Lead-Lag Effects When Prices Reveal Cross-Security Information

Patrik Säfvenblad*
Stockholm School of Economics
Department of Finance

September 1, 1997

Abstract

This paper introduces and analyses a model of cross-security information aggregation in a rational expectations equilibrium (REE). The model predicts a well-defined lead-lag structure between securities returns as a result of Bayesian information extraction from realised securities prices. Both leads and lags will be strongest between securities with highly correlated return processes, but only weakly correlated return innovations. Securities whose prices reveal highly precise signals will tend to lead other securities.

The model has important implications for empirical testing of lead-lag effects between financial markets and instruments. As an application of the model, it is demonstrated that stock option returns will tend to lag the corresponding stock prices.

Direct empirical tests of the lead-lag effects between individual stocks on the Paris Bourse provided strong support for the model. In addition to confirming the predicted pattern of leads and lags, the paper demonstrates that the cross-security correlation is higher for short-term returns than for long-term returns for about a third of securities pairs traded on the Paris Bourse. This result is interpreted as strong cross-security correlation of revealed information, which gives the model strong support over alternative specifications of multi-asset securities markets, such as the nonsynchronous trading hypothesis or the Chan (1993) model.

---

*The paper has benefitted from comments made by Clas Bergström, Bevan Blair, Magnus Dahlquist, Gustaf Hagerud, Richard Kihlstrom, Richard Lyons, David Smith, Staffan Viotti and Robert E. Whaley. 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

Although stocks are traded individually, their returns are strongly correlated. This implies the existence of an informational link between the different securities. If the price of one security moves "out of line" relative to other securities, relative prices are corrected in subsequent trading.

While such a mechanism is highly intuitive, its implications for measured stock returns are not well understood. In partial remedy, this paper derives and tests a model of lead-lag effects, or equivalently cross-autocorrelations, between individual securities that trade simultaneously. In the model, information revealed in the price of one security affects the valuation of all other securities. The resulting price adjustments generate a well-defined pattern of cross-autocorrelation in security returns.

The theoretical analysis of the lead-lag relation is basically a study of relative informativeness. If the price of a security is informative for prices of other securities, its returns will lead those of other securities. In general, this lead is reciprocal, so that between closely related securities there will be both leads and lags. If the information revealed in prices is correlated across securities, the cross-security informativeness of realised prices is reduced, and consequently, lead-lag effects are weakened. If the correlation in revealed information is strong, returns may even become negatively cross-autocorrelated.

The lead-lag relation is normally expected to be bidirectional, but empirically there are cases where the lead is virtually unidirectional. One example of such a relation is the case of stocks and stock options. Returns on the underlying stock lead option returns but evidence of the reverse influence is weak. Two reasons for this are discussed in section 4.5. Firstly, an option price is less informative for a stock price than *vice versa*. Secondly, the large number of traded option series results in a one–to–many effect. Each individual option price has low informativeness for the stock price, while the stock price is highly informative for all traded options.

Empirically, long-term returns are usually more strongly correlated across securities than short-term returns. This is also the predicted effect of microstructure effects such as nonsynchronous trading and bid-ask bounce effects. However, as shown in section 5, it is not uncommon that short-term returns are more strongly correlated than long-term returns. This correlation pattern is incompatible with existing models of cross-security price formation,[1] but, according to the model presented in this paper, it is the result of strong cross-security correlation in revealed information.

Most empirical work on lead-lag effects concentrates on the *speed* of price adjustment. A security is said to lead other securities if its price adjustment to a common factor is earlier than that of other securities. This definition is often used in the literature on lead-lag effects between index futures and their corresponding cash indices.[2] In this paper, the mechanism is different. Trading reveals information that *causes* price revisions of securities with correlated underlying values or information.

Five sections follow this introduction. The next section, section 2, provides

[1] Including Admati (1985), Amihud and Mendelson (1987), Lo and MacKinlay (1990a), Chan (1993) and Mech (1993).

[2] Examples include Lo and MacKinlay (1990b), Badrinath et al. (1995), Chan (1992), Brennan et al. (1993) and McQueen et al. (1996).

a theoretical background based on the body of rational expectations equilibrium models. Section 3 presents a formal model of cross-security information aggregation and derives implications for cross-autocorrelation of securities returns. Section 4 discusses the interpretation and estimation of the model and the empirical results are presented in section 5. The conclusion and suggestions for further development of the model are offered in section 6.

2 Rational expectations equilibrium models

2.1 Introduction

The model developed in this paper uses the large literature on asset pricing in noisy rational expectations equilibria (REE) as a starting point. These models share a large number of properties that make them useful for analysing information transmission in financial markets.[3] Furthermore, the REE models are adaptable to various trading environments. This paper focuses on the auction (order driven) market, but given the similarity of REE models, the model can be used to analyse a price formation in, for example, a specialist dealer or multiple dealer market.

In an REE market, profit maximising agents trade risky assets among themselves in a usually centralised marketplace. Prices result from equilibrium demand strategies, where agents use private information and take any exogenous noise, such as liquidity trading or imperfect information, into account.

The rational expectations equilibrium implies that agents' demand strategies are optimal with regard to the demand strategies of all other agents. All demand strategies are thus optimal with regard to the resulting price. Therefore, the REE prices aggregate all publicly available information about future values. Since the REE prices are optimal with regard to rational agents' demand strategies, the information revealed in a security price cannot affect the valuation of the security proper. However, the realisation of a price is still an information event, since it can be used to update estimates of the value of *other* securities.

In addition to information that is revealed through trading, other sources of information also influence asset prices. Examples include scheduled announcements or overnight trading in other markets. Such information will affect asset prices by changing all agents' valuation of securities, and prices change accordingly without trading.[4]

The noise in REE securities prices has two separate sources. Firstly, the information collected by informed speculators may not perfectly reveal the underlying value of securities. In an economy with many speculators, this requires that agents' measurement (signal) errors are correlated across individuals. Secondly, liquidity trading (or any trading not based on the expected returns), introduces additional noise into the price system.

[3] Seminal papers include Grossman (1976), Grossman and Stiglitz (1980), Hellwig (1980), Glosten and Milgrom (1985) and Kyle (1985, 1989). The basic similarities between the various REE models have been demonstrated by several authors, e.g., Krishnan (1992), Paul (1994), Rochet and Vila (1994), Sarkar (1994), and Vives (1995).

[4] In practice, an announcement may also trigger information aggregation through prices, in particular when each agent makes an independent analysis of the news event, or when the announcement affects individual agents' demand functions.

2.2 The competitive REE

The model presented in this paper is based on a competitive REE model. In such models, prices are martingales and the expected profits of a rational, but uninformed, speculator is zero. Hellwig (1980) and Kyle (1985) are examples of competitive REE models.

If the economy is not competitive, individual uninformed speculators have market power over residual supply. They will use their market power to share some of the profits that informed speculators make from trading with liquidity traders. In this kind of equilibrium, returns are negatively autocorrelated (Kyle, 1989).

The market power, and the expected profits, of speculators decrease rapidly with the number of speculators active in the market. With as few as 3–4 informed speculators, pricing will be close to the competitive prices.[5] The non-competitive case is consistent with the model of cross-security information aggregation, but is not discussed in this paper.

2.3 REE in a multi-security market

A multi-security REE model is an analytically complex, but intuitively simple extension of the single-security REE model. This is evident in Admati's (1985) multi-security extension of the Hellwig (1980) model. The main difference is that demand strategies must be optimal with regard to the *whole* price vector, not only the price of a single security.

In the Admati model, agents submit multidimensional demand schedules, conditioned on the full price vector. Such demand schedules require the calculation of N^2 parameters, where N is the number of securities in the market. Theoretically, this may be achieved in markets with only a few securities, but it is obviously an unrealistic assumption for markets where hundreds of securities are traded. Furthermore, stock exchanges only offer rudimentary types of cross-security limit orders, such as conditional trades or basket trades.

If traders cannot condition their orders on the full set of prices, even REE prices will be inefficient across securities. It is this inefficiency that is exploited in the model presented in this paper. The cross-security information aggregation is equivalent to the difference between single-security REE prices, derived by Hellwig (1980), and the multi-security REE prices, derived by Admati (1985).[6]

It must be pointed out that the resulting lead-lag effects do not present arbitrage opportunities. Although returns are cross-autocorrelated, and thus predictable, the price inefficiency cannot be used to make trading profits. Prices will adjust without trading, since the predictability of price changes is known by all agents.

2.4 Earlier cross-autocorrelation models

This paper is close in spirit to the paper of Chan (1993), who studies the pricing problem of a Kyle (1985) style market maker who acts in a market where underlying value innovations are correlated across securities. The market maker is

[5] See Holden and Subrahmanyam (1992) for a multi-period setting, and Caballé and Krishnan (1994) for a multi-security setting.

[6] With some minor caveats regarding optimal information revelation and optimal information acquisition.

a specialist who only observes order flow in "his own" security. As order flows in other securities are unobservable, the market maker must deduce the information content of these flows by observing the *prices* of other market makers. As a result, returns will be positively cross-autocorrelated.

This paper makes two important theoretical extensions to the Chan model. Firstly, it implements the model for a general REE setup. This is conceptually important as it is obvious that a specialist market maker has a physical information monopoly over the incoming order flow. That similar effects persist in an auction market, where order flow is visible to all traders, is far from obvious. Secondly, the extended model allows signals to be correlated across securities. This addition is of both theoretical and empirical importance. Theoretically, it adds the possibility of modelling index arbitrage trading and other multi-security trading strategies. Empirically, it is necessary in order to model observed return patterns on the Paris Bourse. The observed high cross-security correlation in short-term returns is simply not compatible with the basic Chan model.

In another extension of the Chan (1993) model, Shin and Singh (1996) model what the authors call "spurious predictability." Their results are nested by the results of this paper.[7]

3 Model

3.1 General model

Assume a Walrasian stock market of the Hellwig (1980) type. There are two types of traders, 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 a security. Liquidity traders trade for some exogenous reason (e.g. hedging, 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.

There are N securities that are claims to separate underlying values, arranged in the vector $\mathbf{V}$. The underlying values are not observable, but before trading, agents share a prior valuation of securities, denoted $\mathbf{P}^*_{-1}$. The prior reflects all public information available before demand schedules are submitted. The valuation error of the prior is normally distributed, with a publicly known covariance matrix, $\mathbf{\Pi}$.

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

where $\mathcal{F}_{-1}$ denotes the public information set before trading.

Before trading takes place, some of the speculators receive or acquire a (private) noisy measurement of the underlying values of one or several securities. Agents calculate optimal demand schedules using the private signal and the

[7] The term "spurious predictability" is somewhat misleading. While equilibrium returns are predictable, the predictability cannot be traded away by arbitrageurs, because of the imposed informational constraints. The results of Shin and Singh (1996) correspond to establishing when $\omega_{ij} \neq 0$ (equation 23 of this paper).

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 → B → Trading → C → D → 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.

equilibrium covariance structure of signals and returns.[8]

Each security is traded in a separate, frictionless, competitive call auction. Before trading, a Walrasian auctioneer collects demand schedules for individual securities. At a predetermined point in time, the auctioneer sets a price vector, $\mathbf{P}$, clearing supply and demand for all stocks.

Relying on standard REE results, we know that each price realised in trading will reveal a signal, F_i, relative to the underlying value of security i.[9] In this paper, the signal is modelled as a noisy measurement of the error in the prior valuation. For all stocks, we use vectors and matrices 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 $\mathbf{\Phi}$. Basically, the signal is a weighted sum of investors' private information distorted by the extent of liquidity trading and other price noise.[10]

Relying on the competitiveness assumption, the price in each of the N separate markets can be represented by the following equation of Bayesian updating:

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

[8] Under standard assumptions (exponential utility over next period's wealth and normality) demand schedules will be linear in the price.

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

[10] The properties of its covariance matrix, $\mathbf{\Phi}$, are discussed in some more detail in section 4.2.

where Π_{ii} is the ith diagonal element of the covariance matrix $\mathbf{\Pi}$. Under the normality assumption, equation 4 holds for all competitive single-security REE models, i.e., whenever realised prices are unbiased predictors of the underlying value or future price sequence.

Define κ_i as the immediate response of an individual price to new information, F_i. We have,

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

In vector and matrix notation, the κ_i:s are arranged in a diagonal matrix, $\hat{\mathbf{\Omega}}$.

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

to obtain

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

When prices are observed, they will be used to extract the information revealed in trading using equation 4. Next, a posterior estimate of $\mathbf{V}$, denoted $\mathbf{P}^*$, is formed using Bayesian updating and the extracted signal $\mathbf{F}$:

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

where $\mathbf{\Omega}$ is an updating matrix that efficiently updates all N prices using the N individual signals. In the case of normally distributed variables, Bayesian theory provides an explicit solution for $\mathbf{\Omega}$:

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

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

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

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, $\mathbf{P}^*$, use observed returns to extract information about $\mathbf{F}$. We have,

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

where $\mathbf{I}$ is an $N \times N$ identity matrix.

Equation 12 summarises the model. Before interpreting equation 12, define the elements of $\mathbf{\Omega}$ as:

$$\mathbf{\Omega} = \begin{bmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1N} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{N1} & \omega_{N2} & \cdots & \omega_{NN} \end{bmatrix}. \tag{13}$$

As $\hat{\mathbf{\Omega}}$ is a diagonal matrix, $\mathbf{\Omega}\hat{\mathbf{\Omega}}^{-1}$ can be seen as a normalisation of the updating matrix $\mathbf{\Omega}$:

$$\mathbf{\Omega}\hat{\mathbf{\Omega}}^{-1} = \begin{bmatrix} \omega_{11}/\kappa_1 & \omega_{12}/\kappa_1 & \cdots & \omega_{1N}/\kappa_1 \\ \omega_{21}/\kappa_2 & \omega_{22}/\kappa_2 & \cdots & \omega_{2N}/\kappa_2 \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{N1}/\kappa_N & \omega_{N2}/\kappa_N & \cdots & \omega_{NN}/\kappa_N \end{bmatrix}. \tag{14}$$

Returns will thus be cross-autocorrelated whenever there are non-zero off-diagonal elements in the matrix $\mathbf{\Omega}$. From the structure of $\mathbf{\Omega}\hat{\mathbf{\Omega}}^{-1}$ we deduce that only the relative informativeness of signals matters for cross-autocorrelations. First stage returns are already adjusted for the informativeness of signals via the $\hat{\mathbf{\Omega}}$ matrix. Therefore, the matrix $\mathbf{\Omega}\hat{\mathbf{\Omega}}^{-1}$ is not affected by the absolute informativeness of revealed information.

It is straightforward, but complicated due to the matrix algebra, to show that the expected value of $P_i^* - P_i$, can be written as a weighted sum of "unexpected" returns or signals of all other securities.[11] Equation 15 formulates this intuitive result.

$$\begin{aligned} P_i^* - P_i &= \sum_{j=1}^{N} \omega_{ij} \left(F_j - \mathrm{E}\left[F_j | \mathcal{F}_{-1}, F_i \right] \right) \\ &= \sum_{j=1}^{N} \frac{\omega_{ij}}{\kappa_j} \left(r_j - \mathrm{E}\left[r_j | \mathcal{F}_{-1}, r_i \right] \right) \end{aligned} \tag{15}$$

In the first part of equation 15, the adjustment is formulated in terms of the revealed information. The weights ω_{ij} measure the relative informativeness of security j for the pricing of security i. The second part of the same equation provides the results expressed in returns instead. When using returns, the informativeness is normalised by κ_j, security j's initial response to the revealed information.

3.2 A one-factor model

Equation 12 provides an explicit solution for cross-autocorrelation. However, in order to demonstrate the model's implications, a less general setting is needed. Therefore, this section develops a "one-factor model," where priors and signals have both a market component and an individual stock component.

Assume that the covariance matrix of the prior has the structure

$$\Pi_{ij} = \begin{cases} \pi^m + \pi^s & \text{if } i = j \\ \pi^m & \text{if } i \neq j \end{cases} \quad \forall\, i, j, \tag{16}$$

[11] This also follows directly from the optimal signal extraction implied by Bayesian updating.

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

$$\underset{N\times N}{\mathbf{\Pi}} = \begin{bmatrix} \pi^m+\pi^s & \pi^m & \cdots & \pi^m \\ \pi^m & \pi^m+\pi^s & \cdots & \pi^m \\ \vdots & \vdots & \ddots & \vdots \\ \pi^m & \pi^m & \cdots & \pi^m+\pi^s \end{bmatrix}. \tag{17}$$

Similarly, we let the revealed information have a one-factor structure with the variance of the market signal ϕ^m and the additional variance of individual stock signals, ϕ^s:

$$\Phi_{ij} = \begin{cases} \phi^m+\phi^s & \text{if } i=j \\ \phi^m & \text{if } i\neq j \end{cases} \quad \forall\, i,j. \tag{18}$$

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

$$\mathbf{r} = \kappa^s\mathbf{F}, \tag{19}$$

$$\kappa^s = \frac{\pi^m+\pi^s}{\pi^m+\pi^s+\phi^m+\phi^s}. \tag{20}$$

The posterior returns, $\mathbf{r}^*$, can be interpreted as the sum of a market return and a security-specific return:

$$\mathbf{r}^* = \mathbf{\Omega}\mathbf{F} = \mathbf{\Omega}\frac{1}{\kappa^s}\mathbf{r}, \tag{21}$$

$$\mathbf{\Omega} = \left(\frac{\pi^s}{\pi^s+\phi^s}\mathbf{I} + \frac{\pi^s\phi^s\left(\pi^m/\pi^s-\phi^m/\phi^s\right)}{\left(\pi^s+\phi^s+N\left(\pi^m+\phi^m\right)\right)\left(\pi^s+\phi^s\right)}\mathbf{1}\mathbf{1}'\right), \tag{22}$$

where $\mathbf{I}$ is an $N\times N$ identity matrix, $\mathbf{1}$ is an $N\times 1$ column vector of ones and, consequently, $\mathbf{1}\mathbf{1}'$ is an $N\times N$ matrix of ones. The first term in the definition of $\mathbf{\Omega}$ (equation 22) can be interpreted as the Bayesian response to security-specific information. The second term is the optimal response of individual stocks to revealed market-wide information. All off-diagonal elements, ω_{ij}, in $\mathbf{\Omega}$ are equal, which implies that all securities react similarly to information revealed in all stocks,

$$\omega_{ij} = \frac{\pi^s\phi^s\left(\pi^m/\pi^s-\phi^m/\phi^s\right)}{\left(\pi^s+\phi^s+N\left(\pi^m+\phi^m\right)\right)\left(\pi^s+\phi^s\right)} \quad \forall i,j \quad i\neq j. \tag{23}$$

We can derive most of the direct implications of the model from equation 22. First note that cross-autocorrelation between individual stocks will converge to zero as N, the number of securities, increases. When a large number of securities contribute to the price discovery, individual lead-lag effects are weakened. However, it is easily shown that the cross-autocorrelation with the market return is strengthened as the number of securities grows. The reason is intuitively simple; if many securities share a market factor, the market factor will be better known, but individual securities' contribution to information revelation is reduced. Therefore, individual cross–autocorrelations are reduced, while the cross-autocorrelation with the market return is strengthened.

The sign of cross-autocorrelation is determined by $\pi^m/\pi^s - \phi^m/\phi^s$:

$$\text{sign}\,(\omega_{ij}) = \text{sign}\left(\frac{\pi^m}{\pi^s} - \frac{\phi^m}{\phi^s}\right) \quad \forall i,j \quad i \neq j. \tag{24}$$

This implies that if priors are more strongly correlated across securities than signals, cross-autocorrelation will be positive. If signals are more strongly correlated than underlying returns, the observed returns will be negatively cross-autocorrelated. The reason is straightforward; if the prior has strong cross-security correlation, the market factor is not well known before trading, and prices will therefore be used to identify the current market factor. If, however, there is high cross-security correlation in revealed signals, returns provide a bad measurement of the index level, and the index innovation is more likely to be caused by signal errors. Therefore, stocks react negatively to innovations in other stock prices.

3.3 The two securities case

A second intuitive example of the model's direct effects is the two securities case. Here, cross-effects can be calculated explicitly. As before, $\mathbf{\Pi}$ measures the covariance of the prior estimate of the underlying values:

$$\mathbf{\Pi} = \begin{bmatrix} \pi^m + \pi_1 & \pi^m \\ \pi^m & \pi^m + \pi_2 \end{bmatrix}. \tag{25}$$

If $\pi_1 > \pi_2$, the prior of security 1 is relatively more noisy than the prior of security 2. If π^m is big, the errors in prior valuation are strongly correlated.

Similarly, $\mathbf{\Phi}$ measures the precision of the information revealed in trading:

$$\mathbf{\Phi} = \begin{bmatrix} \phi^m + \phi_1 & \phi^m \\ \phi^m & \phi^m + \phi_2 \end{bmatrix}. \tag{26}$$

If $\phi_1 > \phi_2$, information revealed from trading security 1 is more noisy than information revealed from trading security 2. If ϕ^m is big, signal errors are strongly correlated. $\mathbf{\Omega}$ can be calculated explicitly from equation 9. The off-diagonal elements, ω_{12} and ω_{21} determine the cross-autocorrelation:

$$\omega_{12} = \frac{\pi_1\phi_1}{|\mathbf{\Pi}+\mathbf{\Phi}|}\left(\frac{\pi^m}{\pi_1} - \frac{\phi^m}{\phi_1}\right), \tag{27}$$

$$\omega_{21} = \frac{\pi_2\phi_2}{|\mathbf{\Pi}+\mathbf{\Phi}|}\left(\frac{\pi^m}{\pi_2} - \frac{\phi^m}{\phi_2}\right), \tag{28}$$

where $|\mathbf{\Pi}+\mathbf{\Phi}|$ is the determinant of $(\mathbf{\Pi}+\mathbf{\Phi})$. The term ω_{12} measures how security 1 reacts to unexpected returns in security 2. Similarly ω_{21} measures effect *from* security 1 *to* security 2.

From the symmetric nature of the lead-lag relation, simple comparative static analysis allows the following conclusions to be drawn from equation 27–28.

1. If the prior of the underlying values is highly correlated across securities, both leads and lags will increase ($\pi^m \uparrow \Rightarrow \omega_{12} \uparrow$ and $\omega_{21} \uparrow$).

2. If revealed signals are strongly correlated across stocks, leads and lags will decrease, and may become negative ($\phi^m \downarrow \Rightarrow \omega_{12} \downarrow$ and $\omega_{21} \downarrow$).

3. A security with higher prior variance will exhibit weaker (or more negative) leads to other securities ($\pi_1 \uparrow \Rightarrow \omega_{12} \downarrow$).

4. A security with higher prior variance will be less affected by events in other securities. The absolute value of the positive or negative lag is reduced. ($\pi_1 \uparrow \Rightarrow |\omega_{21}| \downarrow$).

5. A security with noisy revealed signals tends to lag the other securities ($\phi_1 \uparrow \Rightarrow \omega_{12} \uparrow$).

6. A security with noisy revealed signals tends not to lead other securities. Both positive and negative leads are reduced ($\phi_1 \uparrow \Rightarrow |\omega_{21}| \downarrow$).

4 Discussion

The model, as presented in the previous section, is highly stylised, and the adaptation to real world settings may not be altogether intuitive. Therefore, this section discusses some aspects of the model and its applicability. For simplicity, the discussion focuses on the one-factor model and the two securities case. Implications are equally valid for more general factor specifications.

4.1 The cross-security correlation of the prior

The properties of the common prior are critical for predicted lead-lag effects. If the cross-security correlation of the prior is high, the first stage response to common factors is lower than the optimal response after having observed information revealed in other securities prices. The model thus predicts positive cross-autocorrelation between individual securities.

Therefore, whenever the noisiness of the common factor is high relative to the additional noisiness of individual securities, positive lead-lag relations are expected. This is one of the reasons why most empirical studies document positive lead-lag relation, say between the cash index and index futures or between options and underlying stocks. In both cases the common factor carries most noise, and positive lead-lag effects should be expected.

In a multi-period setting, the covariance of the prior has two separate components, the covariance of the last posterior and the covariance of value innovations since the last transaction.[12] In a market with many individual securities, Bayesian updating ensures that the covariance matrix of the posterior is close to diagonal since the common factors are known with much higher precision than individual stock factors. Therefore, the cross-security correlation of the prior valuation will be determined mostly by the covariance of value innovations.

An examination of securities returns reveals that the correlation structure of value innovations is relatively constant. It is clear, for instance, that stocks exhibit consistent and strong cross-security correlation in daily returns. For practical purposes it can be assumed that the covariance structure of value innovations is also similar for other choices of return periods, such as close-to-open or intraday. Therefore, the cross-security correlation in long-term returns

[12] In addition, it is possible that new public information reduces the uncertainty concerning the underlying value of securities.

(e.g. monthly returns) can be used as a simple proxy for the correlation in the prior.

During continuous trading, the common factor prior will be known with relatively high precision, at least in markets with many traded securities, such as stock markets. Therefore, lead-lag effects tend to be relatively weak during continuous trading.[13] However, during periods of high return volatility, normally in early morning trading and before the close, the index level prior can be assumed to be more noisy, relative to the level of uncertainty in individual stocks. Therefore, lead-lag effects can be expected to be relatively strong around open and close.

4.2 Cross-security correlation of signals

Modelling information revelation as a signal $\mathbf{F}$ gives the model both generality and simplicity. The formulation is valid for all REE models. However, depending on the market setting, the interpretation is different. In a market maker framework, the signal will be equal to the realised net order flow facing the market maker. In an auction market setting, however, the signal is the net demand facing *any* agent.

The cross-security correlation of signals plays an important role in the theoretical analysis. If signal correlation is high, cross-autocorrelation will be reduced. If signal correlation is higher than the cross-security correlation in the prior, negative cross-autocorrelation will result. Although the cross-security correlation of revealed information cannot be observed, REE models clearly indicate what to expect.[14]

In a rational expectations equilibrium, the equilibrium pricing rule is known by all agents, and prices thus reveal net demand at all price levels. $\mathbf{\Phi}$, the variance of $\mathbf{F}$, is thus also a direct measure of the covariance of net demand. Net demand is the sum of informed demand and liquidity demand, and the same is true for the $\mathbf{\Phi}$ matrix.[15] We can thus write,

$$\mathbf{\Phi} = \mathbf{\Phi}_I + \mathbf{\Phi}_L, \tag{29}$$

where $\mathbf{\Phi}_I$ is the covariance of the price effect from informed demand and $\mathbf{\Phi}_L$ is the covariance from liquidity demand. If the number of informed agents is relatively large, the covariance of informed trading, $\mathbf{\Phi}_L$, is primarily caused by a correlation of errors in private information across individuals and securities. The covariance of signals across individuals determines the level of aggregate price noisiness while the level of covariance across securities determines the noisiness of the index level. Such covariances can be seen as a "market mood," a signal shared by all informed investors.

The weight of an individual agent's private signal is determined by how aggressively the agent trades based on private information. The weight increases with the precision of the private signal and decreases with the agent's risk aversion.[16]

[13] In addition, they may be impossible to verify using econometric techniques, due to the effects of nonsynchronous trading.

[14] Explicit results concerning the correlation structure of revealed information are found in Admati (1985) and Caballé and Krishnan (1994).

[15] Liquidity trading is, by assumption, uncorrelated with informed trading.

[16] See, e.g., Hellwig (1980) and Admati (1985).

In a real-world trading situation, most traders only trade in a few securities (relative to the total number traded on the stock exchange). Therefore, $\mathbf{\Phi}$ is most easily seen as the correlation of net demand across securities. The correlation in net demand can originate both from informed trading or liquidity trading. Index arbitrage or basket trading are good examples of trading strategies that induce positive cross-security correlation in net demand, regardless of whether trading is informed or not. Another case is when prices reveal a universal "market mood" or investor sentiment related to a common factor of securities prices.

4.3 Cross-security correlation in short-term returns

The cross-security correlation in short-term returns will be approximately equal to the average of cross-security correlation in the prior and the revealed information. This easily seen by examining the covariance matrix of first stage returns. Standard algebra gives us,

$$\mathrm{Var}\left(\mathbf{r}\right) = \mathrm{Var}\left(\hat{\mathbf{\Omega}}\mathbf{F}\right) = \hat{\mathbf{\Omega}}\left(\mathbf{\Pi}+\mathbf{\Phi}\right)\hat{\mathbf{\Omega}}. \tag{30}$$

As $\hat{\mathbf{\Omega}}$ is diagonal, it is clear that the cross-security correlation of one-period returns is approximately equal to the average of $\mathbf{\Pi}$ and $\mathbf{\Phi}$.

As discussed above, the covariance matrix of the prior can be assumed to be "close" to the covariance matrix of long-term returns. When observed one-period returns are more strongly correlated than corresponding long-term returns, the revealed signals must be more strongly correlated across securities than the value innovations. Equivalently, the correlation in $\mathbf{\Phi}$ must be larger than the correlation in $\mathbf{\Pi}$.

4.4 Empirical implications

The empirical implications of the model are mostly straightforward. Prices that are informative for other prices will tend to lead other prices. Empirical testing thus primarily requires a theoretical analysis of which prices are most informative for other prices.

As argued above, securities with highly correlated return processes will have relatively stronger correlation of their prior valuations. Therefore, lead-lag effects will be stronger between strongly correlated securities. This is easily applicable to lead-lag effects between stocks. There are several ways of identifying closely related stocks, some of which are trivial. Stocks of companies that share characteristics such as geographic location, industry, size range, volatility or market factor loading, will tend to be more strongly cross-autocorrelated than other stocks.

In particular, this approach can be used when testing lead-lag effects within various financial markets, including option markets, bond markets and commodity markets. For example, lead-lag effects can be expected to be strong between bonds in the same maturity range and between options on the same stock. The argument can be extended to lead-lag effects between different markets, such as the lead-lag effect between stocks and stock options. Using the Black-Scholes formula, it is evident that the value of out-of-the-money options is less correlated with the value of the underlying stocks than in-the-money options. Therefore, the strongest lead-lag effects can be expected between stocks and in-the-money options.

Cross-autocorrelation decreases in the correlation of revealed signals. The revealed signals are, however, not observable. Therefore, signal correlation must be identified theoretically or econometrically. Testable hypotheses are mainly generated by models of investor behaviour. When traders are more prone to cross-security trading strategies, cross-autocorrelation would tend to be weaker. For example, index arbitrage trading will tend to reduce cross-autocorrelation while pure liquidity trading will tend to increase cross-autocorrelation. Theoretical models of the trading strategies can thus be tested using cross-autocorrelation data.

Using econometric methods it is possible to identify when the return dispersion is high. One way is to estimate the cross-sectional return dispersion across all traded securities. However, this method must be used with care, especially if the estimation of return dispersion is performed in-sample. The risk of spurious effects is high, since low return dispersion automatically implies weak lead-lag effects if the cross-security correlation in long-term returns is kept constant.

Another implication of the model is that stocks that reveal highly informative signals tend to lead other securities. An obvious problem of this proposition is that the informativeness of stock prices and signals is not observable. However, since trading reveals information, we expect the most heavily traded securities to be less noisy estimates of the corresponding true value. Returns on liquid securities should thus tend to lead returns on less liquid securities.

4.5 Application: The lead-lag relation between stocks and stock options[17]

As the leverage effect of options (especially of deep out-of-the-money options) should make these attractive to informed investors, Chan et al. (1993) conjecture that option returns should lead stock returns. However, several empirical studies, including Stephan and Whaley (1990), Easley et al. (1993) and Chan et al. (1993), report that returns in the stock market lead returns in the options market. Although there is a feedback from options to stocks, it is generally much weaker.

The predictions of cross-security information aggregation are broadly in line with the empirical evidence of this lead-lag relation. In the model, two separate effects make it more likely that options lead stocks than *vice versa*. Firstly, the value of the option is "more stochastic" than the value of the underlying stock. Its value depends not only on the price of the underlying stock, but also on interest rates and future return volatility, both of which are stochastic. In an option that is deep out-of-the-money, the stock return volatility provides a larger, albeit still small, proportion of option return volatility.

Secondly, the large number of traded option series will reduce the measured lead-lag effects. As an example, it can be mentioned that in Stockholm, more than 20 options, calls and puts combined, with different strike prices and maturities are traded on Volvo B shares alone. Any lead from options to stocks will be diluted by this one-to-many effect.

[17] I thank Robert E. Whaley for suggesting this extension.

Documenting lead-lag relation would therefore be easier if an "option index" was used. This is exactly the approach used by Easley et al. (1993). By adding trading volume in different option series, the authors create an index that measures stock information innovation in option trading more precisely than any individual option price series could.[18] This option index leads stock returns by at least 5 minutes, but the reverse lead extends to at least 20 minutes. Clearly, as noted by Chan et al. (1993), even when using an option index, non-synchronous trading in options, reduces the probability of observing a lead from options to stocks.

When using the model presented in this paper to analyse the lead-lag relation between stocks and options, we must be cautious since options are rarely traded in the kind of pure auction market modelled in this paper. However, as the pricing of the model is equivalent to the zero-profit market makers' pricing decision, we can feel relatively comfortable that the results will also carry over to other market settings.

The key to the lead-lag effects is the relative information revelation in the respective markets. If less information is revealed in the option market, it will tend to lag the stock market. A simple example shows that this can indeed be the case, even under very strong assumptions of trading efficiency.

The underlying value of the stock 1 is V_1. Create an option portfolio with option δ equal to unity. This implies that the option portfolio has the same marginal price response to the change in the stock value (that is – unity). This also implies that the variance of the option prior will be equal to the variance of stock prior, π_1, plus the variance of the prior of implied volatility (π_σ):

$$\mathbf{\Pi} = \begin{bmatrix} \pi_1 & \pi_1 \\ \pi_1 & \pi_1 + \pi_\sigma \end{bmatrix}. \tag{31}$$

Assume that agents investigate the underlying value and volatility separately. Individuals' private signals will then be more noisy for the option than for the stock. Even if agents trade optimally in stock and options markets, this added noise will persist in the realised market price (Admati, 1985). The covariance of revealed information can therefore be written as:

$$\mathbf{\Phi} = \begin{bmatrix} \phi_1 & \phi_1 \\ \phi_1 & \phi_1 + \phi_\sigma \end{bmatrix}. \tag{32}$$

The off-diagonal elements of $\mathbf{\Omega}$ can now be calculated explicitly as in equation 27–28:

$$\omega_{12} = 0, \tag{33}$$

$$\omega_{21} = \frac{\pi_\sigma \phi_\sigma}{|\mathbf{\Pi} + \mathbf{\Phi}|} \left(\frac{\pi_1}{\pi_\sigma} - \frac{\phi_1}{\phi_\sigma} \right). \tag{34}$$

As seen in equation 33–34, this very simple setup implies that stock returns will lead option returns while option returns will not lead stock returns. Although the same information is available among the *agents* in both markets, the realised *prices* contain different information. The strength of the lead depends on the

[18] The fact that Easley et al. (1993) use volume data instead of price data does not matter for the conclusions, assuming, as in standard REE models, that there is a known mapping from excess demand to option price changes.

stock price's informativeness for option valuation. It thus increases in π_1/π_σ and decreases in ϕ_1/ϕ_σ.[19]

The assumption that all traders trade optimally in both markets is obviously fairly strong. However, it shows that an information-based lead from options to stocks must be based on transaction costs or other market imperfections. Given that the informed investor *can* trade in both markets, it will always be optimal to reveal exactly as much private information in both markets (profits in both markets increase with the amount of private information revealed in trading and are unaffected by the amount of information revealed elsewhere).

5 Empirical tests

5.1 Testing under continuous trading

Most stock markets operate on a continuous basis. Although this is perfectly consistent with the model, continuous trading poses some problems for empirical testing of the model. The simultaneity of price discovery is a key element of the theoretical model. Without this simultaneity, the model becomes very difficult to test for two separate reasons. Firstly, if two stocks trade in random order at different points in time, the cross-autocorrelation effects are very similar to those predicted by cross-security information aggregation.[20] Secondly, a large share of observed lead-lag effects can be attributed to improved knowledge of a market factor. If stocks trade nonsynchronously, realised market returns can be observed between trades in a single security. This will bias estimates of cross-autocorrelation.[21]

Therefore, any test of the model must try to reduce the nonsynchronicity of prices. In addition, it is preferable to test the model just after the occurrence of an "information event" where the high return volatility makes the identification of lead-lag effects less sensitive to measurement errors and other noise. For intraday returns, it is probable that cross-stock adjustments will be too small to be identifiable using econometric techniques.

This demand for exact simultaneity strongly restricts the choice of dataset. Clearly, it is possible to assume that prices realised within, say 10 seconds, are *de facto* simultaneous. This may be a correct interpretation of the actual information processing capacity of the market place, but testing should preferably be carried out using data which is not subject to any nonsynchronicity in trading.

The best real world candidate for testing the model is an opening call auction of the kind used at the Toronto Stock Exchange and the Paris Bourse. This market setting closely resembles the model setup of this paper. All limit orders are submitted to the electronic trading system *before* the opening call auction. Prices are then set *simultaneously* for all stocks, i.e., without any nonsynchronicity.

[19] Intuitively, the ratio π_1/π_σ should be quite high. Using the Black-Scholes formula with a given stock price leaves little additional noise in option prices. The ratio ϕ_1/ϕ_σ depends on the price precision in each market, which is hard to judge in the general case.

[20] See, e.g., Fischer (1966), Scholes and Willams (1977) and, in particular, Lo and MacKinlay (1990a,b).

[21] The sign of the bias depends on the chosen specification of price formation.

5.2 Data

The data chosen for the empirical testing comprises opening and closing prices for all stocks traded in the automated CAC system of the Paris Bourse.[22] In order to minimise possible problems of nontrading and low liquidity, the sample is restricted to the 70 most traded stocks on the monthly settlement list (*Reglement Mensuel*). All stocks used have at least 1000 days of price data during the sample period. For the chosen stocks, there are virtually no nontrading days. The sample covers five trading years, 1991–1995.

Opening prices are set in a simultaneous call auction procedure. During the preopening period, starting 08:30 (09:00 until 1992), orders can be freely added and cancelled. Traders observe an indicative opening price based on limit orders entered into the system. However, most executed orders are entered into the system during the last 5-10 minutes of the preopening period. At 10:00, an opening price is calculated and all crossing orders are executed (approximately 5% of total daily trading volume is traded at the opening prices). After the opening, trading is continuous until the close (17:00). [23]

In the Paris market, closing prices are also suitable for testing the model. As in most continuous stock markets, trading is very active in the last minutes of trading. This almost eliminates the problems of nonsynchronous trading. Average non-trading for the stocks in the sample is only a few seconds, as a number of traders in fact compete to trade at the day's last prices. For practical purposes, closing prices can therefore be treated as simultaneously determined.

The main advantage of the dataset is the absence of nontrading. However, the long time period between open and close makes the observation of lead-lag effects less probable. It is possible that cross-security price effects drown in the noise induced by overday trading and overnight information. Except for stocks listed on US exchanges, it is impossible to observe prices after the close, but it is possible to use prices from, say, 15 minutes into the trading day to test the properties of lead-lag effects at open. However, the available intraday time series are too short for a meaningful analysis.

5.3 Methodology

5.3.1 Returns

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:

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

Overnight returns are similarly measured from close to open:

$$r_{i,t}^{\text{night}} = \log\left(P_{i,t}^{\text{open}}\right) - \log\left(P_{i,t-1}^{\text{close}}\right). \tag{36}$$

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.

[22]The data has been provided by SBF–Paris Bourse.

[23]Detailed accounts of the trading procedures are provided by Biais et al. (1995, 1996) and de Jong et al. (1995).

5.3.2 Cross-autocorrelation estimates

Cross-autocorrelations at open are estimated using the regression model

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

while cross-autocorrelation at close is calculated using the regression model

$$r_{i,t}^{\text{night}} = \beta_0 + \beta_1 r_{j,t-1}^{\text{day}} + \varepsilon_{i,t}. \tag{38}$$

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

5.3.3 Cross-sectional testing

As expected from the model and the large number of traded securities on the Paris Bourse, lead-lag effects between individual securities returns are relatively weak. Therefore, testing requires the aggregation of data. Relative to Chan (1993), this paper provides a methodological innovation in investigating the cross-sectional properties of cross-autocorrelation estimates between individual securities pairs.[24]

The cross-sectional approach provides a total of 4830 estimates of pair-wise cross-autocorrelation. As these estimates are produced from a mere 70 return series, there is strong dependence between individual estimates. To alleviate this problem, the cross-sectional estimation (table 3) is performed in two steps. In the first step, 70 separate regressions are estimated, holding either leading or lagging security constant. This provides 70 sets of parameter estimates, the mean of which is reported along with the cross-sectional standard error (across the 70 regressions). Reported significance levels test whether the mean is different from zero.

5.4 Results

5.4.1 Cross-autocorrelation with market return

One of the model's predictions is that stock returns should be more strongly cross-autocorrelated with the market return than with other individual stock returns. Results in table 2 strongly support this hypothesis. Average cross-autocorrelation between individual stock returns is much lower than cross-autocorrelation with the market return (*0.019*, *0.037* versus *0.098*, *0.114*).

The model also predicts that "noisy" securities should be most strongly cross-autocorrelated with the market return. As trading volume can be used as a proxy for price informativeness, the least traded securities should thus lag the market return more strongly than more liquid securities. Since the sample provides a relatively narrow range of liquidity, the effect should be relatively weak within the sample. Empirical results (not reported) show that cross-autocorrelation at open decreases in trading volume. However, at close the effect of trading volume is weakly positive. The most traded securities lag the

[24] Chan (1993) aggregates returns to index series and uses time series methods on the less noisy index series. This approach is also used in a companion paper, Säfvenblad (1997).

Table 2: Return correlations and cross-autocorrelations

	Mean	Median	Min	Max	N
Pairwise correlations					
Overday returns	0.188	0.162	-0.043	0.505	2415
	(0.101)				
Overnight returns	0.131	0.089	-0.500	0.603	2415
	(0.132)				
20-day returns	0.269	0.258	-0.172	0.863	2415
	(0.169)				
Pairwise cross-autocorrelations					
At open	0.019	0.017	-0.335	0.148	4830
	(0.039)				
At close	0.037	0.033	-0.092	0.185	4830
	(0.043)				
Cross-autocorrelation with market return					
At open	0.098	0.105	-0.124	0.279	70
	(0.071)				
At close	0.114	0.111	-0.056	0.267	70
	(0.057)				

Reported cross-autocorrelations are the estimates of β_1 from the regression model: $r_{i,t} = \beta_0 + \beta_1 r_{j,t}$, $i \neq j$. N is the number of distinct security pairs in the sample. 20-day returns are calculated using closing prices. Cross-sectional standard errors in parentheses.

market return more strongly than less traded securities. This is partly the expected result of less idiosyncratic noise in the most liquid stocks, but other trading-based explanations must be used to explain these effects. This question is not pursued any further in this paper.

5.4.2 Cross-autocorrelation and the cross-security correlation of the prior

As discussed above, the correlation in long-term returns can be used as a measurement of cross-security correlation in the common prior valuation of securities (π^m, in the single-factor model). Table 3, panel *a*, reports the results obtained using the correlation in 20-day returns as a proxy for π^m, in a linear specification of the relation between cross-autocorrelation and cross-security correlation of the prior.[25]

Results show that lead-lag effects are significantly stronger between highly correlated securities, as predicted by the model. Parameter estimates of the correlation effect are similar at open and close (*0.049*, *0.053*). Lead-lag effects are, however, still fairly weak even between strongly correlated securities.

[25] Other return lengths yield similar results. Choosing long return lengths minimises the influence between daily cross-autocorrelation and measured return correlation.

Table 3: Cross-autocorrelation as a linear function of return correlation and trading volume

	Average across 70 regressions		
	$\hat{\beta}_0$	$\hat{\beta}_1$	R^2
Panel *a*: Correlation in 20-day returns			
At open	0.003^{*}	0.049^{**}	0.081
	(0.011)	(0.062)	
At close	0.019^{**}	0.053^{**}	0.059
	(0.018)	(0.044)	
Panel *b*: Trading volume of leading stock			
At open	-0.116^{**}	0.078^{**}	0.081
	(0.102)	(0.069)	
At close	-0.257^{**}	0.171^{**}	0.252
	(0.189)	(0.123)	
Panel *c*: Trading volume of lagging stock			
At open	0.190^{**}	-0.102^{**}	0.134
	(0.149)	(0.086)	
At close	-0.160^{**}	0.114^{**}	0.127
	(0.151)	(0.100)	

This table reports the average of 70 separate regressions of the following type. Panel *a*: $[\textit{Estimated cross-autocorrelation}]_{i,j} = \beta_0 + \beta_1[\textit{Estimated correlation in 20-day returns}]_{i,j} + \varepsilon_j$. Panel *b*: $[\textit{Estimated cross-autocorrelation}]_{i,j} = \beta_0 + \beta_1[\log \textit{trading volume for leading stock} /10]_{i,j} + \varepsilon_j$. Panel *c*: $[\textit{Estimated cross-autocorrelation}]_{i,j} = \beta_0 + \beta_1[\log \textit{trading volume for lagging stock} /10]_{i,j} + \varepsilon_j$. Trading volume is measured in millions of FRF per day. There are 69 observations in each regression. Standard errors in parentheses report the standard error of the parameter estimate across the 70 regressions. Significance levels test whether the mean of estimates is different from zero. $^{**}/^{*}/^{\circ}$ Significantly different from zero at the 0.01/0.05/0.10 level.

Table 4: Same-industry effects on cross-autocorrelation

	Correlation in			
	Returns	20-day returns	Cross-auto-correlation†	Stock pairs
Paris overday (at open)				
Same industry	0.212	0.311	0.020	80
	(0.101)	(0.183)	(0.036)	
Different industry	0.188	0.270	0.018	2335
	(0.102)	(0.168)	(0.039)	
Paris overnight (at close)				
Same industry	0.124	0.311	0.028	80
	(0.138)	(0.183)	(0.042)	
Different industry	0.134	0.270	0.037	2335
	(0.132)	(0.168)	(0.043)	

Cross-sectional standard errors in parentheses. †Average cross-autocorrelation (both ways) across stock pairs.

5.4.3 Cross-autocorrelation and the precision of the prior

In panels *b* and *c* of table 3, similar methodology is used to estimate the effects of price precision on lead-lag effects. Assuming that high trading volume implies better price precision, both in prior and revealed information, the theoretical results of section 3.3 predict that high volume stocks will exhibit stronger leads to other securities.

This prediction is confirmed by data both at open and close. The lead of a stock increases significantly in the stock's trading volume. Judging from parameter estimates and R^2, the volume effect is particularly important at close.

It was also predicted that the most liquid securities would exhibit weaker lags to other securities. However, the empirical evidence is mixed. At open, the lag is significantly decreasing in the lagging stock's trading volume. At close, the relation is reversed. The most liquid securities also exhibit the strongest lags to other securities. This is analogous to the increased cross-autocorrelation with the market return mentioned in section 5.4.1. The result is partly due to the most liquid stocks' higher market factor loading, but it is most probable that other effects contribute to this relation. It should also be pointed out that since all stocks in the sample are very liquid, the range of liquidity may not be large enough to detect the basic effect.

5.4.4 Same industry shares

It is also possible to test whether the cross-autocorrelation is stronger between companies with highly correlated return processes, without using correlation data as explanatory variable. One way is to identify closely related shares using another method of identification such as industry classifications. Returns on stocks within the same industry can be expected to be more strongly correlated than other shares, are therefore also expected to exhibit stronger lead-lag effects than other stocks.

The industry effect seems to be quite small for the stocks in this sample. The 20-day correlation between stock returns for stocks within the same industry is *0.311*, compared to *0.270* between other stocks (table 4). Not surprisingly, the difference in cross-autocorrelation is small, and not statistically significant. At close, it is even slightly higher between firms in different industries.

5.4.5 Cross-security correlation in short-term returns

One important feature of the model is that it is sufficiently rich to model both positive and negative cross-autocorrelations in stock returns. This property is not present in alternative models such as the Chan (1993) model or the nonsynchronous trading specification of Lo and MacKinlay (1990a). Both models predict that the cross-security correlation of short-term returns is lower than the correlation in long-term returns for all security pairs.[26]

In the Chan, model this result is obtained because signals (revealed net demand) are uncorrelated across securities. In the Lo-MacKinlay framework, price measurements are delayed. Therefore, for shorter return intervals, measured returns reflect partly different value innovations.

This difference in predictions makes it possible to test whether the model of cross-security information aggregation developed here is a useful extension of the Chan (1993) model or a good complement to the nonsynchronous trading model.

The effect on measured returns is most easily described using simulations. Figure 1 presents a simulation of the relation between correlations in short-term and long-term returns for four separate specifications of multi-security stock markets. Each dot represents the short- and long-term correlation between a pair of securities. For dots below the dashed 45° line, short-term correlation is lower than long-term correlation.[27] The solid line is a least squares regression line, included to help compare different panels.

Panel *a* presents a simulation of the model of cross-security information aggregation, developed in section 3. The average correlation between signals varies between *0.3* and *0.6*. For many stocks with lower cross-security correlation than *0.4* in long-term returns, the short-term correlation is *higher* than long-term correlation.

Panel *b* presents the results of a similar simulation where the correlation in signals has been set to zero, as in the Chan (1993) model. The regression line does not pass exactly through the origin, but it is clear that short-term returns are less strongly correlated than long-term returns for virtually all security pairs.

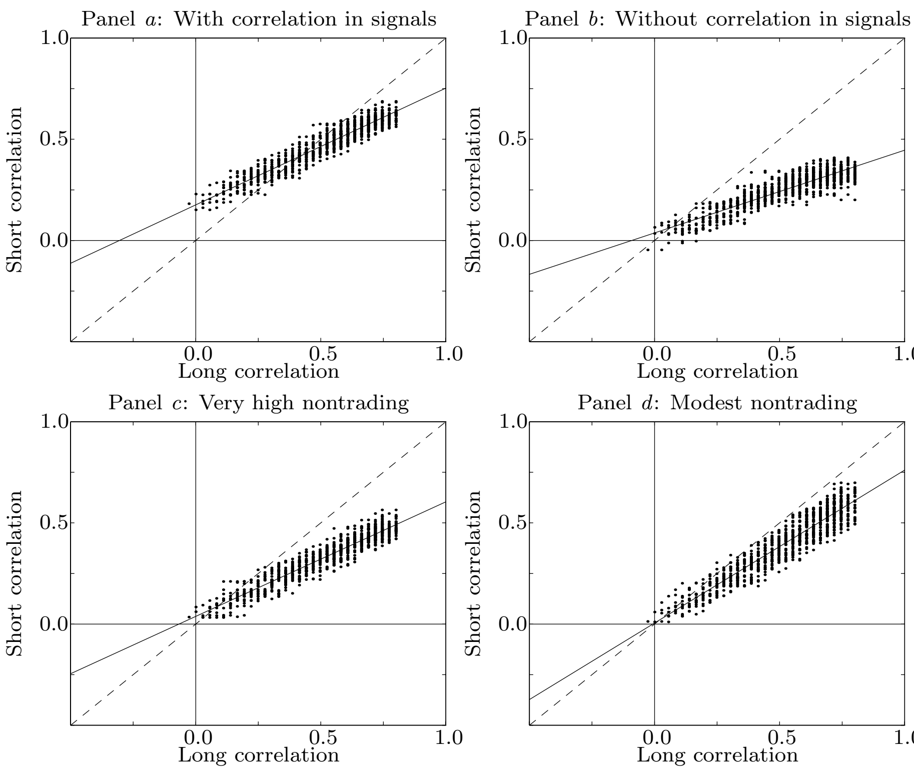
Panels *c* and *d* present simulations of a pure Lo and MacKinlay (1990a) model, where stock prices are nonsynchronous measurements of an underlying value process. In panel *c*, prices are subject to very high nontrading (50% per day); in panel *d*, the nontrading frequency is close to what can be observed empirically for less liquid markets (5%).

The cross-security correlation of overday and overnight returns in the Paris Bourse sample is presented in figure 2. For the overnight returns in panel *b*,

[26] Closer to zero is the exact formulation, but negatively correlated stock returns are very rare. This can, e.g., be seen in figure 2. The prediction can also be formulated in terms of the least squares regression lines in figures 1 and 2. Both models predict a positive slope and an intercept of zero.

[27] An addition of a temporary component, such as bid-ask bounce, does not alter the basic relation between short and long-term correlation for any of the models.

Figure 1: Simulation of the cross-security correlation in short- and long-term returns using four separate specifications of price return behaviour in a multi-security market

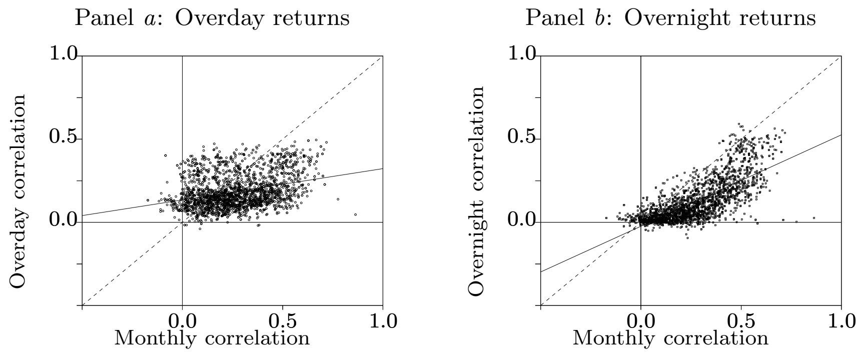


Correlation in 1-period returns plotted against the correlation in 20-period returns for four separate specifications of cross-security return behaviour. Each dot represents the long-term and short-term correlation of a security pair. Dots below (above) the dashed 45° line indicate that securities have higher (lower) long-term than short-term correlation.

Panels *a* and *b* show a simulation of the model of cross security information aggregation. In panel *a*, the cross-security correlation of revealed information varies between *0.3* and *0.6* ($\phi^m/(\phi^m + \phi^s) \in [0.3, 0.6]$). In panel *b*, the correlation of signals is set to zero, similar to the Chan (1993) model. Panel *c* and *d* present the return behaviour under nonsynchronous trading modelled according to Lo and MacKinlay (1990a). In panel *c*, the trading frequency is set to *0.50* per stock per day, spread evenly throughout 25 daily subperiods. In panel *d*, the trading frequency is a more "normal" *0.95* per day.

The solid line is a least squares regression line using all observations. The cross-security correlation of value processes varies between *0.10* and *0.80* for all four panels ($\pi^m/(\pi^m + \pi^s) \in [0.0, 0.8]$).

Figure 2: Cross-security correlation in overday and overnight returns compared to correlation in monthly returns



Cross-security correlation in overday and overnight returns plotted against the correlation in monthly (20-day) returns. Each dot represents a pair of securities (in total 2415). Dots below (above) the dashed 45° line indicate that for this security pair, short-term returns are less (more) strongly correlated than long-term returns. The solid line is a least squares regression line using all observations.

both the Chan model and the nonsynchronous trading model seem to fit the observed relation between short- and long-term returns. However, given the absence of nonsynchronicity in the sample, nonsynchronous trading is obviously not a possible explanation. This conclusion is further strengthened by the large difference between correlation in short- and long-term returns. Nontrading of at least 50% per day is needed to generate a similar pattern.

In the case of overday returns (panel *a*), approximately one third of security pairs exhibit *higher* short-term than long-term correlation. As mentioned above, this cannot be reconciled with the nonsynchronous trading model or the Chan (1993) model.[28]

In order to generate this return structure, revealed information must be strongly correlated across stocks. This implies an initial overreaction to common information for about a third of the stock pairs. These empirical results suggest that order submission in the Paris market is strongly correlated across stocks at close.

To the best of the author's knowledge, this kind of analysis of short-term versus long-term correlation is new. It is therefore impossible to say whether the cross-security return correlations documented in this paper are representative for other stock exchanges.

6 Conclusion

The model of cross-security information developed in this paper provides results that are intuitive and easily adaptable to stock price data. For individual

[28]The same conclusion can be drawn from the fact that the intercept of the least squares regression line is significantly different from zero for overday returns, but not for overnight returns (not reported).

stock returns, the model predicts a well-defined and testable structure of cross-autocorrelations. Securities with informative prices will tend to lead other securities, especially if revealed information is uncorrelated across securities. If the number of stocks sharing a factor is large, the cross-autocorrelation of individual stock returns will be low. However, it can still be significant if stocks share a factor not shared by other securities. Implications can be drawn to lead-lag structures of, for example, stocks, debt, warrants, and options related to a single company. As the model uses information extraction from realised prices, empirical testing is straightforward.

In the empirical section, several of the model's predictions were supported by data from the Paris Bourse. Cross-autocorrelation is higher between securities with highly correlated long-term returns at open and close. More liquid securities, should have less noisy prices, and therefore tend to lead other securities. This conjecture was supported by data both at open and close. However, the most liquid securities were also strongly lagging other securities at close, hinting at very strong reciprocal leads and lags between the most liquid securities. Cross-autocorrelation with the market return is, as predicted, much stronger than cross-autocorrelation between individual securities.

It was also shown that the model presented in this paper can explain observed return patterns better than the alternative models. The Lo and MacKinlay (1990a) model requires unrealistically high nontrading frequencies to explain the big difference between the short-term and long-term correlation of stock returns. The Chan (1993) model, which is nested by the model of this paper, is not sufficiently rich to generate higher short-term than long-term correlation of returns. As this property was observed for roughly one third of Paris overday return pairs, the addition of cross-security correlation in revealed information clearly adds empirical usefulness to the model.

The model of cross-security information aggregation has a large potential for modelling stock price behaviour. The model is set in a general REE framework, which makes it easy to adapt to different trading environments. Moreover, its generality makes it a useful tool for analysing a large number of price discovery issues. In this paper, the lead-lag relation between stock options and corresponding stock returns was discussed. A companion paper, Säfvenblad (1997), studies the implications for stock index returns.

References

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

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

Badrinath, S. G., Jayant R. Kale, and Thomas H. Noe, "Of Shepherds, Sheep, and the Cross-Autocorrelations in Equity Returns," *Review of Financial Studies*, 1995, *8* (2), 401–430.

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.

Brennan, Michael J., Narasimhan Jegadeesh, and Bhaskaran Swaminathan, "Investment Analysis and the Adjustment of Stock Prices to Common Information," *Review of Financial Studies*, 1993, *6* (4), 799–824.

Caballé, Jordi and Murugappa Krishnan, "Imperfect Competition in a Multi-Security Market with Risk Neutrality," *Econometrica*, May 1994, *62* (3), 695–704.

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.

_ , Y. Peter Chung, and Herb Johnson, "Why Option Prices Lag Stock Prices: A Trading-based Explanation," *Journal of Finance*, December 1993, *48* (5), 1957–1967.

de Jong, Frank, Theo Nijman, and Ailsa Röell, "A Comparison of the Cost of Trading French Shares on the Paris Bourse and on SEAQ International," *European Economic Review*, August 1995, *39* (7), 1277–1301.

Easley, David, Maureen O'Hara, and P. S. Srinivas, "Option Volume and Stock Prices: Evidence on Where Informed Traders Trade," 1993. Unpublished Manuscript, Cornell University.

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

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.

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.

McQueen, Grant, Michael Pinegar, and Steven Thorley, "Delayed Reaction to Good News and the Cross-Autocorrelation of Portfolio Returns," *Journal of Finance*, July 1996, *51* (3), 889–919.

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, "Learning the True Index Level: Index Return Autocorrelation in an REE Market," Working Paper No. 190, 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, "The Generality of Spurious Predictability," 1996. Unpublished manuscript, Washington University at St. Louis.

Stephan, Jens A. and Robert E. Whaley, "Intraday Price Change and Trading Volume — Relations in the Stock and Stock Option Market," *Journal of Finance*, March 1990, *45* (1), 191–220.

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