



HAL
open science

Modeling new-firm growth and survival with panel data using event magnitude regression

Frédéric Delmar, Jonas Wallin, Ahmed Maged Nofal

► **To cite this version:**

Frédéric Delmar, Jonas Wallin, Ahmed Maged Nofal. Modeling new-firm growth and survival with panel data using event magnitude regression. *Journal of Business Venturing*, 2022, 37 (5), pp.106245. 10.1016/j.jbusvent.2022.106245 . hal-03910411

HAL Id: hal-03910411

<https://hal.science/hal-03910411>

Submitted on 23 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modeling New-Firm Growth and Survival Using Event Magnitude Regression

Frédéric Delmar
emlyon business school
23, avenue Guy de Collongue
CS 40203 69134 Ecully Cedex, FRANCE

&

Jonas Wallin
School of Economics and Management, Lund University
P.O. Box 7080,
S-220 07 Lund, SWEDEN

&

Ahmed Maged Nofal
emlyon business school
23, avenue Guy de Collongue
CS 40203 69134 Ecully Cedex, FRANCE

Author note

A later version of this paper is published as:

Delmar, F., Wallin, J., & Nofal, A. M. (2022). Modeling new-firm growth and survival with panel data using event magnitude regression. *Journal of Business Venturing*, 37(5), 106245. <https://doi.org/10.1016/j.jbusvent.2022.106245>

Frédéric Delmar <https://orcid.org/0000-0003-4974-9031> (Corresponding Author)
Jonas Wallin <https://orcid.org/0000-0003-0381-6593>
Ahmed Maged Nofal <https://orcid.org/0000-0002-2532-1981>

We have no conflict of interest to disclose.

A previous version of this paper was presented as: “Modelling new firm growth and survival: Some practical solutions” F Delmar, J Wallin - *Academy of Management Proceedings*, 2018

This research is supported by a grant from the Swedish Research Council (Registration Number 2018-01726).

Correspondence according this article should be addressed to Frédéric Delmar, emlyon business school, 23, avenue Guy de Collongue, CS 40203 69134 Ecully Cedex, FRANCE. E-mail: delmar@em-lyon.com, phone: +33 630 83 37 22.

Modeling New-Firm Growth and Survival with Panel Data Using Event Magnitude**Regression****Abstract**

We introduce a new model to address three methodological biases in research on new venture growth and survival. The model offers entrepreneurship scholars numerous benefits. The biases are identified using a systematic review of 96 papers using longitudinal data published over a period of 20 years. They are: (1) distributional properties of new ventures; (2) selection bias; and (3) causal asymmetry. The biases make the popular use of normal distribution models problematic. As a potential solution, we introduce and test an event magnitude regression model approach (EMM). In this two-stage model, the first model explores the probability of four events: a firm staying the same size, expanding, contracting, or exiting. In the second stage, if the firm contracts or expands, we estimate the magnitude of the change. A suggested benefit is that researchers can better separate the likelihood of an event from its magnitude, thereby opening new avenues for research. We provide an overview of our model analyzing an example data set involving longitudinal venture level data. We provide a new package for the statistical software R. Our findings show that EMM outperforms the widely adopted normal distribution model. We discuss the benefits and consequences of our model, identify areas for future research, and offer recommendations for research practice.

Keywords: New firm growth and survival; longitudinal, methods; quantitative.

Modeling New-Firm Growth and Survival with Panel Data Using Event Magnitude

Regression

New ventures differ enormously. Most ventures fail within their first years, and most do not grow. Only a small share of growing firms or high-growth firms have long lasting effect on job creation and productivity growth (Davidsson et al., 2006; Haltiwanger, 2015; Sterk et al., 2021). To explain these differences in new venture growth (expansion and contraction) and survival (termination, and exit), entrepreneurship and organizational scholars have, over the last decades, increasingly relied on longitudinal studies with repeated measures over time for the same subject (e.g., Anglin et al., 2018; Hmieleski and Baron, 2009; Shepherd et al., 2021; Wennberg et al., 2016). These studies employ various perspectives such as strategy (Bennett and Levinthal, 2017; Eberhart et al., 2017; Wennberg et al., 2016), sociology (Barron et al., 1994; Carroll and Hannan, 2000; Mens et al., 2015), and economics (Acemoglu et al., 2018; Sterk et al., 2021), which call for new statistical methods; otherwise, theories may not be adequately tested and data opportunities can be missed out (Zhou et al., 2021).

In fact, systematically probing into extant research, we find that modelling new venture growth and survival effectively poses three significant problems that researchers need to address. First, the specific distributional properties of new ventures make models assuming the normal distribution inadequate. Second, the presence of strong selection, such as self-selection and survival bias, increases the risk of biased conclusions. Particularly, extant work builds on ventures, most of which have a low probability of growing (Hurst and Pugsley, 2011), tend to remain static in size, and go out of business, leading not only to modeling bias (Coad, 2007; Crawford et al., 2015; Erlingsson et al., 2013) but also survival and selection biases (Certo et al., 2016; Holcomb et al., 2010; Yang and Aldrich, 2012). Third, the asymmetry between contraction and growth confounds causal mechanisms. Studies assume that the same causal mechanisms symmetrically explain the different dynamics, such as growth and contraction

(Whetten, 1987), which in turn misguide the independence assumption (i.e. that new venture later dynamics depend on new venture early dynamics) (Holcomb et al., 2010; Keil et al., 2009). The three problems lead to inaccurate variance estimates, deflating/inflating errors, and obscuring of the causal mechanisms that explain new venture growth and survival (Antonakis et al., 2019a; Gielnik et al., 2017; Holcomb et al., 2010; Keil et al., 2009; Williams and Shepherd, 2017).

To mitigate these problems, we introduce a “two-part model” that minimizes the potential for modeling, selection, and survival biases when analyzing new venture growth and survival. We propose a specific approach to separate the likelihood of an event from its magnitude (Farewell et al., 2017). Specifically, we develop an event magnitude model (EMM) in the context of entrepreneurship and compare it to models assuming a normal distribution. We start with estimating the probability of change (i.e., probabilities of exit and growth) and then, conditional on the probability of the change occurring, we estimate its magnitude. We estimate the magnitude in terms of four outcomes: (a) growing, (b) contracting, (c) remaining static, or (d) exiting.

Our paper makes five contributions. First, we systematically review, critique, and synthesize prior literature on new venture growth and survival. Second, we show how to better model new venture growth and survival relative to the standard normal distribution. In this regard, we show that the normal distribution assumption overstates the probability of growth. Third, we relax the assumption of symmetry in growth; that is, we do not assume that growth and contraction are mirror images of each other. By doing so, we can better estimate the effect for firms that expand, compared to the estimate for firms that contract. Moreover, we disentangle the probabilities of growth, and contraction, from their respective magnitudes. Probability and magnitude are confounded in models assuming a normal distribution. We can thereby better understand what is happening when firms change size and allow for rare but

influential outcomes in our data. Fourth, we provide a flexible and comprehensive technique that captures different new venture dynamics including growth, survival, contraction, and exit, and therefore reduces selection and survival biases. Further, EMM lessens the likelihood of inflating/deflating error terms, yielding less-biased estimates. It also minimizes the potential of omitted variable bias. Finally, we have developed an open-source package that scholars can download and use to analyze new venture growth and survival data. The code package is found in the supplementary files as `cams_0.1.0.tar.gz`. To run it, `r-tools` need to be installed and then `“install.packages(path_to_file, repos = NULL, type="source")”`. Details on the coding and simulated data are provided in Appendix IV.

Theoretical Background

New firm growth and survival

Understanding new venture growth and survival is an important topic for multiple strands of entrepreneurship and organizational research. By new venture growth and survival, we refer to the changes in size (growth and contraction) and survival (exit and termination) of new entrants—predominantly *de novo* firms, but *de alio* ones, too. The topic has attracted interest from entrepreneurship scholars using different perspectives as strategy (Bennett and Levinthal, 2017; Eberhart et al., 2017; Wennberg et al., 2016), sociology (Barron et al., 1994; Carroll and Hannan, 2000; Mens et al., 2015), and economics (Coad et al., 2013; Coad et al., 2017; Haltiwanger, 2015; Luttmer, 2010; Sterk et al., 2021). However, this literature is mixed and divergent both in terms of theory and empirics. For example, there have been mixed findings related to the influence of new venture age on growth where some scholars report positive effects, others find that age has no effect or negative effects on growth. For example, Yan and Williams (2021) discusses how venture age at internalization across different studies show contradictory empirical results in its role of predicting growth and survival. Hence, despite laudable efforts, this research has not substantially progressed (Demir et al., 2016;

Garnsey et al., 2006; Wright and Stigliani, 2012); and has raised concerns about whether new venture growth and survival might be random (Coad et al., 2015; Derbyshire and Garnsey, 2014; van Witteloostuijn and Kolkman, 2019). In such circumstances, and to explore whether there are cumulative findings and systematic biases, we need to systematically review this literature (Nofal et al., 2018; Shane, 2009). We do so to identify and highlight some potential reasons and biases for existing contradicting findings, and propose alternatives to mitigate some of these biases.

New Venture Growth and Survival –Systematic Review Findings

We review the literature with an explicit and systematic method to capture and review work on the strands of new venture growth, survival, contraction, and exit. We limit our list to articles published from 2000 onwards, which removed 2 articles published in 1994 (Cooper et al., 1994) and 1997 (Barkema and Vermeulen, 1997). Accordingly, our systematic review retrieved a total of 96 empirical papers. The stop date of the search was May 15, 2021.

The 96 retrieved articles focus on the antecedents of venture growth and survival. In the Appendix I, we describe the method and the papers extracted. Table 1A groups these articles as follows: Table 1Aa contains 57 articles examining venture growth, Table 1Ab contains 25 articles examining survival, and Table 1Ac contains 14 articles examining both survival and growth. There are 57 articles using models assuming a normal distribution: Ordinary Least Square (OLS) (n=21); a fixed effects approach (n=15), a random effects approach (n=9); Heckman/Two-Stage Least Squares (2SLS) estimation (n=7); Generalized Least Square (GLS) (n=5), and Generalized Method of Moments (GMM) (n=10). To explain survival or exit, logit/probit models (n=21) and Cox regressions (n=15) dominate. Of the retrieved articles 59% adopt models assuming a normal distribution to explain venture growth and survival. The share of such models is even higher if we exclude articles only investigating survival and exit (76%).

Most of the retrieved papers (almost 50%) draw on samples from the U.S. and Sweden. There is a significant lack of data covering new venture growth and survival during the last decade (80% of the articles have data with the latest observation in 2011), and none of the articles use data that capture dynamics during the last 5 years. The three most common measures used to assess venture growth and survival are sales (41%), number of employees (27%), and number of years of survival (8%). The rest of the articles (24%) use different productivity and performance measures, such as profits, product development, domestic and foreign investments, and others. This literature shows that social resources (Eberhart et al., 2017; Khaire, 2010; Korunka et al., 2010; Prashantham and Dhanaraj, 2010), on-the-job-embeddedness (Mai and Zheng, 2013a), human capital (Korunka et al., 2010; Shrader and Siegel, 2007; Symeonidou and Nicolaou, 2018; Tzabbar and Margolis, 2017), financial capital (Cooper et al., 1994; Eberhart et al., 2017), tax credits and exemptions (Friske and Zachary, 2019), involvement in R&D (Haeussler et al., 2019), founders participation (Kor, 2003), and self-efficacy (Baum and Locke, 2004) influence new venture growth and survival.

Mixed Findings. Our systematic effort reveals some informative patterns. First, we find mixed evidence. For example, some researchers found that the effect of venture age and size have differential effects on growth and survival, depending on data structure and theoretical perspectives. On the one hand, age and size tend to be connected as younger ventures are smaller, and older ventures tend to be bigger (Davidsson et al., 2006). Statistically, there is also a compositional shift towards larger ventures that are less likely to exit as cohorts of ventures age. On the other hand, theory suggest that venture age and size are double-edged sword, often interacting with each other. For example, liability of newness suggests that age reduces both the probability of exit and growth (Yang and Aldrich, 2017). Here, size functions as a buffer ,where large initial financial resources (Wiklund et al., 2010) can diminish the age effect or even reverse the effect on new venture growth (Yan and Williams, 2021). Size can also decrease

the probability of growth by introducing more rigidity, coordination costs, resource allocation and complexity, thereby making the venture less able to adapt to new growth opportunities (Sine et al., 2006). Age might decrease the probability of growth as ventures become increasingly complacent or as increasing competition erodes profits from initial rents (Aksaray and Thompson, 2018). With age, the ventures learn to become more efficient, and in turn increase its survival chances and better exploit growth opportunities (Soto-Simeone et al., 2020; Wennberg et al., 2016). Not only are the directions and patterns observed mixed across studies, but so are the effect sizes imposed on new venture growth and survival. For example, venture age coefficients ranged from -0.23 to 0.93 in our retrieved articles (e.g., Bertoni et al., 2011; Jennings et al., 2009; Kor, 2003; Wennberg et al., 2016). The venture size coefficients ranged from -0.07 to 0.94 (e.g., Baum and Locke, 2004; Bertoni et al., 2011).

Another central venture-level variable is venture profitability for survival and growth, where early profitability is important for survival and growth (Ben-Hafaïedh and Hamelin; Davidsson et al., 2009; Hvide and Møen, 2010) and related financial performance aspirations (Gimeno et al., 1997; Wennberg et al., 2016). Still though, extant papers show some differences in the magnitude of the effect of profitability (e.g., ROA) on new venture performance indicators (e.g., Fattoum-Guedri et al., 2018; Robinson and Phillips McDougall, 2001; Wennberg et al., 2016). It is worth noting that while these discrepancies could be due to different measures (Davidsson et al., 2006), many of them are due to empirical biases, as we will show below.

Modeling and Selection Biases. Second, we find that existing research shows patterns of modeling, self-selection, and survival biases, which could yield inaccurate estimates, deflate/inflate errors, and obscure the causal mechanisms that explain new venture dynamics (Antonakis et al., 2019a; Gielnik et al., 2017; Holcomb et al., 2010; Keil et al., 2009; Williams and Shepherd, 2017). Table 1A shows that a lion share of the studies use models requiring new

venture data with a normal distribution—a requirement that longitudinal data do not usually satisfy (Douglas et al., 2020)¹. Most new ventures have a low probability of growth (Hurst and Pugsley, 2011), tend to remain static at the same size, and in most cases go out of business (Coad, 2007; Crawford et al., 2015; Erlingsson et al., 2013), resulting in survival and self-selection biases (Certo et al., 2016; Holcomb et al., 2010; Yang and Aldrich, 2012).

Causal Mechanisms. Third, the review shows that prior research treats the different dynamics of new ventures similarly in terms of their causal mechanisms. For instance, prior work assumes that the same causal mechanisms affect both growth and contraction symmetrically (Whetten, 1987). In many instances, exit has been assumed to be the same as failure—and both failure and exit have been assessed as the opposite of success, with an implicit assumption that both failure and success pose the same causal mechanism (Wennberg and DeTienne, 2014). Moreover, both survival and growth—despite their differences (Haeussler et al., 2019; Wennberg et al., 2016)—have commonly been used as a proxy to measure success (Shane, 2003).

Based on our findings, and as Yang and Aldrich (2012: 480) argue, we find that “divergent results partially stem from methodological problems that have systematically biased previous findings”. To progress, we need to mitigate three main issues found in panel data analysis. First, the distribution of new venture data makes normal models a debatable choice. Yet, such models dominate extant research. Second, new venture data are subject to substantial selection biases that are still too rarely corrected for in extant research. Third, empirical research largely assumes causal asymmetry between contraction and growth in their models. It is theoretically more likely that there is asymmetry in causality, both in terms of the likelihood of

¹ Though studies employing logit/probit regressions do not need to satisfy log-normal distribution assumptions, some scholars argue that they fall short of interactive effects that characterize various findings on new venture dynamics.

the contraction or growth happening and in the magnitude of these events. In the next section, we discuss each of these problems in the context of longitudinal data.

The specific distributional properties of new ventures

The dominating models in our review assume normally distributed data. This is seldom the case in longitudinal data on new ventures. For instance, most new ventures never grow because they are founded only to provide employment for the founders themselves (Parker, 2009). Therefore, an important challenge in organizational and entrepreneurship research is to separate “mere” stagnant firms from growth-oriented firms (Henrekson and Sanandaji, 2014; Koudstaal et al., 2016; Levine and Rubinstein, 2017). This distinction is even harder to make when we acknowledge that entrepreneurs’ growth aspirations might form or dissipate as they learn about the value of the business opportunity (Jovanovic, 1982; Jovanovic and Lach, 1989; March and Shapira, 1992). Trying to separate stagnant firms from growth firms through different sampling procedures is rarely efficient, beyond the important distinction between unincorporated and incorporated firms (e.g., Levine and Rubinstein, 2017). Even after we exclude unincorporated firms that have no legal avenue for growth, understanding new venture growth, is challenging when static, unchanging firms dominate samples.

In fact, as most incorporated ventures do not grow, and only a few grow substantially (Haltiwanger, 2015), existing data, including register data, tend to have distributions with low average growth and substantial outliers. In other words, even though research is increasingly turning to longitudinal register data linking employer-employee data to study new-firm dynamics (Coad et al., 2014; Delmar et al., 2013; Haltiwanger, 2015), in such register data, growing ventures tend to drown in the oceans of stagnant micro-sized ventures. When analyzing such data, a critical question is whether there is a subpopulation of ventures that will never grow. The existence of such population leads to an excess of zero observations, because between any two longitudinal observation times, the ventures in this subpopulation will always

be observed to have no dynamics (i.e. an increase, or a decrease, in sales or employees). Statistically speaking, modeling only the population average (of no growth) limits our ability to identify sources of variances, i.e., potential causal mechanisms as suggested by our theories to explain growth and high growth. Similarly, the existence of influential outliers might have disproportionate influence on substantive conclusions regarding relationships among variables for the main population of firms (Aguinis et al., 2013), leading to mixed evidence, and survival and selection biases (Holcomb et al., 2010; Yang and Aldrich, 2012).

From the findings of the systematic review, we observe that organizational researchers often carry out several transformations of the data to manage these distributional properties and increase model fit (e.g., achieving a higher R^2). One example is transforming growth and survival variables using log and ln functions (e.g., Coad, 2010; Singh and Delios, 2017). Meanwhile, as most firms tend to remain stagnant the value of their growth/contraction tend to be zero, and hence using such functions becomes undefined. Another example is winsorizing (e.g., Delmar and Wiklund, 2008), which is a procedure where data at the tails are truncated to limit the effects of outliers. The winsorized firms or observations are only outliers relative to a population mean. However, transforming data to capture mainly a large population of firms that will never grow limits our ability to explain growth, because our distribution of firms becomes so heavily dominated by non-growing firms. Potential causal relationships explaining firm growth become, thus, at risk of being drowned out, and consequently of not being detected (Aguinis et al., 2013). This is a problem that entrepreneurship studies share with other research fields such as strategy (e.g., Meyer-Doyle et al., 2019) and organizational behavior (Aguinis et al., 2013), leading to the likelihood of Type I error (i.e. finding false significance). Hence, existing modeling of firm dynamics fail to explain growing firms because large pools of stagnant firms dominate the distribution. Hence, we ask:

Q1: To what extent do the distributional properties of new venture data with many stagnant firms and few growing firms affect our findings on new venture growth and survival?

Selection bias, exit, and growth

In new venture longitudinal data, many firms are subject to termination and exit in their first years. Over time, only surviving firms remain. This feature of data could lead to even more biases (Xu et al., 2019). Selection bias arises when the units of analysis have different likelihoods of being included in the study compared to the target population, according to important study characteristics—specifically, the exposure and the outcome of interest. Hence, selection bias is caused by some kind of problem in the process of selecting units of analysis, either (i) at the outset, or (ii) in a longitudinal design, when some units of analysis are dropped from the study (Mazumdar et al., 2007; Schmidt and Woll, 2017). The risk is that we might have biased inference and incorrect conclusions because we do not sample a relevant population of units (i.e., firms with growth potential), or that our process, exposure, and outcome are linked, but the link is not fully captured by the model or data. Selection bias is present in two forms in studies of new firm growth and survival: self-selection bias and survival bias.

Self-selection bias. Self-selection bias is defined as the non-random assignment to a treatment or a group (Clougherty et al., 2015). If it occurs, the effect of choosing to participate (self-selection) is confounded with the treatment (the causal mechanism). Self-selection bias may or may not be a problem for research on growth (and exit), depending on the assumptions made by theory. For example, many of the retrieved papers assume that growth and exit are exogenously determined, building upon the assumption that mechanisms outside the firm and the entrepreneur decide the size and the survival of the firm (e.g., Singh and Delios, 2017; Tian et al., 2019), and hence endogeneity may be less problematic. Examples of such theories are those based on rationality, minimum efficient scale or size (MES) (Aksaray and Thompson,

2018; Reichstein et al., 2010), evolutionary theories (Aldrich, 1999; Carroll and Hannan, 2000), and resource-dependence theory (Pfeffer and Salancik, 2003).

Nevertheless, self-selection bias becomes a problem when growth and exit are understood as a strategic decision, and therefore endogenous (Rocha et al., 2018). Endogenous theories of growth and exit promote strategic choice as the primary mechanism in their model (Bradley et al., 2011; Davidsson et al., 2006; Gimeno et al., 1997; Zahra et al., 2006). Here, self-selection bias is an important issue to consider: those firms that expand, contract, or exit *choose* to do so, so we can no longer make the assumption that all firms need or want to expand, contract, or exit. We are dealing with a set of theories where firms choose to undergo a form of treatment.

Survival bias. Another form of selection bias, and perhaps the most well-known, is survival bias. Survival bias can be understood as the problem of the units of analysis not surviving until they reach treatment, independent of their willingness to do so. Hence, we cannot know the causal effect of a mechanism because units have an unequal chance of surviving until the treatment is carried out. Hence, we confound the effect of surviving with the effect of the treatment. New venture entry and growth are also closely connected to firm exit. For example, growing firms take more risks, and hence are more likely to fail and exit until they reach a certain size (Delmar et al., 2013). Many new ventures have a low probability of survival, with only about 40–55% of ventures still active after five years, depending on country and industry (Cefis and Marsili, 2011; Coad et al., 2013). Given that almost all retrieved articles use data that have time spans based on years, rather than months or quarters, existing research is likely to suffer from this survival bias.

Both self-selection and survival biases lead to biased selection and are of potential threat to causality. This is not always a showstopper per se, although there is no simple remedy (Mazumdar et al., 2007; Schmidt and Woll, 2017) if the problem is not tackled at the design

stage. Sampling data, especially register data, can lead to specific distributional properties that generate different forms of bias, such as survival bias (Certo et al., 2015; Elwert and Winship, 2014; Heckman, 1979; Palestini, 2015). In this respect, an important aspect of modeling for register data is understanding the sampling scheme of our observations, and hence potential sources of selection bias. By “sampling scheme,” we mean the decisions about what to sample from the population, which will define the weight of each observation in our analysis. An observation corresponds to a unit at a given point in time—e.g., the firm i in a specific year t . The estimated coefficients (parameters) will depend on the weighting of the observations in the sample.

In sum, existing work may be suffering from different forms of biased sampling (self-selection and survival). Disregarding such biases leads to poor and biased estimates. The importance of addressing such biases is dependent upon the theory used and the assumptions made. Consequently, when analyzing new-firm growth and survival we need to consider growth and survival together, not separately. The assumptions we make about growth and firm survival, and the research question we seek to answer, determine whether or not we have a problem with selection and survival bias. The assumptions and the research question also affect how we design and sample our units of observation (i.e., new firms). Hence, we ask:

Q2: To what extent does the presence of selection such as self-selection and survival bias affect our findings on new venture growth and survival?

Contraction and expansion

The third problem is related to the causal mechanisms of new venture growth and survival. Specifically, many studies assume that expansion and contraction are two sides of the same mechanism—i.e., that contraction is simply a negative value of expansion (Headd, 2003; Whetten, 1987). Yet, many dominating growth theories discuss only the expansion side of new venture dynamics. They seldom discuss contraction (i.e., almost all retrieved articles focus on growth and survival), and whether the same causal mechanisms are at play, or not. However,

our empirical models assume such symmetry when we model both contraction and expansion, and our dependent variables tend to be normally distributed around a mean. Here, we advocate asymmetric causality (Hatemi-j, 2012), where causal mechanisms might not only have differential effects, but might also entail different pathways (Ozcan, 2018).

For example, lack of profits might lead to contraction, but profits might not necessarily lead to expansion. Even if the same causal mechanisms are at play, we can easily assume that the duration of the effect may be different. Theoretically, we can argue that entrepreneurs will react more quickly to profit, gains and invest in expansion to take advantage of a window of opportunity, whereas their reaction to declining profits might not lead to contraction because they might hope for a quick turnaround. Moreover, legislation may affect the speed of expansion and contraction. New ventures can be slower in expanding, i.e., recruiting new employees, when profits or demands are increasing because of strong labor-market laws that make firing difficult (Card et al., 2018; Gideon et al., 2002; Henrekson and Johansson, 2010). Here, the entrepreneur tries to find other solutions to meet an increase in demand to mitigate potential volatility in demand. Similarly, an entrepreneur does not necessarily rush to fire people even if demand and profits diminish, as the setback might be temporary and the loss of dedicated and trustworthy employees might impose a long-term challenge, both economically and emotionally (Wiklund et al., 2003).

Empirically, the distribution of employment growth data can be asymmetric. There is no reason to assume that the distribution of number of employees fired mirrors the distribution of number of employees hired, especially when the potential magnitude of contraction is also different from that of expansion. In terms of expansion, a new venture can expand infinitely from its current size; in terms of contraction, it can only shrink from its current size down to zero. As a result, assuming symmetry could result in various biased results. Thus, we cannot assume symmetry between expansion and contraction. Finally, we ask:

Q3: To what extent does the asymmetry between contraction and growth, reflected in the observed magnitude of contraction relative to expansion, affect our findings on new venture growth and survival?

The EMM

Building on our systematic review findings, we observe a pressing need to investigate alternative user-friendly models. Models addressing the problems raised above and that allow researchers to better analyze venture growth and survival using longitudinal data. We suggest EMM. EMM is a two-part model, similar to those previously developed to analyze machine failures, sexual behavior, nutrition, fertility, ecology, agriculture, and healthcare costs (Farewell et al., 2017). In fact, there are several contexts where such two-part models could be deployed: in strategy, the creation of an alliance and its subsequent performance (Hitt et al., 1998; Meyer-Doyle et al., 2019); in organizational behavior, the probability of mental illness among employees and its consequences (Follmer and Jones, 2018); or in entrepreneurship, the modeling of new-firm growth and survival. What all these areas of application have in common is that they feature two potential subpopulations: one comprising units of analysis that will never change, but rather remain static, and another comprising units that may exhibit substantial changes. In this respect, it is worth re-noting that a large static subpopulation can lead to an excess of zero observations, as no change occurs between observation points. Moreover, two-part models are of further interest for the analysis of repeated measures of an outcome variable over time (Farewell et al., 2017), which are becoming increasingly available and important to entrepreneurship and organizational scholars (Card et al., 2018; Certo et al., 2017; Ployhart and Vandenberg, 2010; Xu et al., 2019).

Our model builds on two steps. In the first step, we estimate logistic regression models for four different outcomes (expansion, contraction, stasis, and termination). In doing so, we model multiple outcomes, not just a binary one. In the second step, we estimate the magnitude of expansion or growth, given one of those outcomes. The model is developed for both changes

in sales and changes in employment. To be consistent within the paper, we present the model with employee changes, but the model specification and the results for sales are in Appendix III. For employee changes—hiring and firing—we use a geometric (hiring) or constrained geometric distribution (firing). The geometric distribution is the discrete counterpart to the exponential distribution. Thus, EMM can be seen as a discrete counterpart to a (asymmetric) Laplace distribution, which has been shown to fit firm growth data well (Coad, 2007). For changes in sales or profit—i.e., continuous growth instead of discrete growth—the researcher can replace the geometric distribution with some non-negative continuous distribution, such as exponential, log-normal, or gamma distribution. EMM differs from other related models in that we allow causal asymmetry in expansion and contraction. It is worth noting that EMM is also adaptable to cross-sectional data where growth is measured retrospectively on active and surviving firms. However, it will not estimate firm termination and subsequently one would need to address the issue of survival bias.

Time to expansion or contraction. The time between each observation is important for the relevance of EMM. If the interval is too short—for instance, one week—then the statement that “most small firms do not contract or expand” is likely false, because a given firm remaining the same size over a week is not a meaningful indication of stasis. On the other hand, if the measures are repeated too seldom—say, every five years—it is possible for several different changes to have occurred in the meantime. The observed difference in size is likely due to many expansions and contractions. Because the sum of many random variables converges to a normal random variable (Central Limit Theorem) (Billingsley, 1995), this will make the normal distribution more suitable. However, size differences might also have accumulated among firms, meaning that previously growing firms are more likely to grow in the future than those that have not grown, so we still observe a distribution that is very different from the normal

(Delmar et al., 2003; Weinzimmer et al., 1998). Thus, one needs to verify the fit of the chosen distribution to the actual data.

Probability of expansion or contraction. EMM assumes venture size (i.e., the number of employees, but any other size measure is possible) at year t , E_t , is a Markov process—one where the future probabilities are determined by the variable's most recent values. That is, the most recent value (i.e., the previous year, in our case) fully captures history. The advantage of making such an assumption is that our model is frugal, and needs only the previous year's data to predict the outcome of the following year. Frugality is an important advantage here, as many new ventures have only been in existence for a limited time, and are likely to be terminated soon afterwards. Hence, having to rely on multiple years to estimate a model becomes a problem when dealing with new-firm growth and survival. The assumption that the previous year's data is sufficient diminishes the demands on data in general—and, in the case of a cohort of new firms, also mitigates problems with survival bias.

More specifically, given the previous year's company size, E_{t-1} , and covariates, the distribution of E_t is independent of previous years $E_{t-2}, E_{t-3}, \dots, E_0$. That is,

$$P(E_t | E_{t-1}, \dots, E_0, X_t, X_{t-1}, \dots, X_0) = P(E_t | E_{t-1}, X_t).$$

Here X_t denotes covariates for year t . This assumption is not unusual in growth modeling—in fact, it is an implicit assumption in the auto-regressive or vector auto-regressive models typically used in such modeling (Coad, 2007, 2010).

Figure 1 describes the first step. We start by saying that E_t can belong to four disjointed events: the firm hires employees, i.e., expands (A_E); the firm remains static (A_S); the firm fires employees, i.e., contracts (A_C); and the firm exits (A_X). Then, given a hiring or firing, we define the magnitude of hiring, H_t , or firing, F_t .

To define our EMM of firm growth, we operate in two stages. In the first stage, we define a probability distribution of the four events, defined above, through the following conditional probabilities:

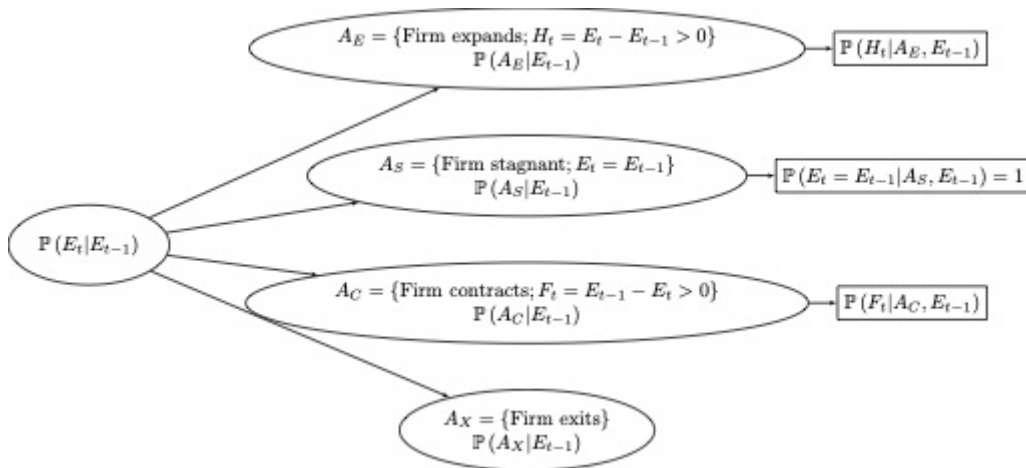
- The probability of a firm exiting: $P(A_X | E_{t-1} X_{t-1})$;
- The probability of a firm retaining the same number of employees as in the previous year (remaining static), given that the firm does not exit: $P(A_S | E_{t-1} X_{t-1}, (A_X)^C)$ (here $(A_X)^C$ is the complement event of A_X);
- The probability that a firm expands by hiring one or more new employees over the year, given that it neither remains static nor exits: $P(A_E | E_{t-1}, X_{t-1}, (A_X)^C, (A_S)^C)$.
- If none of the three above events occur, then the implication is that the firm contracts—i.e., the firm fires one or more employees.

Each of the probabilities can include covariates through the logistic regression formulation². The event of stasis (A_S) might seem redundant for larger companies, but it is crucial for micro-sized firms because 69% of these firms remain static in the data we analyze below (see Table 1). The event of remaining in stasis becomes less important the larger the venture becomes, but it is still a non-negligible factor for small firms (12%).

Figure 1

A model of new-firm dynamics taking into account the probability of an event and the magnitude of such an event. Here, E_t is number of employees at year t , H_t denotes hiring at year t , and F_t firing at year t .

² Two-part models normally only have a binary first stage, and often only one continuous distribution. That is, the distribution of H_t and F_t are handled simultaneously.



Magnitude of expansion and contraction. To complete our EMM, we also need to estimate the magnitude of expansion or contraction, i.e., to model how many people the company actually hires, or fires given that either A_E or A_C occurred. For micro-sized ventures the magnitude is often 1, 2, or 3 employees; therefore, we want a model taking only discrete values, and thus a continuous approximation for this size class is not natural. For larger firms, using a continuous distribution is probably less problematic.

For the number of people hired in the current year (H_t), we use a geometric distribution. The classical formulation of the geometric distribution comes from tossing a coin until “tails” comes up, or anything with a binary outcome (occurring with probability p). The geometric random variable is then the number of trials until the desired outcome, k . Another interpretation is that the geometric distribution is the discrete counterpart to the exponential distribution. Thus, the probability mass function, for H_t , is given by:

$$P(H_t = k; p) = g(k; p) = p(1 - p)^{k-1}, k = 1, 2, \dots$$

For the number of employees hired, the parameter p is not a natural parameter to use because we do not use the classical interpretation of the geometric distribution. Instead, we propose using theta ($\theta = 1/p = E[H_t]$), which corresponds to the expected number of the geometric distribution. To incorporate covariates in the model, we formulate a Generalized Linear Model (GLM) by the following link function: $E[H_t] = \theta = e^{X_t\beta} + 1$. Here the “+1” just

means that the number of employees hired can never be <1 . This formulation also allows us to use the number of employees in the previous year as an offset variable (offset implies a variable with a fixed coefficient, often of 1), i.e.:

$$\mathbb{E}[H_t|E_{t-1}] = E_{t-1} \exp(X_t\beta) + 1.$$

When using this offset, the independent variable effect should be interpreted as a multiplicative effect of the previous year's number of employees.

If the venture fires employees, we assume that the number of firings, F_t , follows a truncated geometric distribution. The reason for this is that we have already conditioned on the company not exiting, and hence the number of people fired must be less than the number of employees in the previous year, E_{t-1} . This results in the following density:

$$P(F_t = k|p, E_{t-1}) = g(k; p)/G(E_{t-1} - 1; p), k = 1, 2, \dots, E_{t-1} - 1,$$

where $G(E_{t-1}-1;p)$ represents the probability that geometric distribution is less than or equal to $E_{t-1} - 1$.

While we are using geometric distribution for the magnitude here, the EMM can easily be extended to other types of densities that perhaps better fit the amplitude of hiring or firing—such as a negative binomial distribution. However, we do not explore more flexible densities for our data because the geometric distribution fits these data well. This would mean that the model can accommodate even more extreme distributions—but we would then need to estimate more parameters.

In summary, the EMM estimates first the probability of change, and second the magnitude of the change, should the change occur. This procedure allows us to better separate expansion, contraction, stasis, and exit, thereby achieving the necessary flexibility to study the dynamic nature of new firm growth and survival.

Method

Data

In the previous sections, we described the problems inherent in longitudinal data on new venture growth and survival. We described what EMM is and why entrepreneurship scholars can benefit from using such a technique. We now shift focus to demonstrate how the use of EMM can potentially influence the interpretation of results. Specifically, we apply our model to a prior study on new venture growth and survival (Reference masked). We chose this paper for three main reasons. First, it uses longitudinal firm level data on new Swedish ventures. Such data have been frequently used in research published in leading management journals (References masked). We illustrate how EMM can help scholars to better understand some common firm level determinants of venture growth and survival, and to benchmark our model compared to a log normal model. Second, the chosen paper reported several limitations that we believe the EMM has the potential to address. Put differently, we will see if the EMM offers a different perspective using the same data, and whether it gives insights into how scholars can overcome some of the limitations related to longitudinal datasets on new ventures. Third, the firm level variables examined in this study yield diverging results across prior studies, and accordingly, this allows us to address the problem of having mixed results in research on new venture growth and survival.

The growth and survival of Swedish firms in terms of size distribution, exit and entry rates, as well as growth, are similar to those of other national economies (Andersson and Klepper, 2013), they represent a large share of published papers in our review, and are therefore suitable for this exercise. The data used to test the EMM covers the full population of incorporated firms founded in Sweden's knowledge-intensive sectors between 1995 and 2002. These firms comprise a significant and growing proportion of all new firms founded, and their entry and expansion are generally seen as an important engine of economic growth and change (Acs and Armington, 2004). This sector was growing rapidly during the period of investigation and had few barriers to entry. Therefore, it was attractive for new ventures because they could

enter a growing market with surplus demand not yet met by incumbent ventures. We are therefore likely to observe a large number of entries with growth potential, and at the same time the firms are subject to intensive competition, leading to a high likelihood of exit. These dynamics allow us to explore the functionality of the EMM.

Measures

Dependent variable. We choose employee growth to model new-firm growth in the main analyses for four reasons. First, this is the dependent variable used in the study we are re-analyzing; therefore, it provides potential comparison. Second, new growing firms create the lion's share of new jobs (Haltiwanger, 2015); employment is relevant to policy (Sterk et al., 2021). Third, while sales growth is most often used by organizational scholars, employment growth is a better indicator of organizational change (McKelvie and Wiklund, 2010). Employees' growth better captures stagnant firms as it is more likely that they stagnate in terms of number of employees and are less likely to achieve the same sales revenues in two subsequent years. In the EMM, employee growth is the count of the number of employees added to or subtracted from the firm, compared to the previous year. It is worth noting that in Appendix III we use sales growth, as well, to validate our model.

Independent variables. We use four independent firm-level variables of particular importance to the study of new venture growth and survival: financial aspiration levels (as a spline), Return on Assets (ROA), firm age, and firm size as number of employees during the previous year, in line with the previous published study using these data. Aspiration level for performance is calculated as a moving two-year average of a venture's past performance in terms of sales per employee. Our data comprises professional services firms, and revenue per employee is an accepted (and probably the most important) way to measure performance in the

professional services industry (Arora and Nandkumar, 2011). Such services are often labor-intensive, and the ability to sell customers a maximum number of hours per employee is a strong indicator of performance. Moreover, it is clearly related to the question of whether a decision-maker should hire more staff, or not. We include an accounting variable for performance (ROA) to ensure that our findings are more than an artifact of variability in performance. Firm age corresponds to the number of years a firm has been active since its year of entry. Firm age is an important variable in entrepreneurship and firm dynamic studies, as firms tend to alter their behaviors with age, but also because their probability of termination tends to diminish with age (Coad et al., 2014; Foster et al., 2016; Haltiwanger, 2015). We discuss firm size separately below.

Estimation procedure

To validate the EMM, we split the data into a training set (70% of the data) and a prediction set (30%). We first estimate the parameters of both the log-normal model and our model on the training set. Then we examine the predictive ability of the two models on the prediction set. We compare the effectiveness of the different models with the Pearson's Chi-squared statistic.

Were we to analyze the data in their raw form—i.e., all firms for all cohorts 1995–2002—we would have had a sampling bias issue, with older firms being overrepresented in terms of firm (i) year (j) observations. This is clear when examining the age distribution of the firms in each year. In the first year (1995) there are only firms of age 0. In the final year (2002) the ages range from 0 to 7 years. Each year, in this data set, we add a new cohort of firms born that year. This means that by the end of the observation period we have relatively more firms (i) year (j) observations from firms in the oldest cohorts versus those in the youngest cohort. That is, we oversample observations from old firms. As a robustness test, we try a different sampling approach to assess how sensitive our results are to sampling.

Firm Size. In our analysis, we separate micro firms from small firms into two samples. Firm size is important both conceptually and empirically. Conceptually, firm size is often related to growth and access to resources—for example, Gibrat’s law (Evans, 1987; Reichstein and Jensen, 2005). Empirically, the size of a coefficient and its statistical significance are driven by the most frequent firm size. In our data, 94% of the firms have 1–9 employees. Firm size is therefore an important covariate with which to compare models, as size class is likely to affect coefficients. A consequence of having data so highly dominated by a size class is that the estimated coefficients will mirror the dominant category in the sample. In our case, micro firms that are mostly static will dominate the estimates. Hence, we are unable to capture the potential causal mechanism leading to growth for other types of firms. Potentially, this underestimates what leads to growth for non-static firms. It also increases the risk of type II errors: we fail to reject the null hypothesis. In other words, we fail to detect a relationship between the dependent and independent variable because the true relationship is drowned in static micro firms.

Distribution. For the normal distribution, the magnitude (the number of employees hired or fired) is not distinguishable from the event of hiring or firing. We parameterize the distribution such that the expected number of hirings/firings is an exponential function of covariates. If the magnitude is connected to the current size of the company, one can add this as an offset in the covariates. For both the EMM and the normal distribution, the parameters are estimated using maximum likelihood on the training data set. The confidence intervals presented are computed using a Normal approximation and the Fisher information matrix. The predictive statistics below are all from the test dataset.

We estimate the parameters for both the EMM and the normal distribution on the training data set. We use maximum likelihood estimates. The confidence interval of the coefficients is obtained through second-order approximation of the likelihood function, i.e., the Hessian. We evaluate the performance of the model on the testing data set.

Results

Distributional properties of longitudinal new venture data

In this section, we address our first question pertaining to distributional properties of longitudinal new venture data, and we examine how the log-normal model and the EMM behave in comparison. To conserve space, we will report only the estimation of our results for the two dominant size classes: micro firms with 1 to 9 employees (94% of our sample, or 13,250 of 14,090 firms) and small firms with 10 to 49 employees (7% of our sample, or 1,056 of 14,090 firms). Note that these two groups overlap because micro firms can expand into small firms and small firms can contract back to micro firms.

Figures 2a and 2b show the ratio of employees in the current year, divided by the number of employees in the previous year (for small firms only), given that they do not exit in the current year, for micro firms and small firms, respectively. The histogram is based on the data from the prediction set. The dashed line is the density given by the log-normal distribution; the full line is the density of our proposed model. The full line is calculated using *density* in R from random samples of the distribution, i.e., a kernel density estimate of the empirical density.

We observe the following from the figures when studying the density dependence of the actual data. First, the log-normal model, represented by the dashed line, overall fits the data badly in both size classes, and for both firing and hiring. We then compare the density functions of our model, the log-normal model, and the data. We examine the fit by observing how well the density function generated by either the log-normal or our own model maps to the distribution of the empirical data. We find that, in the log-normal model, the density function tends to overestimate the dynamics in both firing and hiring, and to underestimate stasis. Meanwhile, the EMM achieves a better fit, as the density function generated is closer to the actual data in both size classes and for expansion as well as contraction.

Second, micro firms are highly likely to remain static (value=1) compared to small firms, which are generally more likely to change in size, in terms of both expansion and contraction. This indicates that the large representation (93% of our sample) of micro firms most likely would lead to biased estimates. Third, we do observe asymmetry in the data in both size classes. In both classes, the likelihood of hiring (17% for micro and 40% for small) is greater than the likelihood of firing (13% micro and 36% for small). This is in line with our previous argument, and is especially apparent for micro firms, who cannot contract below their actual size by firing, but can hire with no upper bound.

Finally, we also observe fat tails in the hiring side of our distribution, where a few firms are likely to grow substantially from one year to another by a factor larger than 2. We conclude that it is important to separate the data into size classes and to consider asymmetry in firing and hiring.

Model fit. To test the difference between the log-normal model and the EMM for small firms (10–49 employees), on the prediction set, we use a Chi-square test³. We first create eight groups divided by firm size, with an equal number of observations in each group in the prediction set. We then compute the expected number of firms that should be contained in each group for the two models. This allows the use of a Pearson's Chi-square test to test the null hypothesis that the empirical data follows the distribution (of firm sizes) imposed by the specific model (Pearson, 1900). In other words, we test whether the bins in a histogram of the empirical distribution of firm sizes can come from either of the two models.

Our results show that the value of the Chi-square statistic for the log-normal model is 323 compared to 4 for the EMM. The null hypothesis that a model generates the data is rejected at 99% level when the Chi-square value is above 18.5. Our findings thus reject that the data are

³ Note that we do not test the null hypothesis of equal distribution for micro firms (1–9 employees) because the mass of the distribution is heavily concentrated at 1–2 employees (76% of the observations). Thus, the test is not relevant.

normal, but do not reject that the data are from the EMM. This confirms the observations made from Figures 2a and 2b.

Figure 2

The ratio represents the number of employees in the current year divided by the number of employees in the previous year for micro firms and small firms, given they do not exit during the current year.

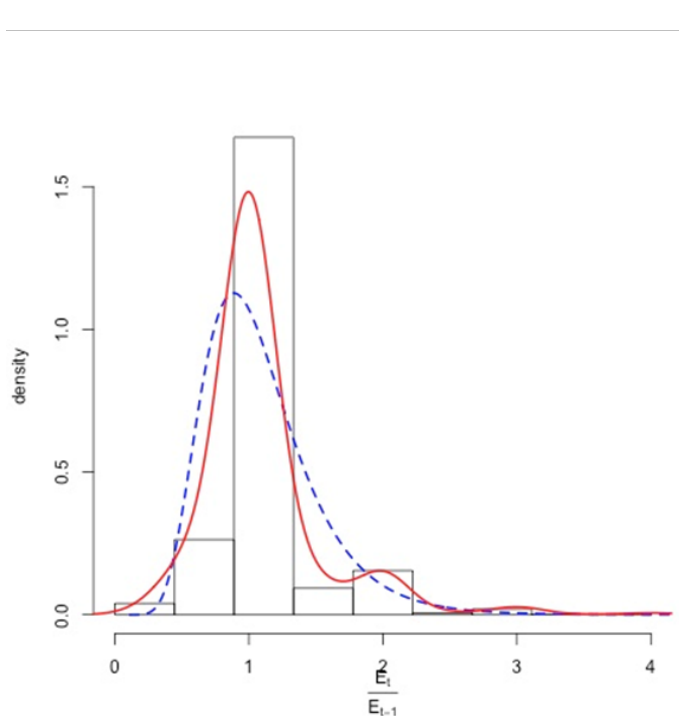


Figure 2a. Micro Firms

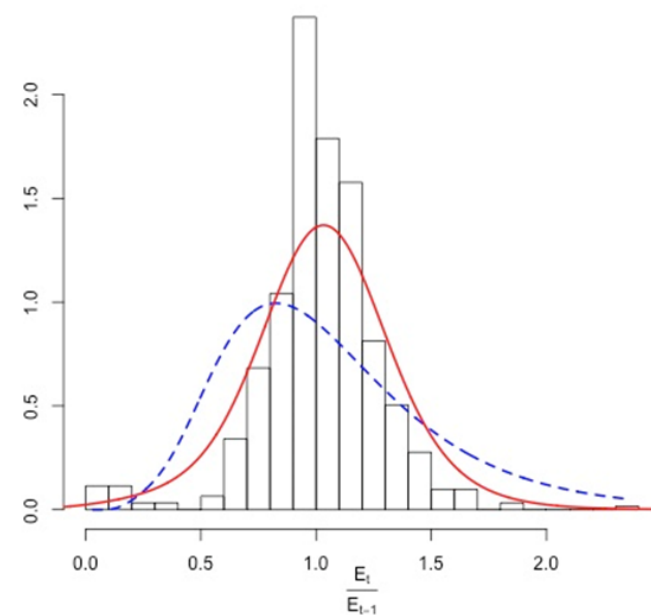


Figure 2b. Small Firms

Table 1

Probabilities of test data, micro and small firms

Prob / model	Empirical	EMM	Normal
Micro firms (1–9 employees)			
P(<i>Hiring</i>)	0.19	0.19 (0.002)	0.25 (0.003)
P(<i>Static</i>)	0.69	0.69 (0.003)	0.62 (0.003)
P(<i>Firing</i>)	0.12	0.12 (0.002)	0.13 (0.001)
Small firms (10–25 employees)			
P(<i>Hiring</i>)	0.54	0.50 (0.013)	0.46 (0.008)
P(<i>Static</i>)	0.12	0.12 (0.009)	0.08 (0.004)
P(<i>Firing</i>)	0.34	0.36 (0.012)	0.46 (0.010)

Note: Observations for micro firms n=25,379; small firms n=1,590. In parentheses are the bootstrap estimated standard deviation.

Table 1 shows the probabilities from the test data for micro firms (1–9 employees) and small firms (10–25 employees), given that the company has not yet exited the market. The three columns show the probabilities for the empirically observed values on the actual data, the proposed EMM, and the normal distribution, including standard errors in parenthesis.

For micro firms, we can see that 69% have the same number of employees as the previous year, with only 19% hiring and 12% firing. When we compare the log-normal to the actual distribution for these firms, we see that this model underestimates stasis (62%) while overestimating both hiring (25%) and firing (13%).

Turning to small firms, we observe that 12% of the firms are static, 54% are hiring, and 34% are firing. Here, the differences between the observed probabilities and probabilities generated by the log-normal model are even greater. Normal predicts equal probabilities for both firing and hiring (46%), while the real data show more hiring than firing (54% and 34% respectively). The log-normal model also underestimates stasis, with 8% compared to 12% in the actual data.

Our findings show that the EMM is much closer to the actual data than the log-normal model, in both size classes. For micro firms, our model estimates stasis at 69%, hiring at 19%,

and firing at 12%. For small firms, our model estimates stasis at 11%, hiring at 54%, and firing at 34%.

Selection Bias

Our second question is related to the presence of selection bias. We look here at survival bias in particular. Regarding the probabilities for exit, around 20% of all firms (both micro and small) will do so in our data. In the prediction dataset, the EMM estimates that for a given year, 15.6% of the micro firms and 10.2% of the small firms will exit the market. This fits the empirical data well, where 16.2% of the micro firms and 11.5% of the small firms exit each year. The probability of a firm exiting over the entire observation period in the data was 43% for micro firms and 23% for small firms.

Estimated models. Table 2 displays the results of the EMM compared to a standard OLS estimation (where the independent variable is the logarithm of the data). In stage 1, we predict the probabilities of exit and hiring (model 1 and 2). In stage 2, we separate the probability of a change (operationalized as hiring or firing of employees) from the magnitude of that change (model 3 for firing) and Model 4 for hiring). These separations are not possible with a normal or an OLS model, our reference model (Model 5)⁴. We introduce age, ROA (logarithm), and aspiration below and above the aspiration points as independent variables. The models are estimated across the classes of micro firms and small firms (Recall that firm size in the previous year is used as an offset variable).

As expected, we find that the coefficients for our independent variables differ for the probability of exit (Model 1) and the probability of hiring (Model 2). The coefficients also differ between micro and small firms. The size difference in coefficients for the same independent variable, in some cases, is substantial. For example, our results show that the coefficient for ROA (logarithm) predicting exit is -2.23 (sign.=0.000) for micro firms and -4.11 (sign.=0.000)

⁴ For the sake of simplicity, we do not include a Heckman correction to control for survival—although our two-stage model can easily incorporate such a feature at the first stage.

for small firms. Put differently, the coefficient is 1.8 times stronger for small firms than for micro firms.

Comparing the coefficients for the probability of expansion (contraction) (Model 2) with the magnitude of expansion (contraction) (Model 3 and 4 respectively), we look at the estimated coefficients for the magnitude. We find that obtaining significant effects is less likely for small firms than for micro firms. Our results also show that the effect of aspirations below and above the aspiration point disappears for small firms. In all, there are strong empirical indications that survival bias is an issue affecting the estimated models.

However, as we argued earlier, sampling strategy can bias the results we obtain. Ideally, we would have all firms in the age span of interest present in all years (Keil et al., 2009), thereby mitigating the problem with young (and hence, often very small and stagnant) firms dominating the sample to be analyzed. The dataset used here only started to follow all newly founded firms after 1991. Hence, for the early years there is no representative sample of firms—only firms of age 0 (in 1991), age 0–1 (in 1992), etc. In order to examine the potential bias of this sampling selection, we constructed two data sets: one containing firms aged 0–3, and one that included only the years containing observations of firms in all year categories (i.e., excluding 1991–1994). In Appendix II, we graphically report the difference between the two sampling procedures and the results of the analyses. The differences across models and the improved fit of our model over the log-normal one remained unchanged. Hence, the conclusions remain the same.

Asymmetry in estimations

Our third question addresses the issues of causal asymmetry in expansion (hiring) and contraction (firing). Our findings reported in table 2 show asymmetry. For the magnitude of firing (Model 3), the coefficient for profitability (*LogRoA*) is negative and significant for micro firms, but not small firms ($\beta=-2.56$, $\text{sign.}=0.02$; $\beta=-0.68$, $\text{sign.}=0.32$ respectively). For the

magnitude of hiring (Model 4), the coefficient for profitability is negative and significant for micro firms, but not small firms ($\beta=-1.36$, $\text{sign.}=0.00$; $\beta=-0.21$, $\text{sign.}=0.63$ respectively). With increasing profitability, the magnitude of hiring and firing decreases for micro firms. For the magnitude of firing, firm age is positive for micro firms but negative for small ones ($\beta=0.04$, $\text{sign.}=0.08$; $\beta=-0.01$, $\text{sign.}=0.60$ respectively). For the magnitude of hiring, firm age is significant and negative for micro and small firms ($\beta=-0.15$, $\text{sign.}=0.000$; $\beta=-0.03$, $\text{sign.}=0.02$ respectively). Our results indicate that with increasing age, the magnitude of firing increases for micro firms but does not affect small firms. With increasing age, the magnitude of hiring decreases for both micro and small firms. The differences between the coefficients are substantial, indicating different forms of asymmetries between hiring and firing, but with firm size also having an effect.

Our results also confirm the difference between the probability of expansion and its magnitude. Comparing the EMM stage 1 (probability of hiring, Model 2) with the EMM stage 2 (magnitude of hiring, Model 4) also reveals important differences in coefficient size between the two, implying that different mechanisms may be at play. Profitability has a positive effect on the probability of hiring for micro firms, but a negative one for magnitude ($\beta=3.20$, $\text{sign.}=0.00$; $\beta=-1.36$, $\text{sign.}=0.00$ respectively). Firm age has a negative effect for both the probability of hiring and its magnitude for micro firms ($\beta=-0.05$, $\text{sign.}=0.01$; $\beta=-0.15$, $\text{sign.}=0.00$, respectively). This indicates that the mechanisms affecting the probability of hiring or firing are likely to be different from those that affect the magnitude of hiring and firing.

Table 2

Estimated coefficients for our two-stage model with an OLS regression. EMM is composed of stage 1, predicting the probability of the event of exit and the probability of the event of hiring, and stage 2, predicting the magnitude of hiring and firing.

Covariates	Model 1: EMM Stage 1 Prob. (Exit)			Model 2: EMM Stage 1 Prob. (Hiring)			Model 3: EMM Stage 2 Magn. (Firing)			Model 4: EMM Stage 2 Magn. (Hiring)			Model 5: OLS (Growth)		
	β	(s.e.)	p-value	β	(s.e.)	p-value	B	(s.e.)	p-value	β	(s.e.)	p-value	β	(s.e.)	p-value
Micro firms (1-9 employees)															
<i>Intercept</i>	-1.23	(0.04)	< 10 ⁻¹⁵	0.37	(0.08)	< 10 ⁻⁰⁵	-2.77	(0.22)	< 10 ⁻¹⁵	-0.85	(0.06)	< 10 ⁻¹⁵	0.09	(0.01)	< 10 ⁻¹⁵
<i>log(roa)</i>	-2.23	(0.19)	< 10 ⁻¹⁵	3.20	(0.35)	< 10 ⁻¹⁵	-2.56	(1.10)	0.02	-1.36	(0.28)	< 10 ⁻⁰⁷	0.23	(0.02)	< 10 ⁻¹⁵
<i>age</i>	-0.09	(0.01)	< 10 ⁻¹⁵	-0.05	(0.01)	< 10 ⁻⁰⁷	0.04	(0.02)	0.08	-0.15	(0.01)	< 10 ⁻¹⁵	-0.02	(0.00)	< 10 ⁻¹⁵
<i>asspl(assp > 0)</i>	0.02	(0.003)	< 10 ⁻¹³	-0.27	(0.01)	< 10 ⁻¹⁵	0.09	(0.02)	0.0001	0.008	(0.01)	< 10 ⁻¹⁵	-0.02	(0.00)	< 10 ⁻¹⁵
<i>asspl(assp < 0)</i>	-0.04	(0.004)	< 10 ⁻¹⁶	-0.10	(0.01)	< 10 ⁻¹⁵	-0.01	(0.03)	0.72	-0.07	(0.01)	< 10 ⁻¹⁵	-0.02	(0.00)	< 10 ⁻¹⁵
Small firms (10-49 employees)															
<i>Intercept</i>	-1.12	(0.18)	< 10 ⁻¹⁰	0.42	(0.14)	< 10 ⁻¹²	-2.33	(0.16)	< 10 ⁻¹⁵	-1.67	(0.23)	< 10 ⁻¹²	0.06	(0.02)	0.003
<i>log(roa)</i>	-4.11	(1.14)	0.0003	3.63	(0.72)	< 10 ⁻⁶	-0.68	(0.69)	0.32	-0.21	(0.63)	0.63	0.19	(0.09)	0.03
<i>age</i>	-0.15	(0.03)	<10 ⁻⁷	-0.01	(0.02)	0.45	-0.01	(0.02)	0.60	-0.03	(0.06)	0.02	-0.00	(0.00)	0.06
<i>asspl(assp > 0)</i>	0.05	(0.01)	<10 ⁻⁴	-0.08	(0.01)	< 10 ⁻⁰⁷	0.09	(0.02)	< 10 ⁻⁵	0.001	(0.01)	0.40	-0.02	(0.00)	< 10 ⁻¹⁵
<i>asspl(assp < 0)</i>	0.01	(0.03)	0.75	-0.07	(0.03)	0.003	-0.05	(0.05)	0.26	-0.02	(0.01)	0.25	-0.00	(0.00)	0.11

Note: Observations for micro firms n=25,379; small firms n=1,590. OLS: For small firms: Adjusted R-squared: 0.6609; Residual standard error: 0.3108. For micro firms:

Adjusted R-squared: 0.7527; Residual standard error: 0.354.

Table 2 displays the results from an OLS regression (Model 5). The coefficients are quite different from those we obtain when we separate expansion from contraction and probability from magnitude in the EMM. As an example, the coefficient for profitability (*LogRoA*) is positive and significant for micro firms in the OLS model ($\beta=0.23$, $\text{sign.}=0.00$), but negative for the magnitude of firing ($\beta=-2.56$, $\text{sign.}=0.02$) and the magnitude of hiring ($\beta=-1.36$, $\text{sign.}=0.00$), and positive for the probability of hiring ($\beta=3.20$, $\text{sign.}=0.00$). A possible interpretation is that the OLS model mostly picks up the probability of expansion, but not its magnitude. The coefficients for age show a similar pattern. The coefficient for age is negative and significant for micro firms in the OLS model ($\beta=-0.02$, $\text{sign.}=0.00$), but positive for the magnitude of firing ($\beta=0.04$, $\text{sign.}=0.08$), negative for the magnitude of hiring ($\beta=-1.15$, $\text{sign.}=0.01$), and negative for the probability of hiring ($\beta=-0.27$, $\text{sign.}=0.00$). Hence, we can observe important differences in how coefficients behave based on the model used. To conclude, we show that taking into the distributional properties with many stationary firms and few growth firms, by separating the probability of an event from its magnitude is important to our understanding of new venture growth and survival. Further, we show that there are reasons to consider asymmetry in expansion and contraction as coefficients differ substantially. Current models fail to make these differences.

We argued earlier in our theory that the EMM is sensitive to how many observations we have per unit over time. The observed difference in firm size is likely due to many expansions and contractions. The central limit theorem suggests that the sum of many random variables converges toward a normal distribution, even if the original variables themselves are not normally distributed. Hence, the normal distribution could be suitable if we have access to an increased number of observations per firm for the observation period. We run a simulation to examine how quickly the central limit theorem kicks in.

We start by simulating 5,000 firm sizes for both small and micro firms. Each firm has a random size of 10 to 25 employees for the small firms and a random size of 1 to 9 employees for the micro firms. We then measure the firms' development each month, where each firm has either a firing or hiring event with 5% probability. If a firm has a hiring month, we generate the number of new employees from a binomial distribution where the size is the current number of employees and probability is 30%. If a firm has a firing month, we generate the number of fired people from a binomial distribution where the size is the current number of employees and probability is 10%. The parameters of the simulation are chosen so that it should approximately match the observed data (on a yearly basis). We run the simulation for each month during four years (a total of $4 \cdot 12 = 48$ monthly observations per firm) and present the result for each year in Figures 3a and 3b for micro and small firms, respectively.

In Figure 3a for micro firms, the events of actual hiring and firing are so infrequent and small that even after four years the validity of the normal approximation is questionable. In Figure 3b for small firms, we can see that because, so few events (firing or hiring) have occurred in the first year (12 observations per firm), the normal approximation is not yet valid. After four years, enough events have occurred for the central limit theorem to kick in. Thus, after four years or 48 firm-month observations, the normal approximation is a good approximation and can be safely assumed. We conclude that the EMM model is quite robust, especially for micro firms and even for small firms, if the number of observations is limited.

Figure 3a

Simulation of increased number of observations for micro firms (1–9 employees)

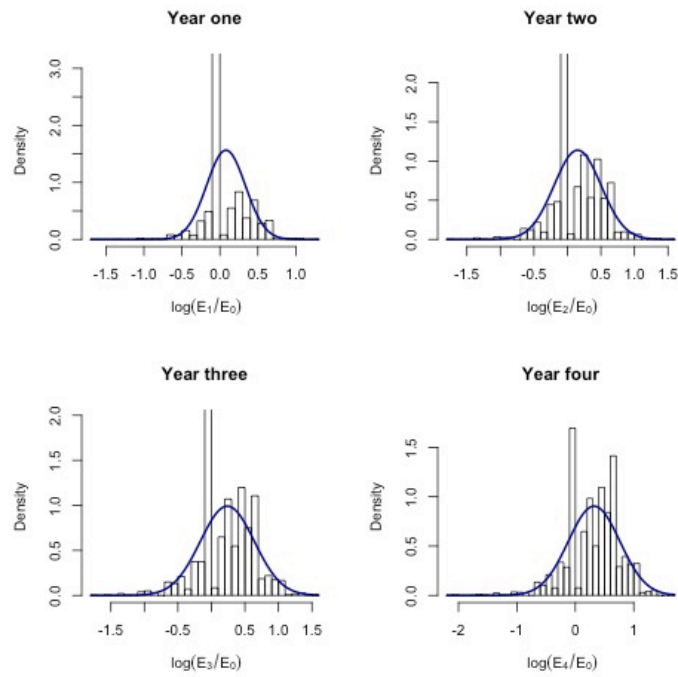
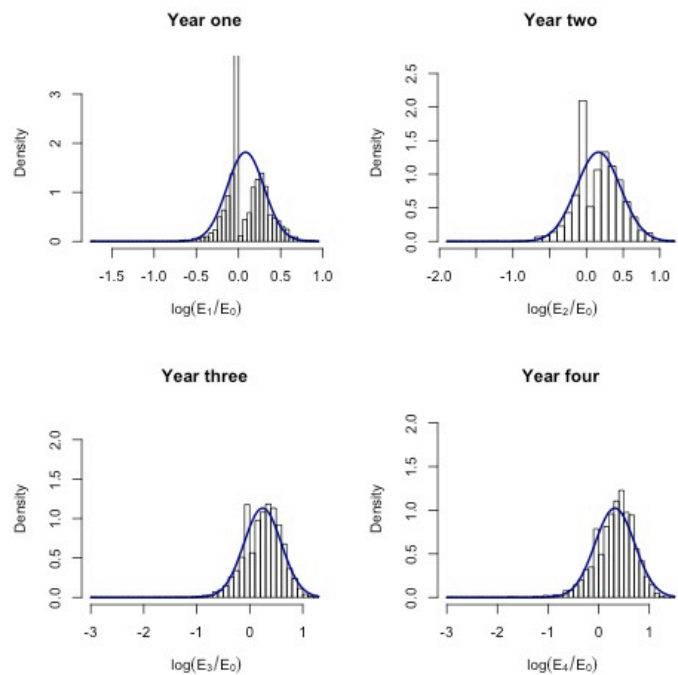


Figure 3b

Simulation of increased number of observations for small firms (10–25 employees)



Discussion

We systematically review extant literature on new venture growth and survival. Our findings show that existing research suffers from the presence of mixed evidence, selection and survival biases, and misguidance in terms of the causal mechanisms governing new venture growth and survival. These problems are connected to three methodological biases: (1) the specific distributional properties of new ventures, often making models assuming the normal distribution inadequate, (2) the presence of strong selection, such as self-selection and survival bias, increasing the risk of biased conclusions, and (3) the asymmetry between contraction and growth confounding causal mechanisms.

To mitigate these biases, this paper introduces a two-part model approach—the Event Magnitude Model, which allows organizational and entrepreneurship researchers to better separate the likelihood of an event from its magnitude. We apply the EMM to the specific problem of new-firm growth and survival. This is an informative context in which to test the EMM, as most new firms do not grow and have high exit rates; if they do grow, there is an important difference in the magnitude of their growth. Further, we argue that expansion and contraction are asymmetric, both empirically and theoretically. We make a case against using a standard normal model to explain new-firm growth and survival, and instead develop a new type of two-stage model. In its first step, the EMM estimates the probability of exit, stasis, firing, or hiring of employees using a logistic regression approach. Then, in a second step, the EMM estimates the magnitude of these changes in a similar way to a negative binomial regression. We illustrate the model using data from published research.

We find strong support for our claim that the normal model generates densities that diverge sharply from the actual data, thereby offering a poor fit to the data. Our EMM model offers a much better fit. In general, the normal model underestimates stasis and overestimates

both firing and hiring. We conclude that the log-normal model limits the development of knowledge of new-firm growth and survival.

We also find that controlling for size class is important to generate unbiased estimates. Because samples of new firms are dominated by micro firms, and those micro firms are most likely to remain in stasis, we are at risk of missing relationships causing firm expansion (e.g., hiring employees) or contraction (e.g., firing employees). For example, it may seem surprising that our results show that the effect of aspirations below and above the aspiration point disappears for small firms. However, there are two potential reasons for this difference between size classes. The first is that because the small-firm size class contains far fewer observations than the micro class, we are less likely to detect differences in this class—i.e., we lack statistical power for small firms compared to micro firms. A second reason could be that there is a dependence between the size of a firm and the aspiration level of its entrepreneurs. We can also theorize that aspiration levels are dependent on firm size and cannot assume here that the effects are purely related to issues of statistical power. Taken together, this finding emphasizes the importance of controlling for size class.

Expansion and contraction are not symmetrical in either theory or data; the data and the models clearly show that they are not mirror images. In our data, new firms are in general more likely to hire than to fire. This is also true for the magnitude of their changes in size. An obvious but still relevant observation is that firms cannot fire more employees than they already employ, but they can hire indefinite numbers of new ones. We also observe that our coefficients have different estimates when predicting probabilities versus magnitude and hiring versus firing, and when comparing our model with the standard OLS model. Other alternatives include two-stage least square models using bootstrapped standard errors (Antonakis et al., 2010; Startz and Wood-Doughty, 2016) and multilevel models (Snijders and Bosker, 2011). However, we still believe that EMM yields more reliable results for two reasons. First, if the data have rare

extreme values (i.e., in our case high growth-oriented new ventures), bootstrapping undervalues these observations—a main drawback solved by the EMM. Second, multilevel modeling follows the assumption of random effects, including the normal distribution (Antonakis et al., 2019b), which is also a problem that EMM addresses.

Implications for research on new firm growth and survival

Our EMM examines new firm growth, separating out expansion and contraction. Therefore, it provides important flexibility to better estimate how new firms develop by asking not whether firms grow or exit, but rather what is the *probability* of a new firm hiring or firing, conditional on its survival, and then what the magnitude of this change is. The EMM approach introduces much-needed flexibility in the study of new firms' growth and exit. We highlight many empirical problems—including the dominance of micro firms, asymmetries in firing and hiring, and the consequence of numerous static firms—that we believe have a strong impact on the estimated coefficients in our models.

There is no doubt that new-firm growth and new-firm survival are central topics within entrepreneurship. Explaining how this population of firms behaves, and how they generate economic value—whether as a group, or as individual firms in competition—is crucial to understanding the function of entrepreneurship in modern economies. However, the vast majority of entrepreneurship literature has neglected the problem of how to best estimate growth and survival. These two topics are rarely tackled jointly within the same paper—even though we have powerful evidence that survival and growth are linked at both the population (Carroll and Hannan, 2000; Geroski and Mazzucato, 2001) and individual-firm levels (Delmar et al., 2013). At the population level, churning (i.e., the rates of firm entry and exit) (Barnett et al., 2003; Luttmer, 2011), especially in a new industry, is correlated to firm growth (Klepper, 1996). At the firm level, under-sampling of firms' failure leads to both success bias and survival bias in our models (Denrell, 2003).

We believe that intermingling the probabilities and the magnitude of change is an important limit to progress in this domain. The relatively simple EMM we have developed and tested here provides a first step for research to be more aware of the natural challenges of the data we seek to analyze. For example, an interesting particularity of a two-stage model is that we can distinguish between the effect of covariates that predict an event and those that predict the magnitude. We can theorize whether, for example, profitability affects the likelihood of an event (firing, hiring, stasis, or exit) and also its magnitude. This is a context where we can imagine the different causal mechanisms that might be at play. The ability to better explain “how” firms grow is suggested as an important step to advance research (Shepherd and Patzelt, 2022). Hence, our model potentially opens up new ways to theorize new-firm growth and survival. We believe that continuing to impose normal distribution on data that is not normally distributed is not the answer; we are merely forcing a square peg into a round hole. Instead, our models need to adapt to the data we wish to analyze. In addition, we show that the EMM can separate out and deal more effectively with zero-inflated distributions, where most units never change or exhibit only small fluctuations over time. Indeed, by developing models that better fit data, we can advance research in this area.

Such research also has important implications for policy. Many studies have pointed to the importance of new high-growth firms to rejuvenate economies—in terms of employment, but also in terms of generating innovation and sharpening competition. Research has had a difficult time developing models that go beyond strong description and actually test causal mechanisms that might account for high-growth firms (e.g., Sterk et al., 2021). While we may not have reached that milestone in this one study, we do point to a possible way forward, using models that are just as easy to implement as normal. We also show that increases and decreases in an outcome variable are not necessarily symmetric, for not only theoretical rationale but also for empirical reasons.

Limitations and Future Research

Our study has several limitations that could give pause to future avenues for research. To start, we have only tested our model on a single sample. Even though we have used both a training and a prediction set, we might still be over-fitting to the dynamics of this sample. Future research will need to expand the EMM to other samples.

Second, we have only partially resolved the problem with stasis versus change and magnitude. We show that micro firms are more static than small firms. This is partly because micro firms have less of a stochastic element tied to their employment size. That is, a larger firm is more likely to experience some change in size than a smaller one, because firms become more subject to random changes in their size as they grow. Specifically, the changes we observe are not due to the causal mechanisms we are interested in. Rather, they are since, with increasing size, the probability of an employee leaving, or of the firm hiring an extra employee to replace a planned departure (e.g., a retirement), just increases. Hence, small changes in firm size might often just be stochastic movement around a static firm size. A next step, therefore, is to try to separate stochastic changes or noisy variation from true changes in firm size.

One potential approach would be to relax the Markov assumption, and instead assume that the companies are in an active or static stage that is modeled by a hidden Markov model (Cappé, Moulines, and Ryden, 2005). This approach has been successfully applied in myriad fields, such as finance (Rydén et al., 1998), driver behavior (Maghsood et al., 2016), and speech recognition (Rabiner, 1989). The idea is that stasis is a hidden state in which firms have a small magnitude of firing and hiring—or, alternatively, natural fluctuations in sales or profits. These small fluctuations are purely due to randomness, and do not indicate any overall growth or decline of the firm.

Third, while we discussed how to best model new-firm growth and survival using an existing sample, we offer little advice on how to design such studies in the future. A

fundamental challenge in new-firm research is that we have few observations per unit of observation. Because new firms on average have short life spans, we will have few observations per firm when modeling new-firm growth and survival. Working with fewer observations leads to a higher stochastic element than studying long-established firms. To mitigate this problem, we probably need data with shorter time intervals than the standard one-year interval found in most register data (e.g., Coad et al., 2013). This was the objective of our simulation. We show that this was difficult to achieve without a substantial number of observations, and especially problematic for micro firms. Hence, the need for an increased number of observations per period of observation has to be balanced against the cost of such data, and the observed fact that most new and micro firms remain relatively static.

Increasing cohort size, or incorporating already-existing firms into a sample, are not perfect solutions either. Looking solely at cohorts of new firms offers several advantages, but also some disadvantages. On the plus side, cohorts provide strong information about age and age-related variation. For example, in many cohorts, the variation across causal mechanisms tends to be large, because causal mechanisms are emerging as the firms create routines and solidify the business model. Consequently, the variation will diminish as a function of time. The age coefficient might be different for cohorts of younger firms than in a sample mixing cohorts of new firms with incumbent firms. However, non-growing micro firms will dominate such designs. Therefore, comprehensive data including both new firm entry and incumbents may provide a better alternative to study growth. Future work will have to examine how to best design such studies.

While we have briefly discussed both survival bias and selection bias, the details of when different forms of bias affect our estimated coefficients have yet to be fully estimated. That is, we have not fully taken account of how the number of new firms per cohort, and their relative survival rates, affect the distribution of observations in a population of new firms.

Conclusion

Current research tends to underestimate the problems related to studying new firm growth and survival. In this paper, we argue that firm survival and firm growth are related, and that firm growth is rare but can be substantial. We also argue that firm contraction and expansion are not symmetrical—neither theoretically nor empirically. Therefore, the standard models assuming the normal distribution are not appropriate. Rather, we suggest a two-stage model called the “Event Magnitude Model”, which explores the probability of an event and then its magnitude. We empirically show that the EMM outperforms the normal model. Our EMM thereby opens new avenues to develop and test theories that are not possible using log-normal distribution models. We have provided a step-by-step example showing that our two-stage model fits data better, changes the value of estimated coefficients, and opens a new way to cut data and thus generate new insights. We hope that our work will contribute to a vital discussion of how we can model rare events with potential large magnitudes for organizations, entrepreneurship, and new-firm dynamics. Finally, we have developed the code for R with simulated data, making our model readily accessible and easy to use.

References

- Acemoglu D., Akcigit U., Alp H., Bloom N., Kerr W. Innovation, reallocation, and growth. *American Economic Review* 2018;108:3450-3491.
- Acs Z., Armington C. Employment growth and entrepreneurial activity in cities. *Regional studies* 2004;38:911-927.
- Aguinis H., Gottfredson R.K., Joo H. Best-Practice Recommendations for Defining, Identifying, and Handling Outliers. *Organizational Research Methods* 2013;16:270-301.
- Aksaray G., Thompson P. Density dependence of entrepreneurial dynamics: Competition, opportunity cost, or minimum efficient scale? *Management Science* 2018;64:2263-2274.
- Aldrich H. *Organizations evolving*. Sage Publications Ltd; 1999.
- Alsos G.A., Isaksen E.J., Ljunggren E. New venture financing and subsequent business growth in men- and women-led businesses. *Entrepreneurship Theory and Practice* 2006;30:667-686.
- Andersson M., Klepper S. Characteristics and performance of new firms and spinoffs in Sweden. *Industrial and Corporate Change* 2013;22:245-280.
- Anglin A.H., McKenny A.F., Short J.C. The impact of collective optimism on new venture creation and growth: A social contagion perspective. *Entrepreneurship Theory and Practice* 2018;42:390-425.
- Antonakis J., Bastardoz N., Rönkkö M. On ignoring the random effects assumption in multilevel models: Review, critique, and recommendations. *Organizational Research Methods* 2019a:1094428119877457.
- Antonakis J., Bastardoz N., Rönkkö M. On Ignoring the Random Effects Assumption in Multilevel Models: Review, Critique, and Recommendations. *Organizational Research Methods* 2019b;0:1094428119877457.
- Antonakis J., Bendahan S., Jacquart P., Lalive R. On making causal claims: A review and recommendations. *The Leadership Quarterly* 2010;21:1086-1120.
- Arora A., Nandkumar A. Cash-out or flameout! Opportunity cost and entrepreneurial strategy: Theory, and evidence from the information security industry. *Management Science* 2011;57:1844-1860.
- Artz K.W., Norman P.M., Hatfield D.E., Cardinal L.B. A longitudinal study of the impact of R&D, patents, and product innovation on firm performance. *Journal of Product Innovation Management* 2010;27:725-740.
- Aspelund A., Berg-Utby T., Skjævdal R. Initial resources' influence on new venture survival: a longitudinal study of new technology-based firms. *Technovation* 2005;25:1337-1347.
- Barkema H.G., Vermeulen F. What differences in the cultural backgrounds of partners are detrimental for international joint ventures? *Journal of International Business Studies* 1997;28:845-864.
- Barnett W.P., Swanson A.-N., Sorenson O. Asymmetric selection among organizations. *Industrial and Corporate Change* 2003;12:673-695.
- Barron D.N., West E., Hannan M.T. A Time to Grow and a Time to Die: Growth and Mortality of Credit Unions in New York City, 1914-1990. *American Journal of Sociology* 1994;100:381-421.
- Bastié F., Cieply S., Cussy P. Does mode of transfer matter for business performance? Transfers to employees versus transfers to outsiders. *Small Business Economics* 2018;50:77-89.
- Basu S., Wadhwa A. External venturing and discontinuous strategic renewal: An options perspective. *Journal of Product Innovation Management* 2013;30:956-975.
- Battisti M., Beynon M., Pickernell D., Deakins D. Surviving or thriving: The role of learning for the resilient performance of small firms. *Journal of Business Research* 2019;100:38-50.

- Baum J.R., Locke E.A. The Relationship of Entrepreneurial Traits, Skill, and Motivation to Subsequent Venture Growth. *Journal of Applied Psychology* 2004;89:587-598.
- Ben-Hafaïedh C., Hamelin A. Questioning the Growth Dogma: A Replication Study. *Entrepreneurship Theory and Practice*;0:10422587211059991.
- Bennett V.M., Levinthal D.A. Firm Lifecycles: Linking Employee Incentives and Firm Growth Dynamics. *Strategic Management Journal* 2017;38:2005-2018.
- Bertoni F., Colombo M.G., Grilli L. Venture capital financing and the growth of high-tech start-ups: Disentangling treatment from selection effects. *Research Policy* 2011;40:1028-1043.
- Bertoni F., Colombo M.G., Grilli L. Venture capital investor type and the growth mode of new technology-based firms. *Small Business Economics* 2013;40:527-552.
- Billingsley P. *Probability and measure*. New York: Wiley; 1995.
- Bjørnskov C., Foss N.J. Institutions, entrepreneurship, and economic growth: what do we know and what do we still need to know? *Academy of Management Perspectives* 2016;30:292-315.
- Bradley S.W., Wiklund J., Shepherd D.A. Swinging a double-edged sword: The effect of slack on entrepreneurial management and growth. *Journal of Business Venturing* 2011;26:537-554.
- Brush C.G., Manolova T.S., Edelman L.F. Properties of emerging organizations: An empirical test. *Journal of Business Venturing* 2008;23:547-566.
- Calvino F. Technological innovation and the distribution of employment growth: a firm-level analysis. *Industrial and Corporate Change* 2018;28:177-202.
- Card D., Cardoso A.R., Heining J., Kline P. Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics* 2018;36:S13-S70.
- Carroll G.R., Hannan M.T. *The Demography of Corporations and Industries*. Princeton, New Jersey: Princeton University Press; 2000.
- Casillas J.C., Acedo F.J., Rodríguez-Serrano M.Á. How does internationalization begin? The role of age at entry and export experience in the early stages of the process. *BRQ Business Research Quarterly* 2020;23:107-119.
- Cefis E., Marsili O. Revolving doors: entrepreneurial survival and exit. *Journal of Evolutionary Economics* 2011;21:367-372.
- Cefis E., Marsili O. Good times, bad times: innovation and survival over the business cycle. *Industrial and Corporate Change* 2019;28:565-587.
- Certo S.T., Busenbark J.R., Woo H.-S., Semadeni M. Sample selection bias and heckman models in strategic management research. *Strategic Management Journal* 2015;n/a-n/a.
- Certo S.T., Busenbark J.R., Woo H.s., Semadeni M. Sample selection bias and Heckman models in strategic management research. *Strategic Management Journal* 2016;37:2639-2657.
- Certo S.T., Withers M.C., Semadeni M. A tale of two effects: Using longitudinal data to compare within- and between-firm effects. *Strategic Management Journal* 2017;38:1536-1556.
- Chen Y., Song M. The persistence and dynamics of new venture growth. *Small Business Economics* 2020:1-20.
- Chrisman J.J., McMullan W.E. Outsider assistance as a knowledge resource for new venture survival. *Journal of small business management* 2004;42:229-244.
- Clougherty J.A., Duso T., Muck J. Correcting for Self-selection Based Endogeneity in Management Research: Review, Recommendations and Simulations. *Organizational Research Methods* 2015;19:286-347.
- Coad A. *Firm growth: A survey*. 2007.
- Coad A. Exploring the processes of firm growth: evidence from a vector auto-regression. *Industrial and Corporate Change* 2010;19:1677-1703.

- Coad A., Daunfeldt S.-O., Johansson D., Wennberg K. Whom do high-growth firms hire? *Industrial and Corporate Change* 2014;23:293-327.
- Coad A., Frankish J., Roberts R.G., Storey D.J. Growth paths and survival chances: An application of Gambler's Ruin theory. *Journal of Business Venturing* 2013;28:615-632.
- Coad A., Frankish J.S., Roberts R.G., Storey D.J. Are firm growth paths random? A reply to "Firm growth and the illusion of randomness". *Journal of Business Venturing Insights* 2015;3:5-8.
- Coad A., Nielsen K., Timmermans B. My first employee: an empirical investigation. *Small Business Economics* 2017;48:25-45.
- Coad A., Srhoj S. Catching Gazelles with a Lasso: Big data techniques for the prediction of high-growth firms. *Small Business Economics* 2020;55:541-565.
- Coeurderoy R., Cowling M., Licht G., Murray G. Young firm internationalization and survival: Empirical tests on a panel of 'adolescent' new technology-based firms in Germany and the UK. *International Small Business Journal* 2012;30:472-492.
- Colombo M.G., Giannangeli S., Grilli L. Public subsidies and the employment growth of high-tech start-ups: assessing the impact of selective and automatic support schemes. *Industrial and Corporate Change* 2013;22:1273-1314.
- Colombo M.G., Murtinu S. Venture Capital Investments in Europe and Portfolio Firms' Economic Performance: Independent Versus Corporate Investors. *Journal of Economics & Management Strategy* 2017;26:35-66.
- Cooper A.C., Gimeno-Gascon F.J., Woo C.Y. Initial human and financial capital as predictors of new venture performance. *Journal of Business Venturing* 1994;9:371-395.
- Crawford G.C., Aguinis H., Lichtenstein B., Davidsson P., McKelvey B. Power law distributions in entrepreneurship: Implications for theory and research. *Journal of Business Venturing* 2015;30:696-713.
- Daunfeldt S.-O., Halvarsson D. Are high-growth firms one-hit wonders? Evidence from Sweden. *Small Business Economics* 2015;44:361-383.
- Davidsson P., Delmar F., Wiklund J. *Entrepreneurship and the Growth of Firms*. Edward Elgar; 2006.
- Davidsson P., Steffens P., Fitzsimmons J. Growing profitable or growing from profits: Putting the horse in front of the cart? *Journal of Business Venturing* 2009;24:388-406.
- Delmar F., Davidsson P., Gartner W.B. Arriving at the high-growth firm. *Journal of Business Venturing* 2003;18:189-216.
- Delmar F., McKelvie A., Wennberg K. Untangling the relationships among growth, profitability and survival in new firms. *Technovation* 2013;33:276-291.
- Delmar F., Shane S. Does business planning facilitate the development of new ventures? *Strategic Management Journal* 2003;24:1165-1185.
- Delmar F., Shane S. Legitimizing first: organizing activities and the survival of new ventures. *Journal of Business Venturing* 2004;19:385-410.
- Delmar F., Shane S. Does experience matter? The effect of founding team experience on the survival and sales of newly founded ventures. *Strategic Organization* 2006;4:215-247.
- Delmar F., Wennberg K., Hellerstedt K. Endogenous growth through knowledge spillovers in entrepreneurship: an empirical test. *Strategic Entrepreneurship Journal* 2011;5:199-226.
- Delmar F., Wiklund J. The Effect of Small Business Managers' Growth Motivation on Firm Growth: A Longitudinal Study. *Entrepreneurship Theory and Practice* 2008;32:437-457.
- Demir R., Wennberg K., McKelvie A. *The Strategic Management of High-Growth Firms: A Review and Theoretical Conceptualization*. Long Range Planning 2016.
- Deng Z., Jean R.-J.B., Sinkovics R.R. Polarizing effects of early exporting on exit. *Management International Review* 2017;57:243-275.

- Denrell J. Vicarious learning, undersampling of failure, and the myths of management. *Organization Science* 2003;14:227-243.
- Derbyshire J., Garnsey E. Firm growth and the illusion of randomness. *Journal of Business Venturing Insights* 2014;1-2:8-11.
- Douglas E.J., Shepherd D.A., Prentice C. Using fuzzy-set qualitative comparative analysis for a finer-grained understanding of entrepreneurship. *Journal of Business Venturing* 2020;35:105970.
- Eberhart R.N., Eesley C.E., Eisenhardt K.M. Failure Is an Option: Institutional Change, Entrepreneurial Risk, and New Firm Growth. *Organization Science* 2017;28:93-112.
- Ebert T., Brenner T., Brixy U. New firm survival: the interdependence between regional externalities and innovativeness. *Small Business Economics* 2019;53:287-309.
- Edelman L.F., Brush C.G., Manolova T.S., Greene P.G. Start-up motivations and growth intentions of minority nascent entrepreneurs. *Journal of small business management* 2010;48:174-196.
- Elwert F., Winship C. Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. *Annual review of sociology* 2014;40:31-53.
- Erlingsson E., Alfarano S., Raberto M., Stefánsson H. On the distributional properties of size, profit and growth of Icelandic firms. *Journal of Economic Interaction and Coordination* 2013;8:57-74.
- Evans D.S. Tests of alternative theories of firm growth. *Journal of Political Economy* 1987;95:657-674.
- Farewell V.T., Long D.L., Tom B.D.M., Yiu S., Su L. Two-Part and Related Regression Models for Longitudinal Data. *Annual Review of Statistics and Its Application* 2017;4:283-315.
- Fariborzi H., Keyhani M. Internationalize to live: a study of the post-internationalization survival of new ventures. *Small Business Economics* 2018;50:607-624.
- Fattoum-Guedri A., Delmar F., Wright M. The best of both worlds: Can founder-CEOs overcome the rich versus king dilemma after IPO? *Strategic Management Journal* 2018;39:3382-3407.
- Follmer K.B., Jones K.S. Mental Illness in the Workplace: An Interdisciplinary Review and Organizational Research Agenda. *Journal of Management* 2018;44:325-351.
- Foster L., Haltiwanger J., Syverson C. The Slow Growth of New Plants: Learning about Demand? *Economica* 2016;83:91-129.
- Frid C.J., Wyman D.M., Coffey B. Effects of wealth inequality on entrepreneurship. *Small Business Economics* 2016;47:895-920.
- Friske W.M., Zachary M.A. Regulation, New Venture Creation, and Resource-Advantage Theory: An Analysis of the US Brewing Industry. *Entrepreneurship Theory and Practice* 2019;43:999-1017.
- Garnsey E., Stam E., Heggeman P. New Firm Growth: Exploring Processes and Paths. *Industry and Innovation* 2006;13:1-20.
- Geroski P.A., Mazzucato M. Modelling the dynamics of industry population. *International Journal of Industrial Organization* 2001;19:1003-1022.
- Gideon K., Barley S.R., Evans J. Why Do Contractors Contract? The Experience of Highly Skilled Technical Professionals in a Contingent Labor Market. *Industrial and Labor Relations Review* 2002;55:234-261.
- Gielnik M.M., Zacher H., Schmitt A. How small business managers' age and focus on opportunities affect business growth: a mediated moderation growth model. *Journal of small business management* 2017;55:460-483.
- Gilbert B.A., McDougall P.P., Audretsch D.B. New venture growth: A review and extension. *Journal of Management* 2006;32:926-950.

- Gimeno J., Folta T.B., Cooper A.C., Woo C.Y. Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. *Administrative Science Quarterly* 1997;42:750-783.
- Gjerløv-Juel P., Guenther C. Early employment expansion and long-run survival: examining employee turnover as a context factor. *Journal of Business Venturing* 2019;34:80-102.
- Grilli L., Murtinu S. Government, venture capital and the growth of European high-tech entrepreneurial firms. *Research Policy* 2014;43:1523-1543.
- Grillitsch M., Schubert T. Does the timing of integrating new skills affect start-up growth? *Strategic Entrepreneurship Journal* 2020.
- Gruenhagen J.H., Sawang S., Gordon S.R., Davidsson P. International experience, growth aspirations, and the internationalisation of new ventures. *Journal of International Entrepreneurship* 2018;16:421-440.
- Haeussler C., Hennicke M., Mueller E. Founder-inventors and their investors: Spurring firm survival and growth. *Strategic Entrepreneurship Journal* 2019;13:288-325.
- Haltiwanger J. Job Creation, Job Destruction, and Productivity Growth: The Role of Young Businesses. *Annual Review of Economics* 2015;7:341-358.
- Hatemi-j A. Asymmetric causality tests with an application. *Empirical Economics* 2012;43:447-456.
- Hayton J.C. Competing in the new economy: the effect of intellectual capital on corporate entrepreneurship in high-technology new ventures. *R&D Management* 2005;35:137-155.
- Headd B. Redefining Business Success: Distinguishing Between Closure and Failure. *Small Business Economics* 2003;21:51-61.
- Heckman J.J. Sample Selection Bias as a Specification Error. *Econometrica* 1979;47:153-161.
- Henrekson M., Johansson D. Gazelles as job creators: a survey and interpretation of the evidence. *Small Business Economics* 2010;35:227-244.
- Henrekson M., Sanandaji T. Small business activity does not measure entrepreneurship. *Proc Natl Acad Sci U S A* 2014;111:1760-1765.
- Hernandez-Vivanco A., Domingues P., Sampaio P., Bernardo M., Cruz-Cázares C. Do multiple certifications leverage firm performance? A dynamic approach. *International Journal of Production Economics* 2019;218:386-399.
- Hitt M., Harrison J., Ireland R.D., Best A. Attributes of Successful and Unsuccessful Acquisitions of US Firms. *British Journal of Management* 1998;9:91-114.
- Hmieleski K.M., Baron R.A. Entrepreneurs' optimism and new venture performance: A social cognitive perspective. *Academy of Management Journal* 2009;52:473-488.
- Holcomb T.R., Combs J.G., Sirmon D.G., Sexton J. Modeling levels and time in entrepreneurship research: An illustration with growth strategies and post-IPO performance. *Organizational Research Methods* 2010;13:348-389.
- Honig B., Karlsson T. Institutional forces and the written business plan. *Journal of Management* 2004;30:29-48.
- Hopp C., Greene F.J. In pursuit of time: Business plan sequencing, duration and intraentrainment effects on new venture viability. *Journal of Management Studies* 2018;55:320-351.
- Hurst E., Pugsley B.W. What Do Small Businesses Do? . National Bureau of Economic Research Working Paper Series 2011;17041.
- Hutzschenreuter T., Horstkotte J. Managerial services and complexity in a firm's expansion process: An empirical study of the impact on the growth of the firm. *European Management Journal* 2013;31:137-151.
- Huynh K.P., Petrunia R.J. Age effects, leverage and firm growth. *Journal of Economic Dynamics and Control* 2010;34:1003-1013.

- Hvide H.K., Møen J. Lean and Hungry or Fat and Content? Entrepreneurs' Wealth and Start-Up Performance. *Management Science* 2010;56:1242-1258.
- Jennings J.E., Jennings P.D., Greenwood R. Novelty and new firm performance: The case of employment systems in knowledge-intensive service organizations. *Journal of Business Venturing* 2009;24:338-359.
- Jovanovic B. Selection and the evolution of industry. *Econometrica* 1982;50:649-670.
- Jovanovic B., Lach D. Entry, exit and diffusion with learning by doing. *American Economic Review* 1989;79:690-699.
- Kang T., Baek C., Lee J.-D. Effects of knowledge accumulation strategies through experience and experimentation on firm growth. *Technological Forecasting and Social Change* 2019;144:169-181.
- Keil T., McGrath R.G., Tukiainen T. Gems from the ashes: Capability creation and transformation in internal corporate venturing. *Organization Science* 2009;20:601-620.
- Khaire M. Young and no money? Never mind: The material impact of social resources on new venture growth. *Organization Science* 2010;21:168-185.
- Kim S.M., Anand G., Larson E.C., Mahoney J. Resource co-specialization in outsourcing of enterprise systems software: Impact on exchange success and firm growth. *Journal of Science and Technology Policy Management* 2019.
- Klepper S. Entry, Exit, Growth, and Innovation over the Product Life Cycle. *American Economic Review* 1996;86:562-583.
- Kolvereid L., Isaksen E.J. Expectations and achievements in new firms. *Journal of Small Business and Enterprise Development* 2017;24.
- Kor Y.Y. Experience-Based Top Management Team Competence and Sustained Growth. *Organization Science* 2003;14:707-719.
- Korunka C., Kessler A., Frank H., Lueger M. Personal characteristics, resources, and environment as predictors of business survival. *Journal of Occupational and Organizational Psychology* 2010;83:1025-1051.
- Koudstaal M., Sloof R., Praag M.v. Risk, Uncertainty, and Entrepreneurship: Evidence from a Lab-in-the-Field Experiment. *Management Science* 2016;62:2897-2915.
- Koufteros X., Verghese A., Lucianetti L. The effect of performance measurement systems on firm performance: A cross-sectional and a longitudinal study. *Journal of Operations Management* 2014;32:313-336.
- Levine R., Rubinstein Y. Smart and illicit: who becomes an entrepreneur and do they earn more? *The Quarterly Journal of Economics* 2017;132:963-1018.
- Liao J., Gartner W.B. The Effects of Pre-venture Plan Timing and Perceived Environmental Uncertainty on the Persistence of Emerging Firms. *Small Business Economics* 2006;27:23-40.
- Löfsten H. New technology-based firms and their survival: The importance of business networks, and entrepreneurial business behaviour and competition. *Local Economy* 2016;31:393-409.
- Lopez-Garcia P., Puente S. What makes a high-growth firm? A dynamic probit analysis using Spanish firm-level data. *Small Business Economics* 2012;39:1029-1041.
- Luttmer E.G.J. Models of Growth and Firm Heterogeneity. *Annual Review of Economics* 2010;2:547-576.
- Luttmer E.G.J. On the Mechanics of Firm Growth. *The Review of Economic Studies* 2011;78:1042-1068.
- Maghsood R., Johannesson P., Wallin J. Detection of steering events using hidden Markov models with multivariate observations. *International Journal of Vehicle Systems Modelling and Testing* 2016;11:313-329 %@ 1745-6436.
- Mai Y., Zheng Y. How On-the-Job Embeddedness Influences New Venture Creation and Growth. *Journal of small business management* 2013a;51:508-524.

- Mai Y., Zheng Y. How On-the-Job Embeddedness Influences New Venture Creation and Growth. *Journal of small business management* 2013b;51:508-524.
- Maliranta M., Nurmi S. Business owners, employees, and firm performance. *Small Business Economics* 2019;52:111-129.
- March J.G., Shapira Z. Variable risk preferences and the focus of attention. *Psychological review* 1992;99:172-183.
- Mazumdar S., Tang G., Houck P.R., Dew M.A., Begley A.E., Scott J. et al. Statistical analysis of longitudinal psychiatric data with dropouts. *Journal of Psychiatric Research* 2007;41:1032-1041.
- McKelvie A., Wiklund J. Advancing Firm Growth Research: A Focus on Growth Mode Instead of Growth Rate. *Entrepreneurship Theory and Practice* 2010;34:261-288.
- Mens G.L., Hannan M.T., Pólos L. Age-Related Structural Inertia: A Distance-Based Approach. *Organization Science* 2015;26:756-773.
- Messersmith J.G., Patel P.C., Crawford C. Bang for the buck: Understanding employee benefit allocations and new venture survival. *International Small Business Journal* 2018;36:104-125.
- Meyer-Doyle P., Lee S., Helfat C.E. Disentangling the microfoundations of acquisition behavior and performance. *Strategic Management Journal* 2019;40:1733-1756.
- Millán J., Millán A., Román C., Stel A. Unraveling the relationship between the business cycle and the own-account worker's decision to hire employees. *International Entrepreneurship and Management Journal* 2015;11:321-342.
- Musso P., Schiavo S. The impact of financial constraints on firm survival and growth. *Journal of Evolutionary Economics* 2008;18:135-149.
- Nadeem M., Gan C., Nguyen C. The importance of intellectual capital for firm performance: Evidence from Australia. *Australian Accounting Review* 2018;28:334-344.
- Naldi L., Davidsson P. Entrepreneurial growth: The role of international knowledge acquisition as moderated by firm age. *Journal of Business Venturing* 2014;29:687-703.
- Navarro J.L.B., Casillas J.C., Barringer B. Forms of growth: How SMEs combine forms of growth to achieve high growth. *Journal of Management and Organization* 2012;18:81.
- Nielsen K. Human capital and new venture performance: the industry choice and performance of academic entrepreneurs. *The Journal of Technology Transfer* 2015;40:453-474.
- Nofal A.M., Nicolaou N., Symeonidou N., Shane S. Biology and Management: A Review, Critique, and Research Agenda. *Journal of Management* 2018;44:7-31.
- Ozcan P. Growing with the market: How changing conditions during market growth affect formation and evolution of interfirm ties. *Strategic Management Journal* 2018;39:295-328.
- Palestrini A. Firm Size Distribution and the Survival Bias. *Economics Bulletin* 2015;35:1630-1637.
- Papangkorn S., Chatjuthamard P., Jiraporn P., Chueykamhang S. Female directors and firm performance: Evidence from the Great Recession. *International Review of Finance* 2019;21.
- Parker S.C. *The economics of entrepreneurship*. Cambridge University Press; 2009.
- Pe'er A., Keil T. Are all startups affected similarly by clusters? Agglomeration, competition, firm heterogeneity, and survival. *Journal of Business Venturing* 2013;28:354-372.
- Pearson K. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 1900;50:157-175.
- Pfeffer J., Salancik G.R. *The External Control of Organizations: A Resource Dependence Perspective*. Stanford: Stanford University Press; 2003.
- Ployhart R.E., Vandenberg R.J. Longitudinal research: The theory, design, and analysis of change. *Journal of Management* 2010;36:94-120.

- Prashantham S., Dhanaraj C. The Dynamic Influence of Social Capital on the International Growth of New Ventures. *Journal of Management Studies* 2010;47:967-994.
- Rabiner L.R. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE* 1989;77:257-286 %@ 0018-9219.
- Reichstein T., Dahl M.S., Ebersberger B., Jensen M.B. The devil dwells in the tails. A quantile regression approach to firm growth. *Journal of Evolutionary Economics* 2010;20:219-231.
- Reichstein T., Jensen M.B. Firm size and firm growth rate distributions—The case of Denmark. *Industrial and Corporate Change* 2005;14:1145-1166.
- Robb A.M., Watson J. Gender differences in firm performance: Evidence from new ventures in the United States. *Journal of Business Venturing* 2012;27:544-558.
- Robinson K.C., Phillips McDougall P. Entry barriers and new venture performance: a comparison of universal and contingency approaches. *Strategic Management Journal* 2001;22:659-685.
- Rocha V., van Praag M., Folta T.B., Carneiro A. Endogeneity in Strategy-Performance Analysis: An Application to Initial Human Capital Strategy and New Venture Performance. *Organizational Research Methods* 2018;1094428118757313.
- Rothaermel F.T., Hill C.W. Technological discontinuities and complementary assets: A longitudinal study of industry and firm performance. *Organization Science* 2005;16:52-70.
- Rutherford M.W., Mazzei M.J., Oswald S.L., Jones-Farmer L.A. Does establishing sociopolitical legitimacy overcome liabilities of newness? A longitudinal analysis of top performers. *Group & Organization Management* 2018;43:906-935.
- Rydén T., Teräsvirta T., Åsbrink S. Stylized facts of daily return series and the hidden Markov model. *Journal of applied econometrics* 1998:217-244 %@ 0883-7252.
- Schmidt S.C.E., Woll A. Longitudinal drop-out and weighting against its bias. *BMC medical research methodology* 2017;17:164-164.
- Schoonjans B., Van Cauwenberge P., Vander Bauwhede H. Formal business networking and SME growth. *Small Business Economics* 2013;41:169-181.
- Segarra A., Callejón M. New firms' survival and market turbulence: New evidence from Spain. *Review of industrial Organization* 2002;20:1-14.
- Shane S. *A General Theory of Entrepreneurship: The Individual-Opportunity Nexus*. Cheltenham, UK: Edward Elgar; 2003.
- Shane S. Introduction to the focused issue on the biological basis of business. *Organizational Behavior and Human Decision Processes* 2009;110:67-69.
- Shao L. Dynamic study of corporate governance structure and firm performance in China. *Chinese Management Studies* 2019;13.
- Shepherd D.A., Patzelt H. A Call for Research on the Scaling of Organizations and the Scaling of Social Impact. *Entrepreneurship Theory and Practice* 2022;46:255-268.
- Shepherd D.A., Souitaris V., Gruber M. Creating New Ventures: A review and research agenda. *Journal of Management* 2021;47:0149206319900537.
- Shrader R., Siegel D.S. Assessing the relationship between human capital and firm performance: Evidence from technology-based new ventures. *Entrepreneurship Theory and Practice* 2007;31:893-908.
- Sine W.D., Mitsuhashi H., Kirsch D.A. Revisiting Burns and Stalker: Formal Structure and New Venture Performance in Emerging Economic Sectors. *The Academy of Management Journal* 2006;49:121-132.
- Singh D., Delios A. Corporate governance, board networks and growth in domestic and international markets: Evidence from India. *Journal of World Business* 2017;52:615-627.
- Sleuwaegen L., Onkelinx J. International commitment, post-entry growth and survival of international new ventures. *Journal of Business Venturing* 2014;29:106-120.

- Snijders T.A., Bosker R.J. *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage; 2011.
- Soto-Simeone A., Sirén C., Antretter T. New venture survival: A review and extension. *International Journal of Management Reviews* 2020;22:378-407.
- Startz R., Wood-Doughty A. Improved Estimation of Peer Effects using Network Data. 2016.
- Sterk V., Sedláček P., Pugsley B. The nature of firm growth. *American Economic Review* 2021;111:547-579.
- Stevenson R.M., Kuratko D.F., Eutsler J. Unleashing main street entrepreneurship: Crowdfunding, venture capital, and the democratization of new venture investments. *Small Business Economics* 2019;52:375-393.
- Symeonidou N., Nicolaou N. Resource orchestration in start-ups: Synchronizing human capital investment, leveraging strategy, and founder start-up experience. *Strategic Entrepreneurship Journal* 2018;12:194-218.
- Tang J., Tang Z. The relationship of achievement motivation and risk-taking propensity to new venture performance: a test of the moderating effect of entrepreneurial munificence. *International Journal of Entrepreneurship and Small Business* 2007;4:450-472.
- Tatikonda M.V., Terjesen S.A., Patel P.C., Parida V. The role of operational capabilities in enhancing new venture survival: A longitudinal study. *Production and Operations Management* 2013;22:1401-1415.
- Tian L., Yang J.Y., Wei L. Speed to legal registration and nascent venture performance: A temporal dilemma for nascent entrepreneurs in an emerging economy. *Journal of small business management* 2019;57:476-495.
- Tranfield D., Denyer D., Smart P. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management* 2003;14:207-222.
- Tzabbar D., Margolis J. Beyond the startup stage: The founding team's human capital, new venture's stage of life, founder-CEO duality, and breakthrough innovation. *Organization Science* 2017;28:857-872.
- van Witteloostuijn A., Kolkman D. Is firm growth random? A machine learning perspective. *Journal of Business Venturing Insights* 2019;11:e00107.
- Vanacker T., Manigart S., Meuleman M., Sels L. A longitudinal study on the relationship between financial bootstrapping and new venture growth. *Entrepreneurship & Regional Development* 2011;23:681-705.
- Waheed A., Malik Q.A. Board characteristics, ownership concentration and firms' performance. *South Asian Journal of Business Studies* 2019;8.
- Wang G., Li L., Jiang X. Entrepreneurial business ties and new venture growth: The mediating role of resource acquiring, bundling and leveraging. *Sustainability* 2019a;11:244.
- Wang L., Tan J., Li W. The impacts of spatial positioning on regional new venture creation and firm mortality over the industry life cycle. *Journal of Business Research* 2018;86:41-52.
- Wang X., Lin Y., Shi Y. The moderating role of organizational environments on the relationship between inventory leanness and venture survival in Chinese manufacturing: IMS. *Journal of Manufacturing Technology Management* 2019b;30:413-440.
- Weinzimmer L.G., Nystron P.C., Freeman S.J. Measuring organizational growth: Issues, consequences and guidelines. *Journal of Management* 1998;24:235-262.
- Welbourne T.M., Neck H.M., Meyer G. The entrepreneurial growth ceiling: Using people and innovation to mitigate risk and break through the growth ceiling in initial public offerings. *Management Decision* 2012;50:778-796.
- Wennberg K., Delmar F., McKelvie A. Variable risk preferences in new firm growth and survival. *Journal of Business Venturing* 2016;31:408-427.

- Wennberg K., DeTienne D.R. What do we really mean when we talk about 'exit'? A critical review of research on entrepreneurial exit. *International Small Business Journal* 2014;32:4-16.
- Wennberg K., Wiklund J., DeTienne D.R., Cardon M.S. Reconceptualizing entrepreneurial exit: Divergent exit routes and their drivers. *Journal of Business Venturing* 2010;25:361-375.
- Whetten D.A. Organizational Growth and Decline Processes. *Annual review of sociology* 1987;13:335-358.
- Wiklund J., Baker T., Shepherd D. The age-effect of financial indicators as buffers against the liability of newness. *Journal of Business Venturing* 2010;25:423-437.
- Wiklund J., Davidsson P., Delmar F. What do they think and feel about growth? An expectancy-value approach to small business managers' attitudes toward growth. *Entrepreneurship Theory and Practice* 2003;27:247-270.
- Williams T.A., Shepherd D.A. Mixed method social network analysis: Combining inductive concept development, content analysis, and secondary data for quantitative analysis. *Organizational Research Methods* 2017;20:268-298.
- Wright M., Stigliani I. Entrepreneurship and growth. *International Small Business Journal* 2012.
- Xu R., DeShon R.P., Dishop C.R. Challenges and Opportunities in the Estimation of Dynamic Models. *Organizational Research Methods* 2019:1094428119842638.
- Yan J., Williams D.W. Timing is everything? Curvilinear effects of age at entry on new firm growth and survival and the moderating effect of IPO performance. *Journal of Business Venturing* 2021;36:106020.
- Yang T., Aldrich H.E. Out of sight but not out of mind: Why failure to account for left truncation biases research on failure rates. *Journal of Business Venturing* 2012;27:477-492.
- Yang T., Aldrich H.E. "The liability of newness" revisited: Theoretical restatement and empirical testing in emergent organizations. *Social Science Research* 2017;63:36-53.
- Zahra S.A., Sapienza H.J., Davidsson P. Entrepreneurship and Dynamic Capabilities: A Review, Model and Research Agenda*. *Journal of Management Studies* 2006;43:917-955.
- Zhou L., Wang M., Zhang Z. Intensive longitudinal data analyses with dynamic structural equation modeling. *Organizational Research Methods* 2021;24:219-250.

APPENDIX I. SYSTEMATIC REVIEW OF NEW FIRM GROWTH AND SURVIVAL

We follow the search protocols of Tranfield et al. (2003) for conducting systematic reviews in the field of management. We started broadly by searching for the keywords “new venture” and “new firm” in Thomson ISI Web of Knowledge. This resulted in 8,220 and 44,667 papers, respectively. We then limited our review to English articles published in the fields of business, management, and economics. This left us with a total of 3,715 papers for the keywords “new venture” and 23,506 papers for “new firm”. Within the 3,715 papers, we used eight other keywords to capture new firm growth and survival: “venture dynamics” (n=580), “venture growth” (n=976), “venture expansion” (n=125), “venture survival” (n=327), “venture contraction” (n=2), “venture termination” (n=7), “venture exit” (n=87), and “venture closure” (n=14). Within the 23,506 papers with “new firm”, we used eight keywords: “firm dynamics” (n=3,992), “firm growth” (n=200), “firm expansion” (n=651), “firm survival” (n=940), “firm contraction” (n=43), “firm termination” (n=52), “firm exit” (n=593), and “firm closure” (n=78). As the aim of this paper is to capture the dynamics of new ventures, we limit the retrieved studies to articles taking a longitudinal perspective. Therefore, within each search term, we used two further keywords “longitudinal”, and “panel”.

We screened the articles using title and abstract analysis to identify every paper that might be relevant to our topic—the papers whose dependent variable is one of the dynamics indicators (i.e., growth, expansion, survival, contraction, termination, and exit). This yielded 60 papers for “new venture”, and 46 papers for “new firm”. We excluded 12 papers for being qualitative, and/or having dependent variables not on venture dynamics, leaving 94 papers. Finally, we included some papers based on experts’ opinions to ensure that we did not miss any relevant articles—we shared the list of the 94 retrieved articles with two experts in the field and asked them to identify any papers that our procedure had failed to identify. This additional step

yielded 5 more papers, three of which were review articles (i.e., Bjørnskov and Foss, 2016; Gilbert et al., 2006; Soto-Simeone et al., 2020). We limit our list to articles published from 2000 onwards, which removed 2 articles published in 1994 (Cooper et al., 1994) and 1997 (Barkema and Vermeulen, 1997). Accordingly, our systematic review retrieved a total of 96 empirical papers. The stop date of the search was May 15, 2021.

Table 1Aa

Articles included in the systematic review (sorted by year) and focused on growth

Paper	Data	Source of Data	Method of Analysis	Dependent Variable (DV)	DV Descriptives	Time Span
1. Kor (2003)	73 US firms	Constructed from Registration statements and Compustat	Random effects GLS regression	Annual rate of sales growth	Mean: 0.83 SD: 2.64	1990 – 1999
2. Delmar and Shane (2003)	223 Swedish firms	Statistics Sweden	Fixed effects and Heckman two-stage estimation	Product development (1-5 scale)	Mean: 4.43 SD: 1.01	1998 – 2000
3. Baum and Locke (2004)	229 US firms	North American architectural woodwork firms	Structural equation modeling	Compound annual sales-growth rate and annual employment-growth rate	Mean: 0.11 SD: 0.36	1993 - 1999
4. Hayton (2005)	237 US high-tech new ventures	Secondary data from IPO prospectuses (Securities and Exchange Commission [SEC], EDGAR Database), the US Patent and Trademark Office database, and company press releases	Hierarchical regression	Venture expansion: Number of acquisitions and number of joint ventures or strategic alliances	Mean: 0.17 SD: 0.29	1994 - 1998
5. Rothaermel and Hill (2005)	556 US incumbent firms	Standard & Poor's COMPUSTAT and DRI databases	Park's estimator and GLS	Venture performance: Incumbent firm ROE and ROA	Mean: 18.36 SD: 17.90 Mean: 7.61 SD: 5.03	1972 - 1997
6. Alsos et al. (2006)	327 Norwegian new firms (sole proprietorships, partnerships with mutual responsibility, and partnerships with shared responsibility, as well as unlisted limited liability companies)	Norwegian business register, the Norwegian Central Coordinating Register for Legal Entities	Hierarchical linear regression	Venture growth: Sales turnover after adding a constant of 10	Mean: 12.44 SD: 1.86	2002 – 2004

7. Shrader and Siegel (2007)	198 US high-tech firms	IPO prospectuses	Hierarchical linear regression	Venture growth: Sales growth	Not Found	1993 – 2003
8. Tang and Tang (2007)	227 firms with female nascent entrepreneurs	Panel Study of Entrepreneurial Dynamics	Hierarchical regression	Venture performance: Founder-reported performance measure (1-5 scale)	Mean: 3.88 SD: 0.56	1998 – 2003
9. Delmar and Wiklund (2008)	176 Swedish firms	Statistics Sweden’s data on incorporated companies in Sweden	Heckman two-stage estimation	Venture growth: Size in terms of employees (converted into full-time equivalents) and annual sales	Mean: 1.17 SD: 0.44 Mean: 1.25 SD: 0.48	1994 – 1999
10. Jennings et al. (2009)	307 Canadian surviving new firms	The British Columbia Business to Business Directory and the Canadian Law List	Pooled OLS, and random-effects GLS regression	Venture productivity: Annual revenue data and annual data on the numbers of lawyers and partners	Mean: 0.55 SD: 0.33 Mean: 1.05 SD: 0.93	1998 – 2003
11. Artz et al. (2010)	272 US firms	COMPUSTAT	Three-Stage least squares regression	Venture performance: ROA and sales growth	Mean: 7.04 SD: 12.61 Mean: 4.51 SD: 9.61	1986 – 2004
12. Coad (2010)	8,503 French firms	EAE databank collected by Service des Études et des Statistiques Industrielles (SESSI) and provided by the French Statistical Office	OLS and LAD regression	Venture growth: Growth rates calculated by taking the differences of the logarithms of size (employees)	Mean: 0.01 SD: 0.14	1996 – 2004
13. Edelman et al. (2010)	442 US firms	Panel Study of Entrepreneurial Dynamics	Structural equation modeling	Venture growth intentions: (1) I want the business to be as large as possible; and (2) I want a size I can manage myself or with a few key employees	Mean: 0.19 SD: 0.39	1998 – 2003
14. Khaire (2010)	137 US advertising agencies	Archival sources: The Standard Directory of Advertising Agencies, The Standard Directory of National Advertisers, and Ad \$ Summary	Fixed-Effects GLS regression	Venture growth: Annual revenues, number of employees, and number of new clients	Mean: 0.46 SD: 0.43 Mean: 28.99 SD: 41.61 Mean: 4.56 SD: 4.38	1997 – 1985
15. Holcomb et al. (2010)	308 US firms	Thomson Financial’s Securities Data Company (SDC) New Issues database	Random and fixed effects multilevel regression modelling	Venture growth: Employment growth and ROA	Not Found	1997 – 2001
16. Huynh and Petrunia (2010)	19,233 Canadian firms	T2LEAP Database	GMM regression	Venture growth: Growth in sales	Mean: 0.07 SD: 0.06	1985 – 1997

17. Vanacker et al. (2011)	214 Belgian Firms	Primary data and the BEL-FIRST database (Bureau van Dijk)	OLS and GEE regression	Venture growth: Value added defined as sales income less materials and services purchased	Mean: 5.48 SD: 0.72	2001 – 2003
18. Lopez-Garcia and Puente (2012)	5,089 Spanish firms	The Bank of Spain Firm Demography Database	Dynamic Probit regression	Venture growth: Employment growth	Not Found	1996 – 2003
19. Navarro et al. (2012)	1,252 Spanish firms	SABI database (System of Analysis of Iberian Balances)	Cluster analysis	Venture growth: Growth in number of employees and sales	Mean: 2.12 SD: 0.84 Mean: 4.30 SD: 2.54	2001– 2005
20. Welbourne et al. (2012)	366 US firms	COMPUSTAT, the Security Data Corporation database, and Going Public: The IPO Reporter	OLS regression	Venture short-, long-term performance: Earnings per share, sales per employee, and year-end stock price	Mean: 20.08 SD: 1.34 Mean: 171.38 SD: 220.38 Mean: 0.13 SD: 0.21	Not Found
21. Colombo et al. (2013)	536 Italian firms	Research on Entrepreneurship in Advanced Technologies (RITA) database	GMM regression	Venture growth: Employees	Nor reported	1994 – 2003
22. Hutzschenreuter and Horstkotte (2013)	91 German firms	HDAX index of the German stock exchange and annual reports	Fixed-effects regression	Venture growth: Firm’s growth rate by the number of its expansion steps in the second period over the number of its subsidiaries at the beginning of that period	Mean: -1.75 SD: 0.94	1985 – 2007
23. Mai and Zheng (2013b)	615 US firms	Panel Studies of Entrepreneurial Dynamics	OLS and logit regression	Venture growth: Sales growth	Mean: 4.07 SD: 16.33	1998 – 2004
24. Schoonjans et al. (2013)	40,882 Belgian firms	Belfirst database of Bureau Van Dijk and the PLATO database provided by VOKA	Fixed-effects regression	Venture growth: Growth in employees, net asset, and added value	Mean: 0.06 SD: 0.01 Mean: 0.09 SD: 0.01 Mean: 0.04 SD: 0.01	1992 – 2008
25. Grilli and Murtinu (2014)	903 EU firms	VICO dataset	OLS, fixed effects, and GMM regression	Venture growth: Sales and number of employees growth	Varies according to different sub-samples	1994 – 2004
26. Koufteros et al. (2014)	137 Italian firms	Amadeus database	Random effect regression	Venture performance: ROTA, ROA, ROE, EBIT	Not Found	Varies

27. Naldi and Davidsson (2014)	138 Swedish firms	Statistics Sweden	Fractional logit regression	Venture growth: Sales growth	Not Found	2000 – 2006
28. Daunfeldt and Halvarsson (2015)	1,210 Swedish firms	IFDB database	Quantile regression	Venture growth persistence: Number that survived as high-growth firms into the next 3-year period using sales and employees	Varies	1997 – 2008
29. Millán et al. (2015)	3,374 EU firms	European Community Household Panel	Random effects logit regression	Venture growth: Hiring employees	Not Found	1994 – 2001
30. Frid et al. (2016)	1,025 US firms	Panel Study of Entrepreneurial Dynamics II	OLS and Cox survival regression	Venture growth: Revenues and hiring	Not Found	2005 – 2012
31. Colombo and Murtinu (2017)	759 European VC-backed and 7,611 non-VC-backed high-tech entrepreneurial firms	The VICO dataset	OLS and instrumental variables regression	Venture growth: Sales, real payroll expenses, and real fixed assets	Varies	1994 – 2010
32. Eberhart et al. (2017)	24,624 Japanese firms	COSMOS2 database from Teikoku Databank, Ltd. (TDB)	Difference in difference OLS	Venture growth: Sales growth	Varies according to reform implementation	1998 – 2007
33. Kolvereid and Isaksen (2017)	207 Norwegian firms	Primary Data	Hierarchical regression	Venture growth: Accumulated sales and employment growth	Not Found	2002 – 2010
34. Singh and Delios (2017)	2,152 Indian firms	Two databases: Prowess database of the Center for Monitoring the Indian Economy merged with Capex database	Random effects GLS regression	Venture growth: Total value of domestic and foreign investments	Mean: 0.25 SD: 1.16 Mean: 0.41 SD: 1.24	2002 – 2009
35. Tzabbar and Margolis (2017)	578 US biotech firms	BioScan and Knowledge Express	Parametric survival models and Huber-White sandwich estimator	Venture success in breakthrough innovation: Discrete probability of a breakthrough innovation occurring in the time interval	Varies according to venture growth stage	1964 – 1999
36. Acemoglu et al. (2018)	9,835 US firms	Longitudinal Business Database (LBD), the Census of Manufacturers (CMF), the NSF Survey of Industrial Research and Development (RAD), and the NBER Patent Database (PAT)	Simulated method of moments	Venture growth: Growth in sales and number of employees	Not Found	1987 – 1997

37. Bastié et al. (2018)	6,948 French firms	New Enterprises Information System (SINE) and the SINE survey	Two-stage regression	Venture performance: Sales turnover	Not Found	2003 – 2007
38. Calvino (2018)	6,561 Spanish firms	Spanish Technological Innovation Panel	GMM, OLS, and fixed-effects regression	Venture growth: Employment growth	Not Found	2004 – 2012
39. Gruenhagen et al. (2018)	688 Australian firms	Comprehensive Australian Study of Entrepreneurial Emergence (CAUSEE)	Logit regression	Venture growth (internationalization): Sales	Varies according to level of internationalization	Not Found (3 year-period)
40. Hopp and Greene (2018)	375 U.S. entrepreneurs	Panel Study of Entrepreneurial Dynamic	Logistic regression	Venture viability: Early-stage profitability	Mean: 0.22 SD: 0.41	2005 – 2011
41. Nadeem et al. (2018)	571 Australian firms	Bloomberg database	OLS and GMM regression	Venture performance: ROA and ROE	Mean: 10.45 Mean: 21.29 SD: Not Found	2005 – 2014
42. Rutherford et al. (2018)	1,128 US firms	Kauffman Firm Survey (KFS)	Logit regression	Venture performance: Sales and ROS	Mean: 0.10 SD: 0.30 Mean: 0.10 SD: 0.30	2004 – 2010
43. Symeonidou and Nicolaou (2018)	520 US firms	Kauffman Firm Survey (KFS)	Fixed effects and 2SLS regression	Venture performance: Constructed measure from revenues, size, and profitability	Mean: 0.32 SD: 1.11	2004 – 2010
44. Hernandez-Vivanco et al. (2019)	247 Portuguese firms	Amadeus Panel Dataset and IPAC Historical Checks	Two-step system-GMM regression	Venture performance: Return on sales, return on capital employed, and return on assets	Not Found	2004 – 2015
45. Kang et al. (2019)	1,151 Korean firms	PATSTAT database	Quantile regression	Venture growth: Sales growth	Not Found	1085 – 2014
46. Kim et al. (2019)	175 US firms	ITSTAR’s sales data, SDC Platinum data for mergers and acquisitions, and firm-level COMPUSTAT data	Heckman correction and two-stage instrumental variable (IV) regression	Venture growth: Number of employees and sales growth	Mean: 55,000 SD: 123,000 Mean: 15B SD: 29B	1990 – 2001
47. Papangkorn et al. (2019)	1,951 US firms	ISS database (Institutional Shareholder Services, formerly RiskMetrics) and COMPUSTAT	Fixed-effects regression	Venture performance: Tobin Q, ROA, BEP ratio, and EBITDA ratio	Mean: 2.05 SD: 1.27 Mean: 0.05 SD: 0.10 Mean: 0.10 SD: 0.09 Mean: 0.14 SD: 0.09	1997 – 2014
48. Shao (2019)	2,545 Chinese firms	Shanghai and Shenzhen stock exchanges	GMM regression	Venture Performance: Tobin’s Q	Mean: 2.71 SD: 2.08	2001 – 2015

49. Stevenson et al. (2019)	676 US firms	Two databases: PWC Moneytree database and US Kickstarter Data	OLS regression	VC growth: Number of state VC investments	Not Found	2006 – 2016
50. Tian et al. (2019)	145 Chinese firms	Chinese Panel Study of Entrepreneurial Dynamics	OLS and logit regression	Venture performance: Whether the venture covered its start-up costs from profitable operations, and number of employees	Not Found	2009 – 2011
51. Waheed and Malik (2019)	309 Pakistani non-financial firms	Directors' report, financial statements, ownership concentration data from shareholding report and the age of the firm is computed from the time the firm is registered at the SCCP	GMM regression	Venture performance: ROA, Tobin's Q, and ROE	Mean: 3.01 SD: 22.43 Mean: 1.32 SD: 1.41 Mean: 7.44 SD: 27.04	2005 – 2016
52. Wang et al. (2019a)	332 Chinese entrepreneurs	Primary data	Optimal scaling regression	Venture Growth: Self-report (scale 1 to 7) on growth in sales, market share, and employment	Not Found	2013 – 2016
53. Coad and Srhoj (2020)	45,465 Croatian firms, and 14,096 Slovenian firms	Census data of the Financial Agency (FINA) of the Republic of Croatia and the Agency of the Republic of Slovenia for Public Legal Records and Related Services (AJPES)	Logit least absolute shrinkage and selection operator regression	Venture growth: Employment and sales turnover	Varies according to database	2003 – 2016
54. Casillas et al. (2020)	106 Spanish firms	Survey of Business Strategies (SBS)	OLS, random, and fixed effects regression	Venture expansion (internationalization): Export volume, export intensity, and growth of exports	Mean: 12.61 SD: 2.86 Mean: 26.76 SD: 27.31 Mean: 75.13 SD: 1,232.94	1990 – 2013
55. Chen and Song (2020)	1,697 US firms	Kauffman Firm Survey (KFS)	GMM regression	Venture growth: Sales revenue	Mean: 0.03 SD: 1.01	2004 – 2011
56. Grillitsch and Schubert (2020)	434,247 Swedish firms	Firms and establishment dynamics database (FAD), business statistics database (FEK), business group register, integrated longitudinal database for health insurance and labor market studies (LISA)	OLS and fixed-effects regression	Venture growth: Sales growth	Not Found	1997 – 2012
57. Sterk et al. (2021)	US Firms (Census data)	Census Longitudinal Business Database	Reduced-form modeling	Venture growth: Employees	Not Found	1979 – 2012

Table 1Ab**Articles included in the systematic review (sorted by year) and focused on survival**

Paper	Data	Source of Data	Method of Analysis	Dependent Variable (DV)	DV Descriptives	Time Span
58. Segarra and Callejón (2002)	12,885 Spanish firms	DIRCE (Directorio Central de Empresas) compiled with the Instituto Nacional de Estadística (INE)	Cox-proportional hazard regression	Venture exit	Not Found	1994 – 1998
59. Chrisman and McMullan (2004)	141 US firms	Longitudinal study of one outsider assistance program	Logit regression	Venture survival: Year the venture started	Mean:0.81 SD: 0.40	1994 – 2001
60. Delmar and Shane (2004)	223 Swedish firms	Primary data	Weibull regression	Venture survival: Venture disbanding	Not Found	1998 – 2000
61. Honig and Karlsson (2004)	396 Swedish nascent entrepreneurs	Primary data	Logit regression	Venture survival: Continuance of a project throughout the time horizon of the study Venture performance: Profitability during the time horizon of the study	Mean: -0.38 SD: -0.45 Mean: 0.57 SD: 0.50	1998 – 2000
62. Aspelund et al. (2005)	80 independent Norwegian (65) and Swedish (15) new technology-based firms	Database of Scandinavian technology-based start-ups	Cox regression	Venture survival: Year of establishment, and the number of years of survival.	Mean: 1997.9 SD: 2.19	1995 – 2002
63. Delmar and Shane (2006)	223 Swedish firms	Statistics Sweden	Event history analysis	Venture survival: Disbanding Venture growth: Average amount of monthly sales	Not Found	Not Found (30-month period)
64. Liao and Gartner (2006)	276 US firms	Panel Study of Entrepreneurial Dynamics	Logit regression	Venture survival: Status of startup efforts: “operating business”, “active startup”, “inactive startup”, and “no longer worked on”	Mean: 0.56 SD: 0.50	1998 – 2003
65. Brush et al. (2008)	237 US firms	Panel Study of Entrepreneurial Dynamics	Cox-proportional hazard regression	Venture survival: (1) Continue organizing and (2) failure time	Mean: 0.65 SD: 0.46 Mean: 100.65	1998 – 2003

					SD: 86.92	
66. Korunka et al. (2010)	354 Austrian firms	Vienna Entrepreneurship Studies	Logit regression	Venture survival: Self-reported, including a variety of business development patterns, ranging from survival despite a considerable financial loss to business expansion and strong business growth	Mean: 4.56 SD: 4.38	1998 – 2005
67. Wennberg et al. (2010)	1,735 Swedish firms	Statistics Sweden datasets: RAMS and LOUISE	Multinomial logit regression	Venture exit: (0) continuation, (1) harvest sale, (2) distress sale, (3) harvest liquidation, and (4) distress liquidation	Mean: 0.34 SD: 0.48 Mean: 0.08 SD: 0.27 Mean: 0.06 SD: 0.24 Mean: 0.26 SD: 0.44 Mean: 0.26 SD: 0.43	1995 – 2002
68. Wiklund et al. (2010)	37,782 Swedish firms	Annual reports of ventures registered in the Swedish patent office	Event history analysis and logit regression	Venture failure: (1) Completed bankruptcies, with or without a surplus, (2) completed liquidations, and (3) closures on the basis of companies' own request	Not Found	Varies (7 years of existence)
69. Coeurderoy et al. (2012)	196 UK and German firms	Dun & Bradstreet in the UK and Creditreform in Germany	Probit regression	Venture survival: A dummy variable for survival	Mean: 0.84 SD: 0.37	1997 – 2003
70. Robb and Watson (2012)	4,016 US firms	Kauffman Firm Survey (KFS)	Cox-proportional hazard regression	Venture closure: Business closure rate	Not Found	2004 – 2008
71. Yang and Aldrich (2012)	1,100 US firms	Panel studies of Entrepreneurial Dynamics II	Cox-proportional hazard regression	Venture survival: Failure rate	Not Found	2005 – 2011
72. Basu and Wadhwa (2013)	477 US firms	COMPUSTAT	Binomial regression	Venture strategic discontinuous: Strategic renewal	Mean: 0.66 SD: 1.47	1990 – 2000
73. Tatikonda et al. (2013)	812 new manufacturing ventures	Affarsdata database	Proportional hazard regression	Venture survival: Survival over 6 years	Mean: 0.36 SD: Not Found	2005 – 2010
74. Deng et al. (2017)	133,626 Chinese firms	National Bureau of Statistics of China	Cox regression	Venture exit (international): Exit from exporting	Not Found	1998 – 2008

75. Fariborzi and Keyhani (2018)	2,330 US firms	Kauffman Firm Survey (KFS)	GEE and cox-proportional hazard regression	Venture survival (internationalization): Business closure	Not Found	2004 – 2012
76. Messersmith et al. (2018)	1,012 US firms	Kauffman Firm Survey (KFS)	Cox regression	Venture survival: Business failure	Mean: 0.16 SD: 0.36	2004 – 2010
77. Wang et al. (2018)	629 Canadian firms	Scott's Corporate Director and Financial Post Markets: Canadian Demographics	Binomial regression	Venture exit: Firm mortality (the number of telecom equipment manufacturing establishments in a municipality that exited the industry or were acquired by another company in each year)	Mean: 0.46 SD: 1.04	1995 -2005
78. Battisti et al. (2019)	245 New Zealand firms	Existing longitudinal survey that tracked the performance of small firms in New Zealand annually over five years	Fuzzy cluster analysis	Venture survival: Resilient firm performance (1-5 scale)	Not Found	2007 – 2011
79. Cefis and Marsili (2019)	2,329 Dutch firms	Netherlands Central Statistics Office (CBS)	Kaplan-Meier estimator	Venture exit: Hazard rate of exit	Not Found	2001 – 2015
80. Ebert et al. (2019)	6,776 German firms	IAB/ZEW Start-up Panel of the Centre for European Economic Research (ZEW) and Creditreform database	Semiparametric Cox regression	Venture survival: Time that elapsed between the entry and the exit of the company	Not Found	2007 – 2011
81. Gjerløv-Juel and Guenther (2019)	95,837 Danish firms	Danish panel dataset: the Integrated Database for Labor Market Research	Logit and GEE regression	Venture exit: Firm exit as two successive years of no activity	Mean: 3.33 SD: 3.04	1994 – 2007
82. Wang et al. (2019b)	283,083 Chinese firms	Chinese Annual Survey of Industrial Firms (CASIF) database	OLS, fixed effects log-normal regression survival analysis	Venture survival: Duration as the number of years from registration to closure	Mean: 5.69 SD: 1.93	1995 – 2005

Table 1Ac

Articles included in the systematic review (sorted by year) and focused on both survival and growth

Paper	Data	Source of Data	Method of Analysis	Dependent Variable (DV)	DV Descriptives	Time Span
83. Delmar et al. (2003)	11,748 Swedish firms	Several registers in Sweden (i.e., the official “Bureau of Census”)	Cluster analysis	Not applicable	Not Found	1987 – 1996
84. Musso and Schiavo (2008)	16,500 French firms	EAE survey and the DIANE database	Cox-proportional hazard regression	Venture survival and growth: Probability of failure and total factor productivity	Not found	1996 – 2004
85. Bertoni et al. (2011)	538 Italian new technology-based firms	2004 RITA directory	OLS, WG, and GMM regression	Venture survival and growth: Sales and number of employees	Mean: 12.98 SD: 1.77 Mean: 2.08 SD: 1.12	1994 – 2003
86. Delmar et al. (2011)	31,602 Swedish firm	Statistics Sweden	Pooled OLS, Cox, quantile, random effects panel regression	Venture survival and growth: Value added growth, number of years the firm is active, and firm growth in turnover	Mean: 0.03 SD:0.52 Mean: 0.92 SD: 4.45 Mean: 0.12 SD: 2.17	1997 – 2002
87. Bertoni et al. (2013)	531 Italian firms	Research on Entrepreneurship in Advanced Technologies (RITA) database	Quantile regression	Venture growth and survival: Sales, employees growth rates, and hazard rate of exit of sample firms in 2000–2003	Not Found	1994 – 2003
88. Delmar et al. (2013)	23,382 Swedish firms	Statistics Sweden: RAMS	Pooled OLS, fixed effects, Cox’s semi-parametric survival regression	Venture survival and growth: Venture exit (hazard ration) and sales	Mean: 0.97 SD: 0.19 Mean: 0.85 SD: 0.67	1995 – 2002
89. Pe'er and Keil (2013)	46,879 Canadian firms	T2-LEAP database	Piecewise exponential model	Venture survival: Failure to continue	Mean: 0.46 SD: 0.30	1984 – 1998
90. Sleuwaegen and Onkelinx (2014)	5,800 Belgian firms	Constructed dataset with the help of the National Bank of Belgium	OLS regression and Heckman two-stage estimation	Venture survival and growth (Internationalization): Exports growth and exit rate per industry	Mean: 0.05 SD: 0.29 Exit rate: Not Found	1998 – 2005

91. Nielsen (2015)	1,151 Danish firms	Register data from IDA (Integrated Database for Labour Market Research) combined with a questionnaire survey	OLS, ordered logit regression	Venture survival and growth: Business closure and number of employees	Not Found	1999 – 2004
92. Löfsten (2016)	131 Swedish small high-technology firms	Primary data	Factor and correlation analyses	Venture survival and growth: Binary, sales, and employment	Varies	2005 – 2014
93. Wennberg et al. (2016)	14,760 Swedish firms	Statistics Sweden	Cox, random, and fixed effects regression	Venture growth and exit: Number of employees growth, and cessation of the venture	Mean: 0.80 SD: 0.80 Mean: 0.91 SD: 0.61	1995 – 2002
94. Aksaray and Thompson (2018)	2,677 US youth born between 1957 and 1964	National Longitudinal Survey of Youth 1979 (NLSY79)	Random effect logit and linear regression	Venture growth and exit: Self-employment earnings growth rate and exit from self-employment	Mean: 4,955 SD: 0.52	1957 – 2012
95. Haeussler et al. (2019)	1,498 German firms	Two databases: Mannheim Enterprise Panel from Creditre form database merged with the PATSTAT database	Cox, IV probit, and fixed effects regression	Venture survival and growth: Business failure and turnover growth	Mean: 0.21 SD: Not Found Mean: 48.84 SD: 81.68	1998 – 2007
96. Maliranta and Nurmi (2019)	27,766 Finnish firms	Finnish Longitudinal OWNer Employer-Employee (FLOWN) database	Logit regression	Venture survival and growth: Employment	Varies	2009 – 2013

* Largest sample size reported in case sample sizes are different within the different estimated models.

APPENDIX II. ROBUSTNESS MODELS

Figure A.2

The circles indicate the used data in our robust selection scheme. The stars represent unused data.

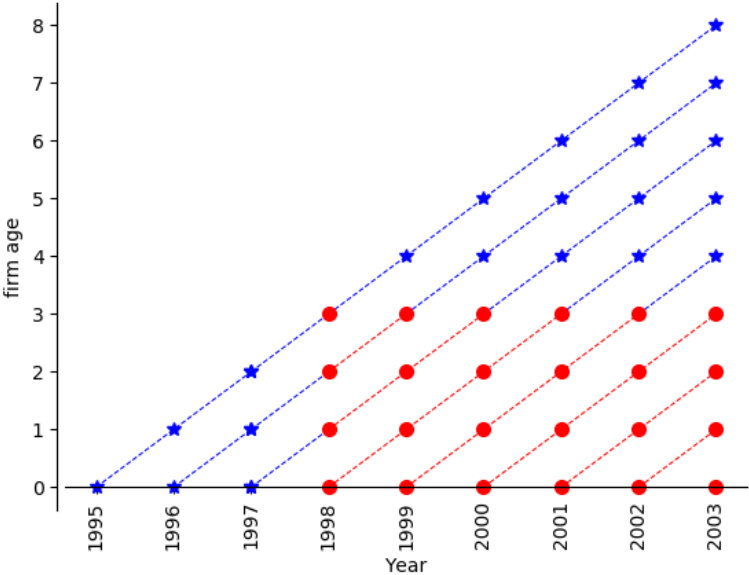


Table 2A

Probabilities of test data, micro and small firms

Prob / model	Empirical	EMM	LN
Micro firms (1-9 employees)			
P(<i>Hiring</i>)	0.18	0.19 (0.002)	0.25(0.003)
P(<i>Stagnant</i>)	0.71	0.71 (0.003)	0.64 (0.003)
P(<i>Firing</i>)	0.11	0.10 (0.002)	0.11 (0.001)
Small firms (10-25 employees)			
P(<i>Hiring</i>)	0.54	0.54 (0.013)	0.49(0.008)
P(<i>Stagnant</i>)	0.13	0.11 (0.009)	0.08 (0.004)
P(<i>Firing</i>)	0.33	0.34 (0.012)	0.43 (0.010)

Table 3A

Estimated coefficients log odds for the probability of hiring and exit

Covariates	β	s.e.	p-value	β	s.e.	p-value
		Micro firms (1-9 employees)			Small firms (10-49 employees)	
		Exit			Exit	
<i>Intercept</i>	-1.07	0.07	< 10 ⁻¹⁵	-0.77	0.37	0.04
<i>log(roa)</i>	-2.33	0.21	< 10 ⁻¹⁶	-4.11	1.51	0.005
<i>age</i>	-0.14	0.02	< 10 ⁻¹¹	-0.24	0.10	0.02
<i>asspI(assp > 0)</i>	0.02	0.003	< 10 ⁻⁰⁸	0.05	0.02	0.87
<i>asspI(assp < 0)</i>	-0.03	0.005	< 10 ⁻¹¹	0.01	0.06	0.37
		Micro firms (1-9 employees)			Small firms (10-49 employees)	
		Above			Above	
<i>Intercept</i>	0.77	0.16	< 10 ⁻⁰⁵	0.61	0.32	0.06
<i>log(roa)</i>	3.00	0.45	< 10 ⁻¹¹	3.50	1.01	0.0006
<i>age</i>	-0.22	0.05	< 10 ⁻⁰⁷	-0.11	0.08	0.19
<i>asspI(assp > 0)</i>	-0.19	0.01	< 10 ⁻¹⁵	-0.07	0.02	0.0002
<i>asspI(assp < 0)</i>	-0.18	0.02	< 10 ⁻¹⁵	-0.22	0.08	0.004

Note: Observations for micro firms n=17,508; small firms n=724.

Table 4A

Estimated coefficients for the magnitude of expansion and contraction, ROBUST

Covariates	β	s.e.	p-value	β	s.e.	p-value
		Micro firms (1-9 employees)			Small firms (10-49 employees)	
		Firing			Firing	
Intercept	-2.76	0.27	< 10 ⁻¹⁵	-2.03	0.34	< 10 ⁻⁷
log(roa)	-4.82	2.11	0.02	-1.58	1.16	0.17
age	0.07	0.16	0.96	-0.09	0.09	0.28
asspl(assp > 0)	0.17	0.04	0.0001	0.11	0.03	0.0001
asspl(assp < 0)	-0.03	0.09	0.74	-0.05	0.08	0.53
		Micro firms (1-9 employees)			Small firms (10-49 employees)	
		Hiring			Hiring	
Intercept	-0.22	0.11	0.03	-1.72	0.23	< 10 ⁻¹²
log(roa)	-2.64	0.33	< 10 ⁻¹⁴	-0.32	0.63	0.62
age	-0.32	0.03	< 10 ⁻¹⁶	-0.04	0.06	0.53
asspl(assp > 0)	0.07	0.01	< 10 ⁻¹³	0.003	0.02	0.86
asspl(assp < 0)	-0.06	0.01	< 10 ⁻¹⁵	-0.05	0.02	0.06

Note: Observations for micro firms n=17,508; small firms n=724.

Table 5A

Estimated coefficients for the log normal, ROBUST

Covariates	β	s.e.	p-value	β	s.e.	p-value
		Micro firms (1-9 employees) Firing			Small firms (10-49 employees) Firing	
Intercept	0.15	0.01	$< 10^{-15}$	0.05	0.34	0.21
log(roa)	0.15	0.03	$< 10^{-09}$	0.32	1.16	0.01
age	-0.04	0.00	$< 10^{-15}$	-0.05	0.09	0.67
asspI(assp > 0)	-0.01	0.00	$< 10^{-15}$	-0.02	0.03	$< 10^{-12}$
asspI(assp < 0)	-0.02	0.00	$< 10^{-15}$	-0.01	0.02	0.01

APPENDIX III. EMM for Firm Sales

An alternative and popular measure of growth is firm sales or revenues. Sales growth is a dependent variable in 41% of the papers examined in our review. Here we adapt our EMM to sales data instead of employee data to show the versatility of our model. The dynamics of an increase or a decrease in sales (compared to previous years) are not substantially different from increases or decreases in employees (hiring and firing). The researcher can apply the same framework introduced here, but with the following modifications.

First, the probability of equal sales compared to the previous year is not needed here, unless the researcher assumes that the firms are in an active or static stage that is modeled by a hidden Markov model (Cappé, Moulines, and Ryden, 2005), as discussed in the main text. In the current version of the EMM, the probability model is irrelevant because sales fluctuate more than employees and are less likely to be the same as in the previous year. Second, and more importantly, sales is not integer value but a continuous value. Hence, we need to change the density distribution of the magnitude of contraction and expansion in sales. Preliminary analysis shows that the magnitude for expansion (even after removing extreme outliers) is very heavy tailed and necessitates a logarithmic transformation. For this transformed variable (the logarithmic value of expansion, or increase, in sales) we use a Gamma distribution. For contracting, or decreasing, sales we choose to model the ratio between sales and previous sales as a Beta distribution.

We compare the fit in distribution below in Figure A2. We see that from a distributional perspective the normal distribution is a poor fit (which does not necessarily mean that coefficients are wrong). We can also see that the small firms (defined as having between 1–4 million euro in annual sales) has a stronger left tail compared to the micro firms (having less than 1 million euro in annual sales). Remember that the coefficients for expansion are based on logarithmic transformed data.

Table 6A is a replication of Table 2, but for sales changes instead of employee changes. Note that models for expansion and contraction are based on data with different scales (logarithmic and ratio, respectively). Hence, one cannot compare the values of the coefficients across the rows in Table 6A.

Figure 3A still shows a significant improvement in fit when using the EMM instead the log-normal model. Table 6A gives some interesting new insights. The normal model would suggest that higher profitability (LogROA) leads to higher sales. However, we find that higher profitability leads to an increased probability of expanding sales but does not predict its magnitude. Hence, the EMM is equally valuable for sales as for employment growth.

Table 6A

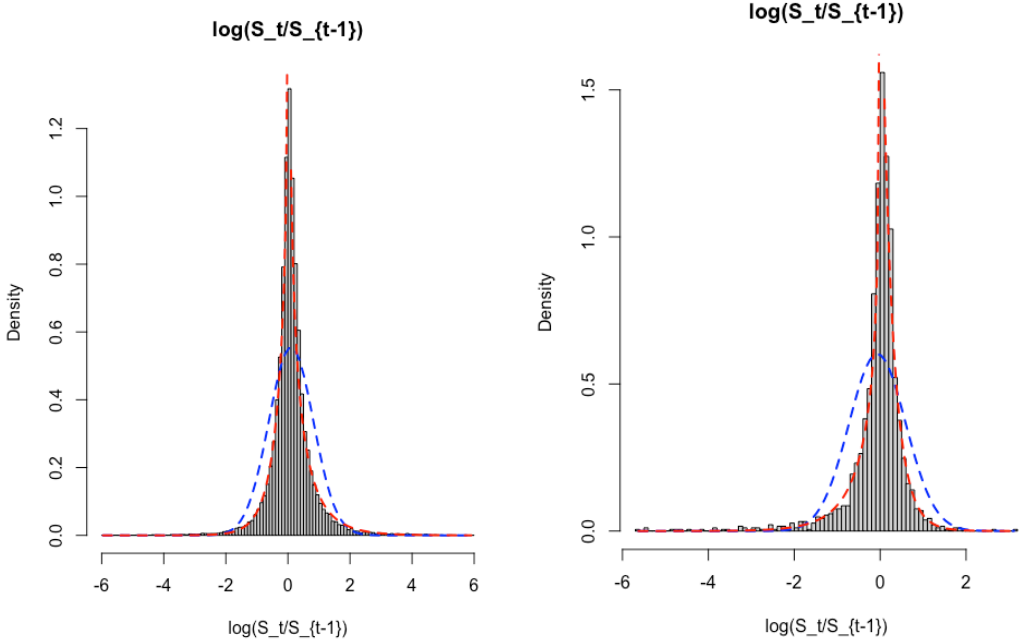
Estimated coefficients for our two-stage model with an OLS regression. EMM is composed of stage 1, predicting the probability of the event of increased sales, and stage 2, predicting the magnitude of increased and decreased sales.

Covariates	Model 1: EMM Stage 1 Prob. (Increased sales)			Model 2: EMM Stage 2 Magn. (Decreasing sales)			Model 3: EMM Stage 2 Magn. (Increasing Sales)			Model 4: OLS Model (Growth)		
	β	(s.e.)	p-value	β	(s.e.)	p-value	β	(s.e.)	p-value	β	(s.e.)	p-value
Micro firms (sales < 1 million euro)												
<i>Intercept</i>	0.45	0.028	< 10 ⁻¹⁵	0.76	0.021	< 10 ⁻¹⁵	-0.44	0.023	< 10 ⁻¹⁵	0.14	0.008	< 10 ⁻¹⁵
<i>log(roa)</i>	3.50	0.133	< 10 ⁻¹⁵	1.76	0.104	< 10 ⁻¹⁵	-0.09	0.093	0.36	0.85	0.037	< 10 ⁻¹⁵
<i>age</i>	-0.04	0.005	< 10 ⁻¹⁴	0.02	0.003	< 10 ⁻¹⁰	-0.09	0.004	< 10 ⁻¹⁵	-0.03	0.001	< 10 ⁻¹⁵
<i>Sales previous (million SEK)/ in Stage 1</i>	-0.09	0.007	< 10 ⁻¹⁵	0.03	0.004	< 10 ⁻¹¹	0.01	0.002	< 10 ⁻⁷	-0.02	0.001	< 10 ⁻¹⁵
<i>Prev Empl in Stage 2</i>												
Small firms (sales 1-4 million euro)												
<i>Intercept</i>	-0.18	0.163	0.27	0.36	0.095	0.0002	-1.16	0.008	< 10 ⁻¹⁵	-0.22	0.036	< 10 ⁻⁹
<i>log(roa)</i>	3.47	0.570	< 10 ⁻⁹	1.86	0.456	< 10 ⁻⁴	-0.26	0.319	0.42	0.81	0.15	< 10 ⁻⁷
<i>age</i>	0.01	0.014	0.88	0.02	0.011	0.06	-0.04	0.009	0.0002	0.006	0.004	0.89
<i>Sales previous (million SEK)/ Prev Empl in Stage 2</i>	0.01	0.006	0.20	0.03	0.003	< 10 ⁻¹⁵	0.05	0.002	0.0002	0.006	0.001	< 10 ⁻¹⁵

Note: Observations for micro firms n= 36375; small firms n= 1868. OLS: For small firms: Adjusted R-squared: 0.04, Residual standard error: 0.71. For micro firms: Adjusted R-squared: 0.05; Residual standard error: 0.65.

Figure 3A

The logarithmic ratio represents the sales in the current year divided by the sales in the previous year for micro firms (figure to the right) and small firms (figure to the left), given that they do not exit during the current year.



APPENDIX IV. Syntax for the R Package

This is a small tutorial on how to install our (A) R-package, (B) simulate data, and then run the (C) EMM for employment and (D) sales, respectively model firm growth using the model proposed in the paper.

A install the R-package

- * Download and install [R](<https://cran.r-project.org/>)
- * Download and install [R-studio](<https://www.rstudio.com/>) (not needed but recommended).
- * Download and install [r-tools](<https://cran.r-project.org/bin/windows/Rtools/>).
- * Download our developed R-package which is found in the supplementary files, and in R run:

```
``r
install.packages("path_to_file", repos = NULL, type="source")
``
```

B simulate data employment

In the `cams` package we can generate simulated data created to be similar to the data for which we produced the files.

This can be simulated in R by: (Here we simulate medium sized companies for large companies set size to `"small"")

```
``r
library(cams)
n <- 10000
size = "medium"
if(size== "small"){
  trunc_emp_prev=Inf
}else{
  name='medium'
  trunc_emp_prev = 15
}
data <- generate.data(n,size)
``
```

C Estimate EMM for employment

Now that we have simulated the data in `data` we first created a dummy variable for truncated employment:

```
``r
data$emp_prev_trunc <- data$emp_prev
data$emp_prev_trunc[data$emp_prev_trunc > trunc_emp_prev] = trunc_emp_prev
data$emp_prev_trunc <- factor(data$emp_prev_trunc)
``
```

Then we setup the actual EMM object:

```
``r
cam_obj <- growthObjInit(data, "emp", "emp_prev", exitIndex= "exit")
dist_list <- list(Above = " y ~ 1 + ln_roa + age + aspp_hist_over +
  aspp_hist_below + offset(log(emp_prev))",
  Below = " y ~ 1 + ln_roa + age + aspp_hist_over +
  aspp_hist_below + offset(log(emp_prev))",

```

```

        ProbEqual = " y ~ 1 + ln_roa + assp_hist_over +
assp_hist_below + emp_prev_trunc ",
        ProbAbove = " y ~ 1 + ln_roa + assp_hist_over +
assp_hist_below ",
        ProbExit  = " y ~ 1 + ln_roa + assp_hist_over +
assp_hist_below + emp_prev_trunc "
cam_obj <- growthEstimate(cam_obj, formula = dist_list)
'''

```

To fit a similar log normal object we can use the built in `lm`:

```

'''r
logNorm_obj <- lm("log(emp) ~ -1 + assp_hist_over + assp_hist_below +
emp_prev_trunc + offset(log(emp_prev))", data = data)
'''

```

D Code for analyzing sales on non-available data

In this part of the code we setup the data for the analysis of sales data.

```

'''r
library(logspline)
library(cams)
library(plotrix)

set.seed(1)

pred_set = 0.3

small=1

#data not aviable
reg_dat <- read.table("../../data/reg_dat.dat")
# remove NA data
reg_dat <- reg_dat[complete.cases(reg_dat),]

# remove negative sales are assumed zero, change scale
reg_dat$sales[reg_dat$sales <0 ] = 0
reg_dat$sales_prev[reg_dat$sales_prev <0 ] = 0
reg_dat$sales <- reg_dat$sales/10^6
reg_dat$sales_prev <- reg_dat$sales_prev/10^6

# large or small previous magnitude of sales
if(small){
  ind <- reg_dat$sales_prev <= 10
}else{
  ind <- (reg_dat$sales_prev > 10) &(reg_dat$sales_prev <= 40)
}
reg_dat <- reg_dat[ind, ]

#remove outliers, these are clear errors of the data
ratio <- abs(reg_dat$sales[reg_dat$sales_prev >0
]/reg_dat$sales_pre[reg_dat$sales_prev >0 ])
ratio_indx <- reg_dat$sales_prev >0 &
abs(log(reg_dat$sales/reg_dat$sales_prev)) < 6
reg_dat <- reg_dat[ratio_indx,]

#change scale for assp
reg_dat$assp_soc_prev <- reg_dat$assp_soc_prev/100000

```

```

reg_dat$assp_hist      <- reg_dat$assp_hist/100000
reg_dat$assp_soc      <- reg_dat$assp_soc/100000

#seperate positive and negative assp
reg_dat$assp_hist_over <- reg_dat$assp_hist
reg_dat$assp_hist_over[reg_dat$assp_hist < 0] <- 0
reg_dat$assp_hist_below <- reg_dat$assp_hist
reg_dat$assp_hist_below[reg_dat$assp_hist > 0] <- 0

```

#should not have negative previous sales either, i.e. don't allow dormant companies.

```

index <- reg_dat$sales_prev>0 & reg_dat$sales>0
reg_dat <- reg_dat[index,]
...

```

Then we run similar analysis to employee data (including log normal file).

Note that the `dist_list` are the regression models for:

- * Positive sales magnitude (`Above`),
- * Negative sales magnitude (`Below`),
- * Probability of equal sales (`ProbEqual`) (not very relevant for sales)
- * Probability of positive sales (`ProbAbove`)
- * Probability of company leaving the market (`ProbExit`)

```

```r
cam_obj <- growthObjInit(reg_dat, "sales", "sales_prev", exitIndex=
"exit_term")
dist_list <- list(Above = " log(sales/sales_prev) ~ 1 + ln_roa + age +
emp_prev ",
 Below = " y ~ 1 + ln_roa + age + emp_prev ",
 ProbEqual = " y ~ 1 + ln_roa + emp_prev ",
 ProbAbove = " y ~ 1 + ln_roa + age +sales_prev ",
 ProbExit = " y ~ 1 + ln_roa + emp_prev + sales_prev
")
#fit everything
cam_obj <- growthEstimate(cam_obj,
 formula = dist_list,
 dist = list(Above = "gamma", Below =
"betamc"),
 difference=F)

logNorm_obj <- lm("log(sales/sales_prev) ~ 1 + ln_roa + age + emp_prev",
data = reg_dat)
...

```

Then we plot the log ratios of the sales and previous sales, and compare the density of the log-normal and EMM model.

```

```r
##
# Empirical data
##
Y_ratio <- log(reg_dat$sales/reg_dat$sales_prev)
seq_ratio <- seq(min(Y_ratio), max(Y_ratio), length.out=100)
hist(Y_ratio,
     main="log(S_t/S_{t-1})",
     xlab="log(S_t/S_{t-1})",
     prob=T,
     breaks=100)

```



```

##
# log normal model
##
l_mean <- predict(logNorm_obj, newdata = reg_dat)
l_sigma <- summary(logNorm_obj)$sigma
pX = rep(0, length(seq_ratio))
for(i in 1:length(seq_ratio))
  pX[i] <- mean(dnorm(seq_ratio[i], mean =l_mean, sd = l_sigma))
lines(seq_ratio, pX, col='blue',lwd=2,lty=2)

##
# EMM model
#
##
ProbA <- predict(cam_obj$estimate$ProbAbove, newdata = reg_dat, type =
'response')
ProbE <- predict(cam_obj$estimate$ProbExit, newdata = reg_dat, type =
'response')
growth_lag <- reg_dat[,cam_obj$growthIndexPrev]
n <- length(growth_lag)
Y <- matrix(NA, nrow = n, ncol = 2)
if(is.null(cam_obj$exitIndex) ==FALSE){
  U_EXIT <- runif(n)
  Y[,1] <- U_EXIT < ProbE
}else{
  Y[, 1] <- 0
}

above <- runif(n)< ProbA
above = above * ( Y[,1]==FALSE)
alpha <- 1/summary(cam_obj$estimate$Above$model.fit
)$dispersion
beta <- alpha/exp(predict(cam_obj$estimate$Above$model.fit, newdata =
reg_dat))
Above <- rgamma(n=length(beta), shape = beta, scale= alpha)
seq_ratio_pos <- seq(0+0.1, max(Y_ratio), length.out=100)
pX = rep(0, length(seq_ratio_pos))
for(i in 1:length(seq_ratio_pos)){
  pX[i] <- mean(ProbA*dgamma(seq_ratio_pos[i], scale =1/beta,
shape=alpha))
}
lines(seq_ratio_pos, pX, col='red',lwd=2,lty=2)

...

### Code for used for the simulation:
The code is available in the package but we put it here for transparency
```r

logit <- function(x){1/(1+exp(-x))}
#
simulate company data that mimics the article

generate.data <- function(n, size="small"){
 if(size=="small"){

```

```

emp_prev <- rgeom(n, 1/2.06)+1

emp_prev <- rgeom(n, 1/2.06)
while(sum(emp_prev>8)>0){
 index = emp_prev>8
 emp_prev[index] = rgeom(sum(index),1/2.06)
}
emp_prev <- emp_prev+1

#regression on the other variables
#model.age = glm.nb(age~1 + emp_prev, data=reg_dat)
age = rnbinom(n, mu = exp(0.9397), size = 1.4462) +2
#lm_roa.model <- lm(ln_roa ~ emp_prev + age, dat= reg_dat)
ln_roa <- 0.1088506 + -0.0021723 * emp_prev + -0.0032622 * age +
rnorm(n, sd = 0.1037)
#assp.model <- lm(assp_hist ~ emp_prev + age + ln_roa , dat= reg_dat)
assp_hist <- -0.24157 + 0.52207 * emp_prev + -0.20382 * age + 7.9261 *
ln_roa + rnorm(n, sd = 6.837)
sales_prev <-exp(-0.498648 -0.022451 * assp_hist + 0.063615 *
emp_prev + 0.087288 * age + 0.919724 * ln_roa + rnorm(n, sd = 0.9459))
sales <-exp(0.0332283+ 0.7949240 * log(sales_prev) + 0.0581012 *
assp_hist + 0.0215592 * emp_prev -0.0100692 * age + 0.5632460 * ln_roa +
sqrt(rgamma(n, shape=1))*rnorm(n, sd = 0.5552))

reg_dat.simulate <- data.frame(emp_prev = emp_prev,
 age = age,
 ln_roa = ln_roa,
 assp_hist = assp_hist,
 assp_hist_over = assp_hist,
 assp_hist_below = assp_hist,
 sales = sales,
 sales_prev = sales_prev)
reg_dat.simulate$assp_hist_over[reg_dat.simulate$assp_hist_over<0] = 0
reg_dat.simulate$assp_hist_below[reg_dat.simulate$assp_hist_below>0] =
0

X = cbind(rep(1,n),
 reg_dat.simulate$ln_roa,
 reg_dat.simulate$age,
 reg_dat.simulate$assp_hist_over,
 reg_dat.simulate$assp_hist_below)

##
simulate model
#
##

#exit_term
beta.exit <- c(2.14, 0.1,0, -0.04, 0.125)
Emp_effect <- -0.5*emp_prev>1
Exit <- runif(n) > logit(X%*%beta.exit + Emp_effect)

#equal
beta.equal <- c(-1.41, -2.39, 0, 0.03, -0.03)
Emp_effect <- -3 + 3*exp(1-emp_prev)
Equal <- runif(n) > logit(X%*%beta.equal + Emp_effect)

#prob above
beta.above <- c(0.23, 3.42, 0, -0.19, -0.22)
Above <- runif(n) > logit(X%*%beta.above)

Above

```

```

beta.Above <- c(-0.86, -1.71, -0.142, 0.06, -0.07)
Above.val = rgeom(n, prob = 1/(1+exp(X%%beta.Above+ log(emp_prev))))
+ emp_prev
Below
beta.Below <- c(-2.71, -4.40, -0.01, 0.33, -0.18)
Belov.val = emp_prev -rgeomc(X, beta.Below, emp_prev, offset
=log(emp_prev))

reg_dat.simulate$exit <- Exit
reg_dat.simulate$emp <- emp_prev*Equal +
(1-Equal) * Above * Above.val+
(1-Equal) * (1-Above) * pmax(Belov.val,1)

}else{

emp_prev <- rgeom(n, 1/8.9)
while(sum(emp_prev>40)>0){
index = emp_prev>40
emp_prev[index] = rgeom(sum(index),1/8.9)
}
emp_prev <- emp_prev+10
#model.age = glm.nb(age-2 ~ 1 + emp_prev, data=reg_dat)
age = rbinom(n, mu = exp(1.458010 + 0.003768*emp_prev), size =
1.821) + 2
#lm_roa.model <- lm(ln_roa ~ emp_prev + age, dat= reg_dat)
ln_roa <- 8.859e-02 + -8.679e-04 * emp_prev + 8.498e-05 * age +
rnorm(n, sd = 0.09403)
#assp.model <- lm(assp_hist ~ emp_prev + age + ln_roa , dat= reg_dat)
assp_hist <- 2.19752 + -0.01359 * emp_prev + -0.09728 * age + 3.36240
* ln_roa + rnorm(n, sd = 6.498)
sales_prev <-exp(2.7169899 + 0.0059802 * emp_prev + 0.0076848 * age +
-0.2055079 * ln_roa + rnorm(n, sd = 0.3671))
sales <-exp(0.9379676 * log(sales_prev) +0.0277109 * assp_hist
+ 0.0044575 * emp_prev + 0.5811246 * ln_roa + sqrt(rgamma(n,
shape=1))*rnorm(n, sd = 0.5552))
reg_dat.simulate <- data.frame(emp_prev = emp_prev,
age = age,
ln_roa = ln_roa,
assp_hist = assp_hist,
assp_hist_over = assp_hist,
assp_hist_below = assp_hist,
sales = sales,
sales_prev = sales_prev)
reg_dat.simulate$assp_hist_over[reg_dat.simulate$assp_hist_over<0] = 0
reg_dat.simulate$assp_hist_below[reg_dat.simulate$assp_hist_below>0] =
0

X = cbind(rep(1,n),
reg_dat.simulate$ln_roa,
reg_dat.simulate$age,
reg_dat.simulate$assp_hist_over,
reg_dat.simulate$assp_hist_below)

beta.exit <- c(-2.27, -3.74, 0, 0.05, 0.015)
Exit <- runif(n) > logit(-X%%beta.exit)
#equal
beta.equal <- c(-1.54, 0.44, 0, -0.01, 0.02)
Equal <- runif(n) > logit(X%%beta.equal)
#prob above
beta.above <- c(0.27, 3.52, 0, -0.09, -0.08)

```

```
Above <- runif(n) > logit(X%%beta.above)
Above
beta.Above <- c(-1.65, -0.2, -0.03, -0.01, -0.02)
Above.val = rgeom(n, prob = 1/(1+exp(X%%beta.Above + log(emp_prev))))
+ emp_prev
Below
beta.Below <- c(-2.31, -0.69, -0.02, 0.11, -0.06)
Belov.val = emp_prev - rgeomc(X, beta.Below, emp_prev, offset =
log(emp_prev))

reg_dat.simulate$exit <- Exit
reg_dat.simulate$emp <- emp_prev*Equal +
(1-Equal) * Above * Above.val+
(1-Equal) * (1-Above) * pmax(Belov.val,1)

}
return(reg_dat.simulate)
}
```