metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Econ

The Economic and Social Review, Vol. 39, No. 1, Spring, 2008, pp. 13–38

The Distribution of Productivity in Irish Manufacturing Between 1995 and 2004 – Determinants, Changes and Implications

STEFANIE A. HALLER*
The Economic and Social Research Institute, Dublin

Abstract: Using plant level data from the Irish *Census of Industrial Production*, this paper documents the extent of the productivity spread in Irish manufacturing industries and its determinants. It looks at changes in the distribution of productivity over the period 1995-2004 and at movements of plants within the distribution. It also examines the relationship between spreads and productivity growth. The annual average productivity growth of 3.9 per cent over the period has rendered plants across the distribution more productive. However, there was less than proportional entry of new plants at the top of the productivity distribution until 2000. Persistence of plants within the productivity distribution is high, although mean convergence is faster for plants with below average productivity. Productivity growth is slower in industries with larger spreads in the lower half of the distribution.

I INTRODUCTION

Underlying aggregate productivity dynamics are huge differences in productivity between plants even in narrowly defined sectors. The increasing availability of micro data sets has directed attention to this productivity 'spread', i.e. the difference between the best and the worst

* I am grateful for valuable discussions with John Fitz Gerald, Laura Malaguzzi Valeri and Eoin O'Malley as well as for comments and suggestions from participants of The Economic and Social Research Institute lunchtime seminars, from an anonymous referee and from the editor. Author's email: stefanie.haller@esri.ie

Paper delivered at The Twenty-First Annual Conference of the Irish Economic Association, Bunclody, Co. Wexford, April 25–27, 2007.

performing firms in an industry as well as to its causes and implications (e.g. Baily *et al.* (1992), Oulton (1998), Bartelsman and Doms (2000), Haskel and Martin (2002), Martin (2005)). There are different perceptions as to whether one should be concerned about the productivity spread (Martin, 2005). First, the productivity spread can be seen as an expression of the co-existence of successful and less successful firms in the competitive selection process in a market economy. Second, "a long tail of underperforming firms" could indicate that the selection process is hampered and, therefore, resources are bound in an unproductive way in firms that do not exit. Third, one could claim that any differences in productivity that we observe are due to measurement problems. The last point applies in particular to labour productivity as a measure of productivity. That is, a plant that substitutes few higher skilled employees for more lower skilled employees will have higher measured labour productivity, but their contributions to welfare need not be any different.

In this context, the aim of the present study is to describe how productivity is distributed in the Irish manufacturing sectors and to document the extent of the productivity spread and its determinants. A further point of interest is to examine how this distribution has changed over the 10-year period from 1995 to 2004 and to document how plants move within the productivity distribution. Implicitly, this gives an indication of the efficiency of resource allocation among plants within sectors. The paper also examines whether there is a correlation between the productivity spread and productivity growth.

For Ireland two studies have looked at productivity at the plant level: Girma, Görg and Strobl (2004) compare the performance of domestic non-exporters, domestic exporters and foreign-owned multinationals using a non-parametric approach based on the principle of first order stochastic dominance. They find that for the year 2000 the distributions of labour productivity for foreign-owned multinationals dominate those of domestic exporters and non-exporters, while they do not find clear differences in plant performance between domestic exporters and non-exporters. Ruane and Uğur (2005) decompose labour productivity into the components attributable to surviving, entering and exiting domestic and foreign-owned plants. Their analysis shows that foreign-owned plants in Ireland contribute a substantial share to overall productivity growth in the period 1991-1999. Most of the productivity growth is generated within surviving plants. The process of entry and exit is also productivity enhancing except in some of the low-tech industries where substantial restructuring is evident.

Labour productivity in the Irish manufacturing sectors grew on average at 3.9 per cent per annum between 1995 and 2004. At the same time the spread between plants at the top and at the bottom of the distribution of labour productivity has decreased somewhat. Comparing the distribution in 2004 to that in 1995 reveals that in the absence of the productivity growth which has rendered plants across the board more productive, there would be fewer plants at the very top of the distribution in 2004 than there were in 1995. This is due to the less than proportional entry at the top of the productivity distribution until 2000. There is evidence of convergence of productivity to an industry mean, and convergence is faster for plants with labour productivity initially below the industry mean. At the same time there is considerable persistence of plants in the productivity distribution with a large share of plants remaining in their relative position over 3-year periods and for plants that do not exit even over the 10-year period. Overall, the reallocation of resources driven by the competitive process appears to be efficient both within and between sectors.

The paper is structured as follows: Section II provides a brief description of the data set. Section III documents the productivity spread and examines its determinants. Section IV compares the distribution of productivity in 2004 to that in 1995 and looks at how plants move within the distribution; it also explores whether there is a correlation between productivity spreads and productivity growth. Section V summarises and concludes.

II DATA

The data set used is the data on the local units from the annual Census of Industrial Production carried out by the Central Statistics Office (CSO).[1] As the data set is a census it covers all local units with 3 or more employees in operation in any one year in the NACE (Rev. 1.1) industrial sectors 10-41. Local units are defined as being primarily engaged in one industrial activity and may be part of larger enterprises. As the deflators[2] for the years prior to 1995 are based on a different industry classification, data for the years 1995-2004 are used. I confine attention to NACE (Rev. 1.1) sectors 15-36, excluding sector 23 (Manufacture of Coke, refined petroleum products and nuclear fuel) for reasons of confidentiality. A summary of the industry groupings used and the corresponding NACE codes is given in Table 1.

[1] The possibility for controlled access to this anonymised micro data set on the premises of the CSO is provided for in the Statistics Act 1993. Assistance with the data by George Hussey, CSO is most gratefully acknowledged. The paper has been screened by the CSO to ensure that no confidential information is revealed.

[2] Turnover is deflated using the wholesale price index at the 2-3-digit NACE level until 1999 and the industrial price index from 2000. The base year is 2000. Both indices are provided by the CSO. Where more detailed information is available these indices are amended with the Annual Total Output Price Index (PRON) provided by Eurostat.

In the absence of an appropriate variable for capital stock, it is not possible to generate a measure of total factor productivity, which would account for differences in input use more appropriately. As a consequence, the measure of productivity used here is labour productivity defined as the natural log of turnover (in €1,000) per employee. The decision in favour of turnover per employee over value added per employee was taken for two reasons: First, while deflators for materials and fuels are available, this is not the case for industrial services. Consequently, real value added has to be obtained by way of single deflation. However, even if double deflation were possible the measure of value added obtained would not capture adequately differences in input prices across firms given that the deflators for materials and fuels are available only at the manufacturing-wide level. Second, using value added per employee in the dataset at hand implies that a larger number of plants which are 'low-productivity' in terms of turnover per employee (80-120 observations per year) cannot be used for the analysis because their value added per employee is smaller than 1 (€1,000) or even negative. Given that I am interested in productivity spreads within industries, I prefer not to cut these low-productivity plants out of the sample. In addition, as most of the analysis is conducted at a rather detailed level of aggregation, the log of turnover per employee captures differences in productivity between plants in the same industry well. Moreover, beyond changes implied by the smaller sample size, the results obtained when using the log of value added per employee as an alternative measure of labour productivity are qualitatively similar.[3]

Before proceeding with the analysis, I screen the data for irregularities. There are roughly 950 plant-year observations over the period that report either turnover or the number of employees to be zero; these observations drop out when taking logs. I drop all observations where there is only one plant in a 4-digit industry-year cell. I treat as outliers those plants that have a negative take-in of materials and that have observations for profit margins (turnover less materials and labour cost over turnover) in the top and bottom percentile of the distribution of profit margins. I also treat as outliers plants whose deviations from the industry-year mean of labour productivity lie in the top or bottom percentile. I accumulate outlier definitions and delete all plants that have 1 or more outliers according to the material and profit margin definition and I omit all plants that have 2 or more outliers according to the deviation from the industry-mean labour productivity definition.

[3] These results are available from the author on request.

III THE PRODUCTIVITY SPREAD AND ITS DETERMINANTS

Table 1 provides summary statistics of the distribution of labour productivity, i.e. the log of turnover per employee. As can be seen from the table, there are substantial differences both within and across sectors. Looking at the median, only plants in food and tobacco, chemicals and in 2004 also electrical and optical equipment have on average higher levels of turnover per employee than the manufacturing sector as a whole. At the other end, textiles, clothing and leather stand out with a median level of turnover per employee well below average, especially in 1995.

In addition to the median, the table shows productivity spreads between the 90th and the 10th percentile, between the 90th percentile and the median, and between the median and the 10th percentile. These spreads vary substantially between industries, in particular the ratio of plants at the upper end of the distribution to plants at the lower end (p90/p10). As labour productivity is a measure in log terms the differences may not seem so large, however, in terms of turnover per employee the ratios range from factors 3.5-4.5. For example, a plant at the 90th decile of the productivity distribution in the chemical industry in 1995 has 4.17 times the turnover per employee of a plant at the 10th decile of the productivity distribution in the chemical sector in 1995.

Whether the spread is larger in the upper or in the lower half of the distribution depends very much on the industry under consideration, with a few exceptions the differences are not very large. At the aggregate level the differences are quite small. For the manufacturing sector as a whole all productivity spreads decreased marginally from 1995 to 2004, with a few exceptions this is also true in most of the individual sectors. This is in contrast to findings for the UK, where Haskel and Martin (2002) document a small increase between 1980 and 2000.

While the median level of those plants that survived the full ten-year period was slightly higher than in the manufacturing sector as a whole in 1995, it is substantially smaller in 2004. The spreads for these plants are nearly unchanged. The Irish-owned plants have lower median labour productivity than the manufacturing sector as a whole, whereas it is higher for the foreign-owned plants. Note that both groups individually have lower spreads than when they are grouped together. Hence, it is the large differences in productivity between the two groups of plants in each sector that are responsible for the spreads at the aggregate level.

To get an idea of the development over time, Table 2 shows the year on year growth rates from 1995 to 2004. Averaging over 3-digit industry means, turnover per employee grew at about 3.9 per cent over the 10-year period.

Table 1: *Summary Statistics of Labour Productivity (Log Turnover Per Employee) and the Productivity Spread 1995-2004*

NACE		*1995* Median	Std Dev	p90/ p10	p90/ p50	p50/ p10	Plants
15,16	Food and Tobacco	5.07	0.75	1.43	1.16	1.23	656
17-19	Textiles, Clothing and Leather	4.12	0.54	1.38	1.14	1.21	386
20	Wood	4.30	0.55	1.37	1.16	1.18	190
21-22	Paper, Printing and Publishing	4.42	0.56	1.34	1.18	1.14	500
24	Chemical	5.10	0.73	1.43	1.17	1.21	200
25	Rubber	4.28	0.55	1.34	1.18	1.14	224
26	Mineral	4.42	0.55	1.36	1.17	1.16	244
27,28	Basic Metal and Metal Products	4.35	0.62	1.36	1.17	1.17	515
29	Machinery	4.33	0.54	1.35	1.13	1.19	314
30-33	Electrical and Optical Equipment	4.32	0.70	1.50	1.24	1.21	374
34,35	Transport	4.21	0.56	1.40	1.17	1.19	129
36	Other	4.10	0.55	1.37	1.15	1.19	355
All Manufacturing (Nace15-36, ex 23)		4.48	0.62	1.39	1.17	1.19	4,087
Survivors		4.50	0.53	1.30	1.13	1.15	2,112
Domestic Plants		4.33	0.56	1.35	1.15	1.18	3,471
Foreign Plants		4.87	0.52	1.19	1.08	1.09	616

NACE		*2004* Median	Std Dev	p90/ p10	p90/ p50	p50/ p10	Plants
15,16	Food and Tobacco	5.19	0.81	1.45	1.15	1.26	542
17-19	Textiles, Clothing and Leather	4.43	0.68	1.43	1.18	1.22	231
20	Wood	4.80	0.55	1.33	1.17	1.14	265
21-22	Paper, Printing and Publishing	4.60	0.50	1.29	1.13	1.14	563
24	Chemical	5.44	0.76	1.38	1.19	1.15	172
25	Rubber	4.76	0.47	1.24	1.12	1.10	264
26	Mineral	4.84	0.67	1.39	1.14	1.22	318
27,28	Basic Metal and Metal Products	4.74	0.63	1.29	1.16	1.11	627
29	Machinery	4.61	0.50	1.30	1.14	1.15	288
30-33	Electrical and Optical Equipment	5.00	0.74	1.42	1.22	1.17	322
34,35	Transport	4.31	0.47	1.28	1.13	1.13	98
36	Other	4.38	0.64	1.41	1.21	1.17	475
All Manufacturing (Nace15-36, ex 23)		4.81	0.65	1.36	1.16	1.17	4,165
Survivors		4.57	0.53	1.29	1.13	1.14	2,112
Domestic Plants		4.72	0.59	1.33	1.14	1.16	3,659
Foreign Plants		5.21	0.62	1.20	1.09	1.09	506

Note: Spreads (p50/p10, p90/p50, p90/p10) calculated at 3-digit NACE level.

Table 2: *Average Annual Growth Rates of Labour Productivity in Manufacturing – 1995-2004*

	1996	*1997*	*1998*	*1999*	*2000*	*2001*	*2002*	*2003*	*2004*	*1995-2004*
	%	%	%	%	%	%	%	%	%	%
Growth	5.8	8.6	–1.5	4.7	6.7	–3.1	1.7	7.9	4.6	3.9
Std. Dev.	24.4	21.7	18.9	22.5	20.8	21.0	22.9	19.3	22.9	22.0

Note: Calculated at 3-digit NACE level.

However, the dynamics from year to year vary considerably. Moreover, the standard deviations reflect sizeable differences in labour productivity growth between industries.

Manufacturing in Ireland has seen a different development from most other European countries in another potentially important dimension for this analysis. The shift towards more service-oriented economies has in many countries led to a decline of the manufacturing sector both in terms of the number of plants as well as in terms of the number of persons employed. While the services sectors in Ireland have certainly seen increases in numbers of enterprises and employees, in the manufacturing sector as a whole the figures are nearly the same in 2004 as they were in 1995. However, both the number of plants and the number of employees in the industrial sector are declining after a peak around the millennium.

Determinants of the Productivity Spread

To get an idea which factors are associated with the productivity spread, I regress the standard deviation (StdDev) and the spread between the 90th and the 10th percentile (p90/p10) on a number of 3-digit industry-mean indicators. The two measures of the productivity spread are calculated as industry means at 3-digit industry level. The estimating equation is the following:

$$SPREAD_{It} = \alpha_0 + \alpha_1\, EMP_{It} + \alpha_2\, HHI_{It} + \alpha_3\, PM_{It} + \alpha_4\, EXP_{It} + \lambda_t + \varepsilon_{It}$$

Among the explanatory factors considered are the industry average of plants' employment levels excluding outside-piece workers in log terms, EMP; the Herfindahl-Hirschman index normalised on the interval [0;1] as a measure of concentration HHI; and profit margins, PM, as a measure of excess profits in the industry calculated as the industry average over the plants' share of turnover less material cost and wage cost in turnover. The share of exporters in industry *I*, EXP, is included as a measure of exposure to foreign markets. In addition to the constant α_0, time dummies are also included. If the spread is greater in less competitive industries, we expect $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_3 > 0$, $\alpha_4 < 0$

as industries with high employment, high concentration and high profit margins are less competitive. In contrast, industries with a large share of exporters might be considered more competitive.

Table 3 shows the results. Except for the exporter share the determinants work largely as expected. Care should be taken when interpreting these results: the estimated effects should not be read as causal effects, indeed causality could just as well work in the opposite direction.

Table 3: *Determinants of the Productivity Spread*

Dep. Variable	*StdDev*	*p90/p10*
EMPloyment	0.108	0.066
	(0.012) ***	(0.009) ***
HHI	0.025	0.153
	(0.056)	(0.042) ***
Profit Margin	0.137	-0.005
	(0.078) *	(0.059)
EXPorter share	0.036	0.025
	(0.039)	(0.030)
Constant	0.262	1.215
	(0.042) ***	(0.032) ***
Year Dummies	yes	yes
N	740	740
Adj. R^2	.17	.15

Notes: ***, **, * indicate significance at 1 per cent, 5 per cent, 10 per cent, respectively. The unit of observation in these regressions are 3-digit industry-year means. Industry-year cells with less than 5 plants are excluded. Industries where data is not available for all years are excluded as well.

Sectors with high average employment levels have greater spreads; this may indicate a lack of competition if such sectors are more difficult to enter. The coefficient on the Herfindahl-Hirschman index is significant with a positive sign only for the spread between the 90th and the 10th percentile.[4] This is plausible if concentration is taken as an indicator of a lack of competition. A similar argument applies to profit margins, if a lack of competition allows plants to extract large rents. This coefficient is weakly significant when the standard deviation is to be explained. Instead for the spread between the 90th and the 10th percentile the coefficient is close to zero and not significant. This is the result of opposing effects in the upper and the lower half of the productivity distribution. Unreported regressions show that

[4] Note the results for the concentration measure are qualitatively identical when using CR5 – the share of the five largest plants in industry turnover – as an alternative measure.

high average profit margins are positively associated with the spread between the 90th percentile and the median, but negatively associated with the spread between the median and the 10th percentile. The coefficient on the share of exporting plants is not significantly different from zero. The reason for this is probably that the share of exporters is larger than 50 per cent in three quarters of all industry-year cells.

IV CHANGES IN AND MOVEMENT WITHIN THE DISTRIBUTION

The Relative Density

To examine the changes in productivity distributions at sectoral level, relative distribution methods are employed. These methods allow a direct comparison of the entire unit-level distribution of a variable at one point in time to the entire unit-level distribution at another point in time. Relative distribution methods were pioneered by Parzen and are described in detail in Handcock and Morris (1999, Chapter 2). I will give an intuitive explanation of this method here, a full algebraic description can be found in the Appendix.

The idea is the following: Taking two random variables that represent measurements of two different populations, Y_0 from a 'reference' population and Y from a 'comparison' population, the distribution of these populations can be expressed by their cumulative density functions (CDFs) F_0 and F, respectively. Assuming continuity and common support, the *relative distribution* of Y to Y_0 can then be expressed as the random variable R, where

$$R = F_0(Y).$$

R is obtained from Y by transforming it with the function F_0 (i.e. the CDF of Y_0). It is continuous with outcome space [0,1]. That is by generating a distribution that transforms the outcomes of the reference distribution to those of the comparison group (R), it becomes possible to compare the two distributions in common space. If the two distributions are identical, it can be shown that the CDF of the relative distribution is a 45° line and the probability density function (PDF) of the relative distribution is the uniform PDF. The relative PDF can be interpreted as a density ratio.

The relative distribution, which summarises differences between the reference and the comparison distribution, can be decomposed further into a 'location' and a 'shape' component. The location component is obtained by constructing a distribution that has the location – in this case the median – of the comparison group but the shape of the reference group. Hence, comparing the relative distribution to the reference distribution will result in a uniform distribution if the comparison and the reference group have the same location. The shape component is obtained when comparing the comparison distribution to the relative distribution. As these two distributions have the same median, a comparison will show differences in shape, a concept that comprises mainly the differences in spread and skewness of the two distributions. That is the shape component shows a uniform distribution if the comparison and the reference distribution have the same shape net of location effects.

As an example take the distribution of labour productivity from 1995 to 2004 for the manufacturing sector as a whole in Figure 1. The first graph in top left corner plots the probability density functions using a Gaussian kernel in 1995 and 2004. One can see that the distribution has shifted to the right, indicating an increase in the average level of labour productivity over the whole range of the distribution. The graph in the top right shows the relative density function (RDF) that is obtained by setting the cumulative density functions of the 1995 and 2004 distributions in relation to each other and grouping the values of this ratio into bins (deciles in this case). If the two distributions were identical the RDF would be equivalent to the uniform distribution, which is indicated by the horizontal dashed line at 1. A segment of the bar plot below this line indicates that the proportion of the distribution of labour productivity in the manufacturing sector as a whole in 2004 in that segment is smaller by the ratio recorded on the vertical axis, than the proportion of that distribution in 1995 in that same segment. The figure thus illustrates the shift of the distribution in that we observe lower frequencies of plants at the bottom of the distribution and higher frequencies of plants at the top. The graphs in the bottom line show, respectively, the location and the shape components of the decomposition of the relative density function. As mentioned above, the location component indicates which fraction of the differences between the two distributions is due to a median shift of the whole distribution, and the shape component shows – adjusted for differences in location – where the two distributions differ in shape. The location component picks up the median shift while the shape component shows that in 2004 in absence of the median shift the frequency of plants in the top decile of the distribution of labour productivity would only be about 70 per cent of what it was in 1995.

Table 4 lists entropy measures for the graphs in Figures 1-6. They give a numerical indication of the size of the change in the distribution from 1995 to 2004,[5] and show in which industries the changes have been strongest. They also highlight the relative importance of the location and shape components in the overall change in each industry. For the manufacturing sector as a whole they confirm the graphical observation that the change in location is responsible for the largest part of the change in the distribution between 1995 and 2004.

[5]A more formal explanation is set out in the Appendix.

Table 4: *Entropy of the Change in the Distribution of Labour Productivity from 1995 to 2004, by Sector and Component*

	Entropy	*Location*	*Shape*	*Share Location %*	*Share Shape %*
Food and Tobacco	0.028	0.003	0.032	9.67	113.27
Textiles, Clothing and Leather	0.512	0.453	0.057	88.44	11.10
Wood	0.495	0.452	0.058	91.35	11.81
Paper, Printing and Publishing	0.208	0.162	0.045	78.02	21.54
Chemical	0.123	0.102	0.023	82.53	18.37
Rubber	0.126	0.089	0.036	70.38	28.60
Mineral	0.155	0.106	0.052	68.29	33.66
Basic Metal and Metal Products	0.194	0.183	0.020	94.41	10.16
Machinery	0.207	0.185	0.025	89.43	12.01
Electrical and Optical Equipment	0.313	0.308	0.018	98.67	5.78
All Manufacturing	0.149	0.140	0.009	93.99	6.35
Surviving Plants	0.130	0.125	0.005	96.19	4.00
All Domestic Plants	0.182	0.173	0.014	94.75	7.78
All Foreign Plants	0.153	0.138	0.019	90.29	12.67

Note: Only the density ratio of the location component is a true density in the sense that it sums to 1, this does not have to be the case for the density ratio of the shape component (see the Appendix for details). As a consequence, also the shares of location and shape in the overall entropy do not necessarily add up to 100.

Figure 2 is equivalent to Figure 1, except that only those plants that survive the full 10-year period are considered. The shape component for these plants indicates at best a minor shortfall of plants in the top decile in 2004 relative to 1995. Hence, it must be the case that there has been exit of plants at the top of the productivity distribution. These plants have not been replaced with equally potent new entrants and/or the new entrants have not been able to increase their productivity quickly enough to make up for this shortfall. Figures 3 and 4 show the decomposition when looking at the domestic and the foreign-owned plants separately. For the domestic plants, Figure 3 exhibits a very similar pattern to that in the manufacturing sector with all plants (Figure 1). Figure 4 shows that for the foreign-owned plants there are fewer plants in the bottom 2 percentiles of their distribution in 2004 relative to 1995 and if anything somewhat more plants at the top of the distribution. This indicates that there is insufficient entry of highly productive domestic plants and those that do enter are unable to move to the top of the productivity distribution in a short time.

Figure 1: *Probability Density Functions, Relative Density, Location and Shape Components Reflecting the Change in the Distribution of Labour Productivity from 1995 to 2004 in All Manufacturing Plants (Reference Group is 1995)*

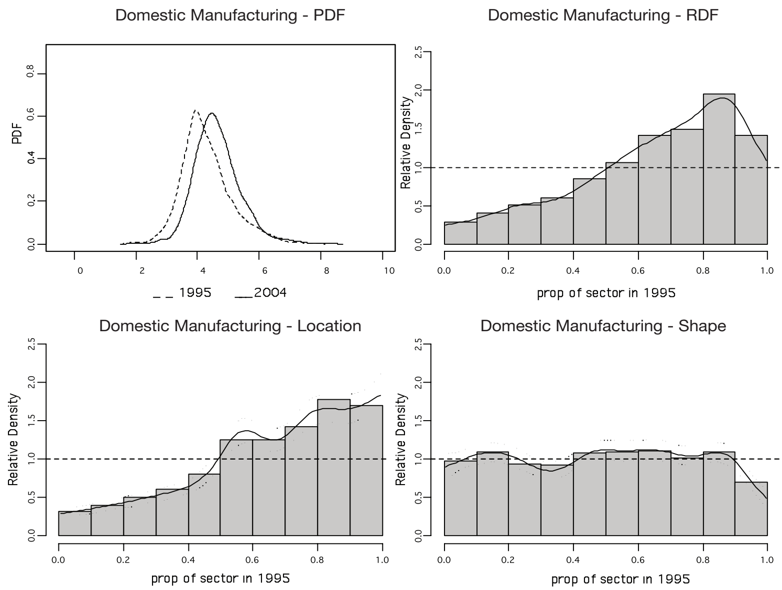


Figure 2: *Probability Density Function, Relative Density, Location and Shape Components Reflecting the Change in the Distribution of Labour Productivity from 1995 to 2004 in All Manufacturing Plants that Survived the Full 10-year Period (Reference Group is 1995)*

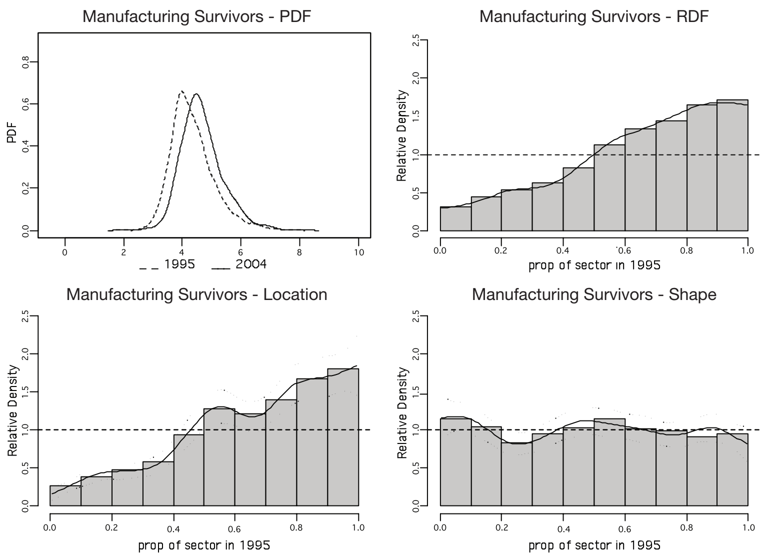


Figure 3: *Probability Density Functions, Relative Density, Location and Shape Components Reflecting the Change in the Distribution of Labour Productivity from 1995 to 2004 in all Irish-Owned Manufacturing Plants (Reference Group is 1995)*

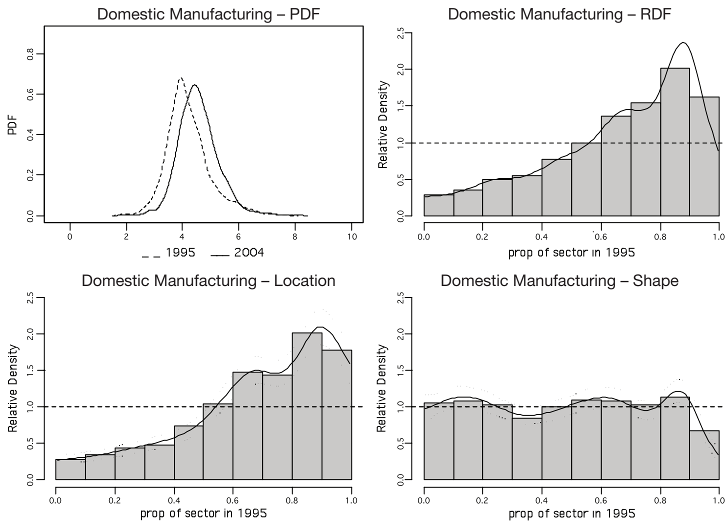


Figure 4: *Probability Density Functions, Relative Density, Location and Shape Components Reflecting the Change in the Distribution of Labour Productivity from 1995 to 2004 in all Foreign-Owned Manufacturing Plants (Reference Group is 1995)*

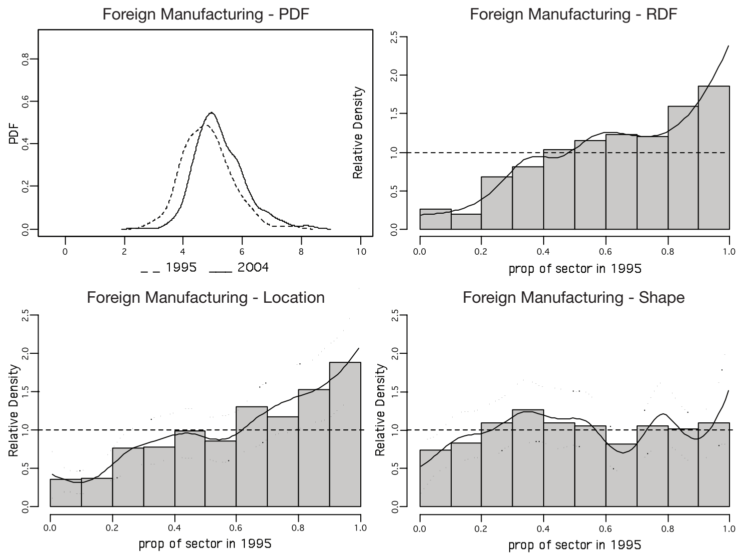


Figure 5: *Probability Density Functions, Relative Density, Location and Shape Components Reflecting the Change in the Distribution of Labour Productivity from 1995 to 2004 in Manufacturing Industries (Reference Group is 1995)*

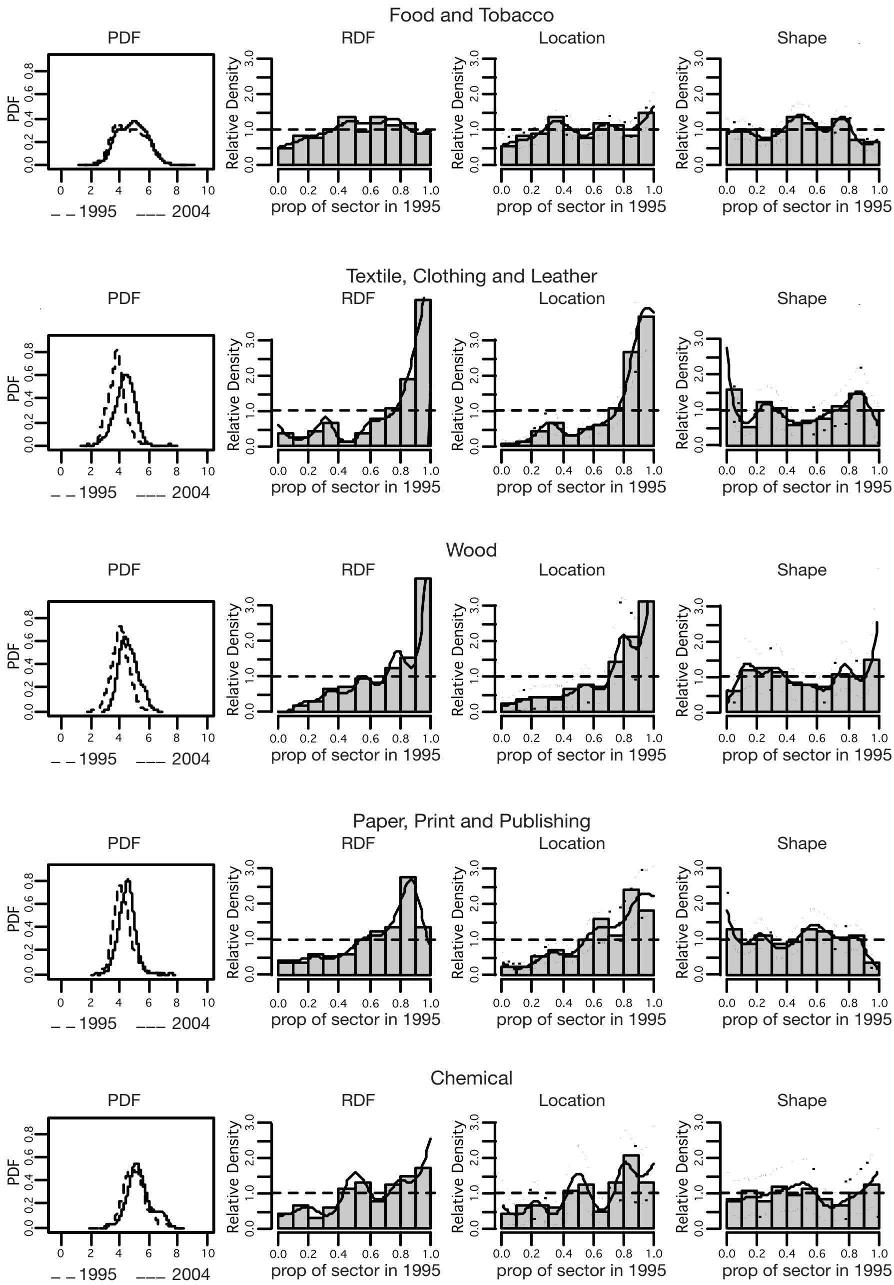
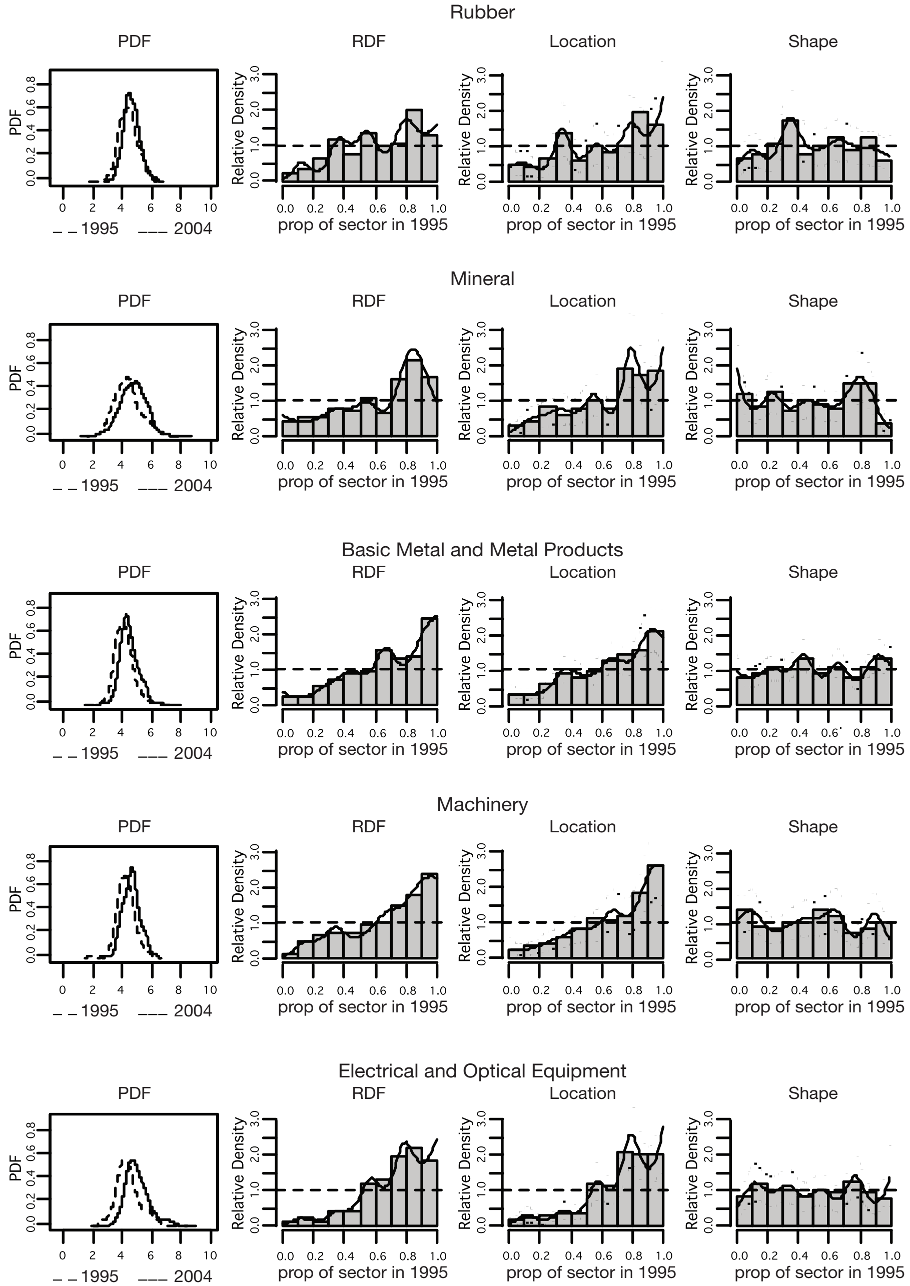


Figure 6: *Probability Density Functions, Relative Density, Location and Shape Components Reflecting the Change in the Distribution of Labour Productivity from 1995 to 2004 in Manufacturing Industries (Reference Group is 1995)*



Figures 5 and 6 illustrate the changes in distribution for most of the subsectors of manufacturing.[6] Paper, printing and publishing, rubber, minerals and electrical and optical equipment show relatively similar dynamics as the overall manufacturing distribution in that median productivity has increased, but frequencies of plants at the very top of the distribution are lower in 2004 than they were in 1995. In chemicals, basic metal and metal production and machinery the main change has been a more or less pronounced median shift, the changes in the shape of the distribution are in all cases around 10 per cent (cf. Table 4).

The two sectors that have seen the largest changes are textiles, clothing and leather (Figure 5, second line) and wood (Figure 5, third line). The number of plants in textiles, clothing and leather taken together has decreased by more than 60 per cent over the 10-year period. The distribution of labour productivity in this sector has changed from very concentrated and high-peaked in 1995 to become somewhat more dispersed at a higher median in 2004. Two changes in the external environment are the likely drivers of this degree of restructuring. One is that the Multi Fibre Agreement that restricted market access for developing country producers of textiles and clothing to developed countries was finally phased out on 1 January 2005. The other factor that has most likely contributed to the restructuring in this rather low-tech sector is the continuous increase in labour costs in absolute terms and also relative to the rest of the EU. Given that the implications of the end of the Multi Fibre Agreement did not hit fully until 2005 when imports from developing countries soared, it is more than likely that continued restructuring including exit of plants has brought about further changes in the distribution of productivity in this sector since.

In the wood sector, in contrast, the number of plants has increased and the distribution of productivity appears more concentrated in 2004 than it was in 1995. The median shift has moved plants from the lower parts of the distribution to the top three deciles relative to 1995 as can be seen from the location component, while the shape component shows that the concentration process moved plants towards the 3rd, 4th and 9th deciles. The main growth in this sector has been in the building and construction-related subsectors NACE 202 (manufacture of veneer sheets, manufacture of plywood, laminboard, particle board, fibreboard and other panels and boards) and NACE 203 (manufacture of builder's carpentry and joinery). Hence, the substantial increase in productivity and in the number of plants in this sector could be due to the construction boom over the past decade.

[6] Sectors Transport (NACE 34 and 35) and Other (NACE 36) are excluded here. The transport sector is too small to make the assumption of continuous support; and the composition of the NACE 4-digit industries captured under 'Other' (not elsewhere classified) has changed too much to allow for a sensible comparison over time.

A sector that has seen very little change, in turn, is food and tobacco; it has the lowest entropy of all sectors in Table 4. From the plot of the two density functions in the first line in Figure 5, it is apparent that the distribution has become somewhat more concentrated in the centre from 1995 to 2004. The concentration is due to a location shift that has left fewer plants in the bottom part of the distribution and a change in shape where plants have moved from the top of the distribution to the centre.

Transition Matrices and Distribution of Entrants

In order to get a better idea of how plants move within the productivity distribution, I calculate transition matrices. The transition probabilities are calculated by quintile as shares of the initial industry size at 3-digit level (where industries with less than 4 plants have been excluded). I calculate one transition matrix for the 10-year period from 1995 to 2004 (Table 5), and another one with the averages of 3-year transitions from 1995 to 2004 (Table 6).

Table 5: *Transition Matrix for All Manufacturing 1995-2004 (10-Year Period)*

			2004			
1995	*Q1*	*Q2*	*Q3*	*Q4*	*Q5*	*Exit*
Q1	18.65	12.01	7.19	4.21	1.61	56.34
Q2	16.11	18.54	12.82	7.32	1.63	43.59
Q3	10.42	16.21	14.93	10.04	5.68	42.72
Q4	10.81	8.69	17.01	13.69	10.31	39.49
Q5	2.19	4.84	6.52	17.70	22.01	46.74

Notes: Row 2 column 3 (Q1Q2) shows the fraction of plants that moved from the first quintile (defined at 3-digit industry level) in 1995 to the second quintile in 2004. 3-digit industry-year cells with less than 4 observations excluded.

As Haskel and Martin (2002) find for the UK as well, there is a high degree of persistence in the productivity distribution. From the main diagonal of Table 5 one can see that over the period between 14 and nearly 22 per cent of all plants that were in a certain quintile are still in this quintile 10 years later. Persistence is high at the bottom, but highest in the top quintile of the distribution. If plants change quintile, they are more likely to change to an adjacent quintile, but rather unlikely to move up or down more than two quintiles. For example, of those plants that were in the first quintile of the productivity distribution in 1995, only 1.61 per cent made it to the top quintile in 2004. Of the plants in the bottom quintile just over 40 per cent are still in operation 10 years later and 70 per cent of the plants that survived are still in the bottom two quintiles. In the higher quintiles the exit rate is not quite so high; although it increases again towards the top.

Table 6 shows the transition matrix based on 3-year averages over the same period. For the shorter periods persistence in the same quintile is even higher, ranging from 25 per cent in the 3rd quintile to 47 per cent in the 5th quintile. Mobility is again higher to adjacent quintiles than to quintiles further away from the initial quintile. Exit is highest at the bottom end of the distribution.[7]

Table 6: *Transition Matrix for all Manufacturing 1995-2004 (3-Year Averages)*

				2004		
1995	*Q1*	*Q2*	*Q3*	*Q4*	*Q5*	*Exit*
Q1	42.90	16.36	7.27	4.21	1.74	27.52
Q2	20.68	28.92	18.26	9.39	2.75	19.99
Q3	8.64	20.82	25.55	15.08	5.30	24.60
Q4	5.13	10.46	21.01	27.53	13.79	22.07
Q5	2.66	3.45	5.77	17.33	46.93	23.87

Notes: Row 2 column 3 (Q1Q2) shows the fraction of plants that moved from the first quintile (defined at 3-digit industry level) in t to the second quintile in t+3. Average over periods 1995-1998, 1998-2001, 2001-2004. 3-digit industry-year cells with less than 4 observations excluded.

The analysis of the relative densities suggests that the shortfall of plants at the top of the productivity distribution is due to insufficient entry in this part of the productivity distribution. This cannot be captured in the transition matrices because the transition probabilities are calculated as shares of the population in the base year and by definition the entrants are not part of the distribution in that year. To examine this further I present the proportions of entrants to the different quintiles in Table 7 distinguishing between two sub-periods and between domestic and foreign entrants. From this table it becomes clear that until 2000 there has been less than proportional entry at the top of the distribution. Only since 2001/2 this trend has reversed to more than proportional entry at the top. In particular, the domestic entrants are responsible for this trend. While the number of foreign entrants has been

[7] A transition matrix A can be shown to be generated by a long-run steady state distribution, if one can find the set of probabilities Z such that Z = AZ, i.e. the set of probabilities such that after the transition matrix is applied to them, one still obtains this same set of probabilities. However, this is only possible for symmetric transition matrices. Given that it is not possible to determine the population from which the plants currently in operation, observed entrants and potential entrants are drawn, such a steady-state distribution could at best be obtained for those plants that survive from one year to the next. I checked for existence of a long-run steady-state distribution in the 9 transition matrices that can be obtained for the subset of plants that survive from one year to the next. However, none of these fully converged to a long-run stable distribution; and those distributions that were nearly stable are not identical for the different years.

relatively constant over the period, their share in all entrants has decreased from about 9 per cent on average until 1999 to 5 per cent on average after 2000. In nearly all years 50 per cent or more of the foreign entrants enter in the top two quintiles, from 2001 it is even 60 per cent on average. In contrast, until 2000 just over 40 per cent of the new domestic plants entered into the top two quintiles, only since 2001 this figure has increased to around 50 per cent. In fact, in this period 15 per cent and more of the domestic entrants entered into the top decile. Since in addition the number of entrants per year has been lower after 2000 than before, the new high-productivity entrants have not been able to fully compensate for the plants that have exited from the top of the distribution.

Table 7: *Proportion of Entrants to Each Quintile (%)*

	Q1	*Q2*	*Q3*	*Q4*	*Q5*	*No. of Plants*
All entrants						
1996-2004	17.55	14.38	21.26	24.07	22.74	2,775
1996-2000	17.23	14.33	25.82	26.48	16.14	1,654
2001-2004	18.02	14.45	14.54	20.52	32.47	1,121
Domestic entrants						
1996-2000	17.45	14.19	26.24	26.95	15.17	1,536
2001-2004	18.53	14.23	14.80	20.53	31.90	1,047
Foreign entrants						
1996-2000	14.41	16.10	20.34	20.34	28.81	118
2001-2004	10.81	17.57	10.81	20.27	40.54	74

Note: Quintiles are calculated at 3-digit industry level based on all active plants in every year. 3-digit industry-year cells with less than 4 observations excluded.

Convergence

To examine whether there is convergence of productivity levels to an industry mean, I run Galton-Markov regressions. Defining the deviation of each plant's labour productivity from its 4-digit industry mean as $p_{it} = lp_{it} - lp_i$, the basic regression is

$$p_{it} = \alpha + \beta\, p_{it-1} + \varepsilon$$

which, if $\beta < 1$, implies convergence of plants to a mean industry productivity level plus α. It also implies that convergence is symmetric for plants below and above the mean, and that the speed of convergence is the same across industries. A more general version of the above equation is therefore

$$p_{it} = \alpha + \beta_1 p_{it-1} + D\beta_2 p_{it-1} + \varepsilon_t, \qquad D = 1 \quad if \quad p_{it-1} > \overline{p}_{it-1}$$

which has the following interpretation. The term $D\beta_2$ allows for a different (industry specific) convergence speed if the plant has previous productivity above its industry average. This is to allow plants below the mean to have a different convergence speed to those above the mean. If competition is important Oulton (1998) has argued that we expect convergence to be faster for plants below the mean and hence β_2 should be positive. The results of estimating these equations are set out in Table 8.

Table 8: *Convergence – Galton-Markov Regressions*

	(1)	*(2)*
β_1 (Lag)	0.823	0.767
	(0.005)***	(0.010)***
β_2 (Lag > Mean)		0.106
		(0.015)***
Constant	–0.004	–0.027
	(0.005)	(0.006)***
Year Dummies	yes	yes
N	34,178	34,170
Plants	6,215	6,215
Adj. R^2	.67	.67

Notes: *** indicates significance at 1 per cent. Dependent variable is the deviation of a plant's labour productivity from its 4-digit industry mean.

The simple regression in column (1) shows that the coefficient on β, is indeed smaller than 1, indicating that there is mean convergence. In column (2) I control for plants with productivity above the mean, thereby allowing plants below the mean to have a different convergence speed. The results confirm that plants with productivity below the mean converge faster. This is consistent with somewhat higher upward mobility in the bottom quintiles of the transition matrices.

The Productivity Spread and Productivity Growth

This section examines whether there is a positive correlation between different measures of the spread and productivity growth. To do this I regress

4-digit industry level average productivity growth on three different measures of the spread, namely, the standard deviation (StdDev), the spread between the 90th percentile and the median (p90/p50) and between the median and the 10th percentile (p50/p10) with and without industry dummies. The results are presented in Table 9.

Table 9: *Spreads and Productivity Growth*

Dep Variable	*Ind lp Growth*	*Ind lp Growth*
p90/p50	0.028	0.017
	0.008***	0.009*
p50/p10	–0.039	–0.026
	0.007**	0.008***
StdDev	0.026	0.002
	0.005***	0.008
4-dig Ind Dummies	no	yes
N	946	946
Adj. R^2	0.05	0.38

Note: ***, **, * indicate significance at 1, 5, 10 per cent. Unit of observation are 4-digit industry-year means.

Without the industry dummies, there is a positive correlation between a high spread in the top half of the distribution and industry-level productivity growth and a negative correlation with a high spread in the bottom half of the distribution. Also a high standard deviation is positively associated with industry-level productivity growth. When the industry dummies are included, only the negative correlation with the spread in the bottom half of the distribution retains full significance.[8] Note that the industry dummies account for quite a large share of the variation. Spreads and productivity growth may well be endogenous; the spread might widen if an industry experiences high growth and good firms do particularly well, but industries may also grow faster when spreads are wider. Thus, these regressions should be taken as indications of a correlation rather than as an attempt to uncover causality.

[8] When only the sample of domestic plants is considered, in the regression with industry dummies only the negative coefficient on p50/p10 remains marginally significant at 10 per cent.

V SUMMARY AND CONCLUSIONS

The average manufacturing plant at the 90th percentile has about 1.36 times the labour productivity of the average plant at the 10th percentile, in terms of turnover per employee this equals a factor 3.9. Spreads are driven mainly by competition. Regression analysis shows that average industry employment, concentration and profit margins are positively correlated with different measures of the productivity spread. We can conclude that over the 10-year period under consideration productivity spreads in the Irish industrial sector are an expression of a competitive process. Overall, the reallocation of resources within and between sectors seems to be efficient, if somewhat slow in some sectors.

However, comparing the distribution of labour productivity in 2004 to that in 1995, all sectors have a higher level of labour productivity at the median in 2004. The analysis shows, however, that in the absence of this median shift there would have been a smaller share of plants at the very top of the distribution in 2004 than there was in 1995. This latter observation is somewhat more pronounced if only Irish-owned plants are considered. This development is not due to plants moving down in the productivity distribution as those plants that remained in operation over the full 10 years have seen little change to their levels of productivity beyond the positive 'Celtic-tiger' shock that has benefited plants at all levels of productivity. It turns out that this shortfall of plants at the top is due to the less than proportional entry in the top deciles of the productivity distribution until 2000. While entry since 2001/2 has been more than proportionally at the top of the productivity distribution, this has not been on a large enough scale to fully replace those plants that have exited at the top of the distribution over the 10-year period.

In most industries the shape of the distribution of labour productivity does not change very much over time. After 3 years between 25 and 47 per cent of plants are still in the same quintile of the productivity distribution. Over 3 years there is a turnover of about 23 per cent of the plants in industrial production. Between 22 and 43 per cent of those plants that do not go out of business are still in the same quintile 10 years later. The analysis shows, however, that there is convergence of productivity to an industry mean, and convergence is faster for plants with labour productivity initially below the industry mean. This is consistent with a somewhat higher mobility of plants in the bottom half of the distribution as documented in the transition matrices. There is some evidence that productivity growth is hampered in industries with large spreads in the bottom half of the distribution.

REFERENCES

BAILY, MARTIN N., CHARLES HULTEN and DAVID CAMPBELL, 1992. "Productivity Dynamics in Manufacturing Plants", *Brookings Papers: Microeconomics,* pp. 187-267.

BARTELSMAN, ERIC and MARK DOMS, 2000. "Understanding Productivity: Lessons from Longitudinal Microdata", *Journal of Economic Literature,* Vol. 38, pp. 569-594.

GIRMA, SOURAFEL, GÖRG HOLGER and ERIC STROBL, 2004. "Exports, International Investment, and Plant Performance: Evidence from a Non-parametric Test", *Economics Letters,* Vol. 83, No. 3, pp. 317-324.

HANDCOCK, MARK S. and MARTINA MORRIS, 1999. *Relative Distribution Methods in the Social Sciences,* New York: Springer.

HASKEL, JONATHAN and RALF MARTIN, 2002. "The UK Manufacturing Productivity Spread", CeRiBA Discussion paper, Centre for Research into Business Activity, London.

MARTIN, RALF, 2005. "Computing the True Spread", CEP Discussion Paper 692, Centre for Economic Performance, London School of Economics, London.

OULTON, NICHOLAS, 1998. "Competition and the Dispersion of Productivity", *Oxford Economic Papers,* Vol. 50, pp. 23-38.

RUANE, FRANCES and ALI UĞUR, 2005. "Labour Productivity and Foreign Direct Investment in Irish Manufacturing Industry: A Decomposition Analysis", *The Economic and Social Review,* Vol. 36, No. 1, pp. 19-44.

APPENDIX – RELATIVE DISTRIBUTION METHODS

Denote the cumulative distribution function (CDF) of an outcome attribute for a reference group by $F_0(y)$. The CDF of the comparison group will be denoted by $F(y)$. Assuming that F and F_0 are continuous with common support, the grade transformation of Y and Y_0 is defined as the random variable

$$R = F_0(Y).$$

R is obtained from Y by transforming it by the function F_0. It is continuous with outcome space [0,1]. That is by generating a distribution that transforms the outcomes of the reference distribution to those of the comparison group (R), it is possible to compare the two distributions in common space. A realisation of R, r, is the 'relative data'. The relative data can be interpreted as the percentile rank that the original comparison value would have in the reference group. The CDF of R can then be written as

$$G(r) = F(F_0^{-1}(r)) \qquad 0 \leqslant r \leqslant 1,$$

where r represents the proportion of values, and $F_0{}^{-1}(r) = \inf_y \{y \mid F_0(y) \geqslant r\}$ is the quantile function of F_0. The probability density function (PDF) of R is

$$g(r) = \frac{f(F_0{}^{-1}(r))}{f_0(F_0{}^{-1}(r))} \qquad 0 \leqslant r \leqslant 1.$$

If the two distributions are identical, then the CDF of the relative distribution is a 45° line and the PDF of the relative distribution is the uniform PDF. The relative PDF $g(r)$ can be interpreted as a density ratio: the ratio of the fraction of the respondents in the comparison group to the fraction in the reference group at a given level of the outcome attribute $Y(F_0{}^{-1}(r))$. This can be seen more easily by expressing the relative PDF explicitly in terms of the original measurement scale, y. Denoting the r^{th} quantile of R by the value y_r on the original measurement scale, the y_r corresponding to r is $F_0{}^{-1}(r)$. The relative PDF is then

$$g(r) = \frac{f(y_r)}{f_0(y_r)} \qquad y_r \geqslant 0.$$

The relative CDF, $G(r)$, can be interpreted as the proportion of the comparison group whose attribute lies below the r^{th} quantile of the reference group.

Assuming that the comparison distribution is a simple shifted version of the reference distribution, i.e. $F(x) = F(x - c)$ or $F(x) = F(x \times c)$ for some constant c, the difference between the two distributions can be summarised by this shift. Differences that remain after a location adjustment are differences in 'shape', a general concept that comprises spread, skew, and other distributional characteristics.

Denote by Y_A a random variable describing the reference population location adjusted to have the same median as the comparison population. For an additive shift, define Y_A as the random variable $Y_0 + \rho$ where $\rho = \text{median}(Y) + \text{median}(Y_0)$. The CDF of Y_A can then be written as $F_A(y) = F_0(y + \rho)$. Y_A defines a hypothetical population that has the location (here: the median) of the comparison group but the shape of the reference group.

From these three distributions – Y, Y_A, and Y_0 – two relative distributions that represent the effects of the location and shape changes can be constructed. Let $R \equiv R_0{}^1 = F_0(Y)$ be the relative distribution of Y to Y_0. The location shift is given by the relative distribution of Y_A to Y_0, denoted by $R_0{}^A = F_0(Y_A)$. $R_0{}^A$ has a uniform distribution when the comparison and the reference group have the same location. The shape change, in turn, is given by the relative distribution of Y to Y_A, denoted $R^1{}_A = F_A(Y)$. $R^1{}_A$ has a uniform distribution when, net of location shifts, the two distributions have the same shape.

These two effects form an exact decomposition of the relative distribution of Y to Y_0 in the sense that $R^1{}_A$ is the relative distribution of $R^1{}_0$ to $R^A{}_0$. This can be expressed in terms of the density ratios from above

$$\underbrace{\frac{f(y_r)}{f_0(y_r)}}_{\substack{\textit{overall relative}\\ \textit{density}}} = \underbrace{\frac{f_A(y_r}{f_0(y_r)}}_{\substack{\textit{density ratio for}\\ \textit{location difference}}} \times \underbrace{\frac{f(y_r)}{f_A(y_r)}}_{\substack{\textit{density ratio for}\\ \textit{shape difference}}}.$$

Note that the density ratio for the location effect is a true density (i.e. it sums to 1). The density ratio for the shape effect is in general not, because of the scale change imposed by using f_A rather than f_0 as the reference distribution for $R^1{}_A$. The density ratio preserves the cut-points, y_r, so that the location and shape effects are applied at the same value of y_r.

While the location and shape components allow a very intuitive graphical analysis of the differences between the reference and the comparison group, a more formal measure of divergence between two distributions is the Kullback-Leibler divergence, given by

$$D(F; F_0) = \int_{-\infty}^{\infty} \log\left(\frac{f(x)}{f_0(x)}\right) dF(x) = \int_0^1 \log(g(r))g(r)dr.$$

It is the negative entropy[9] of the relative density. Note that $D(F; F_0)$ is always nonnegative, hence, higher values of this divergence measure imply that the reference and the comparison distribution differ by more. It can be interpreted as the expected information for discriminating g from a uniform distribution based on a single observation of R.

[9] The entropy of the event x is the sum/integral, over all possible outcomes i of x, of the product of the probability of outcome i times the log of the inverse of the probability of i (which is also called i's *surprisal* – the entropy of x is the expected value of its outcome's surprisal). Entropy measures are continuous, they are largest (but not bounded) when all the outcomes are equally likely, and the amount of entropy should be the same independenlty of how the process is regarded as being divided into parts.