894**	-1.785*	-1.776*	-1.871*	-1.865*	-1.797*
	(0.837)	(0.996)	(0.992)	(1.006)	(1.015)	(0.991)
adv_pybt	-1.131	-0.973	-0.901	-0.843	-0.929	-1.019
	(0.787)	(0.743)	(0.740)	(0.764)	(0.733)	(0.716)
adv_fam	-0.320	-0.315	-0.304	-0.317	-0.339	-0.358*
	(0.212)	(0.214)	(0.215)	(0.215)	(0.215)	(0.215)
adv_oth	-0.0983	0.0297	-0.0110	0.0320	0.0242	-0.0187
	(0.401)	(0.421)	(0.424)	(0.435)	(0.435)	(0.417)
own_xs	-0.245	-0.210	-0.187	-0.140	-0.120	-0.142
	(0.251)	(0.264)	(0.262)	(0.259)	(0.259)	(0.261)
own_male_inv	-0.0282	0.0601	0.0718	0.0794	0.0735	0.0478
	(0.273)	(0.278)	(0.278)	(0.282)	(0.286)	(0.285)
legform2	0.235	0.292	0.296	0.265	0.221	0.201
	(0.348)	(0.351)	(0.351)	(0.346)	(0.343)	(0.348)
legform3	0.550**	0.561**	0.501*	0.456*	0.423	0.432*
	(0.243)	(0.253)	(0.259)	(0.259)	(0.263)	(0.262)
l_vol		-0.437***	-0.415***	-0.383***	-0.365***	-0.382***
		(0.110)	(0.116)	(0.117)	(0.120)	(0.117)
l_odxs		-0.441**	-0.445**	-0.446**	-0.468**	-0.486**
		(0.222)	(0.223)	(0.226)	(0.229)	(0.225)
l_odxs_pc		-0.519*	-0.517*	-0.563**	-0.510*	-0.434
		(0.267)	(0.265)	(0.270)	(0.282)	(0.282)
l_odlim_use		-0.505*	-0.428	-0.384	-0.396	-0.462*
		(0.280)	(0.284)	(0.289)	(0.290)	(0.281)
l_odlim_pc		-0.202	-0.223	-0.185	-0.146	-0.136
		(0.277)	(0.277)	(0.278)	(0.279)	(0.276)
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Region	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,907	1,907	1,907	1,907	1,907	1,907
Pseudo-R^2	0.1008	0.1446	0.1483	0.1588	0.1609	0.1550

lagged size. However, to control for other influences, and to avoid omitted variable bias, we include a collection of control variables that have been shown in prior work to influence new and small firm survival (Storey and Greene, 2010), including some information on the founder observed at start-up (as in Cooper et al., 1994). In addition to the standard variables we also use information on bank account activity (such as bank account volatility and overdraft excess, which constitutes a novel feature of our dataset, and is shown to be a key factor influencing business survival (Frankish et al., 2010)). We do not cluster standard errors at the business level, because our focus on survival in year 6 implies that our final dataset is essentially a cross-section. We also investigate the possible presence of multicollinearity by examining the relevant Variance Inflation Factor (VIF) diagnostics, and we conclude that multicollinearity is not a pressing concern.[12] A correlation matrix is not presented here for reasons of space, but is available from the authors.

Our approach in Table 6 is to distinguish between explanations of duration, with a prime focus on prior size and growth. We begin in Column (1) by examining only sales, together with the 'standard' explanatory variables covering sector, region, the human capital of the founder, sources of advice and legal status of the firm. These show that lagged (log) sales enhances survival, and that of the 'standard' variables, only company status ('leg_form3') has a positive influence, with two of the advice providers (adv_srs and adv_coll) being associated with lower survival (at the 10% level of significance). Some regional and sectoral dummies (not shown) are also significant. Column (2) augments column (1) with trading variables relating to the performance of the business' bank account, that constitute an original set of explanatory variables in our dataset. While the con-

[12]This was done by re-estimating the regression models in Table 6 using an Ordinary Least Squares Linear Probability Model (OLS-LPM). In our original regressions, the only troublesome variables were age-squared and industry dummies. In our original regressions, VIF statistics for age and age^2 were not satisfactory, indicating a high degree of collinearity between these two variables. However, we applied the often-used mean-centering technique to derive a quadratic term for age as (age-mean(age))^2, and with this alternative independent variable we had no problems of multicollinearity for these variables. Some of our industry dummies displayed VIF values above the threshold value of 4.0 mentioned in Hair et al. (1998), so we repeated the analysis without including industry dummies in order to verify the robustness of our results, and observed that our main results did not change. We conclude that multicollinearity is not a problem in our present analysis.

ventional variables included in Column (1) were mainly insignificant, we see that the trading variables included at the bottom of Column (2) are almost all significant, and furthermore that they improve the model fit quite substantially (raising the pseudo-R^2 from 0.1008 to 0.1446). The trading variables in column (2) show that survival chances are significantly reduced by higher volatility, the use and amount of time spent in (unauthorized) overdraft excess, and even the use of an (authorized) overdraft facility. The inclusion of these trading variables is in itself a novel contribution of our paper, and we suggest that future work would benefit from including information on business bank accounts wherever possible.

Column (3) of Table 6 shows that lagged growth ($t-2 : t-1$) has a positive effect on survival, even controlling for lagged size ($t-1$). This is at odds with Gambler's ruin theory and is consistent with our Hypothesis 2c. Columns (4) and (5) includes longer lags of growth, and shows that growth has a long term impact on survival. Column (5) even shows that lagged size (at $t-1$) becomes insignificant once lagged growth is controlled for. Column (6) shows that start-up size has a significant impact on survival, even controlling for lagged size. This suggests that growth helps survival beyond the direct effect of increasing the resource stock available to a firm to endure adversity.

6 Discussion and Conclusion

This paper contributes to the emergent literature on growth path taxonomies by analysing the growth paths of nascent businesses through an examination of their banking records. Our empirical investigations were guided by developments of Gambler's Ruin theory. We find every possible growth path seems to occur with roughly equal probability, pointing to a strong role of chance which is at the heart of Gamblers Ruin. However, we also observed that the probability of survival is not purely a matter of chance since it is influenced by both the growth path and inital size of the enterprise. Even controlling for lagged size and lagged growth, longer lags of growth had a significant effect on survival chances. Start-up size had a significant influence on survival, even when controlling for lagged size.

Previous work on this issue has made attempts to categorize firms into qualitatively different 'species' (such as gazelles, gorillas, mice, elephants, etc.) based largely on their growth performance. These attempts at creating taxonomies of firms implicitly rely on a rather arbitrary choice of cut-off points, according to which firms are placed into different categories. Our data, however, suggest the differences between firms with different growth paths are largely differences of degree rather than kind; that is, different categories of firms arranged according to growth paths should be seen as largely quantitative rather than qualitative arrangements. A 'gazelle' will not always remain a 'gazelle' – in fact, after a period of high growth, 'gazelles' are often observed to revert to average performance, or perhaps experience below-average performance (Parker et al., 2010).

This paper acknowledges that the current models of firm growth have very limited explanatory power, but it does not view this as "a negative state of affairs" (McKelvie and Wiklund, 2010, p. 262) – but rather as the reverse. It argues that it necessary to build theories of new business performance – of which growth is one dimension – around recognising that this is primarily, but not exclusively, a game of chance. In this we are developing the notions of Gibrat (1931) and subsequently empirically developed by Davies and Geroski (1997, p. 385) that "firm growth rates are random and therefore firm size follows a random walk."

What we do not imply is that growth is a pure random walk, but rather that this is the dominant component. The task is then to build theories and conduct testing that acknowledge the dominant role of chance but combining this with some independent variables that do have explanatory power. From our models the temporal locus of the determinants of firm's performance may extend far back into a firm's history (Bamford et al., 2004). What is also clear is that many of the "usual suspects" such as the age, gender, prior experience and education of the founder and sources of advice have a limited role to play – certainly in comparison with observing performance via financial data.

With regard to future work, our ultimate aim is to improve our ability to predict the growth and survival of new enterprises. Given the strong chance elements present we never expect this to be achieved with a high degree of accuracy, but our work points to this task becoming more tractable as new firms increase in scale and size. We see our next task being to investigate whether there is an appropriate point in time, as a new firm evolves, when the power of prediction models increases markedly.

Appendices

A Variables used in regressions

Table 7 provides information on the variables used in our regressions.

Table 7: Variables used in the regressions

Dependent variable	
open	= 1 if start-up is still open at the end of the period
Main independent variables: Size and growth:	
log_sales	log of annual credit turnover of the current account
gr_sales	Growth of annual credit turnover
startup size	Sales in the first year
Structural variables observed at start-up:	
age	(mean) age of start-up owner-manager(s)
age_sq	quadratic function of age, calculated as $(\text{age-mean(age)})^2$ to avoid problems of multicollinearity
education	highest educational attainment of owner-manager(s): none, GCSE, A-level, Degree or higher, according to the UK National Vocational Qualification scale.
bexp_none	dummy variable equal to 1 if the owner-manager(s) has no previous business experience
adv_x	sources of advice and support sought prior to start up: enterprise agency/business link (entbl), accountant (acc), solicitor (sol), college (coll), (Barclays) start right seminar (srs), the princes trust (pybt), family (fam), other (oth) (recoded into dummy variables)
own_xs	= 1 if more than a minimum number of owner-managers of the start-up: company 2+, partnership/LLP 3+
own_male_inv	= 1 if there is at least one male owner-manager of the start-up, 0 otherwise
legform$x$	legal form of start-up, recoded into dummy variables. legform2: partnership; legform3: sole trader. Omitted category is legform1: company (including LLP)
Trading variables:	
vol	ratio of the standard deviation of monthly turnover to the mean monthly turnover, summed over two six-month periods to obtain an annual volatility indicator
odxs	= 1 if in excess of authorised overdraft limit at any time
odxs_pc	proportion of period in excess of authorised overdraft limit
odlim_use	= 1 if authorised overdraft used at any time
odlim_pc	mean proportion of authorised overdraft limit used
Industry and Region dummies:	
Industry	business sector of firm at start-up, recoded into dummy variables: Agriculture; Manufacturing; Construction; Motor trades; Wholesale; Retail; Hotels & catering; Transport; Property services; Business services; Health, education & social work (hesw); and Other services
Region	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

B Changes in data composition across years

Table 8 shows changes in the regional composition of our data, while Table 9 shows changes in the industry composition.

Table 8: Changes in the regional composition of our data over the six-year period.

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
East of England	0.160	0.167	0.170	0.174	0.176	0.181
East Midlands	0.072	0.072	0.067	0.065	0.059	0.056
London	0.220	0.218	0.220	0.215	0.214	0.208
North East	0.038	0.037	0.038	0.037	0.038	0.037
North West	0.065	0.068	0.065	0.067	0.066	0.067
South East	0.126	0.125	0.125	0.128	0.129	0.129
South West	0.101	0.099	0.102	0.102	0.102	0.105
West Midlands	0.093	0.093	0.093	0.093	0.093	0.089
Yorkshire	0.060	0.058	0.056	0.058	0.061	0.063
Scotland	0.000	0.000	0.000	0.000	0.000	0.000
Wales	0.065	0.064	0.064	0.062	0.061	0.064
Northern Ireland	0.000	0.000	0.000	0.000	0.000	0.000
OBS.	4872	3634	2905	2431	2070	1766

Table 9: Changes in the industry composition of our data over the six-year period.

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
Agriculture	0.009	0.010	0.010	0.010	0.011	0.010
Manufacturing	0.054	0.056	0.058	0.056	0.056	0.057
Construction	0.152	0.154	0.158	0.167	0.171	0.168
Motor trades	0.033	0.033	0.032	0.033	0.034	0.033
Wholesale	0.024	0.023	0.023	0.024	0.024	0.022
Retail	0.117	0.111	0.110	0.102	0.093	0.093
Hotels & Catering	0.089	0.083	0.078	0.066	0.060	0.062
Transport	0.037	0.036	0.032	0.031	0.031	0.029
Property services	0.037	0.041	0.042	0.043	0.046	0.050
Business services	0.262	0.266	0.273	0.274	0.277	0.275
Health, education & social work	0.026	0.027	0.029	0.031	0.033	0.034
Other services	0.159	0.158	0.156	0.163	0.164	0.168
OBS.	5192	3878	3092	2575	2184	1867

References

Acs, Z. J., Parsons, W., and Tracy, S. (2008). High impact firms: gazelles revisited. Small Business Administration: Office of Advocacy.

Bamford, C., Dean, T., and Douglas, T. (2004). The temporal nature of growth determinants in new bank foundings: implications for new venture research design. *Journal of Business Venturing*, 19(6):899–919.

Bates, T. (2005). Analysis of young, small firms that have closed: delineating successful from unsuccessful closures. *Journal of Business Venturing*, 20(3):343–358.

Baum, J., Locke, E., and Smith, K. (2001). A multidimensional model of venture growth. *Academy of Management Journal*, 44(2):292–303.

Coad, A. (2007). A closer look at serial growth rate correlation. *Review of Industrial Organization*, 31(1):69–82.

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

Coad, A. and Rao, R. (2008). Innovation and firm growth in high-tech sectors: A quantile regression approach. *Research Policy*, 37(4):633–648.

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

Cowling, M. and Liu, W. (2011). ASBS Barometer report on business growth, access to finance, and performance outcomes in the recession. Department for Business, Innovation and Skills (BIS), London, UK.

Davies, S. W. and Geroski, P. A. (1997). Changes in concentration, turbulence, and the dynamics of market shares. *Review of Economics and Statistics*, 79(3):383–391.

Davila, A., Foster, G., and Gupta, M. (2003). Venture capital financing and the growth of startup firms. *Journal of Business Venturing*, 18(6):689–708.

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

Denrell, J. (2004). Random walks and sustained competitive advantage. *Management Science*, 50(7):922–934.

DeTienne, D., Shepherd, D., and De Castro, J. (2008). The fallacy of "only the strong survive": The effects of extrinsic motivation on the persistence decisions for under-performing firms. *Journal of Business Venturing*, 23(5):528–546.

Frank, M. Z. (1988). An intertemporal model of industrial exit. *Quarterly Journal of Economics*, 103(2):333–344.

Frankish, J., Roberts, R. G., Storey, D. J., Spears, T. C., and Coad, A. C. (2010). Do entrepreneurs really learn? or do they just tell you that they do? Mimeo, University of Sussex.

Fraser, S. (2005). Finance for small and medium-sized enterprises. Bank of England, London.

Garnsey, E. and Heffernan, P. (2005). Growth setbacks in new firms. *Futures*, 37(7):675–697.

Garnsey, E., Stam, E., and Heffernan, P. (2006). New firm growth: Exploring processes and paths. *Industry and Innovation*, 13(1):1–20.

Geroski, P., Mata, J., and Portugal, P. (2010). Founding conditions and the survival of new firms. *Strategic Management Journal*, 31(5):510–529.

Geroski, P. A. (2000). The growth of firms in theory and practice. In Foss, N. and Mahnke, V., editors, *Competence, Governance, and Entrepreneurship*. Oxford University Press, Oxford, UK.

Gibrat, R. (1931). *Les inégalités économiques: applications, aux inégalités des richesses, a la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc.: d'une loi nouvelle la loi de l'effet proportionnel.* Recueil Sirey, Paris.

Gimeno, J., Folta, T., Cooper, A., and Woo, C. (1997). Survival of the fittest? entrepreneurial human capital and the persistence of underperforming firms. *Administrative science quarterly*, 42(4):750–783.

Hair, J., Anderson, R., Tatham, R., and Black, W. (1998). *Multivariate Data Analysis: Fifth Edition.* Prentice Hall: Upper Saddle River, NJ.

Hamilton, R. T. (2011). How firms grow and the influence of size and age. *International Small Business Journal*, forthcoming:DOI: 10.1177/0266242610383446.

Hansen, B. and Hamilton, R. (2011). Factors distinguishing small firm growers and non-growers. *International Small Business Journal*, forthcoming.

Harada, N. (2007). Which firms exit and why? an analysis of small firm exits in Japan. *Small Business Economics*, 29(4):401–414.

Headd, B. (2003). Redefining business success: Distinguishing between closure and failure. *Small Business Economics*, 21(1):51–61.

Henrekson, M. and Johansson, D. (2010). Gazelles as job creators: a survey and interpretation of the evidence. *Small Business Economics*, 35(2):227–244.

Hmieleski, K. and Baron, R. (2009). Entrepreneurs' optimism and new venture performance: A social cognitive perspective. *Academy of Management Journal*, 52(3):473–488.

Jenkins, S. (1995). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics*, 57(1):129–136.

Leitch, C., Hill, F., and Neergaard, H. (2010). Entrepreneurial and business growth and the quest for a "comprehensive theory": tilting at windmills? *Entrepreneurship Theory and Practice*, 34(2):249–260.

Levinthal, D. (1991). Random walks and organizational mortality. *Administrative Science Quarterly*, 36(3):397–420.

Madsen, T. and Servais, P. (1997). The internationalization of born globals: an evolutionary process? *International Business Review*, 6(6):561–583.

Majumdar, S. (2004). The hidden hand and the license raj to an evaluation of the relationship between age and the growth of firms in India. *Journal of Business Venturing*, 19(1):107–125.

McKelvie, A. and Wiklund, J. (2010). Advancing firm growth research: A focus on growth mode instead of growth rate. *Entrepreneurship theory and practice*, 34(2):261–288.

McMahon, R. (2001). Deriving an empirical development taxonomy for manufacturing smes using data from australia's business longitudinal survey. *Small Business Economics*, 17(3):197–212.

Mlodinow, L. (2008). *The drunkard's walk: How randomness rules our lives*. Pantheon.

Parker, S., Storey, D., and van Witteloostuijn, A. (2010). What happens to gazelles? The importance of dynamic management strategy. *Small Business Economics*, 35:203–226.

Phillips, B. and Kirchhoff, B. (1989). Formation, growth and survival: Small firm dynamics in the US economy. *Small Business Economics*, 1(1):65–74.

Politis, D. (2005). The process of entrepreneurial learning: A conceptual framework. *Entrepreneurship theory and practice*, 29(4):399–424.

Pompe, P. and Bilderbeek, J. (2005). The prediction of bankruptcy of small-and medium-sized industrial firms. *Journal of Business Venturing*, 20(6):847–868.

Storey, D. and Greene, F. (2010). *Small business and entrepreneurship*. Financial Times/Prentice Hall: UK.

Storey, D. J. (2011). Optimism and chance: the elephants in the entrepreneurship room. *International Small Business Journal*, forthcoming.:DOI: 10.1177/0266242611403871.

Taylor, M. (1999). Survival of the Fittest? An Analysis of Self-Employment Duration in Britain. *Economic Journal*, 109(454):140–155.

Wennberg, K., Wiklund, J., DeTienne, D., and Cardon, M. (2010). Reconceptualizing entrepreneurial exit: Divergent exit routes and their drivers. *Journal of Business Venturing*, 25(4):361–375.

Wiklund, J., Patzelt, H., and Shepherd, D. (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.