Duration of firms in an infant industry: the case of Indian computer hardware [1]

Sanghamitra Das [a,*], Krishna Srinivasan [b]

[a] *Indian Statistical Institute, 7 SJS Sansanwal Marg, New Delhi 110016, India*
[b] *African Department, International Monetary Fund, 700 19th Street NW, Washington, DC 20433, USA*

Abstract

This study analyses survival of firms in an infant industry in a developing economy. It is found that (a) entry size and the probability of exit are positively related, which is the opposite of the result in most previous studies; (b) similar to previous studies, post-entry size is negatively related to the hazard; (c) firms with longer duration have a more volatile time path of size; (d) the hazard declines with age at an increasing rate suggesting increasing returns to learning and (e) diversification of firms and public ownership have little effect on duration.

JEL classification: L11; L63

Keywords: Duration of firms; Computer industry; Exit; Hazard

* Corresponding author.

[1] We are thankful to Satya Das, Sanjiv Jaggia and participants at the conference on "Post-Entry Performance of Firms" at Lisbon, Portugal in May 1994, for helpful discussions and to two anonymous referees for their helpful comments. The data collection was carried out while the first author was a Fulbright Research Fellow at the Indian Statistical Institute, New Delhi. The support of the Fulbright Foundation, the US Educational Foundation in India and the host institution are gratefully acknowledged by the first author. The usual disclaimer applies. The views expressed here are those of the authors and should not be attributed to the institutions of affiliation.

1. Introduction

This paper examines how the duration (life-time) of firms in an industry and its hazard (probability) of exit from the industry are related to the characteristics of firms. It studies an infant industry in a developing country—the computer hardware industry in India. The product consists of computer systems and peripherals. This is an important industry because it affects, in a major way, the productivity of other sectors of the economy and has received policy favors in recent years. The industry 'took off' at the beginning of our sample period, 1982–83 to 1991–92, with the initiation of computerization in the public sector and substantial lowering of entry barriers in the computer industry leading to a more than ten-fold increase in the number of firms over this period. The industry association for computer hardware firms in India (Manufacturers' Association of Information Technology—MAIT) calls the post-1983 period the new era for the industry (MAIT, 1991). Hence the industry can be interpreted as an infant industry over our sample period.

It may be remarked that we do not know of similar studies of duration or exit in the computer industry in other countries. Most previous studies of exit have used data from the manufacturing sector as a whole (e.g. Dunne et al., 1989, henceforth DRS and Pakes and Ericson, 1990). Needless to say, there are merits in studying a narrowly defined industry in which firms face similar industry characteristics, even though the results for the industry may not be applicable to other industries. The established empirical regularities of previous studies, all relating to developed countries, cannot be assumed to hold for the Indian computer industry. In what follows we use a reduced-form approach to study the patterns in the industry and check to see if the results are consistent with any of the existing theories or if new models are necessary to explain them.

The theoretical models of exit focus on the relationship of exit with size and/or age. This has guided the choice of independent variables in the empirical work on exit to consist of size and/or age. [2] A general finding is that exit is negatively related to current size and age, while the results for entry size are mixed. Using bi-annual US data on small businesses, Audretsch and Mahmood (1992) find exit to be negatively related to entry size, whereas DRS find the opposite for data on US manufacturing plants considering exit over a 5-year period over the census years 1967, 1972 and 1977. For annual Portuguese data, Mata and Portugal (1994) find a negative relationship between exit and entry size.

Duration models provide an alternative method for studying exit. The dependent variable is the duration of a firm in the industry by the end of the sample period. Hence each firm corresponds to one observation, which in turn does not correspond to any particular point in time. Therefore, in duration models the

[2] In addition, industry and firm/plant specific dummies may be present.

independent variables are not measured at a point in time but are taken as averages over the measured duration (e.g. Bandopadhyay, 1994). [3] Hence for studying the duration (or equivalently the hazard) of firms, we take average size over the measured duration as the post-entry size. Instead of relating the probability of exit (hazard) at a point in time to size at that time, duration models relate duration (or hazard) to its average size over the observed survival period.

Similar to some previous studies we also take entry size as an explanatory variable. But unlike all previous studies, both theoretical and empirical, whose focus was not on infant industries per se, we consider an additional aspect of size, i.e. the volatility of size over the duration. Infant industries are likely to be more dynamic than mature industries due to greater uncertainty, changing information flows and learning. These factors are likely to lead to volatility in firm sizes. Hence, we examine whether volatility is related to hazard or the probability of exit. If there is a relationship then future work should attempt to model this. In summary then, we examine the relationship of exit to three facets of size—entry size, post-entry average size and the volatility in size (measured by the covariation in size). Apart from size, we consider the relationship of exit to diversification (whether or not a firm is diversified into other industries besides computer) and ownership (whether the firm is in the public sector or private sector). Finally, the effect of age on hazard (probability of exit) is given by the nature of duration dependence and hence does not explicitly appear in the vector of explanatory variables.

In short, theoretical analyses of exit, which have guided empirical work, relate exit to entry size, post-entry size and age. They do not have any implication for the volatility of size. But in infant industries, which are more common in developing countries than in developed countries, size is likely to be more volatile than in the stable environment of mature industries. Hence, we check whether exiting firms are associated with a more or less volatile size path. Besides volatility of size the other variable not considered in previous empirical studies is whether a firm is in the public or private sector. This variable is relevant for a developing country like India where the government plays an active role in the production process unlike in the developed countries. Of course, the impact of these additional variables on exit is likely to be industry specific.

Our main findings are: (a) entry size is positively related to the probability of exit (hazard); (b) post-entry size is negatively related to hazard; (c) firms with longer duration have a more volatile time path of size; (d) hazard declines with time (age) at an increasing rate and (e) diversification as well as ownership have little effect on hazard. A discussion of these results is given later.

[3] The measured duration of a firm is the true duration if the firm has exited the industry during the sample period. Otherwise, it is a right-censored random variable and the likelihood function is accordingly adjusted.

2. Data

Data were compiled from various issues of a computer magazine, *Dataquest*, published by H.C. Gupta on behalf of Cyber Media (India) Ltd., New Delhi. The magazine reports for each firm its sales in computer hardware in that year. It also provides information on its birth year, the type of ownership of a firm (public or private) and whether it is diversified into other industries besides computer. Sales data are obtained directly from the firms and are expected to be reliable. The data set includes 74 firms over the sample period. These firms make up almost the entire industry sales in every year. A firm is said to 'exit' when its sales become low enough (perhaps 0) for it not to be reported in *Dataquest*. Once a firm is not listed, it is not observed to come back into the listings. Hence it is reasonable to treat non-listing as exit. There were only a couple of exits that occurred due to mergers. Furthermore, one firm changed its name and is treated as a continuing firm.

Table 1 reports the frequency distribution of duration of the 74 firms. Among the 74 firms in the industry, 38 firms entered the industry at different points in time since 1982–83 and exited by the year 1992–93, while 36 firms entered the industry since 1982–83 but did not exit by the year 1992–93. For the firms which exited the industry, the survival time varies from 1 year to 7 years. A large number of firms exit within 5 years, with the maximum number of exits occurring when a firm is 4 years old. The observed hazard of firms varies anywhere from 0 to 21%. Since we have a large number of censored observations the hazard is computed by the procedure described in Kiefer (1988). It may be tempting to conclude from Table 1 that the hazard function is nonmonotonic in age (holding constant other explanatory variables). This cannot be presumed because in Table 1 as age

Table 1
Frequency distribution of duration

Duration of firms or age (in years)	Number of firms exiting	Number of firms censored	Observed hazard
1	7	2	0.09
2	7	3	0.11
3	8	5	0.15
4	9	4	0.21
5	1	4	0.03
6	3	2	0.13
7	3	4	0.16
8	0	3	0.00
9	0	7	0.00
10	0	2	0.00
Total	38	36	

changes so do the other explanatory variables. The hazard function may or may not be nonmonotonic with age, ceteris paribus.

Size of a firm is measured as real sales in 100 000 rupees deflated by the consumer price index (CPI) (base year = 1985). [4] The explanatory variables are entry size, covariation in size, diversification and ownership. The last two are dummy variables which take value 1, respectively, if a firm is diversified and if it is publicly owned.

3. Econometric analysis

Duration analysis typically begins with specifying the hazard function directly. The hazard $h(t)$ represents the instantaneous probability that a firm will exit the industry in period $t+\delta$ for a small δ, given that it has survived up to time t. It is defined as $f(t)/S(t)$, where $f(t)$ is the density function of t or the age and $S(t) = \text{Prob}(T > t)$ is the survivor function, where T is the duration of the firm in the industry. In addition to t, $h(\cdot)$ depends on a vector of explanatory variables X, which in our case consists of entry size (Esize), average size (Asize), covariation in size (Cosize), diversification dummy (Didum), and ownership dummy (Odum). The effect of age on hazard is implied by duration dependence, given as $(\mathrm{d}/\mathrm{d}t)h(t)$. Age has a positive (negative) effect on hazard if the duration dependence is positive (negative).

First we attempt to model the hazard function semiparametrically as Cox's proportional hazard model. This is the most widely used semiparametric hazard function. In this model, $h(\cdot)$ is factored as $h(t; X) = \exp(X\beta)\gamma_0(t)$, where the vector β denotes the coefficients of X and $\gamma_0(t)$ is the 'baseline' hazard corresponding to $X\beta = 0$. This model is semiparametric in that the estimation of β does not require the knowledge of the baseline hazard. However, it assumes that the effect of X on hazard does not change with time or age. In particular, for our model, it implies that the effect of entry size on hazard is the same whether a firm is 1 year old or 10 years old! Hence before discussing the results it is important to check the appropriateness of the specification of the model. To test whether Cox's model is appropriate for our data we use the recently available test of Horowitz and Neumann (1992). The test statistic, which has a Chi-squared distribution with 1 degree of freedom, was found to be 60.806. Hence, the null hypothesis that the model is true is rejected at all the usual levels of significance. Therefore, the estimates from this model are not reported.

[4] Various measures of firm size have been used in the literature, such as the value of assets, employment and sales. When data have been available for the various measures, the results have generally been invariant to the measure of size (see DRS). In our case, data on other measures of size are not available to permit confirmation of our results with the various measures of size.

We now turn to fully parametric models. The log-likelihood function for any distribution of duration is, in general, given by

$$L=\sum_{i=1}^{74}\left[C_i\ln(f(t_i;X_i))+(1-C_i)\ln(S(t_i;X_i))\right] \quad (1)$$

where $C_i=1$ if the ith observation is uncensored and 0 otherwise, $f(\cdot)$ is the density function and $S(\cdot)$ is the survivor function. The results can be quite sensitive to the distribution used because two different distributions may impose quite different restrictions on the hazard function. For example, consider the four commonly used distributions in duration analysis in economics—exponential, Weibull, log-normal and log-logistic (see Greene, 1993, p. 718). The exponential distribution assumes that the hazard does not change with age (holding X fixed) whereas Weibull presumes that it changes monotonically with age. Similarly, log-normal and log-logistic imply a nonmonotonic hazard with a very specific form of nonmonotonicity, namely, the hazard initially increases with age at a nonincreasing rate and then decreases at a nonincreasing rate. For a given data set, all the restrictions cannot be true simultaneously. Hence, as Greene (1993, p. 718) points out, it is likely to be counterproductive to estimate hazard functions assuming different distributions which imply different restrictions on the monotonicity of the hazard about which the researcher is unsure.

In short, a mechanical search for robust results across various distributions is unlikely to be successful due to the sensitivity of the results to the restrictions on the hazard. Hence, in the studies of duration one typically finds that the results are based on one hazard function that seems either economically or statistically reasonable for the data (e.g. see Bandopadhyay, 1994, or Jaggia and Thosar, 1995, or Mata and Portugal, 1994, and the references therein). In our data, for instance, the Cox's model implies that the hazard or the probability of exit goes up with post-entry size whereas the Weibull hazard goes up monotonically with age! Hence, we first consider the nature of the two key features of any hazard function, namely proportionality and monotonicity.

So far we only know that we can rule out distributions that imply a hazard function that is proportional to age (duration), such as exponential, compound exponential, Weibull and Gompertz (see Cox and Oakes, 1984, p. 17) because our data rejected the proportional hazards model. We do not want to impose any restriction on monotonicity because that is equivalent to restricting the relationship of exit to age. Hence, we need a flexible hazard function that does not impose any restriction on monotonicity or the form of nonmonotonicity and has a nonproportional hazard.

In general, mixing two distributions is a parsimonious way of increasing flexibility tremendously as compared with being restricted to either one of the component distributions. But mixing distributions are typically quite complicated. However, the one suggested recently by Jaggia and Thosar (1995) is not only very flexible, it also has a closed-form hazard and survival function, which makes computation quite convenient. In addition, it accounts for unobserved heterogeneity among firms. It is the Weibull–gamma mixture. It provides huge flexibility of the hazard by increasing the number of parameters to be estimated only by 1 as compared with the log-normal or log-logistic distribution. Weibull which has a monotonic hazard and log-logistic which has a nonmonotonic hazard are both special cases of this function (and exponential is a special case of Weibull). Since the log-normal distribution is very similar to the log-logistic distribution, the former may be considered as a close special case so that if the log-logistic distribution is decisively rejected it is unlikely that the log-normal distribution will be true.

As we see below, the estimates of the Weibull–gamma mixture imply that the Weibull and log-logistic distributions are very unlikely to be true. When we assume a log-normal distribution we find that it consistently overpredicts the observed hazard at each duration. At durations 4, 5 and 6 where the observed hazards, respectively, are 0.21, 0.03 and 0.13 (Table 1) the mean of predicted hazard over firms at each of these durations are, respectively, 0.68, 0.81 and 0.99! This suggests that the specific form of nonmonotonicity in the hazards of log-logistic and log-normal is being refuted by our data. We now turn to the Weibull–gamma distribution and present it formally because it is rather recent in the literature and its flexibility will become clear.

The survivor function for the Weibull–gamma mixture is given by (for details, see Jaggia and Thosar, 1995):

$$S(t;X)=\left[1+\sigma^2\exp(X\beta)t^{\alpha}\right]^{-1/\sigma^2} \tag{2}$$

where σ^2 determines the extent of unobserved heterogeneity and α affects the duration dependence. The hazard function equals

$$h(t;X)=\frac{\exp(X\beta)\alpha t^{\alpha-1}}{1+\sigma^2\exp(X\beta)t^{\alpha}} \tag{3}$$

This is the hazard function we use. Note that the hazard above has a nonproportional form. Also, it is not obvious from Eq. (3) how β relates to the effect of a regressor on the hazard. The derivative of the hazard with respect to the jth regressor, X_j whose coefficient is β_j, is given by

$$\frac{\partial h(t;X)}{\partial X_j}=h(t;X)\beta_j\left(1-\frac{\sigma^2\exp(X\beta)t^{\alpha}}{1+\sigma^2\exp(X\beta)t^{\alpha}}\right) \tag{4}$$

Hence, the effect of a regressor on the hazard has the same sign as its coefficient. The effect of age on hazard is the duration dependence, i.e. the derivative of $h(t;X)$ with respect to t, $h'(t;X)$ given as

$$h'(t;X)=\exp(X\beta)\alpha t^{\alpha-2}\frac{\alpha-1-\sigma^2\exp(X\beta)t^{\alpha}}{\left[1+\sigma^2\exp(X\beta)t^{\alpha}\right]^2} \tag{5}$$

The sign of $h'(t;X)$ is the same as the sign of $(\alpha - 1 - \sigma^2 \exp(X\beta)t^{\alpha})$, which, in general, is ambiguous. The second derivative of the hazard function is also ambiguous in sign a priori. This means that the Weibull–gamma hazard does not a priori restrict the form of monotonicity (unlike the Weibull or exponential) or nonmonotonicity (unlike the log-normal or log-logistic). The estimated values of α, β and σ^2 will determine them for a given data set. Hence the hazard has a very flexible form. Note that $\alpha < 1$ is a sufficient condition for negative duration dependence, i.e. probability of exit to decline with age. The estimates of the Weibull–gamma mixture are given in Table 2.

It can be seen that all the three size variables, Esize, Asize and Cosize, as well as the duration dependence parameter α and the unobserved heterogeneity parameter σ^2 are significant at reasonable levels. Before we can use the estimates for any inference a specification test is in order to formally check whether the Weibull–gamma assumption is appropriate. There are many choices such as the likelihood ratio test or moment tests (see Jaggia and Thosar, 1995, for use of such tests in parametric duration models). The test statistic for these tests depend on the number of observations and are biased towards rejecting the null in large samples. Therefore, as in Das (1992), we use the traditional Chi-square goodness of fit test for which the test statistic does not depend on the sample size (see Rao, 1973, chapter 6). The null hypothesis is that the Weibull–gamma specification is true. The test statistic has 2 degrees of freedom (the number of cells, which are the ten observed duration values in the sample, minus the number of estimated parameters (7) minus 1). Its sample value is 1.19. Hence we accept the model at all the usual levels of significance and hence use it for inference. Since σ^2 is very far away from 1 (the value at which the Weibull–gamma mixture specializes to log-logistic) and 0 (the value at which the Weibull–gamma mixture specializes to Weibull) we can confidently conclude that the log-logistic and Weibull distributions are not appropriate for our data.

Now we turn to the discussion of results based on the Weibull–gamma mixture.

(1) The sign on the Esize variable indicates a positive relationship between the entry size of a firm and the probability of its exit and is opposite of the result in the earlier studies of Audretsch and Mahmood (1992) and Mata and Portugal (1994) on developed countries. This result is interesting and is not implied by the widely used Jovanovic (1982) model as well as the Frank (1988) model. In Jovanovic's model, the driving factor for all decisions is the time-invariant managerial ability. Hence, given post-entry size, entry size does not have any explanatory power (see Pakes and Ericson, 1990, for a detailed analysis, especially pp. 18 and 19). However, if Jovanovic's model is extended to also allow for contemporaneous shocks such as fluctuations in market demand or costs to affect current output and profits, a large entry size for a given post-entry size indicates a slow-growing firm facing negative shocks; this in turn hastens exit. [5] In developing countries where most industries are infant, fluctuations in industry-wide shocks are likely to be more common than in the mature environment of developed

Table 2
Estimates

Variable	Coefficient	Standard error	t-ratio
Esize	17.027	9.025	1.89
Asize	−16.812	6.231	2.70
Cosize	−72.298	34.048	2.12
Didum	−1.858	13.518	0.14
Odum	0.368	29.868	0.01
α	89.950	26.891	3.34
σ^2	103.958	18.880	5.54

countries. Timely response to these shocks and learning will be critical for survival. Since rigidities of various kinds abound in developing countries like India, a large entry size may become a liability as it hampers flexibility (e.g. laying off workers is not as easy in India as in the US) and response to new information which is critical at entry time.

(2) Post-entry size (Asize) is negatively related to exit. A large post-entry size for a given entry size implies that a firm has succeeded in learning and adapting in the uncertain environment of an infant industry. Hence, the result is not surprising. It is similar to that found for mature industries in developed countries and is consistent with most theoretical models such as the 'shake-out hypothesis' (see the learning models of Jovanovic, 1982, and Frank, 1988). [6]

(3) There is no 'bench-mark' prediction about the relationship of volatility of size path (given by Cosize) and survival. We find that firms with a volatile time path of size have a lower probability of exit, that is, volatility and exit are negatively related. This may not be surprising for infant industries where there is likely to be a huge variance in information flows to which firms must respond. Responsiveness to a fast changing environment will be reflected in the volatility of sales. Those who cannot respond quickly and hence have a less volatile size path are likely to exit early. The validity of such an explanation should be examined in future theoretical work.

(4) It is clear from Table 2 that diversification into other industries and ownership have little effect on exit. This may suggest that survival in this industry is very much due to industry or product specific factors and failure in these factors cannot be made-up by the advantages of diversification or public sector ownership.

(5) Lastly, we turn to the relationship between age and exit. Since α significantly exceeds 1, the slope of the hazard in Eq. (5) is ambiguous. Hence, it is evaluated for each firm using the data values and the estimates. We find that out of the 74 firms, 58 (nearly 80%) have hazards declining with age.

[5] We are thankful to a referee for this explanation.

[6] The 'stake-out' model of Ghemawat and Nalebuff (1985) has the opposite implication.

In the models of Jovanovic (1982) and Frank (1988), in which firms learn more or less passively over time, exit and age are negatively related whereas there is not much of a relationship in the 'active' learning model of Ericson and Pakes (1995) in which firms invest to learn. On the other hand, in the models of Jovanovic and Lach (1989) and Lambson (1991), age and exit are positively related as younger firms (with more efficient technologies) drive out older firms. For the computer industry, which is infant as well as high-tech, all the above forces are likely to be present. The effect of age on hazard observed for a firm will then depend on the relative strength of these forces.

Hence, our finding that hazards decline with age for 80% of the firms suggests that passive learning over time is the dominant force. Further, for 60 out of 74 the second derivative of the hazard is negative, thereby implying increasing returns to learning! This could arise due to various reasons. First, although firms in mature as well as infant industries keep learning about their own efficiencies over time and find their niches in the product market as they age, the returns to such learning may be increasing in an infant industry while it may be diminishing in a mature industry. Second, in an infant industry, 'brand-effects' or learning by consumers about the existence of a firm's product may increase over the age of the firm and have a positive impact on its survival (see Das, 1995, for similar evidence on the growth of surviving firms).

4. Concluding remarks

This paper studies the duration and hazard (exit) of firms in an infant industry in a developing country. The relationship of hazard with post-entry size and age are similar to those found in earlier studies on mature industries in developed countries. But that of entry size and hazard is different. Entry size and exit are positively related in the Indian computer hardware industry. Size is a liability at entry as it hampers response to learning in the rigid environment of a developing country. Interestingly, we also find that the successful firms have a more volatile size path. Further, the estimated hazard function implies that there is increasing returns to learning in the Indian computer hardware industry. Finally, it is surprising to find that diversification and public-sector ownership do not provide additional advantages for survival, thereby suggesting that a firm's survival in the Indian computer hardware industry depends largely on its product.

Since previous studies of exit have focused on mature industries in developed countries, it would be interesting to study infant industries in developed countries also and compare the growth experience to that in developing countries.

References

Audretsch, D.B., Mahmood, T., 1992. Firm selection and industry evolution: the post-entry performance of new firms. Working Paper, Research Unit Market Processes and Corporate Development.

Bandopadhyay, A., 1994. An estimation of the hazard rate of firms under Chapter 11 protection. Review of Economics and Statistics 76, 346–350.

Cox, D.R., Oakes, D., 1984. Analysis of Survival Data. 1st ed. Chapman and Hall, New York.

Das, S., 1992. A micro-econometric model of capital utilization and retirement: the case of the U.S. cement industry. Review of Economic Studies 59, 277–297.

Das, S., 1995. Size, age and firm growth in an infant industry: the computer hardware industry in India. International Journal of Industrial Organization 13, 111–126.

Dunne, T., Roberts, M.J., Samuelson, L., 1989. The growth and failure of U.S. manufacturing plants. Quarterly Journal of Economics 104 (4), 671–698.

Ericson, R., Pakes, A., 1995. Markov-perfect industry dynamics: a framework for empirical work. Review of Economic Studies 62, 53–82.

Frank, M.Z., 1988. An intertemporal model of industrial exit. Quarterly Journal of Economics, 103, 333–344.

Ghemawat, P., Nalebuff, B., 1985. Exit. Rand Journal of Economics 16, 184–194.

Greene, W.H., 1993. Econometric Analysis, 2nd ed. Macmillan, New York.

Horowitz, J.L., Neumann, G.R., 1992. A generalized moments specification test of the proportional hazards model. Journal of American Statistical Association 87, 234–240.

Jaggia, S., Thosar, S., 1995. Contested tender offers: an estimate of the hazard function. Journal of Business and Economic Statistics 13, 113–119.

Jovanovic, B., 1982. Selection and evolution of industry. Econometrica 50 (3), 649–670.

Jovanovic, B., Lach, S., 1989. Entry, exit, and diffusion with learning by doing. American Economic Review 79 (4), 690–699.

Kiefer, N.M., 1988. Economic duration data and hazard functions. Journal of Economic Literature 26, 646–679.

Lambson, V.E., 1991. Industry evolution with sunk costs and uncertain market conditions. International Journal of Industrial Organization 9, 171–196.

Manufacturers' Association of Information Technology (MAIT), 1991. A Report on the Evolution of IT industry. New Dehli.

Mata, J., Portugal, P., 1994. Life duration of new firms. Journal of Industrial Economics 62, 227–245.

Pakes, A., Ericson, R., 1990. Empirical implications of alternative models of firm dynamics. Working Paper, Yale University.

Rao, C.R., 1973. Linear Statistical Inference and Its Applications, 2nd ed. John Wiley and Sons, New York.