



Catalogue no. 11F0019MIE — No. 220

ISSN: 1205-9153

ISBN: 0-662-36818-5

## Research Paper

Analytical Studies Branch research paper series

# Factors Determining the Success or Failure of Canadian Establishments on Foreign Markets: A Survival Analysis Approach

By Jean Bosco Sabuhoro and Yvan Gervais

Business and Trade Statistics Field  
13-C8, Jean Talon Building, Ottawa, K1A 0T6

Telephone: 1 800 263-1136

*This paper represents the views of the authors and does not necessarily reflect the opinions of Statistics Canada.*



Statistics  
Canada

Statistique  
Canada

Canada

# **Factors Determining the Success or Failure of Canadian Establishments on Foreign Markets: A Survival Analysis Approach**

by

Jean Bosco Sabuhoro and Yvan Gervais

**11F0019MIE No. 220**  
**ISSN: 1205-9153**  
**ISBN: 0-662-36818-5**

Business and Trade Statistics Field  
Statistics Canada  
Jean Talon Building, 13-C8  
Ottawa, Ontario, K1A 0T6

**How to obtain more information :**  
National inquiries line: 1 800 263-1136  
E-Mail inquiries: [infostats@statcan.ca](mailto:infostats@statcan.ca)

**May 2004**

This paper represents the views of the authors and does not necessarily reflect the opinions of Statistics Canada.

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2004

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission from Licence Services, Marketing Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6

*Aussi disponible en français*

# *Table of Contents*

<b>Abstract .....</b>	<b>ii</b>
<b>1. Introduction.....</b>	<b>1</b>
<b>2. Model specification and estimation .....</b>	<b>1</b>
2.1 Survival analysis and exporter duration .....	1
2.2 Determinants of export failure: A multivariate analysis .....	4
<b>3. The data .....</b>	<b>5</b>
3.1 Descriptive statistics .....	6
<b>4. The results .....</b>	<b>8</b>
4.1 Non-parametric analysis of survival and hazard functions: The Kaplan-Meier estimator and the life table method .....	8
4.2 Plots and specification of the survival function .....	13
4.3 Semi-parametric multivariate analysis .....	13
<b>5. Conclusions and implications.....</b>	<b>17</b>
<b>References.....</b>	<b>19</b>

## ***Abstract***

This paper uses the 1993-2000 Exporter Registry of Statistics Canada to study the factors determining the success or exit of Canadian establishments on foreign markets. The survival analysis model is adopted to study the survival and hazard rates of exporting establishments, with the Cox proportional hazards regression used for the econometric analysis. The study found that one third of establishments export for one month only. Application of the Kaplan-Meier Estimator reveals that the probability of exit before 12 months is 42.2% and the median survival time is 20 months. The hazard of exit is found to vary negatively with the relative size of the establishment, number of exported products and destinations, and the proportion of new entries into export episodes. Establishments exporting to the U.S. Eastern Seaboard face a lower risk of exit than establishments exporting to other U.S. regions and any other destinations. Also, hazard rates vary across provinces of residence and host industry. For example, belonging to an industry other than Manufacturing increases the risk of failure on foreign markets, except for "Agriculture and related services" as well as for "Fishing and trapping, logging and forestry". Belonging to a multi-establishment enterprise and being located in a Census Metropolitan Area appears to increase the risk of exit. These results indicate that destination and product diversification, size, and experience matter for Canadian exporting establishments.

JEL Classification: C41, F14

*Key words:* Exports, success, exit, foreign markets, survival analysis, Canada.

## ***1. Introduction***

Canada relies on trade to create jobs and growth more than any other industrialized country (Team Canada Inc. Business Plan). Canada is a relatively small open economy, which needs access to the global marketplace to thrive. Understanding the factors associated with the success or failure of Canadian exporters on foreign markets is needed to meet the challenges of increasing the number of exporters, broadening Canada's export product range and exporting beyond the U.S. market. On the other side, the openness of the Canadian economy makes it vulnerable to monetary and financial crises happening in the rest of the world (e.g. the Pacific Rim at the end of the 1990's). So, the export market characteristics matter when one analyses the success or failure of the Canadian exporter.

We take advantage of a data set available at Statistics Canada—The Exporter Registry—to undertake evidence-based analysis, to initiate an informed debate and to shed some light on the factors conditioning the success or failure of Canadian exporters on foreign markets.

The purpose of the paper is to identify and to quantify the impact of factors conditioning the survival time of Canadian exporters. The study will investigate the role of establishment and industry characteristics, export market characteristics, market structure, product characteristics, and business cycle in modeling the success or failure of Canadian exporters.

Our analysis is innovative in different respects. First, it uses a unique data set, the Exporter Registry that no other researcher has analyzed in relation to the survivability of Canadian exporting establishments. This implies that the results are new. Second, it relies on a methodology that, although well known in health sciences and industrial organization, has rarely been applied in international trade. Finally, the period analyzed coincides with monetary and financial turbulence in some export destinations (South East Asia, Latin America) as well as Team Canada visits to these destinations to promote trade. This strengthens the need to understand the determinants of success or failure of Canadian exporters on foreign markets.

The rest of the study is organized as follows. Section two discusses the methodology of survival analysis and its relation to export performance for Canadian establishments. The data set is briefly described in section three, while the descriptive and multivariate analyses are presented in section four. The last section offers a summary of the results and their implications for policy makers and exporters.

## ***2. Model specification and estimation***

### ***2.1 Survival analysis and exporter duration***

Survival analysis has recently been applied to a broad range of fields in social sciences. Examples include the failure rates for new Canadian firms (Baldwin *et al.*, 2000), strike duration (Greene, 1993), unemployment duration (Kiefer, 1988), the turnover and mobility of firms (Caves, 1998), new firm survival (Audretsch and Mahmood, 1995) and the duration of business

cycles (Abderrezak, 1997). This paper innovates in applying survival analysis to investigate the ability of Canadian establishments to remain active as exporters. The variable of interest is the length of time that elapses from the beginning of exporting, either until its end or until the measurement is taken, which may precede termination. Observations typically consist of cross-sections of duration  $t_1, t_2, \dots, t_n$ . In this study, the unit of observation is an exporter episode, whose length is measured in months.

Let  $T$ , a continuous random variable, denote the length of a successful episode. The probability of a successful episode continuing until time  $t$  is given by the survival function:

$$S(t) = P(T \geq t) = \int_t^{\infty} f(x)dx = 1 - F(t) \quad (1)$$

where  $f()$  and  $F()$  are the density (pdf) and cumulative distribution (CDF) functions for  $T$ , respectively. The CDF is the probability that an exporter selected at random will have a survival time less than some stated value,  $t$ , or  $F(t) = P(T < t)$ . The probability that an exporter episode ends at time  $t$ , given that it has survived until that time, is defined as the hazard or failure rate:

$$h(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)} = \frac{-d \ln S(t)}{dt} \quad (2)$$

Some authors report on the integrated hazard function, which does not have a convenient interpretation. In particular, it is not a probability. It is defined as:

$$H(t) = \int_0^t h(x)dx \quad (3)$$

Given that  $S(t) = e^{-H(t)}$ , it can be shown that the integrated hazard function is inversely related to the log of the survival function:  $H(t) = -\ln S(t)$ .

The hazard function provides a convenient definition of duration dependence. Positive duration dependence exists at point  $t^*$  if  $\frac{dh(t)}{dt} > 0$  at  $t=t^*$ , which means that the probability that an episode will end shortly increases as the episode increases in length. Negative duration dependence exists at point  $t^*$  if  $\frac{dh(t)}{dt} < 0$  at  $t=t^*$ , which means that the probability that an episode will end shortly decreases as the episode increases in length.

Censoring is an unavoidable problem in the analysis of duration data. Censoring comes in many forms and occurs for many reasons. The basic distinction is between left and right censoring. An observation on a variable  $T$  is right-censored if you only know that it is greater than some value  $c$ . Consider analyzing the survival time of a Canadian exporter. At the time the Exporter Registry for 2000 was assembled, some exporters were still exporting. Then, for these exporters, duration or survival time is right-censored because observation is terminated before the event (exit the

export markets) occurs. In the context of survival data, left censoring occurs when at the beginning of the sample (January 1993) some establishments have already experienced the event (i.e. exited the export markets). In social sciences, left censoring often means something quite different. Observations are left censored if the origin time (the starting date of exporting), not the event time, is known only to be less than some value (January 1993 in this study). Allison (1997, p.10) considers these observations as right censored.

The log-likelihood function can be estimated with a semi-parametric technique (often referred to as Cox regression) or parametrically. The primary advantage of semi-parametric estimation is the absence of a distributional assumption regarding the exporter's ending dates, therefore making it robust. The basic Cox regression or proportional hazard rate model specifies the hazard function as:

$$h(t, x, \beta, h_0) = h_0(t)\phi(x, \beta) \tag{4}$$

where  $h_0(t)$  is the baseline hazard rate characterizing the evolution of the hazard function with respect to survival time and assuming  $\phi(x, \beta) = 1$ ;  $\phi(x, \beta)$  is a shift factor that depends on exporter characteristics ( $x$ ) and their coefficients ( $\beta$ ). Note that the hazard ratio for two exporters with different covariate values depends only on the function  $\phi(x, \beta)$ . For example, let  $\phi(x, \beta)$  be exponential and take the ratio of hazards for two exporters  $i$  and  $j$ , then:

$$\frac{h_i(t)}{h_j(t)} = \exp\{\beta_1(x_{i1} - x_{j1}) + \dots + \beta_k(x_{ik} - x_{jk})\} \tag{5}$$

It is easy to see that  $h_0(t)$  cancels out of the numerator and the denominator and the ratio of hazards is constant over time. If the logs of hazards for any two exporters were graphed against time, the proportional hazards property implies that the hazard functions would be parallel. A disadvantage of the proportional hazards model is that, without a distributional assumption, it is possible to identify only the relative impact of covariates on exporters' duration.

Parametric estimation avoids this drawback by relaxing the assumption that the ratio of hazard rates for any pair of exporter episodes is constant through time. However, one must choose a general functional form for the hazard function. The Weibull distribution is often used in empirical studies because it allows for constant, increasing and decreasing hazard rates.

In this study, graphical analysis based on survival, log-survival and log-log-survival functions is used to discriminate between different functional forms. The data analyzed do not support the exponential and Weibull distributions. Consequently, the multivariate analysis relies then on a Cox proportional hazards model that is valid independently of the distribution form.

## ***2.2 Determinants of export failure: A multivariate analysis***

The analysis assesses the contribution of individual factors in conditioning hazard rates after controlling for other factors. The prime factors that condition an exporter's failure are establishment and industry characteristics. The existing empirical literature (e.g. Bernard and Jensen (1997, 2001), Roberts and Tybout (1997)) has shown that successful exporters are substantially larger, pay higher wages, have higher productivity and are more likely to belong to a multi-plant enterprise or be affiliated with a multinational. Size may be a proxy for several effects. Larger enterprises by definition have been successful in the past (experience), but size may be associated with lower average costs (economies of scale). Size can also be taken as a proxy for market experimentation. It is natural to characterize costs associated with foreign market gambles as sunk costs. These might include the cost of gathering and processing information about demand conditions abroad or costs of establishing a distribution system. Size is proxied by the relative size (SIZEREL) of establishment defined as the ratio of the establishment's monthly value of exports and the average monthly value of exports.

A dummy variable for multi-establishment enterprises (MULTIPLANT) is included to capture ownership effects. As success or failure on foreign markets may vary depending upon the establishment's host industry or province residence, dummies for province (PROVINCE) and industry (INDUSTRY) are included in the model. Location in a Census Metropolitan Area (CMA) may have an influence on how establishments fare in export markets. A dummy for CMA captures this influence.

The export market characteristics are also important in determining the hazard rate of Canadian exporters. To account for this influence, dummies for different destinations are created. The destinations included are the U.S. Eastern Seaboard, U.S. Heartland, U.S. Midwest, U.S. South East, U.S. West, European Union, Japan, Mexico, South America and Other Countries. The dummy variables could pick up many influences including the effects of distance, trade policy, market size, exchange rate variations.

The number of destinations (COFDEST) and the number of products exported in accordance with Canadian Export Classification or HS-8 (COUNTOFHS8) are used to capture diversification on the export markets.<sup>1</sup> The more exports are diversified, the less likely exporters are affected by adverse demand shifts on the export markets.

The survival rate for establishments may depend on the number of establishments entering and exiting export markets. To account for this dynamic process of entry and exit, a variable PROPOFNEW—the proportion of new entrants at the beginning of each export episode—is created.

---

<sup>1</sup> HS-8 is the eight-digit harmonized system commodity classification. It classifies goods on the basis of what they are, not according to their stage of production.



Finally, whether the establishment exited the export markets in the past can exert an effect on the survival time. To account for this factor, dummy variables for the number of previous experiences on foreign markets are added as explanatory variables (UNREPEAT, REPEAT1 ... REPEAT4).<sup>2</sup>

### ***3. The data***

The Exporter Registry consists of domestic export data assigned to the exporting establishments and enterprises, from 1993 to 2000. This assignment is done at the transaction level, which is coming from U.S. imports documents (Canadian exports to U.S.) and Canadian export documents (Canadian exports to non-U.S. destinations). Basic information about exporting establishments and enterprises is also stored in the database. In its detailed version, each domestic export transaction is assigned to its exporter. This means that all variables available at the transaction level (value, commodity classification, province of origin, country of destination, etc.) are available along with information about the exporting establishment and enterprise (name, SIC-E and province of residence). The quality of data varies with the exporter size as the probability of linking documents increases with size (see International Trade Division (2001) for more information).

Published aggregates based on the Exporter Registry 1993-2000 include all establishments exporting \$30,000 or more in value in at least one year from 1993 to 2000. The present study includes all establishments regardless of size. This broader population has a significant impact on the descriptive analysis such as the survival distribution function, as the establishments exporting less than \$30,000 tend to experience shorter export episodes. The inclusion of these small exporters should not affect the multivariate analysis.

As for the structure of data, for each episode in the sample, the dependent variable called DURATION, is defined as the number of months the establishment stayed on export markets or, for censored cases, the number of exporting months up to and including the last month in which that establishment's survival time was observed. A status variable called CENSORVAR is created to distinguish censored establishments (CENSORVAR = 0) from uncensored cases (CENSORVAR = 1). To define DURATION and CENSORVAR, the following convention is adopted: if an establishment stays out of export markets for 12 consecutive months, it is considered exited. The rationale behind this convention is to account for seasonality in export behaviour and thus not to mistakenly consider exited establishments which export once or twice a year. To comply with convention, observations for which the end of exportation sequence is less or equal to 12 or greater or equal to 85 are censored. The rest of the variables are made of covariates or explanatory variables created from the Exporter Registry. In all, there are 107,992 observations, of which 47,015 are censored (43.5%).

---

<sup>2</sup> Additional variables on wages and employment might be relevant for the analysis of Canadian exporters' survival time but are absent from the Exporter Registry.

### 3.1 Descriptive statistics

Table 1 presents the list of variables used in the descriptive analysis as well as the multivariate analysis. The variable to be explained is the duration of export measured in months. The average stay on foreign markets is 22.8 months while the median duration is 7 months.<sup>3</sup> The one-way frequency tabulation reveals that 37.3% of establishments exported for only one month while 5.6% of establishments exported continuously each month from 1993 to 2000.

According to the status variable, CENSORVAR, 56.46% of observations are uncensored while 43.54% of observations are censored.

**Table 1.** Variable names, definitions and means

Variable	Definition	Mean
DURATION	Duration of episode in months	22.800
STARTSEQ	Month the establishment started exporting	36.000
ENDSEQ	Month the establishment stopped exporting	57.800
TOTALVALUE	Total value of exports for the whole duration	16,084,900
NUMBEROFPR	Number of previous episodes	0.356
CENSORVAR	Status variable (CENSORVAR =1 for uncensored cases)	0.566
SIZEM	Monthly value of exports	212,542
SIZEREL	Relative size of establishment	0.837
COUNTOFHS8	Number of exported products	6.060
COFDEST	Number of destinations	2.200
PROPOFNEW	Proportion of new exporting establishments at the start of episode (%)	17.200
EUROPEAN UNION	=1 if European Union is the destination of exports	0.122
JAPAN	=1 if Japan is the destination of exports	0.044
MEXICO	=1 if Mexico is the destination of exports	0.040
SOUTH AMERICA	=1 if South-America is the destination of exports	0.0402
OTHER COUNTRIES	=1 if Other countries are the destination of exports	0.168
U.S.USE	=1 if U.S. Eastern Seaboard is the destination of exports	0.487
U.S.USH	=1 if U.S. Heartland is the destination of exports	0.362
U.S.USM	=1 if U.S. Midwest is the destination of exports	0.347
U.S.USS	=1 if U.S. South East is the destination of exports	0.274
U.S.USW	=1 if U.S. West is the destination of exports	0.356
UNREPEAT	=1 if establishment entered the foreign markets once	0.724
REPEAT1	=1 if establishment entered the foreign markets twice or failed once	0.209
REPEAT2	=1 if establishment failed twice	0.057
REPEAT3	=1 if establishment failed three times	0.010
REPEAT4	=1 establishment failed four times	0.0008
MULTIPLANT	=1 if the establishment is from a multi-establishment enterprise	0.161
CMA	=1 if the establishment is located in a CMA	0.739
PROVINCE	=1 if the exporting establishment is located in that province	---
INDUSTRY	=1 if the exporting establishment belongs to that industry	---

<sup>3</sup> These statistics are calculated using weights to account for unlinked establishments, but without accounting for the censoring factor. Estimation procedures that do not account for the censored nature of the data produce biased estimates.

The average monthly value of exports (SIZEM) is \$212,542, which undoubtedly reflects the influence of large establishments.

As shown in Table 1, the average number of previous episodes is less than one, as the majority of establishments experimented with foreign markets only once. More specifically, indicators for number of previous experiences (UNREPEAT and REPEAT<sub>i</sub>,  $i=1..4$ ) reveal that 72.4% of establishments entered the foreign markets once, 20.9% entered the foreign markets twice or failed once, 5.7% failed twice while 1% failed three times and 0.08% failed four times.

It is no surprise that U.S. regions are by far the preferred destination, with 48.7% of exporters selling to the U.S. Eastern Seaboard, 36.2% of exporters selling to the U.S. Heartland, 35.6% of exporters selling to the U.S. West, 34.7% of exporters selling to the U.S. Midwest and 27.4% of exporters selling to the U.S. South East, respectively. Each U.S. region outperforms any other economic bloc as a destination for Canadian exporters, except for the Other Countries bloc that comprises 203 countries.

It can be inferred from the smaller mean number of destinations (2.2) and the mean number of the products exported (6.06) that Canadian exporting establishments are not very diversified in terms of destinations and products.

Inspection of the province of residence variables reveals that four provinces—Ontario, Quebec, British Columbia, Alberta—account for 88.18% of exporting establishments with respective shares of 43.74%, 20.41%, 14.77% and 9.26%. The rest is distributed as follows: 0.6% in Newfoundland, 0.33% in Prince Edward Island, 2.14% in Nova Scotia, 1.89% in New Brunswick, 4.01% in Manitoba, 2.64% in Saskatchewan, 0.11% in Yukon, 0.06% in Northern Territories and 0.02% in Nunavut. Of these exporting establishments, 73.9% are located in metropolitan areas.

As for the type of organization to which the exporting establishments belong, 16.1% belong to a multi-establishment and multi-location enterprise, 12.6% to a multi-establishment, multi-location and multi-provincial enterprise and 8% to a multi-establishment, multi-location, multi-provincial and multi-activity enterprise.

Finally, the distribution of exporting establishments between industries is uneven. Manufacturing (34%) and Wholesale trade (27.1%) are the front runners. In descending order, we then find Retail trade (8%), Business service (7.3%), Construction, transportation and storage (6.3%), Other (5.8%), Agriculture and related services (5.5%), Finance and insurance (2.6%), Mining, quarrying and oil well (1%), Fishing and trapping, logging and forestry (1%), and finally Communication and other utility (0.7%).

## 4. The results

### 4.1 Non-parametric analysis of survival and hazard functions: The Kaplan-Meier estimator and the life table method

The Kaplan-Meier estimator (KME) is a non-parametric method used for estimating survival functions. This estimator incorporates information from all of the observations available both, uncensored and censored, by considering survival at any point in time as a series of steps defined by the observed survival and censored times. At any specified time point ( $t_i$ ), the number of establishments still exporting is called the number at risk. Let  $n_i$  and  $d_i$  be the number of establishments at risk and the number of exporters, which exited the export markets, the survival rate can be expressed as:

$$S(t) = \prod_{t_i \leq t} \left( \frac{n_i - d_i}{n_i} \right) = \prod_{t_i \leq t} \left( 1 - \frac{d_i}{n_i} \right) = \prod_{t_i \leq t} (1 - h_i) \quad (6)$$

where  $h_i = \frac{d_i}{n_i}$  is the hazard rate or the ratio of the number of exits and the number of exporters at risk at  $t_i$ . Survival and exit probabilities are computed for establishments that exported to foreign markets over the January 1993-December 2000 period.

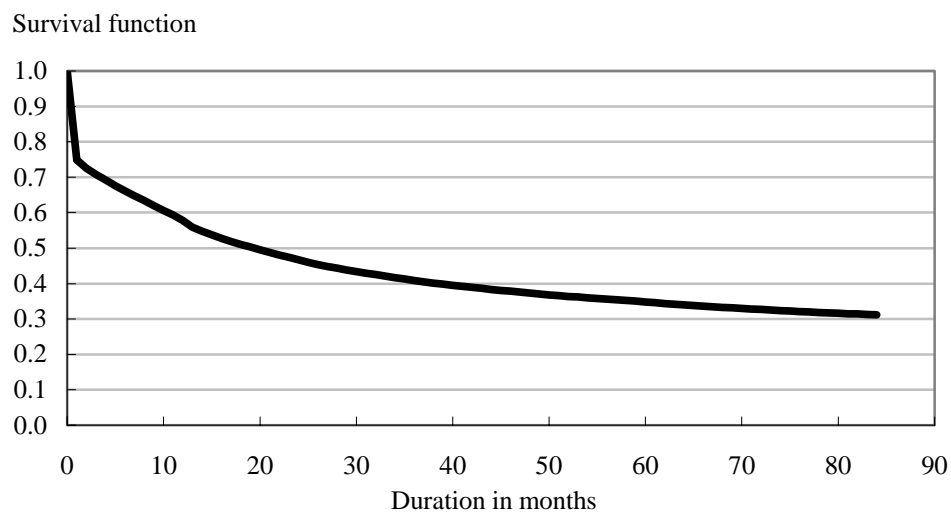
Table 2 shows the Kaplan-Meier (KM) estimates for selected months of export duration. The two-right hand columns give the lower and upper limits for a 95% confidence interval around each survival probability. The first column is the duration of export in months. The second column contains the KM estimates. At 1 month, for example, the estimate is 0.748. This means that the probability that an establishment will survive for 1 month or more is 74.8%. By the same token, the probability that an exporter will survive for 12 months or more is 57.8%. The third column labelled exit probability is just one minus the KM estimate, which is the probability of exiting the foreign markets during the specified number of months. Thus, the probabilities of stopping during the first month and the first 12 months are 25.2% and 42.2%, respectively. The median survival time is a better measure of central tendency than the mean survival time for censored survival data. For the KM estimates, the median survival time for Canadian establishments is 20 months, which is well below the mean survival of 36.7 months. The survival distribution function for the Exporter Registry data is displayed in Figure 1.

A popular alternative to the KME is the life table method in which event times are grouped into intervals of arbitrary widths. The treatment of cases censored within an interval as if they were censored at the midpoint of the interval is particularly problematic when the intervals are wide. The life table for the Exporter Registry data is shown in Table 3, using an interval of 6 months of export. The exporter survival times are strongly skewed toward brief duration, with 50% of all episodes ending within the 18-24 months interval. The conditional probability of exit is an estimate of the probability that an exporter will exit export markets, given that it made it to the start of the interval (i.e., number failed/effective sample size). The effective size of the sample is the number of exporters who have not exited at the start of the interval. For example, the conditional probability for the interval 30-36 is 5.8%.

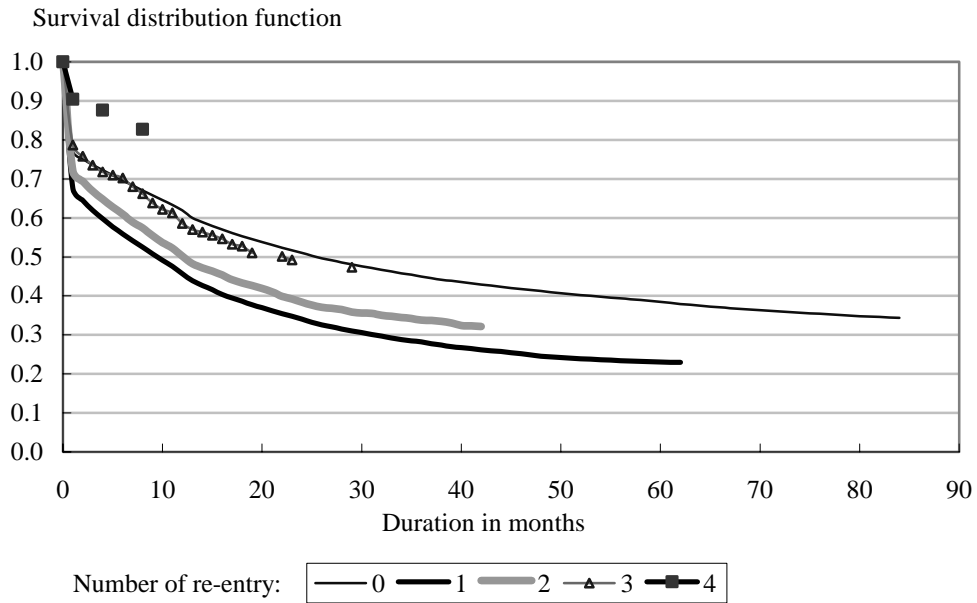
**Table 2.** Kaplan-Meier estimates for the Exporter Registry data

Duration	Survival probability	Exit probability	SDF_LCL	SDF_UCL
0	1.000	0.000		
1	0.748	0.252	0.746	0.751
2	0.725	0.275	0.722	0.728
3	0.707	0.293	0.704	0.709
4	0.691	0.309	0.689	0.694
5	0.676	0.324	0.673	0.679
6	0.661	0.339	0.658	0.664
7	0.647	0.353	0.644	0.650
8	0.633	0.367	0.630	0.636
9	0.620	0.380	0.617	0.623
10	0.606	0.394	0.603	0.609
11	0.593	0.407	0.590	0.596
12	0.578	0.422	0.575	0.581
18	0.511	0.489	0.508	0.514
24	0.468	0.532	0.465	0.471
30	0.434	0.566	0.431	0.437
36	0.409	0.591	0.406	0.412
42	0.389	0.611	0.386	0.393
48	0.373	0.627	0.370	0.377
54	0.360	0.640	0.357	0.364
60	0.348	0.652	0.345	0.352
66	0.336	0.664	0.333	0.339
72	0.327	0.673	0.324	0.333
78	0.319	0.681	0.315	0.322
84	0.312	0.688	0.309	0.316

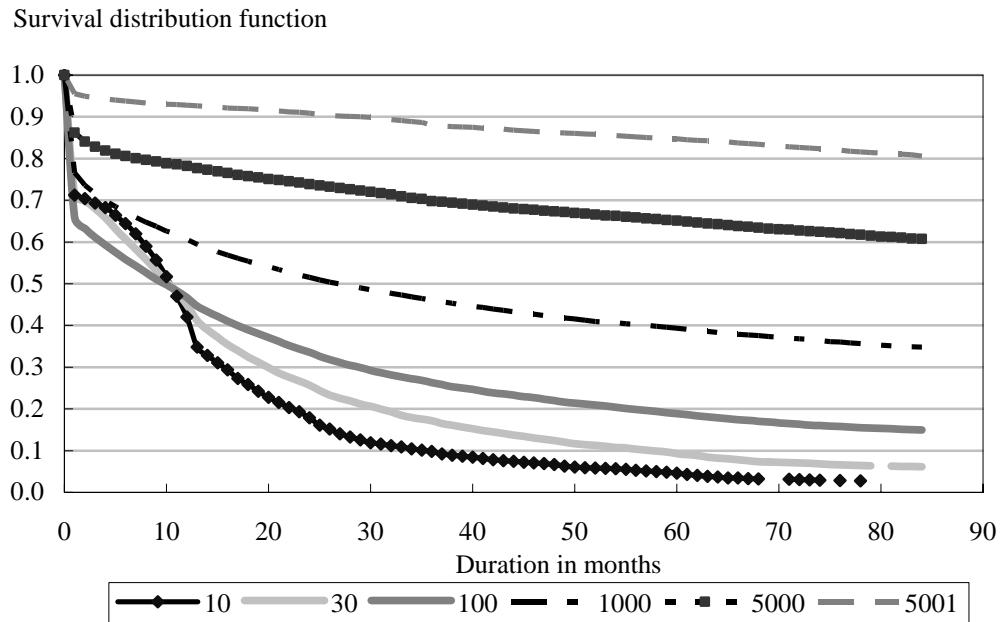
**Figure 1.** Survival function for the Exporter Registry data



**Figure 2. Survival curves by re-entry into export markets**



**Figure 3. Survival curves by size of the establishment**



**Table 3.** Life table for the Exporter Registry data

Interval	Effective size	Conditional probability of exit	Survival	Exit	Hazard
0-6	101,184.0	0.335	1.000	0.000	0.067
6-12	58,654.5	0.123	0.665	0.335	0.022
12-18	48,434.5	0.125	0.583	0.417	0.022
18-24	40,278.0	0.086	0.511	0.490	0.015
24-30	35,151.0	0.076	0.466	0.534	0.013
30-36	31,012.0	0.058	0.431	0.569	0.010
36-42	27,841.0	0.051	0.406	0.594	0.009
42-48	25,197.5	0.042	0.385	0.615	0.007
48-54	22,902.5	0.037	0.369	0.631	0.006
54-60	20,658.0	0.033	0.356	0.645	0.006
60-66	18,641.0	0.036	0.344	0.656	0.006
66-72	16,842.0	0.029	0.332	0.668	0.005
72-78	15,327.0	0.026	0.322	0.678	0.004
78-84	13,919.0	0.021	0.314	0.686	0.004
84-90	12,283.0	0.003	0.307	0.693	0.001
90-96	9,245.5	0.000	0.306	0.694	0.000

The survival column is the life table estimate of the survival function, i.e., the probability that the event occurs at a time greater or equal to the starting time of each interval. For example, the estimated probability that an exporter will not exit until month 12 is of 58.3%, which is close to the probability found with the KME (57.5%). Exit is just one minus the survival column. In this example, the corresponding exit rate is 41.7%, which is again close to the KME (42.2%). The hazard column gives an estimate of the hazard function at the mid-point of each interval. So, the hazard rate for the 12-18 interval is 2.22%. It can be concluded that the KME and the life table with 6-month intervals lead to similar conclusions.

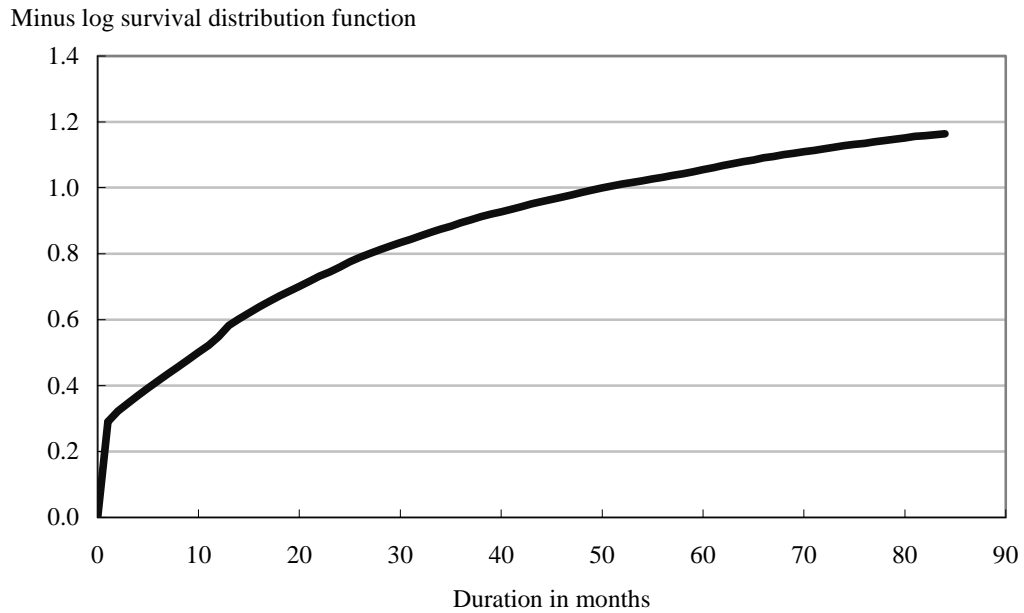
To check for the impact of previous exits on survival time, the number of previous episodes (NUMBEROFPR) is used as a stratification variable in generating the survival curves. These are presented in Figure 2. They suggest that exits exert a negative effect on survival time only for establishments that exited the foreign markets once or twice. One can conjecture that there is learning by exiting since the establishments that exited the export markets more than twice fare better than the ones that entered export markets once.

The annualized average export value is used to create SIZE CLASSES. The classes are:

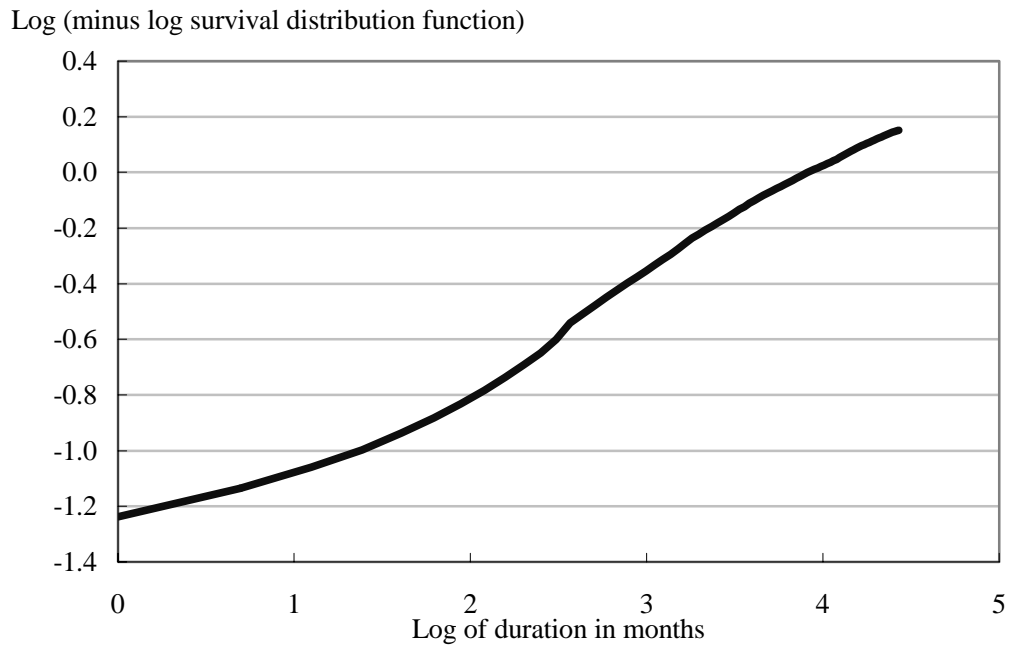
- 10 = < \$10,000
- 30 = \$10,000 to \$30,000
- 100 = \$30,000 to \$100,000
- 1,000 = \$100,000 to \$1,000,000
- 5,000 = \$1,000,000 to \$5,000,000
- 5,001 = >\$5,000,000

Survival curves by size classes are displayed in Figure 3. There is evidence of a positive relationship between size and survival times. However, the survival curves for establishments whose export values are less than \$100,000 are superimposed at earlier times, meaning that they may have the same survival rates or exit rates.

**Figure 4. Log-survival plot for the Exporter Registry data**



**Figure 5. Log-log-survival plot for the Exporter Registry data**





## 4.2 Plots and specification of the survival function

Graphs can be used to display data on duration and to guide the selection of a functional form describing the data. The Log-survival (LS) plot is a plot of  $-\log S(t)$  versus  $t$ . Recall that

$-\log S(t) = \int_0^t h(x)dx$  is the cumulative hazard function or integrated hazard function.

Examination of the LS plot can tell whether the hazard rate is constant, increasing or decreasing with time. A convex integrated hazard implies that the hazard itself is increasing, which is typically referred to as positive dependence. A concave integrated hazard implies a decreasing hazard or negative dependence. The LS plot for the Exporter Registry data is shown in Figure 4. The concave function implies that the hazard rate declines with time. This rules out the possibility that the data be described by an exponential function which would be consistent with a linear relationship through the origin.

The log-log-survival (LLS) plot is a plot of  $\log[-\log S(t)]$  versus  $\log(t)$ . If the survival times followed a Weibull distribution with a hazard given by  $\log h(t) = \beta_0 + \beta_1 \log t$ , then the LLS plot would be a straight line with a slope of  $\beta_1$ . The LLS plot in Figure 5 is convex early on and concave later. Therefore, the Weibull distribution cannot adequately describe the survival times of exporting establishments.

## 4.3 Semi-parametric multivariate analysis

As mentioned in Section 2, the multivariate analysis assesses the contribution of individual factors in conditioning the hazard rates after controlling for other factors. The proportional hazards model developed by Cox is applied to the Exporter Registry data. Equation (4) is estimated with the partial likelihood method and the Efron's method is used to handle ties. When computing the proportional hazards regression, the U.S. Eastern Seaboard is the omitted category for export destinations; Ontario is the reference for province of residence; manufacturing is the reference for industry indicators while "entered foreign markets once" is the omitted category for number of previous experiences indicators.

The estimation results and summary statistics are shown in Tables 4 and 5, respectively. Note that Table 4 contains no intercept—a characteristic of the partial likelihood estimation. The intercept is part of  $h_o(t)$  which appears in both the numerator and denominator in equation (4). This implies that group-specific hazard rates cannot be recovered from the regression output of the proportional hazards model. The second and third columns in Table 4 are the parameter estimates (Betas in equation (4)) and their associated standard errors. The p-values reported in column 4 indicate the probability of obtaining a test statistic whose absolute value is greater than or equal to that of the sample statistic if the null hypothesis of a zero coefficient is true. Thus a p-value lower than the reference significance level is taken as evidence to reject the null of a zero coefficient.

**Table 4.** Influence of selected determinants on hazard: A Cox proportional hazards regression

Variable	Parameter estimate	Standard error	Pr > Chi-square	Hazard ratio
SIZEREL	-0.0078	0.0016	<.0001	0.992
COUNTOFHS8	-0.1286	0.0016	<.0001	0.879
COFDEST	-0.6311	0.0098	<.0001	0.532
PROPOFNEW	-0.0069	0.0002	<.0001	0.993
European Union	0.2394	0.0168	<.0001	1.270
Japan	0.0654	0.0280	0.0195	1.068
Mexico	0.5247	0.0537	<.0001	1.690
South America	0.3429	0.0306	<.0001	1.409
Other countries	0.1756	0.0143	<.0001	1.192
U.S. Heartland	0.0644	0.0137	<.0001	1.067
U.S. Midwest	0.1341	0.0143	<.0001	1.144
U.S. South East	0.2734	0.0153	<.0001	1.314
U.S. West	0.1501	0.0141	<.0001	1.162
Newfoundland	0.0854	0.0490	0.0814	1.089
Prince Edward Island	-0.2213	0.0753	0.0033	0.802
Nova Scotia	-0.1621	0.0296	<.0001	0.850
New Brunswick	-0.2424	0.0313	<.0001	0.785
Quebec	-0.0398	0.0112	0.0004	0.961
Manitoba	-0.1309	0.0232	<.0001	0.877
Saskatchewan	-0.0200	0.0266	0.4518	0.980
Alberta	-0.0058	0.0157	0.7102	0.994
British Columbia, Yukon, Northern Territories and Nunavut	-0.1849	0.0141	<.0001	0.831
Multiplant	0.1670	0.0117	<.0001	1.182
CMA	0.1153	0.0103	<.0001	1.122
Agriculture and related services	-0.2220	0.0204	<.0001	0.801
Fishing and trapping, logging and forestry	-0.1262	0.0410	0.0021	0.881
Mining, quarrying and oil well	0.1756	0.0388	<.0001	1.192
Construction, transportation and storage	0.2965	0.0172	<.0001	1.345
Communication and other utility	0.3170	0.0430	<.0001	1.373
Wholesale trade	0.2087	0.0109	<.0001	1.232
Retail trade	0.3359	0.0162	<.0001	1.399
Finance and insurance	0.3048	0.0252	<.0001	1.356
Business service	0.4701	0.0158	<.0001	1.600
Other	0.4010	0.0179	<.0001	1.493
Repeat1	0.0772	0.0098	<.0001	1.080
Repeat2	-0.1590	0.0186	<.0001	0.853
Repeat3	-0.5511	0.0518	<.0001	0.576
Repeat4	-1.6029	0.3016	<.0001	0.201

All coefficients are significant at the 1% significance level except for the coefficients on Japan and Newfoundland which are significant at the 1.95% and 8.14% levels respectively and for the coefficients on Saskatchewan and Alberta which are insignificant at conventional levels. The parameter estimates suggest that an increase in the relative size of the establishment, the number of exported products and destinations, and the proportion of new establishments entering an export episode, decrease the hazard of exit and, thus, increase the likelihood of survival on foreign markets.

Compared to exporting to the U.S. Eastern Seaboard, exporting to other U.S. regions and other economic blocs or countries increases the hazard of failing on foreign markets.

Having a province of residence other than Ontario lowers the hazard of exit, except for Newfoundland. It is worth emphasizing that the impacts of Saskatchewan and Alberta are not statistically significant. Given that previous research on failure rates for new Canadian firms (Baldwin *et al.*, 2000) has found that firms located in Ontario are more likely to survive than firms located in other provinces, this result must be taken with caution.

Belonging to a multi-plant enterprise increases the hazard of failure on export markets. This runs against the existing empirical literature (Bernard and Jensen (1997, 2001), Roberts and Tybout (1997)). Export standards might require large sunk costs that would encourage firms not to export from all their plants. This is especially true for food exports. Location in a CMA appears to increase the hazard of exit and thus decrease the survival time.<sup>4</sup>

Relative to manufacturing, establishments belonging to other industries face a higher risk of exiting the export markets, except for "Agriculture and related services" as well as for "Fishing and trapping, logging and forestry".

Finally, in accordance with the analysis based on survival curves by re-entry into export markets, exiting more than twice appears to decrease the hazard of exit. So, establishments learn from past failures. The hazard ratio will help out with quantifying these qualitative results.

The last column in Table 4 is the hazard ratio ( $e^{\beta}$ ). For example, the hazard ratio for COUNTOFHS8 is  $e^{-0.12860} \cong 0.879$ . For quantitative covariates like COUNTOFHS8 (number of exported products according to HS8), a more helpful statistic is obtained by subtracting 1 from the hazard ratio and multiplying by 100. This gives the estimated percentage change in the hazard for each one-unit increase in the covariate. This means that a one-unit increase in the number of exported products (COUNTOFHS8) decreases the hazard of exiting by 12.1%. The interpretation for the other quantitative variables is similar. Therefore, a percentage point increase in the relative size of the establishment (SIZEREL) induces a decrease of 0.8% in the hazard of failure on foreign markets. The impact is even stronger for the number of destinations. A one-unit increase in the number of destinations lowers the hazard of exit by 46.8%. The firmer influence of the number of destinations and products exported favour the market diversification hypothesis as a way to meet the challenge of foreign market gambles. A percentage point rise in the proportion of new entrants at the beginning of an export episode reduces the hazard of exit by 0.7%.

For indicator or dummy variables with values of 1 and 0, the hazard ratio is interpreted as the ratio of the estimated hazard for those with a value of 1 to the estimated hazard for those with a value of 0, controlling for other covariates. For example, the hazard for the variable European Union is 1.27. This suggests that the hazard of exit for establishments that exported to the European Union is 27% higher than the hazard of exit for establishments that exported to the

---

<sup>4</sup> Further research is needed to identify differences between establishments located in CMAs and non-CMAs.

U.S. Eastern Seaboard. Similarly, the hazards of exit for establishments that exported to Japan, Mexico, South America and other countries are respectively 6.8%, 69.0%, 40.9% and 19.2% higher relative to establishments concentrating on the U.S. Eastern Seaboard. The hazards of exit for establishments which exported to the U.S. Heartland, Midwest, South East and West are respectively 6.7%, 14.4%, 31.4% and 16.2% higher than that for establishments which exported to U.S. Eastern Seaboard.

As mentioned before, Ontario is the reference province in the proportional hazards regression. In general, establishments located in other provinces are less likely to fail on foreign markets than establishments residing in Ontario, except for Newfoundland, where the hazard of exit is 8.9% higher relative to Ontario. Establishments residing in three Maritime Provinces (New Brunswick, Prince Edward Island and Nova Scotia) have the lowest hazard ratios compared to establishments located in Ontario (i.e. 78.5%, 80.2% and 85.0% respectively). Establishments located in CMAs experience a hazard of exit that is 12.2% higher than that of non-CMA establishments while establishments belonging to a multi-establishment and multi-location enterprise experience a hazard ratio that is 18.2% higher than that of the establishments that do not have this complexity.

After controlling for other factors, the hazard of exit varies across industry groupings. For example, the hazard ratio in business service, retail trade, communication and other utility, finance and insurance, construction, transportation and storage, and wholesale trade is 60.0%, 39.9%, 37.3%, 35.6%, 34.5% and 23.2% higher than that of manufacturing. On the other hand, the hazards of exiting the export markets in "Agriculture and related services" as well as in "Fishing and trapping, logging and forestry" are 80.1% and 88.1% lower than that of manufacturing. It can be concluded that, relative to manufacturing, the hazard of exit is higher in service-providing industries than in goods-producing industries.

As the export activity may be questionable for services industries, the econometric analysis is also implemented to a sample of goods producing industries. There is no sign reversal in the coefficients. Generally speaking, the results are qualitatively the same. It can be concluded that including or excluding services industries does not make a significant difference.

The summary statistics reported in Table 5 include, among others, three tests—a likelihood ratio, a score and a Wald test—designed to assess the overall significance of the coefficients. The null hypothesis is that all coefficients are zero. All three statistics are over a  $\chi^2_{38,0.01}=60.35$ , which means that at least one coefficient is different from zero. The pseudo- $R^2$  is an  $R^2$ -type measure of goodness of fit for the model. Although it is well known that there is no consensus among statisticians or econometricians on the best way to calculate an  $R^2$  for non-linear models including the proportional hazards model, Cox and Snell (1989) proposed a generalized  $R^2$  that was later endorsed by Maggee (1990). The likelihood ratio chi-square statistic that tests the null hypothesis that all covariates have zero coefficients proxies the  $G^2$  in Maggee's formula. The  $R^2$

is computed as  $R^2 = 1 - \exp\left\{-\frac{G^2}{n}\right\} = 1 - \exp\left\{-\frac{70607.94}{107992}\right\} \cong 1 - 0.5200 = 48\%$ .

**Table 5.** Summary statistics

Model fit statistics			Testing global hypothesis: Beta=0			
Criterion	Without covariates	With covariates	Test	Chi-square	DF	Pr > ChiSq
-2 LOG L	1,342,338.7	1,271,730.8	Likelihood ratio	70,607.94	38	<.0001
AIC	1,342,338.7	1,271,806.8	Score	45,334.69	38	<.0001
SBC	1,342,338.7	1,272,149.5	Wald	35,112.02	38	<.0001
Pseudo-R <sup>2</sup> = 0.48						

Unlike for the linear model, this R<sup>2</sup> cannot be interpreted as the proportion of variation in the dependent variable (duration) explained by the covariates. It is just a number between 0 and 1 that is larger when covariates are strongly associated with the dependent variable.

## 5. Conclusions and implications

This paper identifies and quantifies the impact of factors conditioning the survival time of Canadian establishments on foreign markets. Establishment and industry characteristics, export market destinations, market structure, province and CMA locations, and enterprise structure are investigated.

Toward that end, survival analysis techniques are implemented on the 1993-2000 Exporter Registry data set of Statistics Canada. Censoring is an important issue given that many establishments were exporting at the beginning and/or at the end of the sample.

It was found that one third of establishments export for one month only. It is conjectured that they may export to meet an unsolicited request. Based on the Kaplan-Meier Estimator, the probability of exit before 12 months is 42.2% and the median survival time is 20 months. Of all the establishments, 72.4% experimented with foreign markets only once. The main export destinations for Canadian exporters are still U.S. regions. The distribution of exporting establishments is uneven between provinces, with four provinces (Ontario, Quebec, British Columbia and Alberta) picking up 88.2% of establishments. The same can be said for the distribution between industries, with manufacturing (34%) and wholesale trade (27.1%) way ahead of the rest. Beside destinations, the exporting establishments are also concentrated in terms of number of exported products.

A set of conclusions can be drawn from the multivariate analysis. First, the hazard of exiting the foreign markets is found to vary negatively and significantly with the relative size of the exporting establishment, the number of exported products and destinations, and the proportion of new establishments entering export episodes. Second, establishments exporting to the U.S. Eastern Seaboard face a lower risk of exit than establishments exporting to other U.S. regions and to any economic bloc or country. Third, hazard rates vary across provinces of residence and host industry. For example, belonging to an industry other than manufacturing increases the risk of failure on foreign markets, except for "Agriculture and related services" as well as for "Fishing and trapping, logging and forestry". Fourth, there is some evidence that the number of previous exits on foreign markets affects the survival time. Hence, establishments learn from

past failures. Finally, belonging to a multi-establishment enterprise and being located in a CMA does not seem to improve the probability of survival. It remains to be tested whether plants located closer to the border are designated to export.

Understanding the determinants of failure or success of exporting establishments on foreign markets is of paramount importance to a small open economy like Canada that relies on international trade and investment to enhance its economic growth. This paper contributes to this understanding by providing empirical support for the diversification paradigm. At the same time, it highlights the influence of industry, location and destination characteristics on the performance of Canadian exporting establishments. These results could be useful to policymakers and export promotion managers.

## ***References***

Abderrezak, Ali. 1997. Consumer Expectations and Cyclical Durations. *Quarterly Review of Economics and Finance*, vol. 37, pp. 843-857.

Allison, P. 1997. *Survival Analysis Using the SAS System: A Practical Guide*. Cary: SAS Institute Inc.

Audretsch, D. and T. Mahmood. 1995. New Firm Survival: New Results Using a Hazard Function. *Review of Economics and Statistics*, vol. 77, pp. 97-103.

Baldwin, J.R., L. Bian, R. Dupuy and G. Gellatly. 2000. *Failure Rates for New Canadian Firms: New Perspectives on Entry and Exit*. Catalogue no. 61-526-XPE. Analytical Studies Branch. Ottawa: Statistics Canada.

Bernard, A.B. and J.B. Jensen. 1997. Exporters, Skill-upgrading and the Wage Gap. *Journal of International Economics*, vol. 47, pp. 1-25.

Bernard, A.B. and J.B. Jensen. 2001. *Why Some Firms Export*. NBER Working Paper n° 8349.

Caves, R.E. 1998. Industrial Organization and New Findings on the Turnover and Mobility of Firms. *Journal of Economic Literature*, vol. 34, pp. 1947-1982.

Chen, M.-Y. 2002. Survival Duration of Plants: Evidence from the U.S. Petroleum Refining Industry. *International Journal of Industrial Organization*, vol.20, pp. 517-555.

Cox, D.R. and E.J. Snell. 1989. *The Analysis of Binary Data*, Second Edition, London: Chapman and Hall.

Dick, Andrew R. 1993. *Information, Enforcement Costs and Cartel Stability: An Empirical Investigation*. UCLA Department of Economics Working Paper: 698.

Greene, William H. 1993. *Econometric Analysis*, Second Edition. New York: McMillan.

Hoffman, Emily. 1992. Estimation of Length of Job Search by Survival Analysis. *Eastern Economic Journal*, Oct.-Dec. 1991, vol. 17, pp. 393-401.

Hosmer, D.W. and S. Lemeshow. 1999. *Applied Survival Analysis*. New York: John Wiley & Sons.

International Trade Division. 2001. A profile of Canadian exporters. Catalogue no. 65-506-XIE. Ottawa: Statistics Canada.

Kiefer, Nicholas M. 1988. Economic Duration Data and Hazard Functions. *Journal of Economic Literature*, vol. 26, pp. 646-679.

Maggee, L. 1990.  $R^2$  Measures Based on Wald and Likelihood Ratio Joint Significance Tests. *The American Statistician*, vol. 44, pp. 250-253.

Roberts, M. and J. Tybout. 1997. An Empirical Model of Sunk Costs and the Decision to Export. *American Economic Review*, vol. 87, pp. 545-564.

Scully, Gerald W. 1994. Managerial Efficiency and Survivability in Professional Team Sports. *Managerial and Decision Economics*, vol. 15, pp. 403-411.

Team Canada Inc. Business Plan 1999-2002. <http://exportsource.gc.ca>.