Small Business Economics (2006) 26: 39–59
DOI 10.1007/s11187-004-6482-z

Sectoral Differences in Plant Start-up Size in the Finnish Economy

Satu Nurmi

ABSTRACT. New knowledge on the factors behind the choice of plant entry scale is important for understanding the entry process and the recruitment decisions of new plants in different sectors of the economy. The purpose of this paper is to evaluate the effect of various industry attributes on plant start-up size, measured with employment, using data on the entire Finnish business sector. The paper extends previous research by examining the differences between manufacturing and services and by using a richer set of explanatory variables. The relative importance of the covariates for different-sized entrants is taken into account by using a quantile regression approach. The results for manufacturing and services are remarkably similar. However, the findings imply that in the future analysis it is also important to consider the effect of the regional availability of educated and experienced work force on plant start-up size in these two sectors. The findings on the importance of scale economies and industry turbulence in determining start-up size correspond to the earlier studies. The results also show that single plant and multiplant entrants face a rather different entry environment. In addition, the employment share of foreign-owned entrants in the industry has to be taken into account.

KEY WORDS: entry, industry differences, quantile regression, start-up size.

JEL CODE: J21, J23, L11, L25

1. Introduction

New firms play a fundamental role in the economy by creating a large number of new jobs and introducing new product innovations and technology enhancements in the markets. Since entry is a response to perceived profit opportunities, new firms drive prices towards the competitive level and improve efficiency in the industry. Through displacement effects and competitive pressures entrants speed up creative destruction which is often seen as an inevitable prerequisite for the evolution of the markets.

There is a vast literature on the extent and determinants of entry and on the post-entry performance of firms and plants (Geroski, 1991; Caves, 1998). It has been found that regardless of the industry or time period, there is a considerable amount of entry in the economy. Typically, entering firms are small and a large fraction of them exit within a short period after entry. Many empirical studies have also concluded that start-up size, measured with employment, is an important determinant of the likelihood of survival facing a new firm after entry. More precisely, there is a positive relationship between start-up size and the chances of survival (Mata and Portugal, 1994; Audretsch, 1995).[1] Furthermore, entrant's size has an effect on its subsequent growth. Among others Evans (1987) and Dunne et al. (1989) have shown that small firms have higher growth, which may be due to the cost disadvantage induced by operating at a sub-optimal size.

Acs and Audretsch (1989) and Mata (1991) have taken into account the fact that small and large-scale entry may be affected by different factors by using as the dependent variable the small firm entry share or the numbers of small and large entrants. However, to my knowledge, there are only a few earlier studies examining directly the determinants of the scale at which a firm or a plant enters, namely, the studies by Mata and Machado (1996) and Görg et al. (2000). Furthermore, these studies are based solely on manufacturing data with fairly short time periods and limited coverage of explanatory variables. New

Final version accepted on June 17, 2004

Statistics Finland
00022 Statistics Finland
Finland
E-mail: satu.nurmi@stat.fi

information on the determination of plant start-up size in different sectors of the economy may be very useful for the decision making in industrial and labour market policy. Examining the choice of entry scale may also create new knowledge about the factors behind employment creation and recruitment decisions of new plants.

The purpose of this paper is, firstly, to study the industry-level and regional determinants of plant start-up size in the Finnish business sector during the years 1989–2000, i.e. including also the service sector besides manufacturing. Examining non-manufacturing sectors is important since these sectors cover a large share of all small businesses and the effect of industry characteristics on the entry scale may be quite different in this environment. Very small entrants may not respond to market conditions to the same extent as large entrants do. In addition, human capital factors may play a different role in manufacturing and services. Second, a richer data set on manufacturing, containing information on various aspects of industry performance and competitive environment, is used in order to extend the previous empirical literature. Third, the data allows the sensitivity analysis of the results to the different data sets and size thresholds used. In addition, possible business cycle effects on plant start-ups can be controlled for when panel data is used. Plant is chosen as a unit of analysis instead of firm, because decisions regarding the purchase of factors of production, including labour, are usually made at the plant level. More importantly, changes in ownership and legal status do not affect the plant identification code.

Previous studies have found that industry-level variables describing, for example, industry scale economies, growth and turbulence, are important determinants of start-up size. However, according to Mata and Machado (1996), industry attributes may have different effects on small and large-scale entrants. In addition, there are vast differences in firm start-up size even within narrowly defined markets. These findings emphasise the importance of taking into account the size-related plant-level heterogeneity in addition to the industry-level factors. However, the sources of this heterogeneity generally remain unobserved due to the lack of micro-level data that would be comprehensive enough. As a consequence, the previous studies have used quantile regression techniques where the effect of the covariates is allowed to vary across plants over the size distribution. In economics the number of empirical applications using more flexible quantile regression methods has been rapidly increasing in recent years (Koenker and Hallock, 2001).

Theoretical literature on the choice of firm start-up size is scarce. However, there are four alternative approaches explaining the determinants of optimal firm size in general (You, 1995). In the traditional technological approach firm size is determined by technical and allocational efficiency. However, this approach cannot explain the vast heterogeneity in firm size, because it relies on the concept of the representative agent. As a result, more advanced approaches have been developed. In the transaction cost approach (or institutional approach) firm size is determined by transaction cost efficiency, whereas in the industrial organisation approach size distribution is explained by market power. Lastly, the growth literature focuses on the dynamics of the size distribution of firms. This approach includes stochastic models, life-cycle models and evolutionary models of firm growth. It may be argued that the growth literature is more appropriate in explaining differences in start-up size than the theories explaining the size differences of existing firms in equilibrium.

Jovanovic's (1982) life-cycle model, for instance, assumes that firms enter at a similar scale. Since firms have no prior knowledge about their true ability, a typical entry is at a small scale. The model implies that young and small firms show on average higher and more volatile growth rates but also have lower survival rates. Survival is unaffected by start-up size but it is heavily conditioned on growth. According to Cabral (1995), firms enter at a small scale especially if entry involves sunk costs which cannot be recovered in the case of failure. It should be noted that the determinants of entry size may be to some extent related to the determinants of entry, which are described in the theoretical models on firm entry. It has been found, for example, that different-sized entrants respond to entry barriers in different ways (Mata, 1991).

The paper is organised as follows. Section 2 presents the data used and gives some evidence on the inter-industry differences in plant start-up size and average plant size. The empirical framework and explanatory variables related to start-up size are introduced in Section 3. Section 4 presents the empirical results including a comparison of the results for manufacturing and services, and an extended model for manufacturing. Finally, Section 5 concludes and discusses the possibilities for further research.

2. The size of entering plants

2.1. *The data*

The first data source used is the Business Register (BR) of Statistics Finland, which allows the inclusion of other sectors of the economy besides manufacturing in the analysis. The BR covers registered employers and enterprises subject to value added tax in Finland annually over the period 1988–2001.[2] In addition, plant-level data is available. A plant or an establishment is defined as a production unit of an enterprise that produces as similar goods or services as possible, and usually operates at a single location. Only active production plants are included in the analysis, not for example, headquarters or service units. The BR contains basic information on plant location, industry, ownership, sales and employment. The number of employees comprises paid employees and entrepreneurs. The number is given in terms of full-time workers. Information on average employee characteristics in each plant over the period 1988–2000 can be linked to the BR from the PESA data (Plant-level Employment Statistics Data on Average Characteristics) formed by linking the Business Register and the Employment Statistics of Statistics Finland (Ilmakunnas et al., 2001).

However, there are some problems with longitudinal linkages in the BR, due to changes in taxation and statistical practices. These statistical reforms have had an effect especially on the coverage of the smallest plants in a few particular years. Therefore, in order to produce as reliable a series as possible, only plants with at least three employees are included.[3] An entry or a start-up is defined according to the year when a plant appears for the first time in the BR during the period 1988–2000, so entries in 1988 cannot be identified. Hence, the interest is on genuinely new plants but it is possible that they are new subsidiaries of existing firms. The actual birth dates for a firm cannot be identified due to possible changes in ownership structure, whereas plant identification code does not change after entry. Diversified entry, i.e. changing industry, is not considered as a start-up when the data comes from the BR. In addition, it should be noted that due to the cut-off limit, these plants may have existed before the first observation with less than three employees. Entry is thus actually defined according to the time when a plant reaches the size of three employees, which is treated as the plant's birth year. As a consequence, the period analysed covers the start-ups during the period 1989–2000, and the final data set used in the comparison of manufacturing and services consists of 69 322 observations on plant start-ups.[4]

The second data source used in this study is the Longitudinal Data on Plants in Manufacturing (LDPM) of Statistics Finland, which is based on the annual Industrial Statistics surveys over the period 1974–2001 (Ilmakunnas et al., 2001). The Industrial Statistics covers, in principle, all Finnish manufacturing plants with five employees or more. Smaller plants are included only if their turnover corresponds to the average turnover in firms with 5–10 employees. However, over the period 1995–2001 the sample is smaller, i.e. only plants that belong to firms with at least 20 persons are included. Therefore, these years cannot be included in this analysis. The information content of the LDPM is more extensive, i.e. it contains information on various plant-level variables, including employment, output, value added and capital stock.[5] The employment figures are reported as annual averages. The number of hours worked are also reported.[6]

A plant is considered as a start-up when it appears for the first time in the LDPM during the period 1974–1994. Hence, also transitions from outside manufacturing are now considered as entries. However, these broad industry changes are probably not very common. In addition, the fact that only those plants that do not appear in the files during the whole period 1974–1980 are defined as entrants increases the confidence that entrants really are genuinely new plants in the data. The size cut-off of five employees has also been used by Mata and Machado (1996) and Görg et al. (2000). Subsequently, the period analysed covers the start-ups during the period 1981–1994, and the final data set used in the extended model for manufacturing consists of 4,739 observations on plant start-ups.

Table I describes the size distribution of new plants in both data sets separately for manufacturing and services. Plant size is measured with the number of employees. According to the BR data, the number of entries is exceptionally high

TABLE I
Descriptive statistics for plant start-up size 1989–2000 (BR) and 1981–1994 (LDPM)

Year	N	Mean	SD	Median	Min	Max	Skewness	Kurtosis
Manufacturing (the BR data)								
1989	1262	10.96	41.25	4.10	3.00	998.00	16.32	334.23
1990	1156	8.90	23.88	4.00	3.00	400.00	10.08	131.44
1991	690	10.09	24.14	4.10	3.00	276.00	7.60	71.77
1992	585	10.97	30.02	4.30	3.00	415.00	9.16	103.12
1993	978	8.66	23.24	4.10	3.00	367.00	9.57	113.04
1994	820	8.47	19.70	4.10	3.00	306.00	8.71	99.45
1995	856	10.72	36.62	4.00	3.00	555.70	10.01	120.74
1996	817	11.02	43.91	4.00	3.00	857.10	12.37	197.53
1997	821	8.23	18.67	4.00	3.00	274.50	9.31	108.48
1998	637	10.07	30.08	3.90	3.00	568.00	12.08	198.6
1999	541	13.11	45.67	3.90	3.00	592.10	9.44	105.36
2000	651	13.03	44.95	4.80	3.00	779.80	11.77	169.15
All	9814	10.17	32.90	4.00	3.00	998.00	13.54	260.32
Services (the BR data)								
1989	5312	8.41	28.95	3.80	3.00	1300.00	22.78	825.78
1990	5874	7.75	19.55	3.80	3.00	568.50	14.21	285.3
1991	3445	7.50	21.58	3.90	3.00	770.00	21.64	640.6
1992	2974	8.36	16.20	4.00	3.00	241.00	7.48	76.72
1993	3602	6.17	10.43	3.90	3.00	231.00	10.68	168.5
1994	3959	6.33	11.12	4.00	3.00	264.00	11.03	175.73
1995	3908	8.12	37.78	3.80	3.00	1222.00	21.97	589.59
1996	3836	6.01	10.02	3.70	3.00	220.00	9.20	121.77
1997	4007	6.74	20.97	3.80	3.00	658.60	19.73	506.51
1998	3774	6.46	13.12	3.70	3.00	355.10	13.99	276.91
1999	3675	7.88	20.62	3.90	3.00	685.50	15.71	385.8
2000	3977	8.03	17.31	4.00	3.00	355.00	10.52	153.64
All	48343	7.35	20.90	3.80	3.00	1300.00	24.39	1006.97
Manufacturing (the LDPM data)								
1981	299	21.78	33.55	12.00	5.00	321.00	5.40	38.25
1982	775	18.93	47.06	10.00	5.00	746.00	10.83	141.63
1983	280	19.76	29.79	11.00	5.00	336.00	6.37	54.51
1984	360	22.08	36.15	11.00	5.00	313.00	5.05	29.87
1985	201	31.86	59.28	14.00	5.00	489.00	4.87	28.55
1986	222	31.49	53.78	16.00	5.00	508.00	5.12	33.38
1987	354	31.46	78.01	13.00	5.00	744.00	6.66	50.27
1988	243	29.65	70.81	12.00	5.00	824.00	8.07	78.69
1989	256	42.82	105.00	14.00	5.00	995.00	5.58	37.39
1990	271	35.58	57.52	16.00	5.00	548.00	4.72	30.55
1991	825	19.07	32.97	11.00	5.00	586.00	9.01	120.69
1992	270	24.94	57.10	12.00	5.00	813.00	10.45	137.11
1993	207	19.86	31.12	11.00	5.00	287.00	5.53	37.25
1994	176	23.89	32.05	13.00	5.00	237.00	3.88	19.07
All	4739	24.86	53.48	12.00	5.00	995.00	8.51	97.86

Note: [1] Mean start-up size without the largest outliers (start-ups with more than 500 employees) was 9.56 employees in manufacturing and 7.12 employees in services using the BR data. The corresponding value for the LDPM data was 22.89 employees.

in 1989 and 1990 in both sectors, whereas it has decreased during the recession years at the beginning of the 1990s.[7] As expected the mean size of an entering plant is higher in manufacturing than in services. In addition, the standard deviation of size is clearly higher in manufacturing.[8] However, the quantile regression method used is not very sensitive to outliers. Both the high coefficient of skewness and the median which is lower than the mean imply that the size distribution is highly skewed to the right.

It can be seen that the number of entries in the LDPM data has been higher than usual in 1982 and 1991. The latter can be explained by the increased number of small plants in the Industrial Statistics sample in 1991.[9] The average start-up size is considerably higher than in the BR data, 24.86 employees with a standard deviation of 53.48. Naturally, this is mainly due to the higher size cut-off limit.[10] The size distribution is still clearly dominated by the small plants. Median is higher than in the BR data but notably lower than the mean. Differences in the largest start-up sizes in the two data sets are probably due to different definitions of entry.

2.2. *Start-up size by industry*

There are large inter-industry differences in plant size due to different roles played by scale economies, capital intensity and technology in different sectors of the economy. These differences in optimal size have an effect on the typical start-up size in the industry. Table II lists 2-digit industries in the order of average start-up size during the period 1989–2000 using the BR data.[11] The average start-up size can be compared with the average size of plants in each industry. The table reports also the relative size of entrants with respect to all plants and the number of start-ups and plants in the industry. Average start-up size varies from 37.1 employees in the manufacture of pulp and paper to three employees in tobacco products. The average plant size varies from 241.6 employees in tobacco products to 6.8 employees in collection, purification and distribution of water. At the 3-digit industry level used in the regression analysis the variation is likely to be even higher.

The relative size of entrants with respect to all plants is on average 0.43 in the total business sector. This reflects the fact that the entrants' share of the number of plants in the sample is 10.8%, whereas the employment share of new entrants is only 4.6%. On the other hand, the small plants' share of all entering plants is 99% when plant is considered to be small if it has less than 100 employees. The small entrants' share of all new employment is 83%. The last column in Table II gives the entry penetration rate in each industry measured as the share of employment in start-ups relative to total industry employment. It can been seen that in most industries entrants'share of employment is very low. There are only a few industries with entry penetration rate higher than 10%. This corresponds to the earlier findings on entry penetration (see e.g. Geroski, 1991, pp. 16–23).

Figure 1 presents the evolution of average plant start-up size in relation to the average size of all plants over the period 1989–2000, separately for the whole business sector, manufacturing and services.[12] The figure shows that the relative size of entrants in manufacturing is considerable lower than the relative start-up size in the service sector. This reflects the fact that most of the service sector plants are rather small. In addition, the variation of relative start-up size is slightly higher in services. There are some peaks in the relative start-up size, i.e. in 1992 and 1995. In addition, at the end of 1990s there is a clear upward trend in all sectors. These changes may be due to changes in the statistical system or the effect of a temporary increase in the number of larger start-ups.[13]

The relative size of entrants in manufacturing over the period 1981–1994 based on the LDPM data is presented in Figure 2. The earlier years show that there has been an increasing trend in the relative start-up size during the 1980s, which is mostly due to two large jumps in 1985 and 1989. However, this figure gives a completely different picture of the relative start-up size in manufacturing over the years 1989–1994. Contrary to the rather stable pattern of the relative entry size around 25% in Figure 1, Figure 2 shows a relative entry size that is highly fluctuating between 30 and 60%. It should be noted that there are many differences in these two data sets when it comes to the coverage of plants in terms of size threshold, definition of entry and differential

TABLE II
Average start-up size, average size, relative start-up size, the number of start-ups and plants, and entry penetration by industry (averages 1989–2000)

SIC	Industry description (manufacturing 15–37)	Average start-up size	Average plant size	Relative start-up size	Number of start-ups	Number of plants	Entry penetration
21	Pulp and paper	37.05	181.55	0.20	128	2338	0.01
32	Radio, television and communication equipment	35.21	108.68	0.32	206	1977	0.03
61	Water transport	29.71	86.06	0.34	131	1218	0.03
27	Basic metals	28.61	138.40	0.21	87	1368	0.01
40	Electricity, gas etc.	26.85	29.59	0.90	240	5140	0.04
23	Coke, refined petroleum products and nuclear fuel	23.50	198.63	0.12	4	131	0.01
31	Electrical machinery	22.36	52.62	0.41	276	3459	0.04
30	Office machinery and computers	19.89	78.87	0.22	31	293	0.02
64	Post and telecommunication	18.93	22.47	0.85	2708	21683	0.10
35	Other transport equipment	16.65	86.20	0.19	249	2192	0.02
24	Chemicals	16.49	75.07	0.22	156	2762	0.01
66	Insurance and pension funding	16.25	22.34	0.73	305	5721	0.04
13	Mining of metal ores	12.06	113.33	0.11	6	79	0.03
29	Machinery and equipment n.e.c.	11.75	40.57	0.29	1417	15248	0.03
73	Research and development	10.45	38.87	0.28	114	583	0.06
72	Computer and related activities	10.32	19.74	0.52	1829	9476	0.10
25	Rubber and plastic	10.09	35.15	0.28	331	4636	0.02
63	Supporting transport activities, travel agencies	9.82	20.78	0.46	1192	11094	0.05
26	Other non-metallic mineral products	9.51	29.54	0.32	431	5897	0.02
19	Leather	9.40	28.48	0.33	84	1453	0.02
15	Food and beverages	9.23	40.77	0.23	707	11727	0.01
67	Activities auxiliary to financial intermediation	9.01	15.23	0.59	275	944	0.24
65	Financial intermediation	8.78	18.59	0.47	741	23944	0.02
22	Publishing and printing	8.29	23.79	0.35	1135	14729	0.03
62	Air transport	7.91	92.81	0.08	57	666	0.01
33	Medical instruments etc.	7.77	28.58	0.27	296	3275	0.02
74	Other business activities	7.45	14.60	0.51	8888	54576	0.08
60	Land transport	7.39	13.27	0.55	3918	34466	0.07
34	Transport equipment	7.30	43.71	0.17	128	1912	0.01
28	Fabricated metal products	6.80	18.74	0.36	1884	19037	0.04
18	Wearing apparel	6.65	29.13	0.23	287	3693	0.02
20	Wood	6.55	30.28	0.22	1005	10659	0.02

45	Construction	6.33	14.00	0.46	10587	73907	0.07
51	Wholesale trade	6.33	14.07	0.45	6904	57463	0.05
70	Real estate activities	6.24	10.62	0.59	2577	16095	0.09
17	Textiles	6.17	27.65	0.22	245	3377	0.02
36	Furniture	6.05	20.38	0.30	698	8341	0.02
37	Recycling	5.99	7.86	0.76	28	221	0.10
14	Other mining and quarrying	5.80	15.00	0.38	157	1781	0.03
55	Hotels and restaurants	5.75	9.60	0.60	5649	49819	0.07
52	Retail trade	5.61	9.04	0.62	9244	107953	0.05
41	Collection, purification and distribution of water	5.40	6.81	0.80	22	466	0.05
50	Sale and repair of motor vehicles	4.95	8.17	0.61	3327	40089	0.05
71	Renting of machinery and equipment	4.48	7.05	0.64	484	2787	0.11
10	Mining of coal and lignite	4.22	8.04	0.53	153	925	0.09
16	Tobacco products	3.00	241.64	0.02	1	40	0.00
10–74	Total business sector	7.63	17.91	0.43	69322	639640	0.05

treatment of small multiplants in the LDPM.[14] This leads to a larger number of observations in the BR in most years. However, the higher cut-off for the LDPM sample alone does not explain the peaks in 1989 and 1990. The peaks may also be partly due to an exceptionally large number of very large start-ups in these years. Even though these differences cannot be fully explained, this finding puts special emphasis on the importance of taking into account the effect of data set used on the results.

3. The determinants of plant start-up size

3.1. *Industry attributes and start-up size*

The data content in the BR allows the use of similar covariates as the earlier studies. As an extension, the results for manufacturing and services can be compared. Furthermore, the model can be extended by using additional covariates. The BR data allows the inclusion of human capital variables, whereas the richer LDPM data set allows the construction of variables describing industry-level profitability, capital intensity and investments. In addition, industry-level data on R&D activity and foreign trade is available for manufacturing.

Variables describing structural barriers to entry and the possibility of strategic action of incumbents against the entrants are the most important deterrents of entry, whereas higher industry profits and demand attract more entry. These factors are likely to affect entry differently depending on the scale of entry. Earlier studies by Mata and Machado (1996) and Görg et al. (2000) have characterised some important relationships between industry attributes and start-up size. Scale economies are clearly an important determinant for plant start-up size. Higher Minimum Efficient Scale (MES) increases the cost disadvantage faced by sub-optimal entrants and worsens their competitive position. Hence, higher MES is expected to increase the start-up size. Mata and Machado have found that large firms are created only in those sectors which are characterised by considerable economies of scale but at the same time smaller firms are not deterred from entering the market.

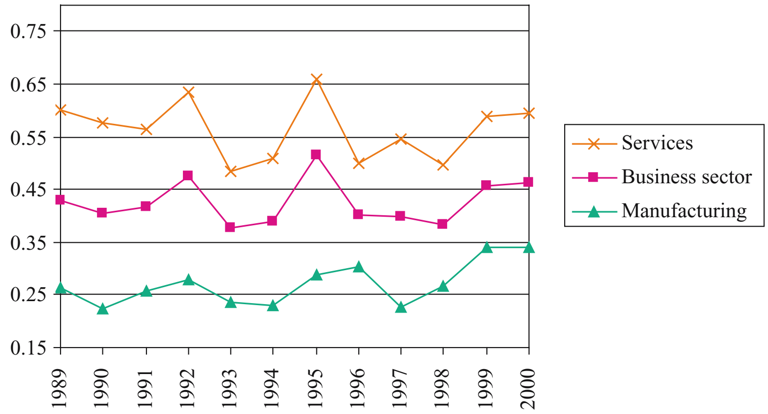

Figure 1. The relative size of entrants in broad sectors 1989–2000 (the BR data).

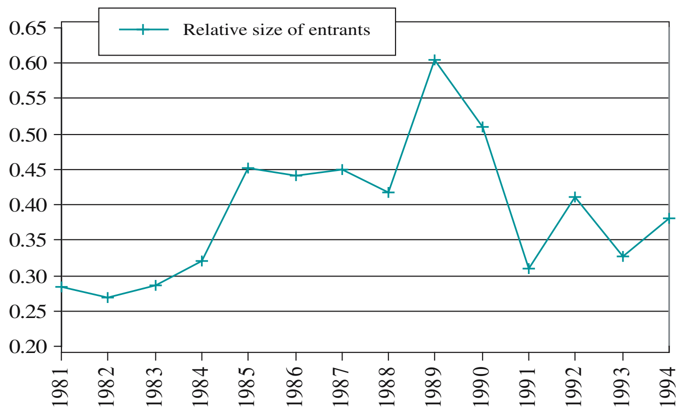

Figure 2. The relative size of entrants in manufacturing 1981–1994 (the LDPM data).

However, in markets where a large number of plants operate below MES small entrants may have better chances of being successful. They may overcome the scale-related disadvantages by, for example, occupying a niche with a differentiated product or taking advantage of the more flexible organisational structure small units often have. Sub-optimal scale, which is measured as the proportion of industry employment in plants smaller than MES, describes this within industry heterogeneity. A large number of plants operating below MES may simply reflect the fact that large and small plants do not compete with each other, so small entrants do not necessarily suffer from any cost disadvantage resulting from inefficient size. Thus, a negative relationship is expected between sub-optimal scale and start-up size.[15]

Choosing to enter at a small scale, entrants intend to reduce the likelihood of aggressive response from incumbents because the payoff of retaliatory behaviour against small entrants is likely to be small. In large markets the competitive and retaliatory action against entrants is likely to be less severe. Furthermore, the potential for increasing market share in the future will be higher. Subsequently, for a given MES, the larger the market size measured with the natural logarithm of industry employment, the larger the expected entry scale. If the markets are growing fast, the likelihood of survival is greater and the reactions from incumbents are weaker, so a higher start-up size is possible. Industry growth is measured as the logarithmic difference in the industry employment between consecutive years. Similarly to earlier studies, industry turbulence is used as an indirect measure for sunk costs faced by entrants. Mata and Machado suggest measuring industry turbulence as the product (instead of the sum) of employment weighted entry and exit rates. The advantage of this new measure is that it attains high values only if both entry and exit are high suggesting that the sunk costs in the industry are low. Since low sunk costs reduce the losses in case of failure and increase the benefits of exploiting the existing scale economies, industry turbulence is expected to be positively related to the scale of entrants.

The regional availability of experienced and educated work force may be important for the choice of start-up scale. Increased supply of high-skilled workers may improve the entrant's opportunity to raise productivity through learning by doing and R&D, which increases the chances of survival and thus start-up size. Hence, the previous studies could be extended by controlling for the regional composition of work force, which could be measured with the average education and experience of the employees and the share of women in each region.

In the second section the focus is on manufacturing only. This allows the introduction of various new variables describing the ownership structure of entering plants as well as industry structure and performance. The choice of start-up scale reflects differences in perceived ability and expectations about efficiency and profitability. Plants belonging to firms with more than one plant may gain from the previous experience, accumulated managerial knowledge, and financial or distribution channels of their parents. This probably makes them more certain about their future performance, which may result in larger start-up size. Hence, the employment share of entering multiplants in the industry is also controlled for. Correspondingly, higher employment share of entering foreign-owned plants may increase the start-up size. It can be argued that foreign-owned firms have better initial conditions and more funding available, which allows larger scale of operations at start-up. On the other hand, the presence of new foreign-owned firms in the markets may have an impact on the entry size of domestic firms through introducing higher levels of competition, lower prices and increasing demand in the markets.[16]

Industry profitability is described with the price-cost margin, which is measured as the difference between the industry value added and costs of labour and raw materials in relation to the industry value added. Thus, it captures differences in technology, demand conditions and the intensity of competition between industries, and is expected to be positively related to entry in general. However, small-scale entry may be less affected by industry profitability because small entrants often rely on new product innovations or market niches. The standard deviation of price-cost margins in the industry is included as a measure of market heterogeneity, which may increase the entry scale as there are higher profit opportunities.

High capital requirements impede especially small-scale entry, so higher capital intensity in the industry is expected to increase start-up size. Investment rate measures capacity expansion in the industry and can be seen as a sign of strategic action against new entrants, which also decreases the expected start-up size. R&D intensity, measured as R&D expenditures per employee, may be thought of as measuring product differentiation effort which is directed towards deterring entry. This may result in entrants preferring smaller start-up size. However, since small plants have an evident scale disadvantage in innovative activity, high R&D intensity may also increase the entry scale. In markets with high export intensity growth potential is greater and economies of scale can be exploited more easily. As a result, a higher start-up size can be expected. On the other hand, import penetration measures the extent of foreign competition in the industry that may decrease the chances for survival of inefficiently scaled entrants.

3.2. *Quantile regression approach*

Ordinary least squares estimation is restricted to measuring the effect of covariates on the conditional mean of the dependent variable. It does not take into account the plant-level heterogeneity which is likely to have an effect on the estimated coefficients. It is very likely that the effect of the industry-level covariates on start-up size will vary across different-sized plants in the industry. The regression quantiles estimator developed by Koenker and Bassett (1978) allows for this neglected heterogeneity and enables the evaluation of the determinants of plant start-up size at various quantiles of the plant size distribution. Thus, it provides a more complete picture of the relationship between industry characteristics and plant start-up size. In addition, quantile regression is robust to outliers and deviations from normality, and equivariant to monotonic linear and non-linear transformations of the dependent variable.

The most common form of quantile regression models is median regression, where the objective is to estimate the median of the dependent variable, conditional on the values of the independent variables. Median regression minimises the sum of the absolute residuals rather than the sum of squared residuals. Generalised quantile regression allows the estimation of the other quantiles of the distribution.

The quantile regression model can be viewed as a location model. Let $\{y_i{:}i = 1, \ldots, n\}$ be a random sample on a random variable Y having distribution function F. Then the θth sample quantile, $0 < \theta < 1$, may be defined as any solution to the minimisation problem:

$$\min_{b \in R} \left\{ \sum_{i:y_i \geqslant b} \theta |y_i - b| + \sum_{i:y_i < b} (1-\theta)|y_i - b| \right\}.$$

The analogue of the linear model for the θth quantile is defined in a similar manner. Let $\{x_i : i = 1, \ldots, n\}$ denote a sequence of (row) K-vectors and suppose $\{y_i : i = 1, \ldots, n\}$ is a random sample on the regression process $u_i = y_i - x_i'\beta$ having distribution function F. Then the θth regression quantile, $0 < \theta < 1$, is defined as any solution to the minimisation problem:

$$\min_{b \in R^K} \left\{ \sum_{i:y_i \geqslant x_i'b} \theta \left|y_i - x_i'b\right| + \sum_{i:y_i < x_i'b} (1-\theta)\left|y_i - x_i'b\right| \right\}.$$

4. Results

4.1. *Comparison of manufacturing and services*

There may be vast differences between the patterns of start-up size in manufacturing and the service sector. One of the main differences between manufacturing and the service sector is the absence of scale economies in services. As a consequence, in services a lower start-up size is possible. This is consistent with the observation of mean plant size of 13.0 employees for services and 38.2 employees for manufacturing. Instead of scale economies, the entry process in services may depend more on other industry-level and regional characteristics.

In order to compare the results for manufacturing and services with the previous results by Mata and Machado (1996) and Görg et al. (2000), a regression model which is comparable to the earlier studies is first estimated. Hence, the model includes as explanatory variables MES, sub-optimal scale, industry size, industry growth and turbulence. It has been suggested that the size distribution is approximately lognormal, although empirical results do not always support this assumption. In addition, there are some extreme observations in the data and the differences in start-up size between manufacturing and services may be large. Hence, in the regressions the logarithm of employment during the year of entry is preferred as a measure for plant start-up size.

The definitions of the variables used in the paper are described more thoroughly in Table III. In the regression analysis industry attributes are calculated at the 3-digit industry level using the Standard Industrial Classification (SIC) adopted in 1995, which amounts to having information on 181 industries (101 manufacturing industries). All explanatory variables are measured at the year of start-up. The descriptive statistics for manufacturing and services are reported in Table IV. As expected, the average start-up size in services is lower than in manufacturing, both according to the mean and the median. Furthermore, the optimal scale according to MES is considerably higher in manufacturing than services, whereas the proportion of sub-optimal scaled plants is higher in services. Industry turbulence is somewhat higher in services.

Table V shows the OLS and quantile regression estimates for manufacturing using the BR data with size cut-off of three employees. All estimations include time dummies in order to control for the cyclical effects.[17] Moulton (1990)

TABLE III
Variable definitions

Variable	Description
Start-up size	Logarithm of the number of employees at start-up
Scale economies	Mean size of the largest plants in each industry accounting for one half of the industry's employment (Comanor and Wilson, 1967)
Sub-optimal scale	Proportion of industry's employment in plants smaller than MES
Small plant cost disadvantage	(Small plant (with less than industry average number of employees) sales or shipments/ Employment)/(Industry sales or shipments/Employment)
Industry size	Logarithm of the number of employees in the industry
Industry growth	Logarithmic change in industry employment in two consecutive periods
Turbulence	(Employment of entrants/Total number of employees in the industry) * (Employment of exits/ Total number of employees in the industry)
Average experience	Average seniority (=number of months in the firm) of the employees in the region in the previous year
Average education	Average number of schooling years in the region in the previous year
Share of women	Average share of women employment in the region in the previous year
Share of multiplant entrants	Share of employment in entering plants that belong to firms with more than one plant
Share of foreign-owned entrants	Share of employment in entering plants where in excess of 50% of the equity is held by non-Finnish residents
Profitability	Price-cost margin = (Value added – Wages – Materials)/Value added in the industry
Variability of profits	Standard deviation of price-cost margins in the industry
Capital intensity	Industry capital–labour ratio
Investment rate	Gross investments/Gross output in the industry
R&D intensity	R&D expenditures/Number of employees in the industry
Export intensity	Exports/Gross output in the industry
Import penetration	Imports/(Gross output + imports – exports) in the industry
Year dummies	Dummy variables for different entry years (reference groups are the first years in the analyses, i.e. 1981 and 1989)

TABLE IV
Descriptive statistics for manufacturing and services using the BR data

Variable	N	Mean	SD	Median	Min	Max
Manufacturing						
Start-up size	9814	1.690	0.779	1.386	1.099	6.906
Scale economies	9814	0.064	0.072	0.048	0.004	1.202
Sub-optimal scale	9814	0.397	0.078	0.387	0.000	0.757
Industry size	9814	8.757	1.069	9.131	1.435	10.366
Industry growth	9797	0.003	0.152	−0.002	−1.211	1.215
Turbulence	9814	0.004	0.017	0.001	0.000	0.458
Average experience	9814	0.124	0.015	0.121	0.091	0.172
Average education	9814	11.740	0.345	11.717	11.036	12.641
Share of women	9814	0.407	0.035	0.409	0.307	0.494
Services						
Start-up size	48343	1.570	0.654	1.335	1.099	7.170
Scale economies	48343	0.022	0.019	0.018	0.003	0.726
Sub-optimal scale	48343	0.519	0.083	0.510	0.000	0.821
Industry size	48343	9.424	1.039	9.686	1.194	10.857
Industry growth	48196	0.023	0.122	0.027	−1.010	1.067
Turbulence	48343	0.013	0.029	0.005	0.000	0.927
Average experience	48343	0.121	0.016	0.118	0.091	0.172
Average education	48343	11.848	0.392	11.798	11.036	12.641
Share of women	48343	0.413	0.036	0.419	0.307	0.494

Note: [1] Scale economies and average experience were divided by 1 000 to generate easily readable coefficients.

TABLE V
Quantile regression results for manufacturing and services using the BR data

Quantiles	OLS	0.15	0.25	0.50	0.75	0.90
Manufacturing (SIC 15–37, 9 797 observations)						
Scale	2.410	0.185	0.532	1.666	5.431	7.249
economies	(0.338)***	(0.082)**	(0.142)***	(0.405)***	(0.678)***	(1.562)***
Sub-optimal	0.049	0.036	0.072	0.210	0.769	0.217
scale	(0.176)	(0.024)	(0.053)	(0.136)	(0.279)***	(0.712)
Industry	0.025	0.001	0.004	0.010	0.027	0.053
size	(0.013)*	(0.001)	(0.003)	(0.007)	(0.019)	(0.039)
Industry	0.007	0.000	0.008	0.019	−0.047	−0.212
growth	(0.088)	(0.011)	(0.022)	(0.069)	(0.159)	(0.252)
Turbulence	6.372	1.693	3.812	7.067	5.724	20.492
	(0.701)***	(1.126)	(1.263)***	(1.472)***	(4.370)	(9.595)**
Constant	1.318	1.086	1.084	1.120	1.063	1.700
	(0.165)***	(0.023)***	(0.042)***	(0.113)***	(0.235)***	(0.573)***
R^2	0.075					
F (H_0: $\beta_i = 0$)	22.16***	306.39***	214.55***	102.22***	148.18***	82.73***
Services (SIC 50–74, 48 196 observations)						
Scale	7.223	0.442	2.340	6.996	17.028	22.344
economies	(1.770)***	(0.358)	(1.015)**	(2.100)***	(4.445)***	(5.531)***
Sub-optimal	−0.419	0.021	0.016	0.104	0.170	−0.841
scale	(0.282)	(0.021)	(0.130)	(0.260)	(0.554)	(0.861)
Industry	0.018	0.000	0.000	0.005	0.015	0.031
size	(0.016)	(0.000)	(0.004)	(0.011)	(0.023)	(0.040)
Industry	0.043	0.001	0.034	0.048	−0.032	0.143
growth	(0.085)	(0.005)	(0.023)	(0.052)	(0.119)	(0.213)
Turbulence	2.625	0.199	0.762	1.638	3.670	3.779
	(0.674)***	(0.152)	(0.303)**	(0.669)**	(1.270)***	(1.352)***
Constant	1.435	1.077	1.061	1.056	1.086	2.009
	(0.208)***	(0.019)***	(0.101)***	(0.200)***	(0.390)***	(0.648)***
R^2	0.073					
F (H_0: $\beta_i = 0$)	15.04***	4.88	145.98***	99.66***	127.34***	185.44***

Notes: [1] All estimations include time dummies. [2] Robust standard errors in parantheses. [3] ***, ** and * indicate significant at 1%, 5% and 10% level, respectively.

showed that standard errors can be seriously biased downwards when the effects of aggregate explanatory variables on micro-level variables are estimated, because the disturbance terms are correlated within groups. Therefore, heteroscedasticity-consistent estimates of the standard errors with clustering on industry have been calculated for the OLS regressions (White, 1982).[18] Similarly, for the quantile regression results the estimated standard errors have been derived by bootstrapping (using 300 replications) and using clustering on industry, which allows observations to be independent across industries, but not necessarily within industries.[19] According to Buchinsky (1998), bootstrapping is an appropriate and well-behaving method for quantile regressions. Especially in cases of heteroscedastic errors the estimated standard errors can be understated if bootstrapping is not used. Following Mata and Machado the lowest quantile estimated is 0.15 due to the low variability of start-up size at the left tail of the size distribution.

The results for manufacturing show that scale economies have a highly significant, positive effect on start-up size and the effect increases towards the higher quantiles. However, according to the OLS estimates the coefficient of sub-optimal scale is positive, but not significant. Surprisingly, it also turns out to be positive and increasing in the quantile estimations, except for the highest quantile. This contradicts the earlier results by Mata and Machado and Görg et al. who find that the effect of sub-optimal scale on start-up size is negative for all quantiles.

However, the result may be affected by the high correlation between MES and sub-optimal scale (−0.53 in the total sample).[20] As a consequence, another variable describing small plant (with less than industry average number of employees) cost disadvantage, measured as the ratio of small plant productivity to the industry average productivity, was used alternatively, but this covariate was insignificant in all quantiles.

Industry size is positive and increasing in all estimations, but significant only in OLS. Previous studies have also found that industry size does

TABLE VI
Quantile regression results on the effects of human capital in manufacturing and services

Quantiles	OLS	0.15	0.25	0.50	0.75	0.90
Manufacturing (SIC 15–37, 9 797 observations)						
Scale	2.404	0.185	0.550	1.656	5.367	7.810
economies	(0.273)***	(0.085)**	(0.094)***	(0.352)***	(0.643)***	(1.016)***
Sub-optimal	0.058	0.039	0.075	0.206	0.829	0.379
Scale	(0.141)	(0.025)	(0.046)*	(0.118)*	(0.242)***	(0.493)
Industry size	0.025	0.001	0.004	0.010	0.034	0.059
	(0.008)***	(0.001)	(0.003)	(0.006)*	(0.013)**	(0.028)**
Industry	0.006	0.003	0.007	0.017	−0.033	−0.198
growth	(0.059)	(0.009)	(0.019)	(0.037)	(0.101)	(0.148)
Turbulence	6.358	1.679	3.797	7.048	5.725	19.887
	(0.894)***	(0.838)**	(0.986)***	(1.153)***	(2.160)***	(6.857)***
Average	0.001	−0.000	0.000	0.000	0.001	0.003
experience	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.003)
Average	0.139	0.002	0.013	0.026	0.155	0.388
education	(0.050)***	(0.004)	(0.017)	(0.033)	(0.076)**	(0.142)***
Share of	−0.516	−0.017	0.069	0.035	−0.383	−1.152
women	(0.549)	(0.027)	(0.124)	(0.258)	(0.697)	(1.449)
Constant	−0.148	1.075	0.904	0.812	−0.685	−2.734
	(0.639)	(0.050)***	(0.220)***	(0.405)**	(0.995)	(2.059)
R^2	0.076					
F (H_0: $\beta_i = 0$)	12.47***	248.70***	224.83***	120.15***	192.99***	176.15***
Services (SIC 50–74, 48 196 observations)						
Scale	7.323	0.424	2.469	7.318	17.179	23.674
economies	(0.852)***	(0.244)*	(0.537)***	(0.815)***	(2.141)***	(2.455)***
Sub-optimal	−0.352	0.021	0.051	0.234	0.262	−0.525
scale	(0.149)**	(0.012)*	(0.057)	(0.102)**	(0.237)	(0.370)
Industry size	0.023	0.000	0.003	0.008	0.021	0.039
	(0.008)***	(0.000)	(0.001)*	(0.004)*	(0.009)**	(0.017)**
Industry	0.032	0.001	0.037	0.042	−0.021	0.062
growth	(0.050)	(0.002)	(0.014)***	(0.032)	(0.065)	(0.147)
Turbulence	2.654	0.199	0.782	1.622	3.623	4.202
	(0.384)***	(0.112)*	(0.171)***	(0.361)***	(0.692)***	(0.783)***
Average	0.001	0.000	−0.000	0.001	0.002	0.003
experience	(0.001)	(0.000)	(0.000)	(0.000)*	(0.001)*	(0.002)
Average	0.165	0.001	0.033	0.107	0.223	0.434
education	(0.041)***	(0.001)	(0.008)***	(0.022)***	(0.058)***	(0.118)***
Share of	0.973	0.002	0.100	0.525	1.279	2.544
women	(0.334)***	(0.004)	(0.050)**	(0.183)***	(0.397)***	(0.862)***
Constant	−1.083	1.068	0.599	−0.584	−2.302	−4.705
	(0.617)*	(0.022)***	(0.129)***	(0.327)*	(0.829)***	(1.631)***
R^2	0.081					
F (H_0: $\beta_i = 0$)	20.84***	7.51	653.92***	315.l06***	406.27***	679.22***

Notes: [1] All estimations include time dummies. [2] Robust standard errors in parantheses. [3] ***, ** and * indicate significant at 1%, 5% and 10% level, respectively.

TABLE VII
Descriptive statistics for manufacturing using the LPDM data

Variable	N	Mean	SD	Median	Min	Max
Start-up size	4739	2.658	0.854	2.485	1.609	6.903
Scale economies	4739	0.119	0.087	0.096	0.007	1.068
Sub-optimal scale	4739	0.418	0.075	0.410	0.090	0.779
Small plant cost disadvantage	4739	0.893	0.169	0.871	0.311	2.384
Industry size	4739	9.259	1.068	9.446	2.565	10.663
Industry growth	4737	−0.026	0.085	−0.020	−0.543	0.660
Turbulence	4739	0.003	0.013	0.000	0.001	0.538
Share of multiplants	4739	0.255	0.272	0.171	0.000	1.000
Share of foreign-owned plants	4739	0.029	0.119	0.000	0.000	1.000
Profitability	4739	0.157	0.057	0.157	−0.132	0.531
Variability of profits	4739	0.002	0.019	0.000	0.000	0.915
Capital intensity	4739	0.121	0.094	0.101	0.003	0.985
Investment rate	4739	0.052	0.025	0.049	0.001	0.439
R&D intensity	4739	0.068	0.118	0.032	0.002	1.049
Export intensity	4739	0.292	0.189	0.256	0.000	1.064
Import penetration	4739	0.294	0.227	0.286	−1.432	1.308

Note: [1] R&D intensity was divided by 100 000 and scale economies, variability of profits and capital.

not have strong explanatory power. Similarly, industry growth turns out to have no significant effect on start-up size in manufacturing. Industry turbulence has a clear positive effect on start-up size and it becomes increasingly important as we move upwards in the conditional size distribution. Overall, these results suggest that large plants enter mostly those industries where scale economies are important or sunk costs are low (turbulence is high). This is in accordance with the previous findings.[21]

In the service sector the effects of scale economies and turbulence correspond to the results for manufacturing. Only scale economies and turbulence have statistically significant effects on plant start-up size. Sub-optimal scale is again positive for most quantiles. On the contrary, the alternative measure, small plant cost disadvantage, turned out to be negative but insignificant in all regressions.[22] It can be concluded that, contrary to the expectations, the results for services and manufacturing seem to be surprisingly similar.[23] This may be due to the limited number of explanatory variables which may lead to the omission of some relevant factors. For example, the availability of skilled labour may be more important for start-ups in services than in manufacturing. For example, high-tech service sectors may be even more dependent on the availability of educated and experienced employees in creating and implementing new innovations than the manufacturing industries using high technology.

Table VI shows the results for manufacturing and services when regional average education, work experience and share of women employees are controlled for. The regional division is based on the regional classification of Statistics Finland, where Finland is divided into 20 regions or provinces corresponding to the so-called NUTS level 3 of the European Union. To decrease the possibility of simultaneity bias, human capital variables are measured in the year previous to start-up. According to the descriptive statistics (Table IV) plant start-ups in services are located in regions where employees are on average more educated but slightly less experienced than in manufacturing. As expected, the share of women is higher in services.

Inclusion of the human capital variables in the model increases the significance of other coefficients. The availability of educated personnel in the region seems to have a positive and increasing effect on plant start-up size. This in accordance with the hypothesis that the availability of qualified work force increases plant's prospects of survival and thus start-up size. Another explanation may be that the shift in demand towards better educated workers due to

TABLE VIII
Quantile regression results for manufacturing using the LDPM data (4737 observations)

	OLS	0.15	0.25	0.50	0.75	0.90
Scale economies	1.719	0.598	0.938	1.565	2.474	2.724
	(0.310)***	(0.264)**	(0.329)***	(0.507)***	(0.554)***	(0.487)***
Small plant cost	−0.247	−0.307	−0.346	−0.266	−0.257	−0.060
disadvantage	(0.102)**	(0.093)***	(0.097)***	(0.117)**	(0.152)*	(0.171)
Industry size	−0.021	−0.010	−0.018	−0.033	−0.023	0.033
	(0.018)	(0.014)	(0.018)	(0.027)	(0.034)	(0.034)
Industry growth	0.333	0.196	0.284	0.232	0.375	0.761
	(0.156)**	(0.207)	(0.191)	(0.201)	(0.256)	(0.331)**
Turbulence	8.398	4.637	7.832	6.163	12.390	26.554
	(2.408)***	(3.603)	(3.514)**	(4.902)	(7.523)*	(14.308)*
Share of	0.616	0.157	0.233	0.432	0.878	1.239
multiplant entrants	(0.056)***	(0.053)***	(0.060)***	(0.097)***	(0.106)***	(0.146)***
Share of foreign-	0.351	0.095	0.256	0.191	0.431	0.665
owned entrants	(0.124)***	(0.171)	(0.147)*	(0.199)	(0.282)	(0.330)**
Profitability	0.064	0.286	0.364	0.125	0.001	0.151
	(0.401)	(0.298)	(0.398)	(0.545)	(0.456)	(0.673)
Variability of	1.560	2.527	2.080	1.643	0.675	0.147
profits	(0.348)***	(5.077)	(5.434)	(3.670)	(3.109)	(4.769)
Capital intensity	0.279	0.101	0.127	0.409	0.553	−0.066
	(0.224)	(0.256)	(0.261)	(0.362)	(0.436)	(0.549)
Investment rate	−0.121	0.232	−0.262	−1.404	−0.256	0.690
	(0.878)	(0.667)	(0.746)	(1.029)	(1.195)	(1.278)
R&D intensity	−0.181	−0.269	−0.395	−0.237	−0.185	0.096
	(0.166)	(0.166)	(0.234)*	(0.339)	(0.411)	(0.531)
Export intensity	0.060	0.106	0.095	0.180	0.124	−0.102
	(0.149)	(0.144)	(0.164)	(0.246)	(0.289)	(0.305)
Import	0.107	0.007	0.031	0.006	0.175	0.297
penetration	(0.075)	(0.088)	(0.095)	(0.129)	(0.182)	(0.201)
Constant	2.561	1.962	2.149	2.676	2.888	2.577
	(0.205)***	(0.190)***	(0.237)***	(0.288)***	(0.371)***	(0.447)***
R^2	0.146					
F (H_0: $\beta_i = 0$)	49.05***	151.16***	265.96***	354.13***	458.22***	614.44***

Notes: [1] All estimations include time dummies. [2] Robust standard errors in parantheses. [3] ***, ** and * indicate significant at 1%, 5% and 10% level, respectively.

recent changes in firms' production technology, i.e. skill-biased technological change, is reflected in the results.[24] It can be argued that this process has increased the average level of education of workers, and at the same time, average plant (start-up) size has increased due to higher productivity and increased competition. However, education turns out to be more important for start-ups in services, whereas in manufacturing the effect is significant only for the largest plants. This may simply describe the different nature of these sectors, i.e. services having a large share of small high-tech and business services plants with highly educated employees. In addition, it may be argued that physical and human capital are more complementary in services where larger plant size increases the demand for skilled work force more than in manufacturing.

Regions characterised by experienced work force may include older firms in the mature phase of their life cycle, which could be expected to increase plant start-up size. However, these regions may also offer market niches for new small plants. The effect of regional work experience, measured with worker seniority, on the start-up size of service sector plants turns out to be positive and somewhat significant in the 0.5 and 0.75 quantiles. In manufacturing the coefficient of experience is insignificant. Using average age of employees instead of seniority to approximate work experience did not change the results much. The average share of female employees in the region also seems to be more closely connected to start-up size in services.[25] The share of women is positively related to the start-up size in services, whereas in manufacturing the relationship is mostly negative and insignificant. Overall, the results seem to indicate that it is important to consider the different role of human capital in manufacturing and services. However, it should be emphasised that these results have to be interpreted with caution because of possible simultaneity problems. In addition, the demand and supply effects of regional skilled work force are difficult to separate.

4.2. *The extended model for manufacturing*

In the second part of the analysis the model for manufacturing is extended by using additional covariates described earlier. For the LDPM data industry attributes are calculated at the 4-digit industry level using the SIC adopted in 1979, which closely corresponds to the level used in the previous analysis with the BR data.[26] Table VII presents the descriptive statistics for the analysis using LDPM data. Comparison of the corresponding variables shows that the mean values for manufacturing are higher for the LDPM data than the BR data, except for industry growth and turbulence. This is most likely due to the higher cut-off limit of five employees in the LDPM data. In order to assess the sensitivity of the results to the data set used, the basic model comparable to Table V was estimated (not reported). However, it should be noted that different size thresholds, time periods and definitions of entry may have some influence on the results. Nevertheless, the results seemed to correspond closely to the earlier findings.[27]

Table VIII reports the estimation results for the extended model for manufacturing. The effects of scale economies and industry turbulence on start-up size remain positive and increasing along the size distribution. The coefficients of industry size and growth are mostly insignificant. Sub-optimal scale is replaced with small plant cost disadvantage, which turns out to be negative and significant for most quantiles. However, the absolute effect is non-monotonically decreasing, which contradicts the earlier results. This finding may be explained by some differences in the industrial structure or business environment in Finland in comparison to Portugal and Ireland, but further analysis is needed. Among the new variables the share of multiplants of all entrants is highly significant and more important for the larger quantiles. The employment share of foreign-owned entrants is also positive and non-monotonically increasing. This gives some evidence that larger start-ups often belong to firms with multiple plants and have a high share of foreign ownership.

Most other variables turn out to be insignificant. Industry price-cost margin does not have a considerable effect on the start-up size. The variability of price-cost margins or industry heterogeneity turns out to be positively related to the start-up size. However, the coefficient for variability in profits is insignificant in all quantile regressions. Capital intensity has a weak positive effect on start-up scale. As expected, investment rate has a negative sign in most parts of the size distribution. R&D intensity is also negative and significant only for the 0.25 quantile. This would suggest that R&D intensity decreases the start-up size, especially for small entrants. However, this variable is not measured very accurately. In accordance with the expectations, both higher export intensity and higher import penetration seem to be positively related to the start-up size, but they turn out to be insignificant.[28]

Since it may take some time for some industry attributes to affect the entry decisions of new plants, the model was reestimated using one-year lags of industry profitability, growth and turbulence. These variables turned out to be mostly insignificant and sometimes negative (not reported). Past profits have not been found to have much predictive power in explaining expected post-entry profitability in earlier studies either. The problem of simultaneity, i.e. an individual entry affecting industry performance and growth, is not likely to very severe because the entry penetration rates measured with employment were very low in most industries (see Table II).

It seems that the variable describing multiplant status is rather significant in explaining plant start-up size. It can be hypothesised that a new plant belonging to a firm with other plants has better financial and other connections, more experience and more realistic expectations about its future possibilities than a new single plant firm. In this environment a higher entry size can be expected. In fact, it was found that the average start-up size for multiplants in the sample is 64 employees, whereas the average start-up size for single plants is only 17.4 employees. In addition, the industry-level factors determining start-up size may differ substantially between these two groups. Tables IX and X include the estimates separately for multi- and single plant entries, i.e. new plants belonging to firms with more than one plant or only a single plant. Unfortunately, it was impossible to trace reliably the number of plants of the parent firm in the year preceding entry due to possible changes in the firm ownership structure, and hence, firm identification code. For the same reason, firms that have entered by establishing several plants at the same time cannot be reliably identified.

The results show that very different industry characteristics determine the start-up size of multi- and single plants. Scale economies seem to be more important for multiplants. However, the effect of MES is still non-monotonically increasing for both groups. Small plant cost disadvantage does not have a clear effect on the start-up size of multiplants, while for single plants its coefficient is always negative and mostly significant. Industry size and growth turn out to be mostly insignificant, but for the highest single plant quantiles the positive coefficient for growth

TABLE IX
Quantile regression results for start-ups belonging to firms with more than one plant (758 observations)

Quantiles	OLS	0.15	0.25	0.50	0.75	0.90
Scale	3.408	2.437	3.140	4.011	4.418	3.284
economies	(0.535)***	(1.232)**	(0.979)***	(0.915)***	(0.734)***	(0.994)***
Small plant cost	0.075	−0.432	−0.043	0.289	0.423	0.773
disadvantage	(0.243)	(0.490)	(0.454)	(0.484)	(0.389)	(0.506)
Industry size	0.045	−0.036	−0.019	−0.046	0.094	0.135
	(0.044)	(0.087)	(0.080)	(0.077)	(0.069)	(0.085)
Industry growth	−0.373	−0.016	−0.355	−0.271	0.047	−1.028
	(0.514)	(0.659)	(0.599)	(0.698)	(0.870)	(1.236)
Turbulence	13.258	11.747	12.875	22.533	11.635	18.631
	(5.799)**	(16.394)	(16.494)	(15.971)	(19.846)	(30.339)
Share of foreign-	0.465	0.229	0.207	0.062	0.788	0.519
owned entrants	(0.231)**	(0.392)	(0.366)	(0.360)	(0.472)*	(0.555)
Profitability	−1.068	−1.490	−1.077	−0.626	−0.888	−1.216
	(0.553)*	(1.279)	(1.218)	(1.169)	(1.001)	(1.233)
Variability of	2.699	3.229	3.038	2.536	1.658	1.776
Profits	(0.469)***	(12.662)	(11.302)	(9.202)	(11.322)	(12.609)
Capital intensity	−0.400	−0.128	−0.524	−0.837	−0.125	−0.278
	(0.284)	(0.771)	(0.686)	(0.696)	(0.636)	(0.800)
Investment rate	0.408	−1.320	−0.888	0.287	−2.066	1.856
	(1.465)	(2.703)	(2.469)	(2.396)	(2.507)	(3.861)
R&D intensity	0.333	−0.181	0.056	−0.122	0.619	1.203
	(0.478)	(0.957)	(0.980)	(1.080)	(0.822)	(0.917)
Export intensity	0.054	0.095	0.857	0.566	−0.463	−0.512
	(0.333)	(0.742)	(0.624)	(0.511)	(0.489)	(0.644)
Import	0.225	0.119	−0.005	0.416	0.412	0.455
penetration	(0.226)	(0.506)	(0.482)	(0.423)	(0.451)	(0.562)
Constant	2.143	2.505	2.108	2.522	2.184	2.323
	(0.495)***	(0.949)***	(0.952)**	(1.008)**	(0.926)**	(0.986)**
R^2	0.215					
F (H_0: $\beta_i = 0$)	30.87***	42.47**	81.54***	186.48***	156.22***	66.88***

Notes: [1] All estimations include time dummies. [2] Robust standard errors in parantheses. [3] ***, ** and * indicate significant at 1%, 5% and 10% level, respectively.

TABLE X
Quantile regression results for start–ups belonging to firms with one plant (3 979 observations)

Quantiles	OLS	0.15	0.25	0.50	0.75	0.90
Scale	0.737	0.395	0.593	0.924	0.819	1.052
economies	(0.286)**	(0.232)*	(0.241)**	(0.409)**	(0.548)	(0.603)*
Small plant cost	−0.208	−0.221	−0.372	−0.196	−0.282	−0.056
disadvantage	(0.136)	(0.095)**	(0.098)***	(0.101)*	(0.140)**	(0.224)
Industry size	0.005	0.001	−0.015	−0.013	0.002	0.042
	(0.016)	(0.013)	(0.016)	(0.024)	(0.032)	(0.038)
Industry growth	0.416	−0.094	0.161	0.120	0.544	1.253
	(0.177)**	(0.213)	(0.211)	(0.202)	(0.257)**	(0.338)***
Turbulence	7.852	3.627	5.918	6.731	19.022	30.382
	(2.335)***	(4.149)	(4.612)	(3.672)*	(9.910)*	(15.134)**
Share of foreign-	0.441	0.119	0.276	0.209	0.511	0.931
owned entrants	(0.142)***	(0.164)	(0.161)*	(0.207)	(0.333)	(0.440)**
Profitability	0.340	0.308	0.779	0.474	−0.064	0.384
	(0.441)	(0.266)	(0.346)**	(0.598)	(0.565)	(0.672)
Variability of	−0.707	−0.474	1.530	−0.184	−4.002	−7.465
profits	(2.240)	(5.204)	(6.150)	(6.244)	(5.974)	(8.093)
Capital intensity	0.369	−0.127	−0.065	0.175	0.867	1.091
	(0.227)	(0.217)	(0.262)	(0.435)	(0.433)**	(0.533)**
Investment rate	−1.124	0.714	−0.059	−1.793	−2.425	−1.435
	(0.769)	(0.608)	(0.612)	(0.977)*	(1.118)**	(1.673)
R&D intensity	−0.185	−0.052	−0.370	−0.344	−0.094	0.075
	(0.177)	(0.167)	(0.196)*	(0.351)	(0.439)	(0.590)
Export intensity	0.186	0.076	0.183	0.209	0.370	0.297
	(0.140)	(0.121)	(0.136)	(0.211)	(0.287)	(0.293)
Import	0.132	−0.026	0.006	0.052	0.118	0.375
penetration	(0.090)	(0.094)	(0.096)	(0.145)	(0.178)	(0.194)*
Constant	2.442	1.825	2.179	2.538	2.970	2.678
	(0.199)***	(0.179)***	(0.223)***	(0.276)***	(0.330)***	(0.491)***
R^2	0.060					
F (H_0: $\beta_i = 0$)	12.32***	91.69***	199.84***	286.40***	166.61***	193.55***

Notes: [1] All estimations include time dummies. [2] Robust standard errors in parantheses. [3] ***, ** and * indicate significant at 1%, 5% and 10% level, respectively.

is highly significant. Industry turbulence and the share of foreign-owned plants have positive, non-monotonically increasing effects on start-up size, but both covariates seem to be more important for single plants. According to the quantile regression results multiplants do not seem to be significantly affected by the rest of the industry attributes, with a few exceptions. On the contrary, for single plants capital intensity turns out to be positive and significant for the highest quantiles, whereas investments reduce the start-up size especially in the 0.5 and 0.75 quantiles. The effect of R&D activity is negative and significant in the 0.5 quantile. Export intensity turns out to be insignificant and the coefficient for import penetration is only significant in the highest quantile.

5. Conclusions

The first aim of this paper was to examine whether the earlier findings based on data from the manufacturing sector also hold for services. It was found that the results are remarkably similar for these two sectors of the economy when only a restricted set of explanatory variables is used. Further analysis is needed to shed some light on this surprising result. For example, the regional work force composition seems to have a different role in the choice of start-up size in manufacturing and services. The results show that the availability of educated personnel is positively related to start-up size both in manufacturing and services, but the effect is more important in the service sector. Work experience and share of women seem to be more closely connected to the start-up size in services. However, when considering the effects of human capital on the choice of start-up scale, simultaneity of these decisions may cause problems.

The second purpose was to test the effects of some additional covariates on plant start-up size using data on manufacturing only. The findings suggest that the effects of some industry characteristics on plant start-up size differ considerably depending on the scale of entry, so the OLS estimates can be highly misleading. It is found that large plants are less likely than small plants to enter industries with low scale economies or high sunk costs. In addition, belonging to a multi-unit firm or having a large share of foreign ownership increases the start-up size. There is also some evidence on the negative effect of R&D intensity on the entry scale of small plants. In addition, it is found that the start-up scale of single plant entrants is more affected by the industry attributes than that of multiplant entrants, who face a rather different entry environment.

In each industry there are barriers to entry which may increase the optimal start-up size. The results suggest that large-scale entry is preferable when the need for sunk investments or the possibility of strategic action against entrants is low. Being small allows entrants to minimise the capital requirements and to avoid aggressive responses from the incumbents. In particular, the start-up scale of single plant entrants seems to be more affected by the strategic behaviour in the industry than that of entrants belonging to multiplant firms. Through a learning process small entrants become more certain about their performance and their prospects of growth and survival improve.

In the further analysis, the effect of financial constraints on entry scale would be worth studying. In public policy debate financial factors are often seen as a very central and interesting determinant of firm start-ups and job creation. In addition, the effects of regulation and business environment on the entry patterns in different industries could be studied more extensively. For example, pharmacies and the banking sector are affected by very different rules and legislation regarding start-ups. In addition, entry patterns may vary considerably over the phase of the industry life cycle. The data would also allow the analysis of cyclical effects on start-up size. It could be studied whether the business cycle has had similar effects on the rate and scale of entry in different industries.

Acknowledgements

I am grateful to Ari Hyytinen, Pekka Ilmakunnas, Mika Maliranta, Otto Toivanen and participants in various seminars for helpful comments. Valuable suggestions of the two referees are gratefully acknowledged. I would also like to thank the Academy of Finland for funding and Statistics Finland for allowing access to the data.

Notes

1 However, there are also opposite findings of a negative relationship between start-up size and the probability of survival (e.g., Disney et al.,1999; Mata et al., 1995).

2 To illustrate, in 1998 the BR covered, in terms of personnel, 99% of the non-agricultural enterprise sector in Finland, i.e. over 200 000 business sector plants.

3 Furthermore, plants with less than three employees are likely to be small family businesses where the choice of start-up size may be strongly influenced by factors which are very different from the industry or regional attributes used in this analysis.

4 Agriculture and public sector are excluded from this analysis.

5 Capital stock is estimated as the real value of machinery, equipment, transportation equipment, buildings and structures using the perpetual inventory method.

6 The number of employees includes persons who are, for example, on maternity leave, on annual leave or temporarily laid-off, which may bias some of the results.

7 Higher number of start-ups in 1989 and 1990 may be partly explained by the business cycle effects, partly by the changes in the statistical system. However, when these years were excluded from the estimations, the results were consistent across the two alternative samples.

8 The largest plant start-up size in the BR is 1300 employees, which is very high and may have been recorded incorrectly.

9 However, the main results did not change when these years were dropped from the estimations.

10 There seem to be some large start-ups in 1989, which temporarily increase the mean and standard deviation of the size distribution.

11 It should be noted that the average start-up size according to the mean may be influenced by only a few large outliers. However, the quantile regression method used in the further analysis relies on the median, which is less suspicious to outliers than the mean.

12 The business sector is defined as the total economy minus agriculture, hunting, forestry and fishing, and

community social and personal services. Manufacturing does not include mining and quarrying or electricity, gas and water supply. Services include wholesale and retail trade, restaurants and hotels, transport, storage and communication, finance, insurance, real estate and business services.

[13] The exclusion of the largest outliers (start-ups with more than 500 employees) decreased the relative start-up size in 1995.

[14] Small firms (having less than 20 employees) with multiple plants are treated as a single plant in the Industrial Statistics survey. In addition, industrial plants located in the same area or plants engaged in secondary activities in another industry may be combined.

[15] In addition, an alternative measure can be used to characterise the small plant cost disadvantage, i.e. small plant (plants with less than industry average number of employees) productivity in relation to the average productivity in the industry.

[16] Görg and Strobl (2002) have studied the impacts of foreign multinationals on the start-up size of domestic entrants in Irish manufacturing.

[17] The model assumes that the impact of the different covariates remains constant over the period, which may not be a valid assumption. However, the main findings did not change when the model was reestimated using data on separate years.

[18] Using clustering has a clear effect on the standard errors and some coefficients loose their significance compared to the analysis without clustering (Nurmi 2003).

[19] In the models with regional human capital effects, clustering is based on a grouping variable which combines industry and region.

[20] When plants in the information and communications technology (ICT) sectors were excluded from the analysis, the relationship between sub-optimal scale and start-up size became more negative in services. This finding may suggest that ICT-plants are affected differently by the industry's size distribution. The average start-up size for ICT-plants in the sample is 9.6 employees, whereas for non ICT-plants it is 7.5 employees.

[21] It may be argued that there are endogeneity problems in some of the covariates, for example, MES and sub-optimal scale, which describe the size distribution of plants in the industry. In addition, some variables may affect the entry decisions with a lag. However, when the basic model was reestimated using one-year lags of all the variables, the most notable change was that the coefficient for turbulence lost its significance in services.

[22] It can be tested whether the effects of the covariates are the same between different quantiles. *F*-tests showed that there are significant differences between different quantiles for most of those coefficients, which were significant in the estimations (not reported).

[23] In order to take into account the cut-off limit of three employees, we also estimated a truncated regression model with robust standard errors for start-up size with truncation from below ln 3. The coefficient signs were as expected and the absolute effects increased considerably. However, while for services most relationships were statistically significant, for manufacturing they were not.

[24] Chennells and Van Reenen (1999) provide a survey of the effects of technical change.

[25] It should be noted that variables for seniority and the share of women are to some extent correlated (−0.65).

[26] This amounts to having 79 industries in manufacturing. R&D expenditures and imports are available at the 4-digit industry level only for some industries, others are aggregated to the 2- or 3-digit levels.

[27] The results with truncation from below ln 5 were very similar to the OLS estimates when the LDPM data set was used.

[28] It should be noted that due to high correlations between some of the variables there may be some degree of multicollinearity in this specification. Due to a high correlation between export intensity and import penetration (0.45), these variables were included in separate models (not reported). In addition, R&D intensity was excluded because of its high correlation with imports (0.48). However, the effects of foreign trade remained insignificant and other findings did not change.

References

Acs, Z. J. and D. B. Audretsch, 1989, 'Small-firm Entry in US Manufacturing', *Economica* **56**, 255–265.

Audretsch, D., 1995, *Innovation and Industry Evolution*, Cambridge, MA: The MIT Press.

Buchinsky, M., 1998, 'Recent Advances in Quantile Regression Models: A Practical Quideline for Empirical Research', *Journal of Human Resources* **33**, 88–126.

Cabral, L., 1995, 'Sunk Costs, Firm Size and Firm Growth', *Journal of Industrial Economics* **43**, 161–172.

Caves, R., 1998, 'Industrial Organization and New Findings in the Turnover and Mobility of Firms', *Journal of Economic Literature* **36**, 1947–1982.

Chennells, L. and J. V. Reenen 1999, *Has Technology Hurt Less Skilled Workers? An Econometric Survey of the Effects of Technical Change on the Structure of Pay and Jobs*. Institue for Fiscal Studies Working Paper No. W99/27, London.

Comanor, W. S., and T. A., Wilson, 1967, 'Advertising Market Structure and Performance', *Review of Economics and Statistics* **49**, 423–440.

Disney, R., J. Haskel and Y. Heden, 1999, 'Entry, Exit and Establishment Survival in UK Manufacturing', Center for Research on Globalisation and Labour Markets, Research Paper 99/9. University of Nottingam.

Dunne, T., M. Roberts and L. Samuelson, 1989, 'The Growth and Failure of U.S. Manufacturing Plants', *Quarterly Journal of Economics* **104**, 671–698.

Evans, D., 1987, 'Tests of Alternative Theories of Firm Growth', *Journal of Political Economy* **95**, 657–674.

Geroski, P., 1991, *Market Dynamics and Entry*, Oxford: Basil Blackwell.

Görg, H. and E. Strobl, 2002, 'Multinational Companies and Entrant Start–up Size: Evidence from Quantile Regressions', *Review of Industrial Organization* **20**, 15–31.

Görg, H., E. Strobl and F. Ruane, 2000, 'Determinants of Firm Start-Up Size: An Application of Quantile Regression for Ireland', *Small Business Economics* **14**, 211–222.

Ilmakunnas, P., M. Maliranta and J. Vainiomäki, 2001, 'Linked Employer–Employee Data on Finnish Plants for the Analysis of Productivity, Wages and Turnover', in T. P. Jensen and A. Holm (eds.), *Nordic Labour Market Research on Register Data*, Copenhagen: TemaNord 2001: 593, Nordic Council of Ministers. pp. 205–246.

Jovanovic, B., 1982, 'Selection and Evolution of Industry', *Econometrica* **50**, 649–670.

Koenker, R. and G. Bassett, 1978, 'Regression Quantiles', *Econometrica* **46**, 33–50.

Koenker, R. and K. Hallock, 2001, 'Quantile Regression', *Journal of Economic Perspectives* **15**, 143–156.

Mata, J., 1991, 'Sunk Costs and Entry by Small and Large Plants', in P. Geroski and J. Schwalbach (eds.), *Entry and Market Contestability: An International Comparison*, Oxford: Basil Blackwell, pp. 49–62.

Mata, J. and J. A. F., Machado, 1996, 'Firm start-up size: A conditional quantile approach', *European Economic Review* **40**, 1305–1323.

Mata, J. and P. Portugal, 1994, 'Life Duration of New Firms', *Journal of Industrial Economics* **42**, 227–245.

Mata, J., P. Portugal and P. Guimarães, 1995, 'The Survival of New Plants: Start-up Conditions and Post-entry Evolution', *International Journal of Industrial Organization* **13**, 459–481.

Moulton, B. R., 1990, 'An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units', *Review of Economics and Statistics* **72**, 334–338.

Nurmi, S., 2003, *Sectoral Differences in Plant Start-up Size*. Helsinki School of Economics, Working Papers No. W-335.

White, H., 1982, 'Maximum Likelihood Estimation of Misspecified Models', *Econometrica* **50**, 1–25.

You, J. I., 1995, 'Small Firms in Economic Theory', *Cambridge Journal of Economics* **19**, 441–462.