



Data-driven definitions of gazelle companies that rule out chance: application for Russia and Spain

Ivan Savin^{1,2,3,4}  · Maria Novitskaya⁵

Received: 11 November 2021 / Revised: 19 December 2022 / Accepted: 23 December 2022
© The Author(s) 2023

Abstract

The phenomenon of fast-growing companies exhibiting sustained growth and creating disproportionately many new jobs, so-called “gazelles”, has been widely analyzed in the literature. The criteria defining “gazelles”, however, lack a consensus, while it cannot be ruled out that superior performance of these companies is just good luck. We use large firm-level datasets for Russia and Spain and conduct a Monte Carlo experiment with first-order Markov chains to derive a definition of “gazelle” companies and ensure that their existence cannot be explained by chance only. Our results demonstrate that the definitions of “gazelle” companies differ between the two countries warning against using same definition for different countries. We find that the “gazelles” account for about 1–2% of the companies in our datasets and are responsible for approximately 14% of employment growth in Russia and 9% in Spain. These companies are concentrated in economic sectors like retail trade, real estate and construction.

Keywords Fast-growing firm · Markov Chain Monte Carlo · SME · Sustained job creation · Employment

JEL Classification C63 · D22 · J60 · L25

✉ Ivan Savin
ivan.savin@uab.cat

- ¹ Institute of Environmental Science and Technology, Universitat Autònoma de Barcelona, Barcelona, Spain
- ² Graduate School of Economics and Management, Ural Federal University, Yekaterinburg, Russian Federation
- ³ ESCP Business School, Madrid, Spain
- ⁴ ICTA, Edifici Z, UAB Campus, 08193 Bellaterra, Spain
- ⁵ Institute of Economics, of the Ural Branch of the Russian Academy of Sciences, Yekaterinburg, Russian Federation

1 Introduction

Employment rate is one of the main indicators of labor market, which characterizes the economy as a whole. Unemployment growth leads to a decrease in the quality of life, as it entails an increase in poverty, social conflicts, and crime. Therefore, over the last few decades job creation has been one of the key priorities for policy makers around the world (Economist, 2012; Vértesy et al., 2017). Especially in times of economic instability, governments are taking active measures to stimulate employment growth.

The issue of job creation by firms is also widely studied in the scientific literature (Delmar et al., 2003; Guilmi & Mokhtari, 2016; Henrekson & Johansson, 2010). Most articles focus on a small fraction of fast-growing firms, namely the so-called “gazelle” companies, or simply “gazelles” (Rocha & Ferreira, 2021). These companies consistently create a disproportionately large number of new jobs in different countries and economic sectors. Many studies tried to formulate criteria to define these “gazelle” companies and identify factors explaining their high growth rates (Coad et al., 2014). Succeeding in this would allow to develop more effective policy measures to promote sustained employment growth (Brown et al., 2017).

As it turns out, however, the definition of “gazelle” companies in the literature varies a lot depending on the country and the period under consideration. While some researchers looked on employment growth (e.g., Ahmad, 2006; Flackenecker et al., 2020; Schreyer, 2000), others defined “gazelles” based on their revenue growth (Autio et al., 2000; Birch, 1987). They also used different threshold values for growth rates and different number of years of consecutively high performance necessary to be counted as a “gazelle” (see Sect. 2). Furthermore, while earlier studies suggested that “gazelles” are small companies (Brüderl & Preisdörfer, 2000; Lotti et al., 2003), later studies could not confirm this (Brown et al., 2017; Henrekson & Johansson, 2010). It has been further argued that it is possible that periods of high growth rates of firms are associated neither with external conditions (e.g. macroeconomic indicators) nor with individual characteristics of the companies, and thus are not predictable (Bianchini et al., 2017; Coad et al., 2013; Gibrat, 1931). If the latter is true, policies aiming to support “gazelles” will not be effective in achieving their target but only increase volatility in firm performance (Comin & Philippon, 2005). Furthermore, given the inconsistent results on the drivers of success of the “gazelle” companies, Barney (1986) argued that success of “gazelles” may be purely due to good chance. If his assumption is right, looking for factors explaining persistent growth is a waste of time.

The present study aims to derive the definition of “gazelle” companies from empirical data by asking the question about how long a firm should continuously exhibit superior performance in terms of employment growth to be recognized as a “gazelle”. To answer it, we analyze large samples of companies from Russia and Spain provided by Bureau van Dijk (BvD) and employ a data-driven Markov Chain Monte Carlo (MCMC) method that allows us to abstract from distributional

assumptions regarding firm performance and ensure that high growth rates of companies in those countries cannot be explained by chance alone. As a result, we derive definitions of “gazelle” firms for both countries depending on the number of years we observe them in the data and estimate their contribution to employment growth in each of the two considered economies. Furthermore, we discuss in which economic sectors the “gazelle” companies tend to concentrate over time.

Methodologically our approach is closest to Henderson et al. (2012), who studied “gazelle” companies in the US observed between 1965 and 2008 by means of MCMC approach and to Korzinov (2018), who applied the method to a set of European countries (the UK, France, Italy and Spain) in the period 2004–2011. There is no study, to the best of our knowledge, doing anything similar for the case of Russia, one of the largest European economies that is relatively understudied in the scientific literature. Moreover, Russia—frequently referred to as an emerging economy or economy in transition—represents a good contrast to a developed economy like Spain to test if the data-driven definition of “gazelle” companies differs between countries. Furthermore, since we analyze a more recent period of 2010–2018, our findings are compared to those obtained earlier for Spain to test if the definition of “gazelles” is also sensitive to the period of investigation. Last not least, more recent period we analyze can be considered as more relevant for current policy making.

Our findings are valuable for several reasons. First, as we show in the paper, there is no “one size fits all” definition of “gazelle” companies, and performance being exceptional in Spain is not necessarily deemed the same in Russia. This calls for more nuanced analysis of high-growth firms and, respectively, differentiated policy measures to support these companies. Second, using the novel data-driven definitions of “gazelles” we demonstrate that the contribution of the “gazelle” companies in employment growth in the two economies differs. In Russia they are responsible for 13–15% of employment growth, while in Spain—for about 9%. The larger estimate for Russia may attract more attention both from the side of policy makers and economic researchers to study the phenomenon of “gazelle” companies in the country more thoroughly. Third, we show that while in Russia “gazelle” companies tend to concentrate in information and communication, wholesale and retail trade, and transportation and storage, in Spain sectors with disproportionately more “gazelles” are real estate, construction, and education. These results indicate which sectors can potentially serve as a source of employment growth in the two economies.

To summarize, while many previous definitions of “gazelles” were based on arbitrary thresholds with regard to firm size, age, employment and revenue growth, none of them could ensure that the (temporary) success of those firms is driven by good luck only. We apply the methodology originally used by Henderson et al. (2012) based on MCMC approach to derive a definition of “gazelles” that rules out chance. In doing so, we focus on firm performance in terms of employment and not return on assets since this is a more relevant indicator for macroeconomic policy. We use data from Russia and Spain as our two test cases: the first allows us to test the methodology for an emerging economy setting, while the second—to compare our results to those obtained earlier for Spain by Korzinov (2018) and to check if the definition of “gazelles” is also sensitive to the period of investigation. As a result, we do

not only derive definitions of “gazelles” and quantify their number and contribution to country’s employment growth, but also identify sectors where “gazelles” tend to concentrate.

The remainder of this paper is organized as follows. Section 2 provides a literature overview on various definitions of “gazelle” companies in the past and the role of randomness in explaining high rates of firm growth. Section 3 describes our data and the Markov Chain Monte Carlo approach used. Section 4 presents our results for Russia and Spain, while Sect. 5 concludes.

2 Literature review

2.1 Definition of “gazelle” companies and their growth factors

There are many studies in the literature devoted to “gazelle” firms and their growth factors. Studying growth trajectories of US firms Birch (1987) classified them into three categories: large ones he named as “elephants”, small—“mices”, and steady and rapidly growing “gazelles”. In the latter category, Birch (1987) included firms that exhibited revenue growth of 20% or higher for at least five years in a row. At that time only 4% of firms, which provided about 70% of employment growth, were classified as “gazelles”. According to Birch (1987), these were young consistently high-performing firms that increase their staff much faster than their competitors and contribute greatly to the employment rate in the country.

The study by Birch (1987) sparked scientific discussion and many more studies analyzing “gazelle” firms, although the definitions employed varied considerably (see Table 1). First, Birch and Medoff (1994) and Birch et al. (1995) added an additional restriction that a base year revenue should exceed 100 thousand USD excluding very small companies. Autio et al. (2000) considered firms as “gazelles” if they increased sales by at least 50% consecutively over three years. Analyzing 367 Finnish “gazelle” firms from more than 15 industries between 1994 and 1997, Autio et al. (2000) concluded that “gazelle” firms are predominantly small and young companies.

Schreyer (2000) was one of the first who classified “gazelles” companies not based on their sales growth but employment. He defined the fast-growing firms as those that were in the top 10% of the fastest growing companies in terms of employment in the first five years of operation and had more than 10 employees. That is, a priori, Schreyer considered only young firms. According to Schreyer (2000), the fast-growing companies provide 50% to 60% of all jobs in G7 countries. Halabisky et al. (2006) used a similar approach determining two types of “gazelles”: a “hyper growth” firm increasing its employment by 150% in the four-year period, and a “strong-growth” firm with 50–150% growth rate over the same period. These two types of “gazelles” accounted for 7% of the firm population but for 56% of the net employment creation in the private sector over the period 1985–1999.

Today, one of the widely used definitions of “gazelle” firms is the one provided by the Organization for Economic Cooperation and Development (Ahmad, 2006) stating that “gazelle” firms are enterprises with an average employment growth rate

Table 1 Definitions of “gazelle” companies used in the literature

Authors	Definition of “gazelle” companies	Country/industry	Period	Some key findings
Birch and Medoff (1994) Birch et al. (1995)	Companies with revenue growth of 20% or more every year for five years, and base-year revenue exceeding 100 000 USD	USA/all USA/all	1988–1992 1990–1994	4% of firms generated 60% of all new jobs in the US “Gazelles” account for all new jobs in the US economy
Autio et al. (2000)	Companies with revenue growth of 50% or more every year for three years	Finland/all	1994–1997	“Gazelles” account for 0.2% of firms and are not concentrated in high-tech sectors
Brüderl and Preisendörfer (2000)	A firm with one employee must exhibit a 500% employment growth rate over the first four years (i.e. 5 more employees); a firm with five or more jobs—at least 100% growth in four years	Upper Bavaria (Germany)/all but crafts, agricultural businesses, physicians, architects, and lawyers	1985–1986	“Gazelles” account for 4% of all firms and over one-third of all new jobs
Schreyer (2000)	Companies among the top 10% fastest growing in terms of employment in the first five years of operation	Germany, Italy, the Netherlands, Spain, Sweden and Quebec (Canada)/manufacturing	1985–1995	“Gazelles” account for 50–60% of all new jobs
Halabisky et al. (2006)	Companies with an employment growth of 50% or more for four consecutive years (1985–1989)	Canada/all private sector	1985–1999	“Gazelles” account for 7% of all firms and 56% of new jobs
Deschryvere (2008) Bravo-Biosca (2010) Flachenecker et al. (2020)	Companies with an employment growth of 20% or more for three consecutive years with more than ten employees	Finland/all 8 European countries, USA, Canada, and New Zealand/all 21 EU Member State/all	2003–2006 2002–2005 2012–2017	“Gazelles” account for 5% of all firms and more than all new net jobs The share of “gazelles” is 37% higher in the US than in the average European country (5.9% and 4.3% respectively). “Gazelles” account for between 35 and 50% of all new jobs “Gazelles” account for 7 to 20% of all firms (on average 10.7%) and responsible for 53% of net new jobs

Table 1 (continued)

Authors	Definition of “gazelle” companies	Country/industry	Period	Some key findings
Hözl (2014)	Comparison of two approaches: (i) belong to 5% best performing in terms of a product between absolute job growth and relative job growth and have more than ten employees (HGF), (ii). have an employment growth of 20% or more for three consecutive years and have more than ten employees (HBF)	Austria/all private sector	1972–2007	HBFs have a larger probability of repeating their fast growth than HGFs. Little difference between HBF and HGF for survival
Bianchini et al. (2017)	Companies with an annual growth rate belonging to the top 10% of the yearly cross-sectional distribution of either sales or employment growth (or both) for at least four (out of five) years	Italy, Spain, France and the UK/all	2004–2011	“Gazelles” account for 1–1.5% of all firms in manufacturing and 1.5–2% in services
Erhardt (2021)	Comparison of three approaches: i) belong to 1% of firms with the highest absolute increase in employees; ii) belong to 1% of firms with highest percentage growth rate; iii). have an employment growth of 20% or more for three consecutive years and have more than ten employees	Bulgaria/all	2001–2010	“Gazelles” according to the three definitions vary considerably in their ability to repeat their fast growth

of more than 20% per year over a three-year period and with ten or more employees at the beginning of each reporting period. An additional restriction by OECD is that the term “gazelle” firm should only be applied to young, fast-growing firms, or more specifically to businesses under five years old. Deschryvere (2008), using the definition by Ahmad (2006), shows that in Finland in the period 2003–2006 high-growth firms account for 5.4% of the total number of firms with more than ten employees. He concludes that while majority of fast-growing firms are small (i.e. under 50 employees) companies, fast-growing medium-sized (i.e. 50–249 employees) firms create more jobs. Recently this definition has been used in several cross-country studies of “gazelles” (Bravo-Biosca, 2010; Flachenecker et al., 2020). Erhardt (2021) and Mogos et al. (2021) compared the definition of Ahmad (2006) with several alternative approaches that select firms belonging to the top 1% or 5% in terms of absolute and relative employment growth. They conclude that the results differ a lot between the definitions. Hölzl (2014) earlier compared the definition of Ahmad (2006) to the product of absolute and relative job growth and came to a similar finding.

Despite a plethora of complementary theories, economists disagree on the factors that drive sustained high growth rates for a company. Researchers distinguish two main types of factors that determine high growth rates, namely, idiosyncratic characteristics of firms and external factors that indirectly affect their performance. For example, Acs and Mueller (2008) find that “gazelle” companies in the US are typically located in large diversified metropolitan regions. Cabral and Mata (2003) studying Portuguese manufacturing firms argue that a lack of financial resources hinders business growth, especially in case of small or recently established companies. Becchetti and Trovato (2002) using data for Italy come to the same conclusion. More recently, Abbate and Sapio (2019) demonstrated that stock markets do not stimulate faster growth of “gazelle” companies, but merely increase firm growth dispersion. Feindt et al. (2002), St-Jean et al. (2008) and Moreno and Coad (2015) stress the role of the entrepreneur and her management skills in seizing opportunities for the firm growth. Many studies, including Brüderl and Preisdörfer (2000), Lotti et al. (2003) and Voigt and Moncada-Paternò-Castello (2012), find that smaller companies tend to grow faster than their larger competitors. Furthermore, Brüderl and Preisdörfer (2000) studying “gazelles” in Germany conclude that new firms grow faster if they offer a more innovative product. Later Coad and Rao (2008) and Ciriaci et al. (2016) generalized this argument to innovation more generally (proxied, e.g., by investments in R&D or patents) and confirmed it based on the US and Spanish data stressing that only innovative companies are able to sustain high growth over time. As for firm age, Evans (1987) and Yasuda (2005) have shown that younger firms are more likely to exhibit high employment growth. Henrekson and Johansson (2010) did a meta-analysis of existing studies on “gazelles” concluding that age of a company is a more important factor in determining the firm growth than their size. The authors argue that “gazelle” firms can be of all sizes and that large “gazelles” remain an important source of job creation. Henrekson and Johansson (2010) also argue that “gazelles” are present in most industries, but particularly in high-tech industries and services. Bianchini et al. (2017) comparing Italy, Spain, France and the UK find that persistent high-growth firms do not differ in terms of

their economic and financial characteristics from their competitors which exhibit only temporary growth.

Speaking more generally, Pugliese et al. (2021) did a large literature analysis of drivers of startup growth. After examining 316 studies and 66 drivers grouped in five categories, they come to conclusions similar to the ones above: financial resources, firm age, technological capabilities, R&D investments and the role of personal characteristics of the entrepreneur including her experience and education are among the most important factors of startup growth (measured either by sales or employment).

According to Guilmi and Mokhtari (2016), who conducted a systematic review of the scientific literature on the fast-growing firms, at the time of their publication 7% of existing studies were about “gazelle” companies in Spain. Spain was the fifth most popular country after Sweden (17%), the US (16%), the UK (13%), and Canada (8%). Unlike the Spanish “gazelle” companies and other developed countries more generally, which have been analyzed in a series of studies, there is almost no research of Russian “gazelle” companies. One of the few exceptions is the study by Yudanov (2007), who argues that in the Russian economy many market niches are not yet filled which creates opportunities for fast growth. However, his study covered only few industries and a small sample of firms identifying in total only 50 “gazelles”. Later, Yudanov (2010) estimated the share of “gazelles” in Russia with 7–8% in the period between 2000 and 2007. One explanation for the larger share of “gazelles” in Russia provided by Yudanov and Yakovlev (2018) and also confirmed by Pletnev and Barchatov (2019) was that high-growth firms in Russia may benefit from a patronage by the state and large corporations.

We have chosen Spain to test if the data-driven definition of “gazelles” obtained by the MCMC approach is not only country but also time specific. Even though we took the same country as Korzinov (2018), the period we analyze is different from the one used earlier: 2010–2018 in our case vs. 2004–2011 by Korzinov. Spain was chosen since out of the four countries analyzed by Korzinov it has the largest number of observations which is beneficial for the MCMC analysis. In Sect. 4 we report on this comparison.

To summarize, the literature review presented above shows that “gazelle” firms are mainly represented by small young companies that provide a large share of total employment in the economy. These companies received their name due to the nature of their development. They are distinguished not by gradual, but by rapid growth. Among the most common growth indicators, various researchers distinguish employment and sales. It has been demonstrated that firm size, age and industry affiliation tend to play a significant role for firm growth and must be taken into account. However, what we also find is that the studies devoted to “gazelle” firms tend to use different definitions of these companies (see Table 1¹) and typically look on a particular country (see also literature reviews by Delmar et al., 2003 and

¹ Note that the studies reviewed in Table 1 are not necessarily the first in the literature to introduce those definitions of “gazelles”. Table 1 is not meant to be an exhaustive but illustrative summary of different definitions of high-growth firms in the literature.

Henrekson & Johansson, 2010). As a result, it is difficult to make general conclusions on which firms are “gazelles” and how large their contribution to the economy is.

2.2 Randomness as an explanation for the fast firm growth

In contrast to many studies presented in the previous subsection, which argued that the firm’s sustained growth can be explained by firm idiosyncratic characteristics or external factors, Barney (1986) was one of the first who suggested that it cannot be ruled out that some periods of sustained high growth performance may be explained just by firm’s good fortune or luck. Earlier, Gilovich et al. (1985) and Tversky and Kahneman (1971) stressed that people mistakenly try to fit patterns to completely random observations. This is because people tend to regard a small sample of observations randomly drawn as highly representative, which explains some misperceptions of random sequences like “hot hand fallacy” and “gambler fallacy” (Roney & Trick, 2009). In fact, stochastic processes can produce a long sequence of seemingly extraordinary performance which is nothing else than a byproduct of chance—take the arcsine law of random walks—but are often overlooked in studies on sustained superior performance (Henderson et al., 2012). The relevant question therefore is: would a firm that was increasing its number of employees in the last few years keep increasing it further, or was it just by chance? We believe that it is worth to ensure that the company’s growth rates are not purely random, and in the following we review the literature which addressed this research question.

Perhaps the most popular way to model random process of firm growth is the law of proportional effects, or the so-called Gibrat’s law (Gibrat, 1931). The idea of Gibrat’s law is the stochastic nature of the company’s growth rates. According to this law, it is assumed that the growth of a firm in period t is some function of its results in the previous period $t-1$ and a random error, which takes into account unobservable factors. Gibrat’s law is described in Eq. (1):

$$\ln S_{i,t} = a_i + \beta \ln S_{i,t-1} + e_{i,t} \quad (1)$$

If the estimated coefficient β equals 1, Gibrat’s law holds, that is, the size of the company follows a random walk, and the growth rate does not depend on its size. When $\beta < 1$, growth rates decrease with increasing firm size. Accordingly, for $\beta > 1$, the opposite dynamics is observed.

To test the Gibrat’s law, Greene (1993) suggested to study the regression residuals to separate the effect of observed factors from the effect of unobservable variables. This approach formed the basis for many studies including the present one. Lotti et al. (2009) examining Spanish manufacturing firms and Coad et al. (2013) analyzing startups in the UK both conclude that the firm growth trajectories are consistent with a random process supporting the Gibrat’s law.² There

² The approach by Coad et al. (2013) is also known as Gambler’s Ruin theory where random growth is represented as a function of size and survival.

are however many studies which do not find empirical support for the Gibrat's law. Pirogov and Popovidchenko (2010) review results of more than fifty scientific studies that consider Gibrat's law and conclude that in most cases the law is fulfilled only for larger and older companies. For smaller and younger firms, the growth rates depend on firm characteristics and are not random. A similar conclusion has been reached by Becchetti and Trovato (2002) and Lotti et al. (2003) for Italian firms. Capasso et al. (2013) find that persistent high-growth firms are very rare and coexist with a large number of the so-called "bouncing" firms (experiencing alternately highly positive and highly negative growth rates).

However, testing Eq. (1)—and logically all models based on that (including Gambler's Ruin theory)—requires imposing restrictive assumptions about how firm performance is distributed. As McKelvey and Andriani (2005) have shown, while Gaussian distribution assumes independent events, real-world processes are much more interdependent with extreme observations occurring more often than normal Gaussian-based statistics would expect. Another example demonstrating the risk of misleading results from applying Gaussian assumptions to Eq. (1) has been presented in Henderson et al., (2012). To deal with this problem, Henderson et al. (2012) suggested to use a Markov Chain Monte Carlo (MCMC) approach when studying whether randomness can explain firm success. In particular, the authors derive from the data how many consecutive years a firm must rank among the top fastest growing ones to be called a "gazelle", and then compared this number with the number of false positive results generated by a—random in nature—MCMC experiment. Looking on the firm growth in the United States from 1965 to 2008, Henderson et al. (2012) conclude that a firm that has been observed, for example, for 7 years, must be in the top 10% of the fastest growing companies for at least five consecutive years to be called a "gazelle" firm. Korzinov (2018) applied a similar approach on data from four European countries (the UK, France, Italy and Spain) from 2004 to 2011. It turns out that in Spain a firm observed for seven years, must be in the top 10% of the fastest growing companies for only two years to be classified as a "gazelle" firm. This demonstrates that abnormal growth in Spain and the US has different criteria. In the present work, we apply a similar empirical strategy, but use a more recent data for Spain and Russia, with the latter been hardly investigated on the presence of "gazelles" before.

Finally, it is worth mentioning that a recent article by Esteve-Perez et al. (2022) employs discrete-time duration models to study the probability of experiencing a high-growth episode and firms' persistence in high growth rates. Among advantages of the approach is the possibility to establish determinants of transitions to and from the high-growth state and how these change over time. These features are unique advantages of this approach, since MCMC assumes transition probabilities to be random and constant over time. However, Esteve-Perez et al. (2022) also have to make restrictive assumptions on the distribution of transition probabilities. Therefore, the approach by Henderson et al. (2012) to date is still the only one to our knowledge that uses the distribution of firm performance from the data instead of making any ad hoc assumptions.

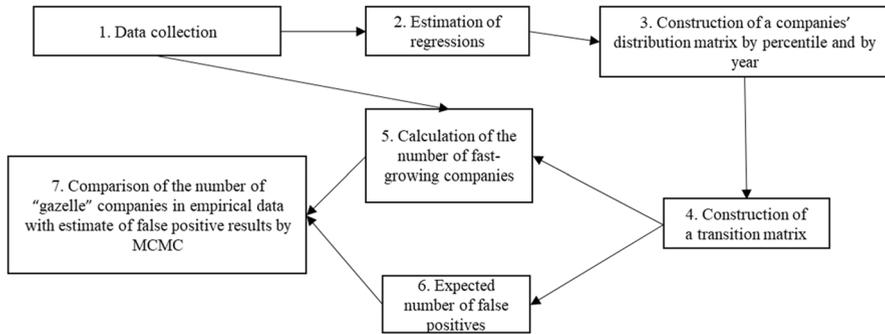


Fig. 1 Summary of research steps

3 Data and methods

The literature review in Sect. 2 demonstrated that there is no consensus among researchers on the definition of “gazelle” companies. Therefore, in this study we derive the definition in Russia and Spain from the underlying data employing a Monte Carlo experiment with first-order Markov chains. Our methodology is summarized in Fig. 1. After collecting the data, we isolate the part of unexplained variation in the firm growth in step 2 making observations comparable in terms of firm age, size, year of observation as well as industry affiliation. Next, we construct a matrix of typical transitions between different performance percentiles (steps 3–4) allowing us to use the distribution of firm performance from the data instead of making ad hoc assumptions. We run 1000 Markov chain simulations for the population of our firms to establish criteria firms must meet—how long companies must rank among the 10% fastest growing companies in their country in terms of employment growth—to be considered as a “gazelle” and count their number in our dataset (step 5). To exclude the possibility that the sustained growth can be explained by a stochastic process, we count the number of false positives that can be expected to be generated by a random Markov process (step 6). Finally, we compare how many fast-growing firms we see in real data and how many are expected based on the MCMC experiment (step 7). In the following we explain each step in more detail.

The methodology described above closely resembles the one used in Henderson et al. (2012) and Korzinov (2018). In particular, the research steps summarized in Fig. 1 are virtually identical to the ones introduced by Henderson (2012) and followed by Korzinov (2018). There are, however, a few differences worth mentioning. First, in contrast to us, Henderson et al. used return on assets (ROA) and Tobin’s q . ROA as a dependent variable. Similar to Korzinov (2018) we instead concentrate on employment as a measure of firm growth. While employment and sales are both feasible and have been widely used in the literature (Delmar et al., 2003), employment—in contrast to sales—is not sensitive to inflation and currency exchange rates. Furthermore, as our study is more interested in macro-oriented job creation, measuring growth in terms of employment is a natural choice. Second, Henderson et al. (2012) used in step 2 as controls firm size, industry, and year dummies but not firm

age. We added this variable in line with Korzinov (2018) since it appears often in the literature (see Sect. 2). Third and finally, neither Henderson et al. (2012) nor Korzinov (2018) after obtaining results in step 7 make a further effort in estimating in which sectors “gazelles” are concentrated and how much employment growth they generate in their respective countries, which makes our insights more policy relevant (see Sect. 4.2).

We use data on Russian and Spanish companies for the period from 2010 to 2018 obtained from the Bureau van Dijk (our step 1 in Fig. 1).³ In total, 220,166 unique Russian and 188,855 unique Spanish companies with at least one employee are considered. Since existing studies tend to use sometimes absolute and sometimes relative measures of firm growth (Henrekson & Johansson, 2010; Korzinov, 2018), we also employ both measures to assess firm growth rates and ensure robustness of our results. In particular, relative growth is expressed in percentage points, calculated using the following formula:

$$Rel\ Growth = \frac{(Employees_t - Employees_{t-1})}{Employees_{t-1}} \quad (2)$$

while the absolute growth indicator is measured as a difference between the logarithms⁴ of the number of firm employees (*Employees*) for two consecutive years:

$$Abs\ Growth = \log(Employees_t) - \log(Employees_{t-1}) \quad (3)$$

To isolate the part of the variation in firm growth that is not explained by observable characteristics (step 2), we use a limited set of variables designed to ensure comparability of observations in accordance with Korzinov (2018). First, to consider the impact of possible macroeconomic shocks, the regression includes dummy variables for each year under consideration. Second, to account for intra-industry differences, the regression includes categorical variables for each industry, according to the two-digit NACE 2 classifier. In addition, indicators such as company age and size preceding the respective growth rates are used since they are frequently mentioned as typical features of “gazelles” (see Sect. 2.1). Company age is measured in years, while company size is a categorical variable introduced by BvD ranging between 1 (small company) and 4 (very large company). See Appendix A for more details. As a result, the following two regression models will be tested:

$$Rel\ Growth_{it} = \beta_1^r age_{it-1} + \beta_2^r size_{it-1} + \beta_3^r industry + \beta_4^r year + e_{it} \quad (4)$$

$$Abs\ Growth_{it} = \beta_1^a age_{it-1} + \beta_2^a size_{it-1} + \beta_3^a industry + \beta_4^a year + e_{it} \quad (5)$$

Table 2 provides descriptive statistics of the data using the variable specified earlier. It is worth noting that the final sample includes only firms that

³ <https://www.bvdinfo.com/> (last accessed 20.07.2020).

⁴ We use difference in logs and not difference in absolute values following Korzinov (2018). As Delmar (2003) described, absolute growth measure has been often logarithmized to adjust for skewness.

Table 2 Descriptive statistics

Variable	Russia				Spain			
	number of observations	Min	Max	Mean	Number of observations	Min	Max	Mean
Age	2,201,660	0	317	10.22	1,857,465	0	176	16.62
Size	2,201,660	1	4	2.15	2,107,285	1	4	2.18
Relative growth	1,187,207	-0.99	16,073	0.62	1,440,664	-0.99	3774.15	0.18
Absolute growth	1,187,207	-9.35	9.69	-0.14	1,440,664	-8.08	8.24	0.05

Table 3 Distribution of companies by the number of consecutive years observed

	Russia (%)	Spain (%)
At least 2 years	100.00	100.00
At least 3 years	89.68	97.85
At least 4 years	79.43	95.34
At least 5 years	69.53	92.50
At least 6 years	58.26	89.04
At least 7 years	43.31	85.06
At least 8 years	34.81	77.65
At least 9 years	29.73	72.78

were observed in at least two consecutive periods to be able to calculate their growth. Comparing the two countries, we see that Russian companies are typically younger and grow faster (in relative terms) than their Spanish counterparts. Among the observed Spanish companies, more than 72% are observed over the entire period under consideration (nine years), while there are only about 30% of Russian companies in the sample observed consequently through the entire period (Table 3). This may be because the Russian economy is still in a transition phase with many firms entering and exiting the market (Savin, 2020). Note, however, that the BvD data do not allow to distinguish “true” entry/exit from missing values due to any other reason, which is a common problem for firm-level analysis. For example, Dosi et al. (2015, p. 647) report that for France, Germany, and the UK they also had “about 50% of the firms [...] observed for at least 6 years”. Korzinov (2018) applying methodology similar to ours to the data from Spain, Italy France, and the UK in the period 2004–2011 had even only about 10–25% of firms observed in all nine years. Thus, quality of our data is not any worse than in the previous literature. Furthermore, as can be seen from Fig. 2, the distribution of relative firm growth in both countries is skewed with few firms increasing their size by much more than 100% in a single year. For completeness of our analysis, these observations cannot be excluded from the sample. Moreover, Waring (1996) argued that by excluding outliers one deletes the most relevant data for the study of sustained long-term firm growth, since abnormal firm growth is an outlier per definition.

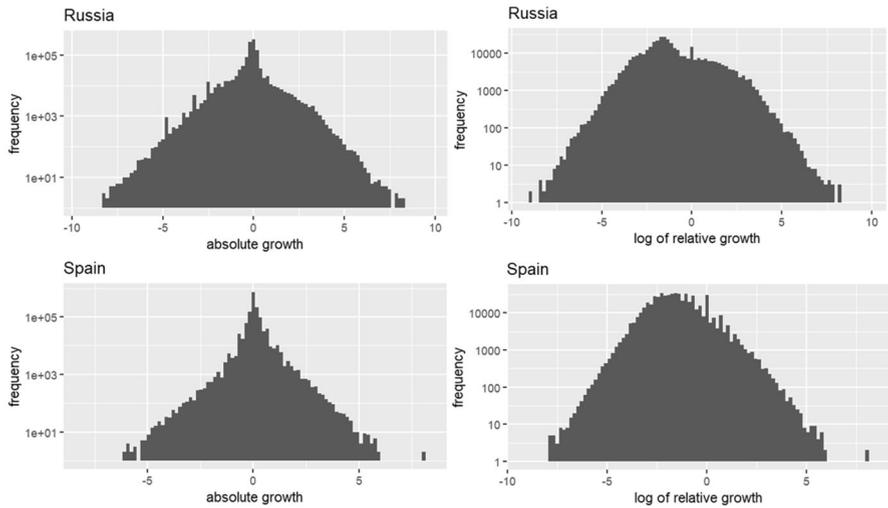


Fig. 2 Density plots of absolute and relative growth for companies in Russia and Spain. The y-axis is in log scale

Because of the presence of many outliers, regression Eqs. (4–5) are estimated using not the ordinary least squares but the least absolute deviation method. The latter requires neither a normal distribution of the error term nor the assumption of homoscedasticity and has been previously used in the firm growth literature by Bottazzi et al. (2011).⁵ The residuals obtained from this estimation are the data we subsequently are working with. This is because these residuals are free from the role of observable characteristics and can be compared across industries and years and for firms having different size and age.

To derive the definition of “gazelle” companies and make sure that firms observed empirically cannot be explained just by chance, we conduct an MCMC experiment. Before explaining the experiment, we should clarify that within this study we analyze a random process with discrete time characterized by a Markov chain with matrices of transition probabilities, which record the dynamics of the process on the basis of empirical data instead of making any ad hoc assumptions about its distribution. The Markov chain is specified by a finite set of states and transition probabilities between all states. The randomness of the process lies in the fact that over time the process passes from one state to another in a random order, which is not known in advance. A random process is called Markov if the probability of any state in the

⁵ One should note that since the number of employees is a discrete random variable, a large proportion of growth states has a zero value (see Fig. 2). As discussed by Pennacchio and Sapio (2020), the number of employees behaves like a continuous random variable only when it is “large”, which does not hold for samples with primarily small companies. However, our dataset encompasses firms of all sizes. Furthermore, the least absolute deviation method is less sensitive to this problem and was used earlier by Colombelli et al. (2014) to the growth of patent applications, which has a similar peculiarity of many zero observations.

future depends only on its state in the present and does not depend on when and how the process ended up in this state.

The following parameters are used to describe a Markov process with discrete states:

- a list of states that will be implemented in step 3 (Fig. 1) when constructing a matrix of distributions of companies by percentile by year;
- a transition matrix describing transitions between states. This is a matrix of transition probabilities for processes with discrete time, which will be implemented in step 4.

The list of states of Markov chains is determined by calculating the 100 percentiles for each period and assigning each firm its percentile. That is, the states are determined by the percentiles of the firm growth rate and the probability of the transition between percentiles i and j between two consecutive periods t and $t+1$ is estimated. It is assumed that the transition probability p_{ij} remains constant over time, the companies under consideration are homogeneous in their resources, and they maintain their position at a certain percentile for one year. Thanks to the steps 3 and 4 we thus abstract from distributional assumptions regarding firm performance and measure the distributions directly from the data.

In our step 5, the number of fast-growing companies will be calculated by comparing the actual data with 1000 MCMC simulations “replaying” the observed history of firms in our data many times and allowing randomness to generate many different outcomes mirroring the true data (Henderson et al., 2012). The MCMC method reproduces a random process multiple times for distribution parameters and the initial state of companies. That is, each MCMC run is random, but its outcome is driven by the initial current state of the company (e.g. in the first year of its observation it belonged to 45th percentile) and the transition matrix estimated from the data. The MCMC path of each firm is randomly generated with a length equal to the number of years this company is observed. Each step is based on the uniform random generation of a number from the interval $[0, 1]$, which is compared with the cumulative probability obtained by summing the probabilities over the row of the transition matrix. In this study MCMC implementation was simulated 1000 times. The MCMC experiment is implemented in the Python programming language using the PyMC3 package (Salvatier et al., 2016), which is used for Bayesian statistical modeling and probabilistic machine learning and focuses on advanced Monte Carlo Markov Chain algorithms and variational approximation algorithms. The algorithms of parallel computation allow us to reduce the computation time of the experiment from a week to one-two days.

To identify “gazelle” companies, we measure for how many consecutive years a company must rank among the 10% fastest growing in its country in terms of employment growth to be considered abnormal.⁶ This number of consecutive

⁶ The choice of the 10% cut-off is consistent with the values used by Henderson et al. (2012) and Korzinov (2018).

Table 4 Regression results

	Relative growth		Absolute growth	
	Russia	Spain	Russia	Spain
Intercept	-0.1563*** (0.002)	-0.0069 (0.009)	-0.1674*** (0.002)	-0.0068 (0.008)
Age	0.0037*** (<0.0001)	-0.00000005 (<0.0001)	0.0040*** (<0.0001)	-0.00000005 (<0.0001)
Size	0.0347*** (0.000)	0.0032*** (<0.0001)	0.0377*** (0.001)	0.0032*** (<0.0001)
Industry dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
R ²	0.01	0.01	0.03	0.01

Standard errors are in brackets. ***, **, * denote 1%, 5% and 10% significance level, respectively. Coefficients indicate whether prevalence of respective topics changes with the value of the covariates

years depends on how many years the corresponding firm is observed in our dataset. This is because firms observed for more years have more chances of accidentally reaching the top growth performance. Therefore, the definition of a “gazelle” firm will be formulated differently for each country, measure of growth (relative or absolute) and period under consideration. The number of consecutive years necessary to be attributed to “gazelles” is then determined by the minimum number of years to satisfy the 5% significance threshold to be counted as least likely outcomes among 1000 MCMC realizations.

To give an illustration, let us consider 65 thousand companies observed consecutively in Russia over all 9 years. Having simulated the history of each of these firms in 1000 runs, we may find that the ninety-fifth percentile of the number of years among the top 10% fastest growing across the 65 million simulated outcomes is five. We then set a threshold sufficiently high to keep the likelihood of a false positive below 5%, i.e. we minimize the chance that the firm we define as a “gazelle” is in fact consistent with the random Markov process on the percentile state space described in steps 3–4. Hence, the threshold we should choose for a firm observed in Russian dataset for all nine years must be six years consecutively belonging to top 10% fastest growing firms.

In step 6, the expected number of fast-growing companies satisfying the thresholds derived in step 5 will be calculated based on the implementation of the MCMC experiment. After that we compare the results obtained, to conclude whether the number of “gazelles” observed empirically is significantly higher than the number that can be explained by random process (step 7). If the observed number of companies falls into the 5% (or less) of the least probable realizations of the MCMC experiment, we can conclude that this number of “gazelles” cannot be explained by chance alone.

Table 5 Percentile distribution by years for four Russian companies in terms of absolute growth

Company name	2010	2011	2012	2013	2014	2015	2016	2017	2018
PJSC «Nizhnekamskshina»	96.0	96.0	94.0	95.0	95.0	95.0	93.0	94.0	94.0
CJSC «RuzOvo»	54.0	86.0	42.0	51.0	48.0	54.0	60.0	88.0	96.0
LLC «Uralmechanika»	43.0	5.0	53.0	66.0	57.0	70.0	24.0	57.0	88.0
JSC «Stroyservis»	46.0	85.0	77.0	70.0	79.0	83.0	64.0	80.0	79.0

4 Results and discussion

4.1 Defining “gazelle” companies

We start from presenting results of a regression analysis for Eqs. (4, 5) in Table 4. The firm size and age have a positive and significant association with firm growth rate (both absolute and relative) in Russia. In particular, older and larger⁷ companies tend to demonstrate a larger growth rate. For Spain, in contrast, only size of the company has a positive and statistically significant association. One explanation may be stronger incentives and support programs in Spain to start and develop new businesses (Barba-Sanchez et al., 2019), while in Russia the share of small enterprises in total employment is more than two times lower compared to developed countries (Chigrin, 2018; Savin et al., 2019, 2020). In the following we work with residuals from Table 4, that is the observations free from the effects of our controls (Greene, 1993).

In step 3 we build a matrix containing the distribution of companies by percentile and by year. A total of four such matrices were compiled: for absolute and relative growth rates and for each of the two countries. To construct such a matrix, we classify each company in each year to one of the percentiles. Table 5 presents a small sample of such a matrix for Russia and absolute growth.

If we denote by the percentile i the state in which a particular firm is in the period t , and by the percentile j the state in which the same firm is in the next period $t+1$, according to a first-order Markov chain the transition from i to j depends only on these two states and we can calculate the conditional transition probability based on our data. Thus, the transition matrices indicate the proportion of companies from each state (percentile) either moving in the next year to a different percentile (i.e. in 2011 CJSC «RuzOvo» was in 86th percentile but in 2012 its performance in terms of employment growth deteriorated and dropped to 42nd percentile) or staying in the same state. These matrices are constructed by averaging the results over all years available. In the calculated transition matrix (step 4 in Fig. 1) each cell contains the

⁷ The reason for the positive and significant association with firm size may be the fact that instead of sales or number of employees (typically used in the literature), we use a categorical variable introduced by BvD. This categorization is very rough with majority of firms in our samples falling into the second group, and could explain the positive and significant association between firm size and employment growth.

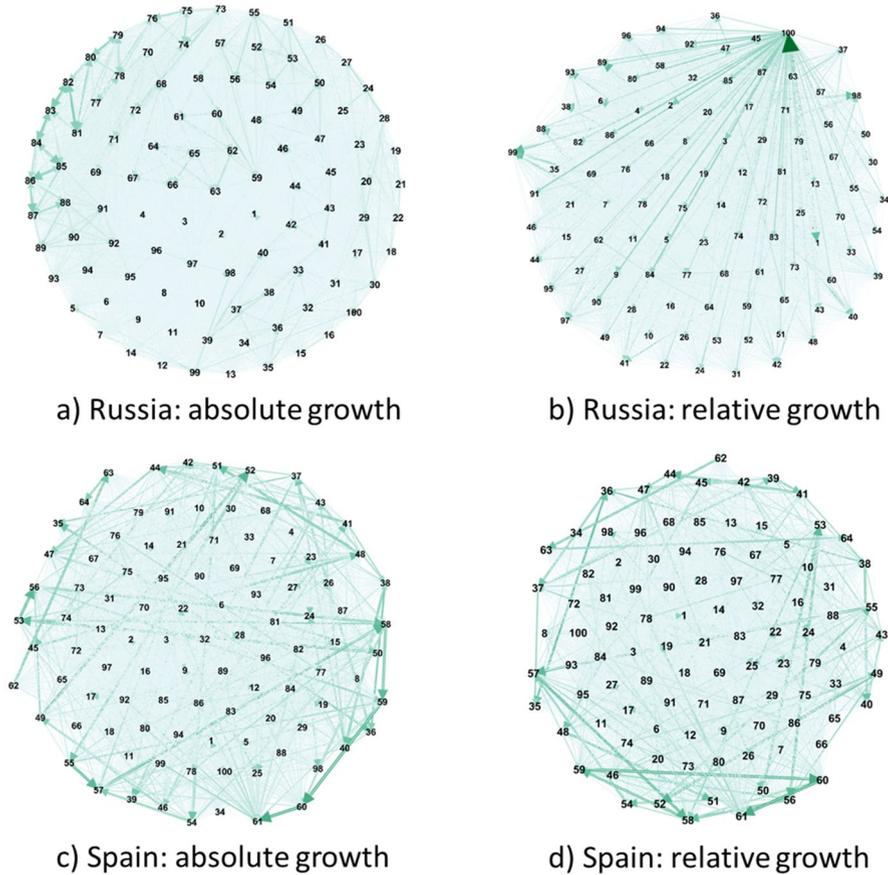


Fig. 3 Transition matrices for Russia and Spain. Nodes represent percentiles of firm performance, while edge thickness—the likelihood of transition from one state to the other. The graphs are arranged using Fruchterman-Reingold algorithm in Gephi software plotting nodes (percentiles) between which firms tend to “transit” more often closer together

average share of companies that made the transition from state i to state j in two consecutive periods of time. It is logical that the sum of each row of the transition matrix equals to unity. Again, we build four transition matrices: for the absolute and relative growth rates for both countries. The resulting transition matrices are visualized as graphs by means of Gephi software in Fig. 3 using Fruchterman-Reingold algorithm. Each graph contains one hundred nodes. Each node corresponds to one of the percentiles (states). The thickness of the edge on the graph reflects the probability of a company’s transition from one state to another. In other words, the thicker the edge on the graph, the greater the share of companies that followed that transition path. It is worth noting that these graphs are directed ones, that is, a transition from one state to another does not mean a reverse transition to the same state. At the same time, one can move virtually from any state to any other one (i.e. the graphs

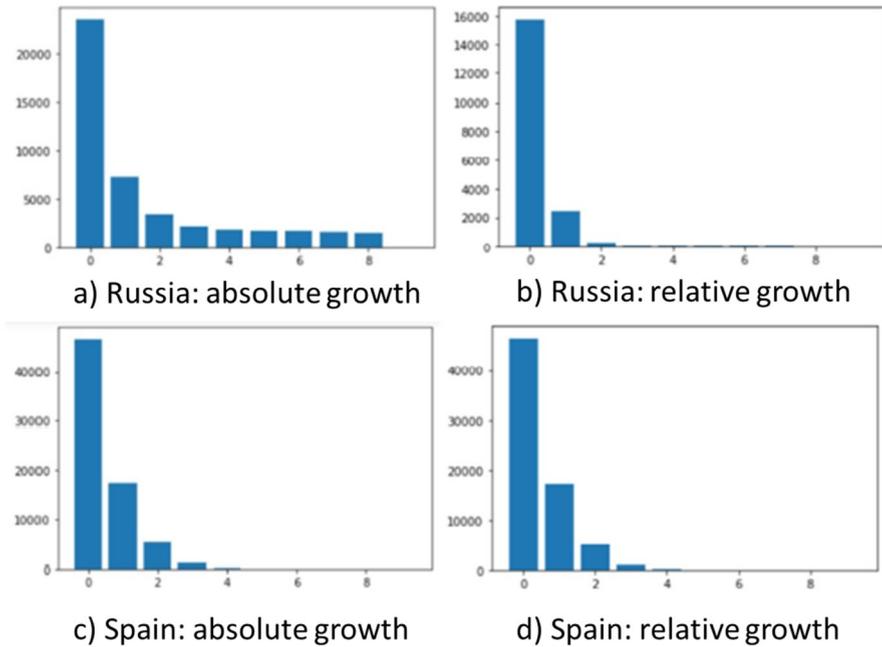


Fig. 4 Number of companies belonging to the top 10% of fastest growing over nine years

are nearly fully connected), since we have so many observations and observe firms reaching very different growth rates. Hence, from any percentile a firm can potentially achieve any growth rate, but with different probabilities.

We conclude that in the period from 2010 to 2018 there were many Russian companies maintaining their performance in terms of employment growth in absolute numbers either by increasing or reducing by one percentile their position from the 79th to the 88th percentile. That is, companies whose growth rate is higher than the average growth rate of their competitors in the sample preserve their relative performance from year to year. The graph of the transition matrix in terms of relative growth looks different. We can observe a few thick edges from the top (100th) percentile to much lower ranked states, including the 1st percentile. Thus, the transitions are more abrupt suggesting that companies which showed record growth in any given year can rapidly become the worst performing ones. In Spain in terms of absolute growth, the most likely transitions are not always as gradual as for Russian companies studied over the same period (e.g. next to 58th to 59th or 60th to 61st percentiles there are also frequent transitions from 58 to 38th, 50th to 53rd or from 37 to 35th). In terms of relative growth rates, the transition matrix for Spain is similar to the one for absolute growth displaying mostly gradual shifts with some exceptions (e.g. from 80th to 22nd percentile).

In step 5 (Fig. 1) we now formulate a definition of “gazelle” firms based on the actual data and 1000 MCMC runs. To formulate a data-driven definition, we look on how many years each firm has been consecutively ranked in the top 10% fastest

Table 6 Number of years in the top 10% of fast-growing companies to be classified as a “gazelle”

Number of consecutive years observed in the sample	Absolute growth		Relative growth	
	Russia	Spain	Russia	Spain
2	1	1	1	1
3	2	1	2	1
4	2	2	2	2
5	2	2	2	2
6	2	2	2	2
7	3	3	2	3
8	3	3	2	3
9	8	3	2	3

growing in terms of absolute or relative employment growth. Since this value is directly related to the number of years that a given company is observed in the original database, we formulate the definition of “gazelles” separately for each group of firms according to how many years we observe them in the sample. For each group of firms, we plot the number of companies belonging to the 10% fastest growing. In Fig. 4 histograms for Russian and Spanish companies observed over nine years are presented for indicators of absolute and relative growth, respectively (histograms for firms observed consecutively from 2 to 8 years are presented in Appendix B).

Comparing the histograms in Fig. 4, it is worth noting that among the Spanish companies observed during the entire period, there are very few companies belonging to the top 10% fastest growing in terms of absolute growth for four or more years out of nine. At the same time, there is approximately the same number of Russian companies which are ranked among the top 10% in absolute growth for five to eight years out of nine considered. While for Spanish companies the graphs based on the rate of absolute and relative growth are approximately the same, for Russian companies they differ considerably. This shows that there are relatively many companies in Russia that can develop steadily and increase the absolute number of employees over the years, e.g. increasing their staff by five employees. But if one looks at them in terms of relative growth, their performance is more modest.

To be sure that the observed number of consecutive years during which the company belonged to the 10% fastest growing is not a false positive result, in line with Henderson et al. (2012) we calculate the minimum number of years to satisfy the 5% significance threshold based on 1000 MCMC realizations (as explained in Sect. 3). In particular, based on the transition matrices and using a random number generator, the growth path of each firm is simulated as a percentile position in each year 1000 times. We initialize the firm path by putting it in the same growth percentile to which the firm belonged in the first of the observed periods, and let it move within the transition matrix the same number of years it is observed in our dataset. This creates a pseudo-random growth trajectory. Using the simulated data and for each number of years under consideration, we then formulate the definition of a “gazelles” to keep the likelihood of incorrectly calling a firm as a “gazelle” under 5%. The results for both countries and for both growth indicators (absolute and relative) are

Table 7 Number of “gazelle” companies

	Absolute growth		Relative growth	
	Russia	Spain	Russia	Spain
Observed number of “gazelle” companies	4245	2333	2225	2322
Observed number of “gazelle” companies in % of the total number of companies	1.9	1.2	1.0	1.2
Average number of “gazelle” companies expected from the MCMC experiment	1812 (54.1) [0.31]	856 (15.2) [0.54]	919 (32.4) [0.62]	847 (14.8) [0.56]

Standard deviation (δ) are in circle brackets, while p-values of the Kolmogorov–Smirnov test—in square brackets

summarized in Table 6 in terms of the minimum number of years a firm must exhibit consecutively outstanding (10% fastest growing) performance, while Table 7 reports the number of “gazelle” companies observed in real data.

From Table 6 we see that for Spanish companies the criteria to be classified as a “gazelle” are the same for the absolute and relative rate of growth. For Russian companies, these criteria are different. Considering the absolute growth rate of Russian companies, we can conclude that a company can be called a “gazelle” if it belongs to the 10% of the fastest growing consistently for eight years out of nine possible. At the same time, considering the relative growth rate of Russian companies, only 2 years would be sufficient. Furthermore, we find differences between the definitions of “gazelle” companies in Russia and Spain: a firm consequently observed in Russia for seven–nine years should belong to the 10% fastest growing in terms of relative growth for only two years to be considered as a “gazelle”, while for Spain it should do so for three years in a row.

Looking on the number of “gazelle” companies in the two countries (Table 7), we see that for Spain the numbers are approximately the same and are about 1.2% of the entire sample. For Russia the number of “gazelles” in terms of absolute growth is about twice larger and constitutes almost 2% of the sample.

Now we need to ensure that the number of abnormally performing “gazelle” companies we observed in step 5 cannot be explained by pure chance. To this end, we predict from the MCMC experiment how many firms could be expected to satisfy the definitions we derived (step 6). That would allow us to compare the observed number of “gazelle” companies with the number of “gazelle” companies generated randomly (step 7).

Using the simulated results over 1000 restarts and employing the same definitions of “gazelles” presented in Table 6 we calculate the expected number of “gazelles” generated by the MCMC experiment. The results are presented in Table 7 both in means over 1000 restarts (hereinafter μ) and the corresponding standard deviations (hereinafter δ). If the number of “gazelle” companies observed on real data exceeds $\mu + 3\delta$, we can conclude that the observed number of “gazelle” companies is statistically larger than the number of false positives obtained from the MCMC experiment

Table 8 Number of employees in “gazelles” in Russia and Spain

	Absolute growth		Relative growth	
	Russia	Spain	Russia	Spain
Average number of employees	42	21	47	21
Median number of employees	37.5	6	35	6
Share of SMEs (≤ 249 employees), in %	98.0	99.2	98.0	99.2

at the 1% significance level.⁸ The range of $\mu \pm 3\delta$ is used based on the assumption that the distribution of 1000 MCMC restarts reporting the expected number of “gazelles” is normal. To make sure that this is the case, we use the Kolmogorov–Smirnov test, which indicates that we cannot reject the null hypothesis that the distributions are normal for both countries and both growth indicators (the p-values of the test are reported in Table 7).

Based on the results in Table 7 we conclude that for both countries and for both considered growth criteria the number of “gazelles” observed in real data significantly exceeds the number of fast-growing companies one could expect to be present just by pure chance. In other words, the fact that we observe so many abnormally growing companies in the data must have an explanation other than pure luck. These results are consistent with those obtained by Korzinov (2018) for four European economies in the period 2004–2011 and by Henderson et al. (2012) for the US for the period 1965–2008. In the following section we are going to explore the industries where these “gazelles” are concentrated and their contribution to the employment growth of their economies.

Comparing our results for Spain with the ones from Korzinov (2018), we find that in the more recent years firms have to grow relatively faster to be classified as a “gazelle” (see Table 11 in Appendix B). In particular, if they are observed for four (seven) years, two (three) out of them should be among top 10% of fastest growing, while before one and two years, respectively, were sufficient. At the same time, the fraction of companies our MCMC approach classified as gazelles has marginally increased: from 1% by Korzinov (2018) to 1.2% in our study (Table 12 in Appendix B). This can be interpreted as an indication that persistent high growth in Spain has become more popular in the last decade,⁹ and that the definition of “gazelles” may vary not only by country but also time.

4.2 Specialization and contribution of the “gazelle” companies

Having identified “gazelle” companies in Russia and Spain, we want to see how large these companies are and in which industries they were concentrated over

⁸ Note that for a 5% significance level $\mu + 2\delta$ would be sufficient, but we apply here a stricter threshold.

⁹ This result is in line with our Table 1 if we compare earlier results for the EU countries by Bravo-Biosca (2010) with the more recent ones by Flachenecker et al. (2020).

Table 9 Industries with highest concentration of “gazelles” in Russia and Spain over time

	Absolute growth		Relative growth	
	Russia	Spain	Russia	Spain
2011	1. Information and communication (19.5) 2. Research and development (12.5) 3. Wholesale and retail trade (10.0)	1. Construction (21.0) 2. Real estate (14.6) 3. Wholesale and retail trade (10.6)	1. Wholesale and retail trade (19.5) 2. Information and communication (12.4) 3. Research and development (7.8)	1. Construction (21.8) 2. Real estate (14.6) 3. Wholesale and retail trade (10.6)
2012	1. Wholesale and retail trade (13.2) 2. Information and communication (12.4) 3. Research and development (9.4)	1. Construction (18.4) 2. Real estate (15.4) 3. Wholesale and retail trade (12.3)	1. Wholesale and retail trade (13.2) 2. Information and communication (12.4) 3. Research and development (9.6)	1. Construction (18.4) 2. Real estate (15.4) 3. Wholesale and retail trade (12.6)
2013	1. Information and communication (18.3) 2. Transportation and storage (13.6) 3. Wholesale and retail trade (11.4)	1. Wholesale and retail trade (23.0) 2. Construction (14.0) 3. Real estate (12.6)	1. Information and communication (18.5) 2. Transportation and storage (13.9) 3. Wholesale and retail trade (11.4)	1. Wholesale and retail trade (23.0) 2. Construction (14.0) 3. Finance and insurance (12.6)
2014	1. Information and communication (20.6) 2. Transportation and storage (12.9) 3. Wholesale and retail trade (11.5)	1. Education (16.3) 2. Wholesale and retail trade (14.8) 3. Finance and insurance (11.8)	1. Information and communication (20.2) 2. Transportation and storage (12.5) 3. Wholesale and retail trade (11.5)	1. Education (17.8) 2. Wholesale and retail trade (14.7) 3. Finance and insurance (11.8)
2015	1. Information and communication (21.8) 2. Transportation and storage (19.5) 3. Wholesale and retail trade (15.4)	1. Construction (17.8) 2. Wholesale and retail trade (14.9) 3. Education (12.8)	1. Information and communication (21.8) 2. Transportation and storage (19.6) 3. Wholesale and retail trade (15.2)	1. Construction (17.8) 2. Wholesale and retail trade (14.9) 3. Education (12.8)
2016	1. Information and communication (27.3) 2. Transportation and storage (22.0) 3. Wholesale and retail trade (16.8)	1. Education (20.0) 2. Wholesale and retail trade (12.2) 3. Real estate (10.8)	1. Information and communication (26.6) 2. Transportation and storage (22.4) 3. Wholesale and retail trade (16.9)	1. Education (20.0) 2. Wholesale and retail trade (11.7) 3. Real estate (10.2)
2017	1. Information and communication (27.7) 2. Wholesale and retail trade (21.2) 3. Transportation and storage (12.3)	1. Education (18.8) 2. Wholesale and retail trade (15.8) 3. Real estate (13.3)	1. Information and communication (27.3) 2. Wholesale and retail trade (18.3) 3. Transportation and storage (13.5)	1. Education (18.7) 2. Wholesale and retail trade (15.7) 3. Real estate (10.9)
2018	1. Information and communication (27.6) 2. Wholesale and retail trade (21.2) 3. Transportation and storage (20.5)	1. Education (18.8) 2. Wholesale and retail trade (15.6) 3. Real estate (12.5)	1. Information and communication (27.2) 2. Transportation and storage (18.3) 3. Wholesale and retail trade (17.5)	1. Education (18.7) 2. Wholesale and retail trade (15.5) 3. Real estate (11.1)

Percentage of “gazelle” companies among all companies in the industry are in brackets

Table 10 Percentage of employment growth provided by the “gazelles”

	Absolute growth		Relative growth	
	Russia	Spain	Russia	Spain
2011	34	9	28	9
2012	18	13	23	13
2013	8	23	9	23
2014	10	5	10	5
2015	20	4	11	4
2016	9	5	9	5
2017	11	5	7	5
2018	10	6	10	6
Mean	15	8.75	13.37	8.75
Median	10.5	5.5	10	5.5

the period under consideration. From Table 8 we see that Russian “gazelles” are on average (and in median) larger than their Spanish counterparts, and that most of these companies are small and medium-sized enterprises (SME).

Now we calculate the percentage of “gazelle” companies for each industry and each year, and present in Table 9 three economic sectors with the highest concentration of “gazelles”. In other words, we show in which industries the “gazelles” are observed disproportionately more often. In Russia “gazelle” companies were concentrated mostly in the information and communication sector, wholesale and retail trade, and transportation and storage. A few times in 2011–2012 research and development sector was also among the top three most populated by the “gazelles”. In Spain, “gazelle” companies are relatively more common in wholesale and retail trade, construction, and real estate. In 2013–2014 financial and insurance activities and in 2015–2018 education were also among the sectors with the largest share of the “gazelles”. It is worth mentioning that the ranking of the countries is rather consistent for absolute and relative growth rates of firms stressing that our results are robust to the choice of the growth measure.

To assess the contribution of “gazelle” companies to the employment growth in the respective economies, we calculate the percentage of employment growth generated by the “gazelles” compared to the total sample of observations in our dataset. The results are presented in Table 10. While making up around 1–2% of all companies in our sample, the “gazelles” provide on average about 14% of new jobs in Russia and about 9% in Spain. Since among Russian companies one can more often find a “gazelle” among medium-sized or large firms (Table 8), they generate a larger share of employment growth in the economy compared to their counterparts in Spain. Furthermore, as we know that “gazelles” in Russia

measured in terms of absolute growth are on average smaller than those measured in relative growth, their total contribution to employment is rather similar, even though the number of “gazelles” in absolute growth is almost twice larger.

5 Conclusions

In this study we focus on the phenomenon of fast-growing “gazelle” companies that exhibit sustained employment growth over time. Such companies typically create a relatively large number of jobs and make a significant contribution to the development of their economies. In the literature many researchers tried to define these companies considering different countries and periods. As a result, many definitions have been given, while comparison of these companies over countries has been largely complicated. Furthermore, given earlier inconclusive results on the drivers of success of the “gazelle” companies, it has been argued that their success may be purely due to good luck.

To address these problems, we follow a different approach. Based on large firm-level datasets for Russian and Spanish economies for the period 2010–2018 we conduct a Monte Carlo experiment with first-order Markov chains to derive a definition of “gazelle” companies from the data and ensure that their existence cannot be explained by chance only. In doing this, we account for firm size, age, industry, year of observation and use both absolute and relative employment growth measures as an indicator of firm growth. As a definition of “gazelle”, we estimate the number of years a company should consequently belong to the top 10% fastest growing to keep the likelihood of a false positive results under 5%. The empirically observed numbers of “gazelle” firms significantly exceed the expected values derived from 1000 MCMC simulation runs rejecting the hypothesis that chance alone can explain the existence of “gazelles” in Russia and Spain. This implies that the identified “gazelles” must have competitive advantages, which may be reflected among others, in the form of better resources, technologies, management routines, partners along supply value chain that allow them to expand on the expense of their competitors and demonstrate sustained superior performance (Cantner et al., 2019; Eisenhardt & Martin, 2000; Savin & Egbe-tokun, 2016; Savin & Mundt, 2022; Teece, 2007).

We find that the “gazelle” firms account for about 1–2% of the total number of companies in our dataset and are responsible for approx. 14% of employment growth in Russia and 9% in Spain. Most of these companies are concentrated in economic sectors like wholesale and retail trade, information and communication, real estate, construction, and transportation and storage. The relatively small share of gazelles we identified (compared to 4–5% on average reported in the literature reviewed in Table 1) indicates that genuine “gazelles” are rare species. This can be interpreted as indirect support of arguments by Nightingale and Coad

(2013) that policy makers could be more selective in stimulating market entry and increasing number of startups in different sectors (Savin et al., 2022), since most of new businesses do not perform well, and subsidizing all of them may be counterproductive. Instead, policy should focus on fostering innovative startups that are more likely to generate considerable economic and societal impacts (Colombelli et al., 2016), and specifically on identifying and supporting high-growth firms. The latter task is not easy as one has to avoid following popular myths about “gazelles” (e.g., that they are concentrated in high tech industries) and think of time-specific and peer-based support measures, such as consultations and networking with experienced entrepreneurs in periods of market turbulence or business transition (for details, see Brown et al., 2017).

Note that the analysis presented here does not substitute but complements earlier studies searching for factors explaining sustained fast growth of companies. Traditional econometric analysis employed in the literature tries to establish a statistical association or a causal link between firm performance and other firm characteristics or external conditions. As we showed in our literature review, some studies failed to find any particular features of “gazelles” (except for their growth), but they cannot rule out the possibility that some variables omitted in the analysis have a role in explaining this exceptional performance. Nor can these studies ensure that the exceptional performance of some companies can be pure luck only. While our results do not specify factors explaining sustained firm growth, they demonstrate which firms can be called “gazelles” and therefore show that factors justifying their success exist with certainty. Finding these factors is a task for future research, which would allow policy makers to derive programs to foster the share of “gazelle” companies and, thus, stimulate employment and economic growth.

Our study has important implication for the theory and practice on analyzing high-growth firms. As already noted by Churchill and Lewis (1983), small businesses vary widely in the problems they encounter and the growth patterns they exhibit. Therefore, it is not so surprising that we find that data-driven definition of “gazelles” that rules out chance is sensitive to the country and period it is applied to. Future theoretical research on “gazelle” companies should account for that by avoiding simple “one size fits all” definitions and being more attentive to the context in which these firms are analyzed. Regarding empirical research, we hope our results will stimulate more work on data-driven definitions of “gazelle” companies as it will be interesting to see how much variation these definitions can produce for companies in Asia, South America, and Africa.

Appendix A: Further details on the data and methods used

Firm size is a categorical variable introduced by BvD ranging between 1 and 4. Firm with value 4 (very large company) should match at least one of the following conditions:

- Operating Revenue \geq 100 million EUR (140 million USD),
- Total assets \geq 200 million EUR (280 million USD),
- Employees \geq 1000,
- Listed.

Firm with value 3 (large company) should match at least one of the following conditions:

- Operating Revenue \geq 10 million EUR (14 million USD),
- Total assets \geq 20 million EUR (28 million USD),
- Employees \geq 150,
- Not very large.

Firm with value 2 (medium sized company) should match at least one of the following conditions:

- Operating Revenue \geq 1 million EUR (1.4 million USD),
- Total assets \geq 2 million EUR (2.8 million USD),
- Employees \geq 15,
- Not very large or large

Firm with value 1 (small company) is a company which is not included in other categories.

Appendix B: Further results

See Figs. 5, 6, 7, 8, Tables 11, 12.

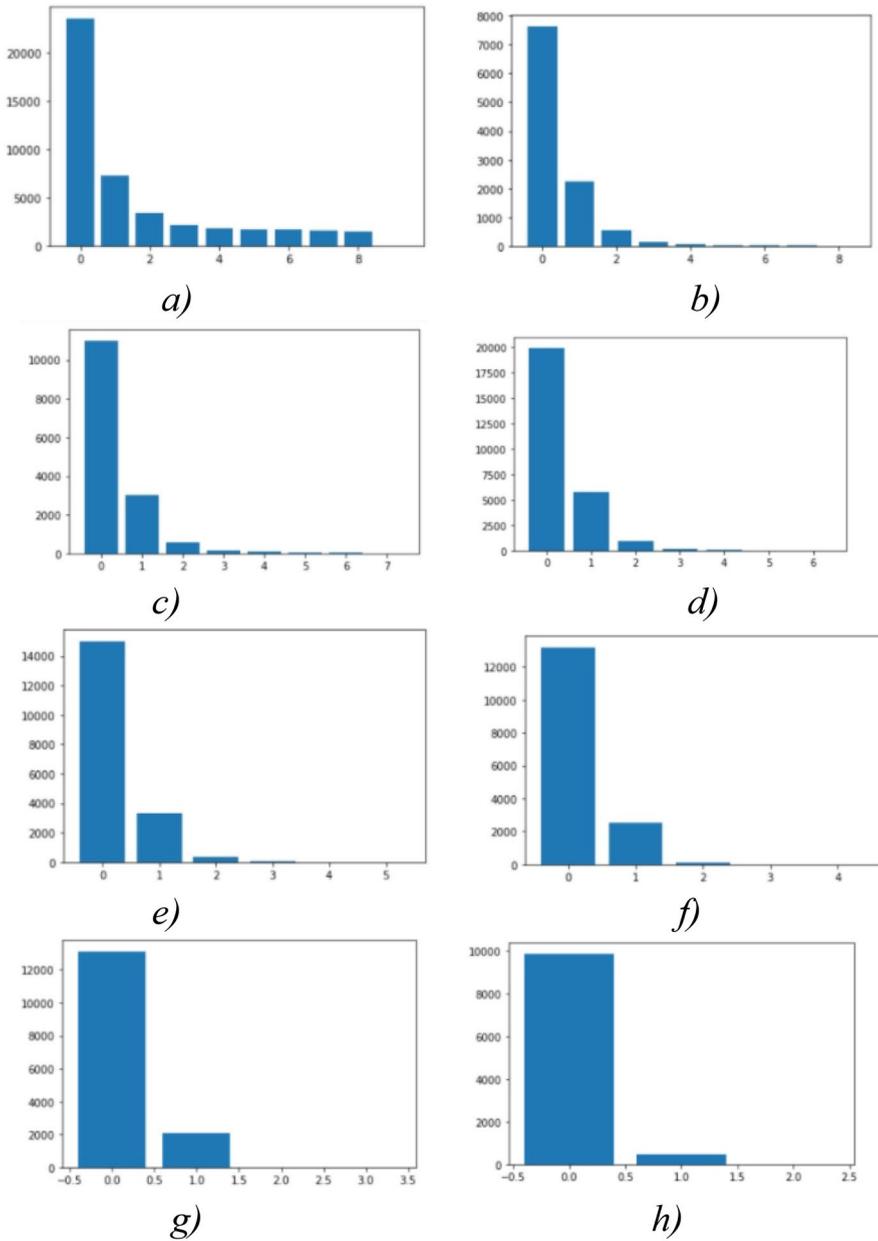


Fig. 5 Number of Russian companies belonging to the top 10% of fast-growing companies in terms of absolute growth over **a** 9 years, **b** 8 years, **c** 7 years, **d** 6 years, **e** 5 years, **f** 4 years, **g** 3 years, **h** 2 years

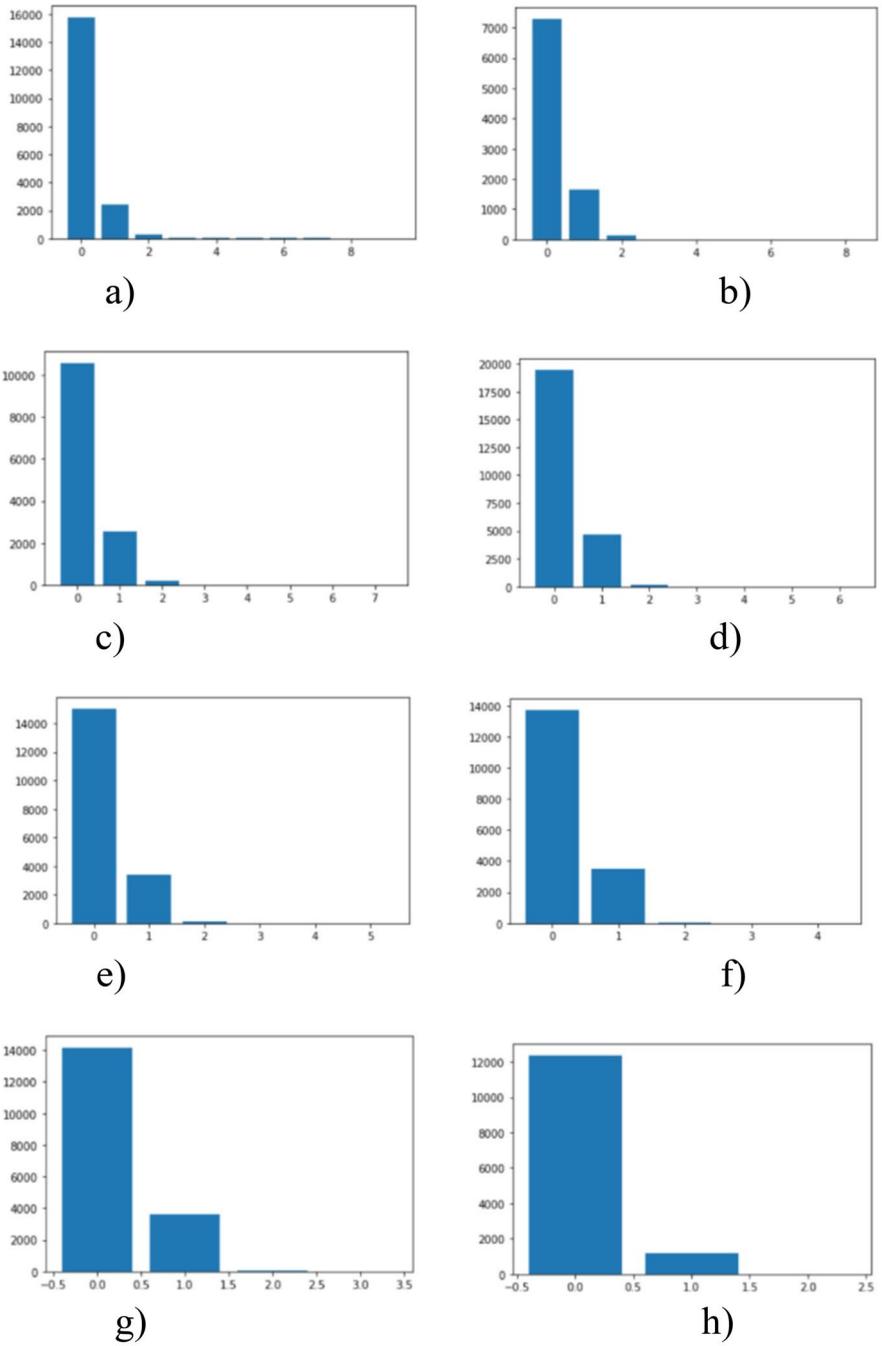


Fig. 6 Number of Russian companies belonging to the top 10% of fast-growing companies in terms of relative growth over **a** 9 years, **b** 8 years, **c** 7 years, **d** 6 years, **e** 5 years, **f** 4 years, **g** 3 years, **h** 2 years

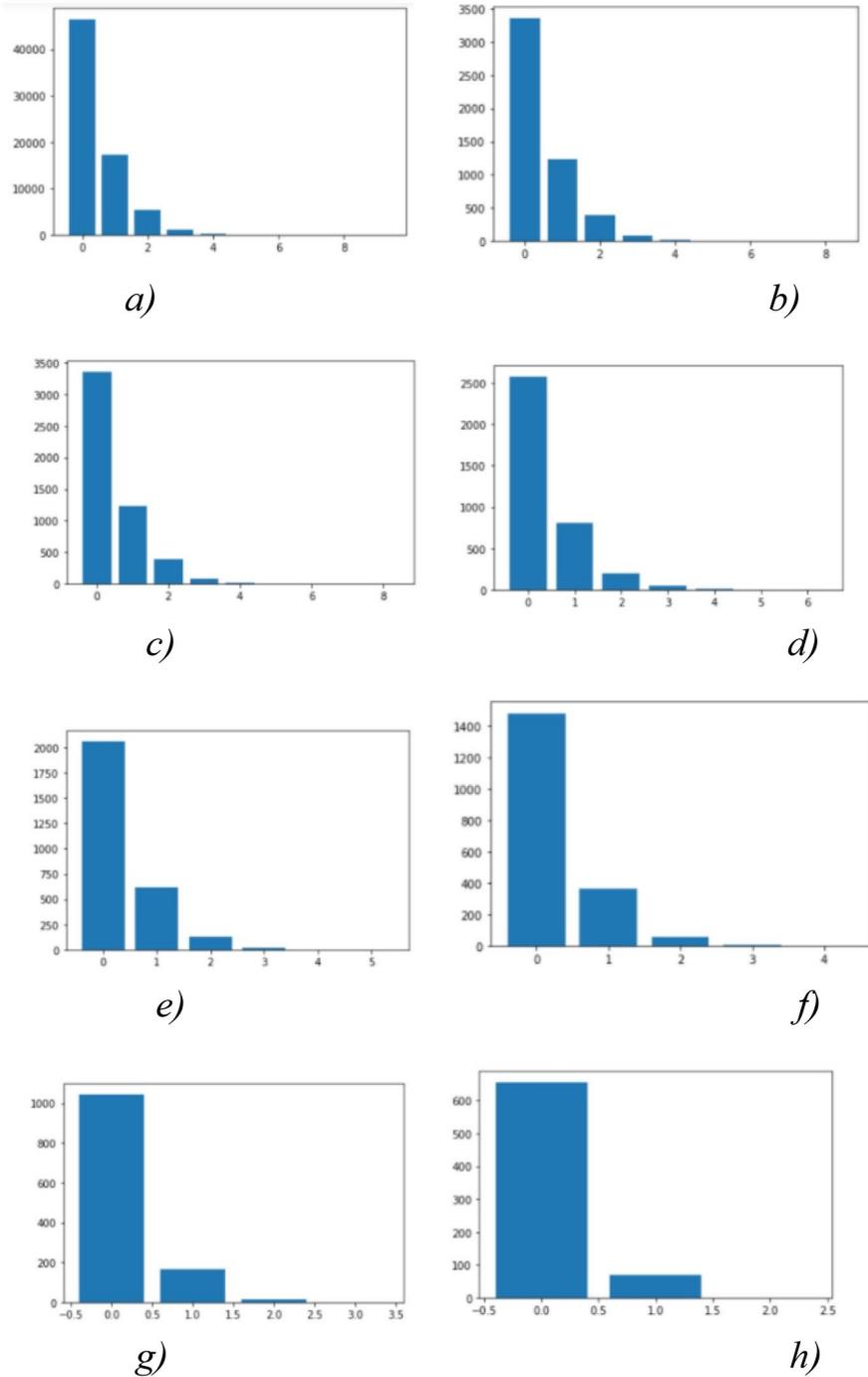
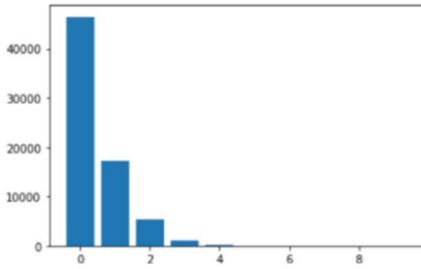
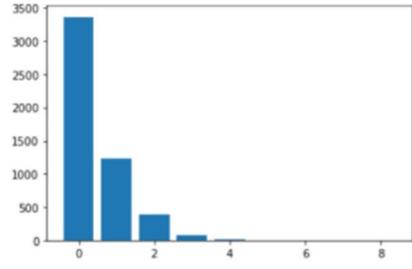


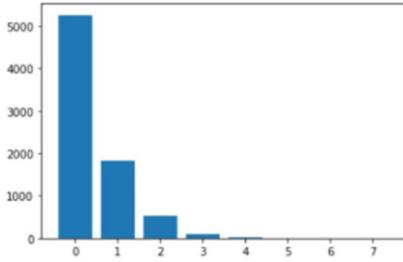
Fig. 7 Number of Spanish companies belonging to the top 10% of fast-growing companies in terms of absolute growth over **a** 9 years, **b** 8 years, **c** 7 years, **d** 6 years, **e** 5 years, **f** 4 years, **g** 3 years, **h** 2 years



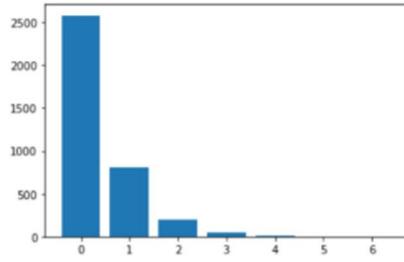
a)



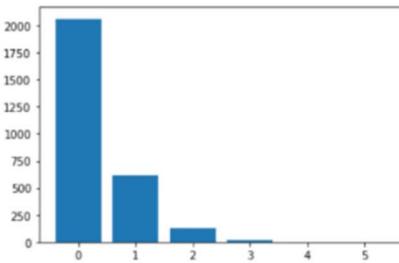
b)



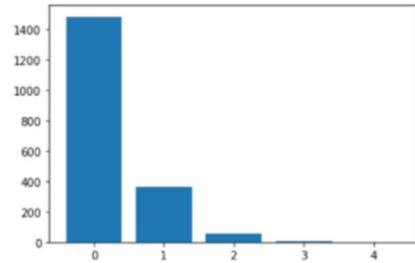
c)



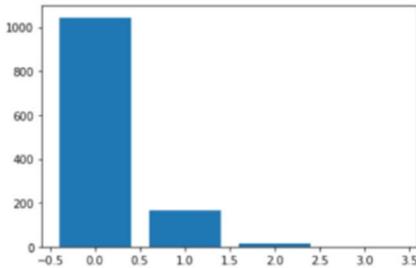
d)



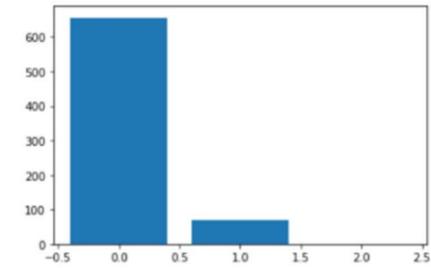
e)



f)



g)



h)

Fig. 8 Number of Spanish companies belonging to the top 10% of fast-growing companies in terms of relative growth over **a** 9 years, **b** 8 years, **c** 7 years, **d** 6 years, **e** 5 years, **f** 4 years, **g** 3 years, **h** 2 years

Table 11 Number of years in the top 10% of fast-growing companies to be classified as a “gazelle”. Comparison of our results with those by Korzinov (2018) for Spain

	Absolute growth		Absolute growth	
	Our results	Korzinov (2018)	Our results	Korzinov (2018)
2	1	1	1	1
3	1	1	1	1
4	2	1	2	1
5	2	2	2	2
6	2	2	2	2
7	3	2	3	2
8	3	–	3	–
9	3	–	3	–

Differences are marked with bold

Table 12 Number of “gazelle” companies. Comparison of our results with those by Korzinov (2018) for Spain

	Absolute growth		Relative growth	
	Our results	Korzinov (2018)	Our results	Korzinov (2018)
Average number of “gazelle” companies expected from the MCMC experiment	856 (15.2)	2209.7 (41.0)	847 (14.8)	2155.4 (50.3)
Observed number of “gazelle” companies in % of the total number of companies	1.2	1.0	1.2	1.0
Observed number of “gazelle” companies	2333	4295	2322	4110

Standard deviations (δ) are in circle brackets

Acknowledgements Ivan Savin acknowledges support from the Russian Science Foundation (RSF grant number 19-18-00262). We thank Vladimir Korzinov and Philipp Mundt for many helpful comments and suggestions on the earlier drafts of the paper. The usual disclaimer applies.

Funding Open Access Funding provided by Universitat Autònoma de Barcelona.

Data availability Data and code are available upon request.

Declarations

Ethics approval and consent to participate This research did not involve Human Participants and/or Animals.

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the

material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abbate, C., & Sapio, A. (2019). Gazelles and muppets in the city: Risk sharing and firm growth quantiles in a junior stock market. *Industrial and Corporate Change*, 28(6), 1405–1427. <https://doi.org/10.1093/icc/dtz024>
- Acs, Z. J., & Mueller, P. (2008). Employment effects of business dynamics: Mice, gazelles and elephants. *Small Business Economics*. <https://doi.org/10.1007/s11187-007-9052-3>
- Ahmad, N. (2006). A proposed framework for business demographic statistics. OECD Statistics Working Paper Series. <https://doi.org/10.1787/18152031>.
- Autio, E., Arenius, P., Wallenius, H. (2000). Economic impact of gazelle firms in Finland. Helsinki University of Technology. ISIB Working Papers 3.
- Barba-Sánchez, V., Arias-Antúnez, E., & Orozco-Barbosa, L. (2019). Smart cities as a source for entrepreneurial opportunities: Evidence for Spain. *Technological Forecasting and Social Change*, 148, 119713. <https://doi.org/10.1016/j.techfore.2019.119713>
- Barney, J. B. (1986). Strategic factor markets: Expectations, luck, and business strategy. *Management Science*, 32(10), 1231–1241.
- Becchetti, L., & Trovato, G. (2002). The determinants of growth for small and medium sized firms. *The Role of the Availability of External Finance*, *Small Business Economics*, 19, 291–306. <https://doi.org/10.1023/A:1019678429111>
- Bianchini, S., Bottazzi, G., & Tamagni, F. (2017). What does (not) characterize persistent corporate high-growth? *Small Business Economics*, 48, 633–656. <https://doi.org/10.1007/s11187-016-9790-1>
- Birch, D. L. (1987). *Job creation in America*. Free.
- Birch, D. L., Hagerty, A., & Parsons, W. (1995). *Who's creating jobs?* Cognetics Inc.
- Birch, D. L., & Medoff, J. (1994). Gazelles. In L. C. Solmon & A. R. Levenson (Eds.), *Labor markets, employment policy and job creation* (pp. 159–167). Boulder: Westview.
- Bottazzi, G., Coad, A., Jacoby, N., & Secchi, A. (2011). Corporate growth and industrial dynamics: Evidence from French manufacturing. *Applied Economics*, 43(1), 103–116. <https://doi.org/10.1080/00036840802400454>
- Bravo-Biosca A., (2010). *Growth dynamics: exploring business growth and contraction in Europe and the US*. NESTA research report, Nov 2010. NESTA: London
- Brown, R., Mawson, S., & Mason, C. (2017). Myth-busting and entrepreneurship policy: The case of high growth firms. *Entrepreneurship & Regional Development*, 29(5–6), 414–443. <https://doi.org/10.1080/08985626.2017.1291762>
- Brüderl and Preisendörfer. (2000). Fast-growing businesses: Empirical evidence from a German study. *International Journal of Sociology*, 30(3), 45–70. <https://doi.org/10.1080/15579336.2000.11770218>
- Cabral, L., & Mata, J. (2003). On the evolution of firm size distribution: Facts and theory. *The American Economic Review*, 93(4), 1075–1095. <https://doi.org/10.1257/000282803769206205>
- Cantner, U., Savin, I., & Vannuccini, S. (2019). Replicator dynamics in value chains: Explaining some puzzles of market selection. *Industrial and Corporate Change*, 28(3), 589–611. <https://doi.org/10.1093/icc/dty060>
- Capasso, M., Cefis, E., & Frenken, K. (2014). On the existence of persistently outperforming firms. *Industrial and Corporate Change*, 23(4), 997–1036. <https://doi.org/10.1093/icc/dtt034>
- Churchill, N. C., & Lewis, V. L. (1983). Five Stages of Small Business Growth. *Harvard Business Review*, 61(3), 30–51.
- Ciriaci, D., Moncada-Paternò-Castello, P., & Voigt, P. (2016). Innovation and job creation: A sustainable relation? *Eurasian Business Review*, 6, 189–213. <https://doi.org/10.1007/s40821-015-0031-3>
- Chigrin, A. (2017). Small business and the competitiveness of Russia. *An Unconventional View, Journal of the New Economic Association*, 3(39), 110–126. (in Russian).

- Coad, A., Frankish, J., Roberts, R. G., & Storey, D. J. (2013). Growth paths and survival chances: An application of Gambler's Ruin theory. *Journal of Business Venturing*, 28(5), 615–632. <https://doi.org/10.1016/j.jbusvent.2012.06.002>
- Coad, A., & Rao, R. (2008). Innovation and firm growth in high tech sectors: A quantile regression approach. *Research Policy*, 37(4), 633–648. <https://doi.org/10.1016/j.respol.2008.01.003>
- Coad, A., Daunfeldt, S. O., Hölzl, W., Johansson, D., & Nightingale, P. (2014). High-growth firms: Introduction to the special section. *Industrial and Corporate Change*, 23(1), 91–112. <https://doi.org/10.1093/icc/dtt052>
- Colombelli, A., Krafft, J., & Quattraro, F. (2014). High-growth firms and technological knowledge: Do gazelles follow exploration or exploitation strategies? *Industrial and Corporate Change*, 23(1), 261–291. <https://doi.org/10.1093/icc/dtt053>
- Colombelli, A., Krafft, J., & Vivarelli, M. (2016). To be born is not enough: The key role of innovative startups. *Small Business Economics*, 47(2), 277–291. <https://doi.org/10.1007/s11187-016-9716-y>
- Comin D., Philippon T. (2005). *The rise in firm-level volatility: causes and consequences*, NBER Working Paper. (pp. 11388)
- Delmar, F. (1997). Measuring growth: Methodological considerations and empirical results. In R. Donckels & A. Miettinen (Eds.), *Entrepreneurship and SME research: On its way to the next millenium* (pp. 190–216). Aldershot: Avebury.
- Delmar, F., Davidsson, P., & Gartner, W. B. (2003). Arriving at the high-growth firm. *Journal of Business Venturing*, 18(2), 189–216. [https://doi.org/10.1016/S0883-9026\(02\)00080-0](https://doi.org/10.1016/S0883-9026(02)00080-0)
- Deschryvere, M. (2008). *High-growth firms and job creation in Finland*. Discussion Paper 1144, Helsinki: Research Institute of the Finnish Economy (ETLA)
- Dosi, G., Moschella, D., Pugliese, E., & Tamagni, F. (2015). Productivity, market selection, and corporate growth: Comparative evidence across US and Europe. *Small Business Economics*, 45, 643–672. <https://doi.org/10.1007/s11187-015-9655-z>
- Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: What are they? *Strategic Management Journal*, 21(10–11), 1105–1121.
- Erhardt, E. C. (2021). Measuring the persistence of high firm growth: Choices and consequences. *Small Business Economics*, 56(1), 451–478. <https://doi.org/10.1007/s11187-019-00229-7>
- Esteve-Pérez, S., Pieri, F., & Rodriguez, D. (2022). One swallow does not make a summer: Episodes and persistence in high growth. *Small Business Economics*, 58, 1517–1544. <https://doi.org/10.1007/s11187-020-00443-8>
- Evans, D. E. (1987). The relationship between firm growth, size and age: Estimates from 100 manufacturing industries. *Journal of Industrial Economics*, 35(4), 567–581.
- Feindt, S., Jeffcoate, J., & Chappell, C. (2002). Identifying Success Factors for Rapid Growth in SME E-commerce. *Small Business Economics*, 19, 51–62. <https://doi.org/10.1023/A:1016165825476>
- Flachenecker, F., Gavigan, J. P., Goenaga, X., Pasi, G., Preziosi, N., Stamenov, B., & Testa, G. (2020). High Growth Enterprises: demographics, financing & policy measures. JRC Technical Report. Joint Research Centre. Brussels, Belgium. <https://ec.europa.eu/jrc/en/publication/high-growth-enterprises-demographics-finance-policy-measures>
- Gibrat R. (1931). Les inégalités économiques: applications, aux inégalités des richesses, a la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc.: d'une loi nouvelle la loi de l'effet proportionnel Recueil Sirey, Paris (1931).
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17(3), 295–314. [https://doi.org/10.1016/0010-0285\(85\)90010-6](https://doi.org/10.1016/0010-0285(85)90010-6)
- Greene, W. H. (1993). *Econometric Analysis* (2nd ed.). New York.
- Guilmi, R., & Mokhtari, B. (2016). The phenomenon of high growth SMEs (gazelles): A systematic review of literature and implications for further research (Part I). *International Journal of Economics & Strategic Management of Business Process*, 45, 100–124.
- Halabisky, D., Dreessen, E., & Parsley, C. (2006). Growth in firms in Canada, 1985–1999. *Journal of Small Business and Entrepreneurship*, 19(3), 255–268. <https://doi.org/10.1080/08276331.2006.10593370>
- Henrekson, M., & Johansson, D. (2010). Gazelles as job creators: A survey and interpretation of the evidence. *Small Business Economics*, 35, 227–244. <https://doi.org/10.1007/s11187-009-9172-z>
- Henderson, A. D., Raynor, M. E., & Ahmed, M. (2012). How long must a firm be great to rule out chance? Benchmarking sustained superior performance without being fooled by randomness. *Strategic Management Journal*, 33(4), 387–406. <https://doi.org/10.1002/smj.1943>

- Hölzl, W. (2014). Persistence, survival, and growth: A closer look at 20 years of fast-growing firms in Austria. *Industrial and Corporate Change*, 23(1), 199–231. <https://doi.org/10.1093/icc/dtt054>
- Korzinov V. (2018). *On technology emergence and pattern detection in aggregate innovative activity*. PhD thesis, Karlsruhe Institute of technology. <https://doi.org/10.5445/IR/1000084152>.
- Lotti, F., Santarelli, E., & Vivarelli, M. (2003). Does Gibrat's law hold among young, small firms? *Journal of Evolutionary Economics*, 13, 213–235. <https://doi.org/10.1007/s00191-003-0153-0>
- Lotti, F., Santarelli, E., & Vivarelli, M. (2009). Defending Gibrat's law as a long-run regulatory. *Small Business Economics*, 32, 31–44. <https://doi.org/10.1007/s11187-007-9071-0>
- McKelvey, B., & Andriani, P. (2005). Why Gaussian statistics are mostly wrong for strategic organization. *Strategic Organization*, 3(2), 219–228. <https://doi.org/10.1177/1476127005052700>
- Mogos, S., Davis, A., & Baptista, R. (2021). High and sustainable growth: Persistence, volatility, and survival of high growth firms. *Eurasian Business Review*, 11, 135–161. <https://doi.org/10.1007/s40821-020-00161-x>
- Moreno, F., Coad, A. (2015). *High-growth firms: Stylized facts and conflicting results*. In: Entrepreneurial growth: individual, firm, and region. Emerald Group Publishing Limited, pp 187–230. <https://doi.org/10.1108/S1074-754020150000017016>
- Nightingale, P., & Coad, A. (2014). Muppets and gazelles: Political and methodological biases in entrepreneurship research. *Industrial and Corporate Change*, 23(1), 113–143. <https://doi.org/10.1093/icc/dtt057>
- Pennacchio, L., Sapio, A. (2020). *Financing, selection, and value-adding effects of venture capital: a review of econometric methods and issues*. In: Handbook of Quantitative Research Methods in Entrepreneurship, Saridakis G. and Cowling M. (Eds.) <https://doi.org/10.4337/9781786430960.00016>
- Pirogov, N. K., & Popovidchenko, M. G. (2010). Testing Gibrat's law: A review of studies. *Korporativnye Finansy*, 1(13), 16–119. (in Russian).
- Pletnev, D., & Barkhatov, V. (2019). Ups and Downs of High-Growth Firms in Russia. In G. Prosteian, J. J. Lavios Villahoz, L. Brancu, & G. Bakacsi (Eds.), *International Symposium in Management Innovation for Sustainable Management and Entrepreneurship* (pp. 127–136). Berlin: Springer.
- Pugliese, R., Bortoluzzi, G., & Balzano, M. (2021). What drives the growth of start-up firms? A Tool for Mapping the State-of-the-Art of the Empirical Literature. *European Journal of Innovation Management*, 25(6), 242–272. <https://doi.org/10.1108/EJIM-03-2021-0163>
- Rocha, R. G., & Ferreira, J. J. (2021). Gazelles (high-growth) companies: A bibliometric science map of the field. *Journal of the Knowledge Economy*, Forthcoming. <https://doi.org/10.1007/s13132-021-00828-4>
- Roney, C. J., & Trick, L. M. (2009). Sympathetic magic and perceptions of randomness: The hot hand versus the gambler's fallacy. *Thinking & Reasoning*, 15(2), 197–210. <https://doi.org/10.1080/13546780902847137>
- Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using pyMC. *PeerJ Computer Science*, 3(2), e55. <https://doi.org/10.7717/peerj-cs.55>
- Savin, I. V., Mariev, O. S., & Pushkarev, A. A. (2019). Survival of the fittest? Measuring the strength of market selection on the example of the Urals Federal District. *The HSE Economic Journal*, 23(1), 90–117. <https://doi.org/10.17323/1813-8691-2019-23-1-90-117>. (In Russian).
- Savin, I. V. (2020). Studying market selection in Russia and abroad: Measurement problems, national specificity and stimulating methods. *Journal of the New Economic Association*, 48(4), 197–204. <https://doi.org/10.31737/2221-2264-2020-48-4-9>. (In Russian).
- Savin, I., Chukavina, K., & Pushkarev, A. (2022). Topic-based classification and identification of global trends for startup companies. *Small Business Economics*. <https://doi.org/10.1007/s11187-022-00609-6>
- Savin, I., & Egbetokun, A. (2016). Emergence of innovation networks from R&D cooperation with endogenous absorptive capacity. *Journal of Economic Dynamics and Control*, 64, 82–103. <https://doi.org/10.1016/j.jedc.2015.12.005>
- Savin, I. V., Mariev, O. S., & Pushkarev, A. A. (2020). Measuring the strength of market selection in Russia: When the (firm) size matters. *Voprosy Ekonomiki*, 2, 101–124. <https://doi.org/10.32609/0042-8736-2020-2-101-124>. (In Russian).
- Savin, I., & Mundt, P. (2022). Drivers of productivity change in global value chains: reallocation vs. innovation. *Economic Letters*, 220, 110878. <https://doi.org/10.1016/j.econlet.2022.110878>
- Schreyer, P. (2000). High-growth firms and employment. OECD Science, Technology and Industry Working Papers, 2000/3, Paris. <https://doi.org/10.1787/861275538813>

- St-Jean, É., Julien, P., & Audet, J. (2008). Factors associated with growth changes in “Gazelles.” *Journal of Enterprising Culture*, 16(2), 161–188. <https://doi.org/10.1142/S0218495808000089>
- Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, 28(13), 1319–1350. <https://doi.org/10.1002/smj.640>
- The Economist. (2012). Les misérables, 2012. <https://www.economist.com/briefing/2012/07/28/les-miserables>
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76, 105–110. <https://doi.org/10.1037/h0031322>
- Vértesy, D., Del Sorbo, M., & Damioli, G. (2017). *High-growth, innovative enterprises in Europe*. EUR 28606 EN. Luxembourg: Publications Office of the European Union. <https://doi.org/10.2760/328958>
- Voigt, P., & Moncada-Paternò-Castello, P. (2012). Can Fast Growing R&D-Intensive Smes Affect the Economic Structure of the Eu Economy? A Projection to the Year 2020. *Eurasian Business Review*, 2, 96–128. <https://doi.org/10.14208/BF03353814>
- Waring, G. F. (1996). Industry differences in the persistence of firm-specific returns. *The American Economic Review*, 86, 1253–1265.
- Yasuda, T. (2005). Firm growth size, age and behavior in Japanese manufacturing. *Small Business Economics*, 24(1), 1–15.
- Yudanov, A. (2007). Fast growing firms (“gazelles”) and the evolution of Russian economy. *Voprosy Ekonomiki*, 2, 85–100. <https://doi.org/10.32609/0042-8736-2007-2-85-100>
- Yudanov, A., & Yakovlev, A. A. (2018). “Unorthodox” fast-growing firms (gazelles) and North’s limited access order. *Voprosy Ekonomiki*, 3, 80–101. <https://doi.org/10.32609/0042-8736-2018-3-80-101>. (In Russian).
- Yudanov, A. (2010). Embodiments of Entrepreneurial Spirit: «Gazelle-Firms» in Russia. *Journal of the New Economic Association*, 5, 92–109. (In Russian).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.