FIRM-LEVEL PRODUCTIVITY DYNAMICS IN SPANISH MANUFACTURING: A NON-PARAMETRIC APPROACH

Sonia Ruano*

Universidad Carlos III de Madrid
PIE-FEP

May, 1999

Abstract

This paper examines total factor productivity differences among continuing, entering and exiting firms, using a sample of Spanish manufacturing firms over the period 1991-1997 drawn from the *Encuesta Sobre Estrategias Empresariales (ESEE)*. Evidence favours the predictions of theoretical models of market dynamics in the sense that the productivity distribution for surviving firms is clearly better than that for failing firms. This study also concludes that productivity distribution of entry cohorts is stochastically dominated by that of incumbents but tends to improve over time, in part due to the exit of less productive members. Finally, this paper examines the degree of persistence in productivity differentials across firms. It finds signs of persistent heterogeneity, because the higher past productivity is, the better current productivity distribution is.

Keywords: Total Factor Productivity, firm turnover, stochastic dominance and manufacturing.

* I would like to thank Jose Carlos Fariñas, Rebeca de Juan, Ana Martin Marcos and Daniel Miles for useful comments and suggestions.

Fundacion Empresa Pública, Pza. Marques de Salamanca, 8; 28006-Madrid. Tel: +34 91 577 79 29.

e-mail: sruano@funep.es

1. Introduction

Recent theoretical models of market evolution relate the decision to enter or exit with productivity, in the sense that entry and exit costs act as a selection mechanism that pressure inefficient firms to exit. The model credited to Hopenhayn (1992) explains these decisions in a competitive framework in which output is a function of a productivity shock uncorrelated across firms. The model assumes that the initial distribution of productivity levels is unknown for potential entrants which, in every period, choose to enter by paying a sunk entry cost. Simultaneously, the installed firms decide to stay or exit the market.

Industry equilibrium will involve simultaneous flows of entering and exiting firms: less efficient firms learn about their relative inefficiency and choose to exit while more efficient firms stay in business. The decision rule in equilibrium for an individual firm consists of comparing its productivity with a critical level that is related to the entry cost. The higher the entry costs, the higher the level of profits needed to make entry profitable will be, but the minimum productivity level needed for incumbents to survive will be lower. Therefore, the relationship between entry cost and firm turnover will be negative.

Empirical literature focussing on the relationship between productivity and firm performance includes Baily, Hulten and Campbell (1992), who study productivity in U.S. manufacturing plants, and Aw, Chen and Roberts (1997), who study productivity in Taiwanese manufacturing firms. Common findings in both studies are the following. Turnover patterns reflect differences in total factor productivity, but entering and exiting firms are themselves heterogeneous groups. Average productivity of new firms increases more rapidly and converges to the productivity level of incumbents. It is because, on average, less productive members exit the market. Baily, Hulten and Campbell (1992) also find that there is strong persistence in relative productivity across plants.

Additionally, these studies measure the effects of the evolution of incumbent's productivity and firm turnover on the evolution of aggregate productivity. They find two opposing results. Baily, Hulten and Campbell (1992) point out that the growth of output shares in high productivity plants was a major factor in the productivity growth of U.S. manufacturing. But the contribution of entry and exit was not great. On the contrary, Aw, Chen and Roberts (1997) find that in Taiwanese manufacturing industries, firm turnover has been an important source of productivity growth. Different levels of entry and exit costs have justified these two opposing results.

Our contribution to this literature can be summarised as follows. First, the empirical analysis of productivity dynamics is based on the entire distribution of productivity for continuing, entering and exiting firms. This approach is consistent with the wide heterogeneity in firm level performance observed in data. Second, the compared cumulative distribution functions for different groups of firms are ranked using the concept of stochastic dominance. Finally, the significance of productivity differentials among productivity distributions is tested using non-parametric procedures.

This paper explores and tests productivity differences among continuing, entering and exiting firms using a sample of Spanish manufacturing firms over the period 1990-1997. The data set comes from the *Encuesta sobre Estrategias Empresariales (ESEE)*. It concludes that firm level productivity distribution improved in Spanish manufacturing in the period studied and, at the same time, productivity distribution tended to concentrate. Evidence drawn from comparison among productivity distributions of continuing, entering and exiting firms favours the predictions of theoretical models of market dynamics in the sense that the productivity distribution for surviving firms is clearly better than that for failing firms. This study concludes also that productivity distribution of entry cohorts is stochastically dominated by that of incumbents but tends to improve over time, in part due to the exit of less productive members. Finally, this paper finds signs of persistent heterogeneity in productivity levels across firms.

The rest of the paper is organised as follows. Section two explains the theoretical issues behind multilateral productivity indexes. Section three describes the data set used in this study. Section 4 studies the evolution of firm level productivity distribution over the period 1990-1997. Section 5 presents the productivity differences among entering, exiting and continuing firms. Section 6 analyses the persistence in productivity differentials across firms. Section 7 concludes.

2. Productivity measurement

Total factor productivity at the firm level is measured by a multilateral productivity index. The advantage of this kind of measure is that the parameters of the production function are not required to compute productivity. The index we use is based on the *translog multilateral productivity index* proposed by Caves, Christensen and Diewert (1982). The expression of this index at time *t* for firm *f* is:

$$\ln l^{*}_{ft} = \ln y_{ft} - \overline{\ln y} - \frac{1}{2}\sum_{n=1}^{N}(v_{ntf} + \overline{v}_{n})(\ln x_{nft} - \overline{\ln x_{n}}) \qquad [1]$$

where y_{tf} is the output level of the firm *f* at time *t*; v_{ntf} and x_{ntf} are, respectively, the cost share and the quantity of input *n* corresponding to firm *f* at time *t*; *N* denotes the total number of inputs[1] and the bars denote the average value over the relevant variable (e.g., $\overline{v}_{n}$ indicates the average of the cost share of input *n* across all observations (firm-year) in the sample.

This index can be interpreted as a bilateral comparison between every firm and a hypothetical representative firm, which has as output $\overline{\ln y}$, as cost share vector ($\overline{v}_{1}$, $\overline{v}_{2}$, ..., $\overline{v}_{N}$) and as input vector ($\overline{\ln x_{1}}$, $\overline{\ln x_{2}}$,..., $\overline{\ln x_{N}}$), that is, the average firm in terms of output, inputs and cost shares over the period. The main properties of this index are: *circularity* (the use of various indexes provides transitive comparisons among any subset of observations) and *superlativity* (the translog production function from which the expression of the index is derived can be interpreted as a second order approximation to the real function).

The weakness of this productivity measure, when applied for comparisons based on panel data, is the necessity to recalculate all the indexes each time the sample is extended in its temporal dimension. Good (1985) solves this problem by constructing different reference firms for each cross-section and then chaining the bilateral indexes for the reference firms corresponding to adjacent years. The transitivity of this index is based on the fact that the sequence of comparisons is not ambiguous, given that all comparisons are indirectly referred to the reference firm of the base year (in this study, 1990). The expression of this *chained multilateral total factor productivity index* is the following

$$\ln l^*_{ft} = \ln y_{ft} - \overline{\ln y_t} - \frac{1}{2}\sum_{n=1}^{N}(v_{ntf} + \overline{v_{tn}})(\ln x_{nft} - \overline{\ln x_{nt}}) + \sum_{s=2}^{t}\left[\overline{\ln y_s} - \overline{\ln y_{s-1}} - \frac{1}{2}\sum_{n=1}^{N}(\overline{v_s} + \overline{v_{s-1}})(\overline{\ln x_{ns}} - \overline{\ln x_{ns-1}})\right] \quad [2]$$

where the average firm in terms of output ($\overline{\ln y_t}$), inputs ($\overline{\ln x_{1t}}$, $\overline{\ln x_{2t}}$,..., $\overline{\ln x_{Nt}}$) and cost shares ($\overline{v_{1t}}$, $\overline{v_{2t}}$, ..., $\overline{v_{Nt}}$) varies over time. Total factor productivity change between t_0 and t_1 of firm f is given by the expression[2]

$$\Delta \ln l^*_{ft_1} = \ln l^*_{ft_1} - \ln l^*_{ft_0}$$

Finally, the reference firm is also considered to vary across industries[3]. As such, this index provides a measure of the proportional difference of total factor productivity for firm f at time t relative to the average firm of the industry-group this firm belongs to in the base year.

3. Data

The data set considered in this study is drawn from the *Encuesta Sobre Estrategias Empresariales (ESEE)*, an annual survey referred to a representative sample of Spanish manufacturing firms over the period 1990-1996. In 1990, the base year, 2188 firms participated in the survey according to a selective sampling scheme, in which the probability that a firm is asked to participate depends on its size category. All firms with more than 200 workers (large firms) were asked to participate, and the rate of participation reached 67.6 percent. Firms that employed between 10 and 200 workers

[1]For the construction of the index, we consider three inputs: labour, capital and materials. For more details see the appendix.

[2] For more details about total factor productivity indexes see Good, Nadiri and Sickless (1996).

(small firms) were chosen according to a random sampling scheme, and the sampling fraction was 3.9 percent of the number of firms in the population[4]. In subsequent years the survey maintained its initial characteristics, minimising attrition and annually incorporating newly created firms selected with the same sampling criteria as in the base year.

TABLE 1 ABOUT HERE

The estimations in the following sections will take into account the different percentages of coverage in both size-groups. Table 1(*a*) reports the number of firms in the manufacturing population, distinguishing by size-group. The number of entries and exits in the base year are not provided by the *ESEE*, although they can be estimated from the data of the following years. The average rates of entering and exiting over the period 1991-1996 are, respectively, 6 percent and 3.2 percent in small firms; and 0.3 percent and 0.8 percent in large firms (see Table A1 in appendix). The total number of entering and exiting firms presented in Table 1(*a*) have been estimated using these average rates.

4. The analysis of productivity growth

4.1. The approach

The empirical analysis of productivity evolution is based on the study of the behaviour of the entire distribution of productivity, which is characterised by the cumulative distribution function. Let $F_t(z)$ denote the cumulative distribution function of firm-level total factor productivity (z) at time t; then the evolution of productivity at the firm level over the period t_0 - t_0+k is fully described by the sequence of functions: $F_{t_0}(z)$, $F_{t_0+1}(z)$, ..., $F_{t_0+k}(z)$.

The estimation of cumulative distribution functions requires considering the different percentages of coverage of the sample for small and large firms. The selective

[3]Firms have been grouped in eighteen industries corresponding to NACE-CLIO R-25 classification.
[4] See Table 1(*b*)

sampling scheme in the *ESEE* implies that the cumulative distribution function of the whole population of firms cannot be directly estimated. However, we have a random sample of firms for every size-group, which permits us to estimate the two conditional cumulative distribution functions separately. Let τ denote a dummy variable taking the value 1 when the firm employs more than 200 workers and 0 otherwise. The conditional cumulative distribution functions of the two size groups at time *t* are: $F_{St}(z) = F_t(z|t = 0)$ and $F_{Lt}(z) = F_t(z|t = 1)$ for small and large firms, respectively. The conditional cumulative distribution function of the whole population at time *t* is related to the conditional cumulative distribution functions of the two size groups in the following way

$$F_t(z) = \quad P_t(t = 0) \times F_{St}(z) + P_t(t = 1) \times F_{Lt}(z) \qquad [4]$$

where $P_t(.)$ represents the conditional probability of being a small or a large firm at time *t*. Thus, the expression of the estimator of the cumulative distribution function of productivity (*z*) at time *t* is given by the following expression

$$\hat{F}_t(z) = \quad \hat{p}_t \times \hat{F}_{St}(z) + (1 - \hat{p}_t) \times \hat{F}_{Lt}(z) \qquad [5]$$

where $\hat{F}$ represents the estimator of F, and $\hat{p}_t$ represents the estimated probability of being a small firm at time *t*. We will call this method the *two-stage procedure of estimation*

The estimations of cumulative distribution functions have been obtained from gaussian kernel density estimators, such as:

$$\hat{F}_{St}(z) = \int_{-\infty}^{z} \hat{f}_{St}(x)dx \text{ and } \hat{F}_{Lt}(z) = \int_{-\infty}^{z} \hat{f}_{Lt}(x)dx .$$

For each of these estimations, the smoothing parameter of kernel estimators is obtained as: $h_{opt} = 0.9An^{-1/5}$, where *n* denotes the size of the sample, and $A = min(s_n, riq)/1.34$. In this expression, s_n denotes the standard deviation, and *riq* the interquartile range (see Silverman, 1986). Finally, the probabilities have been calculated from the information provided by the *ESEE* for the base year (1990), and we assume the same probabilities for the rest of the period[5].

4.2. Empirical results

Figure 1 reports kernel estimators of the cumulative distribution functions of productivity corresponding to four cross-sections: 1991, 1993, 1995 and 1997. It shows that there has been a systematic shift in the productivity distributions over time in the direction of higher productivity. A different way to summarise the movement in the productivity distributions is to present the quartiles of each cross-sectional estimated distribution, which are reported in Table 2. The median has increased 6.6 percent over the period 1990-1997; that is, the mean annual rate of growth of productivity has been 0.9 percent in this period. The magnitude of the rightward shift in distribution varies over years. In particular, the main improvement in productivity occurs after the year 1993 when the business cycle in Spanish manufacturing entered a stage of expansion.

FIGURE 1 AND TABLE 2 ABOUT HERE

The rightward shift of the distribution is accompanied by the narrowing of distribution. Table 2 exhibits that over the studied period the dispersion of the distribution, measured by the interquartile range, decreases. Most of the narrowing of the interquartile range comes from the relative large increase of the lower quartile, which increases 9.7 percent over the period 1990-97, while the median and the upper rise only 6.6 and 2.3 percent, respectively. This fact indicates that the narrowing of productivity differentials in the manufacturing industry is due to a reduction in the mass of low productivity firms, which could be related to the effect of firm turnover.

5. Productivity and firm turnover

This section analyses whether firm turnover in Spanish manufacturing reflects the underlying differences in productivity, as the models of market selection predict.

[5]In the base year (1990), small firms represent 97.3 percent. The ESEE does not provide information for the other years; thus, we assume the same probabilities for the rest of the period, that is, p_t=0.973 for all t.

Differences in total factor productivity among entering, exiting, and continuing firms are examined by comparing the entire distributions of firm-level productivity in groups of firms with different transition patterns.

We are interested not only in the existence of differences between groups of firms in terms of their cumulative distribution functions, but also in ranking the compared functions. For this purpose we use the concept of stochastic dominance. Let X and Y denote random variables with respective cumulative distribution functions $F_X(.)$ and $F_Y(.)$. *X stochastically dominates Y (X s.d. Y)* if $F_X(z) - F_Y(z) \leq 0$ for all $z \in Z$, where $P(X \in Z \cap Y \in Z) = 1$. The equality of the cumulative distribution functions corresponding to different groups of firms and the sign of existing differences are checked by using non-parametric tests.

The rest of this section is organised as follows. Firstly, we describe the tests of equality and stochastic dominance that we use to rank distributions. Secondly, we study differences among productivity distributions among entering, exiting and continuing firms.

5.1. Testing procedures

The test of stochastic dominance we are interested in consist of a two-sample problem that can be formalised as follows. Let $X_1,..., X_m$ denote a random sample of size m, which corresponds to a group of firms, from the cumulative distribution functions $F_X(.)$, and let $Y_1,..., Y_n$ denote a random sample of size n, which corresponds to a different group of firms, from cumulative distribution functions $F_Y(.)$. Then, X stochastically dominates Y if two complementary conditions are statistically satisfied.

i) Differences between the cumulative distribution functions of compared groups of firms are significant, i.e., we want to test whether the null hypothesis H_0: $F_X(z) = F_Y(z)$ all $z \in Z$ can be rejected (two-sided test);

ii) The sign of the differences between compared distributions is that expected, i.e., we want to test the null hypothesis H_0: $F_X(z) \leq F_Y(z)$ all $z \in Z$ (one-sided test).

The null hypothesis in (i) can also be written as

$$H_0\text{: } \sup_{z\in Z}\left|F_X(z) - F_Y(z)\right| = 0$$

and the statistic proposed by Smirnov (1939) to test this hypothesis is

$$\hat{d}_N = \sqrt{\frac{n.m}{N}} \sup_{z\in Z}\left|\hat{F}_X(z) - \hat{F}_Y(z)\right|,$$

where $\hat{F}_X$ and $\hat{F}_Y$ are the empirical distribution functions corresponding to F_X and F_Y, respectively, and $N=n+m$[6].

Similarly, the null hypothesis in (ii) can also be written as

$$H_0\text{: } \sup_{z\in Z}\left(F_Y(z) - F_X(z)\right) \geq 0$$

which can be tested using the Smirnov (1939) statistic for the two-sample problem[7]; that is

$$\hat{h}_N = \sqrt{\frac{n.m}{N}} \sup_{z\in Z}\left(\hat{F}_Y(z) - \hat{F}_X(z)\right).$$

5.2. Incumbents, entrants and exiting firms

[6] Kolmogorov (1933) showed that the limiting distribution of this statistic is given by $\lim_{N\to\infty} P\left\{\hat{d}_N > u\right\} = -2\sum_{k=1}^{\infty}(-1)^k \exp(-2k^2u^2)$.

[7] Smirnov (1939) showed that the limiting distribution of this statistic is given by $\lim_{N\to\infty} P\left\{\hat{h}_N > u\right\} = \exp(-2u^2)$.

For each cross-section, firms in the sample can be classified in three groups: entering (in the year they begin to work), exiting (in the last year they work) and continuing firms. The data set allows us to identify six cohorts of entering firms (1991, 1992, 1993, 1994, 1996 and 1997)[8] and six cohorts of exiting firms[9] (1991, 1992, 1993, 1994, 1995 and 1996). We will examine the differences in terms of productivity among these groups of firms by comparing their estimated cumulative distribution functions. We will also test whether observed differences are significant or not.

To estimate the cumulative distribution function for each group of firms, we have to recall that the sample covers different percentages of the total number of firms in the population of small and large firms. To estimate the cumulative distribution function for continuing firms, we apply the two-stage procedure of estimation described in section 4. That is, we firstly estimate the conditional cumulative distribution function for each size group of continuing firms, and then, we estimate the marginal cumulative distribution function of the whole population of continuing firms according to equation [5] [10]. To estimate the cumulative distribution functions for entering and exiting firms, we cannot apply the same method. It is because the number of large entering or exiting firms is not large enough to obtain kernel estimators[11]. However, 99.9 percent of the entering firms and 99.3 percent of the exiting firms employ fewer than 200 workers (see Table 1*a*). It suggests that, in both cases, we can estimate the cumulative distribution function for the whole population from data about small firms.

TABLE 3 ABOUT HERE

Table 3 reports the quartiles of kernel estimators of cumulative distribution functions of productivity for the pooled samples corresponding to entering, exiting and continuing firms. The median productivity of entrants is 6.1 percent lower than the median productivity of incumbents. The differential between the median productivity of incumbents and exiting firms reaches 13.5 percent. Then it indicates that the median productivity of entering firms is 7.4 percent higher than that of exiting firms.

---

[8] Entering firms in 1995 were not included in the sample.
[9] We consider an exit year the last year a firm is in the market.
[10] In 1990, small firms represent 96.7 percent of continuing firms; this percentage is assumed to remain fixed over the period. That is, p_t=0.967 for all *t*.

The comparison in terms of the upper or the lower quartiles ranks the three groups of firms in the same way. However, the lower the quartile is, the higher the productivity differentials are. To summarise, the position of the distribution for continuing firms is to the right of the distribution for entering firms, which is to the right of the distribution of exiting firms.

FIGURE 2 ABOUT HERE

Figure 2 reports the kernel estimators of the cumulative distribution functions of productivity for entering, exiting and continuing firms. The relative position of kernel estimators of distributions gives prime evidence of two facts. First, the productivity distribution of incumbents stochastically dominates the productivity distribution of entrants. Second, the productivity distribution of entrants stochastically dominates the productivity distribution of exiting firms. This implies that the productivity distribution of exiting firms is stochastically dominated by the productivity distribution of incumbents. We will apply for each time period[12] the two-sided and the one-sided tests to the following bilateral comparisons of cumulative distribution functions of productivity: i) continuing fims vs entrants, ii) continuing vs exiting firms, and finally, iii) entrants vs exiting firms, to test whether the productivity differentials are significant.

Let F_{1t} (z), F_{2t} (z) and F_{3t} (z) denote the cumulative distribution functions of productivity (z) at time t for continuing, entering and exiting firms, respectively. Then, the null hypotheses we have to test are the following.

i) F_{1t} (z) stochastically dominates F_{2t} (z).

ii) F_{1t} (z) stochastically dominates F_{3t} (z).

ii) F_{2t} (z) stochastically dominates F_{3t} (z).

TABLE 4 ABOUT HERE

---

[11] The annual number of large entering firms is three at the most and, in most cases, the annual number of exiting firms is below four. For more details see the appendix

[12] The distribution of the statistics for the two-sample problem is derived under the assumption of independence between the observations from the two compared distributions.

i) Productivity differences between continuing and entering firms

The quartiles of kernel estimators of the cumulative distribution functions of productivity for different cohorts of continuing firms are reported in Table 4. Table 5 presents the quartiles of the distributions of entering firms, distinguishing by cohorts of entrants, for the entry year and for the years that follow.

TABLE 5 ABOUT HERE

Two conclusions that can be drawn from Table 4 and Table 5 are the following. First, the quartiles for the continuing firms are higher than the quartiles for the recent cohort of entrants for each time period. Second, the magnitude of the differences between the two groups varies over time.

Table 6 presents the statistics and p-values corresponding to the two-sided and the one-sided tests (at the first and fourth columns, respectively) applied to the comparison between the distributions of entrants and continuing firms. The null hypothesis of stochastic dominance of continuing firms cannot be rejected at any reasonable level of significance. However, in 1994 and 1996, the null hypothesis of equality between the productivity distributions of continuing and entering firms can be rejected, respectively, at the levels 0.31 and 0.47 which are relatively high.

TABLE 6 ABOUT HERE

A detailed examination of Table 5 allows us to study the evolution of the productivity distribution of the cohorts of entrants over the earlier years they operate in markets. In general, in the earlier years the distribution of cohorts of entrants shifts to the right and narrows. This narrowing is mainly due to the relatively higher improvement of the lower quartiles. One possible explanation for this fact is that the members of a cohort are heterogeneous; however, market forces operate selecting the relatively high-productivity members. Additionally, in contact with the market, surviving firms could improve their productivity and converge to the productivity levels of incumbents.

Figure 3 shows the estimated cumulative distribution functions for the pooled samples of the surviving and the failing members of the six cohorts of entrants at the entry year. The quartiles for both estimated distributions are reported in Table 7, which also reports the statistics and p-values for the test of stochastic dominance of survivors. In the entry year, the median productivity of failing entrants is 19 percent lower than the median productivity of survivors. Equality of both distributions is rejected for any level of significance and the null hypothesis of stochastic dominance of the surviving-entrants cannot be rejected.

FIGURE 3 ABOUT HERE

Whether the improvement in productivity of surviving entrants in contact with the market is higher than for the rest of the firms and, therefore, there is a convergence process, is an open question for further research.

ii) Productivity differences between continuing and exiting firms

Table 6 reports the hypothesis test statistics for the null hypotheses of equality of distributions of productivity for continuing and exiting firms (second column) and stochastic dominance of the former group (fifth column). The null hypothesis of equality of both distributions is clearly rejected for each time period and, the null hypothesis of stochastic dominance of continuing firms can never be rejected at any reasonable level of significance.

iii) Productivity differences between entering and exiting firms

Finally, we test whether the productivity differential between entering and exiting firms, in favour of the former group, is significant. Hypotheses test statistics and p-values are reported in the third and the last columns of Table 6. The null hypothesis of equality can only be rejected in 1994 and 1996, at the levels 0.14 and 0.02, respectively. In these years, the stochastic dominance of entrants cannot be rejected at any reasonable level of significance.

Summarising, we find significant differences in productivity among continuing, entering and exiting firms. Empirical results show that exiting firms are less productive than survivors, in the sense that the productivity distribution of exiting firms is stochastically dominated by that of continuing firms, which also stochastically dominates the distribution of entering firms. Productivity differences between entering and exiting firms in favour of the former are not significant in general. Finally, we find that, over the earlier years, productivity distribution of entering firms improves, and the heterogeneity among members of entry cohorts tends to decrease as a result of the exiting of the less productive members of the cohort.

6. Productivity and persistence

In previous sections we have found evidence in favour of the improvement of firm-level productivity distribution over time, and of the relationship between firm turnover and productivity. However, the comparison of productivity distributions over years cannot reveal the individual patterns of behaviour that determine the way in which distribution evolves over time. In this section we will study the degree of persistence in firm-level productivity.

Now we are concerned about the relationship between the productivity level of firm *f* at time *t* (z_t) and its productivity level in the past (z_{t-1}). This goal involves the analysis of bivariate data; hence, we are compelled to study the joint distribution of z_t and z_{t-1}.

Let $\hat{f}(z_t, z_{t-1})$ denote the kernel estimator of the bivariate probability density function of productivity at time *t* and productivity at time *t-1*. Then the kernel estimator of the conditional probability density function of productivity at time *t* given productivity at time *t-1* is defined as

$$\hat{f}(z_t \mid z_{t-1}) = \frac{\hat{f}(z_t, z_{t-1})}{\hat{f}(z_{t-1})},$$

where $\hat{f}(z_{t-1})$ is the kernel estimator of the marginal probability density function of productivity at time *t-1*, which is defined as

$$\hat{f}(z_{t-1}) = \int_{-\infty}^{\infty} \hat{f}(z, z_{t-1}) dz \, .$$

Then the kernel estimator of the conditional cumulative distribution function of productivity at time *t* given productivity at time *t-1* is defined as

$$\hat{F}(z_t \mid z_{t-1}) = \int_{-\infty}^{z} \hat{f}(z \mid z_{t-1}) dz$$

We estimate this conditional cumulative distribution function for the pooled sample of firms[13].

FIGURE 4 ABOUT HERE

Figure 4 consists of two panels that show the kernel estimator of the conditional cumulative distribution function $\hat{F}(z_t \mid z_{t-1})$ for the pooled sample of firms in two different ways. The first panel reports the function $\hat{F}(z_t = z \mid z_{t-1})$ for seven different values of z. For each value of z, this function gives the probability that productivity at time t is lower than z for different productivity levels at time t-1. The second panel depicts the estimator of the conditional cumulative distribution function of productivity at time *t*, given that productivity at *t-1* equals, that is, $\hat{F}(z_t \mid z_{t-1} = z)$, for seven different values of z.

The stylised facts that can be drawn from Figure 4 are the following. First, the function $\hat{F}(z_t = z \mid z_{t-1})$ is decreasing in z_{t-1}, for all values of z. It implies that the higher the productivity at time *t-1*, the higher the probability of having a "high"

---

[13]To estimate the conditional cumulative distribution function we apply the two-stage procedure of estimation described in section 4. That is, we firstly estimate the conditional cumulative distribution function for each size group, and then we estimate the conditional cumulative distribution function of the whole population according to equation [5], where p_t=0.973.

productivity at time *t*. Second, the position of the cumulative distribution function of productivity at time *t* depends positively on the productivity level at time *t-1*. That is, the distribution of productivity at time *t* shifts to the right when the productivity level at time *t-1* increases. It implies that differences in terms of productivity are persistent over time. Finally, the probability of productivity improvements is strictly positive, whatever the initial productivity level, but this probability is inversely related to the initial level.

7. Concluding remarks

This paper employs a data set referring to a sample of Spanish manufacturing firms over the period 1990-1997, which comes from the Spanish survey *Encuesta sobre Estrategias Empresariales (ESEE)*, to study the stylised facts of firm-level productivity dynamics. It concludes that firm-level productivity distribution in Spanish manufacturing improved over the studied period and, at the same time, productivity distribution tended to concentrate.

The panel structure of the sample is used to explore and test productivity differences among continuing, entering and exiting firms. Evidence drawn from comparison among productivity distributions of continuing, entering and exiting firms can be summarised in the following facts. First, the productivity distribution of incumbents stochastically dominates the productivity distribution of entering and exiting firms. Second, cohorts of entrants are themselves a heterogeneous group whose distribution tends to improve over time, in part due to the exit of less efficient members. Third, differences between productivity distributions for entering and exiting firms are not significant. Finally, this paper finds signs of persistent heterogeneity in productivity levels across firms.

Appendix: definitions of output and inputs

Output: measured by annual gross production of goods and services expressed in real terms using individual price index for each firm drawn from the *ESEE*.

Labour input: measured by the number of effective yearly hours of work, which is equal to normal yearly hours plus overtime yearly hours minus non-working yearly hours. The cost of labour is measured by the sum of wages, social security contributions, and other labour costs paid by the firm.

Materials: measured by the cost of intermediate inputs; it includes 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*.

Stock of capital: it is calculated according to Martín Marcos and Suárez (1997), who use the perpetual inventory formula: $k_t^* = I_t + k_{t-1}^*(1-d_t)\frac{P_t}{P_{t-1}}$, where I_t represents investment in equipment, d_t are depreciation rates obtained from Martin Marcos (1990) and P_t corresponds to price indexes for equipment published by the Spanish National Institute of Statistics (INE). 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.

INSERT TABLE A1 AND TABLE A2

References

Aw, B. Y.; Chen, X. and Roberts, M. (1997), "Firm level evidence on productivity differentials, turnover, and exports in Taiwanese manufacturing", Pennsylvania State University, Working Paper.

Baily, M.N.; Hulten, C. and Campbell, D. (1992), "Productivity dynamics in manufacturing plants", *Brooking Papers: Microeconomics 1992*, pp. 187-225.

Good, D. H. (1985), " The effect of deregulation on the productive efficiency and cost structure of the airline industry", PH.D. dissertation, University of Pennsylvania.

Good, D.H.; Nadiri, M.I. and Sickless, R. (1996), "Index Number and Factor Demand Approaches to the Estimation of Productivity", NBER Working Paper 5790.

Hopenhayn, H. (1992), "Entry, Exit, and firm dynamics in long run equilibrium", *Econometrica*, 60, September, pp. 1127-1150.

Kolmogorov, A.N. (1933), "Sulla determinazione empirica di une legge di distribuzione", *Giorn. dell´Istit. Degli att.*, Vol 4, pp. 83-91.

Martin Marcos, A (1990), "Estimación del stock de capital para los sectores de la EI", Documento Interno nº 4. *Programa de Investigaciones Económicas. Fundación Empresa Pública.*

Martin, A and Suárez, C (1997), "El stock de capital para las empresas de la Encuesta Sobre Estrategias Empresariales". Documento Interno nº 13. *Programa de Investigaciones Económicas. Fundación Empresa Pública.*

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", *Bull. Math. Univ. Moscow*, vol. 2, nº 2, pp.3-14.

Table 1(*a*)

Number of firms in manufacturing population

1990	Total	Size (number of workers)	
		≤ 200 workers	> 200 workers
Population	38503	37453	1050
Continuing firms	35046	96.7%	3.3%
Entering firms	2250	99.9%	0.1%
Exiting firms	1207	99.3%	0.7%

Table 1(*b*)

Number of firms in the sample and sampling fractions

1990	Total	Size (number of workers)	
		≤ 200 workers	> 200 workers
Sample	2188	1478	710
Sampling fractions		3.9%	67.6%

Table 2

Evolution of productivity distribution: descriptive statistics

	Lower quartile	Median	Upper quartile	Interquartile range
1990	-0.235	-0.105	0.041	0.277
1991	-0.231	-0.101	0.036	0.266
1992	-0.213	-0.082	0.042	0.255
1993	-0.208	-0.076	0.043	0.252
1994	-0.184	-0.058	0.059	0.242
1995	-0.153	-0.040	0.069	0.222
1996	-0.136	-0.031	0.082	0.218
1997	-0.138	-0.039	0.064	0.202
Mean annual rate of growth	0.014	0.009	0.003	-0.011

Table 3
Productivity distributions of continuing, entering and exiting firms: descriptive statistics

	Continuing firms	Entering firms	Exiting firms
Lower quartile	-0.171	-0.256	-0.377
Median	-0.054	-0.115	-0.189
Upper quartile	0.061	0.021	-0.018
Interquartile range	0.232	0.277	0.359

Table 4

Total Factor Productivity of continuing firms: descriptive statistics

	Lower quartile	Median	Upper quartile	Interquartile range
1991	-0.224	-0.098	0.036	0.260
1992	-0.202	-0.077	0.043	0.245
1993	-0.198	-0.069	0.048	0.246
1994	-0.174	-0.052	0.063	0.237
1995	-0.149	-0.037	0.070	0.219
1996	-0.131	-0.027	0.088	0.219
1997	-0.128	-0.033	0.069	0.197
Mean annual rate of growth	0.016	0.011	0.005	-0.011

Table 5

Total Factor Productivity of entering-firm cohorts

Cohort of entrants	Descriptive statistics	1991	1992	1993	1994	1995	1996	1997
1991	Lower quartile	-0.36	-0.31	-0.24	-0.28	-0.19	-0.22	-0.18
	Median	-0.17	-0.14	-0.11	-0.09	-0.01	-0.09	-0.08
	Upper quartile	0.01	0.00	0.02	0.07	0.01	0.04	0.01
1992	Lower quartile		-0.37	-0.37	-0.30	-0.26	-0.22	-0.18
	Median		-0.18	-0.19	-0.15	-0.11	-0.09	-0.08
	Upper quartile		0.02	-0.01	-0.01	0.01	0.03	-0.01
1993	Lower quartile			-0.30	-0.20	-0.19	-0.15	-0.15
	Median			-0.15	-0.09	-0.08	-0.04	-0.05
	Upper quartile			0.01	0.04	0.05	0.07	0.05
1994	Lower quartile				-0.23	-0.23	-0.12	-0.20
	Median				-0.12	-0.07	-0.02	-0.07
	Upper quartile				0.01	0.10	0.09	0.05
1996	Lower quartile						-0.15	-0.13
	Median						-0.05	-0.04
	Upper quartile						0.05	0.06
1997	Lower quartile							-0.26
	Median							-0.13
	Upper quartile							-0.01

Table 6
Productivity differences among continuing, entering and exiting firms.
Hypotheses test statistics (and p-values)

Null hypotheses	Test of equality of distributions			Test of stochastic dominance (*sd*)		
	$F_1 = F_2$	$F_1 = F_3$	$F_2 = F_3$	F_1 *sd* F_2	F_1 *sd* F_3	F_2 *sd* F_3
1991	1.09 (0.18)	1.21 (0.10)	0.25 (0.99)	0.01 (0.99)	0.06 (0.99)	0.21 (0.91)
1992	1.83 (0.00)	1.33 (0.06)	0.44 (0.99)	0.20 (0.93)	0.29 (0.85)	0.38 (0.75)
1993	1.52 (0.02)	1.20 (0.11)	0.51 (0.96)	0.05 (0.99)	0.08 (0.99)	0.15 (0.96)
1994	0.96 (0.31)	2.16 (0.00)	1.15 (0.14)	0.10 (0.98)	0.01 (0.99)	0.00 (0.99)
1995		1.54 (0.02)			0.05 (0.99)	
1996	0.85 (0.47)	1.74 (0.01)	1.53 (0.02)	0.19 (0.93)	0.05 (0.99)	0.00 (0.99)
1997	2.36 (0.00)			0.01 (0.99)		

F_1, F_2 and F_3 denote, respectively, the cumulative distribution functions of continuing, entering and exiting firms.

Table 7

Total Factor Productivity: surviving and failing members of entry cohorts

	Surviving-entering firms	Surviving-failing firms
Lower quartile	-0.237	-0.449
Median	-0.106	-0.291
Upper quartile	0.023	-0.068
Interquartile range	0.260	0.381

Hypotheses test statistics	Statistic	(p-value)
• Equality of distributions	1.82	(0.00)
• Stochastic dominance of surviving firms distribution	0.09	(0.98)

Table A1

Percentages of continuing, entering and exiting firms

	≤ 200 workers				> 200 workers			
	Total	Continuing firms (%)	Entering firms (%)	Exiting firms (%)	Total	Continuing firms (%)	Entering firms (%)	Exiting firms (%)
1991	1425	93.6	3.0	3.4	755	99.7	0.1	0.1
1992	1305	89.9	5.9	4.2	671	98.5	0.3	1.2
1993	1310	90.0	7.0	3.0	566	98.1	0.5	1.4
1994	1305	93.0	3.7	3.3	618	99.2	0.2	0.6
1995*	1259				558			
1996	1195	87.6	10.5	1.9	522	99.0	0.4	0.6
Mean		90.8	6.0	3.2		98.9	0.3	0.8

* Entering firms in 1995 were not included in the sample.

Table A2

Productivity of entering firms and exiting firms

(>200 workers)

	1991	1992	1993	1994	1995	1996	1997
Entering firms (productivity at the entry year)	-0.124**	-0.245* -0.035	-0.455* -0.094** 0.189	-0.015		-0.241* 0.029	-0.091** 0.181 0.282
Exiting firms (productivity at the exit year)	-0.130**	-0.898* -0.210** -0.178** -0.176** -0.049** -0.046 0.051 0.114	-0.377* -0.141** -0.066 0.122 0.147 0.251 0.338	-1.278* -0.109** 0.231	-0.231*	-0.601* -0.022 -0.013	

* Lower than the lower quartile of productivity distribution.
** Between the lower quartile and the median of productivity distribution.

Figure 1

Evolution of productivity distributions over the period 1990-1997 (Kernel cumulative distribution estimators)

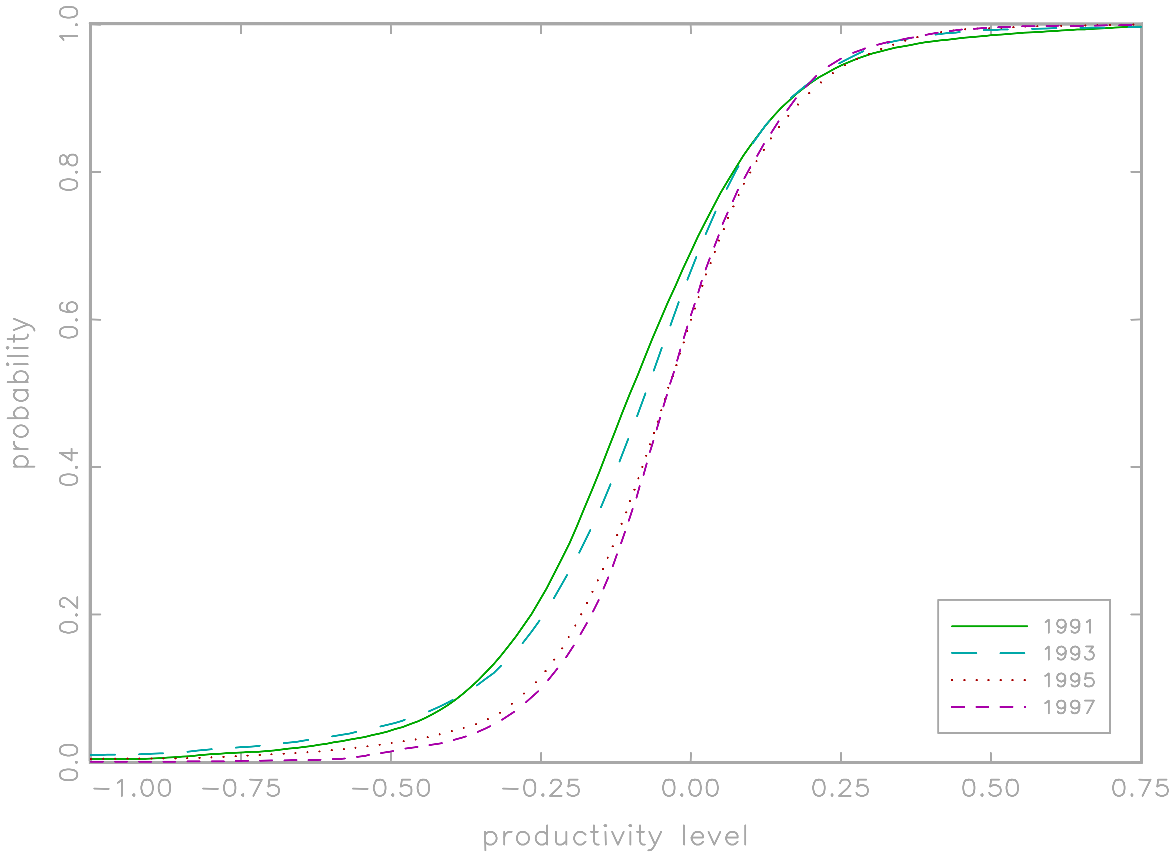


Figure 2

Total Factor Productivity differences among continuing, entering and exiting firms (Kernel cumulative distribution estimators)

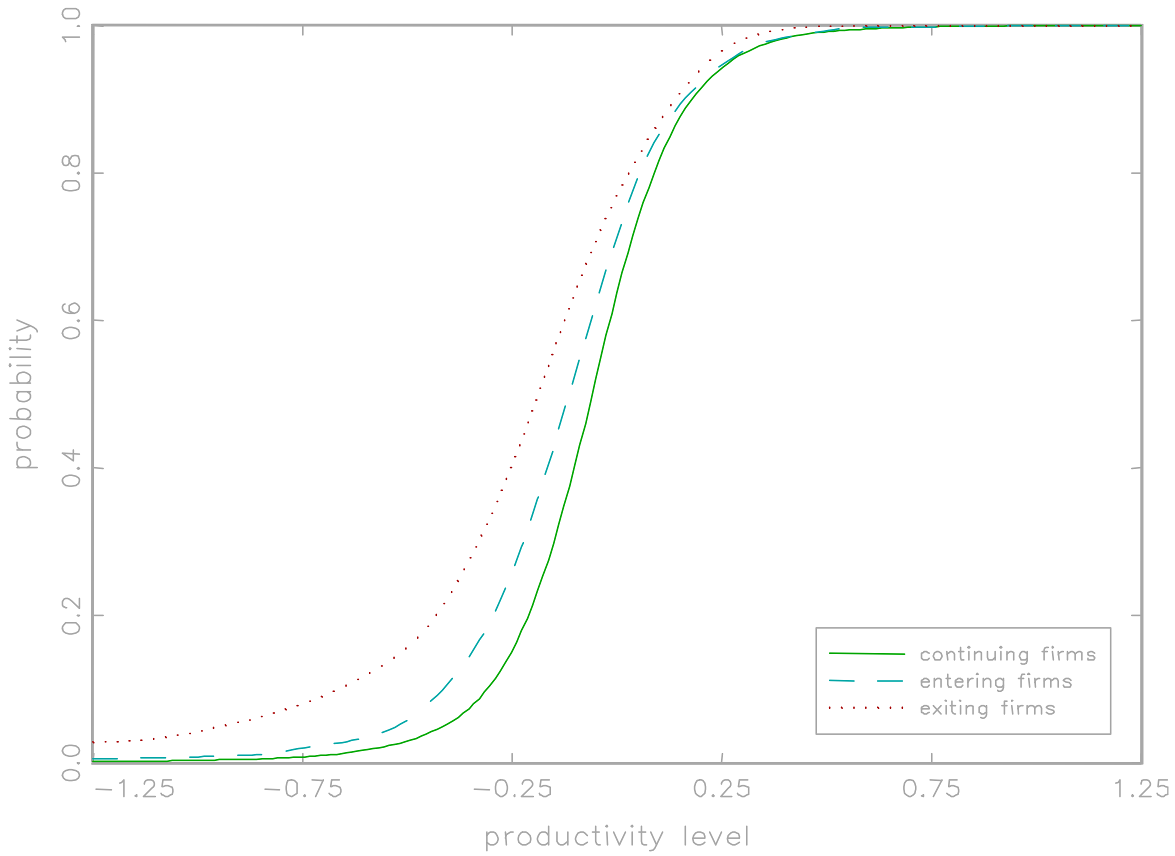


Figure 3

Selection mechanism: productivity differences between surviving and failing entrants (Kernel cumulative distribution estimators)

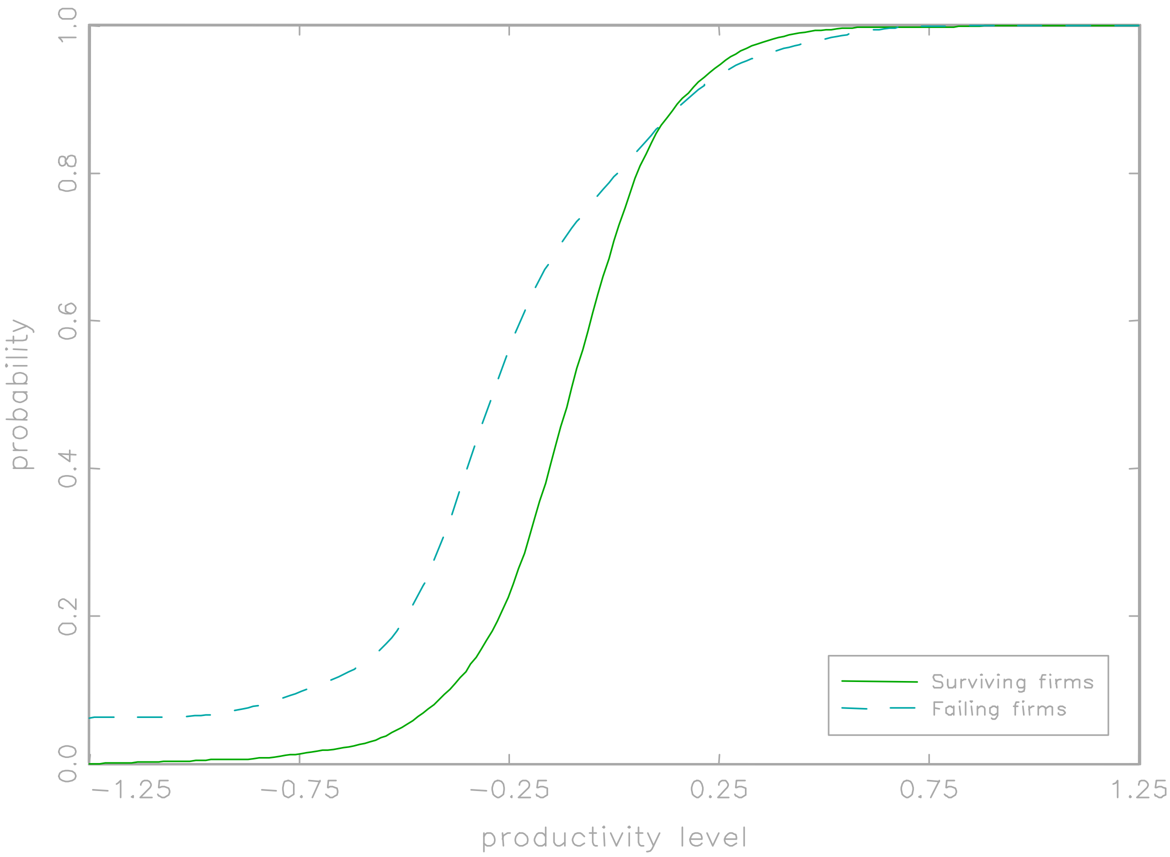
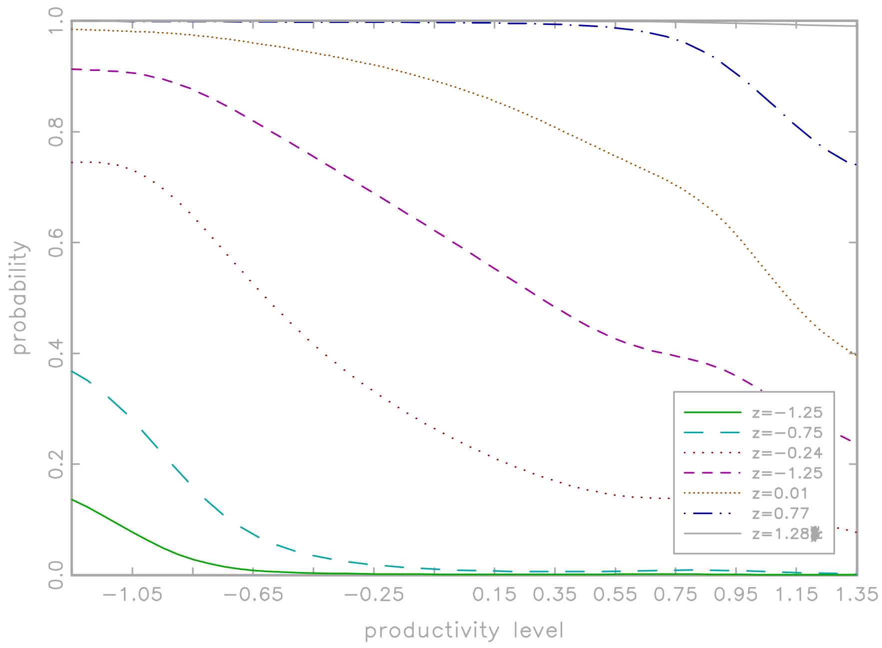


Figure 4

Persistence of Total Factor Productivity at the firm level
(Kernel conditional cumulative distribution estimator)

$$\hat{F}(z_t = z \mid z_{t-1})$$



$$\hat{F}(z_t \mid z_{t-1} = z)$$