

# Innovation and the Survival of New Firms Across British Regions

Christian Helmers  
Wolfson College  
University of Oxford

Mark Rogers  
Harris Manchester College  
University of Oxford

December 2007

## ABSTRACT

*This paper analyses the survival of the complete cohort of more than 162,000 limited companies incorporated in Britain in 2001 over the subsequent five-year period. For this purpose, we estimate firms' hazards of failure and survival functions using nonparametric and semi-parametric techniques. The paper focuses on two important policy-related issues. The first is to what extent survival rates vary across regions in Britain. The Regional Development Agencies (RDA) Act 1998 has led to the establishment of 12 RDA's in Britain and it is interesting to ask if and how firm survival varies across them. A second, and related, policy issue concerns innovation. The data available allows us to look at the intellectual property (IP) activity of all British firms, including that of the 163,000 new firms in 2001. The results indicate substantial differences in survival rates across RDAs. IP activity increases probability of survival. These differences across regions and the importance of IP activity remain even when we control for a large range of industry and firm-level characteristics shaping firms' hazards of failure.*

**KEYWORDS:** Start-ups, firm survival, IP, RDA.

**JEL Classification:**

# 1 Introduction

The objective of this paper is to analyze the survival of the complete cohort of more than 162,000 limited companies incorporated in the United Kingdom in 2001 over the subsequent five-year period. For this purpose, we estimate firms' hazards of failure and survival functions using non-parametric and semi-parametric techniques. The paper focuses on two important policy-related issues. The first is to what extent survival rates vary across regions in Great Britain. The Regional Development Agencies (RDA) Act 1998 has led to the establishment of 12 RDA's in Britain. The RDAs coordinate substantial assistance to firms with the core aim of encouraging enterprise, employment and competitiveness (in 2005/6 the budget for the English RDA's was £ 2.2 billion). Given the existence of these RDAs, it is interesting to ask if and how firm survival varies across them. This gives insight into the different economic conditions facing the RDAs and the resulting challenges. We also disaggregate RDAs into county and unitary authority level data to provide further insight into spatial patterns of survival.<sup>1</sup>

A second, and related, policy issue concerns innovation. The data available allows us to look at the intellectual property (IP) activity of all British firms, including that of the 162,000 new firms in 2001. Schumpeter (1934) suggested a distinction between inventions, describing new discoveries, and innovation, describing the successful implementation of an invention into a commercial product. IP in form of patents and trade marks captures both aspects of Schumpeter's typology. Since Schmookler (1966), many researchers have used patents as an indicator of invention and knowledge creation. Use of trade mark data is less common and more recent, but there are indications it has significant explanatory power in understanding firm performance (Greenhalgh and Rogers, 2006, 2007). These papers argue that trade marks proxy product innovation. Trade marks also have the advantage of covering firms in service-based activities. Given the importance attributed to inventions and innovation by Schumpeter and many others in determining firm performance and shaping industry dynamics, IP should be expected to exert an influence on the most fundamental measure of firm performance, namely survival.

While both the regional and IP aspect of firm survival are of interest to economists and policy makers alike, we are also interested in combining these two aspects. For example, does a firm that is IP active in North East England have the same chance of survival as an IP active firm in South East England? These types of questions focus on one of the critical functions of RDAs: to support innovative firms.

---

<sup>1</sup>Due to data restrictions, we exclude Northern Ireland from our analysis.

The paper is structured as follows. First, the literature on firm survival and IP is reviewed. Second, the Oxford Firm Level IP database, which makes the analysis possible, is discussed. Third, we provide an overview of survival rates across the different British regions. Fourth, we provide Lifetable estimates for IP-active and inactive firms and across RDA regions. We then estimate a Cox proportional hazard model to control for other covariates and hence check the regional non-parametric results. This section also includes the IP variables as covariates.

## 2 Related literature

### 2.1 Survival analysis

Survival analysis has been used in several occasions in the economics literature to analyze the determinants of firm failure. These papers are strikingly similar in their estimation approaches, but differ substantially in terms of the type and depth of the data studied. Overall, the existing papers mostly use small, in part selected, industry- or firm-level panels. Usually, papers that study firms' innovative activity use even smaller data sets of industry-specific subsets of firms, hence are unable to take account of both firm-specific and industry-level characteristics. The main conclusion derived from the existing literature is that firms' innovativeness matters for firm survival. At the same time, it appears that firms in highly innovative environments are exposed to higher risks of exit. Other variables that have been found to play a role are (start-up) firm-size, firm-structure/ownership, industry growth, and the number of firms entering and leaving the industry. Overall, most papers are unable to include firm-level data for IP in the form of patents, and none of the papers reviewed here takes account of trade marks.

Disney, Haskel and Heden (2003) provide an analysis of around 140,000 manufacturing establishments in the UK from 1996 to 1991. They find that only around 35 percent of new entrants will survive after five years. Those that do survive are around four times larger than new entrants (in terms of employment). They use non-parametric estimates and a Cox proportional hazards model and, although they are limited in terms of possible explanators, find that establishments that are part of groups have lower exit rates. This represents, to our knowledge, the only recently published evidence on UK firms (although the data is for manufacturing firms from 1986 to 1991).

Mata and Portugal (1994) were among the first to use survival analysis to study firm exit. They track a cohort of Portuguese manufacturing firms born in 1983 to analyze determinants of their eventual failure. Mata and Portugal's data set is broad in coverage including all manufacturing firms with at least five employees. They use

the Cox proportional hazards model to find that start-up size of the new-born firms, industry growth and the number of plants operated by the new-born firms reduce the likelihood of failure, while entry into the industry increases the likelihood. They do not account for innovative activities of the firms studied.

Mahmood (1992) and Audretsch and Mahmood (1995) use a Cox proportional hazards model to study firm survival of about 12,000 US firms founded in 1976. While still not accounting for firm-level IP, these authors do take account of firms' innovative environment at the industry-level. Audretsch and Mahmood (1995) use the approach adopted by Audretsch (1991), who analyzed firm survival at the industry level, and apply it to firm-level data by adding firm-specific determinants of survival, such as ownership and firm size, although not IP firm-level information. Audretsch (1991) identified three factors influencing firm survival. Firstly, how much the new firm's output is below the minimum efficiency scale (MES). Secondly, firm survival is influenced by the firm's technological environment. Audretsch and Mahmood construct this measure by dividing the total number of innovations by industry employment (innovation data come from expert, subjective assessments of industries). Finally, market structure is captured through industry growth and competition variables. Applying this setting to the firm level, Audretsch and Mahmood's main finding is that technological conditions matter. Firms in highly innovative environments face higher risk of exit. Firms in industries with a large degree of scale economies, captured by the capital intensity variable, have a higher likelihood to exit. However, similar to Mata and Portugal, the larger the start-up size of the firm, the lower is the likelihood to exit. In addition, they find some indication that also the macroeconomic environment matters, with unemployment increasing the risk of failure. For firm-specific characteristics, the authors find that stand-alone companies have a lower likelihood to fail compared to branches or subsidiaries. This finding may appear counterintuitive, as one might expect subsidiaries to receive support from parents. However, subsidiaries may be under pressure to perform and their parents may be quick to close them down if they do not.

Suarez and Utterback (1995) use survival analysis to test whether firm survival is directly influenced by a firm's entry timing with respect to the evolution of technology in the industry, measured by the emergence of a 'dominant design'. This idea is very similar to the notion of shakeout and escalation put forward by Jovanovic and MacDonald (1994) and Sutton (1991, 1998). Suarez and Utterback conjecture that the probability of survival is greater for firms entering the industry before the emergence of a dominant design than for firms entering later. They claim that this relationship obtains as a dominant design results in barriers to entry erected by those firms that acquire the dominant design. These firms have the possibility to accumulate assets such as market knowledge, distribution networks and reputation; they can also exploit

economies of scale and raise entry barriers, forcing others to exit. On the other hand, the earlier a firm enters the market, the more it learns about the market, hence the higher its chances to come up with the dominant design and to belong to the ‘winner’ group of firms. They find that the time before and after the emergence of a dominant design in the industry indeed affects a firm’s survival function. For each year a firm entered the industry prior to the emergence of the dominant design reduces the hazard of failure by 6-15 percentage points.

Agarwal and Audretsch (2001) also consider the product life cycle. They use a data set which is organized by products covering 33 products produced by 3,431 US-American firms. Given the level of disaggregation of their data set, Agarwal and Audretsch are able to determine the stages of products within their life-cycles. The authors derive a similar result to Suarez and Utterback using Lifetable analysis and by estimating a Cox proportional hazard model. They show that the empirical relationship between firm size and survival is influenced by technology and the industry product life cycle. While it is still true in general that smaller entry size warrants higher risk of failure, this is no longer true at later stages of the product life cycle and for technologically intensive products. The authors conjecture that firms enter a more mature and technologically intensive market mainly to fill a niche; hence, the usual size-survival relationship does not apply. Unlike any of the previous papers discussed, Agarwal and Audretsch control for exit through merger and acquisition (M&A). If firms merged during the period observed, the smaller firm in the merger is recorded as exited, while the larger one is kept in the data base.

Cefis and Marsili (2005) use firm-level dummies to distinguish between innovating and non-innovating firms in a cohort of new-born Dutch firms over the period 1996-2003. They find that innovators benefit from an innovation premium giving them higher life expectancy (11 percent higher survival time for innovating firms). As they also distinguish between product and process innovations,<sup>2</sup> they find that firms introducing process innovations experience a 25 percent increase in survival time, while product innovations do not have any statistically significant effect. Using Pavitt’s (1984) sector classification according to their technological intensity, they also find that firms in technologically more intensive sectors have higher chances of survival, which stands in direct contrast to the findings by Audretsch and Mahmood (2005). According to their results, firm survival is also influenced by age and size. However, size has decreasing effects on firm survival. Moreover, they find that growth is more important for survival than initial size of a firm. Cefis and Marsili observe in their data set only the initial

---

<sup>2</sup>The information on innovation comes from the Community Innovation Survey (CIS-2).

conditions of the firm (i.e., production and technological conditions). They are unable to distinguish in exit due to failure or M&A. In contrast to our analysis, Cefis and Marsili define their cohort of firms as all firms in their data set alive in 1996. It is, however, not clear whether they accounted for the resulting problem of left truncation in their estimation procedure.

Cockburn and Wagner (2007) use an approach that is, in some ways, the closest to the approach pursued in this paper. Cockburn and Wagner directly measure the effect of patenting on firm survival. They use a small sample of 356 internet-related firms that made an IPO on the NASDAQ during the dot-com boom between 1998 and 2001. All firms belong to one of the categories Internet services, Internet software, or computer software. During this boom period, the USPTO made it possible to apply for patents on software and notably on business methods. The authors therefore also test specifically whether patents of that category had any different effect on survival compared to patents in ‘traditional’ categories. They find that patenting is positively related with firm survival. Firms that applied for more patents were less likely to exit the market. However, these results do not obtain for patenting of business methods - patents in that category did not provide firms with higher chances of survival. Yet, firms which held business method patents that attract more forward citations per claim appear to be more attractive targets for M&A. Cockburn and Wagner explicitly allow for exit through M&A by estimating a competing risks Cox’s proportional hazard model allowing for two modes of exit, bankruptcy/minimal market value or M&A. Cockburn and Wagner constructed their data in a very similar way to the data set used in this paper, which will be described in the following section.<sup>3</sup> However, due to the specific sample used by Cockburn and Wagner, the authors are unable to control appropriately for the environment in which the firms operate by using industry-level variables.

## 2.2 Patents and trade marks

Patents have been recognized to represent a measure of innovation output by firms. As such, patents have been found to be positively correlated with firms’ R&D, both across firms and over time for given firms (Griliches, 1990; Klette and Kortum, 2004). Thus, firms with a larger number of patents should be expected to be more innovative and, therefore, have a competitive advantage. Another advantage conferred by patents relates to their being regarded as intangible assets. As shown by Griliches (1981), firms accumulating assets in the form of patents, therefore, experience positive effects with regard to their market value. Finally, patents may also serve strategic purposes, such as

---

<sup>3</sup>They use the commercial Compustat database to retrieve US firms’ financial information and match these firms with data on patents from USPTO, Delphion Inc., EPO, and JPO.

detering and blocking competitors from entering a certain market (Hall, 2007). Patent data have been widely used in applied work for very diverse purposes. Yet, the use of patent data has many more or less well-documented pitfalls. One of the problems when using national and EPO patents arises from the fact that sometimes applicants file the same or very similar applications at both institutions. For example, firms may apply first for a patent at the UK office and then use the obtained priority date to file for an EPO patent making the same, or very similar, claims. Such patents belong to the same patent family and it is unclear whether they should all be counted as single patents.<sup>4</sup> On the one hand, since these patents are based on the same invention, they do not each represent a new innovation by the firm. On the other, as we are interested in whether firms gain any competitive advantage from patenting, such strategic patenting may matter. We therefore count patents pertaining to the same patent family as single patents. When including patent counts as a measure of innovative activity, another concern emerges with regard to possible heterogeneity across patents with respect to their actual value. In particular, it is argued that many patents are of little or no value. Given this, it would be desirable to be able to discriminate among patents. The patent literature has developed a number of possible ways to discriminate among patents, including the fact whether a patent was actually granted, the number of citations received, the respective patent family size, renewals, opposition and litigation and direct values reported by the firms themselves in surveys (van Zeebroeck, 2007). In our situation there are various problems with such methods. Using patent grants for new firms means the firms may already be 2 to 4 years old by the time the patent is granted. In a similar way, the use of citations (or renewals) means waiting 5 to 20 years after the patent is granted, meaning we could only study startups in the 1980s or 1990s. The data has no information on opposition, litigation and patent family size and, in any event, these may be more important for larger firms. Given this, we adopt a rather crude measure by distinguishing between UK and EPO patents, where EPO patents are presumably more valuable. The importance is not only reflected in distinctively higher fees for EPO patents, but due to the international scope of EPO patents, they reflect differences in business strategies.<sup>5</sup> It can be expected that only more valuable patents will be patented in several countries other than the country in which the firm is registered.

---

<sup>4</sup>If the exactly same application is made to both institutions, one of the two has to be withdrawn once one of the institutions grants the patent.

<sup>5</sup>An EPO patent application costs Euro 4000, while a UK patent application costs around Euro 300. In fact, applications for EPO are likely to be much higher since it needs to be submitted in two languages and use of a patent attorney is strongly recommended. For both EPO and UK patents, the application is published, if the application passes an initial examination, after 18 months. Hence, our use of publications means that there is an 18 month delay from submitting original invention.

The reasons why trade marks can be expected to give firms a competitive advantage are as follows. Fundamentally, trade marks have a signalling function indicating a certain level of quality or other characteristics consumers can expect from a product. As such, trade marks help consumers reduce search costs and hence producers are able to sell larger quantities or charge a higher price (Landes and Posner, 1987). It is also argued in Greenhalgh and Rogers (2007) that trade marks proxy product innovation by the firm (see also Mendoca et al., 2004). From this perspective the trade mark data proxies a range of activities that are associated with product innovation, such as marketing, advertising and design. Perhaps more importantly, trade mark data may capture which firms are better at this bundle of activities.

The use of trade mark data is much less well explored. This is surprising as trade marks applications often outnumber patents applications. In the UK, trade marks can be obtained either through an application to the UK Intellectual Property Office for a UK Trade Mark, or through an application to the Office of Harmonization for the Internal Market for a Community Trade Mark. Fees differ substantially, as the UK trade mark costs about 300 Euros while the Community trade mark costs around 2000 Euros.

### **3 The OFLIP database**

The data used for the analysis comes from the Oxford Firm Level IP database (OFLIP). The database draws on the Financial Analysis Made Easy (FAME) data that covers the entire population of registered UK firms (FAME downloads data from Companies House records). OFLIP contains additional information on the IP activity of firms in the form of patents and trade marks. OFLIP has been constructed by matching the FAME database and a number of firm-level IP datasets.<sup>6</sup>

The FAME database is a commercial database provided by Bureau van Dijk.<sup>7</sup> To construct the data set, the December 2006 edition of FAME has been used. It covers around 2.04 million active firms. For all of these firms, basic information, such as name, registered address, firm type, and industry code are available. Availability of financial information varies substantially across firms. The smallest firms are legally required to submit only very basic balance sheet information such as shareholders' funds and total assets. The largest firms provide a large range of profit and loss information as well as detailed balance sheet data. Importantly, the FAME database also lists around 0.9 million so called 'inactive' firms. These inactive firms are those that have

---

<sup>6</sup>For details on the matching process and further details on the database see Rogers, Helmers, and Greenhalgh (2007).

<sup>7</sup><http://www.bvdep.com/en/FAME.html>

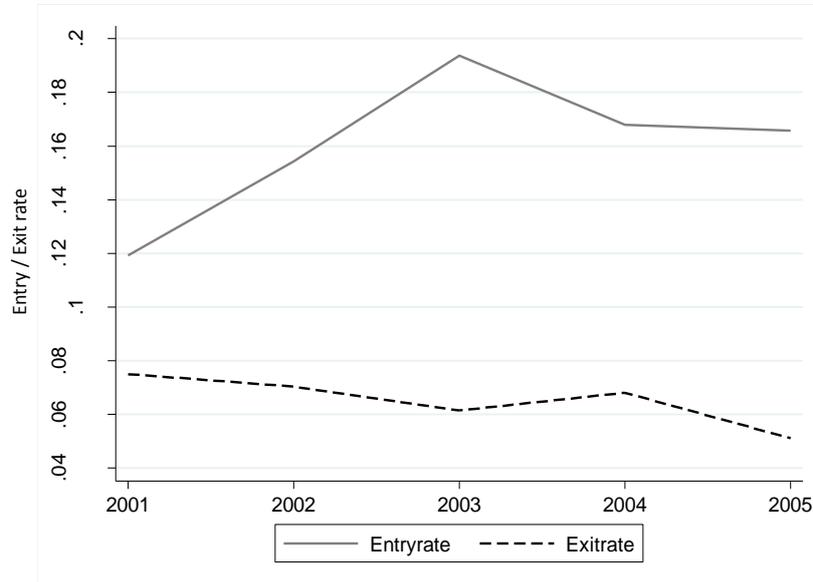
exited the market and belong to one of the following categories: dissolved, liquidated, entered receivership or declared non-trading. The fact that FAME tracks inactive firms allows us to identify all firms entering and exiting the market throughout the five-year period observed.<sup>8</sup> FAME gives exact dates for market entry in the form of a firm's incorporation date. To determine date of exit we use the date that the last set of accounts were filed.<sup>9</sup>

The IP data used for the construction of OFLIP database comes from three different sources: the UK IP Office, Marquesa Ltd. and the European Patent Office (EPO) ESPACE Bulletin. Data on UK patent publications were supplied by the UK IP Office. Marquesa Ltd supplied data on UK trade mark publications and Community (OHIM) marks registered. The Community trade mark data did include International Marks designating the EU (OHIM allowed these since 1st October 2004). Data on EPO publications by British entities was downloaded from ESPACE Bulletin DVD 2006/001. For our analysis, we use publications of UK patents, trade marks and EPO patents, as well as registrations for Community trade marks. For each of these, the data set also includes an associated application date, although it is only available for IP that has been published. Equally, for those patents which succeeded in being granted also the grant year is available. IP is commonly regarded as an output measure of innovation. Hence, if the objective is to measure whether innovation output conveys a firm any competitive edge over its competitors, it seems appropriate to use publication date (registration in case of Community trade marks) as the reference point.

The data has been complemented through information on M&A extracted from the Bureau van Dijk ZEPHYR database. This allows use to identify if a company exited due to M&A. Ideally, we would estimate a competing risks model, where the possibility of firms exiting the market in various ways is modeled explicitly. However, there were only 14 of the 2001 cohort of firms exited due to a M&A. Due to this small number, these firms are simply counted as survivors.<sup>10</sup>

In summary, the OFLIP database contains the population of UK registered companies together with their IP activity for the period 2001 to 2005.

Figure 1: Entry / Exit rates for population of UK firms 2001-2005



## 4 Overview of the data

Entry and exit rates for the population of registered firms in the UK between 2001 and 2005 are plotted in Figure 1. The exit rate slightly decreased over the period observed, while the entry rate substantially increased between 2001 and 2003. The reason for this was a change in the tax law in 2002. The UK Government introduced a zero per cent rate of corporation tax for registered companies with profits up to £ 10,000. The result of this was to rapidly increase the number of sole traders that formed companies (such as tradespeople and consultants). Over the 5-year period shown, entry rates exceed exit rates on average by 9.5 percentage points. This implies an average annual net increase of around 170,000 firms in the UK. The influence of the tax change is therefore substantial and means that entry rates should be treated with caution.<sup>11</sup> In particular, it is thought that many IT consultants and business service sole traders converted into

<sup>8</sup>As such, the data set is much wider in coverage than for example the one used by Bridges and Guariglia (2006) who also use the FAME database for their firm survival analysis.

<sup>9</sup>This is an important advantage of OFLIP over census data sets used in previous work, such as Dunne et al. (1987). In cases where the exit date was missing, the date of last annual return or the date of last transaction at companies house is used instead.

<sup>10</sup>Disney, Haskel and Heden (2003) also found very low numbers for their UK firm-level data set.

<sup>11</sup>It is not clear whether exit rates would also be biased downwards. In general, we might expect some exit because a firm wants to convert back to a sole trader (avoiding the slightly higher administrative costs of a registered company). The tax change after 2002 would appear to outweigh these administrative costs, hence we might expect exit to be slightly lower.

registered companies in 2002 and 2003. To counter this somewhat unintended effect, the government introduced new legislation taxing all the company's profit at 19 percent if the profits earned were distributed to shareholders. Hence, the zero percent tax rate applied only up to £ 10,000, if the profits were retained.<sup>12</sup>

In general, as Geroski (1991) has recorded, there is a positive correlation between entry rates and exit rates across industries in the UK. Using data from 1987 he reported a correlation coefficient of 0.79 for a sample of 95 industries. Here, we find a correlation coefficient of only 0.20 for the five year period for a sample of 252 industries. Looking at individual years, the correlation coefficients vary between 0.004 in 2003 and 0.28 in 2001. Given the tax driven increase in entry rates in 2002 and 2003, the low correlation might be expected. The possibility that the tax changes had a non-uniform effect on entry across industries, also means that it is not possible to compare the correlation coefficients obtained here with Geroski's.

Our analysis focuses on the cohort of firms incorporated in 2001, this avoids the problem arising from increased entry rates due to the change in tax regime in 2002. Using only firms incorporated in 2001 also avoids problems of left truncation as in Cefis and Marsili (2005), i.e., all firms are observed from the onset of failure risk, and ensures that there are no unobserved effects driving survival stemming from the time elapsed previous to the observed period.

As is well known in the literature, the rate of firm failure is largest during the first few years of existence. Mata and Portugal (1994), for example, found for a data set of Portuguese manufacturing firms that 20 percent of the new-born firms fail during their first year of existence and only half of the firms survive for four years. For our data, there were a total of 162,469 new firms registered in 2001. The survival rates for this cohort are as follows. In 2002, 161,493 or 99.4 per cent were still in business. The high figure simply reflects that a registered company almost always survives to file its first set of accounts. By 2003, 140,215 or 86.3 per cent were still in business, with 76.4 per cent left by 2004, and 70.24 per cent by 2005.<sup>13</sup>

#### 4.1 Regional differences

In order to obtain a first understanding of regional differences, Figure 4 plots the number of new incorporations by regional development authority (RDA). Darker areas show RDAs with a higher number of new incorporations in 2001. The darkest region

---

<sup>12</sup>These somewhat confusing regulations are now to be abolished and only a small company tax rate of 19 percent will apply (see HM Revenue and Customs website <http://www.hmrc.gov.uk/>).

<sup>13</sup>When the data is used for formal survival analysis, we encounter the problem of right-censoring. This means that for all firms of the 2001 cohort that have not exited by the end of 2005, failure remains unobserved. The only thing that is known about these firms is that failure occurs some time between  $[t, \infty)$ .

on the map is the London RDA region, which had 46,255 incorporations in 2001. After London, it is the South East and North West RDAs that have the most new firms in 2001. Wales, the North East and the Highlands have the lowest numbers of start up firms. Table 1 shows the exact numbers of firms. The purpose of Figure 4 is to show the substantial geographical differences and to indicate the regions we are considered. The boundaries shown within each RDA in the map are counties and unitary authorities.

Table 1: Number of new firms and failure rates by RDA in 2001

RDA	No. of			Failure rate			
	new firms	IP-active	High-tech	IP-inactive		IP-active	
				all firms	high-tech	all firms	high-tech
South West	11,687	183	46	0.254	0.143	0.082	0
South East	26,273	442	88	0.284	0.218	0.104	0.1
London	46,255	808	66	0.357	0.339	0.113	0
East of England	11,763	198	66	0.288	0.236	0.091	0
East Midlands	7,873	147	25	0.261	0.208	0.122	0
Yorkshire	9,018	169	23	0.273	0.286	0.101	0
North West	14,957	223	50	0.301	0.298	0.148	0
West Midlands	17,482	265	61	0.289	0.241	0.094	0.286
North East	3,112	47	13	0.257	0.385	0.043	n.a.
Wales	4,387	76	25	0.268	0.261	0.079	0
Scotland	9,218	164	21	0.261	0.263	0.128	0.5
Highland	444	4	0	0.221	n.a.	0.5	n.a.

Figure 5 plots the failure ratio, which is computed as one minus the fraction of newly incorporated firms still alive in 2005 i.e.,  $1 - \frac{Firms_{2005,c}}{Firms_{2001,c}}$ . Specifically, the figure shows the failure rate for non-IP active firms (159,743 in 2001) by RDA.<sup>14</sup> London and North West RDAs have the highest failure rate: in London and in the North West, 35.7 percent and 30 percent respectively of 2001 firms fail by 2005. In contrast, the lowest failure rates are in the Scottish Highlands and the South West regions. Figure 6 also shows failure rates but this time according to the average for counties and unitary authorities (Britain has 142 of these covering the country). The figure shows five different ‘shades’ of failure rates, based on the quintiles of the distribution. For example, the darkest shaded counties have a failure rate of between 31 percent and 47 percent; the white regions have the lowest failure rate (between 5 percent and 23 percent). Clearly, these are major differences in failure rates. Note also that there are differences across counties within RDAs. This indicates that either the support given to new firms varies within RDAs or, more likely, that there are a great many other mechanisms at work in driving failure rates. Figure 6 indicates that the high failure rates occur in London, Kent, Buckinghamshire and the M4 corridor. One might suggest that

<sup>14</sup>A firm is counted as IP active if it has had any form of IP within the period 2001-2005 and correspondingly a firm is counted as inactive if it did not have any form of IP during the entire period observed.

the high rates of new firm formation in these regions make this an expected result, but clearly these regions also have greater demand and other characteristics that drive the high entry rates. An alternative explanation is to say that competition is more intense in these regions but this is, in many ways, a tautology. Explaining why competitive intensity varies across regions is a more difficult task. However, it is not only regions with high rates of firm formation that have high failure rates, the figure also shows some other high failure rate areas: Weston super Mare (directly east of London and M4 corridor); some authorities in the South Wales valleys, and also Pembrokeshire; Herefordshire; Liverpool to Manchester corridor; and South Ayrshire in Scotland.

Figures 7 and 8 show differences in failure rates for the subset of 2001 incorporated IP active firms. The unique aspect of the database used here is that it contains full (population) data of IP active firms. There are 2,726 IP active firms incorporated in 2001 (these firms had one or more trade marks or patent publications over the period 2001 to 2005). Since financial data on R&D are not available for SMEs or micro firms, using IP data is, perhaps, the only proxy for innovation with population coverage (newly incorporated firms overwhelmingly belong to the SME or micro two firm-size groups).

Figure 7 shows failure rates by RDAs for IP active firms born in 2001. Note that, in all regions, the failure rates for IP active firms are lower than for IP inactive firms (see Table for exact numbers). For example, for London, the IP active firms failure rate is 11.3 percent compared to 35.7 percent for IP inactive firms. However, comparing Figure 5 with Figure 7, there are some interesting differences in relative survival rates across RDAs. The Scottish Highlands have a (relatively) high failure rate for IP active firms, as does the rest of Scotland (called "Scottish Enterprise" RDA). The East Midlands also has a higher failure rate for IP active firms. The East of England and Wales seem to have lower failure rates for IP active firms.

Figure 8 shows failure rates by counties and unitary authorities, again by quintile. This figure shows a more interesting pattern. London no longer has a high failure rate, but some surrounding counties do (Kent, Surrey and Hertfordshire). The Highlands of Scotland also shows a high failure rate, but this result is entirely driven by the fact that two out of the only four IP active firms in this region failed by 2005. In addition there are various counties and unitary authorities around Britain that exhibit high failure rates. Conversely, some areas such as the M4 corridor have lower failure rates for IP active firms than IP inactive firms.

The differences shown in the previous figures are interesting as background to innovation policy but we are, ultimately, interested in why differences occur. One major aspect may be that failure rates differ across industries and that regions differ according to their mix of industries. This is something we analyse in more detail in the next

section. However, before this we plot failure rates for firms in high tech industries.<sup>15</sup> Not all new firms in high tech need be innovative but it is more likely than on average.

Figures 9 and 10 plot the failure rate of 2001 incorporated high tech firms by RDAs (there are in total 484 of such firms). Figure 9 shows IP inactive high tech firms. The highest failure rates are in London and North West, followed by East Midlands and East of England. Figure 10 shows high tech firms that are also IP active. These are likely to be the most innovative firms and we might be especially interested in their outcomes. Figure 10 shows that their failure rates do differ from those in Figure 9, although there are few IP active high tech firms on which to base this comparison.

Figures 4 to 10 have given an overview of the pattern of failure rates in Britain. It is clear that there are substantial differences across regions, counties and unitary authorities. This is the case whether we look at IP inactive or active firms, or whether we look at the high tech sector. It is also clear that the differences between two regions depend on the sub-sample of firms being studied. Table 4 summarises much of the information for RDAs.

## 4.2 Firm-level summary statistics

Table 2 summarizes the basic characteristics of firms at the beginning of the period and those that survived by 2005. Due to the enormous variance in the data with regard to financial information on firms, we focus on median values. Firms that survived until 2005, had quadrupled median assets and doubled median turnover. This suggests either strong growth of surviving firms or a drop-out of smaller firms. We also included the number of directors of a firm. This variable proxies the managerial skill pool available to a firm. Somewhat surprisingly, we observe that the median number of directors falls from 4 to 3 for surviving firms. For the binary variables, we look at mean values to obtain the percentage share within the sample. As such, the number of firms being a subsidiary of a holding company increased by 2.4 percentage points between 2001 and 2005 indicating a positive relationship between survival and being part of a holding company. Foreign ownership increased only slightly.<sup>16</sup> The university variable indicates whether a firm has a link with a local British university.<sup>17</sup> This variable captures possible effects arising from cooperation agreements with research institutes, as for example knowledge spillovers, but also the availability of skilled human resources. Overall, the number of firms linked to universities is very small for the newly

---

<sup>15</sup>We follow the OECD definition: high tech firms as those in UK SIC 244, 353, 33, 32, and 30.

<sup>16</sup>Subsidiary and foreign ownership data are only available for the last set of financial accounts (e.g. 2005 for survivors), hence these figures do not capture any changes in a firm's status through time.

<sup>17</sup>This is derived from searching all of the firm's address for the word university, hence any university business park is likely to be picked up by this method.

incorporated firms. Yet, Table 2 shows that the share of firms attached to universities slightly increases of the observed time period. Finally, the percentage shares of firms that patent or take out a trade mark is tiny for the entire population of firms. This, of course, masks enormous differences across sectors. Yet, it becomes evident from Table 2, that IP activity increases substantially over time if a firm survives.

Table 2: Firm characteristics of the 2001 cohort (2001 vs. 2005)

	Year 2001	Year 2005
Variable	Median	
Total assets (1,000 £)	7	28
Turnover (1,000 £)	55	111
No. of directors	3	4
Variable	Mean	
Patent	0.005%	0.2%
Trade Mark	0.02%	0.5%
Patent / Trade Mark	0.02%	0.7%
Subsidiary	6.6%	9.0%
Foreign owned	2.4%	2.8%
University	0.08%	0.10%

To gain further insight about the IP active firms, Table 3 presents a summary of their characteristics and IP activity. The number of IP active firms more than doubles within the five-year period. Yet, the share of IP active firms within the 2001 cohort remains tiny at around 0.7 per cent in 2005. UK trade marks are the most common form of IP used by the firms. The relatively high number of firms with EPO patents is surprising.<sup>18</sup> The increase in the number of firms obtaining Community trade marks outpaced the increase firms taking out UK trade marks. This might point towards stronger international business orientation of surviving firms.

## 5 Survival Analysis - Empirical Strategy

The previous sections have shown that about 30 percent of all newly incorporated firms failed within the five year period 2001-2005. We have also pointed to substantial regional differences in failure rates and provided some descriptive evidence hinting at the importance of IP in shaping failure rates. In this section, we analyse firm failure rates using survival analysis.

<sup>18</sup>Obtaining an EPO is much more expensive than a UK patent and offers protection throughout the EU, hence smaller firms might be thought to use EPO patents infrequently.

Table 3: Number of IP active firms / Number of patents and trade marks 2001-2005

Year	No. of firms				
	IP active	UK TM	Community TM	UK patent	EPO patent
2001	372	262	57	50	36
2002	940	744	112	77	58
2003	918	611	186	128	81
2004	759	451	187	108	103
2005	761	436	216	115	114
Total	3,750	2,504	758	478	286

Year	Average no. of TM		Average no. of Patents	
	UK TM	Community TM	UK patent	EPO patent
2001	1.47	1.69	1.62	1.78
2002	1.43	1.50	1.75	1.47
2003	1.47	1.35	1.37	1.49
2004	1.54	1.35	1.65	1.38
2005	1.53	1.52	1.74	1.58
Total	1.48	1.44	1.61	1.50

Survival analysis describes the time that elapses from a certain starting point, for example birth, until a specific event occurs, for example death. Hence, the dependent variable in survival regression analysis is time. Since time is obviously positive or zero, the data is usually not normally distributed. This could easily be dealt with using standard econometric techniques. However, besides the conceptual problem of assuming that time is normally distributed, in survival analysis, there is another problem. The data observed is usually right-censored as already pointed out above. This means that individuals are observed only during a certain interval and the event of interest does not occur for all individuals in the sample within that period. Standard statistical techniques are unable to deal with this kind of censoring. In contrast, survival techniques deal well with this problem and address also the problem of non-negativity.

The objective of our survival analysis is to estimate the hazard rate of a firm to exit the market, denoted by  $h(t)$ . The hazard rate is interpreted as the conditional probability density that failure occurs at time  $t$  conditional on the firm having survived up to that point in time.

The hazard function is derived as follows. Denote the time until failure as  $T \geq 0$ .  $T$  has a cumulative distribution function  $F(t) = \Pr(T \leq t)$  and a probability density function

$$f(t) = \lim_{\delta \rightarrow 0} \left[ \frac{\Pr(t \leq T < t + \delta t)}{\delta t} \right]$$

Then the hazard function  $h(t)$  is defined as

$$h(t) = \lim_{\delta \rightarrow 0} \left[ \frac{\Pr(t \leq T < t + \delta t | T \geq t)}{\delta t} \right] = \frac{f(t)}{1 - F(t)} \quad (1)$$

The hazard function represents the probability that failure occurs within a certain interval, conditional on having survived up to the beginning of that interval divided by the width of the interval. It is important to note that  $h(t)$  can vary between  $[0, \infty)$ . The cumulative hazard function  $H(t)$  is defined as

$$H(t) = \int_0^t h(t) \quad (2)$$

The cumulative hazard function measures the amount of risk accumulated up to a certain time, while the hazard rate measures the rate at which risk is accumulated. Instead of expressing the hazard of failure, one can alternatively estimate the survival function  $S(t)$ , which is given by the reverse cumulative distribution function of  $T$ ,

$$S(t) = 1 - F(t) = \Pr(T > t) \quad (3)$$

The survival function represents the probability of surviving beyond  $t$ . Hence, the function is equal to one at  $t = 0$  and falls monotonically to zero as  $t \rightarrow \infty$ . We will repeatedly switch in our analysis between the hazard and the survival function as both are directly linked.<sup>19</sup>

The hazard rate derived above is independent of covariates influencing a firm's likelihood of failure. To take account of covariates, survival analysis provides a range of possible tools. By far most popular technique is the semi-parametric Cox proportional hazard model, which we use in our analysis in what follows.

We first estimate simple survival and hazard functions using nonparametric tech-

---

<sup>19</sup>This direct link becomes evident from the following expression

$$H(t) = \int_0^t \frac{f(t)}{S(t)} = \int_0^t \frac{1}{S(t)} \frac{d}{dt} S(t) dt = -\ln S(t)$$

niques, only distinguishing between firms that are IP active and those that are IP inactive. In a second step, we move on to include covariates, using a semi-parametric model, namely the Cox proportional hazard model and use more refined IP indicators. It appears that it is more appropriate to estimate a proportional hazard model than a accelerated lifetime model (AFT). In a proportional hazard model covariates shift the hazard function, while in the AFT model, the hazard function is identical across firms, but firms move faster along it according to their covariates. For our purpose, shifts in the hazard function lend themselves more easily for interpretation.<sup>20</sup>

## 5.1 Non-parametric Approach

We start by estimating the survival function  $S(t)$  using the Kaplan-Meier (1958) estimator. The Kaplan-Meier estimator is a non-parametric estimator, i.e., it does not make any assumptions about the distribution of failure times or how covariates shift the hazard function. It is nevertheless possible to estimate different survival functions for subsets of the data. We will use this to estimate separate survival functions for IP active and IP inactive firms and for all British regions and test whether they are statistically different.

The Kaplan-Meier estimator is given by

$$\hat{S}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{n_i}\right) \quad (4)$$

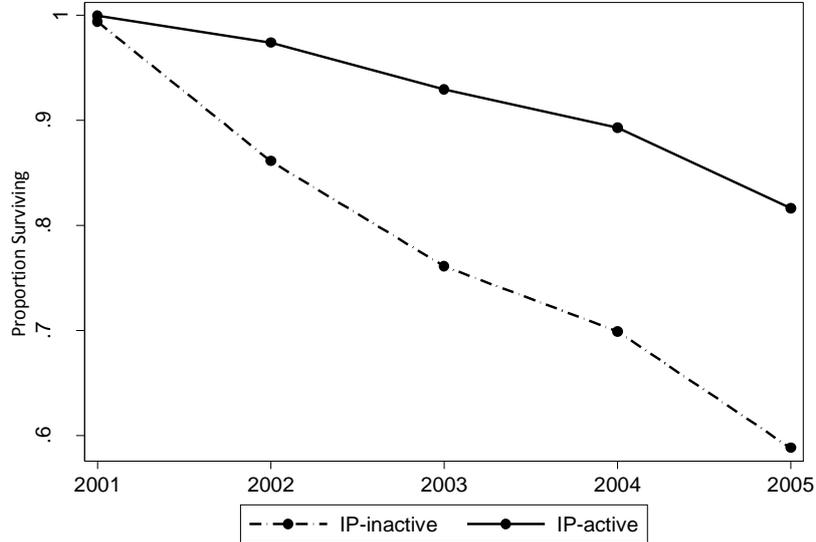
where  $n_i$  denotes the number of firms in the risk set at  $t_i$  and  $d_i$  the number of failures at  $t_i$ . The product is over all observed failure times less than or equal to  $t$ .

Since the Kaplan-Meier estimator estimates the hazard or survival function for each period of risk, we first group firms into year intervals for ease of interpretation. However, since we are grouping continuous data into discrete intervals, we use the so called Lifetable estimator to adjust for grouping. The Lifetable estimator produces an estimate centered on the midpoint of the interval in order to account of firms leaving at different times within the year-interval (Jenkins, 2005). As one major objective of this paper is to estimate the effect of firm's innovative activity on its survival probability, we group the data set into IP active and IP inactive firms and estimate the corresponding survival functions for each group. Since IP activity is a proxy for innovation this can be viewed as a test of survival of innovative firms versus non-innovative firms. We

---

<sup>20</sup>There are very few examples in the literature employing AFT models. For example Cefis and Marsili (2005) use the AFT model but only because their main variable of interest, the innovation variable, violates the proportional hazards assumption.

Figure 2: Survival rates for IP-active and IP-inactive firms



summarize firms' overall IP activity with a single dummy variable taking the value of one if a firm has obtained any form of IP.

Table 4 shows the results for the Lifetable estimates for both groups. To test more formally, whether there is a difference across the two groups, we use the log rank test. The null hypothesis of the log rank test is that there is no difference between groups.<sup>21</sup> It clearly rejects the null hypothesis at the 1 per cent level for the survival function of IP active and IP inactive firms to be equal.<sup>22</sup> In addition, Figure 2 plots the survival functions for IP-active and IP-inactive groups. For both groups of firms, however, the figure shows that risk of failure increases rapidly during the first two years of existence, which matches our observation from section 4 that most firms drop out between their second and third year of life. Subsequently, the survival rate risk slightly flattens to become steeper again during the fourth year after incorporation. Table 4 and Figure 2 provide strong indicative evidence that IP may have an important impact on firms' chances to survive.

Figure 3 plots the survival rates by region. It shows that there exist differences

<sup>21</sup>One common criticism of the use of the log rank test is that it gives too much weight to later event times as the number of observations in the risk sets become relatively small. This is not the case in our sample as 70 percent of the population survive the five-year interval studied which gives more than 114,000 observations in 2005. We also expect the test to be appropriate as it is best suited as a test for differences between groups when the hazards of the groups are proportional to each other, which can be seen in Figure 2 to approximately hold.

<sup>22</sup>Note that the log rank test yields the same result using continuous rather than grouped data.

Table 4: Lifetable estimates for IP active and IP inactive firms

Year	Total no. of firms	No. of failures	S(t)	Std. error
IP inactive				
2001	159,743	975	0.9939	0.0002
2002	158,768	21,208	0.8611	0.0009
2003	137,560	15,972	0.7611	0.0011
2004	121,587	9,895	0.6992	0.0011
2005	111,688	9,644	0.5881	0.0014
IP active				
2001	2,726	1	0.996	0.0004
2002	2,725	70	0.9740	0.0031
2003	2,655	121	0.9296	0.0049
2004	2,534	100	0.8929	0.0059
2005	2,432	109	0.8163	0.0089
Log-rank test for equality of survivor functions				
$\chi^2(1) = 499.02$ $\text{Pr} > \chi^2 = 0.0000$				

across regions. Most noticeably, London and the Highlands stand out. London exhibits are markedly lower survival rate than any other RDA region. Its survival rate for 2005 is 0.5355, while the one for the Highlands is 0.7164. The survival rates for the other regions vary between 0.5895 for the North West and 0.6479 for the North East. We also perform a log rank test for differences in survival rates across RDAs. Also in this case, the rejects the null hypothesis at the 1 per cent level for the survival function across regions to be equal.

Using year intervals implies that the only thing that is known about entry and exit is that it occurred at some point within the year interval. This problem is known as interval censoring. To avoid this problem in the subsequent analysis, we construct a continuous time measure by using complete entry and exit dates for all the analysis that follows.

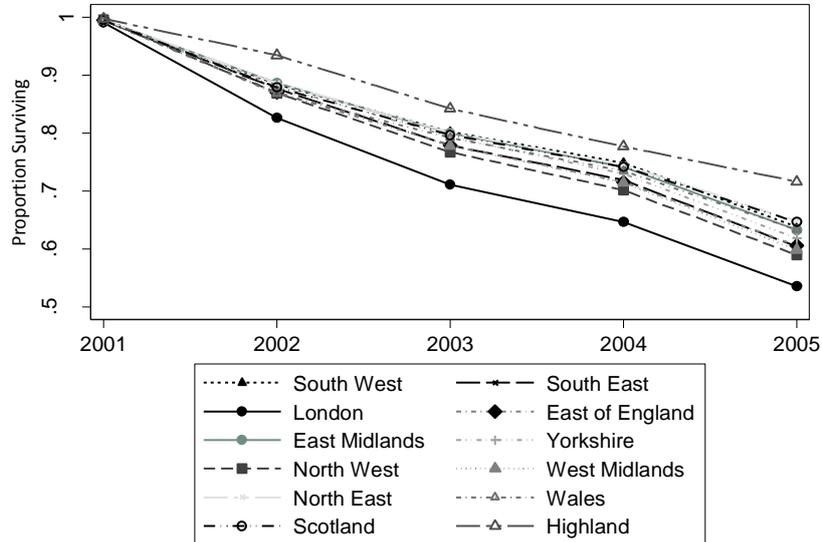
## 5.2 Proportional hazard Cox model

### 5.2.1 The model

The descriptive evidence of Section 4, and the non-parametric estimates of Section 5.1, point to substantial differences in firm survival across regions and between IP-active and inactive firms. An important question is what drives these differences?

The underlying process is driven by entrepreneurs that generate a constant stream

Figure 3: Survival rates across RDA regions



of new business ideas and set up firms to capitalize on these. These ideas are then tested in the market place and there is a large failure rate. We can think of failure stemming from one, or both, of two aspects:

- i) The underlying quality of the new idea (in other words its market demand)
- ii) The resources available to the entrepreneur to capitalize on the idea.

Resources could include finance, and the related capital, labour and materials; but they could also include knowledge about production methods or markets. There is also an important role for competitive pressure, by either incumbents or other new entrants, in affecting a firm's survival chances. This competitive pressure can be in various forms, such as substitute goods causing low demand for a rival, or competitor firms using up, or raising costs of, resources (e.g. skilled labor). In our empirical analysis we can think of IP as acting as a proxy for a 'better quality' idea hence, *ceteris paribus*, IP should increase the probability of survival. Finding variables that capture the resources available to entrepreneurs is difficult, however we can use information on whether the firm is located near a university and also the number of directors. Regional differences may be important as availability of resources and market conditions may vary across regions. To identify the specific mechanisms at work, we have to control for a range of other variables which also drive market entry and exit. Such variables are proxies for

the competitive environment, such as concentration ratios and entry/exit costs, or for demand growth, such as industry growth rates.

In order to take account of such covariates influencing a firm's hazard function, we write the hazard function for firm  $i$  in a general way as

$$h_i(t) = g(t, \beta_0 + \beta_x x_i) \quad (5)$$

where  $x_i$  is a vector of covariates and  $\beta$  are the coefficients to be estimated.

The most popular way to estimate the hazard function is to parameterize it as follows

$$h_i(t) = h_0(t) \exp(\beta_0 + \beta_x x_i) \quad (6)$$

which is called the proportional hazard model. It consists of two components,  $h_0$  is the baseline hazard which depends only on  $t$  and not on  $x_i$  and the covariates  $\exp(\beta_0 + \beta_x x_i)$ . The hazard of firm  $i$  with covariates  $x_i$  is, therefore, multiplicatively proportional to the baseline hazard  $h_0$ . Hence, the covariates  $x_i$  shift the hazard by a constant proportion relative to the baseline hazard. This also implies that the ratio of firm  $i$ 's hazard is proportional to any firm  $j$ 's hazard in the sample

$$\frac{h(t|x_i)}{h(t|x_j)} = \frac{\exp(\beta_x x_i)}{\exp(\beta_x x_j)} = \exp[\beta_x(x_i - x_j)] \quad (7)$$

where  $\exp[\beta_x(x_i - x_j)]$  is constant if the covariates do not vary over time. It is crucial to test whether the proportional hazard assumption holds in the specified model.

The functional form of the Cox model (Cox, 1972) is written in the proportional hazard formulation shown above. However, in the Cox model, no specific functional form for the baseline hazard  $h_0$  is assumed. More specifically, due to the way the hazard is computed,  $h_0$  drops out of the calculations. Therefore, the intercept is non-identifiable and the relationship between the hazard rate and the covariates is estimated without taking account of  $h_0$ . This implies that only the relative hazard between firms, not the absolute hazard rate, can be estimated. The relative hazard at time  $t$  between firms  $i$  and  $j$  is defined by (8). The coefficients  $\beta$  are estimated using the Partial Likelihood (PL) method instead of the Maximum Likelihood (ML) method. The PL considers only the ordered failure times while the ML estimator focuses on firm spells. Hence,

the Cox model is notably ignoring all information available at times when no failure occurs. This is done because the Cox model assumes that spells in which no failure occurs contain little information on the incidence of failure. Cox (1972) has shown that ignoring those spells results only in little efficiency loss.

### 5.2.2 Results for IP variables

We start with a parsimonious specification, only including separate dummy variables for a firm having a patent - UK and/or EPO - or a trade mark - UK and/or Community. This is a rather crude measure, but it provides insight before we use counts of patent/trade mark in the second specification. The use of counts lets us distinguish between UK and EPO patents and UK and Community trade marks.

The results in the first column of Table 5 yield the results consistent with the non-parametric estimations. The Cox estimates show that both IP-active firms, either having patents or trade marks, experience a statistically significantly lower hazard rate of failure than IP-inactive firms. Having a patent, therefore, reduces the hazard rate by 55 percent, and having a trade mark, lowers the hazard by slightly more than 52 percent.<sup>23</sup> These are substantial effects. The second column of Table 5 presents the results for patent and trade mark counts. These results allow to distinguish between the effect of the different forms of IP on the hazard of failure. UK and EPO patents have very similar effects, lowering the hazard rate by about 40 percent for each patent. However, there is a marked difference in the effect of UK and Community trade marks. While each UK trade mark lowers the hazard rate by about 40 percent, Community trade marks have an effect of only little more than 24 percent. Note, however, that most IP-active firms in the sample only have one patent or one trade mark (see Table 3). In summary, these initial results not controlling for covariates, yield a strong effect of IP on lowering the hazard rate of failure.

### 5.2.3 Regional effects and results for other variables

In our third specification, we add the RDA dummy variables to the specification including trade mark and patent counts. We use the Highlands as the benchmark category, as we know from Section 5.1 that this was the RDA region with the highest survival rate. Hence, all other regions are measured according to how much more likely firms are to fail than in the RDA region with the lowest failure rate.

---

<sup>23</sup>Coefficients have to be exponentiated to obtain this interpretation of their magnitude, e.g.,  $1 - \exp(-0.795) = 0.55$ .

The results are reported in the third column of Table 5. There are substantial differences among the RDA coefficients. As would be expected given the evidence provided above, the coefficient for London is the largest in magnitude, pointing to an almost 86 percent increased failure rate relative to the Highlands. Another region with a substantially higher failure rate is the North West. Firms in the North West experience an upward shift in the hazard rate of 55 percent vis-à-vis firms in the Highlands. Also note that the magnitude and statistical significance of the patent and trade mark coefficients remain almost unaffected from the inclusion of RDA dummies.

In a fourth step, we control for firm-level and industry-level characteristics while still including either IP dummies, or IP counts, and RDA dummies. The literature reviewed in Section 2 provides some guidance with regard to other factors influencing firm survival. All industry-level variables have been calculated on the SIC 3-digit level using the entire OFLIP database (i.e., all firms in the economy). Therefore, the industry-level variables completely reflect the environment in which the new firms started up and operated. This is a major strength of the data set used, as it allows for a nearly complete reflection of the economic environment in which firms operate.

At the industry level, we include a proxy for market entry and exit costs as proposed by Bernard and Jensen (2007).<sup>24</sup> While other authors, as for example Mata and Portugal (1994), are limited to including the actual number of entering firms in an industry, the OFLIP dataset allows us to construct entry and exit rates at the industry level. Furthermore, we use sector-level capital intensity, computed as the ratio of the amount of firms' assets and labor within each industry. To measure how important size is within each sector, we computed the minimum efficiency scale (MES) as the ratio of average first-year firm size to average firm size within the industry as proposed by Baldwin and Gellatly (2003). It measures the distance between the average entrant and the average firm size within each industry. This provides a better measure of firm size in the industry than the traditional MES measure as used by, for example, Fritsch et al. (2005). To account for competition within sectors, we include the four-firm concentration ratio, which measures the share of the four largest firms within a sector. We also explored alternative measures of the degree of competition, such as the Herfindahl index (Mata and Portugal, 1994) or the price-cost-margin (Aghion et al. 2005); all measures yield very similar results. Finally, we also use the industry growth rate measured by growth in industry-level assets. We use assets because data coverage is substantially higher than for turnover or employment. This is also a variable often used in survival analysis (see for example Audretsch and Mahmood, 1995).

<sup>24</sup>The proxy is constructed as follows  $\text{Entry/Exit costs} = 1 - \min[\text{entryrate}; \text{exitrate}]$ .

At the firm level, we use mainly dummy variables. The choice of variables is mainly motivated by data availability, as small firms do hardly report any financial data (and therefore would be dropped from the analysis by including financial variables). The first dummy indicates whether a firm was foreign owned. The variable is the same as the one used by Audretsch and Mahmood (1995) and simply captures foreign ownership as a determinant of survival. The second binary variable indicates whether the firm is a subsidiary of a holding company. We choose to include a dummy variable rather than estimating separate hazard functions for independent firms and those belonging to a group of firms (as in Disney, Haskel and Heden, 2003). The third dummy indicates whether a firm is located at a university address. Apart from these dummies, we also include the number of directors of a firm (in log) as a proxy for managerial skills available to the firm.

We then tested whether the proportional hazard assumption holds for the different specifications presented in Table 5. Performing a range of tests, including graphical tests plotting the log Kaplan-Meier estimator,  $\log(-\log \hat{S}(t))$ , against  $\log t$  and tests based on the residuals, we find the model specification to be appropriate.

The results of the different specifications, which have been estimated with adjusted standard errors to account for within group correlation, are shown in Table 5.<sup>25</sup>

Interpreting the industry and firm-level covariates for the two specifications without industry dummies, we find that the coefficients for the industry and firm level variables are very similar for both specifications. The coefficient of the proxy for entry and exit costs is in both specifications, using IP dummies or counts, negative, statistically highly significant and very large in magnitude. Bernard and Jensen (2007) also found higher entry/exit costs to lower the failure rate of firms. This is consistent with the idea that high entry/ext costs restrict competition. The coefficients of the level of sector capital intensity are positive, which means that increases in sector capital intensity shift the failure hazard up. This can be explained by the fact that new firms may be credit-constrained and therefore find it more difficult to survive in more capital intensive industries. The coefficient of the MES variable is negative and statistically significant. The higher the MES variable, the closer the average first-year firm is to the MES and the less it should be at a size disadvantage. The four-firm concentration ratio provides an interesting result. According to the negative and statistically significant coefficient, industries with higher concentration warrant a lower hazard of failure. We interpret this result as the outcome of less competition and possibly higher prices (through for example collusion) in more concentrated industries. As in Audretsch (1991), industry

---

<sup>25</sup>Note that the number of firms falls by 28,674 once we include sector-level variables as the corresponding information is missing in FAME for these firms.

growth does not have any statistically significant effect on the hazard rate. This is most likely explained by the fact that industry growth is mostly driven by few very large firms in the sector. Hence, the effect on new-born firms should be expected to be negligible.

At a firm-level, firms linked to universities have a lower hazard rate than other firms. The effect is sizeable, with a downward shift of the hazard function by around 55 percent. This variable seems to capture effects stemming from collaboration with universities, possibly giving above all small start-ups essential assistance during their early years of existence. Similar to Audretsch and Mahmood (1995) and Bernard and Jensen (2007), we find that subsidiaries are less likely to exit the market. The finding that the hazard for subsidiaries is lower than for stand-alone firms, possibly reflect the fact that for subsidiaries financial constraints may be less binding. Similarly, firms owned by foreign firms are equally less likely to exit. Finally, increasing the log of the number of directors lowers the hazard of failure by about 28 percent. More directors could indicate more skill or management resources for the firm, or possibly that the quality of the business idea is higher.

When we control also for industry-specific effects, nearly all of the coefficients retain their signs with the exception of the coefficient of the variable measuring market concentration, which becomes positive, but is no longer statistically significant. In addition, now the coefficient of capital intensity is statistically significant, implying that firms in industries with higher capital intensities experience an upward shift in their hazard rate.

It is important to note that, despite adding the above variables, the coefficients on the IP dummies remain statistically significant, although coefficient magnitude falls. Yet, once we include the industry and firm level variables, only the coefficients of UK IP remains statistically significant. Both EPO patents and Community trade marks have no longer a statistically significant effect on the hazard of failure. The coefficients on the RDA dummies remain statistically significant. Hence, despite controlling for a large range of industry and firm-specific characteristics, IP variables and regional location exert an important effect on firms' hazards of failure.

## 6 Conclusion

This paper uses a new database to analyse the survival of the complete cohort of all British companies registered in 2001. The database is able to track the outcome of these 162,000 firms to 2005 and also, uniquely, their intellectual property (IP) activity during 2001-2005. The paper focuses on two important policy-related issues. The first is the extent to which IP activity alters the survival outcomes of these start-up firms.

The second is whether, and to what extent, survival of start-up firms varies across regions.

Intellectual property activity is captured by four measures: UK patent publications, EPO patent publications, UK trade mark publications and Community trade mark registrations. Over the five year period, 3,750 (2.3 per cent) of the 2001 start-up firms use one or more of these forms of IP. The most common form of IP used is a UK trade mark, followed by Community trade marks and then UK patents. The dominance of trade marks reflects two factors. First, obtaining a trade mark is cheaper and easier than patent protection. Second, trade marks are used by all sectors in the economy, hence they capture IP activity by service sector firms. The use of trade marks in empirical research is novel, hence we should ask what trade mark activity means? Our expectation is that a trade mark proxies the launch of a new, or upgraded, product; hence, we expect a trade mark to indicate innovation. The use of patent data in empirical research is common, although there are few studies on small, start-up firms.

The Regional Development Agencies (RDA) Act 1998 led to the establishment of 12 RDAs in Britain. The RDAs coordinate substantial assistance to firms with the core aim of encouraging enterprise, employment and competitiveness. Given the existence of these RDAs, it is interesting to ask if and how firm survival varies across them. The answer is that there are large differences. The failure rate of the 2001 firms by 2005 varies between 36 percent in London to 22 percent in the Highlands (Table 1). Looking at firms in high-tech sectors we also find wide variations in failure rates across RDAs. Table 1 also indicates that the failure rates of IP-active firms varies across RDAs, although in this case the highest rate is 15 percent in the North West (London is now 11 percent). This also indicates that the failure rate for IP active firms are substantially lower than IP-inactive firms. This is a result that runs through all of our analysis.

Why might failure rates vary so much across regions? A short answer is that competitive conditions vary. This, in turn, is due to differing industrial structures and economic conditions within each region. Section 5 uses a Cox proportional hazards model to investigate these issues and, at the same time, it includes firm-level IP and other characteristics that might influence survival. The results indicate, as expected, that competitive conditions are important. Variables for exit/entry costs, minimum efficiency scale and concentration (all defined at 3-digit SIC level) are all significant. IP activity is now disaggregated into its four components and we find that only national patents and trade marks are significant in improving survival, with UK trade marks having a similar impact to UK patents. The results also show that being located near a university reduces failure rates, as does being foreign owned or part of a larger group of firms. We also find that firms with more directors have higher survival rates.

The Cox model also includes a set of dummies for the RDAs. If the industry-level and firm-level variables had completely explained the regional differences we would have expected the coefficients on these RDA dummies to be insignificant. This is not the case: there are still significant differences in survival rates across RDAs despite controlling for a range of factors. These differences are due to unobserved factors, some of these could relate to very specific industry factors, but others are likely to relate to the availability of resources or support to start-up firms.

## References

- [1] Agarwal Rajshree and Audretsch David (2001): *Does entry size matter? The impact of the life cycle and technology on firm survival*, Journal of Industrial Economics, Vol. 49, No. 1, pp. 21-43.
- [2] Aghion Philippe, Bloom Nick, Blundell Richard, Griffith Rachel, and Howitt Peter (2005): *Competition and Innovation: An Inverted-U Relationship*, Quarterly Journal of Economics, Vol. 120, No. 2, pp. 701-728.
- [3] Audretsch David (1991): *New-Firm Survival and the Technological Regime*, Review of Economics and Statistics, Vol. 60(3), pp. 441-450.
- [4] Audretsch David and Mahmood Talat (1995): *New Firm Survival: New Results Using a Hazard Function*, The Review of Economics and Statistics, Vol. 77, No. 1., pp. 97-103.
- [5] Baldwin John and Gellatly Guy (2003): *Innovation Strategies and Performance in Small Firms*, Edward Elgar.
- [6] Bernard Andrew and Jensen Bradford (2007): *Firm structure, multinationals, and manufacturing plant deaths*, Review of Economics and Statistics, Vol. 89, No. 2, pp. 193-204.
- [7] Bridges Sarah and Guariglia Alessandra (2006): *Financial Constraints, Global Engagement, and Firm Survival in the UK: Evidence from Micro Data*, Working Paper Nottingham University.
- [8] Cefis Elena and Marsili Orietta (2005): *A matter of life and death: innovation and firm survival*, *Industrial and Corporate Change*, Vol. 14, No. 6, pp. 1167-1192.
- [9] Cleves Mario, Gould William, and Gutierrez Roberto (2004): *An introduction to Survival Analysis Using Stata*, STATA Press 2nd edition.
- [10] Cockburn Iain and Wagner Stefan (2007): *Patents and the Survival of Internet-related IPOs*, NBER Working Papers 13146.
- [11] Cox David (1972): *Regression Models and Lifetables*, Journal of the Royal Statistical Society, Vol. 34, 187220.
- [12] Disney, R., Haskel, J. and Heden, Y. (2003). *Entry, exit and establishment survival in UK manufacturing*, Journal of Industrial Economics, 51(1), pp. 91-112.

- [13] Dunne Timothy, Roberts Mark and Samuelson Larry (1989): *The Growth and Failure of U.S. Manufacturing Plants*, The Quarterly Journal of Economics, Vol. 104(4), pp. 671-98.
- [14] Fritsch Michael, Brixey Udo and Falck Oliver (2006): *The Effect of Industry, Region, and Time on New Business Survival - A Multi-Dimensional Analysis*, Review of Industrial Organization, Vol. 28, No. 3, pp. 285-306.
- [15] Geroski Paul (1991): *Domestic and foreign entry in the United Kingdom: 1983-1984*, in Geroski P. and Schwalbach J. (Eds.) *Entry and Market Contestability*, Basil Blackwell, Oxford.
- [16] Greenhalgh Christine and Rogers Mark (2006): *The value of innovation: The interaction of competition, R&D and IP*, Research Policy, Vol. 35(4), pp. 562-580.
- [17] Greenhalgh Christine and Rogers Mark (2007): *Trade marks and performance in UK firms: Evidence of Schumpeterian competition through innovation*, Oxford University, Economics Department Working Paper: <http://www.economics.ox.ac.uk/Research/wp/pdf/paper300.pdf>.
- [18] Griliches Zvi (1981): *Market Value, R&D, and patents*, Economic Letters, Vol. 7, pp. 183-187.
- [19] Griliches Zvi (1990): *Patent Statistics as Economic Indicators: A Survey*, Journal of Economic Literature, Vol. 28(4), pp. 1661-1707.
- [20] Hall Bronwyn (2007): *Patents and Patent Policy*, Oxford Review of Economic Policy, Forthcoming.
- [21] Jenkins Stephen (2005): *Survival Analysis*, University of Essex.
- [22] Jovanovic Boyan and MacDonald Glenn (1994): *The Life Cycle of a Competitive Industry*, Journal of Political Economy, Vol. 102(2), pp. 322-47.
- [23] Kaplan E. L. and Meier Paul (1958): *Nonparametric estimation from incomplete observations*, Journal of the American Statistical Association, Vol. 53, pp. 457-48.
- [24] Klette Tor Jakob and Kortum Samuel (2004): *Innovating Firms and Aggregate Innovation*, Journal of Political Economy, Vol. 112, Issue 5, pp. 986-1018.
- [25] Landes William and Posner Richard (1987): *Trademark law: An economic perspective*, Journal of Law and Economics, Vol. 30(2), pp. 265-309.

- [26] Mahmood Talat (1992): *Does the Hazard Rate for New Plants Vary between Low- and High-Tech Industries?*, Small Business Economics, Vol. 4, pp. 201-209.
- [27] Mata Jose and Portugal Pedro (1994): *Life Duration of New Firms*, *The Journal of Industrial Economics*, Vol. 42, No. 3., pp. 227-245.
- [28] Mendoca, Pereira, and Godinho (2004): *Trademarks as an indicator of innovation and industrial change*, Research Policy, Vol. 33, pp. 1385-1404.
- [29] Pavitt, K. (1984). *Sectoral patterns of technical change*, Research Policy 13: 343-373.
- [30] Rogers Mark, Helmers Christian and Greenhalgh Christine (2007): *An analysis of the characteristics of small and medium enterprises that use intellectual property*, Oxford Intellectual Property Research Centre.
- [31] Schumpeter Joseph (1934): *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*, Harvard University Press.
- [32] Schmookler Jacob (1966): *Invention and economic growth*, Cambridge, Mass., Harvard University Press.
- [33] Suarez Fernando and Utterback James (1995): *Dominant Designs and the Survival of Firms*, Strategic Management Journal, Vol. 16, No. 6, pp. 415-430.
- [34] Sutton John (1991): *Sunk Costs and Market Structure*, MIT Press.
- [35] Sutton John (1998): *Technology and Market Structure*, MIT Press.
- [36] van Zeebroeck Nicolas (2007): *The puzzle of patent value indicators*, CEB Working Paper.

Table 5: Cox Regression

Variables	Coefficient					
Sector level						
Entry/Exit costs	..	..	..	-14.341*** (0.237)	-14.337*** (0.237)	-13.136*** (0.506)
Capital intensity	..	..	..	0.002 (0.003)	0.002 (0.003)	0.013*** (0.003)
MES	..	..	..	-3.927** (1.711)	-3.922** (1.710)	-9.409*** (2.622)
4-Firm Concentration Ratio	..	..	..	-0.074** (0.032)	-0.074** (0.032)	0.004 (0.043)
Industry Growth Rate	..	..	..	-0.007 (0.010)	-0.007 (0.010)	-0.011 (0.017)
Firm level						
Patent Dummy	-0.795*** (0.182)	..	..	-0.439** (0.211)	..	..
UK Patent Count	..	-0.515*** (0.087)	-0.515*** (0.182)	..	-0.508** (0.229)	-0.541** (0.233)
EPO Patent Count	..	-0.543** (0.221)	-0.534** (0.220)	..	-0.067 (0.217)	-0.094 (0.223)
Trade Mark Dummy	-0.741*** (0.083)	..	..	-0.648*** (0.100)	..	..
UK TM Count	..	-0.515*** (0.087)	-0.522*** (0.182)	..	-0.518*** (0.096)	-0.531*** (0.096)
CTM Count	..	-0.276** (0.119)	-0.302** (0.122)	..	-0.209 (0.151)	-0.225 (0.152)
University	..	..	..	-0.817*** (0.294)	-0.821*** (0.295)	-0.827*** (0.295)
Ln no. of Directors	..	..	..	-0.334*** (0.018)	-0.334*** (0.018)	-0.295*** (0.018)
Foreign Owned	..	..	..	-0.817*** (0.048)	-0.817*** (0.048)	-0.827*** (0.048)
Subsidiary	..	..	..	-2.042*** (0.056)	-2.042*** (0.056)	-2.031*** (0.055)
RDAs						
South West	..	..	0.245** (0.947)	0.252** (0.121)	0.252** (0.121)	0.289** (0.122)
South East	..	..	0.366*** (0.094)	0.313*** (0.120)	0.312*** (0.120)	0.353*** (0.121)
London	..	..	0.618*** (0.935)	0.692*** (0.119)	0.691*** (0.119)	0.733*** (0.121)
East of England	..	..	0.371*** (0.095)	0.403*** (0.121)	0.403*** (0.121)	0.431*** (0.122)
East Midlands	..	..	0.267*** (0.095)	0.298** (0.122)	0.297** (0.122)	0.320*** (0.123)
Yorkshire	..	..	0.318*** (0.095)	0.403*** (0.122)	0.403*** (0.122)	0.424*** (0.122)
North West	..	..	0.429*** (0.094)	0.491*** (0.121)	0.491*** (0.121)	0.522*** (0.122)
West Midlands	..	..	0.383*** (0.094)	0.431*** (0.121)	0.431*** (0.121)	0.465*** (0.121)
North East	..	..	0.234** (0.099)	0.297** (0.126)	0.297** (0.126)	0.319** (0.127)
Wales	..	..	0.288*** (0.097)	0.289** (0.124)	0.289** (0.125)	0.303** (0.125)
Scotland	..	..	0.249*** (0.095)	0.396*** (0.122)	0.396*** (0.122)	0.414*** (0.122)
Highland	..	..	dropped	dropped	dropped	dropped
Industry dummies	..	..	..	..	..	included
Number of firms:	162,455	162,455	162,455	133,781	133,781	133,781
Number of observations:	701,850	701,850	701,850	591,999	591,999	591,999
Number of failures:	58,076	58,076	58,076	35,798	35,798	35,798

Note: Adjusted Standard Errors for Within Correlation

Figure 4: No. of new firms by RDA region

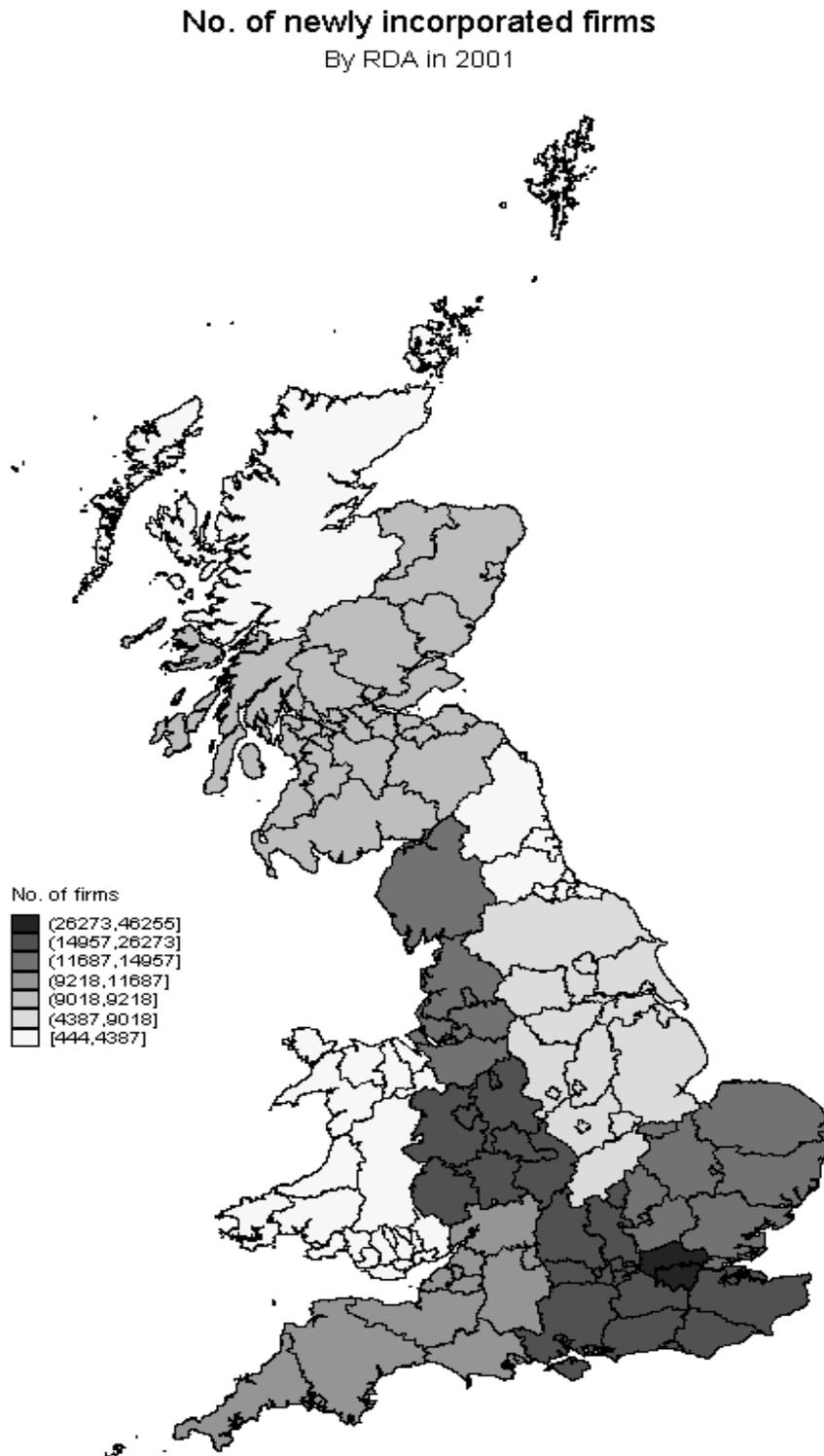


Figure 5: Failure rates of IP-inactive firms by RDA region

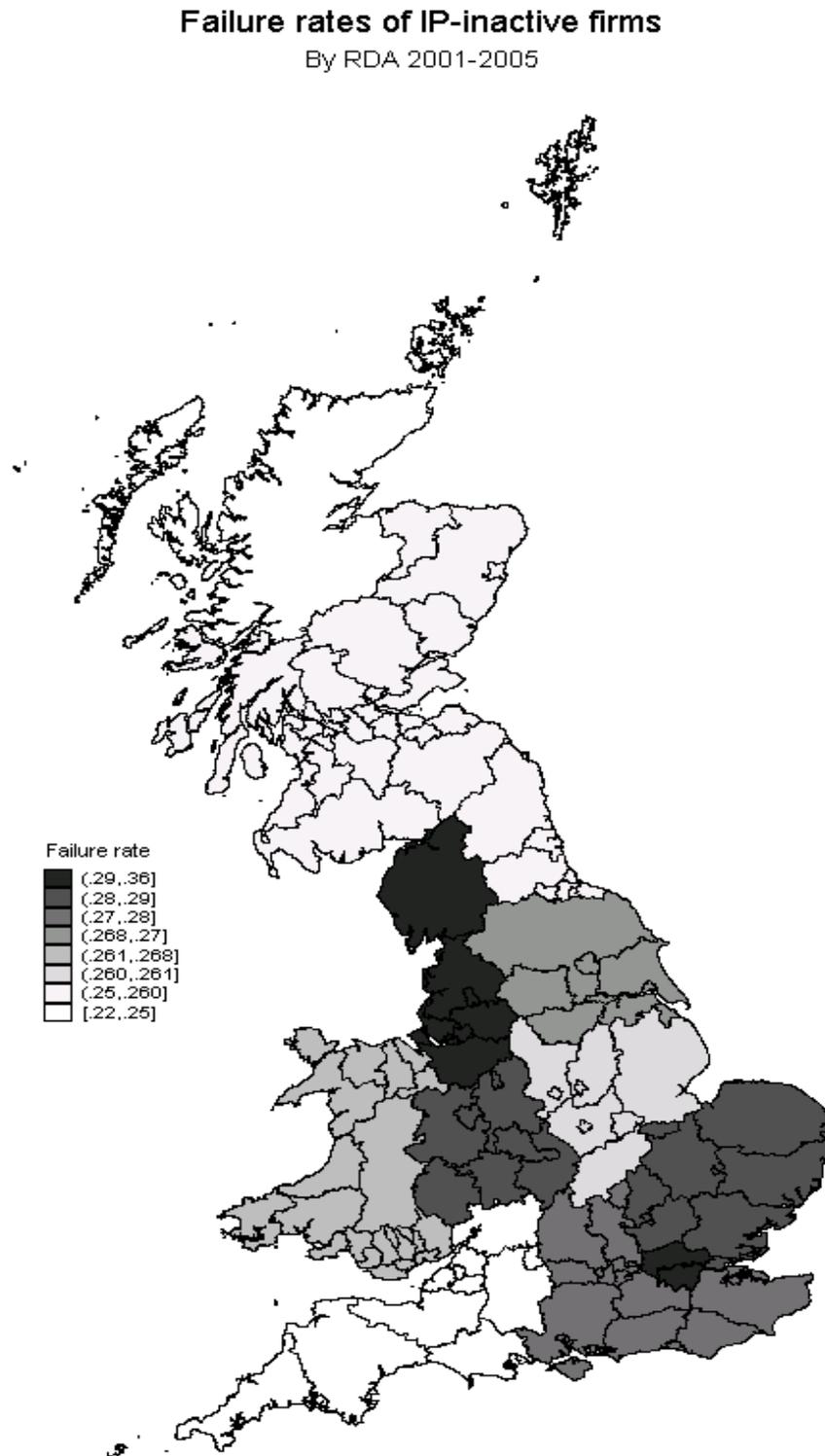


Figure 6: Failure rates of IP-inactive firms by county / unitary authority

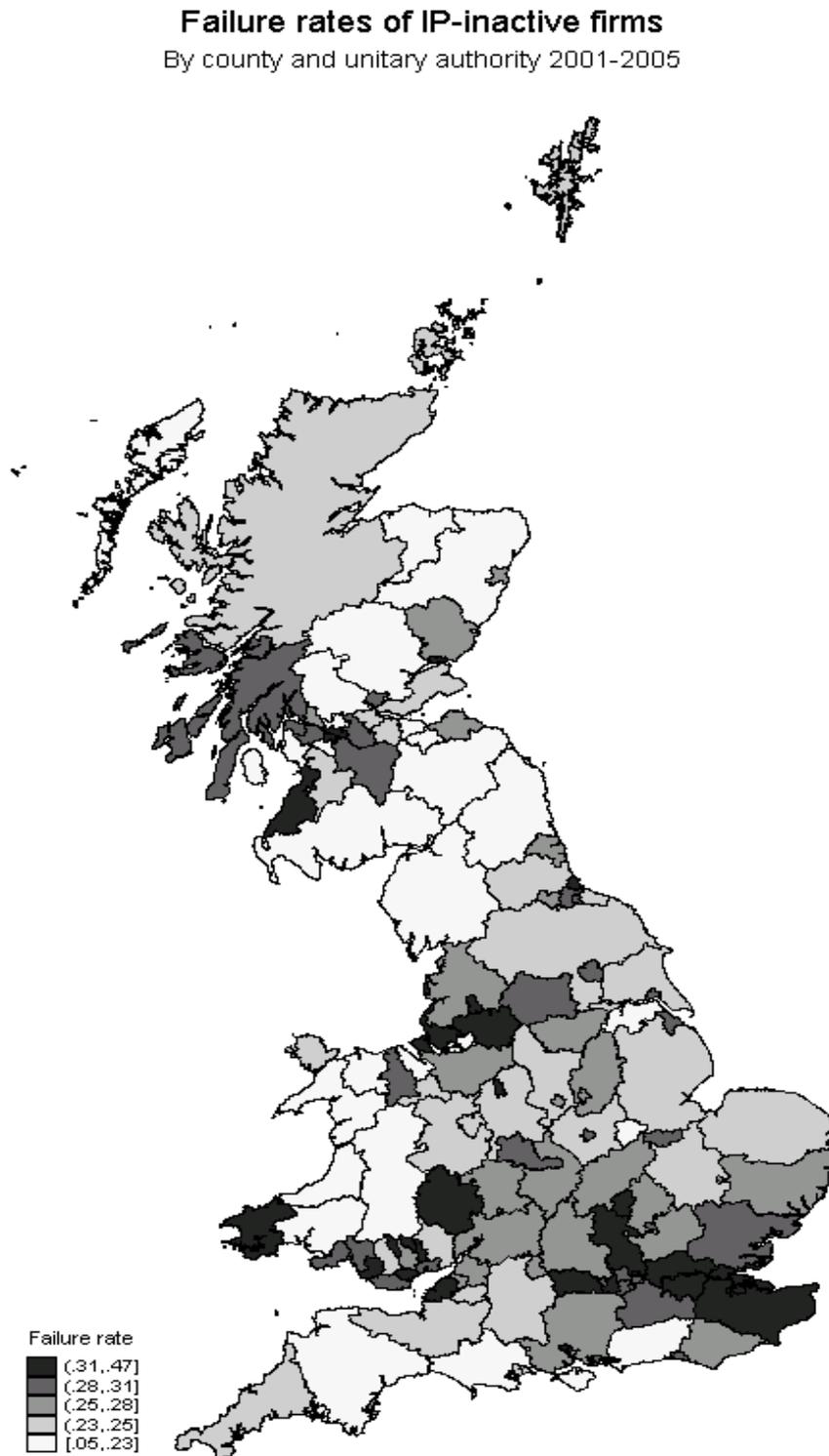


Figure 7: Failure rates of IP-active firms by RDA region

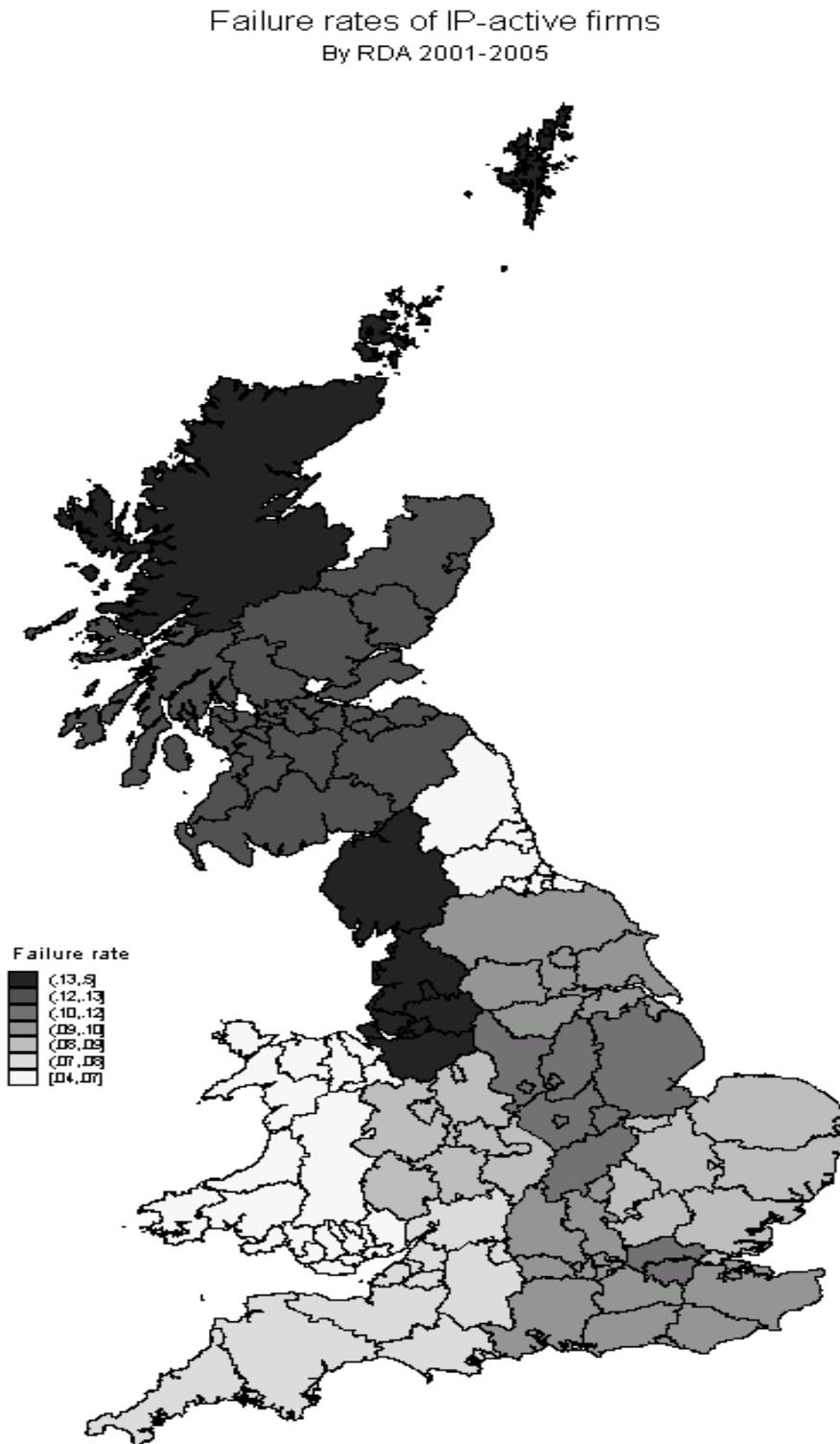


Figure 8: Failure rates of IP-active firms by county / unitary authority

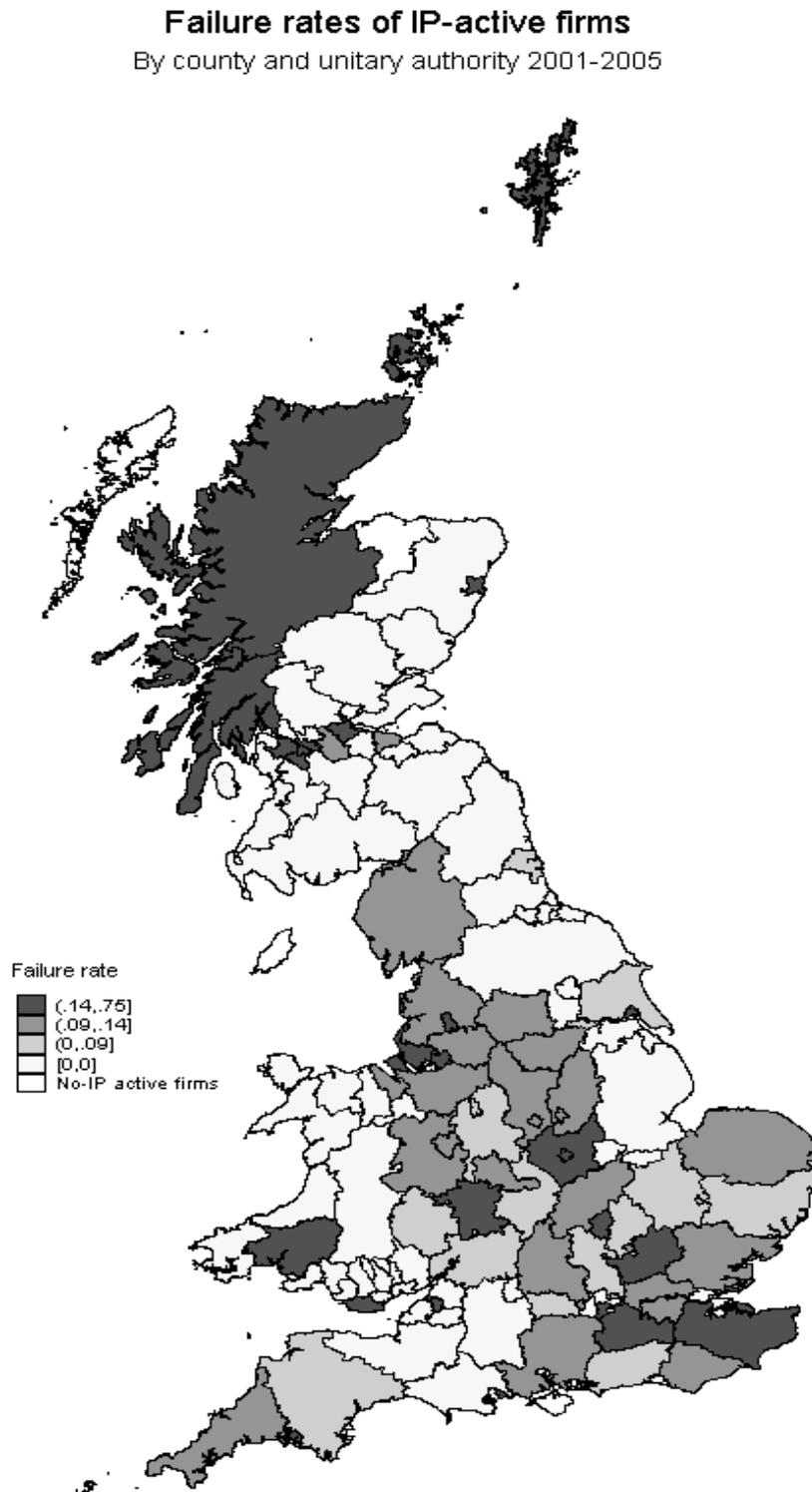


Figure 9: Failure rates of IP-inactive high-tech firms by county / unitary authority

**Failure rates of IP-inactive high-tech firms**  
By RDA 2001-2005

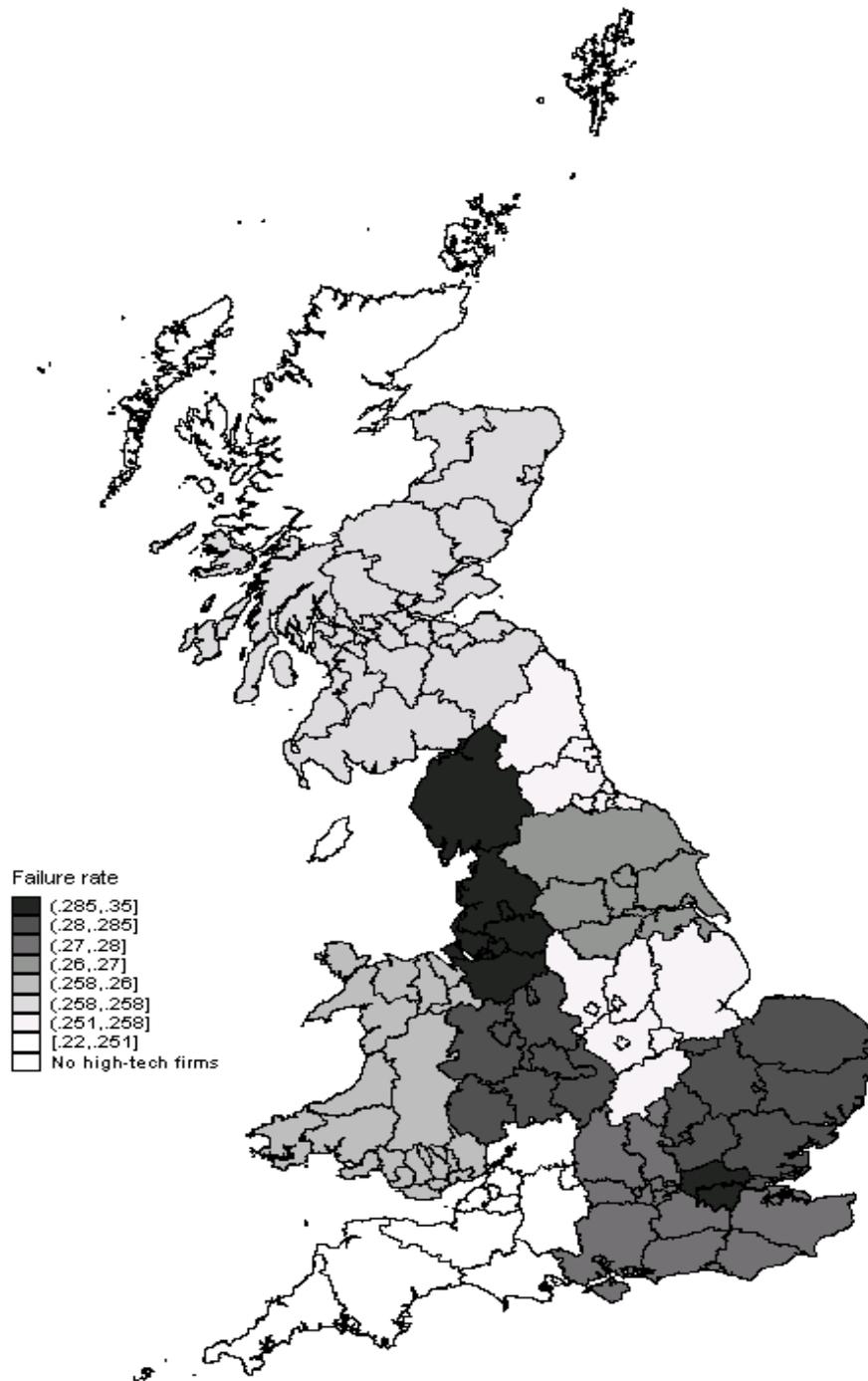


Figure 10: Failure rates of IP-active high-tech firms by county / unitary authority

