Probability of Survival of New Manufacturing Plants: the case of Chile[*]

José Miguel Benavente[†] **Christian Ferrada**[‡]

Abstract

This paper studies the probability of survival of the manufacturing plants that start producing in Chile in the period 1979-1999 using a proportional hazards model. Opposing previous empirical international evidence, the survival diminishes with age, initial size, and with the rate of growth of the plant. It also diminishes with the regional unemployment rate. Comparing companies by size, the rate of risk of death of small and medium businesses is 20.3% greater than large companies, but more interesting, this rate is unaffected by the public subsidies received by the smaller ones, after controlling simultaneously with other explanatory variables. Using four different samples from the same data, the consequences of applying different criteria that validate and/or eliminate inconsistent data are studied, obtaining significant differences in the resulting coefficients.

Keywords: Chilean manufacturing; Entry; Survival; Duration.
JEL classification: L11; L60.

1. Introduction

The study of the determinants of the survival of plants has been a recent area of theoretical and empirical research. However, the study of these determinants considering size of plants is more recent and scarce. Traditionally, this topic has been motivated by the decisions that are made by the possible entrants, the competitive behavior of the established companies and the later performance of the incoming companies. In the case of Chile, recent research suggests that the firm life, especially for small and medium businesses (SME's) has been surprisingly promoted by explicit public support (Cabrera et. al, 2002). In this paper, we aim to make a contribution to this discussion by analyzing differences in survival rates between two categories of firm sizes when receiving subsidies, contrasting their results with international evidence.

Considering the variables affecting the survival of the plants, evidence shows some results that agree among studies and other results that are contradicted. For example, the survival increases in some studies and diminishes in others with age, initial size of the

[*] This document is part of Christian Ferrada's Economics Master thesis and benefited in great part with the motivation in the topic and suggestions from his thesis's advisor: Andrea Repetto. We are also grateful to Claudia Allendes, Raphael Bergoeing, David Bravo, Andrés Hernando, Oscar Landerretche, Gerhard Reinecke, and Andrea Sánchez for their useful comments and suggestions.
[†] Economics Department, University of Chile, email: jbenaven@econ.uchile.cl.
[‡] Economics Department, University of Chile, email: cferrada@econ.uchile.cl.

plant and the competition of the industry[1]. Since the results are varied and not absolutely conclusive, it is not possible to establish with certainty what happens to a company after it has entered the industry, the effects that it can have on the dynamics of the sector which it enters, or the differences in the performance of the plants for different categories of size. This motivates the study of the determinants of survival of the plants and the differences between SME's in Chile.

Econometric methods that have been used in empirical studies to estimate the effect of the explanatory variables on the survival of the manufacturing plants include probit, logit and duration models. Duration models are used in this paper, they have been recently developed and used primarily in economics on the job and unemployment duration research. Among the duration models, two procedures are considered: the non-parametric (that analyzes the effect of the explanatory variables by qualitative form) and the semi-parametric (that analyzes by quantitative form).

The behavior of the plants is studied from the moment they are born, that is to say, the only data used were that of plants where the moment of its birth can be identified. The considered explanatory variables include three categories: variables of the plant, variables that measure the performance of the industry and variables that measure the macroeconomic performance. Next section describes the data and the four different samples that are considered in the estimations. Section 3 details the models of duration and their properties when being applied to the data. Section 4 presents and discusses the empirical results and finally, the conclusions appear in section 5.

2. Data

The data used come from the National Industrial Annual Survey (ENIA) conducted by the National Institute of Statistics (INE) for the variables associated to the performance of the plants and the industry. This survey is carried out at plant level and includes almost all the manufacturing companies in Chile with 10 or more employees. The observations are annual. Period 1979-1999 is considered and the following data are used: information on employment, worked days, added value, gross value of the production, sales, costs of raw materials and wages, capital, investments, subsidies, exports, demand by electricity and industrial sector (4-digits ISIC code).

The data used in this paper only consider the plants whose moment of birth can be identified. A plant is born when it is registered in a year different from 1979 and has not been registered in the previous period. In order to identify the death of a plant, the plant must be registered in a year different from 1999 and not be registered in the following years. This measurement of death overestimates the rate of death of the plants, because in addition to the closing or bankruptcy of the plants, it is considered that a plant has died when the number of employees fell below 10, when they were eliminated, when they changed their production activity, when they were not able to be located at the time of the survey, when there was no movement of capital, when operations were paralyzed, when they were shut down, and when under investigation by the Internal Tax Service (SII) or

[1] For more details see Jovanovic (1982), Evans (1987a,b), Dunne, Roberts and Samuelson (1989), Baldwin and Gorecki (1991), Wagner (1994), Audretsch and Mahmood (1994 and 1995), Doms, Dunne and Roberts (1995), Mata, Portugal and Guimaraes (1995), Konings, Roodhooft and Van de Gutch (1996).

because they had merged with another plant. With the purpose of diminishing some of this source of bias and improve errors of measurement and management, to the first sample the filters based on the work of Micco (1995)[2] are applied. In addition, in the first sample only 8 industrial sectors to 3-digits ISIC code are considered to make the results comparable with the work of Bergoeing, Hernando and Repetto (2003). Three other samples from the same data base are used with the purpose of contributing to the discussion of considering different filters and management of the data to make them more consistent. The second sample of data amplifies the 8 sectors of the first sample to all the sectors in which the total factor productivity of the plants can be calculated, which corresponds to 23 sectors. In the third sample the filters based on the work of Micco (1995) are not applied, but we tried to maintain the totality of the data, fixing them when possible with the data of the same plant in the previous year, following year or both[3]. Finally, the fourth sample is equal to the third, but it is catalogued as deaths of the plants that the INE has classified as closed down or bankrupt, this is the reason why the sample is reduced to the period in which this information was controlled (1996-1999).

In the following paragraphs the explanatory variables in the estimations will be described. The age of the plants is measured in years, because the data of the ENIA corresponds to annual observations. A plant is one year of age if it is not registered in the survey the previous year. Since the observations cover period 1979-1999, the maximum registered age corresponds to 20 years (for a company that was born in 1980, that is, that was not registered in 1979 but registered in 1980, and which remained alive until 1999). The size of the plant is measured by the logarithm of the total number of workers (skilled and unskilled) fitted by the number of worked days in the year (quotient between worked days and 365). The initial size of the plant corresponds to the size of the plant when it enters the sample. The growth corresponds to the rate of growth that is calculated following the method used by Davis, Haltiwanger and Schuh (1996):

$$g_t = \frac{E_t - E_{t-1}}{(E_t + E_{t-1})/2} \qquad (1)$$

where E_t corresponds to the fitted employment in t and E_{t-1} to the fitted employment in $t-1$. The total factor productivity is estimated using the method based in Olley and Pakes (1996), where the problems of simultaneity between the demand of factors and the productivity unobserved term and the selection problem originated from observing plants with productivity greater than the productivity of dead plants. In addition, the productivity estimation is complemented by the method proposed by Levinsohn and Petrin (1999), where the demand of electricity instead of the investment of the plant is used to consider the registries that have no positive investment. This variable is adimensional and it is obtained from the used data of the study of Bergoeing, Hernando and Repetto (2003). The exports correspond to the percentage that represent within the total sales. The profits of the plant are measured like the price-cost margin, corresponding to the gross value of production minus the costs of wages and of raw materials, divided by the gross value of production. The investment of the plant corresponds to the investment made in three types of capital (real estate, machineries and equipment, and vehicles and transport). The obtaining method is based on Liu (1993) and Olley and

[2] For a detailed enumeration of these filters see Annex No. 1.
[3] Annex No. 2 details the adjustments made to the data in the third sample.

Pakes (1996), where it is necessary to obtain an estimation of the stock of capital by means of the following rule of accumulation:

$$k_{t+1} = (1-\delta)k_t + i_t \qquad (2)$$

where k_t corresponds to the stock of capital in period t, δ is the depreciation rate of the capital (that is different for each type of capital) and i_t is the level of investment in period t. The investment and capital data also are obtained from the data used in the study of Bergoeing, Hernando and Repetto (2003) and are measured in real terms based in 1985. The favorable policies correspond to the percentage that represents subsidies (by tributary tax exemptions or other types of subsidies) within the total sales. With respect to the characteristics of the industry, the barriers of entrance are measured in two forms: by the average of spend publicity of the entrants to the industry (economic sector to 3-digits ISIC code) per year and region where the plant is entering and by the average of the stock of capital of the entrants to the industry, per year and region where the plant is entering. These two variables are supposed constant for each plant during their life and are measured in real terms based in 1985. The competition in the industry is measured by the growth rate of the industry and by the rate of entrance to the industry, in the year and region to which the plant belongs. Relative to the characteristics of the macroeconomic ambient, the considered interest rate corresponds to the average interest rate for collocations from one to three years in the financial system, nonreadjustable and annualized, and the unemployment rate corresponds to the annual desoccupation rate from the INE. The regional rate is used, to have a near measurement of the economic cycle that faces the plant in its surroundings. These data are obtained from the Central Bank of Chile.

Because of the nature of the data, in terms of the concepts used in the duration literature, truncated and right censored datum exist, because there are data that are not considered because it does not fulfill the filters of Micco (1995) and therefore is eliminated from the sample although the plant has valid data for other years, and because if a plant survives until year 1999, it is not known what happens to it in the following year (if it dies or it continues surviving).

3. Duration models

In order to analyze the time elapsed until the death of the plants (dependent variable) and how it is affected by the explanatory variables models of duration will be used, which allow the calculation of the probability of survival (or analogous the hazard rate), to consider explanatory variables that change in time and to treat formally censored observations (observations that are not known if they result in death or not), this is the reason why the properties of the conventional estimators[4] improve.

The hazard rate is defined as:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr(t < T \leq t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(t)}{S(t)} \qquad (3)$$

where $\lambda(t)$ corresponds to the hazard rate, $f(t)$ to the density of probability and $S(t)$ to the survival probability. $\lambda(t)\Delta t$ corresponds to the probability that the individual or set of

[4] Least Squares, probit or logit models lead to biased and inconsistent estimations.

individuals in study have an equal duration to t, that is to say, that happens the event in analysis (the death of the manufacturing plants) in the interval of time $[t,t+\Delta t]$, but conditional on surviving until moment t. As the density of probability can be written in terms of the survival probability, a unique relation between hazard rate and probability of survival exists, this is the reason why modeling one of these functions is equivalent to doing it for the other function. The relation between $f(t)$ and $S(t)$ is given by:

$$f(t) = \lim_{\Delta t \to 0} \frac{\Pr(t < T \leq t + \Delta t)}{\Delta t} = -\frac{\partial S(t)}{\partial t} \qquad (4)$$

In order to consider the hazard rate and the probability of survival without taking into account quantitatively the differences observed between the explanatory variables, the non-parametric procedure is used, calculating the Kaplan-Meier estimator[5]. In order to consider these differences and to do a more complete analysis, the semi-parametric procedure is used, where the most popular specification is the one of the model of proportional hazards of Cox (1972), where the hazard rate is affected by the time and the explanatory variables by means of the following functional form:

$$\lambda(t, X) = \lambda_0(t)\exp(\beta' X) \qquad (5)$$

where $\exp(\beta' X)$ corresponds to the term by means of which the explanatory variables affect the hazard rate, and where $\lambda_0(t)$ it is the base hazard rate, which only depends on the time (age of the plants), it does not have a pre-established functional form and it is the same one for all the individuals. The global hazard rate is estimated by partial maximum likelihood, which allows to indeterminate the base hazard rate and to consider only the effect of the explanatory variables. The base hazard rate is recalculated after which the coefficients are estimated. The advantage of this is that it does not force the assumption of a specific form for the base hazard rate (in contrast with parametric procedures) that if it is chosen unsuitably it produces unreliable and unstable estimators (Heckman and Singer, 1986)[6]. The main characteristic of the semi-parametric procedure is that the differences in the explanatory variables lead to proportional changes in the hazard rate independent of time. For example, for a certain time t, a change in the X_k variable from a value of X_{k1} to X_{k2} takes to a change in the hazard rate given by:

$$\frac{\lambda(t, X_{k2})}{\lambda(t, X_{k1})} = \exp[\beta_k (X_{k2} - X_{k1})] \qquad (6)$$

To this procedure an additional term of heterogeneity is added, that allows the considering differences between the individuals attributed to variables non-observed and that improves problems of omitted variables and errors of measurement in the survival times or in the explanatory variables (Lancaster, 1990). Considering this term, the hazard rate is written like:

$$\lambda(t, X \mid \theta) = \theta\lambda_0(t)\exp(\beta' X) \qquad (7)$$

where the heterogeneity effect is given by the term θ. The no inclusion of this term leads to the overestimation of the fall of the hazard rate in the time, to obtain proportional hazards that are not constant (they should be by construction) and to overestimate the obtained coefficients (Lancaster, 1990).

---

[5] For more details, see Cox and Oakes (1984) or Lancaster (1990).

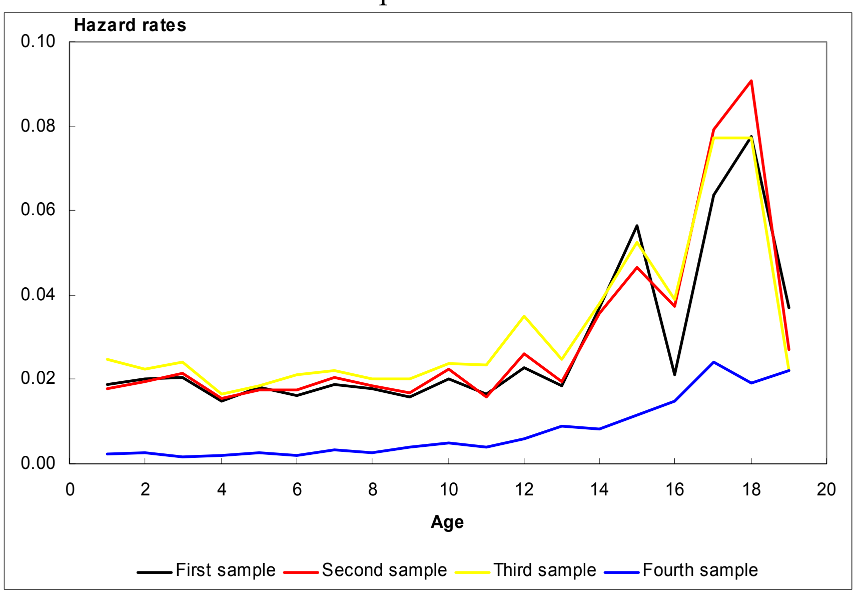
[6] Ferrada (2004) shows parametric and semi-parametric estimations of the survival of manufacturing plants in Chile, obtaining that the second method produces more efficient estimations.

4. Empirical results

Before obtaining empirical results and with the purpose of looking at the data consistency of the ENIA and comparing it with another Chilean data, the results of the rates of exit in the manufacturing sector between ENIA and SII data bases are compared, considering periods with common data (1996-1999). This last data base had to represent better dynamics of what happened in Chile, because it considers all the companies that maintain commercial activities[7]. The rate of exit is greater in the data of SII for every year, except in 1999. The average rate of exit in the SII data is of 19%, whereas in the data of the ENIA it is of 14%. The previous statement is consistent when considering 2-digits ISIC sectors: the rate of exit is greater in the data of SII in 32 of the 36 sectors-year considered and the remaining 4 sectors-year that differ from this behavior correspond to the year 1999.

The results begin by analyzing how the hazard rate obtained by means of the non-parametric procedure (estimator of Kaplan-Meier) based on the age of the plants varies.

Figure No. 1: Hazard rates of the manufacturing plants based on the age for the four different samples from the same data base.



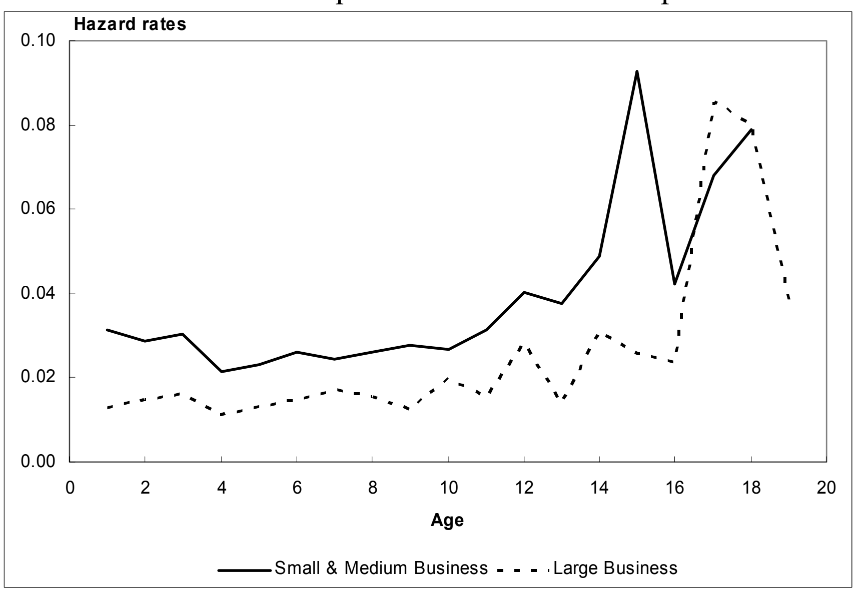
The rate of risk is considerably smaller for the fourth sample [Figure No. 1], because in this case there are fewer considered deaths (considering only plants that went out of business or closed during period 1996-1999 by means of information given by the INE on status of the plant). The other 3 samples have similar hazard rates, being relatively constant until 13 years old (something superior to 2%)[8], age from which they behave on irregular and increasing form (reaching 9% for the second sample at 18 years old), which

[7] Remember that ENIA data covers only plants with 10 or more employees, while SII data covers all plants, including plants with 10 or less employees.

[8] This means that 2% of the plants die every year within the first 13 years of life, conditional on surviving until the considered year.

contradicts most of the theoretical and empirical studies[9]. An explanation for the volatile behavior is in that from this age there are few plants that survive in the sample, this is the reason why the performance of these plants can dramatically affect the calculated rates. Using the non-parametric procedure also the hazard rate by groups of data can be obtained, for example, for the companies cataloged by category of size (small and medium and large companies). On average, small and medium companies have a hazard rate of 20.3% greater than the large companies. For the four samples these effects vary considerably: 17.1%, 26.8%, 52.0% and -15.8% for the first, second, third and fourth samples, respectively. It is important to notice that the rate of risk is smaller for the small and medium companies in the fourth sample, which reinforces the idea that different samples give different results. Figure No. 2 shows the differences in the hazard rates by size of companies for the second sample, which is the one that has the greater difference between groups.

Figure No. 2: Hazard rates of the manufacturing plants based on the size of the plants for the second sample.



It is observed that in most of the ages, the hazard rate is greater for the small and medium companies, except at 17 years. Since the hazard rate varies considerably according to the size of the company, it is supposed that the other variables also have a similar effect, this is the reason why it is important to study the effect of the other determinants on the rate of risk. For this, the semi-parametric method of Cox (1972) is used, whose results appear in Table No. 1.

---

[9] Most of international empirical studies show that survival increases with age.

Table No. 1: Determinants of the hazard rate of the manufacturing plants.
Results for the four different samples from the same data base.

Variables	Sample			
	1st	2nd	3rd	4th
Characteristics of the plant				
Initial size	6.43E-01	6.60E-01	6.79E-01	6.79E-01
	(11.16)	(13.85)	(16.54)	(16.54)
Current size	-7.61E-01	-7.90E-01	-8.59E-01	-8.59E-01
	(-13.93)	(-17.74)	(-21.36)	(-21.36)
Growth	2.74E-01	3.82E-01	3.47E-01	3.47E-01
	(2.65)	(4.58)	(4.42)	(4.42)
Productivity	-5.60E-05	-4.37E-05	-3.32E-05	-3.32E-05
	(-1.92)	(-2.85)	(-3.64)	(-3.64)
Exports	2.56E-01	2.68E-01	2.38E-01	2.38E-01
	(1.74)	(2.00)	(1.91)	(1.91)
Profits	-5.55E-01	-6.16E-01	-5.16E-02	-5.16E-02
	(-4.70)	(-6.94)	(-3.77)	(-3.77)
Investment	-5.88E-07	-5.25E-07	-5.00E-07	-5.00E-07
	(-2.57)	(-3.02)	(-2.98)	(-2.98)
Subsidies	2.92E-01	-3.06E-01	4.69E-02	4.69E-02
	(0.86)	(-0.67)	(0.61)	(0.61)
Characteristics of the industry				
Publicity	1.04E-06	7.92E-07	3.47E-06	3.47E-06
	(0.25)	(0.20)	(1.00)	(1.00)
Capital	2.57E-08	2.48E-08	2.73E-08	2.73E-08
	(0.77)	(0.75)	(0.97)	(0.97)
Growth	3.15E-01	4.83E-01	8.90E-01	8.90E-01
	(1.36)	(2.85)	(7.49)	(7.49)
Rate of entrance	1.36E+00	1.20E+00	-1.38E-01	-1.38E-01
	(3.31)	(3.93)	(-0.36)	(-0.36)
Characteristics of macroeconomic ambient				
Interest rate	1.34E-01	1.19E-01	1.29E-01	1.29E-01
	(5.37)	(5.82)	(7.28)	(7.28)
Unemployment rate	-3.02E-02	-3.84E-02	-3.33E-02	-3.33E-02
	(-2.09)	(-3.21)	(-3.14)	(-3.14)
Number of obs.	9640	15872	20403	20403

Asymptotic t statistics in parenthesis.
The estimations include sectorial (3-digits ISIC code) and annual dummies.

As opposed to what is expected, the initial size increases the hazard rate, because companies with greater initial size would have more confidence in their future performance and have a greater possibility of surpassing the economies of scale of the industry. Nevertheless, the barriers to the entrants (measured by publicity and capital) which are closely related to initial size, are not significant for the Chilean case in the determination of the survival. The explanation that is left is that companies that start small have more possibilities of surviving because they incur in smaller sunk costs. The current size diminishes the hazard rate, because plants with a greater size have greater resources to invest in feasibly profitable projects, generally they have more capable managers (Lucas, 1978) and because if the business fails, they have a greater buffer stock of defense that allows them to decrease in size before closing the business. Following the findings with respect to the initial size, the plants that enter with more caution have less hazard rates; growth also increases the hazard rate, indicating that plants which behave in a less volatile way as far as the use variations have greater probabilities of surviving. The total factor productivity increases the probability of survival, like the profits and the investment of the plant, this reinforces the idea of Ericson and Pakes (1995) which says that more profitable plants survive longer and also possibly do this by investing in profitable projects. The effect of the exports and the subsidies received do not have a significant effect on any of the considered samples. The competence of the industry measured by growth of the industry and the rate of entrance, are not significant in all the cases, but when they are, they present the expected effects: The rate of growth increases the survival because it indicates that the plants grow without producing losses in the participation of market of the other companies, and the rate of entrance diminishes the survival because it is bound to a greater competition within the productive sector. The interest rate increases the hazard rate because it increases the price of loans requested by the new entrants and the rate of unemployment diminishes the survival explaining that the macroeconomic recessions agree with a greater death of the plants and therefore with greater spaces of development for the companies that enter the industry.

With respect to the heterogeneity considered in the models, in the first sample the data do not comply to the consideration of this assumption, nevertheless, it does in the other samples, affirming that the estimations are improved when eliminating problems in the data due to errors in the measurement of the explanatory variables or in the dependent variable or by omitted variables (Lancaster, 1990). In the same way, with the specifications of the model, for the four samples, it is observed that the errors are orthogonal, also eliminating the existence of possible endogeneity between age of the plants and variables, like for example, productivity, size or investment.

When comparing the samples, it is observed that the signs of the considered coefficients do not vary between samples, but the magnitudes vary. Then, the researcher has to be careful when selecting the size of the sample or the filters to improve the quality of data. For example, the increase in exit rate for SME's compared with large plants is positive in the first three samples, but negative in the fourth; or an increase of productivity from zero to the average of the fourth sample produces a reduction in the hazard rate of a 12% for the first sample, but of only a 7% for the fourth[10].

The semi-parametric method of Cox (1972) can also be performed for SME's and large companies separately, whose results appear in Table No. 2.

[10] These numbers are calculated using Equation No. 6.

Table No. 2: Determinants of the hazard rate of the manufacturing plants.
Results for the second sample and for SME's and large companies.

Variables	Size of plant	
	Small and medium	Large
Characteristics of the plant		
Initial size	5.33E-01	7.49E-01
	(5.15)	(13.13)
Current size	-8.40E-01	-8.11E-01
	(-7.01)	(-13.55)
Growth	3.25E-01	5.22E-01
	(2.22)	(4.54)
Productivity	-8.99E-05	-2.80E-05
	(-1.44)	(-1.83)
Exports	1.64E+00	-6.48E-02
	(5.48)	(-0.42)
Profits	-1.10E+00	-8.58E-01
	(-5.62)	(-5.83)
Investment	-3.54E-07	-6.38E-07
	(-0.93)	(-2.93)
Subsidies	-6.61E-01	-5.13E-01
	(-0.43)	(-1.02)
Characteristics of the industry		
Publicity	-1.67E-05	1.70E-06
	(-1.08)	(0.40)
Capital	3.18E-06	2.45E-08
	(4.76)	(0.65)
Growth	5.99E-01	3.58E-01
	(2.37)	(1.52)
Rate of entrance	7.73E-01	1.32E+00
	(1.64)	(3.21)
Characteristics of macroeconomic ambient		
Interest rate	7.76E-02	1.33E-01
	(2.47)	(4.77)
Unemployment rate	6.74E-03	-7.71E-02
	(0.40)	(-4.48)
Number of obs.	4937	10787

Asymptotic t statistics in parenthesis.
The estimations include sectorial (3-digits ISIC code) and annual dummies.

As obtained in Table No. 1, public subsidies do not have a significant positive effect over the survival of plants, and this remains valid after controlling by the size of plants. Previous research in Chile (Cabrera et. al, 2002) found that subsidizing the creation of new firms stimulates the entrance of more inefficient companies and delay, but not avoid the exit of the unsuccessful. Also, this policies diminish the growth of productivity of the economy because they delay the exit of more unproductive plants and delay the growth of more potential plants.

5. Conclusions

The decisions made by the possible entrants when starting a business, the competitive behavior of the established companies and the later performance of the entrants motivate the study of the determinants of the survival of the manufacturing plants. The empirical evidence is not conclusive, obtaining different results in different countries and more drastic still, different results in the same countries[11].

In this paper, as opposed to what is expected from the international evidence, the rate of survival falls with age, initial size and growth of the plant. The first effect is associated with the absence of an initial experimental period in the lives of plants. The second effect suggests that companies that start small survive longer, explained mainly by the absence of high sunk costs. The third effect suggests that abrupt behaviors in the change of size diminish the survival. In agreement with what is expected, survival increases with variables that reflect the value of the business: total factor productivity, profits and investment. Variables like exports, subsidies and barriers to entrance are not significant in any of the four analyzed samples. The competition in the industry has strong effects on the survival, increasing it with the rate of growth of the industry and diminishing it with the rate of entrance. The interest rate diminishes the survival, reflecting the difficulties in the access to credits and the rate of unemployment increases the survival, showing that after recessions, more spaces exist to develop businesses, partly explained by the continuous process of creation and destruction.

The use of semi-parametric models of duration is an efficient way to consider all the explanatory variables of the model simultaneously, including terms of heterogeneity among companies and improving problems of omitted variables or errors in the measurement. The endogeneity presence, although it could be motivated by theoretical reasons, is not in the data because of the errors of the models are all orthogonal.

Finally, this work notably contributes to the discussion of public support to SME's, because it is obtained that subsidies do not affect the survival of plants in a positive way, and this remains valid after controlling by the size of plants. This supports the idea developed by Cabrera et. al (2002) where they stated that SME's public support must focus and be more selective on good projects and good plants and not to be directed to small firms motivated only for their size (to prevent market failures like credit restrictions).

References

- Audretsch, D., 1995. "Innovation, Growth, and Survival", *International Journal of Industrial Organization*, **13**, 441-457.
- Audretsch, D. and Mahmood, T., 1994. "The Rate of Hazard Confronting New Firms and Plants in U.S. Manufacturing", *Review of Industrial Organization*, **9**, 41-56.
- Audretsch, D. and Mahmood, T., 1995. "New Firm Survival: New Results Using a Hazard Function", *The Review of Economics and Statistics*, **77**, 97-103.

[11] For example, see Audretsch and Mahmood (1994) and Audretsch and Mahmood (1995) for the manufacturing sector in US.

- Baldwin, J. and Gorecki, P., 1991. "Firm Entry and Exit in the Canadian Manufacturing Sector, 1970-1982", *Canadian Journal of Economics*, **24**, 300-323.
- Bergoeing, R., Hernando, A. and Repetto, A., 2003. "Idiosyncratic Productivity Shocks and Plant-Level Heterogeneity", *Working Paper 173*, Centro de Economía Aplicada, Universidad de Chile.
- Cabrera, A., de la Cuadra, S., Galetovic, A., and Sanhueza, R., 2002. "Las pyme: quiénes son, cómo son y qué hacer con ellas", SOFOFA, Chile.
- Camhi, A., Engel E. and Micco, A., 1997. "Dinámica de Empleo y Productividad en Manufactura: Evidencia Micro y Consecuencias Macro", *Working Paper 19*, Centro de Economía Aplicada, Universidad de Chile.
- Cox, D. R., 1972. "Regression Models and Life-Tables (with Discussion)", *Journal of the Royal Statistical Society*, Series B, **34**, 187-220.
- Cox D. R. and Oakes, D. 1984. Analysis of Survival Data, London: Chapman & Hall.
- Davis, S., Haltiwanger, J. and Schuh, S., 1996. *Job Creation and Job Destruction*, The MIT Press.
- Doms, M., Dunne, T. and Roberts, M., 1995. "The Role of Technology Use in the Survival and Growth of Manufacturing Plants", *International Journal of Industrial Organization*, **13**, 523-545.
- Dunne, T., Roberts, M. and Samuelson, L., 1989. "The Growth and Failure of U.S. Manufacturing Plants", *Quarterly Journal of Economics*, **104**, 671-698.
- Ericson, R. and Pakes, A., 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work", *Review of Economic Studies*, **62**, 53-82.
- Evans, D., 1987a. "The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries", *Journal of Industrial Economics*, **35**, 567-581.
- Evans, D., 1987b. "Tests of Alternative Theories of Firm Growth", *Journal of Political Economy*, **95**, 657-674.
- Ferrada, C., 2004. "Supervivencia de las Plantas Manufactureras: Evidencia de Chile", *mimeo*, Universidad de Chile.
- Geroski, P., *Market Dynamics and Entry*. Oxford: Blackwell, 1991.
- Gibson, J. and Harris, R., 1996. "Trade Liberalization and Plant Exit in New Zealand Manufacturing", *The Review of Economics and Statistics*, **78**, 521-529.
- Heckman, J. and Singer, B., 1986. "Econometric Analysis of Longitudinal Data", in Zvi Griliches y Michael Intriligator (eds.), *Handbook of Econometrics*, 3, Capítulo 29, Elsevier Science Publishers, 1690-1763.
- Hopenhayn, H. y Rogerson, R., 1993. "Job Turnover and Policy Evaluation: A General Equilibrium Analysis", *Journal of Political Economy*, **101**, 915-938.
- Hedges, L., 1992. "Modeling Publication Selection Effects in Meta-Analysis", *Statistical Science*, **7**, 246-255.
- Jovanovic, B., 1982. "Selection and the Evolution of Industry", *Econometrica*, **50**, 649-670.
- Konings, J., Roodhooft, F. and Van de Gucht, L., 1996. "The Life Cycle of New Firms and its Impact on Job Creation and Job Destruction", *Working Paper No. 9669*, Faculty of Economics and Applied Economics, Katholieke Universiteit Leuven.
- Lancaster, T., 1990. *The Econometric Analysis of Transition Data*, Cambridge: Cambridge University Press.

- Levinsohn, J. and Petrin, A., 1999. “When Industries Become More Productive, Do Firms? Investigating Productivity Dynamics”, *NBER Working Paper 6893*.
- Liu, L., 1993. “Entry-Exit, Learning and Productivity Change: Evidence from Chile”, *Journal of Development Economics*, **42**, 217-242.
- Lucas, R., 1978. “On the Size Distribution of Business Firms”, *Bell Journal of Economics*, **9**, 508-523.
- Mata, J., Portugal, P. and Guimarães, P., 1995. “The Survival of New Plants: Start Up Conditions and Post-Entry Evolution”, *International Journal of Industrial Organization*, **13**, pp. 459-481.
- Micco, A., 1995. “Creación, Destrucción y Reasignación de Empleos en Chile”, *Magíster de Economía Aplicada Thesis*, Universidad de Chile.
- Olley, S. and Pakes, A., 1996. “The Dynamics of Productivity in the Telecommunication Equipment Industry”, *Econometrica*, **64**, 1263-1298.
- Pavcnik, N., 2000. “Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants”, *NBER Working Paper 2893*.
- Wagner, J., 1992. “The Post-Entry Performance of New Small Firms in German Manufacturing Industries”, *Journal of Industrial Economics*, **42**, 141-154.

Annex No. 1: Filters based on the work of Micco (1995)

Following the methodology of Micco (1995) to eliminate some of the inconsistencies and errors of measurement in the data provided in the ENIA, three types of criteria are used to filter the data: “level criterion”, “size criterion”, and “ISIC criterion”. The level criterion is not fulfilled for one observation that has some of the following inconsistencies:

- Worked Days <= 0.
- Gross Value of the Production < 0.1.
- Added value < 0.1.
- Size = 0.
- ISIC < 3000.
- Remuneration of Workers = 0.
- Sales < Exports.
- Added value > Gross Value of the Production.
- (Administrative Employment + Productive Process Employment = 0) & (Remuneration of Workers + Remuneration of Workers on a Commission Basis <> 0).
- (Administrative Employment + Productive Process Employment <> 0) & (Remuneration of Workers + Remuneration of Workers on a Commission Basis = 0).
- Total Remuneration <= 0.
- Total employment - Productive Process Employment - Administrative Employment = 0.

The employment criterion is not fulfilled if the employment is less than 15 employees. This criterion has the purpose of improving the specification of the deaths and of the entrances, since the ENIA only considers plants that have more than 10 employees in the year and it is possible to register by births continuous plants that have increased their employment more than 10 employees, or it is possible to register by deaths continuous plants that have diminished their employment below the 10 employees. ISIC criterion is not fulfilled if 3-digits ISIC code is equal to 372 or the 2-digits code is equal to 39. These three criteria differ in the form that invalidate the data depending on if the registry of the plant corresponds to a birth, a death or to a continuous plant.

A registry corresponds to a valid birth when the plant is registered in a year greater than 1979, is not observed in the previous year and fulfills the three criteria (level, employment and ISIC) in the year when it is registered for the first time. A registry corresponds to a death when a plant is registered in a year previous to 1999, is not observed in the following year and fulfill the three criteria in the year in which it is registered. A registry corresponds to a continuous plant when it is registered in the current and previous years and fulfills the "level criterion" in both years, fulfills the "employment criterion" in at least one of both years and fulfills "ISIC criterion" in the current year.

Annex No. 2: Adjustments made to the data in the third sample

In the third specification the data is fixed that do not fulfill the criteria considered in Micco (1995). Since the explanatory variables are not considered in all these criteria, the data of the variables are only fixed when they are subject to the filters and only for the data where it is possible to calculate the age of the plants. Next the adjustment of the data for each considered filter will be detailed:

- Worked Days <= 0: They do not fulfill 33 registries, of which 32 are fixed (1 is eliminated from the sample). 17 registries are updated with the average of the days worked between the previous and the posterior year, 4 registries are updated with the worked days of the previous year and 11 registries are updated with the worked days of the posterior year.
- Gross Value of the Production (VBP) < 0.1: They do not fulfill 10 registries, of which 9 are fixed (1 is eliminated from the sample). 8 registries are updated with the VBP of the previous year and 1 registry is updated with the VBP of the posterior year.
- Added Value (VA) < 0.1: They do not fulfill 308 registries, of which 291 are fixed (17 are eliminated from the sample). 114 registries are updated with the average of the VA between the previous and the posterior year, 147 registries are updated with the VA of the previous year and 30 registries are updated with the worked days of the posterior year.
- Employment = 0: They do not fulfill 2 registries, of which 1 is fixed with the employment of the previous year (1 is eliminated from the sample).
- ISIC < 3000: All the registries fulfill this filter.
- Sales < Exports: They do not fulfill 4 registries, of which 3 are fixed with the average of the sales and the exports between the previous and the posterior year (1 is eliminated from the sample).
- Added Value (VA) > Gross Value if the Production (VBP): They do not fulfill 31 registries, of which 31 are fixed (3 are eliminated from the sample). 17 registries are updated with the average of the VA and VBP between the previous and the posterior year and 4 registries are updated with the VA and VBP of the posterior year.
- Total Remuneration <= 0: They do not fulfill 4 registries. 2 registries are updated with the average of total remuneration between the previous and the posterior year, 1 registry

is updated with the total remuneration of the previous year and 1 registry is updated with the total remuneration of the posterior year.