Modelling the price dynamics of a stock market

Asher Mullokandov*

May 08, 2015

Abstract

We describe a model of daily stock market return dynamics based on investor behavior. This model assumes a simple set of underlying dynamical equations describing the total wealth and cash of investors, as well as their assets, and introduces parameters that describe investor behavior. Using methods from statistical physics, as well as empirical analysis of market data, we find that the daily returns for a wide range of financial institutions obey the second order differential equation which describes the dynamics of a damped harmonic oscillator. We analyze this model and describe "calm" and "frantic" regimes separated by a phase transition which is captured by the frequency term of the model. Furthermore, we introduce a general method for probing the influence of external factors on market dynamics via the influence on investor behavior.

1 Introduction

The question of what drives stock price movements is a fundamental one in the theory of financial markets, and one which has profound implications for forecasting and managing financial crises, as well as for foundational economic issues. Understanding the dynamics of a stock price can offer insights into the fundamental rules which govern the operation of a market, and furthermore, may provide insight into the internal and external factors that influence market behavior. In this paper we describe a theoretical and empirical study of daily returns and propose a general model of stock market return dynamics based on investor behavior which accurately describes the daily return responses observed in real-world markets. Moreover, our model naturally incorporates a way of precisely specifying how external factors, in particular the news, enter into the system and allows us to easily test the relation between exogenous noise and various parameters and features.

2 Preliminary Tests of News Influence

A possible first step is to directly test the correlation between news data and price data. It’s reasonable to suppose that changes in news sentiment would have some measurable influence

*asher@bu.edu
Figure 1: On the left, the sentiment time series for Bank of America, and on the right the sentiment time series for JP Morgan

on price movement, though it is certainly not the case that we can explain the majority of price dynamics simply by reading the news. The news data which we will use is a very large, newly created, database of English language news from around the world. The data include the number of occurrences per day in the English language news of thousands of financial institutions (eg, banks, insurance companies, national banks) as well as economic terms, currencies, countries, etc. To each article is also associated a sentiment score, calculated as the number of positive words (as defined by the Harvard IV-4 sentiment dictionary) minus negative words, divided by the total sentiment per article. Figure 1 displays two example sentiment time series, for Bank of America and JP Morgan. We can clearly see, for example, days of extremely poor news, which correspond to the announcement of severe fines imposed by the US government.

The simplest preliminary analysis would be to check the same-day and lagged correlation between the news data and the daily returns. We do this and note that it is possible for news to drive price changes, but also for price changes to drive the news. These two cases would be distinguished by the peak correlation corresponding to a 0 or 1 day forward lag in the first case, and a -1 day lag in the second case.

Figure 2 displays two examples of such a comparison between sentiment and return, on the left is Bank of America and on the right is the Australia and New Zealand Banking Group. We can see that the largest correlation is for the same day, which achieves a maximum of $\sim30\%$.

We wish, however, to model the influence of the news in a much more sophisticated and detailed way, and also to understand the interplay between the external noise and internal dynamics which affect stock markets. To do this, we must move beyond simple comparisons and begin to probe deeper into the underlying rules that govern this complex system.
Figure 2: The lagged correlation between the sentiment and return time series for Bank of America and the Australia and New Zealand Banking Corporation

3 The model

To begin, we consider a very simple general equation for the returns,

$$\Delta_t \partial_t p \approx \mathcal{F}(N)$$

where $\Delta_t$ denotes a differential operator, $p$ the price of an individual stock, and $\mathcal{F}(N)$ a function of external parameters (e.g., the news). The operator $\Delta_t$ determines the time evolution of the returns, and we now wish to specify its structure. We may do so by empirically determining the lagged autocorrelation of the returns, $C(t, \Delta t)$. Figure 1 displays the lagged autocorrelation for a number of major international banks, calculated for a window of 100 days. The figure on the left is the average of the banks with the largest one-day drop, and the figure on the right is the average for the bank with the smallest one day drop. Fitting this data yields the following function,

$$C(t, \Delta t) = \int dt \partial_t P(t + \Delta t)P(t) \approx \exp \left[-\frac{\Delta t}{2\tau}\right] \cos(\omega \Delta t)$$

which we recognize as the Green’s function for the damped harmonic oscillator:

$$(\tau \partial_t^2 + \eta \partial_t + \omega^2) \partial_t P(t) = F(t)$$

This empirical analysis suggests that the daily return’s are subject to a damped harmonic oscillator. Moreover, the parameters in the left hand side of the above equation are critical in determining the details of the dynamics, and offer a possible way for external influences such as the news to impact the system.

This is, however, a slightly incomplete picture. We have extracted the form of the operator from data, but we would like to understand why the returns exhibit this particular structure at a more “microscopic level.” To accomplish this, we propose a model which describes the market at a deeper level: the investors who ultimately make trading decisions, and on whom external news actually acts.
Figure 3: Return autocorrelation. On the left is the average of the banks with the largest one-day drop in autocorrelation, on the right the average of the ten banks with the smallest one-day drop.

Consider an investor with equity $E$, who owns assets $A$ with prices $p$, and who has access to cash $c$. Below are the dynamical equations which we assume the investor obeys, with $\tau_A$ and $\tau_E$ the response times of the assets and equity, and $S$ and $O$ representing external news sentiment and occurrence or, in fact, any other influence which impacts investor decisions:

$$\frac{\delta A(t + \tau_A)}{A} = \beta \left( \frac{\delta E(t + \tau_E)}{E} + S \right)$$

$$\frac{\delta p(t + \tau_p)}{p} = \alpha \left( \frac{\delta A(t)}{A} + S \right)$$

$$\delta E(t) = A \delta p(t)$$

We assume that $S$ and $O$ affect both the trading of the brokers (first equation) and the bidding on the prices. In addition, we introduce the "behavioral" parameters $\alpha$ and $\beta$ which can themselves be affected by the news:

$$\alpha = \alpha(O, S, \partial_t O), \quad \beta = \beta(O, S, \partial_t O)$$

For small $\tau_x$ the first two equations become second order equations in $\partial_t$. Our goal is to find an effective equation for $p$, and we will accomplish this via a "mean field" approximation for the equity $E$.

### 3.1 $\tau_E = 0$

First suppose that $\tau_E = 0$. The equity $E$ is comprised of assets $Ap$ plus some cash $c$:

$$E = Ap + c, \quad \partial_t c = \partial_t Ap$$

Let’s also assume that $Ap \gg c$, i.e. most of the capital is in the form of these assets $E \approx Ap$. We can now write the variational equations we assumed above as second order differential
equations:

\[
\left( \frac{\partial_t + \tau_A \partial_t^2}{A} \right) A = \beta \frac{\partial_t E}{E} + \beta S = \beta \frac{A \partial_t p}{A p} + \beta S
\]  \hspace{1cm} (6)

\[
\frac{\left( \frac{\partial_t + \tau_p \partial_t^2}{p} \right) p}{p} = \alpha \frac{\partial_t A}{A} + \alpha S
\]  \hspace{1cm} (7)

We now derive the equation for the returns. Consider \( \alpha \)Eq(6) + Eq(7) + \( \tau_A \partial_t E \)Eq(7):

\[
- \alpha \tau_A p \partial_t S + \frac{\alpha \tau_A (\partial_t A)^2}{A^2} p
- (\alpha + \gamma)Sp - \frac{\tau_A u^2}{p} - \frac{\tau_A \tau_p u \partial_t u}{p}
+ \left[ \tau_A \tau_p \partial_t^2 + (\tau_A + \tau_p) \partial_t + (1 - \gamma) \right] u = 0
\]  \hspace{1cm} (8)

where \( u = \partial_t p \).

Using Eq(7) we can also get rid of \( \partial_t A/A \)

\[
\tau_A \tau_p \partial_t^2 u + (1 - \gamma - 2 \tau_A S) u
+ (\tau_A + \tau_p - 2 \tau_A \tau_p S) \partial_t u
= (-\alpha \tau_A S + \alpha + \gamma)Sp + \alpha \tau_A p \partial_t S
+ \frac{\tau_A}{\alpha p} \left( (\alpha - 1)u^2 + (\alpha - 2)\tau_p u \partial_t u - (\tau_p \partial_t u)^2 \right)
\]  \hspace{1cm} (9)

The resulting equation is precisely the differential equation for a damped harmonic oscillator:

\[
[\tau \partial_t^2 + \eta \partial_t + \omega^2] u = aSp + b(\partial_t S)p + O \left( u^2, \partial_t u^2 \right)
\]  \hspace{1cm} (10)

This is a remarkable finding: using our microscopic model of investor behavior, we have derived the very differential operator which we had previously extracted from market data.

The variables and parameters above are,

\[
\frac{1}{\tau} = \frac{1}{\tau_A} + \frac{1}{\tau_p}, \hspace{0.5cm} \eta = 1 - 2 \tau S, \hspace{0.5cm} \omega^2 = \frac{1 - \gamma - 2 \tau_A S}{\tau_A + \tau_p},
\]

\[
a = \frac{\gamma + \alpha - \alpha \tau_A S}{\tau_A + \tau_p}, \hspace{0.5cm} b = \frac{\tau_A \alpha}{\tau_A + \tau_p}, \hspace{0.5cm} p(t) = \int^t u(t') dt'
\]  \hspace{1cm} (11)

Additionally, we have found a non-linear term,

\[
O \left( u^2, \partial_t u^2 \right) = \frac{\tau_A ((\alpha - 1)u^2 + (\alpha - 2)\tau_p u \partial_t u - (\tau_p \partial_t u)^2)}{\alpha p(\tau_A + \tau_p)}
\]  \hspace{1cm} (12)
With the microscopic model, the dynamics are richer and we have a damped oscillator with a driving force coupled to $p$ and nonlinearities of type $\sim (\partial_t p)^2$. Taking the return $u \equiv \partial_t p$ as the fundamental variable, the nonlinearities are roughly of type $u^2 + a\partial_t u^2$. A key feature to note is that the frequency term $\omega^2$ depends on $(1 - \gamma - 2\tau_A S)$. For $\gamma < 1 + 2\tau_A S$ this term has the usual stable damped oscillator solutions, while $\gamma > 1 + 2\tau_A S$ will result in instabilities and divergent solutions. In a market interpretation, the stable phase would be one of relative "calm" during which gains and losses would be comparatively small and the returns should display clear oscillatory behavior, and the unstable phase would be a "paniced" one during which we would expect to see crashes or bubbles. In this phase, the decay or growth of the Green’s function would be exponential. We are very much looking forward to further investigating this fascinating behavior, and to incorporating the effects of other investors via couplings.

4 Conclusions and future work

We have described a theoretical model of investor behavior which produces an equation for the dynamics of daily returns which is familiar to physicists: a damped harmonic oscillator. We confirm this behavior by analyzing stock market data for a wide range of financial institutions across different time periods. Our model is micro-economic in nature and accounts for the network of investors in the market, provides systemic information about the macro-economic behavior, and incorporates in a natural way both endogenous and exogenous factors which influence market behavior. In particular, we plan to use this model to probe quantitatively the impact of external financial news on price dynamics, and develop a theoretical framework for testing the efficient market hypothesis. The model has natural "calm" regimes (corresponding to a real valued $\omega$), where market movements are slow and losses and profits are small, and "frantic" regimes (corresponding to an imaginary $\omega$), in which returns are exponential and either bubbles form or crashes happen. As in real markets, these regimes are distinct and separated by a phase transition. In addition to providing fundamental insight into the dynamics of prices, our model can identify parameters which serve as an early warning tool for detecting system-wide dynamics which lead to crashes.

5 Acknowledgements

I thank Nima Dehmamy for close collaboration on this project, and Irena Vodenska for many fruitful conversations. This paper was prepared for Gene Stanley’s Econophysics course, Spring 2015.

A Green’s Function

Let us briefly review Green’s function methods as applied to a related but slightly simpler operator. Given a linear equation of the form:

$$\tau \partial_t^2 P + \partial_t P \approx F$$
We can solve for $\partial_t P$ using the Green’s function:

$$\partial_t P(t) = \int dy G(t - y) F(y)$$

The Green’s function can be found from inverting the operator in the equation.

$$G(t) = \int \frac{dk}{2\pi i\tau k + 1} e^{ikt} = \frac{1}{\tau} \exp[-t/\tau] \approx \int dy \langle \partial_t P(t - y) \partial_t P(y) \rangle = \exp[-|t|/\tau] \quad (13)$$

If we want to check if only the past values of $F(t)$ affect $\partial_t P(t)$, we need to use the “retarded Green’s function” $G_R$. This means that if $t - y < 0, G_R(t - y) = 0$. Therefore, we need to include a step function into $G(t)$:

$$G_R(t) = \theta(t) G(t)$$

Using this, the lagged inner product of the prices becomes:

$$\int dt \partial_t P(t) \partial_t P(t + \Delta t) = \int dt dy dy' G_R(t - y) F(y) G_R(t + \Delta t - y') F(y')$$

$$= \frac{1}{\tau^2} \int dt dy dy' \exp[-(2t + \Delta t - y' - y)/\tau] \theta(t - y) \theta(t + \Delta t - y') \quad (14)$$

This integral can be broken down to two pieces: one where $y > y' - \Delta t$ and one for $y < y' - \Delta t$. redefining $t - y \rightarrow y$ we get:

$$\int dy F(y) G_R(t - y) = \int_0^{\infty} dy G(y) F(t - y)$$

and similarly for $t + \Delta t - y' \rightarrow y'$. Thus we have:

$$\int dt \partial_t P(t) \partial_t P(t + \Delta t)$$

$$= \frac{1}{\tau^2} \int dt \int_0^{\infty} dy dy' F(t - y) F(t + \Delta t - y') e^{-(y + y')/\tau}$$

$$= \frac{1}{\tau^2} \int dt \int_0^{\infty} dy dy' \exp[-(y + y')/\tau] F(t) F(t + \Delta t + y - y')$$

$$= \frac{1}{\tau^2} \int dt \int_{-\infty}^{\infty} dx \int_0^{\infty} dy \exp[-(2y + \Delta t - x)/\tau] F(t) F(t + x)$$

$$= \frac{e^{-\Delta t/\tau}}{2\tau} \int_{-\infty}^{\infty} dx e^{x/\tau} \int_{-\infty}^{\infty} dt F(t) F(t + x)$$

$$= C \frac{e^{-\Delta t/\tau}}{2\tau} \quad (15)$$

Where we defined $x = \Delta t + y - y'$ which gives $y + y' = y' - y + 2y = 2y - x + \Delta t$. The constant $C$ is the result of the remaining integrations, but as long as it is finite and nonzero the above is suggesting that the convolution of price returns with itself should fall as an exponential. If we also normalize the convolution by dividing out $\int P^2(t) dt$ the constants $C/2\tau$ go away and only the exponential part remains.