CAMBRIDGE
Cambridge University Press
978-1-107-17548-8 — The Rise and Fall of Business Firms
S. V. Buldyrev , F. Pammolli , M. Riccaboni , H. E. Stanley
Excerpt
More Information

1

Introduction

This book aims at characterizing some fundamental features of processes of Schumpeterian innovation and "creative destruction" (Schumpeter 1934) driven by firms through the development of new ideas and the launch of new products (see Nelson 1959; Ziman 1977; Freeman 1982; Winter 1984; Dosi 1988; David 1992; Aghion and Howitt 1992; Aghion and Durlauf 1998; Mokyr 2005).

Firms are the key actors of our representation of the relationship between innovation, growth, and instability.

In particular, we found inspiration from a provocative thought experiment by Herbert Simon:

> Suppose that [. . .] a mythical visitor from Mars [. . .] approaches the Earth from space, equipped with a telescope that reveals social structures. The firms reveal themselves, say, as solid green areas with faint interior contours marking out divisions and departments. Market transactions show as red lines connecting firms, forming a network in the spaces between them. [. . .] A message sent back home, describing the scene, would speak of *large green areas interconnected by red lines*. It would not likely speak of *a network of red lines connecting green spots.*
>
> *(Simon 1991, page 30)*

Despite the indisputable dominant role of firms in modern economies, and notwithstanding a few notable exceptions, the mainstream economic thought still considers firms as green peas (see Penrose 1959; Teece 1982; Teece et al. 1994). To fill this research gap, we point our telescope right at the core of the large green areas and their interior contours, and we focus on how small and large multiproduct firms shape the evolution of industry structure and contribute to economic growth.

The top 10 largest employers in the United States account for more than 5% of the total employment in the private sector, and the top 20 US companies combined have more employees than all job losses caused by the Great Recession at its peak in 2009. Furthermore, firm-level events of growth and decline related to the top 100 U.S. firms explain about one-third of aggregate fluctuations in output growth (Gabaix 2009), while industry-specific events account for at least half of macroeconomic volatility (Atalay 2017).

Firm turnover and innovation shape economic dynamics at higher levels of aggregation, and the growth and decline of business firms associated with technological innovations, market selection, and demand evolution affect real business cycles and economic growth (see Fu et al. 2005; Gabaix and Ibragimov 2011; Acemoglu et al. 2012; Carvalho and Gabaix 2013; Acemoglu and Cao 2015).

A crucial aspect of real-world firms vividly painted by Simon is that they are not elementary units of analysis, but they have "interior contours marking out divisions" (Simon 1991). Firms producing and selling multiple products across markets and countries dominate the global economy. Although the majority of firms produce a single product, multiproduct firms represent more than 90% of the total output in the United States (Bernard et al. 2010). Walmart, the largest company in the world with 2.3 million employees, sells more than 4.2 million products in 11.7 thousand retail units. Amazon, the second largest company in the world in terms of workforce, sells around 356 million products all around the world.

Data availability at the micro level of individual products and narrowly defined markets make it now possible to investigate the nuts and bolts of firm size dynamics, measuring the relative contribution of relevant elements, such as product adding, growing, declining, dropping, and switching.

Against this background, we present here a set of facts and a theoretical framework on the growth and decline of business firms.

We build on our own work in the field from the past two decades (Stanley et al. 1996; Amaral et al. 1998; Lee et al. 1998; Bottazzi et al. 2001; De Fabritiis et al. 2003; Fu et al. 2005; Buldyrev et al. 2007; Riccaboni et al. 2008; Growiec et al. 2008, 2018) and on a research tradition that dates back to Robert Gibrat to introduce a parsimonious theoretical framework, which, we believe, accounts for a vast set of observed empirical regularities and their interplay (Gibrat 1931; Ijiri and Simon 1977; Sutton 1997, 2007).

Specific features of our approach have been outlined in earlier publications (see, for exemple, Fu et al. 2005; Growiec et al. 2018). However, this book represents an original contribution in that, together with significant revisions and extensions of our previous work on the subject, it provides a comprehensive view of our framework for the analysis of the growth and decline of business firms.

We consider both the scale and scope of firms (Chandler 1990). Firms grow by increasing the size and scope of their activities. One mechanism of growth, which has been largely investigated in the literature, concerns expansion and productivity gains of already existing activities, while a second fundamental driver is associated with entry of new firms as well as with innovation and diversification by established firms across new products and markets (see Carsten and Neary 2010; Arkolakis and Muendler 2012; Autor et al. 2017).

In our journey, we combine analytical techniques that are well established in statistical physics and in economics.

On the one hand, we present firms as systems composed of units and products, and we are interested in establishing which relationships display scaling phenomena (Stanley 1971; Montroll 1987; Kadanoff 1991; Mantegna and Stanley 1995; West 2017) across multiple levels of aggregations of the economic system (i.e., products, firms, industries, and national economies). The identification of consistent patterns, statistical regularities, and scaling laws, i.e., common properties of a set of plots of one quantity against another one across data sets, time frequencies, and time periods, has sustained our process of theory formation by providing discipline on the shape of invariant measures and on their interplay (see Brock 1999).

On the other hand, we aim at estimating some key conditional densities/predictive distributions that are significantly influenced by relevant explanatory mechanisms so that they can discriminate across otherwise equally plausible data generating processes (Barro 1991, 1996; Brock and Durlauf 1999; Durlauf and Quah 1999).

Overall, this book aims at providing a unifying interpretative framework for a set of otherwise disconnected empirical facts and regularities, which hold across empirical domains and timeframes, irrespective of differing features (Schmalensee and Willing 1989; Sutton 1997; Caves 1998; Klepper and Thompson 2006; Sutton 2007). In particular, while focusing on the key role of firms in shaping innovation, we also account for a list of relevant and related "stylized facts" (analogously to Klette and Kortum 2004) that have been observed in a number of quite diverse settings:

(I) *Size distributions of firms are highly skewed in their upper tails*. An extensive literature has investigated the properties of size distributions of varying skewness (Schmalensee and Willing 1989; De Wit 2005; Rossi-Hansberg and Wright 2007*a*). While exact shapes of size distributions are still debated (De Fabritiis et al. 2003; Luttmer 2010; Bottazzi et al. 2011), the Pareto and the lognormal distributions are retained as benchmarks. In some cases, following Gibrat's Law, the size distribution of firms has been found to be approximately lognormal for a broad range of data (Gibrat 1931; Sutton 1997). Alternatively, following Pareto, other scholars have shown that in a double logarithmic scale, empirical size distributions are well approximated by a straight line in the upper tail (Simon and Bonini 1958; Ijiri and Simon 1977). Other studies have found that the size distribution of firms is a power law distribution with a specific exponent, so that the number of firms with a size greater than a given value S is inversely proportional to S

(the so-called Zipf's Law, see Steindl 1987; Gabaix 1999; Axtell 2001; Solomon and Richmond 2001; Saichev et al. 2009; Johansen and Sornette 2010; Luttmer 2011, 2012; Carvalho and Grassi 2015).

(II) *Small, young firms have a lower probability to capture economic opportunities and, as a consequence, to grow and survive; but those small firms that do survive tend to grow faster than large firms.* Among larger firms, growth rates tend to be unrelated to past growth or to firm size (Mansfield 1962; Jovanovic 1982; Evans 1987b; Hall 1987; Dunne et al. 1989; Perline et al. 2006; Rossi-Hansberg and Wright 2007*b*; Luttmer 2011), while growth rates for the small surviving firms decrease with size.

(III) *The distribution of firms' growth rates is not Gaussian, but "tent-shaped."* While, for many decades, the literature on the growth of firms has focused on the relation between size and the corresponding expected growth rate, a scaling law was recently observed, which reveals rare events involving extremely large positive and negative shocks, uncovering a double-exponential distribution close to the origin, with power law tails (Stanley et al. 1996; Bottazzi et al. 2001; Fu et al. 2005; Perline et al. 2006; Buldyrev et al. 2007; Bottazzi et al. 2011; Growiec et al. 2018). In particular, Fu and colleagues (Fu et al. 2005) have shown that once the tent-shaped distribution emerges at the product and firm level, it becomes a stable feature of processes of growth at higher levels of aggregation within the economy up to the level of a country's gross domestic product (GDP) (Lee et al. 1998; Fu et al. 2005).

(IV) *Conditioning on firm size, the volatility of growth rates exhibits a decreasing pattern* (see also the seminal papers by Hymer and Pashigian 1962; Mansfield 1962; Evans 1987b; Stanley et al. 1996; Bottazzi et al. 2001; Sutton 2002; Riccaboni et al. 2008; Bottazzi et al. 2011; Koren and Tenreyro 2013). In recent years, it has been found that the standard deviation of firms' growth rates decays with firm size as a power law, with a power of approximately $-0.15/-0.25$. In other words, large firms have a lower volatility of growth but they seem to be more unstable than the central limit theorem would predict. The negative size-variance relationship has been observed at different aggregation levels up to the GDP level (Lee et al. 1998; Fu et al. 2005; Castaldi and Dosi 2009; Koren and Tenreyro 2007, 2013). This fact, together with the heavy tail size distribution of firms, is the basis of the "granularity hypothesis" put forward by several researchers (Gabaix 2009; Gabaix and Ibragimov 2011; Carvalho and Gabaix 2013), according to which the individual shocks experienced by large firms can propagate aggregating into large GDP shocks (see Aoki and Hawkins 2010; Barro and Jin 2011; Aoki and Yoshikawa 2012; Solomon and Golo 2014; Carvalho and Grassi 2015) and contribute to the evolution of economic cycles.

Obviously, departures from these stylized facts will be found in the literature (Perline et al. 2006). Nevertheless, the fact that these empirical patterns hold for multiple and diverse data sets, contexts, time frames, and levels of aggregation suggests that context-specific modelling is of little help and that some primitive and general mechanisms should be investigated instead (Ijiri and Simon 1964, 1977; Sutton 1997).

In particular, we propose that any candidate model of firm growth should be tested simultaneously along all of the dimensions I–IV. Any coherent interpretative framework on the size and growth of business firms should match at least this minimum set of robust empirical facts since they characterize relevant restrictions on the space of plausible interpretative hypotheses. In fact, as stated by Brock, scaling and universality characterize “unconditional objects” (Brock 1999). The robustness of stable laws becomes then an issue, since a very large class of stochastic processes can lead to the same steady state distribution. By simultaneously considering multiple facts and their relations, we reduce the set of plausible generative processes. We then evaluate the accuracy of our conclusions through a considerable amount of statistical tests and simulations on the interplay between different predictions in order to account for as many consistent patterns as possible.

In the literature, the choice of the regularities to be matched by theoretical models has been rather partial, while facts have been somehow vaguely defined, e.g., “the size distribution of firms is highly skewed” or “the variance of growth rates is higher for smaller firms” (see Klette and Kortum 2004, p. 989). When authors have aimed at endogenizing economic forces behind firm dynamics in a general equilibrium setting, they have focused on specific predictions, along with a subset of the relevant dimensions (see Lentz and Mortensen 2008; Akcigit and Kerro 2010).

The framework that we introduce here is statistical in nature (Ijiri and Simon 1977; Slanina 2014). We do not explicitly model firms as entities obeying a maximizing behavior (Lucas 1978; Jovanovic 1982; Ericson and Pakes 1995; Luttmer 2007; Rossi-Hansberg and Wright 2007*b*; Luttmer 2011, 2012; Carvalho and Grassi 2015). Instead, we derive some coherent predictions on facts and relations from a core of “robust and primitive” features of the stochastic environment in which firms operate (Sutton 2007).

In particular, we aim to unravel some fundamental features of processes of index innovation, competition, and growth by means of assumptions that are both parsimonious and plausible. To generate falsifiable hypotheses, we begin by introducing the simplest possible model, which represents the growth and decline of business firms in relation to innovation “capture and loss” of elementary business opportunities. The neoclassical tradition within the stochastic framework focuses on equilibrium properties and on steady-state distributions (see, e.g., Jovanovic 1982; Klette and Kortum 2004; Koren and Tenreyro 2013). Our work is complementary to this tradition, since our objective is to study the properties of a growing economy in which the number of firms and products increases in time.

In our approach, time plays a key role as in any evolutionary process and, as a consequence, a steady state is present only as a limiting case.

Following John Sutton (2007), we assume that strategic interaction, product interdependence, and market forces are at work at the local level of specific markets, while firms diversify their activity across multiple lines of business and products. Each individual business opportunity is like an "island," large enough to fit just a single firm.

We come to a simplified representation of innovation and competition as the process through which firms conquer and lose "green islands" of variable sizes. A rule of proportional growth is then applied to both the number of units/products taken by each firm and to the size of each business unit, measured by its sales performance.

To summarize, we rely on two key first-approximation assumptions based on the seminal ideas presented by Ijiri and Simon, and by Gibrat (Gibrat 1931; Ijiri and Simon 1964, 1977):

(A) *Innovation: entry and exit (Simon's model)*. The number of units (product lines) in a firm grows in proportion to the number of already exisisting units, while there is a positive probability that a new unit is captured by an entrant firm;

(B) *Proportional growth of existing units (Gibrat's model).* Each unit of a firm fluctuates in size at a proportional rate, which is independent of its size, while it operates in a given, independent market.

As we will show in Chapter 3, Simon's and Gibrat's models, taken separately, cannot account for the stylized facts I–IV listed earlier.

The combination of the two models in a single generalized proportional growth (GPG) framework is the key idea of this book (see also Growiec et al. 2018).

We introduce the smallest possible number of parameters, which play a straightforward role and allow direct checks of model consistency across various domains, featuring different regimes of innovation and growth.[1] Then, through a process of successive approximations, we derive relevant sub cases. Although GPG is a minimalist benchmark, it has a rich parametric space. Simon's model relies on three essential parameters: the rate of entry of new units, λ, which is a proxy of

[1] For a complementary approach, see Nelson and Winter (2002) on evolutionary theorizing in economics change, and Malerba et al. (2018) on the role of rich, history-friendly models in accounting for the evolution of industries. In addition, Robert Axtell (1999; Axtell and Guerrero 2019) relies upon agent-based approaches to account for the evolution of industry structure.

the arrival of innovation; the rate of exit of existing units, μ, which provides a simplified representation of the other side of "creative destruction" and the rate of entry of new firms, ν, which maps new small companies pursuing innovative ideas. The Gibrat side of the GPG framework, in turn, has three essential parameters: the mean growth rate of individual units, m_η; the dispersion of unit growth rate, V_η; and the dispersion of unit sizes, V_ξ, which in the simplest benchmark case have a lognormal distribution as a result of Gibrat's multiplicative growth process. The mean size of units, m_ξ, which in the benchmark model is the same for all units, is not an essential parameter as it can be regarded as a unit size.

We explore the impact of each parameter of GPG on the stylized facts I–IV presented earlier, and investigate the case of multiple levels of aggregation, i.e., when firm divisions are made up by multiple products sold in independent submarkets, or when firms aggregate to contribute to industrial sectors and national economies.

The framework and the evidence presented in this book show how the growth of complex economic organizations can be accounted for by a primitive representation of the arrival of new products and the underlying competitive environment in which they grow, shrink, and die.

In our mapping between models and data, we rely on different techniques, including simulations, econometric analysis, and data visualitazion (Tukey 1977).

The book is organized as follows. In Chapter 2, we present the key motivating evidence for the four stylized facts I–IV that have been discussed earlier. In Chapter 3, we develop our framework and derive its predictions. Statistical tests to compare predictions with empirical data are discussed in Chapter 4, where we rely on multiple data sources. In particular, we focus on PHID (PHarmaceutical Industry Database), a unique dataset which covers about one million products and thousands of firms in the pharmaceutical industry worldwide (for detailed investigations on the evolution of the pharmaceutical industry, see Orsenigo 1989; Gambardella 1995; Pammolli 1996; Sutton 1998; Matraves 1999; Henderson et al. 1999; Mc Kelvey et al. 2004).

PHID has allowed us to study how the size and growth of each individual firm is influenced by the rate of arrival, number, size, and growth of its products. Even though similar data are now available for different industries (see Argente et al. 2018), the pharmaceutical industry is an ideal setting to test a model on size and growth of firms. In particular, as it has been shown by Sutton (1998, chapter 8), the long-term evolution of the pharmaceutical industry has been constantly shaped by innovation, while the industry consists of a large number of products and therapeutic classes, which identify almost independent submarkets (Sutton 2007).

In Chapter 5, we complement the analysis of Chapter 4 and test the empirical accuracy and plausibility of our assumptions, as well as the sensitivity and robustness of our predictions. We focus on some unrealistic components of our framework and outline the corresponding violations of the benchmark, which is obviously an approximation that does not aim to reproduce the entire fine structure of the observed phenomena. Our analysis in Chapter 5 shows the key function of the benchmark to sustain a strategy of successive approximations. In fact, we have designed our framework starting from the simplest possible assumptions, followed by building upon its simplicity, parsimony, generality, and falsifiability (Popper 1961; Simon 1968) to encompass additional mechanisms that can lead to finer approximations, auxiliary hypotheses, and more accurate representations within specific domains.