The New Approach to Estimation of the Hazard Function in Business Demography on Example of Data from New Zealand

Zając, Paweł

AGH University of Science and Technology

2013

Paweł Zając*

The New Approach to Estimation of the Hazard Function in Business Demography on Example of Data from New Zealand

1. Introduction

In the recent years the subject of new-firm survival has become very popular among academics all over the world. There are multiple reasons behind those studies. The most common motivation for researchers is direct link between number of enterprises, unemployment and job growth. Moreover, studies about the survival of new enterprises are related to the bankruptcy and its social as well as economic consequences ([1], [11]). On the other hand according to the concept of creative destruction suggested at the beginning of the last century by Joseph Schumpeter [18] new born enterprises are increasing competition on the market and by innovative ideas improve efficiency of environment and stimulate economy to growth. Finally, the Organisation for Economic Co-operation and Development (OECD) in some recent analyzes done by Foster, Haltinwager and Krizan [8] found that business demography explains up to 30% of the productivity growth.

Literature includes a lot of studies about the survival of new enterprises. Among factors which most often affect survival rates are age ([7], [9]), industry ([12], [13]), location ([2], [10]), size ([3], [4], [20]), capital and firm's individual characteristics ([5], [19]).

Researchers agree that initial firm size is important determinant of survival rates (e.g. [6], [21]). Smaller firms die more often than big ones. Among reasons behind that theory those which are the most frequently pointed are superior management in large enterprises and the fact that large firms most often have

* AGH University of Science and Technology, Department of Applications of Mathematics in Economics, Krakow, Poland, e-mail: pzajac@zarz.agh.edu.pl

bigger budget that can be used to survive tough periods of economic crisis. Nunes and Sarmento [17] suggest that there is a relation between bigger start up and expected high profits. The reduction of employment could also be used as a tool to survive for large enterprises. In addition, the limited amount of available capital could lead to insufficient investments in small enterprises.

There is a discussion between academics about the best measurements of enterprise's size. Mata, Portugal and Guimarães [15] claim that current size should be used in analyzes. Moreover the fact that firm has grown in the past suggests that this firm is in a good shape and probability of firm's death is small. On the other hand if a firm's size is decreasing that may indicate financial problems and a relatively high probability of death. Although others like López-Garcia and Puente [14] argue that a decrease in the size of an enterprise is more like a part of process of death than just a change in the probability group to which the firm belongs and points at start up size as a better statistic than current size. An enterprise can eventually increase the size using some financial credit which could in the end be the reason of bankruptcy.

The main purpose of this paper is to analyze and build a new mathematical model describing long term survival rates for new enterprises. As a main determinant in this study current enterprise size measured by number of employees is taken. For this purpose the author makes use of data from Statistics New Zealand [16] and estimates the hazard functions using simulations based on the stochastic process.

The paper is structured as follows. The next section presents dataset from Statistics New Zealand used in further simulation. The third section presents the non parametric analysis of survival rates. In the fourth section Financial Income Regime Model is presented. The fifth section describes the Monte Carlo simulation and prediction of the hazard function. The final section provides conclusions.

2. The dataset

The dataset used for simulations in this paper was provided as a business demography statistics by Statistics New Zealand Tatauranga Aotearoa (http://www.stats.govt.nz). It is a government's department and the New Zealand's national statistical office. In order to fulfill international standards in this area the methods used to identify business births and deaths by Statistics New Zealand are align well with practice study and recommendations from Eurostat and the OECD. The OECD study on business start-up rates found that enterprise birth rates are considered key economic indicators. Business births and deaths are defined by criteria based on a combination of factors of production (land, labour, capital). The Enterprise which assembly new factors of production is treated as new, disassembled of factors of production is equivalent to enterprise's death.

Statistics include data of nearly 500 000 enterprises born in the period between 2001–2010. Only economically significant enterprises are taken into study. According to Statistics New Zealand they all meet at least one of the following criteria:

– annual expenses or sales subject to GST of more than $30 000,
– 12-month rolling mean employee count of greater than three,
– part of a group of enterprises,
– registered for GST and involved in agriculture or forestry,
– over $40,000 of income recorded in the IR10 annual tax return (this includes some units in residential property leasing and rental).

For this paper study, it is very important that all other enterprises are excluded from statistics. In further research of this article enterprises are split into groups defined by number of employees. In the data from Statistics New Zealand the employee count is sourced from the Employer's Monthly Schedule tax form. Special awareness is needed to group with no employees. This count size category may have working owners, work provided by other businesses or contractors and business activity that requires no work (example: passive investment). Table 1 holds information about the amount of new born enterprises in New Zealand.

Table 1

Number of enterprise births in years 2001–2010 by employee count size group.

Reference period	Employee count size group							
	0	1–5	6–9	10–19	20–49	50–99	100 or more	Total births
2001	40 066	3 575	533	352	171	31	8	44 736
2002	35 844	3 115	503	329	154	25	7	39 977
2003	43 704	6 895	712	455	128	15	10	51 919
2004	60 588	7 309	736	440	169	26	12	69 280
2005	56 424	7 294	702	448	170	27	8	65 073
2006	55 464	6 991	663	366	145	21	10	63 660
2007	54 618	7 228	637	395	128	22	7	63 035
2008	53 390	6 999	618	366	136	27	12	61 548
2009	49 043	5 635	495	298	100	10	7	55 588
2010	41 272	4 931	410	309	101	15	5	47 043

Source: New Zealand Business Demography Statistics: At February 2011 and 2012

Only about 0.1% of all new born enterprises have more than 50 employees. The Monte Carlo simulation of such a small group of individuals most probably would generate high errors and would lead to false implications. That's the main reason why the author decided to focus on enterprises with less than 50 employees, actually that group is known in literature as small and medium establishments (SAMEs). In this paper all enterprises from SAMEs are segregated to clusters with 0, 1–5, 6–9, 10–19 and 20–49 employees.

3. Non parametric analysis

The problem addressed by this paper is to conduct research on the "lifetime" of enterprises. In classical non-parametric approaches researchers work on a group of n enterprises. Moreover, T_1, ..., T_n are positive independent and identically distributed variables and they represent time between the birth and death of each enterprise (lifetime). The following notation is used:

– cumulative distribution function:

$$F(t) = Prob(T_i \leq t) \tag{1}$$

– density function:

$$f(t) = \frac{dF(t)}{dt} \tag{2}$$

– survival function:

$$S(t) = 1 - F(t) = Prob(T_i > t) \tag{3}$$

– hazard function:

$$h(t) = \frac{f(t)}{S(t)} \tag{4}$$

In practice $F(t)$ represents the probability that enterprise will die until time t, the survival function $S(t)$ is the probability of survival past time t. In this paper study is made on discrete (annual) dataset thus as by $f(t)$ we understand the probability of death at the time t. Although, the most often conditional probability is requested, probability that an enterprise will die at time t, under the condition that it still exists in $t - 1$ is called the hazard function $h(t)$. The hazard function is the ratio of density and survival function.

The most commonly used estimator of the survival function $S(t)$ is simple frequency non parametric Kaplan–Meier estimator. It is defined as:

$$\hat{S}(t) = \prod_{i|t_i \leq t}\left(1 - \frac{d_i}{n_i}\right) \tag{5}$$

where n_i is the number of existing enterprises at time t_i and d_i is the number of deaths at time t_i.

In the dataset from New Zealand we got information about the survival rate of enterprises. In our calculations we will separate them into groups by employees count size. Figure 1 presents survival rates of enterprises born in the year 2002 by employee count size. Unfortunately, only approximate numbers are available.
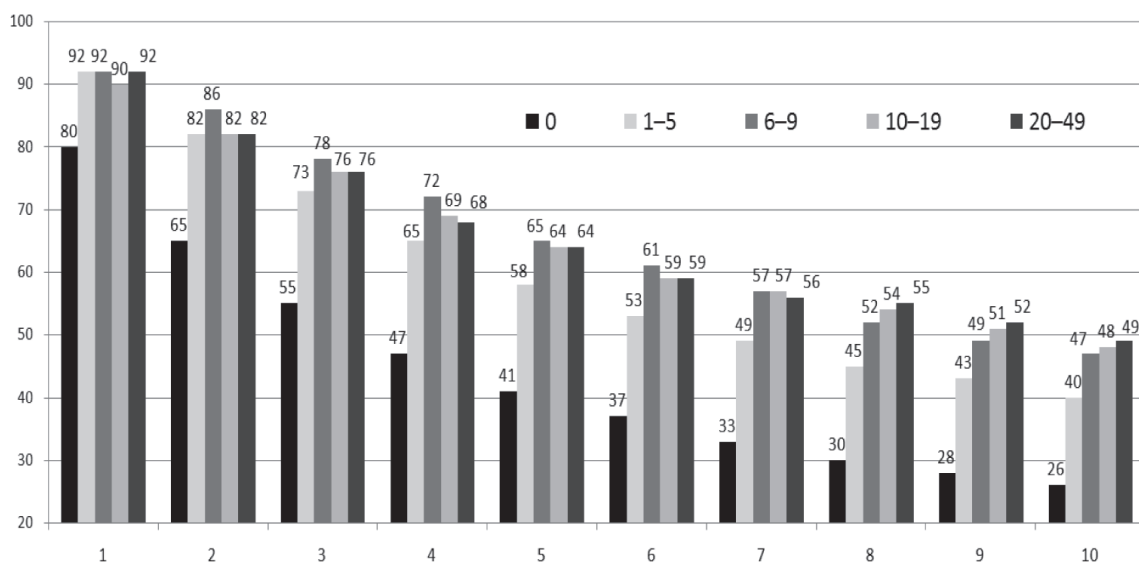

Figure 1. Survival rate of enterprises born in 2002 by employee count size group.
Source: New Zealand Business Demography Statistics, At February 2012.

While studying the figure above it is worth noting that the number of employees is an important factor in the survival rate analyses. After 10 years almost half of enterprises with more than 10 employees and only one from four enterprises with no employees still exists. Firms with employees have increased chance to survive.

The author decides to calculate one hazard function for each cluster. Hazard functions from years 2001–2010 are calculated and in order to produce final hazard functions their weighed combinations are taken. The amount of new

born enterprises from each year and cluster are taken as weights. Results are presented in Figure 2.
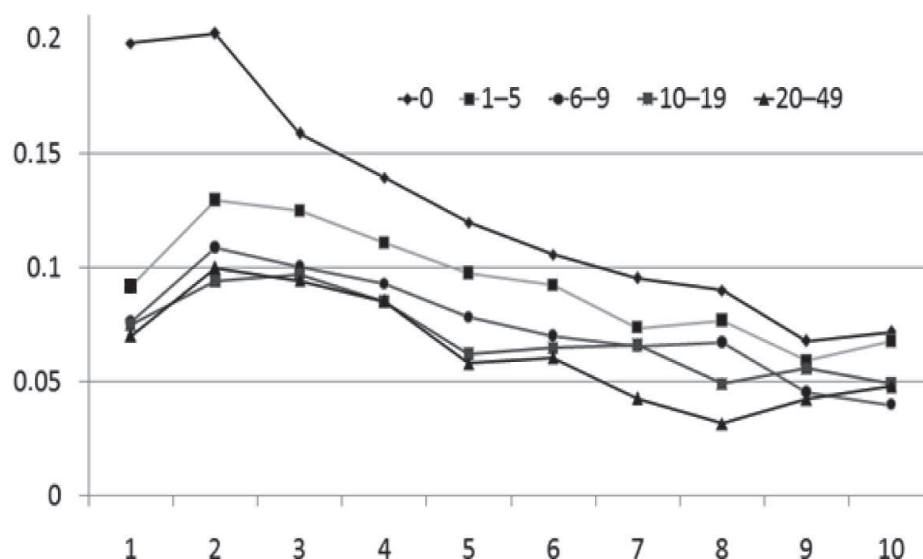
Figure 2. Combined hazard function for enterprises born in years 2001–2010.

The probability of surviving another year also seems to depend on the number of employees. Despite the age smaller enterprises are more likely to die. But differences in probabilities of death are decreasing with time.

4. The Financial Income Regime Model

The main goal of Financial Income Regime Model (FIRM) is prediction of the hazard function. In particular the author believes that evolution of the financial situation of enterprise could be described as stochastic process, which is, in some aspect, similar to first-order autoregression. The most basic postulation of FIRM is that each enterprise is made to generate income. Since there are no simple and general relations between income in current and past periods the model assumes that the available income in every period is random and independent from time and environment. Similar to previous section there are n enterprises. X_t^i are independent variables which represent the financial condition of an enterprise i in time t, $i = 1, ..., n$. It is important to mention that financial assets not necessarily are non-negative as enterprises can have debts. An enterprise dies only when

certain level of debt is reached. The financial condition of an enterprise can be expressed using recursive algorithm:

$$\begin{cases} X_0^i = 0 \\ X_t^i = \alpha X_{t-1}^i + \varepsilon_t^i, & \text{when} \quad X_{t-1}^i \geq 0 \\ X_t^i = \beta X_{t-1}^i + \varepsilon_t^i, & \text{when} \quad X_{t-1}^i < 0 \end{cases} \tag{6}$$

with

$$\varepsilon_t^i \sim N\left(\mu, \sigma^2\right). \tag{7}$$

As t we understand time expressed with positive integer values. The model above contains four indicators α, β, μ, σ. The parameter α represents the savings, of an enterprise, this parameter is used in a situation when a firm has some financial reserve, and describes what part of that reserve would be kept to the next time period. On the other hand, the parameter β can be described as information how enterprises handle their debts. Two regimes are separated according to the financial situation of a firm, change in regime depends only on financial conditions in the most recent time period. Although, the recursive structure of the model implies that X_t^i depends on the full history. Parameters μ and σ are mean and standard deviation of random income in a single time period. They are taken as constant and independent from time. The construction of definition (6) implies that X_t^i behave like a random walk with two additional leverages. Even though, it is a simplification zero seems to be the natural starting point for the financial state of a new born enterprise. In order to make the model as easy as possible we name the level of debts that kills the enterprise as -1. That means that when $X_t^i < -1$, the enterprise dies.

5. Simulations

In this section a Monte Carlo simulation is applied to the recursive formula (6). In simulation the author works with 250000 independent enterprises. Moreover, for each enterprise there is generated and set a series of random values Z_t^i from standard normal distribution using random number generator implied in R-program as function rnorm(). Starting value for financial condition X_t^i for each enterprise is set to zero. The values of random incomes ε_t^i are obtained from Z_t^i using formula:

$$\varepsilon_t^i = \left(Z_t^i + \mu\right) \cdot \sigma. \tag{8}$$

As it is possible to relate the standard normal to all normal random variables we get:

$$Z_t^i \sim N(0,1) \cdot \sigma \Rightarrow \varepsilon_t^i \sim N(\mu, \sigma^2). \tag{9}$$

After setting parameters the hazard function can be simulated.

5.1. Goodness of fit criterion

Clearly the set of (α, β, μ, σ) determine fitted values. In this study we take the assumption that

$$\begin{cases} \alpha \in <0,1> \\ \beta \in <0,2> \\ \mu \in <-1,1> \\ \sigma \in <0,2> \end{cases} \tag{10}$$

All possible combinations are checked with precision to 0,01. More precise calculations in author's opinion are not necessary because of rounding errors in real data from Statistics New Zealand. The method of least squares is taken as a measurement of goodness of fit. The procedure is applied to enterprises divided on clusters. For each cluster coefficients of determination are also computed. The results are presented in Table 2.

Table 2

Estimated FIRM parameters and coefficients of determination for fitted hazard functions according to numbers of employee.

	0	1–5	6–9	10–19	20–49
α	0.85	0.93	0.99	0.85	0.96
β	1.95	1.16	1.04	1.08	1.075
μ	0.08	−0.01	0	0.11	0.09
σ	1.3	0.72	0.68	0.74	0.72
R^2	0.974022	0.948772	0.935394	0.871511	0.891479

Source: own calculations

Estimated values of α are mixed and there is no straight dependency between the size of enterprise and consumption of savings. Although estimated values of β are strongly decreasing with size of enterprise.

According to the results in table 2 the best values are fitted to hazard function for enterprises with no employees. The coefficient of determination is about 97%. The worst result in the simulation of the hazard function is achieved for enterprises with 10 to 19 employees with coefficient of determination about 87%. Although this value is still significant and shows that this approach is very suitable to real data. Figure 3 shows real and fitted hazard functions in the best and in the worse situation.
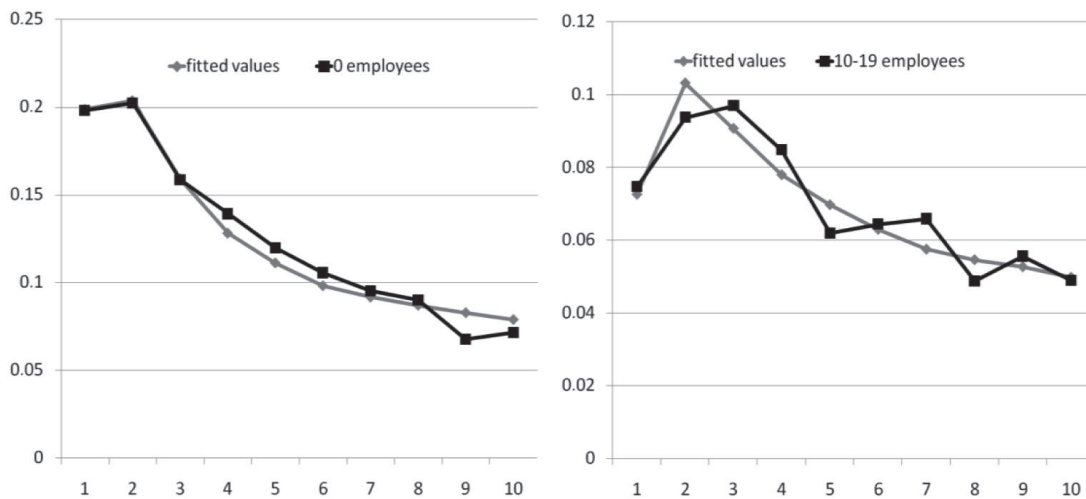

Figure 3. The hazard functions and fitted values using FIRM.

Source: own calculations

It can be observed that in general FIRM works better for small enterprises. Nevertheless, it can be explained by the fact that the amount of enterprises with no employees is about 6 times bigger than amount of all enterprises with at least 1 worker. In fact if the lifetime of enterprise behaves like the random independent stochastic process, then implications based on too small a sample could be misleading.

5.2. Long term simulation

The computational implementation of FIRM gives the possibility to simulate the hazard function for further years. Simulation like this can be also treated as a test for the method. Figure 4 illustrates predictions of basic functions used in the non-parametric analysis of a lifetime for enterprises with no employees.
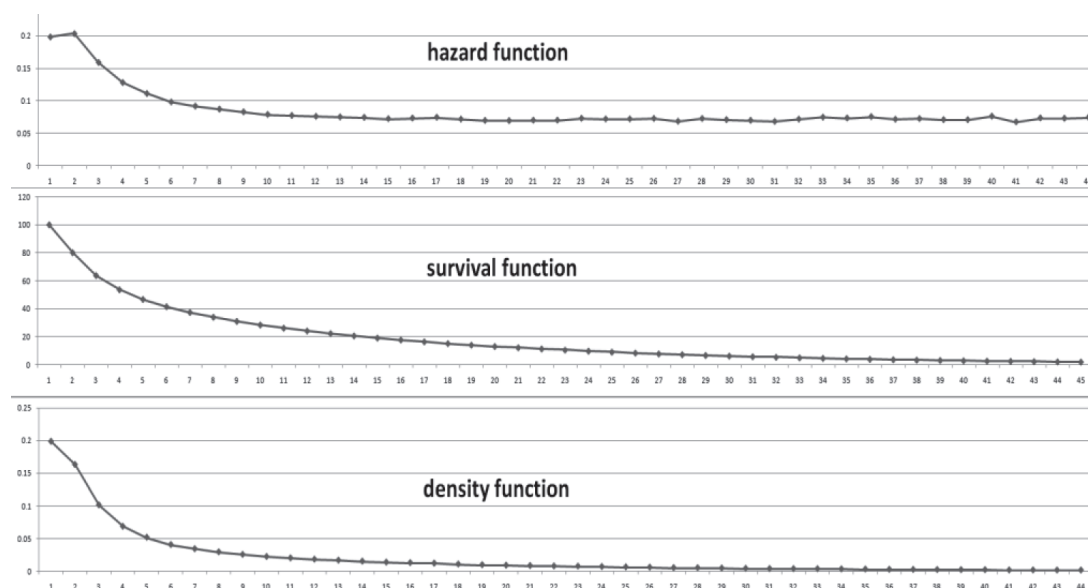
Figure 4. The hazard function, survival function and density of lifetime for enterprises with no employees.

Source: own calculations

Some important implications can be derived from this prediction. The hazard function after rapid changes during first years of an enterprise's existence with time become stable. For enterprises with no employees the hazard function stabilizes at level 7.2%. The probability of the death of the enterprise becomes independent from age at a certain point. This conclusion is in agreement with common sense, it is easily acceptable that the probability of survival of "old" enough enterprises should be the same. Similar effects are observed for other clusters of enterprises in New Zealand. Moreover, the FIRM gives the possibility for approximation of mean, standard deviation for the lifetime of enterprises in clusters. Results are presented in Table 3.

Table 3

Approximation of lifetime summaries for enterprises according to numbers of employee.

	0	**1–5**	**6–9**	**10–19**	**20–49**
Mean	8.7	14.4	17.1	18.2	34.6
Standard deviation	11.5	18.8	20.1	21.4	49.2
Stabilization level for the hazard function	7.2%	4.4%	3.3%	4.28%	1.7%

Source: own calculations

Nevertheless, the probability of death of enterprises becomes stable with time, it can be observed that large enterprises are much less likely to die.

6. Conclusions

The Simulation study confirms that the FIRM can be used as a tool to estimate the hazard function for new born enterprises. Evaluated parameters and implications from the model are in agreement with economic theory. The empirical results confirm that the size of an enterprise measured by the number of employees is a very important determinant of the probability of survival. Bigger firms usually can survive longer then small units. Moreover, the study indicates that the probability of death of the enterprise is decreases with age and finally after about 15 years becomes constant in time.

More research is needed to investigate the implicit formula of the hazard function based on the theoretical distribution of the stochastic process of X_t^i.

Some of results of this study could be used by policy-makers to help small and medium establishments. Acceptance that firms' yearly income is random implies that enterprises should have some financial help from government when they experience market turbulence. It would have direct impact on a firm's survival. Thanks to that help a lot of enterprises could get back to business with new experiences and in future pay back received help. This suggest that not only recently popular policies designed to promote new enterprises are important. The promotion of new firms should be accompanied by a system which helps already existing enterprises in reducing unemployment and aiming at economic growth.

References

[1] Acs Z., Audretsch D., *Small Firms and Entrepreneurship: An East-West Perspective*, Cambridge University Press, 1993.

[2] Acs Z., Armington C., *The Impact of Geographic Differences in Human Capital on Service Firm formation Rates*, "Journal of Urban Economics" 2004, vol. 56, pp. 244–278.

[3] Agarwal R., Audretsch D.B., *Does Entry Size Matter?: The Impact of the Life Cycle and Technology on Firm Survival*, "Journal of Industrial Economics" 2001, vol. 49, pp. 21–43.

[4] Angelini P., Generale A., *On the Evolution of Firm Size Distributions*, "American Economic Review" 2008, vol. 98(1), pp. 426–438.

[5] Astebro T., Bernhardt I., *The Winner's Curse of Human Capital*, "Small Business Economics" 2005, vol. 24, pp. 63–78.

[6] Audretsch D.B., Mahmoon T., *New Firm Survival: New Results Using a Hazard Function*, "Review of Economics and Statistics" 1995, Vol. 77(1), pp. 97–103.

[7] Caves R.E., *Industrial organization and new findings on the turnover and mobility of firms*, "Journal of Economic Literature" 1998, vol. 36(4), pp. 1947–1982.

[8] Foster L., Haltinwager J.C., Krizan C.J., *Aggregate Productivity Growth. Lessons from Microeconomic Evidence*, in: *New Developments in Productivity Analysis*, University of Chicago Press 2001, pp. 303–372.

[9] Geroski P., *What do we Know About Entry*, "International Journal of Industrial Organization" 1995, vol. 13(4), pp. 413–614.

[10] Gilbert B.A., McDougall P.P., Audretsch D.B., *New Venture Growth: A review and extension*, "Journal of Management" 2006, vol. 32, pp. 926–950.

[11] Gurgul H., Zając P., *The dynamic model of birth and death of enterprises*, "Statistics in Transition: new series" 2011, vol. 12, pp. 381–400.

[12] Jovanovic B., MacDonald G.M., *The life cycle of a competitive industry*, "Journal of Political Economy" 1994, vol. 102, pp. 322–347.

[13] Lin P., Huang D., *Technological Regimes and Firm Survival: Evidence Across Sectors and Over Time*, "Small Business Economics" 2008, vol. 30, pp. 175–186.

[14] López-Garcia P., Puente S., *Business Demography in Spain: Determinants of Firm Survival*, "Documentos de Trabajo" 2006, vol. 608, Banco de España.

[15] Mata J., Portugal P., Guimaraes P., *The Survival of New Plants: Start-up Conditions and Post-entry Evolution*, "International Journal of Industrial Organization" 1995, vol. 13(4), pp. 459–481.

[16] *New Zealand Business Demography Statistics: At February 2012*, Statistics New Zealand Tatauranga Aotearoa, 2012.

[17] Nunes A., Sarmento E., *Business Demography Dynamics In Portugal: A Non-Parametric Survival Analysis*, GEE Papers 0022, Gabinete de Estratégia e Estudos, Ministério da Economia e da Inovação, Lisbon 2010.

[18] Schumpeter J.A., *Theory of Economic Development*, published in 1934, Harvard Economic Studies, Cambridge 1912.

[19] van Praag C.M., *Business Survival and Success of Young Small Business Owners*, "Small Business Economics" 2003, vol. 21(1), pp. 1–17.

[20] Wagner J., *The post-entry performance of new small firms in German manufacturing industries*, "The Journal of Industrial Economics" 1994, vol. 42(2), pp. 141–154.

[21] Wicker A.W., King J.C., *Employment, ownership and survival in micro business: a study of new retail and service establishments*, "Small Business Economic" 1989, vol. 1, pp. 137–152.