Regional survival in Portugal
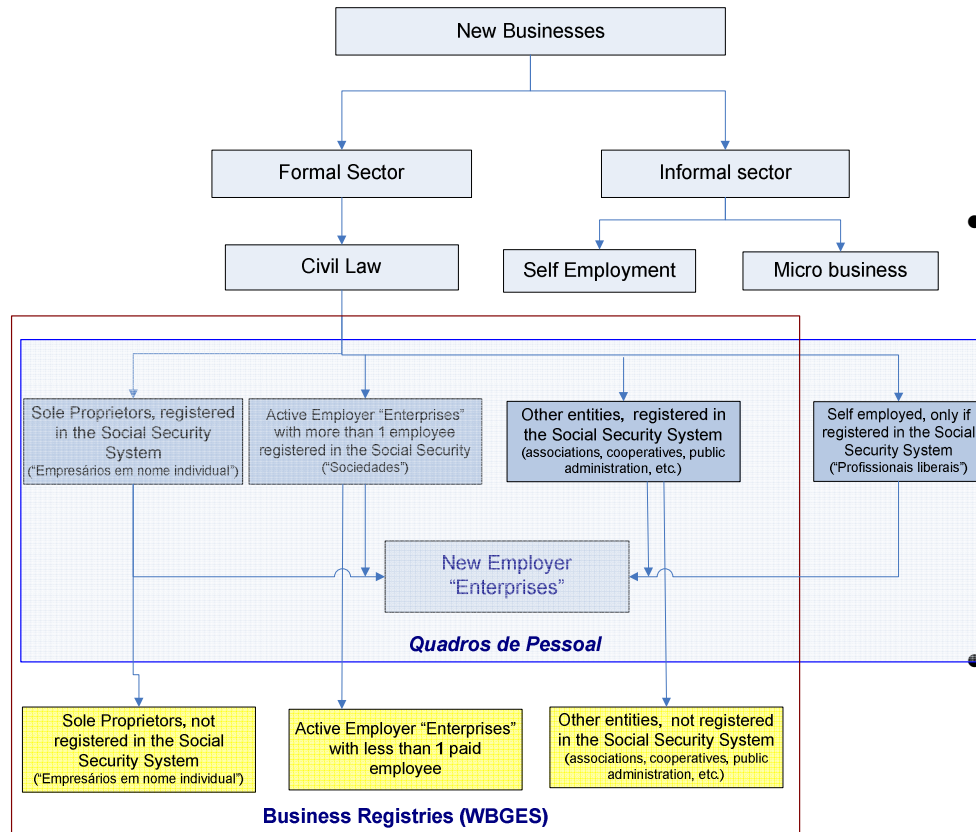
Alcina Nunes

Instituto Politécnico
de Bragança

(alcina@ipb.pt)

Elsa Sarmento

Universidade de Aveiro

Gabinete de Estratégia e
Estudos, MEID

(esarmento@ua.pt)

Conselho Superior de Estatística
25 de Março de 2011
Lisboa

- **Aims**:
  - Approach the determinants of firm survival for 7 NUTII regions, using non-parametric (Kaplan-Meier and Nelson-Aalen) and semi-parametric (Cox proportional hazard model, cloglog) procedures, considering the period 1985 to 2007, based on the performance indicators obtained by applying the Eurostat/OECD methodology "Manual of Business Demography Statistics" to the dataset "Quadros de Pessoal".

- **Motivation**:
  - Approach the prevalence of some of the main stylised facts and determinants of firm survival at the national and regional level in Portugal.

- **Target**:
  - Employer Enterprise Population (enterprises with more than one paid worker according the registrars of the Portuguese Social Security System), using the entrepreneurship dataset obtained from *Quadros de Pessoal.*

- **Contribution**:
  - Application of a methodology that allows comparability at the international and regional level, to a rich matched employer-employee dataset, over a period of 20 years.

"Quadros de Pessoal"

Diagram:

New Businesses
- Formal Sector
  - Civil Law
- Informal sector
  - Self Employment
  - Micro business

Quadros de Pessoal
- Sole Proprietors, registered in the Social Security System ("Empresários em nome individual")
- Active Employer "Enterprises" with more than 1 employee registered in the Social Security ("Sociedades")
- Other entities, registered in the Social Security System (associations, cooperatives, public administration, etc.)
- Self employed, only if registered in the Social Security System ("Profissionais liberais")

New Employer "Enterprises"

Business Registries (WBGES)
- Sole Proprietors, not registered in the Social Security System ("Empresários em nome individual")
- Active Employer "Enterprises" with less than 1 paid employee
- Other entities, not registered in the Social Security System (associations, cooperatives, public administration, etc.)

- Its importance relies on its longitudinal national coverage, at a microeconomic level, and on its uniqueness as a linked employer-employee data source in Portugal since 1983;

- It is a comprehensive administrative dataset obtained through a mandatory employment survey, conducted annually by the Portuguese Ministry of Labour and Social Security;

  Enterprises and establishments with wage earners are legally obliged to fill in a standardize annual questionnaire;

  It covers, virtually, all enterprises and establishments with wage earners with the exception of public administration, non-market services and activities with workers not included in the "common" social security regime;

- Focus on the formal sector of the economy.

Enterprise birth

Employer enterprise birth
A birth occurs when and enterprise starts from scratch and actually starts activity. An enterprise creation can be considered an enterprise birth if new production factors, in particular new jobs, are created.

If a dormant unit is reactivated within two years, this event is not considered a birth. This caused enterprise births to be effectively accounted for from 1987 onwards, instead of 1985 .

Thus, this population consists of enterprises that have at least one paid employee in its birth year and also of enterprises that, despite existing before the year in consideration, were below the one employee threshold.

The employer enterprise birth rate is based on a numerator which follows the above definition for employer enterprise births, and a denominator which consists of the population of active enterprises with one or more employees during the reference period.

Enterprise death

A death amounts to the dissolution of a combination of production factors with the restriction that no other enterprises are involved in the event. Deaths do not include exits from the population due to mergers, take-overs, break-ups or restructuring of a set of enterprises. It does not include exits from a sub-population resulting only from a change of activity.

An enterprise is included in the count of deaths only if it is not reactivated within two years. Equally, a reactivation within two years is not counted as a death.

This caused enterprise births to be effectively accounted up to 2005, instead of 2007 .

Enterprise churn rate

Given by the sum of enterprise birth and death rates.

The entrepreneurship database obtained from *Quadros de Pessoal*, following the Eurostat/OECD (2007) methodology, consists of an annual average of 215,903 active employer enterprises over the period 1985-2007, with an annual average of 36,803 births and 23,743 deaths.

**Population of Active Employer Enterprises, 1987-2007**
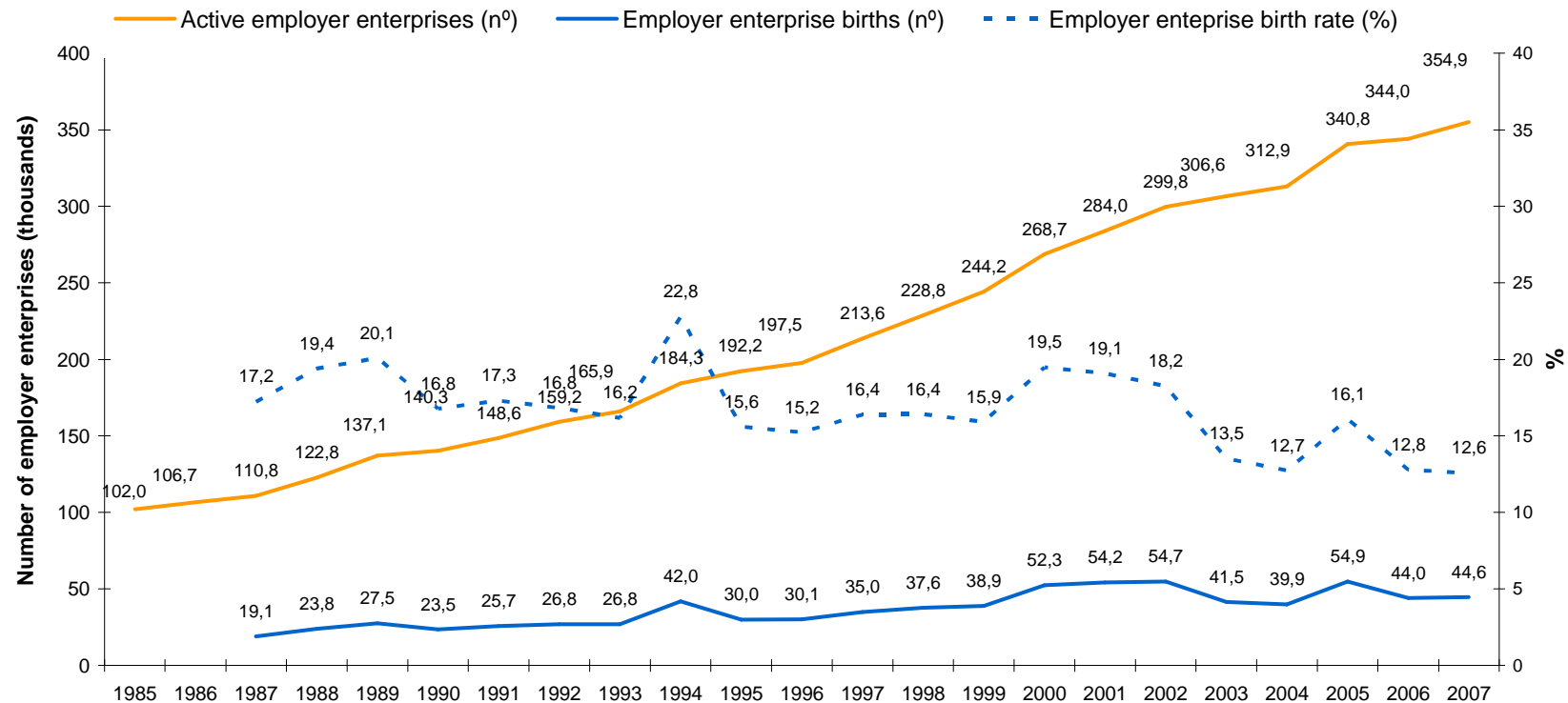
Source: Own calculations based on Quadros de Pessoal, GEP, Ministério do Trabalho e da Solidariedade Social.
Note: Employer Enteprises are enterprises which have at least one paid worker. Birth rates are calculated as the ratio of the number of employer enterprise births over the population of employer enterprises during the reference period.

Among the European countries, Portugal has one of the highest records of new firm creation, even when other universes and methodologies are considered.

Birth rates according to different sources

■ 2005 ■ 2006
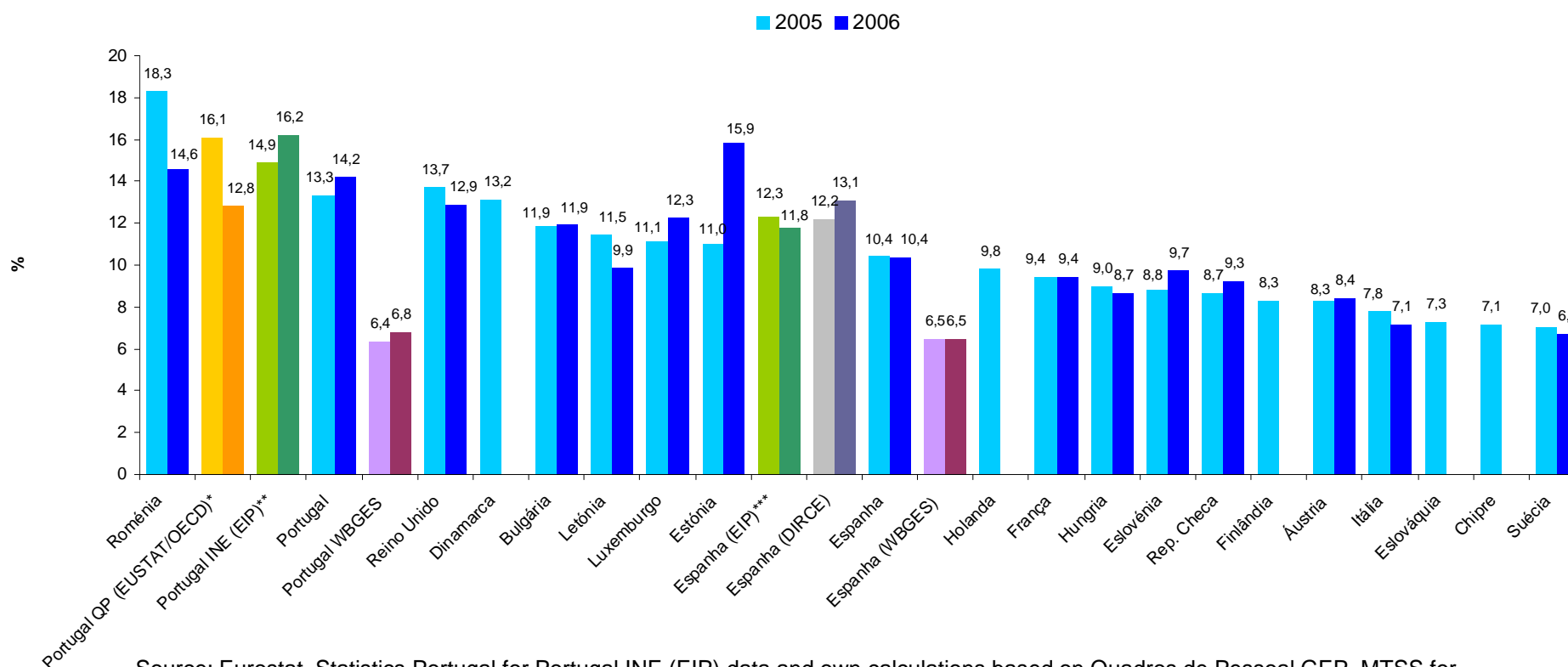
Source: Eurostat, Statistics Portugal for Portugal INE (EIP) data and own calculations based on Quadros de Pessoal GEP, MTSS for Portugal QP (Eurostat/OECD)and SDBS Business Demography Indicators from the OECD (EIP).
Notes: Preliminary version of 2005 for Bulgaria, Romania, Portugal and Slovenia.
    * Employer enterprises according to the Eurostat/OECD methodology, based on Quadros de Pessoal.
    ** Statistics Portugal data, for enterprises with more than 1 paid employee (employer enterprises).

Regional business characteristics and dynamics

Norte is the region with the highest number of births and share of active enterprises in the country. Algarve is the region with the highest growth in active enterprises and rate of birth in Portugal between 2000 and 2007, where firms are created with the smallest average size.

Active Employer Enterprises by NUT II (1987, 1997, 2004 and 2007)

0-20.000	(cyan)
20.001 to 50.000	(blue)
> 50.000	(dark blue)

NORTE

CENTRO

LISBOA e VALE DO TEJO

ALENTEJO

ALGARVE

AÇORES

MADEIRA
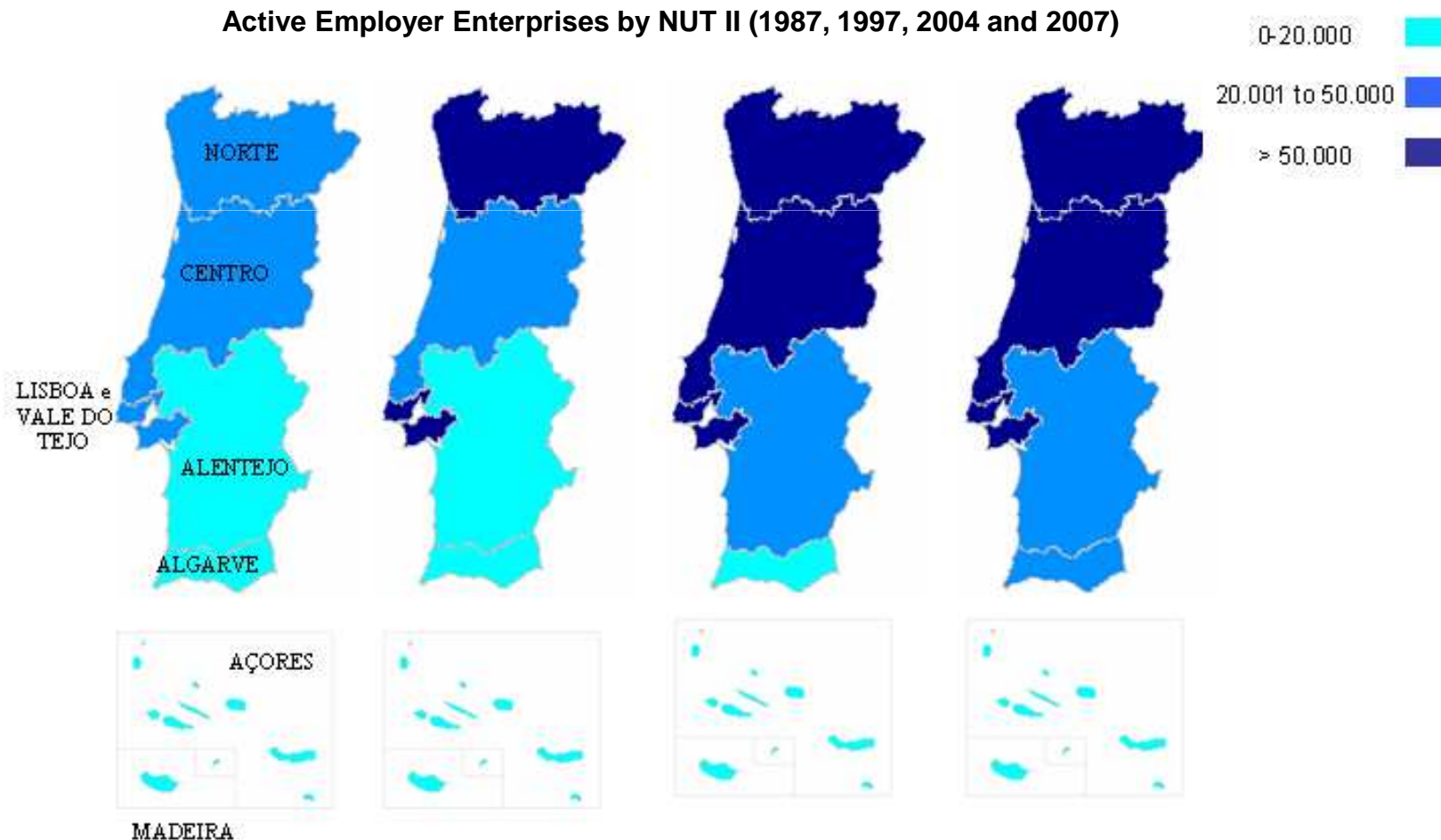
Annual average growth rate of employer enterprises, births and deaths by NUT II

	Active employer enterprises			Births		Deaths	
NUTII Regions	**1985-2007**	**1995-2000**	**2000-2007**	**1987-2007**	**2000-2007**	**1985-2005**	**2000-2005**
Norte	6.2	7.1	4.4	4.3	-1.5	6.4	-1.0
Algarve	9.2	7.9	6.7	6.2	1.0	7.4	3.4
Centro	6.6	8.6	4.0	4.6	-5,3	6.3	1.3
Lisboa e Vale do Tejo	4.5	5.1	3.4	4.3	-1.1	4.8	0.8
Alentejo	5.8	8.3	3.1	3.1	-4.2	5.3	2.1
Açores	3.7	3.9	3.1	1.1	-1.0	0.4	-4.9
Madeira	6.4	7.1	4.4	4.9	-1.1	6.3	-3,3
Portugal	5.8	6.9	4.1	4.3	-2.3	5.7	0.2

Average churn rates of Portuguese NUT II regions

	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%	28,0%	25,5%
Madeira	28,6%	29,4%	27,3%
Portugal	28,5%	29,3%	27,3%

Death rates of active employer enterprises by NUT II regions (1987, 1997 e 2004)

Legend:
- 0 - 9.9%
- 10% - 10.9%
- 11% - 11.9%
- > 12%

Deaths of active employer enterprises by NUT II regions

Legend: ■ Norte ■ Lisboa e Vale do Tejo ■ Centro ■ Alentejo ■ Algarve ■ Açores ■ Madeira

Year	Norte	Lisboa e Vale do Tejo	Centro	Alentejo	Algarve
2005	14807	9165	7851	2869	2148
2004	11501	8490	6545	2355	1974
2003	11471	8761	6819	2681	1812
2002	12298	9978	7272	2709	1888
2001	14074	9672	7728	2882	1904
2000	15534	8825	7363	2580	1813
1999	10295	7548	5455	2088	1486
1998	7966	6698	4671	1880	1320
1997	7883	6317	4376	1863	1244
1996	7286	5977	4077	1523	1117
1995	7856	5805	4118	1510	1072
1994	7572	6236	3920	1707	1133
1993	8311	7422	4910	2295	1254
1992	6441	4998	3362	1581	1005
1991	5798	4912	3191	1593	907
1990	6000	4630	2901	1713	835
1989	6287	5237	3539	1616	990
1988	4929	4057	2455	1320	587
1987	3921	3704	2156	1009	496
1986	4393	3914	2870	1207	536
1985	4301	3605	2326	1023	512

Regional dynamics

Norte

Legend: Taxa de Natalidade, Taxa de Mortalidade, Taxa de turbulência para Portugal

Ano	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Total	29,1	32,1	33,7	29,8	28,4	27,7	31,6	35,3	27,9	26,6	27,5	27,2	28,7	35,9	35,1	30,9	25,2	24,0	31,9
Mortalidade	10,5	11,7	13,3	12,3	11,3	11,9	14,9	11,9	11,8	11,1	10,7	10,1	12,0	16,6	14,4	12,1	10,9	10,6	12,1
Natalidade	18,6	20,5	20,4	17,4	17,1	15,9	16,7	23,4	16,1	15,5	16,8	17,1	16,7	19,3	20,6	18,8	14,3	13,4	19,9

Nota: O somatório das taxas de natalidade e mortalidade de empresas empregadoras, no topo das barras, corresponde à taxa de turbulência.

Centro

Legend: Taxa de Natalidade, Taxa de Mortalidade, Taxa de turbulência para Portugal

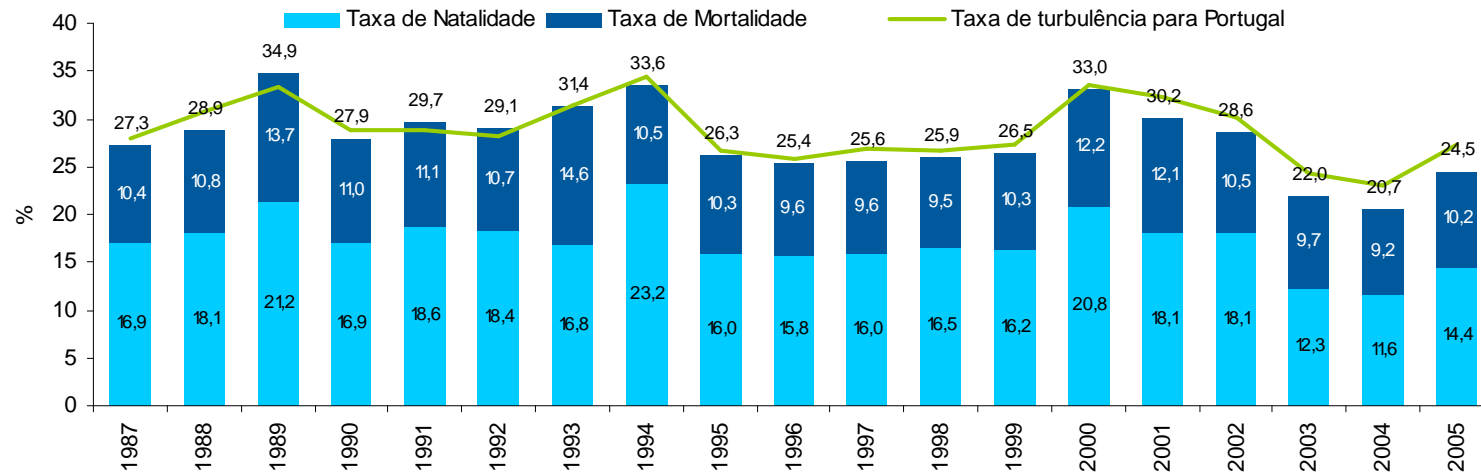| Ano	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Total	27,3	28,9	34,9	27,9	29,7	29,1	31,4	33,6	26,3	25,4	25,6	25,9	26,5	33,0	30,2	28,6	22,0	20,7	24,5
Mortalidade	10,4	10,8	13,7	11,0	11,1	10,7	14,6	10,5	10,3	9,6	9,6	9,5	10,3	12,2	12,1	10,5	9,7	9,2	10,2
Natalidade	16,9	18,1	21,2	16,9	18,6	18,4	16,8	23,2	16,0	15,8	16,0	16,5	16,2	20,8	18,1	18,1	12,3	11,6	14,4

Nota: O somatório das taxas de natalidade e mortalidade de empresas empregadoras, no topo das barras, corresponde à taxa de turbulência.

Birth, death and churn rates in 1987, 2000 and 2005
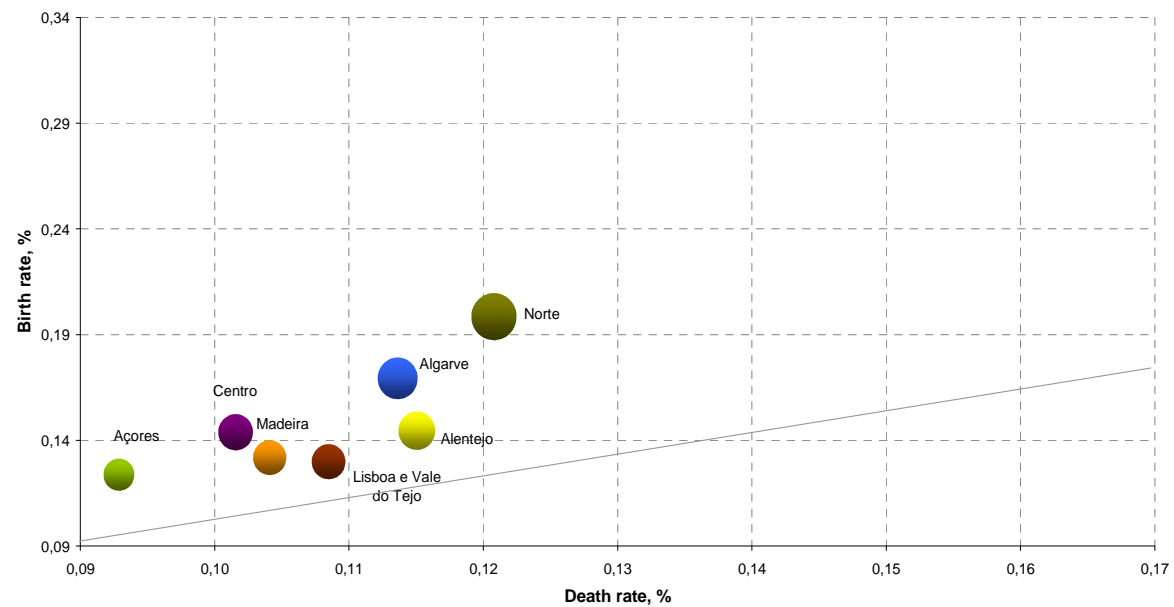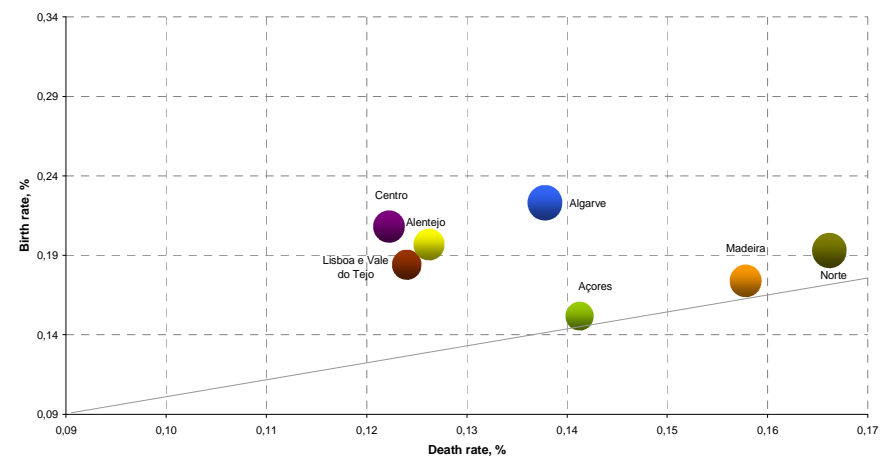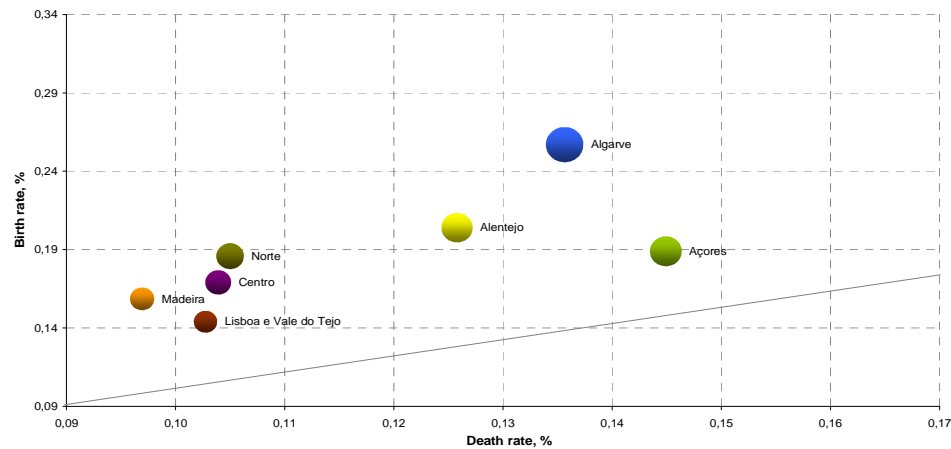
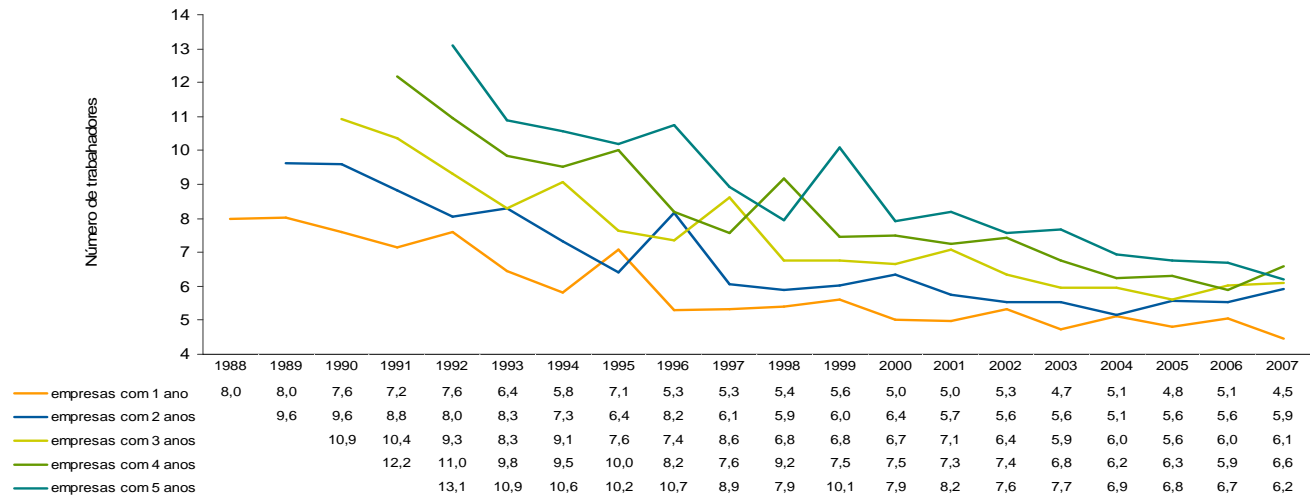## Entrants and exiters average firm size, by NUTII

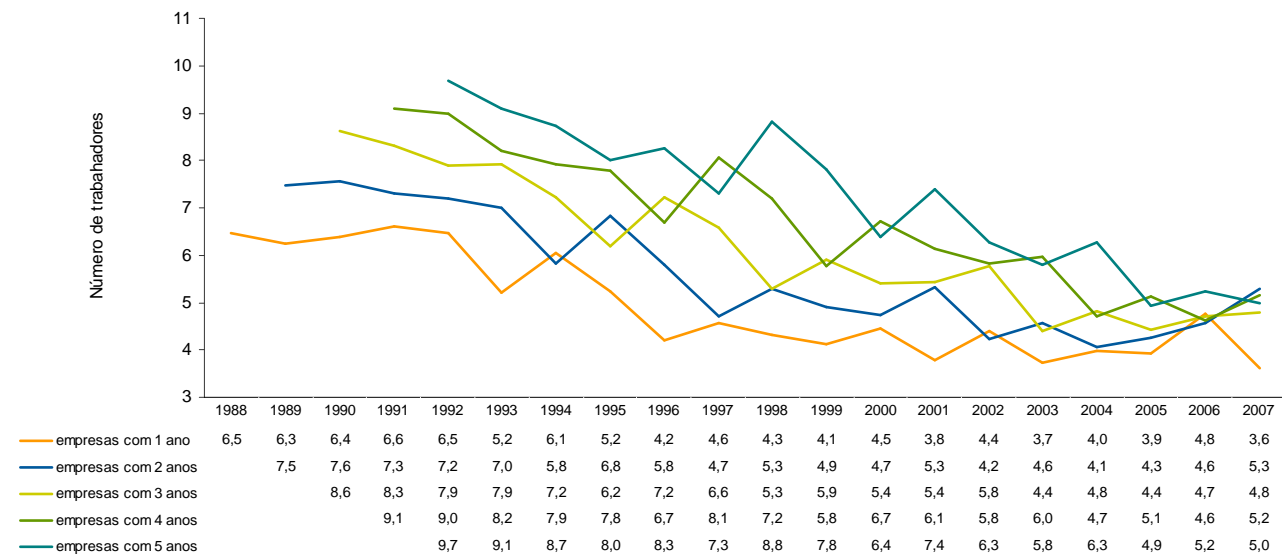|	Average firm size of entrants								Average firm size of exiters							
	Norte	Algarve	Centro	Lisboa	Alentejo	Açores	Madeira	Portugal	Norte	Algarve	Centro	Lisboa	Alentejo	Açores	Madeira	Portugal
1985									7.9	6.3	6.0	6.8	5.7	3.1	7.0	6.8
1986									7.5	5.7	6.0	6.1	6.3	3.7	9.1	6.5
1987	6.2	4.7	5.2	5.1	4.5	3.6	6.9	5.4	8.4	5.1	5.9	6.3	5.2	3.3	8.1	6.7
1990	6.1	4.1	5.3	6.3	4.3	3.0	4.8	5.7	7.7	4.9	5.7	6.1	4.6	3.4	7.1	6.4
1991	6.2	4.4	5.3	6.8	4.0	2.8	6.2	5.8	8.0	4.1	6.3	6.7	4.4	2.7	4.5	6.6
1992	5.6	4.0	4.6	5.5	4.2	3.4	4.4	5.1	7.6	4.5	5.9	6.3	4.9	3.2	5.9	6.4
1993	4.9	3.8	5.1	6.5	4.1	4.4	5.0	5.2	7.6	4.6	6.1	12.4	5.4	3.6	5.4	8.2
1994	5.9	4.1	4.5	8.6	4.3	3.8	5.4	6.1	6.5	3.8	4.6	5.5	3.9	4.0	4.3	5.4
1995	4.6	3.4	3.6	4.3	3.3	2.8	5.2	4.1	6.4	3.6	4.6	4.9	4.3	3.1	3.2	5.2
1996	4.3	3.3	3.8	4.5	3.9	2.4	4.7	4.1	5.5	3.4	4.6	5.0	3.9	2.4	4.5	4.9
1997	4.5	3.4	3.5	4.4	3.6	2.5	3.8	4.1	5.9	3.5	4.3	5.6	3.9	2.1	3.7	5.1
1998	4.6	3.3	3.4	5.1	3.3	3.2	4.4	4.3	6.0	3.0	4.1	4.8	3.2	2.5	3.4	4.8
1999	4.2	3.4	3.6	4.8	3.6	2.6	4.0	4.1	6.4	3.2	4.0	6.1	3.5	2.9	5.6	5.4
2000	4.0	2.9	3.1	6.6	3.1	3.9	4.3	4.3	5.2	3.1	4.3	5.7	3.0	3.0	4.3	4.8
2001	4.6	3.4	3.7	4.6	3.3	4.1	4.8	4.2	5.0	4.1	4.2	5.7	3.6	3.3	4.9	4.8
2002	4.1	3.2	3.3	4.4	3.3	3.9	4.4	3.9	5.0	3.0	3.8	4.8	3.4	3.2	3.9	4.4
2003	4.2	3.3	3.3	4.6	3.1	3.5	3.6	4.0	4.2	2.9	3.3	5.0	2.8	3.3	3.2	4.0
2004	3.8	3.3	3.3	4.5	3.2	3.9	4.2	3.8	3.9	2.9	3.0	4.5	3.0	3.7	3.2	3.7
2005	3.9	3.5	3.8	4.9	3.5	3.3	3.3	4.0	3.5	2.9	3.3	4.4	2.6	3.0	4.0	3.6
2006	3.4	3.2	2.9	3.8	2.9	3.3	4.1	3.4								
2007	3.4	3.0	2.7	3.9	3.1	3.5	3.7	3.4								

Norte, average firms size (nº employees)

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
empresas com 1 ano	8,0	8,0	7,6	7,2	7,6	6,4	5,8	7,1	5,3	5,3	5,4	5,6	5,0	5,0	5,3	4,7	5,1	4,8	5,1	4,5
empresas com 2 anos		9,6	9,6	8,8	8,0	8,3	7,3	6,4	8,2	6,1	5,9	6,0	6,4	5,7	5,6	5,6	5,1	5,6	5,6	5,9
empresas com 3 anos			10,9	10,4	9,3	8,3	9,1	7,6	7,4	8,6	6,8	6,8	6,7	7,1	6,4	5,9	6,0	5,6	6,0	6,1
empresas com 4 anos				12,2	11,0	9,8	9,5	10,0	8,2	7,6	9,2	7,5	7,5	7,3	7,4	6,8	6,2	6,3	5,9	6,6
empresas com 5 anos					13,1	10,9	10,6	10,2	10,7	8,9	7,9	10,1	7,9	8,2	7,6	7,7	6,9	6,8	6,7	6,2

Centro

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
empresas com 1 ano	6,5	6,3	6,4	6,6	6,5	5,2	6,1	5,2	4,2	4,6	4,3	4,1	4,5	3,8	4,4	3,7	4,0	3,9	4,8	3,6
empresas com 2 anos		7,5	7,6	7,3	7,2	7,0	5,8	6,8	5,8	4,7	5,3	4,9	4,7	5,3	4,2	4,6	4,1	4,3	4,6	5,3
empresas com 3 anos			8,6	8,3	7,9	7,9	7,2	6,2	7,2	6,6	5,3	5,9	5,4	5,4	5,8	4,4	4,8	4,4	4,7	4,8
empresas com 4 anos				9,1	9,0	8,2	7,9	7,8	6,7	8,1	7,2	5,8	6,7	6,1	5,8	6,0	4,7	5,1	4,6	5,2
empresas com 5 anos					9,7	9,1	8,7	8,0	8,3	7,3	8,8	7,8	6,4	7,4	6,3	5,8	6,3	4,9	5,2	5,0

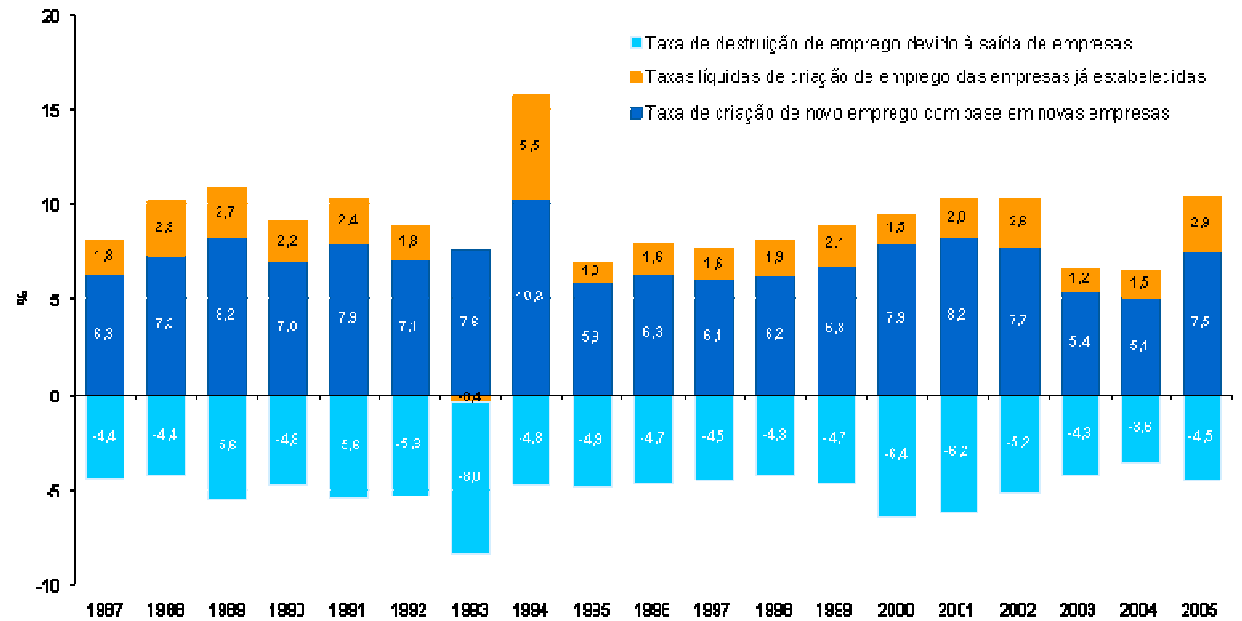# Regional dynamics: employment

Norte
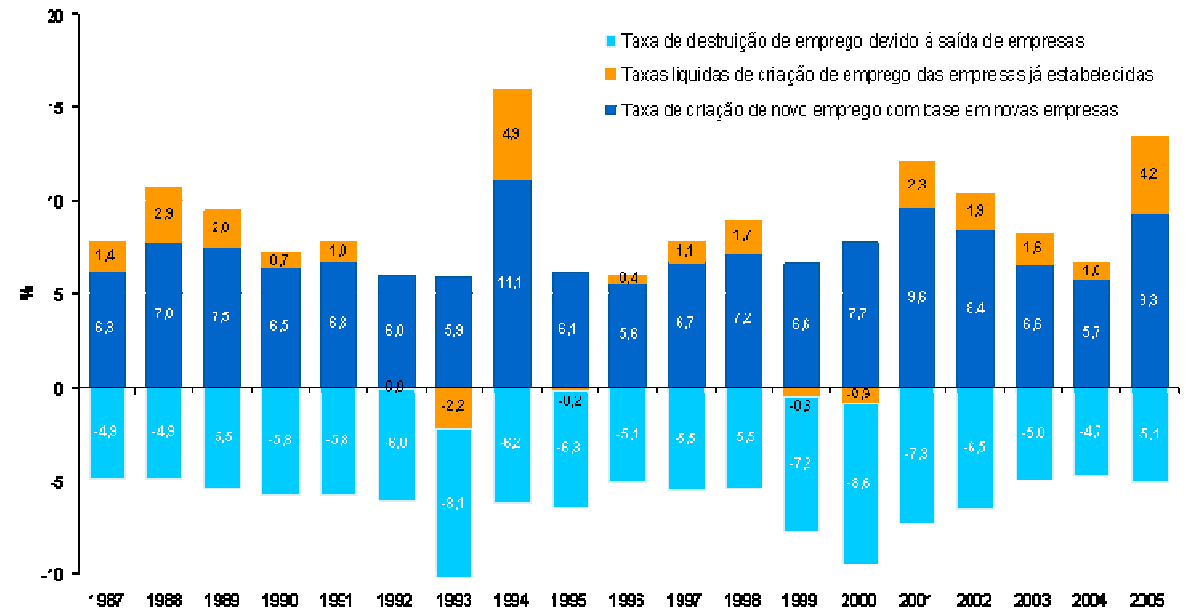
Centro

Non-parametric survival analysis

The survivor function reports the probability of a firm of surviving beyond time t (the moment of observation), that is the probability that there is no failure event (a "death") prior to t. The function is equal to 1 at time $t=0$ and decreases towards 0 as time (t) goes to infinity.
The survivor function is thus represented by:

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

The non-parametric Kaplan-Meier estimator was applied. For a dataset with observed failure times, $t_1, \ldots, t_k$, where k is the number of distinct failure times observed in the data, the **Kaplan-Meier** estimate at any time t is given by:

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

The most common estimator for the cumulative hazard rate is the non-parametric **Nelson-Aalen** estimator, which is defined by the sum of the instantaneous ratio of the failures over the number of enterprises at risk:

$$\overline{H}(t) = \sum_{j|t_j \le t} \frac{d_j}{n_j}$$

Life Table for Employer Enterprise Births, 1987-2005

Time	Observations	Deaths	Censured Observations	Kaplan-Meier		Nelson Aalen	
				Survivor Function	Failure Function	Hazard Rate	Cumulative Hazard Rate
Years	n°	n°	n°	%	%	%	%
				P(S)	100-P(S)	P(D)	$\sum P(D)$
1	451.041	63.088	24000*	86,0%	14,0%	14,0%	14,0%
2	364.233	46.351	22000*	75,1%	24,9%	10,9%	26,7%
3	295.786	32.973	28000*	66,7%	33,3%	8,4%	37,9%
4	235.002	23.655	24000*	60,0%	40,0%	6,7%	47,9%
5	187.102	17.353	19000*	54,4%	45,6%	5,6%	57,2%
6	150.840	12.966	12000*	49,7%	50,3%	4,7%	65,8%
7	125.525	10.059	11000*	45,8%	54,2%	4,0%	73,8%
8	104.121	7.735	9.613	42,4%	57,6%	3,4%	81,2%
9	86.773	6.089	7.943	39,4%	60,6%	3,0%	88,3%
10	72.741	5.068	7.491	36,6%	63,4%	2,8%	95,2%
11	60.182	4.172	11000*	34,1%	65,9%	2,5%	102,2%
12	45.130	3.037	6.150	31,8%	68,2%	2,3%	108,9%
13	35.943	2.422	5.626	29,7%	70,3%	2,2%	115,6%
14	27.895	1.681	5.546	27,9%	72,1%	1,8%	121,7%
15	20.668	1.133	4.733	26,4%	73,7%	1,5%	127,1%
16	14.802	805	5.361	24,9%	75,1%	1,4%	132,6%
17	8.636	490	4.418	23,5%	76,5%	1,4%	138,2%
18	3.728	228	3.500	22,1%	77,9%	1,4%	144,4%

Non-parametric survival analysis

In Portugal, the estimated median duration of a new born enterprise lies between **5 and 6** years.

Comparing with other countries (US, UK, Germany ,Italy, Spain), the maximum of the unconditional hazard function is reached before the sixth year, indicating that Portuguese firms keep on failing for a longer period, before the hazard rate starts declining.
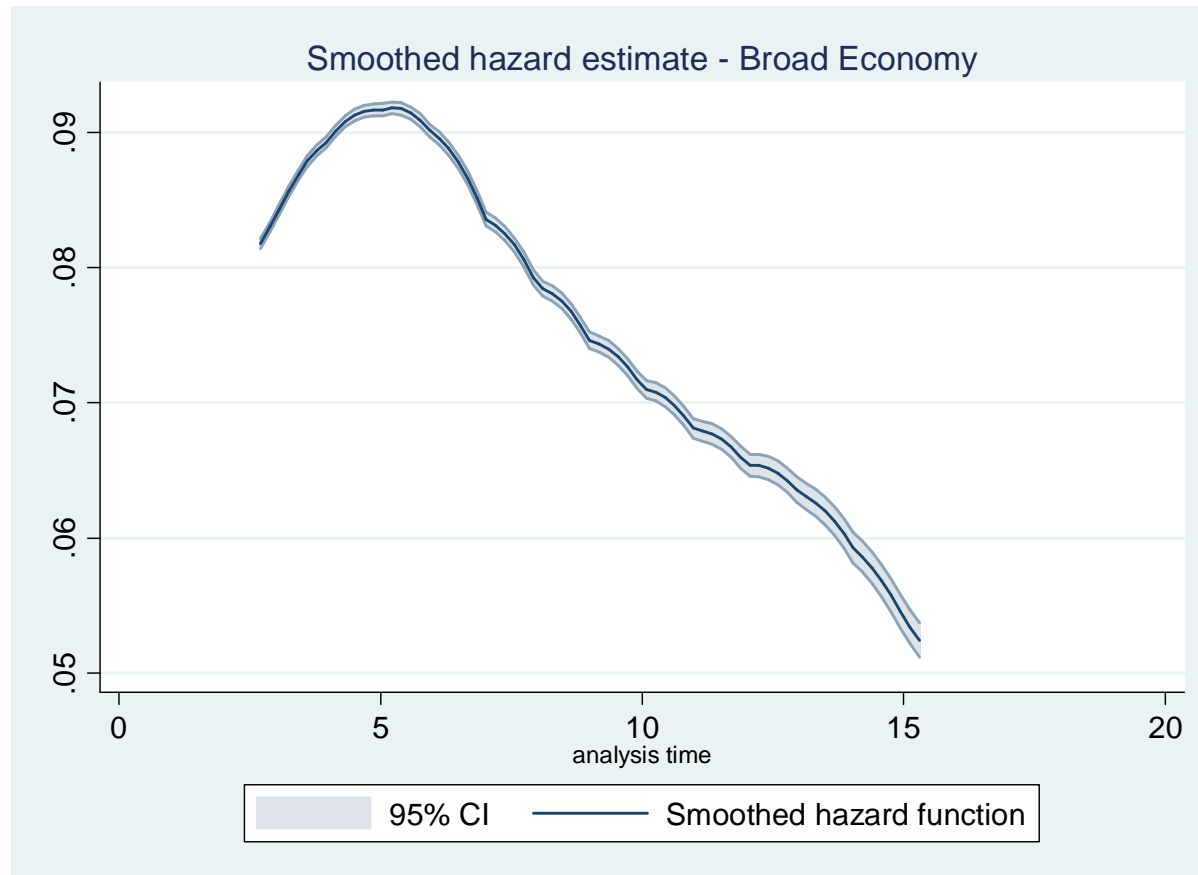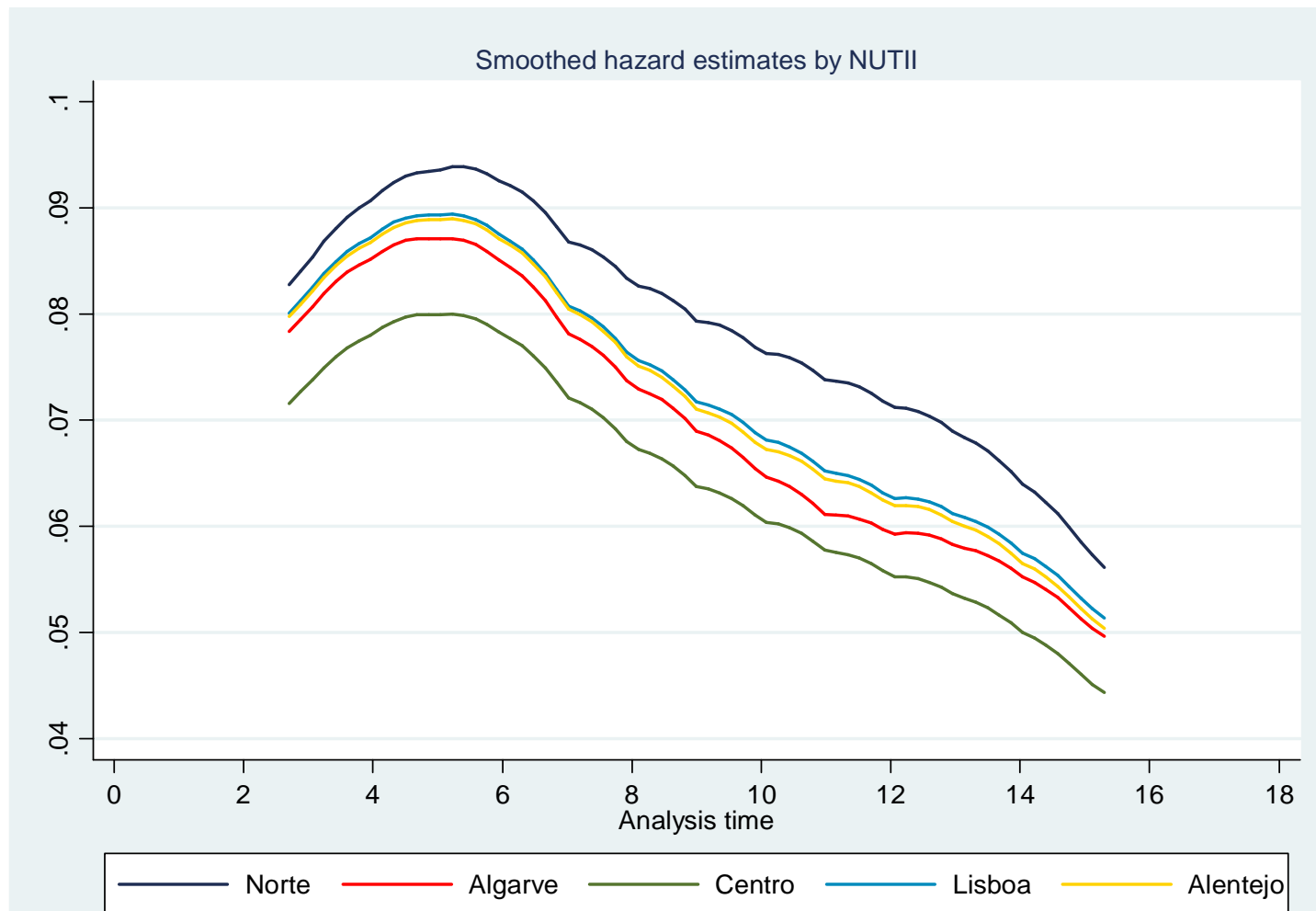
At the end of the analysis period, Norte is the region that presents the lowest survival rate, with only 20,7% of the firms' population managing to survive after eighteen years of activity. In Centro, 27,4% of active start-ups are still alive after 18 years. Centro has a higher survival rate than the economy's average, it is where more firms manage to survive longer.

Survival Table for Employer Enterprise Births by NUTII region, 1987-2005

Time	Norte	Centro	Lisboa e Vale do Tejo	Alentejo	Algarve	Açores	Madeira	
1	85,6%	87,4%	85,5%	85,8%	85,6%	85,1%	86,1%	86,0%
2	75,1%	77,7%	75,1%	75,5%	75,5%	74,2%	76,0%	75,1%
3	66,5%	70,1%	67,0%	67,0%	67,7%	67,0%	68,3%	66,7%
4	59,8%	64,0%	60,5%	60,4%	61,2%	59,9%	61,3%	60,0%
5	54,1%	58,9%	55,0%	54,9%	55,8%	54,5%	56,2%	54,4%
6	49,4%	54,5%	50,4%	50,7%	51,1%	50,5%	51,6%	49,7%
7	45,3%	50,7%	46,6%	46,9%	47,2%	46,7%	47,5%	45,8%
8	41,7%	47,5%	43,2%	43,4%	44,2%	43,7%	44,6%	42,4%
9	38,7%	44,5%	40,2%	40,5%	41,1%	41,2%	41,7%	39,4%
10	35,8%	41,9%	37,6%	37,7%	38,5%	38,9%	38,6%	36,6%
11	33,0%	39,5%	35,1%	35,2%	36,2%	36,3%	36,6%	34,1%
12	30,5%	37,4%	32,8%	33,0%	34,0%	33,9%	34,3%	31,8%
13	28,1%	35,3%	30,8%	31,0%	32,0%	31,3%	31,7%	29,7%
14	26,4%	33,4%	29,0%	29,3%	30,2%	29,4%	29,9%	27,9%
15	24,8%	31,8%	27,4%	27,8%	29,0%	28,2%	28,2%	26,4%
16	23,2%	30,4%	26,1%	26,2%	27,8%	26,4%	26,9%	24,9%
17	21,9%	28,9%	24,6%	24,9%	25,4%	25,4%	26,6%	23,5%
18	20,7%	27,4%	22,9%	23,2%	23,9%	23,8%	25,4%	22,1%

The disparities among the Portuguese regions are confirmed by equality tests. Both Log-rank and Wilcoxon (Breslow) tests allow for the rejection of the hypothesis of survival equality among regions.

Smoothed hazard estimates by NUTII

Non-parametric survival analysis

Firm survival rates of different enterprise cohorts in Norte

Survival years	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
1	87,3%	84,7%	86,8%	87,0%	85,7%	81,5%	87,3%	85,6%	85,9%	87,1%	88,3%	85,9%	79,6%	82,2%	87,1%	87,3%	86,7%	85,7%
2	75,9%	75,0%	77,4%	77,2%	72,7%	72,8%	77,9%	76,8%	76,2%	77,6%	77,1%	70,0%	66,6%	71,3%	77,5%	77,9%	76,9%	
3	67,6%	67,7%	69,3%	67,1%	65,6%	64,8%	70,7%	69,2%	68,7%	68,3%	64,1%	59,6%	58,4%	62,7%	69,7%	69,9%		
4	62,0%	61,4%	61,3%	61,2%	59,7%	59,4%	64,9%	63,9%	61,7%	57,5%	55,1%	53,0%	52,6%	56,6%	63,2%			
5	56,8%	54,6%	56,1%	56,4%	55,6%	55,0%	59,7%	57,9%	52,9%	50,2%	49,7%	48,0%	47,6%	51,4%				
6	51,7%	50,8%	52,1%	52,7%	51,9%	51,4%	54,9%	50,5%	46,6%	45,2%	45,6%	43,9%	43,4%					
7	48,0%	46,8%	48,8%	49,4%	48,5%	47,0%	48,5%	45,2%	43,0%	41,6%	41,7%	39,9%						
8	44,2%	43,9%	45,8%	46,1%	44,8%	41,1%	43,4%	41,6%	40,0%	38,4%	38,8%							
9	41,7%	41,3%	43,2%	42,5%	39,8%	37,2%	40,3%	38,9%	37,1%	35,5%								
10	39,3%	39,3%	39,5%	37,7%	36,0%	34,7%	38,0%	36,0%	34,4%									
11	36,9%	36,6%	35,1%	34,1%	33,5%	32,4%	35,4%	33,5%										
12	34,3%	32,7%	31,8%	31,8%	31,6%	30,3%	33,5%											
13	30,8%	29,6%	29,5%	29,8%	29,7%	28,0%												
14	28,5%	27,6%	27,8%	27,9%	28,0%													
15	26,7%	26,4%	26,1%	26,1%														
16	25,2%	24,6%	24,4%															
17	23,7%	23,3%																
18	22,4%																	

Centro

Anos de Sobrevivência	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
1	88,4%	85,7%	88,3%	87,2%	88,4%	83,3%	89,0%	88,6%	88,4%	88,3%	88,3%	88,1%	85,5%	85,9%	87,2%	88,4%	88,0%	86,9%
2	76,1%	76,2%	80,1%	78,6%	75,6%	74,1%	79,8%	80,2%	78,2%	78,8%	79,1%	76,3%	74,1%	76,6%	77,9%	79,7%	77,4%	
3	69,2%	69,7%	73,7%	67,7%	69,0%	67,3%	73,5%	73,3%	70,8%	70,5%	69,7%	67,6%	66,7%	69,2%	70,8%	72,3%		
4	63,0%	63,8%	64,9%	62,2%	63,8%	62,6%	68,6%	67,9%	64,0%	63,0%	62,4%	61,3%	60,8%	63,4%	64,5%			
5	58,7%	56,8%	60,7%	57,7%	59,4%	59,0%	63,9%	62,7%	58,8%	56,8%	56,9%	56,8%	55,8%	58,1%				
6	53,1%	52,7%	56,8%	53,8%	55,5%	55,1%	59,7%	57,6%	53,3%	52,1%	52,7%	52,9%	51,2%					
7	50,0%	49,0%	52,8%	50,8%	52,4%	51,2%	55,0%	52,8%	49,6%	48,7%	49,1%	48,9%						
8	47,1%	46,7%	49,8%	48,4%	49,2%	47,3%	50,9%	49,6%	46,8%	45,6%	45,0%							
9	44,2%	44,1%	47,1%	45,6%	45,5%	43,8%	48,3%	46,4%	44,1%	42,8%								
10	41,8%	41,2%	44,3%	43,3%	42,3%	41,4%	45,5%	43,8%	41,3%									
11	39,7%	38,9%	41,2%	40,9%	40,1%	39,3%	43,2%	40,9%										
12	37,8%	36,7%	38,7%	38,2%	37,9%	37,6%	41,3%											
13	35,1%	34,2%	36,6%	36,4%	36,1%	35,7%												
14	32,2%	32,5%	34,8%	34,7%	34,3%													
15	30,2%	30,9%	33,4%	33,4%														
16	28,7%	29,7%	31,9%															
17	27,4%	28,1%																
18	26,0%																	

Semi-parametric survival analysis

- The statistical representation of the relation between the survival time of a firm and specific variables is known as the hazard rate model of the duration of the life of a firm.

- According to the model a given firm j faces a hazard rate h_j, that is a function of a baseline hazard rate h_0, which all firms face, transformed by a set of explanatory variables (X) through a vector of parameters (β). In this model, two firms with the same birth date will face a different hazard function if, and only if, their other characteristics are different. This model seems a natural solution to understand the temporal pattern of survival and to identify the covariates that could be related significantly to survival, being also appropriate for working with longitudinal datasets, characterized by right censored data.

- The complementary log-log (cloglog) model incorporates the main features of discrete duration models, where the logarithm of the probability that a firm exits at time t, given that it survived in t-1, is explained by a series of explanatory covariates X_t, plus a set of parameters identifying the baseline hazard function, according to the following specification :

$$\log h\left(t \mid x_t, x_0\right) = \lambda_t + \beta x_t + \gamma x_0, \text{ para } t = 1, ..., k$$

Explanatory Variables Considered in the Model

Variable	Definition	Measurement
Start-up Size	Number of employees at the birth year of the firm.	Logarithm of the number of employees.
Current Size	Number of employees at the current year.	Logarithm of the number of employees.
Industry Entry Rate	Industry entry rate calculated for sectors defined at a 2-digit CAE level.	Logarithm of the industry entry rate, defined as the number of entrants divided by the total number of firms in industry.
Concentration (HHI)	Herfindhal-Hirschman Index (HHI) calculated for industries at a 2-digit CAE level.	Logarithm of the HHI.
Growth	Logarithmic difference of industry employment in two consecutive periods.	Logarithm of the number of employees at year t minus the logarithm of the number of employees at year $t-1$.
Entry Rate x Growth	Interaction variable, defined as the product of entry and growth.	Product of logarithms.
Turbulence	Sum of entry and exit rates calculated for sectors defined at a 2-digit CAE level.	Sum of logarithms of the industry entry rate with the industry exit rate.
Sector Dummies	Dummies for 4 broad sectors: Agriculture, Construction, Manufacturing and Services.	
Regional Dummies	Dummies for 7 NUTII regions: Norte, Centro, Lisboa e Vale do tejo, Alentejo, Algarve, Açores and Madeira	
Year Dummies	Dummies for each current year.	

Variable	Model (1)	Model (2)
Log of Start-up Size	0.692 *** (0.002)	1.339 *** (0.007)
Log of Current Size	----	0.588*** (0.002)
Industry (2 digit) Start-up entry rate	1.353 *** (0.013)	1.42*** (0.018)
Start-up Industry HHI (2 digit)	0.986 *** (0.001)	0.881*** (0.001)
Industry Growth (log)	----	1.326*** (0.018)
Growth x Entry rate	----	1.966** (0.017)
Turbulence	----	4.195 *** (0.193)
Sector Dummies Agriculture	(a)	(a)
Construction	1.130 *** (0.012)	1.405 *** (0.021)
Manufacturing	1.123 *** (0.013)	1.567 *** (0.024)
Services	0.951 *** (0.010)	1.115*** (0.016)5
Year Dummies	Yes	Yes
Number of firms	447772	447772
LR X2	42744.42 ***	46506.55 ***
Log likelihood	-655716.05	-423334.96

The values presented are the hazard ratios, when the variable increases by one unit.

hazard ratio >1: an increase in the given explanatory variable increases the probability of exit.

a hazard ratio <1: an increase in the variable decreases the hazard probability.

The estimation is made for the period 1995 to 2007.

Variables	Portugal	Regions						
		Norte	Centro	Lisboa	Alentejo	Algarve	Açores	Madeira
Log of start-up size	1.334 ***	1.310 ***	1.342 ***	1.414 ***	1.250 ***	1.286 ***	1.139 ***	1.267 ***
	(0.007)	(0.011)	(0.016)	(0.014)	(0.023)	(0.028)	(0.041)	(0.044)
Log of current size	0.459 ***	0.479 ***	0.418 ***	0.463 ***	0.451 ***	0.444 ***	0.501 ***	0.440 ***
	(0.002)	(0.004)	(0.005)	(0.004)	(0.008)	(0.009)	(0.018)	(0.014)
Industry (2 digit) start-up entry rate	1.24 ***	1.233 ***	1.181 ***	1.160 ***	1.515 ***	1.290 ***	1.429 ***	1.613 ***
	(0.018)	(0.028)	(0.038)	(0.032)	(0.077)	(0.08)	(0.152)	(0.164)
Start-up industry HHI (2 digit)	0.988 ***	0.986 ***	0.986 ***	0.992 ***	0.988 ***	0.997	1004.0	0.975 ***
	(0.001)	(0.002)	(0.003)	(0.003)	(0.005)	(0.006)	(0.011)	(0.009)
Turbulence rate	5.29 ***	11.444 ***	3.664 ***	3.406 ***	1.698 ***	3.574 ***	1.285	14.441 ***
	(0.222)	(0.791)	(0.344)	(0.336)	(0.22)	(0.699)	(0.341)	(4.792)
Industry Growth (log)	1.122 ***	1.087 ***	1.140 ***	1.159 ***	1.144 ***	1.113 *	1.326 ***	1.036
	(0.018)	(0.029)	(0.044)	(0.036)	(0.065)	(0.07)	(0.146)	(0.104)
Growth x Entry rate	1.082 ***	1.069 ***	1.116 ***	1.091 ***	1.083	1.015	1.272 ***	1.033
	(0.017)	(0.028)	(0.042)	(0.033)	(0.061)	(0.065)	(0.141)	(0.109)
Sector dummies								
Agriculture	0.612 ***	0.443 ***	0.706 ***	0.643 ***	0.795 ***	0.823 *	1.178 *	0.841
	(0.01)	(0.016)	(0.024)	(0.033)	(0.034)	(0.069)	(0.11)	(0.156)
Construction	0.895 ***	0.858 ***	0.931 ***	0.919 ***	1.155 ***	1.013	1.742 ***	1.343 ***
	(0.009)	(0.012)	(0.021)	(0.023)	(0.049)	(0.616)	(0.18)	(0.116)
Manufacturing	(a)	(a)	(a)	(a)	(a)	(a)	(a)	(a)
Services	0.705 ***	0.631 ***	0.770 ***	0.799 ***	0.884 ***	0.869 ***	0.905	0.792 ***
	(0.006)	(0.007)	(0.015)	(0.017)	(0.032)	(0.048)	(0.077)	(0.062)
Regional dummies								
Norte	(a)	-----	-----	-----	-----	-----	-----	-----
Centro	0.847 ***	-----	-----	-----	-----	-----	-----	-----
	(0.006)							
Lisboa	1.027 ***	-----	-----	-----	-----	-----	-----	-----
	(0.007)							
Alentejo	0.926 ***	-----	-----	-----	-----	-----	-----	-----
	(0.01)							
Algarve	0.939 ***	-----	-----	-----	-----	-----	-----	-----
	(0.011)							
Açores	0.942 ***	-----	-----	-----	-----	-----	-----	-----
	(0.02)							
Madeira	1.014	-----	-----	-----	-----	-----	-----	-----
	(0.019)							
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of firms	447772	164599	97606	109405	33692	25802	7523	9140
LR X2 (34) / LRX2 (28) for regions	47329.9 ***	17871.03 ***	10705.95 ***	11192.01 ***	369.5.93 ***	2698.05 ***	860.98 ***	142147 ***
Log likelihood	-422915.7	-152494.34	-91424.886	-106768.52	-33112.584	-23204.949	-6.744.716	-8283.6

hazard ratio >1: an increase in the given explanatory variable increases the probability of exit.

The correlation between turbulence (sum of the entry and exit rates) and growth rate is indeed positive (58%) and statistically significant at 5% confidence level, corroborating our previous argument.

	Log of Start-up Size	Log of Current Size	Industry (2digit) start–up entry rate	Start-up Industry HHI (2 digit)	Log of Industry Growth	Growth X Entry rate	Turbulence
Log of Start-up Size	1						
Log of Current Size	0.8253*	1					
Industry (2digit) start–up entry rate	0.0570*	0.1036*	1				
Start-up Industry HHI (2 digit)	0.1122*	0.1417*	0.6323*	1			
Industry Growth (log)	-0.0523*	-0.0409*	0.3552*	0.2619*	1		
Growth x Entry rate	---	---	0.0044*	---	-0.0383*	1	
Turbulence	-0.0268*	-0.0571*	0.5349*	0.3057*	0.5797*	---	1

Note: * refers to the correlations coefficients with 5% statistical significance.

- The instantaneous probability of exit is monotonically decreasing with firm size. After entry, the conditional probability of failure increases continuously up to the sixth year of activity. Breaking down by region, sector and class dimension, we identify statistically significant disparities.

- We find that firms that start small and experience faster post-entry growth, face a higher probability of survival. Firm's current size dimension is extremely important to determine the probability of survival, particularly in the Services sector.

- Firms that experience more competition from entrants, also face higher probabilities of failure.

- A higher entry rate combined with fast growth rates for any given industry also generates a shorter duration of firms. It might seem easier to enter the market in earlier stages of the product life-cycle, when markets are expanding, but it becomes particularly difficult to survive.31

- Regions such as Norte, Algarve and Madeira show the highest growth rates in enterprise births and deaths but also of firm churn. After 2000, a new tendency seems to emerge, where both births and deaths of employer enterprises are relatively less turbulent. This is verified for all regions, sectors and size classes.

- In the non-parametric analysis for Portugal, we find that around 25% of enterprises entering the market fail within the first 2 years of activity and that more than 50% fail within a period of six years. Breaking down the analysis, we identify statistically significant disparities among regions. Norte has the lowest survival rate and Centro is where firms manage to survive longer. The survival gap between these two regions gets amplified over time.

- However, we find a different result from the literature, for the effect of industry growth in survival rates. Firms operating in industries which are growing faster seem to suffer from a higher probability of failure. The combined effect of entry and growth can also contribute to explaining this unexpected effect of industry growth on survival probabilities.

- This has to do with turbulence and the high rates of entry and exit verified in Portugal throughout this period. Industry growth, by sector, reinforces this view, and the interaction between high industry growth rates and entry rates at the start-up moment, seems to unfold the reasons for the negative impact on firms´ survival.

- By correcting heterogeneity, we obtain stronger magnitudes of the hazard ratios.