UCD GEARY INSTITUTE DISCUSSION PAPER SERIES

The Firm Size Distribution in a Small Open Economy: Theory and Evidence

Qi Li , SPIRe and Geary Institute

Patrick Paul Walsh, SPIRe and Geary Institute

Abstract

We construct a theoretical model of the dynamic processes (firm entry, growth, decline, and exit) that underpin the determination of a limiting firm size distribution (FSD). In particular, we model such dynamic processes using key structural parameters; sunk cost, exogenous entry constraints, and opportunity values of *finite* duration. The limiting FSD we derive, in steady state, turns out to be a combination of a *Logarithmic* and *Zipf* distribution. We estimate these structural parameters using long periods of Irish company data for defined cohorts of firms, in terms of trade orientation, within narrowly defined industries. Within non-exporting and exporting samples of companies our model fits the actual FSD well with a good return to the *Zipf* distribution in the upper tail, that is less dependent on the estimated structural parameters, and a good return at the lower tail, where the *Logarithmic* effects are endogenously driven by firm heterogeneity in estimated structural parameters.

JEL Classification System: L11 F15

*We thank John Sutton for very useful comments on an earlier draft.

1 Introduction

For nearly a century, economists have studied firm size distribution (FSD). To industrial organization economists, the importance of understanding firm size distribution is similar to understanding income inequality in development economics, or wage heterogeneity in labour economics.

The size distribution of firms tends to have one general property. The bulk of the firms are clustered around small sizes, while there are also a number of larger firms whose size distribution tends to closely resemble the upper tail of the log–normal distribution or the related Pareto and Yule distribution or even Zipf's distribution.[1] Such patterns in FSD have been modelled using the mathematics of stochastic processes, most notably by Simon and Bonini (1958) and Ijiri and Simon (1964). These models do not have much economic content as firm growth is a stochastic realization from a distribution whose mean and variance is independent of firm size. Consequently, they do not tell us much about the fundamental drivers of firm growth or about the moments of the firm size distribution, which are the key determinants of market structure. (Klepper and Thompson 2005).

[1] The most elementary firm dynamics model is that of Gibrat (1931). Gibrat showed that the firm size distributions examined were approximately log – normal in form and explained this by postulating a "law of proportionate effect". Accordingly, he finds that FSD has an ever increasing mean and variance over time. Hence, the model has no real steady state.

On the other hand, the increased availability of company level data has enabled us to have a better understanding of the empirics of firm growth which have turned out to be more complex than what the early stochastic growth models assumed. Evans (1987), Hall (1987), and Dunne, Roberts and Samuelson (1989) have shown that Gibrat's law fails to hold, if small firms are included in the datasets. Empirical findings have also shown that FSD for given cohorts are stable (see Axtell, 2001), and among surviving firms, both the mean and variance of firm growth decline with firm size, even after controlling for the other factors (Klepper and Thompson, 2003). The possibility of exit also declines with size. FSD is different across sectors and even within a single industry the FSD can evolve with business cycles (Marsili, 2006).

Understanding differences in FSD within industries remains a challenge.[2] We wish to look closer at the fundamental determinants of firm entry, growth and exit that can be different by certain firm cohorts, for example by trade orientation, and the emerging market structure. The modelling approach in this paper is a variant of the "Islands Model" first proposed by Simon (1955) and reconsidered by Steidl (1965), Ijiri and Simon (1977), Levy and Solomon (1996), Sutton (1998), and Geroski (2000). In line with this literature, luck is the principal factor that distinguishes winners from losers among the pool of contenders. However, rather than imposing simple statistical rules, we model entry, growth, decline and exit endogenously - jointly determined by structural factors. Those determinants provide us with the economic content that allows us to estimate such structural parameters by firm cohorts within industries and hence enable a better understanding of the heterogeneity in limiting FSD.

The innovations in our modelling are motivated by two strands of literature. The first strand is concerned with firm heterogeneity and hypothesizes that different

[2] Hymer and Pashigian (1962), analyzing more disaggregated data, find a high degree of heterogeneity in FSD across sectors. They conclude that it is quite unclear whether any simple statistical regularity in terms of the FSD can actually exist. Similar concerns are found in Singh and Whittington (1975), and Lotti et al. (2003). Some researchers find that FSD evolves along with industry life–cycles and dynamics (Dinlersoz and Macdonald 2008). These findings reflect that models introducing simple statistical regularities cannot reveal the true properties of the dynamics of firms in specific sectors.

structural primitives underpin different groups of firms. Using 2003 firm data from Irish manufacturing, we find that size distribution of exporting firms is very different from that of non–exporters. From Figure I we see that non–exporting firms tend to be much smaller.[3] This evidence suggests that size distributions vary across firms with different trade orientation. Analysis of the FSD in an open economy has received relatively little and only recent attention (Nocke and Yeaple, 2008). Hence trade specific structural primitives could help us account for the evolution of FSD.

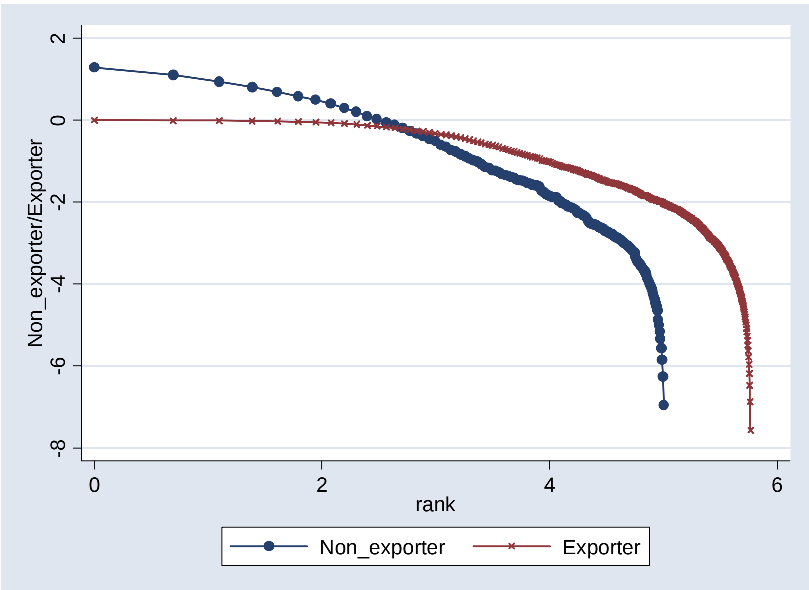


Figure I: Firm Size Distribution of Non–Exporters and Exporters in year 2003
(In cumulative distribution function (CDF) and log–log scale)

Firm heterogeneity has already been the subject of recent theorizing about industry evolution. Cabral and Mata (2003) use differences in financing constraints between young and old firms as the key factor that explains the skewness observed in FSD of Portuguese manufacturing firms. Lutter (2005) emphasizes sunk cost and the ease to imitate other firms as a mechanism that creates firm heterogeneity across industries. More recently, Nocke and Yeaple (2008) analyze the effects of

[3] Firm size distribution is described by cumulative distribution function (CDF) in log–log scale; x axis denotes rank in log scale; y axis denotes cumulative probability in log scale. And cumulative probability is the frequency of observations larger than this rank.

globalization on the distributions of firm size by inducing heterogeneity in firms' product range. The FSD is also linked to the life cycle of the industry (Klepper and Graddy, 1990; Agarwal and Gort, 1996), or the technological regime (Dosi, Marsili, Orsenigo and Salvatore, 1995). These literatures suggest that the size distribution is influenced by firm cohort characteristics within industries. We follow the above literature by grouping firms by trade orientation and industry. The evolution of their size distribution will depend on a set of group–specific primitives, such as a sunk cost, exogenous entry constraint, and opportunity (sub-market) values. High sunk cost and exogenous entry constraints will reduce the chances of new firms winning new opportunities and at the same time gives more chances to incumbents to build up their portfolio of opportunities (size). On the other hand, opportunities with higher values will attract more potential wins from entrants and leads to more competition for incumbents. Such primitives will vary the evolution of firm size and skew the FSD differently. No single independent distribution will be able to fit all industries.

The second innovation in our modelling is related to a strand of literature that is concerned with issues of finite modelling. The island model was originally proposed to describe a growing industry (where industry size and the number of firms keep growing) without consideration of decline and/or exits of firms[4]. As a result, many researchers have modelled various decline processes to reflect shrinkage of industries and firms. This allows them to obtain different types of limiting firm size distributions depending on the way the decline process is imposed or the way the steady–state distribution is calculated. For instance, Simon (1955) introduces size–independent firm exits and obtains a Waring distribution in a steady–state. Later, he induced a Katz distribution by introducing a constant rate of decline in his model (Simon 1960). Steidl (1965) imposes the possibility of firm decline and obtains hypergeometric distribution. Levy and Solomon (1996) adds to the model a minimum firm size below which firms cannot decline and obtains a Pareto distribution. The

[4] Simon (1955) introduces a model in that at each time interval one size unit is added to the industry and each new unit could be either a new firm of size 1 (with probability *p*) or a unit of growth of an incumbent firm (with probability 1 - *p*). If the unit of growth accrues to an incumbent firm, the probability that it accrues to a particular firm is proportional to the firm's size.

literature is impressive for the many ways it can model decline processes and generate different explicit forms of the limiting distribution. However, one drawback is the lack of economic content; growth and decline of firms are only explained by some simple statistic regularity (e.g. exit rate, or decline rate) and the regularity is arbitrarily imposed. Another drawback is that all firms face the same regularity (e.g. firms face same exit rate, or same size–independent decline rate) without consideration for firm heterogeneity. In addition, the models only describe a stationary industry or growing/declining industry and cannot capture the dynamics of industries (e.g. life-cycle).

One exception to this is the modelling approach of Klepper and Thompson (2005). In their setup, all firms belong to a set of submarkets and the submarkets are being created with finite life. Firms expand when they are able to exploit new opportunities that arrive in the form of submarkets; they contract and ultimately exit when the submarkets in which they operate are destroyed. The central force for industry evolution is the creation and destruction of submarkets. The framework can explain a wide range of well–known regularities about industry dynamics. Our model is similar in spirit to Klepper and Thompson's (2005) in that we assume all opportunities for growth have finite duration. However, similar to the traditional island model we also build explicit forms of underlying processes to describe industry evolution. The finite duration of the opportunity assumption makes our model closer to reality without losing generality. The vanishing of existing opportunities is the central force for firm decline or exit and it can be explained in many ways, such as expiry of contracts, obsolescence of technology, and decline of geographic areas. Indeed, a firm is referred to as exiting when all the opportunities in that firm die. The assumption mainly has three advantages. First, it allows for firm heterogeneity to come from different opportunity dynamics for defined cohorts of firms. Secondly, with finite duration the model can describe different stages in the industrial life–cycle. If the duration is longer than the time interval the model represents a growing or stationary industry; otherwise, the model represents a declining industry. Finally, decline or exit is linked to entry and growth. Because, the decline of a firm depends when it took the opportunities and the durations of the opportunities - namely, its growth history. This is consistent with many empirical findings, such as Simon (1997).

Our model identifies a set of structural parameters, such as sunk cost, exogenous entry constraint, opportunity values within finite duration; and provides structure on how they are expected to influence the FSD. This explicit and structural model can transparently explain the mechanisms underlying firm growth and also the accumulating size–growth regularity found in many of the earlier empirical works. With our theory in hand we need evidence about the quantitative importance of our primitives to judge their relevance. Laincz and Rodrigures (2005) analyses how the structural parameters in their model affect the FSD by simulation. We use an empirical approach, based on Irish manufacturing data and use the dynamics of underlying processes observed to restore structural primitives. Our analysis suggests that the structural parameters vary across industries and cohorts of firm by trade orientation within industries. Such differences among cohorts in their structural parameters should help us account for an important cause of heterogeneity in the FSD. In a small open economy, sunk costs and entry constraints are different for non-exporters and exporters. The size and duration of contracts can also be expected to be different. Hopefully, we capture an important source of heterogeneity in the FSD.

Furthermore, we predict the limiting distribution based on the estimated parameters and compare it with Zipf's distribution, which is widely used (e.g. Axtell, 2001). We find that the predicted model fits better to real data, especially in the low tail of the distribution, which illustrates the explanatory power of our model relative to other distributions. The paper is organized as follows. In Section 2, we model the processes underlying firm size evolution, including entry, exit, growth and decline. Section 3 derives the limiting firm size distribution from the model. In Section 4, we estimate primitives in the model using Irish manufacturing data over a long time span. In Section 5, we predict the limiting firm size distribution using the primitives estimated and compare them with the actual FSD in our exporting and non-exporting firm populations. Finally, Section 6 contains some concluding remarks.

2 The Model

In the literature about the stochastic models of firm size dynamics, there exists a well-established tradition describing the modification of firm size as the cumulative effect of a number of shocks generated by the diverse accidents that affected the firm's history (Simon, 1955; Steidl, 1965; Ijiri and Simon, 1977; Levy and Solomon, 1996; Sutton, 1998; Geroski, 2000; and Amaral et al, 2001). We apply a variant of the framework as follows.

Imagine that a sequence of discrete and independent investment opportunities arise over time. Each opportunity is of the same size, in terms of the profit it yields to any single firm that takes it up. Each opportunity will be taken up by exactly one firm and this firm could be an active firm or a potential firm. If a new opportunity is taken by a potential firm, this firm enters the market. In the meanwhile, each opportunity is of the same duration and the opportunity is taken off from system when it expires after its duration. And if all the opportunities taken by a firm die, the firm is referred to exit. Once a new opportunity is taken up by a firm, the firm will monopolize the opportunity until this opportunity expires. Thus a firm's current size is measured by the number of opportunities it has taken up to date.

These opportunities represent all sorts of "accidents": the exploitation of technological novelties, the reaction to demand shocks, or the effects of managerial reorganization. As the reasons showed in previous section, we assume that all the accidents can only last a finite duration. Thus the death of opportunities becomes the driving force of the decline or exit of firms, instead of imposing a decline rate in the traditional models. To capture the forces that affect the firm size distribution in a structural model, we introduce a set of exogenous variables to describe firm cohort–specific characteristics, including:

(1) *I*, the present value of a new opportunity

(2) *a*, duration of an opportunity

(3) *Cs*, sunk cost for new entrant to enter market

In the following, we demonstrate that the above exogenous variables are the key determinants for the dynamic forces that underpin industry evolution.

2.1 Entry

Assumption 1: *The probability that a new opportunity is taken by any currently active firm is proportional to the size of that firm.*

This assumption can be seen in a bulk of traditional models (Simon, 1955; Steindl, 1965; Levy and Solomon, 1996; and more recently Bottazzi and Secchi, 2005). This is the application of Gibrat's law to incumbent firms. Many early empirical works were only using the datasets of large firms and most of them confirmed that Gibrat's law seem to work well enough. The main reason that Gibrat's law holds in those works is that they only focus on surviving firms, but not on small firms, entrants or exiting firms. Cabral and Mata have shown that the firm size distribution in Portuguese manufacturing overtime can evolves toward a log-normal distribution, which gives evidence that Gibrat's law holds for firms in a mature industry. Suppose that incumbents are larger and more mature than new entrants; then it would be natural to link the incumbents' growth to Gibrat's law. As pointed by Bottazzi and Secchi (2005), if one switches off the entry and exit dynamics, as the empirical investigations show, most of those theoretical models are driven by Gibrat's law.

We denote the probability that a new opportunity is allocated to a potential firm at period t as P_t. If a potential firm enters, it pays entry fee *Cs,* earns an opportunity value of *I*, and can expect *E,* a *chance* to earn more opportunities during *a* periods that follow the winning of *I*.[5] Once a firm enters, it is protected by the opportunity (e.g. contract or patent for technology exploitation) and can compete with

[5] A potential firm could expect earn more opportunities after the first *a* periods, but for risk aversion, depreciation and simplicity reasons we only focus upon the first *a* periods.

all incumbents for other opportunities during *a* periods, without risk of exit. If a potential firm enters the market, the expected *Entry Profit* is;

E (*Entry Profit*) = *I - Cs + E(Chance)* **(1)**

For a given potential firm *i*, the expected *chance* can be solved as:

E(Chance)=

$$E(\frac{S_{i,t+1}}{\sum_i S_{i,t+1}} I(1-E(P_{t+1}))) + E(\frac{S_{i,t+2}}{\sum_i S_{i,t+2}} I(1-E(P_{t+2}))) + .. + E(\frac{S_{i,t+a}}{\sum_i S_{i,t+a}} I(1-E(P_{t+a}))) \quad \textbf{(2)}$$

Where $S_{i,t}$ denotes the size of firm *i* in period *t* , where the firm enters in period *t*. Expectation of future equals to current value if no additional information,[6] so we have

$$E(P_{t+1}) = E(P_{t+2}) = ... = E(P_{t+a}) = E(P_t) = P_t$$

$$E(\sum_i S_{i,t+1}) = E(\sum_i S_{i,t+2}) = E(\sum_i S_{i,t+3}) = ... = E(\sum_i S_{i,t+a}) = \sum_i S_{i,t}$$

From the model, we also have $E(S_{i,t+1}) = I$, and the following reduced form

$$E(S_{i,t+j}) = E(S_{i,t+j-1}) + \frac{E(S_{i,t+j-1})}{E(\sum_i S_{i,t+j-1})}(1-E(P_{t+j-1}))I$$

Combining the equations above, we have:

$$E(\textit{Entry profit}) = I - Cs + I((1+\frac{(1-P_t)I}{\sum_i S_{i,t}})^a - 1) \quad \textbf{(3)}$$

And the expected value of the *chance* is

$$E(Chance) = I((1+\frac{(1-P_t)I}{\sum_i S_{i,t}})^a - 1) \quad \textbf{(4)}$$

Assumption 2: *For a given cohort of firms, the probability to allocate the next opportunity to a new firm,* P_t*, is proportional to the Entry profit.*

In modelling firm dynamics as an exogenous stochastic process, we do not intend to discount the role of rational firm choice. On the contrary, we assume that the stochastic process is driven by a well–defined maximization problem for each firm.

[6] The underlying assumption is that there will be an equilibrium entry and market size. For any potential firm, it believes in the equilibrium entry and market size. There is a similar assumption in Klepper and Thompson (2003).

But whatever the details of the maximization problem, it will yield an equilibrium entry rate and the entry must be positively related to the entry profit. Thus, we have:

$$P_t = h(I - Cs + I((1 + \frac{(1-P_t)I}{\sum_i S_{i,t}})^a - 1)) \quad \textbf{(5)}$$

where h is constant and at group level.

In equation **(5)**, the positive association between entry rate P_t and entry profit is represented by h. h can be interpreted as some constraint for entry, for example a financial constraint (Cabral and Mata 2003), an imitation constraint (Lutter 2005) or a technology constraint as mentioned in many models. h could vary among groups and reflects group-specific primitives. Entry rate P_t can be solved implicitly from equation **(5)** and it is a function of current market size and group–level structural parameters.

$$P_t = f\,(I, Cs, a, h, \sum_i S_{i,t}) \quad \textbf{(6)}$$

Whatever the details of the market structure, it will yield an equilibrium market size, $\sum_i S_{i,t}$. In our model, we have the equilibrium market size defined as:

$$\sum_i S_{i,t} = a*I.^{7} \quad \textbf{(7)}$$

Substitute equilibrium market size equation **(7)** into equation **(6)**; then entry rate P_t now only depends on group–level parameters. As the parameters represent the characteristics of a market in equilibrium and do not vary over time, we can drop the subscript t for P_t and rewrite equation **(6)** as:

$$P = f\,(I, Cs, a, h) \quad \textbf{(8)}$$

It means that for a given group of firms in market equilibrium,[8] the entry process is determined by a set of group–level primitives, such as the sunk cost, exogenous entry constraint, opportunity value, and duration. A sector with low sunk cost or a low entry constraint will attract more potential entrants and a relatively large entry rate will be observed, which will make its FSD different when compared to others. Opportunity duration enables entrants to compete for new opportunities as

[7] If the duration is longer than the time interval, after the first *a* periods the market *size* reaches a steady state. In a steady state, an old opportunity is taken and a new opportunity is born in each period but the market size remains constant.
[8] The definition of group ensures that the firms within a group has the same group–level primitives.

incumbents, without risk of shrinkage, during the period - so it is positively associated to firm entry. Such structural characteristics can vary by firm cohorts and can account for most of the firm size distribution heterogeneity coming from entry dynamics.

2.2 Growth and Decline

In each period, any active firm has chance to obtain a new opportunity but also has the chance to lose the opportunity it obtained *a* (duration) periods ago.

Thus, for any given firm *i*, one can write:

$$S_{i,t+1} - S_{i,t} = \textit{E (opportunity gain) – E (opportunity lose)}$$

$$= I \frac{S_{i,t}}{\sum_i S_{i,t}}(1-P) - I \frac{S_{i,t-a}}{\sum_i S_{i,t-a}}(1-P)$$

$$= I(1-P)\left(\frac{S_{i,t}}{\sum_i S_{i,t}} - \frac{S_{i,t-a}}{\sum_i S_{i,t-a}}\right) \quad \textbf{(9)}$$

We rewrite equation **(9)** as:

$$S_{i,t+1} - S_{i,t} =$$

$$I(1-P)[(\frac{S_{i,t}}{\sum_i S_{i,t}} - \frac{S_{i,t-1}}{\sum_i S_{i,t-1}}) + (\frac{S_{i,t-1}}{\sum_i S_{i,t-1}} - \frac{S_{i,t-2}}{\sum_i S_{i,t-2}}) + \ldots + (\frac{S_{i,t-a+1}}{\sum_i S_{i,t-a+1}} - \frac{S_{i,t-a}}{\sum_i S_{i,t-a}})$$

Combining equation **(7),** we have:

$$\Delta S_{i,t} = \frac{1}{a}(1-P)\sum_{j=1}^{a} \Delta S_{i,t-i} \quad \textbf{(10)}$$

Equation **(10)** predicts the relationship between firm size and growth found in many empirical works, such as autocorrelation and impeded growth, which had been presented as a challenge for theory to explain. First, it implies that there exists autocorrelation for growth and current growth is positively associated with the past of firm. Simon (1977) estimates the autocorrelation of firm growth from US industry data and finds the one year coefficient is positive, reaching 0.35. Similar findings can also be seen in Dunne et al. (1988), Lotti et al. (2003) and more recently Bottazzi and Secchi (2006a). On the other hand, growth only depends on the performance in the finite past, that is, the last *a* periods, which means autocorrelation length is determined by the opportunity duration.

Second, the autocorrelation is stationary and a steady state exists. Using equation **(10)** we see the mean and variance of $\Delta S_{i,t}$ tends to a finite constant. As $\Delta S = gS$, the model implies impeded growth, which matches the findings of Evans (1987), Hall (1987), Kumar (1985) and Hart and Oulton (1996). Impeded growth has been the biggest challenge of Gibrat's law. We show that though Gibrat's law is imposed on surviving firms, firm size distribution as a whole deviates from log–normal distribution when we allow for entry, exit, and finite duration of opportunities. Gibrat's law represents symmetry of firm growth. However, in our model, the presence of opportunity duration favours those new entrants and protects them from shakeout until the end of their first *a* periods. Entry depends on the group–level primitives and is not necessarily equal to the average growth of firms within the cohort. We do not necessarily have symmetry of firm growth.

2.3 Exit

If all opportunities taken by a firm expires, the firm exits. From above we have:

If $S_{i,t} + \frac{1}{a}(1-P)\sum_{j=1}^{a} \Delta S_{i,t-i} < 0$, firm *i* exits. **(11)**

Some early works (e.g. Steidl, 1965) introduced minimum efficient size (MES), which means firms can not decline below this size. We do not specify a MES because in our model, the basic unit is opportunity value *I* (e.g. contract value) and in theory a firm can make profit and survive even if there is only one contract left, as the sunk cost has been paid.[9]

Equation (**11**) shows that the possibility of firm exit is negatively related to current size, which is consistent with many works, such as Hughes (1994), Wagner (1994), Audretch and Mahmood (1995), Audretch et al. (2000) and Mahmood (2000). The possibility of firm exit is also negatively associated with firm's average growth in

[9] However, in the empirical part below, we allow for the existence of *MES*, as we are estimating the average levels of cohort–level parameters. It is the case that some firms cannot run in sectors with small *I*.

last a periods. It implies that efficiency or profitability in recent history is an important determinant for firm survival and this matches findings in Austin and Rosenbaum (1990), Dunne and Roberts (1991), Mayer and Chappel (1992), Doi (1999), and Audretch et al. (2000), which have proposed a negative relationship between firm survival and growth in various aspects. At the same time, the association depends on duration of opportunity, a, and equilibrium entry rate, P. The model predicts that even for given firm size and growth, higher entry rates result in higher exit rate, which is consistent with Cable and Schwalbach (1991), Dunne, Roberts and Samuelson (1988), and Evans and Siegfried (1994). As both P and a are at group level, the decline dynamics of firms would vary across groups and is a function of group specific primitives.

3. Steady State

In this section, we derive an explicit form of the steady state using the mechanisms presented in the previous section. We will show that the steady state will also vary across groups of firms and depend on group level primitives. In steady state, we can predict the limiting FSD for a given group of a firm's primitives.

As there exists an equilibrium entry rate, P, for each group of firms and I only influences firm growth/decline and exit through P; the opportunity value I is normalized to 1, without losing generalization.[10] We define the steady state as the size distribution for a given group of firms that does not change over time in terms of its shape and total numbers of firms. Thus, we have:

$$f_t(S) = f_{t-c}(S)$$

where $f_t(S)$ denotes the number of firms with size S in period t; c is any finite constant. Thus we drop the subscript t below.

In a steady state, the flows for the firms with size S, $f(S)$ cohort have the flowing structure:

Expected Inflow = Expected growth of a firm from (S-1) to S + Expected decline of a firm from (S+1) to S

[10] We use P as exogenous variable for the following deduction but we should note that P is determined by a set of group–level exogenous variables.

Expected Outflow = Expected decline of a firm from *S* to *S-1*+ Expected growth of a firm from *S* to (*S+1*)

$$Expected\ Inflow = Expected\ Outflow$$

This condition can be written as:

$$\frac{(1-P)(S-1)f(S-1)}{\sum_s Sf(S)} + \frac{(S+1)f(S+1)}{\sum_s Sf(S)} = \frac{Sf(S)(1-P)}{\sum_s Sf(S)} + \frac{Sf(S)}{\sum_s Sf(S)} \quad (12)$$ [11]

In the meanwhile, analyzing $f(1)$ cohort, we have the boundary condition:

Expected Inflow = *Entry* + Expected decline of a firm from *S = 2* to *S = 1*

Expected Outflow = *Exit* + Expected growth of a firm from *S = 1* to *S = 2*

$$Expected\ Inflow = Expected\ Outflow$$

The boundary condition can be written as:

$$P + \frac{2f(2)}{\sum_s Sf(S)} = \frac{f(1)(1-P)}{\sum_s Sf(S)} + \frac{f(1)}{\sum_s Sf(S)} \quad (13)$$

Because *I* has been normalized, the equilibrium market size equation **(7)** above can be rewritten as:

$$\sum_s Sf(S) = a \quad (14)$$

Combining equation **(12), (13)** and **(14),** we have the limiting distribution:

$$f(S) = A\left(\frac{1}{S} + K\frac{(1-P)^S}{S}\right) \quad (15)$$

where *A* is the normalization factor and $K \equiv \dfrac{-f(1) + \dfrac{aP}{1-P}}{f(1) - aP}$.

Equation **(15)** shows that the limiting firm size distribution is the combination of *Logarithmic* distribution and *Zipf* distribution.[12] Most empirical

---

[11] From above, expected growth of a firm from *S* to *S+1*can be written as $\frac{Sf(S)(1-P)}{\sum_s Sf(S)}$. Expected decline of a firm from *S* to *S-1* can be written as $\frac{Sf(S)}{\sum_s Sf(S)}$, because all the opportunities in the system have the same duration distribution, so they have an equal chance to expire.

[12] *Zipf* distribution states that the frequency of any observation is inversely proportional to its rank in the frequency table. For example, the most frequent observation will occur approximately twice as often as the second most frequent observation, which occurs twice as often as the fourth most frequent observations etc.

works describe the firm size distribution as how the probability density appears on a log-log scale. Zipf distribution as a special case of Pareto distribution seems to be the most robust empirical finding for FSD (see Axtell, 2001; and, Fujiwara, 2003) and it also describes surprisingly diverse natural and social phenomena, including city size distribution (Knudsen, 2001), word frequency in text (Simon, 1960) and internet traffic. Especially for the upper tail of distribution, Zipf distribution mimics the data very well. But in the lower tail, a non-linear property is found rather than the straight line in Zipf law, for example Ijiri and Simon (1977); and, Hart and Prais (1956).

The combination of logarithmic and Zipf distribution has nice properties. From equation **(15),** at the upper tail where S is large, $f(S) \to A\frac{1}{S}$ with $P < 1$; the model returns to Zipf distribution and the upper tail of FSD is independent of cohort characteristics. However, at the lower tail where S is small, the logarithmic takes effect and distribution shows concavity if K is negative or convexity if K is positive. And the direction and degree of bending are now endogenously determined by cohort characteristics, such as a and P, which accounts for most of the heterogeneity of FSD. This is consistent with the finding that heterogeneity and deviation from log–normal are found only when small firms are included in data.

Additionally, in principle any forms of distribution between Zipf distribution and logarithmic distribution can be obtained by specifying a particular K in our model (e.g. Waring distribution and Extended Katz distribution). This incorporates most of the previous theoretical works, which have proposed explicit forms of limiting distribution. Not only will our K reflect cohort primitives but they will be different for different groups of firms inside industries. In particular, we will see that non-exporting and exporting firms in the same industries and country have very different estimated primitives and hence K values.

4. Empirical Estimation

If the underlying processes (entry, growth, decline, and exit) are endogenously determined by group level primitives, we can estimate the key primitives quantitatively given data on the dynamics of these processes. We estimate such primitives in this section. This will further help understand how the structural parameters influence firm size distribution.

4.1 Data

Our main data source is the Annual Employment Panel Survey carried out by Forfás over the period 1972 to 2003, covering all Irish manufacturing companies. The unit of observation is employment (permanent staff) annual data at the firm level. Although the number of firms in operation in a given year varies from between 4,000 and almost 7,000, the total number of firms tracked in the data is 27,407. These are identified by Sector (*4-digit NACE industry codes*), Start-up date, Close date and a set of characteristics, such as ownership, location, and trade information.

A main advantage of the survey data is that it covers the entire universe of firms, unlike some of the publicly available data that is especially biased to large firms and is not a random sample from an industry's firm size distribution, and so offers a biased view of the actual firm size distribution. For instance, the COMPUSTAT database, frequently used for several industry studies, covers only publicly traded firms in the US and heavily under–represents the population of small firms (Dinlersoz and Macdonald, 2008).

Moreover, we use employment as the measure of firm size, rather than sales. While both employment and sales have traditionally been used as measures, relatively little is known about the relationship among different measures. However, sales can be difficult to deflate at the level of the firm.

4.2 Definition of a group

To accurately estimate primitives, we need to define a group first. We use two measures to define a group of firms; one is to allocate firms to 4–digit industries and another is to group firms by trade orientation within the same 4–digit industries. From an empirical standpoint, there is no perfect definition of a group in our model. The 4–digit level of aggregation reflects an attempt to balance two forces. Too broad a definition, e.g. 1- or 2-digit classification, will make firms within the group face different demand shocks and primitives so that the model cannot make correct predictions. For instance, machinery for mining and construction (2952), and machinery for food and tobacco processing (2953) are virtually unrelated, but they are within a broad 3–digit level. On the other hand, too narrow a definition causes industry evolution to coincide with the evolution of a very particular product. This is similar to Dinlersoz and Macdonald (2008), which gives an example that manganese dioxide and silver oxide batteries are very closely related products, but that, nevertheless, are classified as two separate 7–digit codes. At the same time, a narrow definition will decrease the number of observations within a cohort and increase the volatility. This can be a problem for Irish manufacturing even at the 4-digit level as it is a small economy.

Another factor is that in our analysis we can classify our data into *Home* and *US* employment dominated industries.[13] There are 43 US industries (e.g. with the majority of jobs in a 4-digit industry being US owned between 1972 and 2003), and 58 Home industries (e.g. with the majority of jobs in a 4-digit industry being Irish owned between 1972 and 2003).[14] A foreign versus indigenous dualism has been at the centre of most research on Irish manufacturing. Irish industrial policy since the 1950s clearly targeted green field and export-oriented FDI, mainly US companies, to locate in high-technology 4-digit sectors and away from traditional manufacturing (Li, Walsh and Whelan, 2008). The economic factors that govern such FDI entry, survival and exit should be expected to be different to Home industries.

Trade orientation is another dimension of our grouping. Within all industries there are both exporting and non-exporting firms. Using trade information for firms in the Forfás annual expenditure survey 1983-2003, we label all firms as *Exporting* if a

[13] Firms do not operate in 90 4-digit industries in Ireland. In addition, we find a small number of firms in another 70 4-digit industries. This has lead researchers to aggregate up to 3-digit industries. Our strategy is to work with the 4-digit industries that explain 99 per cent of employment in each of the periods between 1972-2003. This excludes 70 small industries and about 300 firms.

[14] Each sector has either a majority of Irish owned or majority of US owned firms in terms of their contribution to sector employment. While clearly there are UK and other non-US Foreign owned firms, these do not aggregate up to a majority in any sector.

firm has any exports over the survey period.[15] For incumbent firms in 1972 that exited before 1983, we classify as exporting or non-exporting using exporting grant information. *Exporting* is treated as a fixed effect for the period 1972 – 2003 to control for pre-selection effects.[16] We perform our analysis using 4-digit groupings only and extend it to allow for trade orientation within 4-digit groupings.

[15] The expenditure survey excludes plants with less than 19 employees up to 1999, although they are included from the year 2000. The annual employment survey generally has the same firm identification number as the annual expenditure survey. We used phone numbers, address and name to match any outstanding firms. Based on an analysis of the expenditure survey of all firms in the year 2003 we work with the assumption that exporting was a rare feature of small firm activity (less than 19 employees) for the period of our study.

[16] There are very few observations where exporters become non-exporters over the period while only 10 per cent of exporters in the 1983-2003 survey came from non-exporting history. The nature of industrial policy encouraged exporting from start-up.

4.3 Parameters and Methodology

For a given group of firms, both entry and exit probabilities of a firm are determined by group–specific primitives. We employ a variant of a maximum–likelihood estimation procedure to estimate the primitives at firm level within groups using our rich data on firm entry and exit. The data from Irish manufacturing employment survey provides a long time span, up to 32 years. In the long span, with a large number of small firms the dataset allows us to observe strong micro turbulence in terms of high turnover rate, and lots of entry and exit, which make the estimates of primitives possible.

We need the following cohort–level structural parameters by proxy or by estimation:

1. ***Sunk Cost, Cs_t.***

We proxy this using the median size of all new firms in a grouping, which can be seen in many early works and is discussed in Sutton (1998).

2. ***Entry Constraint, h:*** This will be estimated.

3. ***Opportunity Value, I:*** This will be estimated.

4. ***Equilibrium Entry Possibility, P_t:*** This is approximated using the *entry rate* which is done in most of the early work (see Simon 1960 and more recently Sutton 1998).

5. ***The number of opportunities, N_t:*** This is approximated by the number of firms each year.

6. Duration, a: This will be estimated.

Mathematically, if we use *Entry Rate* to estimate the probability that a new opportunity is taken by a potential firm, a deduction is that the number of opportunities is the number of firms each year.[17] Intuitively, we can assume that there

[17] *p1 = Entry Rate,*
p2 = probability that a new opportunity is allocated to a new firm
p1=p2
By definition, we have
p1= New Entrants / Number of Firms, (each year)
*p2 * Number of opportunities = New Entrants*
This implies the Number of opportunities = Number of firms

is only one opportunity for each firm in a year, on average. These opportunity value could be positive or zero, which give firms different growths. A large number of firms imply that the opportunities for the group are more decentralized, whereas the value of them could be small.

Duration of opportunities is an important primitive in determining all the underlying processes and also the limiting FSD. However, we cannot analyze duration *a* directly, because in reality there are various time intervals (the unit for duration *a*) for cohorts of firms, which depends on the arrival speed of opportunities. Thus, duration *a* has to be adjusted by the number of opportunities in a year. Besides, we assume that duration *a* is drawn from some distribution and we are trying to estimate the mean for each cohort of firms.

At this point, we have three unknowns: *a, I,* and *h;* and three conditions: the entry process, exit process and equilibrium market size. Rewrite equation **(7), (8)** and **(11)** as:

1 Entry Process: $P_t = f\,(I, Cs_t, a, h, N_t)$ **(16)**

2 Exit Process: $Z = S + \textit{Expected Growth} - \textit{Expected Decline} = g\,(I, a, N_t)$ **(17)**

$Z > MES$ [18] *Firm survives*

$Z < MES$ *Otherwise*

3 Equilibrium Market Size: $\sum_{j=1}^{a} IN_{t-j} = \sum_{s} S_t$ [19] **(18)**

We employ a variant of a maximum–likelihood estimation procedure to estimate parameters at firm level within groups (for more detail on the methodology, see ***Appendix A***).

4.4 Estimation Results

First, we run regressions at the firm level within each industry. Thus, a group here means an industry. We dropped all the industries which exit before year 2003 and only focus on large industries with employment share over 0.01%, avoiding the

[18] We assume the existence of *MES* (minimum efficient size) but do not specify it.

[19] Total value of the market is the total opportunities obtained in recent duration *a* periods and all opportunities obtained before last *a* periods have vanished.

volatility. The regression result seems a little messy at first. But after dropping the insignificant coefficients, we find almost all the coefficients left are consistent with the equations in our model. For example, in the exit process, if expected growth and current size (ΔS_t and S_t) of a firm is controlled for, a growth in history will result in a current decline, which proved the existence of the duration of opportunities. Furthermore, in a few exceptions this only happened in some very small industries with share less than 5%. For some small industries with low turnovers, all the coefficients are insignificant, so that we can not estimate the parameters for those industries in the presence of the sample size problem. Thus, finally, we have estimates for 30 Home industries and 11 US industries.

In Table 1 we present the results. Different primitives across industries can be found. Industries that have estimates of opportunities with large duration *a*, are bodies for motor vehicles (3420 a=8) or ships and boats (3510 a=7). These would be expected to have longer production cycles. In the industries related to fashion, such as wearing accessories (1820) or games and toys (3650), the production cycle are not expected to be so long and their *a* values are only 1, which is what we expect. The highest entry profit (about 59) occurred in paper industry (2121) and computers and other information processing equipment (3002). They have high opportunity values (about 30) but also high sunk cost (about 14) and entry constraints.[20] Industries in television and radio receivers sound or video recording or reproducing apparatus (3650) has the highest opportunity value up to 80, but its entry profit is not the highest, although it is high (about 54). This could mainly result from its short duration in opportunities - the rapid advancement of technology increases competitive pressure and the risk of exit for new entrants.

If we take a weighted average, using employment share, of the estimated characteristics within Home industries and US industries respectively, we find for US industries, opportunities of longer duration, higher opportunity value and entry profit, but also we see higher sunk cost and entry constraints. The following table reports the

[20] The higher *h* is, the lower entry constraint is.

ratio of those characteristics for Irish Industries dominated by either Home or US companies.

	a	*I*	*Cs*	*Entry Profit*	*Entry Constraint*[21]
Home vs US	1 : 1.145	1 : 2.04	1 : 1.53	1 : 2.36	1 : 1.4

We next split the sample further by trade orientation and do regressions within each industry in each sub-sample. Thus, a group means all exporters or non–exporters in a given industry.

Table 2 (i) and (ii) documents the primitives for non–exporters and exporters, respectively. After dropping some groups without any significant coefficient, we have estimates for 21 Home industries and 9 US industries for non–exporters and 10 Home industries and 4 US industries for exporters.

The highest entry profit (121.68) is given for exporters in computers and other information processing (3002) and it is more than twice of the entry profit (about 59, the highest in previous estimation) of that industry without differentiating by trade orientation. At the same time, comparing the primitives of exporters with that estimated by industry only, we see 1.74 times higher opportunity values; 1.65 times higher durations, but also 2.7 times higher sunk costs and 2.6 times higher entry constraints. All this indicates is that we should differentiate the growth/decline and exit of firms in exporting with other non–export firms within industries. There is a common pattern in most industries. For instance, in production and preserving of meat industry (1511), compared to non–exporting counterparts (export: non–export), the exporter cohort has higher entry profit (57.22: 38.57), higher opportunity value (40.03: 24.76), higher sunk cost (38.97: 13.4) and longer durations of opportunities (3: 2). We take the weighted average of the primitives, using employment shares, within export and non–exporting populations of industries.

[21] The ratio of *h* of Home industries vs US industries is 1.4:1.

	a	*I*	*Cs*	*Entry Profit*	*Entry Constraint*
Non-exporters vs Exporters	1 **:** 1.34	1 **:** 3.18	1 **:** 4.45	1 **:** 4.47	1 **:** 5.86

An evident difference can be observed. *Cs, Entry Profit,* and *Entry Constraint* for exporters are more than four times of those for non–exporters. In addition, values of *a* and *I* value are much higher for exporting companies. This confirms conventional wisdom that exporting firms have higher sunk costs and other entry constraints but acquire access to bigger opportunities that have a longer life span. The primitives that drive non-exporting firm dynamics are very different when compared to exporters that co-exist in the same country and industry.

The patterns in the table above could play an important role to explain the heterogeneity of firm size distribution in Figure I (same result but shown in log–log scale).

5. Prediction

We have shown that the model is capable of explaining a wide array of regularities pertaining to industry dynamics. In this section, we simulate the limiting distribution in the model based on estimated primitives for non–exporting and exporting firms and compare the results with the actual data and the *Zipf* distribution. On the one hand, it gives a visual impression of the fit to actual data and shows the prediction power of the model. On the other hand, we aim to provide further support for the importance of group specific primitives that underpin heterogeneity in firm size distributions.

5.1 Methodology

First, we look at our limiting distribution derived in equation **(15)**. As mentioned above, empirically, duration *a*, should be adjusted by the number of opportunities each year; in our case it has been adjusted by the number of firms of a given cohort. We rewrite equation **(15)** as:

$$f(S) = A\left(\frac{1}{S} + K'\frac{(1-P)^S}{S}\right) \quad \textbf{(19)}$$

where $K' \equiv \dfrac{-f(1) + \dfrac{aNP}{1-P}}{f(1) - aNP}$ (or $K' \equiv \dfrac{-\dfrac{f(1)}{N} + \dfrac{aP}{1-P}}{\dfrac{f(1)}{N} - aP}$) and N denotes the equilibrium number of firms for a given cohort.

Thus, the limiting size distribution is determined by duration a, equilibrium entry rate P (determined by cohort characteristics, such as Cs, a, I), and, $\dfrac{f(1)}{N}$, the share of first rank.[22]

A tricky issue is how to define the first rank. This could be the minimum size in theory. As noted, high volatility is the feature of small firms and the definition of "small" is also different across groups, such as industries, so we have to setup various thresholds for the first ranks. We chose sunk cost as the threshold. New entrants tend to be small and the median size of new entrants (Cs) is a good measure for effective entry size in a given cohort (Sutton 1998), which should reflect the definition of "small" for the cohort. In the meanwhile, as the sample size is small we can see high year-dependent volatility of very small firms' behaviours, especially in the first ten years. Thus, if "small" is defined by a firm with size of 1, large fluctuations of "small" firms' share over years would be expected. Cs as a cohort–specific characteristics can smooth the fluctuation to some extent. Hence, the way we employ to estimate $\dfrac{f(1)}{N}$ is by the number share of firms with size less than sunk cost averaged over years.[23]

In contrast to probability mass function (PMF) in log–log scale[24] widely used in traditional works, we describe the simulation result and data by cumulative

---

[22] $f(1)$ is the number of firms with size 1 and, of course, is the number of firms in first rank in theory; N is the total number of all firms.

[23] We employ the average estimated $\frac{f(1)}{N}$ over last twenty years to avoid the high volatility observed in first ten years. And if no entry occurred, we define the minimum size as the first rank.

[24] x axis denotes rank; y axis denotes frequency of observations in each rank (not cumulative probability). In this case, function of Zipf law is a downward straight line with slope = -1.

distribution function (CDF) in log–log scale.[25] The main reason is that the sample size of Irish data is too small to do by the PMF. For example, we have about 5,000 firms in 2003, and about 300 ranks. If the firms are differentiated, on average there are only 7-8 firms in each rank. Taking account that many firms are within first three ranks, from the 5th or 6th rank above there is only 1-2 firms in each rank so that it is hard to describe data using PMF.

We compare the simulation results with FSD in year 2003 and Zipf distribution. Year 2003 is the last year in our data and we assume that distribution in the year is the closest to limiting distribution. And Zipf law, as the most robust empirical finding (Axtell 2001), provides an important reference for our comparisons. We emphasize the heterogeneity of cohorts and show that the model determined by our primitives can fit actual limiting distribution better than independent models.

5.2 Simulation Results

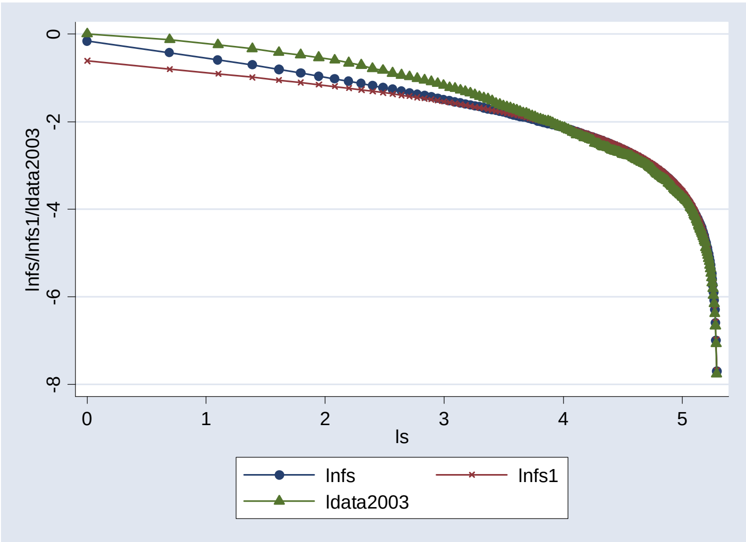
5.2.1 All firms

In theory, the most accurate way to simulate is within each group, such as a 4–digit industry, or a group of exporters within a 4-digit industry. However, as mentioned above, high volatility of FSDs of such small groups over time has been observed with there being small sample size in Irish manufacturing. Hence, for comparison purposes, we acquired the weighted average primitives of small groups (a group is defined as a 4 – digit industry) in the whole sample, and simulate the model based on the average primitives.[26] Then, we compare the simulated model with FSD of the whole sample and Zipf distribution.

As shown in Figure II, in the upper tail with large *S*, the model converges to a Zipf distribution and both the simulated model and Zipf distribution can perfectly fit the data. In the lower tail, our model performs better than the Zipf distribution, but neither fits the data very well. There are more small firms in the data than in either of the distributions. On the one hand, the parameters in the model reflect the

[25] x axis denotes rank in log scale; y axis denotes cumulative probability in log scale. And cumulative probability is the frequency of observations larger than this rank.
[26] Weighted average *a* and *P* of industries in Table 1 to obtain the average *a* and *P*.

characteristics of the sample to some extent, which gives us a better fit than the Zipf distribution. However, on the other hand, the defined group (the whole sample) is too broad without the consideration of heterogeneity and the weighted average can not exactly represent the common characteristics for the sample involving various industries and trade orientations. This implies that it is hard to fit FSD of all firms by any single distribution (consistent with Schmalensee, 1989), unless those firms have similar structural primitives.



(*a*=3.53, *P*=0.0668, $\frac{f(1)}{N}$ =0.24)

Figure II: Simulation Results for all firms

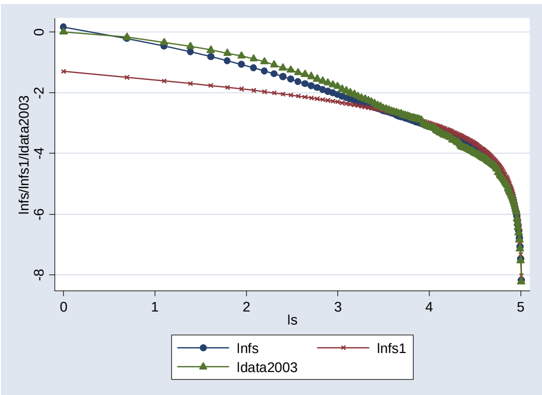
5.2.2 Exporters and Non–exporters

We split the sample by trade orientation and use the weighted average of primitives of small groups (a group is defined as firms with same trade orientation within same 4–digit industry) within each sub–sample.[27] Then, we compare the simulated model with FSD of each sub–sample, respectively.

The results are shown in Figure III. In both of the sub-samples, the model and Zipf distribution can fit the data well in the upper tail. However, it is impressive

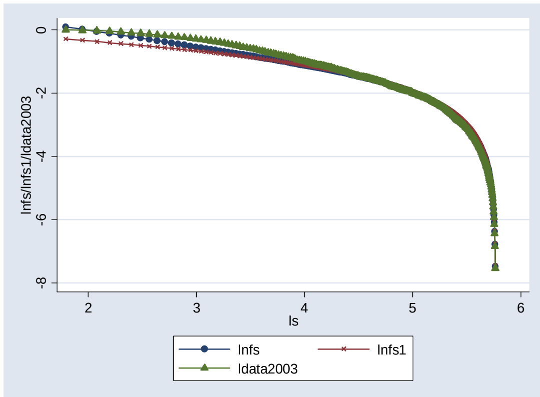
[27] Weighted average *a* and *P* of cohorts in Table II (i) and (ii) respectively to obtain the average *a* and *P* for each sub–sample.

that the simulated distribution fits much better than the Zipf distribution in the lower tail. Especially for non–exporters, a large gap can be observed between Zipf distribution and the actual data, compared to our model. Though exporters and non–exporters have very different FSD, as shown in Figure I, the model is able to capture the difference because it is endogenously determined by the primitives that drive exporter and non-exporter dynamics within industries. These happen to be fundamentally different.



(*i*) Non – Exporters

($\boldsymbol{a}$=2.66, $\boldsymbol{P}$=0.069, $\frac{f(1)}{N}$=0.187)



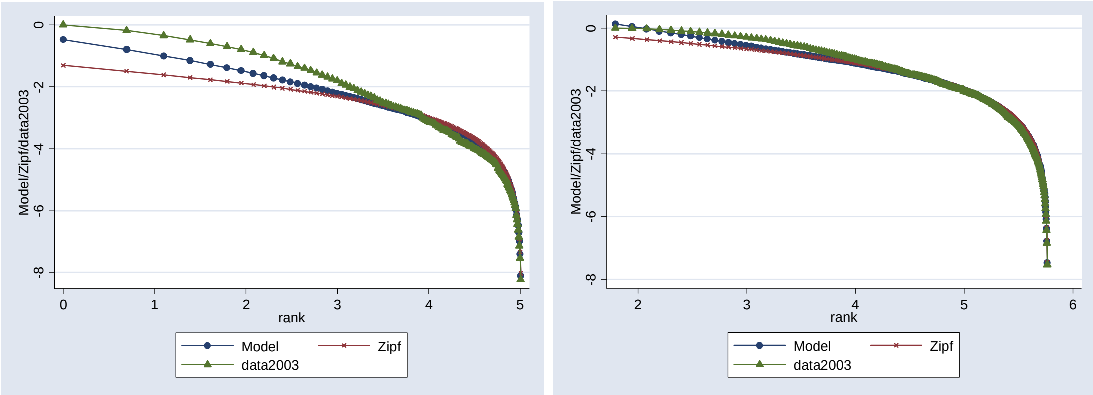
(*ii*) Exporters

($\boldsymbol{a}$=3.64, $\boldsymbol{P}$=0.053, $\frac{f(1)}{N}$=0.195)

Figure III Simulation Results for Non – Exporters and Exporters

To provide more evidence supporting the idea that exporters and non-exporters firm dynamics are driven by different primitives, we simulate the model by using the weighted average of primitives from the whole firm population sample (used in Figure II) and compare it with FSDs of exporters and non–exporters, respectively (as shown in Figure IV). We find that in the upper tail the simulated model fits well with the actual distributions of both non–exporters and exporters, especially for exporters, which means that for FSD of large firms, the dependence on structural primitives is low. However, in the lower tail, deviation of the model from FSD of non–exporters can be observed, though it shows better performance than Zipf distribution. Comparing with the well–fit for Non-Exporters in Figure III (i), Figure

IV (i) implies that when we are trying to fit the FSD of the whole sample, the deviation observed in the lower tail mainly comes from small non–export firms, the number of which is much more than expected.



(*i*) Non – Exporters (*ii*) Exporters

$$(\boldsymbol{a}=3.53,\ \boldsymbol{P}=0.0668,\ \frac{f(1)}{N}=0.24)$$

Figure IV Simulate sample parameters and compare with Non–Exporters and Exporters

6 Conclusions

This paper has proposed an endogenous model in which a firm's fortunes are determined by the success they have in exploiting valuable opportunities and the time that elapses before those opportunities vanish. Turning our focus away from purely statistical patterns we try to learn more about the determinants of market structure. We introduced a set of structural primitives, sunk cost, entry constraints, opportunity value, and opportunity duration, which are the determinants of the entry, growth and exit processes that underline industrial evolution. We show how different primitives result in different distributions of firm size for defined groups of firms, under the roof of the same stochastic processes. This implies that no single exogenous model can fit all firm cohorts and no model should aggregate over firms with different primitives unless those structural primitives have a linear property. Our focus was to show that exporting and non-exporting firm dynamics inside industries are driven by different structural primitives which generate different limiting size distributions.

In addition to explaining heterogeneity of firm size distribution, we showed that the model can explain a large number of empirical regularities that have been the subject of intense theorizing. The model predicts that (i) firm growth is decreasing in size; (ii) firm survival is increasing in size; (iii) the autocorrelation of firm growth; (iv) conditional on firm size and growth history, the exit rate is increasing in the entry rate. These predictions match precisely with empirical observation.

Jovanovic's (1982) selection model famously predicts impeded growth and shakeouts of new entrants. His omitted variable is the discovery of a firm's quality which rises with age due to incomplete information. Our model and his are not mutually exclusive. We proposed another mechanism from the demand side. Finite duration of opportunities reflects discrete demand shocks to firms. A small firm has good growth under the protection of opportunity it obtained; however, small firms and new entrants underwent shakeouts as they can not afford demand shocks (losing opportunities), even though they know exactly their true costs and profits. Therefore, once finite demand shock is accounted for, we can generate size, growth, and survival patterns that appear consistent with the empirical regularities using a simple model with complete information.

Much of the early stochastic models on firm size distributions and growth was directed toward what exogenous statistical regularity suits and what kind of distribution fits data well. In principle, our model can generate any concavity in the lower tail of distribution and roughly fit distributions between logarithmic and Zipf distribution by specifying parameter K and equilibrium entry rate P. The model predicts Zipf distribution in the upper tail, which is broadly observed for large firms. To avoid pre–defining statistical relationships, Sutton (1998) derives a geometric distribution to model a lower bound Lorenz curve by imposing a symmetry principle. Taking consideration of finite duration of opportunities, the geometric distribution will be broken in the presence of opportunity outflow. As a result, the lower bound could be bent in both the lower tail and the upper tail and it depends on equilibrium entry rate *P*, which is a function of firm cohort characteristics.[28] This extension implies a lower bound could be heterogeneous across groups of firms.

On the empirical side, we estimated the structural primitives to quantitatively determine the sources of heterogeneity in firm size distribution. A thick upper tail in the FSD of exporters compared with that of non–exporters can be explained by higher value opportunities of longer duration, higher sunk (trade) cost and entry constraints. High sunk cost and entry constraint lower entry rates and long duration lowers the risk of firm failures. Low turnover and high opportunity value give good growth for the exporters, which skew the firm size distribution right. We did not expect that the exact values of primitives can be obtained - we only wanted to assess the quantitative power of determinants in a comparative sense. Indeed, one can obtain more accurate estimates by simulation after imposing relations to primitives of some specific firm–cohorts.

The structure of the model is intentionally very general and in reality there are a wide variety of group–specific primitives which may affect entry and exit processes and, finally, the firm size distribution. Our group–specific structural primitives are estimated by trade orientation within narrowly defined industries. Hard for us to say whether our key grouping and model would work well in all economies but it is highly likely that the model may serve well for some. We hope it can provide a way forward to understanding the forces behind the evolution of firm size distribution.

---

[28] Sutton (1998) derives geometric distribution $(1 - \lambda)\lambda^n$ by assuming that every incumbent has the same possibility to capture a new opportunity. In the limit, the firm size distribution is restricted to a lower bound Lorenz Curve. If considering finite duration of opportunities, each opportunity has equal chance to exit as all the opportunities have the same duration distribution. Thus, opportunity outflow will break geometric distribution by imposing outflow distribution of $n(1 - \lambda)\lambda^n$. And if $\lambda > 0.5$ there are more exits of small firms, which will flatten Geometric distribution; if $\lambda < 0.5$, the highest possibility of exit is for firms with size of $[-\frac{1}{\ln \lambda}]$. It implies that the lower bound is endogenously determined by λ, which is a function of firm cohort primitives.

Appendix A:

We have three unknowns, that is, *a*, *I*, and *h*, and three conditions, the entry process, exit process and equilibrium market size.

(1) Exit Process

Let $Z = S + Expected\ Growth - Expected\ Decline$ **(20)**

where *S* denotes the current size of firm.

We have

$Z > MES$ *Firm survive*

$Z < MES$ *Otherwise*

For a given firm *j* in a given group,

$$Expected\ Growth_t = \frac{S_{t,j}}{\sum_j S_t} N_t I(1-P_t) \quad \textbf{(21)}$$

where $S_{t,j}$ denotes the current size of firm *j* in period *t*; $\sum_j S_t$ denotes market size in period *t*; N_t denotes the number of firms in period *t* and it reflects the number of opportunities each period as mentioned above; P_t denotes entry rate in period *t*.

$$Expected\ Decline_t = \sum_k (Expected\ Growth\ in\ t - k\ Period) * P(a=k)$$

$$= \sum_k \frac{S_{t-k,j}}{\sum S_{t-k}} N_{t-k} I(1-P_t) P(a=k) \quad \textbf{(22)}$$

where $P(a=k)$ denotes the probability of *duration*, $a=k$.

Combining equation **(20), (21)** and **(22)**, we have:

Exit Condition: $Z - MES > 0$ *Firm Survive*

$Z - MES < 0$ *Otherwise*

$$Z - MES = S_{t,j}\left(1 + \frac{1}{\sum_j S_{t,j}} N_t I(1-P_t)\right) - \sum_k \frac{S_{t-k,j}}{\sum_j S_{t-k,j}} N_{t-k} I(1-P_t) P(a=k) - MES$$

We employ logit model at firm level on panel data to estimate within each firm–cohort and have the following regression:

$$y = x_0\beta_0 + \sum_k x_k\beta_k + c \qquad \textbf{(23)}$$

y = 1 *Firm Survive*

y = 0 *Otherwise*

where $x_0 = S_{t,j}$; $x_k = \dfrac{S_{t-k,j}}{\sum_j S_{t-k,j}} N_{t-k}(1-P_t)$; and c = - *MES*

And the latent equation is

$$Z - MES = \quad x_0\left(1 + \frac{1}{\sum_j S_{t,j}} N_t I(1-P_t)\right) + \sum_k x_k IP(a=k) + c \qquad \textbf{(24)}$$

Comparing equation **(23)** and **(24)**, it is easy to find the relation between $IP(a=k)$ and β_k . Note that $\hat{\beta}_k$ is not the estimated value of $IP(a=k)$; because in the maximizing- likelihood method, the regression results depend on the error variance in the corresponding underlying latent models. For example, in the logit model the error variance is normalized to $\frac{\pi^2}{3}$, but 1 in the probit model. If do the same regression in both models, we can find the estimates obtained from the logit model are roughly a factor $\frac{\sqrt{\pi}}{3}$ larger than those obtained from the probit model.

By the logic, we introduce a normalized factor *m* for the estimates and have

$$\hat{\beta}_k = -mI\hat{P}(a=k)$$ [29] for k = 1, 2, …,K [30]

Combining $\sum_k P(a=k) = 1$, we have $\hat{P}(a=k) = \dfrac{\hat{\beta}_k}{\sum_{i=1}^{K}\hat{\beta}_i}$

Thus, $$\hat{E}(a) = \sum_{k=1}^{K} k\,\hat{P}(a=k) = \sum_{k=1}^{K} k\,\frac{\hat{\beta}_k}{\hat{\beta}_1 + \hat{\beta}_2 + \ldots + \hat{\beta}_K} \qquad \textbf{(25)}$$

[29] We can see almost all the significant coefficients are negative; only few exceptions happened in some very small industries with a share less than 5% and we drop them. This confirmed the existence of duration.

[30] In our estimation, we impose K=8.

(2) **Equilibrium Market Size**

Rewrite equation **(18)** as: $\sum_{j=1}^{E(a)} IN_{t-j} = \sum_{s} S_t$

As mentioned above, N_t opportunities arrive in period t and the total value of the market, S_t, is the total opportunities obtained in recent duration a periods.

Thus, we have the estimated $$\hat{E}(I) = \hat{E}\left(\frac{\sum_{s} S_t}{\sum_{j=1}^{E(a)} N_{t-j}}\right) \approx \frac{E\left(\frac{\sum_{s} S_t}{N_t}\right)}{\hat{E}(a)} \qquad \textbf{(26)}$$

(3) **Entry Process**

Rewrite equation **(5)** as **(27)**, we can easily estimate another unknown h,

$$P_t = h{*}E(Entry\ Profit) = h(\hat{E}(I) - Cs_t + \hat{E}(I)((1 + \frac{(1-P_t)\hat{E}(I)}{\sum_{j} S_{j,t}})^{\hat{E}(a)} - 1)) \qquad \textbf{(27)}$$

where P_t is the entry rate of cohort.

References

Agarwal, Rajshree, and Gort, Michael (1996). First – mover advantage and the speed of competitive entry, 1987 – 1996. *Journal of Law and Economics,* 44: 161 – 177.

Amaral, L. A. N., Gopikrishnan, P., Plerou, V. and Stanley, H. E, (2001). A model for the growth dynamics of economic organization. *Physica A*, Vol. 299, pp. 127 – 136.

Audretsch, D.B., and Mahmood, T. (1995). New firm survival: new results using a Hazard function. *Review of Economics and Statistics*, 77(1): 97 – 103.

Audretsch, D.B., Houweling P, Thurik AR (2000). Firm survival in the Netherlands. *Review of Industrial Organization,16*: 1–11.

Axtell, R.L.,(2001). Zipf distribution of US firm sizes. *Science 293,* 1818 – 1820 September 7th.

Bottazzi, G. and Secchi, A (2005). Explaining the distribution of firms growth rates. LEM paper series, number 2005/16.

Bottazzi, Gilulio, and Angelo Secchi, (2005). Explaining the distribution of firms growth rates, LEM Working Paper Series.

Cable, J. and Schwalbach, J. (1991). International comparisons of entry and exit. In P. A. Geroski and J. Schwalbach (Eds), *Entry and Market Contestability: an International Comparison.* Blackwell Press.

Cabral, Luis M. B. and Mata, J, (2003). On the evolution of the firm size distribution: facts and theory. *The American Economic Review*, Vol. 93, No. 4 (Nov., 2003), pp. 1075-1090.

de Wit, Gerrit (2004). Firm size distribution An overview of steady – state distribution resulting from firm dynamics models. *International Journal of Industrial Organization* 23 (2005) 423 – 450.

Dinlersoz, E. M. and Macdonald, G. (2008). The industry life – cycle of the size distribution of firms. *Review of Economic Dynamic,* 07 – 107.

Doi N (1999). The determinants of firm exit in Japanese manufacturing industries. *Small Business Economics 13*: 331–337.

Dosi, G., Marsili, O., Orsenigo, L. and Salvatore, R. (1995). Learning, turbulence, and the dynamics of market shares. *The Review of Economics and Statistics*, LXXIX (3).

Dunne, T., Roberts, M.J., (1991). Variation in producer turnover across U.S. manufacturing industries. In Geroski PA, Schwalbach J (eds). *Entry and Market Contestability: an International Comparison*. Blackwell, London 187–203.

Dunne, T.., Roberts, M..J., and Samuelson, L.(1995). The growth and failure of U.S. Manufacturing plants. *Quarterly Journal of Economics*, Vol. 62, pp. 53 – 82.

Evans, David S., (1987 a). The relationship between firm growth, size, and age: estimates for 100 manufacturing industries. *Journal of Industrial Economics* 35 (4), 567 – 581.

Evans, Laurie Beth and John J. Siegfried (1994). Entry and exit in United States Manufacturing from 1977 to 1982. In David B. Audrestch and John J. Siegfried (eds), *Empirical Studies in Industry Organization: Essays in Honor of Leonard W. Weiss*. Boston: Kluwer Academic Publisers, 253 – 74.

Geroski, P.A (2000). Growth of firms in theory and in practice. *New Directions in Economics Strategy Research*. Oxford: Oxford University Press.

Hall, Bronwyn H., (1987). The relationship between firm size and firm growth in the U.S. manufacturing sector, *Journal of Industrial Economics* **35** (1987), pp. 583–606.

Hart, O., and S.J. Prais (1956). The analysis of business concentration: a statistical approach. *Journal of the Royal Statistical Society* **119** (1956), pp. 150–191.

Hart, Peter E. and Nicholas Oulton, (1996). Growth and size of firms. *Economic Journal*, 106(3), pp. 1242-1252.

Hymer, S. and Pashigian, P (1962). Firm size and rate of growth. *Journal of Political Economy*, Vol. 70, pp. 556 – 569.

Ijiri, Yuji, and Simon, Herbert, (1975). Some distributions associated with Bose – Einstein statistics.

Ijiri, Yuji, and Simon, Herbert, (1977), Skew distribution and the sizes of business firms. North – Holland, Amsterdam.

Jovanovic, Boyan, (1982). Selection and the evolution of industry. *Econometrica*, 50 (3), pp. 649 – 670.

Klepper, S. and Thompson, P., (2005). Submarkets and the evolution of market structure. *RAND Journal of Economics, winter, 2006.*

Klepper. S. and Graddy, E. (1990). The evolution of new industries and the determinants of market structure. *RAND Journal of Economics, 21*, 27 – 44.

Knudsen, Thorbjorn, (2001). Zipf's law for cities and beyond: The case of Denmark. *American Journal of Economics and Sociology*, Jan 2001.

Kumar, Manmohan S., (1985). Growth, acquisition activity and firm size: evidence from the United Kingdom. *Journal of Industrial Economics, 33(3)*, pp. 327-338.

Laincz, C. and A. Rodrigues, (2005). The impact of Cost – Reducing R&D spillovers on the Ergodic distribution of market structures. Working Paper.

Laincz, Christopher A. and Ana Sofia Domingues Rodrigues, (2005). A theoretical foundation for understanding firm size distributions and Gibrat's law. *The RAND Journal of Economics*, L11, L13.

Levy, M. and Solomon, Sorin, (1996). Power laws are logarithmic Boltzmann laws. *International Journal of Modern Physics C* **7** (1996), p. 595.

Li, Q., Walsh, P.P., and Whelan, C., (2008). Jobless growth through creative destruction: Ireland industrial development path 1972 – 2003. IZA Discussion Paper, No 3211.

Lotti, Francesca and Enrico Santarelli, (2001). Industry dynamic and the distribution of firm size: a non-parametric approach, LEM Papers Series, number 2001/14.

Luttmer E. G. J. (2004). The size distribution of firms in an economy with fixed and entry costs. Working paper 663, Federal Reserve Bank of Minneapolis.

Mahmood T (2000). Survival of newly founded businesses: a log-logistic model approach. *Small Business Economics 14*:223–237.

Malcai, Ofer, Biham, Ofer, and Solomon, S., (1999). Power – law distribution and Levy – stable intermittent fluctuation in stochastic systems of many autocatalytic elements. *The American Physical Society*, *Volume 60, Number 2*, August 1999.

Marsili O, (2006). Stability and turbulence in the size distribution of firms: evidence form Dutch manufacturing. *International Review of Applied Economics, Vol. 20, No.2,* 255 – 272, *April 2006*.

Mayer, W.J., Chappell, W.F. (1992). Determinants of entry and exit: an application of the compounded bivariate Poisson distribution to U.S. industries. *South Economics 58*:770 –778.

Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity, *Econometica*, Volume 71 Issue 6, Pages 1695 – 1725.

Nocke, V and Yeaple, S. (2008). Globalization and the size distribution of multiproduct firms. CEPR Discussion Paper No. DP6948.

Schmalensee, Richard (1989). Inter–industry studies of structure and performance. In R. Schmalensee and R. Willg, eds., *Handbook of Industrial Organization, Vol. 2*. Oxford: North – Holland, pp. 952 – 1009.

Simon, H. A. and Bonini, C. P. (1958). The size distribution of business firms. *American Economic Review* **48** (1958), pp. 607–617 (Also printed in Ijiri and Simon (1977, ch. 7)).

Simon, H. A., (1960). Some further notes on a class of skew distribution functions. *Information and Control* **3** (1960), pp. 80–88 (Also printed in Ijiri and Simon (1977, ch. 2)).

Simon, H.A., (1955). On a class of skew distribution functions. *Biometrika* **52** (1955), pp. 425–440 (Also printed in Ijiri and Simon (1977, ch. 1)).

Singh, A. and Whittington, G. (1975). The size and growth of firms. Review of Economics Studies, Vol. 48 (1958), pp. 607 – 617.

Steindl, J., (1965). Random processes and the growth of firms. A Study of the Pareto Law, Griffin and Company (1965).

Sutton, J., (1997). Gibrat's legacy, *Journal of Economic Literature* **35** (1997), pp. 40–59.

Sutton, J., (1998). Technology and market structure. Cambridge, MA: MIT Press

Wagner J., (1994), Small firm entry in manufacturing industries. *Small Business Economy 5*:211–214.

Table 1: **Estimation Results by Industries**

FDI	Nacecode	Nace Description	Duration	Opportunity	Sunk Cost	Entry Rate	Entry Profit	h
0	1551	Operation of dairies and cheese making	8.00	8.38	21.16	0.030	12.08	0.00252
0	1730	Finishing of textiles	2.99	5.14	6.83	0.080	5.65	0.01418
0	1810	Leather clothes	1.00	16.80	5.89	0.123	8.54	0.01446
0	1820	Other wearing apparel and accessories	1.00	32.12	8.55	0.064	33.93	0.00189
0	1822	Other outerwear	5.00	8.87	10.17	0.101	14.73	0.00682
0	1920	Luggage handbags and the like saddlery and harness	4.00	2.81	4.27	0.063	3.08	0.02041
0	2010	Sawmilling and planing of wood; impregnation of wood	3.27	4.64	5.77	0.037	8.87	0.00413
0	2121	Corrugated paper and paperboard and of containers of paper and paperboard	3.00	31.07	13.29	0.042	59.57	0.00071
0	2222	Printing n.e.c.	3.00	7.82	6.38	0.034	14.65	0.00235
0	2522	Plastic packing goods	3.00	37.17	11.19	0.094	38.24	0.00245
0	2620	Non-refractory ceramic goods other than for construction purposes; refractory ce	1.00	5.71	1.94	0.128	4.33	0.02959
0	2621	Ceramic household and ornamental articles	3.00	6.49	4.38	0.082	8.44	0.00969
0	2670	Cutting shaping and finishing of stone	1.39	9.83	3.67	0.058	12.49	0.00463
0	2800	Fabricated metal products except machinery and equipment	3.00	6.33	9.92	0.102	6.11	0.01664
0	2810	Structural metal products	1.00	19.06	9.23	0.175	14.36	0.01219
0	2811	metal structures and parts of structures	4.70	4.24	11.04	0.048	3.60	0.01341
0	2812	Builders carpentry and joinery of metal	3.00	6.46	8.92	0.087	9.24	0.00940
0	2821	Tanks reservoirs and containers of metal	2.00	25.06	10.14	0.059	26.93	0.00219
0	2822	Central heating radiators and boilers	1.00	44.28	21.74	0.095	33.74	0.00281
0	2851	Treatment and coating of metals	1.00	9.57	7.95	0.083	7.76	0.01070
0	2875	Other fabricated metal products n.e.c.	2.45	8.09	4.42	0.056	13.08	0.00431
0	2920	Other general purpose machinery	1.00	11.52	5.00	0.148	11.12	0.01334
0	2922	Lifting and handling equipment	1.00	19.15	10.50	0.101	18.23	0.00552
0	2932	Other agricultural and forestry machinery	3.89	3.44	5.43	0.065	4.44	0.01466
0	3410	Motor vehicles	6.05	6.48	8.34	0.057	10.19	0.00557
0	3420	Bodies (coachwork) for motor vehicles; trailers and semi-trailers	8.00	1.69	6.97	0.025	-1.13	-0.02179
0	3510	Building and repairing of ships and boats	7.00	0.32	2.50	0.152	-1.71	-0.08896

0	3610	Furniture	6.00	2.90	4.26	0.124	5.56	0.02222
0	3614	Other furniture	2.59	4.26	5.73	0.050	6.93	0.00726
0	3663	Other manufacturing n.e.c.	8.00	2.08	2.97	0.123	5.52	0.02222
1	1589	Other food products n.e.c.	3.00	12.74	10.57	0.103	23.22	0.00443
1	1823	Underwear	2.97	31.79	19.57	0.081	54.56	0.00149
1	2440	Pharmaceuticals medicinal chemicals and botanical products	5.00	17.46	32.11	0.144	25.80	0.00560
1	2924	Other general purpose machinery n.e.c.	4.00	9.58	8.43	0.081	17.79	0.00454
1	2952	Machinery for mining quarrying and construction	2.00	26.80	8.21	0.034	47.45	0.00072
1	2956	Other special purpose machinery n.e.c.	7.00	3.46	14.83	0.093	1.67	0.05609
1	3002	Computers and other information processing equipment	4.00	28.59	14.87	0.164	59.00	0.00278
1	3150	Lighting equipment and electric lamps	2.00	18.01	4.50	0.064	26.02	0.00247
1	3162	Other electrical equipment n.e.c.	8.00	5.85	7.20	0.104	8.17	0.01275
1	3230	Television and radio receivers sound or video recording or reproducing apparatus	1.00	80.35	12.41	0.110	53.96	0.00149
1	3650	Games and toys	1.00	18.07	6.12	0.087	16.31	0.00533
Home		**Overall**	**3.46**	**9.38**	**7.20**	**0.061**	**12.49**	**0.00486**
Us		**Overall**	**3.96**	**19.11**	**10.62**	**0.103**	**29.47**	**0.00349**

***Table 2 (i):* Estimation Results by Industries for Non-exporters**

FDI	Nacecode	Nace Description	Duration	Opportunity	Sunk Cost	Entry Rate	Entry Profit	h
0	1511	Production and preserving of meat	2.00	24.76	13.40	0.049	38.57	0.00126
0	1513	Production of meat and poultry meat products	3.96	3.48	5.88	0.140	2.84	0.04927
0	1551	Operation of dairies and cheese making	8.00	5.37	8.43	0.034	8.73	0.00394
0	1722	Woollen-type weaving	2.00	11.96	7.05	0.066	8.73	0.00751
0	1754	other textiles n.e.c.	8.00	8.26	12.53	0.158	9.49	0.01665
0	1820	Other wearing apparel and accessories	1.00	20.92	7.48	0.069	19.93	0.00348
0	1920	Luggage handbags and the like saddlery and harness	4.00	2.29	4.19	0.068	2.00	0.03398
0	2224	Composition and plate-making	2.15	5.76	9.06	0.029	2.69	0.01073
0	2524	Other plastic products	1.00	14.81	5.42	0.088	12.67	0.00694
0	2610	Glass and glass products	1.00	20.87	3.29	0.090	19.22	0.00467
0	2811	Metal structures and parts of structures	1.00	15.54	7.17	0.049	17.78	0.00278
0	2812	Builders carpentry and joinery of metal	3.99	2.94	8.60	0.089	-0.64	-0.13938
0	2851	Treatment and coating of metals	1.00	7.73	8.26	0.079	5.18	0.01516
0	2875	Other fabricated metal products n.e.c.	3.00	4.38	4.27	0.058	5.50	0.01060
0	2920	Other general purpose machinery	1.00	7.93	5.56	0.158	6.38	0.02470
0	2932	Other agricultural and forestry machinery	1.00	8.67	4.76	0.065	7.63	0.00859
0	3410	Motor vehicles	1.00	36.51	7.80	0.052	29.72	0.00176
0	3420	Bodies (coachwork) for motor vehicles; trailers and semi-trailers	8.00	1.40	6.00	0.023	-1.29	-0.01772
0	3612	Office and shop furniture	1.00	9.59	5.67	0.089	7.92	0.01129
0	3614	Other furniture	2.78	3.14	3.05	0.051	4.60	0.01106
0	3663	Other manufacturing n.e.c.	8.00	1.32	2.42	0.136	3.95	0.03444
1	1751	Carpets and rugs	1.00	36.20	7.35	0.060	16.37	0.00366
1	1772	Knitted and crocheted pullovers cardigans and similar articles	1.00	12.52	5.33	0.067	9.66	0.00698
1	1823	Underwear	4.00	18.49	14.91	0.095	28.57	0.00332

1	2924	Other general purpose machinery n.e.c.	4.99	2.52	6.29	0.091	1.48	0.06161
1	2956	Other special purpose machinery n.e.c.	7.00	1.56	5.18	0.088	-0.74	-0.11962
1	3162	Other electrical equipment n.e.c.	5.89	3.01	8.22	0.108	2.89	0.03720
1	3430	Parts and accessories for motor vehicles and their engines	1.00	10.10	12.40	0.078	5.03	0.01558
1	3622	Jewellery and related articles n.e.c.	2.00	2.63	2.33	0.070	3.58	0.01952
1	3650	Games and toys	1.00	7.94	5.71	0.094	4.54	0.02064
		Residuals	2.00	9.26	6.84	0.075	8.64	0.00868
0		Home Overall	2.72	9.45	5.82	0.067	9.93	0.00679
1		US Overall	2.51	7.24	5.95	0.079	6.03	0.01309
		Overall	**2.66**	**9.07**	**5.84**	**0.069**	**9.25**	**0.00750**

Table 2 (ii): **Estimation Results by Industries for Exporters**

FDI	Nacecode	Nace Description	Duration	Opportunity	Sunk Cost	Entry Rate	Entry Profit	h
0	1511	Production and preserving of meat	3.00	40.03	38.97	0.035	57.22	0.00061
0	1551	Operation of dairies and cheese making	3.00	45.17	60.79	0.019	65.15	0.00029
0	1740	Made-up textile articles except apparel	4.00	16.87	30.50	0.037	12.37	0.00298
0	1822	Other outerwear	2.00	31.29	28.71	0.048	36.10	0.00132
0	2112	Paper and paperboard	3.56	18.91	17.29	0.029	23.93	0.00123
0	2213	Publishing of journals and periodicals	4.46	41.39	13.00	0.011	81.79	0.00013
0	2522	Plastic packing goods	3.00	39.28	13.71	0.084	43.02	0.00196
0	2811	Metal structures and parts of structures	8.00	6.13	19.41	0.035	-2.02	-0.01709
0	2875	Other fabricated metal products n.e.c.	2.00	29.14	17.81	0.044	49.78	0.00088
0	2932	Other agricultural and forestry machinery	3.00	12.14	13.50	0.062	18.15	0.00342
1	1823	underwear	1.00	116.51	36.38	0.062	94.36	0.00066
1	2513	Other rubber products	3.00	21.79	17.96	0.053	35.68	0.00149
1	3002	Computers and other information processing equipment	4.94	49.87	39.69	0.130	121.68	0.00106
1	3162	Other electrical equipment n.e.c.	8.00	18.41	24.50	0.119	56.35	0.00211
		Residuals	3.50	24.88	23.05	0.051	29.11	0.00174
0		Home Overall	3.39	27.68	26.22	0.045	36.24	0.00123
1		US Overall	4.43	32.82	27.24	0.081	58.20	0.00139
		Overall	**3.64**	**28.88**	**26.46**	**0.053**	**41.38**	**0.00128**