Available online at www.sciencedirect.com
SCIENCE DIRECT
International Journal of Industrial Organization
23 (2005) 505–534
International Journal of Industrial Organization
www.elsevier.com/locate/econbase

Firm productivity, heterogeneity, sunk costs and market selection

Jose C. Fariñas [a,*], Sonia Ruano [b]

[a] Universidad Complutense de Madrid, Spain
[b] Bank of Spain, Spain

Received 31 January 2002; received in revised form 3 March 2004; accepted 25 February 2005
Available online 13 June 2005

Abstract

The main objective of this paper is to explore whether or not patterns of entry and exit are systematically related to productivity differences at the firm level, as suggested by models of industry dynamics Hopenhayn, (1992) [Hopenhayn, Hugo (1992). Entry, Exit, and firm dynamics in long run equilibrium, Econometrica, 60, September, pp.1127-1150]. The comparisons of productivity distributions for entering, exiting and continuing firms are performed by non-parametric procedures and for a large scale firm-level panel data set of Spanish manufacturing firms. The main empirical findings can be summarized as follows. First, heterogeneity in productivity levels across firms is persistent through time. Second, entry and exit decisions are systematically related to differences in firm productivity. In particular, the productivity distribution of continuing firms stochastically dominates the distributions of entering and exiting firms. Third, at the moment of entry, the group of failing members of any entry cohort has lower productivity than the group of surviving members of the same entry cohort. Fourth, the post-entry productivity level of entering firms grows more rapidly than the productivity of incumbent firms, although this pattern is not always highly significant. Finally, we find that sunk costs are an important source of heterogeneity across firm productivity. The evidence we find is consistent with models of industry dynamics predicting lower productivity for firms operating in markets with a higher level of sunk entry costs.

JEL classification: D24; M20; L10

Keywords: Total factor productivity; Entry; Exit; Market selection

* Corresponding author.
E-mail address: farinas@ccee.ucm.es (J.C. Fariñas).

0167-7187/$ - see front matter
doi:10.1016/j.ijindorg.2005.02.002

1. Introduction

This paper uses a micro panel data set of Spanish manufacturing firms to examine productivity differences between groups of entering, exiting and continuing firms. To organize our empirical work, we rely on the emerging theoretical literature that seeks to account for the observed heterogenous productivity across individual producers. The field of industrial economics has contributed to the analysis of this heterogeneity through models of industry dynamics proposed by Jovanovic (1982), Hopenhayn (1992), Ericson and Pakes (1995), where the path of growth and failure that characterize micro data is driven to a large extent by differences in productivity. They allow for industry equilibriums with simultaneous flows of firm entry and exit and heterogeneity in productivity at the firm level. In particular, we rely on the Hopenhayn's (1992) model that emphasizes the notion of sunk cost to explain productivity differences. On the empirical side, Tybout (1996), Caves (1998), Bartelsman and Doms (2000), Haltiwanger (2000) and Foster et al. (2001) provide excellent reviews of the literature. We examine productivity differences between firms that are similar to those comparisons performed by Aw et al. (2000, 2001, 2002). A related literature reports decompositions that try to measure the contribution of entry, exit and share effects to productivity growth as in Baily et al. (1992), Griliches and Regev (1995), Olley and Pakes (1996) and Foster et al. (2001), among others. Baldwin (1993), Audretsch (1995) and Roberts and Tybout (1996) analyze different aspects of the dynamic process by which firms evolve over time, including productivity.

The main purpose of this paper is to explore if entry and exit patterns are related to productivity differences at the firm level, as suggested by models of market dynamics. In particular, we focus on the exam of productivity differences between continuing, entering and exiting firms. We take as reference a sample of Spanish manufacturing firms over the period 1990–1997. The paper makes three contributions to the literature. First, we add another national perspective to the available evidence based on firm level productivity. Second, we use the methodology proposed by Delgado et al. (2002) that compares the entire distribution of firm productivity rather than just marginal moments, typically means. In particular, we compare the cumulative distribution function of total factor productivity for continuing, exiting and entering firms. The paper implements a testing procedure based on the concept of stochastic dominance for ranking differences between productivity distributions. Third, we test if productivity differences between groups of firms are consistent to market selection forces and, also, if sunk costs, that we measure at the firm level, influence the distribution of productivity levels across incumbent firms.

Our results are largely similar to those found in parallel studies (see Aw et al., 2001, 2002; Foster et al., 2001) and, therefore, provide evidence favorable to some of the implications of models of industry dynamics. The main empirical findings can be summarized as follows. First, heterogeneity in productivity levels across firms is persistent through time. Second, entry and exit decisions are systematically related to differences in firm productivity and, in particular, the productivity distribution of continuing firms stochastically dominates the distributions of entering and exiting firms. Third, for any entry cohort, the initial productivity distribution of the failing members is stochastically dominated by the distribution of surviving members. Fourth, the behavior of post-entry productivity growth indicates that the productivity level of surviving entrants grows more rapidly and tend to reach the level of incumbent firms. Finally, we confirm that the level of sunk costs is associated to productivity differences across cohorts of continuing firms. In particular, as predicted by Hopenhayn's (1992) model, firms operating in industries with high sunk costs have lower productivity than firms within industries of low sunk costs.

The paper is organized as follows. Section 2 describes a theoretical framework relating firm turnover and productivity levels. The testing procedure that has been used throughout the paper is also explained. Section 3 describes the data set and the index used for measuring firm productivity. Section 4 presents the main empirical results and Section 5 concludes.

2. A framework for analysis

2.1. Entry, exit and productivity differences

To organize our empirical work, we rely on the model of firm and market dynamics developed by Hopenhayn (1992). This model characterizes the long-run equilibrium properties of an industry composed of a large number of price-taking firms producing a homogeneous output. According to this model, firm output is a function of input quantities and a random variable, which can be interpreted as a firm specific productivity shock. These productivity shocks, denoted by θ, are the source of heterogeneity in the model, and they are assumed to be independent across firms. At each time period, potential entrants may decide to enter by paying a sunk entry cost (c_e). Before entry, these firms have no knowledge of their productivity shock. Simultaneously, incumbents decide either to exit the industry or to stay in business. Finally, after productivity shocks are observed, firms choose their output level to maximize their expected discounted profits.

The initial productivity shock of potential entrants is assumed to be drawn from the same unknown cumulative distribution function. In subsequent years, productivity shocks follow a conditional cumulative distribution function, which is assumed to be strictly decreasing in the last productivity shock, θ_{t-1}. This Markov process implies that the higher the current productivity is, the better the distribution of future productivity shocks. Therefore, firms with a high current productivity are more likely to experience high productivity levels in the future. Once the current productivity shock has been realized, the firm updates the expectation about the future trajectory of its productivity.

The first result produced by the model is the existence of a reservation level for the productivity shock, θ^*, that defines the decision rule for the firm to exit or to remain in the market. Firms with $\theta_{t-1} < \theta^*$ will exit the market at period t and firms with $\theta_{t-1} \geq \theta^*$ will continue their activity. At time $t-1$, the model predicts that the productivity distribution of firms that will exit the market in future periods (exiting firms) should be stochastically dominated by the productivity distribution of firms that will remain in the market (continuing firms). As low productivity firms are more likely to exit, at each point in time, exiting firms will be concentrated among the least productive units.

The mechanism of market selection described above, makes the productivity distribution of continuing firms to improve over time. On the entry side of the market, this selection process gives raise to the second empirical implication of the model: any cohort of continuing firms at time t stochastically dominates the cohort of entering firms at t. This productivity difference occurs because older firms have had more occasions to experience the selection process, i.e. productivity levels higher than the critical level associated to the failure rule. As the cohort of continuing firms ages, more selection takes place and surviving members concentrate a higher proportion of high productivity units.

The third prediction comes from the three elements the model integrates: initial heterogeneity of productivity levels, selection based on productivity and persistence of productivity shocks. These three elements together imply that the productivity distributions of surviving and failing members of the same entry cohort should be different at the moment of entry. On one hand, firms with an initially low productivity would experience a low productivity in the future and they would have a greater probability of exit. On the other hand, firms with an initially high productivity are more likely to experience high future productivity that permit them to exceed the minimum level of productivity required for incumbent firms to survive. Therefore, at the moment of entry, surviving members of the entry cohort had a higher productivity than members of the cohort that would exit. Market selection, heterogeneity and persistence imply that exit patterns reflect initial productivity differences.

The fourth prediction refers to the role that sunk entry costs play in the failure process. Hopenhayn's (1992) model demonstrates that a higher sunk entry costs will lower the minimum level of productivity, θ^*, needed for incumbent firms to survive. Sunk entry costs are an important source of productivity heterogeneity as they influence the distribution of productivity levels across incumbent firms. For the group of continuing firms, with productivity level $\theta \geq \theta^*(c_e)$, the productivity distribution $F(\theta \mid \theta \geq \theta^*(c_e))$ is non-decreasing in the level of sunk costs. Similarly, the productivity distribution of exiting firms, $F(\theta \mid \theta < \theta^*(c_e))$, is a non-decreasing function of the sunk entry costs. In short, sunk entry costs are a key determinant of the distribution of efficiency levels across firms.

In the next section, we describe a testing procedure for examining productivity differentials between groups of entering, exiting and continuing firms. The objective of these comparisons is to test if productivity differences reflect market selection forces as predicted by Hopenhayn's (1992) model.

2.2. Testing procedure

The exam of the empirical implications we are interested in can be formulated as comparisons between productivity distributions corresponding to different groups of firms. In this section, we describe a procedure for testing differences between distribution functions, previously proposed by Delgado et al. (2002). The procedure relies on the concept of first order stochastic dominance and permits to establish a ranking for the compared distributions. Let F and G denote the cumulative distribution functions of productivity corresponding to two groups of firms to be compared, then (first order) stochastic dominance of F relative to G is defined by the following condition: $F(z) - G(z) \leq 0$, with strict inequality for some z, where $P(z \in \mathbb{R}) = 1$.

We have a random sample of size n for the group of firms that corresponds to the cumulative distribution function F. At the same time, we have a second random sample of size m, independent of the first one, that has been drawn from the distribution G. Then stochastic dominance of F relative to G requires that two statistical conditions have to be satisfied. First, the two compared distributions are not identical, i.e. the null hypothesis H_0: $F(z)-G(z)=0$ for all $z \in \mathbb{R}$ can be rejected. This two-sided test can be formulated as:

$$H_0 : \sup_{z \in \mathbb{R}} |F(z) - G(z)| = 0 \quad \text{vs.} \quad H_1 : \sup_{z \in \mathbb{R}} |F(z) - G(z)| \neq 0. \tag{1}$$

Second, the sign of the difference is as expected, i.e. the null hypothesis H_0: $F(z)-G(z) \leq 0$ for all $z \in \mathbb{R}$, with strict inequality for some z, cannot be rejected. This one-sided test can be formulated as:

$$H_0 : \sup_{z \in \mathbb{R}} \{F(z) - G(z)\} = 0 \quad \text{vs.} \quad H_1 : \sup_{z \in \mathbb{R}} \{F(z) - G(z)\} > 0. \tag{2}$$

The Kolmogorov–Smirnov test statistic for these two-sided test is, respectively,

$$\delta_N = \sqrt{\frac{n.m}{N}} \sup_{z \in \mathbb{R}} |T_N(z)|, \tag{3}$$

and

$$\eta_N = \sqrt{\frac{n.m}{N}} \sup_{z \in \mathbb{R}} T_N(z), \tag{4}$$

where $T_N(z)=F_n(z)-G_m(z)$ and $N=n+m$. F_n and G_m represent the empirical distribution functions for F and G, respectively. The limiting distributions of both test statistics, δ_N and η_N, are known under independence.[1]

For some comparisons, we provide also the bootstrap approximation of the p-values for both test statistics. Although there is no theoretical reason for justifying a better performance of the bootstrap test in our context, the fact that some sample sizes are small has lead us to provide also bootstrap p-values for checking the robustness of our conclusions.

Bootstrap p-values are computed in our context in two steps. In the first step, we resample by random sampling with replacement, under the null hypothesis, in the available sample of $n+m$ observations. In the second step, we compute the bootstrap analogs for δ_N and η_N, say δ_N^* and η_N^*, based on the resample. The bootstrap p-values can be obtained by Monte Carlo repeating the two previous steps B times and computing δ_N^{*b} and η_N^{*b}, $b=1,\ldots,B$. Then, bootstrap p-values are approximated by

$$P_B - \text{value}(\delta_N) = \frac{\sum_{b=1}^{B} 1\left(\delta_N^{*b} \geq \delta_N\right)}{B}, \tag{5}$$

and similarly for η_N. Under H_1, bootstrap p-values converge to zero.

[1] These test statistics were proposed by Smirnov (1939). Kolmogorov (1933) and Smirnov (1939) showed the limiting distributions of both test statistics under the assumption that all the observations are independent. For more details see Darling (1957).

Finally, we mention two estimation issues related to the application of one and two-sided Kolmogorov and Smirnov tests to our data set. First, the application of the previously defined testing procedure requires independence of the observations. Given that the data set is a panel of firms, observations on productivity for two consecutive years may correspond to a set of firms that are surveyed repeatedly and, therefore, cannot be considered either independent or stationary. As a consequence, we apply the testing procedure separately for each time period. Second, to test for stochastic dominance requires the use of empirical distributions of two compared groups of firms. Given the sampling properties of our data set (see Section 3), we use cumulative distribution functions for the groups of small firms and large firms. Direct application of the testing procedure to the whole population of firms is avoided since this would have required the estimation of a mixture of two distributions. Consequently, we compare groups of firms (continuing, exiting and entering firms) within a given size (i.e. continuing vs. exiting firms in the group of small firms) or between size categories (i.e. large continuing vs. small exiting firms).

2.3. Graphical comparisons

To further illustrate the comparisons between different groups of firms, we have graphed estimates of the distribution functions. Since the purpose is to produce graphical representations of the productivity differences between two groups of firms, we represent these differences for the whole population of firms (i.e. small and large firms) by means of the smooth sample distribution function. In this section, we present the procedures followed for the estimation of the cumulative distribution function and the relative distribution function.

First, we discuss the approach that has been followed for the estimation of a univariate cumulative distribution function, F, based on a sample $Z_1,\ldots Z_N$ of size N, where the sample is a combination composed of two random samples of small firms and large firms of sizes N_S and N_L ($N_S+N_L=N$), respectively. Let τ be a dummy variable, which is equal to 0 for small firms and equal to 1 for large firms, which Section 3 defines explicitly. The proposed estimate is based in the idea that, for a given group of firms in the population (say continuing firms) and time period, the cumulative distribution function for the whole population, $F(.)$, and the conditional cumulative distribution functions for the two size groups, $F(.|\tau)$, are related by the following expression:

$$F(.) = p \times F(.|\tau = 0) + (1 - p) \times F(.|\tau = 1), \tag{6}$$

where p represents the probability of being a small firm ($\tau=0$) in the considered group and time period. Time indexes are omitted to simplify the notation. Therefore, the cumulative distribution function of the whole population of firms is a mixture of the conditional cumulative distribution functions corresponding to the two size groups, where the parameter of the mixture is the probability of being a small firm in the corresponding population group. Then, the univariate cumulative distribution function for the whole population of firms can be estimated according to expression (6) as a weighted average of some estimators of the cumulative distribution functions corresponding to both size groups.

The weigthed kernel distribution estimate, $\hat{F}_h$, of a univariate cumulative distribution function for the whole population of firms, F, can be expressed as:

$$\hat{F}_h(z_0) = \hat{p} \int_{-\infty}^{z_0} \left(\frac{1}{N_S} \sum_{i=1}^{N} K\left(\frac{z - Z_i}{h} \right) \right) dz + (1 - \hat{p}) \int_{-\infty}^{z_0} \left(\frac{1}{N_L} \sum_{i=1}^{N} K\left(\frac{z - Z_i}{h} \right) \right) dz, \tag{7}$$

where h and $K(.)$ denote, respectively, the bandwidth and the kernel function; and $\hat{p}$ represents the estimated probability of being a small firm in the considered group and time period. Expression (7) can be rewritten as:

$$\hat{F}(z_0) = \int_{-\infty}^{z_0} \left(\sum_{i=1}^{N} \omega_i K\left(\frac{z - Z_i}{h} \right) \right) dz, \tag{8}$$

where weights ω_i are given by

$$\omega_i = (1 - \tau_i) \frac{\hat{p}}{N_S} + \tau_i \frac{1 - \hat{p}}{N_L}, \tag{9}$$

satisfying that $\sum_{i=1}^{N} \omega_i = 1$, where τ_i is a dummy variable, which is equal to 0 if firm i is small and equal to 1 otherwise.[2]

3. Productivity measurement

The data set of this study is drawn from the Encuesta sobre Estrategias Empresariales (ESEE), an annual survey on a representative sample of Spanish manufacturing firms, which contains 15,087 observations corresponding to an average number of 1886 firms over the period 1990–1997.

A first characteristic of the data set is that firms participating in the survey were chosen according to a selective sampling scheme. Firms participate with different probabilities depending on their size category. In the base year, all firms with more than 200 workers (large firms) were asked to participate. The rate of participation in this size category achieved nearly the 70% of the population. Additionally, approximately the 4% of the group of firms that employed less than 200 workers (small firms) were chosen according to a random sampling scheme. The same selection criterion was applied to the 18 industries considered in this study. Therefore, the coverage of the data set is different depending on the size category, although it can be considered that, separately for both size groups, the data set is a random sample that permits the estimation of the distribution of any firm characteristic in Spanish manufacturing.

A second characteristic refers to the process of entry and exit of firms and the way this process has been recorded in the data set over the period 1990–1997. On one hand, newly

[2] The estimates of the proportions of small and large firms in manufacturing population are 0.97 and 0.03, respectively. Analogously, these proportions are: 0.97 and 0.03, in the group of continuing firms; 0.98 and 0.02, in the group of entering firms; and 0.99 and 0.01, in the group of exiting firms.

created firms have been added annually to the original sample. This entering firms have been selected according to the same sampling criteria as were selected the original firms in the sample. On the other hand, exiting firms in the sample have been recorded each year and, given the sampling properties of the data set, they can be considered a random sample representing the cohort of firms leaving the market in the population. This information allows the classification of firms according to their trajectories over any interval of time. In particular, for two consecutive years, t and $t+1$, three groups of firms can be defined: (i) firms entering the market at $t+1$; (ii) exiting firms that leave the market at $t+1$; and (iii) continuing firms staying in business both years t and $t+1$.

In the rest of this section, we present an index for measuring firm productivity, which follows the framework developed by Aw et al. (2001). Total factor productivity at the firm level is measured by a multilateral productivity index that is an extension of the multilateral total factor productivity index proposed by Caves et al. (1982). Our index uses as the reference point the average firm of the firm's size group and then chain-links the reference points to preserve transitiveness between firms of different size groups within the same industry. This extension takes into account the characteristics of the data set and, in particular, the fact that sampling proportions are different for small and large firms. A similar extension of the index can be found in Good et al. (1996). One of the advantages of this kind of measures is that the parameters of the production function are not required to compute productivity.

The expression of total factor productivity index at time t, for the firm f, which belongs to the size group τ and to the industry s is:

$$\begin{aligned}\ln\lambda_{ft} &= \ln Y_{ft} - \overline{\ln Y}_{\tau s} - \frac{1}{2}\sum_{k=1}^{K}\left(W_{ft}^{k} + \overline{W}_{\tau s}^{k}\right)\left(\ln X_{ft}^{k} - \overline{\ln X}_{\tau s}^{k}\right) + \overline{\ln Y}_{\tau s} - \overline{\ln Y}_{s} \\ &\quad - \frac{1}{2}\sum_{k=1}^{K}\left(\overline{W}_{\tau s}^{k} + \overline{W}_{s}^{k}\right)\left(\overline{\ln X}_{\tau s}^{k} - \overline{\ln X}_{s}^{k}\right),\end{aligned}$$

where Y_{ft}, is the output of firm f at time t, W_{ft}^{k} and X_{ft}^{k} are, respectively, the cost share and the quantity of input k corresponding to firm f and time t. Firms are classified in two size groups ($\tau=1$, 2) and 18 industries corresponding to the NACE-CLIO R25 classification ($s=1$,. . ., 18). This index measures the proportional difference of total factor productivity for firm f at time t relative to a given reference firm. The reference firm varies across industries and, for a given industry s, it is defined as a firm such that: (i) its output is equal to the geometric mean of firm's output quantities in industry s over the entire period; (ii) the quantities of inputs are equal to the geometric means of firm's input quantities in industry s over the entire period; and (iii) the cost shares of inputs are equal to the arithmetic mean of firm's cost shares in industry s over the entire period. Notice that the reference firm varies across industries and, therefore, when observations of different industries are pooled productivity differences across industries are removed.

The first three terms of the right hand side of the index, compares firm's productivity with a different reference firm depending on its size group. Hence, comparisons between observations corresponding to the same size group are transitive. The last three terms measure productivity differences between the reference firm of any size group and a common reference firm, which is the average firm over the entire sample of small and large firms in industry s.

Finally, the total factor productivity index for each firm is computed using information drawn from the Encuesta sobre Estrategias Empresariales (ESEE). The variables have been defined as follows. The output Y_{ft} is measured by annual gross production of goods and services expressed in real terms using individual price index for each firm drawn from the ESEE. The estimation considers three inputs, X_{ft}^{k} ($k=1,\ldots, 3$), labor, materials and the stock of capital. Labor input is measured by the number of effective hours of work per year, which is equal to normal yearly hours plus overtime yearly hours minus non-working yearly hours. Material inputs are measured by the cost of intermediate inputs and they include raw materials purchases, energy and fuel costs and other services paid by the firm. It is expressed in real terms using individual price indexes of intermediate inputs for each firm drawn from the ESEE. The stock of capital is calculated using the perpetual inventory formula: $k_t^{*}=I_t+k_{t-1}^{*}(1-d_t)(P_t)/(P_{t-1})$ where I_t represents investment in equipment, d_t stands for depreciation rates and P_t corresponds to price indexes for equipment published by the Instituto Nacional de Estadistica. With respect to input cost shares, W_{ft}^{k}, each share is the fraction of the cost of the input on total input costs. Total cost is the sum of labor cost, intermediate input cost and the cost of capital. The cost of labor is measured by the sum of wages, social security contributions and other labor costs paid by the firm. The user cost of capital is measured by the cost of long-term external debt of the firm plus depreciation rates (d_t) minus the variation of the price index for capital goods.

4. Empirical results

In this section, we present a set of empirical results on productivity differentials between groups of entering, exiting and continuing firms. Testing procedures are applied to the data set to see if the observed productivity differences reflect market selection forces as indicated by predictions summarized in Section 2.1. Results are presented as follows. Section 4.1 investigates persistence in firm productivity levels. Section 4.2 compares productivity level distributions corresponding to entering and exiting firms relative to continuing firms. In Section 4.3, we examine if the productivity distributions of surviving and failing members of the same entry cohort are different at the moment of entry, as suggested by Hopenhayn's model. Section 4.4 compares productivity growth for entrants and incumbent firms to test for convergence in productivity levels following entry. Finally, Section 4.5 provides an exploratory empirical assessment of the relationship between productivity levels and sunk costs. In particular, we test if the productivity distribution of incumbent firms is non-decreasing in the level of sunk entry costs.

4.1. Persistence in firm productivity levels

We begin the analysis of the relationship between productivity and firm turnover verifying the accuracy of the underlying assumption of Hopenhayn's model that productivity differences across firms tend to persist over time. In particular, the model assumes that the productivity distribution is stochastically increasing in past productivity levels. This means that the probability for a firm to have higher productivity at time t, increases with firm's productivity at time $t-1$. We apply kernel techniques for estimating the conditional distribution function of productivity level at time t given the productivity level at time $t-1$.

For the whole population of large and small firms, the (weighted) kernel estimate of the conditional cumulative distribution function of productivity at time t given productivity at time $t-1$ is defined as:

$$\hat{F}_{t|t-1}(z_t|z_{t-1}) = \frac{\hat{F}_{t,t-1}(z_t,z_{t-1})}{\hat{F}_{t-1}(z_{t-1})},$$

where $\hat{F}_{t-1}(z_{t-1})$ is the (weighted) kernel univariate distribution estimate of productivity at time $t-1$ obtained according to expressions (8) and (9); and $\hat{F}_{t,t-1}(z_t, z_{t-1})$ denotes the (weighted) kernel bivariate distribution estimate of productivity at time t and at time $t-1$, which is estimated as:

$$\hat{F}_{t,t-1}(z) = \int_{-\infty}^{z_t^0} \int_{-\infty}^{z_{t-1}^0} \left(\sum_{i=1}^{N} \omega_i K_2\left(\frac{z - Z_i}{h}\right)\right) dz_t dz_{t-1},$$

where z is equal to the vector (z_t, z_{t-1}) $K_2$is the bivariate kernel function obtained as the product of identical gaussian univariate kernel functions, i.e. $K_2(z)=K(z_t)K(z_{t-1})$; weights ω_i are given by expression (9); and the bandwidth parameter h has been calculated using the expression given by Silverman (1986) for bandwidth selection in k-variate kernel estimation. The conditional cumulative distribution functions have been estimated for the period 1991–1997.

Fig. 1 depicts the kernel estimate of the conditional cumulative distribution function of productivity at time t, given that productivity at $t-1$ equals z, $\hat{F}_{t|t-1}$ $(z_t|z_{t-1}=z)$, for some arbitrarily chosen values of z. The level 0.00 corresponds to firms with a productivity that equals the level of an average firm of the same industry and size group, 0.26 corresponds to firms with a productivity that is 26% higher than the average and similarly for the rest of values. The estimates correspond to year 1996 and are similar for the rest of years. The graphs show evidence of persistence with the productivity ranking in t reproducing the same pattern as in $t-1$. Overall, these findings confirm the assumption of positive serial correlation in firm productivity level and coincide with empirical evidence previously found in the literature. In analyzing persistence, Baily et al. (1992) and Bartelsman and Dhrymes (1998) report transition matrices of individual producers in the relative productivity distribution. Year-to-year transition probabilities show a high degree of persistence, and similar results are also obtained over larger periods of 5 and 10 years. Our non-parametric estimation of the distribution of current productivity conditional on past productivity also confirms the existence of large and persistent differences across firms.

4.2. Entry, exit and productivity differences

In this section, we examine the magnitude of productivity differentials between continuing, entering and exiting firms. The data set we analyze includes information on

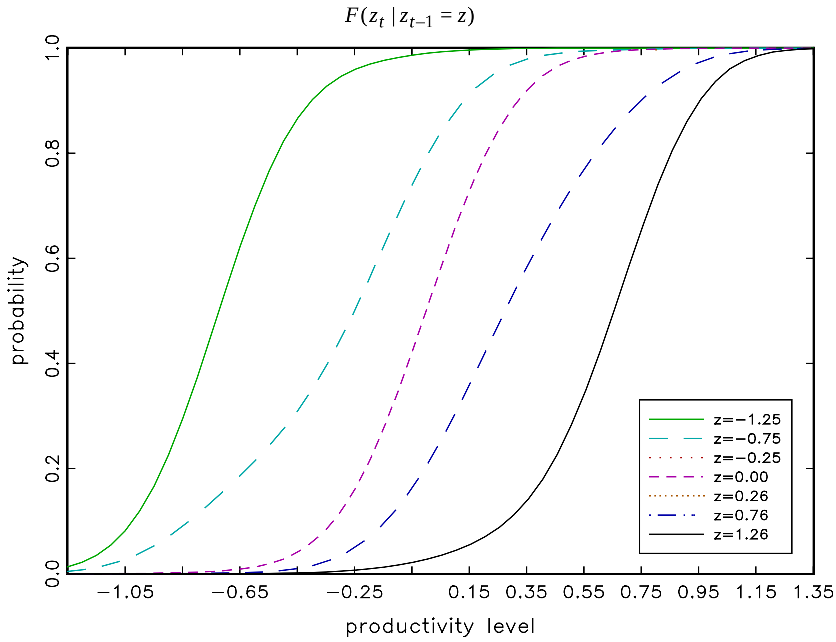

Fig. 1. Persistence in firm level productivity (weighted Kernel conditional distribution of productivity in year 1996 given productivity in year 1995).

Spanish manufacturing firms over the period 1990–1997. We are able to define seven cohorts of firms entering the market from year 1990 to 1997.[3] Since these cohorts of firms enter the market from the beginning of year t, they permit comparisons that are based on its productivity in year t. The cohorts of exiting firms correspond to individual producers in period t that exit the market in period $t+1$. Seven cohorts of exiting firms can be identified from year 1991 to 1997. These cohorts are defined by the exit year $t+1$ but the comparison is based in its productivity level in year t. As members of the 1990 cohort were not recorded by the ESEE, this cohort is excluded from the analysis.

Fig. 2 illustrates the differences between the productivity of exiting firm and continuing firms for years 1991–1997. The graphs correspond to the weighted kernel estimates of the cumulative distributions of both small and large firms. The position of the cumulative distributions of continuing firms is to the right of the distributions of exiting firms. This position indicates that for any quantile of the distribution, continuing firms have higher productivity than exiting firms. Therefore, the position of both distributions suggests that the productivity distribution of exiting firms is stochastically dominated by the distribution of continuing firms. At the median of the distribution, continuing firms have an average 12.8% higher productivity than exiting firms over the period.

[3] Entering firms for year 1995 were not surveyed by the ESEE.

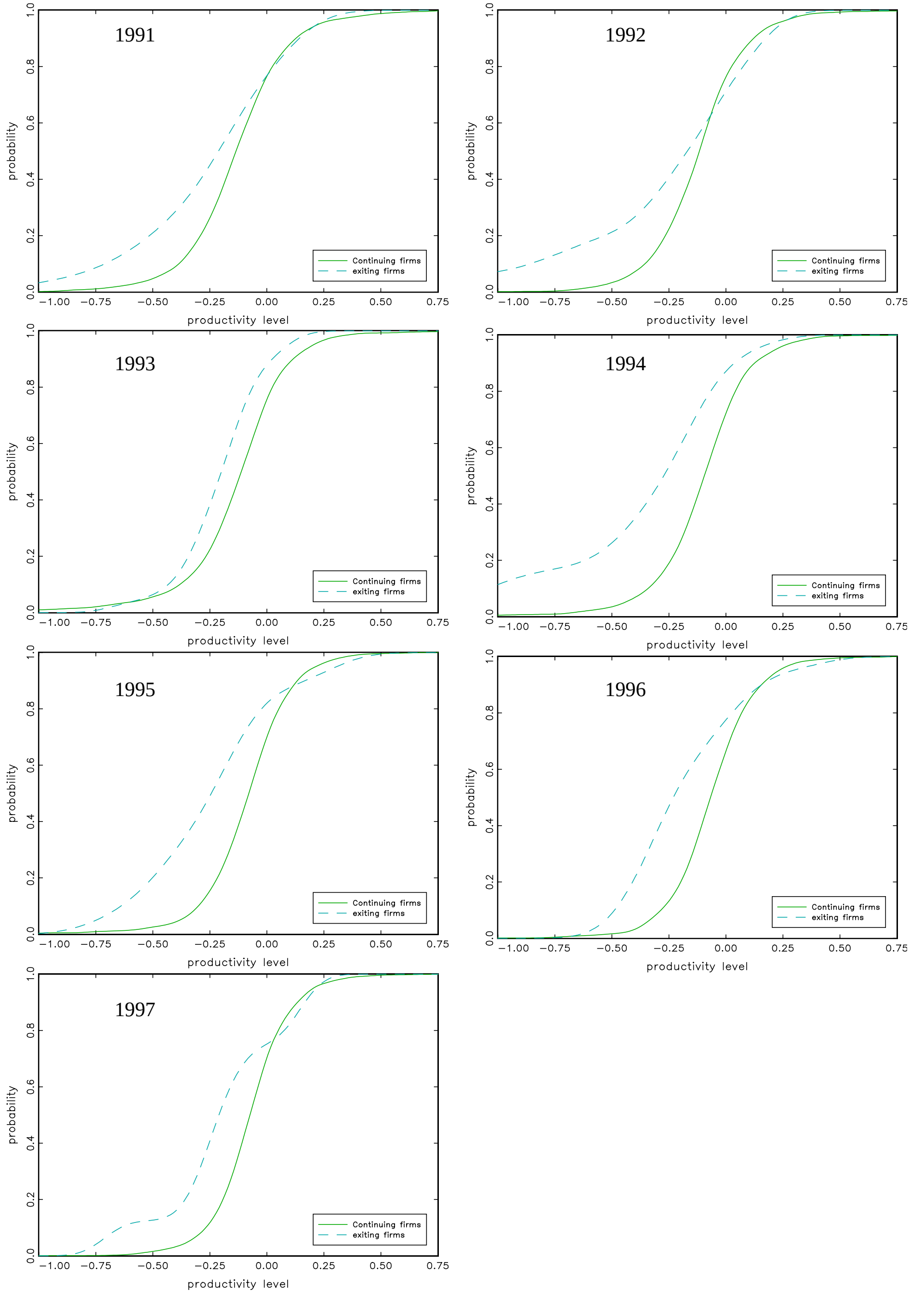

Fig. 2. Productivity differences of exiting firms vs. continuing firms (smooth sample distribution functions).

Analogously, Fig. 3 shows the estimates of the cumulative distribution of productivity for entering firms and continuing firms over the period 1990–1997. The positions of the cumulative distributions indicate that continuing firm productivity distributions stochastically dominate the productivity distributions of entering firms. Consequently, continuing firms have higher productivity than entering firms. The median productivity of continuing firms is 9% higher than entering firms over the period.

Given the assessed differences, now we apply the testing procedure described in Section 2 to formally test if productivity differences between continuing, entering and exiting firms are as expected. Although we are interested in the comparison of distributions for the whole population of firms (small and large), the characteristics of the data set constraint us to apply the testing procedure separately for the group of small and large firms, i.e. comparing the cumulative distribution functions in the groups of small and large firms. Therefore, differences between continuing and entering firms, on one hand, and differences between continuing and exiting firms, on the other hand, are tested separately in the populations of small and large firms. For each time period, one- and two-sided tests described in Section 2.2 are applied to compare

$$F_t(.|\tau = \tau_0) \quad \text{vs.} \quad G_t(.|\tau = 0), \quad t = 1990,..,1997 \text{ and } \tau_0 = 0,1,$$

where τ is a dummy variable equal to 0 for small firms and equal to 1 for large firms; F_t denotes the productivity level ($\ln \lambda_{ft}$) distribution for continuing firms at time t and G_t denotes the distribution for the target group, i.e. entering or exiting firms. As the groups of entering and exiting firms are exclusively composed of small firms, we perform two comparisons. The first one is the comparison: $F_t(.|\tau=0)$ vs. $G_t(.|\tau=0)$, where $F_t(.|\tau=0)$ refers to the productivity distribution of small continuing firms and $G_t(.|\tau=0)$ corresponds to the groups of entering or exiting firms. The second comparison is: $F_t(.|\tau=1)$ vs. $G_t(.|\tau=0)$, where $F_t(.|\tau=1)$ refers to the productivity distribution of large continuing firms and $G_t(.|\tau=0)$ is for entering or exiting firms.

Table 1 summarizes the hypotheses test statistics on differences between the productivity distributions of continuing and exiting firms. As indicated, tests are applied separately to the groups of small and large continuing firms over the period 1991–1997. For all comparison, Table 1 provides the one and two-sided Kolmogorov–Smirnov test statistics and their corresponding asymptotic p-values. Bootstrap p-values are not reported given that they are very similar to their asymptotic counterparts and consequently do not change the results. First, the null hypothesis of equality between the distributions of small continuing firms and exiting firms can be rejected for all years. Furthermore, the hypothesis that the sign of the difference is favorable to incumbent firms cannot be rejected at any reasonable significance level. Second, for the groups of large continuing firms vs. exiting firm, the equality of both distributions can be rejected at the 0.01 level for all years. Furthermore, the hypothesis that incumbent firms have greater productivity than exiting firms cannot be rejected. Therefore, the productivity distribution of surviving firms stochastically dominates the distribution of failing firms. This result confirms that failing firms are concentrated among the least productive units.

Table 2 shows the results of testing differences between the productivity distributions of continuing and entering firms for the period 1990–1997. Table 2 gives test statistics and p-values separately for the comparisons of entering firms relative to small continuing firms,

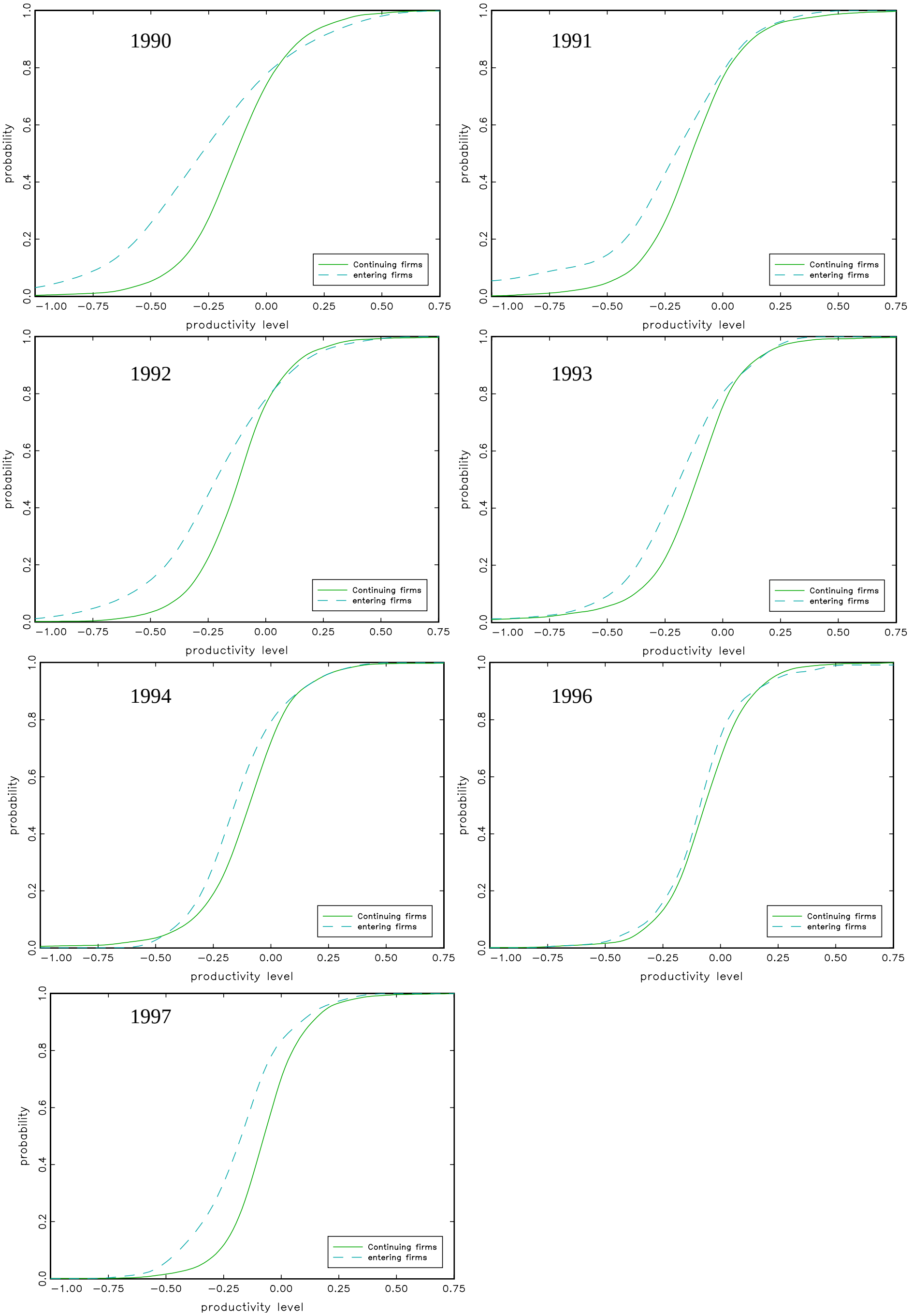

Fig. 3. Productivity differences of entering firms vs. continuing firms (smooth sample distribution functions).

Table 1
Productivity differences between continuing and exiting firms

Year	Small continuing firms vs. exiting firms						Large continuing firms vs. exiting firms					
	Number of observations[a]		Equality of distributions		Differences favorable to continuing firms		Number of observations[a]		Equality of distributions		Differences favorable to continuing firms	
	n_S	m	Statistic	p-value[b]	Statistic	p-value[b]	n_L	m	Statistic	p-value[b]	Statistic	p-value[b]
1991	1136	33	1.319	0.062	0.354	0.779	727	33	2.044	0.000	0.264	0.870
1992	1123	47	1.488	0.024	0.575	0.516	637	47	2.402	0.000	0.482	0.628
1993	1100	27	1.560	0.015	0.140	0.962	528	27	2.642	0.000	0.067	0.991
1994	1099	35	2.253	0.000	0.011	1.000	585	35	3.419	0.000	0.000	1.000
1995	983	17	1.803	0.003	0.335	0.799	502	17	2.709	0.000	0.291	0.844
1996	1002	21	2.038	0.000	0.175	0.940	507	21	2.675	0.000	0.170	0.944
1997	1038	8	1.278	0.076	0.433	0.687	512	8	1.825	0.003	0.214	0.913

Hypotheses test statistics.
[a] n_S, n_L and m denote sample sizes for small continuing firms, large continuing firms and exiting firms, respectively.
[b] Based on the limiting distribution. p-values based on the bootstrap approximation (10,000 replications) are not reported.

on one hand, and entering firms vs. large continuing firms, on the other hand. For small continuing firms, the null hypothesis of equality between both distributions can always be rejected at significance levels less than 0.05, except for 1996 with a significance level that is less than 0.10. The null hypothesis that the sign of the difference is as expected, i.e. the productivity distribution of continuing firms is greater than the productivity of entering firms, cannot be rejected at any reasonable significance level in the period. For the groups of large incumbent firms and small entrants, the null hypothesis that both productivity distributions are identical is always rejected at any significance level, and test statistics do

Table 2
Productivity differences between continuing and entering firms

Year	Small continuing firms vs. entrants						Large continuing firms vs. entrants					
	Number of observations[a]		Equality of distributions		Differences favorable to continuing firms		Number of observations[a]		Equality of distributions		Differences favorable to continuing firms	
	n_S	m	Statistic	p-value[b]	Statistic	p-value[b]	n_L	m	Statistic	p-value[b]	Statistic	p-value[b]
1990	841	39	1.921	0.001	0.259	0.874	474	39	2.586	0.000	0.246	0.886
1991	1136	38	1.378	0.045	0.197	0.925	727	38	2.029	0.001	0.068	0.991
1992	1123	68	1.964	0.001	0.222	0.906	637	68	3.256	0.000	0.144	0.959
1993	1100	82	1.670	0.008	0.231	0.898	528	82	3.173	0.000	0.072	0.990
1994	1099	36	1.343	0.054	0.193	0.928	585	36	2.464	0.000	0.149	0.956
1996	1002	121	1.239	0.093	0.250	0.882	507	121	3.773	0.000	0.229	0.900
1997	1038	83	2.553	0.000	0.017	0.999	512	83	5.031	0.000	0.017	0.999

Hypotheses test statistics.
[a] n_S, n_L and m denote sample sizes for small continuing firms, large continuing firms and entering firms, respectively.
[b] Based on the limiting distribution. p-values based on the bootstrap approximation (10,000 replications) are not reported.

not permit to reject the null hypothesis that continuing firms stochastically dominate entering firms. Overall, on the entry side of the market, we observe productivity differences with respect to incumbent firms that are consistent with the mechanism of market selection.

The evidence already presented indicates that the productivity of continuing firms is above the productivity level of entering and exiting firms. All these patterns are consistent with market selection forces as predicted by Hopenhayn's (1992) model. Although models of industry dynamics do not contain predictions concerning comparisons between entering and exiting firms, we conclude this section examining productivity differences between both groups of firms. Table 3 presents the hypothesis test statistics of productivity differentials between contemporaneous entering and exiting firms. The general pattern indicates that entering firms have greater productivity than exiting firms but these differences are rather small and they are not statistically significant over the whole period. The null hypothesis of equality between both distributions can only be rejected in years 1994 and 1996 at usual significance levels. The stochastic dominance of entering firms relative to exiting firms cannot be rejected in both years; and for the rest of years, differences between both distributions are favorable to entering firms but they cannot be considered statistically significant. Are these results consistent with the literature reporting micro patterns of productivity? We can point to some comparable evidence that have quantified how firm entry, growth and exit contribute to aggregate productivity growth. As Haltiwanger (2000) and Foster et al. (2001) indicate, the productivity gap between entering and exiting firms, at the point of simultaneous entry and exit, is rather small and it increases over a longer horizon. When the productivity gap between exiting and entering firms is quantified evaluating the productivity of exiting firms at t and the productivity of surviving entrants after a period of time, say $t+1$, selection and learning effects increase the gap substantially. The general pattern indicates that the gap is greater the greater is the distance between t and $t+1$. Since we are comparing productivity levels that correspond to a period of simultaneous entry and exit, our result of no significant differences in the gap between entering and exiting firms is consistent with patterns described in the previous literature.

Table 3
Contemporaneous productivity level differences between entering and exiting firms

	Number of observations[a]		Equality of distributions			Differences favorable to entering firms		
	n	m	Statistic	p-value[b]	p-value[c]	Statistic	p-value[b]	p-value[c]
1991	38	33	0.560	0.913	0.816	0.560	0.534	0.447
1992	68	47	0.952	0.325	0.257	0.952	0.163	0.130
1993	82	27	0.708	0.697	0.591	0.379	0.751	0.684
1994	36	35	1.337	0.056	0.042	0.007	0.799	0.757
1996	121	21	1.841	0.002	0.001	0.221	0.907	0.855
1997	83	8	0.582	0.887	0.791	0.545	0.552	0.469

Hypotheses test statistics.
[a] n and m denote the sample sizes for entering firms and exiting firms, respectively.
[b] Based on the limiting distribution.
[c] Based on the bootstrap approximation (10,000 replications).

4.3. Firm exit decisions and market selection

Market selection based on productivity in combination with initial heterogeneity and persistence in productivity shocks implies that exit patterns should reflect initial productivity differences among firms of the same entry cohort. Therefore, the productivity distribution of surviving and failing members of the same entry cohort should be different at the moment of entry. In particular, the initial productivity distribution of surviving members of any entry cohort should stochastically dominate the initial productivity distribution of failing members of the same entry cohort.

In this section, we compare and test for differences between the productivity distributions of surviving and failing entrants. We define five cohorts of entering firms corresponding to $t = 1990, \ldots, 1994$. These cohorts of firms can be observed in our data set up to year 1997. Therefore, for each entry cohort, the group of failing entrants corresponds to firms leaving the market before 1997. In order to increase somewhat the number of observations available for the analysis, we define the cohort of entering firms at t as the group of those firms that are ≤ 3 years old at t. For this purpose, we compute the age of the firms in year t as equal to ($t -$ birth year $+ 1$).

The inferences we are able to draw from the comparisons of the productivity distributions of surviving and failing entrants are affected by two types of biases. On one hand, the group of surviving entrants includes those entering firms exiting the market after 1997. Therefore, under the null hypothesis that the productivity distribution of failing entrants is stochastically dominated by the distribution of surviving entrants, the selectivity bias reduces the true difference between both distributions. On the other hand, the sample of failing members of the entry cohort excludes those firms entering the market but failing before year t. Again, under the same null hypothesis, the observed distribution of failing entrants has greater productivity than the true distribution. Therefore, both biases operate against the null hypothesis we want to test, i.e. that the productivity distribution of failing entrants is stochastically dominated by the distribution of surviving entrants.

Fig. 4 shows the kernel estimates of the initial productivity distribution functions for surviving and failing entrants. The cumulative distribution of surviving firms is always to the right of the distribution of failing firms. These comparisons indicate that the productivity distribution of failing entrants is stochastically dominated by the distribution of surviving entrants, as predicted by models of industry dynamics.

Table 4 exhibits the test statistics and p-values. The reported results point out to two conclusions. First, the null hypothesis that both distributions are identical can be rejected at significance levels that range between 0.03 and 0.20. Second, the null hypothesis that the productivity of surviving entrants is greater than the productivity of failing entrants can never be rejected at reasonable significance levels. These findings are broadly consistent with selection forces. In a given entry cohort, exiting firms are those individual producers with the lowest initial productivity level.

4.4. Post-entry productivity growth of entrants vs. incumbents

The evidence presented in Section 4.2 indicates that the productivity of entering firms is lower than the productivity of the average incumbent. This result coincides with individual

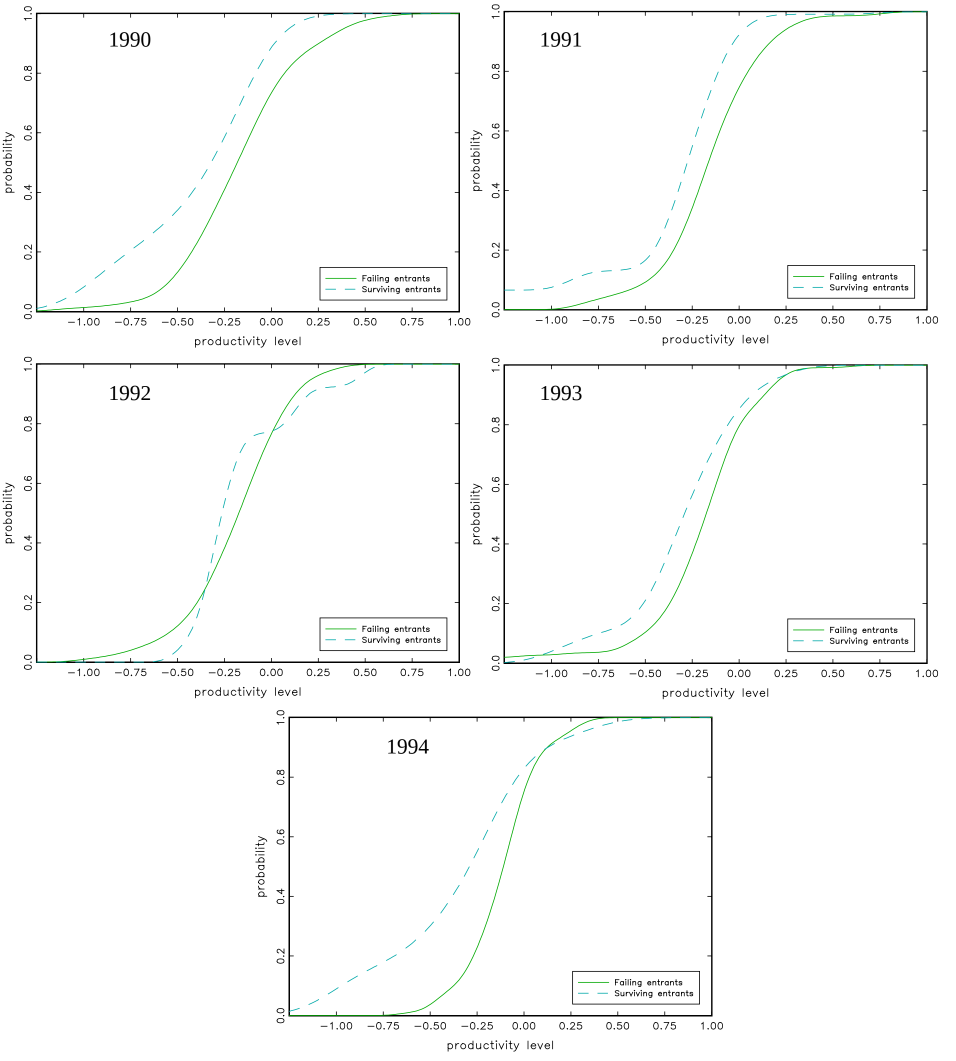

Fig. 4. Productivity differences of surviving entrants vs. failing entrants (smooth sample distribution functions).

productivity patterns documented by the empirical literature that uses longitudinal micro-level data sets. This literature suggests that surviving members of entry cohorts improve their productivity until they approach the average level of the productivity of incumbent firms (Bartelsman and Doms, 2000; Foster et al., 2001). This pattern has been interpreted by Aw et al. (2001) as consistent with selection and learning forces. First, the performance of surviving members of the entry cohort incorporates selection effects through the failure rule that induces firms with a low level of productivity to exit the market. Second, the pattern is also consistent with learning effects playing a role through investment, exploitation of scale economies, etc.

Table 4
Productivity differences between surviving and failing members of various entry cohorts

	Number of observations[a]		Equality of distributions			Differences favorable to surviving entrants		
	n	m	Statistic	p-value[b]	p-value[c]	Statistic	p-value[b]	p-value[c]
1990	78	14	1.193	0.116	0.080	0.044	0.996	0.977
1991	72	15	1.096	0.181	0.136	0.000	1.000	1.000
1992	79	13	1.005	0.264	0.200	0.423	0.699	0.621
1993	119	18	1.047	0.223	0.173	0.120	0.972	0.944
1994	89	12	1.358	0.050	0.032	0.234	0.896	0.838

Hypotheses test statistics.
[a] n and m denote the sample sizes for surviving entrants and failing entrants, respectively.
[b] Based on the limiting distribution.
[c] Based on the bootsrap approximation (10,000 bootstrap replications).

In this section, we compare the post-entry productivity growth of entering firms with the productivity growth of incumbent firms. We are interested in testing the existence of convergence between the productivity levels of surviving entrants and incumbent firms. The hypothesis of convergence between both productivity distributions can be examined by testing if the productivity growth distribution of surviving entrants stochastically dominates the distribution of continuing firms. Productivity growth for firm f between years t and $t+k$ is given by $\ln\lambda_{ft} - \ln\lambda_{ft+k}$. Three cohorts of surviving entrants are defined: firms entering in 1990, 1991 and 1992 and surviving until 1997. The three entering cohorts that we consider permit to observe a period of post-entry performance of 5 or more years. One and two-sided tests are applied to the comparison: $F_t(.|\tau_0)$ vs. $G_t(.)$, $t=1990$, 1991, 1992 and $\tau_0=0$, 1; where $F_t(.|\tau_0)$ denotes the productivity growth distribution of small and large continuing firms between t and 1997, and $G_t(.)$ denotes the productivity growth distribution of cohorts of firms entering at t and surviving until 1997.

Fig. 5 depicts the estimates of the cumulative distributions of productivity growth for surviving entrants and for continuing firms. Results vary across cohorts of entrants. For the 1990 and 1992 entry cohorts, both distributions cross the distribution of continuing firms at the 0.5 quantile. Therefore, both distributions cannot be ranked according to the criterion of (first order) stochastic dominance. The productivity distribution of the 1991 entry cohort is to the left of the distribution of continuing firms, indicating that the productivity growth of the latter is higher than the productivity growth of surviving entrants.

We have formally tested for differences between the productivity growth distribution of continuing and surviving entering firms. Test statistics and p-values are given in Table 5 separately for small and large continuing firms. There are remarkable differences in results corresponding to small and large continuing firms. For large continuing firms, test statistics indicate that differences in the distribution of productivity growth between surviving entering firms and (large) continuing firms are significant. However, the null hypothesis that differences are favorable to entering firms can be rejected. The comparison between small continuing firms and surviving entrants provides evidence more favorable to the hypothesis of convergence in productivity levels between both groups of firms. After entry, firms have greater

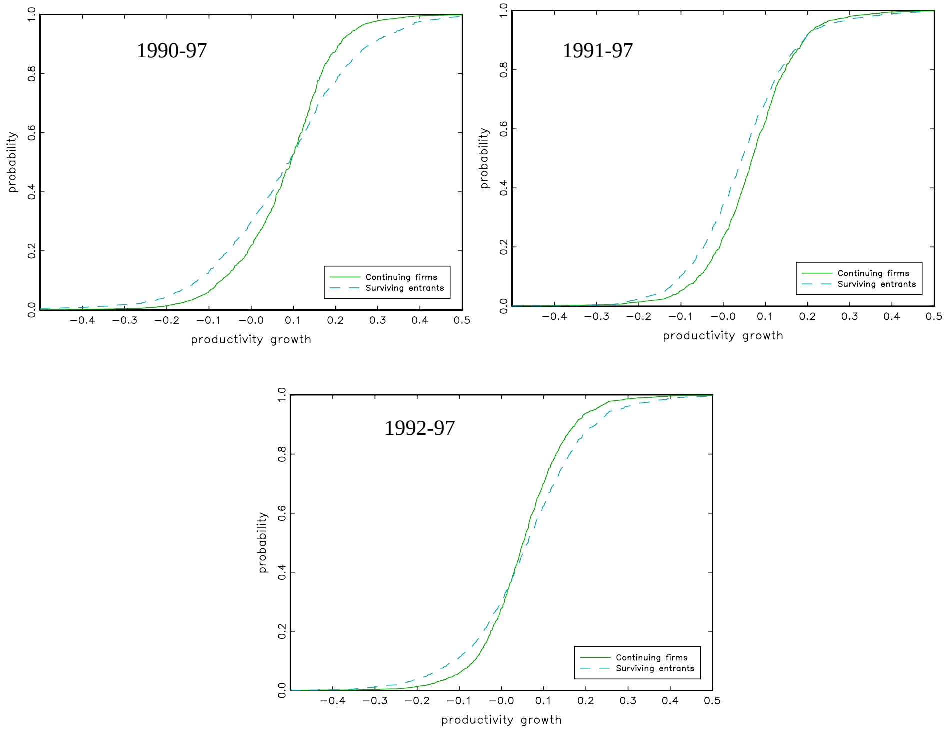

Fig. 5. Post-entry productivity growth differences of surviving entrants vs. continuing incumbent firms (smooth sample distribution functions).

productivity growth than incumbents, although differences are not significant (with the exception of the 1990 entry cohort). Overall, we find some evidence that the productivity of surviving entrants grows more rapidly and tend to reach the level of incumbent firms. This pattern reduces to the comparison between surviving entrants and small continuing firms.

According to Hopenhayn's (1992) model, only selection forces play a role in the evolution of productivity after entry, given the assumption that firm productivity follows an exogenous Markov process. The fact that for some groups of firms the initial difference in the productivity level between surviving entrants and incumbent firms is reduced after entry, does not exclude the existence as well of learning effects.

4.5. Productivity differences and sunk costs

The existence of a failure boundary in the distribution of firm productivity levels is a basic prediction of models of industry dynamics (Hopenhayn, 1992). This boundary, denoted in Section 2.1 by θ^*, is the lowest level of productivity that will enable the firm to have, over future periods, positive discounted expected profits. Firms with productivity levels less than θ^* will exit the market and firms with productivity levels greater than θ^* will remain in the market. According to this prediction, the failure boundary θ^* must

Table 5
Productivity growth differences between surviving entering firms and continuing firms

Period	Surviving entrants vs. small continuing firms								Surviving entrants vs. large continuing firms							
	Number of observations[a]		Equality of distributions			Differences favorable to surviving entrants			Number of observations[a]		Equality of distributions			Differences favorable to surviving entrants		
	n_S	m	Statistic	p-value[b]	p-value[c]	Statistic	p-value[b]	p-value[c]	n_L	m	Statistic	p-value[b]	p-value[c]	Statistic	p-value[b]	p-value[c]
1990–97	568	27	1.577	0.014	0.012	0.745	0.329	0.305	292	27	1.617	0.011	0.008	1.103	0.088	0.070
1991–97	685	25	0.860	0.450	0.394	0.860	0.228	0.198	437	25	1.215	0.104	0.087	1.215	0.052	0.044
1992–97	697	37	0.928	0.356	0.319	0.840	0.244	0.223	429	37	1.528	0.019	0.015	1.528	0.009	0.007

Hypotheses test statistics.

[a] n_S, n_L and m denote sample sizes for small continuing firms, large continuing firms and surviving entrants, respectively.

[b] Based on the limiting distribution.

[c] Based on the bootstrap approximation (10,000 replications).

define a separating line between the productivity distribution function of surviving firms and the productivity distribution of exiting firms. However, as Fig. 2 indicates, both distributions overlap over a large range of productivity levels suggesting that the failure boundary separating the productivity distributions of continuing and exiting firms is missing. In this section, we explore this empirical fact and we offer an explanation based on the existence of sunk costs as an important source of heterogeneity of productivity distributions for different groups of firms.

Hopenhayn's (1992) model demonstrates that a higher sunk entry cost (c_e) will lower the minimum level of productivity, $\theta*$, needed for incumbent firms to survive. High entry costs reduce the intensity of market selection, as they provide a barrier to the entry of new firms. The insulation of incumbent firms from the effects of market selection will result in a distribution of surviving firms with a higher proportion of low productivity units. In short, sunk entry costs are a key determinant of the distribution of productivity levels across incumbent firms in the industry. Fig. 6 considers the case of two failure boundaries: $\theta*(c_e^H)$ is the boundary for industries with high sunk entry cost and $\theta*(c_e^L)$ is the boundary for industries with low sunk entry costs. The model predicts that the critical level of productivity corresponding to high sunk costs is to the left of the productivity level corresponding to low sunk costs: $\theta*(c_e^H)<\theta*(c_e^L)$. Each boundary defines a separating line for the productivity distributions of both surviving and exiting firms. As illustrated for the group of industries with low sunk entry costs, the productivity distribution of surviving firms is to the right of the boundary defined by $\theta*(c_e^L)$ and the distribution of exiting firms is to the left of the boundary. The same pattern can be predicted for the group of firms with high sunk entry costs.

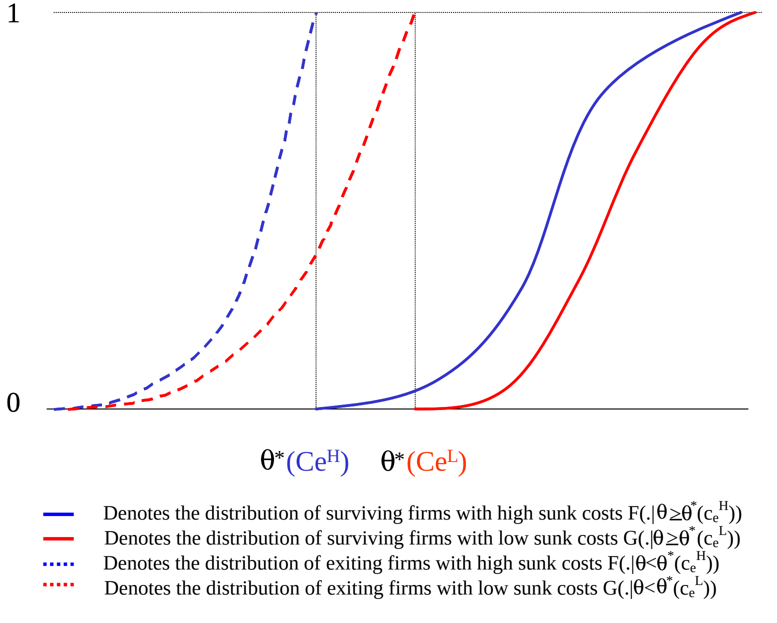

Fig. 6. Productivity distributions and sunk entry costs.

The productivity distributions of surviving and exiting firms that are drawn in Fig. 2 are the result of the aggregation of many productivity distributions, corresponding to groups of firms with different sunk entry costs and various boundary failures. The aggregation over these groups of firms may explain the fact that the distributions of surviving and failing firms overlap over a large range of productivity levels, covering the effects of the failure boundaries. Models of firm dynamics that seek to account for heterogeneity across individual productivity point to the heterogeneity of sunk entry costs as one of the idiosyncratic factors determining the distribution of efficiency levels across incumbent firms.

Although productivity thresholds, θ^*, are not directly observable, we can take advantage of the hypothesis that they are related to the magnitude of sunk entry cost so as to test for some of their implications. In particular, the ranking of productivity distributions for the group of surviving firms has to be related to the magnitude of sunk entry costs. Fig. 6 illustrates this prediction considering two productivity thresholds: $\theta^*(c_e^H)$ for the group of firms with high sunk entry cost and $\theta^*(c_e^L)$ for the group of firms with low sunk costs. The distribution $F(\theta\,|\,\theta \geq \theta^*(c_e^H))$ denotes the productivity distribution of surviving firms with high sunk entry costs and $F(\theta\,|\,\theta \geq \theta^*(c_e^L))$ denotes the distribution of surviving firms with low sunk entry costs. The prediction of Hopenhayn model is $F(\theta\,|\,\theta \geq \theta^*(c_e^H)) \geq F(\theta|\theta \geq \theta^*(c_e^L))$. This condition states that the productivity distribution of surviving firms with low sunk entry costs stochastically dominates the distribution of firms with high entry costs. More generally, the productivity distribution of surviving firms $F(\theta\,|\,\theta \geq \theta^*(c_e))$ is non-decreasing with sunk entry costs. A similar prediction can be stated for the group of exiting firms with distribution function $F(\theta\,|\,\theta \geq \theta^*(c_e))$.

Our objective is to provide an exploratory empirical assessment of the relationship between productivity levels and sunk costs. We adopt the methodology outlined in Sutton (1991) and Kessides (1990) to measure sunk costs. This methodology has been recently used by Ghosal (2002). The index proposed by Sutton (1991) is given by $\sigma/S = \mu^* K/S$ where σ is the minimum level of sunk cost an entrant must incur and S denotes the size of the market. The index assumes that sunk costs are proportional to the cost of establishing a new plant of minimum efficient scale. This minimum scale is the output of the median firm relative to the industry output (μ). The index uses as K/S the aggregate capital-sales ratio for the industry, under the assumption that this ratio equals the capital-sales ratio of the median firm. We calculate the index, which is an industry-level measure, for two-digit NACE-CLIO R44 manufacturing industries using data drawn from the Encuesta Industrial. We label the index σ/S as SUNK(Sutton).

The second measure of sunk costs is based on the theory of contestable markets. Kessides (1990) draws on this literature to emphasize that the share of sunk capital outlays committed by the potential entrant is a function of the capital's durability, degree and type of specificity and mobility. While these characteristics are unobservable, Kessides (1990) provides meaningful proxies. The first proxy, RENT, denotes the extent to which capital can be rented by the firm. The ESEE provides information on this characteristic. RENT is equal to one for firms partially renting their capital requirements, and zero otherwise. Sunk cost will be low for firms using capital that can be readily leased. The second proxy is a measure of the intensity of the second-hand market for the capital employed by the firm. USED denotes for each firm the ratio of sales of used capital equipment over the capital stock. This information is provided by the ESEE. The higher USED is the lower sunk costs are. The third proxy is based on the conjecture that the share of sunk outlays will be low for firms using capital that depreciates very rapidly. DEPR, provided by the ESEE, is a ratio that measures depreciation charges during the year against the fixed assets of the firms. A higher DEPR implies a lower sunk cost.

As Hopenhayn's (1992) model makes a long-run equilibrium prediction across industries or markets that differ in the level of sunk costs, we classify firms according to an industry level measure of sunk costs. We follow the strategy used by Ghosal (2002) to segment the sample in two groups of industries with high and low sunk costs. First, we use the SUNK(Sutton) industry-level measure and, for the rest of variables, we define additional sunk cost measures by the median value of USED and DEPR for each two-digit NACE-CLIO R44 manufacturing industry. For the variable RENT, the industry-level considered corresponds to the proportion of firms partially renting their capital requirements in the industry. Second, all firms within an industry with the level of SUNK(Sutton) $<$50th percentile were classified in the group of low sunk cost and high if SUNK(Sutton) $\geq$50th percentile. Similarly, sunk costs are low if RENT, USED or DEPR $\geq$50th percentile and high if RENT, USED or DEPR $<$50th percentile. Third, we combine the three characteristics suggested by Kessides (1990) to define additional industry groupings. The group of low sunk costs firms corresponds to those firms within industries where the following condition holds: RENT and USED and DEPR $\geq$50th percentile. The group of high sunk costs is defined by the condition: RENT and USED and DEPR $<$50th percentile. According to this classification, the group of low sunk costs corresponds to firms within industries with a high intensity of rental capital, a frequent use of second hand capital and a high rate of depreciation. We expect that this combination of characteristics may produce a stronger separation between the sub-samples of low and high sunk cost firms. We label as SUNK(Kessides) the variable that combines the three characteristics. All variables have been estimated for 3 years: 1991, 1994 and 1997.

Table 6 reports the hypotheses test statistics of comparing the productivity distributions of continuing firms that fall in the two groups of high and low sunk costs. We compare $F_t(.|\tau=\tau_0)$ vs. $G_t(.|\tau=\tau_0)$ for three time periods, $t=1991$, 1994, 1997, and two size groups, $\tau_0=0$, 1, using the one and two-sided tests described in Section 2.2, where F_t and G_t denote the productivity level distribution for high and low sunk cost firms in year t, respectively. The statistics summarized in Table 6 permit us to test if the productivity distribution of low sunk cost firms stochastically dominates the productivity distribution of high sunk costs. Tests are performed separately for the five measures of sunk costs outlined above.

The variable SUNK(Sutton) does not produce productivity differences between the group of high and low sunk cost firms. As a general pattern, the null hypothesis that the sign of the difference is favorable to the group of low sunk cost firms relative to high sunk cost firms can be rejected at standard significance levels, excluding the group of large firms in 1991.

Table 6
Productivity differences between low and high sunk costs firms

Year/Sunk-cost measure	Small low cost firms vs. small high cost firms						Large low costs firms vs. large high cost firms					
	Number of observations[a]		Equality of distributions		Differences favorable to low cost firms		Number of observations[a]		Equality of distributions		Differences favorable to low cost firms	
	n_H	n_L	Statistic	p-value[b]	Statistic	p-value[b]	m_H	m_L	Statistic	p-value[b]	Statistic	p-value[b]
1991												
SUNK(Sutton)	557	491	1.800	0.003	1.800	0.002	263	348	0.841	0.480	0.685	0.391
DEPR	451	577	0.829	0.498	0.262	0.871	318	293	0.786	0.568	0.565	0.528
USED	350	698	0.974	0.331	0.120	0.972	212	399	1.812	0.003	0.033	0.998
RENT	255	793	0.812	0.524	0.064	0.992	255	356	0.652	0.789	0.555	0.540
SUNK(Kessides)	159	445	0.795	0.552	0.506	0.534	197	213	1.183	0.122	0.246	0.886
1994												
SUNK(Sutton)	513	469	1.585	0.013	1.585	0.007	216	283	1.064	0.208	1.064	0.104
DEPR	447	535	1.006	0.263	0.058	0.993	262	237	1.440	0.032	0.089	0.984
USED	340	642	0.929	0.354	0.174	0.941	168	331	0.716	0.685	0.716	0.359
RENT	228	754	1.386	0.043	1.386	0.021	214	285	0.574	0.896	0.574	0.517
SUNK(Kessides)	166	399	1.420	0.036	0.431	0.689	161	164	1.499	0.022	0.218	0.909
1997												
SUNK(Sutton)	495	482	1.119	0.164	1.119	0.082	184	253	1.071	0.202	1.071	0.101
DEPR	455	522	0.846	0.472	0.361	0.771	234	203	2.251	0.000	0.198	0.925
USED	326	651	0.735	0.652	0.735	0.339	146	291	0.807	0.533	0.807	0.272
RENT	242	735	0.852	0.462	0.852	0.234	188	249	0.703	0.706	0.703	0.372
SUNK(Kessides)	170	400	0.532	0.940	0.357	0.705	138	140	1.703	0.006	0.413	0.712

Hypotheses test statistics.

[a] n_H and n_L denote sample sizes for small high and low sunk costs firms, respectively; similarly m_H and m_L for large firms.

[b] Based on the limiting distribution. p-values based on the bootstrap approximation (10,000 replications) are not reported.

The rest of the variables – DEPR, USED, RENT, and SUNK(Kessides) – produce evidence more favorable to the hypothesis that sunk costs are an important source of heterogeneity of firm productivity distributions. First, for the group of large firms, the null hypothesis that the productivity of low sunk cost firms is greater than the productivity of high sunk cost firms cannot be rejected as expected. This result applies

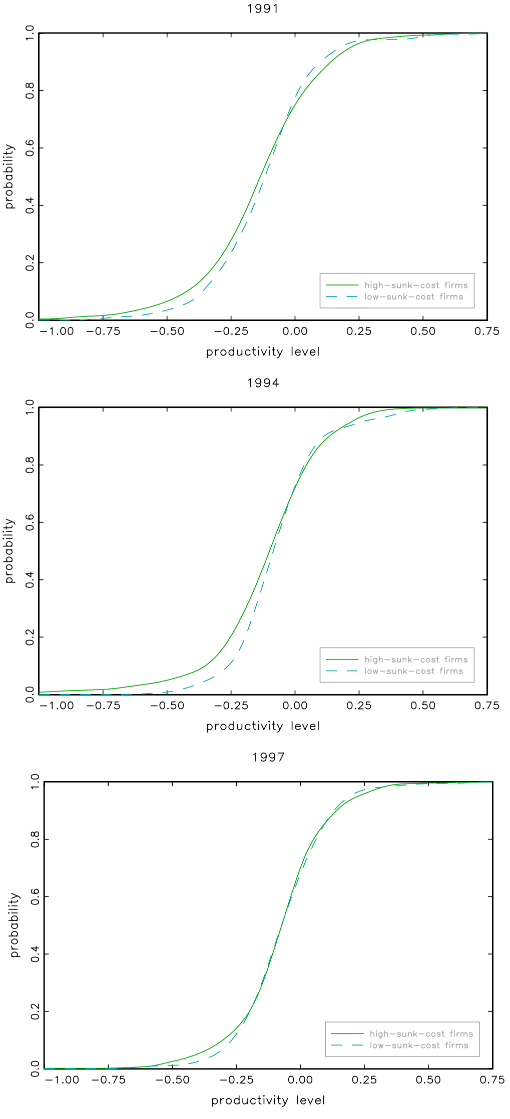

Fig. 7. Productivity differences of continuing firms with low sunk costs vs. high sunk costs. (Smooth sample distribution functions).

to the four sunk measures defined following the approach of Kessides (1990). With respect to the hypothesis of equality between the distributions of low and high sunk costs, the variable SUNK(Kessides) permits rejection of the null hypothesis for the three years considered. With respect to the rest of variables (DEPR, USED, RENT), the null hypothesis of equality cannot be rejected in all time periods. Second, for the group of small firms, the null hypothesis that the sign of the difference is as expected, i.e. low sunk cost firms have greater productivity than high sunk cost firms, cannot be rejected at the usual significance levels. However, according to two-sided tests, the magnitude of the productivity differences tends to be not significant for the group of small firms. Overall, the results indicate that for the group of large firms productivity differences are significant and favorable to low sunk cost firms. Within the group of small firms, the magnitude of productivity differences is rather small (not significant) but yet in the expected direction.

Finally, Fig. 7 illustrates the differences between the group of high and low cost firms. The variable used to separate the sample in two groups of firms is SUNK(Kessides), which is an industry level measure of sunk costs based on a combination of DEPR, USED and RENT. In 1991 and 1994, the lower tail of the distribution of low sunk cost firms is to the right of the distributions of high sunk cost firms. Test statistics in Table 6 confirm that the productivity distribution of low sunk costs stochastically dominates the distribution of high sunk costs. In 1997, both distributions overlap each other and cannot be ranked according to the criterion of stochastic dominance. With the exception of 1997, the results indicate that sunk costs are an important source of heterogeneity across individual firm productivity. These productivity differences are associated to the height of sunk costs as suggested by models of industry dynamics. In particular, firms operating in industries with low sunk costs have higher productivity than high sunk cost firms. This evidence is obtained when we use the methodology outlined in Kessides (1990) to quantify sunk costs.

5. Conclusions

In this paper, we use a micro panel data set of Spanish manufacturing firms to examine productivity differences between groups of entering, exiting and continuing firms. To account for observed heterogeneous productivity, we rely on models of industry dynamics, in particular Hopenhayn's (1992) model, to see if systematic differences across firms reflect market selection forces. Our empirical strategy is to compare the entire productivity distribution of groups of firms. These distributions are ranked using non-parametric procedures. Our empirical study reveals a set of productivity differences that can be summarized as follows.

First, there are persistent differences in productivity across firms. As in papers reporting transition matrices of individual producers in the relative distribution (Baily et al., 1992; Bartelsman and Dhrymes, 1998), our non-parametric estimation of the distribution function of current productivity conditional on past productivity also confirms the existence of large and persistent differences across firm productivity.

Second, the evidence presented indicates that the productivity of incumbent firms is greater than the productivity level of entering and exiting firms. The productivity distribution of the former group stochastically dominates the distribution of entering and exiting firms. At the median of the distribution, the productivity of surviving firms is 9% higher than the productivity of entering firms and 12.8% higher than the productivity of exiting firms. These patterns are consistent to market selection as predicted by Hopenhayn's (1992) model.

Third, market selection and persistence in productivity shocks imply that firm exit patterns should reflect initial productivity level differences among firms of the same entry cohort. We confirm that the initial productivity distribution of surviving members of any entry cohort dominates he distribution of failing members of the same entry cohort. Aw et al. (2001) and Foster et al. (2001) also report that the less productive plants from the entering cohort are those that exit the market. This pattern is broadly consistent with selection effects.

Fourth, we compare the post-entry productivity growth of entering firms with the productivity growth of incumbent firms. The evidence we find indicates that the productivity of surviving entrants grows more rapidly and tends to reach the level of incumbent firms. Although this pattern is not always highly significant, it is clearly present when the comparison is between surviving entrants and small continuing firms. Similar results have been reported by Bartelsman and Doms (2000), Aw et al. (2001) and Foster et al. (2001), who have interpreted this result as consistent with selection and learning forces.

Finally, we confirm that the level of sunk costs is associated to productivity differences across cohorts of continuing firms. In particular, firms with high sunk costs have, on average, lower productivity than low sunk costs firms. This evidence is consistent with Hopenhayn's (1992) model, which predicts that an increase in the level of sunk costs lowers the minimum productivity level for incumbent firms to survive. As a consequence, these firms are subject to less market selection. Similar results have been obtained by the empirical study conducted by Aw et al. (2002) on Taiwan and Korea. The analysis reveals a set of productivity differences between firms of both countries that are consistent with institutional characteristics making sunk entry costs higher in Korea than in Taiwan. Similarly, in our study, firms are classified according to an industry-level measure of sunk costs and we confirm that the productivity distribution of continuing firms is non-decreasing with sunk costs.

Acknowledgements

We are grateful to Miguel A. Delgado, Norbert Janz, Jordi Jaumandreu, François Laisney, Silvio Rendón, Peter Schmidt, three anonymous referees and to Lars-Hendrik Röller and Stephen Martin for their useful comments and suggestions. Earlier versions of this paper benefited from presentations in Mannheim (ZEW Summer Workshop), Oviedo (Workshop on Efficiency and Productivity), Zaragoza (Encuentro de Economía Aplicada), Lausanne (EARIE Conference), Madrid (Jornadas Economía Industrial) and Barcelona (Workshop on Demography of firms and industries). This research has been partially funded by projects SEC2000-0268 and SEJ2004-02525.

References

Audretsch, David, 1995. Innovation and Industry Evolution. MIT Press, Cambridge.

Aw, Bee Yan, Chung, Sukkyun, Roberts, Mark J., 2000. Productivity and turnover in the export market: micro evidence from the Republic of Korea and Taiwan. World Bank Economic Review 14, 65–90.

Aw, Bee Yan, Chen, Xiaomin, Roberts, Mark J., 2001. Firm level evidence on productivity differentials and turnover in Taiwanese manufacturing. Journal of Development Economics 66, 51–86.

Aw, Bee Yan, Chung, Sukkyun, Roberts, Mark J., 2002. Productivity, output and failure: a comparison of Taiwanese and Korean manufactures. NBER Working Paper 8766.

Baily, Martin N., Hulten, Charles, Campbell, David, 1992. Productivity dynamics in manufacturing plants. Brooking Papers: Microeconomics, 187–225.

Baldwin, John, 1993. The Dynamics of Industrial Competition. Cambridge University Press, Cambridge.

Bartelsman, Eric J., Dhrymes, Phoebus, 1998. Productivity dynamics: U.S. manufacturing plants 1972–1986. Journal of Productivity Analysis 9 (1), 5–34.

Bartelsman, Eric J., Doms, Mark, 2000. Understanding productivity: lessons from longitudinal microdata. Journal of Economic Literature 38, 569–594.

Caves, Richard E., 1998. Industrial organization and new findings on the turnover and mobility of firms. Journal of Economic Literature XXXVI, 1947–1982.

Caves, Douglas W., Christensen, Laurits R., Diewert, Erwin, 1982. Multilateral comparisons of output, input, and productivity using superlative index numbers. The Economic Journal 92, 73–86.

Darling, D.A., 1957. The Kolmogorov–Smirnov, Cramér–Von Mises tests. Annals of Mathematical Statistics 28, 823–838.

Delgado, Miguel A., Fariñas, José C., Ruano, Sonia, 2002. Firm productivity and export markets: a non-parametric approach. Journal of International Economics 57, 397–422.

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

Foster, Lucia, Haltiwanger, John, Krizan, C.J., 2001. Aggregate productivity growth: lessons from microeconomic evidence. In: Dean, E., Harper, M., Hulten, C. (Eds.), New Contributions to Productivity Analysis. University of Chicago Press.

Ghosal, Vivek, 2002. Impact of uncertainty and sunk costs on firm survival and industry dynamics. Mimeo.

Good, David, Nadiri, M. Ishaq, Sickless, Robin, 1996. Index number and factor demand approaches to the estimation of productivity. NBER Working Paper 5790.

Griliches, Zvi, Regev, Hain, 1995. Firm productivity in Israeli industry, 1979–1988. Journal of Econometrics 65, 175–203.

Haltiwanger, John, 2000. Aggregate growth: what have we learned from microeconomic evidence, Economics Department Working Paper no. 267 OECD.

Hopenhayn, Hugo, 1992. Entry, exit, and firm dynamics in long run equilibrium. Econometrica 60, 1127–1150 (September).

Jovanovic, Boyan, 1982. Selection and the evolution of industry. Econometrica 50, 649–670.

Kessides, Ioannis, 1990. Market concentration, contestability and sunk costs. Review of Economics and Statistics LXXII, 614–622.

Kolmogorov, A.N., 1933. Sulla determinazione empirica di une legge di distribuzione. Giorn. DellIstit. Degli Att. 4, 83–91.

Olley, Steven, Pakes, Ariel, 1996. The dynamics of productivity in telecommunications equipment industry. Econometrica 64 (6), 1263–1297.

Roberts, Mark J., Tybout, James R., 1996. Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity and Market Structure. Oxford University Press.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman and Hall.

Smirnov, N.V., 1939. On the estimation of the discrepancy between empirical curves of distribution for two independent samples. Moscow University Mathematics Bulletin 2 (2), 3–14.

Sutton, John, 1991. Sunk Costs and Market Structure. The MIT Press.

Tybout, James R., 1996. Heterogeneity and productivity growth: assessing the evidence. In: Roberts, Mark J., Tybout, James R. (Eds.), Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity and Market Structure. Oxford University Press.