

Universidade de Aveiro
Departamento de Economia, Gestão e Engenharia Industrial

Documentos de Trabalho em Economia
Working Papers in Economics

Área Científica de Economia
E/nº 53/2010

Análise comparativa de sobrevivência empresarial: o caso da região Norte de Portugal

Elsa Sarmento
Alcina Nunes

Submission of Papers for Publication (Para submissão de artigos para publicação): Prof. Sílvia Jorge (sjorge@ua.pt). Universidade de Aveiro, DEGEI, Economia, Campus Universitário de Santiago. 3810-193 Aveiro. Portugal.

Análise comparativa de sobrevivência empresarial: o caso da região Norte de Portugal

Elsa Sarmento (esarmento@egi.ua.pt)

Universidade de Aveiro, Departamento de Economia, Gestão e Engenharia Industrial

Alcina Nunes (alcina@ipb.pt)

Instituto Politécnico de Bragança, Departamento de Economia e Gestão

Resumo: Este estudo aborda a capacidade de sobrevivência de empresas na região Norte de Portugal, de 1985 a 2007, fornecendo comparações regionais e nacionais, com recurso a funções de risco e sobrevivência e a métodos não paramétricos e semi-paramétricos. A análise assenta na criação de uma base de dados de empreendedorismo baseada na informação dos Quadros de Pessoal, ao qual foi aplicada a metodologia da OCDE e do Eurostat patente no “Manual Estatístico da Demografia das Empresas”. Os resultados apontam para um encerramento precoce da actividade empresarial no Norte, sendo a duração mediana da vida das empresas inferior à das restantes regiões. A elevada taxa de turbulência de entrada e saída de empresas no mercado é apontada como a variável com o efeito mais significativo nas probabilidades de sobrevivência.

Palavras-Chave: Empreendedorismo, Dinâmica Empresarial, Sobrevivência, Análise Regional.

Classificação JEL: C14, C41, L25, L26, R11

Title: Comparative Survival Analysis of Firms: the case of the Portuguese North region

Abstract: This study approaches the survival ability of firms in the North region of Portugal from 1985 to 2007, allowing regional and national comparisons through the application of non-parametric and semi-parametric methods. This analysis is based on the creation of a specific entrepreneurship database, based on *Quadros de Pessoal*, where the methodology of Eurostat/OECD’s “Manual of Business Demography Statistics” was applied to. In the North, firms close earlier on during their infancy as their median duration is below the median duration of firms located in the rest of the country. The high turbulence of firm turnover is pointed out as a fundamental determinant of the survival probabilities of firms in this region.

Key-Words: Entrepreneurship, Firm Dynamics, Survival, Regional Analysis.

JEL Classification: C14, C41, L25, L26, R11

1. Introdução

A análise de sobrevivência cuja origem remonta ao ramo da epidemiologia, foi adoptada pela análise de sobrevivência empresarial, complementando métodos de estudo anteriores sobre a demografia de empresas. Esta é uma área de estudo onde eram sentidas, até há pouco tempo, as consequências das dificuldades de acesso e de disponibilização de dados longitudinais, que fornecessem informação quantitativa sobre a diferenciação espacial das probabilidades de sobrevivência das empresas (Brixy e Grotz, 2006). Esta, associada à escassa divulgação de métodos econométricos específicos para a análise de sobrevivência e à falta de atenção dada à componente regional, não permitiu produzir os avanços e os resultados empíricos desejáveis.

Em Portugal, a disponibilização dos Quadros de Pessoal do Gabinete de Estratégia e Planeamento (GEP) do Ministério do Trabalho e da Segurança Social (MTSS), tem-se relevado fundamental, tendo contribuído para o fomento de um conjunto importante de investigação aplicada sobre a análise de empresas e sobrevivência (Mata e Portugal, 1994; Mata e Guimarães, 1995; Mata, 1993; Mata e Machado, 1996; Baptista *et al.*, 2008; Cabral, 2007; Cabral e Mata, 2003; Baptista e Carias, 2007; Baptista e Mendonça, 2007). A presente análise decorre de trabalhos anteriores sobre a demografia de empresas (Nunes e Sarmento, 2010a e 2010b; Sarmento e Nunes, 2010), onde se procurou adoptar uma perspectiva regional sobre os fenómenos de sobrevivência de empresas.

Este trabalho assenta na criação e exploração de uma base de dados específica de empreendedorismo, baseada nos Quadros de Pessoal. Decorrente da aplicação da metodologia baseada no *Manual of Business Demography Statistics* publicado conjuntamente pelo Eurostat e pela OCDE (2007), considerou-se um subconjunto da base de dados dos Quadros de Pessoal, onde estão integradas apenas empresas activas¹, que empreguem pelo menos um trabalhador remunerado. Este subconjunto de empresas constitui a grande fonte de criação de emprego (dependente) em Portugal. Complementarmente a outras metodologias e universos já existentes na análise da demografia empresarial, baseadas por exemplo no Sistema de Contas Integrado do Instituto Nacional de Estatística (2009a e 2009b; GEP 2010) e nas Estatísticas Estruturais da União Europeia (Eurostat, 2009), a metodologia Eurostat/OCDE permite obter indicadores comparáveis, não só a nível regional e nacional, como também a nível internacional (OCDE, 2008 e 2009).

¹ Consideraram-se apenas as empresas classificadas nas secções A a Q da CAE-Rev.2.1.

A definição de empresa (Eurostat/OCDE, 2007), corresponde à mais pequena combinação de unidades legais, isto é, a uma unidade organizacional que produz bens ou serviços, que auferir de uma certa autonomia de decisão, particularmente no que diz respeito à alocação dos seus recursos correntes. As empresas empregadoras (*employer enterprises*) são empresas existentes na população de empresas activas, que empregam pelo menos um trabalhador remunerado. Os encerramentos ou “mortes” de empresas ocorrem porque estas deixam de estar presentes na base de dados, durante pelo menos dois anos, ou porque deixaram de ter pelo menos um trabalhador remunerado, conforme registo nos Quadros de Pessoal, de acordo com a aplicação da metodologia Eurostat/OECD (2007). A base de dados obtida contém numa média anual de 215 903 empresas empregadoras activas no período de 1985 a 2007, com uma média anual de 36 803 nascimentos e 23 743 mortes de empresas.

Na sequência da aplicação da definição de encerramento patente na metodologia Eurostat/OECD, a utilização da análise de sobrevivência a nível regional mostra-se particularmente relevante para ajudar a compreender os movimentos do encerramento de empresas, em particular numa região que apresenta especificidades bastante distintas a nível da sua dinâmica empresarial. Segundo Nunes e Sarmiento (2010b), a região Norte do país apresenta uma probabilidade de encerramento com o maior ritmo de crescimento ao longo do período de actividade das empresas considerado (1985 a 2007). Simultaneamente, é aquela cuja duração mediana é mais baixa. O encerramento de actividade, ou por outras palavras, a “morte” das empresas ocorre, em termos medianos, entre os 5 e os 6 anos de idade, estando este valor compreendido entre os 6 e os 7 anos nas restantes regiões do país. Esta é ainda a região que, no final do período de observação, apresenta uma menor taxa de sobrevivência empresarial.

No desenrolar da análise, serão apresentados resultados empíricos que permitem caracterizar o movimento de encerramento de empresas na região Norte e analisar os determinantes da sua capacidade de sobrevivência, recorrendo a métodos não paramétricos (Kaplan-Meier) e semi-paramétricos (modelo complementar log-log e modelo de risco proporcional de Cox), que permitem parametrizar a caracterização empresarial e sectorial.

Este trabalho desenrola-se em cinco partes distintas. A próxima secção sintetiza os principais aspectos da dinâmica regional, a secção 3 expõe a análise de sobrevivência não paramétrica, a secção 4 desenvolve a análise semi-paramétrica e finalmente a última secção apresenta as principais conclusões.

2. Dinâmica empresarial

A região Norte apresenta a maior proporção de empresas activas no país e o maior número de empresas com menos de 20 trabalhadores². Em 2008, esta região concentrou cerca de 35,6% das empresas em território nacional, mantendo-se como detentora do maior número de empresas activas do país (GEE, 2010; GEP, 2010).

O Norte apresenta a segunda maior taxa média de crescimento anual do total de empresas activas entre 2000 e 2007 (4,4%) (Sarmiento e Nunes, 2010). É responsável pela maioria dos nascimentos de empresas entre 1985 e 2007³ (36% em média), apresentando uma taxa de natalidade superior à média nacional (excepto nos anos de 1991, 1992 e 2000). Mas esta é também a região onde ocorrem a maioria dos encerramentos (35,4%) (Figura 1). A volatilidade da dinâmica empresarial desta região⁴ (Figura 2), em particular entre 1993 e 1998, é a mais elevada do país.

Figura 1: Empresas activas encerradas com um ou mais trabalhadores, 1987-2005



Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

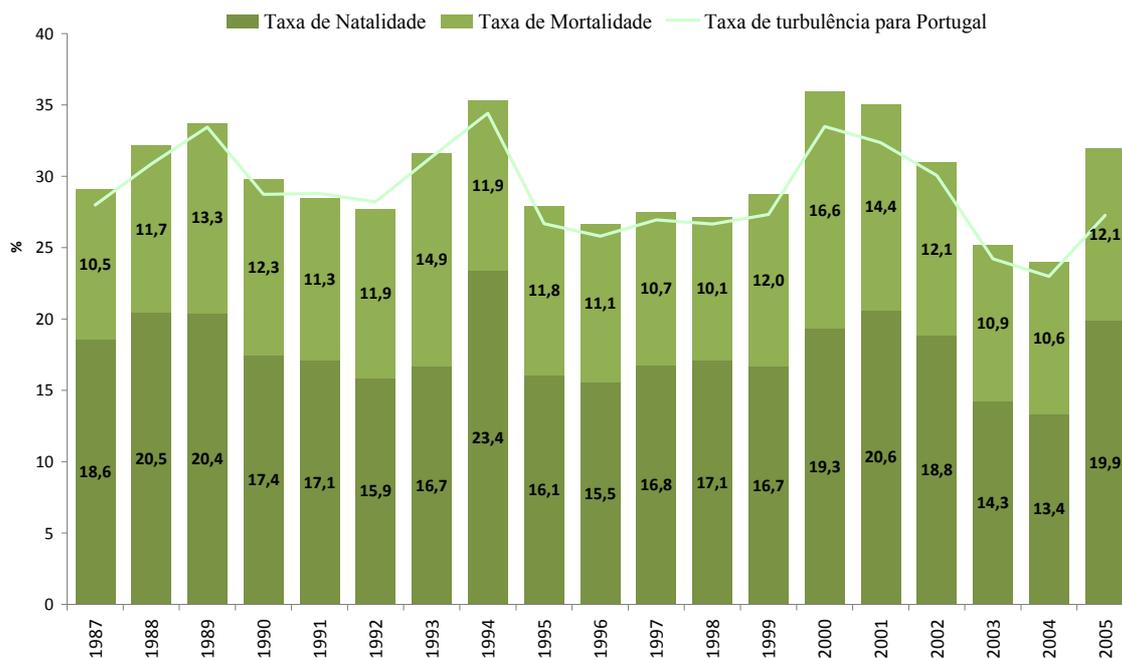
² No entanto, a proporção do número de empresas activas com menos de 20 trabalhadores é mais elevada no Algarve, onde se regista a maior proporção de pequenas empresas do país (Sarmiento e Nunes, 2010).

³ Tendo atingido, em 2005, o valor mais elevado, com uma proporção de 44,4% do total de nascimentos.

⁴ O Norte apresenta a maior volatilidade regional quando medida através do desvio-padrão.

Os “picos” de criação líquida de empresas no Norte, verificados de 1986 a 1989, em 1994⁵ e entre 2001 e 2002, coincidem com os períodos de maior turbulência (dada pela soma das taxas de natalidade⁶ e mortalidade⁷ de empresas), acompanhando de forma mais exacerbada a dinâmica empresarial registada a nível nacional (Figura 2).

Figura 2: Taxa de natalidade, de mortalidade e de turbulência na região Norte, 1987-2005



Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

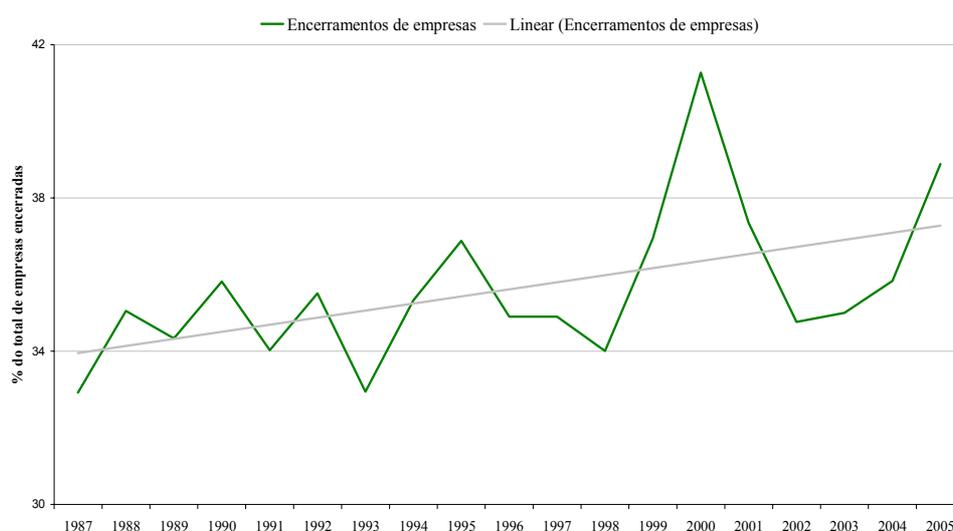
A proporção de encerramentos de empresas no Norte representa mais de um terço do total de encerramentos e empresas no país, atingindo os 39% em 2005 (33% em 1987), superior à proporção de empresas activas na região para a grande maioria dos anos. Os encerramentos de empresas apresentam uma tendência crescente (Figura 3), apesar da inflexão registada no período entre 2000 e 2002, associada ao “pico” a criação de empresas, à elevada correlação dos encerramentos com a taxa de criação de novas empresas e desta última com o ciclo económico (Sarmiento e Nunes, 2010c).

⁵ A região Norte forneceu o maior contributo para o “pico” de criação de empresas empregadoras, ocorrido em 1994, o mais elevado do período em análise (20,8 p.p. de um total de 56,6% de crescimento de novas empresas) (Sarmiento e Nunes, 2010).

⁶ A taxa de natalidade é o rácio do total de nascimentos de empresas empregadoras pelo total de empresas empregadoras activas no período de referência. Um nascimento de uma empresa empregadora ocorre quando esta inicia actividade. Estes não incluem entradas na população devido a fusões, aquisições ou reestruturações de empresas ou reactivações de unidades que estejam adormecidas durante um período de mais de 2 anos. Esta população é também composta por empresas que, embora existindo em anos anteriores, estavam abaixo do limiar de um trabalhador, de acordo com a metodologia Eurostat/OECD (2007).

⁷ A taxa de mortalidade é o rácio do total de mortes em empresas empregadoras, pelo total de empresas empregadoras activas no período de referência. As mortes ocorrem porque as empresas deixam de estar presentes na base de dados (durante pelo menos dois anos) ou porque deixaram de ter pelo menos um trabalhador remunerado, conforme registo nos Quadros de Pessoal, de acordo com a metodologia Eurostat/OECD, 2007.

Figura 3: Proporção de encerramentos de empresas na região Norte



Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

O Norte é a segunda região do país (depois do Algarve), com maior turbulência empresarial, apresentando valores superiores à média nacional entre 1987 e 2005 e entre 1987 e 2000 (Quadro 1). No período compreendido entre 2001 e 2005, verifica-se uma tendência generalizada para a diminuição da turbulência em Portugal (Sarmiento e Nunes, 2010). No entanto, o Norte, a par do Algarve, continua a apresentar a maior taxa de turbulência média do país.

Quadro 1: Taxas de turbulência médias em regiões NUTII, 1987-2005

Taxas de Turbulência	1987-2005	1987-2000	2001-2005
Norte	29,9	30,1	29,4
Algarve	32,5	34,7	29,4
Centro	27,3	28,8	25,1
Lisboa	26,8	27,1	26,2
Alentejo	29,9	31,8	26,6
Açores	27,1	28	25,5
Madeira	28,6	29,4	27,3
Portugal	28,5	29,3	27,3

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

A região Norte apresenta ainda um diferencial a nível de mortalidade de empresas em cerca de 1 ponto percentual superior à média nacional no sector da Indústria, com diferenciais mais

acentuados no sector dos Serviços e da Agricultura e Pescas, em particular entre 2000 e 2005 (Quadro 2).

Quadro 2: Taxas de mortalidade médias por sector económico principal, 1995-2005

1995 a 2000	Agricultura e Pescas	Indústria	Construção	Serviços
Norte	21,6	13,2	21,1	17,9
Portugal	17,5	12,4	21,0	17,0
2000 a 2005				
Norte	45,9	12,8	19,6	17,8
Portugal	23,6	11,8	19,1	16,5

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

3. Análise de Sobrevivência Não Paramétrica

A análise seguinte alicerça-se na estimação não paramétrica, para descrever o comportamento da sobrevivência das empresas no Norte do país, ao longo de um período de 18 anos.

3.1. Modelização teórica das funções de sobrevivência e risco

A aplicação dos tradicionais modelos de sobrevivência ao estudo da demografia de empresas, traduz-se na interpretação da denominada função de sobrevivência, como a probabilidade de uma empresa sobreviver para além de um determinado período de tempo t , condicionada pelo facto de ter sobrevivido até esse mesmo momento (Cleves *et al.*, 2004).

A função de sobrevivência assume o valor 1 na origem, ou seja quando $t=0$, e decresce, tendendo para 0, à medida que o tempo tende para o infinito. Considerando T uma variável não negativa, que representa o período de tempo até ao momento de risco (no caso, o tempo decorrido até que uma empresa abandone o mercado, por encerramento, a partir do momento de entrada), a função de sobrevivência é representada por $S(t) = 1 - F(t) = \Pr(T > t)$. A função de distribuição cumulativa $F(t)$ é então representada por $F(t) = \Pr(T \leq t)$.

No estudo da probabilidade de sobrevivência de uma empresa até um período de tempo t ou da probabilidade de risco desta encerrar após t , aplica-se comumente o estimador não-paramétrico de Kaplan-Meier. Para períodos de risco observados t_1, \dots, t_k , onde k é o número de

períodos distintos de risco observados no conjunto dos dados, o estimador Kaplan-Meier, num dado momento t é dado por:

$$\hat{S}(t) = \prod_{j|t_j \leq t} \left(\frac{n_j - d_j}{n_j} \right), \quad (1)$$

onde n_j é o número de empresas em risco no momento t_j e d_j é o número de encerramentos até t_j . O produto é realizado para todos os períodos de risco desde o momento de criação da empresa em t .

No cálculo da taxa de risco cumulativa, o estimador mais utilizado é o estimador de Nelson-Aalen, que é definido como o somatório das taxas de risco instantâneas do número de empresas em risco de encerrar⁸. O estimador é dado pela seguinte expressão:

$$\bar{H}(t) = \sum_{j|t_j \leq t} \frac{d_j}{n_j}. \quad (2)$$

3.2. Resultados Empíricos para a região Norte

Este método permite observar a performance das empresas da região Norte, no que diz respeito à sua capacidade de sobrevivência, ao longo do seu período de actividade (neste caso compreendido entre 1985 e 2007), bem como estimar qual o risco de sobrevivência associado, quer em termos globais para a região, quer em termos de dimensão e de sector de actividade.

O Quadro 3 apresenta os resultados para a probabilidade de sobrevivência e para o risco de encerramento de uma empresa no Norte do país, ao longo do período 1987-2005,

⁸ O objectivo do cálculo deste estimador é relativamente simples. Se o estimador acumular todos os “riscos” que existem em todos os possíveis instantes entre t_0 e t_j , obter-se-á uma estimativa razoável do risco total que existe entre esses dois pontos de tempo. Assim, por definição, o estimador tem início no momento 0 e cresce ao longo do tempo (nunca decrescendo). Na prática, o estimador não possui uma métrica que seja directamente interpretável (não devendo ser interpretado como uma probabilidade), sendo que a função de risco cumulativa fornece uma ligação importante entre a função de risco (estimada para tempo discreto e que, neste trabalho, corresponde a períodos anuais de tempo) e a função de sobrevivência, cuja natureza de dupla limitação a torna insensível a alterações no risco. Dadas as variações aleatórias que ocorrem em períodos de sobrevivência discretos, a estimação do risco deve ser ajustada de forma a distinguir tendências de “ruídos”. De facto, apesar da função de risco cumulativa ser informativa, quando se considera a estimação de funções de risco contínuas, é importante que se possa visualizar a forma da função de risco num período contínuo de tempo. A solução para obter tal fenómeno visual passa por adoptar a função ajustada de kernel que converte qualquer conjunto estimado de pontos erráticos numa forma funcional “bem comportada” e ajustada. No presente estudo, o estimador ajustado de kernel para a função de risco baseia-se no estimador Nelson-Aalen e na sua variância, isto é, utiliza a taxa de variação das variações do risco acumulado para calcular *pseudo* estimadores de risco, que depois transforma em médias, de forma a estabilizar a função de risco.

condicionada pela sua idade, ou seja pelo facto de estar em actividade há um determinado número de anos.

Quadro 3: Tabela de sobrevivência e risco para as novas empresas criadas, 1987-2005

Tempo	Observações	Encerramentos	Observações censuradas	Função de Sobrevivência Kaplan-Meier	Função de Risco Kaplan-Meier	Taxa de Risco Acumulado Nelson Aalen
Anos	Nº	Nº	Nº	%	%	%
1	160972	23122	19000*	85,6	14,4	14,4
2	119075	14687	5760	75,1	24,9	26,7
3	98628	11213	9401	66,5	33,5	38,1
4	78014	7866	9087	59,8	40,2	48,2
5	61061	5818	5797	54,1	45,9	57,7
6	49446	4345	4025	49,4	50,6	66,5
7	41076	3407	3698	45,3	54,7	74,8
8	33971	2669	3144	41,7	58,3	82,6
9	28158	2075	2295	38,7	61,4	90,0
10	23788	1785	2360	35,8	64,3	97,5
11	19643	1522	3424	33,0	67,0	105,2
12	14697	1107	1962	30,5	69,5	112,8
13	11628	897	1597	28,1	71,9	120,5
14	9134	580	1709	26,4	73,7	126,8
15	6845	406	1585	24,8	75,2	132,8
16	4854	306	1675	23,2	76,8	139,1
17	2873	159	1483	21,9	78,1	144,6
18	1231	69	1162	20,7	79,3	150,2

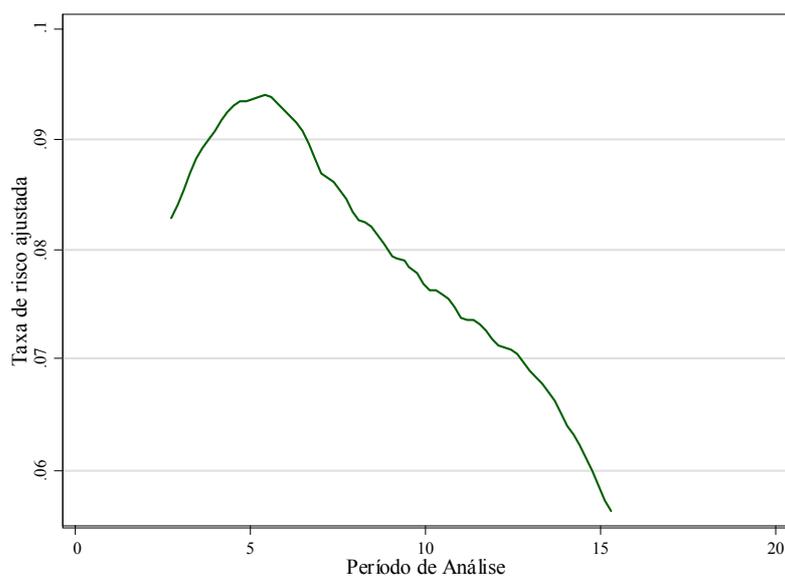
Nota: * Valores aproximados

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

Constata-se que cerca de 86% das novas empresas que nascem em determinado ano (ao longo do período entre os anos de 1987 e 2005), permanecem activas após um ano de actividade. A probabilidade de sobrevivência diminui gradualmente com o aumento do número de anos de “vida” da empresa. Entre o quinto e o sexto ano de actividade, estima-se que cerca de 50% das empresas encerra a sua actividade. Após 18 anos de actividade, apenas 20,7% das empresas permanecem efectivamente activas.

A Figura 4 mostra que, tendo em conta a evolução da taxa de crescimento do risco, existe um aumento da probabilidade do risco de encerramento das empresas até ao sexto ano de actividade. A partir desse momento, o risco de encerramento diminui linearmente.

Figura 4: Estimativa de risco ajustada para o Norte, 1987 – 2005



Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

As taxas de sobrevivência das empresas no Norte do país, por dimensão da empresa, em número de trabalhadores, podem ser visualizadas no Quadro 4.

Quadro 4: Tabela de sobrevivência (%) para empresas nascentes, por dimensão, 1987-2005

Tempo	Dimensão (nº trabalhadores)					
	1-4	5-9	10-19	20-49	50-249	250
1	84,3	89,6	89,7	89,2	90,6	93,2
2	72,5	81,8	81,9	81,7	82,3	83,9
3	63,1	74,8	75,2	74,6	75,5	76,3
4	55,7	69,2	69,6	68,8	70,9	72,4
5	49,4	64,3	64,6	64,3	66,3	68,2
6	44,1	59,9	60,8	60,6	62,4	64,9
7	39,4	56,0	57,6	57,5	60,2	62,6
8	35,5	52,6	54,9	54,4	57,0	61,1
9	32,0	49,7	52,2	52,1	54,3	57,8
10	28,8	46,8	49,6	49,8	52,4	54,4
11	25,9	43,8	47,1	47,7	49,7	51,7
12	23,3	41,0	44,5	45,8	48,4	48,5
13	20,9	38,4	42,1	43,5	46,4	47,2
14	19,0	36,4	40,3	42,1	45,8	45,5
15	17,4	34,6	38,7	41,0	44,2	43,6
16	15,7	33,1	37,3	39,9	41,9	39,6
17	14,4	31,7	36,2	38,7	40,2	39,6
18	13,0	30,6	35,5	37,1	39,2	26,4

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

De acordo com a literatura existente sobre este tema, é natural que empresas de menor dimensão apresentem menores probabilidades de sobrevivência (Mata *et al.*, 1995; Mata e Portugal, 1994; Bartelsman *et al.*, 2004; López-Garcia e Puente, 2006).

Verifica-se que após o primeiro ano de actividade, as empresas com um máximo de 4 trabalhadores, apresentam uma probabilidade de sobrevivência de cerca de 84%, enquanto as empresas de maior dimensão (com mais de 250 trabalhadores), apresentam uma probabilidade de sobrevivência superior, de cerca de 93%. As conclusões são idênticas quando se considera a mediana dos anos de actividade das empresas. As empresas com menos de 4 trabalhadores apresentam uma mediana de idade entre os 4 e os 5 anos. No escalão de dimensão seguinte, esse valor sobe para os 8 a 9 anos e no escalão de maior dimensão (+250), a mediana dos anos de actividade situa-se entre os 11 e os 12 anos.

No Quadro 5, observam-se os resultados da sobrevivência empresarial, desagregados sectorialmente. Os sectores da Construção e da Indústria são aqueles onde as empresas mais encerram prematuramente, no período em análise. Nestes dois sectores, o valor mediano para a actividade de uma empresa situa-se entre os 4 e os 5 anos de vida, aumentando em mais dois anos para os sectores dos Serviços e da Agricultura e Pescas.

Quadro 5: Tabela de sobrevivência (%) para empresas nascentes, por sector, 1995-2005⁹

Tempo	Sectores			
	Agricultura e Pescas	Construção	Industria	Serviços
1	89,3	83,3	85,5	86,2
2	78,9	71,0	74,3	75,8
3	71,0	60,4	64,9	67,5
4	64,4	52,4	56,9	60,7
5	58,8	45,3	49,7	55,0
6	52,8	40,0	44,5	50,3
7	48,1	35,8	40,5	46,5
8	45,2	32,5	37,4	43,5
9	41,9	29,5	34,4	40,6
10	38,3	26,9	31,7	37,6

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

4. Análise de Sobrevivência Semi-Paramétrica

De seguida, complementa-se a análise anterior, abordando os determinantes da sobrevivência empresarial e o seu comportamento de acordo com as características específicas do mercado e da demografia empresarial de cada região, recorrendo a métodos semi-paramétricos.

⁹ Consideram-se apenas os sectores a uma letra (CAE Rev. 2.1) apenas a partir de 1995, devido à introdução do SEC 1995.

Nesta secção do estudo, estimação será efectuada apenas para o período compreendido entre 1995 e 2005, devido à introdução do Sistema Europeu de Contas (SEC95) em 1995, e até 2005, devido à metodologia Eurostat/OECD (2007). A aplicação desta metodologia obriga a utilizar para efeitos de confirmação de possíveis reactivações, os dois anos seguintes após o encerramento das empresas. Consequentemente, as mortes de empresas são apenas calculadas até 2005, ao invés de 2007.

4.1. Modelização teórica da função de risco

A aplicação do método que utiliza um conjunto de variáveis explicativas do comportamento de sobrevivência e do risco de encerramento de uma empresa e a representação estatística da relação entre o seu tempo de sobrevivência e estas variáveis, é denominada como modelo de taxa de risco da duração de vida da empresa.

De acordo com este modelo, uma empresa específica j , enfrenta uma taxa de risco (h_j) , que é função de uma taxa de risco base (h_0) , que todas as empresas a operar no mercado enfrentam. Esta é transformada através de um vector de parâmetros (β) por um conjunto de variáveis explicativas (X) . O modelo pode então ser escrito da seguinte forma:

$$h_j(t) = f(h_0(t), \phi(X, \beta)) \quad (3)$$

Consequentemente, duas empresas com a mesma data de criação, enfrentarão uma diferente função de risco se, e apenas se, as suas características forem distintas.

A aplicação empírica do modelo implica a especificação de uma forma funcional para a função de sobrevivência. Uma das mais comuns é dada pelo modelo de risco proporcional:

$$h_j(t) = h_0(t)\phi(X, \beta) \quad (4)$$

A designação do modelo resulta do facto do risco enfrentado pela empresa ser proporcional ao risco base. Ou seja, o formato da função de risco é o mesmo para todos os indivíduos, sendo

que as variações ocorridas nas variáveis explicativas se traduzem em deslocamentos paralelos dessa função base afectando, assim, apenas a escala da função original e não o seu formato.

Face ao facto do risco se constituir como uma probabilidade condicionada e, por isso, ser positivo, a forma mais conveniente para o termo $\phi(X, Y)$ consiste na forma exponencial. Em resultado, o risco que uma empresa j enfrenta, pode ser escrito da seguinte forma: $h_j(t) = h_0(t)e^{(X, \beta)}$. Note-se que esta forma particular de representação da função de risco oferece a vantagem de permitir uma interpretação conveniente dos coeficientes estimados, dado que

$\beta = \frac{\partial \ln \phi(X, \beta)}{\partial X}$. A expressão anterior significa que o coeficiente de uma variável explicativa é o efeito proporcional constante resultante do acréscimo em uma unidade, desta variável na probabilidade condicionada de encerramento.

A hipótese assumida para a forma funcional de $\phi(X, Y)$ é amplamente aceite, mas o mesmo não acontece para a forma funcional da função de risco base. Assim sendo, a melhor opção será a de não considerar qualquer hipótese acerca da forma funcional da função de risco base. Esta foi a solução apontada por Cox (1972), sendo que os modelos daí resultantes são conhecidos como semi-paramétricos. Cox também sugeriu que os modelos de risco proporcional poderiam ser facilmente alargados, de forma a integrarem variáveis explicativas que variassem ao longo do tempo. Essa será uma das aproximações apresentadas do longo do trabalho.

De facto, o modelo incorpora as principais características dos modelos discretos de duração, onde o logaritmo da probabilidade de uma empresa encerrar no momento de tempo t , dado que sobreviveu até $t-1$, é explicado por um conjunto de variáveis X_t e um conjunto de variáveis que identificam a função de risco base, de acordo com a seguinte especificação:

$$\log h(t | x_t, x_0) = \lambda_t + \beta x_t + \gamma x_0, \quad \text{para } t = 1, \dots, k \quad (5)$$

Na análise seguinte, são aplicadas duas especificações concretas do modelo proporcional de Cox, a especificação contínua do modelo e a especificação discreta. Esta opção é plausível, pois o fenómeno de sobrevivência, sendo contínuo, é reportado nos Quadros de Pessoal, numa base anual, transformando o tempo numa variável discreta. Assim sendo, na especificação discreta, os dados são agrupados em 11 variáveis de tempo (uma para cada período em risco,

entre 1995 e 2005), onde se aplica o modelo complementar log-log (conhecido como modelo Cloglog). Na especificação contínua será adoptado o modelo de risco proporcional de Cox.

4.2. Resultados Empíricos para a Região Norte

Foram seleccionadas sete variáveis explicativas para a estimação semi-paramétrica da função de risco de encerramento da empresa: a dimensão inicial da empresa, a dimensão actual da empresa, a taxa de entrada em actividade da empresa, o índice de concentração (*Herfindahl-Hirschman Index*), a taxa de turbulência, o crescimento específico de cada sector económico, o produto entre crescimento, a taxa de entrada no mercado e finalmente variáveis dicotómicas (*dummies*) sectoriais e temporais.

A descrição de cada variável e das estimativas obtidas através dos modelos Cloglog e proporcional de Cox, estão sistematizados no Quadro 6. Os resultados são apresentados sob a forma de taxas de risco, onde valores superiores a 1 indicam uma taxa de risco crescente e valores inferiores a 1, uma taxa de risco decrescente.

Apesar das diferenças na magnitude dos estimadores, estes apresentam o mesmo padrão em termos de evolução do risco, nos dois modelos. Embora o modelo proporcional de Cox perca significância estatística para algumas variáveis, ambos os modelos permitem concluir que existe um maior risco de encerramento associado a empresas que têm menor dimensão (actual), com actividade em sectores que verificam elevadas taxas de entrada de empresas e particularmente, de turbulência. Este efeito é particularmente explícito no sector industrial.

Relativamente ao sector da indústria, o Norte apresenta uma taxa de risco mais baixa nos restantes sectores económicos, indicando portanto uma menor probabilidade de sobrevivência das empresas que operam no sector industrial, conclusões já suportadas pela análise não paramétrica de sobrevivência e pela demografia empresarial (GEE, 2010; Sarmiento e Nunes, 2010). A indústria é efectivamente o sector no qual as empresas da região Norte apresentam a maior probabilidade de abandonar o mercado.

A taxa de turbulência é a variável com o efeito mais significativo nas probabilidades de sobrevivência das empresas, em todas as regiões do país, mas com especial ênfase na região Norte de Portugal (11,5% de risco).

Quadro 6: Resultados das estimações semi-paramétricas para o Norte

Variáveis	Descrição	Modelo Cloglog	Modelo Proporcional de Cox
Dimensão Original	Logaritmo do número de empregados no momento do nascimento da empresa	1,310 *** (-0,011)	1,081 *** (-0,013)
Dimensão Actual	Logaritmo do número de empregados no momento de realização da análise	0,479 *** (-0,004)	0,639 *** (-0,007)
Taxa de Entrada na Industria (2 dígitos)	Logaritmo da taxa de entrada na indústria definida como o número de empresas que entram dividido pelo número total de empresas existentes na indústria (o valor é calculado tendo em conta a CAE Rev.2.1 a 2 dígitos)	1,233 *** (-0,028)	1,078 ** (-0,036)
HHI da Indústria (2 digit)	Índice Herfindhal-Hirschman (HHI) calculado para indústrias ao nível da CAE Rev.2.1 a 2 dígitos	0,986 *** (-0,002)	0,980 *** (-0,004)
Taxa de Turbulência	Soma das taxas de entrada e saída de empresas por indústria ao nível da CAE Rev.2.1 a 2 dígitos	11,444 *** (-0,791)	9,394 *** (-1,097)
Crescimento Industrial	Logaritmo do número de empregados no ano t menos o logaritmo do número de empregados no ano t-1.	1,087 *** (-0,029)	0,953 (-0,01)
Crescimento Industrial x Taxa de Entrada na Industria	Variável de interacção definida como o produto da taxa de entrada e o crescimento industrial.	1,069 ** (-0,028)	0,997 (-0,063)
Sectores (<i>dummies</i>)			
Agricultura e Pescas		0,443 *** (-0,016)	0,554 (-0,024)
Construção	<i>Dummies</i> para quatro sectores genéricos	0,858 *** (-0,012)	0,81 (-0,018)
Industria		(a)	(a)
Serviços		0,631 *** (-0,007)	0,622 (-0,01)
Anos (<i>dummies</i>)		<i>Dummies</i> para cada ano em análise.	Sim
	Número de Empresas	164599	164599
	LR X^2 (28) / LR X^2 (18)	17871,03 ***	5526,57 ***
	Log likelihood	-152494,34	-348970,76

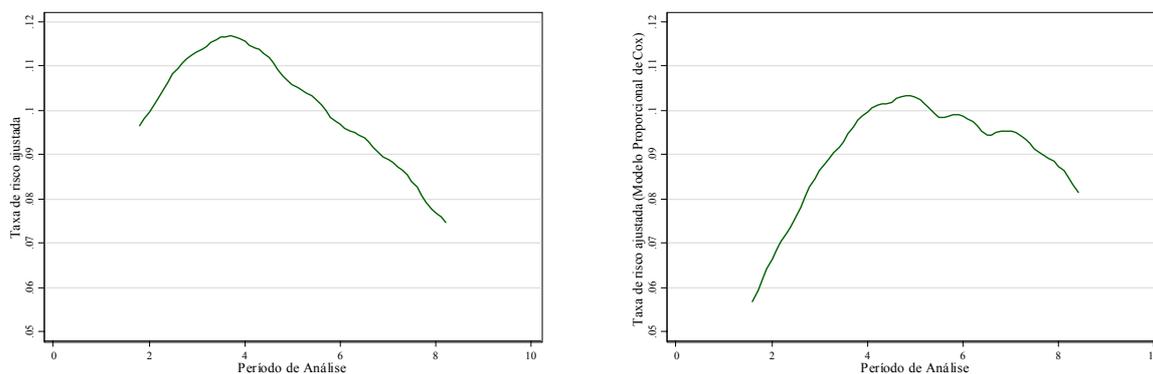
Notas: Nas *dummies* para os anos a expressão “Sim” significa que foram incluídas na estimação. Os desvio padrão são apresentados entre parêntesis. (a) refere-se ao sector de referência. *, **, *** significa respectivamente, 10%, 5% e 1% de nível de significância.

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

A introdução de variáveis explicativas na estimação semi-paramétrica (modelo de risco proporcional de Cox), permite observar uma diminuição nas taxas de risco de encerramento, na evolução da sobrevivência das empresas ao longo do tempo (neste caso compreendido entre 1995 e 2005), quando comparada com a estimação não paramétrica de Kaplan-Meier (Figura 5).

Por outro lado, observa-se que, embora com um valor mais reduzido, o período em que o risco aumenta é superior, no modelo de Cox (aproximadamente 5 anos, em comparação com o pico de 4 anos observado na estimação não paramétrica), decrescendo menos rapidamente e de forma menos linear.

Figura 5: Taxas de risco ajustada não paramétrica (Kaplan Meier) e Semi-paramétrica (Modelo Proporcional de Cox), 1995-2005



Fonte: Cálculos próprios baseados nos Quadros de Pessoal, GEP, MTSS.

5. Comentários finais

A importância económica da região Norte, bem como a preponderância do seu tecido empresarial no contexto da dinâmica nacional, concretamente a nível sectorial, justificam um retrato cuidado sobre a sobrevivência empresarial nesta região, dado que se caracteriza por aspectos que a tornam particular no panorama empresarial do país. O Norte apresenta a maior taxa de crescimento do número de empresas entre 2000 e 2007 e a maior proporção de empresas activas no país, sendo também responsável pela maioria da criação e encerramento de empresas entre 1985 e 2007. É também a região do país onde é visível uma maior volatilidade em termos de dinâmica empresarial. A turbulência registada nas entradas e saídas de empresas do mercado contribuiu significativamente para explicar o elevado comportamento de risco relacionado com a capacidade de sobrevivência das suas empresas.

Na estimação não paramétrica, observa-se que cerca de 86% das novas empresas criadas na região Norte, permanecem activas após um ano de actividade. A probabilidade de sobrevivência diminui com o aumento do número de anos de actividade da empresa. Verifica-se que entre o 5º e o 6º ano de actividade, cerca de 50% das empresas encerram a sua actividade. Após 18 anos de actividade, apenas 20,7% do total de empresas empregadoras permanecem activas na região Norte. A probabilidade de sobrevivência é relativamente mais baixa nos sectores da Construção e da Indústria e nas empresas de menor dimensão.

A introdução de variáveis explicativas na estimação semi-paramétrica obtida para a região Norte (modelo de risco proporcional de Cox), permite observar uma diminuição do risco na evolução da sobrevivência das empresas ao longo do tempo, quando comparada com a estimação não paramétrica de Kaplan-Meier. No entanto, ambos os modelos permitem

concluir pela existência de um maior risco de encerramento, associado a empresas com menor dimensão e com actividade em sectores que verificam elevadas taxas de entrada de empresas, e particularmente de turbulência. Este efeito é particularmente explícito no sector industrial.

A taxa de turbulência associada à entrada e saída de empresas no mercado é o principal determinante das probabilidades de sobrevivência das empresas em todas as regiões, mas assume na região Norte um efeito particularmente expressivo.

Bibliografia

Baptista, R., Escária, V. e Madruga, P. (2008), "Entrepreneurship, Regional Development and Job Creation: the Case of Portugal" in *Small Business Economics*, nº 30, pp. 49-58.

Baptista, R. e Carias, C. (2007), "Job Creation and Destruction by Small Businesses in the Portuguese Regions", IN+, presented at the European Regional Science Association, ERSA 2007.

Baptista, R. e Mendonça, J. (2007), "Short and Long Term Survival of New Businesses: the Role of Human Capital", IN+, Draft version of 2007.

Bartelsman, E., Haltiwanger, J. e Scarpetta, S. (2004), "Microeconomic evidence of creative destruction in industrial and developing countries", Policy Research Working Paper Series 3464, The World Bank.

Brixy, U. e Grotz, R. (2006), "Regional Patterns and Determinants of New Firm Formation and Survival in Western Germany", IAB Discussion Paper, nº 5/2006.

Cabral, L. (2007), "Small Firms in Portugal: A Selective Survey of Stylized Facts, Economic Analysis and Policy Implementation" in *Portuguese Economic Journal*, Vol. 6, nº 1, pp. 65-88.

Cabral, L. e Mata, J. (2003), "On the Evolution of the Firm Size Distribution: Facts and Theory", *The American Economic Review*, Vol. 93, nº 4, pp. 1075 - 1090.

Cleves, M. A., Gould, W.W. e Gutierrez, R. G. (2004), "An Introduction to Survival Analysis Using Stata", Stata Press.

Cox, D.R. (1972), "Regression Models and Life Tables" in *Journal of the Royal Statistical Society, Series B*, Vol. 34, nº 2, pp. 187-202.

- Eurostat (2009), "Business Demography: employment and survival", *Statistics in focus*, 70/2009.
- Eurostat/OECD (2007), "Eurostat/OECD Manual on Business Demography Statistics".
- GEE (2010), "Síntese Estatística da Dinâmica Empresarial: Região Norte", Gabinete de Estratégia e Estudos do Ministério da Economia, da Inovação e do Desenvolvimento.
- GEP (2010), "Quadros de Pessoal, Estatísticas em Síntese", Gabinete de Estratégia e Planeamento, Ministério do Trabalho e da Solidariedade Social.
- INE (2009a), "Demografia de Empresas 2004-2007, O Empreendedorismo em Portugal - Indicadores sobre a Demografia das Empresas", Destaque do Instituto Nacional de Estatística.
- INE (2009b), "Empresas em Portugal 2007", *Temas de Economia e Finanças*, Instituto Nacional de Estatística.
- López-García, P. e Puente, S. (2006), "Business demography in Spain: determinants of firm survival", *Documentos de Trabajo n. 608*, Banco de España.
- Mata, J. (1993), "Entry and Type of Entrant: Evidence from Portugal" in *International Journal of Industrial Organization*, Vol. 11, nº 1, pp. 101-122.
- Mata, J. e Machado, J. (1996), "Firm Start-up Size: a Conditional Quantile Approach" in *European Economic Review*, Vol. 40, pp. 1305-1323.
- Mata, J. e Portugal, P. (1994), "Life Duration of New Firms" in *The Journal of Industrial Economics*, Vol. 42, nº 3, pp. 227-245.
- Mata, J., Portugal, P. e Guimarães, P. (1995), "The Survival of New Plants: Start-up Conditions and Post-Entry Evolution" in *International Journal of Industrial Organization*, nº 13, pp. 459-481.
- Nunes, A. e Sarmiento, E. de Morais (2010a) "Business demography dynamics in Portugal: a semi-parametric survival analysis", *Working Papers do Grupo de Estudos Monetários e Financeiros*, nº 9/2010, Universidade de Coimbra (2010).
- Nunes, A. e Sarmiento, E. de Morais (2010b), "Business demography dynamics in Portugal: a non-parametric survival analysis", *Working Papers do Grupo de Estudos Monetários e Financeiros*, nº 10/2010, Universidade de Coimbra (2010).

OECD/Eurostat (2009), "Measuring Entrepreneurship A Collection of Indicators 2009 Edition", OECD-Eurostat Entrepreneurship Indicators Programme.

OECD/Eurostat (2008), "Measuring Entrepreneurship: A Digest of Indicators", OECD-Eurostat Entrepreneurship Indicators Programme.

Sarmiento, E. de Moraes e Nunes, A. (2010), "Entrepreneurship Performance Indicators for Active Employer Enterprises in Portugal", Temas Económicos nº 9, Ministério da Economia, da Inovação e do Desenvolvimento.