Analytical study of index-coupled herd behavior in financial markets

YONATAN BERMAN, YOASH SHAPIRA and MOSHE SCHWARTZ

School of Physics and Astronomy, Tel-Aviv University - Tel-Aviv, Israel

received 30 August 2016; accepted in final form 11 January 2017
published online 1 February 2017

PACS 89.65.Gh – Economics; econophysics, financial markets, business and management
PACS 89.75.-k – Complex systems
PACS 71.45.Gm – Collective effects: Exchange, correlation, dielectric and magnetic response functions, plasmons

Abstract – Herd behavior in financial markets had been investigated extensively in the past few decades. Scholars have argued that the behavioral tendency of traders and investors to follow the market trend, notably reflected in indices both on short and long time scales, is substantially affecting the overall market behavior. Research has also been devoted to revealing these behaviors and characterizing the market herd behavior. In this paper we present a simple herd behavior model for the dynamics of financial variables by introducing a simple coupling mechanism of stock returns to the index return, deriving analytic expressions for statistical properties of the returns. We found that several important phenomena in the stock market, namely the correlations between stock market returns and the exponential decay of short-term autocorrelations, are derived from our model. These phenomena have been given various explanations and theories, with herd market behavior being one of the leading. We conclude that the coupling mechanism, which essentially encapsulates the herd behavior, indeed creates correlation and autocorrelation. We also show that this introduces a time scale to the system, which is the characteristic time lag between a change in the index and its effect on the return of a stock.

Introduction. – In general, capital markets are assumed to be dominated by exogenous changes rather than by endogenous variables such as human behavior. However, advances in microeconomics and finance have led to the understanding that in practice, financial markets are not perfectly efficient and the effects of human behavior on financial markets are gaining more and more attention [1–4]. Specifically, herd behavior in financial markets became a central research topic in the past few decades [1,5–9], mainly supported by the major advances in the field of behavioral economics [1,10,11], as well as by empirical evidence [12,13].

The importance of human behavior and its effect on the stock market are also supported by the novel approaches of “big-data”, in particular, the analysis of large-scale data regarding financial topics, appearing in the media [14–16]. These examples demonstrate the correlation between stock market moves and behavioral trends reflected in various media channels. In addition, as the recent financial crises brought more attention to the validity of mainstream economic theory, the effects of social tendencies of investors on their financial decisions should be taken into account.

Here we present a model that qualitatively describes some of the short-term behaviors of stock markets, taking into account such effects. The model consists of a set of coupled equations describing the dynamics of the stock market index return and the stock returns. The time series building process involves two terms: The first is a random one, which represents the uncoordinated influence of many economic variables. The complexity of those variables makes them unpredictable and thus expressed in the process by a random variable. The second term is the coupling of the return of the specific stock to the stock index return at an earlier time. The time lag represents the delayed collective response of the investors to the index return. The index could have been taken as a weighted average of stock prices, with time-independent weights. To be concrete and to slightly simplify the analysis, we prefer, however, to take it to be the arithmetic average of stock prices. Our analysis can be taken over directly to the case of general weights.
The model also allows us to explain the index auto-
correlation through herd behavior, along with its expo-
nential decay. Our analysis demonstrates that the herd
behavior notion, which lies at the core of the model is the
assumption allowing for these results. The fast decaying
autocorrelation, for example, that can be found in stocks
intraday returns and is described by the model, cannot
be explained through flow or aggregation of information,
due to the short time lag between consecutive trans-
cessions and the lack of new information. However, it can
be explained by herd behavior. The fact that we obtain
such clear results with very simple and elegant expres-
sions, which all fit to various empirical observations on
stock returns, confirms these qualitative observations and
the model assumptions.

A primitive version of the model was proposed in ref. [9].
In this paper we present a much generalized model with
the model assumptions.

– The model includes a time lag parameter, enabling the
control of the velocity of information flow from the
stock market index, as an aggregate of information,
to the single stocks. Such a new degree of freedom
allows us to detect a complex competing effect of the
time lag and the coupling strength to index on the
stock returns cross-correlations and autocorrelations,
which could not have been detected previously.

– The way our model is now formulated allowed us for
the first time to provide a formal description and
the fourier transform of \( \eta_t \) and \( \Xi_t \) is the
fourier transform of \( \xi_t \). It is possible to write eq. (2) as

\[
A \hat{\eta} = \sigma \Xi,
\]
where

\[
A = \begin{bmatrix}
1 - \frac{\lambda e^{-i\omega \Delta}}{N} & \frac{\lambda e^{-i\omega \Delta}}{N} & \cdots & \frac{\lambda e^{-i\omega \Delta}}{N} \\
-\frac{\lambda e^{-i\omega \Delta}}{N} & 1 - \frac{\lambda e^{-i\omega \Delta}}{N} & \cdots & \frac{\lambda e^{-i\omega \Delta}}{N} \\
\vdots & \vdots & \ddots & \vdots \\
-\frac{\lambda e^{-i\omega \Delta}}{N} & -\frac{\lambda e^{-i\omega \Delta}}{N} & \cdots & 1 - \frac{\lambda e^{-i\omega \Delta}}{N}
\end{bmatrix}
\]

(4)

\( \eta \) is the \( \eta \) column vector and \( \sigma \Xi \) is the \( \sigma_i \Xi_i \) column vector.

By taking the inverse of \( A \), we can now obtain

\[
\hat{\eta} = A^{-1} (\sigma \Xi),
\]
which yields the following equation:

\[
\hat{\eta}_t = \sigma_i \xi_t + \frac{\lambda \varepsilon^{-i\omega \Delta}}{N (1 - \lambda e^{-i\omega \Delta})} \sum_{j=1}^{N} \sigma_j \Xi_j (\omega)
\]

(6)

Tacing the inverse fourier transform we obtain

\[
\eta_t (t) = \sigma_i \xi_t (t) + \frac{\lambda}{N} \sum_{j=1}^{N} \sigma_j \xi_{j-1} \left[ \sum_{k=0}^{\infty} \lambda^k \varepsilon^{-i(k+1)\omega \Delta} \right].
\]

Assuming \( \lambda < 1 \), \( \frac{1}{e^{\omega \Delta} - \lambda} \) can be expanded to the power
series

\[
\frac{1}{e^{\omega \Delta} - \lambda} = \sum_{k=0}^{\infty} \lambda^k e^{-i(k+1)\omega \Delta}.
\]

Substituting this power series in eq. (7) we get the
explicit form of the main equation of the model:

\[
\eta_t (t) = \beta_i = \sigma_i \xi_t (t) + \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{\infty} \sigma_j \lambda^k \xi_j (t - k \Delta).
\]

Results. – Equation (9) demonstrates that the return
of a single stock can be expressed by the noise signals of
the whole market, and is dependent also on the past val-
ues of the noise. However, this depends geometrically on
\( \lambda \). Naturally, when \( \lambda = 0 \), the stocks are uncoupled from
each other completely, and each simply follows a geometric
Brownian motion. We wish to continue from eq. (9)
and use it for describing the correlations and autocorrela-
tions of the stock returns and of the index, induced by the
model.
We assume that

$$\langle \xi_i(t_1) \xi_j(t_2) \rangle = \delta_{i,j} \delta(t_1 - t_2),$$

(10)

reflecting a fast decay of the autocorrelation in the random noise of the stock returns compared to $\Delta$.

Following eq. (9) we obtain that $\langle \eta_i(t_1) \eta_j(t_2) \rangle$ follows:

$$\langle \eta_i(t_1) \eta_j(t_2) \rangle = \sigma_i \sigma_j \delta_{i,j} \delta(t_1 - t_2)$$

$$+ \frac{1}{N} \sum_{k=1}^{\infty} \lambda_k \left[ \sigma_i^2 \delta(t_1 - t_2 + k\Delta) + \sigma_j^2 \delta(t_1 - t_2 + k\Delta) \right]$$

$$+ \frac{1}{N^2} \sum_{m=1}^{N} \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \lambda_k^m \lambda_l \delta(t_1 - t_2 + (l - k)\Delta).$$

(11)

Equation (11) also allows us to write

$$\langle \eta_i(t_1) \eta_j(t_2) \rangle = \sum_{n=-\infty}^{\infty} c_{i,j,n} \delta(t_1 - t_2 + n\Delta).$$

(12)

In addition, for simplicity, we will also assume that $\sigma_i = \sigma, \forall i$. Now, we will examine the two cases—one in which $i \neq j$ and $t_1 = t_2$, and one in which $i = j$ and $t_2 = t_1 + m\Delta$ ($m \in \mathbb{N}$), representing cases of cross-correlation and autocorrelation, respectively.

Cross-correlations. Following eq. (11), it is possible to calculate the cross-correlation between the stock returns and between the stock returns and the index. For that purpose we use the Pearson correlation.

When $i \neq j$ and $t_1 = t_2$ and since $\langle \eta_i(t_1) \rangle = 0$ we obtain

$$\rho_{i,j} = \frac{\langle \eta_i(t_1) \eta_j(t_1) \rangle}{\sqrt{\langle \eta_i(t_1)^2 \rangle} \sqrt{\langle \eta_j(t_1)^2 \rangle}} = \frac{\sigma^2 N \lambda^2}{\sigma^2 \left(1 + \frac{1}{N} \sum_{k=1}^{\infty} \lambda_k^2 \right)^2} = \frac{\lambda^2}{N (1 - \lambda^2) + \lambda^2}.$$  

(13)

We will also calculate the correlation between the return of an arbitrary stock to the index return which follows

$$\rho_i^I = \frac{\langle \eta_i(t_1) + \sum_{j=1}^{N} \eta_j(t_1) \rangle}{\sqrt{\langle \eta_i(t_1)^2 \rangle} \left(\frac{1}{N} \sum_{j=1}^{N} \langle \eta_j(t_1)^2 \rangle \right)^{1/2}}.$$  

(14)

We can simplify the different expressions as follows:

$$\langle \eta_i(t_1) + \frac{1}{N} \sum_{j=1}^{N} \eta_j(t_1) \rangle = \frac{1}{N} \sum_{j=1}^{N} \langle \eta_i(t_1) \eta_j(t_1) \rangle$$

(15)

and

$$\left( \frac{1}{N} \sum_{j=1}^{N} \eta_j(t_1) \right)^2 = \frac{1}{N} \sum_{i=1}^{N} \langle \eta_i(t_1) \frac{1}{N} \sum_{j=1}^{N} \eta_j(t_1) \rangle$$

$$= \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \eta_i(t_1) \eta_j(t_1) \rangle.$$  

(16)

Fig. 1: (Color online) The dependence of the return correlation on the coupling strength. The correlation between the return of two different stocks with respect to $\lambda$ (blue) for $N = 30$. Also presented is the correlation between the return of one stock to the index return (red).

Following the above we obtain

$$\rho_i^I = \frac{1}{\sqrt{\sigma_i^2 \sum_{j=1}^{N} \sum_{k=1}^{N} \lambda_k^2}} \frac{1}{\sqrt{\sigma^2 \sum_{j=1}^{N} \sum_{k=1}^{N} \lambda_k^2}} \frac{1}{\sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_k^2}}$$

$$= \frac{1}{\sqrt{N (1 - \lambda^2) + \lambda^2}} = \frac{1}{\sqrt{N (1 - \lambda^2) + \lambda^2}}.$$  

(17)

In both cases, the correlations between the stock returns and between the stock returns and the index return are independent of $\Delta$. In addition, these correlations naturally increase as the coupling strength increases as illustrated in fig. 1. What is also depicted in fig. 1 is that $\rho_i^I \geq \rho_{i,j}$. Typically, the correlation between the stock returns in financial markets is about 0.2–0.5, corresponding to $\lambda$ values which are close to 1. We will elaborate on this in the “Discussion” section.

Using the obtained correlations we can also analyze the correlation matrix of the stock returns. It is well known that the correlation matrix is characterized by a single large eigenvalue followed by substantially smaller eigenvalues [17,18]. Following our model the correlation matrix has two distinct eigenvalues $-\alpha_1 = 1 + \frac{N \lambda^2}{\lambda^2 + N (1 - \lambda^2)}$ and $\alpha_2 = 1 - \frac{\lambda^2}{\lambda^2 + N (1 - \lambda^2)}$. $\alpha_1$ is obviously larger than $\alpha_2$, as $\alpha_1 - \alpha_2 = \frac{(N+1) \lambda^2}{N (1 - \lambda^2)} > 0$. Figure 2 demonstrates the difference between the two eigenvalues and illustrates that as $\lambda$ increases, the difference between the eigenvalues can be better identified. In practice, the real correlations between the stock returns correspond to relatively large $\lambda$ values. This, along with the model implications, supports the observation of one large eigenvalue in the data.
**Partial correlations.** The observed effects of the index on the behavior of stocks also reveal a characteristic behavior of partial correlations. It is well known that the partial correlation between stock returns with respect to another stock is high compared to the low partial correlations with respect to the index [19–21]. Following eqs. (13) and (17), the partial correlation between the returns of any two stocks with respect to an arbitrary third stock in our model is

$$ρ_{i,j,k}^{\text{partial}} = \frac{ρ_{i,j} - ρ_{i,k}ρ_{j,k}}{\sqrt{(1 - ρ_{i,j}^2)(1 - ρ_{j,k}^2)}}$$

$$= \frac{ρ_{i,j} - ρ_{i,k}ρ_{j,k}}{\sqrt{(1 - ρ_{i,j}^2)(1 - ρ_{j,k}^2)}} \frac{1}{N(1 - λ^2) + 2λ^2}.$$  

Similarly, the partial correlation between the returns of two stocks with respect to the index would be

$$ρ_{i,j,I}^{\text{partial}} = \frac{ρ_{i,j} - ρ_{i,I}ρ_{j,I}}{\sqrt{(1 - ρ_{i,j}^2)(1 - ρ_{j,I}^2)}}$$

$$= \frac{ρ_{i,j} - ρ_{i,I}ρ_{j,I}}{\sqrt{(1 - ρ_{i,j}^2)(1 - ρ_{j,I}^2)}} \frac{1}{1 - (λ^2)} = \frac{1}{1 - N}.$$  

**Autocorrelations.** While the Pearson correlation between the stock returns and between the stock returns and the index return is independent of Δ, we can also calculate the autocorrelation of the return of an arbitrary stock or of the index. These would likely depend on Δ, the time delay between the change in the index and its effect on the stock price, as originally defined in eq. (1).

Following eq. (11), if $i = j$ and $t_2 = t_1 + mΔ$ $(m \in \mathbb{N})$, we obtain

$$⟨\dot{η}_i (t_1) \dot{η}_i (t_1 + mΔ)⟩ = \sum_{n=-\infty}^{∞} c_{i,i,n} δ((n + m) \Delta).$$  

Therefore, we can now calculate the Pearson autocorrelation $ρ_i (mΔ) = \frac{⟨\dot{η}_i (t_1) \dot{η}_i (t_1 + mΔ)⟩}{\sqrt{⟨\dot{η}_i (t_1)^2⟩⟨\dot{η}_i (t_1 + mΔ)^2⟩}}$, in the same way previously done for the cross-correlations. It is implied by eq. (20) that it does not depend on $t_1$ but only on $t_2 - t_1 = mΔ$ and in addition $⟨\dot{η}_i (t_1)^2⟩ = ⟨\dot{η}_i (t_1 + mΔ)^2⟩$,

$$ρ_i (mΔ) = \frac{⟨\dot{η}_i (t_1) \dot{η}_i (t_1 + mΔ)⟩}{\sqrt{⟨\dot{η}_i (t_1)^2⟩⟨\dot{η}_i (t_1 + mΔ)^2⟩}}$$

$$= \frac{σ^2}{1 + σ^2} \frac{⟨\dot{η}_i (t_1)^2⟩(⟨\dot{η}_i (t_1 + mΔ)^2⟩)}{N(1 - λ^2) + 2λ^2} = \frac{λ^m}{N(1 - λ^2) + 2λ^2}.$$  

Fig. 2: (Color online) The eigenvalues ($α_1$ in blue, $α_2$ in red) of the stock returns correlation matrix with respect to $λ$, the coupling strength, for $N = 30$ (top) and $N = 500$ (bottom).

Fig. 3: (Color online) The dependence of the partial correlation on the coupling strength. The partial correlation between the return of any two different stocks with respect to another arbitrary stock for different values of $λ$ with $N = 30$ (blue). Also presented is the partial correlation between the return of two stocks with respect to the index return (red).

These expressions demonstrate that indeed $ρ_{i,j,k}^{\text{partial}} > ρ_{i,j,I}^{\text{partial}}$, as expected. This is also depicted in fig. 3. However, $ρ_{i,j,I}^{\text{partial}}$ is negative for any $N > 1$. In practice the partial correlation with respect to the index is usually positive but very close to 0, and can also be negative [19]. Adding variability to the stock weights in the index and to the value of $λ$ would allow $ρ_{i,j,I}^{\text{partial}}$ to be also positive. We also note that in practice, the partial correlations between stock returns with respect to the index have large deviations, with standard deviation far greater than the mean.
Analytical study of index-coupled herd behavior in financial markets

Fig. 4: (Color online) The dependence of the return autocorrelation of a single stock on the lag. The calculations were done for \( N = 30 \) for different values of \( \lambda \), considering \( \frac{\lambda}{\Delta} \) running from 0 to 100.

Equation (21) can be simply generalized for any lag \( \tau \):

\[
\rho^I (\tau) = \frac{\lambda^2}{N(1 - \lambda^2) + \lambda^2}. \tag{22}
\]

Equation (22) implies that the autocorrelation of any stock return decays exponentially in time, with correlation time \( \Delta \). In addition, with \( \lambda \) close to 0, the autocorrelation is very small, though such a case is practically irrelevant, as the values of \( \lambda \) in actual systems are closer to 1 (see fig. 1).

We can conclude that the coupling to the index, or the herd behavior encapsulated in the model can create the exponentially decaying autocorrelation. In addition, it provides a time scale to the system dynamics, which is the characteristic time in which investors and traders look upon, either literally or by using high-frequency trading tools.

The propagation of the autocorrelation of a single-stock return in time for different values of \( \lambda \) is given in fig. 4.

One can also repeat the calculation of the return autocorrelation for the index rather than for a single stock:

\[
\rho^I (m\Delta) = \frac{\lambda^2}{N(1 - \lambda^2) + \lambda^2}. \tag{23}
\]

Since it follows from the above definitions that

\[
\left\langle \left( \frac{1}{N} \sum_{j=1}^{N} \eta_j (t_1) \right)^2 \right\rangle = \left\langle \left( \frac{1}{N} \sum_{j=1}^{N} \eta_j (t_1 + m\Delta) \right)^2 \right\rangle \tag{24}
\]

and

\[
\rho^I (m\Delta) = \left( \frac{\lambda^2}{N} \right) \left\langle \sum_{j=1}^{N} \eta_j (t_1) \eta_j (t_1 + m\Delta) \right\rangle \tag{25}
\]

we obtain the simple expression

\[
\rho^I (m\Delta) = \lambda^m. \tag{26}
\]

Similarly to eq. (22), we can generalize eq. (27) for any lag \( \tau \):

\[
\rho^I (\tau) = \frac{\lambda^2}{N(1 - \lambda^2) + \lambda^2}. \tag{27}
\]

This derivation demonstrates again that the coupling mechanism, or the herd behavior described by the model generates the exponential decay of the autocorrelation. Numerical results for the autocorrelation of a single stock and the index are presented in fig. 5.

**Discussion.** We presented a simple dynamic model of herd behavior in financial markets, in which the return of a single stock is coupled to the index return added by random noise, and assuming a fixed time lag between the change in the index and the change in the stocks. The model consists of a system of coupled stochastic differential equations, which allows us to calculate the correlations and autocorrelations of stock returns, without the necessity to solve the equations. Our results demonstrate the existence of high correlations and autocorrelations in the stock and index returns. Moreover, we obtain an exponential decay of the return autocorrelation, similar to the intraday autocorrelation found in real markets [9,22,23]. This joins other works by suggesting that herd behavior is the key contributor to this important characteristic of financial markets. While we use herd behavior to detect specifically correlations and autocorrelations, other models which discuss herd behavior identify other phenomena, such as the contribution of herd behavior to bubbles [5] and financial crises [24,25], and the Epps effect [9]. Others, however, argue that correlated trades found in financial markets “is mainly due to the common reaction of
investors to new public information and should not be misinterpreted as herd behavior" [26]. We also note that other models and in particular the capital asset pricing model (CAPM), traditionally used, cannot detect such correlations and autocorrelations, even in its extended forms [27]. We note, however, that the cross-sectional absolute deviation method, introduced in [13] as an extension of CAPM, is used for statistically testing the existence of herd behavior in data [6,8,13].

Our analysis takes into account several simplifying assumptions that enable the analytic derivation of the correlations, but also affect the results. In order to understand the limits of the analysis one should address the effect of these assumptions:

- $\sigma$ was considered as fixed, while in practice each stock has different volatility. In practice it will not affect the index correlation and autocorrelation, but will have an effect on the single stocks. For example, if one stock is extremely volatile compared to the rest, this stock will show lower correlation to the index and to the rest of the stocks.

- $\lambda$ was considered as fixed, while in practice, stocks may be coupled differently to the index. For example, stocks of oil companies may be strongly coupled to the oil price and less strongly to the index price [28], while others, like bank stocks are likely to be strongly coupled to the stock market index. The effect of varied coupling strengths will make large deviations in the correlations between different stocks and between specific stocks and the index, and it will not be possible to obtain expressions as simple as eq. (13) in such a case. However, the exponential decay of the autocorrelation in time will still occur.

- We considered the index return as the arithmetic mean of the single-stock returns. This assumption may be valid for some indices, however, in most cases stocks have different weights according to their market cap. This may change quantitatively the value $\lambda$ required for the same correlation values and it will also make the results dependence on $N$, the stocks number, much less dominant. However, it will likely have no qualitative effect on the results.

- We also considered the time delay $\Delta$ to be uniform. Mathematically, $\Delta$ is used for describing the time it takes to a change in the index to reflect in single stocks originated in herd behavior, which can be identified also as the characteristic time between trades. Our analysis shows that this time lag becomes a characteristic time of the system in general, also determining the relaxation time of the return autocorrelation. Assuming different $\Delta$ values can be empirically fitted for different stocks, the qualitative behavior of the autocorrelation exponential decay in time will still be applicable. This was demonstrated numerically in [9].

This analysis also provides a possible explanation for the faster time scales characterizing stock markets today when compared to those of a few decades ago.

In addition, many argue that fast automatic intraday trading, and in particular high-frequency trading (HFT), led to a faster decay of autocorrelations of stock returns [29–31]. This can be explained simply using the model by claiming that nowadays machines are looking at the trend of the index or of the collective behavior more rapidly than in the past. Our study therefore hints that HFT generally increases market efficiency by closing fast arbitrage gaps and by effectively removing autocorrelation [29,32,33].

Following the above, we should note that there is a qualitative difference between the effect of coupling in stock returns and in stock prices. These two coupling mechanisms may result in very different long-term behaviors and may be independent of one another [28]. For example, the stock price of two oil companies traded in two different stock markets may be similar in the long run due to the global effect of the oil price. However, on the short run, they are likely to behave differently and closer in their daily return to their relevant stock market index return. The model described in this paper corresponds to the shorter time scales, in which the overall trend reflected by the index return affects the behavior of single stocks. Sectorial external forces, such as commodity prices, are known to have a substantial effect on longer time scales on stock prices [28].

In fig. 1 we have shown that in order to obtain correlations that are close to realistic for the US stock markets, for example, it would require a very large coupling parameter. In practice, applying our analysis to data and specifically estimating the parameters in order to fit real market data, requires allowing $\Delta$, $\lambda$, $\sigma$ and the weight of each stock in the index to be different. This is left for future work. Establishing the above, an important direction in which this work can be continued is its integration with the long-term coupling model of stock prices and sector-specific driving forces [28].

REFERENCES

Analytical study of index-coupled herd behavior in financial markets


