EFI, EKONOMISKA FORSKNINGSINSTITUTET
VID HANDELSHÖGSKOLAN I STOCKHOLM

Learning the True Index Level: Index Return Autocorrelation in an REE Auction Market

Patrik Säfvenblad

Working Paper No. 190

September 1997

Working Paper Series
in Economics and Finance

Learning the True Index Level: Index Return Autocorrelation in an REE Auction Market

Patrik Säfvenblad*
Stockholm School of Economics
Department of Finance

September 1, 1997

Working Paper Series in Economics and Finance No. 190

Abstract

This paper develops and tests implications of cross-security information aggregation on index return autocorrelation. In the model, prices are realised individually and simultaneously in REE auction markets, then realigned to take information revealed in other prices into account. This adjustment is symmetric across stocks, leading to index return autocorrelation of MA(1) type.

Autocorrelation will be high if the index level prior is noisy, for example, at Monday open and after high volatility in overnight trading. Autocorrelation will also be higher in portfolios of highly correlated securities. Overnight information revelation and high trading volume reduces the noisiness of the index level prior and, consequently, return autocorrelation.

Index return autocorrelation will be low, or even negative, if there is high cross-security correlation in revealed information, due to, for example, index arbitrage trading or profit taking.

All major predictions are supported by tests using data from the Paris Bourse. In contrast to earlier models of index return autocorrelation, the model can generate both positive and negative index return autocorrelation. This paper also documents instances of negative index return autocorrelation.

*The paper has benefitted from comments made by Clas Bergström, Bevan Blair, Magnus Dahlquist, Richard Lyons and Staffan Viotti. I thank SBF–Paris Bourse, especially Marianne Demarchi and Solenn Thomas, for its hospitality and for providing the data used. I gratefully acknowledge research funding from Bankforskningsinstitutet. Address: Box 6501, S-113 83 Stockholm, Telephone: +46-8-736 9000, Fax: +46-8-31 23 27, Internet E-mail: finpsa@hhs.se.

1 Introduction

Short-term stock index returns are, in most observed cases, positively autocorrelated. This apparent breach of the efficient market hypothesis has attracted much attention, both theoretically and empirically. Three explanations dominate the theoretical literature: nonsynchronous trading, transaction costs and time-varying expected returns.[1]

The oldest and most widely accepted of these hypotheses is the nonsynchronous trading hypothesis, originated by Fischer (1966) and Scholes and Willams (1977). Prices are assumed to be informationally efficient, but only measurable when stocks actually trade. If the component stocks of a stock index trade at separate points in time (nonsynchronously), last recorded prices reflect partly old market-wide information and a delay will occur in the index's reaction to new information even if the "true" index level is known by all market participants.

It is possible to fine-tune the nonsynchronous trading model to generate high levels of index return autocorrelation by assuming extreme nontrading in certain stocks or assuming that thinly traded stocks have high betas. However, even with these enhancements, nonsynchronous trading only accounts for parts of the observed index return autocorrelation.[2]

The transaction cost hypothesis, as formulated in Cohen et al. (1980) and Mech (1993), conjectures that stocks periodically trade at prices that do not reflect all available information, the reason being that transaction or information costs make additional cross-security price analysis non-profitable. This may be a reasonable assumption in a specialist or market maker environment where most traders only trade in a single security. However, if speculators trade in several securities, the effects from imperfect information should be small. Also, order submission and withdrawal are in fact close to costless on automated exchanges.

The advocates of time-varying expected returns argue that risk premia can be time-varying, following predictable patterns, typically mean reversion, and that, as a result, observed returns will be positively autocorrelated (See, for example, Campbell et al., 1993). However, under time-varying risk premia the same autocorrelation is expected to be visible in related asset returns such as individual stock returns or index futures returns. However, both these asset classes have consistently proved *not* to have serially correlated short-term returns.[3]

In addition, the shorter the return frequency considered, the less likely it is that time variation in expected returns is caused by time variation in risk premia. In fact, the observed autocorrelation in intraday index returns is usually strong enough to generate *negative* expected returns. Since plausible expected returns must be higher than the risk-free rate, time-varying expected returns could possibly explain autocorrelation in long-term returns (weekly, monthly, etc.). However, time-varying risk premia do not provide a sensible explanation for the behaviour of intraday returns.

---

[1] See Boudoukh et al. (1994).

[2] See, e.g., Atchison et al. (1987), Berglund and Liljeblom (1988), Lo and MacKinlay (1990a), Mech (1993).

[3] This observation was probably first made by Boudoukh et al. (1994). If individual stock returns are autocorrelated, they are normally *negatively* autocorrelated, probably as a result of bid-ask bounce. That index futures returns have very low, if any, serial correlation is reported by, amongst others, Stoll and Whaley (1990), Chan et al. (1991) and Chan (1992).

In contrast to these earlier models of index return autocorrelation, this paper uses a model of cross-security information aggregation first developed in Säfvenblad (1997). Trade is modelled as taking place in separate, simultaneous call auctions of the Hellwig (1980) type. There are no transaction costs, all stocks trade at efficient prices, and speculators may trade in all securities. The only restriction used to generate the theoretical results is the time and information constraint imposed by a simultaneous call auction procedure.

Each equilibrium transaction price will reveal information, i.e. a signal, to the rest of the market. As stocks trade separately, resulting prices may be inefficient with regard to the other realised prices, simply because other prices were not observable when demand schedules were formulated. A large part of the lead-lag relation can be captured using the index level. When aggregate stock prices are known, the index level is known with greater precision and prices will be revised to take this into account. The autocorrelation in index returns results from this price revision.

Any cross-stock price inefficiencies will, of course, be corrected in subsequent trading. Therefore, the model's scope is limited to the analysis of short-term lead-lag effects and information inefficiencies. The return horizons considered ranges from minutes to hours. Questions related to long-term returns, such as autocorrelation in weekly and monthly stock index returns, require other modelling approaches.

The paper is organised as follows. In the next section (section 2) the model of cross-security information aggregation is presented and testable hypotheses with regard to index return autocorrelation are provided. Section 3 starts by discussing some earlier empirical evidence on index return autocorrelation, and continues by testing the model using data from the Paris Bourse. Section 4 summarises and concludes.

2 A model of cross-security information aggregation

2.1 Background

The model of cross-security information aggregation used in this paper was originally developed in Säfvenblad (1997) to adapt the market maker model of Chan (1993) to a noisy rational expectations equilibrium (REE) auction market.[4] This implies that the model can serve as multi-security extension to several REE models, including Grossman (1976), Grossman and Stiglitz (1980), Hellwig (1980) , Glosten and Milgrom (1985) and Kyle (1985, 1989).[5]

The stochastic environment of the Chan model is extended by introducing cross-security correlation in revealed information. Säfvenblad (1997) shows that this richer information structure is needed to model open-to-close returns on the Paris Bourse. It is also a necessary condition for negative index return

[4]The Chan (1993) model is based on Kyle (1985). Information is modelled as the net order flow observed by a stock specialist.

[5]These, and other REE models, share a large number of basic properties. The similarities are demonstrated in several papers, including Krishnan (1992), Paul (1994), Rochet and Vila (1994), Sarkar (1994) and Vives (1995).

autocorrelation, and allows for common cross-security trading strategies such as index arbitrage trading.

The model is a two-stage mechanism for cross-security information aggregation within a market with many securities. In the first stage, an equilibrium price is reached for each security in an auction market of the Hellwig (1980) type. In the second stage, prices are realigned across securities using information revealed in first stage prices. As a result, observed returns are autocorrelated, following an MA(1)-process with, depending on parameters, positive or negative autocorrelation.

The model clearly shows why index returns are mostly positively autocorrelated. As all stock prices are noisy, their response to new information is proportional to the precision of the information. For an individual stock, the signal precision is low. However, if a large number of prices are observed, these can be used to construct a less noisy estimate of the market factor. The improved precision of the market level estimate justifies a stronger response. As stocks react symmetrically to the new information, index return autocorrelation results.

Normally, the index return autocorrelation is positive, but an important feature of the model is that it can also generate negatively autocorrelated index returns. This will be the case when the error in revealed information has higher cross-security correlation than the error in the prior valuation of securities. Under such circumstances, index returns rather than cross-security information will be used to identify price errors. The result is negatively autocorrelated index returns.

The model's implications are not limited to index return autocorrelation. In fact, the model may also be used whenever there is *simultaneous* information revelation in several securities or in several markets. Some examples are the lead-lag effects between stocks and stock options, between index futures and cash index returns, between prices for the same asset on two exchanges. Säfvenblad (1997) specifically analyses the model's implications for cross-autocorrelation among individual stock returns, and the lead-lag relation between stock returns and stock option returns.

In this model, security prices reflect all historical information, but not information made available simultaneously (or later) in other securities. Individual securities will therefore exhibit a delayed response *only* to information revealed simultaneously in *prices* of other securities. Public information releases, or other information not revealed in prices, will be reflected in all prices simultaneously, and therefore do not result in index return autocorrelation.[6]

In contrast to both the nonsynchronous trading and the transaction cost hypotheses, there is no lagged response to public information. All prices are efficient and react instantly to public information. The index return autocorrelation is not a result of lagging returns, but of causality. Realised prices *cause* a revaluation of all other securities. When securities trade simultaneously the causality will be symmetric and reciprocal, resulting in index return autocorrelation.

Although returns are cross-autocorrelated, and thus predictable, the price inefficiency cannot be used to make trading or arbitrage profits, since price

[6]Public information may, however, generate increased uncertainty about the current value of stocks. If this uncertainty is resolved in trading, autocorrelated index returns may result. Compare with the empirical tests in section 3.5.3.

revisions are predicted by all agents. Prices will therefore adjust *without* trading.

Several authors have discussed the questions of price informativeness and information acquisition in the context of REE models.[7] Those questions however, are not addressed in this paper. Price informativeness is taken to be exogenously determined, but the results of the model have clear implications for information acquisition. As market-wide information is imperfectly reflected in stock prices, it is relatively more profitable to trade on market-wide information in this setting than in the Admati (1985) framework.[8]

The model is best adapted to a simultaneous opening call auction in an electronic limit order book market, such as the Paris Bourse or the Toronto Stock Exchange. Under this trading arrangement, no (or very little) cross-stock information is available until *after* the morning call, and only then can prices adjust to the information revealed in opening prices. Clearly, prices can adjust to any information available *before* the opening procedure.

2.2 Basic model

In the economy, N securities trade separately in simultaneous call auction markets of the Hellwig (1980) type. Each security is a claim to an unobservable underlying value. The fundamental value can, for example, be interpreted as the value of all securities in the absence of private information. The N values are arranged in the $N \times 1$ vector $\mathbf{V}$.

There are two types of traders in the market, namely speculators and liquidity traders. Speculators are rational and profit maximising agents, some of whom have received or acquired information, "a signal," relative to the underlying value of securities. Liquidity traders trade for some other exogenous reason (e.g. hedging or liquidity constraints) and their demand is independent of the expected value of securities. Liquidity trading can be correlated across securities, but is assumed to be independent of past liquidity trading, value innovations and any private signals.

Before trading, agents share a normally distributed, noisy prior belief, $\mathbf{P}^*_{-1}$, about the underlying value of all securities:

$$\mathbf{P}^*_{-1} \ = \mathrm{E}\left[\mathbf{V} \mid \mathcal{F}_{-1}\right] \sim \mathrm{N}\left(\mathbf{V}, \mathbf{\Pi}\right), \tag{1}$$

where $\mathcal{F}_{-1}$ is the public information set available before trading. The covariance matrix of the common prior, $\mathbf{\Pi}$, is known by all agents.

Agents calculate optimal demand schedules using any private information and the equilibrium covariance structure of signals and returns. Standard assumptions (normal distribution and exponential utility over next period's wealth) provide optimal demand schedules that are linear in price.

Demand schedules for individual securities are collected by a Walrasian auctioneer who sets a price vector $\mathbf{P}$ that clears supply and demand for all stocks simultaneously.

Relying on standard REE results, it is known that each price realised in trading reveals a signal, F_i for each stock.[9] Here, the signal is modelled as a

[7] See, e.g., Grossman and Stiglitz (1980) and Holden and Subrahmanyam (1992).

[8] For a formal development in a market maker environment of the Kyle (1985) type, see Shin and Singh (1996).

[9] See, e.g., Hellwig (1980), proposition 5.2. It also follows directly from the martingale property of prices.

Table 1: Sequence of events and information in the model of cross-security information aggregation

Event	Description
A	All agents share a common prior, $\mathbf{P}^*_{-1}$, which is a normally distributed measurement of underlying values $\mathbf{V}$ with covariance matrix $\mathbf{\Pi}$
B	Each agent submits N, optimally calculated, linear demand schedules to a Walrasian auctioneer. Submitted demand schedules are not revealed to other traders.
Trading	The Walrasian auctioneer simultaneously sets market clearing prices, P_i, in all N markets. Orders are executed immediately.
C	Agents observe the realised price vector, $\mathbf{P}$.
D	Agents use the equilibrium noisiness of prices to deduce a signal, $\mathbf{F}$, from realised prices, $\mathbf{P}$. The signal has covariance matrix $\mathbf{\Phi}$.
E	Agents calculate posterior beliefs $\mathbf{P}^*$ of $\mathbf{V}$ using Bayesian updating and the revealed signal $\mathbf{F}$.

Sequence of events: A $\rightarrow$ B $\rightarrow$ Trading $\rightarrow$ C $\rightarrow$ D $\rightarrow$ E. All agents know the information structure of underlying returns, that is, the true value of $\mathbf{\Phi}$ and $\mathbf{\Pi}$. They also know the precision of their own, and other agents' signals. Transaction costs are zero.

noisy measurement of the error in the prior valuation. For all stocks, vectors and matrices are used to write

$$\mathbf{F} \sim \mathrm{N}\left(\mathbf{V} - \mathbf{P}^*_{-1}\ ,\ \mathbf{\Phi}\right). \tag{2}$$

For an individual stock we write

$$F_i \sim \mathrm{N}\left(V_i - P^*_{i,-1}\ ,\ \Phi_{ii}\right), \tag{3}$$

where Φ_{ii} is the ith diagonal element of the covariance matrix $\mathbf{\Phi}$. Intuitively, the aggregate signal can be seen as a weighted sum of all investors' private information distorted by the extent of liquidity trading.[10]

The price in each of the N separate markets can be represented by the following equation of Bayesian updating:

$$P_i = \frac{\Phi_{ii}}{\Pi_{ii} + \Phi_{ii}} P^*_{i,-1} + \frac{\Pi_{ii}}{\Pi_{ii} + \Phi_{ii}} \left(F_i + P^*_{i,-1}\right). \tag{4}$$

Equation 4 holds for all *competitive* single-security REE models, i.e., in all cases where realised prices are unbiased predictors of the underlying value or future price sequence. Otherwise the realised prices must be adjusted for the predictable part of future returns.

Although P_i reflects information available in F_i, it does not reflect all information available in the full signal vector, $\mathbf{F}$. Therefore, stock prices, and the index level, will be adjusted to take this information into account. First, define κ_i as the price's responsiveness to new information in F_i:

$$P_i = P^*_{i,-1} + \kappa_i F_i \qquad \kappa_i = \frac{\Pi_i}{\Pi_i + \Phi_i}. \tag{5}$$

[10] See Admati (1985) for a formal derivation and Säfvenblad (1997) for a short discussion.

Rewrite equation 5 to vector and matrix notation by arranging the κ_i:s in the diagonal matrix, $\hat{\mathbf{\Omega}}$:

$$\hat{\mathbf{\Omega}} = \begin{bmatrix} \kappa_1 & 0 & \cdots & 0 \\ 0 & \kappa_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \kappa_N \end{bmatrix}. \tag{6}$$

The realised price can then be written as the prior plus the signal vector, $\mathbf{F}$, premultiplied by $\hat{\mathbf{\Omega}}$:

$$\mathbf{P} = \mathbf{P}^*_{-1} + \hat{\mathbf{\Omega}}\mathbf{F}. \tag{7}$$

As price and signals are normally distributed, standard Bayesian theory can be used to calculate a closed form solution for the posterior, $\mathbf{P}^*$, given the extracted signal vector **F:**[11]

$$\mathbf{P}^* = \mathrm{E}\left[\mathbf{V} \,|\mathcal{F}_{-1},\ \mathbf{F}\ (\text{or } \mathbf{P})\right] = \mathbf{P}^*_{-1} + \mathbf{\Omega}\mathbf{F}, \tag{8}$$

$$\mathbf{\Omega} = \mathbf{\Pi}\left(\mathbf{\Pi} + \mathbf{\Phi}\right)^{-1}. \tag{9}$$

The posterior is linear in $\mathbf{F}$ with $\mathbf{\Omega}$ as an updating matrix which maps N signals into N efficient prices.

Two sets of returns are defined. The first stage returns, $\mathbf{r}$, are calculated as the difference between recorded prices and the prior:

$$\mathbf{r} = \mathbf{P} - \mathbf{P}^*_{-1} = \hat{\mathbf{\Omega}}\mathbf{F}. \tag{10}$$

Secondly, posterior returns, $\mathbf{r}^*$, which take all information in $\mathbf{F}$ into account are defined. Posterior returns are thus simply the difference between posterior and prior valuation:

$$\mathbf{r}^* = \mathbf{P}^* - \mathbf{P}^*_{-1} = \mathbf{\Omega}\mathbf{F}. \tag{11}$$

It is easy to see that returns will be cross-autocorrelated whenever $\mathbf{P}^* \neq \mathbf{P}$ or, equivalently, when $\mathbf{r}^* \neq \mathbf{r}$. The cross-autocorrelation results because the price adjustment from the observed price, $\mathbf{P}$, to the posterior valuation, $\mathbf{P}^*$, depends on earlier returns, which are used to extract information about $\mathbf{F}$:

$$\mathbf{P}^* - \mathbf{P} = \mathbf{r}^* - \mathbf{r} = \left(\mathbf{\Omega} - \hat{\mathbf{\Omega}}\right)\mathbf{F} = \left(\mathbf{\Omega}\hat{\mathbf{\Omega}}^{-1} - \mathbf{I}\right)\mathbf{r}, \tag{12}$$

where $\mathbf{I}$ is an $N \times N$ identity matrix. This price adjustment is the core of the cross-security information aggregation model. The adjustment return of security i is a weighted sum of "unexpected" returns on all other securities.[12] The weights are determined by the off-diagonal elements in the updating matrix $\mathbf{\Omega}$, normalised by the strength of securities' initial response to information κ_i. We can therefore write:

$$r_i^* - r_i = \sum_{j=1}^{N} \frac{\omega_{ij}}{\kappa_j}\left(r_j - \mathrm{E}\left[r_j | \mathcal{F}_{-1}, r_i\right]\right), \tag{13}$$

where ω_{ij} is the jth element on row i in the matrix $\mathbf{\Omega}$, and κ_j is the jth diagonal element in $\hat{\mathbf{\Omega}}$.

[11] In the more general case, such an explicit solution may not be available; in addition, the optimal updating rule need not always be linear.

[12] From the equilibrium condition, it follows that the expected value of $P_i^* - P_i$, conditional on the realised r_i, is zero. It can be shown that equation 13 satisfies this condition.

2.3 Index returns

Let r_m denote the first stage return of an equally weighted stock index. The index return can be written as the weighted average of the information revealed in each stock:

$$r_m = \frac{1}{N}\mathbf{1}'\mathbf{r} = \frac{1}{N}\mathbf{1}'\hat{\mathbf{\Omega}}\mathbf{F} = \frac{1}{N}\sum_{i=1}^{N}\kappa_i F_i, \tag{14}$$

where $\mathbf{1}$ is an $1 \times N$ column vector of ones. The weights of individual signals are determined by the κ_i:s, the stocks' first stage response to the revealed information in the stock proper. The information revealed in stocks with precise signals (high κ_i) will therefore be more strongly reflected in stock prices and the index level.

Also define the posterior index returns, r_m^* as the difference between the prior and posterior index level:

$$r_m^* = \frac{1}{N}\mathbf{1}'\mathbf{r}^* = \frac{1}{N}\sum_{i=1}^{N} r_i^* = \frac{1}{N}\mathbf{\Omega}\mathbf{F}. \tag{15}$$

The posterior returns, unlike the first stage returns, cannot be rewritten as a sum of signals, since all posterior returns depend on all signals with varying weights, determined by $\mathbf{\Omega}$.

However, from Bayesian theory it is known that for an average of individual values, the average of individual signals is the most efficient aggregate signal. Therefore, an index level signal can be defined as the average of individual stock signals:

$$F_m = \frac{1}{N}\mathbf{1}'\mathbf{F}. \tag{16}$$

Similarly, the index level prior can be defined as:

$$P_{m,-1}^* = \frac{1}{N}\mathbf{1}'\mathbf{P}_{-1}^*. \tag{17}$$

Since the stochastic properties of signals and priors are well-known, it is easy to calculate the variance of the index level prior and signal. Denote the variances by π_m and ϕ_m, respectively:

$$P_{m,-1}^* \sim \mathrm{N}\left(V_m, \pi_m\right), \tag{18}$$

$$\pi_m = \frac{1}{N^2}\mathbf{1}'\mathbf{\Pi}\mathbf{1}, \tag{19}$$

$$F_m \sim \mathrm{N}\left(V_m - P_{m,-1}^*, \phi_m\right), \tag{20}$$

$$\phi_m = \frac{1}{N^2}\mathbf{1}'\mathbf{\Phi}\mathbf{1}, \tag{21}$$

where V_m is an equally weighted index of underlying values. From the definitions above, it follows that π_m is the average of all N^2 elements in $\mathbf{\Pi}$. It will thus be close to the average cross-security correlation in the prior valuation of component stocks.

Likewise, the variance of the market signal, ϕ_m, is approximately equal to the average cross-security correlation in revealed information. Intuitively, it can be seen as the variance of the "market mood," unfounded optimism or pessimism, or just the covariance of liquidity trading across securities.

Using the above development, the optimal market response to information, r_m^*, can be expressed as a constant, κ_m^*, multiplied by the market signal,

$$\begin{aligned} r_m^* &= \kappa_m^* F_m, & (22) \\ \kappa_m^* &= \frac{\pi_m}{\pi_m + \phi_m}. & (23) \end{aligned}$$

The parameter κ_m^* measures how agents' beliefs react to new market-wide information F_m. If signals are only weakly correlated across securities (ϕ_m small), it is possible to know the index level with high precision when the number of securities is large. In this case, κ_m^* will be close to unity.

Now, define a parameter κ_m as a parallel to κ_m^*, measuring the first stage response to index level information, in order to compare first stage and posterior returns:

$$\kappa_m = \frac{\frac{1}{N}\sum_{i=1}^{N} \kappa_i F_i}{F_m} \approx \frac{1}{N}\sum_{i=1}^{N} \kappa_i. \tag{24}$$

In general, κ_m will be approximately equal to the average of κ_i:s. When $\kappa_m \neq \kappa_m^*$ index returns will be autocorrelated, following an MA(1)-process. If $\kappa_m < \kappa_m^*$ the market return underreacts to new information, resulting in positive autocorrelation. On the other hand, if $\kappa_m > \kappa_m^*$ the market overreacts to new information resulting in negative autocorrelation.

Whether index return is positive or negative is determined by the off-diagonal elements in $\mathbf{\Omega}$. If they are "mostly" positive, index returns will be positively autocorrelated; if they are mostly negative, index returns will be negatively autocorrelated. The intuition behind this result will be made clearer in the next section.

2.4 A one-factor model

A direct and simple way to analyse the model's implications is to set up a "one-factor" model, where priors and signals have both a market component and an individual stock component. For an individual security, assume that the prior has the structure

$$\Pi_{ij} = \begin{cases} \pi^m + \pi^s & \text{if } i = j \\ \pi^m & \text{if } i \neq j \end{cases} \quad \forall i, j, \tag{25}$$

where π^m is the variance of the market level prior and π^s is the additional variance for individual securities. π^s is equal for all securities. The covariance matrix of the prior priors can be visualised as:

$$\underset{N\times N}{\mathbf{\Pi}} = \begin{bmatrix} \pi^m + \pi^s & \pi^m & \cdots & \pi^m \\ \pi^m & \pi^m + \pi^s & \cdots & \pi^m \\ \vdots & \vdots & \ddots & \vdots \\ \pi^m & \pi^m & \cdots & \pi^m + \pi^s \end{bmatrix}. \tag{26}$$

Let the revealed information have a similar structure with the variance of the market signal ϕ^m, and the additional variance of individual stock signals, ϕ^s:

$$\Phi_{ij} = \begin{cases} \phi^m + \phi^s & \text{if } i = j \\ \phi^m & \text{if } i \neq j \end{cases} \quad \forall i, j. \tag{27}$$

Using this simplified structure, it is possible to calculate explicit returns. The returns in excess of the prior, $\mathbf{r}$, are simply κ^s, equal for all stocks, multiplied by the revealed signal, $\mathbf{F}$:

$$\mathbf{r} = \kappa^s \mathbf{F}, \tag{28}$$

with

$$\kappa^s = \frac{\pi^m + \pi^s}{\pi^m + \pi^s + \phi^m + \phi^s}. \tag{29}$$

In the first stage of trading, the index level reacts to market-wide information *exactly as* individual stocks react to stock specific information. The realised index return r_m, is therefore the same constant, κ^s, multiplied by the aggregate signal:

$$r_m = \kappa_m F_m, \tag{30}$$

$$\kappa_m = \kappa^s = \frac{\pi^m + \pi^s}{\pi^m + \pi^s + \phi^m + \phi^s}. \tag{31}$$

However, the optimal response of the market level to the same information is different:

$$r_m^* = \kappa_m^* F_m, \tag{32}$$

$$\kappa_m^* = \frac{N\pi^m + \pi^s}{N\pi^m + \pi^s + N\phi^m + \phi^s}. \tag{33}$$

Index returns will be autocorrelated whenever $\kappa_m \neq \kappa_m^*$. Some necessary conditions for index return autocorrelation are immediately visible from equation 31 and 33. There must be several securities ($N > 1$) and prior or signals must be correlated across securities ($\pi^m \neq 0$, or $\phi^m \neq 0$). Index return autocorrelation will be positive if $\pi^m/\pi^s > \phi^m/\phi^s$, that is, when the prior has higher cross-security correlation than the revealed signals. Consequently, provided that $\kappa_m/\kappa_m^* > 0.5$, autocorrelation increases in π^m and ϕ^s and decreases in ϕ^m and π^m.

If signals are more strongly cross-correlated than underlying returns ($\pi^m/\pi^s < \phi^m/\phi^s$) the observed index returns will be negatively autocorrelated. If signal noise is strongly correlated across securities, any common return movements are more likely to be the result of noise than of underlying returns. Therefore, prices will react negatively to any common price movement.

2.5 Implications for a market with continuous trading

The formal model rules out continuous trading, but it is possible to adapt the model to the continuous trading case. In the case of frictionless trading, prices can and will react instantly to new information. Any cross-security price error is eliminated immediately and index return autocorrelation would be observed only over infinitely short time intervals. The model therefore approximates the Admati (1985) model.

The model will still have some effect under continuous, but *nonsynchronous* trading. As stocks trade at irregular intervals, there will be a delay in information about the market factor that adds to the delayed reaction imposed by the nonsynchronicity itself. However, in such a model, the effect of delayed information will be relatively small compared with the effect of nontrading. Also, the index level innovation will be relatively well known after only a small number of securities have traded (say 10–20); the additional information from the remaining (normally 100+) securities is small.

Therefore, measured index return autocorrelation would mostly be attributed to nonsynchronous trading effects. However, the model can generate significantly higher estimates of index return autocorrelation if there is a time restriction on information transmission. Under an explicit information restriction, autocorrelation will be highest when intraday volatility is high, that is, when cross-security price errors can be expected to be large.

2.6 Autocorrelation in intraday returns

Financial markets normally exhibit strong "U-shapes" during the trading day, with both volatility and trading volume at their highest at the opening and closing.[13] Therefore, nonsynchronous trading should add less to index return autocorrelation in early morning and late afternoon trading. However, this is clearly not the case. As shown by McInish and Wood (1991), autocorrelation can even be *higher* during periods of active trading. As transaction costs can be expected to be small when trading volume is high, this is also contrary to the transaction cost hypothesis of Mech (1993).

From the perspective of cross-security information aggregation, this result is not at all surprising. The high volatility and trading volume immediately before closing make transitory cross-security price inaccuracies more probable. As the high volatility is coupled with high trading volume, there is also a real possibility that traders suffer from information overload.

The model not only explains the results of McInish and Wood (1991), but also provides a good explanation for the high observed index return autocorrelation in daily data. Most studies use closing prices to calculate returns. Closing prices are less subjected to nonsynchronous trading but realised in a period of very high volatility. Therefore, cross-security information aggregation could account for the large difference between the autocorrelation expected from nonsynchronous trading and observed levels of autocorrelation.

2.7 Testable hypotheses

The model of cross-security information aggregation provides two main testable predictions of index return autocorrelation. Firstly, autocorrelation increases in the ratio π^m/π^s, that is, the variance of the index level prior divided by the average additional variance of the prior of individual stock prices, or, equivalently, the level of cross-security correlation in the prior.

Secondly, autocorrelation decreases in the ratio ϕ^m/ϕ^s, the level of index level signal noise divided by the additional noise in individual stock signals. This ratio can also be interpreted as the level of cross-security correlation in revealed information.

Index return autocorrelation will thus be strong if there is high uncertainty about the true index level and much security-specific noise in revealed information. On the other hand, index return autocorrelation will be low if uncertainty

[13] See Chan et al. (1991) and Chan (1992), for some empirical evidence and Admati and Pfleiderer (1988) and Foster and Viswanathan (1993) for a theoretical discussion.

about individual stock values is high and revealed information is strongly correlated across securities. In the case of strong cross-security correlation in revealed information, index return autocorrelation may be negative.

An econometrician cannot observe the covariance structure of the prior valuation. However, the covariance of the prior can be estimated relatively easily, since the error in the prior largely consists of value innovations since the last trade. If the correlation structure of value innovations can be believed to be constant, it can be estimated using the correlation in realised long-term returns. Therefore, index return autocorrelation will tend to be higher for portfolios of highly correlated stocks.

When trading is closed, no information is revealed in trading, but both value innovations and new information will be revealed outside trading. Depending on the proportions between new information and value innovations, the prior may become both more and less noisy when trading is closed. Although information events, such as scheduled macroeconomic announcements, may reduce the noise in the index level, nontrading will generally increase noisiness of the index level prior. This effect will be strong at Monday open, when index level noise has accumulated over two nontrading days.

However, overnight developments can also enhance (at least relatively) the precision of the index level prior and thus reduce autocorrelation. This will be the case when there is new information revealed to the market during the night. Overall prior precision will improve, leading to less autocorrelation when the market opens. The empirical test of section 3.5.3 uses changes in US interest rates that are assumed to relatively improve the precision of the index level prior. Similar results would be expected from, for example, earnings announcements and exchange rate changes.

Any index level innovation revealed outside trading will not result in index return autocorrelation. In the model, such information will enter directly into agents' prior valuation before the next round of trading. This implies that changes in the index level that are the result of macro announcements, interest rate changes and so on, should be reflected in stock prices faster than market-wide information that is revealed through trading, primarily reflecting changes in investor valuations or preferences.

The index level prior will be particularly noisy when the index level volatility is high. Using various volatility estimates, it is relatively straightforward to identify when the index prior is more noisy than otherwise. As an example, we know that volatility exhibits a U-shape over the trading day, and we therefore expect that autocorrelation will also be U-shaped. This prediction is thus consistent with the empirical evidence reported by McInish and Wood (1991).

It is much harder to measure the correlation in revealed signals. As these are derived from realised prices, they cannot be used to explain return patterns. In order to capture the cross-security correlation in signals, it is necessary to use other data besides prices. For intraday returns, an example is the index arbitrage trading, which will induce cross-security correlation in revealed information. It may also be possible to analyse the order book movements at opening and closing to identify index arbitrage trading. For the daily data used in this paper, the possibilities are limited to theoretical arguments. Section 3.6 tests an argument based on short-selling restrictions.

Table 2: Selected empirical evidence on index return autocorrelation

Source	Series	Sample period	Return frequency	First auto-correlation
Campbell et al. (1993)	CRSP^{vw}	1950-62	daily	0.13**
		1962-74	daily	0.28**
		1975-87	daily	0.17**
Atchison et al. (1987)	CRSP^{ew}	1978-81	daily	0.17**
	CRSP^{vw}	1978-81	daily	0.31**
Lo and MacKinlay (1990b)	CRSP^{ew} *small stocks*	1962-87	daily	0.35**
	CRSP^{ew} *large stocks*	1962-87	daily	0.17**
Berglund and Liljeblom (1988)	Helsinki^{vw}	1977-82	daily	0.49**
McInish and Wood (1991)	NYSE^{ew} open	1984-85	daily	0.15*
	NYSE^{ew} midday	1984-85	daily	0.06
	NYSE^{ew} close	1984-85	daily	0.27**
Abhyankar (1995)	FT-SE 100	1986-90	60 min.	0.14**
Stoll and Whaley (1990)	S&P 500	1982-87	5 min.	0.51**
	MMI	1984-86	5 min.	0.24**
Chan et al. (1991)	S&P 500	1984-85	5 min.	0.45**
	S&P 500	1988-89	5 min.	0.27**
Chan (1992)	MMI	1984-85	5 min.	0.31**
Abhyankar (1996)	FT-SE 100	1992	5 min.	0.48**

Significance levels as reported or calculated from reported standard errors. ^{ew}Equally weighted index. ^{vw}Value-weighted index. **/*/° Significantly different from zero at the 0.01/0.05/0.10 level.

3 Empirical evidence

3.1 Some earlier empirical evidence

Table 2 presents a selection of published evidence on index return autocorrelation, and clearly shows that, index returns are positively autocorrelated, for most return frequencies and markets. As discussed in section 2.5, the results of McInish and Wood (1991) provide support for the model, documenting high index return autocorrelation under high trading activity.

Similar results are also reported by Chan (1992), who finds that cash index returns lag index futures returns more strongly when the trading intensity is high (the marginal impact is small, but statistically significant). Similar to the results of McInish and Wood (1991), this implies higher index return autocorrelation when trading is active, contrary to the predictions of nonsynchronous trading. Chan also shows that the futures lead is stronger when there are large changes in the index level. This result is also consistent with cross-security information aggregation. High index return volatility implies a combination of a noisy index prior (π^m/π^s high) and a precise index signal (ϕ^m/Φ^s low), both leading to high index return autocorrelation.

3.2 Choice of data

In continuous trading, cross-security information aggregation and nonsynchronous trading have similar implications for observed cross-autocorrelation. It is therefore important to test for cross-security information aggregation in a setting with minimal nontrading. In addition, the physical trading arrangements should be as close to the theoretical model as possible.

Trading at the opening call auction at the Paris Bourse fulfils both these criteria. There is no nonsynchronicity in recorded prices and the trading arrangements are very close to the theoretical model. Data was kindly made available by SBF–Paris Bourse. The dataset provides opening prices, closing prices and trading volume for all stocks and other instruments traded through the CAC electronic trading system. The sample period is five years (1991–1995), comprising 1302 daily observations.[14]

During the sample period, the electronic order book is opened for order submission at 08:30 (09:00 until 1992). The orders are accumulated until the opening call at 10:00, when matching orders are executed at the price that maximises the number of shares traded.[15] This price is recorded as the opening price in the dataset. Approximately 5% of the daily trading volume is executed at opening prices.

To minimise problems of low liquidity, the sample is restricted to the 70 most traded stocks on the monthly settlement list (*Reglement Mensuel*). All selected stocks have an average daily trading volume of at least 5 million FRF per day during the sample period. Summary statistics on all the individual stock series used are reported in table 8 (p. 25–26).[16]

The closing price is realised in continuous trading and is the last price at which a transaction is executed (trading closes at 17:00). As trading is very active during the last minutes of the day, the average nontrading is only a few seconds for sample stocks. Any return spill-over from nonsynchronous trading should thus be negligible.

With the exception of the exact simultaneity in the SBF dataset, the data is thus similar in character to the NYSE data used by Amihud and Mendelson (1987).

3.3 Portfolio construction

As a measure to further reduce effects of low liquidity or other measurement errors, three portfolios of 23-24 stocks are created on the basis of trading volume (*High Volume*, *Medium Volume* and *Low Volume*). The portfolio containing the most liquid stocks, *High Volume*, contains only very liquid stocks. The time series results for this portfolio should thus be considered most robust.

To test hypotheses relative to the cross-security correlation of the prior cross-sectionally, three correlation sorted portfolios of 23-24 stocks are created (*High Correlation*, *Medium Correlation* and *Low Correlation*, see section 3.5.1 for the test). Stocks were ranked according to the average correlation in monthly returns between the stock and all other stocks. The *High Correlation* portfolio has very high cross-security correlation in monthly returns (*0.499*), and is thus created from securities with low idiosyncratic risk. A seventh portfolio, *All*

[14] Only stocks with more than 1000 trading days during the sample period were considered.

[15] There is, however, one important difference from the model setup. During the preopening stage, an indicative price is available to the market. Biais et al. (1996) study the information content at the preopening stage using mainly a single-security perspective. They show that preopening prices are not very informative. Most limit orders are submitted in the final minutes before opening. Detailed accounts of the trading structure are also found in Biais et al. (1995).

[16] The same data is used in Säfvenblad (1997).

Table 3: Summary statistics for stock portfolios

Portfolio	Autocorrelation at open	Autocorrelation at close	Cross-security correlation in monthly returns§	Trading volume‡
High Volume	0.066	0.157**	0.336	87.4
	(0.042)	(0.022)	(0.187)	(1.3)
Medium Volume	0.023	0.052*	0.212	21.4
	(0.043)	(0.026)	(0.133)	(0.3)
Low Volume	0.185**	0.045	0.227	8.6
	(0.054)	(0.033)	(0.142)	(0.1)
High Correlation	0.070	0.123**	0.499	57.4
	(0.043)	(0.025)	(0.077)	(0.9)
Medium Correlation	0.151**	0.097**	0.283	24.6
	(0.048)	(0.032)	(0.089)	(0.3)
Low Correlation	0.080°	0.066**	0.079	35.8
	(0.042)	(0.024)	(0.096)	(0.5)
All Stocks	0.150**	0.122**	0.248	39.1
	(0.043)	(0.027)	(0.156)	(0.5)

Regressions use least squares estimation with asymptotic GMM standard errors (in parentheses) that are robust to heteroskedasticity (Hansen, 1982). §Cross-sectional standard errors in parentheses. ‡Average per day, per stock, trading volume in million FRF. Standard deviation across trading days in parentheses. **/*/° Significantly different from zero at the 0.01/0.05/0.10 level.

Stocks, contains all 70 stocks in the sample. Opening prices, closing prices and daily transaction volume are calculated for each portfolio. Summary statistics for the portfolios are reported in table 3.

3.4 Methodology

In order to use both opening and closing prices, two types of returns are calculated. Overday returns are calculated as the log difference between opening and closing prices:

$$r_{i,t}^{\text{day}} = \log\left(P_{i,t}^{\text{close}}\right) - \log\left(P_{i,t}^{\text{open}}\right). \tag{34}$$

Overnight returns are similarly measured from close to open:

$$r_{i,t}^{\text{night}} = \log\left(P_{i,t}^{\text{open}}\right) - \log\left(P_{i,t-1}^{\text{close}}\right). \tag{35}$$

Overnight returns are dated with the day when the return period ends. For example, Monday overnight return measures the return from Friday close to Monday open. Using these two types of return observations, autocorrelation at open is estimated using the regression model

$$r_{i,t}^{\text{day}} = \beta_0 + \beta_1 r_{i,t}^{\text{night}} + \varepsilon_{i,t}, \tag{36}$$

while autocorrelation at close is calculated using the regression model

$$r_{i,t}^{\text{night}} = \beta_0 + \beta_1 r_{i,t-1}^{\text{day}} + \varepsilon_{i,t}. \tag{37}$$

Regressions use least squares estimation with heteroskedasticity consistent GMM standard errors (Hansen, 1982). Regressions do not exclude or control for outliers.

3.5 Testing effects from the variance of the index level prior

This section presents five separate tests of the prediction that index autocorrelation increases in the variance of the index level prior (π^m in the theoretical model).

3.5.1 Highly correlated return series

We start by testing whether portfolios of highly correlated securities exhibit higher return autocorrelation. Parts of the error in the prior result from value innovations. If innovations are strongly correlated across securities, so will the errors of prior estimates. Industry portfolios are thus expected to exhibit stronger autocorrelation than "mixed" portfolios.[17]

At open the hypothesis is supported for *Medium Correlation* and *Low Correlation* but rejected for *High Correlation* (results in table 3). *High Correlation* has lower autocorrelation than both the other portfolios (*0.070* versus *0.151*, *0.080*). The most probable reason for this rejection is the significantly higher liquidity of stocks in *High Correlation*. Trading volume of stocks in *High Correlation* is about twice that of stocks in *Medium Correlation* and *Low Correlation*.

Given the results at open, results at close are surprisingly well in line with predictions. The autocorrelation increases in the level of cross-security correlation (*0.066*, *0.097*, *0.123*) and the difference between *High Correlation* and *Low Correlation* is statistically significant. This result must be considered particularly strong as the high liquidity of stocks in *High Correlation* should tend to reduce autocorrelation.

3.5.2 A Monday effect

As a second test, we analyse day-of-the-week effects on autocorrelation. The index prior can be expected to be particularly noisy when markets open after the weekend, that is, at Monday open. Private information and other uncertainty have then accumulated during two nontrading days. The highest autocorrelation should thus be observed at the Monday open. For closing returns, day-of-the-week effects should be less pronounced as the closing is always preceded by a full trading day during which index level uncertainty can be reduced to "normal" levels.

Both these conjectures are supported by the results presented in table 4. At close, the null of all days having the same autocorrelation cannot be rejected for any of the portfolios. However, at open there is a strong, significantly positive, Monday effect in all stock portfolios.[18]

[17] Using industry portfolios however, is not a good way to test this proposition for two reasons. Firstly, to obtain a reasonable number of securities, industry portfolios must include less liquid stocks. Secondly, the average correlation between same-industry shares is not much higher than the average correlation among all stocks.

[18] The strong Monday effect motivates the use of a dummy for Monday open in the remaining regressions.

Table 4: Index return autocorrelation conditional on day-of-the-week

Portfolio	$\hat{\beta}_1$ *Mon.*	$\hat{\beta}_2$ *Tue.*	$\hat{\beta}_3$ *Wed.*	$\hat{\beta}_4$ *Thu.*	$\hat{\beta}_5$ *Fri.*	Wald
At Open						
High Volume	0.218^{**}	0.076	-0.066	-0.019	-0.001	8.1°
	(0.066)	(0.110)	(0.116)	(0.077)	(0.117)	
Medium Volume	0.173^{*}	0.090	-0.158	-0.090	0.049	10.1^{*}
	(0.080)	(0.109)	(0.103)	(0.063)	(0.092)	
Low Volume	0.328^{**}	0.156°	0.147	0.048	0.186^{*}	4.3
	(0.114)	(0.088)	(0.112)	(0.079)	(0.076)	
All Stocks	0.302^{**}	0.193°	-0.001	0.014	0.158	12.4^{*}
	(0.068)	(0.106)	(0.111)	(0.058)	(0.108)	
At Close						
High Volume	0.160^{**}	0.100^{*}	0.183^{**}	0.162^{**}	0.169^{**}	1.7
	(0.050)	(0.047)	(0.061)	(0.039)	(0.048)	
Medium Volume	0.041	0.032	0.105	0.061°	0.014	1.2
	(0.063)	(0.048)	(0.077)	(0.037)	(0.055)	
Low Volume	-0.074	0.040	0.079	0.108^{*}	0.066	4.4
	(0.078)	(0.049)	(0.099)	(0.043)	(0.051)	
All Stocks	0.094	0.084°	0.163°	0.145^{**}	0.114^{*}	1.4
	(0.060)	(0.051)	(0.085)	(0.037)	(0.054)	

Results are similar for the correlation-sorted portfolios (not reported). Trading volume is approximately equal across days of the week, with the exception of Mondays, when trading volume is approximately 25% lower than on other days (not reported). Regression model: $r_t = \beta_0 + (\beta_1 D_{1,t-1} + \ldots + \beta_5 D_{5,t-1})r_{t-1} + \varepsilon_t$. $D_{1,t}, \ldots, D_{5,t}$ are dummy variables for the day of the week (1=Monday). The time indices refer to return periods (overnight or overday). The Wald statistic tests the restriction $\beta_1 = \ldots = \beta_5$. $\chi^2(4)$ critical values: 13.2/9.4/7.7 at the 0.01/0.05/0.10 level. Regressions use least squares estimation with asymptotic GMM standard errors (in parentheses) that are robust to heteroskedasticity (Hansen, 1982). $^{**}/^{*}/^{\circ}$ Significantly different from zero at the 0.01/0.05/0.10 level.

3.5.3 High overnight volatility

Large overnight changes in foreign stock market values result in a noisier than usual index prior at opening. Although investors observe information about overnight events, they cannot judge the full impact on French stock values. As discussed earlier, this should result in higher autocorrelation at open.

The results presented in panel *a* of table 5 support this prediction. Autocorrelation at open is higher following large absolute index returns in overnight US trading for all portfolios, although the statistical significance for individual portfolios is weak.

Another important observation from table 5 is that index return autocorrelation is *negative* conditional on low overnight volatility (-0.078, -0.161, -0.077, -0.047). Although this result is not statistically significant, it shows that a comprehensive model of index return autocorrelation must be capable of modelling both positive and negative autocorrelation.

3.5.4 Improved index level precision

The argument above can also be extended to include new information about the value of French equities released during the night. As an overnight change in the US index level implies a probable change in the unobservable French "fundamentals", US index returns always increase the level of noise in the index level prior. However, a change in the US interest rates should have less influence on the fundamental values of French stocks, although it certainly affects the discounted value of these fundamentals.

There will consequently be relatively low uncertainty about the valuation of French stock after the realisation of large overnight interest rate changes. The change in interest rates provides an additional signal that can be used to update the common index level prior. Large large changes in interest rates should *lower* the uncertainty of the index prior and thus reduce autocorrelation at open.

The results presented in panel *b* of table 5 support this prediction. Autocorrelation at open is significantly lower, and in some cases even negative, following large absolute changes in US interest rates. The statistical significance for individual portfolios is not strong, but the point estimates indicate the same pattern for all portfolios.

3.5.5 High trading volume

As argued in the theoretical section, trading increases the precision of prices by releasing private information to the market.[19] If trading is intense, the index level will be a better estimate of the "true" index level than otherwise. Consequently, high trading volume should be associated with low autocorrelation, in particular at close.

The results reported in table 6 support the model's prediction. Autocorrelation at close is lower after days of high trading volume for all reported portfolios. For three out of four portfolios, the difference is statistically significant. There is also a weak spill-over of reduced autocorrelation at the following day's opening (not statistically significant, not reported).

[19] This has also been demonstrated empirically by, e.g., Amihud and Mendelson (1987).

Table 5: Index return autocorrelation at open conditional on overnight stock returns and interest rate changes

Portfolio	$\hat{\beta}_1$ *small*	$\hat{\beta}_2$ *large*	$\hat{\beta}_3$ *Monday*	Wald
Panel *a*: Conditional on index returns				
High Volume	-0.078	0.033	0.241^{**}	1.5
	(0.084)	(0.056)	(0.085)	
Medium Volume	-0.161^{*}	0.002	0.229^{*}	3.6°
	(0.077)	(0.048)	(0.093)	
Low Volume	-0.077	0.205^{**}	0.222^{*}	13.0^{**}
	(0.068)	(0.048)	(0.097)	
All Stocks	-0.047	0.118^{*}	0.244^{**}	4.2^{*}
	(0.076)	(0.049)	(0.080)	
Panel *b*: Conditional on interest rate changes				
High Volume	0.011	-0.059	0.234^{**}	0.7
	(0.067)	(0.068)	(0.084)	
Medium Volume	-0.027	-0.130^{*}	0.247^{**}	1.8
	(0.052)	(0.064)	(0.090)	
Low Volume	0.255^{**}	0.028	0.175°	7.8^{**}
	(0.063)	(0.057)	(0.096)	
All Stocks	0.122^{*}	-0.015	0.240^{**}	3.4°
	(0.059)	(0.062)	(0.077)	

Both panels: Regression model, $r_t = \beta_0 + (\beta_1 D_{1,t-1} + \beta_2 D_{2,t-1} + \beta_3 D_{Monday,t-1})r_{t-1}$. $D_{Monday,t}$ is a dummy variable for Monday open. Panel *a*: $D_{1,t}$ is a dummy variable for the central 50% of return observations, while $D_{2,t}$ is a dummy variable for the remaining observations. Uses daily S&P 500 returns collected by Findata. Panel *b*: $D_{1,t}$ is a dummy variable for the central 60% of interest rate changes, while $D_{2,t}$ is a dummy variable for the remaining observations. Uses changes in 10 year US Treasury bond rates collected by Sveriges Riksbank. The regressions use raw returns (not filtered for the expected effect of overnight US index returns and interest rate changes). Regressions using filtered data give similar parameter estimates and test statistics. The Wald statistic tests the restriction $\beta_1 = \beta_2$. $\chi^2(1)$ critical values: 6.6/3.8/2.7 at the 0.01/0.05/0.10 level. Regressions use least squares estimation with asymptotic GMM standard errors (in parentheses) that are robust to heteroskedasticity (Hansen, 1982).

Table 6: Index return autocorrelation at close conditional on trading volume

Portfolio	$\hat{\beta}_1$ *Low*	$\hat{\beta}_2$ *High*	Wald
High Volume	0.232^{**}	0.097^{**}	9.0^{**}
	(0.037)	(0.026)	
Medium Volume	0.118^{**}	0.009	4.2^{*}
	(0.045)	(0.030)	
Low Volume	0.090	0.005	1.6
	(0.056)	(0.038)	
All Stocks	0.199^{**}	0.065^{*}	5.4^{*}
	(0.050)	(0.029)	

Regression model: $r_t = \beta_0 + (\beta_1 D_{1,t-1} + \beta_2 D_{2,t-1}) r_{t-1}$. $D_{1,t}$ ($D_{2,t}$) is a dummy variable for low (high) trading volume in the portfolio *All Stocks*. The Wald statistic tests the restriction $\beta_1 = \beta_2$. $\chi^2(1)$ critical values: 6.6/3.8/2.7 at the 0.01/0.05/0.10 level. Regressions use least squares estimation with asymptotic GMM standard errors (in parentheses) that are robust to heteroskedasticity (Hansen, 1982). $^{**}/^{*}/^{\circ}$ Significantly different from zero at the 0.01/0.05/0.10 level.

3.6 Testing effects from cross–security correlation in revealed information

The results of section 3.5.3 (table 5), show that the autocorrelation at open is negative conditional on a "quiet night," that is, small US index returns. In the model, this implies that there is non-zero cross-security correlation of the information revealed in the opening call auction (high ϕ^m/ϕ^s), contrary to the assumption of Chan (1993).

Unfortunately, the signals are not observable and the cross-security correlation in revealed information is not measurable. For empirical testing, theoretical arguments must be used to identify situations where the cross-security correlation of revealed information is particularly high or low.[20] As the correlation depends on events in the trading process itself, and not on the events prior to the auction, it is hard to find good testable cases of high and low correlation in revealed information. The two possible tests discussed here are based on index arbitrage trading and short-selling restrictions.

3.6.1 Index arbitrage

If some agents buy or sell several securities simultaneously, as in the case of index arbitrage trading, realised returns and revealed information will be more strongly correlated across securities. Therefore, index return autocorrelation will be reduced or even negative after index arbitrage transactions.

This hypothesis is supported by the empirical results of Harris et al. (1994). The authors use NYSE intraday data to study return behaviour close to large index arbitrage transactions. Index returns are strongly positively autocorrelated, but following on index arbitrage transactions, returns reversals are documented

[20] It is also possible to measure cross-security correlation of revealed information using the cross-security correlation in short-term returns. However, for empirical testing of the model, in-sample measures cannot be used (in the presence of measurement errors, the null would tend to be rejected). However, it is possible to identify cross-security correlation using matching samples or out-of-sample techniques.

(returns are thus negatively autocorrelated). Conditional on non-arbitrage program trades, index return autocorrelation is close to zero, i.e. significantly lower than normal levels of autocorrelation.

3.6.2 Short-selling restrictions

In a market with explicit or self-imposed short-selling restrictions, downward price pressure originates (mostly) from the owners of the security in question. In contrast, upward price pressure may originate from any market participant, owners and non-owners alike. Investors can thus be "stock-picking" in a rising market, but must sell whatever stocks they already hold in a falling market.

As investors in aggregate hold the market portfolio, cross-security correlation in revealed information is higher conditional on an index level decrease. Consequently, there will be lower index return autocorrelation conditional on an index level decrease.

This conjecture is supported by the results in table 7 where return autocorrelation is conditioned on the preceding index return. At open, autocorrelation results are mixed. For the least liquid portfolio, autocorrelation is somewhat higher following on negative index returns, but all other portfolios exhibit higher autocorrelation conditional on index level increases. At close, the pattern is in line with predictions, and index return autocorrelation is consistently higher after days of above average stock market performance. In three out of four reported cases, the difference is statistically significant.

To test whether this asymmetry is, in fact, a result of an asymmetry in investor behaviour, is outside the scope of this paper. It is, however, an interesting topic for future research.

3.7 Interpreting empirical results within the model

The following example is not a formal test of the model. Instead it is intended to show how the model can be used as an analytical tool to interpret observed return phenomena and generate testable hypotheses.

In table 3, we can compare index return autocorrelation at open and close. It is clear that the point estimate of autocorrelation is higher at close than at open for *High Volume*, the portfolio of most liquid stocks (*0.066* at open , *0.157* at close). For the *Medium Volume* portfolio there is no large difference between autocorrelation at open and close (*0.023*, *0.052*), but for *Low Volume* the relationship is the reverse with high autocorrelation at open but low at close (*0.185*, *0.045*).

It seems reasonable to believe that the index level uncertainty is lower at close than at open as overday trading reduces the uncertainty about the true index level.[21] Still, autocorrelation is stronger at close for *High Volume*. In terms of the model of cross-security information aggregation, this effect must be a result of reduced cross-security correlation in revealed information at close.

What may be the cause of this reduction in cross-security correlation of the revealed information? The obvious suspect is the trading behaviour of individual investors. It is well known (albeit from anecdotal evidence) that many investors

[21] This conjecture is supported by results reported in table 8. While individual stock returns are strongly negatively autocorrelated at open (average: -0.125) they are only weakly autocorrelated at close (-0.044). This result is consistent with higher price precision at close.

Table 7: Index return autocorrelation conditional on high and low returns

Portfolio	$\hat{\beta}_1$ *Low*	$\hat{\beta}_2$ *High*	$\hat{\beta}_3$ *Monday*	Wald
At open				
High Volume	-0.085	0.055	0.253^{**}	1.4
	(0.087)	(0.069)	(0.092)	
Medium Volume	-0.098	0.003	0.239^{**}	1.1
	(0.069)	(0.059)	(0.091)	
Low Volume	0.147°	0.104	0.197°	0.1
	(0.079)	(0.066)	(0.112)	
All Stocks	0.010	0.126^{*}	0.255^{**}	1.3
	(0.076)	(0.060)	(0.089)	
At close				
High Volume	0.071°	0.242^{**}	...	6.0^{*}
	(0.040)	(0.042)		
Medium Volume	-0.046	0.137^{**}	...	5.4^{*}
	(0.048)	(0.046)		
Low Volume	-0.027	0.094°	...	1.3
	(0.069)	(0.056)		
All Stocks	0.026	0.205^{**}	...	4.4^{*}
	(0.048)	(0.052)		

Regression model: $r_t = \beta_0 + (\beta_1 D_{1,t-1} + \beta_2 D_{2,t-1} + \beta_3 D_{Monday,t-1}) r_{t-1}$. $D_{1,t}$ ($D_{2,t}$) is a dummy variable for low (high) realised index returns in the preceding trading period. $D_{Monday,t}$ is a dummy variable for Monday open. The Wald statistic tests the restriction $\beta_1 = \beta_2$. $\chi^2(1)$ critical values: 6.6/3.8/2.7 at the 0.01/0.05/0.10 level. Regressions use least squares estimation with asymptotic GMM standard errors (in parentheses) that are robust to heteroskedasticity (Hansen, 1982). $^{**}/^{*}/^{\circ}$ Significantly different from zero at the 0.01/0.05/0.10 level.

on the Paris Bourse prefer to trade at the close or as near as possible to the close. This is probably because most performance evaluation is carried out against closing prices as these are most readily available. Some speculators may also want to close open positions before the trading day ends.

In an REE, market such trades are considered to be liquidity trading since they are not based on expected future returns. If the closing transactions are uncorrelated across securities, this can explain the empirical results. This proposition is clearly testable using intraday data.

Why, then, is the same result not observed for the *Low Volume* portfolio? Although component stocks are less liquid than the *High Volume*stocks, they are still quite liquid. If these less liquid stocks are not subject to position closing liquidity trading at close it could be that positions are closed in earlier trading, or that traders do not let positions grow too big during the day. Both hypotheses are testable using dealer inventory data.

4 Conclusion

This paper derives a model of autocorrelation in stock index returns based on information aggregation across stocks that trade individually.

The first main implication of the model is that increased cross-security correlation in the market prior of security prices increases index return autocorrelation. This proposition was tested using several different approaches. High cross-security correlation of the prior is expected, for example, at Monday open, after overnight US market volatility and in portfolios of highly correlated stocks. All tests support the model at varying levels of statistical significance.

The other main implication of the model is that autocorrelation decreases in the cross-security correlation of the information revealed in trading. This aspect of the model is significantly more difficult to test as signals and signal correlations are unobservable.

One empirical test supported the prediction that index return autocorrelation should be lower after days with negative index returns. Empirical evidence of Harris et al. (1994) supports the prediction of lower index return autocorrelation conditional on index arbitrage trading. They even report return reversals, i.e. negative return autocorrelation.

An important advantage of the model is that information is extracted from prices using standard REE theory. The model is therefore formulated from the point of view of an econometrician who only observes realised prices.

Relative to earlier work on index return autocorrelation, the paper provides a methodological innovation by using a narrow, carefully selected, dataset instead of market-wide stock portfolios. This makes it possible to eliminate other possible sources of index return autocorrelation such as nonsynchronous trading or transaction costs.

Although testing is carried out in a controlled environment, it is highly probable that the same price adjustments are present in intraday trading. There, the resulting autocorrelation will be lower, but cross-security information aggregation can help to account for the index return autocorrelation not explained by nonsynchronous trading, especially during active trading. The model thus explains the findings of McInish and Wood (1991), that is, the U-shape in index return autocorrelation. High autocorrelation at open and close is consistent with the high uncertainty of the index level prior present at open and close.

An important application of the model is as a tool to analyse empirical evidence of cross-security information aggregation. For example, the model can be used to analyse lead-lag effects between index futures returns and cash index returns. The example in section 3.7 shows how testable hypotheses can be generated from "stylised results." Hypotheses are left as suggestions for future research as they require intraday data for the empirical testing.

Two major conclusions can be drawn from the model. Firstly, index return autocorrelation is consistent with efficient markets and prices. Transaction costs, measurement errors and other inefficiencies may increase autocorrelation, but autocorrelation need not be zero in the absence of such imperfections. Secondly, the model can generate both positive and negative autocorrelation in index returns. Almost all earlier theoretical and empirical work has focused on positive return autocorrelation, but as the empirical results show, negative index return autocorrelation is observed in several cases. This is an interesting field for further empirical study.

Table 8: Summary statistics of stocks used for portfolio construction

Corr No.†	Name	Autocorrelation At open	Autocorrelation At close	Trading volume‡	Volatility × 100 close-to-open	Volatility × 100 open-to-close
2	Alcatel Alsthom	0.163*	0.082**	217.2	0.900	0.536
1	Elf Aquitaine	−0.050	0.065**	174.1	1.039	0.627
3	Total	0.002	0.006	116.8	1.035	0.739
3	LVMH Moët Vuitton	0.008	0.019	116.0	0.926	0.717
3	Générale des Eaux	0.015**	0.309	114.1	1.046	0.671
1	Groupe Danone	−0.169**	0.000	107.4	0.847	0.483
1	Société Générale A	−0.012	0.043*	96.3	1.036	0.607
1	Suez (Compagnie de)	−0.081	0.102**	93.3	1.127	0.676
1	Saint-Gobain	−0.074	0.058**	90.8	1.101	0.638
1	Peugeot	−0.041	0.088**	90.2	1.066	0.585
3	Carrefour	−0.053*	0.093	80.9	0.964	0.605
1	l'Oreal	−0.046	−0.003	80.2	1.085	0.630
1	Paribas (fin.)	−0.050	0.069**	74.0	1.228	0.726
1	Lafarge	−0.099*	0.040°	66.3	1.135	0.640
1	Air Liquide	−0.116*	−0.008	65.0	0.931	0.525
3	UAP	−0.006	−0.051	63.2	1.425	0.905
1	Michelin B	0.001	0.047*	62.7	1.248	0.689
2	Axa	0.018**	0.062**	60.6	1.241	0.807
1	Lyonnaise des Eaux	−0.116*	0.029	60.3	1.034	0.604
2	Eurotunnel	−0.004	0.112**	56.5	1.589	1.022
1	Accor	0.018	0.056**	46.3	1.017	0.590
1	Havas	−0.192**	0.037°	39.5	1.188	0.672
3	Canal+	−0.024	0.004	37.0	1.097	0.627
3	Sanofi	−0.014	−0.013	34.9	1.082	0.737
2	CLF	−0.082°	0.014	33.9	1.145	0.713
3	EuroDisney sca	−0.140**	0.139**	31.4	1.579	1.095
3	Pernod-Ricard	0.005	0.087	28.9	1.105	0.800
1	Bouygues	−0.195**	−0.006	28.2	1.112	0.683
1	CCF	−0.183**	−0.070**	25.9	1.190	0.733
1	Bancaire (Cie)	−0.275**	−0.016	25.0	1.524	0.856
3	Pinault-Printemps	−0.286**	−0.121**	25.6	1.104	0.722
2	Carnaud Metalbox	−0.302**	−0.020	24.1	0.996	0.599
1	Schneider	−0.279*	−0.124**	24.4	1.440	0.980
1	Thomson-CSF	−0.223**	−0.019	24.6	1.508	0.858
3	Valeo	−0.012	−0.044	21.4	1.229	0.823

The table continues on the next page.

Table 8 continued

Corr No.†	Name	Autocorrelation At open	Autocorrelation At close	Trading volume‡	Volatility × 100 close-to-open	Volatility × 100 open-to-close
3	AGF	−0.044	−0.084**	21.3	1.086	0.753
2	Eridania Beghin	−0.241**	−0.062*	20.2	1.002	0.598
3	Promodes	−0.026	−0.177	19.3	1.080	0.767
2	CFF	−0.061*	−0.059*	17.9	1.395	0.874
3	Legrand	−0.014**	−0.215**	16.7	1.107	0.846
2	Chargeurs	−0.180**	−0.059*	15.0	1.157	0.694
3	Castorama Dubois	−0.037**	−0.133**	14.3	1.093	0.663
1	Saint Louis	−0.364**	−0.085**	14.8	1.044	0.649
2	Pechiney Intl	−0.155**	−0.102**	13.9	1.276	0.764
3	Rhone-Poulenc cip	−0.158*	−0.356**	12.7	1.836	1.795
3	Roussel-Uclaf ord.	−0.019	−0.344*	12.6	1.182	0.944
2	Credit Lyonnais ci	−0.386**	−0.095**	11.0	1.431	1.022
2	Pechiney cip	−0.292**	−0.081**	11.3	1.337	0.790
2	Credit National	−0.044	−0.183°	11.5	1.055	0.716
2	Casino Guichard	−0.392**	−0.125**	11.6	1.424	0.855
2	Club Mediterranée	−0.204**	−0.040	11.7	1.155	0.747
3	Synthelabo	−0.013	−0.147**	10.3	1.140	0.811
3	Docks de France	0.012**	−0.066	10.7	0.970	0.703
2	Cap Gemini Sogeti	−0.216**	0.008	10.1	1.315	0.824
2	Matra Hachette	−0.152*	−0.138**	10.2	1.759	1.194
2	Cetelem	−0.071	−0.076°	9.5	1.245	0.774
2	Lagardere groupe	−0.197**	−0.075**	9.8	1.434	0.883
3	Groupe Ass. Nat.	−0.026	−0.082*	9.5	1.291	0.887
3	Bic	−0.055°	−0.004	9.1	1.071	0.791
1	CGIP	−0.219**	−0.120**	9.4	1.179	0.779
1	Schneider	−0.361**	−0.071**	7.9	1.510	0.902
1	Imetal	−0.347**	−0.128**	7.7	1.330	0.820
2	Sodexho	−0.355**	−0.118**	6.1	0.975	0.579
3	Rhone-Poulenc Rorer	0.045**	−0.268**	6.8	0.847	0.892
2	Cerus-Europ. Reun.	−0.311**	−0.045°	6.0	1.312	0.879
2	Comptoirs Modernes	−0.341**	−0.153**	6.4	0.982	0.630
2	Sovac	−0.010	−0.121**	5.9	1.164	0.840
2	Essilor Intl	−0.327**	−0.116**	5.9	1.152	0.727
2	Ecco	−0.254**	−0.103**	5.5	1.207	0.763
3	Seb	−0.008	−0.131*	5.9	1.155	0.861
Average		−0.125	−0.044	39.0	1.186	0.771
Sample standard error		(0.131)	(0.108)		(0.203)	(0.184)

Regressions use least squares estimation with asymptotic GMM standard errors (not reported) that are robust to heteroskedasticity (Hansen, 1982). †Included in correlation portfolio, 1=*High Correlation*, 2=*Medium Correlation*, 3=*Low Correlation*. ‡Average per day, per stock, trading volume in million FRF. **/*/° Significantly different from zero at the 0.01/0.05/0.10 level.

References

Abhyankar, Abhay H., "Return and Volatility Dynamics in the FT-SE 100 Stock Index and Stock Index Futures Markets," *Journal of Futures Markets*, 1995, *15* (4), 457–488.

_ , "Does the Stock Index Futures Market Tend to Lead the Cash? New Evidence from the FT-SE 100 Stock Index Futures Market," 1996. Working Paper 96-01, University of Sterling, Department of Accounting and Finance.

Admati, Anat R., "A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Markets," *Econometrica*, May 1985, *53* (3), 629–657.

_ and Paul Pfleiderer, "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies*, 1988, *1* (1), 3–40.

Amihud, Yakov and Haim Mendelson, "Trading Mechanisms and Stock Returns: An Empirical Investigation," *Journal of Finance*, July 1987, *42* (3), 533–555.

Atchison, Michael D., Kirt C. Butler, and Richard R. Simonds, "Nonsynchronous Security Trading and Market Index Autocorrelation," *Journal of Finance*, March 1987, *42* (1), 111–118.

Berglund, Tom and Eva Liljeblom, "Market Serial Correlation on a Small Security Market: A Note," *Journal of Finance*, December 1988, *43* (5), 1265–1274.

Biais, Bruno, Pierre Hillion, and Chester Spatt, "An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse," *Journal of Finance*, December 1995, *50* (5), 1655–1689.

_ , _ , and _ , "Price Discovery and Learning During the Preopening Period in the Paris Bourse," 1996. Unpublished manuscript, University of Toulouse.

Boudoukh, Jacob, Matthew P. Richardson, and Robert F. Whitelaw, "A Tale of Three Schools: Insights on Autocorrelations of Short-Horizon Stock Returns," *Review of Financial Studies*, Fall 1994, *7* (3), 539–573.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, "Trading Volume and Serial Correlation in Stock Returns," *Quarterly Journal of Economics*, November 1993, pp. 905–939.

Chan, Kalok, "A Further Analysis of the Lead-Lag Relationship Between the Cash Market and Stock Index Futures Markets," *Review of Financial Studies*, 1992, *5* (1), 123–152.

_ , "Imperfect Information and Cross-Autocorrelation among Stock Prices," *Journal of Finance*, September 1993, *48* (4), 1211–1230.

_ , K. C. Chan, and G. Andrew Karolyi, "Intraday Volatility in the Stock Index and Stock Index Futures Market," *Review of Financial Studies*, 1991, *4* (4), 657–684.

Cohen, Kalman J., Gabriel A. Hawawini, Steven F. Maier, Robert A. Schwartz, and David K. Whitcomb, "Implications of Microstructure Theory for Empirical Research on Stock Price Behaviour," *Journal of Finance*, May 1980, *35* (2), 249–257.

Fischer, Lawrence, "Some New Stock Market Indexes," *Journal of Business*, 1966, *39*, 191–225.

Foster, F. Douglas and S. Viswanathan, "Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models," *Journal of Finance*, March 1993, *48* (1), 187–211.

Glosten, Lawrence R. and Paul R. Milgrom, "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 1985, *14*, 71–100.

Grossman, Sanford J., "On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information," *Journal of Finance*, May 1976, *31* (2), 573–585.

_ **and Joseph E. Stiglitz**, "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 1980, *70*, 393–408.

Hansen, Lars Peter, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, July 1982, *50* (4), 1029–1054.

Harris, Lawrence, George Sofianos, and James E. Shapiro, "Program Trading and Intraday Volatility," *Review of Financial Studies*, 1994, *7* (4), 653–685.

Hellwig, Martin F., "On the Aggregation of Information in Competitive Markets," *Journal of Economic Theory*, 1980, *22*, 477–498.

Holden, Craig W. and Avinadhar Subrahmanyam, "Long-Lived Private Information and Imperfect Competition," *Journal of Finance*, March 1992, *47* (1), 247–270.

Krishnan, Murugappa, "An Equivalence between the Kyle (1985) and the Glosten-Milgrom (1985) Models," *Economics Letters*, November 1992, *40* (3), 333–38.

Kyle, Albert S., "Continuous Auctions and Insider Trading," *Econometrica*, 1985, *53* (6), 1315–1335.

_ , "Informed Speculation with Imperfect Competition," *Review of Economic Studies*, 1989, *56*, 317–356.

Lo, Andrew W. and A. Craig MacKinlay, "An Econometric Analysis of Nonsynchronous Trading," *Journal of Econometrics*, 1990, *45*, 181–211.

_ **and** _ , "When Are Contrarian Profits Due to Stock Market Overreaction?," *Review of Financial Studies*, 1990, *3* (2), 175–205.

McInish, Thomas H. and Robert A. Wood, "Autocorrelation of Daily Index Returns: Intraday-to-Intraday versus Close-to-Close Intervals," *Journal of Banking and Finance*, 1991, *15*, 193–206.

Mech, Timothy S., "Portfolio Return Autocorrelation," *Journal of Financial Economics*, 1993, *34*, 307–344.

Paul, Jonathan, "Information Aggregation Without Exogenous 'Liquidity' Trading," March 1994. Unpublished manuscript, University of Michigan.

Rochet, Jean Charles and Jean Luc Vila, "Insider Trading Without Normality," *Review of Financial Studies*, 1994, *61* (1), 131–52.

Säfvenblad, Patrik, "Lead-Lag Effects When Prices Reveal Cross-Security Information," Working Paper No. 189, Working Paper Series in Economics and Finance, Stockholm School of Economics 1997.

Sarkar, Asani, "On the Equivalence of Noise Trader and Hedger Models in Market Microstructure," *Journal of Financial Intermediation*, March 1994, *3* (2), 204–212.

Scholes, Myron and Joseph T. Willams, "Estimating Betas from Nonsynchronous Data," *Journal of Financial Economics*, December 1977, *5* (3), 309–327.

Shin, Jhinyoung and Rajdeep Singh, "Endogenous Informed Trading in the Presence of Trading Costs: Effect on Observed Security Prices," 1996. Unpublished manuscript, Washington University at St. Louis.

Stoll, Hans R. and Robert E. Whaley, "The Dynamics of Stock Index and Stock Index Futures Returns," *Journal of Financial and Quantitative Analysis*, December 1990, *25* (4), 441–468.

Vives, Xavier, "Short-Term Investment and the Informational Efficiency of the Market," *Review of Financial Studies*, 1995, *8* (1), 126–160.