HAGFRÆÐISTOFNUN HÁSKÓLA ÍSLANDS

Hagfræðistofnun Háskóla Íslands
Odda v/Sturlugötu
Sími: 525-4500/525-4553
Fax nr. 552-6806
Heimasíða: www.hag.hi.is
Tölvufang: ioes@hag.hi.is

Skýrsla nr. R97:01

Infrequent trading and the stock-index: A Kalman-filter approach to estimation

October 1997

Foreword

Presently the financial environment in Iceland is undergoing momentous changes. Although a gradual process of reform and liberalisation can be traced back to the mid-eighties, it is only during the nineties that the trend has gained decisive momentum. An important ingredient of this ongoing transition is the appearance and rapid growth of an equity market. Because of the small size of the Icelandic economy, its fledgling stock market is facing its own particular problems, some of which have not attracted much attention in the context of larger and more mature markets. This study is intended as a contribution to an emerging discussion of efficiency, information, and risk in the Icelandic stock market which are rapidly becoming topics of great practical relevance.

The present report is based on research sponsored by the Research Contribution of the Icelandic Banks and carried out by Sigurður Ingólfsson. It has derived great benefit from the inspired supervision of Professor Guðmundur Magnússon and Associate Professor Helgi Tómasson. Sigurður Pétur Snorrason of the *Icelandic Stock Exchange* and Sigurgeir Örn Jónsson of *Kaupþing hf.* provided help with the data and Einar Hrafnsson of the *Icelandic University Library* made an invaluable contribution to more than one aspect of the work.

Some erroneous graphs contained in the first ten copies printed in October 1997 have been replaced in this second printing and a few other minor modifications were made by the author.

Institute of Economic Studies
University of Iceland
January, 1998

Tryggvi Þór Herbertsson
Director

Table of Contents

Overview

The present inquiry is organised in three parts. In the first, we provide a background discussion of index numbers, illustrate the basic mechanism that may lead to spurious statistical effects in a stock index when trading is sparse and introduce some problems pertaining to the microstructure of securities markets. In the second part we examine some important studies of infrequent trading effects to the time-series properties of returns in securities markets. The purpose of this is to gain some understanding of the principal sources of measurement error in stock markets. The third part studies some aspects of trading in the *Icelandic Stock Exchange* with an eye to symptoms of infrequent trading problems. Finally it is demonstrated how optimal estimates of a stock index can be obtained by means of a Kalman-filter technique. The way the present estimator is implemented possibly represents an improvement with respect to methods previously suggested. The reason is that it makes it possible to exploit the particularities of continuous market transaction data to reinforce estimates of the value of individual stocks. This is achieved by using a continuous-time filtering framework which eliminates the infrequent trading problem as it is usually defined, and also by taking advantage of simultaneous observations to determine the correct measurement error variance in the state space model. A concluding section suggests ways in which theoretical and practical aspects of this approach may be elaborated.

1 Background

1.1 Index numbers

Up to a point the index-number problem is not dissimilar to problems encountered in descriptive geometry. In either case, the attempt is to define an object by its projection...

Alexander Gerschenkron[1]

When the subject of index numbers is brought up, many are likely to think first of the consumer price index (CPI) while others may think of some security price index of the sort that is published in the newspapers every day. These are probably the most conspicuous among the great variety of economic index numbers in use. Economists use index numbers extensively in dealing with aggregates, usually for the purposes of comparison of some kind. To name but a few of these different uses, and consequently different types of index numbers, economists rely on price deflators to calculate real national income and growth and use quantity indices of such entities as industrial production. Purchasing power parities between currencies are used to compare cost of living, relative to a *numeraire*, and factor productivity indices and indices of import and export prices help to assess the competitiveness of industries and countries. Thus it seems that whenever an economist compares something to something else, she'll be using an index of some sort or another. In view of this important role of index numbers, it is not surprising that for more than a century index number theory has attracted the attention of many an eminent economist, including Irving Fisher, Ragnar Frisch, J.M.Keynes and Paul Samuelson.

A concise definition of index numbers may be found in Selvanathan and Prasada Rao (1994). They see an index number as "...an abstract concept ... used to measure the change in a set of related variables over time or to compare general levels in these variables over countries and regions." [2] This definition underlines some important aspects of index numbers. First, we note that an index is essentially an instrument of comparison. Second, this definition doubly emphasises the *abstract* nature of an index

[1] "Soviet Heavy Industry: A dollar Index of Output 1927-1937" *Economic Backwardness in Historical Perspective: A Book of Essays*. Cambridge, Massachusetts, 1962. p.250.

number. Being an abstraction, the index number neither occurs naturally, nor can it ever be observed. It is also evident that a *single* measure of some characteristic belonging to of a *set* of variables is necessarily the result of an inference or a chain of inferences.

From the point of view of the practitioner, the process of abstraction or the chain of inferences used to arrive at the index number is subsumed under an index number formula. The validity of this formula, then, is obviously of great practical importance. The question of how to choose an appropriate index number formula and how to justify its use in economic inference, is sometimes referred to as *the index number problem.*[3]

In what follows we will briefly outline two major directions of research in the field of index number theory and introduce some of the most commonly used index formulae. The point of departure is provided by a fairly general idea of a price-index number, but a convergence to stock-price indices should be implicit in the pattern of emphases.

1.1.1 A 'naive' approach

Before delving any deeper into index number theory, we show how it is possible to obtain some important index numbers, by means of a 'naive' or 'common-sense' approach. If we want to find some measure of changes in 'the cost of living' over time, e.g. between some base period and the current one, a logical way to proceed seems to be to choose a basket of goods and measure their change in prices between the endpoints of the period. Obviously, the selection of goods in the basket and their relative proportions will pose some problem if the group of consumers whose 'cost of living' we want to measure is heterogenous. We assume, however, that a reasonable basket of goods, typical of the budget of some 'representative consumer', has already been chosen. In the calculation of the index, a *vector of reference quantities,* $\mathbf{q}'_R = (q_1, q_2, \ldots, q_n)$ is taken to represent the basket of goods. In the base period, the prices of the reference goods can be defined by a base period price vector $\mathbf{p}'_0 = (p_{01}, p_{02}, \ldots, p_{0n})$ and correspondingly, we can think of the current price vector

---

[2] Selvanathan and Prasada Rao (1994), p.1. Henceforth: S&PR (1994). Most of the subsequent exposition of index numbers is based on this volume and explicit page references are only given for the most important points.

[3] E.g. S&PR (1994), p.8 and 12.

as $\mathbf{p}'_1 = (p_{11}, p_{12}, \ldots, p_{1n})$. Then the following formula may well be intuitively reasonable as a measure of change in the cost of living:

Equation 1

$$I_{01} = \mathbf{p}'_1 \mathbf{q}_R (\mathbf{p}'_0 \mathbf{q}_R)^{-1} = \frac{\sum_{i=1}^{n} p_{i1} q_{iR}}{\sum_{i=1}^{n} p_{i0} q_{iR}}$$

Thus if the cost of the basket of goods were 100 IKR in the base period and 180 IKR in the current period, we would feel justified in saying that the cost of living had increased by 80% in the interval. Two index numbers that are often considered more fundamental than others, the Laspeyres and Paasche index formulae, follow immediately from this definition by taking the reference quantity vector $\mathbf{q}_R$ as the base and current period quantity vectors, respectively:

Equation 2

$$I_{01} = \mathbf{p}'_1 \mathbf{q}_0 (\mathbf{p}'_0 \mathbf{q}_0)^{-1} = Laspeyres_{01}, \quad I_{01} = \mathbf{p}'_1 \mathbf{q}_1 (\mathbf{p}'_0 \mathbf{q}_1)^{-1} = Paasche_{01}$$

If the distance between the periods is great, some important consumption goods may be replaced by others in the meantime and even if all goods are available in both periods, it is possible that representative consumption patterns change over time. Therefore the Laspeyres and Paasche indices may be seen as extreme positions and some kind of average reference quantity vector be thought to be more appropriate than either extreme. This way of thinking leads to the Edgeworth-Marshal index, where the reference quantity vector is taken as the arithmetic average of the base and current period vectors, the Drobisch index, which is based on a geometric average of the two and the Geary-Khamis index, that uses the harmonic average of the base and current baskets as the reference quantity vector.

Still another way of looking at arithmetic mean-based index numbers is to define the index as an expenditure-share weighted average of *price relatives*, i.e.

Equation 3

$$I_{01} = \sum_{i=1}^{n} w_i \frac{p_{i1}}{p_{i0}}.$$

Taking the weights to represent expenditure shares in the base period, i.e.

Equation 4

$$w_i = w_{io} = \frac{p_{i0}q_{i0}}{\sum_{i=1}^{n} p_{i0}q_{i0}},$$

we obtain the Laspeyres index by direct substitution, and the Paasche index formula can also be derived without great difficulty. If the geometric average is taken, we obtain a class of 'Cobb-Douglas type' index numbers, where different ways of averaging the two period quantity vectors through the expenditure-share weights lead to a number of important index formulae, notably the Theil-Tornqvist index.

1.1.2 The index number problem

The Laspeyres and Paasche index number formulae are more than a century old, originating in the search for an intuitively satisfying measure of inflation, somewhat like our 'naive' reflections above.[4] However, attempts at formalising the field of index number theory also have a long history within economics, apparently beginning in 1922 with. Irwing Fisher's book *The Making of Index Numbers*. Historically, this endeavour has taken two main directions. One is the *functional approach*, which assumes that that the prices and quantities that enter an index formula are related to each other through the actions of rational economic agents. The assumption of a functional relationship between the price of a commodity on one hand, and the quantity consumed on the other, is fundamental in economic theory. However, taken literally, this means that keeping a reference vector of quantities fixed as prices vary is a logical inconsistency. Consequently, on the premises of the functional view the index number problem is conceptualised in terms of a relationship between expenditure and utility, instead of prices and quantities. Apart from the obvious theoretical appeal of such a unified view, which aims to build the microeconomic foundations of index number theory, practical issues such as 'the quality problem' can be treated in the more general framework of classical microeconomic quantification of consumer behaviour if a functional view is taken.

Already in 1924, Konus formulated a theory of the consumer price-index as a measure of the change in expenditure required to maintain a given level of utility, although the first comprehensive presentation of the functional approach appears to be due to

[4] Laspeyres' article appeared in 1871, the one of Paasche in 1874. For exact reference see bibliography in S&PR (1994)

Ragnar Frisch in 1936. Konus showed that for a fixed reference level of *utility*, some of the most important index formulae can be derived by varying the assumptions regarding the form of the utility function. Notably, the Laspeyres and Paasche indices, follow immediately in the expenditure-utility framework when the utility function is of the fixed proportions (Leontief) form. Which index number results depends on whether the base or the current quantity vector is used to form the reference vector.

This notwithstanding, the method of comparing the price of a fixed reference basket in different periods is a simple, direct, and intuitively appealing way of obtaining a price index number. In many cases the implied inconsistency may be less serious than it appears at first sight, e.g. if demand is inelastic and price changes small, leading to a situation where the interdependency of prices and quantities can safely be neglected. A different case where a functional relationship may be ignored in practice, is that of the stock index. As the reference quantity for each listed firm is the number of issued shares, there is no obvious economic relationship between changes in prices and changes in the composition of the basket.[5]

Assuming prices and quantities to be independent of one another in this sense, leads to an *atomistic approach* to the problem of formalising index number theory. This general assumption is less restrictive as a particular method of imposing structure on index number analysis, and consequently has resulted in a number of ramifications.

The position taken by Fisher in his original work is based on a set of *tests*, revelatory of the arithmetic properties of different index number formulae, providing some guidance to the choice of an appropriate formula from a multitude of possible ones. Obviously, the atomistic view can never replace the functional conception of index numbers in the context of economic theory. However, seen as a complementary way of dealing with the problem, it may be expected to yield useful additional insight and, to some extent, even rigorous guidelines for practical work.

1.1.3 The stochastic approach

On the assumptions of the test approach, the index number itself is considered as a single statistical measure of an underlying 'central tendency' in a particular set of observations. The tests then serve to provide some indication as to the precision and

[5] 'Splitting the stock will of course lower the price of each share in the same proportion, but changes in price will not influence quantities. Stock-index numbers that are corrected for splits and dividend payments are called 'yield indices' and the present study focuses on this type of stock index.

consistency of this measure. But strictly speaking the traditional tests only apply to the properties of the index formula itself, and do not provide any measure of the amount or quality of the information yielded by the individual observations. Now the purpose of the index is to identify the value of an abstract entity, like the 'cost of living', on the basis of a concrete set of data. The cost of living itself is unobservable and the individual price observations can only be thought to express it to a limited and varying extent. This implies that it would be desirable to possess some kind of indicator of the amount of 'data-related error' associated with a particular value of the index, as opposed to, e.g. arithmetic bias. Such a measure results immediately if each price movement is seen as a *measurement* of the unobservable central tendency of prices; then it is logical to interpret their degree of deviation from the estimate as its measurement error variance.

Emphasising this aspect of the index number problem leads to what has been called the *stochastic approach*. Pursuing this line of reasoning, Selvanathan and Prasada Rao show that it is possible and indeed desirable, to tackle the economic index number problem in a statistical perspective. In their view, it is in fact a *signal extraction problem*.[6] This means that the appropriate formal framework within which index numbers should be analysed, is neither basically microeconomic nor algebraic, as in the functional and test approaches respectively, but statistical. One benefit is that results of probability theory and rigorous statistical methodology are made available for the interpretation of index numbers, their estimation, and inference about their properties. A further advantage is that it becomes possible to obtain an objective measure of the *quality* of an estimate of the underlying abstract variable at each point in a well established sense of the word, e.g. as a standard error of estimate. In addition, a large number of interesting questions can be formulated as statistical tests and resolved in a rigorous manner. Explicitly defining the index number as a statistical estimator opens the way for research into its 'econometric' properties, which can serve as the basis for selection of the appropriate formula for a given purpose. In many situations, then, the economist will be in a position to claim that a resulting index value is the *optimal estimate* of the variable of interest, be it the 'true cost of living' or 'the rate of inflation'.

---

[6] S&PR (1994), p.5

will now be heteroscedastic due to possibly unequal weight of the error terms. If this problem is countered by assuming that the price-variance of a commodity is inversely proportional to its weight in the representative budget, a GLS estimator is BLUE and yields the Laspeyres index formula introduced earlier, as well as an estimator of its variance, which is approximately proportional to the degree of relative price variability. The Paasche index is easily derived in a similar way.[9]

Now this formulation may be thought to be unduly restrictive in that it does not allow for estimation of commodity specific components of the general movement in prices. Extending the model by adding dummy variables for the commodities, the following model is obtained:

Equation 7

$$p_{it}^{0} = \alpha_t + \beta_i + \zeta_{it}, \quad i = 1,2,\ldots,n \quad t = 1,2,\ldots,T.$$

Here the error terms are assumed to have zero mean as before, and be independent between commodities and over time. The error covariance matrix in each time period is diagonal, with each commodity specific variance term equal to a (possibly time dependant) constant, inversely weighted by the respective budget shares. As it stands this model is not identified. However, by adding the constraint that all commodity specific changes in relative prices sum to zero in each period, i.e.

Equation 8

$$\sum_{i=1}^{n} w_{io}\beta_i = 0,$$

and the parameters of this model, as well as their variances, can be consistently estimated by the method of maximum likelihood.

Now if we want to estimate the index as it changes over more than one period, the question arises whether a more efficient estimator can be obtained by treating the index as a system of simultaneous equations, as compared to the one resulting from separate estimation of each regression equation. If the error terms are correlated, either over time or over commodities, the answer is in the affirmative. In this case, in S&PR(1994), it is suggested that SURE estimation may be applied.[10] Opposite to the basic SURE model usually presented in introductory econometrics textbooks, here

[9] S&PR(1994), p.52-54

[10] Seemingly Unrelated Regression Equations. Judge et al. (1988), chapter 11, provides an exposition.

The basic idea of a stochastic perspective is of course no novelty in index number theory. It can be traced all the way back to the work of Edgeworth, and it is briefly treated by Frisch in his survey article in 1936. But systematic development of the idea that individual prices are *measurements* of an underlying unobservable variable, and that therefore index numbers are essentially stochastic in nature, irrespective of their sampling aspect, seems only to have taken place over the last 10-15 years.[7] Problems of sampling or "design", by contrast, concerning the selection of an appropriate reference basket, have traditionally recieved more attention in the literature, and it is important not to confuse these two ways in which a stochastic element may be thought to enter the index.

Drawing on a number of other authors, Selvanathan and Prasada Rao attempt to relate the fields of index number theory and regression analysis in a systematic way. They argue by a series of steps, gradually establishing the equivalence between increasingly sophisticated index numbers and corresponding statistical estimators.

To lay a foundation, they present a statistical model, based on particular assumptions about the process generating price data.

Equation 5

$$p_{it}^0 = \gamma_t + \varepsilon_{it}, \quad i = 1,2,\ldots n.$$

$$\text{where } p_{it}^0 = \frac{p_{it}}{p_{i0}}, \quad E(\varepsilon_{it}) = 0, \quad \mathrm{E}(\varepsilon_{it}^2) = \sigma_t^2 \mathbf{I}$$

Multiplying both sides by the original prices, and using these assumptions about the error structure, it is easily seen that the OLS estimator takes the form

Equation 6

$$\hat{\gamma}_t = \frac{1}{n}\sum_{i=1}^{n} p_{it}^0$$

and that it is the BLUE. Now the formula in **Equation 6** is just an equally weighted index of price relatives. The interpretation of $\hat{\gamma}_t$, is that it equals 'one plus the true inflation rate' in the intervening period, in other words a reasonable definition of a 'common trend in prices'.[8]

To derive expenditure-share weighted price index numbers it is possible to proceed in exactly the same way, only noting that a regression model corresponding to the index

---

[7] S&PR(1994), p.48-49
[8] S&PR (1994), p.50

correlation is assumed to exist over time but not between contemporaneous observations. The underlying model assuming such a set-up is the regression equation

Equation 9

$$p_{it}q_{i0} = \gamma_t p_{i0}q_{i0} + u_{it}, \text{ with } E(u_{it}) = 0, \quad E(u_{it}^2) = \sigma_t^2 p_{i0}q_{io}.$$

As indicated earlier, the Laspeyres price index is the GLS estimator of this model. After the appropriate transformation has been made (i.e. dividing both sides by $\sqrt{p_{i0}q_{i0}}$) this model can be rewritten for each period as an (OLS) regression equation in the transformed variables

Equation 10

$$y_t = \gamma_t x_t + u_t^*.$$

Stacking the T n-vectors in this equation one on top of the other, yields a SURE system:

Equation 11

$$\mathbf{Y} = (\mathbf{X} \otimes \mathbf{I})\Gamma + \mathbf{U}^*.$$

In this case the resulting point estimates are the same as those obtained by applying the Laspeyres index formula to the data in each period separately, because all the regressors in $\mathbf{X}$ are the same in all periods, but as demonstrated in S&PR (1994) on the basis of a particular data set that exhibits considerable correlation between disturbances in different periods, standard errors are significantly lower for simultaneous estimation.[11] A further advantage is that it allows the testing of cross-equation restrictions, such as $H_0: \gamma_t = \gamma_{t-1}$, $H_1: \gamma_t \neq \gamma_{t-1}$, i.e. to answer the question whether the unobservable inflation rate is constant over time or not. Tests for structural change can also be applied to investigate the effect of changes in the reference basket of commodities. A weakness of the SURE framework as presented here, is that the number of time periods included must be less than the number of commodities in the basket, otherwise the system covariance matrix becomes singular and the estimation technique breaks down.

[11] S&PR (1994), p.134

1.1.4 Fixed- and chain-base index numbers

Until now, we have only considered indices that yield estimates of price-change relative to some fixed base period and in fact, the SURE technique was introduced on the implicit assumption that the base-period weights are invariant. Under many circumstances however, especially if index number time series are being compiled over many periods or if there are other reasons to believe that the basket is changing extensively between the current- and base-periods, this may not be felt to be a reasonable assumption. When this is the case, *chain-base* index numbers are often employed, where the base period prices and quantities at each estimate are taken as the last period's current prices and quantities. Thus if I_{0t}^{Fixed} is taken to be the Laspeyres fixed-base index, then the corresponding chain-base index is

Equation 12

$$I_{0t}^{Chain} = I_{01}^{Fixed} I_{12}^{Fixed} \cdots I_{t-1,t}^{Fixed} = \prod_{s=1}^{t} I_{s-1.s}^{Fixed} = \prod_{s=1}^{t} \frac{\sum_{i=1}^{n} p_{is} q_{i,s-1}}{\sum_{i=1}^{n} p_{i,s-1} q_{i,s-1}}.$$

From the point of view of classical index number theory, there are strong grounds for preferring a chain-base to a fixed-base index, when the distance between the base and the current period is great, even if there is no quality problem associated with the choice of the commodity basket. The reason for this is that if the stochastic underlying 'true price' that the index number is supposed to measure is continuous, i.e. if it exists at every instant, then the value of an index number reported at discrete intervals will depend on the true path of the underlying process, and the bias thus incurred will be an increasing function of the interval size. To shed some light on this matter, we will look at an axiomatic approach to the index number problem presented by F. Divisia in 1925.[12]

Assuming that the underlying prices and quantities are continuous functions with respect to time, Divisia derives a continuous index number as a measure of the movement in prices by application of infinitesimal calculus. If we define the total aggregate value of consumption at time t as

[12] For a more detailed presentation and references, see S&PR (1994), p.115-120

Equation 13

$$V(t) = \sum_{i=1}^{n} p_i(t) q_i(t)$$

then the basic axiom underlying the Divisia index is that the following equality holds:

Equation 14

$$V(t) = P(t)Q(t),$$

with $P(t)$ and $Q(t)$ equal to the continuous-time price- and quantity-indices respectively. This assumption leads to the formulation of the continuous price index at each point in time as

Equation 15

$$\frac{dP(t)}{P(t)} = d[\ln P(t)] = \sum_{i=1}^{n} w_{it} d[\ln p_i(t)] = \frac{\sum_{i=1}^{n} q_i(t) dp_i(t)}{\sum_{i=n}^{n} p_i(t) q_i(t)}.$$

Strictly speaking, this formula is only valid for infinitesimal changes, whereas in practice, one is always interested in movements of prices over intervals of finite size. Substituting to obtain a formulation for a discrete interval we get

Equation 16

$$\ln P(t) - \ln P(t-1) = \int_{t-1}^{t} \sum_{i=1}^{n} w_{is} d[\ln p_i(s)] ds$$

which must be exponentiated to obtain the relative change in the price index. The expression on the right hand side of **Equation 16** is the integral of the path traced by the price movements in the interval and so the value of any index number obtained as a discrete approximation to the Divisia index will be *path dependent.* If we approximate dP by ΔP and dp_i by Δp_i in the expressions on the extreme left and right hand sides of **Equation 15**, we obtain the Laspeyres price index formula by some algebraic manipulation, in the form

Equation 17

$$I_{t,t+1}^{Laspeyres} = \frac{P_{t+1}}{P_t} = 1 + \frac{\sum_{i=1}^{n} q_{it} \Delta p_{i,t+1}}{\sum_{i=1}^{n} q_{it} p_{it}}.$$

In a similar way, it can be shown that the Paasche index is also equivalent to a discrete approximation to the Divisa time-continuous index. We have thus established that the value of both the Laspeyres and Paasche index numbers at each point in time, depends on their history, i.e. the *path* along which the underlying prices have attained this value. The degree of bias introduced into the index in this way, would then seem to be a function of its 'volatility', lending some further support to the view that index numbers should be analysed in a statistical framework. Indeed, an interesting early result, due to von Bortkiewicz in 1923, provides a decomposition of the so-called *Laspeyres-Paasche index gap*, i.e. the degree of divergence of the two formulas for a given data set, in terms of their covariance and their respective coefficients of variation.[13]

In general, economic analysis often depends on the assumption that prices can adjust continuously, and in some cases, as in research concerning the the stock market, explicit models based on the continuity assumption are in widespread use. This is a strong argument in support of the use of chain-base index formulae when prices are obtained by discrete sampling of such processes; the shorter the sampling interval, the closer the approximation to the Divisia theoretical index value at each point.

1.1.5 Stock-price index numbers

Stock price index numbers are similar to CPI-type indices in many ways. Just like the CPI, a stock-index is a single number, chosen to represent an unobservable abstract value, and used to derive period-changes in value. What exactly is thought to be measured by it, depends on the interpretation. At the very least, a stock-index number should express relative change in value of the actual portfolio used to compute it, or a larger one, including shares of all listed companies. In many of its uses, however, far greater demands are made.

---

[13] For a more detailed presentation and references, see S&PR (1994), p.25-27

Table 1: The intended purposes of a stock index[14]

1.	Reflect the trends in a stock market
2.	Make comparison between markets possible
3.	[Serve] as a benchmark [of investment performance]
4.	[Serve] as a basis for scientific research
5.	[Serve] as an indicator [of] the economic situation and [the value] of real capital
6.	[Serve to define] derivatives for speculation and hedging purposes

Thus, implicitly, the stock-index tends to be taken as a measure of changes in the value of all enterprises in the economy, or even, as in some asset pricing models, used to measure the value of a 'market portfolio', meaning *all assets* in the economy. By straightforward extension, then, volatility in the stock-index can be taken to represent investment risk.

On all of these interpretation, there will be a sampling problem involved in choosing a representative basket of stocks, just as with consumer goods in the case of the CPI. What is not as obvious, but no less real, is the existence of a quality problem. Even if a stock's quality is thought to be determined only by its return, it will be affected not only by business prospects, but also by factors such as dividend payments and stock splits. Thus one share of some company is far from being a standard commodity unit. Including a risk dimension, as implied by modern portfolio theory, complicates this quality issue still further.

Normally the prices occurring in stock transactions are thought to exist not only during the instant when the transaction takes place, but to represent discrete samples of an ongoing continuous price process. For this reason the Divisia continuous-time index is the only ideal measure of changes in a stock portfolio's value In practice, any discrete stock index will be path dependent, with the potential bias increasing with the length of the interval between the current and base periods.

Some efforts are made to tackle these problems by index compilers. Various 'investment-performance-indices', designed to minimise the quality-problem have been and still are calculated on a regular basis at stock exchanges around the world.[15] Many exchanges calculate chained index numbers at increasingly small intervals,

---

[14] *European Indices* (1997). Words in square brackets represent modifications by the present author.
[15] Fisher (1966) represents one of the first expeditions into this area.

supposedly among other things as a remedy against potential 'path bias'.[16] The advantages of different sampling schemes for stock index numbers have continually incited a great deal of attention and debate. But these issues all primarily concern the first moment of the index number: its expected value. In stark contrast to other index numbers, such as the CPI, the second moment of stock index numbers, their volatility, is a concept that has an important direct interpretation as investment risk. This aspect of stock-index numbers is of immediate concern in this report. By itself, focusing on the volatility dimension of the index number series leads quite naturally to a stochastic view of the index. To understand the relationship between index volatility and market risk, however, as well as the traditional framework for treating this relationship, we have to introduce a particular market model, and devote some discussion to the theory of stock markets at this point.

1.1.5.1 Markets and market institutions

Stock markets exist to facilitate the allocation of the economy's capital stock. Ideally, then, they should allocate capital resources to their most profitable uses at any given time. The secondary market for shares in limited liability companies plays an important role in this process, as the market valuation of company stock channels investment funds toward the most successful entrepreneurs and away from less successful ones. But for the entrepreneur the stock market represents a means of reallocating some of the *risk* of his venture, at the price of a share in eventual profit. Looking at things from the investor's viewpoint there is the opposite trade-off. Consequently, a different but no less appropriate view of stock markets is that they exist for the purposes of economic risk allocation. This view has become an increasingly important part of the economic theory of stock markets over the last few decades.

To make explicit the abstraction involved in the concept of securities markets, we could choose to talk about *the* stock market, defining it as the set of all opportunities to buy or sell shares in enterprises. In the actual occurrence, however, each concrete stock transaction takes place at a particular time and geographical location and within a given institutional framework. Furthermore, although it is seemingly advancing at a rapid pace, global integration of securities markets is far from accomplished and for

---

[16] In Stoll and Whaley (1990), the Standard and Poor 500 index is used, that is calculated on the

some time to come it will remain justified to treat most national securities markets in isolation.

The most important single institution in the stock market is the stock exchange. We may note that a stock exchange usually doesn't deal exclusively in company equity, but as a rule it will also trade in other forms of securities such as bonds. Neither are stock exchanges the only institutional structures where stocks change hands. Investment banks and other large investors may deal directly with one another, and in most countries an important 'over the counter' market exists for company stock and other securities. To take a specific example, the *Icelandic Stock Exchange* (ISE) only handled a little over half of the total trading volume in listed stock in 1996, a figure that also applies on average to other securities listed on the exchange.[17] Although many of the largest companies and most important types of bonds are listed, evidently only a fraction of the economy's total capital flow passes through the exchange. But looking only at trading volume would lead one to seriously underestimate the role of exchanges. Their relative importance as providers of information and liquidity or immediacy to the capital market probably far exceeds their relative size and therefore the study of the equity market in financial economics can often be conveniently simplified by using trading in stock exchanges as a proxy for the market itself.

Prices in the secondary market are most often assumed to be entirely decided by the forces of supply and demand and in many exchanges the trading process takes the form of a *continuous auction*. Consequently, in traditional finance theory, stock markets are regarded as a close approximation to a perfectly efficient market. In the traditional model the price of a share in a company is some fraction of its total market value, which in turn equals the net present value of the company's future profit flow. This quantity is unobservable by definition, but it can be estimated at any given time, on the basis of currently available information. If new and relevant information appears in every period, the estimated net present value can be continually updated. Adding the assumption that information innovations are reasonably approximated by a zero mean white noise process, this leads to a statistical model of the behaviour of individual stock prices as a *random walk*. Essentially, this means that future stock prices are always unpredictable on the basis of current information only.[18] The

---

average four times a minute.

[17] *ISE Annual Report 1996* , p.24

[18] The random walk model will be explicitly formulated in the next section.

random walk hypothesis leads to a number of empirical tests of the efficiency of financial markets, but as there are innumerable ways in which the pricing process could possibly deviate from a random walk, what they all have in common is that they can never definitely establish the random walk character of prices (i.e. the efficiency of the market). All they can be expected to do, is to fail to reject it with respect to some specific alternative.

A basically equivalent variation of the original random walk model can be derived by assuming a multiplicative and log-normally distributed innovation process and taking the *logarithms* of prices to follow a random walk. If the market is fairly stable and the differencing interval is small, the first differences of price logarithms can then be thought of as a good approximation to continuously compounded returns to investment in the stock. Evidently this important model is based on the idea that the underlying price adjustment process is continuous with respect to time, although only a finite number of discrete-time observations are available.[19]

1.1.5.2 A glimpse of the real-world

If we look at some major existing indices in many countries, we will see that an overwhelming majority is calculated by some slight variation of a Laspeyres or Paasche formula. Prominent exceptions exist, in particular the Dow Jones Industrial Average (DJIA) and the Value Line Composite Index (VLIC). The DJIA is an arithmetic average, calculated as the value of a portfolio composed of one share of each of a selection of 30 'blue-chip' companies. The Major Market Index (MMI) is an exact copy of the Dow Jones. The VLIC is a *geometrically* weighted average of portfolio prices. But these are the exceptions. The two remaining 'world famous' U.S. index numbers, i.e. the NYSE Composite index and S&P 500, follow in principle the value weighted Laspeyres-type formula discussed earlier. So do, as it seems, *all* the major European stock-index numbers.[20] Such indicators as the FT-SE100, compiled by the *Financial Times* and the *London Stock Exchange*, the DAX of the *Deutsche Börse*, HEX, compiled in the *Helsinki Stock Exchange* are all calculated according to some variation of the Laspeyres scheme, which for this reason may be taken as a benchmark. All the exchanges seemingly use rather *ad hoc* schemes for modification

---

[19] The basic reference in the continuous theory of finance is Merton (1990). We will enter into more details in a later section, introducing a rudimentary CAPM in continuous time.

[20] Source: *European Indices*: (1997).

of the index to account for such factors as new listings and delistings, splits, bonus issues, mergers, etc., which invariably differ between exchanges. The principal index of the exchanges is usually not adjusted for dividends, but often yield indices ar also published. Another common feature is that the principal index number in a European stock exchange is typically not an all-share index, although they are often published on the side, but an index containing a selection of stocks, e.g. 100 stocks in the FT-SE100. The main selection criteria are usually *value*, in terms of market capitalisation, and *activity* of the market in the particular stock, in terms of some measure such as volume or number of trades per day. The exact details of the selection criteria and the procedure for changing stocks in the index vary across exchanges. In all the exchanges included in the survey quoted earlier, some kind of chaining of bases is employed, but in general the procedure represents a compromise, where the base only changes at quite long intervals, such as annually or quarterly. Most European exchanges publish new index values at very short intervals, such as 15 or 30 seconds. The ISE all-share yield index differs from the index numbers in the major European stock exchanges in a number of ways.[21] Thus it is calculated from all stocks, only once each day and chained by changing the base in every period. It is corrected for both splits and dividends and the Paasche formula is used to obtain each link in the chain. New listings are taken into the calculation of the all-share index already on their second day of trading, which is atypical.

1.2 Infrequent trading

Infrequent trading, as a problem concerning stock-market indices is often thought to be formulated originally by Lawrence Fisher in 1966, and consequently it is sometimes referred to as the 'Fisher effect'.[22] If last reported prices are used to compute the index and there are stocks that do not trade in every period, the index number will be obtained on the basis of prices that originate in different periods. Instead of being an average of the cross section of stocks, the index number is then partly a time-average of observed stock prices. A clear and concise statement of the type of statistical problem involved, although in a simpler form, can be conveniently

[21] "Þingvísitala hlutabréfa". Radical changes in the principal index are planned to take effect in 1998, bringing it closer to the major European index numbers, but they will not be discussed here.

[22] Although Fisher. does mention the problem, in a subsequent example he seems to be talking about path dependency rather than infrequent trading. He dismisses the problem as minor in monthly data and does not relate it to time aggregation, although he quotes Working's article in another context.

taken from a note by Holbrook Working, appearing in *Econometrica* in 1960. Working is concerned with the time series analysis of commodity prices, but his result is only dependent on the random walk character of the time series in question, so it will apply to stock prices as well assuming the efficient market hypothesis.

Essentially, the infrequent trading problem is a special case of a more general 'time-aggregation problem', that occurs when a flow variable is sampled at time increments larger than those of its 'natural incidence' and the resulting subseries is treated as if it were the true series. Both issues are treated in a general way in Wei (1990), chapter 16, and a number of interesting results are derived for ARIMA models. However, in the context of stock prices and stock-price index numbers, as implied by the efficient market hypothesis, we are interested only in the simplest model of this class, the ARIMA(0,1,0). As this special case is in fact all we need to introduce this problem, it is most expedient to look at it in isolation. The following exposition is based on the original illustration by Holbrook Working:

Let X be a random walk process, or a *'random chain'* in Working's terminology, so that

Equation 18

$$X_t = X_{t-1} + \varepsilon_t \quad (t = 1,2,\ldots;\ E(\varepsilon_t) = 0,\ E(\varepsilon_t^2) = 1,\ E(\varepsilon_t \varepsilon_s) = 0 \text{ when } t \neq s$$

Now imagine that we have a number of observations on this series, that are split up into subgroups in the following way:

$t = 0$	1	2	3	4	5	6	7	8	9
$\varepsilon_t =$	+2.0	-1.1	-0.6	+0.3	+1.3	-1.0	+0.1	+0.7	-0.3
$X_t = 2.4$	4.4	3.3	2.7	3.0	4.3	3.3	3.4	4.1	3.8

The number of observations in each subgroups is m=3 in this example and could be taken to represent monthly prices, in which case the subgroups themselves stand for a quarter of a year. To obtain the price change over a quarter, as well as their correct variance, we would then take first differences between corresponding monthly observations, e.g. $X_9 - X_6$, $X_8 - X_5$, $X_7 - X_4$ etc. We would then obtain the series

Equation 19

$$\Delta X_{t(m)} = X_t - X_{t-m}, \text{ with } Var(\Delta X_{t(m)}) = m$$

If, instead of proceeding this way, *first* average the observations within each subgroup and *then* take first differences, the result will not be the same. For an arbitrary number of disaggregate intervals m, we then obtain:

Equation 20

$$\Delta X^{*}_{t(m)} = \tfrac{1}{m}\left(X_t + X_{t+1} + \cdots + X_{t+m-1}\right) - \tfrac{1}{m}\left(X_{t-m} + X_{t-m+1} + \cdots + X_{t-1}\right)$$

By repeated substitution, using the definition of a random walk in **Equation 18**, we obtain

Equation 21

$$\begin{aligned}\Delta X^{*}_{t(m)} &= \tfrac{1}{m}\left[X_t + (X_t + \varepsilon_{t+1}) + \cdots + (X_t + \varepsilon_{t+1} + \varepsilon_{t+2} + \cdots + \varepsilon_{t-m+1})\right] \\ &- \tfrac{1}{m}\left[(X_t - \varepsilon_t) + (X_t - \varepsilon_t - \varepsilon_{t-1}) + \cdots + (X_t - \varepsilon_t - \varepsilon_{t-1} - \cdots - \varepsilon_{t-m+1}\right] \\ &= \tfrac{1}{m}\left[(m-1)\varepsilon_{t+1} + (m-2)\varepsilon_{t+2} + \cdots + \varepsilon_{t+m-1} + m\varepsilon_t + (m-1)\varepsilon_{t-1} + \cdots + \varepsilon_{t-m+1}\right]\end{aligned}$$

Because the error terms are assumed to be mutually uncorrelated with unit variance, we can derive the variance of the first differences from this result in the following way:

Equation 22

$$\begin{aligned}Var(\Delta X^{*}_{t(m)}) &= \tfrac{1}{m^2}\left[(m-1)^2 + (m-2)^2 + \cdots + 1^2 + m^2 + (m-1)^2 + \cdots + 1^2\right] \\ &= \frac{2m^2+1}{3m}.\end{aligned}$$

On the random walk assumption, the theoretical autocovariance of first differences of the disaggregate series is zero. In the case of the aggregate series, proceeding in the same way as in **Equation 22** we obtain an expression for its autocovariance at lag one, as

Equation 23

$$Cov(\Delta X^{*}_{t(m)}, \Delta X^{*}_{(t-m)(m)}) = \frac{m^2-1}{6m},$$

implying that the theoretical first order autocorrelation of the aggregate difference series is

Equation 24

$$Corr(\Delta X^{*}_{t(m)}, \Delta X^{*}_{(t-m)(m)}) = \frac{m^2 - 1}{2(2m^2 + 1)}.$$ [23]

Clearly m, the number of periods averaged or aggregated, does not have to be very large to induce serious distortion of the time series properties of the resulting first differences. If bimonthly data is averaged to obtain monthly figures, m equals 2 and the variance of the series of first differences will be 1.5, 25% lower than that of the first differences of the corresponding point sampled series, and the autocorrelation will be 0.167 instead of zero. For a monthly series averaged to form a quarterly one, as in our example, the theoretical reduction in variance is nearly 30% and the autocorrelation induced by the averaging procedure is close to 0.21. As m gets larger the variance converges to $\frac{2}{3}m$, while the autocorrelation approaches a value of $+\frac{1}{4}$.

To understand why this happens, we need only to suppose that we correctly took the ('seasonal') first differences over the original disaggregate (or monthly) series, and only *then* averaged the results to obtain our aggregate (or quarterly) series. This would obviously yield a series of moving averages, in which reduced variance and positive autocorrelation would not be surprising. Arithmetically, however, the two procedures give exactly the same result.

Now let us take a fictious example, to link this conclusion to our study of stock index numbers. Imagine a stock exchange on which only three issues are listed, all representing equal proportions of total market capitalisation. Suppose further that all three stocks are perfectly positively correlated. Trading takes place in such a way that at 9 a.m. there is an auction, where shares in one of the three companies are bought and sold at the current price. At noon shares in a second company are auctioned off, and at closing time, say at 3 p.m., trading takes place in the shares of the remaining company. Subsequently the daily stock index is calculated as a weighted average of the most recent price information available for each issue.

To be explicit, we can suppose that weighted price relatives are averaged to obtain a Laspeyres index number for each day as a link relative and then a chain-base index is

---

[23] It may be noted for future refernce that the definition of autocorrelation of order k is $\rho_{xk} = \frac{Cov(x_t, x_{t-k})}{Var(x_t)}$ if x is a stationary series. The term will be used in this sense in what follows, but it is discussed more thoroughly in 3.2 below.

calculated. For each link relative, we are using the base period value shares (capitalisation) as weights, so

Equation 25

$$I_{t,t+1} = \sum_{i=1}^{3} w_i \frac{p_{i,t+1}}{p_{it}} = \frac{\sum_{i=1}^{3} p_{i,t+1} q_{it}}{\sum_{i=1}^{3} p_{it} q_{it}} = \frac{P_{t+1}}{P_t} = 1 + \frac{\Delta P_t}{P_t}$$

where the notation aims to make explicit the differencing involved in the calculation of the link relatives. The series ΔP_t will now have exactly the same properties as the series $\Delta X^*_{t(m)}$ in Holbrook Workings example. From this simple illustration, it is therefore quite clear what the effects on the covariance structure of the resulting index series would be. Admittedly, weighting of each observation by the inverse of the previous period level that yields the link relative would have to be accounted for, but the resulting effect is basically the same.

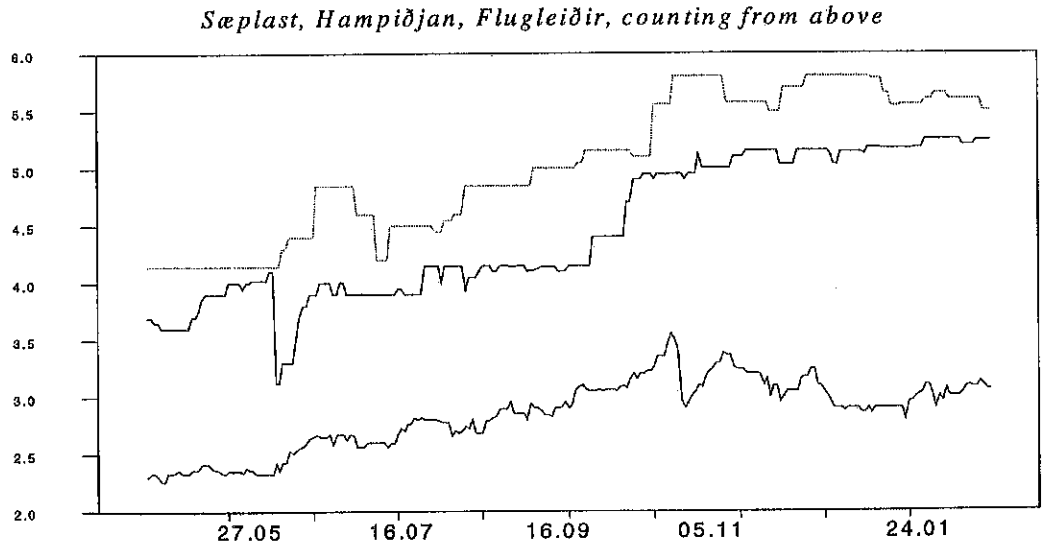
To make our imaginary stock exchange gradually more realistic, we might start by increasing the number of stocks and auctions, which would correspond to an increase in the aggregation parameter m in the example presented by Working. Further, relaxing the assumption that all m stocks are equally weighted in the index, we would still obtain the exact theoretical autocovariance structure of ΔP_t from a slightly modified version of the same formulas as long as the weighting scheme is constant over time. Relaxing the constraint that correlation between stocks is perfect, changes only the degree of the effect as long as the correlation is positive. This, however, is about as far as we can go with the basic model of time aggregation in 'first differences of averages in random chains' due to Working. The reason is that infrequent trading is a problem that arises in continuous-auction markets, where trading times are essentially unequally spaced and nonsynchronous across stocks. In other words, a transaction *can* occur in any stock at basically any instant, but it only *does so* at discrete unequally spaced intervals that vary across stock issues. In this context time aggregation effects arise, just as in the example, from the discrepancy between actual and hypothetical sampling times of the price observations used to calculate the daily index. The occurrence of transactions in such an environment obviously has a random character and ideally would have to be modelled as a stochastic process. Thus the

aggregation parameter m in the model above will be stochastic and this simple model breaks down.

Indeed, the major contributions towards modelling infrequent trading effects in stock markets go along these lines, assuming a directing process for the arrival rate of transactions and analysing the relationship between the true data generating process implied by the market model and the resulting subordinate process governing reported prices or, equivalently, returns.

As a final illustration to facilitate an intuitive grasp of the nature of likely infrequent trading effects on a stock index, we can look at a graph of the price series of three ISE stocks over a period of roughly 200 days.

Figure 1: Daily closing prices of three ISE companies May '96 to Feb '97



The lowest of the three price paths, that of shares in *Flugleiðir hf.* is obviously the most jagged one, and the topmost path, which belongs to *Sæplast* shares has extended flat parts where the price seems to be constant. The reason is that *Sæplast* is among the least frequently trading stocks in the ISE, while those of *Flugleiðir* lie at the other extreme. *Hampiðjan hf.* can be considered an intermediate issue in this respect. In general, a sum or an average of last reported prices of the three stocks at a given date will be composed of the price of *Flugleiðir* on that particular date, but older prices for the two other companies, and on the average the *Sæplast* price will be the oldest and *Hampiðjan* in between. If positive contemporaneous correlation is non-negligible, as is it is as a rule in the stock market, then the result will be a time average in the spirit of our previous illustration. It is shown in Holbrook Workings original paper that the

time averaging phenomenon will not affect the expected value of the of the aggregate series, i.e. it is unbiased. As we will see in later sections, this result carries over to the index. However, as will be readily understood from this graph, an index which in reality is a time average will exhibit *local* or *time dependent* bias at each point. Thus, if stocks in the index are positively correlated, but trading in some of them lags that in others, the index will exhibit downward bias when market prices are rising and upward bias when the market is falling. Another way to express this is to say that as an estimator of *changes* in the value of the underlying portfolio, the index is *biased toward zero.*

1.3 Market microstructure

In the preceding section infrequent trading was presented in a time series framework and shown to be likely to result in an index number with statistical properties that stand in direct conflict with the commonly accepted securities market model. But while sometimes, notably in the investigation that Holbrook Working's note aimed to correct, spurious properties of this kind may be attributed to sloppy handling of time series, this is obviously not the case where stock markets are concerned. Non-synchronicity is an essential feature of the way security markets work, and consequently an infrequent trading effect will be present whenever there is an attempt to sample a cross section of market prices as if they were simultaneous, although its magnitude will in general depend on how closely trading approximates a continuous process, in other words: how frequently stocks trade.

Looking at the problem in this way means seeing it as one of a larger class of issues that collectively make up the field of market microstructure theory. This relatively recent branch of financial economics springs from the observation that the actual pricing process in securities markets substantially deviates from what classical microeconomics would predict. The traditional theory of finance, based on the assumption that prices are set by some kind of a Walrasian auction process in the absence of transaction costs, regards observed transaction prices as true equilibrium prices and predicts that they will adjust instantaneously to new information, i.e. that securities markets will be perfectly efficient. Among the pioneers of market microstructure research are Cohen, Maier, Schwartz and Whitcomb (henceforth: CMSW), and their book *The Microstructure of Securities Markets* can still be taken as a basic reference, although other comprehensive surveys have become available since

its appearance in 1986.[24] According to CMSW, this research agenda gradually came into being in the late seventies, as a consequence of profound structural change in securities markets in the United States. The changes were in part institutional, brought about by the passage in 1975 of new legislation known as the Securities Act Amendments, which had the effect of deregulating the industry to a certain extent. This resulted in greater competition, growth, and increasing national integration of the securities trade in the U.S. In part, however, these changes sprang from technological progress. With the advent of the digital computer, stock exchanges gradually became more automated and increasingly capable of handling a "virtual explosion of the order flow" that ensued.[25] As recognized by many other authors, in the eighties increasing availability of high frequency transaction and quotation data on stock prices also served as an impetus to research in empirical finance, revealing an increasing number of anomalies and deviations from the traditional efficient markets model.

The aim of microstructure analysis as defined by CMSW, is to examine and explain pricing aberrations in a comprehensive framework, without abandoning the assumption of a basically efficient underlying securities market. Combining theoretical models with careful empirical scrutiny and comparison of existing market structure, it is hoped that positive results of microstructure analysis can yield important normative implications.

CMSW refer to the reasons for observed deviations from the classical model collectively as "frictions in the trading process", with an explicit reference to Newtonian mechanics. This epistemological position is based on a literal analogy made between the relation of market microstructure results to the market efficiency model on one hand, and the way in which the basically valid abstraction of the Newtonian model is only apparently refuted by the erratic path exhibited by "the flight of a feather from the leaning tower of Pisa". Because friction in the microstructure is only 'operational', i.e. caused only by the *way securities trade*, its presence does not necessarily entail a rejection of the efficient markets hypothesis. Thus markets can be 'informationally efficient', in the sense that there are no arbitrage opportunities, without necessarily being 'operationally efficient', meaning

---

[24] E.g. O'Hara (1994).
[25] CMSW (1986) p.vii

that new information about expected future profit flows are immediately and accurately expressed in observed prices.

Market microstructure has been a very active field of research for over twenty years now, and even a superficial survey of this literature is beyond the scope of the present work. Nevertheless, a few points may be made about the major trends within this field.

Early microstructure work tended to focus on modelling the role of a 'market maker' that provides liquidity to the market and immediacy to individual investors, covering her costs by maintaining a 'bid-ask spread', i.e. a margin between her selling and her buying price for a particular stock. Although the archetype of such an agent is the NYSE specialist, the resulting models apply to other exchanges as well to some extent, as in most markets some type of authorised dealers provide similar services. By itself, the mere existence of a bid-ask spread in the market introduces an error into observed prices, implying that the true value of a security is never observed.

Furthermore, the dealer maintains an inventory, which is costly. This means that at any given time, her inventory position is likely to influence the way she sets her bid and ask quotes, representing a further effect on observed prices that is not accounted for in the classical model. If inventory imbalances are only gradually worked down over a number of subsequent periods, this will smooth out the price effect of new information, imparting distortions to the price process with respect to the classical model. If dealers have a role concerning the stability of the market, as in the NYSE, where the specialist is charged with 'maintaining a fair and orderly market' in his particular stock or stocks, then a part of their actions will be directly aimed at reducing price volatility and smoothing out sudden changes in response to new information. Even when dealers have no such responsibility, almost all exchanges have some sort of structural mechanism that imposes limits on all too sudden price jumps.

Recent research has increasingly stressed informational aspects of trading, asymmetric information and strategic considerations of dealers and investors.[26] If transactions are costly, agents are likely to accumulate new information until they reach some threshold level before trading. This imparts one kind of price-adjustment delay on observed prices. In the same way there can be a delay in updating limit

[26] This aspect is briefly considered by CMSW, but surveyed extensively in O'Hara (1994)

orders, due to the costs involved in continually monitoring the market. This may result in transactions being executed at a stale limit-order when the market moves suddenly. Still another kind of delay will occur if investors sometimes choose to split up large trades with a view to reducing their impact on prices. In the presence of transaction costs, arbitrage opportunities have to exceed a given threshold to be profitable, and this imposes limits on the finer adjustment of transaction prices. In particular, it will not always be optimal to act on each bit of information instantaneously and thus investors may prefer to execute trades only when some critical information level is attained.

To summarise, market microstructure theory predicts at least four different types of distortions of the actual pricing process with respect to the ideal postulated by an efficient markets model. One of them, the bid-ask spread, is essentially a measurement error with respect to the contemporaneous underlying 'virtual' or 'true' price, i.e. the stock's *value*. What the other three all have in common is that they represent a way in which the true price is expressed in reported prices *with a lag*. One of them is infrequent trading in the proper sense, i.e. the absence of truly contemporaneous point samples of stock prices. When this occurs it is common practice to report the most recently observed transaction price in place of the missing observation. The remaining two are impediments to the continuous updating of *quotation prices* originating with dealers and investors. An important distinction emerges immediately. While infrequent trading and the bid-ask spread result in delay and contemporaneous measurement error of *transaction* prices, respectively, the other types of delay affects the arrival rate and magnitude of updates to *quotation* prices. While the presence of the latter type of price-adjustment delay is predicted by microstructure models, direct empirical verification seems to require knowledge of the unobservable true price. Indirectly, however, the idea of important quotation price lags can be rejected empirically. This will happen if transaction lags can be adequately modelled, and turn out to leave no residual symptoms of price adjustment problems in the appropriately corrected time-series.

2 An outline of previous research

Stóð þá á mörgum fótum fjárafli Skalla-Gríms.

Egils saga Skalla-Grímssonar [27]

2.1 Infrequent trading and diversifiable risk

During the seventies Sharpe and Lintner's Capital Asset Pricing Model (CAPM), came to be increasingly applied in practical portfolio management. If an investor is willing to trade off expected return for a reduction in return variance, she is said to be *risk-averse*. When this applies, the CAPM proposes a simple way to construct an efficient portfolio in the sense of the Markovitz mean-variance criterion. In other words, if the covariance of individual returns with a hypothetical market return factor are known, the variance of a portfolio relative to its return can be reduced to a minimum by diversification. This minimum or residual risk is sometimes called 'systematic' or 'undiversifiable'. The true covariances, or correlation coefficents are of course never known and in practice sample estimates must be obtained from the data. In the context of the CAPM, this is customarily done on the basis of the following line of reasoning: Under the efficient market hypotheses, the first differences of dividend adjusted prices can be expected to yield a constant mean return, or a drift term, plus idiosyncratic noise. If there is contemporaneous correlation between stocks in the market, the constant mean return term can be further decomposed into a company specific mean return and a 'common trend component' which depends on 'the market factor'. To yield an estimate of the covariance of security n's return with the market, this leads to the following regression model, where returns on a stock index representing the market portfolio, are usually taken as a proxy for the market factor.

Equation 26

$$r_{nt} = \alpha_n + \beta_n r_{Mt} + \varepsilon_{nt}$$

with

[27] *Egils Saga Skalla-Grímssonar*, Íslenzk fornrit, II. Bindi, Sigurður Nordal gaf út. Hið íslenzka fornritafélag, Reykjavík, MCMXXXIII. Literally: "Then the income of Skalla-Grimr had many feet to stand on"

Equation 27

$$\alpha_n = \mu_n - \beta_n \mu_M; \qquad \beta_n = \sigma_{nM} / \sigma_M^2 .$$

Based on the continuous-time log-normal model of asset prices introduced in an earlier section, given the appropriate assumptions, a continuous-time version of the CAPM came to be developed early on.[28] The assumptions of the regression model in **Equation 26**, can be easily stated in terms of the continuous price process. They are that continuously compounded returns, r_{nt}, are jointly normal, each with constant mean μ_n, constant variance σ_n^2 and constant covariance σ_{nm}, for $n \neq m$, $n,m = 1,2,\ldots N$. To obtain a measure of the return on the market portfolio a market index can be defined as

Equation 28

$$r_{Market,t} = \sum_{n=1}^{N} r_{nt} x_n ,$$

where the x_n is the constant percentage weight of stock n in the market portfolio, i.e. the relative market value of firm n.[29] From the joint-normality assumption we see that all the assumptions about individual returns continue to hold for the average, with parameters μ_M, σ_M^2 and σ_{nM}. In that case OLS estimates of the parameters in **Equation 26** are valid.[30]

2.1.1 The intervaling effect in 'beta'.

Roughly speaking the CAPM-based approach to portfolio management consists in using the OLS estimate of β_n for each stock n, usually referred to as the particular stock's *beta*, to construct an optimal portfolio given the risk preferences of investors. However, in the course of practical work, an anomaly soon appeared that came to be called 'the intervaling effect on beta'. Dimson (1979) defines the intervaling effect in the following words: "This is a tendency for the explanatory power of the regression

---

[28] In Scholes and Williams (1977), reference is made to Merton (1973)

[29] We note that this definition of a market index, used by Scholes and Williams, is equivalent to a *geometric* average of price relatives, i.e. a 'Cobb-Douglas type' index. Recalling that the weights are the product of price and quantity, we see that keeping them constant implies continuous portfolio reallocation if prices are changing continuously. Thus it may not be practical to buy and hold the market portfolio in this sense. Using an approximation and adding unity, we can derive the Laspeyres link-relative of a chained index from this return-index, however.

[30] The normality assumption is of course not necessary for this, but it is a standard feature of the continuous-time model.

equation and the mean value of beta, estimated from value weighted indexes, to rise as the differencing interval is increased."[31]

Examining this effect in some detail for portfolios of stocks grouped by trading frequency, Dimson finds that frequently traded stocks have on the average an upward bias in their OLS beta estimates while in portfolios of very inactive issues beta estimates exhibit a severe downward bias when monthly returns are considered. Theoretically, the average beta of any reasonably well diversified portfolio should be equal to unity. As the return interval is increased to quarterly and semi-annual, the average OLS betas of the different groups converge to unity, and the explanatory power of the beta regression increases considerably, with the strongest effect appearing for the most infrequently traded stocks.

It was already clear to many, that beta estimates are not necessarily stable over time, nor should the true beta be expected to be invariant, because there is nothing to prevent the character of a business enterprise from changing in such a way that its covariance with the market is affected. When daily data started to be available in the mid-seventies, it was therefore hoped that, among other advantages, this would make quick re-estimation of beta feasible, e.g. in the wake of important changes in company structure.[32] However, these hopes were largely thwarted by the corresponding increase in 'intervaling bias' of OLS beta estimates from daily data. Infrequent or nonsynchronous trading was soon identified as the main culprit, and articles started appearing, the first in 1977, suggesting ways to circumvent or eliminate this problem.

2.1.2 The Scholes and Williams model

The pioneering article concerning the econometrics of infrequent trading was the one by Scholes and Williams in 1977, and it sprang from the desire to make up for the intervaling effect in beta.[33] At the time, daily data on stock markets had only become available quite recently and as the authors point out, almost all previous estimates of the systematic risk coefficient had used monthly return data. While the potential error introduced into monthly returns by nonsynchronuos daily trading of stocks is likely to

---

[31] Dimson (1979), p.179. In fact, as we will see, this effect is not limited to value weighted indices.
[32] This is discussed in CMSW (1986), ch.7
[33] Scholes and Williams (1977). Other early approaches included a trade-to-trade returns approach, based on a then rare dataset of transaction-time-stamped intraday data. For discussion and references, see CMSW (1986), ch.7.

be small in relative magnitude if the stocks trade 'almost daily', the opposite obviously applies to daily returns.

Scholes and Williams present the resulting problem essentially as an 'errors in variables'. If reported daily closing price is in fact set at a last transaction that occurs at a random moment during the day, in general it will be observed with error, and the same is true of the resulting daily returns. But if infrequent trading is widespread in the market, this implies that the index is also measured with error, and then the situation is more serious. In this case, the OLS beta estimator will not only be inefficient, as if the error affects only the dependent variable, but also both biased and inconsistent, the latter meaning that no amount of additional data can cure the problem. At the mention of errors in variables, a student's mind is prone to drift towards the idea of an instrumental variables approach to the estimation of a regression model. As it happens, the consistent estimator proposed by Scholes and Williams, can be expressed as an IV estimator with the series of moving sums of market returns at lag 1, 0 and -1 as instruments.

To model the effects of infrequent trading on individual returns explicitly, Scholes and Williams take the continuous-time market model stated above as their point of departure. They define a random variable, s_{nt} to express the time interval between the last trade in security n in period t and end of the period, so that the reported return in period t is actually the return generated over the interval $\left[(t-1)-s_{n,t-1}, t-s_{nt}\right]$, instead of a return over the assumed interval, $[t-1,t]$. Thus, instead of a series of true returns $\{r_{nt}\}$, what is observed is another sequence $\{r_{nt}^s\}$ that in general is not identical to the former. This also leads to a corresponding 'reported index' in each period, $r_{Market,t}^s = \sum_{n=1}^{N} r_{nt}^s x_n$, which will in general be different from the true index stated above.

As a consequence, observed returns are now generated by a composite stochastic process, where the distribution of observed returns $\{r_{nt}^s\}$, depends on the characteristics of both the true underlying return process $\{r_{nt}\}$, and the process $\{s_{nt}\}$ controlling the length of the nontrading subperiods. Imposing some simplifying restrictions on the properties of $\{s_{nt}\}$, Scholes and Williams set about deriving the moments and comoments of the resulting distribution of observed returns.

One of their restrictions states that all stocks trade at least once in every period $[t-1,t]$. If in fact some stock sometimes misses a day, then strictly speaking this means that a sampling interval of at least a week must be imposed. However, in their own words: "This greatly simplifies the subsequent estimators."[34]

The other main simplification consists in the assumption that the multivariate process $S_t = (s_{1t}, s_{2t}, \ldots, s_{Nt})$, i.e. the vector of trading lags in different stocks in a given period, is identically and independently distributed over time. This of course does not preclude persistent differences in trading frequency between stocks, and one case where the simplifying restrictions of Scholes and Williams would hold, is a market where the arrival of trades in each stock is governed by a time invariant Poisson process with a stock specific parameter. On these assumptions, they obtain the following results:

Equation 29

$$E[r_{nt}^s] = (1 - E[s_{nt} - s_{n,t-1}])\mu_n = \mu_n,$$

i.e. the observed mean return is an unbiased estimator of the true mean. As regards the variance, the situation is different. There Scholes and Williams obtain

Equation 30

$$Cov(r_{nt}^s, r_{mt}^s) = (1 - E[\max\{s_{nt}, s_{mmt}\} - \min\{s_{n,t-1}, s_{m,t-1}\}])\sigma_{nm} + Cov(s_{nt} - s_{mmt}, s_{n,t-1} - s_{m,t-1})\mu_n\mu_m$$

Equation 31

$$Cov(r_{nt}^s, r_{m,t-1}^s) = E[|s_{n,t-1} - s_{m,t-1}|]\sigma_{nm} + Cov(s_{nt} - s_{n,t-1}, s_{m,t-1} - s_{m,t-2})\mu_n\mu_m$$

and in general

Equation 32

$$Cov(r_{nt}^s, r_{m,t-\tau}) = Cov(s_{nt} - s_{n,t-1}, s_{m,t-\tau}, {}_{m,t-\tau-1})\mu_n\mu_m$$

While **Equation 29** is simple enough as it stands, some simplification may be in order regarding the variance and covariance terms. In particular it emerges that

Equation 33

$$Var(r_{nt}^s) = \left\{1 + 2\frac{Var(s_n)}{\nu_n^2}\right\}\sigma_n^2$$

---

[34] p.311n

where $\nu_n \equiv \sigma_n / \mu_n$ is the coefficient of variation of the underlying true return process. The autocovariance of order one reduces in a similar way to the expression

Equation 34

$$Cov(r_{nt}^s, r_{n,t-1}^s) = -\left\{\frac{Var(s_n)}{\nu_n^2}\right\}\sigma_n^2$$

As regards the formulas for the cross-covariances between securities at lag zero and one, they can also be made a little more pleasing to the eye by rearranging to form the expressions:

Equation 35

$$Cov(r_{nt}^s, r_{mt}^s) = \left\{1 - E\left[\max\{s_{nt}, s_{mmt}\} - \min\{s_{n,t-1}, s_{m,t-1}\}\right] + \frac{2Cov(s_n, s_m)}{\rho_{nm}\nu_n\nu_m}\right\}\sigma_{nm}$$

and

Equation 36

$$Cov(r_{nt}^s, r_{m,t-1}^s) = \left\{E\left[\left|s_{n,t-1} - s_{m,t-1}\right|\right] + \frac{2Cov(s_n, s_m)}{\rho_{nm}\nu_n\nu_m}\right\}\sigma_{nm}.$$

Here $\rho_{nm} \equiv \sigma_{nm}/\sigma_n\sigma_m$ is the correlation coefficient between stocks n and m. Covariances at lags greater than one will vanish in this model, because of the assumption that all securities trade at least once in the interval.

From this analysis of the observed distribution we learn that expected returns on individual stocks are unaffected by infrequent trading if the expected value of the arrival time of the last trade in the period isn't changing over time. Furthermore, from **Equation 33** we see that the observed variance of individual returns under infrequent trading is always larger than their true variance, provided that the true return is not zero and trading times are stochastic. The first order autocorrelation of the observed returns process will be negative under the same circumstances from **Equation 34**. We see that variability in the trading-lag process positively affects the magnitude of both parameters and if the arrival rates are a Poisson process, this means, *ceteris paribus*, that less active stocks will suffer greater deviation in the parameters of their observed distributions. Substituting a plausible value "...roughly in the range of 30 to 40", for the coefficient of determination in the formulas, Scholes and Williams are able to draw some specific conclusions for daily data. Recalling that nontrading durations that exceed one period are "ignored", they infer that measured variances and first

order autocovariances will closely approximate true variances and zero, respectively. Contemporaneous covariances between prices will understate their true covariances, as was observed for the beta estimates of less frequently traded securities, and first order cross covariances will have the same sign as the contemporaneous ones, but be smaller in magnitude. In the same way, by inspection of **Equation 31** and **Equation 32** both parameters will deviate more from their true values if the difference in trading frequency between the two securities is great than they will if the two securities trade on the average with a similar frequency.

To look at the implications of these results for large portfolios, e.g. the stock-index portfolio, Scholes and Williams invoke two fundamental ingredients of the CAPM. These assumptions, which appear to hold reasonably well in practice, are that stock returns are on the whole predominantly positively correlated and that the variance of a large portfolio is predominantly determined by the covariances of the individual components. Combining this with the formulas in **Equation 33** and **Equation 34**, they infer that for such portfolios measured variances will understate true variances, and all the more so if less active stocks are given any considerable weight. This can also be seen by proceeding in a more formal way. In an appendix, Scholes and Williams show that the formula for contemporaneous covariances of individual securities in **Equation 30** implies that

Equation 37

$$Cov(r_{nt}^{s}, r_{mt}^{s}) = Cov(r_{nt}, r_{mt}) - Cov(r_{nt}^{s}, r_{m,t-1}^{s}) - Cov(r_{n,t-1}^{s}, r_{mt}^{s})$$

Multiplying both sides of this expression *once* by the index weights x_{nM} and summing over n, yields the expression that is the key to the Scholes-Williams beta estimator,

Equation 38

$$Cov(r_{nt}^{s}, r_{Mt}^{s}) = Cov(r_{nt}, r_{Mt}) - Cov(r_{nt}^{s}, r_{M,t-1}^{s}) - Cov(r_{n,t-1}^{s}, r_{Mt}^{s})$$

as is immediately evident from **Equation 30**. The only missing piece is an expression for the observed variance of the index, and this follows by repeating the same operation as before one more time. Thus we have

Equation 39

$$Var(r_{Mt}^{s}) = Var(r_{Mt}) - 2Cov(r_{Mt}^{s}, r_{M,t-1}^{s})$$

From the results in **Equation 38** and **Equation 39**, taken together with **Equation 30**, the Scholes-Williams consistent and "computationally convenient" beta estimator emerges without further ado, as

Equation 40

$$\hat{\beta}_n \equiv \frac{b_n^- + b_n + b_n^+}{1 + 2\hat{\rho}_M}.$$

where b_n^-, b_n and b_n^+ are the OLS estimates resulting from regressing the individual returns of stock n on the index at lags 1, 0 and -1 respectively, and $\hat{\rho}_M$ is the sample estimate of the first order autocorrelation coefficient of the index. In this context, Scholes and Williams also derive a concise expression for the theoretical first order autocorrelation of the index under infrequent trading, in terms of the ratio of its true and observed variances:

Equation 41

$$\rho_M^s = \frac{1}{2}\left[\frac{Var(r_{Mt})}{Var(r_{Mt}^s)} - 1\right]$$

Given that $\hat{\rho}_M$ is a consistent estimator of the index autocorrelation under infrequent trading, this formula explains why empirical time series analysis reveals positive sample autocorrelation even though the true index value, as an average of 'true returns', is a random walk.

An interesting corollary of the Scholes-Williams model is that under plausible assumption about the coefficient of variation (see above) in a sample of daily data, the observed distribution of returns will be leptokurtic even when the true distribution is normal. On the same assumptions as before, they prove the following expression of the extent of deviation that can be expected:

Equation 42

$$\kappa(r_{nt}^s) = 3(1 + 2Var(s_n)) + O(1/v_n^2).$$

We have seen that the availability of daily data soon spurred research leading to a number of interesting results about the properties of individual returns and the market index in the presence of infrequent trading, admittedly based on somewhat restrictive assumptions. The issue at stake is the availability of a consistent measure of diversifiable risk, which is essential to rational portfolio management.

Scholes and Williams were not the only ones to delve into this problem, at roughly the same time a number of consistent beta oriented studies tried to solve the problem on the basis of trade-to-trade return data, but without reference to an explicit model.[35] Also, independently, CMSW were working on a statistical model of infrequent trading, although publication followed only in 1979.

2.1.3 A market microstructure perspective: the CMSW model

Independently of Scholes and Williams, but at the same time, CMSW were working on a solution to the intervaling-bias problem, that was also based on a stochastic model of infrequent trading. The results started appearing in 1978, and articles on related issues continued appearing in leading journals for nearly a decade. In 1986 a book was published, organising the foursome's most important results to form a comprehensive view of security market microstructure.[36] In the light of considerations similar to those that were sketched in section **1.3** above, CMSW criticise the early research of Scholes and Williams, as well as much of the intervening beta literature on two accounts.

One is to treat the infrequent trading problem in isolation, ignoring its connection with a number of other empirical anomalies of stock markets, most of which seem to relate to small-size firms. Examples of this kind are an increase in both the bid-ask spread and volatility of returns for small firms relative to larger ones, as well as their higher propensity to trade only intermittently, other things being equal.

The other main weakness identified in CMSW (1986) is the assumption, either implicit or explicit in most earlier research, that the transaction time lag that occurs under nonsynchronous trading is the *only* source of spurious autocorrelation in stock market return series. As pointed out by CMSW, this is equivalent to the view that if all stocks traded every day exactly at the closing bell, then both the index as well as individual returns would be absolutely free of spurious elements.

The question whether this is really the case, can be posed in more than one way that is empirically testable. Thus, if transaction lags alone were responsible for index autocorrelation, then an index of frequently traded stocks, such as the Dow Jones

[35] Some of these attempts are discussed in CMSW (1986), ch.7
[36] Some of the articles were co authored by Gabriel Hawawini.

Industrial Average would be almost free of autocorrelation, which it isn't.[37] Also, calculating returns on a trade-to-trade basis would lead to a consistent estimator of the CAPM beta, which it doesn't.[38]

In the view of CMSW, modelling price-adjustment delays solely in terms of differing transaction and sampling times, means ignoring the effect of a host of unobservable but very real delays that affect the update of *quotation-prices*, and can be predicted on the basis of economic microstructure models. As we have seen, the possible sources of quotation adjustment delay highlighted by CMSW mainly fall into two categories. They are (1) impediments originating with dealers, such as inventory imbalances and stabilisation measures, and (2) adjustment lags for individual traders, originating in transaction and information costs. If these two are significant sources of friction, then spurious properties will persist in the index and individual returns, no matter how closely trading approximates continuity. In this case a transaction-based model will not make the theoretical covariance structure of returns fully explicit and as a result, estimators derived from such a model will fail to account for some of the bias in beta.

True to their comprehensive view of market microstructure effects, CMSW include a term accounting for 'bid-ask bouncing' in their model of the relationship between true and observed prices, or as they call it, "the frictions equation". Here we will state the equation in this form for the record only, and then go on to drop the bid-ask spread term. This can be justified by arguing that bid-ask effects cancel out in large portfolios because at any given time some stocks trade at a bid and others at an asked price, and as we will see later on, this is what most authors do assume, at least when focusing on the stock index. Presently, however, we need only appeal to the fact that this is the way that CMSW proceed themselves.[39] With the bid-ask spread component, their *friction equation* looks like this:

Equation 43

$$r^{o}_{j,t} = \sum_{l=0}^{N} (\gamma_{j,t-l,l} r_{j,t-l} + \theta_{j,t-l,l})$$

---

[37] Further on we will briefly discuss a paper by Perry (1985), which sets up an 'experiment' of this kind.

[38] Cf.CMSW, p.113

[39] In fact the presence of a heavily subscripted theta *inside* the parentheses in the frictions equation can cause the student of their model some anguish, given that a bid-ask shock is an observation error that by definition only occurs at the time of a transaction (i.e. measurement). The authors offer some justification in an appendix.

Here the γ terms are random weights, each signifying the fraction of the total observed return over the aggregate period t that is generated in a particular current or earlier period $t-l$, l=1,2,...,N. The underlying assumptions are (1) that the gammas sum to one over l, (2) that the lag structure is stationary over time and (3) that the weights are independent between securities at all lags. The effect of the bid-ask spread is expressed in the random variable θ. As the bid-ask spread is essentially an error of measurement due to transaction costs, thus occurring only when a transaction takes place, it seems somewhat curiously formulated here and as it happens, the authors add a footnote setting it to zero for $l > 0$.[40] The fact that the sum is over N terms implies an assumption regarding a maximum trading delay of N periods. Further, contrasting this specification with that of Scholes and Williams discussed in the previous section, we note that this is a discrete-time model, based on a regular period structure, while the earlier one sought to model trading delays in continuous terms within the period, assuming trading in all periods.[41] As opposed to the Scholes-Williams model, this one is not based directly on transaction times. Instead it takes a stochastic 'delay weighting structure' as primitive. The Scholes and Williams model can be stated in the same way, and shown to be the special case where N=1.

Equation 44

$$r^{o}_{j,t} = \gamma_{j,t,0} r_{j,t} + \gamma_{j,t-1,1,l} r_{j,t-1,1}$$

which is easily seen to be the special case of 1.1 when $N = 1$.

Basically, the frictions equation is a model of time aggregation in a flow variable, i.e. individual returns. In this respect it is not unlike our simple example due to Holbrook Working, except that it assumes that the proportion of the price innovation incurred to the disaggregate process in each subperiod is random, unknown and unknowable. But while we brought this example to bear on an index in section **1.2**, interpreting each subperiod as an auction, CMSW are modelling an individual price process with a 'true

---

[40] Rather perplexingly, the CMSW theta parameter is not mean-zero, but zero at the bid, and further on the authors just drop it, implicitly assuming that observed returns are *exactly* defined as a randomly weighted sum of the disaggregate series of true returns. For a similar procedure see Jukivuolle (1995).

[41] Making a distinction between *nontrading* and *nonsynchronous trading*, as in Miller et al.(1991), where nontrading is said to occur when a stock does not trade in every period, but nonsynchronous trading when stocks do not trade simultaneously within periods, we note, that as opposed to the earlier model, what we have here is a model of nontrading. Obviously, as pointed out by Miller et al., nonsyncronous trading becomes nontrading when holding periods are reduced.

price innovation' of unknown magnitude occurring in each subperiod. Thus they see the observed return in each period as a randomly weighted sum of innovations to the true price process, that may have occurred earlier without resulting in a transaction.

On the basis of the formulation of the price-adjustment delay process in **Equation 43**, CMSW proceed to derive the theoretical covariance structure of individual securities. If a zero bid-ask spread is assumed, equation **Equation 43** leads directly to the following formula:[42]

Equation 45

$$Cov(r^o_{j,t}, r^o_{k,t-n}) = Cov\left(\sum_{l=0}^{N} \gamma_{j,t-l,l} r_{j,t-l}, \sum_{l=0}^{N} \gamma_{k,t-n-l,l} r_{k,t-n-l}\right)$$

Taking $j = k$ yields an expression for the autocovariance of the reported price process of individual securities that can be stated in terms of the moments of the original stochastic processes r_{nt} and γ_{nt}.

Equation 46

$$Cov(r^o_{j,t}, r^o_{j,t-n}) = Var(r_{j,t}) \sum_{l=0}^{N-n} E(\gamma_{j,l} \gamma_{j,l} \gamma_{j,l+n}) + E^2(r_{j,t}) \sum_{l=0}^{N} \sum_{m=0}^{N} Cov(\gamma_{j,t-l,l} \gamma_{j,t-n-m,m}).$$

Under nontrading this expression will be nonzero, but indeterminate in sign. In the presence of a bid-ask effect, by contrast, individual autocorrelation is quite likely to be negative, especially as the microstructure literature offers some theoretical support for both positive and negative autocorrelation due to price adjustment delays, opposing effects that may tend to outweigh one another.

To derive an expression for the serial cross-covariance between securities, CMSW refer to the standard CAPM market equation as stated in **Equation 26**. Imposing the additional restriction that the random weights $\gamma_{j,t,l}$ are independent of the market factor as well as the error term of this equation at all leads and lags, the theoretical cross-covariance structure can be stated making use of true CAPM betas or correlations of individual security returns with the market, expected degree of nontrading, as well as the true variance of the market portfolio.

Equation 47

$$Cov(r^o_{j,t}, r^o_{k,t-n}) = \beta_j \beta_k Var(r_{M,t}) \sum_{l=0}^{N-n} E(\gamma_{j,t}) E(\gamma_{k,l+n})$$

---

[42] The same holds assuming a zero mean i.i.d process to replace it.

This formulation can be interpreted as the product of the two true market covariances divided by the market portfolio variance times a sum of expected delay weights. If the expected delay weight in period *l-n* is zero, the covariance reduces to zero. If not, for the sign of this quantity to be determinate, even in a relative sense, CMSW need a 'regularity condition' saying that

Equation 48

$$E(\gamma_{j,l}) \geq 0 \quad \text{for} \quad \forall\, j,l.$$

This restriction will guarantee that the expression in **Equation 46**, defining the theoretical cross-covariance between individual security returns, is nonzero under infrequent trading, and has the sign of the product of the two securities' true CAPM beta. As beta expresses correlation with the market factor which is predominantly positive, the same will apply to spurious cross serial correlation due to infrequent trading..

Defining the market index in the same way as above (see equation **Equation 28** in the last section) its autocovariance can now be decomposed. This is done by proceeding in the same way as when deriving formula **Equation 39** above.

Equation 49

$$Cov(r^o_{M,t}, r^o_{M,t-n}) = \sum_j \sum_k x_j x_k Cov(r^o_{j,t}, r^o_{k,t-n})$$

Now this quantity can be stated in terms of 'true betas' as before, but first the terms with $j = k$, must be assumed away. Granting this and with the help of the regularity condition that guarantees nonnegative mean weights over time and securities, we obtain

Equation 50

$$Cov(r^o_{M,t}, r^o_{m,t-n}) = Var(r_{M,t}) \sum_j \sum_k x_j x_k \beta_j \beta_k \sum_{l=0}^{N-n} E(\gamma_{j,t}) E(\gamma_{k,l+n})$$

which makes explicit a positive serial correlation in the observed index as a consequence of price adjustment delays in individual securities if betas are mostly positive. The reason that the individual autocorrelation terms must be ignored is that when $j = k$ the independence assumption about $\gamma_{j,t}$ and $\gamma_{k,t-n}$ implicit in **Equation 46** does not hold and **Equation 45** must be used to substitute for these terms in **Equation 49** to obtain **Equation 50**. This leaves the sign of the index autocorrelation

indeterminate. The justification for ignoring these terms is that their effect vanishes in the limit as the number of securities in the portfolio increases, if no particular security dominates the others. This is due to the 'diversification effect' that is the cornerstone of the CAPM, i.e. that the variance of a large portfolio approximates the average covariance of its components. It will be intuitively clear that the weights x_k, given to less active stocks positively affect the magnitude of serial correlation, and that for this reason, taken together with the inverse relationship between firm size and trading intensity, an equally weighted stock-index will have greater serial correlation than a value weighted one, if everything else is equal.

On the basis of these results CMSW are able to derive a consistent beta estimator using the same kind of decomposition as used by Scholes and Williams (leading to equation. **Equation 40**). Not surprisingly, given the close relationship between the two models, this estimator can be seen as a generalisation of the Scholes and Williams estimator stated in **Equation 40**. In the CMSW notation, this estimator is

Equation 51

$$\hat{\beta}_j = \frac{b_j^o + \sum_{n=1}^{N} b_{j,n+}^0 + \sum_{n=1}^{N} b_{j,n-}^o}{1 + 2\sum_{n=1}^{N} b_{M,n+}^o}$$

where the terms in the numerator are the OLS estimates of beta with respect to lags 0 to $\pm N$, and $b_{M,n+}^o$ is what the authors call an "observed intertemporal market beta", meaning: index serial correlation. As it turns out, the advantage of this generalised estimator is more apparent than real, because the problem of selecting N correctly involves a dilemma. With N large, this estimator could be expected to capture the effects of protracted lags in the quotation price adjustment process, the existence of which is predicted by microstructure considerations. At the same time, however, a large value of N results in loss of estimation efficiency, which may quickly outweigh the gain in 'realism'. There may be no simple way to select an optimal value of N with respect to this trade-off, and for this reason CMSW concede that their consistent beta estimator may not be of great practical value.

To obtain a 'feasible consistent estimator', CMSW suggest an indirect approach, using the fact that under intervaling bias, the observed OLS parameter converges to the true parameter as an asymptote when the holding period (i.e. the differencing

interval) increases without bound. This is the first step in a two stage procedure, leading to the 'asymptotic beta estimator', defined as

Equation 52

$$\beta_j^* = \lim_{L\to\infty} b_j^0(L)$$

where $b_j^0(L)$ stands for the OLS estimator based on non-overlapping periods of length L. This estimator is shown by CMSW to consistently estimate true beta. To implement it, they suggest regressing the sample estimates $\tilde{b}_j^0(L)$ on an appropriate function of L'.[43] Obviously, though, this may not be a very practical procedure either, for a great number of observations is required. Consequently, CMSW suggest exploiting proxy for the degree of nontrading of a particular stock to obtain a correction factor that can then be used to estimate β_j^* from $\tilde{b}_j^0(L)$ for increasing but finite differencing intervals L. The estimator obtained in this way is the so-called "inferred asymptotic beta".

Concluding their analysis of the statistical relationship between true and observed return variance-covariance structure, CMSW state their results in an organised manner:[44]

1. Individual returns are serially correlated.
2. A market index composed of a large number of securities will have positive serial correlation.
3. A value weighted market index will have smaller serial correlation than a similarly composed equally weighted market index.
4. OLS beta estimates will be biased, with the absolute bias
   a) going to zero as the differencing interval increases without bound;
   b) greater for otherwise-identical securities with greater expected price-adjustment delays.

CMSW raised an important objection to the approach taken by Scholes and Williams towards correcting for intervaling bias in beta. Based on an economic model of the microstructure of securities markets they suggested that a transaction based statistical model of infrequent trading will not succeed in removing all spurious elements from the observed price process, because of the existence of quotation update delays, that may differ across securities. Although CMSW do model

[43] This is expected to be a nonlinear relationship. Details of this procedure are described in CMSW (1986), p.132-133.

'quotation returns' on the basis of shifts in demand generated by a compound Poisson process, they do not model quotation-update *delays* in a way that is either quantifiable or testable.[45] Their suggested remedy in the case of CAPM beta is to extend the 'span' of its estimator enough to pick up serial cross covariance at up to N lags and leads, reallocating all covariance to the current interval, so to speak. Evidently, for this to be justifiable, the efficient market hypothesis must hold in reality. If quotation lags are important, transaction prices will partly occur on the basis of stale quotes and reflect out of date information. Therefore N has to be larger, in general, than the maximum nontrading interval in the dataset. This requirement may well lead to a value of N that makes efficient estimation infeasible for a given data set. As an alternative way of obtaining consistent beta estimates, CMSW suggest using an 'asymptotic beta estimator'. The idea is that in the limit any microstructure effects contaminating returns data will be small relative to long run relationships. Before looking at more recent models of nontrading, which requires a shift in emphases we will briefly review some empirical results pertaining to this first generation of infrequent trading models.

2.1.4 Some empirical results

The Scholes and Williams and CMSW estimators are far from being the only attempts to correct for CAPM beta bias in the presence of infrequent trading. In contrast to other early research on beta inconsistency, however, they both offer models of infrequent trading phenomena.

One early paper that resulted in a 'consistent beta estimator' was that of Dimson (1979). Dimson does provide some interesting empirical results on the intervaling effect, but he makes no attempt to model its sources explicitly. The Dimson estimator is essentially an extension of the Scholes-Williams estimator to an arbitrary number of lags and leads, much like the CMSW N-lag estimator, but based on a multiple regression approach. Apart from suffering from the same trade-off problems concerning the selection of an appropriate value of N, as the latter, the Dimson

---

[44] CMSW (1986), p.125-6

[45] CMSW, Chapter 4. They model the effect of the presence of 'designated market-makers' on the relationship of (quotation) returns variance and thinness of trading. In fact it is hard to think of any other direct measure of this delay component than the time elapsed between news arrival and quotation update, something that may extremely hard to quantify in a simple way. This is because 'relevance' of business information is inevitably a somewhat circular concept in the context of the efficent market hypothesis.

estimator did not yield encouraging empirical results. In experiments reported in Fowler and Rorke (1983), it was allegedly outperformed not only by the simpler one of Scholes and Williams, but also frequently by OLS. In their article, Fowler and Rorke prove that the Dimson estimator is incorrect.

At the outset of infrequent trading research, a number of scholars pursued a more direct alternative line of reasoning, obtaining exact transaction times and calculating trade to trade returns to serve as a basis for the estimation of CAPM beta.[46] While this approach will minimise the error in the dependent variable of the CAPM regression, we recall that the real errors-in-variables problem lies with the regressor. If the market index is autocorrelated, even under approximately continuous trading, as it may well be when quotation-price delays are considerable, then the trade-to-trade approach will not solve the problem of biased beta estimates.

Still earlier, and somewhat apart from the mainstream of the beta-bias literature, a consistent beta estimator based on a Bayesian approach had been suggested by Vasicek.[47]

Various beta adjustment procedures were finally tested and compared in by McInish and Wood (1985) and Ord, McInish and Wood (1984).[48] Their results indicate that none of the above approaches, results in adequate estimators. Thus the Scholes-Williams, Dimson, and Fowler-Rorke-Jog (1984) estimators only reduce the spread of average beta estimates between portfolios with differing average expected price adjustment delay by some 20-30%.[49] This means that even when 'consistent estimators' were used, consistent estimates were not obtained, and average beta in experimental portfolios of infrequently traded (thin) securities remained less than one and thick portfolio beta remained larger than one. The trade-to-trade method was also tested and performed no better by this measure. Finally the CMSW 'inferred asymptotic beta' was tested with the same result.[50] The Bayesian procedure was not included.

---

[46] See CMSW (1986), ch.7

[47] Vasicek, (1973). This estimator is not examined in the present study.

[48] Working papers. The results are reported in CMSW (1986), ch.7.

[49] Results are reported in CMSW (1986), p.144

[50] In their opinion the experimental setup in this investigation is biased against the inferred estimator. CMSW (1986), p.145

Three of the above approaches were compared by Berglund, Liljeblom and Löflund (1989) in the context of the thinly trading Helsinki Stock Exchange.[51] As a representative of the 'aggregated coefficients' class of estimator, that includes the Scholes-Williams, Dimson and CMSW estimators discussed above a variation of that of CMSW was tested. The other two approaches tested on Finnish data were trade-to-trade based estimation and the Vasicek estimator. The conclusion can hardly be said to be encouraging: "(Our) results indicate that none of the corrections as such are likely to produce much improvement compared to OLS betas."[52]

In view of this it is maybe not surprising that the substantial wave of intervaling effect literature that started in 1977 and persisted well into the late eighties seems to have waned. In part, it was replaced during the late eighties by a line of research into infrequent trading induced distortions in the covariance factor structure that has to be determined in applications of the Arbitrage Pricing Theory. This research originated with Shanken (1987). The APT is sometimes summarised as a multifactor version of the CAPM. It does not depend on the market index to the same extent as the CAPM, but instead APT based portfolio management assumes knowledge of the market covariance-factor structure, i.e. the pattern of cross-covariances between individual stocks. In view of the intervaling effect in beta, and theoretical cross-correlation between individual stocks under infrequent trading, it seems natural to expect OLS estimates of this matrix to be biased as well.

This, in fact, is Shanken's conclusion. Applying an 'aggregated coefficients' estimator of the type proposed in CMSW (1986), he finds that average covariance estimates roughly double for NYSE stocks as compared to results obtained by OLS. As it seems that this line of research has not resulted in any new contributions to the statistical theory of infrequent trading, it is of marginal importance in the present context. It may be noted that fairly intensive research of this kind has been conducted during the nineties in the context of the thinly trading Finnish market. Not surprisingly, the results indicate that this problem is even more severe at the *Helsinki Stock Exchange* than in the NYSE; in particular, instead of doubling, the estimated average covariance increases four- to fivefold when estimates are adjusted for intervaling bias!

---

[51] Berglund et al. (1989)
[52] Berglund et al. (1989), p.61

2.2 Infrequent trading and systematic risk

The impetus that set rolling what I would like to term the second wave of research into infrequent trading problems, just like the first one, was twofold. In fact, the forces at work were exactly the same as earlier. On one side changes in the form and frequency of available market data, on the other: new ways of dealing with investment risk. The latter were represented by the possibility of trading in stock index futures, that opened in the U.S. in 1982.While futures positions can in principle be taken in single stocks, futures contracts based on the index portfolio are somewhat particular in that they permit investors to hedge (or speculate on) changes in the price of the market portfolio itself. This in fact establishes a market in the residual or 'undiversifiable' market risk component discussed earlier in the context of the CAPM. The consequence is enhanced efficiency in financial markets, but at the same time, greater demands are made to the index. The risk management aspect as of infrequent trading research is a convenient organising principle for purposes of exposition. In reality, however, it is inseparable from other aspects of historical development in this field, above all that of technological change. At roughly the same time as the advent of futures trading, new kinds of data became available, making new demands on research. Before attempting to outline some major contributions to research, we will briefly discuss these two sides of the situation.

2.2.1 The data, the stock index and empirical research.

As we saw, the Scholes-Williams paper appeared immediately in the wake of the first attempts to estimate CAPM-beta from daily data. This time around it was the appearance of high-frequency intraday transaction data that spurred research developments. Of the five models of nonsynchronous trading effects that we shall examine in this section, one compares 60, 30 and 15 minute return intervals, two use only 5 minute intervals and a fourth investigates a two-day period breaking it up into one-minute periods. The fifth stands somewhat conservatively apart by investigating daily index returns.

New data breeds new inquiries, there is nothing surprising in that. But while increasing the data frequency from monthly to daily implies increasing it by a factor of a little more than twenty, going from daily to minute-by-minute data is more like a 360-fold increase. As a consequence, one might expect the relative magnitude of

nonsyncronous trading induced error to increase considerably, even if trading volume and frequency had also increased, say, doubled, tripled or increased by a factor of ten.[53]

The eighties also brought an increase in the options available to investors for dealing with risk in the stock market and one of the most important innovations in this field was the stock index futures contract. Like transaction data, this phenomenon was made feasible by computerised trading mechanisms. As an aspect of risk management (or speculation), futures trading involves buying and selling large portfolios of stocks simultaneously. To keep things simple, we can define a stock index futures contract as a financial instrument involving the obligation to buy or to sell an entire portfolio of stocks, composed exactly as the corresponding index portfolio, at a later time and at the prevailing price at that later time. We have seen that CAPM beta is a yardstick that can be exploited to eliminate all company specific risk from a particular portfolio, leaving only a residual level of market risk. The futures contract represents a way to reduce, or eliminate, this residual systematic risk. If for example an investor holds the index portfolio and believes that there will soon be a downturn in the market, she can sell the futures contract for later delivery, receiving its current value, which will be close to the cash value of the portfolio. If she is right, the value of the portfolio she delivers will be less than that of the one she sold and she will make a profit which compensates his losses on the portfolio itself. In most cases, of course, no delivery ever takes place, and only net profits and losses change hands, making the stock index futures market a near perfect paper - or rather: digital - market.

Now the question arises how futures contracts should be priced, and the theory of finance provides an answer to that, based on the efficient market hypothesis that we have already become acquainted with. Ideally, the price of an index futures contract relative to the price of the underlying index portfolio, should be determined by the so called cost-of-carry equation

---

[53] Indeed, looking at a fourfold increase in the number of observations going from an hourly sampling interval to one observation per hour, like in Table I of MM&W (1991), we see that this holds in the case of the earliest series analyzed, dating from 1982 and 1983. In this case the resulting increase in first order autocorrelation of the S&P500 index is nearly fourfold. Oddly enough, however, in the last years of the sample period covered by MM&W, 1988-90, this relation is reversed and average first order index autocorrelation decreases with a decrease in the differencing interval. In particular, reducing the holding period from 60 to 15 minutes, reduces the value of the autocorrelation coefficient roughly by half.

Equation 53

$$F_t = S_t e^{(r-d)(T-t)},$$

where F_t and S_t represent the price of the futures contract and the index portfolio, respectively, *(r-d)* is the net cost of carry, i.e. the difference between the portfolio dividend yield and the riskless interest rate and *(T-t)* defines the remaining life span of the futures contract. Absent transaction costs, so the theory goes, if the difference between the left and right hand sides of **Equation 53** is nonzero, a riskless arbitrage profit can be earned, assuming the possibility of borrowing and investing at the riskless rate of interest. If the futures price is higher, this is achieved by buying the index portfolio and selling the futures contract, but when the futures contract is underpriced the opposite strategy applies. Whichever the case, in practice the necessary transactions would be executed by means of program trading, meaning that a single computer-generated order is used to buy or sell an entire portfolio.

To state the pricing equation **Equation 53** in terms of returns, we note that under the assumptions of a perfectly efficient and continuous market we can take logarithms on both sides and differentiate w.r.t. t to obtain

Equation 54

$$R_{S,t} = (r-d) + R_{F,t}$$

where $R_{S,t}$ and $R_{F,t}$ are instantaneous rates of return on the index portfolio and the corresponding futures contract, respectively. This relation has a number of implications for the second moments and co-moments of rate of return of the futures contract and the index portfolio. In particular,

1. Their variances are equal,
2. They are perfectly positively correlated at lag zero
3. They are uncorrelated at lags other than zero
4. They are serially uncorrelated

In practice however, based on what little forays into the field of market microstructure we have already undertaken, we would expect all of these hypotheses to be rejected by the data, were they to be tested empirically.

A concept intimately related to the cost-of-carry relation is the *futures basis*. This is simply the difference, at any given time, obtained by subtracting the price of the futures contract from the cash value of the index portfolio, i.e. the corresponding

stock index number. Algebraically, it can be formulated as follows, using **Equation 53**:

Equation 55

$$S_t - F_t = \left(1 - e^{(r-d)(T-t)}\right).$$

It is essential, for efficient management of systematic risk to be feasible, that the futures basis should be unpredictable on the basis of current information. However, a glance at **Equation 55** will reveal that it converges to zero deterministically as the futures contract approaches maturity. However, the amount of mean reversion due to this can be shown to be small at reasonable distances away from expiry and 'forced' convergence due to convergence of dividends to the real rate of interest can be tackled by looking only at intraday returns, as both interest and dividends are paid overnight.[54] The empirical evidence, on the other hand, is not easily argued away. What we see are far too frequent violations of the cost-of-carry relation, strong tendency of price changes of the futures contract to lead those in the index, and, consequently, substantial deviations in the futures basis. This situation has led a number of scholars to perform statistical analysis of both index and futures contracts price series, in search of reasonable explanations for these anomalies.

One of the studies examined in this section, that of Stoll and Whaley (1990) seeks to establish why the cost-of-carry relation is violated in such a systematic way as is observed. A second one, Jukivuolle (1995) simplifies the Stoll-Whaley approach, at the same time making it more rigorous, at least in a particular sense of the term. This leads to an elegant indirect derivation of a 'a true stock index'. The other three are primarily concerned with the futures basis. Miller, Muthuswamy and Whaley (1994) investigate mean reversion commonly thought to be incurred by arbitrage reactions, finding that it is more likely to be a 'statistical illusion'. They use data over the period 1982-90, excluding the 1987 crash week. Harris (1989) and Bassett, France and Pliska (1991), on the contrary, examine the behaviour of the futures basis during the days of the crash in extreme detail. They all assume that infrequent trading is to blame, either in part or wholly, for the observed anomalies. The other thing they all have in common, is that what is at stake is an efficient way of transferring *systematic* risk.

---

[54] For the reasoning, see Miller, Muthuswamy and Whaley (1994), p.481-3.

Before going on to look at this research, it is of interest to set a backdrop, so to speak, by mentioning some empirical results that were published in the period separating the two successive waves of interest in the nonsynchronous trading problem that we have identified as pertaining to diversifiable and systematic risk.

Perry (1985) performs an experiment. In an earlier chapter we noted that there seems to be a strong positive relationship between the market value of firms and the frequency of trading in their shares. As a result it is common practice to use firm size as a (negative) proxy for nontrading propensity. For this reason Perry selected thirty of the largest U.S. firms and thirty of the smallest. He then constructed pairs of large-firm and small-firm portfolios, and compared daily return autocorrelation for each pair. The result revealed surprisingly high autocorrelation in portfolios of the more frequently trading firms, and increasing with the number of stocks in the portfolio. In fact, positive first order autocorrelation in returns on the large-firm portfolios surpassed that observed for the small firm portfolios for all portfolio sizes except the largest, i.e. the full thirty stocks. Perry's conclusion: "The implication of this finding is that nonsynchronous trading is not the only cause of correlation in daily market indices."[55]

Two years later, Atchison, Butler and Simonds (1987) followed up on the Perry experiment by another method. Correcting for index autocorrelation by a formula derived from the Scholes-Williams model examined in section *2.1.2*, they compare theoretical with the observed autocorrelation level in an equally-weighted/value-weighted index of 280 stocks, randomly selected from the NYSE. The results are on a par with those of Perry. The infrequent trading model only explains about 16% of the reported autocorrelation for the equally weighted index, and some 13% for the value weighted one. In their discussion, ABS refer to the CMSW quotation-update theory as a possible explanation for this discrepancy, but do not seem over-enthusiastic about their delay model: "..., these efforts presently have not resulted in a model formulation that quantifies the index autocorrelation induced from [...] other frictional sources. This paper demonstrates the need for such a model, which may be more descriptive than the nonsynchronous trading model presented here."[56]

---

[55] Perry (1985) p.517

[56] Atchison, Butler and Simonds (1987), p.117

2.2.2 The Stoll-Whaley model

Stoll and Whaley choose to pose the index autocorrelation problem, in terms of 'cost-of-carry violations', (see **Equation 53**). Quoting empirical studies by other scholars, as well as earlier ones of their own, they conclude that although the situation seems to be 'improving' with time, the frequency of cost-of-carry violation is far too high to be compatible with the efficiency assumption, and that this is "indeed one of the puzzles in stock index futures".[57]

Two of the reasons for this that are suggested by Stoll and Whaley, i.e. *infrequent trading* of stocks within the index portfolio and *bid-ask bouncing*, will be familiar from our treatment above. A third reason, time delays in the computation and reporting of the stock index value, is new. In the interval separating the pioneering studies presented in CMSW (1986) and Scholes and Williams (1977) on one hand, and the Stoll and Whaley paper on the other, great advances had been made in the 'computerisation' of securities trading, resulting in the availability of high frequency transaction data. Thus the empirical studies of the cost-of-carry relation discussed earlier had used half-hour and fifteen minute holding periods to obtain transaction based returns, and in the study we are presently considering (1990), Stoll and Whaley use five minute return intervals. To convert the transaction based data into a return series for a fixed holding period, they pick out the first transaction in every five-minute interval. As the S&P500 is calculated about four times every minute and the futures contract trades more frequently than that, they assume that any potential errors-in-variables effects are sufficiently mitigated by a five minute interval.

Assuming a continuous-time log-normal model of security prices, which implies the assumption that stock returns are independent and identically distributed through time, the true return on stock i in period t is given by

Equation 56

$$R_{i,t} = \mu_i + \eta_{i,t}$$

where μ_i is the expected return and $\eta_{i,t}$ a mean-zero innovation. If none of the stocks skip a trading period, the effect of the bid-ask spread can be modelled separately as follows: The bid-ask error is assumed to affect the return over each interval in two places, once in the beginning of a period and once at the end. For this

reason Stoll and Whaley suggest that we look at the bid-ask effect as an MA(1) process, writing the resulting returns equation as

Equation 57

$$R_{i,t}^{*} = \mu_i + \eta_{i,t} + \vartheta_{i,t} - \delta_i \vartheta_{i,t-1}.$$

The authors do not explicitly state the meaning of the variable δ_i, but it seems logical that it be thought of as a Bernoulli variable taking the values 1 and -1, depending on whether the transaction executes at the bid or the ask. Then, if the markup is changing only slowly with time, this error component would roughly cancel out if both transactions are at the ask or both at the bid, but add up otherwise.

This expression is now summed over the portfolio to obtain an expression for the index value in the presence of a spread, but free of infrequent trading effects:

Equation 58

$$R_{S,t}^{*} = \sum_{i=1}^{m} X_i R_{i,t}^{*} = \mu_S + \eta_{S,t} + \sum_{i=1}^{m} X_i \left(\vartheta_{i,t} - \delta_i \vartheta_{i,t-1}\right).$$

Here the *Xi* represent the index weights, and μ_S and $\eta_{S,t}$, are the index return and error term, respectively. If certain conditions apply, the summation term on the far right can reasonably be ignored as *m* grows large, yielding an approximation to a 'true index' on the assumption that all stocks trade in every period according to this model and quotation update is instantaneous. This depends on the bid-ask errors being independent across stocks and the weights being of order 1/*m*. If there is a small number of stocks in the portfolio or some firm is disproportionately large, this does not hold, something that should be kept in mind when studying a particular stock market or index number.

In an additional step a model of the index under nonsynchronous trading is presented using the index return term in the presence of a bid-ask spread from **Equation 58**. All stocks belonging to the portfolio are assumed to trade at least once every *n* periods for some *n*, a feature that this model has in common with the one in CMSW (1986) discussed earlier.

---

[57] Stoll and Whaley (1990), p.444. The frequency ranges from 80% for the June 1982 S&P500 contract, to around 15% on average in the period 1983-87.

Equation 59

$$R^{o}_{S,t} = \sum_{k=0}^{n-1} \omega_{S,k} R^{*}_{S,t-k} + v_{S,t}$$

The constraints imposed on the weights $\omega_{S,k}$, are that they should (a) be positive constants, (b) decline with k and (c) sum to one, and the disturbance term is assumed to behave nicely. The obvious interpretation of this formula is a backward looking one, of the type $\omega_S(B)R^{*}_{S,t}$, but it must be noted that it can also be interpreted looking forward. In that case we see $\omega_{S,0}$ as the fraction of the true portfolio return in period t that is revealed simultaneously, $\omega_{S,1}$ as the fraction of the t-period return that only comes to be observed in period t+1, and so on, until the return due to the least frequently trading stocks is finally revealed in period t+n. By application of a moderate dosis of algebra to the formula in **Equation 59**, Stoll and Whaley now derive the following theoretical expression of index portfolio returns in the presence of infrequent trading:

Equation 60

$$R^{o}_{S,t} = \omega_{S,0} R^{*}_{S,t} + \sum_{k=0}^{\infty} \gamma_{S,k} R^{o}_{S,t-k} + v_{S,t} + \sum_{k=1}^{\infty} \gamma_{S,k} v_{S,t-k} \,.$$

This expresses the observed portfolio return in period t as a function of lagged observed returns and lagged observation errors, in other words as an ARMA(p,q) process. The coefficients on the AR and MA parts are the same, and the order of the process is infinite. Substituting for the unobservable term $R^{*}_{S,t}$ from the bid-ask spread equation in **Equation 59**, a neater expression is obtained for the observed return on the index portfolio at time t, as

Equation 61

$$R^{o}_{S,t} = \omega_{S,0}\mu_S + \sum_{k=0}^{\infty} \gamma_{S,k} R^{o}_{S,t-k} + \varepsilon^{*}_{S,t} - \gamma_1 \varepsilon^{**}_{S,t-1} - \sum_{k=2}^{\infty} \gamma_k \varepsilon^{***}_{S,t-k}$$

$$\approx \omega_{S,0}\mu_S + \sum_{k=0}^{\infty} \phi_k R^{o}_{S,t-k} + \varepsilon_{S,t} - \sum_{k=1}^{\infty} \theta_k \varepsilon_{S,t-k}$$

where the second step requires an assumption of homoskedasticity.

The three distinct error terms in the identity are linear combinations of the original error terms introduced at various stages of Stoll and Whaley's reasoning as the bid-ask index error component $\vartheta_{S,t}$ (obtained as the weighted sum of the composite error

in **Equation 57**), the observation error $\nu_{S,t}$ from **Equation 59**, and the true process index innovation $\eta_{S,t}$, which is obtained as a weighted sum over *i* from the error term in **Equation 56**. The error terms in the first line will thus be i.i.d. and mean-zero, from the assumptions on the original error terms, but they will not necessarily share the same variance and the authors remark that passing to the second line requires the assumption of homoskedasticity.

Were we to take a closer look at the way these starred error terms are related to the original well-behaved ones, here is what we would find,

Equation 62

$$\varepsilon^{*}_{S,t} = \omega_{S,0}\left(\eta_{S,t} + \vartheta_{S,t}\right) + \nu_{S,t},$$

Equation 63

$$\varepsilon^{**}_{S,t-1} = -\gamma_1 \nu_{S,t-1} - \omega_{S,0}\delta_S \vartheta_{S,t-1}$$

and

Equation 64

$$\varepsilon^{***}_{S,t-k} = \nu_{S,t-k}$$

if I have managed to disentangle the Greek without slipping too grossly.

The important equation here is **Equation 62**, which represents the residual that would be obtained were we to estimate the approximate identity in **Equation 61**, while **Equation 63** and **Equation 64** are 'swallowed up' in the MA term. As it turns out, Stoll and Whaley's whole argument hinges on **Equation 62**. They say: "In the absence of infrequent trading and bid/ask price effects, $\varepsilon_{S,t} = \eta_{S,t}$. Hence, the error term measures the true return innovation in the stock portfolio *S* in period *t*. In the presence of infrequent trading and bid/ask price effects $\varepsilon_{S,t}$ is a *noisy but unbiased proxy* for the true return innovation $\eta_{S,t}$ ..."[58]

In what follows, Stoll and Whaley estimate the model implied by the second line of **Equation 61**, fitting an ARMA(2,3) specification to transaction data, and subsequently apply the resulting 'true innovation process proxy' to the analysis of futures pricing relationships. Somewhat unsurprisingly, they find that autocorrelation

[58] Stoll and Whaley (1990), p.453, my emphasis.

has practically vanished from the resulting index and the lead-lag relationship between cash and futures prices is considerably reduced.

2.2.3 The Miller, Muthuswamy and Whaley model

In Miller, Muthuswamy and Whaley (1994) (henceforth, MM&W), the index autocorrelation problem itself is approached in an essentially similar way as in Stoll and Whaley (1990) and the motivation is exactly the same: deviations from the theoretical futures pricing relationship

MM&W invoke the well known fact that observed futures basis changes are negatively autocorrelated, along with the popular explanation for this phenomenon, i.e. *index arbitrage*. According to this view, when the futures basis widens, say because the price of the futures contract on the index falls relative to the spot price of the underlying index portfolio of stocks, profit seeking arbitrageurs will respond by buying the futures contract and selling the index portfolio in the spot market. This will cause a shift in relative demand for the two products that moves their prices closer together again. This of course works both ways, so the same inference obtains when the price of the index portfolio in the cash market exceeds that of the futures contract, causing the futures basis to revert to its mean. *Prima facie*, this process seems to reflect a financial economist's dream come true, an efficient market where all deviations from a theoretical relationship are quickly exploited and eliminated by an 'invisible hand'.

MM&W ask whether this isn't too good to be true? And in fact, a number of considerations support the view that it is. First, there are the time-series properties of the two components whose difference is defined as the futures basis. A, stock index, as we have seen, is likely to be positively autocorrelated. The futures contract, on the other hand, is a single financial commodity, subject to a bid-ask spread, and if it is autocorrelated, it is more likely to be negatively so, though perhaps to a lesser degree than the index. This, by itself, is likely to induce some negative autocorrelation in the basis changes, even without any arbitrage ever taking place. In particular, if trading is infrequent in some stocks of the portfolio and the futures contract trades continuously, then the price of the index portfolio is updated with a lag as compared to the futures contract. This lagged updating will appear as mean-reversion in the difference between the two quantities, i.e. the futures basis, and negative autocorrelation in its first differences, the basis changes series.

Another point made by MM&W in this context, is that negative autocorrelation is also significant in the basis changes of the futures contract based on the VLIC index already mentioned earlier. In this case the index portfolio is impossible to buy and hold in practice, because the VLIC is a geometric average, implying a continuously reallocated portfolio.

To summarise, although index arbitrage may well be a sufficient condition for negative autocorrelation in the basis changes, it is by no means a necessary one. A simple experiment conducted by MM&W, further corroborates this result. After eliminating from the basis change series all pairs of consecutive price changes exceeding in absolute value a conservative 'transaction cost band', autocorrelation is checked again and compared to that of the original series. The idea is to exclude all instances of mispricing that could have represented a profit opportunity to arbitrageurs in the presence of transaction costs. The resulting reduction in negative autocorrelation turned out to be in the range of one sixth, leading MM&W to conclude that most of the reported mean reversion in the futures basis must be attributed to causes other than index arbitrage. The same conclusion is obtained in a more direct way, by examining available data on index arbitrage trading volume. As a particular example of extreme basis mean-reversion unaccompanied by any index arbitrage at all are the days of the 1987 market crash. This confirms in two different ways that index arbitrage is *not sufficient* to explain the reported amount of negative autocorrelation in basis changes.

As a first step towards establishing the extent to which reported mean-reversion can be attributed to an alternative explanation, i.e. that it is a "statistical illusion", MM&W derive a model of the stock index and the futures price. Although it is relegated to an appendix by MM&W, its derivation is sufficiently central to the subject of this report, and sufficiently different from other approaches that we have seen so far, to justify a statement in some detail. Starting with what is essentially the Scholes-Williams model, MM&W define the individual observed price process as a function of the true process:

Equation 65

$$s_{it}^{o} = (1-\phi)s_{it} + \phi s_{i,t-1}.$$

In this notation s_{it}^{o} represents the reported change in a stock's price at time t, s_{it} the true price change occurring in that period and $s_{i,t-1}$ the true price innovation in period

t-1. The parameter ϕ, $0 \leq \phi \leq 1$, determines the weight of each in the observed price innovation at the end of period t. To obtain an expression for the index, a weighted sum is taken over the index portfolio as before, yielding

Equation 66

$$s_{pt}^{o} = (1-\phi)s_{pt} + \phi s_{p,t-1}$$

where the p indicates an innovation to the portfolio price.

Now this expression is rather limited in scope because it cannot account for *nontrading*, as defined by MM&W, i.e. when stocks skip a trading period entirely, a point already made by a number of critics of the Scholes-Williams nonsynchronous trading formula. For this reason they suggest that it may be extended by assuming that the weight attributed to the true return lagged one period can be distributed over an infinite number of lags, declining exponentially with the order of the lag. This idea leads to the formulation

Equation 67

$$s_{pt}^{o} = (1-\phi)s_{pt} + \left[(1-\phi)\phi s_{p,t-1} + (1-\phi)\phi^{2} s_{p,t-2} + (1-\phi)\phi^{3} s_{p,t-3} + \cdots\right]$$

where the weights on lags one to infinity of the index innovation sum to ϕ by the formula for a geometric series. Now the term in square brackets is simply equal to $\phi s_{p,t-1}^{o}$. This yields an expression for infrequent trading effects in the stock index as the 'modified' AR(1) process in reported returns,

Equation 68

$$s_{pt}^{o} = \phi s_{p,t-1}^{o} + (1-\phi)s_{pt}$$

where the modification involved refers to the non-standard coefficient on the contemporaneous innovation term. This ingenious transformation opens the way to the estimation of a 'true index proxy', just as in the earlier article of Stoll and Whaley, but using a more parsimonious model. While the way this is achieved in MM&W may seem audacious, it certainly leads to a very convenient model.

This model of nonsynchronous trading leads to the expression

Equation 69

$$\sigma_{s^o}^{2} = \left(\frac{1-\phi}{1+\phi}\right)\sigma_{s}^{2},$$

which is a formula for the observed variance of the index in terms of the nontrading parameter ϕ and the true variance of the index portfolio.

To use this model to explain futures basis autocorrelation, MM&W now use basis change time series to estimate the model implied by **Equation 68** in the form

Equation 70

$$s_t^o = \alpha + \phi s_{t-1}^o + \varepsilon_t .$$

Comparing **Equation 68** to **Equation 70**, we notice that it will yield an estimate $\hat{\varepsilon}_t$ of the true return innovations times a constant, as well as an estimate $\hat{\phi}$ of the constant. On the assumptions of the model, we then have a consistent estimator of the true index return innovations as

Equation 71

$$e_t = \frac{\varepsilon_t}{1-\hat{\phi}}$$

Comparing the inferred index innovation series $\hat{e}_t$ to the observed changes in the index level, MM&W find that the overall autocorrelation is reduced by more than two thirds, from 0.128 to 0.039. Replacing the observed index level change series by $\hat{e}_t$ to calculate the basis changes reduces the absolute value of its first order autocorrelation by about one third in their sample, from 0.369 to 0.252. MM&W maintain that this is more of an improvement than it seems, interpreting the change as a 47% drop in the explanatory power of past basis changes. They hypothesise that a more elaborate model would pick up some of the remaining spurious elements. One suggested extension is to allow for intraday patterns in nontrading instead of forcing the nontrading parameter $\hat{\phi}$ to be constant.

2.2.4 Three recent attempts to obtain a 'true index returns process'

In this section we will examine three different ways of obtaining a 'true stock index', free of the effects of infrequent trading. They all are motivated by research into the behaviour of the futures basis and two of them focus primarily on a single event: the crash on 19th October 1987. The methods proposed, however, will apply whenever extensive nontrading results in an autocorrelated index, whether this is on a temporary or a permanent basis.

2.2.4.1 Weighted least squares: Harris (1989)

The stock market crash on 19th October 1987 spurred a great deal of research, initiated both by academics and government regulatory bodies. Obviously, given that 'an orderly market' is in the best interest of the public, the latter had an interest to find out what happened, why it happened, and how it might be made less likely to happen again. One of the things that happened was that some stocks, including the largest capitalisations, ceased trading for extended periods of time, in markets where trading is otherwise nearly continuous. Also, starting on Friday the 16th, and during the week that followed, the bases of the major futures contracts gravely misbehaved, growing out of proportion as changes in the stock index lagged developments in the futures markets. In the cash market order imbalances prevailed, making execution of sell orders difficult, but congestion also appeared in the futures market, notably on Tuesday, when the S&P500 futures contract was temporarily suspended by the Chicago Mercantile Exchange.

With the wealth of recorded high frequency transaction data, it was now technically feasible to examine in detail the course of events within the trading day. An interesting aspect of such analysis, adressed in a number of studies and government sponsored reports, is the question to what extent the observed anomalies in the basis under such extreme conditions are due to infrequent trading? [59] Harris (1989) assumes that the abnormally large futures basis observed during this period could also have resulted from 'disintegration' of these existing to trade what is fundamentally the same risk. Although it may be difficult to separate the nontrading and disintegration on a philosophical level, as nontrading in the cash market already represents one kind of 'market disintegration', Harris suggests a way to quantify that portion of the large cash-futures spread which is due to infrequent trading. The true value of the index portfolio is unknown, but an estimator superior to the conventional index number may be constructed on the basis of information contained in the prices of trading stocks, using the fact that covariances between stocks are not zero. Comparing this estimator of the 'true index' to the futures price would then lead to an estimate of the basis, corrected for infrequent trading effects.

---

[59] Some references to such studies can be found in Harris (1989), Stoll and Whaley (1990), Miller, Muthuswamy and Whaley (1994)

As an estimator of the true value of the index portfolio, Harris proposes to use a weighted-least squares method. He makes the customary distinction between the *price* of the index portfolio, as expressed in the index formula,

Equation 72

$$S_t = \sum_{i=1}^{N} q_i P_{it}$$

and it's *value*,

Equation 73

$$S_t^* = \sum_{i=1}^{N} q_i V_{it}$$

which is assumed to be an underlying unobservable continuous process, while prices are only observed intermittently, and differ from the current value, in general, if they are old. Now the difference between the two can be expressed as follows:

Equation 74

$$S_t^* - S_t = \sum_{i=1}^{N} q_i (V_{it} - P_{it}) = \sum_{i=1}^{N} q_i \Delta_{k_{it}} V_{it} \equiv A_t$$

where k is the number of periods since the stock last traded (zero if it traded at t), and Δ is the difference operator. Now A_t is the 'nonsynchronous trading adjustment' that must be estimated to obtain the corrected index value.

A true returns generating process is assumed, of the form

Equation 75

$$\Delta \log(V_{it}) = f_t + e_{it}$$

with f_t a common return factor and e_{it} is a zero mean idiosyncratic error. For the situation where a set of common factor estimates $\{\hat{f}_t\}$ are available, Harris proposes to estimate the unobserved change in stock value over multiple non-trading periods, by

Equation 76

$$\Delta_{k_{it}} \hat{V}_{it} = P_{k_{it}} \exp\left\{\sum_{i=1}^{k_{it}} \hat{f}_{t-i+1}\right\}$$

where $P_{k_{it}}$ is the price observed k_{it} periods ago. This means in fact, that S_t^* can be calculated directly from the factor estimates. For a value weighted index, the problem then is to minimise the sum of squares

Equation 77

$$\sum_{i=1}^{N} w_i (\%\Delta P_i - f_t)^2$$

with respect to f_t, where the w_i, are the index value weights. Seen as a regression formula, this corresponds to the model

Equation 78

$$\%\Delta P_{it} = f_t + e_{it},$$

with the assumption that the variance of the firm specific error e_{it} is inversely proportional to the value weight w_i , if the approximation on the left hand side is valid, this is equivalent to **Equation 75**. Now the minimising value of f_t is equal to the observed percentage change in portfolio value over the period.

Equation 79

$$\hat{f}_t = \frac{\sum_{i=1}^{N} w_i \%\Delta P_{it}}{\sum_{i=1}^{N} w_i} = \sum_{i=1}^{N} \frac{q_i \Delta P_{it}}{S_{t-1}} = \%\Delta S_t .$$

In principle this formulation can be used to estimate the 'true index return' in each period, $\%\Delta S_t^*$, whenever more than two stocks do trade. Using the most recent reported trade to calculate the index leads to the nontrading problem that we have grown to know and like, because zero-returns are unjustifiably substituted for the unobserved ones in periods when stocks do not trade. In the regression framework outlined above. Thus it seems that excluding these 'observations' that are known to be false (in general), say j_t in number, and estimating the 'general change in prices' from the remaining N-j_t observations, having accounted for the change in weighting, would result in an unbiased estimate at each time t. **Equation 76** would then be used to obtain the portfolio change in value over any given period. The variance-covariance matrix of the regression would then be an indicator of it's true efficiency, permitting confidence intervals to be attached to each point estimate. A weakness of this

procedure may be thought to be that strong assumptions are needed about the form of heteroskedasticity.

The above, however, is not exactly what Harris proposes to do, in order to obtain an adjusted index for the days of the 1987 stock market crash. Instead, he suggests a multiperiod generalisation of **Equation 78** to read

Equation 80

$$\Delta_{k_{it}} \log(P_{it}) = \sum_{j=1}^{k_{it}} f_{t-j+1} + e_{it}$$

for all observed P_{it} in a cross-section over N stocks and T-t_0 time periods. To the extent that the approximation $\Delta \log P \approx \%\Delta P$ can be thought to hold, for each period t, we then have $\hat{f}_t = \%\Delta S_t^*$ and the corresponding relation will hold for multiple nontrading periods, resulting in a series of adjusted returns on the market portfolio.

The conclusions reached by Harris on the basis of this procedure, are that the size of the basis spread is significantly reduced by the use of the adjusted index, but far from eliminated. He infers that other factors besides nontrading, i.e. 'market disintegration', played an important role in the large observed basis spread during the crash, and he specifically mentions the specialist's obligation to 'walk' prices up and down for orderliness. Using Harris' whole sample of two business week (726 observations), the reduction in the first order autocorrelation of the adjusted index as compared to the S&P500, is by some 34%. On some days, it is much larger, including the day of the crash, when it is reduced from 0.554 to 0.248, i.e. by 55%. On other days, however, the reduction in autocorrelation of estimated index returns is surprisingly small. Thus non-trading was extensive on Tuesday following the crash, but effective reduction in the first order index autocorrelation is negligible, from 0.873 to 0.819, or 6%. In this respect Harris' empirical results are equally puzzling as those of in Perry (1985) and Atchison, Butler and Simonds (1987), reinforcing the idea that there it is indeed necessary to delve deeper into the relationship between nontrading and index autocorrelation than has been done so far.

2.2.4.2 The Kalman-filter way: Bassett, France and Pliska (1991)

This paper, just as the one of Harris, is based on the idea that the abnormally large futures basis during the 1987 crash can be better understood by applying a better

estimator of the value of an index portfolio than the customary stock index. The time span investigated by BF&P is somewhat shorter than the full two business weeks of Harris, including only Monday the 19th and Tuesday the 20th October, but their period grid is denser with each period representing one minute of real-time trading instead of five.[60] Instead of the broadly based S&P500, BF&P investigate the futures contract based on the Major Market Index (MMI), which at the time of their study was composed of only 20 of the largest capitalisation U.S. stocks.[61] Like Harris, and in contrast to Stoll and Whaley who consider the index as the primitive time-series, they take the price observations of individual stocks as their point of departure. The basic idea of their approach is the same as in Harris (1989), i.e. that because stocks are contemporaneously correlated in the market, observations missing because of nontrading can be estimated on the bases of information obtained from trading stocks. The most important difference, however, between the methodology of the Harris and the BF&P papers lies in the way estimates are obtained to substitute for missing observations. While Harris uses Weighted Least Squares, BF&P use a Kalman-filter estimation procedure.

The Kalman filter is a recursive algorithm that computes a one period forecast of a stochastic variable x on the basis of it's present level and its conditional distribution at time t. Once an observation becomes available, say at time $t+k$, the prediction error is used to update a set of weights used to form the next prediction. If the variable is observed in every period, k is always equal to one and the Kalman-filter estimator is always updated but if some observations are missing, the updating steps are simply omitted, although forecasting continues on the basis of the last available update. This will be reflected in the mean-square-error (MSE) of the estimate at time t, which provides a measure of estimate precision. In other words, missing observations will lead to larger confidence intervals, everything else being equal. As the Kalman-filter only depends on the conditional distribution of the process, i.e. the distribution of the increments, it can be easily applied to logarithms of stock-prices directly, even though they represent a non-stationary process, and if the log-normal model holds, it can be shown to be optimal in the sense that it is the minimum-mean square error (MMSE)

---

[60] In what follows we will refer to both the authors and their paper, i.e. Bassett, France and Pliska (1991) as BF&P, confident that it will be inferred from the context which applies in each case.

[61] As was indicated earlier, this index is an exact replica of the DJIA.

estimator of $x_{..}$.[62] It is straightforward to extend the Kalman-filter algorithm to take account of contemporaneous correlation in a multivariate time-series framework, so that a partial update takes place in the estimate of nontrading stocks, whenever that one or more of the other stocks' prices are observed.

The Kalman-filter is based on the so called 'state-space representation' of a dynamic system. The idea is that at any given time the system has an underlying stochastic state that is expressed with error in the observations. This idea can be sketched by the following set-up, which is quite similar to that in BF&P:

Equation 81

$$\text{measurement equation:} \quad Y_t = X_t + Z_t$$
$$\text{state equation:} \quad X_{t+1} = X_t + W_t$$

Here the state is a random walk, driven by the white noise error process W_t and the measurement error process Z_t is the difference between the observation Y_t and the current state X_t. The implementation in BF&P consists in interpreting X_t as a 1x20 vector of true values of the MMI stocks, W_t as the 'true price innovation', Y_t. As not all prices are observed because of nontrading, X_t is preceded by a selection matrix in the BF&P formulation of the measurement equation, designed to pick out the trading stocks. The 'equally weighted' index is then calculated from the most recent estimates of the state vector X_t using an unweighted average, and the MSE-matrix of the system is used in the same way to obtain confidence intervals on the basis of average root-mean-square-error (RMSE) an each point estimate.[63]

Major determinants of the behaviour of the Kalman-filter, are the variances of the two error terms in **Equation 81**. This relationship is assumed to be known in theory, but in practice it isn't, and must be decided by the investigator, usually on the basis of some *a priori* considerations. In intuitive terms, if the measurement equation variance is large, the current observation will be reflected in the estimate to a lesser extent, other things equal. This means that the estimates tracing out the transitions of the state between periods will be smoother than otherwise. They shouldn't be *too* smooth, though, as this means true changes in the state are being mistaken for measurement

---

[62] If the increments are not normal, it is still optimal, but only within the class of linear estimators. This result, though, assumes that the relevant true variances are known. For the purposes of the present stude, the basic reference on the Kalman-filter is Harvey (1989)

[63] The MMI, just like the DJIA is in fact implicitly 'price-weighted' as it is the average value of one share of each company.. A truly equally weighted index portfolio has to be continually reallocated to keep the value shares of all firms constant as prices vary.

noise. As for measurement error, BP&F provide a quantitative estimate of the measurement error covariance as a diagonal matrix with all the elements equal to 0.005. The underlying statistical model is a simple random walk without drift, which implies that all remaining variability results from changes in the state-itself.

As the Kalman filter can be run in real time even when no observations are being made, the question arises whether time periods when the exchange is closed should be included or not. If they are, the MSE of the estimate will increase proportionately to the number of periods, causing the confidence interval to balloon at opening. If the intervening time periods are not included, the confidence intervals for opening price estimates will be unrealistically tight, based on dense trading at the previous close. In this case they will almost surely not 'capture' the opening trades. In the approaches treated earlier, this question does not arise because the methods used do not accommodate missing observations. Thus Stoll and Whaley and MM&W, for example, estimate each day's parameters separately, and then average them over days to produce overall results. BF&P, however, discuss this problem and choose a compromise solution, based on setting the 1050 minutes (17.5 hours) when the NYSE exchange is closed equal to some smaller number of 'equivalent trading minutes', like 60 or 120. With a large number of days in the sample, it is likely that such a compromise might be made on some more or less solid grounds, but with a two day sample as in BF&P, it seems a little bit arbitrary. This, however, is a good example of the kind of 'calibration' decisions that have to be made when implementing a Kalman-filter.

The use of the BF&P Kalman-filter index resulted in a reduction in the size of the futures basis as expected, mostly around 50%, but it did not eliminate the excess spread. The overall conclusion, was that "...non-synchronous trading explains a small, but significant portion of the cash-futures spread that prevailed during these days."[64] In their view, the results are consistent with other attempts that were made to quantify the role of infrequent trading in basis anomalies during the crash, such as that of Harris (1989), considered above, their own earlier results and those of a number of other scholars. They speculate on a number of reasons why the Kalman-filter procedure is not more efficient in eliminating excess spread. Some possible reasons suggested by BF&P have to do with the underlying statistical assumptions, such as

[64] BF&P (1991), p.135

normality, and the assumption that nontrading and the price change are independent of one another. If the reason that stocks do not trade is precisely that of large order imbalances caused by a sharper than average change in prices, then any resulting estimator of that particular stock's value that is based on prices of trading stocks will be biased toward zero. Other concievable reasons, according to BF&P, include the idea that certain institutional structures guaranteeing the equivalence of the two financial products may have been undermined by the severity of the financial crisis. This is a 'real' reason for cash and futures prices to diverge, as the two products are no longer close substitutes under looming insolvency and non-negligible default risk on existing contracts. When this occurs markets are not efficient anymore, and there is no reason to expect the theoretical pricing relations to hold.

2.2.4.3 An elaborate ARMA approach: Jukivuolle (1995)

Jukivuolle (1994) grapples the problem of "assessing the true stock index value when some constituent stocks of the index do not trade in every period..."[65] He invokes the Stoll-Whaley model of infrequent trading discussed above, but proposes three major modifications. They are aimed at making the chain of reasoning in that paper at once simpler, more consistent, and more rigorous. As we will remember, the Stoll-Whaley model aims to account not only for infrequent trading, but also the effects of the bid-ask spread. Jukivuolle drops the bid-ask component of the model, and although he does not discuss any particular justification for doing so, it may be assumed that that decision is implicitly based on the idea that these effects will cancel out in large portfolio which is either equally weighted or has value-shares are of order $1/n$, with n equal to the number of stocks in the portfolio. The index he considers, the *Russell 2000*, is in fact equally weighted, based on the 2000 smallest from the 3000 largest capitalisation U.S. companies. Thus the bid-ask spread is as unlikely to have an effect on index autocorrelation as can reasonably be expected for an actual index portfolio.
The second modification concerns the long-run relationship between reported and true returns. Disregarding the bid-ask spread, the Stoll-Whaley model of index returns under infrequent trading can be stated as the following two equations:

Equation 82

$$R_t = \mu + \varepsilon_t \qquad \text{with} \quad \varepsilon_t \propto N(0, \sigma_\varepsilon^2)$$

and

Equation 83

$$R_t^o = \sum_{k=0}^{n-1} w_k R_{t-k} + \nu_t \qquad \text{with} \quad \nu_t \propto N(0,\sigma_\nu^2),$$

which together yield an expression from the observed index returns in terms of the true index innovation process and the measurement error ν_t:

Equation 84

$$R_t^o = \mu + \sum_{k=0}^{n-1} w_k \varepsilon_{t-k} + \nu_t$$

Now this formulation implies that true and observed returns are not cointegrated, and consequently the true and the observed index levels will be able to drift arbitrarily far from each other. To see this, consider the system of equations involving true and observed returns

Equation 85

$$\begin{bmatrix} R_t \\ R_t^o \end{bmatrix} = \begin{bmatrix} X_t - X_{t-1} \\ X_t^o - X_{t-1}^o \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \sum_{k=0}^{n-1} w_k B^k & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \nu_t \end{bmatrix}.$$

Now the two series are cointegrated only if the matrix

Equation 86

$$\begin{bmatrix} 1 & 0 \\ \sum_{k=0}^{n-1} w_k & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

is singular, which is evidently not the case. This is an inconsistency in the Stoll-Whaley formulation, and Jukivuolle suggests that it be removed before proceeding any further. A simple way to do this is to set the lower-right element of **Equation 86** to zero, making its determinant zero too. This is equivalent to dropping the error term ν_t from **Equation 83**. This seems objectionable, if one thinks of this term as a measurement error, as it is hard to imagine anything being in fact observed without error. However, the idea of reported an true returns drifting arbitrarily far apart as a result of measurement error is no less counterintuitive. Jukivuolle, sees this as "the simplest and most straightforward way to deal with the problem", and he

---

[65] Jukivuolle (1995), p.455

believes that "making the relationship in [Equation 83] deterministic does not change the empirical plausibility of the model much."[66] As it turns out, this is quite convenient, because on this assumption **Equation 84** reduces to a "standard infinite order MA process" for the observed index returns.[67]

The third departure from the Stoll-Whaley model concerns the way in which the 'true process' is derived from the observed one. While Stoll and Whaley claim that the white-noise residual process from an ARMA(p,q) model of index returns will be "a noisy but unbiased proxy" for the true return innovations, Jukivuolle uses a decomposition theorem presented in Beveridge and Nelson (1981) to obtain a random walk 'true index' as the so-called permanent component of the observed index portfolio price process in levels. His result is much stronger than the earlier one, in that the resulting estimated index proxy is not only "noisy, but unbiased" like the return innovations of Stoll and Whaley, but "perfectly correlated" with the random walk 'true index level' postulated by the efficient market model. This, of course, is hardly surprising, given the restrictive assumptions of Jukivuolle's model. Having eliminated two out of three random error processes that went into the observed error in the Stoll-Whaley model, the resulting observed error will of course coincide with the remaining true error.

Turning the Stoll-Whaley model in **Equation 83** into levels, relaxing the assumption of a maximum degree of nontrading equal to n, dropping the measurement error term v_t and gathering terms in k, Jukivuolle now writes the basic model as

Equation 87

$$X_t^o = \sum_{k=0}^{\infty} w_k X_{t-k} + \left(\sum_{k=1}^{\infty} k w_k \right) \mu$$

where the weights are non-negative, non-increasing in k and sum to one. Differentiation gives the 'corrected model'

Equation 88

$$R_t^o = \sum_{k=0}^{\infty} w_k R_{t-k}$$

Substituting from **Equation 82** for the true process yields **Equation 84** all over again, but now in corrected form

---

[66] Jukivuolle (1995), p.458n and 457, respectively.

Equation 89

$$R_t^o = \mu + \sum_{k=0}^{\infty} w_k \varepsilon_{t-k}$$

and this leads to an expression for the observed return process as an infinite moving average in the *observed* error terms, i.e.

Equation 90

$$R_t^o = \mu + \sum_{k=0}^{\infty} c_k \varepsilon_{t-k}^o$$

where the crucial step consists in showing that.

Equation 91

$$w_0 \varepsilon_t = \varepsilon_t^o,$$

which is reminiscent of the relationship between the true and the observed error in the Miller, Muthuswamy and Whaley formulation. If this process is invertible, which depends on the coefficents c_t in the lag polynomial **Equation 90**, then it has an ARMA(p,q) representation, which can be estimated from the data. We note that this is where the theoretical considerations presented in Stoll and Whaley (1990) ended, in fact a little earlier, even, because they didn't present any justification for the step from their ARMA(∞,∞) to an ARMA(2,3).

Now the Beveridge-Nelson theorem is obtained by writing the k-period forecast of an ARIMA(p,1,q) process z_t as the present observation plus an accumulation of the expected value of a stationary series of its first differences w_t, i.e.

Equation 92

$$\hat{z}_t(k) = z_t + \hat{w}_t(1) + \cdots + \hat{w}_t(k)$$

where

Equation 93

$$\hat{w}_t(i) = \mu + \lambda_i \varepsilon_t + \lambda_{i+1} \varepsilon_{t-1} + \cdots = \mu + \sum_{j=1}^{\infty} \lambda_j \varepsilon_{t+1-j}$$

by another decomposition theorem, the one due to Wold. Substituting **Equation 93** into **Equation 92**, and gathering terms in the appropriate lags of ε_t results in

---

[67] Jukivuolle, p.457-458

Equation 94

$$\hat{z}_t(k) = k\mu + z_t + \left(\sum_{i=1}^{k} \lambda_i\right)\varepsilon_t + \left(\sum_{i=2}^{k+1} \lambda_i\right)\varepsilon_{t-1} + \cdots$$

and using the stationarity assumption to assure convergence of the sums in λ_i an approximation to an 'asymptotic forecast profile' is obtained as

Equation 95

$$z_t(k) \cong k\mu + z_t + \left(\sum_{i=1}^{\infty} \lambda_i\right)\varepsilon_t + \left(\sum_{i=2}^{\infty} \lambda_i\right)\varepsilon_{t-1} + \ldots$$

when k goes to infinity. This leads to the definition of a *permanent* or *trend component* of z in terms of its current level, z_t and a *transitory, or cyclical* component c_t

Equation 96

$$\bar{z}_t = z_t + \lim_{k\to\infty}\left\{\left[\hat{w}_t(1) + \cdots + \hat{w}_t(k)\right] - k\mu\right\} = z_t + c_t,$$

interpreted by Beveridge and Nelson as "the current observed value of z plus all forecastable changes in the series beyond the mean rate of drift."[68] It is easily shown that on the assumptions the permanent component is a random walk.

It is interesting that this result exploits the ubiquitous 'intervaling effect' to derive the trend as an asymptote.

Jukivuolle now writes the index price process in the Beveridge-Nelson form **Equation 96** as

Equation 97

$$\bar{X}_t^o = X_t^o + \lim_{T\to\infty}\left[\sum_{j=1}^{T-t} \hat{R}_t^o(j) - (T-t)\mu\right]$$

where $\hat{R}_t^o$ are the optimal time t forecasts of the stationary process as in **Equation 96**. This leads to a relatively straightforward proof that the Beveridge-Nelson permanent component of the log of the observed index level equals the log of the true index level, i.e.

Equation 98

$$\bar{X}_t^o = X_t$$

[68] Beveridge and Nelson (1981), p.156

2.3.1 Campbell, Lo and MacKinley (1997)

Continuing a series of articles by Lo and MacKinley in 1988-1990, the authors devote a chapter of their textbook of empirical finance to the econometric analysis of market microstructure phenomena. A subsection of this chapter, *Nonsynchronous Trading*, is devoted to the treatment of infrequent trading. The other three treat the bid-ask spread, the particularities of working with high-frequency transaction data and provide a survey of empirical research.

Despite the name of the subsection, the CLM model is a *nontrading model*, as it assumes a discrete period structure and analyses the results of some securities' skipping periods. As to its basic structure, this model has much in common with the other nontrading models we have seen. Thus it assumes a market model of the CAPM type and defines continuously compounded returns on the index by approximation as the sum of the log-differences of individual security prices. As in the CMSW and Stoll and Whaley models, to name but two, the CLM model is designed as a model of time aggregation over periods when price is not observed. The CLM-model, however, differs from these studies in that a great deal of attention is given to details of the stochastic weighting process that controls time-aggregation. This leads to some interesting results; one example is that CLM are able to derive estimates of 'nontrading probabilities' in individual portfolios from the autocorrelation matrix of a group of portfolios. Such results, however, are unlikely to be directly applicable, because of the restrictive assumptions involved.

The essential idea of the CLM model can be stated in a simple and intuitive way, starting with the by now familiar components, true and observed returns of single security at time t, (r_{it} and r_{it}^{o}) as well as their particular *nontrading probabilities* π_i, where the i, as usually, refers to a security i. If security i does not trade in period t, no price update occurs and hence no return is observed. If it does, the observed return depends on for how many periods preceding t it hasn't traded, for then the observed return will be a time-aggregate of the true return innovations for all skipped periods since the last trade. As the security either trades or it doesn't, the process controlling the relationship of true and observed returns is binomial. This leads to the following:

---

[70] Abbreviated to CLM, in what follows.

Equation 102

$$r_{it}^{o} = \begin{cases} 0 & \text{with probability } \pi_i \\ r_{it} & \text{with probability } (1-\pi_i)^2 \\ r_{it} + r_{i,t-1} & \text{with probability } (1-\pi_i)^2 \pi_i \\ r_{it} + r_{i,t-1} + r_{i,t-2} & \text{with probability } (1-\pi_i)^2 \pi_i^2 \\ \vdots & \vdots \\ r_{it} + \cdots + r_{i,t-k} & \text{with probability } (1-\pi_i)^2 \pi_i^k \\ \vdots & \vdots \end{cases}$$

This leads to serial correlation in observed returns because there is a positive probability that they will in fact represent a sum of some number of past virtual returns. To formalise this, CLM define a stochastic 'weighting variable' $X_{it}(k)$ so that the observed return process can be expressed in a way consistent with the idea in **Equation 102**, as

Equation 103

$$r_{it}^{o} = \sum_{k=0} X_{it}(k)\, r_{i,t-k} \quad i = 1,2,\ldots,N$$

To define X in a way consistent with **Equation 102**, an indicator variable δ_{it} , is used. It is zero if stock i trades in period t, (i.e. with probability $(1-\pi_i)$) and one otherwise (with probability π_i). Then for $k > 0$,

Equation 104

$$X_{it}(0) \equiv (1-\delta_{it})$$

and

$$X_{it}(k) \equiv (1-\delta_{it})\delta_{i,t-1}\delta_{i,t-2}\ldots\delta_{i,t-k}$$

$$= \begin{cases} 1 & \text{with probability } (1-\pi_i)\pi_i^k \\ 0 & \text{with probability } 1-(1-\pi_i)\pi_i^k \end{cases}$$

The trade-indicator sequences $\{\delta_{it}\}$ and $\{\delta_{jt}\}$ are assumed independent for $i \neq j$ and over time, as well as being identically distributed over time. This set-up implies that the probability of $X_{it}(k)$ being zero in **Equation 104** for large $\bar{k}$ is high, and true returns earlier than this $\bar{k}$ will not influence the observed return in period t. To complete the statement of the basic nontrading model, CLM derive an alternative expression for reported returns in terms of the *duration* of nontrading, defined as

Equation 105

$$\tilde{k}_{it} \equiv \sum_{k=1}^{\infty}\left\{\prod_{j=1}^{k}\delta_{i,t-j}\right\}.$$

Then **Equation 104** can be translated into the simple formula

Equation 106

$$r_{it}^{o} = \sum_{k=0}^{\tilde{k}_{it}} r_{i,t-k} .$$

which expresses observed returns as the sum of a random number of random terms. This formula in a discrete model, infrequent trading is essentially random time-aggregation. It also suggests a natural way to think about *nonsyncronous trading* when an underlying continuous process is thought to be sampled at random discrete intervals, i.e. by replacing the summation operator by an integral.

To see the relationship between the *duration* of nontrading and the nontrading *probability* we note that the first two moments of $\tilde{k}_{it}$ are

Equation 107

$$E\left[\tilde{k}_{it}\right] = \frac{\pi_i}{1-\pi_i} \quad \text{and} \quad Var\left[\tilde{k}_{it}\right] = \frac{\pi_i}{(1-\pi_i)^2}.$$

The advantage of the formulation in **Equation 103**, over that in **Equation 106**, is that it permits to make explicit the exact way in which observed returns and their moments depend on the nontrading process $X_{it}(k)$, and in particular, on each stock's nontrading probability π_i. As we might expect, in the light of earlier nontrading models, it can be shown that the expected value of returns is not affected

Equation 108

$$E\left[r_{it}^{o}\right] = \mu_i .$$

The conclusions concerning higher theoretical moments of individual returns are also consistent with the analysis of Scholes and Williams and CMSW, but can now be stated quite clearly in terms of nontrading probabilities. As before, the reported variance of individual returns is an increasing function of the extent of nontrading and decreasing in the coefficient of variation:

Equation 109

$$Var\left[r_{it}^{o}\right] = \sigma_i^2 + \frac{2\pi_i}{1-\pi_i}\mu_i^2 .$$

The autocovariance of individual reported returns is negative, decreasing with the length of the lag and increasing with expected return, as

Equation 110

$$Cov\left[r_{it}^{o}, r_{i,t+n}^{o}\right] = -\mu_i^2 \pi_i^n \quad \text{for } n > 0 .$$

Dividing **Equation 109** by **Equation 110**, we obtain the *n*-th autocorrelation coefficient of the observed return series, in terms of the nontrading parameter and the square of the coefficient of variation. Given that the autocorrelation function is continuous and nonpositive, zero for $\pi_i = 0$ and zero in the limit as the probability of nontrading goes to unity, for a fixed coefficient of variation, it will have a minimum with respect to π_i somewhere between zero and one. For first order autocorrelation (*n* = 1), CLM show that this minimum, i.e. the maximum absolute value of the autocorrelation coefficient is reached when

Equation 111

$$\pi_i = \frac{1}{1+\sqrt{2}|\xi_i|}, \qquad \xi_i \equiv \mu_i/\sigma_i ,$$

i.e. the inverse of the coefficient of variation. An absolute minimum is shown to be equal to $-\frac{1}{2}$, but it is "virtually unattainable for any empirically plausible parameter values", according to CLM.[71]

Cross-covariances for a pair of securities can be expressed in terms of their true betas, i.e. their covariance with the market factor *f*, the variance of the market factor σ_f^2, as well as their respective nontrading probabilities π_i, π_j.

Equation 112

$$Cov\left[r_{it}^{o}, r_{j,t+n}^{o}\right] = \frac{(1-\pi_i)(1-\pi_j)}{1-\pi_i\pi_j}\beta_i\beta_j\sigma_f^2\pi_j^n .$$

where an infinite geometric sum of nontrading probabilities has gone to form the divisor, leaving only the lagged one in the power of its lag as a multiplier.

---

[71] CLM(1997), p.90

This expression immediately entails one quite impressive result. If the *N*x*N* autocovariance matrix of returns up to order *n* is defined as

Equation 113

$$\Gamma_n = E\left[(\mathbf{r}_t^o - \mu)(\mathbf{r}_{t+n}^o - \mu)'\right], \quad \mu \equiv E(\mathbf{r}_t^o)$$

then the *ij*'tht element of this matrix, $\gamma_{ij}(n)$, is given by the formula in **Equation 112**. As the gammas are functions of nontrading probabilities, it is clear that the matrix Γ_n will be asymmetric if these differ across securities. This leads to the observation that

Equation 114

$$\frac{\gamma_{ij}(n)}{\gamma_{ji}(n)} = \left(\frac{\pi_j}{\pi_i}\right)^n$$

which means that we only have to estimate the nontrading probability of one security in a given portfolio, to be able to calculate the rest from the autocovariance matrix.

The asymmetry of the autocovariance matrix also formalises an intuitive consequence of varying trading frequency across stocks, i.e. that the history of the more frequently trading stocks permits to forecast future prices of less frequently trading ones, but not the other way around.

To take a look at the properties of portfolio returns with respect to nontrading, CLM sum over *N* individual returns as is customary. However, they are much more cautious about this step than authors of other papers we have as yet examined. Thus they define the return on portfolio κ as an approximation

Equation 115

$$r_{\kappa t}^o \approx \frac{1}{N_\kappa} \sum_{i \in I_\kappa} r_{it}^o$$

because the logdifferences of portfolio prices do not equal the sum of logdifferences of individual prices as we have noted earlier.[72]

CLM do not analyse the type of index portfolio we are most likely to encounter in practice, i.e. a value-weighted arithmetically averaged portfolio. Instead they consider hypothetical equally weighted portfolios, each containing a large number of different

stocks. This restriction allows them to apply a law of large numbers to the portfolios, deriving elegant *asymptotic* results where the effect of idiosyncratic error vanishes. One problem with studying actual index numbers in small stock exchanges like the Icelandic one, is that the index portfolios are unlikely to be well-diversified in this sense. Another fateful constraint imposed by CLM is that components of the portfolios have *equal* nontrading probabilities. This obviously somewhat reduces the practical applicability of any results obtained. This of course does not preclude that some insight may be gained into the behaviour of the moments and comoments of observed returns under infrequent trading, and they do allow interesting empirical experiments to be conducted with artificially constructed portfolios.

One interesting result is that the n'th order autocorrelation coefficient of such a portfolio is asymptotically equal to its common nontrading probability, raised to the power of n.

Equation 116

$$Corr\left[r_{\kappa t}^{o}, r_{\kappa,t+n}^{o}\right] \overset{asy}{=} \pi_{\kappa}^{n}.$$

As expected, observed portfolio variance is lower than the true variance, by a function of its nontrading probability and average beta:

Equation 117

$$Var\left[r_{\kappa t}^{o}\right] \overset{asy}{=} \beta_{\kappa}^{2}\left(\frac{1-\pi_{\kappa}}{1+\pi_{\kappa}}\right)\sigma_{f}^{2}.$$

The lag n cross-serial covariance between two portfolios of this kind exactly replicates the relationship for two individual stocks in **Equation 112**, with the individual betas replaced by portfolio average betas.

As we remarked earlier, the infrequent trading models we have seen are essentially time-aggregation models, where we cannot observe the disaggregate series of true returns. This is because the 'disaggregate interval' is smaller than the sampling interval, which is the case for all sampling intervals if the underlying process is continuous. As an alternative way of looking at such phenomena, CLM study the effect of imposing time-aggregation when the disaggregate series is known, by increasing the sampling interval. Thus if the original series contains daily returns, then

---

[72] This is what they do in the book, (p.92), in the corresponding article (1990), they only remark that this is the return of a geometrically averaged portfolio.

it can be studied on a weekly, monthly or quarterly basis, with the order of time-aggregation q equal to 5, 22 and 66, respectively. The basic conclusions of CLM indicate that expected reported returns time-aggregate linearly, but their variances and covariances do not. In particular, as we may expect from the existence of the intervaling bias in beta estimates, the effect of the nontrading probability π_i on reported return moments is monotonically decreasing in q for both portfolios and individual securities.

3 Dealing with the problem

Nobody would contest the practical difficulty of measuring composite prices and quantities.

Wassily Leontief[73]

Part three is divided into three main chapters. As indicated by the heading, here the main emphasis is on empirical and practical aspects of the infrequent trading problem. Ultimately, this more general problem is left behind to tackle a narrower one, that of estimating an all-share index in a particular market, the *Icelandic Stock Exchange* (ISE). Hopefully the results can be useful in other small markets with similar microstructure features. The first section of part three attempts to provide a descriptive account of some aspects of trading in the ISE, which should allow the likely degree of nontrading in this market to be inferred. To help form an idea of expected nontrading, the situation in the ISE is compared with that in another small stock market, the *Helsinki Stock Exchange* (HSE). The second section aims to analyse the all-share index calculated and published by the ISE with respect to evidence of infrequent trading effects to returns such as those predicted by theoretical and empirical research studied in previous sections. The analysis principally involves calculating, graphing and interpreting the autocorrelation function of index return time-series. To provide a frame of reference we also inspect the properties of returns on the all-share index portfolios in two other stock exchanges, one where no infrequent trading effects should be expected because trading is virtually continuous (the NYSE), and one in which research has indicated that severe nontrading persists (HSE). Finally, an alternative way of estimating a stock index is proposed and explained. The suggested estimator, based on a Kalman-filter approach, was implemented for an index number published regularly by the ISE, the *Index of Transportation Firms* (ITF). The index numbers resulting from the Kalman-filter method are then compared graphically to the actual index.

[73] *Essays in Economics*, Vol II. Oxford, Basil Blackwell, (1977), p.126

3.1 The situation in the ISE

The *Icelandic Stock Exchange* can properly be called a fledgling stock market, because it is young and growing very fast. End of year 1993 there were only 17 listed companies, but at the end of 1996 they were 32. During that year 5 new companies were listed and turnover in listed stocks more than doubled from the year before.[74] This trend persists, as until the end of August 1997, ten further companies were listed, making the total number of listed firms 42. Still, it is one of the smallest stock exchanges in the world, and in the present context this suggests that market microstructure effects may hamper its apparent efficiency of trading.

3.1.1 Circumstantial Evidence.

A common measure of the size of a stock exchange is the total market capitalisation of listed firms. **Figures 2** and **3** illustrate this fact to a certain extent, and the figures are summarised in **Table 2.**[75] On this scale the ISE surely ranks among the smallest in the world.

Table 2:International Comparison of Stock Exchanges (in millions of US$)

Stock Exchange	No. of companies	Market value	Avg. firm size	Turnover velocity
NYSE	1.969	3.797.687	1.929	43%
Tokyo	11.651	2.397.371	206	20%
London	1.918	928.393	484	41%
Zurich	180	189.117	1.051	40%
Korea	688	107.661	156	108%
Oslo	115	17.840	155	57%
Helsinki	61	12.205	200	18%
Iceland[76]	29	1.243	43	6%

Although market value of firms is sometimes used as an inverse proxy for their degree of price adjustment lag or infrequent trading[77], it may seem rash to conclude that a

[74] Source: *ISE annual report 1996.*
[75] Source : Hawawini (1994). Figures for Iceland are for 1996, but for all other countries for 1992. This will bias the comparison so as to make ISE look relatively bigger than it is, as we can safely assume that all the exchanges grow over extended periods of time..
[76] All figures exclude firms that were not listed throughout 1996.
[77] E.g. in CMSW (1986), p.131. Also CLM (1997), p.130.

small *market* will necessarily be an inactive one. However, knowing the number of firms in the market and total market capitalisation, we can calculate average firm size.

Figure 2: Total capitalisation of listed firms[78]

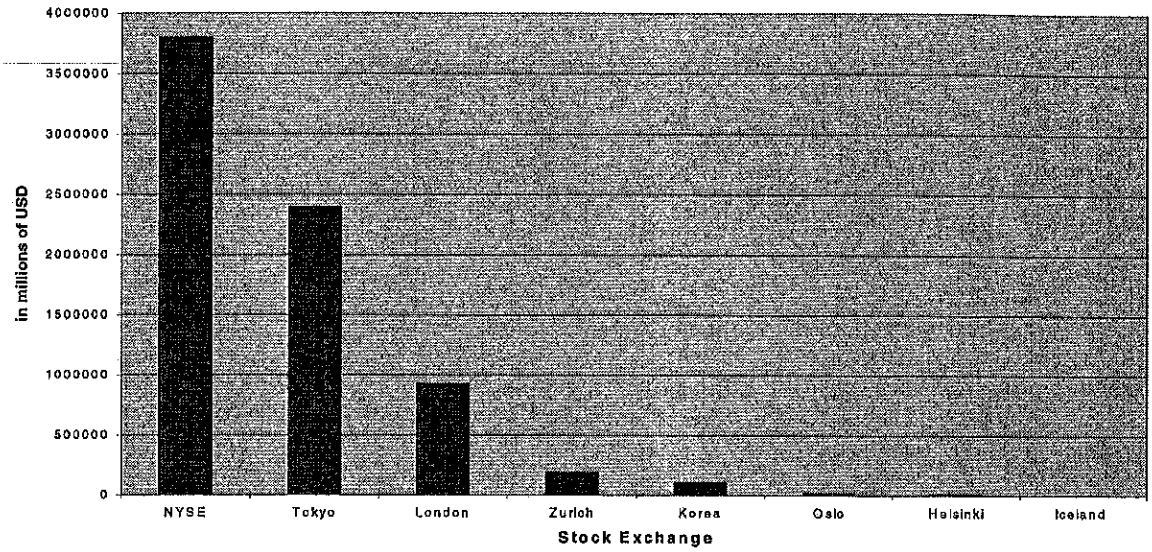
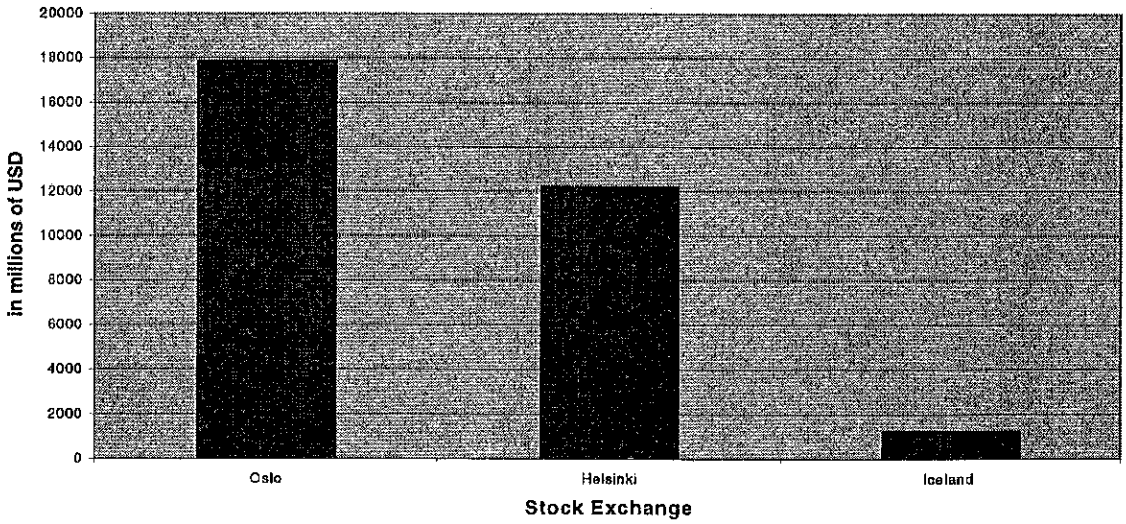


Figure 3:A closer look at the relative size of the ISE[79]

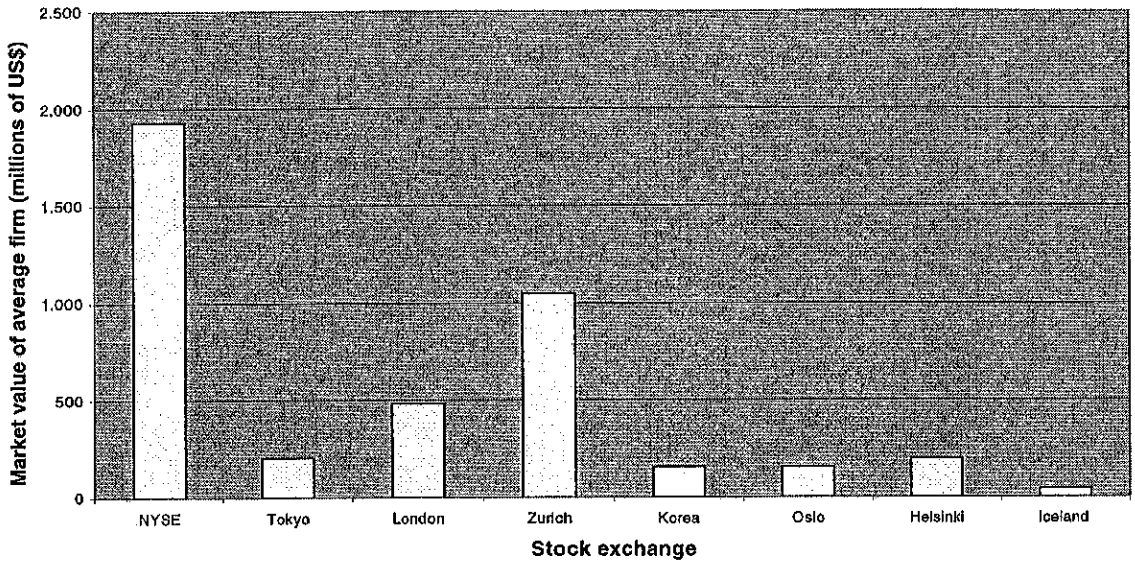


It turns out that the average value of listed firms in the ISE is only a little more than one-fiftieth of that in the exchange that has the largest firms (NYSE), and a little less than one-third of the average firm-size in the Oslo exchange which has the second

[78] Source: *Table 3-1*
[79] Source: *Table 3-1*

smallest firms. **Figure 4** provides a graphical representation of this comparative relationship.

Figure 4:Average market value of listed firms[80]



It can be interpreted as lending some indirect evidence to the idea that infrequent trading problems should be expected in the ISE. Another 'inverse indicator' of thin trading is the *turnover velocity* in a stock exchange, defined as the ratio of the value of shares changing hands in a particular year to total market capitalisation at the end of that year. With everything else equal (in particular the average relative volume of each trade), this measure would be a good indicator of the arrival rate of trades in shares of the average company in the market. Consequently low relative trading value may provide some circumstantial evidence that infrequent trading effects are to be expected. We see from **Table 2**, that the turnover velocity in ISE is one third of that in the stock exchange with the second lowest figure, i.e. the Helsinki Stock Exchange. But as we have no guarantee that a *ceteris paribus* clause is justified even approximately, data on turnover velocity hardly amounts to conclusive evidence.

A substantial amount of research has already been undertaken, exploring infrequent trading effects in the HSE. From any of the resulting papers it can be readily inferred that they are considerable.[81] Whereas in the U.S. market Shanken (1987) found that covariance between stocks approximately doubles when price adjustment delays of up

[80] Source: *Table 3-1*
[81] For examples, see list of references under Martikainen et al. and Kallunki et al.

to three days are accounted for on the basis of the CMSW approach discussed earlier, Kallunki and Martikainen (1997) report an increase by a factor of 4 to 5 for the least frequently trading stocks in the HSE. Although they do not discuss the market index, probably because their research is primarily motivated by the question of applicability of APT models in the HSE, their results indicate substantial positive cross-serial correlation between stocks, which in turn implies an autocorrelated index. This suggests comparing the characteristics of the ISE and the HSE, with an ultimate view to guessing the likely extent of nontrading in the ISE.

The relation between average firm size and relative turnover in the two exchanges, as expressed in **Table 2** lends some support to the view that infrequent trading may be non-negligible in the ISE as well, as we see that firms are both larger in the HSE and shares more liquid in terms of turnover velocity. Having formed a working hypothesis, we will proceed toward a more direct examination of the extent of nontrading in the ISE.

3.1.2 A comparison of nontrading probabilities

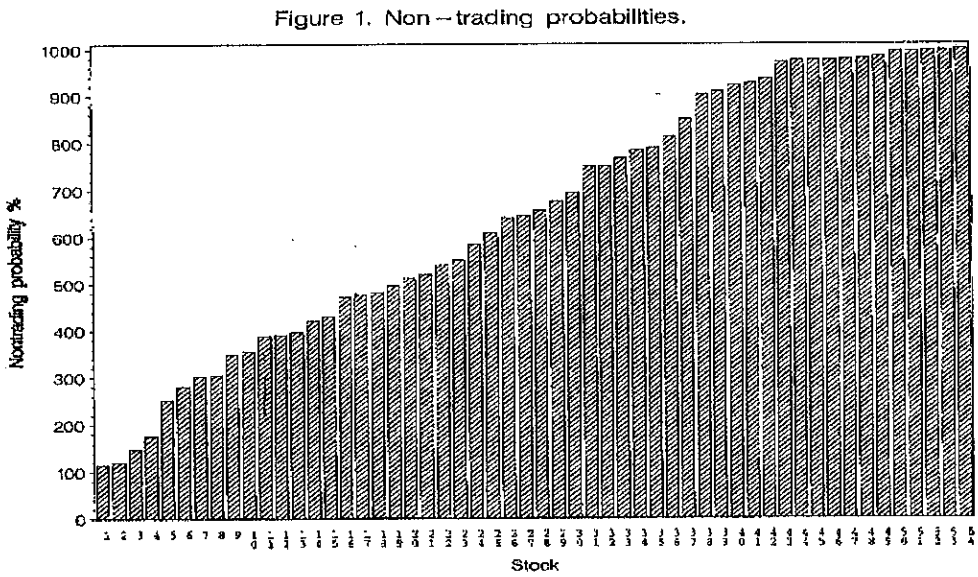
In Kallunki and Martikainen (1997) we find estimates of so called "daily non-trading probabilities" for all 54 stocks listed in the HSE throughout 1990-93. In their sense of this term, the daily nontrading probability of a particular stock means the proportion of trading days on which no trade occurred in that stock, in other words, the *relative frequency* of nontrading days over a period.[82] **Figure 5** reproduces a column graph of the nontrading probabilities in the HSE, that is presented in this paper. To illustrate the problem, Kallunki and Martikainen note that for the Finnish market "the nontrading probabilities in daily return intervals of the 25 most frequently traded stocks are much higher than those observed for the U.S. Major Market Index stocks in five-minute return intervals."[83]

---

[82] This definition is consistent with the one used in CLM (1997), (cf. e.g. p.87), although the perspective is different because CLM are thinking in terms of *à priori* probabilities, not empirical proportions.

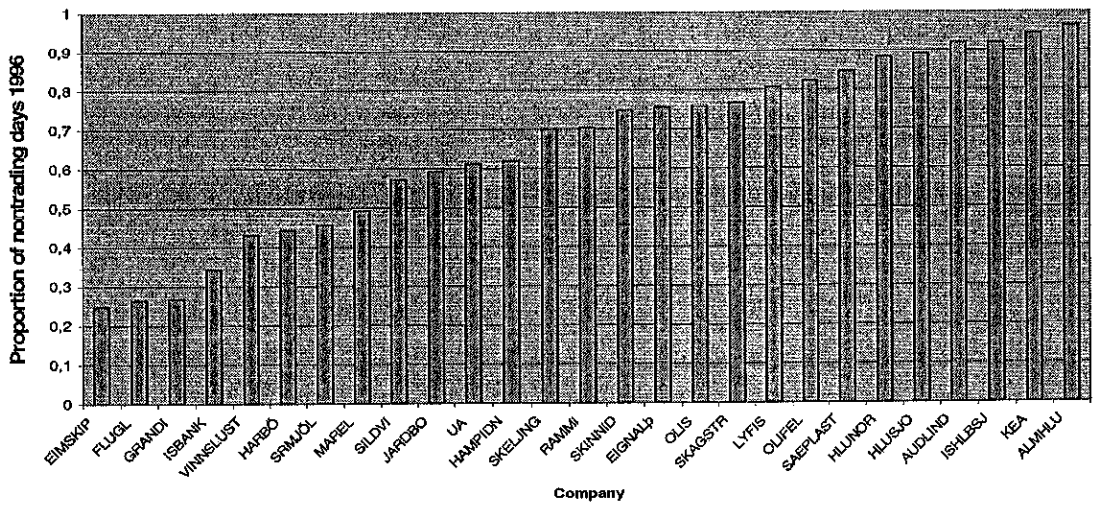
[83] Kallunki and Martikainen (1997), p.4.The MMI was introduced in 1.1.5.2

Figure 5: Relative frequency of nontrading days in the HSE 1990-93[84]



Calculating nontrading probabilities for the 27 stocks that were listed throughout 1996 in the ISE, we obtained the results illustrated in **Figure 6.**

Figure 6: Relative frequency of nontrading days in the ISE in 1996



Comparing the two, although it may be illustrative to some extent, is not entirely straightforward. One fact that has to be kept in mind, is that the data are for 1990-93 in the case of the Finnish exchange, but for 1996 in the case of the Icelandic

[84] The numbers on the *x*-axis represent listed firms. This figure is reproduced exactly as it appears in Kallunki and Martikainen (1997). It seems not to have been proof-read but we will assume that the probabilities are intended to range from 0 to 1. We recall that in Kallunki and Martikainen (1997) a "nontrading probability" is taken to mean the number of days with no trade in a given stock relative to the total number of trading days.

one.However, if the degree of nontrading turns out to be similar, we might be tempted to conjecture that research conducted in the HSE in 1990-93 may give some idea of what to expect in the ISE in 1996. In Kallunki and Martikainen (1997), the data underlying the column graph is not provided, so an exact comparison of the average nontrading probabilities is impossible.

Even so, some indication of the central tendency of nontrading probability in the Finnish market can be inferred from the graph. Thus, in the HSE, judging from the column graph, the median nontrading probability of stocks is approximately 0.65, while it is 0.7 in the ISE. The *mean* nontrading probability in the ISE also lies in this region, it is 0.658. Even if we conclude from this that in some sense of the word nontrading probabilities are *similar* in the two markets, this does not necessarily mean that we should expect the same degree of spurious effects, e.g. as evident in an all-share index. For one, it is a necessary condition for positive autocorrelation effects to appear that the index portfolio be "well-diversified".[85] One way in this may fail to hold is that the number of stocks in the portfolio may be to small. In this case bid-ask spread effects will not cancel out in general. The bid-ask spread in the ISE is probably large because firms are small and trading volume is low, precluding economies of scale. Also, as authorised dealers are few, compared to larger markets, collusion may prevail. Furthermore, if expected returns are large and positive, this entails non-negligible effects from negative serial correlation in individual returns of infrequently trading stocks. This situation is likely to hold in the Icelandic market in this period. The fact that the market value of the ISE all-share index portfolio increased by some 60% during 1996 leaves no doubt that returns are in general large and positive in this market. An additional source of uncertainty is that weights attributed to the least frequently trading stocks may also differ in the two indices, and this will affect the observed properties of the index.[86] A further market characteristic that might cause differences in index return time-series properties between markets with similar mean or median nontrading probabilities, is the dispersion of these probabilities across stocks. On the basis of a casual inspection of the column graphs we get the impression that dispersion is somewhat greater in the HSE, as nontrading probabilities are more extreme there at both extremes and a relatively larger number of stocks has extremely

[85] See the discussion towards the end of 2.3.1 above. Also conclusion 2 of CMSW above (end of 2.1.3).
[86] Refer to CMSW, conclusion 3

high and low nontrading probabilities there than in the ISE. However, this difference is not striking and, more importantly, the results of previous research do not suggest any precise interpretation of such differences. By way of conclusion we may assert that median nontrading probabilities are similar in the two markets. The most obvious difference between the two exchanges that has a bearing on the problem at hand, is that the ISE all-share index has only half as many stocks in 1996 as the HSE all-share in 1990-93. On the basis of previous research we may infer that the small number of stocks in the index will tend to obscure the predicted positive index autocorrelation effects as the time-averaging element in the index may be overshadowed by negative autocorrelation of individual stocks due to infrequent trading and the bid-ask spread.

3.2 Index autocorrelation

3.2.1 The autocorrelation function

Following Wei (1990), we can define the autocovariance of a stationary time-series $\{Z_t\}$ with fixed mean $E(Z_t) = \mu$ and variance $Var(Z_t) = E(Z_t - \mu)^2 = \sigma^2$ as

Equation 118

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E[(Z_t - \mu)(Z_{t+k} - \mu)].^{87}$$

By the stationarity assumption, this quantity will depend only on the time difference between observations, i.e. the value of k and not on the value of t. Normalising the expression **Equation 118**, we obtain the autocorrelation of the series as a function of k, i.e.

Equation 119

$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0}.$$

As we have assumed that the series has the same variance everywhere, the denominator reduces to the series variance, i.e. the covariance for k=0. The expression in **Equation 119** is consistent with the use of the concept of autocorrelation in the work we have studied so far. Also, it is commonly assumed that stock returns, i.e. the first differences of prices, yield a stationary time-series. It will be clear from the formula, that this concept of autocorrelation, just as covariance in general, can only

[87] Wei (1990), p.10. See also Harvey (1993), p.42ff.

identify *linear* relationships between observations. However, as we have seen earlier, this is exactly the type of relationship predicted by existing theories of infrequent trading effects on observed stock index returns.

3.2.2 Tests of hypotheses

It seems a logical extension of our reflections and conjectures on the subject of infrequent trading effects in the ISE in the preceding section, to examine the sample autocorrelation structure of the index return time-series. To do this we will perform a set of hypothesis tests, one for each value of k at which we think nonzero nontrading-induced autocorrelation plausible and present the results in graphical form. This cut-off value of k may be assessed by determining expected *nontrading durations* on the basis of the observed nontrading probabilities in the market.[88] After we have determined the appropriate maximum value of k, we calculate the value of $\hat{\rho}_k$, the sample autocorrelation coefficient and test the null hypothesis that it is equal to zero. To implement such a test, we need some assumptions on the distribution of the underlying returns series, as well as the sampling characteristics of the autocorrelation coefficients given that distribution. Preparing the data we have, applied a logarithmic transformation to the stock price series before differencing it once, thus obtaining a continuously compounded return series that we will assume to be normally distributed. Then the sample autocorrelation coefficient $\hat{\rho}_k$ is normally distributed for large sample sizes and we may approximate its sample standard deviation as

Equation 120

$$S_{\hat{\rho}_k} = \sqrt{\tfrac{1}{n}(1+2\hat{\rho}_1^2+\cdots+2\hat{\rho}_m^2)},$$

where we assume that k=0 for $k>m$.[89] Thus, testing the null hypothesis

Equation 121

$$H_0 : \rho_k = 0$$

against the double alternative $H_1 : \rho_k \neq 0$ at the 5% level of significance, we will reject the null in the case when $|\rho_k| > 1.96 S_{\hat{\rho}_k}$. The reason for using the two-sided alternative is that in view of the low number of stocks entering the calculations of the ISE index, it has not been convincingly established that only positive autocorrelation

[88] Following CLM (1997), these durations may be obtained by using the formula in equation **2-82.**

should be expected in the return series. Now the mean expected nontrading duration in the ISE is 4.64 days, with the highest equal to 26.7 days and the second highest equal to 16.9 days. Only seven stocks have nontrading durations in excess of the mean and together they represent a small fraction of total ISE market capitalisation. Therefore, in a weighted index, infrequent trading induced autocorrelation should be negligible at lags larger than k=5.

3.2.3 The data

A number of considerations affected the choice of the data sample. First, as the object of principal interest is the present situation in the ISE, the data set has to be as recent as possible. Second, the size of the sample has to be large enough to allow for asymptotic results to be applied.[90] The third, and final, consideration taken into account was the possibility of comparison of ISE results with similar ones for foreign stock exchanges. It is obviously of interest to obtain some comparison with the HSE, which is known in the literature as a thin market and we have already discussed in the previous section. The NYSE can also be said to be an obvious choice for comparison, as it is in many respects 'the archetypal stock market'. It is certainly the stock exchange on which the greatest research effort in empirical finance has concentrated. It is large and trade can safely be assumed to be very active in most of its between two and three thousand listed stocks. Therefore, in comparison with the ISE, it can be taken to represent the opposite end of the spectrum of market microstructure effects. The NYSE composite index is a capitalisation-weighted all-share yield index of the market calculated by a Laspeyres-type formula like the other two, and they should all be comparable in that respect. The available data for NYSE ends on the 20th of November 1996. The resulting data sample spans the period 3.1.1996 to 20.11.1996 for all three exchanges. This amounts to nearly 200 daily observations for each series. Thus we can in principle afford to select a large value of k but the final choice was 24 lags, because that covers more than a trading month and is certainly large enough to detect all nontrading induced autocorrelation in all three index numbers. Contemporaneous comparison of these three stock exchanges, however, may not be the most interesting path of analysis in this context. One reason for this is that no

---

[89] See Wei (1990), p.21

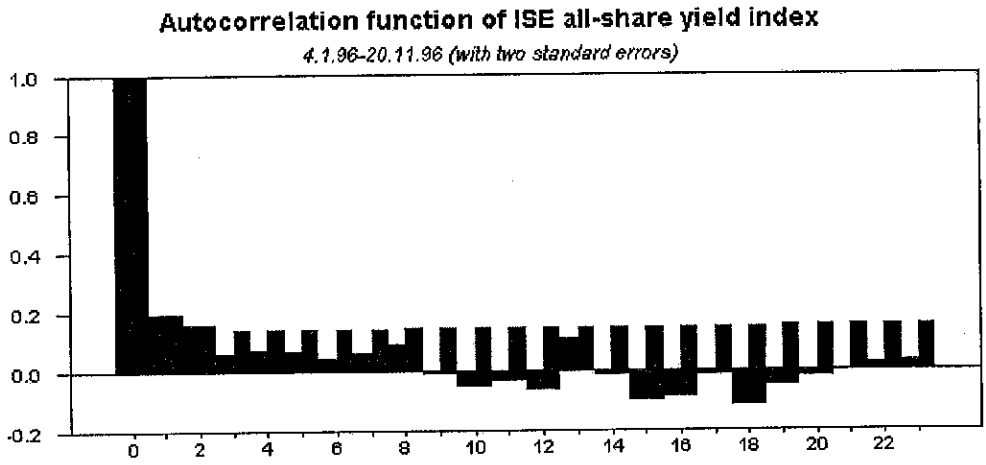
detailed information is accessible on the present state of the Finnish market in terms of capitalisation, turnover, trading frequency etc., while we have established that the Icelandic market in 1996 shares some features with HSE in the period 1990-93. As for comparison with the NYSE, it is likely that the gap between it and the ISE is too large for comparison to be very interesting. For this reason we will supplement our comparison of three exchanges at the same point in time with an additional one, comparing autocorrelation in index returns in the two exchanges in different time periods. In a sense this may be said to offer a dynamic view of the problem. We can safely assume that both stock exchanges have been growing fast in the period, by any common measure. The comparison of different periods may then provide a basis for reflection on how nontrading-symptoms in the index develop over time as total market capitalisation, number of firms and turnover velocity increase. The correlograms may thus be helpful in forming working hypotheses for further research, although it must be emphasised that the simple approach that follows cannot be expected to yield definitive results

3.2.4 Correlograms

The correlogram for daily ISE index returns in the period 3.1.1996-20.11.1996, calculated in the manner described above, turned out to exhibit significant autocorrelation at lags k=1 and k=2, meaning that the size of the sample correlation coefficients at these lags exceed twice the size of the standard error.

[90] It is customary to regard n=50 observations as a rule of thumb to determine the minimum necessary sample size. Another such rule is that the number of lags for which autocorrelations are computed should not exceed $n/4$.These issues are treated in Wei (1990) p.105

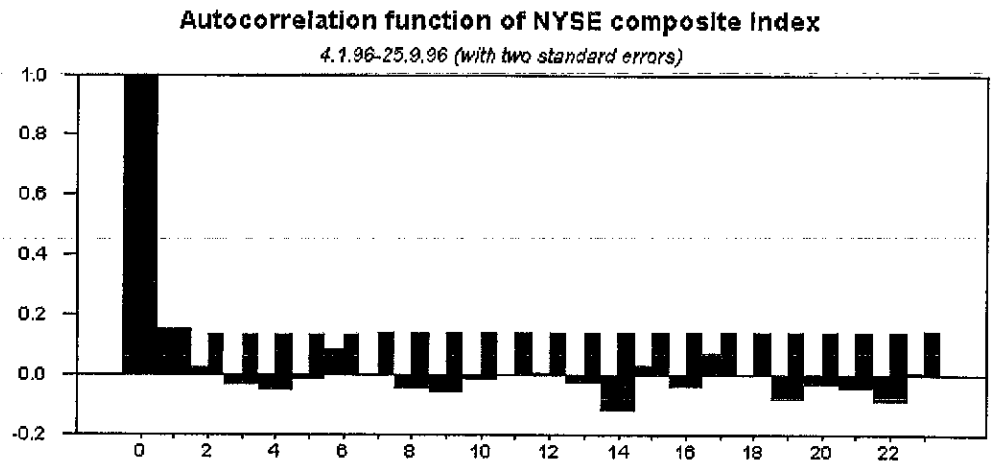
Figure 7



This result is not surprising in itself, given that it agrees with theoretical models predicting the existence and positive sign of index autocorrelation under infrequent trading. Its magnitude, however, is lower than one would expect, recalling that the *Russell 2000* index studied in Jukivuolle (1995) had a first order autocorrelation coefficient of 0.27, for here it is only about 0.2. It is not likely that nontrading is more severe among even the 2000 smallest listed firms in the USA than in the ISE. We have seen, however, that we can expect a complex pattern of interplay between different microstructure effects in the index return series. The high returns on the ISE all-share index in 1996 and its far lower number of stocks may be part of the explanation.

Obtaining a correlogram for the NYSE Composite Index for the same time-span we observe that first-order autocorrelation is significant there too.

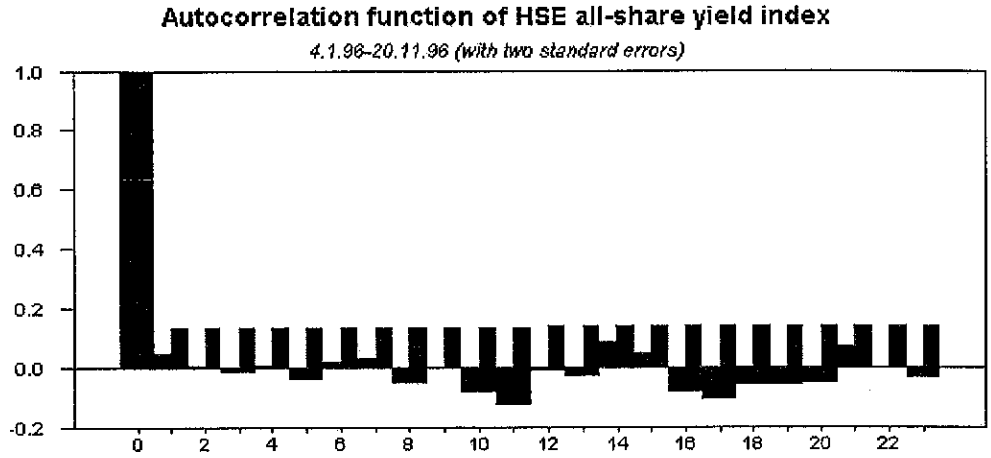
Figure 8



Although this can hardly be because of nontrading, recalling the experiment in Perry (1985) it need not be so surprising either. Theoretically, however, the phenomenon identified by Perry, i.e. positive return autocorrelation in portfolios of frequently traded stocks, has not been well explained. A possible line of reasoning is the one suggested by CMSW (1986), that price-adjustment delays may be caused not only by sparse transaction based price-updates, but also by sporadic quotation-updates. Economically, this hypothesis makes good sense in the presence of information costs and it may even be consistent with the view that markets are fundamentally efficient, given the right assumptions about market structure and investment behaviour. In principle, it seems that the quotation-lag theory could be elaborated much in the same way as nontrading, and empirical investigation carried out on the basis of quotation price data. Relative to the research discussed here however, quotation-lag induced index-return autocorrelation appears as a residual category when the phenomenon can not be attributed to infrequent trading. This is a somewhat unsatisfactory state of theory and if it is in fact the actual state of the art, it is paradoxical that more should be known of the properties of index portfolio returns in small inactive markets than in the worlds most important ones.

Repeating the same experiment for the third time, now for the case of HSE index returns, we obtain another puzzling result.

Figure 9

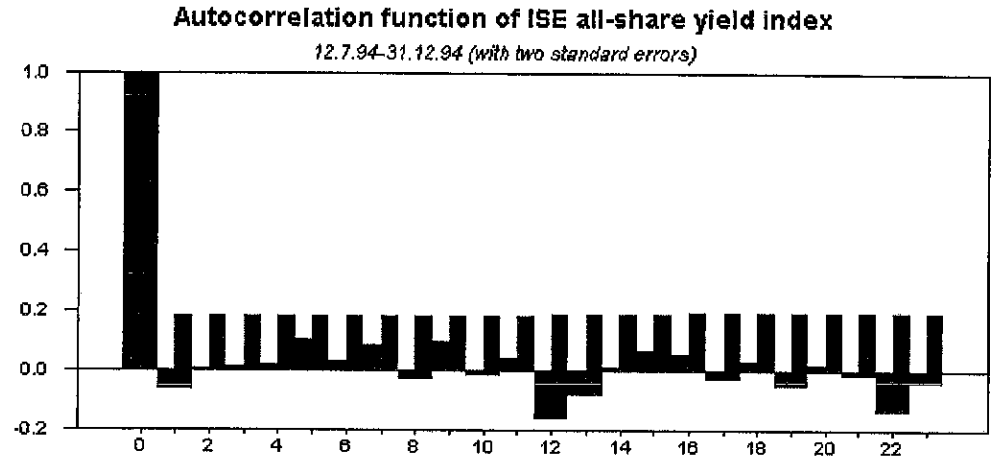


It is reasonable to expect all relevant market size and activity parameters in the HSE to lie somewhere between those in the ISE and the NYSE in 1996.Therefore it is tempting to conclude that the same will apply to autocorrelation, i.e. that it will be positive, but less than in the ISE. Recalling that factors independent of market size, such as the coefficients of variation of individual stocks also play an important role we see that an inference based on size alone may be overhasty. This is demonstrated by the HSE correlogram, that exhibits insignificant autocorrelation at all lags. The relation of the three correlograms to one another is inconsistent with a simple view of infrequent trading symptoms in the index. This result in no way undermines the validity of the theories presented, it should serve to underline the concluding remarks of the last subsection, concerning the need to systematically analyse the interplay of microstructure effects. In particular, it is important not to conclude from the apparent white noise character of the HSE index returns that this index reflects the true returns process more faithfully than the others, or even that it is the market in which trading is most efficient. The 'white noise' HSE index returns will certainly be an aggregate of unobserved microstructure generated component processes which have a dynamic structure, and if the series of true price innovations to the market is white noise, this implies that the two will be imperfectly correlated.

We now add a historical dimension. It has already been indicated that the ISE has grown quite fast since the calculation of the market index commenced in 1994. Thus it may perhaps be justified to expect some sort of identifiable pattern of development if correlograms are inspected year by year. The same could quite possibly hold for the HSE. We have suggested that the present day ISE has some microstructure features in

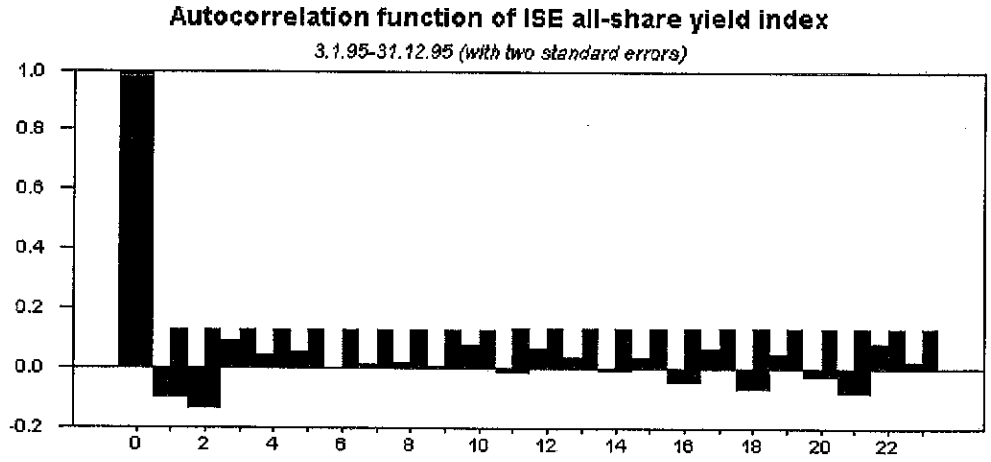
common with the HSE in 1990-93, notably high nontrading probabilities and both exchanges trade by computerised continuous auction in the periods in question.[91] Starting with the ISE in 1994, we have access to index data for the last five and a half months of the year. An inspection of the correlogram reveals the interesting fact that autocorrelation is insignificant at all lags

Figure 10



Specifically, the autocorrelation coefficient is negative at the first lag.

Figure 11



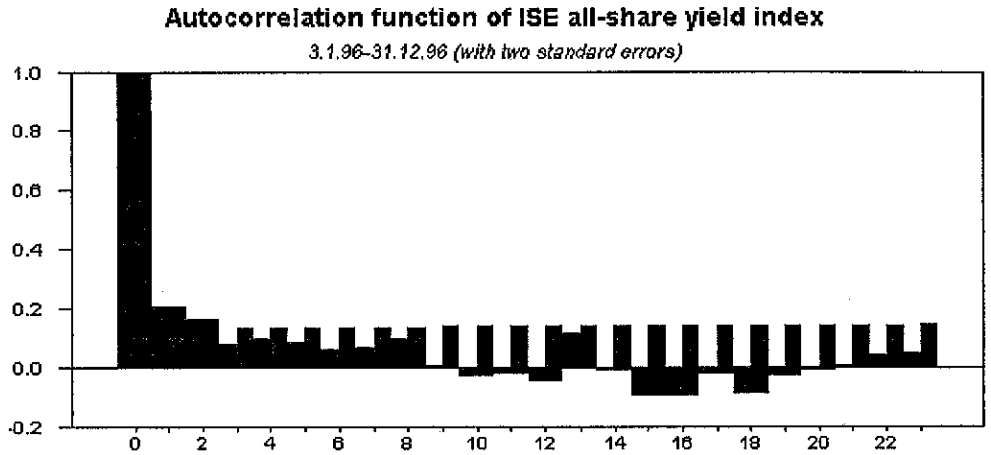
Although it is quite small and allows of no rigorous interpretation as such, the sign of this autocorrelation coefficient suggests that negative autocorrelation effects in

[91] Until 1990 the HSE was a 'call market'. Kallunki and Martikainen (1997) p.3

individual returns due to nontrading and the bid-ask spread may still affect the index because the number of stocks is very low.

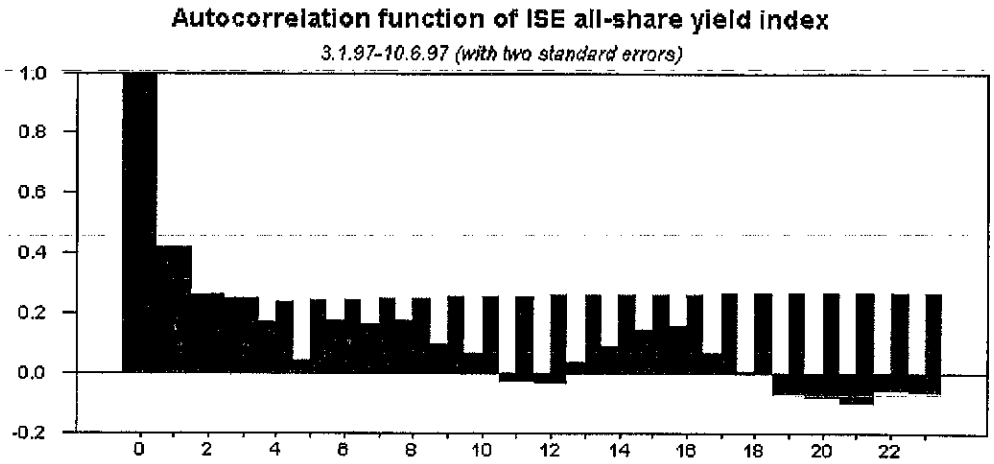
The correlogram for 1995, the first year where we have a full set of data, seems to be consistent with this idea. Here autocorrelation coefficients are negative at the first two lags, and significantly so at the second. In 1996, drawing the correlogram for the whole year, we get practically the same graph as in **Figure 9**, only with a little lower critical values as we have added some 20 observations.

Figure 12



As before, we now find significant *positive* autocorrelation at the first two lags, with $\hat{\rho}_1 \cong 0.2$. In 1997 we have data corresponding to a little under half a year. Inspection of the correlogram reveals significant positive autocorrelation at the first three lags, with $\hat{\rho}_1 \cong 0.4$ and the coefficient for k=2 and 3 both larger than 0.2. In summary, autocorrelation at the lowest lags goes from being negative in 1994 and 1995 to positive in 1996 and 1997 with the size of the first order coefficient doubling between the latter years. One possible reason for this pattern is that the time-averaging effect due to nontrading increasingly outweighs negative individual stock effects as the number of stocks in the index increases.

Figure 13



Although this explanation is consistent with existing microstructure models we are not in a position to quantify the opposing effects and their interaction may well be quite complex.[92] Therefore, as far as the present inquiry is concerned, this idea remains unsubstantiated.

Looking at correlograms for the HSE all-share yield index for the years 1991, 1992 and most of 1993 we can get a rough idea of the index return series that corresponds to the nontrading probabilities discussed in the previous section. We recall that these figures were averages over a four year period, 1990-93 and that there were 54 stocks in the index throughout the whole period. Our assumption of a roughly monotonic growth in size and liquidity of the stock exchanges in question implies that the number of stocks in the HSE all-share will have increased over this period and trading will also have increased in volume and frequency. However, the pattern of growth seems to have been somewhat different in the HSE than in the ISE. In particular the rate of new listings is less dramatic in the HSE. From the above and **Table 2** we may infer that between 1990 and 1992 only seven new firms were listed, making the total number of firms in the index 61 at the end of 1992. This is a 13% increase in three years while in the ISE 15 new firms were added to the 27 listed in January 1996 during approximately 20 months until August 1997, which is a 55% increase.

[92] CLM (1997), p.89-99, take some steps towards disentangling these relationships, as well as the relationship between stock nontrading probabilities and index autocorrelation.Their results are difficult to apply in practice.

Figure 14

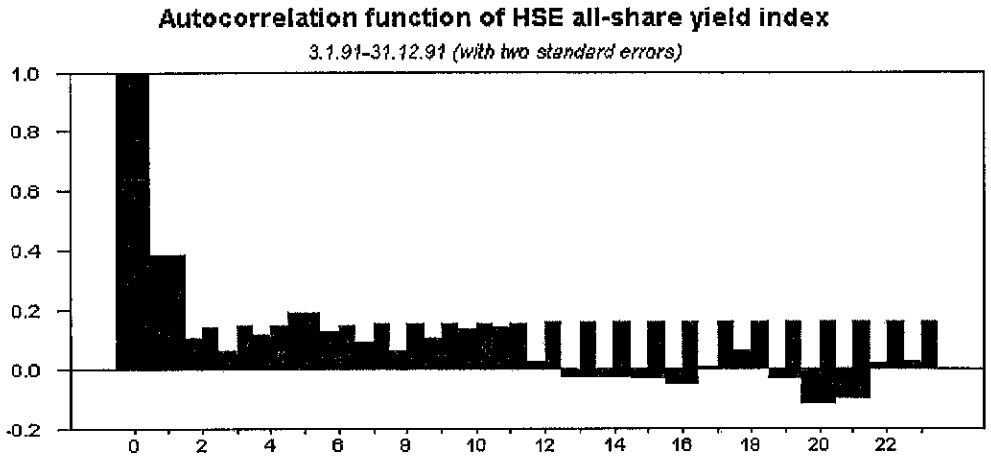
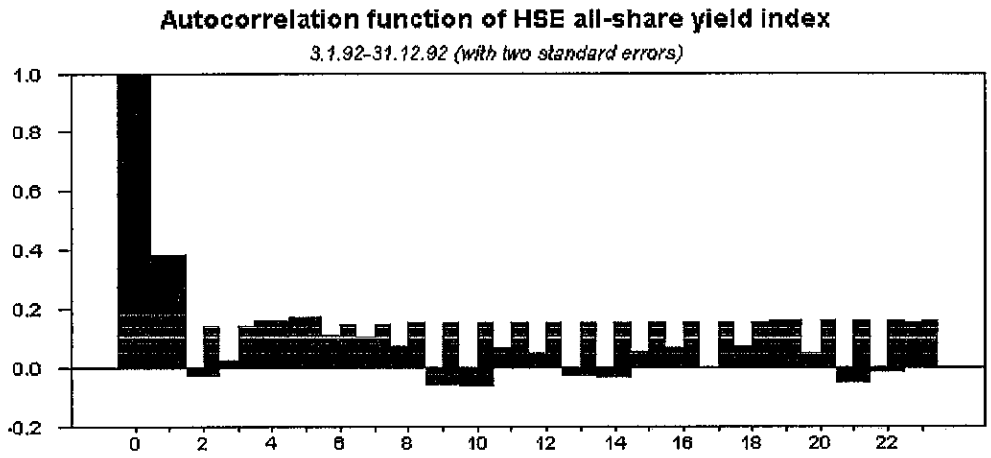


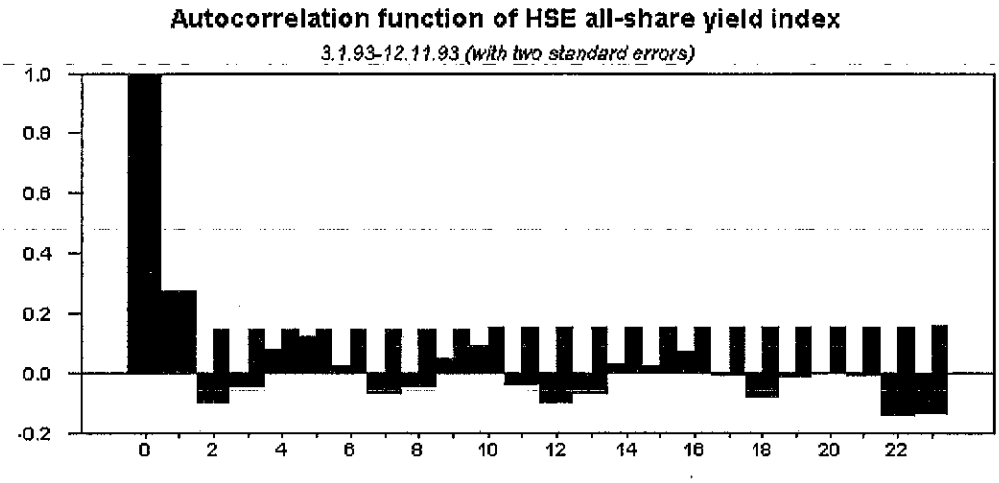
Figure 15



This indicates that the HSE, despite its persistent nontrading problems, may already have passed its 'fledgling phase' in the early nineties and achieved a certain degree of stability. The correlograms do not contradict this idea.

The changes in the autocorrelation structure between years are much less dramatic than in the ISE in 1994-97. First order autocorrelation is everywhere significant. In the first two years, 1991 and 1992, it is close to 0.38, which is greater than in the ISE in 1996, but a little less than ISE in 1997.

Figure 16



In 1993 it is 0.27, a decrease that may signify a development leading to the surprisingly low coefficient observed in the HSE 1996. An interesting common feature in all three HSE correlograms is a large positive autocorrelation coefficient at the fifth lag, which tests significant in the first two years. This phenomenon is probably indicative of a strong 'day of the week effect' as there are five days in the business week.

This section aimed to provide a partial description of some characteristics of the ISE that may be considered relevant to nontrading, and a discussion of observed index autocorrelation in the ISE and two other stock exchanges. Although a discussion like the preceding one can be a useful preliminary to more rigorous analysis, it can never replace it. It may be, however, that something is achieved if we have managed to outline a problem that subsequently can be tackled in a systematic way. It would be interesting to perform a somewhat more elaborate time series analysis of the ISE index returns series. This would require an elaborate theoretical model, making it possible to quantify the expected effects of microstructure phenomena such as infrequent trading and the bid-ask spread given the value of market parameters such as the number and size of firms, size of the effective spread, nontrading probabilities etc. Many of the necessary components have already been provided in the contributions covered in part 2 above. One way to approach this task would be to employ a simulation approach. It would also have to be decided whether modelling the index in terms of *unobserved components* would be useful and whether spectral

analysis could yield otherwise unavailable insights.[93] Such an exercise, however, is beyond the scope of this report.

3.3 Kalman-filter estimation of the stock-index.

The logical way to view microstructure effects in the context of stock index estimation is as various types of measurement error. In fact this idea is implicit in all the approaches we have reviewed so far, as all regard microstructure effects as *spurious* with respect to the true value-generating process of stocks. Another common feature is that they all assume the true value of stocks to be a continuous stochastic process of such a character that time-increments of its logarithm can be interpreted as normally distributed continuously compounded returns to investment. Furthermore, all assume that markets are fundamentally efficient, which in statistical terms can be interpreted as the assumption that in a discrete framework true returns are a white noise process. In dealing directly with spurious elements in the index series, as opposed to e.g. targeting consistent estimators of CAPM beta, two distinctly different approaches can be discerned. One of them, which we have called the *ARMA* approach, works backwards, so to speak, by purging the index returns obtained by conventional techniques of all (presumably) spurious effects *post factum*. The other approach, which we can call *transaction-based* for lack of a better term, directly tackles the problem of efficient index estimation from price data in the presence of microstructure induced measurement error. Examples of the former include the Stoll-Whaley (1990) and the Jukivuolle (1995) models, while the latter is represented here by the work of Harris (1989) and Bassett, France and Pliska (1991).

In this section we wish to suggest a way to estimate the stock index that is consistent with the assumptions of earlier research. Hopefully, though, it may lead to improvements in some respects. It is transaction-based, and in fact it can be considered an extension of the approach in Bassett, France and Pliska (1991), because it is based on simultaneous time series equation estimation of the index under infrequent trading by Kalman-filter techniques. It differs, however, from the method employed in that paper, in two important respects. First, here the Kalman-filter is implemented in a continuous-time framework. Secondly, the present estimator takes

---

[93] See Harvey: *Time Series Models* (1993), p. 30-32 for a few words on unobserved components and

advantage of a situation particular to computerised continuous time auction markets to obtain frequent on-line estimates of the measurement error associated with reported prices. Why these two additional features should be considered an improvement is explained in the next subsection. Subsequently the Kalman-filter, which was introduced in *2.2.4.2* above will be discussed in some detail with respect to the present implementation. A third subsection presents results of the application of this method to the estimation of the ISE *Index of Transportation Firms*.

3.3.1 Continuous markets and measurement error.

As we have seen, a continuity assumption concerning the underlying stochastic stock-value process of which price is the measure, is common to all infrequent trading models discussed above. This view is not only attractive from a theoretical point of view, but also entirely consistent with the structure of continuous auction securities markets in practice. In addition, strong arguments can be presented to support econometric modelling of other economic variables in continuous time.[94] In this perspective, the practice of imposing an arbitrary period grid on market time, which in reality is practically a continuum, appears inconsistent. This practice is probably in part a matter of convenience, as most popular time series methods were developed with a view to data organised as equally spaced observations, and may not be very flexible in this respect. If this is the case, the Kalman-filter methodology is something of an exception. Although it is admittedly somewhat more complicated to carry out in continuous time than discrete, it is still entirely feasible. But there are other strong arguments in support of a continuous time approach in the present case. For one, it makes better use of the available transaction data, which is always advantageous and may be crucial when trading is sparse. Even when the chosen period grid is dense, in a continuous market distinct transactions will often occur sufficiently close together to be reported in the same period. In that case, all but one are discarded if a period structure has been adopted. Furthermore, as has been sufficiently emphasised in this report, an error will be introduced when observations that occur early in a period are interpreted as end-of-period data. In general, when a period structure is assumed, observations reported to occur in the same period can really be spaced further apart in time than ones that are reported in adjacent periods. In a continuous-time framework,

---

the structural approach to the modelling of time series.

by contrast, infrequent trading problems in the narrow sense of this term will simply vanish, as all observations are reported exactly at the time they are observed.

It may be argued that problems like the above can be minimised by selecting a fine enough period grid. This is in fact the strategy employed in BF&P (1991), where a one-minute grid is imposed to estimate the MMI over three days. Such tactics, however, are quickly reduced to absurdity when the aim is to estimate an index over more extended periods. Proceeding in this way for the ISE data set estimated here by a continuous Kalman-filter, the time span of some 200 trading days between May 1996 and February 1997 would yield 72.000 periods, with 99% of the observations missing, as the number of transactions for each series is around 600. Using the finest possible time grid of one second would yield over 4 million periods, with practically all observations missing. In a continuous implementation, on the other hand, there is hardly any practical difference between measuring time in minutes and seconds, and the choice of units only involves a decision about the desired degree of precision. Which may be different for each particular estimation problem.

One basic feature of the Kalman-filter estimation technique, is that it is based on a separate conceptualisation of the unobservable 'signal' and 'noise' components of the data. It is such a perspective that yields the state-space form of an underlying statistical model, which was briefly sketched earlier (*2.2.4.2*). This means that separate assumptions about the degree of measurement error and the degree of variability of the true state of the process (here: the unobservable true value of stocks), must be introduced as basic components of the model. Such *a priori* assumptions, taking the form of a choice of values for the true variance of the measurement and state equations respectively, will have a decisive impact on the resulting estimates. To make this clearer, we can look at extreme choices. A hypothetical statistician that believes his data to be measured without error, would consequently set the prior variance in the measurement equation equal to zero, and the corresponding quantity in the state equation positive. The resulting estimates of the state would simply, and logically, be equal to the measurements at each point. In the opposite case, if the statistician is convinced that the state is fixed throughout the sample period, she will set the prior variance in the state equation to zero and the relevant parameter of the measurement equation positive. In this case, e.g. if the

---

[94] See Bergstrom (1984)

underlying statistical model implies a constant level, the resulting estimate will be the horizontal straight line that minimises the sum of squared errors, given that the filter converges to a steady state. In other words, it will be exactly the same as the OLS estimate of the same model. This comparison demonstrates the crucial importance of correctly deciding the values of these two a priori parameters. However, no rigorous and general methodology to tackle this question exists within the framework of classical statistics.[95] This means that the applicability of Kalman-filter techniques is greatly enhanced when there are strong grounds for setting the variance parameters in a specific way.

Continuous-time estimation of the true value of stocks in continuous auction markets can be assumed to eliminate the particular type of measurement error that we have primarily focused on in this report, i.e. resulting from the non synchronicity of trades. To partly compensate for the scarcity of information implied by a low trading frequency, we propose to use contamporaneous correlation in the market to update estimates of stock value more frequently than the particular stock actually trades. The method is similar to the approach in BF&P.[96] Thus the continuous-time approach should take care of the measurement error due to infrequent trading effects in the narrow sense, i.e. insofar as it is due to false timing of transactions in the reported price data with respect to their true time of occurrence, and the simultaneous equation approach is aimed at aleviating the problem of data scarcity as such.

A source of measurement error emphasised in the market microstructure literature and briefly introduced in this report, is bid-ask bouncing. We have suggested that this effect may be greater in the ISE than elsewhere because of small size and thinness, precluding economies of scale. Furthermore, its effect on the index number series may be considerable, because there is a small number of stocks in the index portfolio. It is not necessarily a trivial task to estimate the size of the bid-ask spread, even when quotation data are available, and a number of different approaches have been suggested.[97] Hardly any research has been undertaken concerning the size of the

---

[95] Following a Bayesian line of argument introduced by Akaike, some authors advocate an approach that involves *estimating* these parameters on the basis of the data. For an exposition and bibliography, see Harvey (1989). We may add that this problem is of course not specific to the Kalman filter approach, as all statistical techniques depend on correct assumptions about the underlying true value and error processes.

[96] See Harvey (1989), Chapter 8 for detailed discussion of Kalman-filter application to multivariate time-series models.

[97] Two are discussed in CLM (1997), p.99-107 and 134-135.

spread for ISE stocks.and it is beyond the scope of this report to attempt an estimation of the ISE spread.[98] Consequently no attempt is made to account for bid-ask induced error in the present model.

Examination of ISE transaction data reveals an interesting feature that it may have in common with data originating in other computerised continuous auction markets. This is that from time to time more than one transaction takes place in the same stock at the same moment, but not necessarily at the same prices. As "a moment" in this context means *one second*, this may seem quite puzzling at first sight. The reason is that when a dealer has a large order, she is often unable to find a single quote on the opposite side that exactly matches it. When this happens, she will use her computer terminal to select a sufficiently large number of smaller offers to fill the order, finally executing all the individual trades at the same time. When these price-quotes differ among themselves, there is good reason to interpret this set of data as informative concerning that dealer's subjective estimate of the instantaneous error of measurement. This is because the different prices can all be assumed to lie within an 'acceptable distance' from the agents estimate of the true value the corresponding stock. This phenomenon can therefore be used to obtain improved estimates of the true measurement equation variance for the stock in question, whenever such an event occurs.

Now measurement error in the market for stocks is essentially a subjective phenomenon, in exactly the same sense as prices of stocks are a subjective measure of their true value. The fact that two or more subjective estimates (i.e.: the buy- and sell-quotes) have to coincide for a price observation to occur does not change this. Thus, when we use transaction prices as the best available measure of the value of stocks, we are relying on the personal opinion of a agents. Instantaneous measurement variance as defined here, then provides us with exactly analogous information about the uncertainty associated with the level of the random variable in question, as the price observation does about its mean. As we assume that the true state cannot change in a time interval of length zero, the resulting variance estimate can be interpreted as measurement equation error variance in the state-space form.

It must be empasised, however, that this estimate can hardly account for the effects of bid-ask bouncing, and may therefore underestimate the true error of measurement at any given time if a significant bid-ask spread exists in the market for the stock in

---

[98] Some steps are taken to analyse the composition of the spread in Sigurgeir Örn Jónsson (1997).

question. This is because each cluster of simultaneous trades contains either only bid or only ask prices, depending on whether the large order is a buy or a sell. If it is known in each case which of the two is taking place, as can in principle be inferred by inspection of the corresponding quotations, it may not be too difficult to take the bid ask spread into account in determining the uncertainty of price measurement. This task must be relegated to an ulterior study and in the present one we will act as if the measurement error were not affected by a bid-ask spread. We note that the mere fact of assuming some degree of error of measurement implies a smoother index, counteracting spurious bouncing between the bid and the ask to a certain extent, although this problem is not attacked directly.

In the implementation described here we experiment with incorporating information obtained on the basis of multiple trades into the estimation process in two alternative ways. One is to replace the a priori measurement equation variance assumption with a new one each time more than two simultaneous observations occur in the same stock. This has the advantage of allowing the measurement variance in the model to change with time, allowing for greater realism. Its disadvantage is that it is sensitive to outliers in the data set of multiple trades. Another approach to the present implemantation problem is to keep this assumption fixed over the sample period, at a value which results from averaging all instantaneous variance estimates, weighted by the number of observations entering each. Which one is appropriate, or whether some combination of the two should be adopted must count as one of the issues to be treated in a sequel to this work.

The sample variance, s_x^2, is obtained by the standard formula.[99] This estimator may be sensitive to departures from normality in the underlying sample and possibly other methods of obtaining estimates of the dispersion should be considered at a later stage .

It is interesting to compare the present approach to the problem of setting the measuremant equation variance with the way this problem is dealt with in the earlier attempt at index estimation by Kalman-filter techniques in Basset, France and Pliska (1991).There, , the authors identify two sources of measurement error, one due to the bid ask spread, and the other to the discreteness of prices, which in U.S. stock exchanges are measured in 'ticks' equal to 1/8$. No such restrictions apply in the Icelandic market and prices are reported in units of 1/100 of IKR, so it is likely that

the latter issue can safely be ignored in the estimation of ISE index numbers. In the paper in question, BPF do not present a separate analysis of each component. Instead they fix the measurement equation variance in the sample period at a specific value, 0.005, implying a standard deviation of 0.07, and remark that this is "consistent with the bid-ask spread and the $1/8^{th}$ tick size".[100] As the data in this study are price logarithms, the standard deviation is a measure of percentage error. This means that it should be expected to be larger for smaller, lower priced and less frequently traded stocks as their spread will be larger and the tick size is larger relative to the price of the stock. It is possible, though, that percentage error can safely be treated as constant across firms in BF&P (1991) on the grounds that the MMI index portfolio represents a relatively homogenous sample of firms. In that paper there is no mention of any source of measurement error variance information similar to that discussed above. This may indicate that either this phenomenon does not occur in the Chicago Mercantile Exchange, or that this information was discarded in transforming the continuous-time data set into one-minute end-of-period observations.

3.3.2 The suggested model and its implementation

The general state-space form appropriate to the application of the Kalman-filter is the following[101]

Equation 122

$$\mathbf{y}_t = \mathbf{Z}_t \alpha_t + \varepsilon_t$$

Equation 123

$$\alpha_t = \mathbf{T}_t \alpha_{t-1} + \eta_t$$

Here $\mathbf{y}_t$ in **Equation 122** represents a vector of stock prices at time t and α_t is the vector of unobservable true values of the stocks at time t. The matrix $\mathbf{Z}_t$ can be thought of as a selection matrix specific to time t, much like in the BFP implementation, which is needed because a full vector of prices is not observed simultaneously in general. Last but not least, ε_t is the error of measurement the

---

[99] Discussed in Newbold (1995), e.g. p.243ff
[100] p. 141
[101] The subsequent discussion is based on Harvey (1989), mainly chapters 3,8 and 9 unless otherwise indicated. The notation closely follows that of Harvey.

likely composition of which is discussed in the previous subsection. **Equation 123** is the state or transition equation, where the matrix $\mathbf{T}_t$, along with the stochastic innovation term η_t determines the way the true state α_t changes with time. Thus if α_t represents a vector of simple value levels, and $\mathbf{T}_t$ is fixed and equal to the identity matrix, the implied model is one where each stock's value follows a random walk. The two error terms are presumed to have zero mean, and be uncorrelated over time and with one another. This means that the following equations hold:

Equation 124

$$E(\varepsilon_t) = E(\eta_t) = E(\varepsilon_t \eta_s) = 0 .$$

Furthermore, the following quantities have to be defined:

Equation 125

$$Var(\varepsilon_t) = \mathbf{H}_t, Var(\eta_t) = \mathbf{Q}_t$$

We can assume the meaning of the time-subscripts on **y**, **Z** and the error terms to be evident. The reason that α, **T**, **H** and **Q** are also time subscripted is that in a general formulation of the state-space model in the context of Kalman-filter estimation, provision is made for these quantities to change with time. This is obviously true of the value of stocks. In the present implementation **T** will be time-varying if a nonzero slope or drift-term is included in the continuous-time model, making the value of each forecast depend on a varying forecast horizon. Otherwise it is the identity matrix. **Q** will be time dependent because in a continuous model, the variability of the state is a function of time. **H**, on the other hand, has to be time-varying to accommodate updates to estimates of the measurement equation variance as described in the previous subsection.

Before an unobservable stochastic process can be estimated, assumptions have to me made concerning its nature. In other words, a statistical model has to be formulated on the basis of an idea about the underlying data generating process. In the context of the Kalman-filter technique, this problem takes the form of assumptions about the true state, i.e. the vector α in the state-space representation. It is convenient to state these assumptions in the form of a structural time series model. [102] Roughly speaking, this means that each observation is thought of as generated by a number of different state

[102] See Harvey (1989), chapters 1 and 2 for a detailed discussion of univariate structural time-series models.

components, each representing a structural element of the aggregate series. In the notation introduced above, a univariate additive structural model may be stated as

Equation 126

$$y_t = \mathbf{z}'\alpha_t + \varepsilon_t = \alpha_{1t} + \alpha_{2t} + \cdots + \alpha_{kt} + \varepsilon_t$$

where **z** is a 1 by k column vector of ones.

The components are selected to represent the composite structure of the observation series in the best possible way, but usually the advantages of the increase in flexibility achieved by a introducing a large number of structural components will have to be weighed against the loss in efficiency resulting from estimation of many parameters. To make this idea clearer, we may state a model of this type in words, as

Equation 127

OBSERVED SERIES = LEVEL+ LOCAL TREND + SEASONAL + CYCLE + DAILY EFFECT + IRREGULAR

which may then be reduced to

Equation 128

OBSERVED SERIES = LEVEL+IRREGULAR

on the basis of parsimony considerations, if this is thought to be sufficiently flexible to provide a good description of the dynamics in the data. In **Equation 128** the state vector is one-dimensional. Assuming that the true state follows a random walk, which is equivalent to particular restrictions imposed on the state equation that describes transition in time, leads to the model of each stock's price used by BF&P. Expressed in the present notation, it is contained in the following two equations

Equation 129

$$y_t = \alpha_t + \varepsilon_t$$

Equation 130

$$\alpha_t = \alpha_{t-1} + \eta_t .$$

In the present study, a two-dimensional state, corresponding to a random walk with nonzero drift was assumed to describe the evolution of the true value of stocks best. The line of reasoning was that stocks are only one of many instruments available to investors. The drift term in a logarithmic random walk model can be interpreted as continuously compounded mean return on investment in stocks and it was thought that this parameter would be subject to economic restrictions conflicting with the assumption that it were identically zero. Such restrictions can then be incorporated

into the assumptions of the model. In retrospect it may be said that a simpler model, more similar to the one in BF&P would have been quite adequate, especially in the context of a simultaneous system, and it would probably have somewhat reduced the required computer programming effort. However, the simple random walk model is a special case of the random walk plus drift model, and once the two-dimensional model is implemented, the two models can easily be compared by fixing the drift term prior variance at zero and setting the appropriate initial values. Thus the continuous model that was assumed is the one that corresponds to the following discrete-time state-space form:

Equation 131

$$y_t = \mathbf{z}'\alpha_t + \varepsilon_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} + \varepsilon_t$$

Equation 132

$$\alpha_t = \mathbf{T}\alpha_{t-1} + \nu_t = \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \zeta_t \end{bmatrix}.$$

Here μ_t is the *local level*, or mean, of the process at time t, β_t the drift parameter, termed *local linear trend* in the vocabulary of structural time-series analysis, and η_t and ζ_t are the corresponding error processes that form the state error vector, ν_t, in each time period. We note that the observed system is univariate, even though the structural model of the underlying state has two components.

The next step is to state this basic model in a continuous-time framework.[103] A general form of a continuous-time state space model of a single time series takes the form

Equation 133

$$\frac{d}{dt}[\alpha(t)] = \mathbf{A}\alpha(t) + \nu(t)$$

with the variance of the vector process $\nu(t)$ of state innovations given by

Equation 134

$$Var[\nu(t)] = E\left[\int_r^s \nu(t)dt\right]\left[\int_r^s \nu(t)dt\right]' = (s-r)\mathbf{Q}$$

where r, s are points in time and $\mathbf{Q}$ covariance matrix of the innovation process, yet to be specified. Here we assume $\nu(t)$ to be a Wiener process with uncorrelated increments. If discrete sampling of the continuous process is irregularly spaced, as in the case we are interested in, we change the notation a little, using τ to indicate the time point, define $s-r=\delta_\tau$ and write the state space representation as

Equation 135

$$y_\tau = \mathbf{z}'\alpha_\tau + \varepsilon_\tau$$

Equation 136

$$\alpha(t_\tau) = e^{\mathrm{A}\delta_\tau}\alpha(t_{\tau-1}) + \int_0^{\delta_\tau} e^{\mathrm{A}(\delta_\tau - s)}\nu(t_{\tau-1}+s)ds\,.$$

We note that in its form **Equation 136** parallels the discrete-time transition equation

Equation 137

$$\alpha_\tau = \mathbf{T}_\tau\alpha_{\tau-1} + \nu_\tau, \quad \tau = 1,2,...,T, \quad \mathbf{T}_\tau = \exp(\mathrm{A}\delta_\tau)\,.$$

We now obtain the covariance matrix $\mathbf{Q}$ at time t_τ as

Equation 138

$$\mathbf{Q}_\tau = \int_0^{\delta_\tau} e^{\mathrm{A}(\delta_\tau - s)}\mathbf{Q}\, e^{\mathrm{A}'(\delta_\tau - s)}ds\,.$$

Roughly speaking, this apparatus is sufficient to set up a functional continuous model of the stochastic state process of a single stock's value. It is a special case of **Equation 133**, with

Equation 139

$$\frac{d}{dt}\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix} + \begin{bmatrix}\eta(t)\\ \zeta(t)\end{bmatrix}.$$

Solving this differential equation and evaluating the resulting matrix exponential yields the state equation

Equation 140

$$\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix} = \begin{bmatrix}1 & \delta_\tau\\ 0 & 1\end{bmatrix}\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix} + \begin{bmatrix}\eta(t)\\ \zeta(t)\end{bmatrix}$$

and direct evaluation of the matrix exponential integral in **Equation 138** gives

---

[103] The exposition that follows draws extensively on chapter 9 in Harvey (1989)

Equation 141

$$Var\begin{bmatrix}\eta_\tau \\ \varsigma_\tau\end{bmatrix} = \mathbf{Q}_\tau = \delta_\tau \begin{bmatrix}\sigma_\eta^2 + \frac{1}{3}\delta_\tau^2\sigma_\varsigma^2 & \frac{1}{2}\delta_\tau\sigma_\varsigma^2 \\ \frac{1}{2}\delta_\tau\sigma_\varsigma^2 & \sigma_\varsigma^2\end{bmatrix}.$$ [104]

This shows that in this model the discrete-sample innovations to the state will be correlated even though the corresponding continuous-time innovation processes themselves are not. However, according to Harvey (1989), "this difference is unlikely to be of any great practical importance."[105] As a tentative step, the results in **Equation 140** and **Equation 141** were used to experiment with Kalman-filter estimation of the true value of shares in some ISE companies for each stock separately in a univariate framework. To proceed to a multivariate system, however, some further theoretical results are needed.

The multivariate state-space representation we will assume can be stated in the following way[106]

Equation 142

$$\mathbf{y}_\tau = (\mathbf{z}' \otimes \mathbf{I}_N)\alpha_\tau + \varepsilon_\tau$$

Equation 143

$$\alpha_\tau = (\mathbf{T}_\tau \otimes \mathbf{I}_N)\alpha_{\tau-1} + \nu_\tau$$

where N is the dimension of the vector of observations $\mathbf{y}_\tau$. If the state is two-dimensional as above, then $\mathbf{z}' = [1,0]$, ε_τ an N-vector, and α_τ and ν_τ $2xN$ matrices. Taking N=2 and using the evaluation of $\mathbf{T}_\tau$ obtained in the univariate case, i.e.

Equation 144

$$\mathbf{T}_\tau = e^{\mathbf{A}\delta_\tau} = \begin{bmatrix}1 & \delta_\tau \\ 0 & 1\end{bmatrix},$$

we see that the form of the transition matrix in the simultaneous system for two stock price series will be

[104] These calculations are performed in Harvey (1989), p.487
[105] P. 487
[106] This formulation is a slight variation of that in Harvey (1989), p.432 for a discrete-time multivariate system.

Equation 145

$$(e^{\mathbf{A}\delta_\tau} \otimes \mathbf{I}_2) = \begin{bmatrix} 1 & 0 & \delta_\tau & 0 \\ 0 & 1 & 0 & \delta_\tau \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

As the basic idea of this exercise is to use simultaneous covariance in the system to obtain updates of the nontrading stock whenever one trades, we need a theoretical expression for the system covariance matrix of the state. This is

Equation 146

$$\mathbf{Q}_\tau = \int_0^{\delta_\tau} (e^{(\mathbf{A}\delta_\tau - s)} \otimes \mathbf{I}_2) Var(\nu) (e^{\mathbf{A}'(\delta_\tau - s)} \otimes \mathbf{I}_2) ds$$

which can be evaluated in the same way as for the univariate case, yielding

Equation 147

$$\mathbf{Q}_\tau = \int_0^{\delta_\tau} \begin{bmatrix} 1 & 0 & s & 0 \\ 0 & 1 & 0 & s \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_{\eta 1}^2 & \sigma_{\eta 12} & 0 & 0 \\ \sigma_{\eta 21} & \sigma_{\eta 2}^2 & 0 & 0 \\ 0 & 0 & \sigma_{\zeta 1}^2 & \sigma_{\zeta 12} \\ 0 & 0 & \sigma_{\zeta 21} & \sigma_{\zeta 2}^2 \end{bmatrix} \begin{bmatrix} 1 & 0 & s & 0 \\ 0 & 1 & 0 & s \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}' ds$$

$$= \begin{bmatrix} \sigma_{\eta 1}^2 \delta_\tau + \frac{1}{3}\sigma_{\zeta 1}^2 \delta_\tau^3 & \sigma_{\eta 12}\delta_\tau + \frac{1}{3}\sigma_{\zeta 12}\delta_\tau^3 & \frac{1}{2}\sigma_{\zeta 1}^2 \delta_\tau^{\ 2} & \frac{1}{2}\sigma_{\zeta 12}\delta_\tau^{\ 2} \\ \sigma_{\eta 12}\delta_\tau + \frac{1}{3}\sigma_{\zeta 12}\delta_\tau^3 & \sigma_{\eta 2}^2 \delta_\tau + \frac{1}{3}\sigma_{\zeta 2}^2 \delta_\tau^3 & \frac{1}{2}\sigma_{\zeta 12}\delta_\tau^{\ 2} & \frac{1}{2}\sigma_{\zeta 2}^2 \delta_\tau^{\ 2} \\ \frac{1}{2}\sigma_{\zeta 1}^2 \delta_\tau^{\ 2} & \frac{1}{2}\sigma_{\zeta 12}\delta_\tau^2 & \sigma_{\zeta 1}^2 \delta_\tau & \sigma_{\zeta 12}\delta \\ \frac{1}{2}\sigma_{\zeta 12}\delta_\tau^{\ 2} & \frac{1}{2}\sigma_{\zeta 2}^2 \delta_\tau^2 & \sigma_{\zeta 12}\delta_\tau & \sigma_{\zeta 2}^2 \delta_\tau \end{bmatrix}$$

$$= \delta_\tau \begin{bmatrix} \Sigma_{\eta\zeta}(\delta_\tau) & \Sigma_\zeta(\delta_\tau) \\ \Sigma_\zeta(\delta_\tau) & \Sigma_\zeta \end{bmatrix}.$$

As this is a rather unwieldy expression, not least because we aim to extend the system later to incorporate an arbitrary number N of different stocks, we will simplify it at the implementation stage, and approximate the a priori VCV-matrix by the block diagonal matrix

Equation 148

$$\mathbf{Q}_\tau \approx \delta_\tau \begin{bmatrix} \Sigma_{\eta\zeta}(\delta_\tau) & \mathbf{0} \\ \mathbf{0} & \Sigma_\zeta \end{bmatrix}.$$

The effect of this simplification on estimation, will be negligible in practice, because the effect of the empirical prediction error on the Kalman-filter MSE-matrix will

dominate that of $\mathbf{Q}$ after a few runs through the recursions. This can be simplified further if there is reason to believe that the drift term is fixed, e.g. equal to zero as in the simple random walk models. The resulting specification will be

Equation 149

$$\mathbf{Q}_\tau \approx \delta_\tau \begin{bmatrix} \Sigma_\eta(\delta_\tau) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where

Equation 150

$$\Sigma_\eta(\delta_\tau) = \delta_\tau \begin{bmatrix} \sigma_{\eta 1} & \sigma_{\eta 12} \\ \sigma_{\eta 12} & \sigma_{\eta 2} \end{bmatrix}.$$

To complete the specification of the variance components in the system that we aim to implement, the measurement equation VCV-matrix will be defined as

Equation 151

$$\mathbf{H}_\tau = \begin{bmatrix} \sigma^2_{\varepsilon_{1\tau}} & 0 \\ 0 & \sigma_{\varepsilon_{2\tau}} \end{bmatrix}.$$

Here we recall that this matrix is thought of as time-dependent here only because we want to be able to update the estimates of the measurement error variance in the market when simultaneous price observations occur in the same stock as described above. We recall that the variable δ_τ expresses the time elapsed since the last observation at time τ. It is clear from the above definitions, that when multiple transaction observations occur simultaneously in the same stock, this variable will be zero and as the Kalman-filter runs through the recursions for each observation, $\mathbf{Q}_\tau$ will then be zero as well, but $\mathbf{H}_\tau$ will be nonzero. This is consistent with our view that the state of the system is not changing in such cases, even when simultaneous price observations differ, because the variability is due to measurement error.

In the actual implementation of the system, the Kalman-filter was programmed in a way quite similar to the basic algorithm presented in Harvey (1989).[107] In order to increase the proportion of observations that occur simultaneously, thus obtaining stronger estimates of the covariance between the two stock-prices without increasing the sample size, the degree of precision in timing trades was scaled down from one

[107] P.105-106.

second to one minute. Thus transactions that occur in the same minute are considered simultaneous. As for justification we argue that it is unrealistic to think that the true value of a stock changes in a smaller time-interval than a minute, and therefore timing to the nearest second can be thought of as redundant and excessive precision in this context. Despite this measure, simultaneous observations in both stocks' prices are of course the exception rather than the rule. For the majority of cases when only one of the stocks trades, a slightly different algorithm was used, i.e. a procedure suggested in Harvey (1989) to deal with the problem of delayed observations.[108] In practice this is equivalent to using an algorithm that treats observations as missing when trades do not occur simultaneously, but the 'delayed observations' method proved more convenient to program.

3.3.3 Tentative results

3.3.3.1 The data

The official index number to be estimated here, is the ISE *Index of Transportation Firms* (ITF). One major advantage of this particular index number for the present purposes is that it is calculated from the prices of only two stocks, those of Flugleiðir hf. and those of Eimskip hf. This reduces the programming effort required at this early stage. Although the portfolio is too small for positive nontrading induced autocorrelation to be expected to occur in its return series, we recall from the last section that this does not imply that the index number calculated by the conventional method will correctly reflect the true value of the stocks at any given moment under infrequent trading. However, as these are the largest and most frequently trading issues in the market on average, the correction for nontrading effects that is implicit in the present approach can be expected to have a smaller effect than it would for any other portfolio of ISE stocks. Whether or not the index calculated from filtered prices differs substantially from the conventional one, an interesting sequel to this experiment will be to estimate an index of less frequently trading stocks and compare the effects. An obvious candidate is the *Index of Oil Distribution Firms*, which is calculated on the basis of a portfolio of three stocks in which trade is quite thin.

---

[108] P.465-466

We may note that as both companies in the portfolio of transportation firms belong to the same sector, positive contemporaneous correlation may be fairly strong, increasing the efficiency gain from simultaneous estimation. The period from 2.05.1996 to 28.02.1997 was chosen as it is the largest and most recent continuous sample period in which prices can be assumed to be free of distortion from stock splits and dividend payments.[109] The data sample contains 627 irregularly spaced and (mainly) non-synchronous price observations for Eimskip and 549 for Flugleiðir. This means that if the filtering experiment is successful we can expect to obtain more than a thousand estimates of each stock's price and the index, but some observations will always be lost while the filter is converging.

The time unit used in this experiment is a day. In the original data set, transactions are timed to a fraction of a day corresponding to one second, but here this superlative degree of precision is reduced to a resolution that corresponds to one minute. The Kalman filter is then applied to the logarithms of prices. At the estimation stage the ¾ of each business day which is not business time, are 'removed' from the transaction time-series, as well as whole days on which neither stock trades. This is a rough way to conform with the idea that price-uncertainty (measured by the mean-square-error (MSE) of the state estimate) should not be affected in the same way by the passing of time during closing hours, as when the exchange is operating. It is quite possible that removing this time altogether is too radical a measure. Some compromise may be more realistic, and this is one aspect of the implementation that requires further study.[110] Also, this means that we are assuming that the exchange is closed if neither stock trades. While the error induced by this is likely to be negligible in the case of the two transportation stocks as they are the two most frequently trading ones in the exchange, a list of trading holidays will have to be used in other cases.

After the estimates of the value of each stock are obtained by filtering the logarithms are turned back into levels and the index is calculated by the same formula as used to calculate the official index on the basis of 'closing prices'. At this stage an additional advantage of the method appears, which is that chaining takes place at much shorter

---

[109] For ISE companies such decisions are made at shareholder meetings that are held in March or April each year.

[110] BF&P discuss this issue in their study cited earlier, (p.144-145) but without suggesting a way of dealing with this problem in general.

intervals than in the daily index, permitting closer approximation to the ideal of a continuous Divisia index.

3.3.3.2 Initial values and variance components.

The initial values of the state vectors were set equal to the first observation for the level term and zero for the local trend (or drift). The initial MSE matrix was set to as a diagonal matrix with all terms equal to a 'finite but very large' value.[111] This is consistent with the assumption that the parameters of the state-vector follow a random walk process.[112]

If the state is assumed to follow a simple random walk with zero drift (henceforth, SRW), the state-vector is one-dimensional and the total variance in the sample will simply be the sum of measurement error variance and the variance of the innovations to the level component. Then a reasonable estimate of the latter can be obtained by subtracting the estimated measurement error variance from the estimated total variance in the sample, which can easily be obtained in the usual way. Calculating a weighted average of measurement-error variance estimates over the whole period, we obtained $\sigma_{\varepsilon 1}^2 = 2.75\times10^{-4}$ for the *Flugleiðir* series and $\sigma_{\varepsilon 2}^2 = 1.9\times10^{-4}$ for *Eimskip*. The total variance of price logarithms for *Flugleiðir* was 0.002975 for the whole period, which implies a daily state variance of $\sigma_{\eta 1}^2 = 1.3\times10^{-5}$, after accounting for the contribution of measurement error to total variance. Total variance for the *Eimskip* series in this period is 0.008039. Subtracting the measurement error contribution and dividing by the number of days we obtain an implied state variance of $\sigma_{\eta 2}^2 = 3.8\times10^{-5}$. For the SRW model, these are all the required parameter settings.

If, however, the state also has a local linear trend term, the issue of how to set the relative contribution of each term to total variance of the state arises. Recalling that the reason for including such a term was that it represents a 'mean return' parameter and we believed that mean return on stocks would be constrained to a certain extent by the mean return offered on other investment opportunities in the economy. This means that the variance component due to this term, should be set at some "very low" value, or even fixed, if the 'mean return on investment' does not fluctuate greatly. In principle, estimates consistent with this argument can be derived on the basis of a

---

[111] Here: 1.000.000.000

study of interest rates and other return indicators for the period in question. For the present purposes it was assumed, somewhat arbitrarily, that a standard deviation of 1.0 percentage points would be plausible for 'mean return' on both stocks over a 208 day period. This implies a daily state variance of $\sigma^2_{\zeta 1} = \sigma^2_{\zeta 2} = 5.0\times10^{-7}$ for the mean return parameters.

To set the state covariance assumption for the two stocks, correlations were estimated for the period 2.05.96 until the end of 1997 from daily data. Because nonsynchronous trading effects can be expected to increase correlation at the first lead and lag at the expense of contemporaneous correlation, these three coefficients were summed to get an estimate of the true contemporaneous correlation. This gave a contemporaneous correlation estimate of 0.27, that was then used to assess the correct *a priori* state covariance for both the level and the local trend terms. This yielded the following figures: $\sigma_{\eta 12} = 6.0\times10^{-6}$ and $\sigma_{\zeta 12} = 1.35\times10^{-7}$.

Table 3: Summary of assumptions about the variance terms.

Source of Variance	Flugleiðir hf.	Eimskip hf.
Measurement error	$\sigma^2_{\varepsilon 1} = 2.75\times10^{-4}$	$\sigma^2_{\varepsilon 2} = 1.9\times10^{-4}$
State variability: level	$\sigma^2_{\eta 1} = 1.3\times10^{-5}$	$\sigma^2_{\eta 2} = 3.8\times10^{-5}$
State variability: local trend	$\sigma^2_{\zeta 1} = 5.0\times10^{-7}$	$\sigma^2_{\zeta 2} = 5.0\times10^{-7}$
State covariance: level	$\sigma_{\eta 12} = 6.0\times10^{-6}$	
State covariance: local trend	$\sigma_{\zeta 12} = 1.35\times10^{-7}$	

3.3.3.3 Three estimation experiments

Using these values the ITF was then estimated by a simultaneous system Kalman-filter. First a SRW (simple random walk) was assumed and the measurement equation variance premises were allowed to change each time a new estimate could be obtained using the method described above. Second, the SRW model was maintained, but measurement equation variance kept fixed throughout the sample period for each firm, using the weighted average of estimates. Third, a 'random walk with drift'

[112] See Harvey (1989), p.107-108.

(RWD) model was estimated, with the measurement equation variance fixed. The results that were obtained are presented graphically in **Figures 17**, **18** and **19**.

Figure 17: Log of reported Flugleiðir price with confidence intervals (SRW model, variable variance assumptions)

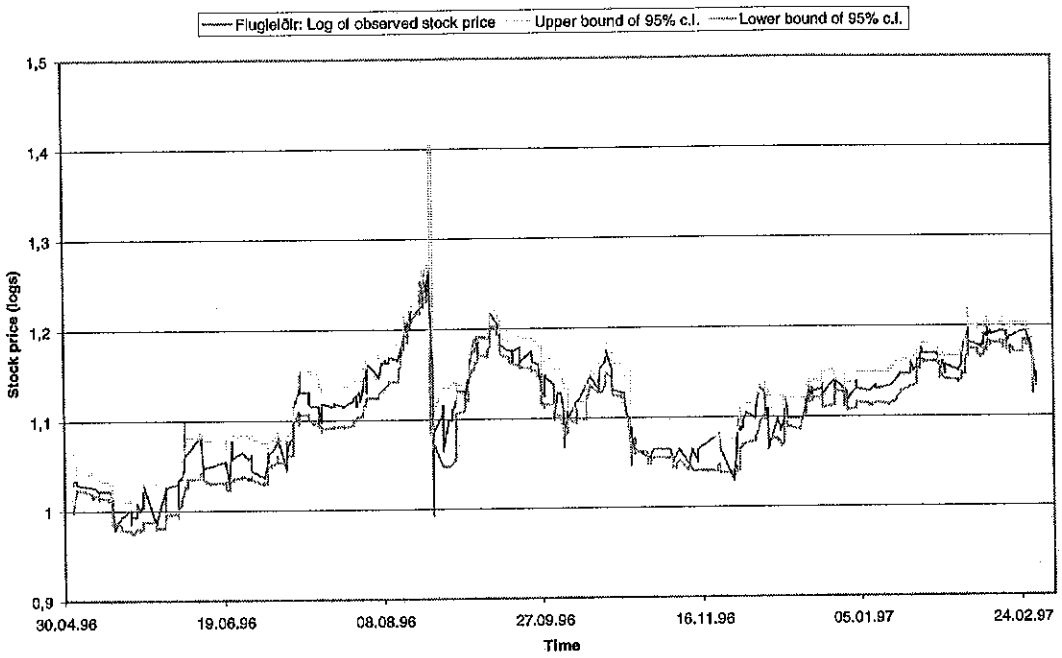


Figure 18: Log of reported Eimskip price with confidence intervals (SRW model, variable variance assumptions)

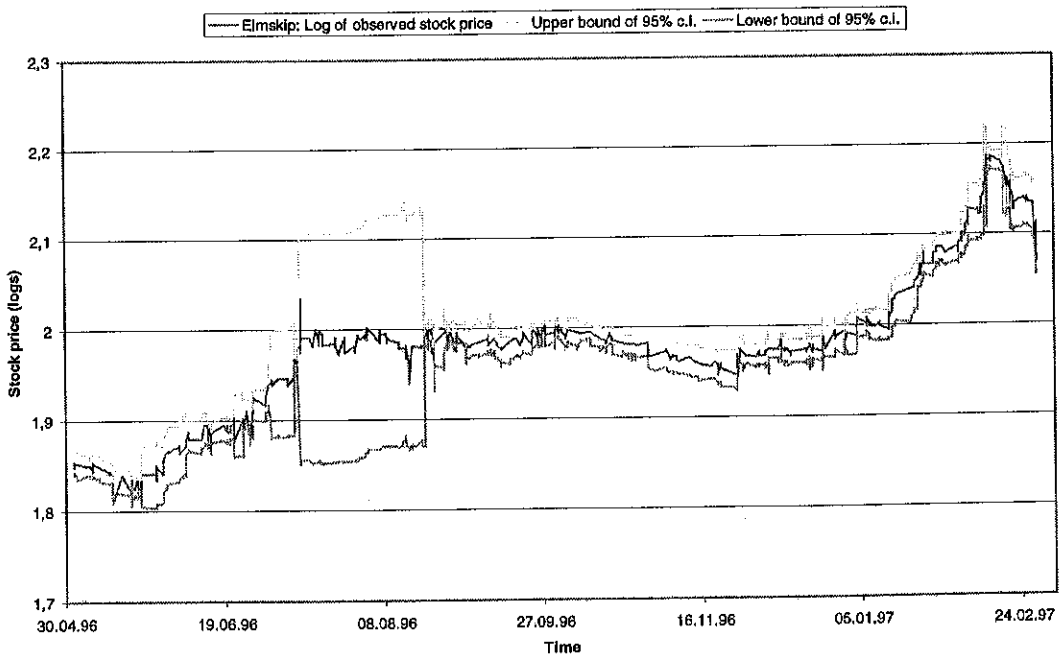


Figure 19: Official ITF index with KF confidence intervals. SRW model, variable assumptions. (in levels)

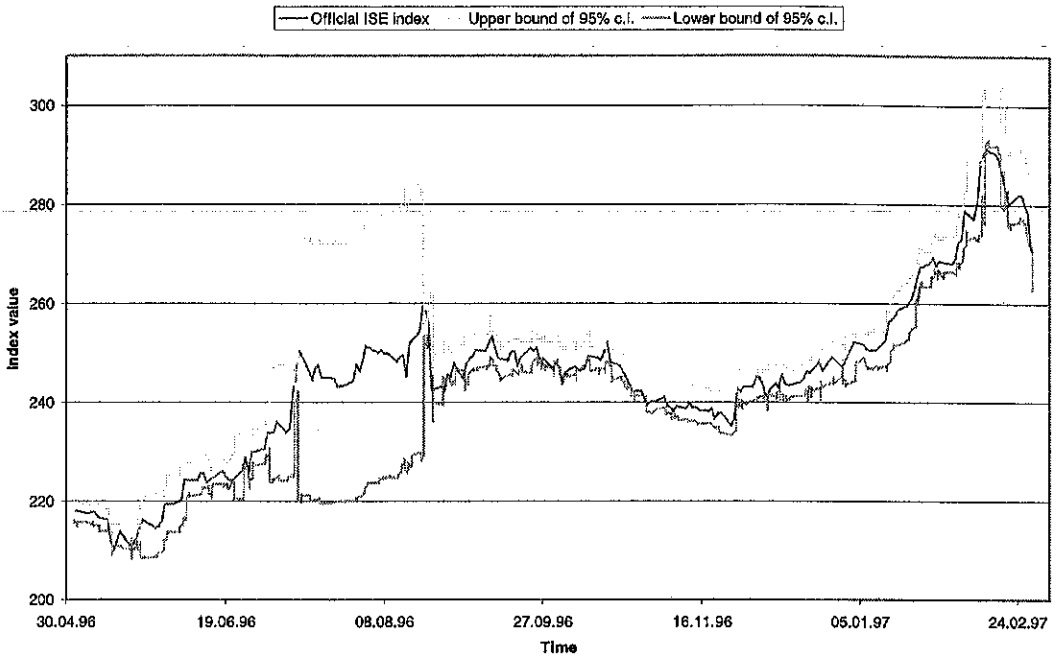


Figure 18 shows the effects of 'outliers' in the estimates in two readily identified cases which carry over to the index estimate as evident in **Figure 19**. The former (and smaller) notable effect is caused by the arrival of an estimate of 0.00383, and the latter (larger and more extended) anomaly is the result of the occurrence of an estimate of 0.00637 for a fixed value of the state equation variance. This phenomenon may serve to underline the importance of the correct choice of these parameters for the efficiency of estimation. Estimation results for *Flugleiðir hf.* do not exhibit symptoms of outliers. Contemplating the likely sources of exceptionally high measurement error variance estimates, it seems possible that in some cases dealers may have ulterior motives in filling large orders that lead them to override their own views about the acceptable uncertainty margin about their bid- or ask price. One such motive could be to obtain or keep the business of an important client in view of long term gain.

The confidence intervals in the graphs are based on the root mean square error of the filtered prices at each point, assuming lognormally distributed returns. The dark line represents the logarithm of observed prices for the individual stocks and the official index value published daily by the ISE. Inspection reveals that the official index lies almost everywhere within the 95% confidence intervals around the filtered values. In

words, this can be expressed as implying that the two estimates do not differ 'radically'. In the present graph, disregarding the anomalies, this confidence margin typically represents some 5 index points in each direction which is roughly equivalent to ±2% of the index value.

The next experiment involves fixing the measurement equation variance assumption for each firm throughout the period, at the value stated in **Table 3**, which is a weighted average of the figures obtained at different times. No attempt was made at discounting the estimates we have termed as outliers. All other assumptions remain the same as before and in particular, this is still an SRW model.

Figure 21 shows that the anomalies apparent in **Figure 18** have now disappeared, but 95% confidence intervals now seem wider over most of the period for *Eimskip hf.* The latter also holds for *Flugleiðir hf.* (**Figure 20** compared to **Figure 17**) and this is a logical consequence of averaging the expected measurement variance, because sporadic increases in expected measurement error have now been distributed more equally over the whole period.

Figure 20: Log of reported Flugleiðir price with confidence intervals (SRW model, fixed variance assumptions)

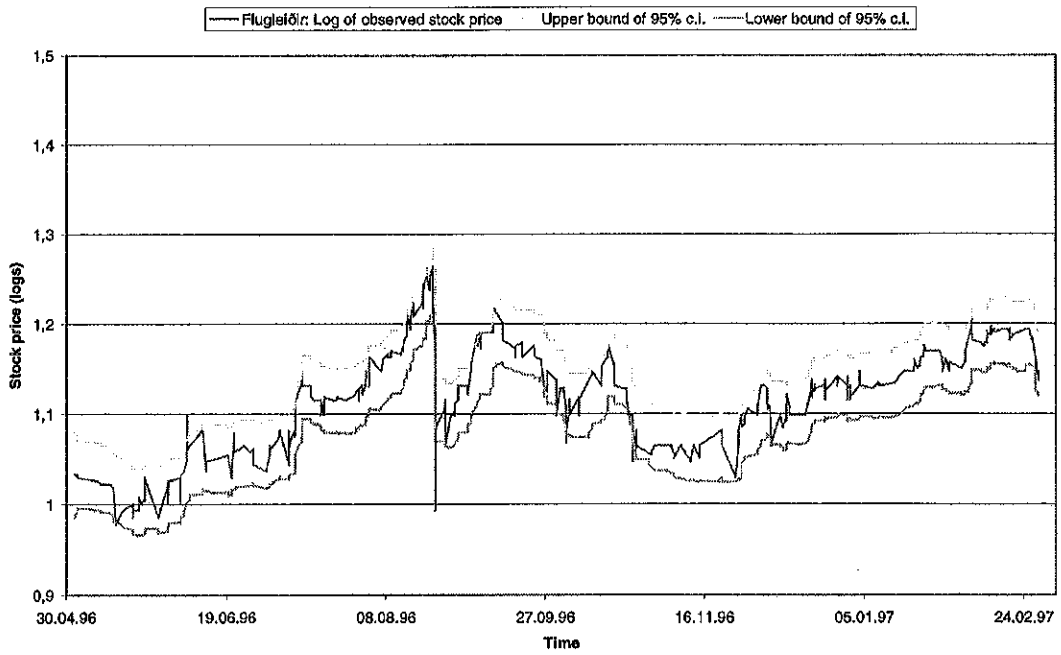


Figure 21: Log of reported Flugleiðir price with confidence intervals (SRW model, fixed variance assumptions)

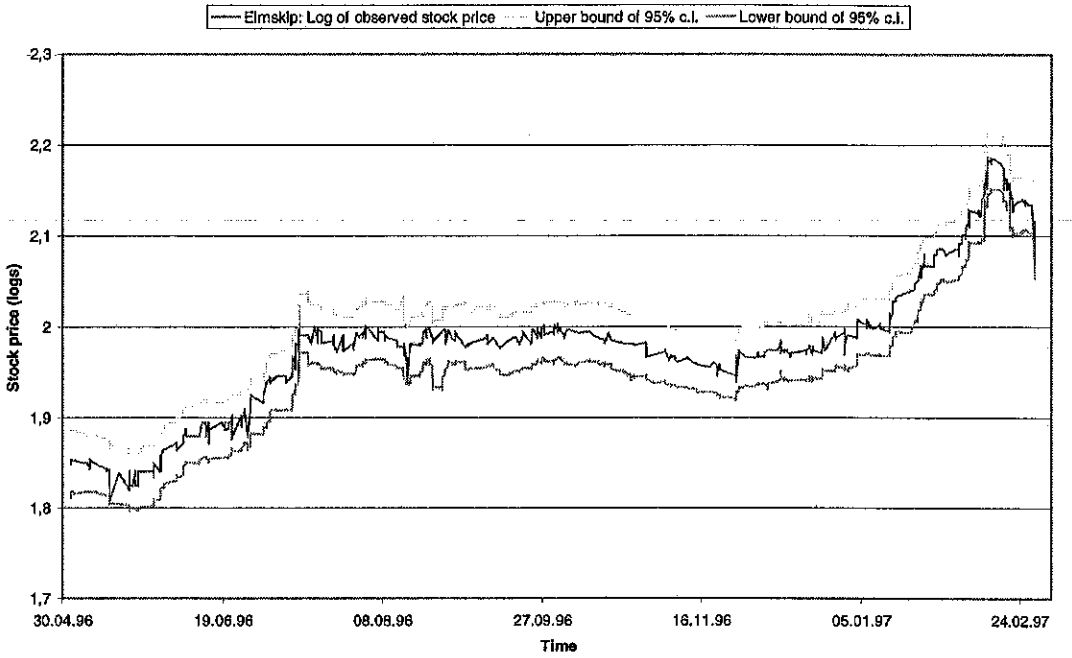
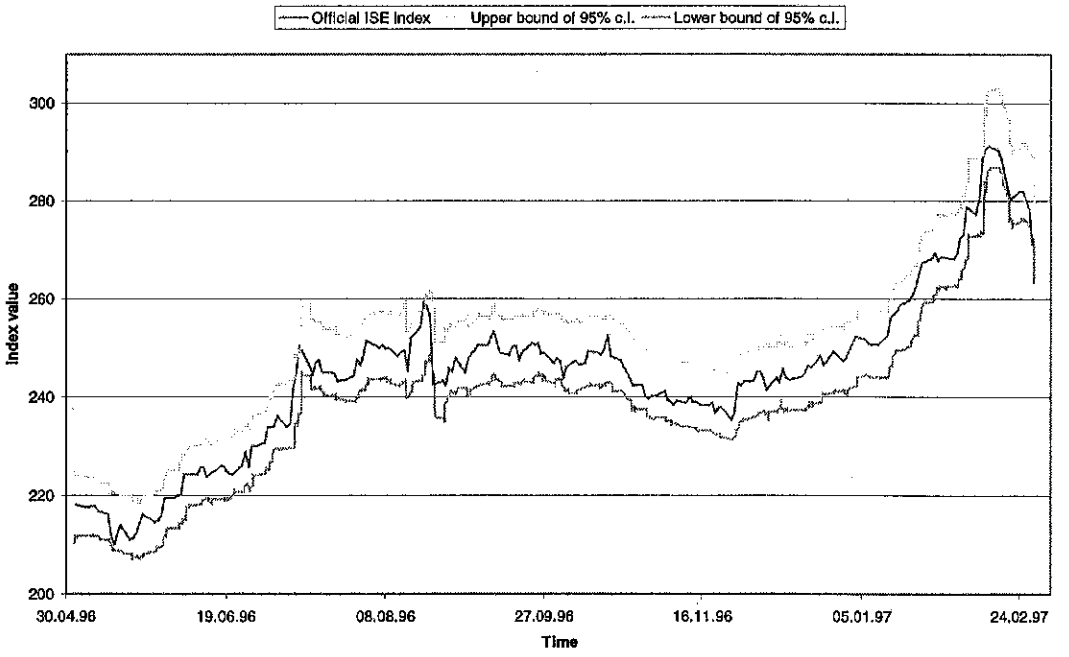


Figure 22: Official ITF index with KF confidence intervals. SRW model, fixed assumptions. (in levels)



A close look at the confidence intervals in **Figure 22** in comparison to those of **Figure 19** also yields the impression of a smoother filtered index estimate than before, which is a consequence of the increased amount of assumed uncertainty in measuring prices over the whole period which allows the state estimate to change less readily in response to observed price changes. Comparing the confidence intervals of

the two index graphs the filtered index in the former appears to approximate the true value with roughly twice the precision of the latter in terms of root mean square error almost everywhere. This suggests that it is important to analyse the reasons for exceptional estimates of the measurement variance estimates and to devise more sophisticated methods of incorporating this information into the index estimation procedure if the analysis does not indicate that such estimates should be discarded altogether.

The third and final ITF estimation experiment consists in setting the variance and covariance of the drift terms of both firms in the two-dimensional state space model to a nonzero value.[113] This means that we are now estimating two aspects of the state, its *level* and its *local trend*, the latter having a possible interpretation as a return parameter. To illustrate the effect of this on the performance of the filter, graphs are presented as before.

Figure 23: Log of reported Flugleiðir price with confidence intervals (RWD model, fixed variance assumptions)

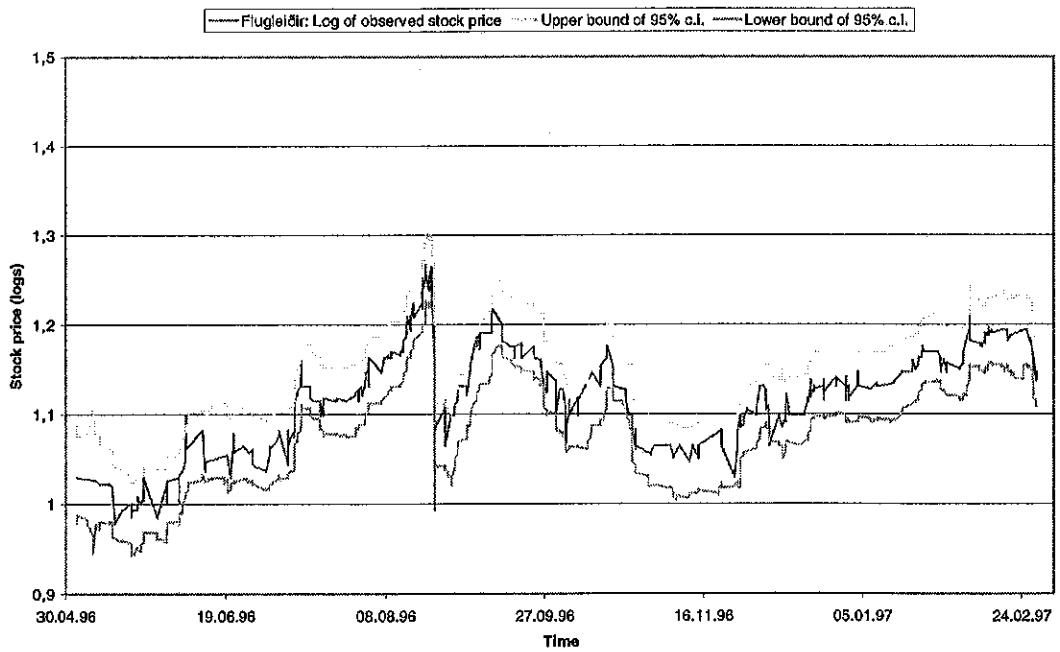


Figure 24:Log of reported Eimskip price with confidence intervals (RWD model, fixed variance assumptions)

[113] See *Table 3-2*

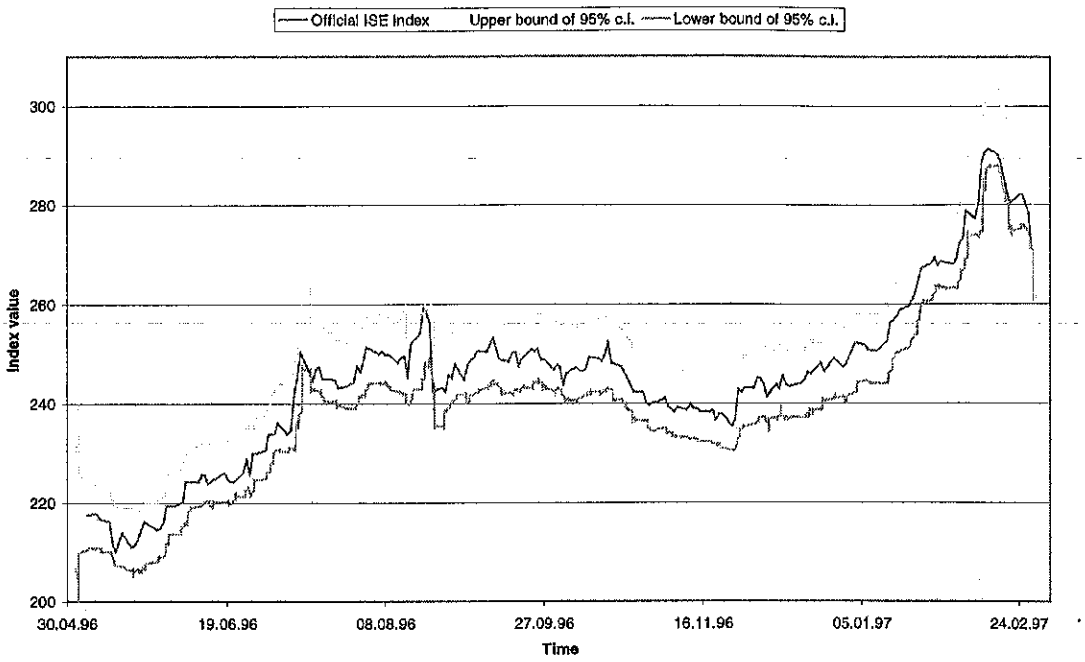
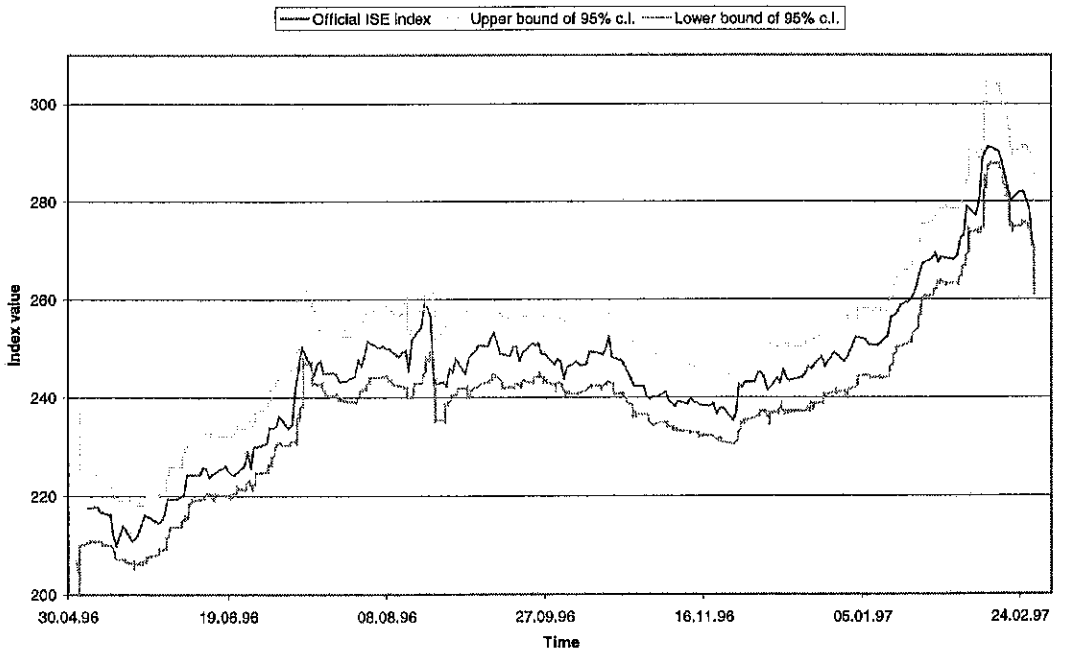


Figure 25: Official index with KF confidence intervals. RWD model and fixed assumptions (in levels)



The extremely wide confidence intervals at the start of the period are symtomatic of slower convergence of the filter due to the load of estimating twice as many parameters. For the rest of the period confidence intervals around the filtered index look more jagged than in **Figure 22** reflecting the greater flexibility of the

formulation of the state when it is two-dimensional. However, there is no clear indication that this represents an improvement in estimation. Thus, by the principle of parsimony, we conclude that this model should be rejected and the simple random walk model explored more fully before the addition of a slope parameter is given any serious consideration.

The overall impression resulting from the informal analysis conducted in this section is that the filtered estimates do not differ substantially from the index as calculated from observed prices in this case. As we are not in a position to assess a figure such as the MSE of estimate for the official index, a direct comparison of the filtering results with the official 'estimator' is not feasible, but for lack of a better measure of their difference or similarity, we could point to the fact that judging from the graphs, the official index falls inside the 95% confidence intervals around the filtered one. This is consistent with the fact that infrequent trading leaves the expected value of a stock index unaffected, which is a standard conclusion in the infrequent trading litterature. In other words, *on the average* the official index returns will approximate true returns, or, to put it differently still: the true value of the underlying stocks will not drift arbitrarily far from the officially reported index value. How far it is likely to stray, depends on the degree of non-trading. Thus, for the two stocks that form the transportation index, the likely deviation between official and filtered estimates represents a minimum among all possible ISE portfolios.

Before concluding this exposition of the filtering experiments that were undertaken at this tentative stage of research, two last graphs should be inspected, showing the officially reported and the filtered index together. The underlying assumptions in the first graph are those of the simple random walk with fixed measurement variance (same as the one underlying **Figure 22**). In the second graph we look at the one with the assumption of a variable level of measurement error, as in **Figure 19**. Here are inspecting roughly half of the whole period, starting with a famous date and including the time span of the largest estimated measurement uncertainty in **Figure 19** which approximately corresponds to the left half of the present subperiod (11.07.96 to 20.08.96 to be exact).

Figure 26: Filtered and official ISE index over 100 days (SRW, fixed assumptions)

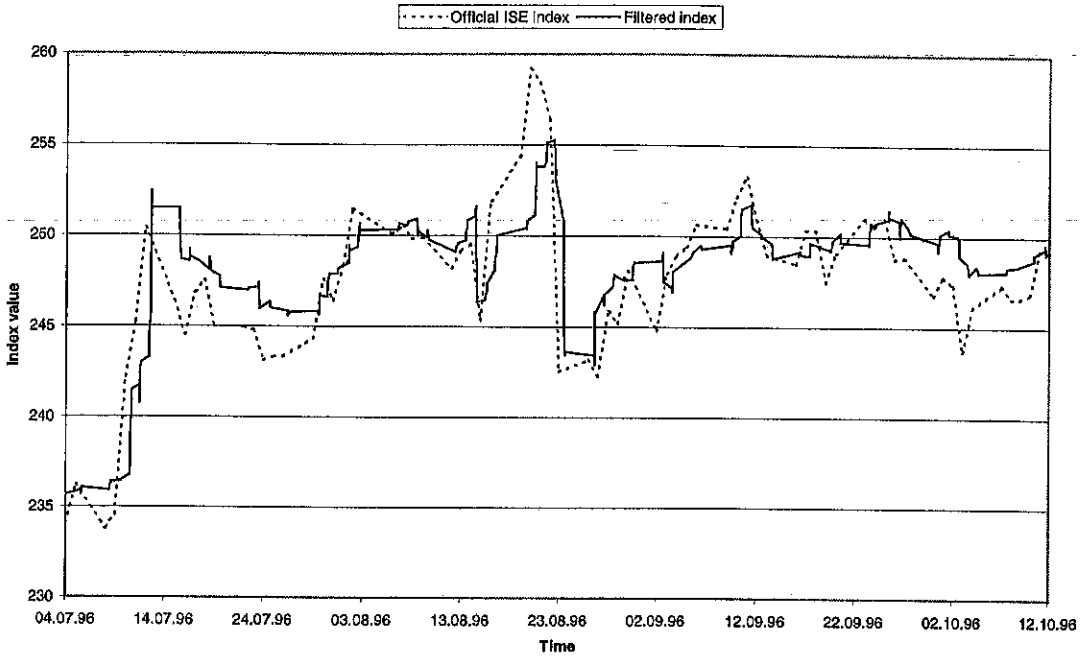
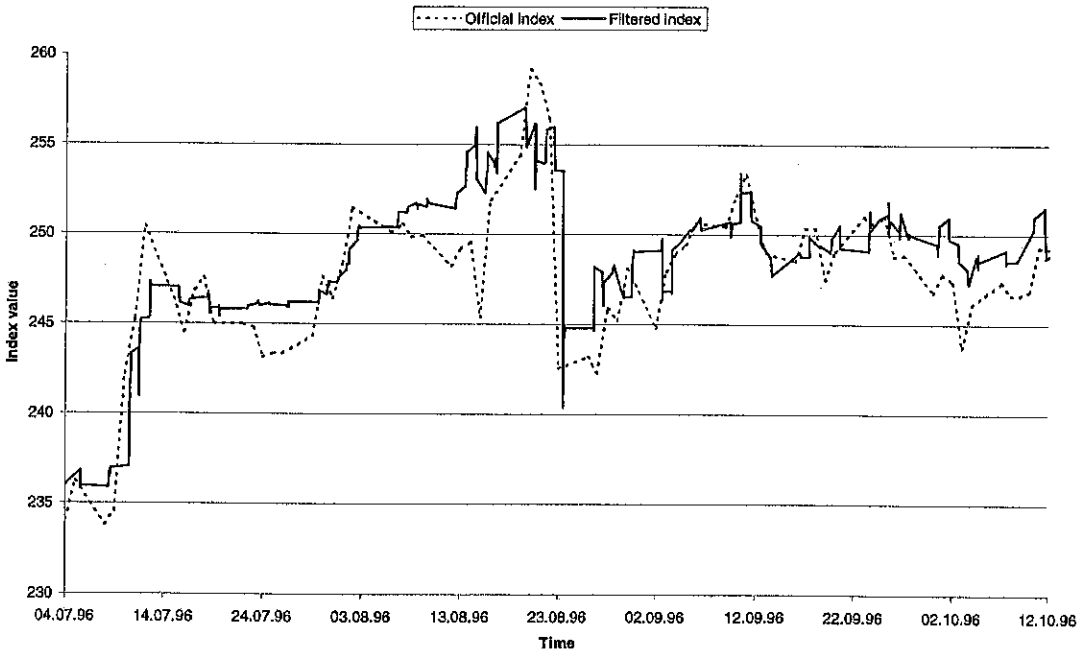


Figure 27: Filtered and official ISE index over 100 days (SRW, variable assumptions)



Apart from revealing to a certain extent how a filtered index compares to the official one in the case at hand, these graphs give some idea of the effects of the different treatment of price uncertainty. Thus we see that compared to the model that assumes a fixed average measurement error, the variable assumptions yield a less volatile index in the subperiod of high estimated price uncertainty in the market and a more volatile

one in the latter half-period when this estimate is lower. By whichever definition, the filtered series is much more jagged, reflecting intraday variation in the underlying stocks and the far greater number of observations, those of the filtered series exceeding those of the observed index by a factor of five on the average. Although we will not analyse the volatility of the filtered series in detail on this occasion it must be noted that inspection of the graphs suggests that in daily terms the filtered series has lower variance in this period. This is consistent with theoretical results because in this case the index portfolio is not 'well diversified' in the sense of infrequent trading theory. Thus negative serial correlation of the individual components predicted in the presence of infrequent trading should not be expected to be dominated by the time-averaging effect of summing over stocks in the index.

It is urgent to analyse the volatility of filtered index series in more detail and compare it to that of the index calculated directly from observed prices. One reason for this is that reliable estimates of volatility in the true value of stocks under infrequent trading can be of great practical importance for enhanced efficiency of risk management in smaller and less mature markets. Another reason is that this task fits well into the framework of further research into the optimal index filter in the presence of microstructure effects, providing one possible way to check the success of a proposed estimator based on an efficient market model. For the time being, however, as is the case for so many related issues touched upon in this report, we must be content with asserting that a door has been opened.

Conclusions

At the outset we discussed index numbers in a general framework, before restricting our focus to stock index numbers and infrequent trading as a specific type of measurement error affecting the reliability of the index measuring portfolio value. We have argued that in the latter situation a Kalman-filter approach to estimation will represent an improvement with respect to current practices. However, this approach is also likely to be applicable to the problem of estimating other types of economic index numbers, e.g. the CPI. In view of the important role of index numbers in the theory and practice of economics, it would certainly be a worthwhile endeavour to explore this possiblity in some detail.

Although the estimated portfolio consists the two the most frequently traded stocks in the Icelandic market, some important advantages of the filtering method emerged on inspection of the resulting graphs. However, the most important result emerging from the present monograph may well be that it demonstrates in principle the feasibility of tackling microstructure problems in the framework of continuous-time state space models. In other words, if the line of reasoning presented here is valid, a whole agenda of new problems emerges. By way of conclusion, it is worthwhile to summarise the two aspects of the matter separately.

The main advantages offered by the use of a filtered stock index similar to the ones suggested here are the following:

- A filtered index eliminates the 'errors-in-variables' aspect of the infrequent trading problem by explicitly accounting for unequally spaced, non-synchronous observations in a continuous market.
- It is calculated from optimal estimates of value instead of directly from observed prices. Although the model and its prior assumptions on which optimality is conditioned in this case are certainly open to further discussion, this general approach is a matter of principle in the presence of measurement uncertainty.
- It yields local estimates of the precision of the reported index value, which is also a matter of principle. This additional information must also be quite useful to

practitioneers, given that a filtered index will eventually be implemented and published on a regular basis.

- It yields a value of the index far more frequently than the usual index can, providing a denser stream of information to investors. The arrival rate of new index estimates is the sum of those of the individual constituent stocks. This also minimises path bias by offering more frequent chaining and thereby a closer approximation to a continuous Divisia index.

Although the present implementation has many weaknesses, some of which have been indicated along the way, it is a worthwhile task to extend it to other sector indices in the ISE as well as to the ISE all-share yield index, preferably in a real-time environment. This will provide a valuable benchmark against which empirical and theoretical results can be assessed in the course of further research. Developing the approach will involve analysis of the following problems in the first stages.

- Ways of incorporating stock specific empirical information about uncertainty of measurement into the underlying prior assumptions. As a first step, quotation data must be inspected in order to better understand the sources of outliers as were encountered in the last section.
- Augmenting the current SRW model of the state in such a way as to account for a possible bid-ask spread in the market. This is especially important because the observations used to derive estimates of measurement uncertainty may in most cases originate with traders that depend on a spread. A way of doing this might be by making the SRW model two-dimensional, consisting of a bid-value and an ask-value. The measurement equation would then have to involve a Bernoulli indicator variable, i.e. $y_t = (1-I_t)\alpha_{1t} + I_t\alpha_{2t} + \varepsilon_t$ because bid and ask transactions should not be expected to occur simultaneously. The two alphas are then considered as unobservable structural components of the data in the spirit of Harvey (1990). If this proves feasible it would seem to amount to an alternative to presently existing ways of estimating the market spread itself.[114]

[114] Some existing approaches to this problem are described in Campbell, Lo and MacKinley (1997), chapter 3.

- An undesireable aspect of a small and thin markets that has not been treated in this report is that trades involving a very small fraction of a company's stock can have a disproportionate effect on the official total value estimate of that company. One aspect of this problem is that traders can move the market in particular stocks or the index at a relatively low cost. Thus it is impractical to offer derivatives in locally traded stock to investors.The state space approach seems to offer a way out in this situation; by taking account of the volume of each trade in determining its associated measurement uncertainty, small transactions can be discounted relative to large ones.

Other interesting problems and avenues of research could be enumerated at this point. In particular, exploitation of quotation data, modeling of the stochastic liquidity specific to each stock. and more elaborate nonlinear methods of filtering could be suggested as ways of making the present approach more complete. While all of these do indeed offer great possibilities, this author feels that the limits of the simplest set of means should be in sight before stepping on to radically more complex methods.

List of References

Atchison, M. D.; Butler, K. C. and Simonds, R. R. "Nonsynchronous Security Trading and Market Index Autocorrelation." *Journal of Finance* 42 (March 1987), 111-118.

Bassett, G.W., Jr.; France, V.G. and Pliska, S.R. "Kalman Filter Estimation for Valuing Nontrading Securities, with Applications to the MMI Cash-Future Spread on October 19 and 20, 1987." *Review of Quantitative Finance and Accounting,* 1 (1991), 135-151.

Berglund,T., Liljeblom, E. and Löflund, A., 1989, "Estimating Betas on Daily Data for a Small Stock Market", *Journal of Banking and Finance* 13, 41-64.

Bergstrom, A.K., "Continuous time stochastic models and issues of aggregation over time", *Handbook of Econometrics*, Vol. II, North-Holland: Elsevier Science Publishers, Amsterdam, (1984),1145-1212.

Beveridge, S. and Nelson,C.R., "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle." *Journal of Monetary Economics,* 7 (1981), 151-174.

Campbell, John Y.; Lo, Andrew W. and MacKinley, A. Craig. *The Econometrics of Financial Markets* Princeton, NJ: Princeton University Press (1997).

Cohen, K.J., Hawawini, G.A., Maier, S.F. , Schwartz, R.A. and Whitcomb, D.A. "Implications of Microstructure Theory for Empirical Research on Stock Price Behaviour." *Journal of Finance* 35 (May 1980), 249-257.

Cohen, K.J., Hawawini, G.A., Maier, S.F. , Schwartz, R.A. and Whitcomb, D.A. "Estimating and Adjusting for the Intervalling-Effect Bias in Beta" *Management Science*, 29 (1983), 135-148.

Cohen, K.J.; Maier, S.A.; Schwartz, R.A. and Whitcomb, D.K. *The Microstructure of Securities Markets* Englewood Cliffs, NJ: Prentice-Hall (1986)

Fisher, L. "Some New Stock-Market Indexes" *Journal of Business*, XXXIX (January 1966), 191-218

Harvey, A. C., *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press (1989).

Harvey, A. C., *Time Series Models.* Second edition, Hemel Hempstead, Hertfordshire, Harvester Wheatsheaf (1993).

Harris, L. "The October 1987 Stock Futures Basis" *Journal of Finance*, XLIV, 1 (March, 1989)

Hawawini, G.A. "The Intertemporal Cross Price Behaviour of Common Stocks: Evidence and Implications." *Journal of Financial Research* 3 (Fall 1980), 153-167.

Hawawini, G.A. "Equity price behaviour: Some evidence from markets around the world." *Journal of Banking and Finance* 18, (1994), 604-620.

ISE Annual Report 1996 , Reykjavík, VÞÍ (1997, in Icelandic),

Jokivuolle, Esa, "Measuring True Stock Index Value in the Presence of Infrequent Trading." *Journal of Financial and Quantitative Analysis* 30 (September 1995), 455-464.

Judge, G.C., Hill, R.C.,Griffiths, W.E., Lütkepohl, H., Lee, T.-C., *Introduction to the Theory and Practice of Econometrics*, New York, John Wiley & Sons Inc. (1988)

Kallunki, J.P., "Using accounting variables as instrumental risk measures in event studies: Evidence from a thinly traded stock market." Working Paper, Department of Accounting and Finance, University of Waasa, (1997). Forthcoming in the *Journal of Financial Management*

Kallunki, J.P., "Handling missing prices in a thinly traded stock market: Implications for the specification of event study methods", Working Paper, Graduate School of Finance and Financial Accounting, University of Waasa, (1997) Forthcoming in the *European Journal of Operations Research.*

Kallunki, J.P. and Martikainen, T., "The covariance factor structure of daily returns in a thinly traded stock market", Working Paper, Department of Accounting and Finance, University of Waasa, (1997). Forthcoming in the *Journal of Multinational Financial Management*

Kitagawa, G. "A Nonstationary Time Series Model and its Fitting by a Recursive Filter." *Journal of Time Series Analysis*, 2. (1981)103-116

Kitagawa, G. "Non-Gaussian State-Space Modeling of Nonstationary Time Series." *Journal of the American Statistical Association*, 82. (1987) 1032-1063

Lo, A. and MacKinley, A.C., "Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test." *Review of Financial Studies*, 1 (1988) 41-66.*

Lo, A. and MacKinley, A.C., "An Econometric Analysis of Nonsynchronous Trading." *Journal of Econometrics*, 45 (1990), 181-211.

Martikainen, T., Pertunen, J., Yli-Olli, P. and Gunasekaran, A., "On the impact of infrequent trading on the APT systematic risk components – Evidence from a thin security market", *European Journal of Operational Research*, 88 (1996), p.23-27.

Merton, Robert C., *Continuous-Time Finance*, Cambridge, MA: Basil Blackwell (1990)

Miller, Merton H.; Muthuswamy, Jayaram and Whaley, Robert E., "Mean Reversion of Standard & Poor's 500 Index Basis Changes: Arbitrage Induced or Statistical Illusion ?" *Journal of Finance* XLIX (June 1994), 479-507.

Martikainen, T. and Perttunen, J., "Return intervals, systematic risk estimates and firm size. Evidence from a thin security market." *Economics Letters* 36/3, (1991), p.311-315.

Martikainen, T., Perttunen, J., Yli-Olli, P. and Gunasekaran, A., "The Impact of the Return Interval on Common Factors in Stock Returns: Evidence from a Thin Security Market." *Journal of Banking and Finance* 18 (1994) 659-672.

Martikainen, T., Perttunen, J. and Puttonen, V. "The lead-lag effect between large and small firms. Evidence from Finland." *Journal of Business Finance and Accounting*, 22/3, (1995), p.449-454.

Martikainen, T., Perttunen, J., Yli-Olli, P. and Gunasekaran, A., "On the Impact of Infrequent Trading on the Systematic Risk Components of the APT.", *European Journal of Operational Research*, 88 (1996), 23-37.

Newbold, P. *Statistics for Business and Economics*, Englewood Cliffs, NJ: Prentice-Hall, Inc. (1995)

O'Hara, M., *Market Microstructure Theory*, Cambridge, MA: Blackwell Publishers, (1994)

Parzen, E. (Ed.), *Time Series Analysis of Irregular Observations*, New York: Springer Verlag (1984)

Perry, P.R. "Portfolio Serial Correlation and Nonsynchronous Trading." *Journal of Financial and Quantitative Analysis* 20 (December 1985), 517-523.

Scholes, M. and Williams, J. "Estimating Betas from Nonsynchronous Data" *Journal of Financial Economics* 5 (December 1977), 309-327.

Sigurgeir Örn Jónsson, "Efficiency and price-formation in securities markets: a study of the Icelandic stock market", B.Sc.-thesis, Faculty of Business and Economics, University of Iceland. (1997, in Icelandic)

Selvanathan, E.A. and Prasada Rao, D.S. *Index Numbers - a Stochastic Approach.* London: Macmillan Press (1994)

Smith, Tom, "Econometrics of Financial Markets and Market Microstructure Effects" *Journal of Financial and Quantitative Analysis* 29 (September 1994), 519-540.

Shanken, J., "Nonsynchronous Data and the Covariance-Factor Structure of Returns" *Journal of Finance* 42 (1987), 221-232.

Stoll, H.R., and Whaley, R.E. "The Dynamics of Stock Index and Stock Index Futures Returns" *Journal of Financial and Quantitative Analysis* 25 (1990), 441-468.

Vasicek, O., (1973), "A note on using cross-sectional information in Bayesian estimation of security betas", *Journal of Finance* 28, Dec., 1233-1239.

Wei, William W.S. *Time Series Analysis - Univariate and Multivariate Methods.* Redwood City, CA: Addison-Wesley Publishing Company. (1990)

Working, Holbrook "Note on the Correlation of First Differences of Averages in a Random Chain" *Econometrica* XXVIII (October 1960), 916-918.