An Exploratory Analysis of the Structure of the FT-SE 100 Index

D. C. Bowie[1] and R. S. Clarkson[2]

Abstract

In this paper we describe an exploratory analysis of the cross-sectional structure of the FT-SE 100 Index and discuss how the construction of the index will impact on any differences in the performance between the FT-SE 100 index and the All Share index. The structure of the index is analysed primarily in terms of the relative sizes of the index constituents at consecutive size ranks using "gap factors". We show that a simple hyperbolic model provides an encouragingly good fit to the gap factors and suggest that it can form the basis for the development of a more sophisticated model of growth in the market capitalisations of securities.

Keywords

Stock index, performance.

[1] Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, EH14 4AS (Scotland); Tel: + 44-131-451 3204, Fax: + 44-131-451 3249, E-mail: D.C.Bowie@ma.hw.ac.uk

[2] Cherrybank, Manse Brae, Dalserf, Lanarkshire ML9 3BN (Scotland); Tel: + 44-1698-882451, Fax: + 44-141-204 1584.

1 Introduction

The FT-SE 100 index is possibly the most widely quoted and followed index on the London market, being a real-time proxy for the more widely based All Share index. The FT-SE 100 index comprises the 'one hundred leading shares' (which correspond rather loosely to the 100 largest companies traded on the London market) compared with a base of nearly 900 shares in the All Share index. Updating the All Share on a real time basis would be impracticable and hence the smaller FT-SE 100 index is used for the specification of traded options and futures on LIFFE. Furthermore, index-tracking funds often follow the FT-SE 100 index rather than the All Share, presumably because of the practical difficulties and costs associated with tracking nearly 900 shares if using a replication stategy.

Past studies have attempted to measure the performance of the FT-SE 100 index to quantify any systematic deviations in its performance away from the All Share index (e.g. Dimson & Marsh (1984)) and the effect on the stock price of a share entering or leaving the index (Goodacre & Lawrence (1994)). In this paper we analyse, at an exploratory level, the cross-sectional structure of FT-SE 100 index.

The structure of the paper is as follows. In Section 2 we briefly describe the construction of the FT-SE 100 index. Section 3 is a description of an empirical study of the relative performances of the FT-SE 100 and All Share indices and of the structure of the FT-SE 100 index. Section 4 contains a summary of the findings and some suggestions for further research.

2 Construction of the FT-SE 100 Index

The FT-SE 100 index was created on 1 January 1984 and comprised, at that date, the 100 largest companies traded on the London Stock Exchange as measured by market capitalisation at the end of 1983. It was designed to be updated continuously in real time and to facilitate the introduction of stock index futures and options trading in the UK market. Dimson & Marsh (1984) suggest that the index was not intended for investment performance management, unlike the All Share index, but it is clear that many portfolio managers do use it for that purpose.

The constituents of the index are generally reviewed on a quarterly basis, in March, June, September and December of each year. A detailed description of the construction of the indices can be found in Brumwell (1984) and the *Guide to Calculation Methods* published by the FT-SE Actuaries Indices Steering Committee (1996). For our purposes the outline below will suffice for the discussion and analysis which follows.

In the description and the ensuing discussion, the rank of the security is the rank based on market capitalisation, i.e. the stock with rank k is the kth largest security. A rank of one is given to the largest security. We talk of ranks 'above k', say, as meaning all those securities with market capitalisations larger than that of the security with the kth largest market capitalisation. Conversely, stocks with ranks 'below k' are stocks with smaller capitalisations than the stock with the kth largest stock.

(a) Stocks on the date of a review which are ranked between 1 and 90 are automatically included in the FT-SE 100 for the next quarter. We will refer to the stocks which enter the FT-SE 100 in this way at a particular review as "IN" stocks.

(b) All stocks ranked at or below 111 are automatically excluded from the FT-SE 100 index for the next quarter. These stocks are referred to as "OUT" stocks.

(c) Stocks ranked from 91 to 100 inclusive at the date of a review are eligible for inclusion in the index. All those ranked between 91 and 100 which were in the index at the previous review remain in the index. Those which were not in the index at the previous review are included if their rank, r, at the time of the current review is such that $(r-90) \leq (\#OUT - \#IN)$, where the notation # is used so that $\#OUT$ is taken to mean the number of "OUT" stocks created at the current review.

(d) Stocks ranked between 101 and 110 are eligible for possible exclusion from the index. If a stock was not in the index at the previous review then it is NOT included at this review. Stocks which are currently in the index will be excluded for the next quarter (at least) if $(111-r) \leq (\#IN - \#OUT)$.

Other stocks which are listed on the exchange for the first time between review dates and which are large enough for immediate inclusion may be added immediately (for example, large privatisations such as British Gas have been included at dates other than review dates). Other stocks whose capitalisations are sufficiently enlarged by virtue of a merger or takeover will also be considered immediately for inclusion. On the other hand if a stock is delisted because of a merger or takeover, then it will be removed from the index immediately and replaced by the largest stock which is not included in the index at the time of the delisting.

The above description indicates that just after any review date, the FT-SE 100 index consists of the 90 largest stocks and 10 other stocks which have actual size ranks of between 91 and 110. We refer to the ranks based purely on size as 'actual' ranks to distinguish them from the ranks of the securities within the index, the 'nominal' ranks. Actual and nominal ranks 1 to 90 will be identical, but this relationship need not hold for other nominal and actual ranks.

The index is then computed as the aggregate [2] of all the market capitalisations of stocks in the index. The one-period rate of return on the index, i.e.

$(I_{t+1}/I_t - 1)$, where I_t is the value of the index at time t, can then be interpreted as the weighted arithmetic average of the one-period rates of return of the constituent securities, with weights equal to the market capitalisation of the securities.

At times beween reviews, the discrepancy between actual and nominal ranks will be even larger. For example, if the price of a share in the index drops to the extent that if a review were to take place immediately then the share would be excluded, then the FT-SE 100 index will reflect that drop and any further decreases until the next review. Additionally, a share which was not in the index at the start of a quarter may experience very high growth between reviews. However, its growth will not have been reflected in the FT-SE 100 even though that growth may have made the company large enough that it would have been included in the index if a review were to have taken place. It might therefore be argued that the FT-SE 100 will underperform the All Share index.

On the other hand, because reviews take place every quarter, stocks which are in the index at the start of a very long period of underperformance of the market will be excluded from the index before they depress the index very much. Also, stocks which experience long term outperformance of the market will be included fairly rapidly in the index, replacing those stocks which have decreased in size. The combination of the two effects just described may result in an outperformance of the All Share index by the FT-SE 100 index. The net effect of the underperformance caused by the discrete nature of the review process and any outperformance caused by the 'winners replacing losers' effect is a question for empirical investigation.

3 Empirical study

3.1 Historical comparison of FT-SE 100 index and All Share index

In order to get some feel for how the All Share and FT-SE 100 indices performed relative to each other, we used a set of daily index values obtained from 31 December 1987 to 1 July 1993. Two analyses were performed: (1) a regression of the (continuously compounded) FT-SE 100 returns (measured as $\log I_{t+1}/I_t$ where I_t is the value of the index at time t) against the All Share returns, and (2) a comparison of the number of 22-day[3] periods when the FT-SE 100 index outperformed the All Share and *vice versa*.

A regression of the FT-SE 100 returns (non-overlapping, 22-day returns) on the All Share returns provided the following relationship:

$$\delta_{FT-SE} = 0.0008 + 0.9575\delta_{FTA} \tag{1}$$

where δ is the continuously compounded return over the 22-day period. The R^2 was 0.983 and the correlation between the indices was over 0.99. The estimated parameters of these regressions are very similar to those obtained by Dimson & Marsh (1984)) in their comparison of the FT-SE 100 and the All Share Index. The fact that the slope coefficient in the relationship is significantly less than unity (as the estimated standard error of the slope coefficient is 0.0160) would seem to imply that the FT-SE 100 will, on average, underperform the market in a bull run and outperform the All Share in a declining market.

The All Share index outperformed the FT-SE 100 index in 639 (46.7%) of the 1369 overlapping 22-day intervals which can be formed from the available data. The overlapping intervals are clearly not independent of each other which makes it difficult to compute confidence limits for the proportion of intervals of outperformance. If we use a set of the non-overlapping intervals and assume that the returns over these intervals are serially independent, then we can obtain an approximate 95% confidence interval for the proportion of times that the All Share index outperforms the FT-SE 100 index, viz: 0.492 ± 0.125. The interval includes 0.5 and so indicates that there is no evidence to support the hypothesis that either index dominated the other when measured over intervals of 22 trading days during the period of investigation. There are 22 ways of constructing different series of non-overlapping 22-day returns, starting the series on days one to 22 of the available data. Repeating the above analysis for all possible ways of constructing the returns confirmed that the number of times that the All Share outperformed the FT-SE 100 index could be attributed to randomness.

The second part of this second analysis comprises a runs[4] test, i.e. a test of whether there existed any sub-periods within the period of investigation during which the All Share consistently outperformed the FT-SE 100, or vice versa. The intuition behind the test is based on testing whether there is any systematic grouping of the intervals of outperformance of the one index by the other. The tests revealed that in eight out of the possible 22 sets of 22-day returns, there was significant grouping of the intervals of outperformance at a 5% significance level. Again, the interpretation of the proportion of 'significantly grouped' data sets is confounded by the fact that the 22 sets of returns are not independent. Nevertheless, there would appear to be an indication that the sub-periods of outperformance are signficantly grouped in some cases. This interpretation is further supported by the observation that in all 22 sets, the maximum number of consecutive non-overlapping 22-day intervals when the one index outperformed the other was *at least* seven.

The analyses reported in this section have been preliminary ones into the relative performances of the FT-SE 100 and All Share indices. The results indicate that the two indices are highly correlated and that, over the whole period of investigation, there was no evidence to suggest that the one index outperformed the other. However, runs tests showed that there might be a grouping of the intervals of outperformance which would indicate that over shorter periods, the indices might show different rates of return.

3.2 Data used

In the remainder of the analyses described in this paper we have used the index constituents and market capitalisations of the FT-SE 100 index as reported by Brumwell (1984-94) each year in the Transactions of the Faculty of Actuaries, Journal of the Institute of Actuaries and the British Actuarial Journal. The data therefore relate to the index as at the last trading day of each of the eleven years from 1984 to 1994, after the changes implemented in the relevant December review. Using this data set does of course exclude any securities which entered and left the index entirely within any one calendar year. The effect of these omissions is of course not measurable within the scope of this study. However, because the focus of this study is more on the cross-sectional structure in the index, rather than on any specific dynamic effects, the practical effects are likely to be negligible.

The total market capitalisation of the FT-SE 100 index has varied around a level of just over 70% of the total market capitalisation of the All Share index.

3.3 Relative sizes of index components

Because the index represents the sum of the market capitalisations of its constituents, its performance (ignoring capital or constituent changes) over time is a weighted average of the constituent security performances, with the weights equal to the market capitalisations of the securities. The very large companies therefore have more influence on the movements in the index than the smaller companies. Abnormal, or unusual performance of the smaller securities might therefore not be reflected in the index movements. (See Adams, Bloomfield, Booth & England (1993, chapter 15).)

Table 1 gives a summary of the distribution of relative market capitalisations for the constituent securities in each year. Although the changes in relative sizes over time are necessarily period specific and cannot be extrapolated into the future, the extent of variability of the sizes provides a measure of the historical instability in the cross-sectional structure of the index. The most notable feature is that the maximum weighting varied considerably between 1984 and 1994 and indeed decreased by nearly one-half, viz. the maximum in 1984 was 8.14% compared with 4.65% in 1994. However, it is also noteworthy that the third quartile (i.e. the market capitalisation of the share ranked at number 25) increased from 1.02% to 1.21% over the eleven years. This increase would indicate that the moderately large companies had grown relative to the index as a whole over the period in contrast with the very largest companies. The median security's weight varied quite extensively between 0.51% and 0.66% with an average of 0.595%. Similarly, the relative size of the first quartile company varied between 0.35% and 0.48%. The weight associated with the smallest constituent of the index, i.e. the minimum in each year, remained relatively constant at around 0.28%.

This evidence would then indicate that the distribution of weights in the index has flattened out in that the difference in market capitalisation between the largest and smallest securities in the index has narrowed. However, the flattening out between maximum and minimum is counteracted to some extent by a widening of the differences in the relative sizes of the stocks ranked between, say, 25 and 75. It will be interesting to determine whether this effect is permanent and ongoing or whether it reverses itself, perhaps when/if the current bout of demergers reverses itself.

Figure 1 plots the distribution of the relative market capitalisations across ranks of the index. It is apparent both from the columns of Table 1 and from

	1984	1985	1986	1987	1988	1989
Min	0.28	0.28	0.28	0.27	0.30	0.29
1st Quartile	0.40	0.42	0.44	0.48	0.47	0.44
Median	0.60	0.62	0.60	0.66	0.64	0.58
3rd Quartile	1.02	1.06	1.01	1.01	1.10	1.00
Max	8.14	8.42	6.23	6.69	6.24	5.48

	1990	1991	1992	1993	1994
Min	0.26	0.22	0.26	0.28	0.29
1st Quartile	0.42	0.35	0.35	0.40	0.40
Median	0.63	0.53	0.51	0.57	0.59
3rd Quartile	1.19	1.16	1.27	1.19	1.21
Max	5.82	6.74	5.58	5.30	4.66

Table 1: Distribution of relative market capitalisations for consituent securities in the FT-SE 100 index.

shape of the curves in Figure 1 that the relative market capitalisations (and hence influence on the index) drops off rapidly as the rank of the constituent share approaches 100. Taking 1984 as an example, the largest security's market capitalisation constituted over 8% of the total capitalisation of the index, whereas the smallest security represented under 0.3% and the median security had a weight of much less than 1%.

An important implication of the very large difference in market capitalisation between the largest and smallest securities is that even systematic outperformance of the smaller securities will not necessarily lead to outperformance of the market by the FT-SE 100 index. For example, Goodacre & Lawrence (1994) have shown that securities exhibit, on average, a 3% abnormal return when entering the index. In addition they demonstrated that there is no corresponding, permanent effect when securities leave the index. Given that approximately 12 securities enter the index each year, the new en-

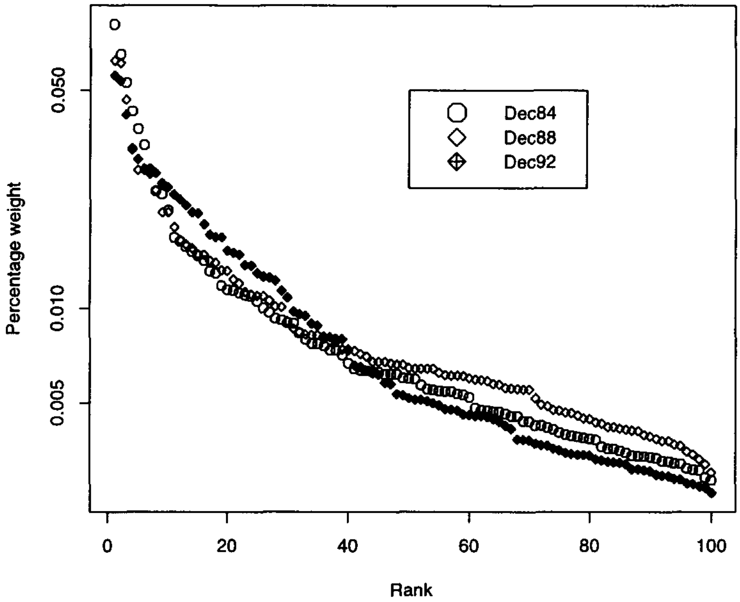

Figure 1: Distribution of relative market capitalisations for constituent securities in the FT-SE 100 index for 1984, 1988 and 1992.

trants would together imply an outperformance of 36% each year. However, as we show later, most of the new entrants enter the index at ranks below 75 so their weighting will be very low. In fact, if we take 0.3% as the average weighting of a new entrant, then the 36% abnormal performance translates into a 0.108% abnormal performance for the index as a whole in that year. Even over eleven years, the cumulative, compound effect of 0.108% abnormal return per year is considerably less than 1.5%. Such a very small outperformance would be extremely hard to detect. In fact the outperformance will be very much less than this even: Goodacre & Lawrence's study indicated that much of the 3% abnormal return occurs *before* the security enters the index as investors speculate on whether the security will enter the index. Because much of the outperformance occurs before the security enters the index, the abnormal performance will not be reflected in the index. Moreover, the abnormal performance which occurs before the security enters the FT-SE 100 index might cause the FT-SE 100 index to underperform (by an extremely small margin) the All Share index, and conversely lead to a tiny outperformance of the All Share by the FT-SE 250. [5]

An alternative display of the concentration of market capitalisations within the index is via the Lorenz curve which is a widely used graphical display in the area of income distribution and equality. A Lorenz curve is a plot of the maximum proportion of companies which can constitute a given proportion of the total market capitalisation of the index. If all companies within the index were of equal size, then the plot would be a straight line from (0%, 0%) to (100%, 100%). This is clearly not the case with the FT-SE 100 index and the Lorenz curve is bowed away from the line of equal market capitalisation. This is clearly evident in the Lorenz curves from the FT-SE 100 constituents for 1984, 1987 and 1990 shown in Figure 2.

The area between the line of equal market capitalisations and the actual Lorenz curves represents a measure of the concentration of market capital-

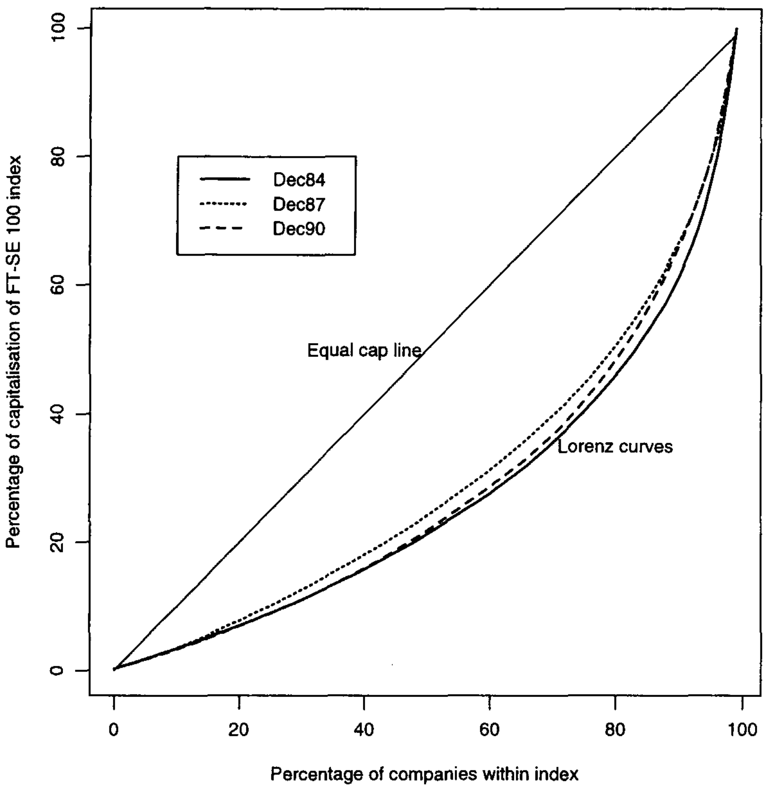

Figure 2: Lorenz curves for the constituent securities of the FT-SE 100 index for 1984, 1987 and 1990.

1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
0.469	0.455	0.443	0.418	0.408	0.432	0.448	0.490	0.471	0.447	0.437

Table 2: Gini coefficients for the FT-SE 100 constituents

isation within the index. Larger areas correspond to situations of greater concentration within the index, while smaller areas correspond to situations of greater equality. The area (as a proportion of the whole area under the line of equal market caps) is referred to as the Gini coefficient and can be calculated in a sample $(x_1, x_2, \ldots, x_n)$ using the formula:

$$g = \frac{\Delta}{\bar{x}} \tag{2}$$

where

$$\Delta = \frac{\sum_i \sum_{j>i} |x_i - x_j|}{n^2}$$

and $\bar{x}$ is the sample mean. The Gini coefficients for the FT-SE 100 index constituents in each of the eleven years is given in Table 2. The coefficients vary between a low of 0.408 in 1988 and a high of 0.490 in 1991, which reflect fairly high levels of concentration. Although the comparison is not necessarily very useful, a feel for the level of concentration implied by the measured Gini coefficients can be obtained by noting that the Gini coefficient for family income in the US in 1980 was about 0.35 (see McDonald (1984)). There are a couple of caveats associated with this analysis, which could be the basis for further investigation. First, we have made no attempt to measure the sampling error in the estimator of the coefficients and so have not constructed confidence intervals to assess the 'significance' of the Gini coefficient. Secondly, the particular set of securities which we are analysing are the largest listed securities in the UK. It might therefore be expected that if a more representative sample of securities across the full range of sizes of companies in the UK were taken, then the measure of concentration may decrease.

3.4 Gap factors between market capitalisations

A more detailed aspect of the structure in the index is revealed when the differences between the market capitalisations of securities at adjacent ranks are considered. We refer to the logarithm of the relative market capitalisations at one rank and its immediately higher rank as the gap factor for the rank. If we denote the gap factor for a rank r at time t by g_{rt} and the market capitalisation of the security at rank r at time t by C_{rt} then

$$g_{rt} = \log(C_{(r-1)t}/C_{rt}) \qquad (3)$$

The gap factors can be thought of as the annualised continuously compounded rate of growth needed for a security at a given rank to move to an immediately higher rank, provided the structure of the index remained constant.

Figure 3 contains the average gap factors at each rank at each year end from 1984 to 1994. The most striking feature of Figure 3 is that, apart from ranks 91 to 100, the gap factors increase with firm size, i.e. large firms have to experience relatively greater market capitalisation growth to change ranks than smaller firms. For firms ranked higher than about 20, the average gap factors are all above 0.03 and range up to over 0.20. For firms ranked from about 21 to 40, the gap factors seem fairly constant at about 0.03. For ranks lower than 40 down to 90, the gap factors are again relatively level at around 0.02. From ranks 91 to 100, the average gap factors rise again. This small rise with decreasing size of firm appears anomalous unless one takes into account the fact that a stock nominally ranked between 91 and 100 may not actually have a market capitalisation which corresponds to that rank. The discrepancy between nominal and actual ranks arises because of the way the index constituents are added and deleted. For example, a stock in the index just after a review may, after some time, decrease in size to have an actual rank of over 100, but the stock will not necessarily be excluded from the index at the next review and its nominal rank will therefore be recorded as being at under 100. Consequently, the gap factors estimated from the nominal ranks will be an overestimate of the gap factors which would be obtained from using the actual ranks.

Based on the structure revealed in Figure 3, we might posit a breakdown of the FT-SE 100 stocks into four subsets: (1) the 'superlarge firms', i.e. the top 20 very large firms which are very unlikely to move out of their subset because the change in share price would have to be very large for the firm to change ranks; (2) the 'seasoned index stocks', i.e. a band of fairly large firms ranked between 21 and 40 which are likely to remain in that band unless something dramatic happens to their stock price; (3) 'marginal stocks', i.e. the stocks which rank between 41 and 90 which can move up or down between ranks (and hence in and out of the index as well) with quite small changes in stock price and (4) 'just in' stocks, whose nominal and actual ranks do not necessarily correspond.

3.4.1 An empirical model for the gap factors

Although we are not aware of a model which would explain the observed cross-sectional distribution of market capitalisations in the FT-SE 100 index, it would be useful to have some sort of parameterisation of the relationship between the rank and market capitalisation of the security within the index.

The relationship between rank and size which is evident in Figure 1 is far from linear or even exponential in nature. A simple hyperbolic function has been suggested by Zipf (discussed in Gell-Mann (1994, pp92-97) and Patel *et al.* (1976)) for predicting size from rank in studies of city-size and many other 'populations' and which was thought may hold for market size as well, viz:

$$C_{rt} = \alpha_t \times \frac{1}{r} \tag{4}$$

where C_{rt} is the market capitalisation of the share ranked rth at time t and α_t

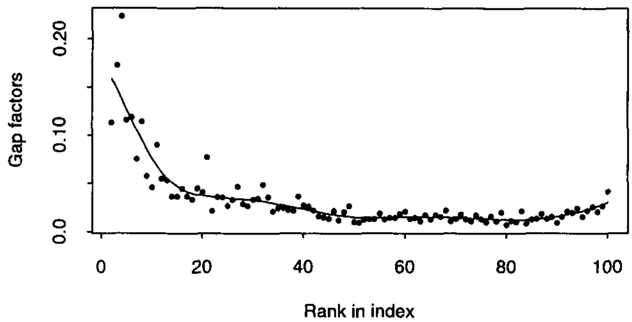

Figure 3: Average gap factors at each rank

is cross-sectionally constant. If the model holds exactly over the full range of ranks then of course α_t would correspond to C_{1t}, the market capitalisation of the largest security at time t. The model, it should be emphasized, is built up purely on an empirical basis and is deterministic in nature. If the relationship gives a reasonable fit, however, it might be useful for measuring systematic differences from that relationship, or even assessing any time-varying nature of the shape of the market cap-rank relationship.

To assess the reasonableness of the model, we again consider the gap factors, g_{rt}. Rewriting equation 3 having substituted C_{rt} by its model equivalent from equation 4, we obtain:

$$\begin{aligned} g_{rt} &= \log(1 - \frac{1}{r})^{-1} \\ &= 1 + \frac{1}{r} + \frac{1}{2r^2} + \frac{1}{3r^3} + \dots \end{aligned} \tag{5}$$

where the last equality is obtained using the familiar Taylor expansion. This last equality indicates that we would expect the gap factors to be approximately linear in $1/r$ for large r. For smaller r, the relationship would become quadratic, and as r became even smaller, the relationship would become cubic, then quartic, etc.

An empirical version of the model in equation 5 might be

$$g_{rt} = \beta_{0t} + \beta_{1t} \log(1 - \frac{1}{r})^{-1} + e_{rt} \tag{6}$$

where e_{rt} is an 'error' term which reflects deviations from the proposed model. Because of the largely descriptive nature of the model, the properties of the error terms are not easy to specify. However, we assume that the error terms will have zero means for all ranks, r, but that the variance structure is more complicated. Specifically, we might expect the variance of the error term to increase at a similar rate to $\log(1 - \frac{1}{r})^{-1}$ as r became smaller. Hence, we weight each observation by the square root of this factor so that the errors will have a constant variance across observations. In addition we restricted the companies used for the fitting of the model to be such that $20 \leq r \leq 90$. The lower bound for r was chosen so as to exclude the superlarge firms which could be considered to be qualitatively different from the other constituents.[6] The upper bound for r is necessary because, as explained earlier, the companies with nominal ranks 91 to 100 do not necessarily correspond to the companies with those actual ranks. Under the hypothesis that the model is a reasonable representation of the structure of the index, the parameter β_{0t} would be expected to be zero and β_{1t} would be expected to be unity.

Table 3 summarises the estimated parameters of the empirical hyperbolic model as fitted by weighted least squares regression. The evidence in Table 3 suggests that the simple hyperbolic model fits the data reasonably well for the majority of ranks in most years. It is only in the years 1984 and 1987 that the estimated coefficients are significantly different from the expected values, as measured by a slope coefficient, β_{1t}, which is significantly different

	1984	1985	1986	1987	1988	1989
β_{0t}	0.01	0.01	0.00	0.01	0.00	-0.01
β_{1t}	0.49	0.55	0.67	0.56	0.72	1.55
Std. error for β_{1t}	0.23	0.33	0.22	0.22	0.24	0.33
R^2	0.52	0.38	0.55	0.56	0.51	0.48

	1990	1991	1992	1993	1994
β_{0t}	0.00	0.00	0.00	-0.00	-0.00
β_{1t}	0.86	1.02	1.03	1.03	1.14
Std. error for β_{1t}	0.33	0.26	0.31	0.27	0.27
R^2	0.43	0.60	0.51	0.56	0.55

Table 3: Summary regression results of the fit of the empirical gap factor model.

from unity at a 5% significance level. [7] A fruitful line of further research would be to try to find an explanation for why the empirical model appears to work reasonably well.

It should further be noted that the market capitalisations themselves can be regarded as being measured with substantial error. The complicated capital structures of many of the very large groups and companies necessitate somewhat subjective decisions to be made about which components of a group to include, and which not to, when measuring the size of the security. So, for example, Shell is measured without adding in the capitalisation of its sister company Royal Dutch (and the quoted market capitalisation can therefore be argued to represent only about 40% of the total relevant capitalisation). This stands in contrast with Vendome and Eurotunnel, for example, which comprise both the UK-registered plc companies and the corresponding companies registered in Europe.

3.4.2 Time varying nature of the gap factors

An important question as regards the performance of the FT-SE 100 over time is whether the structure of the index changes, e.g. whether the gap factors change with time. If the gap factors show an increase over time then it would suggest that the index were being 'stretched', with the larger stocks becoming larger faster than the smaller stocks. The stretching would imply long term outperformance by larger stocks which in turn will usually be stocks which had previously outperformed the market. Clearly trying to perform a time series analysis on the gap factors of each rank would be a large scale statistical exercise beyond the scope of this preliminary investigation, but we can get some impression of any changes by averaging the ranks across subsets of the index and then monitoring those averages. In particular we measured the median gap factors at each year end for the superlarge stocks, the seasoned index stocks and the marginal stocks. It is evident from Figure 4 that, on average, the marginal stocks' gap factors were fairly constant over time, varying around 1%. The seasoned stocks' gap factors showed an increase over time having moved up from around 2% before 1990 to just under 3% after 1990. In contrast, the gap factors for the super large stocks showed a marked decrease from around 9% to around 5%, indicating that the gaps had narrowed at the very largest end of the market.

3.5 Some dynamic aspects - how shares move within the index

Although the data which were available to us were not strictly suitable for any detailed analysis of dynamic elements of the FT-SE 100, we provide some preliminary results which may help in focusing further research. The dynamic aspects which we do consider in this section include measurement of the average changes in rank made by securities within the index. In addition

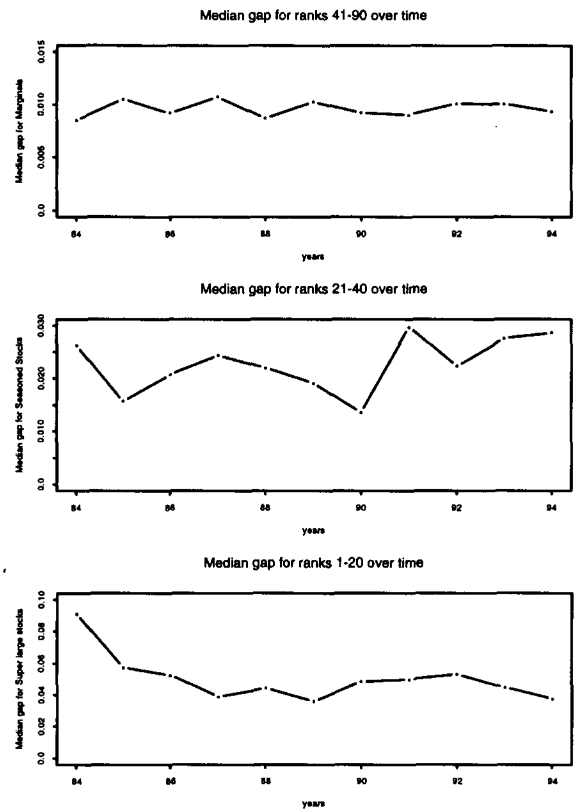

Figure 4: Time series of median gap factors for subsets of the index

we examine the rate of market capitalisation growth associated with the securities in each of the ranks.

3.5.1 Transition matrix

As a first step to analysing movements within the index we constructed an empirical transition matrix showing the distribution of ranks to which securities with a given rank at the start of the year had moved by the end of the year. A similar analysis was conducted by Hart & Prais (1956) and later followed up by Hart (1960) in their studies of business concentration in the UK. They, too, used transition matrices to investigate the changes in firm size, and also analysed the concentration using Lorenz curves.

Because of data restrictions we were unable to monitor the exact ranks of securities which had moved out of the index in a particular year, or into the index. This means that the analysis is somewhat incomplete, particularly for securities with low ranks, i.e. the marginal and just in stocks. Some of the securities moving from 'out of index' to within the index will have existed prior to the date of review and will have grown large enough to be included. Other companies will not have existed (e.g. privatisations) or will have acquired size through merger activity rather than any growth in the value of the issued stock. Those securities moving from within the index to 'out of the index' will be a similarly heterogeneous group of companies.

Securities within the index were grouped into bands of ten based on their ranks within the index, i.e. group 1 consisted of those securities ranked 1 to 10 (inclusive), group 2 consisted of those securities ranked 11 to 20, etc. The movements of securities between groups from the beginning of the year to the end of the year can be summarised in a transition matrix. At the start of each year each group of ranks comprised ten companies. Because there were eleven years of data available to us, we were able to compute ten one-year transition matrices. Hence, in the aggregate of the transition matrices there

Group at start	*Group at end of year*										
	1	2	3	4	5	6	7	8	9	10	Out of index
1	87	13									
2	10	74	14	2							
3	1	13	57	18	4	1	1		1		4
4			18	49	19	6	4	2			2
5			4	18	39	23	10	3	1	1	1
6			1	10	18	28	22	10	5	5	1
7				1	9	28	22	17	7	9	7
8			1	1	6	2	18	29	22	11	10
9					1	6	14	15	24	20	20
10					1	1	2	9	13	24	50
Out of index	2		5	1	3	5	7	15	27	30	

Table 4: Aggregate transition matrix between groups of ranks.

are 100 securities in each group of ranks. The aggregate distribution of those 100 securities across the groups of ranks at the end of the year is given in Table 4.

It is evident from Table 4 that the very large securities in group 1 tend to stay in group 1 from one year to the next, viz. of the 100 securities in group 1 at the start of a year, 87 of them remained within group 1 and 13 moved to group 2 by the end of the year. The movements between groups are much more pronounced for the smaller securities, e.g. for the 100 securities in group 10, only 24 of them stayed in group 10 until the end of the year, with one of the securities moving as high as group 5, and 50 of them moving out of the index. It also evident that 95 securities moved into (and out of) the index in the ten years, an average of just under 10 each year. However, this figure understates the number of entrants as it excludes those securities which entered the index after the start of a calendar year and then left the index before the end of that calendar year. We can also note that of the 95 'new entrants', $30+27+15=72$ of them were ranked below 80 at the end of the year of entry. Securities entering the index at higher ranks are generally the result of mergers, or privatisations, e.g. British Gas.

One possible explanation for the fact that highly ranked securities tend to stay highly ranked, whereas lower ranked securities are more volatile in their relative position is that the gap factors are larger for ranks close to one. Hence smaller securities will move relatively faster up and down the ranks for a given rate of price return than will larger stocks. Based on the empirical model described earlier for the gap sizes at various ranks, and using information about the distribution of stock returns for individual shares, it is conceptually possible to construct an expected distribution of the changes in rank which securities will experience.

3.5.2 Rates of growth in market capitalisation

Although we do not have rates of price return for the individual securities, we can approximate these rates of return by rates of market capitalisation growth, i.e.

$$\delta_{rt} = \log(\frac{C_{rt}(t+1)}{C_{rt}(t)}) \tag{7}$$

where $C_{rt}(s)$ represents the market capitalisation at time s of the security that was ranked r at time t and δ_{rt} is the 'force' of market capitalisation growth. Provided there are no capital changes in the security over the year, the rate of market capitalisation growth and rate of price return will be the same. Because we do not have the market capitalisation figures for securities which leave the index, simply measuring the market capitalisation growths for securities which are in the index both at the start and end of the year will produce biased estimates of market capitalisation growth. In the ten one-year market capitalisation growths, for example, there will be several ranks, especially those near 100, for which all ten market capitalisation growth rates are not computable because the security market capitalisation is not observable at the end of the year. To remove the bias we omitted from the sample of ten possible growth figures at each rank the x highest of the figures, where x is equal to the number of growth figures out of the ten which were not computable. The average growth figures for ranks below about 70 are therefore subject to very high sample variation because the samples are effectively much smaller than the samples at the higher ranks from which fewer stocks left the index in any one year.

The trimmed average rates of market capitalisation growth over one, four and seven years are given in Figure 5. We observe that the average market capitalisation growth across the ranks is fairly constant. However, there does appear to be some indication of slightly higher than average growth for stocks ranked just under 20.

4 Summary and possible areas of further research

This study has provided the preliminary evidence needed to start exploring the FT-SE 100 index and its behaviour relative to the market as a whole. We initially showed that although the FT-SE 100 and All Share indices were highly correlated in their rates of growth, the two indices do diverge on occasion. The divergence may be attributable to random fluctuations or to systematic differences caused by the construction and review process of the FT-SE 100 index. Determining a model for assessing which of the causes explains the divergences is an important area of future research.

The structure of the index was explored in more detail using the differences in size between the market capitalisations of securities at adjacent ranks. This gap factor analysis suggested that there may be structural sub-

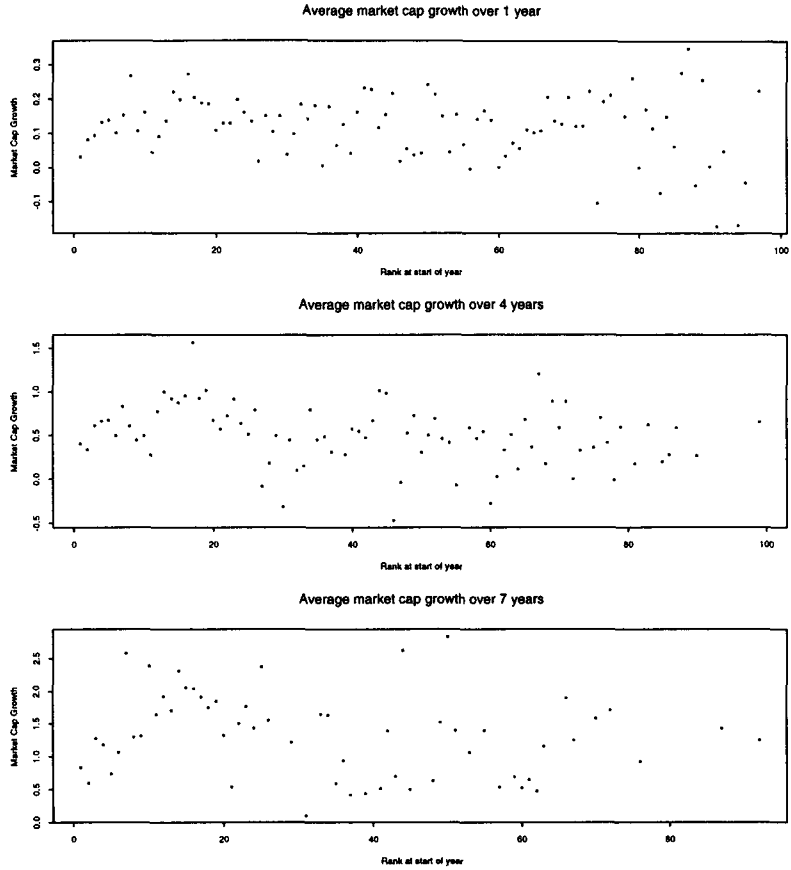

Figure 5: Average growth rates in gap factors over periods of 1, 4 and 7 years

sets within the index based on the market capitalisation differentials. The very large companies have, in general, very large gaps between them. The gaps decrease with size, although not completely smoothly. The gaps at ranks 20 to about 40 are somewhat larger than those at ranks 40 to 90.

Over time, the gap factors for the largest securities have shrunk quite markedly. In contrast, the gap factors between ranks 20 and 40 have increased slightly. Although this pattern of shrinking and stretching within the index may be period specific, a more detailed analysis of this phenomenon may generate greater insights into possible structural anomalies within the index. As a first step to modelling the structure of the index, we parameterised the general shape of the index assuming that the decrease in gap factors was smooth over the ranks. The particular form of model we used was a simple, empirically-based hyperbolic function relating market capitalisation to rank. A more complete stochastic model of the evolution of market capitalisations for individual securities would establish a more formal and useful basis for testing whether any structural subsets exist. We hope that the empirical model may provide the necessary insights for helping to develop that stochastic model.

Notes

[1]The authors would like to acknowledge the support and helpful advice of the other members of the Working Party, namely Messrs Sos Green, Andy Adams and Eric Lambert. In addition, we would like to thank Professor David Wilkie for his useful comments and suggestions on various aspects of the paper.

[2]The index was scaled so that it took on a value of 1000.0 at the start of 1984

[3]22 days represents approximately one trading month.

[4]The runs test is also referred to as Stevens test or the Wald-Wolfowitz test.

[5]A similar process might also occur at the boundary between the FT-SE 250 index and Small Cap index. This effect will probably be very much smaller than that between the FT-SE 100 and FT-SE 250 as there is possibly less speculation about securities moving in/out of the two smaller indices.

[6]Trial plots suggested that it was actually only the largest five or six companies which differed markedly from the suggested model. Nevertheless, we have omitted the largest 20 because they may be considered to be different on *a priori* grounds.

[7]It should be noted that the results presented in this section relate to the fitting of the hyperbolic model to the gap factors in the index. The gap factors represent a fairly detailed level of analysis of the model and indeed the relevant regression analyses (which we do not report here) indicate that fitting the model to the market capitalisations themselves produces measures of goodness of fit which are much higher. The practical usefulness of the model, however, depends on its fit at the more detailed level, which is why the focus remains at this level.

References

[1] Adams, A.T., Bloomfield, D.S.F., Booth, P.M. & P.D. England (1993), *Investment Mathematics and Statistics*, Graham & Trotman (Kluwer Academic Press)

[2] Brumwell, J. (1984 to 1994), Notes on the Financial Times-Actuaries UK Equity Share Indices, *Journal of Institute of Actuaries, Transac-*

tions of the Faculty of Actuaries and British Actuarial Journal

[3] Dimson, E. & P. Marsh (1984), Hedging the Market: The Performance of the FTSE 100 Share Index, *Journal of Institute of Actuaries*

[4] FT-SE Actuaries Share Indices Steering Committee, *Guide to Calculation Methods*, Version 2

[5] Gell-Mann, M. (1994), *The Quark and the Jaguar*, Little, Brown and Company, London

[6] Goodacre, A. & M. Lawrence (1994), Price and Volume Effects Associated with Changes in the Constituents of the FTSE 100 Index, paper presented at INQUIRE (UK) Autumn Seminar, Edinburgh, 2-4 October 1994

[7] Hart, P.E. & S.J. Prais (1956), The Analysis of Business Concentration: a Statistical Approach (with comments), *Journal of the Statistical Society*, Vol 119, 150-191

[8] Hart, P.E. (1960), Business Concentration in the United Kingdom, *Journal of the Statistical Society*, Vol 123, 50-58

[9] McDonald, J.B. (1984) Some Generalized Functions for the Size Distribution of Income, *Econometrica*, Vol 52, No. 3, 647-663

[10] Patel, J.K., Kapadia, C.H. & D.B.Owen (1976), *Handbook of Statistical Distributions*, Marcel Dekker, Inc., New York and Basel