This file is a manuscript of a paper which went on to appear as:

Ajay Shah and Susan Thomas. Market microstructure considerations in index construction. In *CBOT Research Symposium Proceedings*, pages 173–193. Chicago Board of Trade, Summer 1998

Market Microstructure Considerations in Index Construction

Ajay Shah Susan Thomas*

15 March 1998

Abstract

Illiquidity of securities in a market index generates noise in the index, increases the tracking error experienced by index funds and increases basis risk on an index futures market.

This paper proposes methods where liquidity of alternative stocks explicitly influences index construction. These methods are particularly appropriate on markets where liquidity is accurately observed, such as the open electronic limit order book. The results obtained in applying these ideas in the construction of India's NSE-50 index are reported.

Keywords: market index construction, market microstructure, index autocorrelation, non-trading, non-synchronous trading, index futures, spot-futures arbitrage, basis risk, contract design.

*Indira Gandhi Institute for Development Research, Goregaon (E), Bombay 400065. Email `ajayshah@igidr.ac.in` and `susant@igidr.ac.in`. We gratefully acknowledge the financial assistance provided by the USAID Financial Institutions Reform (FIRE) project. These ideas were influenced by interactions with Ashish Chauhan and other members of the National Stock Exchange Index Committee. Useful comments and suggestions were received from seminar participants at IGIDR and the CBOT Research Symposium.

Contents

1 Introduction

The ownership of a market index which is widely used in research and in financial products, like index funds or index derivatives, is valuable. This has led to significant investments, worldwide, into developing new indexes. Hundreds of market indexes exist in the world today. In this article, ideas from market microstructure are used to obtain improved methods for index construction.

Of the various kinds of indexes which can be constructed, the market–capitalisation weighted index has come to the fore in recent decades. Two major reasons can be offered for this. The capitalisation–weighted index is easier to implement as a tradeable portfolio, because price changes do not necessitate portfolio rebalancing. In addition, the capitalisation–weighted index is the 'market' portfolio which occupies a central place in the traditional one–period CAPM. Some of the most successful market indexes of the world, such as the S&P 500 in the US and the FTSE-100 in the UK, are capitalisation–weighted.

The actual construction of a capitalisation–weighted market index involves confronting two additional questions:

1. How large should the index be?

   Larger market indexes would seem to offer greater diversification, but diminishing returns rapidly set in. The alternatives presently in use in major markets range from small numbers like 40 (e.g. the CAC–40 in Paris) to 500 (the S&P 500 in the US).

2. What securities should be included in the index?

   The choice of stocks in an index is presently done using a variety of heuristics, based on the market capitalisation and trading volume of alternative securities, and the subjective judgement of the 'index committees' which are commonly constituted to oversee changes to an index.

Each index construction effort has to face these questions. The financial economics literature, however, is largely silent on these questions. It is well known that all capitalisation weighted indexes for a given market are highly correlated with each other. For example, the S&P 500 and S&P 100 indexes in the US have a correlation of 98% over the period 1976–95. Hence the differences between alternative market indexes are sometimes considered unimportant. However, alternative indexes can be quite different as far as their *liquidity* is concerned. On many markets, the transactions costs involved in portfolio trades for some indexes prove to be five times larger than those seen with other indexes.

The market microstructure literature has explored many consequences of imperfect liquidity for the informativeness, and for practical applications of market indexes:

- At the level of information, illiquidity is associated with non–trading or nonsynchronous trading. These problems generate a lag between the observed market index and the current state of the market. This leads to spurious autocorrelations and an upward bias in volatility estimates.

- In terms of practical applications, index illiquidity implies enlarged transactions costs when the index is traded as a portfolio. This serves to increase the tracking error of index funds, which reduces the usefulness of index funds. On index futures markets, index illiquidity widens the "no–arbitrage band" inside which mispriced futures do not yield profitable arbitrage. Fluctuations of futures prices inside the no–arbitrage band yield basis risk for users of the futures market, which makes the index futures market less useful.

  Index illiquidity can hence impact upon the usefulness of an index as a foundation for financial products based on the index.

This article offers an approach towards index construction which is aimed at creating a highly liquid index. High liquidity would result in reduced noise in the index time–series, and help enable financial products based on the index. The approach described here relies on measures of liquidity derived from 'snapshots' of the limit order book. This approach was developed in course of the creation of the NSE-50 index on India's National Stock Exchange (NSE), an open electronic limit order book market. The NSE-50 index has proved to be the most liquid of the indexes which aim to represent the Indian equity market.

This experience sheds new light upon the role of asset liquidity in designing financial instruments. One strand of the literature has treated the illiquidity of a market index as given, and worked towards discovering optimal research methods and trading strategies in the presence of poor liquidity. The approach taken here, instead, is to attempt the creation of an index which would be minimally contaminated by illiquidity. While the empirical results apply for India, this approach would be applicable in any index construction effort where a liquid index is desirable, and liquidity of the underlying cash market is observed. Examples include markets where the limit order book is observed, or where firm quotes from a set of dealers can be used to measure liquidity.

The remainder of this paper is organised as follows. Section 2 reviews the literature on liquidity and the uses of a market index, on the issues of informativeness and use in financial products. From this perspective, Section 3 surveys the ideas presently used in index construction. Section 4 describes the methods which were developed in the construction of the NSE-50. Section 5 describes the results that were thereby obtained. Section 6 concludes the paper.

2 Implications of Illiquidity of an index

When some or all components of an index are illiquid, two consequences follow: noise in the index (Section 2.1), and consequences of this for financial products based on the index (Section 2.2).

2.1 Informativeness of the index

The problems that derive from nonsynchronous trading of index components have been known since Fisher (1966). The more liquid components of the index trade frequently, rapidly impounding news, while the less liquid components of the index exhibit 'stale prices'. This phenomenon generates an upward bias in index returns variance, and spurious positive autocorrelations of index returns (Lo & MacKinlay 1990, Atchison, Butler & Simonds 1987). Miller, Muthuswamy & Whaley (1994) suggest that mean–reversion observed in stock index basis changes, which is presumed to be the outcome of index arbitrage, is significantly caused by nonsynchronous trading.

One common method used in index calculations uses the latest transaction price (LTP) for each security. Every time a trade takes place, the LTP is updated, giving a fresh value for the index. Under these conditions, the index will exhibit fluctuations even if true prices are constant, when the LTP for a security fluctuates between bid and ask prices. This generates the problem of the 'bid–ask bounce'.

In an index with N stocks, let stock i have a weight w_i in the index and a bid–ask spread of $s_i\%$. The trades which continually take place on individual stocks, moving prices randomly between bid or ask, yield noise in the computed index within a band which is $S\%$ wide where $S = \sum w_i s_i$. If we view the midway quote, (bid+ask)/2, as the true value of a stock, then each transaction price is higher or lower than the true value by $s_i/2$. If we assume that bid and ask prices are equiprobable, and that outcomes on all stocks are independant, the variance of the noise in the index works out to $\sum w_i^2 s_i^2/2$. When program trades are placed for the index portfolio, this independance assumption is violated yielding increased variance of the noise – e.g. when a program trade to buy the index is placed, all LTPs line up to the ask yielding an index value which is $0.5S\%$ above the true level.

These difficulties are present when working with daily returns; they gain much greater importance when working with intra–day data. In many studies, statistical procedures have been explicitly designed to account for this noise. Studies of lead–lag relationships between the spot index and the index futures market, such as Stoll & Whaley (1990), attempt to purge the spot index time–series of spurious autocorrelations before conducting tests. Another approach that has been explored in the literature (Jokivuolle 1995) consists of deriving econometric procedures for estimating the true value of an index when illiquid securities are present. For the present objective, it is useful to observe that when an index has highly liquid components, these issues become less important.

2.2 Use of index in financial products

Index illiquidity has significant practical implications when financial products based on an index are considered.

2.2.1 Tracking error in index funds

Index funds which invest in the underlying securities suffer transactions costs when trading in the index. Index funds which are based on illiquid indexes generally experience a larger 'tracking error' between fund returns and index returns. This diminishes the usefulness of index funds to investors, and reduces the ease with which index fund investors can monitor the behaviour of index fund managers.

2.2.2 Basis risk with index futures

The literature on the design of futures contracts (Tashjian 1995, Holland & Vila 1997) has emphasised the importance of hedging effectiveness in enabling success of a contract. The differences in the market model R^2 between equity portfolios and alternative market indexes prove to be fairly small, which suggests that the identity of a stock market index is relatively unimportant. However, actual hedging effectiveness is also influenced by *basis risk*. Illiquidity of the spot index exacerbates basis risk, and diminishes the hedging effectiveness of a futures contract based on the index. This serves to link the focus upon spot index liquidity of this paper with the emphasis upon hedging effectiveness of the literature on futures contract design.

Kook, Kwon, Lee & Choe (1992) evaluate the choice of index for an index futures market in Korea using the market model R^2 between portfolios and simulated index futures prices, where it is assumed that the futures would trade at their theoretical prices. This approach assumes away basis risk, and the issues connected with liquidity of the spot index. In their approach, the choice of index appears unimportant, since alternative market indexes yield values of R^2 which are 0.01 apart (Table 7, page 439 in their article).

Index futures markets rely on index arbitrageurs to ensure that the futures trade at close to their fair prices.[1] As discussed in Figlewski (1992), three difficulties faced by index arbitrageurs are directly related to illiquidity of index components: (a) the reported index level may be out of date, through 'stale prices', (b) purchasing or selling the index incurs transactions costs, and (c) arbitrageurs suffer execution risk through the time elapsed in executing trades in all index components. Each of these problems generates costs for arbitrageurs, and widens the 'no–arbitrage band' surrounding the fair value on the index futures market.

Speculators are the other major source of order flow on an index futures market. Basis risk reduces the usefulness of the index futures market as a vehicle for speculators who seek to express views about the spot index.

2.2.3 Market manipulation

Market–based manipulation of the spot index is less feasible with a highly liquid index. This helps inspire confidence of users, and improves the probability of success of a contract (Gorham 1994). Cornell (1997) illustrates these problems in futures on the Bond Buyer Municipal Bond Index (BBI), an index which is calculated using prices sampled from six brokers. The futures are traded on the CBOT with cash settlement. The spot index is composed of 40 fairly illiquid securities. This illiquidity leads to basis risk for users of the futures market. The paper describes an event, at 1:15 PM on 19 December 1995, where the BBI spot and futures both failed to move in response to an announcement by the Fed which strongly influenced every other interest rate futures contract. This could be owing

[1] Neal (1993) shows that there is little program trading associated with index options arbitrage. Hence we focus on arbitrage between the cash index and index futures when we analyse the role of index liquidity in the cash–derivatives linkage.

to sampling errors, manipulation of the spot market by futures traders, or misreporting of the prices used in computing the spot index.

2.2.4 Information processing by traders

The standard assumptions of neoclassical economics do not recognise a role for limitations of human information processing. The empirical experience of contract design suggests that these limitations might matter. The noise implicit in reported index values discussed in Section 2.1 above could adversely impact upon the ability of traders to process information, and trade efficiently, in realtime. Some of the sources of execution risk described by Figlewski (1992) would not exist in a market if traders could adequately process the information available to them. Similarly, the displacement of silver futures based on the metric system on US futures exchanges by contracts expressed using British units is an example where traders favoured a contract which involves the reduction of just one floating point computation. The time–series analysis required to extract the true signal from an index contaminated with stale prices and bid–ask bounce requires much more computation.

2.3 Heuristics used in working with illiquid indexes

Practitioners have developed a variety of heuristics which help cope with these problems:

- Instead of trading in the full index, index arbitrageurs often use a small basket of highly liquid stocks which is highly correlated with the index.[2] This strategy is a tradeoff between reduced transactions costs and the unsystematic risk of the less–diversified basket.

  Figlewski (1984) and Peters (1985) show evidence suggesting that stock index futures markets might exhibit a 'disequilibrium' in the early months after commencement of trading, where skills of arbitrageurs are not fully well established. This period might be one in which arbitrageurs on a given market are building human capital in these nuances of index arbitrage.

- Similarly, index funds often try to track a desired index by owning a portfolio which does not fully replicate it.

These procedures do have their limitations. There is some evidence which suggests that the cost of trading an index has an impact upon the success of products based on the index. Examples of this include futures on the Value Line index in the US (Thomas 1995)[3], and index futures at the South African Futures Exchange (SAFEX) (Risk 1996).[4] It is clear

[2]Neal (1993) reports that in 1989, the cash leg of an S&P 500 futures arbitrage used 375 stocks on average. The average for arbitrage involving the the NYSE composite index was 292, even though the index contains over 1600 stocks.

[3]The Value Line index, which was used for the first index futures market in the world, is hard to replicate since it is calculated using the geometric mean of prices. This is considered to be one of the factors underlying the demise of the contract.

[4]Until 15 March 1996, SAFEX index contracts were based on an index which had 80% of the market capitalisation of the market. This index was not adequately liquid for efficient replication. In 1995,

that these problems would be relatively muted, and these heuristics would be less relevant, if an index were highly liquid in the first place.

3 Existing Methods of Index Construction

There is a broad consensus today that market–capitalisation weighted index is the most important kind of index. For example, Table 1 of Canina & Figlewski (1995) lists twenty major stock index futures contracts of the world: fifteen of them are market capitalisation weighted.

The procedures adopted by many prominent indexes predate the modern market microstructure literature. Size has been an important factor driving selection of securities for indexes like the S&P 500, and the strong correlation between size and liquidity in markets like the US implies that criteria based on size are also, to some extent, criteria on liquidity. However, in other markets, the correlation between size and liquidity is often diminished, which renders size a poor proxy for liquidity.[5]

In this context, the rules used by two relatively recent efforts in index construction are reviewed below: that of the Morgan Stanley Country Indices (MSCI) and the IFC Indexes.

- The US financial company, Morgan Stanley, produces a widely used family of country indices called the Morgan Stanley Country Indices (MSCI) covering 45 countries. The MSCI methodology[6] aims to capture 60% of the (world) market capitalisation subject to (a) industry representation, (b) maximising liquidity, (c) maximising "floating stock", and (d) minimising cross-ownership. The rationale underlying the choice of 60% is not disclosed.

  Of the four constraints, "floating stock" is known to be a determinant of liquidity (Demsetz 1968, Kothare 1997) and is likely to be unimportant once liquidity has been accurately quantified. Morgan Stanley uses trading volume as a measure of liquidity, which is limited by the fact that trading volume does not measure transactions costs.

- The International Finance Corporation (IFC) produces a set of market indexes covering 45 countries (IFC 1997). IFC creates three groups of indexes: the IFC Global Indexes which aim to capture 70–75% of market capitalisation, the IFC Investible Indexes which focus on the limits on foreign investment placed by various countries, and the IFC Tradeable indexes which are designed to be convenient for use in financial products.

  The liquidity filters that IFC uses for the tradeable indexes are: (a) trading frequency (the fraction of days for which atleast one trade took place) above 95%, (b) trading volume above $5 million a month and above $100 million a year and (c) market capitalisation above $200 million.

---

the March 1996 futures contract traded at a price which was 4.55% away from the fair value. Hence SAFEX created a smaller index, the Alsi-40, based on the top 40 shares by trading activity and market capitalisation. With this index, basis risk dropped considerably.

[5] India serves as a useful illustration of this problem. The largest company in India, by size, is one of the least liquid securities in the country.

[6] Accessible at `http://www.ms.com/MSCI` on the world wide web.

As is the case with the MSCI, the metrics of liquidity used, trading frequency and trading volume, do not directly measure transactions costs.

The research described in this article was motivated by the silence of the existing financial economics literature on the heuristics which are being used, and refined on an ongoing basis, by index construction efforts in the real world. The next section will focus on a conception of liquidity that follows from the modern market microstructure literature, and apply it in index construction.

4 Issues in Index Construction

4.1 Measuring index liquidity

A variety of metrics such as trading volume, trading frequency, etc. are often used in measuring liquidity. The market microstructure literature has focussed upon *transactions costs* as being the essence of illiquidity. This emphasises the relevance of metrics of liquidity such as the bid–ask spread and market impact cost.

In this paper, we focus on the open electronic limit order book market. This form of market organisation has attractive theoretical properties (Glosten 1994), and has proved to be extremely successful in practice. All but two derivatives exchanges that have been started worldwide after 1986 are electronic limit order book markets (Melamed & Tamarkin 1996). Of the 23 stock exchanges in India, fifteen (which account for 99.9% of the trading volume) are electronic limit order book markets.

The metric of liquidity which directly measures the transactions costs faced by traders on a limit order book market is *market impact cost*. This is defined as the percentage degradation, as compared with the midway quote, that is faced by a market order of a given size. If a security trades at 99/101, and if a market order for 1000 shares is executed at 105, then the market impact cost at 1000 shares is 5%.

If the full limit order book book is observed at a point in time, impact cost can be accurately calculated at all transaction sizes. A dataset which captures the entire limit order book, at a point in time, is termed a "snapshot" of the limit order book. Using a limit order book snapshot, a *liquidity supply schedule* can be calculated. Examples of this, for a security and for a portfolio, are shown in Figure 1.[7]

On markets like the NYSE, measuring transactions costs using the quoted spread has two deficiencies: (a) the true spread could often be tighter than the quoted spread, and (b) the execution cost faced for large transaction sizes is unpredictable. The open electronic limit order book market generates reduced execution risk for index arbitrageurs owing to *transparency of liquidity*, whereby the costs associated with a trade can be accurately known before the trade is placed.

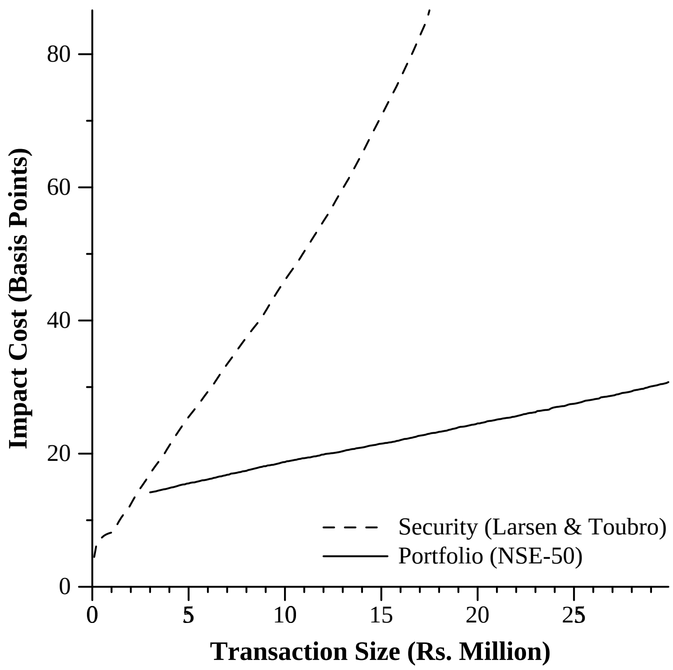
Even where electronic limit order book markets exist, comprehensive data about the limit order book is often unavailable to researchers and practitioners (Goodhart & O'Hara

[7] In the notation of Glosten (1994), the liquidity supply schedule is $R(q)$ where q is the value of shares purchased.

Figure 1: Liquidity Supply Schedule: Two Examples

The liquidity supply schedule shows the one–way market impact cost at all transaction sizes. Impact cost is the percentage degradation, as compared with the midway quote, that is faced by market orders. The average of impact cost on buy and sell transactions is shown in the graphs. The graph shows the liquidity supply schedule at one point in time, 2 PM on 6 June 1996, on India's National Stock Exchange for a security (Larsen & Toubro) and a portfolio (the NSE-50 index). The display for the NSE-50 index starts at Rs.3 million because that is the minimum transaction size required to meaningfully buy the index portfolio.

A market order for Rs.15 million of Larsen & Toubro faced an impact cost of 70.7 basis points, while a market order for Rs.15 million of the NSE-50 index faced an impact cost of 21.5 basis points. Impact cost for Larsen & Toubro was not observed beyond Rs.17.5 million because the limit order book did not support both buy and sell transactions larger than this.



1997).[8] As a consequence of the methods of this article, in March 1996, NSE established procedures to take 'snapshots' of the limit order book on all securities, three times a day.[9] This data is used in maintenance of indexes, and in other situations where accurate measurement of liquidity is required.[10]

4.1.1 Combining information across order book snapshots

The market impact cost observed for a given security is sampled thrice a day over a period of time. The market impact cost varies from one snapshot to the next, and some snapshots may yield a failure of execution of the required market order, if the limit order book is not adequately populated. The occurence of incomplete execution in some snapshots implies execution risk in index trades.[11] Rules governing modifications to the index set require summarisation of the impact cost observed over a host of snapshots over a range of dates. This requires a method for reducing numerous samples into a measure of location.

When a snapshot yields an incomplete execution, traders *can* access off–market sources of liquidity. While the impact cost that would be actually suffered is unobserved, it is estimated to be higher than what would be obtained in most snapshots where complete execution is obtained. The sample median is attractive for this purpose, both because it is non–parametric, and it is estimable as long as complete executions are obtained in more than half the snapshots.[12]

4.2 Number of securities in the index

The theoretical ideal for a market index is one that contains all stocks in the economy. This suggests that larger, more inclusive market indexes would work better. However, the inclusion of illiquid assets into the index leads to the difficulties listed in Section 2. Hence every index construction effort faces a tradeoff between liquidity and inclusiveness (Ross 1992).

For a given equity portfolio, the hedging effectiveness of an index futures market is determined by (a) the market model R^2 faced with respect to the index, and (b) the basis risk faced on the market. Larger market indexes are likely to yield improved R^2, though this may also go along with worsened basis risk.

[8] A recent paper which also uses the idea of snapshots of the limit order book is Jarnecic & McInish (1997).

[9] Trading takes place at NSE from 10 AM to 3:30 PM. The snapshots are taken at 12 noon, 1 PM and 2 PM every day.

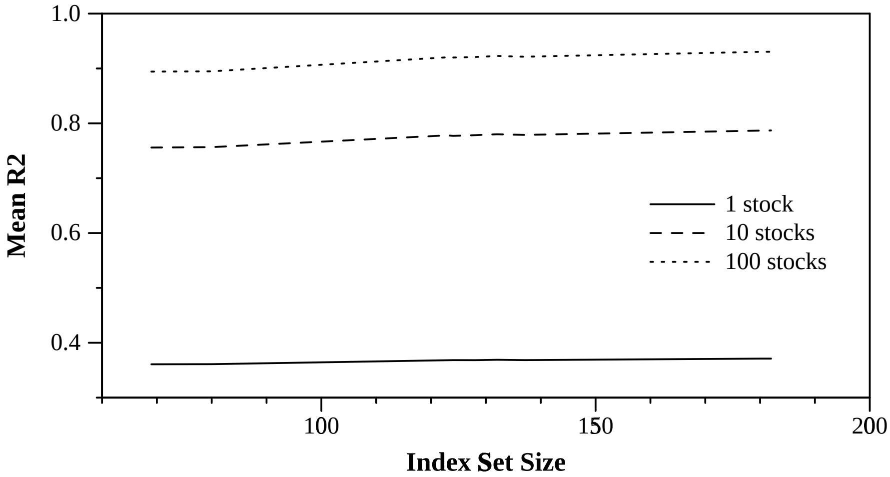
[10] There are 23 stock exchanges in India (Shah & Thomas 1997). Hence traders can exploit liquidity on other exchanges to supplement the liquidity available on the NSE. The NSE-50 effort takes a conservative view in only paying attention to liquidity in the largest market, the NSE.

[11] It should be observed that there are 23 stock exchanges in India, so market orders that fail on NSE are likely to obtain execution on other markets. Alternatively, traders could "work the trade" by placing limit orders at attractive prices.

[12] In the implementation of the NSE-50 index, heuristics were developed to impute impact cost when incomplete execution was obtained using the limit order book. Once this was achieved, the sample mean

Figure 2: Market model R^2 vs. Size of Index

For the nine market indexes which were created, measurement of hedging effectiveness was sought. Equally weighted portfolios of size 1, 10 and 100 were created at random from a universe of 1000 securities in India. The mean R^2 across 1000 random portfolios, in market model estimation between portfolio returns and index returns is shown in the graph. Monthly returns data was used.



In order to obtain empirical evidence about the hedging effectiveness implied by a variety of different alternatives, a set of nine market indexes were constructed over the period 1990–1995, using defined entry and exit criteria instead of fixed index set sizes.[13] For each of these indexes, the average market model R^2 obtained with respect to 1000 randomly chosen equally–weighted portfolios of different sizes is shown in Figure 2. These results are relatively uncontaminated by the statistical difficulties of Section 2 since they are based on monthly returns.

These results, of a modest improvement in hedging effectiveness with larger indexes, are not unlike those from many other countries. Individual securities have an average R^2 of 0.361 for the smallest index (which had 69 stocks) which rises to 0.371 for the largest index (with 182 stocks). At ten–stock portfolios, the R^2 ranges over 0.756 to 0.787, and at hundred–stock portfolios, the R^2 ranges over 0.894 to 0.93.[14]

is used as the measure of location. Details of these rules are available from the authors on request.

[13]For example, one of these indexes admits a security when it has size above Rs.5 billion and trading frequency above 90%, and drops a security from the index when the size drops below Rs.3 billion or the trading frequency drops below 80%. The number of stocks in the index is hence variable; the numbers reported in the text refer to the number of stocks in the index as of December 1995. The nine indexes were computed by varying these entry and exit rules. Details are available from the authors on request.

[14]The BSE–30 is a market index which is widely used in India. It obtains an average R^2 of 0.353 with individual securities, 0.737 for ten–stock portfolios, and 0.872 for hundred–stock portfolios.

4.3 Index computation in realtime

One common method used in index calculations uses the latest transaction price (LTP) for each security. Every time a trade takes place, the LTP is updated, generating a fresh value for the index. This generates noise owing to the bid–ask bounce, as described in Section 2.1.

This problem would be avoided if index computations are based on the midway quote, instead of the LTP. This has the additional advantage that quotes (or the best limit orders in a limit order book market) are likely to be updated before any trades take place, which diminishes the nonsynchronous trading problem. Some indexes already use such a policy, like the FTSE-100 in the UK. But many others, like the S&P 500 in the US, report values based on LTPs.

In order to gauge the economic significance of the noise induced by the bid–ask bounce, a simulation study was undertaken using trades and quotes on 8 January 1998. The noise induced by using LTPs proves to have a standard deviation of 0.026% in the case of the NSE-50 and 0.1% in the case of the NSE Midcap index. These values are large as compared with the resolution to which the index is reported (which is of the order of 0.005%). This noise is also significant when compared with a typical daily index standard deviation of 1.3%.

As with other microstructure–related problems in indexes, the choice of LTP versus the midway quote matters less for an index which is highly liquid. For less liquid indexes, using the midway quote is an easily implemented method to improve the signal-to-noise ratio of the index.

4.4 Calculation of the expiry value

On similar lines, the expiry value of an index is of unusual importance for financial products. For products like the CME's S&P 500 futures, the opening price of the next day (which is based on a call auction) gives the index level that is used for calculating payoffs for index derivatives.

In some markets, the reference level used is the LTP at the end of continuous trading (e.g. the CAC-40 in Paris). This is vulnerable to unusual volatility towards the close, and to market manipulation. There are two alternatives which might be preferred: (a) use of a call auction either at the close or at the next opening, or (b) use of the average of index values of the last 30 minutes as the reference level for calculating derivative payoffs. The NSE-50 today uses the latter option.

5 Construction of the NSE-50 index

5.1 Rules governing evolution of the index set

The evidence about hedging effectiveness (of Section 4.2), coupled with evidence about the impact cost of alternative index sets,[15] led to the choice of 50 as the size of the new index, which was disseminated from April 1996 onwards.[16] The ongoing modifications to the index set are based on two rules, which are applied every six months:

1. A minimum liquidity filter is applied for determining eligibility of securities into the index. This is based on the impact cost faced in transactions on the security which would be required if it were in the index, when doing index trades of Rs.5 million.[17] A security is considered eligible if the impact cost is below 1.5% in 90% of the snapshots of the preceding six months. Securities in the index which fail this test are removed from it.

2. An eligible security is admitted into the index, displacing the smallest security in the index, if the incoming security is atleast twice the market capitalisation of the outgoing security. This rule allows the index to evolve steadily, as new securities acquire size and liquidity, but prevents excessive changes which increase the transactions costs faced by index funds and index arbitrageurs.

Snapshots of the limit order book are only observed from March 1996 onwards. An independent effort was made in creating a historical time–series for the index starting from middle–1990, traditional measures of liquidity to guide the evolution of the index set.[18]

5.2 Results

5.2.1 Liquidity

Liquidity of the NSE-50 and other market indexes, over the six–month period from March to August 1997, is summarised in Table 1. This table illustrates the tradeoffs between inclusiveness (measured by index market capitalisation) and liquidity (measured by impact cost in doing portfolio trades). This evidence shows that NSE-50 is the most liquid index, as compared with the other four indexes, two of which are less inclusive than the NSE-50.

[15] At the time, there were 72 securities in India which met the first condition described ahead, i.e. that 90% of the order book snapshots should yield an impact cost below 1.5%. This evidence was based on a small dataset of order book snapshots and hence treated with caution.

[16] The Indian market is considered fairly illiquid, with trading volume concentrated in a few securities. This has prompted some observers, e.g. Barclay (1997), to suggest a market index with 20 to 25 stocks. Such a conclusion might flow from an analysis of trading volumes, but trading volumes do not measure transaction costs.

[17] The constraint in doing small index trades is the discreteness of round market lots. Index portfolios of Rs.4 million have a correlation of 0.9999 with the true index.

[18] Details of the methods used in the back–calculation are available from the authors on request. The full time–series is available from
`http://www.igidr.ac.in/~ajayshah/INDIAFINANCE/DB/nifty-data.html`.

Table 1: Impact cost in portfolio trades for alternative indexes

Average impact cost (in percent), using three snapshots of the limit order book on each trading day, from March to August 1997. The market capitalisation of the index set is in trillion rupees, as of 31 August 1997. All numbers reported here are based on order books for NSE (the largest market), though some indexes actually use prices from less liquid markets.

Index	Market Cap. (Rs. Trillion)	Impact Cost (%) at trans. size (Rs. Mln) 5	10	20
BSE-30	1.96	0.35	0.46	0.67
Barings India Index	1.59	0.29	0.36	0.50
IFC India Index	2.62	0.53	0.65	0.89
MSCI India Index	2.67	0.53	0.64	0.87
NSE-50	2.21	0.29	0.36	0.49

The one–way impact cost of 0.29% that we see with the NSE-50 (a market index with a market capitalisation of $60 billion) compares favourably with the index spread of 0.5% (Sofianos 1993) on the S&P 500 (a market index with a market capitalisation of $4 trillion). This is a satisfactory result given the traditional fears about poor liquidity which characterise the Indian equity market.[19] Since the S&P 500 is the basis of the largest index futures market in the world, this suggests that an index futures market based on the NSE-50 might not face serious constraints insofar as the obtaining a tight no–arbitrage band is concerned.

5.2.2 Comparison against the BSE-30

The index which was most used in India prior to this work was the BSE-30 index. The NSE-50 is a larger index than the BSE-30, with a market capitalisation of Rs.2.21 trillion as opposed to Rs.1.96 trillion. This improved diversification is associated with reduced volatility: the daily standard deviation of returns on the NSE-50 index is 1.66% as opposed to 1.72% for the BSE Sensex. The difference in volatility is more pronounced intra–day, owing to the inferior liquidity of the BSE – the volatility of the NSE-50 is 0.0435% per minute as compared with 0.0516% for the BSE Sensex.

The Indian equity market is composed of over 5,000 listed securities, most of which are extremely small and illiquid. Market model R^2 are hence calculated using random equally–weighted portfolios drawn from three universes: large capitalisation, medium capitalisation,

[19] This comparison is inaccurate insofar as the transaction size required to trade the NSE-50 is only Rs.5 million, i.e. $140,000, whereas index arbitrageurs working on the S&P 500 use larger transactions. However, this comparison is accurate when we think of the no–arbitrage band surrounding the fair value on the index futures market.

Table 2: Market Model R^2

Each cell in the table reports the average market model R^2 seen across 1,000 randomly chosen, equally weighted portfolios from three different size–classes in the Indian economy. Two indexes are used for the comparison, the NSE-50 index and the BSE-30 index.

	Small		Medium		Large	
Portfolio Size	NSE-50	BSE-30	NSE-50	BSE-30	NSE-50	BSE-30
1	0.09	0.09	0.15	0.15	0.31	0.31
5	0.25	0.23	0.38	0.36	0.62	0.61
20	0.40	0.37	0.59	0.54	0.80	0.79
50	0.46	0.43	0.63	0.59	0.86	0.85

and small capitalisation. The results are summarised in Table 2. In each size classification, for all portfolios larger than 1 security, the NSE-50 index has a higher R^2.

The evidence about the reduced volatility and better market model R^2 of the NSE-50 is relatively unsurprising; many 50–stock indexes would have these properties. However, the methodology of measuring impact cost and building it into index construction procedures has given these improvements without diminished liquidity. On the contrary, the market impact cost seen with index transactions on the NSE-50 is lower than that seen with the BSE-30, so that hedging users of futures based on the NSE-50 index are likely to face reduced basis risk as well. Hence the hedging effectiveness of a futures contract based on the NSE-50 index is likely to be higher than that of futures contracts based on the BSE-30. In this sense, the NSE-50 is not merely a different tradeoff between liquidity and inclusiveness. It dominates the BSE-30 on both counts.

5.2.3 Use in products

The first index fund attempted on an Indian index was the 'Bombay Fund', listed in London in 1993, which used the IFC Index. As the evidence about the liquidity of the IFC Index above would forecast, this fund has experienced a large tracking error. In contrast, index funds which use the NSE-50 index have experienced a tracking error (annualised standard deviation of the error in returns) of below 0.5%. In 1998, NSE will commence trading of index futures based on the NSE-50 index. This will be India's first exchange–traded financial futures.

Apart from the 'Bombay Fund' mentioned above and one other product, all publicly announced index–linked products based on India's equity market since the emergence of the NSE-50 index have been based on it.[20]

[20]Products based on the NSE-50 are: (a) UTI's India Access Fund (offshore), (b) BZW's Nifty Note (offshore), (c) SBI Mutual Fund's Magnum Index Fund (announced) and (d) NSE's index futuures and index options market (announced). The other publicly announced index–based products on India's equity market are (a) the 'Bombay Fund' mentioned above, and (b) a bond with returns linked to the BSE-30

5.2.4 Other applications

The ideas of this paper have been used for the creation of two other market indexes. At NSE, they are the basis of the NSE Midcap index, a 50–stock index drawn from securities which are not in the NSE-50 index using weaker requirements for liquidity. At Skindia Finance Ltd., these ideas were applied to the GDR market, a dealer market, in the creation of the Skindia GDR index (Jain 1997).

6 Conclusion

Alternative market indexes on any market are highly correlated, yet practitioners attach considerable importance to the choice of an index for use in financial products. The *liquidity* of alternative indexes, which can differ drastically, is perhaps an important dimension influencing these choices. One strand of the literature on transactions costs has tried to treat the illiquidity of a market index as given, and worked towards obtaining useful research or practical applications from index time–series which are contaminated by illiquidity. This paper treats liquidity at the security level in an economy as a given, and tries to obtain a relatively liquid index.

The ideas shown here could be useful for creating market indexes in any market the limit order book is observed. This includes electronic limit order book markets (e.g. stock exchanges in London, Toronto, Paris, Tokyo, etc.), manually operated exchanges which have limit order books (e.g. the NYSE), and dealer markets where the liquidity supply schedule can be estimated from firm quotes from a set of dealers (e.g. the GDR market).

The lack of spot market liquidity has been a hurdle in the progression of the 'spiral of innovation' in financial products (Merton & Bodie 1995) in many markets. The application of these ideas might help create market indexes which speed up progress through the ladder of sophistication in product development.

done by ICICI in 1996. This enumeration does not cover OTC products which might have taken place without public disclosure.

References

Atchison, M. D., Butler, K. C. & Simonds, R. R. (1987), 'Nonsynchronous security trading and market index autocorrelation', *Journal of Finance* **42**(1), 111–119.

Barclay, W. (1997), Index derivatives trading and the Indian equity market: Impact, regulation and contract design, Technical report, Financial Institutions Reform and Expansion (FIRE) Project, USAID, Price Waterhouse LLP, USA.

Canina, L. & Figlewski, S. (1995), Program trading and stock index arbitrage, *in* R. A. Jarrow, V. Maksimovic & W. T. Ziemba, eds, 'Finance', Vol. 9 of *Handbooks in Operations Research and Management Science*, Elsevier Science B.V., pp. 315–339.

Cornell, B. (1997), 'Cash settlement when the underlying securities are thinly traded: A case study', *Journal of Futures Markets* **17**(8), 855–871.

Demsetz, H. (1968), 'The cost of transacting', *Quarterly Journal of Economics* **LXXXII**, 33–53.

Figlewski, S. (1984), 'Explaining the early discounts on stock index futures: The case for disequilibrium', *Financial Analysts Journal* **40**(4), 43–47.

Figlewski, S. (1992), Stock index arbitrage, *in* Newman, Milgate & Eatwell (1992), pp. 567–571.

Fisher, L. (1966), 'Some new stock market indexes', *Journal of Business* **39**(1), 191–225.

Glosten, L. R. (1994), 'Is the electronic open limit order book inevitable?', *Journal of Finance* **49**, 1127–1161.

Goodhart, C. A. E. & O'Hara, M. (1997), 'High frequency data in financial markets: Issues and applications', *Journal of Empirical Finance* **4**(2–3), 73–114.

Gorham, M. (1994), Stock index futures: A 12-year review, Technical report, Chicago Mercantile Exchange.

Holland, A. & Vila, A. F. (1997), 'Features of a successful contract: Financial futures on LIFFE', *Bank of England Quarterly Bulletin* **37**(2), 181–186.

IFC (1997), *The IFC Indexes: Methodology, Definitions and Practices*, International Finance Corporation, Washington D.C., USA.

Jain, V. (1997), The Skindia GDR index, *in* T. Waghmare, ed., 'The future of fund management in India – 1997', Invest India – Tata McGraw–Hill Series, Tata McGraw–Hill, chapter 12, pp. 81–84.

Jarnecic, E. & McInish, T. H. (1997), An empirical investigation of the option value of the limit order book on the Australian stock exchange, Technical report, University of Sydney.

Jokivuolle, E. (1995), 'Measuring true stock index value in the presence of infrequent trading', *Journal of Financial and Quantitative Analysis* .

Kook, C. P., Kwon, Y.-J., Lee, W. H. & Choe, H. S. (1992), Selection of underlying index for stock index futures in Korea, *in* S. G. Rhee & R. P. Chang, eds, 'Pacific–Basin Capital Markets Research', Vol. III, Elsevier Science Publishers B. V., pp. 427–440.

Kothare, M. (1997), 'The effects of equity issues on ownership structure and stock liquidity: A comparison of rights and public offerings', *Journal of Financial Economics* **43**, 131–148.

Lo, A. W. & MacKinlay, A. C. (1990), 'An econometric analysis of nonsynchronous trading', *Journal of Econometrics* **45**, 181–211.

Melamed, L. & Tamarkin, B. (1996), *Escape to the futures*, John Wiley.

Merton, R. C. & Bodie, Z. (1995), A conceptual framework for analyzing the financial environment, *in* 'The Global Financial System: A Functional Perspective', Harvard Business School Books, chapter 1, pp. 3–32.

Miller, M. H., Muthuswamy, J. & Whaley, R. E. (1994), 'Mean reversion of Standard & Poor's 500 index basis changes: Arbitrage–induced or statistical illusion?', *Journal of Finance* **XLIX**(2), 479–513.

Neal, R. (1993), 'Is program trading destabilizing?', *Journal of Derivatives* **1**, 64–77.

Newman, P., Milgate, M. & Eatwell, J., eds (1992), *The New Palgrave Dictionary of Money and Finance*, The Macmillan Press.

Peters, E. (1985), 'The growing efficiency of index futures markets', *Journal of Portfolio Management* **11**(4), 52–56.

Risk (1996), 'Equity excitement', *Risk* **9**(5).

Ross, S. A. (1992), Stock market indices, *in* Newman et al. (1992), pp. 582–588.

Shah, A. & Thomas, S. (1997), Securities markets, *in* K. S. Parikh, ed., 'India Development Report 1997', Oxford University Press, chapter 10, pp. 167–192.

Sofianos, G. (1993), 'Index arbitrage profitability', *Journal of Derivatives* **1**, 6–20.

Stoll, H. R. & Whaley, R. E. (1990), 'The dynamics of stock index and stock index futures returns', *Journal of Financial and Quantitative Analysis* **25**, 441–468.

Tashjian, E. (1995), 'Optimal futures contract design', *Quarterly review of economics and finance* **35**(2), 153–162.

Thomas, S. (1995), The saga of the first Stock Index Futures contract: Was it a case of the market using the wrong model and not learning?, Technical report, Case Western Reserve University.