ISIGrowth

INNOVATION-FUELLED, SUSTAINABLE, INCLUSIVE GROWTH

Working Paper

Growth and survival of the 'fitter'? Evidence from US new-born firms

Giovanni Dosi
Scuola Superiore Sant'Anna

Emanuele Pugliese
National Reseacrh Council (CNR)

Pietro Santoleri
Scuola Superiore Sant'Anna

3A/2017 February

This project has received funding from the European Union Horizon 2020 Research and Innovation action under grant agreement No 649186

INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM

WORKING PAPER SERIES

Growth and survival of the 'fitter'? Evidence from US new-born firms

Giovanni Dosi °
Emanuele Pugliese §
Pietro Santoleri °

° Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy
§ National Reseacrh Council (CNR), Florence, Italy

2017/06 February 2017
ISSN(ONLINE) 2284-0400

Growth and survival of the 'fitter'? Evidence from US new-born firms

Giovanni Dosi * Emanuele Pugliese† Pietro Santoleri‡

Abstract

We examine market selection mechanisms and their strength for a representative cohort of US new independent firms. In particular, we explore whether and how effectively markets reward newly-born firms according to their 'fitness' in terms of both labour productivity and profitability. Our analysis yields puzzling results in contrast with canonical industry dynamics models. First, we find that selection on differential growth is mainly related to productivity while profitability plays a negligible role. Second, in contrast with the *growth of the fitter* principle, selection appears to be driven by *changes* in firms' relative productivity. Third, we explore how new firms' relative fitness affects their growth performance in different sectors. Our results reveal that market selection operates quite differently across them with higher incidence for new-born firms in services, low-tech and less concentrated sectors. Fourth, concerning selection via exit, our results support the *survival of the fitter* principle with respect to productivity, while relative profitability does not seem to exert any significant effect on survival probabilities. However, the contribution of firm relative 'fitness' to the total firm exit rates variation appears to be modest.

Keywords Market selection · Replicator dynamics · New firm growth · Survival · Shapley decomposition

JEL L11 · L25 · M13

*Institute of Economics, Scuola Superiore Sant'Anna – g.dosi@sssup.it

†National Research Council - Institute for Complex Systems – emanuele.pugliese@gmail.com

‡Institute of Economics, Scuola Superiore Sant'Anna – p.santoleri@sssup.it

[1] **Acknowledgments:** We would like to thank the Ewing Marion Kauffman Foundation for granting us the access to the confidential version of the KFS database and, in particular, Alicia Robb and Joseph Farhat. We are very much indebted with Daniel Lee at NORC for his precious and patient assistance with the data. Support by the ISIGrowth Project (N. 649186) within the Horizon 2020 Programme of the European Union is gratefully acknowledged. Thanks to Marco Capasso, Alex Coad, Nanditha Mathew, Lorenzo Napolitano, Emanuele Russo, Caterina Santi and several participants to the 16th International Schumpeter Society for insightful suggestions. The usual disclaimer applies.

1 Introduction

Competitive selection is generally considered to be one of the core drivers of aggregate economic growth: by weeding out inefficient firms through competition, markets should enhance overall productivity. Indeed, based on neo-classical (Jovanovic, 1982; Hopenhayn, 1992; Ericson and Pakes, 1995) and evolutionary perspectives (Winter and Nelson, 1982; Silverberg et al., 1988; Dosi et al., 1995), one would expect markets to select the 'fitter' firms via differential growth and via exit. In other words, superior firms are expected to gain market shares, while under-performing ones would decrease their participation and, eventually, exit the market. However, recent contributions investigating these market selection mechanisms have found them to be less effective than expected (Bottazzi et al., 2008, 2010, 2011; Bottazzi and Secchi, 2012; Coad, 2007; Dosi et al., 2014, 2015; Federico and Capelleras, 2015). In this paper, we provide new evidence on the strength of market selection forces by focusing on new-born firms. We contribute to the existing literature in several dimensions.

First, most studies focusing on market selection base their evidence on samples of established manufacturing firms. On the contrary, mainly for lack of adequate data, there is a scarcity of empirical works concerning new firms, with only few exceptions (Delmar et al., 2013; Federico and Capelleras, 2015). The investigation of how effective markets are in rewarding or punishing entrants according to their relative 'fitness' represents a compelling topic given that these firms have been shown to play a crucial role for aggregate job creation and, possibly, labour productivity growth, in both the US (Haltiwanger et al., 2015) and other OECD countries (Anyadike-Danes et al., 2015; Criscuolo et al., 2016). Moreover, given that the flow of entrants in the service sector in the US outnumbers manufacturing new firms around 8 to 1 (Kim et al., 2006), focusing only on the latter seriously hampers the aggregate inferences[1]. Here, we base our evidence on the Kauffman Firm Survey (KFS), a dataset representative of the whole population of independent businesses born in 2004 in the US which are tracked for their first 8 years of life.

Second, distinct from the literature investigating market selection using decomposition methodologies[2], we join a small ensemble of studies which resort to the direct estimation of the relationship between firm relative 'fitness' (defined in terms of productivity and profitability) and growth rates through firm-level regressions (Bottazzi et al., 2010; Dosi et al., 2014, 2015). We share with these works the aim of assessing the strength of market selection mechanisms. In particular, we estimate the relative importance of labour productivity and profitability in "explaining" firm growth and survival by applying the Shapley value technique to offer a measure of the marginal contribution of individual and group of regressors to the overall explanatory power of a model (Huettner et al., 2012).

Third, we contribute to the existing literature by examining both components of selection, namely, selection via growth differentials and via exit. Indeed, previous literature has traditionally focused on the former, while rarely these two aspects have been examined jointly (Delmar et al., 2013). While the simplest version of the theories predicts selection forces to be symmetrical, that is, unfit firms exit and fit firms grow, the empirical evidence is still scant concerning new firms.

Our analysis yields a series of challenging findings at odds with the predictions stemming from industry dynamics models. The first one is that, in contrast to the *growth of the fitter* notion, while selection in terms of differential growth appears to be related to labour productivity, profitability plays only a negligible role. Furthermore, selection on differential growth appears to be mainly driven by *changes* in firms' relative 'fitness' rather than relative 'fitness' levels. Indeed, labour productivity and profitability *changes* (and not productivity and profitability levels) are what actually shape firm growth dynamics. We also explore how fitness variables affect new firms' performance in different industries. Our findings reveal that selection appears to be stronger in services, low-tech and less concentrated sectors. Finally, concerning the survival analysis, we do find support for the *survival of the fitter* conjecture in terms of labour productivity while profitability does not seem to exert any significant effect. However, the marginal contributions of the productivity regressors to the total explained variance is modest.

The remainder of the paper is as follows. In the next section, we review a few incumbent contributions to the field. Sections 3 provides an overview of the data and descriptive statistics. In Section 4, we carry out the *growth of the fitter* analysis while Section 5 deals with the *survival of the fitter* analysis. Next, we present some robustness checks, followed by an overall discussion of the findings.

[1] The few studies providing empirical evidence on market selection in the case of new firms are based on samples of manufacturing (Bellone et al., 2008; Federico and Capelleras, 2015) or knowledge-intensive industries (Delmar et al., 2013).

[2] See Hyytinen and Maliranta (2013) for recent empirical evidence on young firms using decomposition methodologies.

2 Related Literature

Empirical investigations relying on longitudinal micro-data have been providing a rich account of industrial dynamics over the last few decades. One of the most robust stylized facts emerging from this strand of literature is the heterogeneity across firms which has been shown to apply independently of the degree of industrial disaggregation and to be very stable over time despite the competition process (Bartelsman and Dhrymes, 1998; Dosi, 2007; Bottazzi et al., 2008). These ubiquitous and persistent differences relate to both corporate characteristics (e.g. size, efficiency, innovativeness, organizational structures, propensity to grow, etc.) and performances (e.g. growth rates, profitability and survival chances). Therefore, research has been focusing on the sources of corporate heterogeneities and tried to disentangle whether there are regularities between corporate features and their performances (Dosi et al., 2010).

From a theoretical standpoint, a series of models have been addressing market selection forces depicting industries as collections of heterogeneous producers explicitly linking their efficiency to their performances and survival chances (Foster et al., 2008). A distinction can be made among 'evolutionary equilibria' models which assume rational expectations (Jovanovic, 1982; Ericson and Pakes, 1995; Hopenhayn, 1992) and models such as Winter and Nelson (1982), Silverberg et al. (1988), Dosi et al. (1995) and Winter et al. (2003) in which the industrial dynamics are the result of learning and selection among boundedly rational agents (Dosi et al., 1995, 2015, 2016).

Within the first group, according to Jovanovic (1982) firms are born with fixed (but unknown) productivity levels and discover them through a process of Bayesian learning from their post-entry profits over time: firms decide to grow when they realize they are efficient, and contract (or exit) when they 'learn' they are not. Consequently, the cohort's average productivity increases as the cohort ages, even if the productivity of individual firms remain constant over time. Contrarily, Ericson and Pakes (1995) allow firms to know their productivity and to change it over time either as the stochastic outcome of their investments or as the result of changes in market conditions. In turn, Hopenhayn (1992) considers the key role played by industry-specific effects within a competitive industry in stationary equilibrium.

In the above family of models, agents feature profit-maximizing behaviour over an infinite time horizon and in each period, based on their equilibrium size, they decide whether to stay in the market or not. Hence, the selection process is never at work since it is anticipated in the heads of the agents. Moreover, these models do not account for the endogeneity of the growth process (Metcalfe, 1994) and for the persistent profitability and growth differentials across firms since, in equilibrium, these should be eroded by competition (Dosi et al., 1995).

The second group of models draws upon the evolutionary framework. According to this approach, firm dynamics are the outcome of learning and market selection between boundedly rational, interacting agents. On the one hand, learning and innovation are endogenous and idiosyncratic to the firm (Dosi et al., 1995). Superior ability in generating and exploiting new knowledge is the primary source of competitive advantage (Metcalfe, 1994) thus allowing for the presence of abiding heterogeneity in firms' 'identity cards' and performances. On the other, the process of competitive selection involves the retainment of firms featuring higher efficiency and profitability with respect to their competitors.

In evolutionary models, these two main drivers of industry dynamics are often represented by means of a 'replicator dynamics' through which growth is imputed according to firm competitiveness (or 'fitness') relative to its environment. While this is implicit in the tradition of Winter and Nelson (1982)[3], in Silverberg et al. (1988) and Dosi et al. (1995), among others, market selection is depicted explicitly in terms of 'replicator dynamics' of the form:

$$\dot{m}_i = \frac{\mathrm{d}m_i}{\mathrm{d}t} = m_i \left(f_i - \frac{\sum f_i}{N} \right)$$

where m_i refers to the market share of firm i in its industry and f_i is the fitness of firm i. Based on this equation, selection forces operate accurately by rewarding 'fitter' firms (in terms of superior efficiency or profitability) with an increase in their market shares while reducing those of less fit firms which, eventually, exit. Consequently, the economic system would be able to allocate resources in a more efficient way since profitable and efficient firms would account for an increasing share of the population whereas less viable ones would decline and, eventually, cease operations (Coad et al., 2013b). From these canonical industry dynamics models one can derive a series of predictions concerning how market selection forces operate via differential growth and via exit.

[3]Specifically, Winter and Nelson (1982) argue that profitable firms will grow and unprofitable ones will shrink. Initial profits would trigger subsequent growth via a selection mechanism that bolsters the firm's relative position by spawning extra profits that will be re-invested again, generating a virtuous circle.

First, a seemingly central feature of both neo-classical and evolutionary heterogeneous firms models is the prediction of a positive and strong association between firm growth and its relative efficiency and profitability. As argued in Bottazzi et al. (2010, p.3), these models imply that "productivity - proxying production efficiency - ought to be positively related to profitability and firm growth, at least on average. Depending on the models, this occurs either through a direct link between efficiency and growth - as relatively more efficient firms gain market shares by setting lower prices - or through an indirect effect via profitabilities - as more productive firms can enjoy higher profit margins which in turn allow them to invest more (in presence of endemically imperfect capital markets) and eventually grow more".

Second, the *growth of the fitter* notion implies that firms with above-average efficiency and profitability should feature above-average growth and gain market shares and vice-versa for relatively less efficient or profitable firms.

Third, as discussed in Hopenhayn (1992), Dosi et al. (1995), Pakes and Ericson (1998) and Audretsch (1995), among others, market selection forces may present different dynamics across industries. These are shaped by the degrees of competition of an industry, where the adequate 'fit' to survive and grow varies as the competitive landscape changes (Winter, 1984). In particular, some industries might represent more favorable environments for new firm growth and survival than others (Audretsch, 1995). This implies that firms' performance is largely contingent on structural differences in a particular industry's evolution (Klepper, 1996a) and to differences in the innovation intensity of a sector (Winter, 1984). Indeed, in highly innovative environments, the ability of new firms to adapt and offer viable products is particularly important for firm survival and growth. For instance, given that firm growth requires resources, one would expect that increasing profitability would lead to more resources as firms would be able to self-finance subsequent growth. A strong relationship between profitability and growth might be expected especially in the case of highly dynamics and innovative sectors (Thornhill, 2006).

A *fourth* and last prediction derived from both families of industry dynamics models is related to firm survival. Indeed, these models explicitly (Jovanovic, 1982) or implicitly (Winter and Nelson, 1982) contend that relative efficiency and/or profitability levels should significantly affect the chances of a firm to stay or exit the market.

The increased availability of firm-level data has provided the opportunity to address empirically the fourth conjectures outlined above especially for samples of established manufacturing firms. However, contrarily to neo-classical and evolutionary theoretical accounts, recent empirical works have been providing puzzling and challenging results regarding the functioning and strength of market selection forces especially regarding the first three predictions[4]. Among them, a series of studies finds that profitability exerts a limited influence on firm growth (Coad, 2007; Bottazzi et al., 2010, 2011; Coad and Broekel, 2012; Dosi et al., 2014). In particular, Bottazzi et al. (2010) and Dosi et al. (2014) show that profitability accounts for a small share of the variance of firm growth rates using Chinese, French and Italian samples of manufacturing firms. Similar results are found in Coad (2007) who, controlling for unobserved firm-specific effects, persistence and endogeneity, shows that profit rates have a small positive influence on subsequent growth for a sample of French manufacturing firms. Markman and Gartner (2002), whose contribution focused on high-growth firms in the US, reveal that profitability does not have a significant association with sales growth. Few studies have directly addressed this specific topic in the context of young firms and start-ups. Among them, a recent analysis employing a sample of South Korean publicly listed firms, has found a positive relation between profitability and growth for mature firms while, in the case of young firms, profitability affects growth negatively (Lee, 2014). Delmar et al. (2013) and Federico and Capelleras (2015) examine the growth-profitability nexus for new firms yielding contrasting results: while Delmar et al. (2013) show that past profitability has a positive effect on current sales growth in young Swedish firms operating in knowledge-intensive sectors, Federico and Capelleras (2015) do not find past profitability to exert a significant impact for a single cohort of new firms in the Spanish manufacturing sector. Overall, as argued in Bottazzi et al. (2010, p.1985), the absence of any strong relationship between profitability and growth militates against the "naively Schumpeterian" or "classical" notion that profits feed growth (by plausibly feeding investments).

Even when productivity is employed as measure of firm 'fitness', findings are not clear-cut and, in many cases, the evidence does not confirm the strong association with firm growth. In studies based on decomposition exercises the strength of market selection is indirectly captured by the so-called *between component*, namely, the reallocation of market shares from less to more efficient firms. This has been

[4]Notice that the two theoretical stances described above have also recurred to different methodological approaches to examine market selection. In fact, studies drawing on 'evolutionary equilibria' models have been opting for 'indirect' assessment of market selection, since the replicator dynamics is not tested explicitly but inferred through the sign and magnitude of the so-called between component. Contrarily, recent empirical contributions based on evolutionary theoretical grounds have directly tested the replicator dynamics by estimating 'fitness'-growth nexus at the micro-level.

usually shown to have a positive but small contribution to aggregate productivity growth (Foster et al., 2001; Dosi et al., 2015) while Disney et al. (2003), Baldwin and Gu (2006), Foster et al. (2008) even report a negative effect of the *between component* for, respectively, the UK, Canada and the US. What actually has a larger contribution to the dynamics of aggregate productivity is the *within component*, namely, idiosyncratic changes in firm productivity levels.

The investigations recurring to the direct estimations of the productivity-growth relation yield similar results: Bottazzi et al. (2002) and Bottazzi et al. (2010) employing samples of French and Italian manufacturing firms suggest equally weak strength of market selection forces. In particular, Bottazzi et al. (2010) examine the contemporaneous relation between firm growth and relative productivity and profitability. They find that relative productivity and profitability "explain" roughly 3-5% of growth rates variation whereas the contribution of unobserved heterogeneity is considerably larger. Drawing on Bottazzi et al. (2010), Dosi et al. (2015) investigate selection mechanisms employing samples of manufacturing firms in the US, France, Germany and the UK while Dosi et al. (2014) propose a similar framework in the case of Chinese firms. The evidence in both studies reveals that selection forces are at best weak along the productivity-growth and profitability-growth links. Indeed, productivity accounts for around 15-20% of the total variance in growth rates while profitability accounts only for 5%. In all the above-mentioned contributions, as for the literature employing decomposition techniques, idiosyncratic firm fixed effects account for the bulk of the explained variance in firm growth rates.

As previously argued, the second prediction is that, according to the 'replicator dynamics' notion, selection should be based on firms' relative fitness levels. This notwithstanding, empirical evidence across different countries yields results in contrast with this prediction. In fact, a seemingly puzzling finding is that firm expansion or contraction appear to be driven by *changes* in relative 'fitness' rather than relative 'fitness' levels (Dosi et al., 2015, 2014). Consistent with this, the studies by Coad and Broekel (2012) and Du and Temouri (2015) provide evidence regarding the importance of productivity growth for high-growth firms using French and UK data.

Table 1: Related Literature on Productivity-Growth relationship

	Sample			Variable		Method	Results
	Country	New Firms	Period	g	π		$\pi \rightarrow g$
Bottazzi et al. (2002)	Italy	No	89-96	Sales	LP	OLS	+
Bottazzi et al. (2008)	Italy	No	96-03	Sales	LP	OLS FE	+
Bottazzi et al. (2010)	France and Italy	No	91-04	Sales	LP	FE	5%
Coad et al. (2011)	Italy	No	89-97	Sales Employees	LP TFP	LAD VAR	+
Coad and Broekel (2012)	France	No	96-04	Sales Employees	LP TFP	LAD VAR	+
Dosi et al. (2014)	China	No	98-07	Sales	LP	CRE	15-20%
Dosi et al. (2015)	Fra, Ger UK and US	No	01-07	Sales	LP TFP	CRE	14-19%
Decker et al. (2016)	US	Yes	81-10	Employment	TFP	FE	+
Du and Temouri (2015)	UK	Yes	01-10	Sales Employees	TFP	Probit	+

Notes: Authors elaboration based on Lee (2014). We include only those contributions recurring to the direct estimation of the productivity-growth nexus via firm-level regressions. g refers to firm growth, while Π indicate productivity. LP stands for labour productivity and TFP for Total Factor Productivity. +, - and 0 refer to positive, negative and not significant effects. Percentages refer to the amount of total firm growth rates variation accounted for by productivity.

From an empirical standpoint, the analysis of how selective pressures change depending on the sector represents an often overlooked aspect since most of the evidence is based on samples of manufacturing firms with limited evidence regarding services[5]. Recent empirical contributions confirm that selection mechanisms operate heterogeneously across sectors more than previous literature acknowledged and this is true especially during the early stages of firms' life-cycle (Hyytinen and Maliranta, 2013). Baldwin and Gu (2011), who examine the differences in how selection operates between manufacturing and the retail sector in the US, find that the reallocation through entry and exit makes a larger effect on productivity growth in the retail trade with respect to the manufacturing sectors. Moreover, the contribution of new firms to labor productivity growth appears to be weaker in manufacturing. The effect of reallocation

[5] Foster et al. (2001), Foster et al. (2006), Baldwin and Gu (2011) and Hyytinen and Maliranta (2013) represent few exceptions.

among incumbent firms differs between retail and manufacturing: the *between component* has a larger contribution to aggregate productivity growth in the retail trade sector if compared to the manufacturing sector. In line with this, Foster et al. (2001, 2006), focusing on the service sector, show that both the reallocation of market shares and the "churning" associated with entry and exit dynamics have a strong and positive contribution to overall productivity growth. Du and Temouri (2015) examine whether TFP growth acts as a driver of high-growth performance in UK manufacturing and services firms. They show that TFP growth is relatively more important for incumbent firms than for new firms in the manufacturing sector while the opposite is true in services. In line with these results, the evidence provided by Delmar et al. (2013) in the case of Swedish new firms indicate that profitability affects more growth and survival of new firms in services as compared to those in the manufacturing sector. Delmar et al. (2013) also test whether the sectoral innovation intensity mediates selective pressures in terms of profitability. Surprisingly, they report that profitability appears to have a stronger effect on growth and survival in sectors with low innovation intensity.

Table 2: Related Literature on Profitability-Growth relationship

	Sample			Variable		Method	Results
	Country	New Firms	Period	g	Π		$\Pi \rightarrow g$
Bottazzi et al. (2008)	Italy	No	96-03	Sales	ROS ROI	OLS FE	+
Bottazzi et al. (2010)	France and Italy	No	91-04	Sales	GOM	FE	3-5%
Coad (2007)	France	No	96-04	Sales Employees	VA OS	OLS GMM	0
Coad (2010)	France	No	96-04	Sales Employees	GOS	LAD VAR	0
Coad et al. (2011)	Italy	No	89-97	Sales Employees	GOS	LAD VAR	0
Coad and Broekel (2012)	France	No	96-04	Sales Employees	GOS	LAD VAR	0 (+)
Cowling (2004)	UK	No	91-93	Sales	Profit	OLS 2SLS	+
Delmar et al. (2013)	Sweden	Yes	95-02	Sales	ROA EBIT	OLS FE	+
Dosi et al. (2014)	China	No	98-07	Sales	GPM	CRE	5%
Federico and Cap. (2015)	Spain	Yes	96-10	Sales	GOS	GMM	0
Jang and Park (2011)	US	No	78-07	Sales	ROS	GMM VAR	+
Lee (2014)	S. Korea	Yes	98-08	Sales Employees	NIS	GMM LAD	-
Markman and Gar. (2002)	US	No	92-28	Sales Employees	Profits	SHR	0

Notes: Authors elaboration based on Lee (2014). g refers to firm growth, while π indicate profitability. ROS stands for return on sales; ROA for returns on assets; GOS for gross operating surplus; OS for operating surplus; GPM for gross profit margin and NIS for net income to sales. SHR refers to Stepwise Hierarchical Regression. NA indicates that the relationship is not examined in the study. Percentages refer to the amount of total firm growth rates variation accounted for by profitability.

The *fourth* prediction stemming from industry dynamics models is related to the *survival of the fitter*. This has been tested (explicitly or not) by a quite abundant series of empirical studies. Even in this case, the predictions of both neo-classical and evolutionary approaches have not found clear-cut support (Carreira and Teixeira, 2011; Nightingale and Coad, 2013). Baily et al. (1992), who examine the productivity dynamics in US manufacturing firms in the 70s, reject the prediction that there is a productivity threshold below which firms inevitably cease operations. On the contrary, they show that, while roughly 50% of exiting firms were from the bottom two quantiles of the productivity distribution, nearly 30% were from the top two quantiles. Moreover, although exiters were mainly drawn from the bottom of the distribution, a considerable amount of low-productivity plants did not actually cease operations. The resilience of 'unfit' firms has been documented by several works which have also tried to shed light on the reasons why under-performing firms do not exit the market (Meyer and Zucker, 1989; Baden-Fuller, 1989; Karakaya, 2000). Gimeno et al. (1997) argue that firm survival is not only a function of economic performance but performance relative to firm-specific thresholds reflecting different motivation and aspiration levels. In particular, owners featuring higher levels of human capital tend not to endure poor performance while entrepreneurs with lower social capital but higher intangible returns to the job are likely to endure poor performance and, hence, survive. By building on this work, DeTienne et al. (2008)

identify a series of factors leading owners to increase their commitment notwithstanding poor performance, including market opportunities, personal investment, personal options, previous entrepreneurial success among other aspects.

Another stream of literature has analyzed how efficiency and profitability affect exit dynamics according to firms' life cycle. Bellone et al. (2008), using a sample of French manufacturing firms, argue that markets punish persistent bad performers and not those featuring temporary losses of efficiency. They also find profitability to be the major driver of survival and that its effect increases with age. However, concerning young firms (defined as firms between 1 and 3 years of activity), it appears that an increase in TFP is not significantly associated with higher chances of survival while it becomes more important as firms age.[6] Consistent with these findings, Warusawitharana (2014), based on a sample of UK businesses, shows that profitability is more relevant for the survival of mature firms if compared with their younger counterparts. Moreover, although profitability increases in the first years of the firm's life-cycle, this is due to within firms improvements implying that selection among firms does not explain this rise (Warusawitharana, 2014). Esteve Pérez et al. (2015) investigate how productivity affects firm survival depending on the product life cycle. They find that firm productivity is associated with lower chances of exit only in the 'old' phase of the product life cycle when market competition is primarily efficiency-driven. This could be related with the fact that in relatively new industries it takes some time to develop a specific competitive setting. Thus, in the early phases of an industry, it is more important the accumulation of experience with new products than carrying out cost-reduction strategies. Contrarily, in mature phases efficiency improvements are more effective. Hence, during the early phase of the life cycle younger managers with a more risk-loving attitude (Cucculelli and Ermini, 2012; Navaretti et al., 2014) may be more effective in introducing innovations and, thus, surviving market competition.

In this article, we provide a systematic analysis considering all of the above aspects in order to understand whether the empirical evidence in the case of newly established firms bear out such predictions.

3 Dataset and variables definition

The Kauffman Firm Survey (KFS) is a dataset tracking a nationally representative panel of 3,140 start-up firms founded in the U.S. in 2004 until 2011 (for a total of 8 waves). It constitutes a random sample from approximately 250,000 businesses listed in Dun and Bradstreet's (D&B) business database. This dataset has several distinguished features. First, it provides unprecedented insights into startups' early years. Second, the KFS avoids issues of survivorship bias by tracking a sample of new firms founded in the same year over time. Third, it covers new firms operating across a wide spectrum of sectors[7]. The KFS employs a stratified sampling scheme that over-samples new firms in high technology industries (for a more detailed discussion of the KFS design and methodology see Farhat and Robb (2014)). The KFS allows us to tackle some of the most frequent shortcomings in growth studies. On the one hand, since our interest lies in young firm growth, we focus on 'organic' growth. This represents the most common growth path followed by new firms (Delmar et al., 2003). Consequently, those firms controlled by another firm were removed from the database, leaving only those which are fully independent. On the other hand, survival bias has been identified as a major drawback in growth studies (Garnsey et al., 2006)[8]. In order to avoid this, we use an unbalanced panel and rely on the probability weights available in the KFS to account both for sampling and for response biases in the baseline survey. These weights are reconstructed with each survey wave and they account for the systematic failure of new businesses, and thus, their permanent exit from the data set (see the Appendix for a more detailed discussion).

We proceed to cleaning the data by removing those firms with missing values in terms of revenues,

[6] Bellone et al. (2008) maintain that, the fact that productivity and profitability seem to have a larger impact on survival as firms age, it could be reflecting "the fact that young firms are more exposed to market selection, so that the relationship between performance and survival becomes looser. In other words, the micro-economic determinants of market selection of young firms lie elsewhere, perhaps in firm size (Audretsch and Mahmood, 1995) and, again, credit constraints (Aghion et al, 2006)". However, notice that Bellone et al. (2008) consider firm age from the first appearance in their dataset.

[7] The KFS is the only longitudinal dataset representative of a cohort of new firms that includes both information on firm-level financial and economic outcomes. As reported by Zarutskie and Yang (2015, p.5) in the US "only 7 out of 26 relevant data sets for research on entrepreneurship provided longitudinal information on new venture creation, but none of the 7 data sets applied selection criteria that would lead to a representative sample of new businesses". For instance, datasets such as the U.S. Census Bureaus Longitudinal Business Database (LBD) and the Bureau of Labor Statistics' Longitudinal Database (LDB) do not track non-employer firms, nor do they contain information on assets, revenues or financing (Zarutskie and Yang, 2015).

[8] Garnsey et al. (2006) point out that "greater consistency could be provided in factors of growth studies by studying cohorts of comparable new firms over the same time period. This would address the charge that factors of growth studies have drawn inappropriate inferences from survivors in samples by excluding less successful firms".

profits, value added and employment in the first two years (i.e. 2004 and 2005). Thus, we are left with 863 firms in 2005 and 3,720 observations.

In order to investigate the first aspect of market selection (the *growth of the fitter*), we focus, on the one hand, on sales growth, and on the other, on the two most common measures of firm 'fitness': profitability and labour productivity. In accordance to a 'replicator dynamics' type of relationship, we consider both dependent and independent variables in deviation from their yearly cross-sectional average at the (one-digit) industry level[9]. In line with previous literature (Bottazzi et al., 2010), we define relative firm growth as the logarithmic difference:

$$\text{Growth}_{it} = \text{s}_{it} - \text{s}_{it-1}$$

where

$$\text{s}_{it} = \log(\text{Revenues}_{it}) - \frac{1}{N}\sum_{i} \log(\text{Revenues}_{it})$$

The sum is computed over the N firms located in the same one-digit NAICS sector. We select this measure since it better captures the ability of a firm to sell its products and services and their acceptance by the market (Gilbert et al., 2006). With respect to relative profitability, we consider Returns on Sales (ROS) which is defined as:

$$\text{ROS}_{it} = \frac{\text{EBIT}_{it}}{\text{Revenues}_{it}} - \frac{1}{N}\sum_{i} \frac{\text{EBIT}_{it}}{\text{Revenues}_{it}}$$

where EBIT is earning before interests and taxes. We choose this proxy for theoretical reasons (Winter and Nelson, 1982) and empirical comparability (Bottazzi et al., 2010). In fact, in order to test the 'replicator dynamics' implicit in Winter and Nelson (1982), we need a measure of cash-flow over sales, proxied by the share of gross profits in total revenues which, in turn, proxies the potential for re-investment.

Moreover, since we are interested in how efficiency affects firm growth, we employ relative labour productivity computed as the ratio of value added and the number of employees[10]:

$$\text{LP}_{it} = \frac{\text{VA}_{it}}{\text{Employees}_{it}} - \frac{1}{N}\sum_{i} \frac{\text{VA}_{it}}{\text{Employees}_{it}}$$

Value added is computed as revenues minus costs of intermediate inputs. Finally, relative firm size is defined as follows:

$$\text{Size}_{it} = \log(\text{Employees}_{it}) - \frac{1}{N}\sum_{i} \log(\text{Employees}_{it})$$

For both relative labour productivity and size variables, before the logarithmic transformation, we add one to the employees since quite a few new firms in their first years do not record any employee. All monetary variables are deflated with 3-digit level production price indexes available from the US Bureau of Labor Statistics (BLS) (base year 2009).

Table 3 presents the mean values for the variables of interest. One important aspect to highlight is the pattern observed during the period 2008-2009: in particular, average revenues, growth rates, profitability and labour productivity experience a considerable decrease during the financial crisis[11] while they bounce back to pre-crisis levels in 2010-2011.

As found in other works, the distribution of firm growth rates (unreported) displays a Laplace or symmetric exponential distribution (Bottazzi and Secchi, 2006). Moreover, it is highly skewed to the right in the first year (i.e. 2005) while, successively, it has a tendency to move towards the left meaning that, on average, firms tend to grow less as they age and that the representative firm experiences little growth which is consistent with empirical evidence (see, among others, Coad et al. (2013a) and Navaretti et al. (2014)).

Regarding the fitness variables, on the one hand, the labour productivity distribution moves towards the right with time indicating how surviving firms, possibly due to selection and to learning-by-doing effects, become more productive on average. On the other hand, the average profitability does not show a clear trend: it experiences an increase in the first two years while it decreases during the period 2006-2009 and it grows again during 2010-2011.

[9] We chose to normalize the variables at this level of disaggregation given that in some two-digit NAICS sectors the number of firms is very low. However, the subsequent analysis was also conducted employing variables normalized at two-digit level and results were not affected.

[10] As argued in Dosi et al. (2016) and Dosi and Grazzi (2006), TFP might be biased in the presence of technologically heterogeneous firms and complementarity among inputs. However, as a robustness check, we run the analysis using a TFP measure yielding qualitatively similar results (see Appendix 3).

[11] See Zarutskie and Yang (2015) for an thorough examination of young firms performance during the Great Recession employing the KFS dataset.

Table 3: Summary statistics per year

Variable	2004	2005	2006	2007	2008	2009	2010	2011
Revenues ($)	301,819	662,857	1,023,084	1,034,818	679,539	1,849,914	1,804,916	2,000,187
	(57,580)	(110,492)	(150,067)	(185,103)	(180,002)	(161,739)	(173,035)	(178,895)
Growth (log)	-	0.97	0.37	0.14	0.07	-0.08	0.06	0.05
	(-)	(0.71)	(0.23)	(0.10)	(0.05)	(-0.01)	(0.04)	(0.03)
Size	3.92	4.77	5.32	5.20	5.04	5.21	5.83	5.71
	(2.00)	(2.00)	(2.00)	(2.00)	(2.00)	(2.00)	(2.00)	(2.00)
ROS	0.38	0.40	0.39	0.37	0.33	0.22	0.31	0.32
	(0.45)	(0.38)	(0.36)	(0.34)	(0.30)	(0.26)	(0.27)	(0.26)
LP (log)	9.23	9.61	9.86	10.01	9.92	9.86	9.93	9.92
	(9.37)	(9.74)	(10.00)	(10.07)	(10.13)	(10.00)	(10.09)	(10.17)
# of firms	968	863	676	575	462	421	374	350

Notes: Population-weighted results using KFS survey weights. Variables in this table are not de-meaned at the sectoral level. Medians in parenthesis. Monetary variables are measured in real terms (base year 2009).

4 Growth of the 'fitter' analysis

In order to provide a first overview concerning the relations among firm growth, on one side, and profitability and productivity, on the other, Table 4 presents contemporaneous correlations for both the pooled sample and for different NAICS sectors at one-digit level. For the pooled sample, the correlations are positive and significant. However, labour productivity features a higher correlation with growth as compared with profitability (25% vs. 15%).

The relatively low overall correlation between growth and profitability is something found also in other studies focusing on new firms: for instance, considering Swedish new firms in knowledge-intensive services, Delmar et al. (2013) shows that the correlation between profitability and firm growth is only 8%. It is worth noticing that, if we consider that knowledge-intensive services correspond mainly to NAICS 5 (Information, Finance, Real Estate and Professional Services), we can see that the correlation between growth and ROS (5%) is similar to the one reported by Delmar et al. (2013). Moreover, pair-wise correlation coefficients show considerable degrees of heterogeneity across industries as observed in other studies (Federico and Capelleras, 2015).

Table 4: Contemporaneous Correlation Matrix

	Growth-ROS		**Growth-LP**	
NAICS	Pairwise	Obs	Pairwise	Obs
1	0.06	20	0.23	20
2	**0.15**	241	**0.16**	241
3	0.01	528	**0.27**	528
4	**0.25**	704	**0.30**	704
5	**0.05**	1,751	**0.24**	1,751
6	**0.23**	91	0.03	91
7	-0.07	150	0.02	150
8	0.10	235	**0.34**	235
Total	**0.15**	3,720	**0.25**	3,720

Notes: Population-weighted results using KFS survey weights. Variables in this table are not de-meaned at the sectoral level. Bold numbers are statistically significant at the 95% level.

Empirical contributions focused on the analysis of market selection mechanisms have mainly recurred to decomposition methodologies. In this study, we follow the approach outlined by a series of empirical contributions (Bottazzi et al., 2010; Dosi et al., 2014, 2015) pursuing the direct estimation of the relationships between fitness variables and growth rates through firm-level regressions. As argued by Dosi et al. (2015) concerning the productivity-growth nexus the empirical literature has rarely followed this route, perhaps influenced by the theoretical agreement that firm growth and firm efficiency are positively and strongly correlated.

Drawing on Dosi et al. (2015), our main specification accounts for the overall explanatory power of current and lagged levels of relative fitness variables (alternatively, productivity and profitability) upon new firms growth. Successively, we examine the explanatory power of relative fitness levels as opposed to *changes* over time of relative fitness. Finally, we disentangle whether selection forces operate according to different degrees of intensity depending on the sector under examination.

4.1 Firm growth and relative 'fitness'

In order to investigate the relationships described in the previous section, we estimate the following equations:

$$\text{Growth}_{i,t} = \alpha + \sum_{k=0}^{1} \beta_k \text{ROS}_{i,t-k} + \gamma \text{Size}_{i,t-1} + \lambda_t + \phi_{s_i} + \eta_{r_i} + \epsilon_{i,t} \tag{1}$$

$$\text{Growth}_{i,t} = \alpha + \sum_{k=0}^{1} \beta_k \text{LP}_{i,t-k} + \gamma \text{Size}_{i,t-1} + \lambda_t + \phi_{s_i} + \eta_{r_i} + \epsilon_{i,t} \tag{2}$$

where $Growth_{i,t}$ indicates relative firm growth rate for each firm i at time t, ROS is relative profitability while LP refers to relative labour productivity. Moreover, r_i and s_i are respectively the State and the industry in which firm i operates[12].

Eq. (1) and (2) represent our growth models which extend the traditional Gibrat type growth equation where firm growth rates are usually regressed on size. In Eq. (1), growth is regressed against current and lagged ROS in order to capture the association between profitability and firm growth. Eq. (2) is employed in order to examine how firms' efficiency, proxied by labour productivity, is associated with firm growth.

Furthermore, our main focus is not just related to the point estimates and significance levels of current and lagged relative labour productivity and profitability. Instead, we are interested in assessing which is the "overall" contribution of relative productivity and profitability to sales growth. Indeed, if markets are able to reward firms based on their relative 'fitness', productivity and profitability should be good predictors of firms' growth and survival; thus accounting for a good part of its variance. Hence, to verify if this is the case, we resort to an R^2 decomposition which allows us to establish the contributions of individual and groups of variables to the total variance of the model. In particular, we compute the Shapley value (Shapley, 1953) which represents the average marginal contribution of individual regressors to the overall goodness-of-fit of the model [13].

In order to get the intuition behind the procedure used to calculate the Shapley values, assume we start from one of the models described above and that "we successively remove regressor variables, one by one and according to a particular ordering of the variables. The difference in goodness-of-fit associated with the elimination of a variable can be regarded as the variable's marginal contribution in this particular ordering of the regressors. Treating all orderings equally probable, the Shapley value of a variable equals the variable's average marginal contribution over all possible orderings" Huettner et al. (2012, p.3).

Before starting to illustrate the results, we must stress that our empirical strategy is not intended to identify causal effects among, on the one hand, profitability and productivity and, on the other, firm growth rates. Instead, we seek to test the 'replicator dynamics' in its purest form and to report associations between these variables. Furthermore, we are not particularly concerned about omitted variable bias given that previous empirical exercises on the determinants of firm growth did not actually find any particular factor that systematically explain firm growth rates to any large extent (Coad, 2009).

Table 5 presents OLS and FE estimates of Eq. (1): they indicate that profitability exerts a significant effect on firm growth (Table 5). However, while current ROS has a positive association, its lagged coefficient is negative. The signs, magnitudes and patterns of statistical significance are quite stable across both specifications. Table 6 presents the estimation for Eq. (2). As for the profitability model, the signs of the current coefficient is positive while the first lag is negative but the magnitude of the point estimates is considerably higher for productivity[14].

These results so far indicate that firms' 'fitness', particularly our proxy for efficiency, do significantly affect firm growth. Still, we cannot say much regarding the strength of market selection mechanisms in rewarding or punishing firms based on their productivity or profitability.

In Tables 5 and 6, we also report the results of the above models for the Shapley R^2 decomposition. Despite the significant coefficients, we can observe that ROS do not turn out to be very important in

[12] In order to control for location-specific factors that could affect the relationships under study, we use the dummy variables corresponding to all 50 US States plus the District of Columbia. With respect to the sectors, we employ two-digit NAICS dummies.

[13] A more detailed description of this method is provided in Appendix 1.

[14] For both models, we checked whether the same signs hold by estimating two separate specifications with current and lagged 'fitness' variables alternatively inserted as the only regressor. The estimates confirm that the sign pattern is not due to "perverse" collinearity between current and lagged productivity and profitability.

Table 5: Growth and Profitability

	OLS			FE		
	Coef.	R^2 share	Expl. variance	Coef.	R^2 share	Expl. variance
ROS_t	0.007***	3.02		0.008***	1.91	
	(0.002)		**1.43**	(0.002)		**1.46**
ROS_{t-1}	−0.011***	5.93		-0.010***	2.89	
	(0.003)			(0.002)		
Size_{t-1}	0.029	0.50	0.08	-0.100**	0.33	0.10
	(0.021)			(0.048)		
Year FE	(7 d)***	81.85	13.09	(7 d)***	41.78	12.79
Sectoral FE	(21 d)***	3.35	0.54	-	-	-
State FE	(51 d)*	5.35	0.86	-	-	-
Firm FE	-	-	-	(863 d)***	53.09	16.25
Constant	−0.874***			-0.293***		
	(0.323)			(0.059)		
Total		100	15.99		100	30.60
N. obs.	3,720			3,720		
N. firms	863			863		
R^2	15.99			30.60		

Notes: Pooled OLS and FE estimates using KFS survey weights. ***, ** and * indicate significance on a 1%, 5% and 10% level for individual variables (t-test) or groups of dummy variables (F-test). For the latter, we provide in parenthesis the number of variables for each group of dummies. In the "R^2 share" column we report the Shapley values while in the "Expl. variance" we present the contribution of each variable to the total variation of growth rates. For instance, in the case of the OLS estimates, the value for both ROS_t and ROS_{t-1} is obtained by multiplying their Shapley values with the model $R^2((3.02 + 5.93) \times 0.1599 = 1.43)$. Sectoral and State dummies are not included in the FE model since they are time-invariant variables.

Table 6: Growth and Productivity

	OLS			FE		
	Coef.	R^2 share	Expl. variance	Coef.	R^2 share	Expl. variance
LP_t	0.390***	34.85		0.356***	22.32	
	(0.022)		**21.48**	(0.033)		**19.26**
LP_{t-1}	−0.361***	28.29		-0.372***	21.06	
	(0.023)			(0.034)		
Size_{t-1}	−0.028	0.22	0.08	-0.256***	0.75	0.33
	(0.019)			(0.046)		
Year FE	(7 d)***	32.89	11.18	(7 d)***	24.62	10.94
Sectoral FE	(21 d)	1.39	0.47	-	-	-
State FE	(51 d)	2.36	0.80	-	-	-
Firm FE	-	-	-	(863 d)***	31.24	13.88
Constant	−0.748***			-0.225***		
	(0.291)			(0.044)		
Total		100	34.01		100	44.41
N. obs.	3,720			3,720		
N. firms	863			863		
R^2	34.01			44.41		

Notes: see notes in Table 5.

terms of explanatory power. In fact, the Shapley values of current and past ROS represent a mere 1.4% of firm growth rates variation in both OLS and FE models. Contrarily, the Shapley values of labour productivity accounts for a higher share of the total goodness-of-fit of the OLS and FE models and reach around 19-21% of firm growth rates.

Finally, moving to the examination of Gibrat's Law, we observe that point estimates, significance and sign patterns of the $Size_{t-1}$ coefficient are not stable when estimating Eq. 1 and 2 by means of OLS, while the coefficient becomes negative and significant when employing the FE estimator [15]. However, although we can reject Gibrat's law, we must stress that the explanatory power of $Size_{i,t-1}$ is very marginal given that it reaches only up tp 0.75% of total growth rates variation.

4.2 Firm growth and relative 'fitness' levels vs. changes

In what follows, we test the second prediction stemming from industry dynamics models. In particular, drawing on this literature, one would expect firm growth dynamics to be strongly related to firm relative 'fitness' levels. In the previous section we show that the Shapley values obtained for both productivity and profitability are actually the results of two conflicting effects, a positive one from contemporaneous variables and a negative one from the lagged ones. Thus, one may suspect that the actual drivers of firm growth are not related to relative 'fitness' levels at any time period, but rather in their change through time. Hence, in line with Dosi et al. (2015), we specify a different regression model allowing us to test the importance of relative 'fitness' levels *vis-à-vis* relative 'fitness' *changes*. We test this hypothesis for both labour productivity and profitability in the case of newly-established firms by estimating the following equations:

$$\text{Growth}_{i,t} = \alpha + \beta_\Delta \Delta\text{ROS}_{i,t-k} + \beta_M \overline{\text{ROS}}_{i,t} + \gamma \text{Size}_{i,t-1} + \lambda_t + \phi_{s_i} + \eta_{r_i} + \psi_{t,s_i} + \epsilon_{i,t} \tag{3}$$

$$\text{Growth}_{i,t} = \alpha + \beta_\Delta \Delta\text{LP}_{i,t-k} + \beta_M \overline{\text{LP}}_{i,t} + \gamma \text{Size}_{i,t-1} + \lambda_t + \phi_{s_i} + \eta_{r_i} + \psi_{t,s_i} + \epsilon_{i,t} \tag{4}$$

where ΔLP_t and ΔROS_t are the log differences of relative labour productivity and profitability over two consecutive time periods, accounting for the dynamics of relative efficiency and profitability, while $\overline{LP}_t$ and $\overline{ROS}_t$ are the within-firm average relative LP and ROS levels computed over t and $t-1$, in turn capturing the absolute differential efficiency and profitability among firms. Following a 'replicator dynamics' depiction of market selection, firms should be selected and grow mostly according to their "static" relative 'fitness' and, thus, one would expect the Shapley values of $\overline{ROS}_t$ and $\overline{LP}_t$ to be grater than those of ΔROS_t and ΔLP_t.

However, consistent with the findings provided in Dosi et al. (2014) and Dosi et al. (2015), our results show that almost all explanatory power of relative productivity and profitability is attributable to *changes* in relative 'fitness' rather than levels thereof. In fact, labour productivity and profitability levels both account only for around 0.8% of total growth rates variation (Tables 7 and 8). On the contrary, the share of total explained variance accounted for by *changes* in 'fitness' levels is around 1.3% and 19-20% for, respectively, profitability and labour productivity.

[15] To obtain a clearer picture, we estimated the models using OLS, FE and GMM-sys inserting exclusively $Size_{i,t-1}$ as independent variable (along with regional, time, sector dummies and year-sector FE). The estimates clearly reject Gibrat's law for the whole sample of firms since the coefficient is negative and significant. To sum up, the instability of the results commented in the main regressions is probably due to the high correlation between labour productivity and firm size.

Table 7: Growth and Profitability changes vs. levels

	OLS			FE		
	Coef.	R^2 share	Expl. variance	Coef.	R^2 share	Expl. variance
ΔROS_t	0.009***	8.56	**1.37**	0.009***	4.54	**1.39**
	(0.001)			(0.001)		
$\overline{ROS}_t$	−0.004	0.39	**0.06**	-0.002	0.26	**0.08**
	(0.004)			(0.002)		
Size_{t-1}	0.029	0.50	0.08	-0.100***	0.33	0.10
	(0.021)			(0.048)		
Year FE	(7 d)***	81.85	13.09	(7 d)***	41.78	12.79
Sectoral FE	(21 d)***	3.35	0.54	-	-	-
State FE	(51 d)*	5.35	0.86	-	-	-
Firm FE	-	-	-	(863 d)***	53.09	16.25
Constant	−0.874***			-0.293***		
	(0.323)			(0.059)		
Total		100	15.99		100	30.60
N. obs.	3,720			3,720		
N. firms	863			863		
R^2	15.99			30.60		

Notes: see notes in Table 5.

Table 8: Growth and Productivity changes vs. levels

	OLS			FE		
	Coef.	R^2 share	Expl. variance	Coef.	R^2 share	Expl. variance
ΔLP_t	0.376***	62.79	**21.35**	0.364***	43.18	**19.18**
	(0.021)			(0.029)		
$\overline{LP}_t$	0.029*	0.36	**0.12**	-0.016	0.20	**0.09**
	(0.016)			(0.033)		
Size_{t-1}	−0.030	0.22	0.08	-0.256***	0.75	0.33
	(0.019)			(0.046)		
Year FE	(7 d)***	32.89	11.18	(7 d)***	24.62	10.94
Sectoral FE	(21 d)	1.39	0.47	-	-	-
State FE	(51 d)	2.36	0.80	-	-	-
Firm FE	-	-	-	(863 d)***	31.24	13.88
Constant	−0.748***			-0.225***		
	(0.291)			(0.044)		
Total		100	34.01		100	44.41
N. obs.	3,720			3,720		
N. firms	863			863		
R^2	34.01			44.41		

Notes: see notes in Table 5.

4.3 Firm growth and ‘fitness’ variables across sectors

After the investigation based on the whole sample of new firms, we present a more disaggregated analysis given that market selection mechanisms might substantially differ depending on the sector examined. Despite this, the great bulk of empirical evidence focuses on the manufacturing sector. The investigation of how selective pressures operate in different sectors is particularly relevant in our case for two additional reasons. The first one is that the number of entrants in the service sector in the US outnumber manufacturing new firms around 8 to 1 (Kim et al., 2006). The second one is due to the fact that new firms are particularly heterogeneous, ranging from self-employed workers to professionals featuring high levels of education, to cutting-edge high-tech firms attempting to compete in global markets (Hurst and Pugsley, 2011).

We, therefore, begin by conducting a split-sample analysis by considering the differences among manufacturing and services. For the entire sectoral analysis, we estimate Eq. (3) and (4) employing the FE estimator given that it accounts for firm unobserved heterogeneity. Moreover, for the sake of brevity, we only show the results for the Shapley R^2 decomposition (the complete set of results is available upon request). Table 9 contains the Shapley values along with the percentage of total explained variance accounted for by our variables of interests. The estimates indicate that there seem to be slightly more selective pressures in terms of productivity for firms in service sectors as compared with manufacturing: the explanatory power of productivity is higher for service firms. This result is in line with evidence regarding new firms (Delmar et al., 2013) and several studies using decomposition methodologies which report a larger contribution of the *between component* to aggregate productivity growth in services when compared to manufacturing (Foster et al., 2006; Baldwin and Gu, 2011).

As already shown in the previous sections of the paper, ROS turns out to account for a negligible (or even irrelevant) share of total firm growth variation. Therefore, although we present the results of the models including this profitability proxy, we restrain from drawing strong conclusions regarding the heterogeneous effects of selection upon profitability across sectors[16].

Table 9: Manufacturing vs. Services - FE R^2 decomposition

	ΔROS_t	$\overline{ROS}_t$	$Size_{t-1}$	Firm FE	R^2	Obs.	Firms
Manufacturing							
R^2 share	5.01	0.35	0.33	41.39	38.23	2,930	677
Explained variance	**1.91**	**0.13**	0.13	15.83			
Services							
R^2 share	1.52	1.99	1.57	83.71	50.65	528	116
Explained variance	**0.77**	**1.01**	0.79	42.40			
	ΔLP_t	$\overline{LP}_t$	$Size_{t-1}$	Firm FE	R^2	Obs.	Firms
Manufacturing							
R^2 share	39.82	0.16	0.68	25.48	52.40	2,930	677
Explained variance	**20.86**	**0.09**	0.36	13.35			
Services							
R^2 share	34.86	1.19	2.26	53.86	51.32	528	116
Explained variance	**22.94**	**0.79**	1.49	35.45			

Notes: FE estimates using KFS survey weights. For each sectoral regressions, we directly report the Shapley values (R^2 share) for the ‘fitness’ variables, lagged firm size and firm fixed effects. For the same set of co-variates we display the contribution to the total variation of growth rates in the row “Explained variance”. The regressions include a full set of time dummies. The complete results are available upon request.

A reason why selective pressures for new firms appear to be stronger in services than in manufacturing could be related to the presence of higher sunk costs in the latter. According to the model by Hölzl (2015), industries with higher sunk costs should exhibit a lower speed of market shares reallocation. Austrian data confirms this hypothesis. Indeed, market share reallocation is less important as source of overall productivity growth in higher-sunk-cost industries: “sunk costs can explain the low contribution of reallocation of productivity growth in manufacturing sector compared to many service sectors” (Hölzl,

[16] As already pointed out, we opted for this proxy for both theoretical and empirical reasons. In robustness tests, we also check the sensitivity of our findings employing an alternative measure, namely, Returns on Assets (ROA). The significance levels and magnitudes of the coefficients for this proxy are slightly higher if compared with ROS. However, the share of the explained variance accounted for by this regressor is still negligible given that it reaches only up to 3%. The estimates regarding the sectoral analysis using ROA confirm that: i) the dynamic term of profitability accounts for a larger share of growth rates variation; ii) selection on profitability appears to be higher in services, low-tech and less concentrated sectors. Results are available upon request.

2015, p.341) which has been documented by several empirical investigations[17].

The analysis performed for the manufacturing sector allows comparison with Dosi et al. (2015). Using similar techniques, they show that the contribution of productivity to the explained variance of firm growth around 15-20%. While there are some differences in the techniques employed among the two papers, the results are vastly comparable up to some percentage point.

We also perform a split-sample analysis to take into account potential differences among sectors depending on their innovation intensity. We make of use of the technological classes present in the KFS: based on Chapple et al. (2004), the KFS considers firms as high-tech if they belong to a sector which is technology generator or technology employer. According to this taxonomy, industries are defined as generators of technology according to the NSF's Survey of Industrial Research and Development as sectors that exceed the US average for both R&D expenditures for employees and the proportion of full-time equivalent R&D scientists and engineers in the industry workforce. A firm is technology employer whether in its industry the employment of these occupations exceeds three times the national average. The list of sectors considered high-tech can be found in Table A1 in the Appendix.

Table 10: High-Tech vs. Low-Tech - FE R^2 decomposition

	ΔROS_t	$\overline{ROS}_t$	$Size_{t-1}$	Firm FE	R^2	Obs.	Firms
High-Tech							
R^2 share	0.38	0.58	0.64	17.96	61.52	625	152
Explained variance	**0.23**	**0.35**	0.39	11.05			
Low-Tech							
R^2 share	4.68	0.25	0.30	58.23	29.16	3,094	711
Explained variance	**1.36**	**0.07**	0.09	16.98			
	ΔLP_t	$\overline{LP}_t$	$Size_{t-1}$	Firm FE	R^2	Obs.	Firms
High-Tech							
R^2 share	15.42	0.39	0.52	13.32	70.05	625	152
Explained variance	**10.80**	**0.27**	0.36	9.33			
Low-Tech							
R^2 share	45.65	0.19	0.81	33.06	43.88	3,094	711
Explained variance	**20.03**	**0.08**	0.35	14.51			

Notes: see notes in Table 9.

The estimates presented in Table 10 reveal an overall weaker power of 'fitness' variables in high-tech industries with respect to firms in low-tech sectors especially for productivity. This appears to be in line with recent evidence documenting a reduction in the strength of the productivity-growth relationship especially for young high-tech firms *vis-à-vis* their young non-tech counterparts starting from the early 2000s (Decker et al., 2016). These results could be also interpreted as signals that market selection might operate on longer time horizons in environments in which the uncertain and risky innovation activities have a crucial role in shaping industrial dynamics (Geroski and Mazzucato, 2002). Therefore, low levels of efficiency or even temporary losses of 'fitness' do not cause an immediate decrease in firm growth rates.

There are several other competing explanations however with respect to the low explanatory power of productivity in high-tech sectors. First, previous contributions showed that productivity improvements are not always beneficial for firm performance and survival in relatively new sectors (Klepper, 1996b, 2002; Esteve Pérez et al., 2015). Indeed, in these contexts, efficiency improvements might matter less than product innovation. Conversely, in mature phases cost-lowering investment are more effective. Therefore, during the early phase of the life cycle younger managers with a more risk-loving attitude (Cucculelli and Ermini, 2012; Navaretti et al., 2014) may be more effective in introducing innovations and, thus, surviving market competition.

A different possible interpretation concerning the results in high-tech and low-tech sectors is the mediating role of finance with respect to market selection. Indeed, for high-tech firms, performance might respond less to productivity improvements given that their competitiveness is based on innovative efforts that usually yield results long after the project's inception. Market selection would then be based on other factors especially during the first years of activity. In this sense, the role of credit might be crucial in easing "the selection pressures on innovative firms by providing them with the means to survive until their innovative products make it to the market" (Geroski and Mazzucato, 2002).

[17] Foster et al. (2006) shows that retail trade establishments feature reallocation rates of outputs and inputs that are roughly 50 percent higher than for manufacturing in the US. Furthermore, "entry and exit of establishments plays a much larger role in retail trade. [...] This pattern of lower reallocation and greater dispersion in manufacturing makes sense given the presumably higher adjustment costs in manufacturing (e.g., higher barriers to entry given minimum efficient scale)".

An additional empirical test is to investigate whether 'fitness' variables are more important in driving selection among firms in sectors with higher competition/lower concentration. We test this hypothesis by classifying firms according to the concentration of their respective sectors. In order to do this, we exploit the data from the Economic Census of 2007 which provide us with the share of revenues accounted for by the largest 20 and 50 firms in each NAICS sector. According to these data we categorized each NAICS sectors at two-digit level into either a high-concentration and low-concentration industry[18]. The R^2 decomposition exercise provides insights on the relative strength of market selection and confirms our intuition according to which in sectors characterized by lower concentration and, hence, higher competitive pressure, selection based on efficiency and profitability is stronger than in sectors with higher concentration and lower competition (Table 11).

Table 11: High-Concentration vs. Low-Concentration - FE R^2 decomposition

	ΔROS_t	$\overline{ROS}_t$	$Size_{t-1}$	Firm FE	R^2	Obs.	Firms
High-Concentration							
R^2 share	1.00	0.46	0.55	71.10	26.43	1,183	282
Explained variance	**0.26**	**0.12**	*0.14*	*18.79*			
Low-Concentration							
R^2 share	5.26	0.38	0.27	42.08	35.82	2,536	581
Explained variance	**1.89**	**0.14**	*0.10*	*15.08*			
	ΔLP_t	$\overline{LP}_t$	$Size_{t-1}$	Firm FE	R^2	Obs.	Firms
High-Concentration							
R^2 share	41.88	0.18	0.99	39.74	39.54	1,183	282
Explained variance	**16.56**	**0.07**	*0.39*	*15.71*			
Low-Concentration							
R^2 share	40.76	0.26	0.62	26.18	49.58	2,536	581
Explained variance	**20.21**	**0.13**	*0.31*	*13.98*			

Notes: see notes in Table 9.

5 Survival of the 'fitter' analysis

We now turn to the investigation of the second aspect of selection, namely, the *survival of the fitter*. While most datasets do not provide information on firm survival and on the exit mode, the KFS allows to distinguish between firms that either survived, permanently closed operations and merged with or were sold to other firms. In order to study how 'fitness' variables relate to exit, we exclude those new independent firms which exit the market through mergers or acquisitions and, consequently, we consider exit only in terms of firms being permanently 'out of business'. A new firm exiting by merger or acquisition is more likely to be a successful firm, or at least one that still has potential. Contrarily, a firm that permanently ceases operations probably features a mismatch between the firm's resources and capabilities and the opportunities in the marketplace[19]. Contrarily to the first part of the analysis, the sample is larger given that we do not lose the first year to compute growth rates. The unique firms are 980 for a total number of 4,625 observations.

We begin by providing descriptive evidence on whether surviving and exiting firms display systematic differences in terms of their productivity and profitability performances. In particular, we examine the specific position of exiters in the productivity and profitability distributions. We ranked firms according to their LP and ROS levels in order to successively compute the corresponding quantiles in each year (Table 12). It is clear that exit is mainly observed for those firms located at the lower bottom of the distribution of both productivity and profitability. For the pooled sample, 52.75% and 50.57% of the exiting firms belong to the two lowest quantiles of the productivity and profitability distributions, respectively. However, 28.91% and 31.82% of exiting firms belong to the top two quantiles of both distributions. The figures related to productivity closely match the evidence presented in Baily et al. (1992) for the manufacturing sector in the US.

It is also clear that there is a great deal of variation in exit patterns across different years. For labour productivity, we can see that the amount of exiting firms drawn from the highest quantile of the

[18]The sources of the data and the list of sectors can be found in Table A2 in the Appendix.

[19]However, we cannot exclude that, among the firms declaring to be permanently 'out of business', some of them can be considered successful exits. In fact, entrepreneurs may voluntarily cease operations if they have better outside options implying that the firm cannot be considered completely unviable.

distribution tends to decrease with time. Indeed, while in 2004 17.24% of the exiting firms belongs to the highest quantile, this figure is considerably lower in 2010 reaching 3.70%. In line with previous empirical evidence (Bellone et al., 2008; Esteve Pérez et al., 2015), this might indicate that market selection based on 'fitness' becomes more and more important as firms age. With respect to profitability, we observe a different pattern since firms in the highest quantile of the distribution do not show a lower fraction of exiting firms with respect to rest of the distribution. In fact, for almost all years, the lowest percentage of exiting firms is drawn from the central part of the distribution (quantiles 3 and 4), thus suggesting a non-linear relation with survival.

Table 12: LP and ROS of exiting firms

	1	2	3	4	5	Total
Exit during:	LP quantiles in the year of exit					
2004	38.39	06.83	15.61	21.94	17.24	100
2005	27.81	19.11	17.61	20.04	15.07	100
2006	32.92	21.24	25.89	9.04	10.91	100
2007	38.48	19.06	12.98	11.94	17.54	100
2008	39.49	24.05	17.01	12.38	7.07	100
2009	50.75	11.02	21.79	13.29	3.16	100
2010	44.50	24.46	19.64	7.71	3.70	100
Whole period	35.89	16.86	18.33	15.94	12.97	100
Exit during:	ROS quantiles in the year of exit					
2004	31.28	26.14	16.99	12.03	13.56	100
2005	23.20	18.92	22.02	16.65	19.21	100
2006	27.09	21.09	17.87	15.36	18.59	100
2007	40.34	20.12	15.06	12.12	12.35	100
2008	22.28	27.59	11.47	20.14	18.53	100
2009	29.98	16.41	12.67	18.75	22.18	100
2010	34.26	18.32	20.03	20.42	6.97	100
Whole period	28.90	21.67	17.61	15.60	16.22	100

Notes: Population-weighted using KFS survey weights. Quantile 1 is the bottom of the distribution while quantile 5 is the top of the distribution. The first cell on the top left means that 38.39% of exiting firms in 2004 were in the bottom quantile of the 2004 productivity distribution.

Additionally, for each quantile we computed both the share of firms that exited until the end of each period and the corresponding percentage of survivors (Table 13). We can observe that a substantial share of low-productivity and low-profitability firms has a significant degree of resilience. Indeed, 93% of the firms belonging to the lowest quantile of the productivity distribution at time $t-1$ do not exit at at time t. We can also notice that a relevant part of high-productivity and high-profitability firms exit the market. In particular, if we carefully analyze the quantiles of the profitability distribution, we see that the most profitable firms (5th quantile) have a share of exiting firms which is higher compared to the central parts of the distribution while this is not the case for labour productivity. Again, this might hint at a non-linearity in the relationship between survival and profitability.

Table 13: Transition rates: LP and ROS and exit

Quantile	Surviving	Exiting	Surviving	Exiting	
	LP_t		ROS_t		
1	90.87	9.13	90.64	9.36	100
2	93.03	6.97	93.69	6.31	100
3	93.44	6.56	94.22	5.78	100
4	94.50	5.50	94.64	5.36	100
5	95.73	4.27	94.34	5.66	100
	LP_{t-1}		ROS_{t-1}		
1	92.75	7.25	91.33	8.67	100
2	93.79	6.21	94.20	5.80	100
3	93.86	6.21	95.42	4.58	100
4	96.21	3.79	95.19	4.81	100
5	95.41	4.59	94.89	5.11	100

Notes: Population-weighted using KFS survey weights. Quantile 1 is the bottom of the distribution while quantile 5 is the top of the distribution. The first cell on the top left means that, at time t, 90.87% of firms in the lowest quantile of the productivity distribution survived in the same year. The last cell on the bottom left means that 95.41% of firms in the highest quantile of the productivity distribution at $t-1$ suvived at time t.

The statistics outlined above suggest that among exiting firms a considerable amount comes from firms located at the lower tails of the productivity and profitability distributions. Notwithstanding this, we also observe that a non-negligible share of under-performing firms is able to remain in the market.

To explicitly test whether 'fitness' variable do significantly drive exit dynamics, while controlling for other covariates, we estimate a complementary log-log model. Most studies conducting survival analysis employ the Cox Proportional Hazard (CPH) model, where the covariates determine differences across firms with respect to the baseline hazard model. However, the CPH assumes continuous survival time and exact ordering of firms with respect to their failure time (Jenkins, 2005). In our case, since we have annual data, we are only able to observe failures at discrete intervals without being able to order them within each year. Hence, we employ a discrete time model to investigate the relationship between 'fitness' variables and new firm survival. A central concept in survival analysis is the hazard rate h, which can be defined as the probability that a firm exits the market at time t given that it has survived until t, conditional on a vector of co-variates x_{it}. The discrete time duration model can be estimated by binary variable methods, and time varying co-variates can be included (Jenkins, 2005). In order to be estimated, the hazard function requires the specification of a functional form. Following Prentice and Gloeckler (1978), we assume the hazard rate h to be distributed as a complementary log-log function, as it has a convenient property of representing the discrete time representation of an underlining continuous time proportional hazard model:

$$h(x_{it}) = 1 - \exp[-\exp(\beta_0 + x'_{it}\beta + \gamma_t)] \tag{5}$$

where x_{ik} is vector of time-varying regressors including the same variables as in the first part of the analysis. We must stress that the complementary log-log model allows to account for unobserved but systematic differences across firms (also known as 'frailty'). However, we run the complementary log-log model with unobserved heterogeneity and we could not reject the null hypothesis that the frialty variance component is equal to zero at the 1% significance level. Therefore, the reported estimates are obtained under the hypothesis of no unobserved heterogeneity. As for the analysis carried out so far, the estimates presented successively should be interpreted as associations and not causal effects.

Before presenting the results, it is important to stress that, due to the relatively low number of exits in some of the sectors in our sample, we are not able to conduct a split-sample analysis for different industries as in the previous sections of the paper.

In interpreting the point estimates, note that the negative coefficients in these models denote an increase in the likelihood of surviving. Quite surprisingly, the estimates indicate that relative ROS does not appear to significantly affect firm survival (Tables 14 and 15). On the contrary, relative labour productivity has a significant relation with new firm failure (Tables 16 and 17). In particular, a 1% increase in $LP_{i,t}$ decreases the probability of failure by 0.087%, *ceteris paribus*[20]. The model considering both changes and levels of labour productivity (Table 15) confirms the importance of firm efficiency levels in determining the probability of failure (a 1% increase in $\overline{LP}_{i,t}$ decreases the probability of going out of business by 0.15%) while *changes* in efficiency are not significant.

As in the analysis focusing on the *growth of the fitter* principle, our goal is to understand how strong the 'cleansing' effect of market selection mechanisms is with respect to the least fit firms. Therefore, we apply the Shapley R^2 decomposition technique employed for the firm growth analysis (Huettner et al., 2012). As argued in Cox and Snell (1989), the R^2 statistic can be applied to other regression models (such as the complementary log-log model or the CPH) where the maximum likelihood is the criterion of fit (Coad et al., 2016). The Cox-Snell R^2 represents one of the alternatives:

$$\text{Cox-Snell } R^2 = 1 - \left\{ L(0)/L(\hat{\beta}) \right\}^{\frac{2}{n}}$$

where $L(0)$ and $L(\hat{\beta})$ indicate, respectively, the likelihood of the 'null' model and the one of the fitted model. Given that Cox-Snell R^2 reaches a maximum value that is lower than unity for discrete models (Nagelkerke, 1991), we opted for the Nagelkerke R^2:

$$\text{Nagelkerke } R^2 = \text{Cox-Snell} R^2 / [\max(R^2)]$$

$$\text{where } \max\,(R^2) = 1 - L(0)^{\frac{2}{n}}$$

Successively, we applied the method by Huettner et al. (2012) to decompose the Nagelkerke R^2 into Shapley values for the different regressors.

[20] This is calculated as follows: $[1 - \exp(-0.091)] \times 1\% = 0.087\%$.

Table 14: Exit and Profitability

	(1)	R^2 share	Explained variance
ROS_t	0.0001	1.68	**0.78**
	(0.007)		
ROS_{t-1}	−0.001	0.07	
	(0.008)		
Size_{t-1}	−0.170*	2.61	1.17
	(0.092)		
Year FE	(7 d)***	23.93	10.75
Sector FE	(21 d)***	26.64	11.96
State FE	(51 d)	45.06	20.23
Constant	−2.155		
	(1.390)		
Total		100	44.91
N. obs.	4,625		
N. firms	980		
Nagelkerke R^2	44.91		

Table 15: Exit and Profitability changes vs. levels

	(1)	R^2 share	Explained variance
ΔROS_t	0.0006	0.67	**0.30**
	(0.004)		
$\overline{ROS}_t$	−0.001	1.08	**0.49**
	(0.012)		
Size_{t-1}	−0.170*	2.61	1.17
	(0.092)		
Year FE	(7 d)***	23.93	10.75
Sector FE	(21 d)***	26.64	11.96
State FE	(51 d)	45.06	20.23
Constant	−2.155		
	(1.389)		
Total		100	44.91
N. obs.	4,625		
N. firms	980		
Nagelkerke R^2	44.91		

Notes: Complementary log-log model estimates using KFS survey weights. The dependent variable is a dummy taking the value 1 if a firm is out of business and 0 otherwise. ***, ** and * indicate significance on a 1%, 5% and 10% level for individual variables (t-test) or groups of dummy variables (F-test). The values in the "Explained variance" column represent the contribution of each variable to the total variation of exit rates. If $ROS_{i,t}$ increases by 1%, then the hazard rate increases by 0.0001%: $[1 - \exp(0.0001)] \times 1\% = 0.0001\%$.

Table 16: Exit and Productivity

	(1)	R^2 share	Explained variance
LP_t	−0.091***	7.07	**6.48**
	(0.026)		
LP_{t-1}	−0.067**	6.02	
	(0.027)		
Size_{t-1}	−0.110	1.58	0.78
	(0.094)		
Year FE	(7 d)***	22.10	10.93
Sector FE	(21 d)***	23.45	11.60
State FE	(51 d)*	39.79	19.68
Constant	−2.229		
	(1.406)		
Total		100	49.47
N. obs.	4,625		
N. firms	980		
Nagelkerke R^2	49.47		

Table 17: Exit and Productivity changes vs. levels

	(1)	R^2 share	Explained variance
ΔLP_t	−0.012	0.06	**0.03**
	(0.021)		
$\overline{LP}_t$	−0.158***	13.03	**6.45**
	(0.032)		
Size_{t-1}	−0.110	1.58	0.78
	(0.094)		
Year FE	(7 d)***	22.10	10.93
Sector FE	(21 d)***	23.45	11.60
State FE	(51 d)*	39.79	19.68
Constant	−2.229		
	(1.406)		
Total		100	49.47
N. obs.	4,625		
N. firms	980		
Nagelkerke R^2	49.47		

Notes: Complementary log-log model estimates using KFS survey weights. The dependent variable is a dummy taking the value 1 if a firm is out of business and 0 otherwise. ***, ** and * indicate significance on a 1%, 5% and 10% level for individual variables (t-test) or groups of dummy variables (F-test). The values in the "Explained variance" column represent the contribution of each variable to the total variation of exit rates. If $LP_{i,t}$ increases by 1%, then the hazard rate decreases by 0.087%: $[1 - \exp(-0.091)] \times 1\% = 0.087\%$.

In line with the findings derived from the survival analysis, the Shapley decomposition indicates that, among the two 'fitness' variables, labour productivity is once again the most important one. In fact, while current and lagged ROS accounts for a negligible 0.8%, current and lagged labour productivity account for roughly 6% of total variance in firm survival.

As for the *growth of the fitter* analysis, we are interested in testing whether exit is related to relative 'fitness' levels or *changes* thereof. In particular, for the *survival of the fitter* conjecture to hold, the static components of 'fitness' variables should have higher explanatory power compared with the dynamic ones. Results in Tables 15 and 17 indicate that almost the entire explanatory power in the case of both profitability and productivity is accounted for by relative 'fitness' levels thus providing support for the *survival of the fitter* principle. In particular, productivity levels represent 6.45% while productivity *changes* reach only around 0.03% of total variation in firm failure. Likewise, the change in ROS over two consecutive time periods only accounts for 0.30% while the average level of profitability reaches 0.49%.

However, it is clear that the explanatory power of firm 'fitness' variables is not strikingly high if compared with other co-variates included in the models. In fact, time, sectoral and State dummies together "explain" roughly 40%.

6 Robustness checks

We run a series of additional exercises in order to check the sensitivity of our results. First, one source of bias could be related to potential feedback effects or reverse causality from growth to both productivity and profitability[21], even if our results are in principle mostly relevant as an upper bound: any reverse causality effect would prove *a fortiori* the lack of selection force. However, we estimated the models presented in Section 3 including one lag of the dependent variable by means of OLS and FE. The coefficients regarding both profitability and labour productivity are slightly lower than the ones previously shown indicating that reverse causality can, as predicted, make us overestimate the effects of 'fitness' variables. We also re-run the R^2 decomposition exercises employing the above-mentioned specifications: results concerning the average marginal contributions of both productivity and profitability variables do not qualitatively change. The estimates regarding the lagged growth rates indicate a significant and negative association with current growth. In terms of its explanatory power, however, past growth does accounts only for around 1.5% of total variation in firm growth rates. Furthermore, in order to tackle additional sources of endogeneity, we estimated these equations via GMM-diff and GMM-sys estimators. Signs, patterns of statistical significance and magnitude of coefficients do not vary considerably when adopting this alternative specification, thus, we do not believe our estimates to be severely biased. If anything, we can argue that our results feature a potential positive bias from omitted variables and reverse causality.

Second, we further check our results by re-running the entire growth analysis employing the balanced panel. We do not observe substantial differences with the results shown in previous sections of the paper. To tackle potential heterogeneous rates of attrition in different sectors, we also performed the same test for the split-sample analysis and the differences among sectors closely match the ones observed using the unbalanced sample.

Moreover, given that firms in the sample undergo the Great Recession, we also employed sector-year fixed effects. By controlling for shocks common to all firms in a sector in a year, they allow us to remove any meso-scale business cycle. The results (available upon request) do not experience any significant change from the one shown in the article.

With respect to the survival analysis, we repeated the exercises by employing the CPH estimator instead of the complementary log-log model and results are practically unaltered. We also employed the Cox-Snell R^2 instead of the Nagelkerke R^2 and the estimates are essentially identical.

The survival models presented so far do not control for unobserved heterogeneity. Also, time duration estimators do not allow us to compute the contribution of firm fixed effects to the goodness-of-fit of the models. Hence, apart from running the baseline complementary log-log model, we augment it by including a series of time-invariant control variables to check the sensitivity of our results. In particular, we employ two types of co-variates considering both owner and firm level factors. At the owner level, we include additional variables related to the characteristics of the owner. In particular, new firms are endowed with knowledge and experience at birth through the human capital of their founder. Empirical evidence suggests that this pre-entry features will affect the firm's chances of survival (Dencker et al., 2009). Hence, we use the principal owner's years of work experience (*Exp*) as a proxy for pre-entry

[21] See Coad (2007), Dosi et al. (2015) and Federico and Capelleras (2015) for a discussion.

experience. We also employ an education measure (*Edu*) to proxy for owner's human capital which is not experience related. The measure takes a value from 1 to 10, where 1 indicates less than 9 years of education up to 10 which indicates a post-graduate degree. Furthermore, we control for the owner's age (*Age*) based on the conjecture that younger owners might have lower risk aversion, be more receptive to new ideas while older ones might be more anchored to traditional routines and more risk adverse (Persson, 2004). At the same time, a negative effect of age could be due to the fact that older owners might be more reluctant to exit given lower outside opportunities. Finally, we add the number of owners as an additional covariate given that there is some evidence that founding team size positively affects new ventures performance (Klotz et al., 2014). For instance, Bates et al. (2013), employing the KFS, find that larger teams of owners provide higher levels of experience and expertise which might foster new firm growth.

At the firm-level, we insert a dummy variable capturing whether the new firm is home-based or not (*Home-based*) given that previous research has shown that these businesses tend to be smaller and have more modest growth aspirations (Coleman and Robb, 2012). Moreover, we include in the estimation an additional dummy variable indicating whether a start-up is incorporated (as a corporation, partnership, or limited liability company) (*Corp*). Recent research has emphasized how the legal status of a firm at birth is an indication of the willingness to grow of its founders and represent a good predictor of significant performance outcomes (Guzman and Stern, 2016). Finally, we also considered a dummy variable representing whether a firm is high-tech (*High-tech*).

Results for these augmented models are consistent with those yielded in the case of the baseline model (see Tables A3 and A4 in Appendix 3). Indeed, the total variance in firm failure accounted for by profitability is still irrelevant, whereas productivity variables only marginally decrease their share to around 5%. However, in line with previous literature emphasizing the importance of non performance factors in explaining firm survival (Gimeno et al., 1997; DeTienne et al., 2008), the owner characteristics represent roughly 11% of the total variation in firm failure, a figure higher than the one for 'fitness' variables and that highlights the importance of pre-entry features. What seems also to be driving firm exit are 'environmental' variables such as macroeconomic shocks and age-related factors (both included in our time dummies[22]) and sectoral dynamics. Moreover, the State dummies represent the most relevant factor in line with previous literature showing the importance of location conditions in determining new firm survival (Fotopoulos and Louri, 2000). These three groups of dummies together account for around 40% of the explained variance in firm failure.

7 Conclusions

This study contributes to the analysis of the workings and strength of market selection in the case of a single cohort of US new independent firms. In particular, we explore whether and how effectively markets select newly-born firms according to their fit in terms of both labour productivity and profitability. In the first part of the analysis, by estimating the micro-relationships between proxies of firm relative 'fitness' and relative growth rates, we show that the latter are mainly driven by productivity while profitability plays a negligible role. Indeed, through an R^2 decomposition exercise, we find that productivity accounts for around 19-20% of total firm growth rates variation. Contrarily, relative profitability appears to play a marginal role accounting for only around 1.4%.

The evidence presented does reinforce the idea that selection forces do not operate according to a naively Schumpeterian or classical notion (Bottazzi et al., 2010) also in the case of newly-born firms. As argued by Coad (2009, p.108), "the mechanism of selection appears to be rather 'sub-optimal' in the sense that its effectiveness is lower than it could conceivably be". In fact, while industry dynamics models predict that firms would compete for growth opportunities and markets would select them rewarding the 'fitter' ones, our findings indicate that selection via growth and exit seems to operate in more roundabout ways. Moreover, the productivity-growth link is much stronger that the profitability-growth one[23], in line with recent empirical evidence regarding young firms (Federico and Capelleras, 2015). The selection process might indeed be more related to behavioural factors concerning the growth orientation of firms' managers and founders (Coad, 2009; Bottazzi et al., 2010) pricing strategies, willingness to invest and to expand might actually matter more than 'fitness'-based selection (Wiklund and Shepherd, 2003).

[22]However, disentangling which one of the two effects prevails represents an arduous task given that we examine only a single cohort and we restrain from doing it since it is beyond the scope of this article.

[23]On this we agree with Coad (2007, p.385) that "evolutionary models in the future would do better to abandon the assumption of a direct linear relationship between profits rates and growth rates, and replace it with an assumption of total independence between the two."

Furthermore, at odds with the *growth of the fitter* principle, we show that firm growth rates are influenced by *changes* in relative productivity rather than relative productivity levels as found in Dosi et al. (2014) and Dosi et al. (2015). As already discussed in the latter work, the fact that relative 'fitness' levels do not appear to drive firm dynamics also in the case of new firms might be due to the fact that selection works at finer levels of sectoral disaggregation that our empirical evidence is not able to capture[24]. The challenge of properly demarcating units and populations (firms vs products, industries vs markets/submarkets) over which selection forces operate has been taken up by Cantner et al. (2012). They propose a shift in the analysis from the firm to the product level of a narrowly defined markets (compact cars in Germany) and employ a measure of 'fitness' based on information over four main product characteristics and prices. Their findings indicate that at their finer resolution selection does operate according to a replicator dynamics mechanism.

It is also worth noticing the inter-sectoral differences in selection dynamics. Indeed, we find that 'fitness' variables tend to have a higher relevance for new firms in the service sector, a result consistent with the evidence that the speed of market selection is higher in services as compared with the manufacturing sector (Foster et al., 2006; Baldwin and Gu, 2011) and that this might be related to the presence of higher barriers to entry and exit in the latter (Hölzl, 2015). Moreover, our findings reveal that relative productivity appears to be more relevant for new firms operating in low tech and less concentrated sectors.

The negligible role of profitability is also confirmed when addressing selection via exit: current and lagged ROS do not have any significant effect on the likelihood of a firm going out of business. Contrarily, the *survival of the fitter* principle is supported in the case of labour productivity, even if modestly. Its average marginal contribution to the overall firm failure variation is only around 6%.

Our evidence adds support for a more nuanced approach to modelling and testing market selection, moving beyond the assumption that selection operates upon a single firm characteristic (productivity or profitability) but rather some combinations of firms and product features (for one of the first attempts in this direction see Holm et al. (2016)).

References

Anyadike-Danes, M., Bjuggren, C.-M., Gottschalk, S., Hölzl, W., Johansson, D., Maliranta, M., and Myrann, A. (2015). An international cohort comparison of size effects on job growth. *Small Business Economics*, 44(4):821–844.

Audretsch, D. B. (1995). *Innovation and industry evolution*. Mit Press.

Baden-Fuller, C. W. (1989). Exit from declining industries and the case of steel castings. *The Economic Journal*, 99(398):949–961.

Baily, M. N., Hulten, C., Campbell, D., Bresnahan, T., and Caves, R. E. (1992). Productivity dynamics in manufacturing plants. *Brookings papers on economic activity. Microeconomics*, 1992:187–267.

Baldwin, J. R. and Gu, W. (2006). Plant turnover and productivity growth in canadian manufacturing. *Industrial and Corporate Change*, 15(3):417–465.

Baldwin, J. R. and Gu, W. (2011). Firm dynamics and productivity growth: a comparison of the retail trade and manufacturing sectors. *Industrial and Corporate Change*, page dtq064.

Bartelsman, E. J. and Dhrymes, P. J. (1998). Productivity dynamics: Us manufacturing plants, 1972–1986. *Journal of productivity analysis*, 9(1):5–34.

Bates, T., Robb, A., and Parker, S. (2013). Utilizing the kauffman firm survey to predict growth in venture size and scope among small firm startups: 2004. *Washington, DC: SBA Office of Advocacyo. Document Number.*

Bellone, F., Musso, P., Nesta, L., and Quéré, M. (2008). Market selection along the firm life cycle. *Industrial and Corporate Change*, 17(4):753–777.

Bottazzi, G., Cefis, E., and Dosi, G. (2002). Corporate growth and industrial structures: some evidence from the italian manufacturing industry. *Industrial and Corporate Change*, 11(4):705–723.

Bottazzi, G., Coad, A., Jacoby, N., and Secchi, A. (2011). Corporate growth and industrial dynamics: Evidence from french manufacturing. *Applied Economics*, 43(1):103–116.

Bottazzi, G., Dosi, G., Jacoby, N., Secchi, A., and Tamagni, F. (2010). Corporate performances and market selection: some comparative evidence. *Industrial and Corporate Change*, page dtq063.

Bottazzi, G. and Secchi, A. (2006). Explaining the distribution of firm growth rates. *The RAND Journal of Economics*, 37(2):235–256.

Bottazzi, G. and Secchi, A. (2012). Productivity, profitability and growth: The empirics of firm dynamics. *Structural Change and Economic Dynamics*, 23(4):325–328.

Bottazzi, G., Secchi, A., and Tamagni, F. (2008). Productivity, profitability and financial performance. *Industrial and Corporate Change*, 17(4):711–751.

Cantner, U., Krüger, J. J., and Söllner, R. (2012). Product quality, product price, and share dynamics in the german compact car market. *Industrial and Corporate Change*, 21(5):1085–1115.

Carreira, C. and Teixeira, P. (2011). The shadow of death: analysing the pre-exit productivity of portuguese manufacturing firms. *Small Business Economics*, 36(3):337–351.

Chapple, K., Markusen, A., Schrock, G., Yamamoto, D., and Yu, P. (2004). Gauging metropolitan high-tech and i-tech activity. *Economic Development Quarterly*, 18(1):10–29.

[24]This could especially true in the case of new firms given that they might be operating in market niches or relatively new sectors where selection pressures might be of less importance (Srholec and Verspagen, 2012; Esteve Pérez et al., 2015).

Coad, A. (2007). Testing the principle of growth of the fitter: the relationship between profits and firm growth. *Structural Change and economic dynamics*, 18(3):370–386.

Coad, A. (2009). *The growth of firms: A survey of theories and empirical evidence*. Edward Elgar Publishing.

Coad, A. (2010). Exploring the processes of firm growth: evidence from a vector auto-regression. *Industrial and Corporate Change*, 19(6):1677–1703.

Coad, A. and Broekel, T. (2012). Firm growth and productivity growth: evidence from a panel var. *Applied Economics*, 44(10):1251–1269.

Coad, A., Frankish, J., Roberts, R. G., and Storey, D. J. (2013a). Growth paths and survival chances: An application of gambler's ruin theory. *Journal of Business Venturing*, 28(5):615–632.

Coad, A., Frankish, J. S., Roberts, R. G., and Storey, D. J. (2016). Predicting new venture survival and growth: Does the fog lift? *Small Business Economics*, 47(1):217–241.

Coad, A., Rao, R., and Tamagni, F. (2011). Growth processes of italian manufacturing firms. *Structural Change and Economic Dynamics*, 22(1):54–70.

Coad, A., Segarra, A., and Teruel, M. (2013b). Like milk or wine: Does firm performance improve with age? *Structural Change and Economic Dynamics*, 24:173–189.

Cochran, W. G. (1977). Sampling techniques. *New York: John Wiley and Sons*.

Coleman, S. and Robb, A. (2012). *A rising tide: Financing strategies for women-owned firms*. Stanford University Press.

Cowling, M. (2004). The growth–profit nexus. *Small Business Economics*, 22(1):1–9.

Cox, D. R. and Snell, E. J. (1989). *Analysis of binary data*, volume 32. CRC Press.

Criscuolo, C., Gal, P. N., and Menon, C. (2016). Do micro start-ups fuel job creation? cross-country evidence from the dynemp express database. *Small Business Economics*, pages 1–20.

Cucculelli, M. and Ermini, B. (2012). New product introduction and product tenure: What effects on firm growth? *Research Policy*, 41(5):808–821.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2016). Changing business dynamism: Volatility of shocks vs. responsiveness to shocks?

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

Delmar, F., McKelvie, A., and Wennberg, K. (2013). Untangling the relationships among growth, profitability and survival in new firms. *Technovation*, 33(8):276–291.

Dencker, J. C., Gruber, M., and Shah, S. K. (2009). Pre-entry knowledge, learning, and the survival of new firms. *Organization Science*, 20(3):516–537.

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

Disney, R., Haskel, J., and Heden, Y. (2003). Entry, exit and establishment survival in uk manufacturing. *The Journal of Industrial Economics*, 51(1):91–112.

Dosi, G. (2007). Statistical regularities in the evolution of industries. a guide through some evidence and challenges for the theory. In Malerba, F. and Brusoni, S., editors, *Perspectives on Innovation*, pages 1110–1121. Cambridge University Press: Cambridge, UK.

Dosi, G. and Grazzi, M. (2006). Technologies as problem-solving procedures and technologies as input–output relations: some perspectives on the theory of production. *Industrial and Corporate Change*, 15(1):173–202.

Dosi, G., Grazzi, M., Lei, J., and Yu, X. a. (2014). Productivity, profitability, investment and corporate growth: Evidence from chinas manufacturing. *DRUID Conference*, 62(1):53–82.

Dosi, G., Lechevalier, S., and Secchi, A. (2010). Introduction: interfirm heterogeneitynature, sources and consequences for industrial dynamics. *Industrial and Corporate Change*, 19(6):1867–1890.

Dosi, G., Marsili, O., Orsenigo, L., and Salvatore, R. (1995). Learning, market selection and the evolution of industrial structures. *Small Business Economics*, 7(6):411–436.

Dosi, G., Moschella, D., Pugliese, E., and Tamagni, F. (2015). Productivity, market selection, and corporate growth: comparative evidence across us and europe. *Small Business Economics*, pages 1–30.

Dosi, G., Pereira, M. C., and Virgillito, M. E. (2016). The footprint of evolutionary processes of learning and selection upon the statistical properties of industrial dynamics. *Industrial and Corporate Change*, page dtw044.

Du, J. and Temouri, Y. (2015). High-growth firms and productivity: evidence from the united kingdom. *Small business economics*, 44(1):123–143.

Ericson, R. and Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, 62(1):53–82.

Esteve Pérez, S., Pieri, F., and Rodriguez, D. (2015). Age and productivity as determinants of firm survival over the product life cycle. Technical report, Department of Applied Economics II, Universidad de Valencia.

Farhat, J. B. and Robb, A. (2014). Applied survey data analysis using stata: The kauffman firm survey data. *Available at SSRN 2477217*.

Federico, J. S. and Capelleras, J.-L. (2015). The heterogeneous dynamics between growth and profits: the case of young firms. *Small Business Economics*, 44(2):231–253.

Foster, L., Grim, C., Haltiwanger, J., and Wolf, Z. (2016). Firm-level dispersion in productivity: Is the devil in the details? *American Economic Review*, 106(5):95–98.

Foster, L., Haltiwanger, J., and Krizan, C. J. (2006). Market selection, reallocation, and restructuring in the us retail trade sector in the 1990s. *The Review of Economics and Statistics*, 88(4):748–758.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *The American Economic Review*, 98(1):394–425.

Foster, L., Haltiwanger, J. C., and Krizan, C. J. (2001). Aggregate productivity growth. lessons from microeconomic evidence. In *New developments in productivity analysis*, pages 303–372. University of Chicago Press.

Fotopoulos, G. and Louri, H. (2000). Location and survival of new entry. *Small Business Economics*, 14(4):311–321.

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

Geroski, P. A. and Mazzucato, M. (2002). Myopic selection. *Metroeconomica*, 53(2):181–199.

Gilbert, B. A., McDougall, P. P., and Audretsch, D. B. (2006). New venture growth: A review and extension. *Journal of*

management, 32(6):926–950.
Gimeno, J., Folta, T. B., Cooper, A. C., and Woo, C. Y. (1997). Survival of the fittest? entrepreneurial human capital and the persistence of underperforming firms. *Administrative science quarterly*, pages 750–783.
Guzman, J. and Stern, S. (2016). Nowcasting and placecasting: Entrepreneurial quality and performance. In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*. University of Chicago Press.
Haltiwanger, J., Jarmin, R., Kulick, R., and Miranda, J. (2015). High growth young firms: Contribution to job growth, revenue growth and productivity. In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*. University of Chicago Press.
Heckman, J. J. (1977). Sample selection bias as a specification error (with an application to the estimation of labor supply functions).
Holm, J. R., Andersen, E. S., and Metcalfe, J. S. (2016). Confounded, augmented and constrained replicator dynamics. *Journal of Evolutionary Economics*, 26(4):803–822.
Hölzl, W. (2015). Sunk costs and the speed of market selection. *Journal of Evolutionary Economics*, 25(2):323–344.
Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150.
Huettner, F., Sunder, M., et al. (2012). Axiomatic arguments for decomposing goodness of fit according to shapley and owen values. *Electronic Journal of Statistics*, 6:1239–1250.
Hurst, E. and Pugsley, B. W. (2011). What do small businesses do? Technical report, National Bureau of Economic Research.
Hyytinen, A. and Maliranta, M. (2013). Firm lifecycles and evolution of industry productivity. *Research Policy*, 42(5):1080–1098.
Jang, S. S. and Park, K. (2011). Inter-relationship between firm growth and profitability. *International Journal of Hospitality Management*, 30(4):1027–1035.
Jenkins, S. P. (2005). Survival analysis. *Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK.*
Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica: Journal of the Econometric Society*, pages 649–670.
Karakaya, F. (2000). Market exit and barriers to exit: Theory and practice. *Psychology & Marketing*, 17(8):651–668.
Kim, P. H., Aldrich, H. E., and Keister, L. A. (2006). Access (not) denied: The impact of financial, human, and cultural capital on entrepreneurial entryin the united states. *Small Business Economics*, 27(1):5–22.
Klepper, S. (1996a). Entry, exit, growth, and innovation over the product life cycle. *The American economic review*, pages 562–583.
Klepper, S. (1996b). Entry, exit, growth, and innovation over the product life cycle. *The American economic review*, pages 562–583.
Klepper, S. (2002). Firm survival and the evolution of oligopoly. *RAND journal of Economics*, pages 37–61.
Klotz, A. C., Hmieleski, K. M., Bradley, B. H., and Busenitz, L. W. (2014). New venture teams a review of the literature and roadmap for future research. *Journal of Management*, 40(1):226–255.
Lee, S. (2014). The relationship between growth and profit: evidence from firm-level panel data. *Structural Change and Economic Dynamics*, 28:1–11.
Litwin, A. S. and Phan, P. H. (2013). Quality over quantity: Reexamining the link between entrepreneurship and job creation. *Industrial & Labor Relations Review*, 66(4):833–873.
Markman, G. D. and Gartner, W. B. (2002). Is extraordinary growth profitable? a study of inc. 500 high-growth companies*. *Entrepreneurship theory and practice*, 27(1):65–75.
Metcalfe, J. S. (1994). Evolutionary economics and technology policy. *The economic journal*, 104(425):931–944.
Meyer, M. W. and Zucker, L. G. (1989). *Permanently failing organizations*. Sage Publications, Inc.
Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78(3):691–692.
Navaretti, G. B., Castellani, D., and Pieri, F. (2014). Age and firm growth: evidence from three european countries. *Small Business Economics*, 43(4):823–837.
Nightingale, P. and Coad, A. (2013). Muppets and gazelles: political and methodological biases in entrepreneurship research. *Industrial and Corporate Change*, page dtt057.
Pakes, A. and Ericson, R. (1998). Empirical implications of alternative models of firm dynamics. *Journal of Economic Theory*, 79(1):1–45.
Persson, H. (2004). The survival and growth of new establishments in sweden, 1987-1995. *Small Business Economics*, 23(5):423–440.
Prentice, R. L. and Gloeckler, L. A. (1978). Regression analysis of grouped survival data with application to breast cancer data. *Biometrics*, pages 57–67.
Shapley, L. (1953). A value for n-person games (1953). *Contribution to the Theory of Games.*
Silverberg, G., Dosi, G., and Orsenigo, L. (1988). Innovation, diversity and diffusion: a self-organisation model. *The Economic Journal*, 98(393):1032–1054.
Srholec, M. and Verspagen, B. (2012). The voyage of the beagle into innovation: explorations on heterogeneity, selection, and sectors. *Industrial and Corporate Change*, 21(5):1221–1253.
Thornhill, S. (2006). Knowledge, innovation and firm performance in high-and low-technology regimes. *Journal of business venturing*, 21(5):687–703.
Van Beveren, I. (2012). Total factor productivity estimation: A practical review. *Journal of economic surveys*, 26(1):98–128.
Warusawitharana, M. (2014). Profitability and the lifecycle of firms. *Available at SSRN 1568965.*
Wiklund, J. and Shepherd, D. (2003). Aspiring for, and achieving growth: The moderating role of resources and opportunities. *Journal of management studies*, 40(8):1919–1941.
Winter, S. G. (1984). Schumpeterian competition in alternative technological regimes. *Journal of Economic Behavior & Organization*, 5(3-4):287–320.
Winter, S. G., Kaniovski, Y. M., and Dosi, G. (2003). A baseline model of industry evolution. *Journal of Evolutionary Economics*, 13(4):355–383.
Winter, S. G. and Nelson, R. R. (1982). An evolutionary theory of economic change. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.*

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Zarutskie, R. and Yang, T. (2015). How did young firms fare during the great recession? evidence from the kauffman firm survey. In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*. University of Chicago Press.

Appendix 1: Survivorship bias[25]

Survivorship (or attrition) bias has been pointed out as a major shortcoming in growth studies (Garnsey et al., 2006). Attrition refers to the permanent loss of sample members from a longitudinal sample. This generates a potential bias if those firms withdrawing from the original sample are systematically different from those that stay in the sample throughout the whole time span covered by the survey. In case this happens, the remaining sample, represents a population different from the original one (Cochran, 1977). Without correcting for it, estimates would not generalize back to the original population of firms, but rather, to the subpopulation of surviving firms.

The literature has traditionally addressed this issue using three main methods (Farhat and Robb, 2014). First, surveyors can attempt to limit attrition by tracking every single participant through repeated attempts to contact them. KFS surveyors do this and the issue of firms truly disappearing is a rather minor one, as just 3% of firms in the original sample truly vanish from the data set without any explanation over the entire data collection time-span. However, our main issue concerns those new firms that surveyors are able to locate and include in the KFS, but whose data are missing because they have voluntarily or involuntarily ceased operations. By the last year of data, this figure reaches around 50% of the original sample of new firms. This basically leaves us with two options: a two-stage sample selection model (Heckman, 1977) or its semi-parametric version (Wooldridge, 2010). Alternatively, one can exploit complex sample weighting to adjust for systematic sample attrition in each subsequent wave of the survey. This procedure entails that weights are recalculated and reassigned to each surviving sample unit in each subsequent round to ensure that each wave's resulting sample continues to represent the original population from which the original sample was drawn. We decided to rely on weights rather than the above alternatives for a number of reasons. First, the effective use of a sample selection model requires, among other things, the identification of at least one variable that drives firm survival without having any systematic relationship with the dependent variables of interest, a condition that is generally arduous to satisfy in practice (Wooldridge, 2010). Moreover, since in our case the dependent variable is firm growth, it is not trivial to find a covariate related to survival but with np systematic relationship with growth. Second, in those cases in which survey designers foresaw the survival issue at the time of survey design, as in the case of the KFS, estimates that rely on weighting and re-weighting with each subsequent survey wave perform better than two-stage estimates on unweighted data (Farhat and Robb, 2014). This occurs since the weights map the sample in each follow-up survey back to represent the original sample from the baseline. In the baseline year, weights are constructed to account for unequal sampling probabilities that derive from intentional oversampling of some businesses and undersampling of others, that is, sample selection bias. Then, with each subsequent administration of the survey, weights are first adjusted to compensate for nonresponse, meaning attrition, and adjusted once more to ensure the surviving population represents the original one. This means that by employing probability weights, we are not only addressing survival bias: the weights ensure that estimates pertaining to the original sample – which strategically oversampled some firms and undersampled others – also generalize to the intended original population of new businesses. Furthermore, we would want to use survey weights to guarantee generalizability to the entire universe of new firms established in 2004. Therefore, relying on weights works to address sampling bias as well as the larger challenge arising from survival bias. Finally, incorporating survey design features also enables us to partially relax the usual assumption of conditional independence across all observations, noted above, which would otherwise have to be done by estimating clustered standard errors.

[25]This part is extensively drawn from Farhat and Robb (2014) and Litwin and Phan (2013).

Appendix 2: Shapley decomposition

In game theory, the Shapley value (Shapley, 1953) is a way to distribute fairly the total gains of a cooperative game among players. It can be proven that this technique, when adapted in a natural way to distribute the contribution to explained variance of different regressor, is the only possible share to have the following desirable properties (Huettner et al., 2012):

- *Proper decomposition*: the model variance is to be decomposed into shares, that is, the sum of all shares has to be the model variance;
- *Non-negativity*: all shares have to be non-negative;
- *Exclusion*: the share allocated to a regressor with coefficient $= 0$, should be 0;
- *Inclusion*: a regressor with coefficient $\neq 0$ should receive a nonzero share.

To illustrate the procedure to compute the Shapley values, let us start with an empty model and let us compute the explained variance R^2 by adding all the k regressors in a given order P until achieving the following full model:

$$y = \alpha + \sum_{i=1}^{k} \beta_i x_i + \epsilon$$

The Shapley value for a particular regressor j is computed as the average marginal contribution to the R^2 of adding regressor j to the model over all $k!$ possible permutations P of the k regressors. Defining P_i as the set of regressors preceding i in the permutation P, the part R^2_j of the total R^2 assigned to regressor j can therefore be written as:

$$R^2_j = \frac{1}{k!} \sum_{P \subseteq K} R^2(P_j \cup \{x_j\}) - R^2(P_j)$$

For a more detailed discussion refer to Huettner et al. (2012).

Appendix 3: Complementary tables

Table A1: High-Tech industries

NAICS code	Industry
3254	Pharmaceutical and medicine manufacturing
3332	Industrial machinery manufacturing
3333	Commercial and Service Industry Machinery Manufacturing
3341	Computer and peripheral equipment manufacturing
3342	Communications equipment manufacturing
3343	Audio and video equipment manufacturing
3344	Semiconductor and other electrnic component manufacturing
3345	Navigational, measuring, electromedical, and control instruments manufacturing
3364	Aerospace product and parts manufacturing
5112	Software publishers
5161	Internet publishing and broadcasting
5179	Other telecommunications
5181	Internet service providers and Web search portals
5182	Data processing, hosting and related services
5413	Architectural, engineering, and related services
5415	Computer systems design and related services
5413	Architectural, engineering, and related services
5417	Scientific R&D services

Notes: Classification drawn from KFS based on Chapple et al. (2004).

Table A2: High-Concentration and Low-Concentration industries

NAICS code	Industry	20 firms	50 firms
22	Utilities	**44,5**	**70,1**
31-33	Manufacturing	**39,6**	**52,1**
42	Wholesale Trade	16,6	24,9
44-45	Retail Trade	25,4	33,3
48-49	Transportation and Warehousing	**34,9**	**42,7**
51	Information	**49,9**	**62,1**
52	Finance and Insurance	**28,5**	**46,1**
53	Real Estate and Rental and Leasing	16,3	**26,1**
54	Professional, Scientific, Technical and Administrative Services	12,4	18,3
55-56	Management of Companies and Enterprises	15,2	23
61	Educational Services	15,3	22,3
62	Health Care and Social Assistance	9,2	15,1
71	Arts, Entertainment, and Recreation	12,5	19,5
72	Accommodation and Food Services	17,4	23,7
81	Other Services	7	11,3
Mean	27,03	37,18	
Median	17,4	26,1	

Notes: Data from US Economic Census 2007 retrievable at `https://www.census.gov/econ/concentration.html`. Columns 3 and 4 report the share of output accounted for by the largest 20 and 50 companies in each sector. Sectors in bold represent those in which these shares are above the median.

Table A3: Exit and Profitability - augmented model

	Coef.	Explained variance
ROS_t	0.000	
	(0.007)	**0.39**
ROS_{t-1}	−0.000	
	(0.008)	
Size_{t-1}	−0.074	0.87
	(0.111)	
$\text{Edu}_{t=1}$	−0.147***	
	(0.043)	
$\text{Exp}_{t=1}$	−0.034***	
	(0.008)	**11.51**
$\text{Age}_{t=1}$	0.015*	
	(0.008)	
N. $\text{owners}_{t=1}$	−0.231	
	(0.144)	
$\text{Home}_{t=1}$	0.271	
	(0.179)	
$\text{Corp}_{t=1}$	0.259	
	(0.182)	**2.18**
High-$\text{tech}_{t=1}$	−0.327	
	(0.305)	
Year FE	(7 d)***	9.58
Sector FE	(21 d)***	9.92
State FE	(51 d)*	19.03
Constant	−1.444	
	(1.578)	
Total		53.49
N. obs.	4,625	
N. firms	980	
Nagelkerke R^2	53.49	

Table A4: Exit and Productivity - augmented model

	Coef.	Explained variance
LP_t	−0.101***	
	(0.028)	**4.89**
LP_{t-1}	−0.051*	
	(0.028)	
Size_{t-1}	−0.050	0.63
	(0.113)	
$\text{Edu}_{t=1}$	−0.151***	
	(0.113)	
$\text{Exp}_{t=1}$	−0.032***	
	(0.008)	**10.82**
$\text{Age}_{t=1}$	0.014*	
	(0.008)	
N. $\text{owners}_{t=1}$	−0.210	
	(0.147)	
$\text{Home}_{t=1}$	0.164	
	(0.183)	
$\text{Corp}_{t=1}$	0.323	
	(0.184)	**1.96**
High-$\text{tech}_{t=1}$	−0.326	
	(0.407)	
Year FE	(7 d)***	9.88
Sector FE	(21 d)***	9.71
State FE	(51 d)*	18.66
Constant	−1.442	
	(1.584)	
Total		56.55
N. obs.	4,625	
N. firms	980	
Nagelkerke R^2	56.55	

Notes: Complementary log-log model estimates using KFS survey weights. The dependent variable is a dummy taking the value 1 if a firm is out of business and 0 otherwise. ***, ** and * indicate significance on a 1%, 5% and 10% level for individual variables (t-test) or groups of dummy variables (F-test). The values in the "Explained variance" column represent the contribution of each variable to the total variation of exit rates.

Appendix 4: Robustness checks with TFP and ROA

We repeated the whole analysis using alternative measures of both productivity and profitability. In particular, for productivity we resorted to a TFP measure calculated as the residual $e_{i,t}$ of the following equation:

$$y_{i,t} = \beta_0 + \beta_l \log L_{i,t} + \beta_k \log K_{i,t} + e_{i,t} \tag{6}$$

where $y_{i,t}$ is the log of value added of firm i, $L_{i,t}$ the log of the number of employees, and $K_{i,t}$ the log of assets[26]. Concerning the alternative measure for profitability, we adopted Returns on Assets (ROA) measured as earnings before taxes and interests over total assets.

The results (see Tables A5 and A6) broadly confirm the picture provided in the previous sections of the paper. As expected, the exercises conducted employing TFP show that productivity accounts for a larger share of the total explained variance of firm growth rates if compared with labour productivity. Indeed, while the latter represents around 20% of the variance in growth rates, TFP reaches around 35%. The estimates using ROA as alternative proxy for profitability yield qualitatively similar results. ROA accounts for a slightly larger share of explained variance in total firm growth variation (3%). The results concerning the heterogeneous pressures of market selection across sectors are largely confirmed also when performing the estimations with these alternative variables. Likewise, ROA variables represent a negligible share of total variation of exit rates while TFP, accounting for around 4%,“explains” even less if compared with labour productivity (see Tables A7 and A8).

Table A5: Growth and ROA changes vs. levels

	OLS			FE		
	Coef.	R^2 share	Expl. variance	Coef.	R^2 share	Expl. variance
ΔROA_t	0.0004*** (4.99e−5)	13.54	**1.93**	0.0005*** (6.96e-5)	5.52	**1.94**
$\overline{ROA}_t$	−0.0007*** (0.0001)	5.75	**0.82**	-0.0006*** (0.0001)	1.91	**0.67**
Size$_{t-1}$	−8.07e−5 (0.023)	0.07	0.01	-0.117*** (0.046)	0.75	0.07
Year FE	(7 d)***	68.72	9.79	(7 d)***	27.50	9.68
Sectoral FE	(21 d)	4.55	0.65	-	-	-
State FE	(51 d)	7.36	1.05	-	-	-
Firm FE	-	-	-	(920 d)***	64.88	22.84
Constant	−0.571 (0.395)			-0.225*** (0.044)		
Total		100	14.24		100	44.41
N. obs.	3,899			3,899		
N. firms	989			989		
R^2	14.24			44.41		

Notes: Pooled OLS and FE estimates using KFS survey weights. ***, ** and * indicate significance on a 1%, 5% and 10% level for individual variables (t-test) or groups of dummy variables (F-test). For the latter, we provide in parenthesis the number of variables for each group of dummies. In the “R^2 share” column we report the Shapley values while in the “Expl. variance” we present the contribution of each variable to the total variation of growth rates. Sectoral and State dummies are not included in the FE model since they are time-invariant variables.

[26] We do not dispose of data on materials required to apply the Levinsohn-Petrin methods. However, as argued by Van Beveren (2012), the simple TFP measure is highly correlated with the TFP derived from more sophisticated estimators. Moreover, although the revenue-based TFP embeds both technical efficiency and demand factors, Foster et al. (2016) show that TFP based on physical quantity are highly correlated with revenue-based TFP measures.

Table A6: Growth and TFP changes vs. levels

	OLS			FE		
	Coef.	R^2 share	Expl. variance	Coef.	R^2 share	Expl. variance
ΔTFP_t	0.670***	78.49	**41.59**	0.656***	55.93	**35.82**
	(0.029)			(0.036)		
$\overline{TFP}_t$	0.074***	3.19	**1.69**	-0.221***	3.09	**1.98**
	(0.019)			(0.036)		
Size_{t-1}	−0.105	0.51	0.27	-0.561***	1.71	1.09
	(0.018)			(0.040)		
Year FE	(7 d)***	15.39	8.15	(7 d)***	12.60	8.07
Sectoral FE	(21 d)	0.84	0.44	-	-	-
State FE	(51 d)	1.59	0.84	-	-	-
Firm FE	-	-	-	(920 d)***	26.66	17.07
Constant	−0.713**			0.080**		
	(0.291)			(0.044)		
Total		100	52.99		100	44.41
N. obs.	3,899			3,899		
N. firms	989			989		
R^2	52.99			44.41		

Notes: see notes in Table A5.

Table A7: Exit and Profitability changes vs. levels

	(1)	R^2 share	Explained variance
ΔROA_t	0.0003**	3.46	**1.69**
	(0.0001)		
$\overline{ROA}_t$	−0.0002	2.20	**1.08**
	(0.0002)		
Size_{t-1}	−0.158**	2.64	1.29
	(0.092)		
Year FE	(7 d)***	22.70	11.11
Sector FE	(21 d)***	25.73	12.59
State FE	(51 d)	43.27	21.17
Constant	−1.374		
	(1.005)		
Total		100	48.93
N. obs.	6,911		
N. firms	1,280		
Nagelkerke R^2	48.93		

Table A8: Exit and Productivity changes vs. levels

	(1)	R^2 share	Explained variance
ΔTFP_t	−0.008	1.92	**0.97**
	(0.043)		
$\overline{TFP}_t$	−0.138***	6.92	**3.50**
	(0.045)		
Size_{t-1}	−0.151**	2.79	1.41
	(0.077)		
Year FE	(7 d)***	22.46	11.37
Sector FE	(21 d)***	24.77	12.54
State FE	(51 d)**	41.13	20.81
Constant	−1.511		
	(0.983)		
Total		100	50.60
N. obs.	6,911		
N. firms	1,280		
Nagelkerke R^2	50.60		

Notes: Complementary log-log model estimates using KFS survey weights. The dependent variable is a dummy taking the value 1 if a firm is out of business and 0 otherwise. ***, ** and * indicate significance on a 1%, 5% and 10% level for individual variables (t-test) or groups of dummy variables (F-test). The values in the "Explained variance" column represent the contribution of each variable to the total variation of exit rates.