UNIVERSITY of GREENWICH

FOUNDATION FOR EUROPEAN PROGRESSIVE STUDIES
FONDATION EUROPÉENNE D'ÉTUDES PROGRESSISTES

GPERC

Greenwich papers in political economy

Ecosystem complexity, firm learning and survival:

UK evidence on intra-industry age and size diversity as exit hazards

Eshref Trushin, Leicester University
Mehmet Ugur, University of Greenwich

Abstract

Firm age or size diversity in an industry is taken for granted but its implications for industry evolution and firm survival have remained below the radars of empirical research. We address this knowledge gap by drawing on an interdisciplinary theoretical framework informed by theoretical biology, organizational ecology and industrial organisation. We hypothesize that firms in more diverse industries are more likely to exit as a result of rugged fitness distributions where a global fitness optimum is less likely to emerge. We also hypothesize that investment in research and development (R&D) may counterbalance the adverse effect of diversity on survival by enabling the firm to engage in active learning about its market and technology niches. Evidence from discrete-time hazard estimators and an unbalanced panel dataset of 35,136 R&D-active UK firms lend support to these hypotheses. The findings remain robust to: (i) a battery of sensitivity checks, including step-wise estimations, different diversity measures and various firm cohorts; (ii) control for frailty and for a wide range of firm, industry, and macroeconomic factors considered in the survival literature; and (iii) taking account of direct effects of age, size and R&D intensity.

Year: 2018

No: GPERC58

GREENWICH POLITICAL ECONOMY RESEARCH CENTRE (GPERC)

Keywords: Diversity, complexity, firm survival, R&D, ecosystem

JEL codes: L2, C4, O32, O33

Acknowledgements: We thank Edna Solomon of the University of Greenwich for her comments on a previous version of this article.

Corresponding author: Mehmet Ugur

Ecosystem complexity, firm learning and survival:

UK evidence on intra-industry age and size diversity as exit hazards

Eshref Trushin, Leicester University

Mehmet Ugur, University of Greenwich

Abstract

Firm age or size diversity in an industry is taken for granted but its implications for industry evolution and firm survival have remained below the radars of empirical research. We address this knowledge gap by drawing on an interdisciplinary theoretical framework informed by theoretical biology, organizational ecology and industrial organisation. We hypothesize that firms in more diverse industries are more likely to exit as a result of rugged fitness distributions where a global fitness optimum is less likely to emerge. We also hypothesize that investment in research and development (R&D) may counterbalance the adverse effect of diversity on survival by enabling the firm to engage in active learning about its market and technology niches. Evidence from discrete-time hazard estimators and an unbalanced panel dataset of 35,136 R&D-active UK firms lend support to these hypotheses. The findings remain robust to: (i) a battery of sensitivity checks, including step-wise estimations, different diversity measures and various firm cohorts; (ii) control for frailty and for a wide range of firm, industry, and macroeconomic factors considered in the survival literature; and (iii) taking account of direct effects of age, size and R&D intensity.

Key words: Diversity, complexity, firm survival, R&D, ecosystem

JEL classification: L2, C4, O32, O33

Ecosystem complexity, firm learning and survival:

UK evidence on intra-industry age and size diversity as exit hazard

Introduction

Significant and persistent diversity of firm characteristics within an industry is well documented and puzzling (Dosi *et al*., 1997). Indeed, some contributors to the field argue that intra-industry firm diversity has been a driver for the development of strategic management (Noda and Collis, 2001:922). Yet little or no attention has been paid to how intra-industry type diversity (measured as size or age diversity) affects firm survival. The oversight is evident in three major disciplines where system (e.g., industry) evolution and constituent (e.g., firm) dynamics are important areas of study: biology, industrial organisation, and organizational ecology.

In biology, the debate on the relationship between species diversity and ecosystem stability is well documented (Elton, 1958; McCann, 2000; McNaughton, 1993). The received wisdom here is that species diversity contributes to eco-system stability either because of complementarity between species and/or due to asymmetric vulnerability of the diverse species to external shocks. What is usually overlooked in this line of research is the following question: even if diversity is conducive to ecosystem stability, what does it imply for the survival of the system's constituent parts?

A similar oversight is observable in the management literature, which tends to focus on the relationship between within-firm diversity and firm performance (Kilduff, Angelmar, and Mehra, 2000; Richard, 2000; Boone and Hendriks, 2009). Here the attention is on the performance of the firm as a micro ecosystem – at the expense of the performance or rewards enjoyed by the diverse actors or strategies that shape the diversity of the firm. The analogy is less evident in the case of industrial organisation (IO) literature as the latter does pay attention to how industry-level factors (e.g., the innovation regime or the level of market concentration in the industry) affect firm survival (Agarwal and Audretsch, 2001); Audretsch, 1991; Geroski, 1995). However, the implications of intra-industry type diversity for firm survival remains below the radars of the IO research effort too.

To address this knowledge gap, we propose and tests two novel hypotheses on the relationship between intra-industry type diversity and firm survival: (1) Intra-industry type diversity measured as age or size diversity is associated with higher exist hazards for constituent firms due to complexity of the fitness distribution in diverse industries, where the probability of suboptimal choices is higher and the selection of the fittest is less likely at the same time; and (2) Investment in research and development (R&D) reduce the risk of exit hazard in diverse industries by enabling firms to engage in active learning about their optimum market and/or technology niches.

Our hypotheses are informed by a theoretical framework that combines insights from four disciplines. From theoretical biology, we draw on Kauffman (1993; 2016) who demonstrates that diversity is a source of complexity that generates a rugged fitness distribution within the ecosystem. A rugged fitness distribution features multiple local optima, separated by shallow troughs and absence of a global optimum. In such settings, some constituents face a higher risk of exit hazard as the co-evolutionary pockets they belong to may be saddled on a local fitness optima that fall short of ensuring resilience to adverse shocks. The higher the complexity of the ecosystem is, the greater is the risk of belonging to a co-evolutionary cohort saddled on a sub-optimal fitness peak and hence the higher is the mortality risk in the face of adverse shocks.

We combine the insights from theoretical biology with those from the organisational ecology literature (Hannan and Freeman, 1989; Barnett and Hansen, 1996; Barnett, 2008). In this work, firm adaptation in highly diverse industries is costly and requires continued learning in dealing with shifting competitive landscapes and changing logics of competition through *coevolution* of different firms. Adaptation becomes increasingly difficult in a context-specific competitive “arms’ race”, especially when firms compete over similar resources. This is particularly clear in Barnett and Hansen (1996) and Barnett (2008), who find that firms become maladaptive when cohort diversity in an industry increases. In such setting, firms tend to suffer from organizational myopia, inertia, and uncertainty of routines, all of which are due to co-evolution of diverse competitors.

From industrial organisation (IO), we draw on Ericson and Pakes (1995) who demonstrates that heterogeneous firms that invest in R&D experience high level of mortality rates in the initial learning period, but this will be followed by longer survival time for R&D-active firms as the latter discover their optimal product/technology niches. This insight form IO resonates with findings from game theory with bounded cognitive abilities, where players (firms) with larger capacity for information processing (Challet and Zhang, 1998) or those with higher cognitive capacity (Strzalecki, 2014) are able to choose better strategies with higher payoffs. It also ties in with two strands of literature organisational learning: (i) the early contribution of Cohen and Levinthal (1990) who demonstrate why prior knowledge is a significant determinant of absorptive capacity; and (ii) later work that focuses on the co-evolution of knowledge environments, organizational forms and combinative capabilities as determinants of acquiring absorptive capacity (Grant, 1996; Bosch, Volberda, and Boer, 1999).

To test these hypotheses, we use two different measures of intra-industry type diversity: the *Theil entropy index* and the *coefficient of variation* of firm age and employment within 3-digit SIC industries. We report that firms within more diverse industries experience higher rates of exit hazard. We also report that intra-industry type diversity and R&D investment are substitutes in that higher levels of R&D intensity are associated with lower exit hazard at each level intra-industry diversity. These findings are robust to: (i) controlling for own (direct) effects of age and size; (ii) a battery of sensitivity checks, including step-wise estimations, different diversity specifications, firm cohorts and control for frailty; and (iii) a wide range of firm, industry, and macroeconomic factors that have been investigated as potential determinants of firm survival in the existing literature.

The rest of the paper is organised as follows. In the next section we review the related literature and derive our testable hypotheses. Then we present our dataset, which consists of 35,136 R&D-active UK firms from 1998-2012 with 158,316 firm-year observations. Here we also discuss the diversity measures, namely the Theil entropy index and coefficient of variation of firm age and size within 3-digit industries, and elaborate on our estimation strategy. The estimation results and their correspondence with existing work are discussed, followed with a conclusions that distil the main findings and discus their implications.

Related literature and research hypotheses

Although criticised by Hodgson (1993) for failing to fulfil the promise of moving from economic statics to dynamics as a necessary step for developing an adequate evolutionary analysis, Alfred Marshall (1898) can be considered as the pioneer of invoking biological analogies in economics. For Marshall, the theory of evolution in biology is useful in understanding how variety leads to a stable equilibrium through selection based on profitability. In this paper, we also invoke biological analogies into economics to investigate the effects of variety (diversity) not on industry stability, but on the survival chances of the constituents (firms).

For example, the "insurance hypothesis" in biology posits that diversity has a stabilizing effect on aggregate ecosystem properties. The effect is due to a-synchronic responses of the diverse species to environmental shocks (Loreau, 2010: 53; Yachi and Loreau, 1999). This argument is in line with earlier work by Elton (1958) and McNaughton (1993), who drew attention to a positive relationship between diversity and ecosystem stability in terms of population densities and compositions.

Applied to industry evolution, the insurance hypothesis would imply that higher levels of intra-industry diversity would be associated with stable firm density and composition. Such conclusions, however, would be contentious for two reasons – one theoretical and one empirical. Theoretically, May (1974) demonstrates that a growing number of differing species leads to larger fluctuations in ecosystem populations and could result in extinctions. Similarly, Yozdis (1980) reports that complex ecosystems might be more fragile due to high specialization and smaller scope for interaction between different species. In between, Lehman and Tilman (2000) argue that biodiversity may stabilize aggregate ecosystem properties such as biomass, but may also destabilize the dynamics of the system's populations.

Further doubt is cast by Huisman and Weissing (2001), who demonstrate that competition between heterogeneous species for three or more resources leads to chaotic fluctuations in species' populations; and the results (i.e., the winners) of such competition are not predictable even in deterministic settings. Moreover, the uncertainty about the results increases with more types of species competing over larger number of resources. Finally, Mougi and Kondoh (2012) demonstrates that the effect of diversity on ecosystem stability is non-monotonic as it is mediated through the types of between-species interactions, which can be antagonistic, competitive, or mutualistic.

The association between biological systems' diversity and stability is also contested empirically. According to Ives and Carpenter (2007), only 23 percent of findings from 13 studies indicate a positive and significant relationship, whereas 46 percent of the findings indicate a negative relationship, and the remaining 31 percent indicate a relationship close to zero. It appears that what matters for ecosystem stability is not species diversity, but the extent of heterogeneity in the responses of the species to environmental fluctuations. If species' responses are convergent, the biological system's diversity is a source of system instability (Ives and Carpenter, 2007).

More importantly, however, the insurance hypothesis is not helpful for the study of the relationship between ecosystem diversity and constituent survival. Irrespective of whether diversity is conducive to system stability, its effects on constituent survival cannot be deduced from its effects on the system as a whole. To uncover the relationship between system diversity and constituent survival, we need to trace the implications of the system diversity for adaptation strategies of the firms. A seminal work in theoretical biology (Kauffman, 1993) provides significant insights into how complexity is related to ruggedness (multi-peakedness) in the fitness distribution (the opportunity space) and how the latter traps the adaptive walks by the constituents.

In his NK(C) *complexity* model, Kauffman (1993) analyses how the self-organised properties of a system with N elements (traits or firms), K random connections (epistatic inputs) for each element, and C co-evolutionary (intercoupling) pockets can enhance or limit the "efficacy of the natural selection" measured as: (i) the fitness level of the elements or their sets that had adapted to past shocks; and/or (ii) the survival prospects of the sets depending on their positions on the fitness distribution. Kaufman (1993) demonstrates that a global maximum of selection efficacy, i.e., the survival of the fittest, is usually not achievable in *complex systems* where complexity lies between 'order' and 'chaos'.

Kauffman (1993: chapters 2 – 5) allows for a rich set of conclusions, *four* of which are the most pertinent for the analysis in our paper. The first states that system complexity is an increasing function of the number of constituents (N) and the number of co-evolutionary pockets (C). The second states that the fitness distribution becomes more rugged and multi-peaked - i.e., it will be characterized with a larger number of local optima and a higher incidence of conflicting constraints - as the random connections between constituents (K) increases. The third states that the average fitness level is lower the more multi-peaked is the fitness distribution. This is because constituents may never attain a global optimum due to combinational explosion of the possibility space. Finally, the fourth conclusion indicates that the survival prospect of constituents in a co-evolutionary pocket depends not only on the magnitude of the next shocks, but also on the past adaptation journey of the pocket and the position it occupies in the rugged fitness space. Adapting elements face moving fitness peaks as other elements change so that search for optimal fitness becomes complex and elusive.

McKelvey (1999) draws on the NK(C) model to demonstrate how firms should avoid excessive complexity by optimising the number of internal and external value-chain competencies. He reports that the firm's competitive advantage is maximised when organizational complexity (internal co-evolutionary density) is moderate; and that the level of external co-evolutionary density sets an upper limit to the benefits from internal complexity. Stated differently, high levels of internal and external value chain complexity is a hazard factor that reduces the firm's competitive advantage.

In this paper, we are interested in the implications of the NK(C) model for the relationship between complexity and firm survival when complexity is proxied by firm-type diversity (measured as intra-industry age or size variety). Two of the NK(C) model parameters - the random connections for each firm (K) and the co-evolutionary pockets within the industry (C) - are not observed directly in the data. As proxies for their product in the NK(C) model, we use two measures of type diversity: the Theil entropy index and the coefficient of variation of firm age and size. In what follows, we first discuss the relevance of age and size as type identifiers that mediate firm performance in general. Then, in section three, we discuss the strengths and drawbacks of the proposed diversity measures.

There is a wealth of literature on how firm age or size may act as a moderating (contingency) variable that conditions the relationship between an explanatory variable and an outcome variable in various models of firm performance. An early study, Hofer (1975), demonstrates that the effect of strategy on firm performance varies by firm size. This is confirmed by Smith *et al.* (1989) who also report that the relationship between strategy and performance varies by firm size. Vaccaro et al., (2012) report that different firm sizes call for different leadership types, with smaller and less complex firms more likely to benefit transactional leadership whereas larger and more complex organizations benefit more from transformational leaders. A moderating role for firm size is also reported in Zona *et al.* (2013) in the relationship between corporate governance characteristics and innovation outcomes; in Dean, Brown, and Bamford (1998) where responses of small US manufacturing firms to competition are faster and more flexible than those of larger companies with structural inertia; and in Hannan and Freeman

(1989) where the effects of company sunk costs, concertation, profitability, excess capacity and R&D on entry decisions differ by firm size groups.

A number of studies also report a moderating role for firm age. For example, Sørensen and Stuart (2000) find a significant association between firm ageing and rates of patenting and innovation novelty. Other studies report that firm (or venture) age has a moderating effect on the relationship between entrepreneurial orientation and firm growth (Anderson and Eshima, 2013; Lumpkin et al., 2006). These findings are in line with an earlier call for a systematic treatment of age as a moderating variable in Quinn and Cameron (1983), who argued that the criteria and methods needed for effective performance vary with age. More recent evidence in Coad et al., (2016) also accords age with a significant moderating role as R&D investment by young firms are significantly riskier than R&D investment by more mature firms.

Given this rich debate concerning the mediating (indirect) effects of age and size diversity on firm-related outcomes, we are surprised by the absence of control for *indirect effects* of age and size in firm survival models. This oversight is also in stark contracts with the theoretical insights of the complexity literature discussed earlier. As summarised by McKelvey (1999), excessive diversity increases system complexity and may undermine the firms' competitive advantage. This is due to uncertainty about the quality of the local optima and the difficulty in identifying the global optimum within rugged fitness landscapes associated with complex systems. This type of opportunity spaces increase the probability of suboptimal states (choices) and limit the selection of the fittest at the same time (Kauffman, 1993; Kauffman, 2016; McKelvey, 1999). Unexpected functional shifts in firm artefacts may lead to discontinuous evolution in an industry and emergent processes challenging adaptation, which is known as exaptation (Andriani and Cattani, 2016, Kauffman, 2016), and this is more likely in diverse industries with a variety of existing artefacts.

Given the discussion above, we argue that firm age or size diversity is a source of uncertainty in the economic environment in that diversity tends to blur the distinction between optimal and sub-optimal strategies for fitness maximisation and survival. Also type diversity may be conducive to the emergence of co-evolutionary pockets (firm groups) saddled on shallow local fitness optima that may make the firms vulnerable to adverse shocks. Hence, the first hypothesis (**H1**) of this paper can be stated as follows: *Higher levels of firm type (age or size) diversity within an industry is associated with higher rates of exit hazard.*

H1 is compatible with, and has the potential of unifying, the insights from organisational ecology literature (Hannan and Caroll, 1992; Hannan and Freeman, 1977, 1989; and Hannan, 2005) and from game theory, particularly the insights from the literature on games with different cognitive hierarchy of players and imperfect information (Arthur, 1994; Camerer, Ho, Chong, 2004; and Kets, 2012).

In the organisational ecology literature, density (the number of firms in the industry) and cohort diversity increase exit hazard (Hannan, 1998; 2000). Similarly, Barnett (2008) demonstrates how cohort variety increases the costs of discovering and adaptation to rivals' behaviour and creates significant uncertainty in the firms' coevolution trajectories. The cost of firm adaptation is an increasing function of the number of distinct rivals and variation in their experience. On the other hand, Barnett and Hansen (1996) emphasizes the role of arms races or Red Queen evolution in competition, which increases the risk of maladaptation when firms compete with highly varied cohorts of rivals. The increased risk of maladaptation is driven by the firm's limited capabilities to deal with the complexity of the rivals' responses, especially with the new threats that emanate from rivals with unshared co-evolutionary histories (Barnett and Sorenson, 2002). These insights are supported by empirical findings in Barnett and Hansen (1996), who utilize piecewise exponential models and report that the diversity of the bank cohorts in Illinois increased the failure rate.

Further empirical evidence is provided by Barnett and Pontikes (2008:1240), who utilize an exponential model of firm failure rate in the US computer industry and report that an organisation's "…alignment with one context limits adaptability into other contexts", and that organizations are more likely to exit if an industry evolves to a different context due to higher industrial diversity. Another empirical study based on long-term data of retail banks in Illinois concludes that "organizations confronted by a widely varying distribution of competitors grow more slowly and are more likely to face new entrants" mainly due to competency traps, search and adaptation costs of facing multiple cohorts of competitors that use different practices and bring a variety of threats (Barnett and Sorenson, 2002: 289). Variety of firm types within an industry makes it difficult to forecast rivals' actions and this increases uncertainty of firms' strategies. Zahra, Neubaum, and El-Hagrassey (2002) find negative relationship between perceived external strategic environmental uncertainty and return on equity in new ventures.

In addition, **H1** also speaks to findings in the literature on games with incomplete information, where players' types are unknown *ex ante* and players act on the basis of their beliefs about the types of their interlocutors. The information processing constraint in such games will be more biting when players (firms) are more diverse/heterogeneous. According to Camerer et al. (2004), informational constraints of this nature induce agents to engage in iterative thinking with k steps (level-k thinking) or with a variety of cognitive hierarchies (Strzalecki, 2014; Alaoui and Penta, 2016). Discontinuities or hierarchies in firm abilities to process the rivals' strategy-related information in a complex industry with a variety of firm (player) types may lead to sub-optimal choices, lower payoffs and hence higher probabilities of bankruptcies. In this setting, highly heterogeneous cohorts in the industry increase the risk of exit because diversity increases the probability of incorrect inferences and predictions of other players' actions.

Finally, **H1** also resonates with work in strategic management, an early contribution to which drew attention to the positive association between environment complexity and perceived uncertainty and difficulties in the managerial decision making (Duncan, 1972). Later on, it was argued that "under uncertainty, traditional approaches to strategic planning can be downright dangerous" and managers can "…suffer from decision paralysis" (Courtney, Kirkland, and Viguerie, 1997: 68). Furthermore, inflexible or maladaptive firms are at disadvantage in uncertain environments and may have to invest more resources in alternative projects for insurance purposes (Wernerfelt and Karnani, 1987). Although firms may rely on heuristics as a basis for strategic decisions in complex systems or under uncertainty, reliance on heuristics may involve departures from first-best optimization outcomes (Gigerenzer, 2014).

The insights above are based on the notion of search on an NK(C) opportunity (fitness) landscape by heterogeneous firms with bounded rationality. The search is based on experiential learning, which enables firms to 'sample' the opportunity (fitness) landscape, receive behavioural feedback and adapt with a view to discover their optimal niches in the fitness landscape (Levinthal, 1997; Felin et al., 2014). This is akin to passive learning in stochastic models of industry evolution and firm dynamics (Jovanovic, 1982; Hopenhayn, 1992). As argued in Felin et al. (2014), however, firms are not 'algorithmic information processors' only; and they are faced not only with a relatively well-defined opportunity (fitness) landscape that is a function of intra-industry type diversity but also with a nebulous but important novelty/innovation space. The firm's position in the latter and the draw(s) it makes from it depend on its position in the fitness landscape and the resources it invests in learning and exploration.

We argue that R&D investment is a good proxy for investment in learning and exploration, which enables the firm to discover new fitness peaks with better payoffs in conditions of complexity. This argument is in line with Fleming and Sorenson (2001), who analyses invention as a search over technology landscapes using patent citations. R&D expenditures is the key supply side component of this search process. Furthermore, the literature review by Tidd (2001) concludes that uncertainty and complexity are main environmental contingencies that affect the organization, nature, magnitude, and management of innovation. Yet, the interaction between complexity (which is a function of intra-industry diversity) and R&D investment is largely ignored in the survival literature. To address this oversight, we state our second hypothesis (**H2**) as follows: *The firm's R&D investment and intra-industry firm size or age diversity are substitutes in that firms that invest more in R&D are likely to survive longer at each level of intra-industry type diversity.*

We test for **H2** in two stages. First, we provide descriptive evidence on whether firms in more diverse industries invest more in R&D. Prior work reports that environmental uncertainty increases the need for fast adaptation and innovation – as reported by Covin and Slevin (1989) with respect to small firms and by Zahra and Bogner (1999) in relation to product and service innovation in the software industry. It is also in line with real option models in finance, which predicts higher levels of price uncertainty to be associated with higher levels of R&D investment (Pyndick, 1990; Jansen et al., 2006). In our context, a positive association between intra-industry diversity and R&D intensity is a preliminary sign that firms in more diverse industries do indeed invest more in active-learning as survival-enhancing strategy.

In the second stage we test whether R&D intensity has an *indirect effect* on exit hazard through interaction with intra-industry diversity in addition to the *direct effect* investigated in prior literature. The hypothesized indirect effect shares common ground with Erickson and Pakes (1995) who treat R&D intensity as investment in exploration and learning about the firm's true fitness. In this literature, heterogeneous firms faced with idiosyncratic productivity shocks invest in R&D to identify their optimum market/technology niches in the industry. However, the position of the firm on the fitness distribution depends on: (i) the stochastic outcome of its investment; (ii) the success of other firms in the industry; and (iii) the competitive pressure from outside the industry. The model predicts high mortality rates in the initial learning period, followed by longer survival for R&D-active firms that discover their optimal fitness levels.

H2 resonates with arguments in Audretsch, Howeling and Thurik (2004), who point out that firm survival is a result of selection and learning processes that involve experimentation with various business ideas. Such learning processes are more beneficial to the firm in more complex industry environments. It also shares common grounds with Teece (2007) who report that survival in a complex environment depends on dynamic capabilities "to sense, seize, and reconfigure" opportunities. Finally, **H2** also shares common ground with the Schumpeterian models of innovation and selection (Aghion, Akcigit, Howitt 2013; Ugur et al., 2016a), where R&D investment enables the firm to escape competition by offering new products in market niches with higher entry costs.

Empirical strategy

We begin by defining our diversity measures – the *Theil entropy index* and the *coefficient of variation* of employment sizes and firm ages within 3-digit industry codes. We elaborate on the advantages and drawback of the proposed diversity measures, and explain how we minimise the risk of potential bias that may be associated with either measure. In the second part of this section, we introduce our dataset and provide descriptive evidence on the relationship between intra-industry age/size diversity and survival times. Finally, we discuss the specification and estimation issues related to discrete-time hazard models, the way in which we choose between estimators, and the range of sensitivity/robustness checks we undertake.

Measuring firm-type diversity

Because diversity measures are associated with strengths and drawbacks (Stirling, 2010; Solanas *et al*., 2012), we utilise two measures with preponderance of desirable properties: the *Theil's entropy index*

(*TI*) and the *coefficient of variation (CV)* of the firm ages and employment size within 265 industries at the 3-digit SIC level. The Theil entropy index for each industry/year (TI_{jt}) is calculated in accordance with (1a) below.

$$TI_{jt}=\frac{1}{N_{jt}}\sum_{i=1}^{N_{jt}}\frac{L_{ijt}}{\overline{L_{jt}}}*ln\frac{L_{ijt}}{\overline{L_{jt}}} \tag{1a}$$

Here L_{ijt} is age or employment size of the i^{th} firm in industry *j* and year *t*; $\overline{L_{jt}}$ is average age or employment in industry *j* in year *t*; and N_{jt} is the number of firms in industry *j* and year *t*.

Our choice of Theil index is informed by its property of being invariant not only to unit of measurement, but also to any scale factor. The *TI* is also comparable over time and between industries; and it is additive, symmetric, decomposable and statistically testable (Theil, 1972; Haughton and Khandker, 2009). Nevertheless, the *TI* is sensitive to the left end of the size distribution - i.e. it may better reflect the diversity among smaller firms compared to larger firms (Haughton and Khandker, 2009). To address any potential bias that may be due to left-tail sensitivity in our model estimations, we control for size and age separately.

Our second diversity measure is the coefficient of variation (CV_{jt}) for firm age or employment within industry *j*. It is calculated in accordance with (1b) below.

$$CV_{jt}=\sqrt{\frac{\sum_{i=1}^{N_{jt}}(L_{ijt}-L_{jt})^2}{N_{ijt}-1}}\frac{1}{\overline{L_{jt}}}=S_{jt}\frac{1}{\overline{L_{jt}}} \tag{1b}$$

Here S_{jt} is the standard deviation of age or employment, and $\frac{1}{\overline{L_{jt}}}$ is the inverse of the mean age or employment in the industry. All other variables are as defined above. Like *TI*, the *CV* is also invariant to multiplicative scale factors and units of measurement. The drawback here is that it is an interaction term between two variables: the standard deviation of size/age and the inverse of the mean employment or age in the industry. Therefore, in our model estimations, we control for mean employment in the industry to avoid the risk of omitted variable bias (Sørensen, 2002; Bedeian and Mossholder, 2000; Solanas et al., 2012). Both *TI* and *CV* are monotonically increasing with the diversity of firm age or employment in the industry.

Data

We merged two firm-level datasets compiled by the Office of National Statistics (ONS) in the UK: the Business Structure Database (BSD) and the Business Expenditure on Research and Development Database (BERD) [1]. The BSD provides annual demographic data on births, deaths, employment, turnover, number of live local units, foreign/domestic ownership, etc. We excluded companies that exited due to mergers and acquisitions. Hence exit in our dataset refers to liquidations or bankruptcies. This was possible because both firm and local unit (plant) references disappear from the register if exit is due to liquidation/bankruptcy. However, only firm identifiers disappear if the exit was due to mergers or acquisitions. We have constructed the exit year as the earliest of the death year recorded by the ONS, or the first year when the firm employment and turnover are zero for three consecutive years[2]. We also excluded firms with birth date before 1974 as firms were given the same birth year of 1973 when the business register was first introduced in 1973.

On the other hand, BERD consists of repeated annual surveys with stratified sampling of firms known to be R&D-active. We merged BERD with BSD, using the unique firm (enterprise) identifier[3]. In our dataset, the ratio of R&D to turnover is greater than 1 from the 95th percentile onwards. We have considered firms in the top 5 percent of the R&D intensity distribution as outliers and set our baseline estimation sample for firms with R&D intensity less than 1.[4] Our estimation sample consists of 35,136 firms and 158,316 observations from 1998 to 2012. Summary statistics for the estimation sample are presented in Table A1 in the *Appendix*, broken down by exiting firms, survivors and all sample.

Scatter plots in Figure 1 are based on data within 3-digit industries and allow for visual inspection of the relationship between average survival times and average measures of age/size diversity within the industry. Scatters in panels (a) and (b) both indicate a negative relationship between age/size diversity and survival time, which is equivalent to the positive association between type diversity and

[1] The standard disclaimer applies: The use of these data does not imply the endorsement of the data owner or the UK Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.
[2] The second criterion is used because we have identified delays in the ONS assignment of a death code in some cases even though the firm's return for employment and turnover is zero for several years.
[3] Further information on BSD and BERD is provided in Ugur, Trushin, and Solomon (2016b).
[4] Our results are robust to different cut-off points for R&D intensity. These results are not presented here to save space, but they are available on request.

exit hazard rate postulated in **H1**. The negative association between intra-industry age/size diversity and survival time is observed when diversity is measured with *TI* (Panel a) and *CV* (Panel b).

'Insert Figure 1 here'

Scatter plots in Figure 2 shed light on a different empirical pattern in our data. Whether we measure diversity with the Theil index (panel a) or the coefficient of variation (panel b), we observe a positive relationship between intra-industry age/size diversity and average R&D intensity in the industry. This finding provides an empirical underpinning for **H2** in that it points to a higher R&D effort in more diverse industries with higher risk of exit hazard. It also confirms findings in prior work, which report that firms do invest more in R&D when environmental uncertainty is higher (Covin and Slevin, 1989; Zahra and Bogner, 1999). If supported with additional findings lending support to **H2**, this empirical pattern in the data indicates that firms do indeed invest in R&D to learn about their product and technology niches and thereby reduce exit hazard in more diverse industries.

'Insert Figure 2 here'

Modeling exit hazard: Main variables of interest and controls

In this section, we conduct hazard estimations with a view to verify if the descriptive evidence in Figure 1 is statistically significant after controlling for a wide range of firm, industry and macroeconomic factors that have been investigated in the prior literature on firm survival. To do this, we follow the general specification for the hazard rate function (Jenkins, 1995), but we use lagged, hence, predetermined or weakly exogenous covariates x_{it} to deal with simultaneity bias. The probability (*Pr*) of exit in year *t+1* conditional on observable covariates can be stated as follows:

$$Pr(y_{it+1}|x_{it,}v_i)= Pr(x_{it}\beta + M_{it}\alpha + v_i + \gamma_{t+1} + \varepsilon_{it+1}) \qquad (2)$$

Here, *i* and *t* are firm and year indices; x_{it} is a vector of observable firm-level covariates that affect firm exit with an estimated vector of β parameters; M_{it} is a vector of industry, technology (Pavitt classes) and macroeconomic variables that affect firm exit with an estimated vector of α parameters; γ_{t+1} are year dummies; and ε_{it+1} is the disturbance term. Unobserved heterogeneity between firms is captured by the independently and identically distributed (i.i.d.) normal random variable $v_i|x_{it}, M_{it+1} \sim N(0,\sigma_v^2)$. The strong and very common assumption in estimation of such models is that unobserved heterogeneity (v_i) and the disturbance term ε_{it+1} are independent of the firm, industry, and macroeconomic covariates.

How important is the assumption that unobserved heterogeneity v_i is i.i.d. Normal? Nicoletti and Rondinelli (2010) have evaluated biases in estimated parameters of the discrete time hazard models caused by omitting or misspecifying the unobserved heterogeneity distribution using Monte-Carlo simulations. Their results demonstrate that neglect or misspecification of the unobserved heterogeneity are unlikely to lead to a significant bias in the estimated parameters.

The variables of main interest in x_{it} include the *TI* or *CV* measures of intra-industry diversity and the interactions of the latter with firm-level R&D intensity. The remaining firm-level covariates in x_{it} and the industry, technology class and macroeconomic covariates in M_{it} are specified in accordance with the best practice in survival analysis. Definitions of all covariates and the literature that justifies their inclusion in the model are presented in Table 1.

'Insert Table 1 here'

Summary statistics for all variables are reported in Table A1 in the *Appendix*. Our main variables of interest are *TI* and *CV* measures of diversity; and the interactions of the latter with firm-level R&D intensity defined as the ratio of R&D expenditures to turnover. The correlation between *TI* and *CV* is 0.44 for firm sizes and 0.37 for firm ages and significant, but their correlation with the Herfindahl index is low and statistically insignificant. The latter property reduces the risk of collinearity and indicates that the informational contents of intra-industry concentration and size/age diversity measures do reflect different ecosystem properties. *TI*, *CV* and their interactions with R&D intensity enable us to test for H1 and H2 proposed above.

We control for direct effects of R&D intensity in line with prior literature, which tends to adopt a linear specification but report conflicting findings. Although survival-enhancing effects of innovation are reported more frequently, Ugur et al. (2016a) demonstrate that a quadratic specification could be more plausible both theoretically and empirically, and this relationship was empirically confirmed by Sharapov *et al*. (2011) and by Zhang and Mohnen (2013). This is why we control for both R&D intensity and squared value.

Another set of firm characteristics that have been investigated in the survival literature consists of firm age and size. Both theoretical and empirical work indicates that age and size are correlated positively with survival, even though the relationship may not be monotonic (Geroski, 1995; Klette and Kortum, 2004; Aghion *et al*., 2013). We control for age and size in line with existing work; but also to disentangle the *direct* effect of age or size on exit hazard for the effect of age or size diversity.

Labour productivity and firm growth rate relative to industry growth are reported as significant determinants of firm survival (Audretsch, 1991, 1995; Hopenhayn, 1992; Ericson and Pakes, 1995; Mata, Portugal, Guimaraes, 1995; Cefis and Marsili, 2005; and Ugur *et al*., 2016a). The fourth set of firm-level characteristics includes number of plants, whether the firm is engaged in civil R&D only, and domestic versus foreign ownership. Inclusion of these observable characteristics in a hazard model is consistent with Audretsch (1991), Audretsch and Mahmood (1995), Mata *et al*. (1995), Fernandes and Paunov (2015), and Sharapov, Kattuman, and Sena (2011).

Of the industry-level covariates, the effect of entry rate has been studied by Baldwin and Rafiquzzaman (1995) who reported that higher entry rates tend to reduce firm survival. Positive association between entry and exit rates at the industry level has been reported by Dunne et al. (1988) and Sidney *et al*. (2003). The effect of market concentration also varies but it tends to insignificant (see Baldwin and Rafiquzzaman, 1995; McCloughan and Stone, 1998; Ugur *et al*., 2016a).

We control for average number of firm employees at 3-digit SIC industry level in order to address the risk of omitted variable bias, particularly when the coefficient of variation is used as diversity measure. We also control for median R&D intensity in the industry to verify whether higher level of creative destruction affects firm mortality (Aghion *et al*., 2013; Fernandes and Paunov (2015); Ugur *et* *al*., 2016a). Finally, we check if technology classes matter using the Pavitt (1984) taxonomy.[5] The final set of covariates relates to macro-level indicators such as onsets of financial crises, real effective exchange rate of the British pound, and GDP growth. Whilst currency appreciation may affect mortality because of decline in international cost competitiveness, the crisis dummy accounts for changes in the business and credit environment. Finally, GDP growth captures the effect of business cycle on firm survival (Goudie and Meeks, 1991; Bhattacharjee *et al*., 2009; Ugur *et al*., 2016).

Estimation methodology

Our estimation methodology follows Wooldridge (2010) on grouped duration data, where firm exit time is known within one year. The discrete-time hazard rate h_{it} that firm *i* exits in *Te* years conditional on survival for *Te-1* years can be expressed as conditional probability of firm survival for T_i years as follows:

$$h_{it} = \frac{\Pr(T_{e-1} < T_i \leq T_e)}{\Pr(T_i > T_{e-1})} \qquad (3)$$

Bauer and Agarwal (2014: 432) provide evidence that discrete-time hazard models are "superior to the alternatives" in the context of estimating bankruptcy hazards. The parameters are estimated by maximizing the logarithm of the likelihood function. Whereas the Logit specification assumes a logistic distribution for the hazard, the Probit assumes a standard Normal distribution. Given the panel structure of the data we choose random effect estimations as it helps to correct for omitted variable bias (Fernandes and Paunov, 2015), whereas fixed-effect estimations often lead to large biases in all estimated parameters with relatively small number of periods in the dataset due to incidental parameters problem (Lancaster, 2000; Wooldridge, 2010, Bester and Hansen, 2009). The dependent variable is an indicator taking the value of 1 if the firm exits in year *Te,* and zero otherwise. To partially eliminate

[5] Pavitt technology classes are from Pavitt (1984), as revised slightly by Bogliacino and Pianta (2010). Pavitt1 consists of firms in science-based industries such as chemicals, office machinery, precision, medical and optical instruments industries, ICT. Pavitt2 includes specialized suppliers of technology - mechanical engineering industries, manufacturers of electrical machinery, equipment, etc. Pavitt3 includes scale-intensive industries such as pulp and paper, transport vehicles, mineral oil refining industries. Pavitt4 includes industries dominated by external suppliers of technology, e.g., textiles & clothing, food & drink, fabricated metals. Finally, Pavitt5 consists of unclassified industries.

competing causal attributions, we use one-year forwarded firm exit as our dependent variable (see model 2 above).

The panel random-effect model estimators control for unobserved firm heterogeneity. Geroski, Mata, Portugal (2010) emphasize the importance of such control. Wooldridge (2010) demonstrates that a $\sqrt{N}$ consistent estimator in this case, the population-averaged parameters, can be obtained by maximization of the log-likelihood function $log\ L$:

$$log\ L = \sum_{i=1}^{N} \sum_{t=1}^{Te} \{y_{it} log\ h_{it} + (1 - y_{it}) log(1 - h_{it})\} \quad (4)$$

For the random effects model the maximum log-likelihood estimations are based on Gauss-Hermite quadrature approximation (see Naylor and Smith, 1982) with a corresponding probability distribution hazard function *Pr(z)*. To check for robustness, we use both non-proportional hazard functions (Logit and Probit) and the proportional specification through Complementary log-log (Clog-log). Although Jenkins (1995) notes that both estimators tend to converge when hazard rates are small, it is appropriate to use both types of estimators as the hazard functions are not known *ex ante*.

Typical distribution specification for the random-effects estimators are given by the standard Normal Φ cumulative density functions in (5a)-(5c) (Wooldridge, 2010):

Probit: $Pr(z) = \Phi(z) = \Phi(x_{it}\beta + M_{it+1}\alpha + v_i + \gamma_{t+1} + \varepsilon_{it+1}) \quad (5a)$

Logit: $Pr(z) = 1/(1 + exp(-z)) \quad (5b)$

In the complementary log-log random-effects estimator, the conditional probability function is given by *Clog-log*: $Pr(z) = 1 - exp(-exp(z)) \quad (5c)$

We use likelihood ratio test to check if the panel random-effects estimators deliver similar results with pooled estimators, i.e. if panel level variance is insignificant and the ratio $\rho = \frac{\sigma_v{}^2}{\sigma_\varepsilon{}^2 + \sigma_v{}^2}$ is different from zero by sampling chance[6]. We also use robust standard errors of the estimated parameters, which provide consistency when the disturbances are not correlated across firms.

Importantly, in nonlinear models the interaction effect is not equivalent to the marginal effect, i.e. the sign of the estimated parameter for the interaction term between R&D intensity and diversity can be misleading (Norton et al. 2004). We numerically estimate the interaction effects by using *margins*

[6] Stata reports panel level variance $\ln(\sigma_v{}^2)$ in form of lnsig2u_const.

(Williams, 2012) and *inteff* (Norton, Wang, Ai,(2004) procedures in Stata based on delta approximation method applied to the Probit model, which is selected by the AIC and BIC information criteria. In this model, the interaction effect for the conditional mean of the dummy variable y is:

$$E[y|x_1, x_2, X] = \Phi(\alpha_1 x_{1t} + \alpha_2 x_{2t} + \alpha_{12} x_{1t} x_{2t} + x_{jt}\alpha_j) = \Phi(z) \qquad (6)$$

According to Norton et al. (2004), the full marginal effect of the interaction term on the conditional mean survival is:

$$\frac{\partial^2 \Phi(z)}{\partial x_1 x_2} = [\alpha_{12} - (\alpha_1 + \alpha_{12} x_2)(\alpha_2 + \alpha_{12} x_1) z]\, \Phi'(z) \qquad (7)$$

Hence, the marginal effect of the interaction term depends on specific levels of all covariates. We also report graphical representations of the estimated interaction effects following Greene's (2010) recommendation for nonlinear models.

Results

Our results are based on the list of variables summarised in Table A1 in the *Appendix*, where we provide separate summary statistics for exiting firms, survivors and full sample. Our sample consist of 35,136 R&D-active firms observed from 1998-2012, of which 28,287 firms are survivors and 6,849 firms exit over the analysis period. We report results form a wide range of discrete-time hazard models: pooled Probit, pooled Logit, pooled Complementary log-log (Clog-log) and their panel random-effects versions. We present the preferred estimation results in the main text; and the additional sensitivity/robustness checks in the *Appendix*. The preferred estimation results are determined by AIC/BIC values, which point in favour pooled Probit compared to pooled Logit or pooled Clog-log; and in favour of random-effects Probit estimator compared to random-effects Logit or Clog-log.

Table 2 reports the results from pooled probit and random-effects probit estimations for both measures of size diversity: the Theil entropy index (columns 1 and 2) and the coefficient of variation (columns 3 and 4) of employment. Results from other pooled and panel estimators are reported in Table A2a (with the Theil entropy index of employment) and A2b (with the coefficient of variation of employment) in the *Appendix*. We have also conducted a likelihood ratio (LR) test to check whether the panel random-effects estimators are preferable to pooled estimators as they allow for taking account of unobserved firm heterogeneity. The test favours the random-effects estimators, which we use to obtain non-linear interaction effects marginal effects for R&D intensity and diversity (Table 3) and conditional effects depending on different levels of R&D intensity (Table 4).[7]

'**Insert Table 2 here'**

Post-estimation results for pooled Probit (bottom three rows in Table 2) indicate that: (i) the model fits the data well as the Pearson χ^2 does not reject the null hypothesis of good fit; (ii) the overall rate of correct classification is high: 95.66 percent in the estimation based on the Theil index and 95.67 percent in the estimation with the coefficient of variation respectively; and (iii) the model has good power to discriminate between exiting and surviving firm as the area under the ROC curve is 0.69 and 0.68, respectively. In addition, there is a high degree of sign and significance consistency across six estimators and two diversity measures (compare Tables 2, A2a, and A2b). The consistency is evident with respect to covariates of main interest (diversity measures, their interactions with R&D intensity, and the latter's linear and quadratic terms) and the wide range of controls discussed in Table 1. Furthermore, estimation results based on age as the diversity measure (reported in Table A3 in the *Appendix*) are consistent with the findings based on size diversity and across two diversity measures: the Theil entropy index and the coefficient of variation.

The evidence we report provides strong evidential support for **H1**, which postulates that higher levels of intra-industry size diversity are associated with higher levels of exit hazard. The coefficient estimates are significant at 1 percent in preferred estimations and in others reported as robustness checks. This finding is observed after controlling for firm size and its square, and for the mean employment in the industry. Hence, the intra-industry size diversity is an exit hazard in its own right. We must also indicate that this finding is obtained after controlling for firm size and mean size in the industry in order to minimize the risk of bias due to the left-tail sensitivity of the Theil index and the omitted variable

[7] Preferred estimation results of age diversity and exit hazards are presented in Table A3 in the *Appendix*. The results are fully consistent with those based size diversity. Therefore, results from pooled logit and clog-log and form their panel equivalents are not reported to save space. Also LR test results are not reported here for the same reason. However, all these results but can be provided on request.

bias that may be caused by the coefficient of variation (Stirling, 2010; Solanas *et al*., 2012). True the parameter estimates for the Theil entropy index are larger than those of the coefficient of variation. This, however, is to be expected because the mean of the coefficient of variation in the sample is 5.0 compared to a mean Theil index of 1.9. Stated differently, the parameter estimates for the diversity measures are consistent in terms of sign have comparable magnitudes when the scale of the diversity measures are taken into account.

The negative and statistically-significant parameter estimates for the interaction between diversity and R&D intensity lend support to **H2**, which posits that higher levels of R&D investment counterbalance the hazard-increasing effects of diversity. However, we take account of the non-linear nature of the hazard estimators and obtain numerical estimates of the parameter using *margins* and *inteff* procedures in *Stata* (Norton et al., 2004; Williams, 2012). We also provide graphical representations of the estimated interaction effects following Greene's (2010) recommendation for nonlinear models.

Drawing on Greene (2010), we present in Figure 2 the sign and significance of the interaction effects from the Probit model. The horizontal axis indicates the predicted probability of exit whereas the vertical axis indicates the associated Z-statistics. The horizontal lines above and below zero demarcates the Z-values that corresponds to statistical significance at 5%. It can be seen that the parameter estimate for the interaction term are always associated with negative Z-statistics (confirming the negative effect on exit hazard) and the Z-statistics are almost always below the demarcation line for significance. These graphical results lend support to H2, which stipulates that the interaction between diversity and R&D intensity ameliorates the adverse effect of diversity on survival.

'**Insert Graph 2 here**'

Table 3 provides further evidence that lends support to **H2**. Using the delta method and routines discussed in Norton et al., 2004 and Williams, 2012), we present numerical estimates of the marginal interaction effects in Table 3. The results indicate that the marginal effects are always negative and significant at mean values of all the covariates. When the full range of the covariate values are taken into account, the marginal effects are predominantly negative and significant at least at 10% with the exception of the third row, where the marginal effect relates to interaction between R&D intensity and the coefficient of variation measure of size diversity.

Hence we can safely conclude that intra-industry diversity and R&D intensity are substitutes in that an increase in R&D intensity reduces exit hazard at each level of diversity. Given that the interaction effect is significant after controlling for the direct effects of R&D intensity and diversity, it indicates that the optimal level of R&D intensity is higher in more diverse industries. This finding ties in with the descriptive evidence presented earlier and is in line with the active learning approach to R&D investment in Ericson and Pakes (1995). R&D investment ameliorates the adverse effects of intra-industry diversity on survival by: (i) increasing the scope for discovering, reconfiguring, and developing capacities (Anderson and Tushman, 2001; Audretsch *et al*., 2004; Cohen and Levinthal, 1990; Probst and Raisch, 2005; and Teece, 2007); (ii) facilitating adaptation to shifts in the technology regime (Nelson, 1995); or (iii) enabling firms to keep up with shifts in the technology frontier (Aghion *et al*., 2013; Ugur *et al*., 2016).

'Insert Table 3 here'

The results from the preferred estimators and samples above are strongly consistent with findings from a battery of sensitivity checks, including results form: (i) alternative pooled and panel hazard estimators (Table A2a and A2b); (ii) different cut-off points for R&D intensity; (iii) stepwise regressions; (iv) samples that avoid left truncation by investigating firms born in 2000 or 2003 and after; (v) samples that exclude firms in the financial and defence industries. Most of these results are not presented here to save space, but they can be provided on request.

It must also be noted that the results discussed above are highly consistent with those we obtain when age variation is used as a measure of intra-industry type diversity (see Table A3 in the *Appendix*). Using both the Theil index and coefficient of variation measures of age diversity, we find a positive relationship between intra-industry age diversity and exit hazard. Our findings indicate that there is a strong case for taking diversity seriously as a predictor of firm mortality. The relationship between diversity and mortality is driven by shallow fitness distributions and higher probabilities of settling on sub-optimal fitness peaks. This characteristic of the fitness landscape in diverse industries indicates that industry evolution is as much about selection in favour of the fittest as the result of uncertainty about the globally optimal fitness is.

In the following paragraphs we discuss the findings concerning the wide range of controls that capture the hazard-related factors at the firm, industry and macro levels. First, we confirm the diminishing scale effect in the relationship between R&D intensity and survival reported earlier in Ugur et al. (2016a). We find a U-shaped relationship between R&D intensity and exit hazard: as R&D intensity increases exit hazard falls at increasingly slower rates, with a turning point beyond which increased R&D intensity increases exit hazard. The diminishing scale effect can be due to increased riskiness of R&D investments at higher levels of R&D intensity (Ericson and Pakes, 1995; Czarnitzki and Toole, 2013). Secondly, R&D-active firms may fail to diversify their revenue streams at the same pace as their investment in R&D (Fernandes and Paunov, 2015; Ugur et al., 2016a).

Our findings for the remaining controls are in line with the theoretical and empirical literature, which indicates that new entrants have shorter survival time, but those that survive grow faster and survive longer (Klette and Kortum, 2004; Aghion *et al*., 2013; Cefis and Marsili, 2005; and Evans, 1987 among others). The U-shaped relationship between size and exit hazard we report is in line with evidence on size distribution and survival among Portuguese firms (Cabral and Mata, 2003), which suggests that a large size beyond an efficient scale may be a hazard factor in firm dynamics. By controlling for the direct effect of size and demonstrating that the direct effect on survival is non-monotonic, we lend added credence to our findings about diversity. Stated differently, our findings concerning the adverse effects of diversity on survival are not mimicking the higher hazard rates among small firms.

Audretsch and Mahmood (1995) and Fernandes and Paunov (2015), we report that multi-plant firms are less likely to exit as they are better able to diversify risk and restructure. We also report that firm productivity (real turnover per employee) and growth rates relative to median growth in the industry are associated with lower exit hazard. These findings are in line with Doms *et al*. (1995), Mata *et al*. (1995), and Griliches and Regev (1995) among others.

Of the industry-level covariates, we report that the relationship between exit hazard and market concentration is insignificant. This is in line with prior studies, which offer the following explanations: (i) market concentration may be less important than market niches in determining monopoly rents (Geroski, 1995; Wagner, 1994); (ii) entry costs associated with concentrated industries depend on actions of hypercompetitive and less predictable firms, but not on the number of companies (concentration) in an industry; and (iii) industries with similar concentration ratios often show significant variation in the overall firm-size distribution (Carroll, 1985: 1264).

Four Pavitt classes are associated with lower exist hazard relative to the excluded class (Pavitt 4), which consists of firms that depend on import of technology from other industries. This result is in line with Agarwal and Audretsch (2001) and Cefis and Marsili (2005), who indicate that the nature of the technology in the industry matters. Our finding concerning the effect of the average R&D intensity in the industry (as a proxy for intra-industry creative destruction) suggests that exit hazard increases as creative destruction increases confirming the Schumpeterian innovation models (Aghion *et al*., 2013; Fernandes and Paunov, 2015; Ugur *et al*., 2016).

With respect to macroeconomic variables, we report that real currency appreciation (reduced competitiveness) and the onset of a financial crisis tend to increase exit hazard; whereas annual GDP growth rates have a negative relationship with exit hazard. These findings are in line with those reported in prior survival studies that control for macroeconomic variables (Bhattacharjee et al., 2009; Goudie and Meeks, 1991; Ugur et al., 2016).

The final set of evidence we present here has significant implications for organisational strategy because it sheds light on the levels of R&D intensity required to counter-balance and eventually reverse the hazard-increasing effect of diversity. In Table 4 below we present the effects of the diversity on exit hazard conditional on different levels of R&D intensity. Fixing the covariates at their sample means, we varied the level of R&D intensity from the bottom 5^{th} to the top 95^{th} percentile values. The results indicate that the hazard-increasing effect of diversity declines as R&D intensity increases. The adverse effect due to firm-size diversity becomes insignificant between the 70^{th} and 75^{th} percentiles (at R&D intensity of approximately 9–11% of firm turnover) and is eventually reversed at the top R&D intensity. In the case of age diversity, the adverse effect is diluted significantly as R&D intensity increases but it is never neutralised or reversed. This finding indicates that R&D investment does indeed reduce the exit hazard associated with diversity but it is more effective in providing a shield against the adverse effects of size diversity compared to that of age diversity.

‘**Insert Table 4 here’**

Conclusions

In this paper, we have addressed two novel research questions: (i) does intra-industry size and age diversity constitute a hazard factor for firms beyond the direct effect of size and age as such? (ii) does firm investment in R&D moderate the effect of diversity on exit hazard by enabling the firm to increase competencies in dealing with complexity in diverse industries? We have provided a wealth of evidence that justifies affirmative answers to both questions. The hazard-increasing effects of age and size diversity and a moderating indirect effect associated with R&D intensity are evident and consistent across six discrete-time hazard estimators applied to four different samples and two diversity measures.

Our findings are informed by and has the potential for unifying insights on industry diversity and learning from four areas of study: theoretical biology, organizational ecology, industrial organisation and game theory. The findings lend consistent support to the hypotheses we derive from a carefully distilled synthesis of those insights. On the one hand, we demonstrate that intra-industry type (age or size) diversity increases the complexity of the opportunity/fitness landscape and thereby leads to higher likelihood of suboptimal choices and exit hazard. On the other hand, we found that firms in more diverse industries do invest more in R&D to better learn and develop competencies to cope with heterogeneous rivals’ responses under the conditions of rugged fitness landscapes and bounded rationality. Therefore we argue that intra-industry diversity is an important indicator of complexity that should be analysed and modelled in the study of industry evolution and firm dynamics. The implication for business decision making is that firms in diverse industries should select their R&D investment levels not only on the basis of its direct effects on firm survival (which is non-monotonic) but also in the light of existing age and size diversity in the industry. Due to active-learning effects within diverse industries, R&D investment does reduce exit hazard and may eventually enable the firm to neutralise the adverse effect of diversity on survival.

References

Agarwal R, Audretsch DB. 2001. Does entry size matter? The impact of the life cycle and technology on firm survival. *The Journal of Industrial Economics* **49**, 21-43.

Agarwal R, Audtretsch D. 1999. The two views of small firms in industry dynamics: A reconciliation. *Economics Letters* **62**, 245-251.

Agarwal R, Gort M. 2002. Firm and product life cycles and firm survival. *The American Economic Review* **92(2)**, 184-190.

Aghion P, Akcigit U, Howitt P. 2013. What do we learn from Schumpeterian growth theory? Penn Institute of Economic Research Working Papers, no. 13-026. http://economics.sas.upenn.edu/sites/economics.sas.upenn.edu/files/13-026.pdf

Alaoui, L., Penta, A. 2016. Endogenous depth of reasoning. *The Review of Economic Studies* **83(4):** 1297–1333.

Anderson, P., Tushman, M. 2001. Organizational environments and industry exit: the effects of uncertainty, munificence and complexity. *Industrial and Corporate Change* **10**: 675-711.

Anderson, P., Eshima, Y. 2013. The influence of firm age and intangible resources on the relationship between entrepreneurial orientation and firm growth among Japanese SMEs. *Journal of Business Venturing* **28(3)**, 413-429.

Andriani, P. and Cattani, G., 2016. Exaptation as source of creativity, innovation, and diversity: introduction to the Special Section. *Industrial and Corporate Change*, ***25*(1)**, pp.115-131.

Ashby RW. 2011. Variety, constraints, and the Law of Requisite Variety, *Emergence: Complexity and Organization* **13**(1-2): 190-207.

Audretsch D. 1991. New-Firm Survival and the Technological Regime. *Review of Economics and Statistics* **60(3)**, 441-450.

Audretsch D.B. 1995. Innovation, growth and survival. *International Journal of Industrial Organization* **13(4)**, 441-457.

Audretsch D.B, Houweling P., Thurik A.R. 2004. Industry evolution: Diversity, selection and the role of learning. *International Small Business Journal* **22(4)**, 331-348.

Audretsch D.B., Mahmood T. 1995. New Firm Survival: New Results Using a Hazard Function. *The Review of Economics and Statistics* **77**(1), 97-103.

Baldwin J., Rafiquzzaman M. 1995. Selection versus evolutionary adaptation: learning and post-entry performance. *International Journal of Industrial Organization* **13**, 501-522.

Barnett, W. 2008. The Red Queen among organizations: how competitiveness evolves. Princeton: Princeton University Press.

Barnett W, Hansen M. 1996. The Red Queen in organizational evolution. *Strategic Management Journal* **17**, 137-157.

Barnett, W., Sorenson, O. 2002. The Red Queen in organizational creation and development. *Industrial and Corporate Change* **12(4),** 673-695.

Barnett, W., Pontikes, G. 2008. The Red Queen, success bias, and organizational inertia. *Management Science* **54(7)**, 1237-1251.

Bauer J, Agarwal V. 2014. Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking and Finance* **40**, 432-442.

Bergh, D., Lawless, M. 1998. Portfolio restructuring and limits to hierarchical governance: The effects of environmental uncertainty and diversification strategy. *Organization Science* **9(1)**, 87-102.

Bester A, Hansen C. 2009. A Penalty Function Approach to Bias Reduction in Nonlinear Panel Models with Fixed Effects. *Journal of Business & Economic Statistics* **27(2)**, 131-48.

Bhattacharjee A, Higson C, Holly S, Kattuman P. 2009. Macroeconomic instability and business exit: Determinants of failures and acquisitions of UK firms. *Economica* **76(301)**, 108-131.

Bogliacino, F. and Pianta, M. 2010. Innovation and employment: A reinvestigation using revised Pavitt classes. *Research Policy*, **39(6)**, 799-809.

Boone, C. and Hendriks, W., 2009. Top management team diversity and firm performance: Moderators of functional-background and locus-of-control diversity. *Management Science*, ***55*(2)**, pp.165-180.

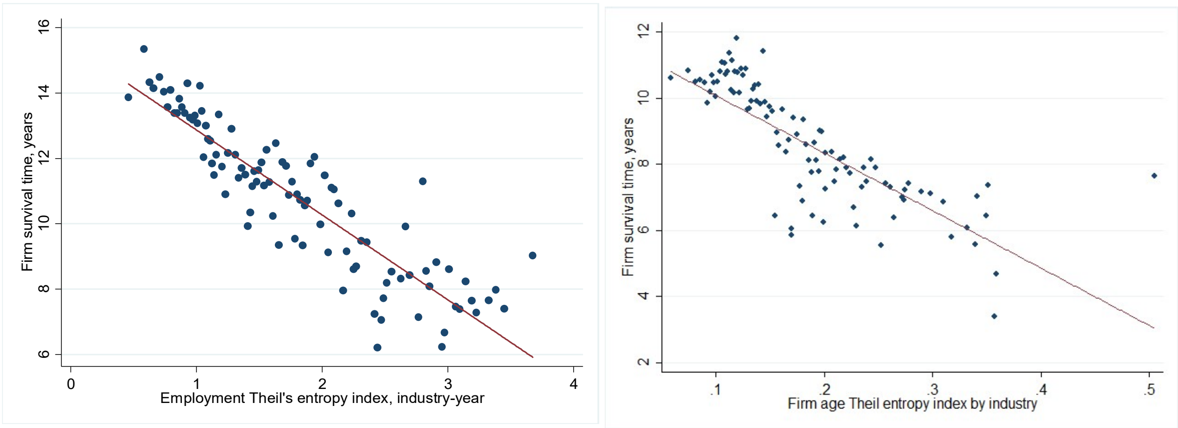
Bosch, F., Volberda, H., Boer, M. 1999. Coevolution of firm absorptive capacity and knowledge environment: Organizational forms and combinative capabilities. *Organization Science* **10(5)**, 551-68.
Cadot O, Carrere C, Strauss-Kahn V. 2013. Trade diversification, income, and growth: What do we know? *Journal of Economic Surveys* **27**(4), 790-812.
Camerer, C., Ho, T.-H., Chong, J.-K. 2004. A cognitive hierarchy model of games. *Quarterly Jounral of Economics* **119(3)**, 861-898.
Carroll GR, Hannan MT 2000. *The demography of corporations and industries*. Chapter 12. Princeton University Press: Princeton.
Carroll GR. 1985. Concentration and specialization: Dynamics of niche width in populations of Organizations. *American Journal of Sociology* **90,** 1262-283.
Cefis E, Marsili O. 2005. A matter of life and death: innovation and firm survival. *Industrial and Corporate Change* **14(6),** 1167-1192.
Challet, D., Zhang, Y.-C. 1998. On the minority game: Analytical and numerical studies. *Physica A* **256**, 514-532.
Cohen W, Levinthal D. 1990. Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly* **35(1)**, 128-152.
Covin, J., Slevin, D. 1989. Strategic management of small firms in hostile and benign environments. *Strategic Management Journal* **10**, 75-87.
Courtney, H., Kirkland, J., Viguerie, P. 1997. Strategy under uncertainty. *Harvard Business Review* **75**, 67-79.
Czarnitzki, D. and A. A. Toole (2013). The R&D Investment–uncertainty relationship: Do strategic rivalry and firm size matter? *Managerial and Decision Economics*, **34(1)**, 15-28.
D'Aveni, R., Dagnino, G., Smith, K. 2010. The age of temporary advantage. *Strategic Management Journa*l **31**: 1371-85.
Dean, T., Brown, R., Bamford, C. 1998. Differences in large and small firm responses to environmental context: Strategic implications from a comparative analysis of business formations. *Strategic Management Journal* **19**, 709-728.
Disney R, Haskel J, Heden Y. 2003. Entry, exit and establishment survival in UK manufacturing. *Journal of Industrial Economics* **51(1)**, 91-112.
Doms M, Dunne T, Roberts MJ. 1995. The role of technology use in the survival and growth of manufacturing plants. *International Journal of Industrial Organization* **13**, 523-542.
Dosi G, Malerba F, Marsili O, Orsenigo L. 1997. Industrial Structures and Dynamics: Evidence, Interpretations and Puzzles. *Industrial and Corporate Change* **6(1)**, 3-25.
Duncan, R. B. 1972. Characteristics of organizational environments and perceived environmental uncertainty. *Administrative Science Quarterly* **17**, 313-327.
Dunne T, Roberts MJ, and Samuelson L. 1988. Patterns of firm entry and exit in U.S. manufacturing industries. *RAND Journal of Economics* **19(4)**, 495-515.
Elton CS. 1958. *The Ecology of Invasions by Plants and Animals*. Methuen & Co. Ltd.: London.
Ericson R, Pakes A. 1995. Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies* **62**(1), pp. 53-82.
Evans DS. 1987. The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *The Journal of Industrial Economics* **35**, 567-581.
Felin, T., Kauffman, S. Koppl, R., Longo, G. 2014. Economic Opportunity and Evolution: Beyond Landscapes and Bounded Rationality. *Strategic Entrepreneurship Journal* 8(4), 269-282.
Fernandes A, Paunov C. 2015. The risks of innovation: Are innovating firms less likely to die? *The Review of Economics and Statistics* **97(3)**, 638-53.
Fleming, L. and Sorenson, O., 2001. Technology as a complex adaptive system: evidence from patent data. *Research policy*, **30(7)**, 1019-1039.
Freel, M.S. 2005. Perceived environmental uncertainty and innovation in small firms. *Small Business Economics* **25**, 49-64.
Geroski, PA. 1995. What do we know about entry? *International Journal of Industrial Organization* **13**(4), 421-440.
Geroski, PA. 2001. Exploring the niche overlaps between organizational ecology and industrial economics. *Industrial and Corporate Change* **10(2)**, 507-540.

Geroski P, Mata J, Portugal P. 2010. Founding conditions and the survival of new firms. *Strategic Management Journal* **31**, 510-529.
Gigerenzer, G. 2014. Risk savvy: How to make good decisions. NY: Viking.
Greene, W. 2010. Testing hypotheses about interaction terms in nonlinear models. *Economic Letters* **107**, 291-296.
Goudie AW, Meeks G. 1991. The exchange rate and company failure in a macro-micro model of the UK company sector. *Economic Journal* **101**, 444-457.
Grant, R. 1996. Toward a knowledge-based theory of the firm. *Strategic Management Journal* **17**, 109-122.
Hannan MT, Freeman J. 1977. The population ecology of organizations. *American Journal of Sociology* **82**(5), 929-964.
Hannan MT, Freeman J. 1989. *Organizational Ecology*. Harvard University Press: Cambridge, MA.
Hannan M. 2005. Ecologies of organizations: Diversity and identity. *Journal of Economic Perspectives* **19**(10), 51-70.
Haughton J, Khandker S. 2009. *Handbook on Poverty and Inequality*. The World Bank: Washington, D.C.
Ho, Y.C., Zhao, Q.-C., Pepyne, D.L. 2003. The no free lunch theorems: Complexity and security. *IEEE Transactions on automatic control* **48(5)**, 783-793.
Hodgekinson, G., Wright, G. 2002. Confronting strategic inertia in a top management team: learning from failure. *Organization Studies* **23(6)**, 949-977.
Hodgson, G. 1993. The Mecca of Alfred Marshall. The Economic Journal **103(417)**, 406-415.
Hofer, C. 1975. Toward a Contingency Theory of Business Strategy. *Academy Management Journal* **18(4)**, 784-810.
Hopenhayn H. 1992. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica* **60(5)**, 1127–1150.
Huisman, J., Weissing, J. 2001. Fundamental unpredictability in multispecies competition. *The American Naturalist* **157(5)**, 488-493.
Hunt, S., Morgan, R. 1996. The resource-advantage theory of competition, dynamics, path dependencies, and evolutionary dimensions. *Journal of Marketing* **60(3)**, 107-114.
Ismail, N., Jenatabadi, H. 2014. The influence of firm age on the relationships of airline performance, economic situation and internal operation. *Transportation Research Part A: Policy and Practice*
67, 212-224.
Ives AR, Carpenter SR. 2007. *Science* **317** (6 July), 58-62.
Jansen, J., Van Den Bosch, F., Volberda, H. 2006. Exploratory innovation, exploitative innovation, and performance: Effects of organizational antecedents and environmental moderators. *Management Science* **52(11)**, 1661-1674.
Jenkins S. 1995. Easy ways to estimate discrete time duration models. *Oxford Bulletin of Economics and Statistics* **57**, 129-138.
Jovanovic B. 1982. Selection and the Evolution of Industry. *Econometrica* **50(3)**, 649-670.
Kahneman, D., Slovic, P., Tvesky, A. 1982. Judgement under uncertainty: Heuristics and biases. NY: Cambridge University Press.
Kauffman, S. 1993. The origins of order. Oxford: Oxford University Press.
Kauffman, S. 2016. A creative universe: The social world. Oxford: Oxford University Press.
Kilduff, M., Angelmar, R. and Mehra, A., 2000. Top management-team diversity and firm performance: Examining the role of cognitions. *Organization science*, **11(1)**, pp.21-34.
Klette TJ, Kortum S. 2004. Innovating firms and aggregate innovation. *Journal of Political Economy* **112(5)**: 986–1018.
Lancaster T. 1990. *The Econometric Analysis of Transition Data*. Cambridge University Press: Cambridge.
Lancaster T. 2000. The incidental parameter problem since 1948. *Journal of Econometrics* **95(2)**, 391-413.
Lehman CL, Tilman D. 2000. Biodiversity, stability, and productivity in competitive communities. *The American Naturalist* **156(5)**, 534-552.
Levitt B, March JG. 1993. Myopia of learning. *Strategic Management Journal* **14**, 95-112.

Loreau, M. 2010. Linking biodiversity and ecosystems: towards a unifying ecological theory. *Philosophical Transactions of the Royal Society* **365**, 49-60. Loreau, M, de Mazancourt, C. 2013. Biodeiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecology Letters* **16**, 106-115.
Lumpkin, G., Wales, W., Ensley, M. 2006. Entrepreneurial orientation effects on new venture performance: The moderating role of venture age. Academy of Management Proceedings 1, N1-N6.
Marshall, A. 1898. Principles of economics. London: Macmillan and Co.
Mata J, Portugal P. 1994. Life duration of new firms. *The Journal of Industrial Economics* **42(3)**, 227-245.
Mata J, Portugal P, Guimaraes P. 1995. The survival of new plants: Start-up conditions and post-entry evolution. *International Journal of Industrial Organization* **13**, 459-481.
Mata J, Portugal P. 2002. The survival of new domestic and foreign owned firms. *Strategic Management Journal* **23**, 323-343.
May RM. 1974. Stability and Complexity in Model Ecosystems. 2nd ed. NJ: Princeton University Press.
McCann KS. 2000. The diversity-stability debate. *Nature* **405**(11 May), 228-233.
McCloughan P, Stone I. 1998. Life Duration of Foreign Multinational Subsidiaries: Evidence from UK Northern Manufacturing Industry 1970-93. *International Journal of Industrial Organization* **16(6)**, 719-747.
McKelvey, B. 1999. Avoiding complexity catastrophe in coevolutionary pockets: Strategies for rugged landscape. *Organization Science* **10(3)**, 294-321.
McNaughton, SJ. 1993. Biodeiversity and functioning of grazing ecosystems. In *Biodiversity and Ecosystem Function*, Schulze E, Mooney H (eds). Springer-Verlag: Berlin; 361-383.
Metcalfe, S. 2005. Evolutionary concepts in relation to evolutionary economics. In: "The evolutionary foundations of economics", ed. Kurt Dopfer, pp.391-430, Cambridge: Cambridge University Press.
Mougi, A., Kondoh, M. 2012. Diversity of interaction types and ecological community stability. *Science* **337**, 349-350.
Naylor J, Smith AF. 1982. Applications of a method for the efficient computation of posterior distributions. *Journal of the Royal Statistical Society, Series C* **31**, 214-225.
Nelson R, 1995. Co-evolution of industry structure, technology and supporting institutions, and the making of comparative advantage. *International Journal of the Economics of Business* **2**(2), 171-184.
Nelson R, Winter S. 1982. *An Evolutionary Theory of Economic Change*. Harvard University Press: Cambridge, MA.
Nicoletti C, Rondinelli C. 2010. The (mis)specification of discrete duration models with unobserved heterogeneity: A Monte Carlo study. *Journal of Econometrics* **159(1)**, 1-13.
Noda T, Collis DJ. 2001. The evolution of intra-industry firm heterogeneity: Insights from a process study. *Academy of Management Journal* **44(4)**, 897-925.
Norton, E., Wang, H., Ai, C. 2004. Computing interaction effects and standard errors in logit and probit models. *The Stata Journal* **4(2)**, 154-167.
Ozsomer, A., Calantone, R., Benedetto, A. 1997. What makes firms more innovative? A look at organizational and environmental factors. *Journal of Business and Industrial Marketing* **12(6)**, 400-416.
Pavitt K. 1984. Sectoral patterns of technical change: towards a taxonomy and a theory. *Research Policy* **13(6)**, 343-373.
Pindyck, R. 1990. Irreversibility, uncertainty, and investment. NBER Working Paper 3307.
Porter M. 1980. *Competitive Strategy*. Free Press: NY.
Probst G, Raisch S. 2005. Organizational crisis: The logic of failure. *Academy of Management Perspectives* **19(1)**, 90-105.
Quinn, R., Cameron, K. 1983. Organizational Life Cycles and Shifting Criteria of Effectiveness: Some Preliminary Evidence. *Management Science* **29(1)**, 33-51.
Russell, R., Russell, C. 1992. An Examination of the Effects of Organizational Norms, Organizational Structure, and Environmental Uncertainty on Entrepreneurial Strategy. Journal of

Management **18(4)**: 639-656.Schumpeter JA. 1942. *Capitalism, Socialism and Democracy*. Harper & Row: New York.
Sharapov, D., Kattuman, P., Sena, V., 2011. Technological environments, R&D investment, and firm survival. MICRO-DYN Working Paper no. 34/10.
Simon, H.A. 1986. Rationality in psychology and economics, In R.M.Hogart and N.W.Reder (eds.), "Rationality Choice", Chicago: Chicago University Press, 25-40.
Smith, K., Guthrie, J., Chen, M. 1989. Strategy, size and performance. *Organization Studies* **10/1**, 063-081.
Sørensen, J.B. and Stuart, T.E., 2000. Aging, obsolescence, and organizational innovation. *Administrative Science Quarterly*, **45(1)**, 81-112.
Stacey, R.D. 1992. Managing the Unknowable: Strategic boundaries between order and chaos in organizations. San-Francisco: Jossey-Bass Publishers.
Stirling A. 2010. Multicriteria diversity analysis: A novel heuristic framework for apprising energy portfolios. *Energy Policy* **38**, 1622-1634.
Strzalecki, T. 2014. Depth of reasoning and higher order beliefs. *Journal of Economic Behavior and Organization* **108,** 108-122.
Teece D. 2007. Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal* **28**, 1319-1350.
Theil H. 1972. *Statistical Decomposition Analysis*. North-Holland Publishing: Amsterdam.
Tidd, J. 2001. Innovation management in context: environment, organization and performance. International Journal of Management Reviews **3(3)**, 169-183.
Ugur M, Trushin E, Solomon E. 2016a. Inverted-U relationship between R&D intensity and survival: Evidence on scale and complementarity effects in UK data. *Research Policy* **45(7)**, 1474–1492.
Ugur M, Trushin E, Solomon E. 2016b. A firm-level dataset for analyzing entry, exit, employment and R&D expenditures in the UK: 1997-2012. *Data in Brief* **8**, 153–157.
Vaccaro, I., Jansen, J., Van Den Bosch, F., Volberda, H. 2012. Management innovation and leadership: The moderating role of organizational size. *Journal of Management Studies* **49(1)**, 28-51.
Yachi, S., Loreau, M. 1999. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proceedings of the National Academy of Science USA* **96**, 1463-1468.
Yates, J.F., Stone, E.R. 1992. The risk construct. In J.F. Yates, ed., "Risk Taking Behavior", West Sussex, England: John Wiley and Son.
Wernerfelt, B., Karnani, A. 1987. Competitive Strategy Under Uncertainty. *Strategic Management Journal* **8(2)**, 187-194.
Williams, R. 2012. Using the margins command to estimate and interpret adjusted predictions and marginal effects. *The Stata Journal* **12(2)**, 308-331.
Wooldridge J. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. The MIT Press: Cambridge, MA.
Zahra S, George G. 2002. Absorptive capacity: a review, reconceptualization, and extension. *Academy of Management Review* **27**, 185-203.
Zahra, S., Neubaum, D., El-Hagrassey, G. 2002. Competitive analysis and new venture performance: Understanding the impact of strategic uncertainty and venture origin. *Entrepreneurship Theory and Practice*, Fall, 1-28.
Zahra, S., Bogner, W. 1999. Technology strategy and software new venture performance: The moderating effect of competitive environment. *Journal of Business Venturing* **15(2)**, 135-173.
Zhang M, Mohnen P. 2013. Innovation and survival of new firms in Chinese manufacturing, 2000-2006. UNU-MERIT Working Paper 2013-057. www.merit.unu.edu/publications/wppdf/2013/wp2013-057.pdf.
Zona, F., Zattoni, A., Minichilli, A. 2013. A contingency model of boards of directors and firm innovation: The moderating role of firm size. British Journal of Management **24(3)**, 299- 315.

Panel a: Theil index as measure of size and age diversity



Panel b: Coefficient of variation as measure of size and age diversity

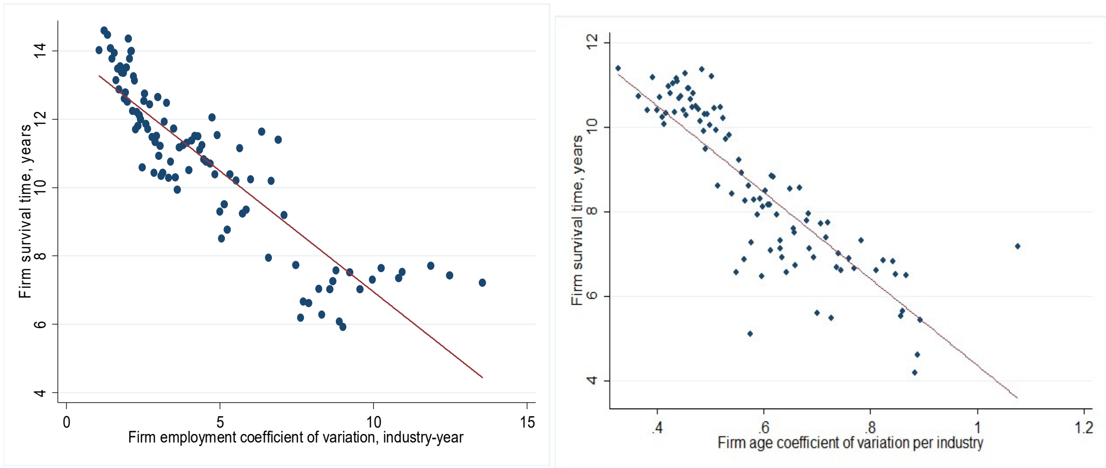
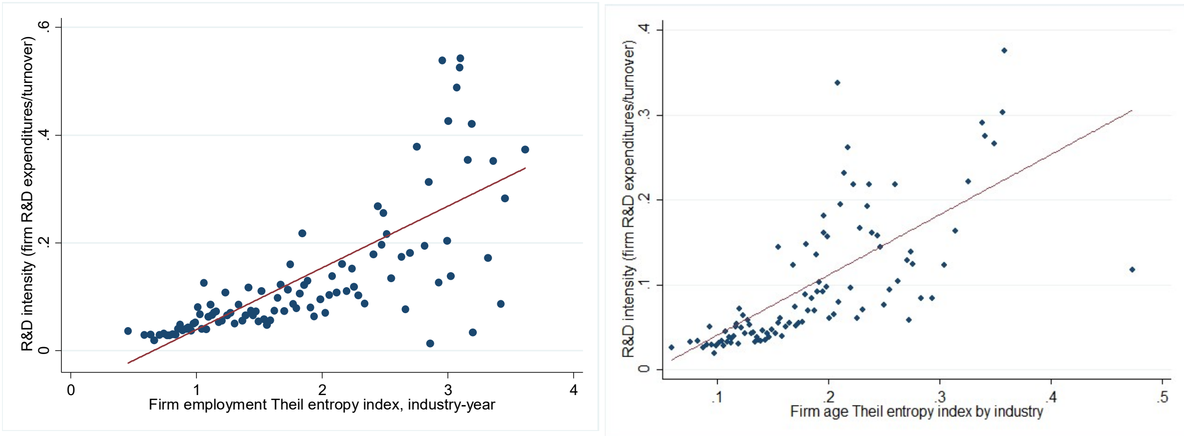


Figure 1: Intra-industry size/age diversity and survival times.

Note: Theil index, coefficient of variation and survival times are averages at the 3-digit industry level.

Panel a: Theil index as measure of size and age diversity



Panel b: Coefficient of variation as measure of size and age diversity

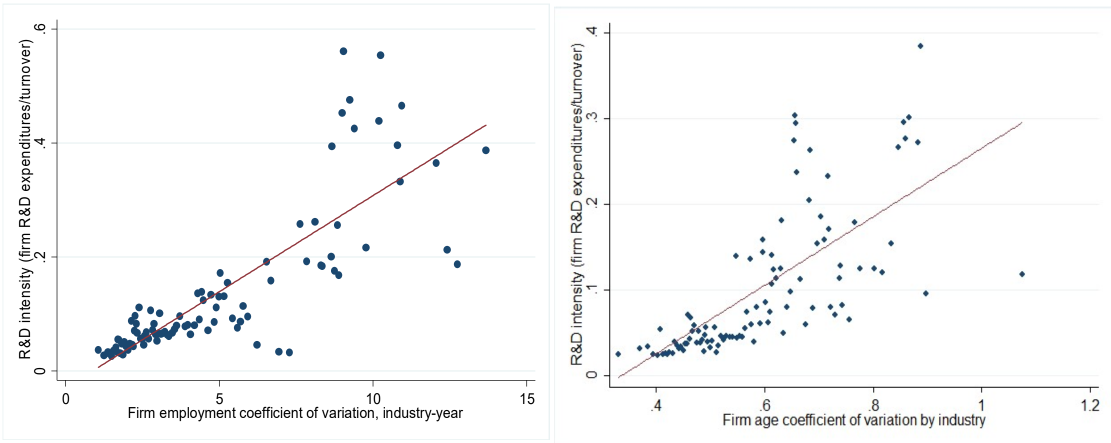
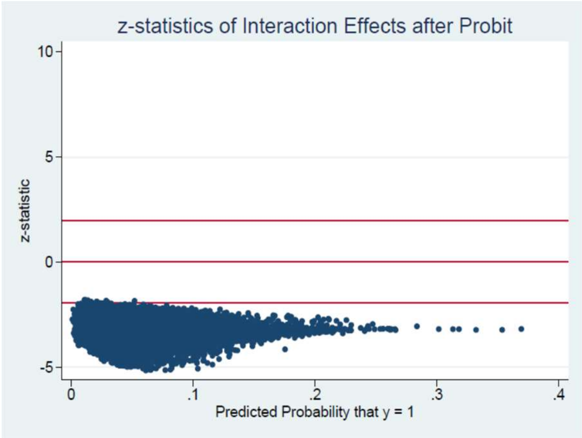
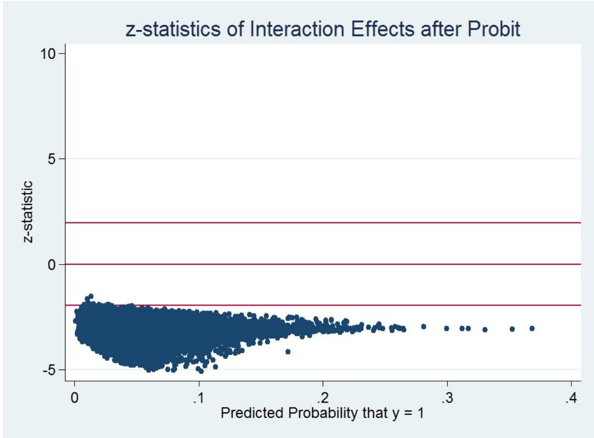
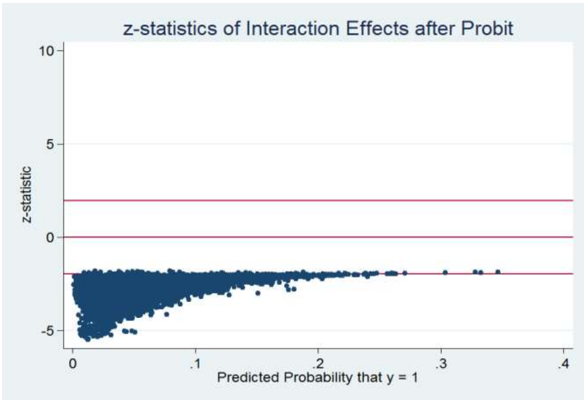


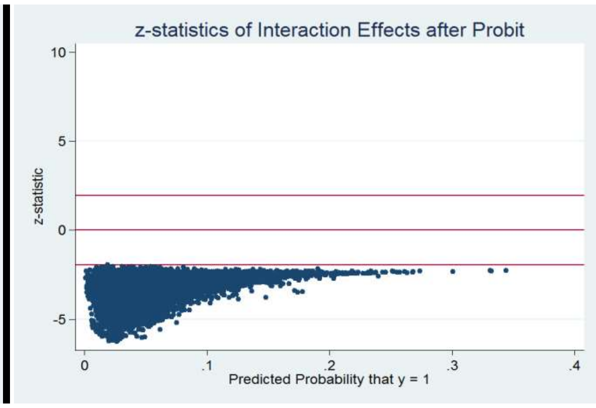
Figure 2: Measures of within-industry size/age diversity and R&D intensity.
Note: Coefficient of variation, survival times and R&D intensity are averages at the 3-digit industry level.





2a. Theil entropy of firm employment. 2b. Coefficient of variation of firm employment.





2c. Coefficient of variation of firm ages. 2d. Theil entropy of firm ages.

Figure 3: The estimated interaction effects between logarithm of R&D intensity and (a) Theil entropy of firm employment, (b) coefficient of variation of firm employment, (c) coefficient of variation of firm ages, and (d) Theil entropy of firm ages.

Table 1: Covariates and expected effects on exit hazard

Covariate	Description (expected sign of effect on exit hazard)	Related literature
	Covariates of main interest	
Firm-size and age diversity (*TI. or C.V.*)	Theil's entropy (*TI*) and coefficient of variation (*CV*) of firm employment or age in 3-digit industries **(+)**	Size diversity is not tested before for firm survival, cohort diversity is tested by Hansen and Barnett (1996) and Barnett (2008)
Interactions *TI*log(R&Dint.+1), CV*log(R&Dint.+1)*	Interaction of firm size and age diversity measures with firm R&D intensity **(-)**	Not tested before for firm survival
	Other firm-level covariates	
R&D intensity *Ln(R&Dint.+1)*	Logarithm of firm R&D intensity **(-)**	Aghion *et al*. (2013); Ericson and Pakes (1995)
Ln(R&Dint.+1) sq.	Squared logarithm of R&D intensity (+)	Aghion *et al*. (2013); Ericson and Pakes (1995); Sharapov et al. (2011); Zhang and Mohnen (2013), Ugur *et al*. (2016a).
Age *log(firm Age)*	Logarithm of firm age in years **(-)**	Hopenhayn (1992); Ericson and Pakes (1995); Geroski, 1995; Cefis and Marsili (2005); Doms *et al*. (1995); Disney *et al*. (2003)
Age squared *log(firm Age) sq.*	Squared logarithm of firm age **(+)**	Agarwal and Gort (2002); Ericson and Pakes (1995); Cefis and Marsili (2005); Evans (1987)
Size *Ln(Employment)*	Logarithm of firm employees **(-)**	Hopenhayn (1992); Ericson and Pakes (1995); Geroski, 1995; Cefis and Marsili (2005); Doms *et al*. (1995); Disney *et al*. (2003)
Size squared *Ln(Empl.) squared*	Squared log. of firm employees **(+)**	Bhattacharjee *et al*. (2009); Cefis and Marsili (2005)
Local units *Ln(Plants)*	Logarithm of firm's local units (plants) **(+)**	Audretsch and Mahmood (1995); Fernandes and Paunov (2015); Audretsch (1991); Griliches and Regev (1995); Mata *et al*. (1995)
Productivity *Ln(Rturn./empl.)*	Logarithm of deflated turnover per employee **(-)**	Audretsch, 1991; Hopenhayn (1992); Ericson and Pakes (1995)
Growth differential *(Growth_dmed)*	Growth rate of firms' deflated turnover minus median industry growth rate **(-)**	Audretsch, 1991; Hopenhayn (1992); Ericson and Pakes (1995); Cefis and Marsili (2005); Mata *et al*. (1995), Audretsch (1995), Ugur *et al*. (2016a)
Civil R&D *(Civilian R&D only)*	Dummy variable indicating that firm is engaged in civilian R&D only **(+/ -)**	Ugur *et al*. (2016a), Sharapov *et al*. (2011)
UK-owned *(UK owned)*	Dummy variable indicating that firm is UK-owned **(+ /-)**	Ugur *et al*. (2016a), Sharapov *et al*. (2011)

Industry covariates		
Concentration *Herfhindahl index*	Herfindahl-Hirschman index of firm shares in industry turnover at 3-digit industry level **(+/-)**	McCloughan and Stone (1998); Baldwin and Rafiquzzaman (1995); Wagner (1994); Geroski (1995)
Pavitt technology class* *(Pavitt #)*	Dummy variables for Pavitt classes 1 to 5, excluded category is Pavitt 4 **(+/-)**	Pavitt (1984); Agarwal and Audretsch (2001); Cefis and Marsili (2005), Ugur *et al.* (2016a)
Entry rate *log(% entry rate)*	Logarithm of firm entry rate (in %) at 3-digit SIC industry level **(+)**	Hannan and Freeman (1989); Fernandes and Paunov (2015)
Median industry R&D intensity *Ln(Med. R&D int.)*	Logarithm of industry median ratio of total R&D to turnover at 3-digit SIC level **(-)**	Audretsch and Mahmood (1995); Fernandes and Paunov (2015), Ugur *et al.* (2016a)
Average firm size in the industry *Ln(Mean empl.)*	Logarithm of average employees per firm in 3-digit SIC industry level **(+/-)**	Fernandes and Paunov (2015); Mata and Portugal (2002); Audretsch *et al.* (2004)
Macroeconomic indicators		
Crisis year *(crisis)*	A dummy variable equal 1 for the Asian crisis year of 1998; *dot.com* bubble crisis of 2001; and the recent financial crisis in 2008 **(+)**	Ugur et al. (2016a); Bhattacharjee *et al.* (2009) report higher hazard rates in periods of crises
Average real effective exchange rate *(Areer)*	Average effective exchange rate against a basket of currencies - an increases in *Areer* indicates appreciation **(+)**	Bhattacharjee *et al.* (2009); Goudie and Meeks (1991)
GDP growth rate	Growth rate of GDP, annual % **(-)**	Business cycle literature; Thompson (2005) for industry output, Mata and Portugal (2002) for employment growth

Note: * Pavitt technology classes are from Pavitt (1984), as revised slightly by Bogliacino and Pianta (2010). Pavitt1 consists of firms in science-based industries such as chemicals, office machinery, precision, medical and optical instruments industries, ICT. Pavitt2 includes specialized suppliers of technology - mechanical engineering industries, manufacturers of electrical machinery, equipment, etc. Pavitt3 includes scale-intensive industries such as pulp and paper, transport vehicles, mineral oil refining industries. Pavitt4 includes industries dominated by technology suppliers, e.g., textiles & clothing, food & drink, fabricated metals. Finally, Pavitt5 consists of unclassified industries.

Table 2: Intra-industry firm size diversity and exit hazard: Preferred estimators

	Theil entropy index		Coefficient of variation	
Dependent variable: exit in year t+1	(1)	(2)	(3)	(4)
Size Diversity	0.0377***	0.0398***	0.0103***	0.0112***
	(0.0109)	(0.0116)	(0.0030)	(0.0032)
Size Diversity*log(R&Dint.+1)	-0.244***	-0.247***	-0.0581***	-0.0588***
	(0.0588)	(0.0617)	(0.0143)	(0.0149)
Log(R&Dint.+1)	-0.578***	-0.599***	-0.755***	-0.779***
	(0.200)	(0.205)	(0.181)	(0.185)
Log(R&Dint.+1) sq.	1.250***	1.266***	1.274***	1.290***
	(0.271)	(0.282)	(0.272)	(0.283)
Log(firm Age)	-0.222***	-0.214***	-0.222***	-0.215***
	(0.0107)	(0.0114)	(0.0107)	(0.0114)
Log(Employment)	-0.197***	-0.217***	-0.197***	-0.218***
	(0.0121)	(0.0137)	(0.0121)	(0.0137)
Log(Employm.) sq.	0.0191***	0.0209***	0.0191***	0.0209***
	(0.0017)	(0.0019)	(0.0017)	(0.0019)
Log(Real turnover /	-0.0992***	-0.104***	-0.0993***	-0.104***
employees)	(0.0079)	(0.0079)	(0.0079)	(0.0079)
Firm growth relative	-0.0674***	-0.0681***	-0.0674***	-0.0681***
to industry median growth	(0.0102)	(0.0096)	(0.0102)	(0.0096)
Log(Plants)	-0.0226	-0.0202	-0.0211	-0.0186
	(0.0157)	(0.0162)	(0.0157)	(0.0162)
Civil R&D only	-0.0920***	-0.0951***	-0.0925***	-0.0957***
	(0.0130)	(0.0138)	(0.0130)	(0.0138)
UK-owned	-0.0672***	-0.0723***	-0.0665***	-0.0717***
	(0.0211)	(0.0222)	(0.0211)	(0.0222)
Log(% entry rate)	-0.0673	-0.0743	-0.0595	-0.0643
	(0.0595)	(0.0629)	(0.0599)	(0.0635)
Log(Mean industry	0.0160**	0.0152*	0.0193**	0.0187**
employment)	(0.0079)	(0.0085)	(0.0078)	(0.0083)
Log(Median R&D int.	0.633***	0.647***	0.630***	0.642***
in industry)	(0.114)	(0.120)	(0.112)	(0.119)
Herfindahl index	-0.0434	-0.0540	-0.0387	-0.0503
	(0.0614)	(0.0651)	(0.0602)	(0.0642)
Pavitt 1	-0.0637***	-0.0578**	-0.0656***	-0.0608***
	(0.0216)	(0.0232)	(0.0220)	(0.0235)
Pavitt 2	-0.101***	-0.103***	-0.101***	-0.104***
	(0.0181)	(0.0195)	(0.0182)	(0.0195)
Pavitt 3	-0.0253	-0.0218	-0.0264	-0.0228
	(0.0234)	(0.0252)	(0.0234)	(0.0251)
Pavitt 5	-0.0494*	-0.0539**	-0.0463*	-0.0505*
	(0.0257)	(0.0274)	(0.0257)	(0.0274)
Average effective real	0.0120***	0.0119***	0.0120***	0.0119***
exchange rate	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Crisis dummy	0.0660***	0.0620***	0.0661***	0.0622***
	(0.0154)	(0.0159)	(0.0154)	(0.0160)
GDP growth (%)	-0.0250***	-0.0266***	-0.0251***	-0.0266***
	(0.0038)	(0.0036)	(0.0038)	(0.0036)
Constant	-1.389***	-1.379***	-1.382***	-1.374***
	(0.116)	(0.120)	(0.116)	(0.120)
Log($\sigma_v{}^2$)		-2.351***		-2.351***
		(0.0460)		(0.0462)

N	158,316	158,316	158,313	158,313
AIC	53,695.9	53,704.9	53,692.6	53,701.1
BIC	53,935.2	53,954.2	53,931.9	53950.4
Log-likelihood	-26,824.0	-26,827.4	-26,822.3	-26,825.6
chi2	2,712.0	2,466.3	2,710.4	2,464.3
3-digit industry dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Correctly classified	95.66%	N/A	95.67%	N/A
p > Pearson χ^2	0.98	N/A	0.81	N/A
Area under ROC curve	0.69	N/A	0.68	N/A

Notes: Top R&D intensity is less than 1. Estimators: (1) and (3) – pooled Probit; (2) and (4) – panel Probit with random effects. The dependent variable is one-year-forward exit indicator, which takes the value of 1 if firm exits in year *t+1*, and zero otherwise. For definitions of Pavitt technology classes. Robust standard errors are in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. The number of firms in the estimation sample is 35,136. N/A: not applicable.

Table 3. Numerical estimations of the non-linear interaction effects in the estimated Probit random effects models.

Interaction indicator	Marginal effect at mean values by delta method (Std.Err.)	Z-stat. (st.dev)	[95% Confidence interval] for the marginal effect	Range of estimated interaction effects and their Z-statistics
Ln(R&D intensity*Theil index of firms' age	-.104** (.050)	-2.949*** (.275)	[-.203;-.006]	Effects: [-.539; -.007] Z-stat.: [-6.240; -1.956]
Ln(R&D intensity*-Coefficient of variation of firms' age	-.047*** (.028)	-2.515*** (.2387)	[-.104;-.008]	Effects: [-.244; -.003] Z-stat.: [-5.454; -1.786]
Ln(R&D intensity*-Coeff. of variation of firms' employment	-.004*** (.001)	-3.240*** (.192)	[-.006;-.002]	Effects: [-.016;-.0001] Z-stat.: [-5.069; -1.535]
Ln(R&D intensity*-Theil index of firms' employment	-.019*** (.004)	-3.393*** (.213)	[-.028;-.010]	Effects: [-.071; -.001] Z-stat.: [-5.157; -1.789]

Note: R&D intensity is less than one. Number of observations: 158,313; number of firms: 35,570. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4. Conditional effects of the intra-industry size and age diversity on exit hazard depending on R&D intensity

Percentile of R&D intensity	R&D intensity	Marginal effects at mean Theil entropy of firm employment	Marginal effects at mean coefficient of variation of firm employment	Marginal effects at mean Theil entropy of firm ages	Marginal effects at mean coef-ficient of vari-ation of firm ages
5	.0009	.0395***	.0111 ***	.9255***	.4819***
		(.0116)	(.0032)	(.1302)	(.0704)
15	.0049	.0385***	.0108***	.9181***	.4781***
		(.0116)	(.0032)	(.1302)	(.0704)
25	.0101	.0373***	.0105***	.9085***	.4733***
		(.0116)	(.0032)	(.1302)	(.0704)
35	.0169	.0356***	.0101***	.8958***	.4669***
		(.0116)	(.0032)	(.1302)	(.0704)
45	.0261	.0334***	.0096***	.8786***	.4582***
		(.0116)	(.0032)	(.1302)	(.0704)
55	.0404	.0298 ***	.0087***	.8518***	.4446***
		(.0116)	(.0032)	(.1302)	(.0704)
65	.0656	.0236**	.0073**	.8047***	.4209***
		(.0116)	(.0032)	(.1302)	(.0704)
70	.0859	.0186	.0061*	.7668***	.4017***
		(.0116)	(.0032)	(.1302)	(.0704)
75	.1122	.0121	.0045	.7176***	.3768***
		(.0116)	(.0032)	(.1302)	(.0704)
80	.1467	.0036	.0026	.6532***	.3443***
		(.0116)	(.0032)	(.1302)	(.0704)
85	.1927	-.0077	-.0001	.5672***	.3008***
		(.0116)	(.0032)	(.1302)	(.0704)
95	.3812	-.0542***	-.0112***	.2149*	.1227*
		(.0116)	(.0032)	(.1302)	(0704)

Note: Other covariates are taken at their mean. Standard errors are in parentheses. Top R&D intensity is less than one. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Conditional effects are based on estimations from panel (random-effect) probit model, reported in columns 2 and 4 of Table 2 above.

Appendix

Table A1. Summary statistics

	Survival firms		Exiter firms		Full sample	
	mean	st.d.	mean	st.d.	mean	st.d.
Theil entropy (TI)	1.8428	0.8095	2.0504	0.8489	1.8632	0.8155
lrdint_theil	0.1917	0.3379	0.2707	0.3932	0.1992	0.3442
empl_cv	4.8792	3.0757	5.6944	3.4126	4.9610	3.1215
empl_cv_lrnd	0.5729	1.1047	0.8235	1.2988	0.5970	1.1262
Theil entropy ages (TI age)	0.1902	0.0806	0.2164	0.0878	0.1875	0.0795
lrdint_theil_age	0.0272	0.0536	0.0385	0.0632	0.0259	0.0520
age_cv (CV age)	0.5975	0.1489	0.6471	0.1578	0.5923	0.1473
age_cv_lrnd	0.0790	0.1461	0.1101	0.1722	0.0753	0.1418
lenterate	0.6409	0.0967	0.6327	0.0999	0.6399	0.0980
lrdint	0.0858	0.1247	0.0812	0.1440	0.0830	0.1270
lrdint2	0.0224	0.0623	0.0213	0.0751	0.0215	0.0637
Log (age + 1)	2.6522	0.6810	2.1294	0.7415	2.6077	0.6244
Log (employment + 1)	3.0155	1.6735	2.3002	1.6126	2.9451	1.6887
Log employment squared	11.8941	12.0226	7.8913	10.2346	11.5252	11.9818
Log (live local units + 1)	0.6941	0.5670	0.5182	0.5503	0.6719	0.5706
Productivity: (LogRturn_empl)	4.3776	.9846	4.0777	1.0315	4.3492	0.9949
growth_dmed	0.0489	0.5867	-0.0297	0.7227	0.0430	0.6063
mean_empl	145.889	505.678	133.277	491.224	145.548	518.905
herfindahl	0.0982	0.1082	0.0967	0.1051	0.0982	0.1080
civil_dummy	0.4304	0.4951	0.3501	0.4770	0.4248	0.4943
uk_owner	0.8747	0.3310	0.9272	0.2596	0.8787	0.3264
RnDint_med	0.0836	0.1118	0.1201	0.1339	0.0973	0.1061
lmean_empl	4.3838	0.9524	4.0748	1.0791	4.3635	0.9655
lRnDint_med	0.0679	0.0882	0.0907	0.1045	0.0701	0.0902
Pavitt1	0.3149	0.4644	0.3974	0.4894	0.3233	0.4677
Pavitt2	0.2244	0.4171	0.2065	0.4049	0.2233	0.4164
Pavitt3	0.1005	0.3007	0.0753	0.2639	0.0981	0.2975
Pavitt5	0.0654	0.2473	0.0694	0.2543	0.0657	0.2478
Average effective exchange rate index	92.082	9.4775	93.931	9.1565	92.2903	9.4681
dummy for crisis years: 1998, 2001, and 2008	0.1539	0.3608	0.1936	0.3952	0.1556	0.3625
GDP growth rate, %	1.5457	2.1824	1.6766	2.3733	1.5571	2.1834
Number of firms	28,287		6,849		35,136	
Observations/counts	151,467				158,316	

* Note: minimum and maximum values are suppressed to comply with non-disclosure requirements of the data hosts, UK Data Service. Pavitt technology classes: 1 - science-based industries; 2 - specialised suppliers of technology; 3 - scale-intensive industries; 4 - industries dominated by suppliers of technology; 5 – unclassified. Onset of crisis dummy takes value of 1 if year is either 1998, 2001 or 2008. Turnover is deflated by 2-digt output deflator with base year at 2010. The Herfindahl index is based on firm turnover at 3-digit industry level. The GDP growth rate and the average effective real exchange rate are from the World Bank Development Indicators 2016 (www.data.worldbank.org/data-catalog/world-development-indicators).

Table A2a. Theil index measure of intra-industry firm size diversity and exit hazard: Different estimators

	(1)	(2)	(3)	(4)
Theil size index (TI size)	0.0820***	0.0793***	0.0827***	0.0795***
	(0.0240)	(0.0234)	(0.0243)	(0.0235)
TI size*log(R&Dint.+1)	-0.522***	-0.500***	-0.522***	-0.499***
	(0.122)	(0.116)	(0.122)	(0.115)
Log(R&Dint.+1)	-1.253***	-1.192***	-1.256***	-1.192***
	(0.428)	(0.412)	(0.430)	(0.411)
Log(R&Dint.+1) sq.	2.623***	2.483***	2.623***	2.481***
	(0.568)	(0.543)	(0.570)	(0.542)
Log(firm Age)	-0.463***	-0.444***	-0.459***	-0.442***
	(0.0231)	(0.0224)	(0.0237)	(0.0230)
Log(Employment)	-0.426***	-0.412***	-0.432***	-0.415***
	(0.0264)	(0.0256)	(0.0282)	(0.0270)
Log(Employm.) sq.	0.0411***	0.0398***	0.0417***	0.0400***
	(0.0037)	(0.0036)	(0.0039)	(0.0038)
Log(Real turnover /	-0.220***	-0.212***	-0.221***	-0.213***
employees)	(0.0173)	(0.0166)	(0.0173)	(0.0166)
Firm growth relative	-0.136***	-0.127***	-0.136***	-0.126***
to industry median growth	(0.0213)	(0.0199)	(0.0214)	(0.0200)
Log(Plants)	-0.0656*	-0.0662*	-0.0637*	-0.0654*
	(0.0353)	(0.0345)	(0.0358)	(0.0347)
Civil R&D only	-0.200***	-0.194***	-0.201***	-0.194***
	(0.0285)	(0.0277)	(0.0286)	(0.0277)
UK-owned	-0.158***	-0.155***	-0.159***	-0.155***
	(0.0484)	(0.0474)	(0.0489)	(0.0477)
Log(% entry rate)	-0.150	-0.146	-0.151	-0.146
	(0.129)	(0.125)	(0.130)	(0.126)
Log(Mean industry	0.0364**	0.0352**	0.0358**	0.0349**
employment)	(0.0173)	(0.0167)	(0.0172)	(0.0165)
Log(Median R&D int.	1.413***	1.364***	1.414***	1.364***
in industry)	(0.236)	(0.225)	(0.239)	(0.226)
Herfindahl index	-0.0779	-0.0718	-0.0824	-0.0736
	(0.137)	(0.133)	(0.138)	(0.134)
Pavitt 1	-0.142***	-0.139***	-0.140***	-0.138***
	(0.0478)	(0.0464)	(0.0486)	(0.0469)
Pavitt 2	-0.213***	-0.205***	-0.214***	-0.205***
	(0.0403)	(0.0392)	(0.0408)	(0.0395)
Pavitt 3	-0.0565	-0.0557	-0.0549	-0.0551
	(0.0531)	(0.0518)	(0.0539)	(0.0523)
Pavitt 5	-0.116**	-0.114**	-0.118**	-0.114**
	(0.0558)	(0.0540)	(0.0561)	(0.0539)
Average effective real	0.0266***	0.0259***	0.0265***	0.0258***
exchange rate	(0.0020)	(0.0020)	(0.0021)	(0.0020)
Crisis dummy	0.140***	0.134***	0.138***	0.133***
	(0.0330)	(0.0318)	(0.0333)	(0.0321)
GDP growth (%)	-0.0589***	-0.0580***	-0.059***	-0.058***
	(0.0087)	(0.0085)	(0.0088)	(0.0086)
Constant	-2.477***	-2.541***	-2.480***	-2.543***
	(0.260)	(0.253)	(0.262)	(0.254)
lnsig2u			-2.648**	-3.569
_cons			(1.184)	(2.595)
N	158,316	158,316	158,316	158,316
AIC	53,709.5	53,715.4	53,710.9	53,717.2

BIC	53,948.8	53,954.7	53,960.2	53,966.5
3-digit industry dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes

Notes: Top R&D intensity is less than 1. Estimators: (1) – pooled Logit; (2) - pooled Clog-log; (3) - Logit random effects; (4) – Clog-log random effects. See Table 2 for other notes.

Table A2b. Coefficient of variation measure of intra-industry firm size diversity and exit hazard: Different estimators

	(1)	(2)	(3)	(4)
Employment CV	0.0224***	0.0217***	0.0227***	0.0218***
	(0.0066)	(0.0064)	(0.0067)	(0.0064)
Employment CV*Log(R&Dint.+1)	-0.124***	-0.119***	-0.124***	-0.119***
	(0.0299)	(0.0286)	(0.0299)	(0.0284)
Log(R&Dint.+1)	-1.633***	-1.556***	-1.637***	-1.556***
	(0.389)	(0.374)	(0.391)	(0.374)
Log(R&Dint.+1) sq.	2.674***	2.534***	2.674***	2.532***
	(0.569)	(0.544)	(0.572)	(0.544)
Log(firm Age)	-0.464***	-0.445***	-0.459***	-0.443***
	(0.0231)	(0.0223)	(0.0238)	(0.0230)
Log(Employment)	-0.427***	-0.413***	-0.434***	-0.416***
	(0.0263)	(0.0256)	(0.0282)	(0.0269)
Log(Employm.) sq.	0.0410***	0.0397***	0.0416***	0.0399***
	(0.0037)	(0.0036)	(0.0039)	(0.0038)
Log(Real turnover /	-0.220***	-0.212***	-0.221***	-0.213***
employees)	(0.0173)	(0.0166)	(0.0173)	(0.0166)
Firm growth relative	-0.136***	-0.126***	-0.136***	-0.126***
to industry median growth	(0.0213)	(0.0199)	(0.0214)	(0.0200)
Log(Plants)	-0.0625*	-0.0632*	-0.0603*	-0.0623*
	(0.0354)	(0.0345)	(0.0360)	(0.0348)
Civil R&D only	-0.202***	-0.195***	-0.202***	-0.195***
	(0.0285)	(0.0277)	(0.0287)	(0.0277)
UK-owned	-0.157***	-0.154***	-0.158***	-0.154***
	(0.0484)	(0.0474)	(0.0489)	(0.0477)
Log(% entry rate)	-0.131	-0.127	-0.131	-0.127
	(0.130)	(0.126)	(0.131)	(0.126)
Log(Mean industry	0.0435**	0.0421**	0.0430**	0.0419***
employment)	(0.0170)	(0.0164)	(0.0169)	(0.0162)
Log(Median R&D int.	1.405***	1.357***	1.406***	1.357***
in industry)	(0.232)	(0.221)	(0.236)	(0.223)
Herfindahl index	-0.0715	-0.0664	-0.0770	-0.0685
	(0.134)	(0.130)	(0.135)	(0.131)
Pavitt 1	-0.148***	-0.145***	-0.145***	-0.144***
	(0.0487)	(0.0474)	(0.0496)	(0.0478)
Pavitt 2	-0.216***	-0.208***	-0.217***	-0.208***
	(0.0404)	(0.0393)	(0.0410)	(0.0396)
Pavitt 3	-0.0591	-0.0583	-0.0574	-0.0577
	(0.0530)	(0.0517)	(0.0538)	(0.0522)
Pavitt 5	-0.110**	-0.107**	-0.111**	-0.108**
	(0.0559)	(0.0541)	(0.0562)	(0.0539)
Average effective real	0.0265***	0.0258***	0.0264***	0.0258***
exchange rate	(0.0020)	(0.0020)	(0.0021)	(0.0020)
Crisis dummy	0.140***	0.134***	0.138***	0.133***
	(0.0330)	(0.0318)	(0.0333)	(0.0321)
GDP growth (%)	-0.0592***	-0.0583***	-0.0595***	-0.0584***

	(1)	(2)	(3)	(4)
	(0.0088)	(0.0086)	(0.0088)	(0.0086)
Constant	-2.463***	-2.527***	-2.466***	-2.529***
	(0.259)	(0.252)	(0.262)	(0.254)
lnsig2u			-2.511**	-3.540
_cons			(1.043)	(2.527)
N	158,313	158,313	158,313	158,313
AIC	53,706.2	53,712.0	53,707.6	53,713.8
BIC	53,945.5	53,951.4	53,956.9	53,963.1
3-digit industry dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes

Notes: Notes: Top R&D intensity is less than 1. Estimators: (1) – pooled Logit; (2) - pooled Clog-log; (3) - Logit random effects; (4) – Clog-log random effects. See Table 2 for other notes.

Table A3. Effects of intra-industry age diversity on firm exit hazard.

	(1)	(2)	(3)	(4)
Firm age coefficient of	0.443***		0.469***	
variation (CV age)	(0.0661)		(0.0706)	
CV age*Log(R&Dint.+1)	-0.902**		-0.876**	
	(0.351)		(0.373)	
Theil firm age index (TI age)		0.853***		0.909***
		(0.123)		(0.130)
TI age*log(R&Dint.+1)		-1.808***		-1.748***
		(0.615)		(0.655)
Log(R&Dint.+1)	-0.489*	-0.681***	-0.536*	-0.723***
	(0.273)	(0.208)	(0.284)	(0.214)
Log(R&Dint.+1) sq.	1.171***	1.170***	1.186***	1.184***
	(0.270)	(0.270)	(0.281)	(0.280)
Log(firm Age)	-0.0689	-0.0642	0.0091	0.0145
	(0.0596)	(0.0596)	(0.0632)	(0.0632)
Log(firm Age) sq.	-0.0308**	-0.0319**	-0.0459***	-0.0471***
	(0.0127)	(0.0127)	(0.0136)	(0.0135)
Log(Employm.)	-0.193***	-0.192***	-0.214***	-0.213***
	(0.0120)	(0.0120)	(0.0137)	(0.0137)
Log(Employm.) sq.	0.0186***	0.0186***	0.0205***	0.0205***
	(0.0016)	(0.0017)	(0.0019)	(0.0019)
Log(Real turnover /	-0.0967***	-0.0967***	-0.102***	-0.102***
employees)	(0.0078)	(0.0078)	(0.0079)	(0.0079)
Firm growth relative	-0.0706***	-0.0705***	-0.0716***	-0.0717***
to industry median growth	(0.0101)	(0.0101)	(0.0096)	(0.0096)
Log(Plants)	-0.0228	-0.0225	-0.0208	-0.0205
	(0.0156)	(0.0156)	(0.0162)	(0.0162)
Civil R&D only	-0.0922***	-0.0938***	-0.0947***	-0.0965***
	(0.0130)	(0.0130)	(0.0138)	(0.0138)
UK-owned	-0.0682***	-0.0680***	-0.0733***	-0.0731***
	(0.0211)	(0.0211)	(0.0222)	(0.0222)
Log(% entry rate)	-0.0394	-0.0404	-0.0475	-0.0480
	(0.0595)	(0.0594)	(0.0627)	(0.0627)
Log(Mean industry	0.0325***	0.0313***	0.0329***	0.0319***
employment)	(0.0081)	(0.0080)	(0.0087)	(0.0086)
Log(Median R&D int.	0.476***	0.469***	0.485***	0.475***
in industry)	(0.112)	(0.112)	(0.117)	(0.117)

Herfindahl index	-0.104*	-0.0993	-0.118*	-0.114*
	(0.0627)	(0.0631)	(0.0659)	(0.0658)
Pavitt 1	-0.0545**	-0.0518**	-0.0486**	-0.0455**
	(0.0214)	(0.0214)	(0.0228)	(0.0228)
Pavitt 2	-0.108***	-0.105***	-0.112***	-0.109***
	(0.0179)	(0.0179)	(0.0194)	(0.0193)
Pavitt 3	-0.0181	-0.0147	-0.0148	-0.0110
	(0.0234)	(0.0234)	(0.0252)	(0.0252)
Pavitt 5	-0.0630**	-0.0555**	-0.0699**	-0.0619**
	(0.0257)	(0.0256)	(0.0276)	(0.0275)
Average effective real	0.0112***	0.0111***	0.0109***	0.0109***
exchange rate	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Crisis dummy	0.0630***	0.0624***	0.0579***	0.0572***
	(0.0154)	(0.0154)	(0.0160)	(0.0160)
GDP growth (%)	-0.0235***	-0.0233***	-0.0251***	-0.0248***
	(0.0038)	(0.0038)	(0.0036)	(0.0037)
Constant	-1.775***	-1.673***	-1.845***	-1.739***
	(0.136)	(0.131)	(0.141)	(0.136)
lnsig2u			-2.344***	-2.343***
_cons			(0.0486)	(0.0489)
N	158313	158316	158313	158316
AIC	53660.3	53661.2	53662.4	53662.6
BIC	53909.6	53910.5	53921.7	53921.8
3-digit industry dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes

Standard errors in parentheses. R&D intensity < 1; by age. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Number of firms - 35,570. (1) and (2) – Pooled probit estimations; (3) and (4) - Probit random-effects estimations.