Institute of Economic Research Working Papers

No. 71/2017

Duration Model of Enterprises – Analysis of Territorial Groups

Iwona Markowicz

Article prepared and submitted for:

9th International Conference on Applied Economics Contemporary Issues in Economy, Institute of Economic Research, Polish Economic Society Branch in Toruń, Faculty of Economic Sciences and Management, Nicolaus Copernicus University, Toruń, Poland, 22-23 June 2017

Toruń, Poland 2017

Iwona Markowicz
iwona.markowicz@wneiz.pl
University of Szczecin,
Mickiewicza 64, 71-101 Szczecin

Duration Model of Enterprises – Analysis of Territorial Groups

JEL Classification: *C10; C14; C41*

Keywords: *model of enterprises duration; Kaplan-Meier estimator; hazard function; cohort tables of enterprises survival*

Abstract

Research background: The popular term of *business demography* or *demography of the firm* denotes a relatively young area of science which focuses on the structures of cohorts of firms and the changes that undergo within these structures. As both the terms suggest, the studies use research methods traditionally applied in demographic studies. Survival analysis is increasingly used in business demography.

Purpose of the article: The purpose of the present study was to build the enterprises survival models for territorial groups in Zachodniopomorskie (Poland). In the first stage the Kaplan-Meier estimator was calculated and the test to verify the similarity of the survival function for poviats was conducted. Poviats were classified into groups. Next, the tables of enterprises survival were built and the business liquidation intensity was analysed in individual groups.

Methodology/methods: In this study the continuous-time non-parametric models were used: Kaplan-Meier estimator, Gehan test and duration table. Those methods were employed to model the survival time and find differences in the survival of firms in the poviats of the Zachodniopomorskie. In keeping with the above scheme five territorial groups with similar enterprises survival time models were distinguished.

Findings & Value added: The study results presented in this article reveal the differentiation of enterprises survival models in the territorial groups. Five groups of poviats were distinguished. These groups, as a result of the study, have been characterized.

Introduction

Survival analysis is increasingly used in business demography. In the literature on the subject, one may come across papers presenting the results of analysis of enterprises based on phenomena modelling methods. For

instance, López-Garcia and Puente (2006) adopt procedures for phenomena analysis to examine firms operating in Spain. The results of many studies (Hannan & Freeman, 1989; Carroll & Hannan, 2000; Caves, 1998) suggest that age, size, growth and mortality of enterprises are interrelated. Santarelli (2000) presented using Cox regression model for analysing the duration of new firms. The analysis is aimed at identifying the relationship between the duration and size of newly established enterprises offering financial intermediation in Italy. In his research on the bankruptcy of quoted companies functioning in Great Britain throughout the period 1965–2002, Bhattacharjee (2005) suggests taking account of initial and current size of firms in the risk model. He has proven that the effect of initial size of a given company on its lifespan is subject to change as the enterprise "grows older". Geroski, Mata and Portugal (2007) conducted survival analysis with reference to new firms established in Portugal. It was based on data collected by Portuguese Ministry of Labour for the period 1982–1995. On the basis of estimated semi-parametric models, it was stated that large enterprises operated longer. Kaniovski and Peneder (2008) used parametric analysis for determining the lifespan of Austrian firms during the period 1975–2004. The researchers proved that the risk of exit was greater in the case of enterprises operating in service sector (compared to production sector) and that larger enterprises were more likely to survive. Nunes and de Morais Sarmento (2012) determined survival function for firms founded in Portugal during the period 1987–2005. 86% of enterprises survived the first year, and only 22% had operated for 18 years.

For the purpose of the survival analysis the parametric, semi-parametric and non-parametric models are used. The construction of the parametric models requires the adoption of a theoretical distribution of the examined variable (Bieszk-Stolorz & Markowicz, 2015a), which is difficult in the case of studies on the duration of enterprises (Markowicz, 2012a; Markowicz, 2014). Therefore, in this study the continuous-time non-parametric models were used: Kaplan-Meier (K-M) estimator, Gehan test and duration table. Those methods were employed to model the survival time and find differences in the survival of enterprises in the poviats of the Zachodniopomorskie voivodeship.

The purpose of the present study was to build the business survival models for territorial groups. In the first stage the K-M estimator was calculated and the test to verify the similarity of the survival function for poviats was conducted. Poviats were classified into groups. Next, the tables of enterprises survival were built and the business liquidation intensity was analysed in individual groups.

This study used data from the registry of REGON, related to enterprises established in 2009-2011. These entities were observed to the end of 2013.

Research Methodology

The present paper presents the findings of studies into duration of enterprise. The initial event is the act of setting up a enterprise (registering it), while the final event is its de-registration. The episode is the time between the initial and the final events, or the duration of enterprise. In the survival analysis the random variable is the time between the events. It is a continuous variable, but it can be occasionally treated as a discrete variable if the observed time series are fixed.

In retrospective research the data can be censored (Bednarski, 2014). The researcher is interested in the probability of the event ending the observation of a given subject in the subsequent time units. If the event has failed to take place by the end of the observation, we call such an observation incomplete or censored (Markowicz & Stolorz, 2009). The main reasons for incomplete data to occur are our inability to continue the observation to the end, i.e. until all the subjects finish the process and the unavailability of some subjects in a fixed period of time. The former reason takes place when the observed process lasts for too long for the research capacity, while the latter – when the subjects have not been observed at the beginning of the process, escape the observation or are eliminated from the sample. The subject's duration can be left-, right- or bilaterally censored. Censoring which is most common in empirical studies is the right-censoring, hence such is the case in the present study (Markowicz, 2012b; Markowicz, 2015).

Random censoring takes place when individual subjects enter the observation at different calendar dates and the observation ends at a fixed date. Such censoring type is the best for the studies on the enterprise survival time. The period of observation is limited by dates. In the present study the observation took place during the period 2009-2011. The observed enterprises were set up and closed down at different moments. In research, Author used the methods of the survival analysis: Kaplan-Meier estimator and survival tables.

A non-parametric model of enterprises survival can be built by means of the K-M method, provided that we assume the presence of censored observations. In contrast to the survival tables, this method does not require grouping the observation times in class intervals.

The K-M estimator is a function that is non-increasing, periodically constant, with leaps at random time points determined by complete observations. The estimator adopts the values (Kaplan & Meier, 1958; Markowicz, 2012b; Bieszk-Stolorz & Markowicz, 2015b):

$$\hat{S}(t_i)=\begin{cases} 1 & \text{for} \quad t_0 \\ \prod_{t\le t_i}\left(1-\dfrac{z_i}{n_i}\right) & \text{for} \quad t_1 \le t \le t_k \\ 0 & \text{for} \quad t > t_k \quad \text{when} \quad \delta_n = 1 \\ \text{undefined} & \text{for} \quad t > t_k \quad \text{when} \quad \delta_n = 0 \end{cases} \tag{1}$$

where:

t_i – moment in which there was at least one event,

z_i – number of events in time t_i,

n_i – number of units of observation at time t_i.

The initial value of the survival function is 1 and it decreases at subsequent points of time t_i, at which at least one analysed event, has occurred. When using the K-M estimator, the probability of survival can be estimated at any time. The statistical relevance of the differences in survival models built for groups can be measured with an adequate non-parametric statistical test, taking into consideration the presence of censored data. The hypothesis that the survival functions for groups are equal (Gehan, 1965; Klainbaum & Klein, 2012) is verified by means of the Gehan test (Namboodiri & Suchindran, 1987; Domański *et al*., 2014).

The next stage of the analysis was the construction of cohort tables of enterprises survival (actuarial method) in individual groups. The tabular model has been built for the three-month models (Markowicz, 2015). The enterprises that did not go into liquidation by the end of 2013 are considered censored.

The number of enterprises that survived (n_t) was given at the beginning of the interval and calculated as $n_t = n_{t-1} - (z_{t-1} + c_{t-1})$, wherein (n_t) for $t = 0\,(n_0)$ denotes the initial number of enterprises in the cohort; z_t marks the number of enterprises liquidated in the interval $\langle t, t+1)$, c_t denotes the number of enterprises which did not experience that event by the end of the cohort observation time.

The paper presents duration tables in actuarial version. Therefore, in next column there is the number of units at risk (at risk of liquidation). This value is represented by means of the formula: $n_t^* = n_t - c_t / 2$.

The next two values in the cohort table of survival were estimated according to the discrete approach because they can be calculated only for the time interval. First of them means the probability of a enterprise to be liquidated in the time interval $\hat{f}_t^*$. The distribution of the business survival cannot be assigned to any known type of probability distribution. This is why the functions describing the process of the enterprises survival are not known, and the

tables of survival contain them their estimates calculated on the basis on empirical data. The probability estimator of business liquidation in the time interval $\hat{f}_t^*$ is a ratio of the number of liquidated economic entities in a given interval of time z_t to the number of enterprises at risk of liquidation by the interval n_t^*. Opposite to the probability of business liquidation in the time interval is the probability of business survival in the time interval:

$$\hat{p}_t^* = 1 - f_t^* = \frac{n_t^* - z_t}{n_t^*} \tag{2}$$

The probability of survival and the hazard intensity are functions continuous in their nature, but in the tables they are presented in a discrete approach. The probability of survival when calculated for interval $\langle t, t+1)$, is the probability that the enterprise will be liquidated after the time $t+1$:

$$\hat{S}_t^* = \prod_{k=1}^{t} p_k^* \tag{3}$$

In the moment $t = 0$ (in moment of setting up a economic entity) $S_t = 1$ and this function is decreasing over time. The rate at which the survival function is decreasing depends on the value of *t* and is defined as a hazard function (a_t – length of time interval):

$$\hat{h}_t^* = \frac{z_t}{\left(n_t^* - \frac{z_t}{2}\right) a_t} \tag{4}$$

Statistical Data

The article presents the results of the cohort analysis. The cohorts comprise the enterprises set up in the poviats of Zachodniopomorskie voivodeship in 2009-2011 that were observed to the end of 2013. Throughout 2009–2011 59587 enterprises were established. By the end of 2013 the number of liquidated enterprises reached 22234 (37%). In that study, data comes from a database REGON (National Official Business Register).

Territorial Groups with Similar Survival Functions

With a purpose to distinguish the territorial groups with similar firm survival times, the relevance of differences in the survival times of firms established in 2009-2011 in 21 poviats was examined. K-M estimators were calculated for each poviat and compared pair-wise. The groups were separated in such a way as to ensure that each of them contains poviats where

the differences among survival time models were not statistically significant. For each pair of poviats the hypothesis $H_0 : S_1(t) = S_2(t)$ was tested. The differences were considered significant when $p \leq 0.05$. Therefore, each group contained only the poviats with similar survival functions. The groups were ordered according to the decreasing probability of the enterprises survival with time. On the other hand, poviats within the groups are ordered depending on the increasing number of similarities to poviats in other groups. The territorial groups are shown in Table 1. In keeping with the above scheme five territorial groups with similar enterprises survival time models were distinguished. The first group consisted of the poviats: Koszalin (city), Szczecin (city) and policki (adjacent to the Szczecin) where the probability of the enterprises survival in the successive months was the highest. Another group comprised a single poviat – drawski. The tests indicated the similarity of the survival function to some poviats in the groups 1 and 3. The remaining groups were characterised with increasingly faster declining survival functions. It needs to be noted that the last group (gryficki) had the lowest probabilities of survival in the successive months. The last column in Table 1 shows the values test for many samples designated for groups of multi-poviat. The results reveal the lack of significant differences in the survival function within the groups.

Table 1. Groups of poviats with similar of enterprises duration model (the Gehan test)

Poviat	Numbers of similarities outside the group	Group	Test (*p*-value)
Koszalin (city)	0	1	1.9565 (0.3760)
Szczecin (city)	1		
policki	2		
drawski	8	2	
wałecki	1	3	6.4037 (0.6021)
kołobrzeski	2		
myśliborski	2		
szczecinecki	2		
świdwiński	2		
goleniowski	3		
gryfiński	3		
Świnoujście (city)	3		
kamieński	5		
sławieński	0	4	4.5315 (0.6051)
łobeski	0		
stargardzki	1		
pyrzycki	2		
choszczeński	4		
białogardzki	7		
koszaliński	8		
gryficki	0	5	

Source: own study

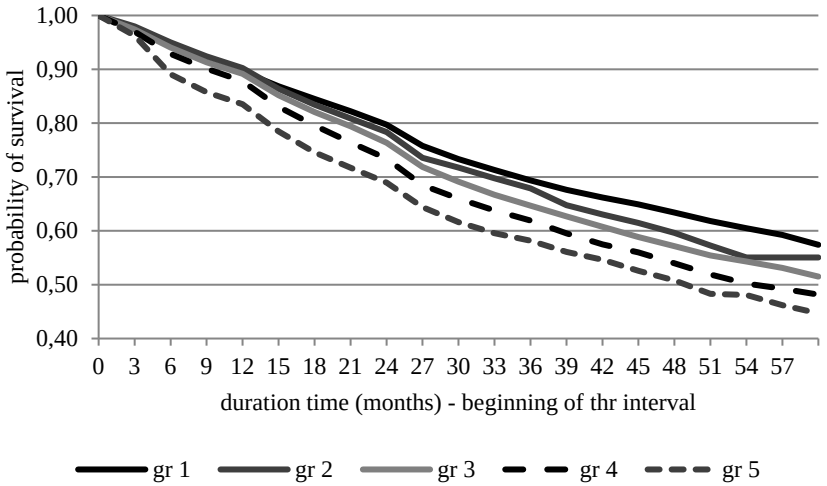
The Cohorts Tables of Enterprises – Actuarial Version

The next stage of the analysis was the construction of cohort tables of enterprises survival in individual groups. Out of the elements of the table, the survival function and the intensity function of enterprises liquidation (hazard function) was analysed.

Figure 1 shows the survival function of enterprises for five territorial groups. The functions of individual groups take a characteristic shape. Having analysed the shape of the survival function in individual groups, the following observations have been made:

- passing from group 1 to group 5 we can observe lower and lower values of survival function in a given period of time,
- the survival functions for individual groups reached the first quartile of the following times: 27.63; 25.48; 24.61; 22.60; 17.56 months,
- only for groups 4 and 5 function survival reaches a value of median (51.27 and 47.36 months).

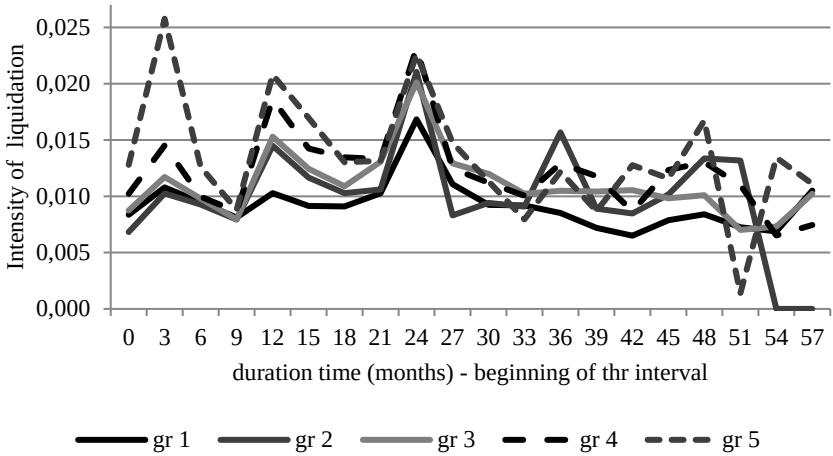
Figure 1. Estimation the probability of enterprises survival (groups 1–5)



Source: own study

Many studies have confirmed that the intensity function of enterprises liquidation adopts the inverted *U*-shape with a fixed maximum, which is in keeping with the theoretical learning model (Wagner, 1994; Audretsch, Santarelli & Vivarelli, 1999; Bartelsmann, Scarpetta & Schivardi, 2005; Markowicz, 2013).

Figure 2. Estimation the intensity of enterprises survival (groups 1–5)



Source: own study

Figure 2 shows the intensity function for five territorial groups. Having analysed the shape of the intensity function in individual groups built according to the survival models, the following observations have been made:

- the intensity function of liquidation (hazard function) for group 1 takes a typical inverted *U*-shape with the maximum marked within 24-27 months; it adopts low values,
- passing from group 1 to 5 we can observe less and less distinct shape of the intensity function in a form of inverted *U*-shape, increasingly higher intensities of liquidation and increasingly stronger fluctuations of this intensity over time.

Conclusions

The study results presented in this article reveal the differentiation of enterprises survival models in the poviats of the Zachodniopomorskie Voivodeship. Five groups of poviats were distinguished. The groups (for 1 to 5) were characterised by:

- decreasing time to firms liquidation (decreasing survival functions),
- less and less distinct shape of the intensity function (inverted *U*-shape),
- growing intensity of enterprises liquidation,
- increasingly stronger fluctuations of intensity in time,

- increasingly higher percentage of enterprise liquidated during two years of activity.

The research has shown that a critical moment is the 24th month of operation. One of the reasons for the enterprises liquidation may be the period of subsidized social security contributions. Since 2005 the start-up owners can pay significantly lower social security contributions for the first 24 months of operation, which definitely helps enterprises to survive. However, the failure of businesses after the period of subsidized contributions comes to an end suggests their weakness.

References

Audretsch, D.B., Santarelli, E., & Vivarelli, M. (1999). Start-up Size and Industrial Dynamics: Some Evidence from Italian Manufacturing. *International Journal of Industrial Organization*, 17(7). DOI: http://dx.doi.org/10.1016/S0167-7187(98)00002-2.

Bartelsmann, E., Scarpetta, S., & Schivardi, F. (2005). Comparative Analysis of Firm Demographics and Survival: Micro-Level Evidence for the OECD Countries. *Industrial and Corporate Change*, 14. DOI: http://dx.doi.org/10.1093/icc/dth057.

Bednarski, T. (2014). On robust causality nonresponse testing in duration studies under the Cox model. *Statistical Papers*, 55(1). DOI: http://dx.doi.org/10.1007/s00362-013-0523-0

Bhattacharjee, A. (2005). *Models of Firm Dynamics and the Hazard Rate of Exits: Reconciling Theory and Evidence using Hazard Regression Models.* St. Salvator's College, School of Economics and Finance.

Bieszk-Stolorz, B., & Markowicz, I. (2015a). Application of survival analysis in efficiency evaluation of professional effectiveness enhancement programme, *Acta Universitatis Lodziensis, Folia Oeconomica*, 1(311). DOI: http://dx.doi.org/10.18778/0208-6018.311.09.

Bieszk-Stolorz B., & Markowicz I. (2015b). Influence of Unemployment Benefit on the Duration of Registered Unemployment Spells, *Equilibrium*, 10(3). DOI: http://dx.doi.org/10.12775/EQUIL.2015.031.

Carroll, G. R., & Hannan, M. T. (2000). *The Demography of Corporations and Industries*, Princeton University Press.

Caves, R. E. (1998). Industrial Organization and New Findings on the Turnover and Mobility of Firms. *Journal of Economic Literature*, 36(4). DOI: http://dx.doi.org/10.4337/9780857938947.00027.

Domański, Cz., Pekasiewicz, D., Baszczyńska, A., & Witaszczyk, A. (2014). *Testy statystyczne w procesie podejmowania decyzji*, Łódź: Wyd. UŁ.

Gehan, E. A. (1965). A Generalized Two-Sample Wilcoxon Test for Double-Censored Data. *Biometrika*, 52(3–4).

Geroski, P.A., Mata, J., & Portugal, P. (2010). Founding Conditions and the Survival of New Firms, *Strategic Management Journal*, 31(5).

Hannan, M. T., & Freeman, J. (1989). *Organizational Ecology*. Cambridge: Harvard University Press.

Kaniovski, S., & Peneder, M. (2008). Determinants of Firm Survival: A Duration Analysis using the Generalized Gamma Distribution. *Empirica*, 35(1). DOI: http://dx.doi.org/10.1007/s10663-007-9050-3.

Kaplan, E.L., & Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association*, 53(282). DOI: http://dx.doi.org/10.1080/01621459.1958.10501452.

Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis. A Self-Learning Text*. New York: Springer. DOI: http://dx.doi.org/10.1007/978-1-4419-6646-9.

López-Garcia, P., & Puente, S. (2006). *Business Demography in Spain: Determinants of Firm Survival*, Documentos de Trabajo 0608, Madrid: Banco de España. DOI: http://dx.doi.org/10.2139/ssrn.901153.

Markowicz, I. (2012a). Application of logistic regression in firmography. *Actual Problems of Economics*, 5(2).

Markowicz, I. (2012b). *Statystyczna analiza żywotności firm,* Szczecin: Wyd. Nauk. US.

Markowicz, I. (2013). Hazard Function as a Tool to Diagnose Business Liquidation. *Folia Oeconomica Stetinensia*, 13(21). DOI: http://dx.doi.org/10.2478/foli-2013-0019.

Markowicz, I. (2014). Business Demography – Statistical Analysis of Firm Duration. *Transformations in Business & Economics*, 13/2B (32B).

Markowicz, I. (2015). Duration Analysis of Firms – Cohort Tables and Hazard Function. *International Journal of Business and Social Research*, 5(11).

Markowicz, I., & Stolorz, B. (2009). The Influence of Censorship on the Results of Data Analysis. *Polish Journal of Environmental Studies*,18(5B).

Namboodiri, K., & Suchindran, C. M. (1987). *Life Table Techniques and Their Applications*. New York: Academic Press Inc.

Nunes, A. & de Morais Sarmento, E. (2012). Business Demography Dynamics in Portugal: A Non-Parametric Survival Analysis. (Chapter 18). In: J. Bonnet, M. Dejardin, A. Madrid-Guijarro (Eds.). *The Shift to the Entrepreneurial Society. A Built Economy in Education, Sustainability and Regulation.* Cheltenham: Edward Elgar Publishing. DOI: http://dx.doi.org/10.4337/9780857938947.00027.

Santarelli, E. (2000). The duration of new firms in banking: an application of Cox regression analysis. *Empirical Economics*, 25(2). DOI: http://dx.doi.org/10.1007/s001810000019.

Wagner, J. (1994). The Post-Entry Performance of New Small Firms in German Manufacturing Industries. *Journal of Industrial Economics*, 42(2). DOI: http://dx.doi.org/10.2307/2950486.