Revisiting stock market index correlations

Mehmet Dalkir
University of New Brunswick,
Department of Economics
P.O. Box 4400
Fredericton, NB E3B 5A3
Phone: (506) 453-4858
Fax: (506) 453-4514
e-mail: mdalkir@unb.ca
Abstract

Comovement of stock market indices increases during crashes, and does not come down when the turmoil settles down. This paper explains upgrade of comovements during turmoil periods with theories from Bayesian learning and dynamical systems involving synchronization. Our main conclusion is that the correlation does not go down because it is learned during the turmoil. This learned level of correlation has high precision, so there is little doubt that it is at a higher level because of a numerical discrepancy. The belief that market movements are loyal to each other turns into a self-fulfilling prophecy. Traders follow other markets closely before making trading decisions. So the belief that interdependence between markets are high during the crash turns into reality by correlated actions of traders in different markets avoiding correlation to fall to its previous level after the crash.

Keywords: stock market index correlations, Bayesian learning, synchronization, stochastic systems.

JEL classification codes: C11, C16, D83, D84, D85, G15.
1 Introduction

Lee and Kim (1993) find that average weekly cross market correlations between 12 major stock markets increase from 0.23 before the October 1987 crash in the U.S. markets to 0.39 afterwards. Even one and a half year after the crash, for the period of June 1989 to December 1990 average weekly cross market correlations were 0.41. Forbes and Rigobon (2002) report that during times of high volatility in the series under consideration, correlation coefficient between these series is biased upwards. Through a detailed common factors analysis, Lee and Kim (1993) conclude that “the phenomenon of closer co-movements among stock markets after the crash existed regardless of the extent of the world stock markets’ volatility”. Forbes and Rigobon (2002) explain why the correlations get high during the crash, but they do not explain the phenomena observed by Lee and Kim (1993). Why the correlations stay high after the crash, regardless of the volatility? This paper is an attempt to explain the stickiness in correlation coefficients between stock markets by referring to theory of Bayesian learning.

Our main conclusion is that the correlation does not go down because it is learned during the turmoil. This learned level of correlation has high precision, so there is little doubt that it is at a higher level because of a numerical discrepancy. The belief that market movements are loyal to each other turns into a self-fulfilling prophecy. Traders follow other markets closely before making trading decisions. So the belief that interdependence between markets are high during the crash turns into reality by correlated actions of traders in different markets avoiding correlation to fall to its previous level after the crash.

Is there a draw back in high level of synchronization between markets over the world? One concern may be diminishing opportunities of cross-country hedging when markets start to
move up and down altogether. Cross-country hedging may be of secondary importance after
cross-industry hedging when the world economies are synchronizing through trade. Also, the
correlation coefficient measuring comovements at higher frequencies may fail to measure the
secular growth in the indices. So hedging opportunities still may exist for investors targeting
long term growth rather than short term speculations.

2. Bayesian Learning

One renovation brought by this paper is the utilization of the phase, rather than the
magnitude of the stock market indices. This approach is convenient in terms of qualitative
classification of the signal relative to the prior distribution of beliefs. By “phase” or “innova-
tion in phase” it is meant the direction of change in the market index. For example, if
we arbitrarily take zero as the most recent peak point of the market index, π (in radians),
or 180° (in degrees) will be the point when it is making the first dip after that. And then
the phase of the index will keep on increasing (sometimes decreasing) and when it reaches
2π (or 360°) we will call it a complete cycle. In converting the index magnitude to phase,
we are losing the amplitude of the index, and gaining the ability to attach a direction to
the perceived index values and the signals about the indices. To our understanding, gains
dominate losses for this case: in practice, the initial concern on the index is usually not
its absolute value, but its direction of change. Phase definition of the index still has this
capability. What the phase series lacking is the amplitude of the index (which is assumed
to be constant through this paper), which has secondary importance beside the phase.

In the literature, Bayesian learning systems usually posed within the framework of the
normal (Gaussian) distribution, or distributions with a support on the line (or on a plane).
In this paper, a distribution with support on a circle (or on a torus) will be considered. The counterpart of the normal distribution on the circle is the von Mises distribution which has some desirable properties similar to the normal distribution. These properties are given by Mardia and Jupp (2000, pp.41-43). In this study, one additional property of the von Mises distribution is crucial: like the normal distribution, it is in the conjugate family of itself. That is, if the prior and the signal $x_t$ are von Mises, then the posterior is von Mises as well.

2.1 The circular case

The setup of the model is similar to the Gaussian case as expressed in Chamley (2004):

1. The value of the nature’s parameter $\mu$ is chosen randomly before the first period according to a von Mises distribution $M(\delta, \kappa_\mu)$, where $\delta$ is the mean direction and $\kappa_\mu$ is the precision.

2. There is a countable number $n$ of individuals that receive a private signal $x_i$, where the stochastic process $\{x_i\}_{i=1,...,n}$ is defined on a probability space $(\Omega, \mathcal{F}, P)$ where $\mathcal{F}$ is a $\sigma$-algebra on the set $\Omega$, and $P$ is a probability measure on the measurable space $(\Omega, \mathcal{F})$. Specifically, $x_i = \mu + \epsilon_i$. $\text{Corr}(\epsilon_i, \epsilon_j) = 0$ for $i \neq j$, and $\epsilon \sim M(0, \kappa)$ which implies $x_i \sim M(\mu, \kappa)$. All individuals have the same payoff function from their actions $a$: $U(a) = -E[(a - \mu)^2]$. Every individual $i$ is assigned to a time $t$ exogenously. Individual $t$ chooses his action $a_t \in [0, 2\pi)$ once and for all in period $t$. Each individual chooses an action only once, and at each time period, there is only one individual acting.

3. The public information at the beginning of period $t$ is made of the prior distribution $M(\delta, \kappa_\mu)$ and the set of previous actions $H_t = \{a_{t-1}, \ldots, a_1\}$.
2.2 Perfect observation of actions

The common payoff function induces the individuals to choose their beliefs $\tilde{\mu}_t \triangleq E_t[\mu]$ as their action: $a_t = \tilde{\mu}_t$ for all $t$. Since all individuals know the payoff function is common, they are aware that the actions reflect beliefs perfectly. So the public belief on $\mu$ one period later, $\mu_{t+1}$ is equal to the individual $t$’s personal belief: $\mu_{t+1} = \tilde{\mu}_t$. The public belief $\beta_t \sim M(\mu_t, \kappa_t)$ is updated according to the Bayesian rule below (Bagchi, 1994) to become the public belief for the next period: $\beta_{t+1} \sim M(\mu_{t+1}, \kappa_{t+1})$.

$$\mu_{t+1} = \tan^{-1} \frac{\kappa \sin x_t + \kappa_t \sin \mu_t}{\kappa \cos x_t + \kappa_t \cos \mu_t},$$ (2.1)

$$\kappa_{t+1}^2 = \kappa_t^2 + \kappa^2 + 2\kappa \kappa_t \cos(\mu_t - x_t).$$ (2.2)

The updating rule in the von Mises case is particularly interesting: unlike the normal case, the updating rule for precision depends on the signal $x_t$ and the prior mean $\mu_t$. Three particular cases are listed below:

1. **Assertive signal**

   When $x_t = \mu_t$ for a certain time interval $t$, then $\kappa_t$ grows by the precision of the signal:

   $$\kappa_{t+1} = \kappa_t + \kappa.$$ This case is identical with the Gaussian case.

2. **Orthogonal signal**

   When $|\mu_t - x_t| = \pi/2$ at a certain time interval $t$, then $\kappa_t^2$ grows by $\kappa^2$: $\kappa_{t+1}^2 = \kappa_t^2 + \kappa^2 \Rightarrow \kappa_{t+1} < \kappa_t + \kappa$.

3. **Negative signal**

   When $|\mu_t - x_t| = \pi$ at a certain time interval $t$, then $\kappa_t$ may increase or decrease:
\( \kappa_{t+1} = |\kappa_t - \kappa| \).

In general, \( \lim_{t \to \infty} P(A(\kappa)\kappa t < \kappa_{t+1} < \kappa t) = 1. \) If the incoming signal has any vector component in the direction of the common belief, definitely there will be gains in terms of precision.

Before continuing with the divergence property in the circular case, the linear case will be summarized for comparison purposes.

The updating of mean and variance of the public belief is referred to as “Gaussian learning rule” in the linear case, and is a consequence of the normal distribution being the *conjugate family* of itself. \(^2\) The Gaussian updating rule is particularly interesting since the updating of the precision \( \rho \) is independent of the mean \( \mu \).

In the linear case, the recursive equation for precision of belief \( \beta_t \) could be identified as a function of time:

\[
\rho_{t+1} = \rho_0 + t \rho_\epsilon. \quad (2.3)
\]

where \( \rho_\epsilon \) is the precision of the signal \( s_t \). So precision grows indefinitely in time, and public belief becomes “almost certain” as \( t \to \infty \). The variance \( \sigma_t^2 = 1/\rho_t \) converges to zero like \( 1/t \).

One difficulty with the linear case is that every signal contributes to the precision of the public belief. A reduction in the precision of beliefs is impossible for this case by signals contradicting the learned value of a parameter. By adopting the circular support for the distribution of the signal, we have the flexibility of qualitatively classifying the signals as

---

\(^1\)see proof of Lemma ??.

\(^2\)That is, if the signal \( s_t \) is normal, then the set of possible common distributions for the prior \( \beta_t \) and the posterior \( \beta_{t+1} \) form the conjugate family of the normal distribution.
assertive, orthogonal, and negative.

Specifically, the following proposition applies for the circular case:

**Proposition 2.1** \( \kappa_{t+1} > \kappa_t \) when \( \cos(\mu_t - x_t) > -\frac{\kappa}{2\kappa_t} \). If \( \kappa_t \) is growing, history is gaining more importance over the private signal through formula (??).

Next, we will propose and prove a lemma showing that \( \kappa_{t+1}/\kappa \) converges to \( t\rho \) in distribution.

**Lemma 2.1** \( \lim_{t \to \infty} P(\kappa_{t+1}/\kappa > t\rho) = 1 \).

**Proof.**

From (??) and (??) we have the recursive relation

\[
\begin{align*}
  z_{t+1} &= z_t + u_t \Rightarrow z_{t+1} = z_o + \sum_{j=1}^{t} u_j. \\
  (2.4)
\end{align*}
\]

for \( z_t \) and \( u_t \) complex where \( z_t \equiv \kappa_t \cos \mu_t + i\kappa_t \sin \mu_t, u_t \equiv \kappa \cos x_t + i\kappa \sin x_t, \) and \( z_o \equiv \kappa_{\mu} \cos \delta + i\kappa_{\mu} \sin \delta. \)

Then we obtain the real recursive equations

\[
\begin{align*}
  \kappa_{t+1} \cos \mu_{t+1} &= \kappa_{\mu} \cos \delta + \kappa \sum_{j=1}^{t} \cos x_j \Rightarrow C_t \triangleq \sum_{j=1}^{t} \cos x_j = \frac{\kappa_{t+1}}{\kappa} \cos \mu_{t+1} - \frac{\kappa_{\mu}}{\kappa} \cos \delta, \\
  \kappa_{t+1} \sin \mu_{t+1} &= \kappa_{\mu} \sin \delta + \kappa \sum_{j=1}^{t} \sin x_j \Rightarrow S_t \triangleq \sum_{j=1}^{t} \sin x_j = \frac{\kappa_{t+1}}{\kappa} \sin \mu_{t+1} - \frac{\kappa_{\mu}}{\kappa} \sin \delta.
\end{align*}
\]

\( C_t \) and \( S_t \) are the sufficient statistics that represent the history \( H_t \) similar to the sufficient statistic \( T_t \triangleq \sum_{j=1}^{t} s_j \) in the linear case. Mardia and Jupp (2000) give the distributions for \( R_t \equiv \sqrt{C_t^2 + S_t^2} \), and \( \bar{\theta}_t \equiv \tan^{-1}(S_t/C_t) \). We will establish divergence of \( \kappa_{t+1}/\kappa \) without
using the distribution function explicitly. We use (2.1) and (2.2) and \( R^2_t = C^2_t + S^2_t \) to derive an expression for \( \kappa^2 R^2_t \):

\[
\kappa^2 R^2_t = \kappa^2_{t+1} + \kappa^2_\mu - 2\kappa_{t+1}\kappa_\mu \cos(\mu_{t+1} - \delta). \tag{2.5}
\]

From Mardia and Jupp (2000, p.75)

\[
E[\kappa^2 R^2_t] = \kappa^2 t^2 \rho^2 + \kappa^2 t(1 - \rho^2) = \kappa^2_\mu + E[\kappa^2_{t+1}] - 2\kappa_\mu E[\kappa_{t+1} \cos(\mu_{t+1} - \delta)] \tag{2.6}
\]

where \( \rho \) is the mean resultant length of the distribution under consideration. Specifically for the von Mises case, \( \rho \equiv A(\kappa) \).

Collecting terms and dividing by \( t^2 \), we get

\[
E[\kappa^2_{t+1}/t^2] = \kappa^2 \rho^2 + \frac{\kappa^2_\mu}{t} [1 - \rho^2] - \frac{\kappa^2_\mu}{t^2} + 2\kappa_\mu E[\frac{1}{t^2} \kappa_{t+1} \cos(\mu_{t+1} - \delta)]. \tag{2.7}
\]

All the terms except the first vanishes as \( t \to \infty \). The last term vanishes since from the case of an assertive signal we know \( \sup_\omega \kappa_{t+1}(\omega) = \kappa_\mu + \kappa t, \forall \omega \in \Omega \) and thus \( \lim_{t \to \infty} \kappa_{t+1}/t^2 = 0 \) with certainty. Since the cosine function is bounded, it is also true that

\[
\lim_{t \to \infty} \kappa_{t+1} \cos(\mu_{t+1} - \delta)/t^2 = 0 \text{ with certainty. We have established that}
\]

\[
\lim_{t \to \infty} E[\kappa^2_{t+1}/t^2] = \kappa^2 \rho^2, \text{ or}
\]

\[
\lim_{t \to \infty} E \left[ \frac{\kappa^2_{t+1}}{\kappa^2 t^2} - \rho^2 \right] = 0 \tag{2.8}
\]
Revisiting stock market index correlations

which implies

\[ \frac{\kappa_{t+1}}{\kappa_t^2} \to_{L_1} \rho^2 \Rightarrow \frac{\kappa_{t+1}}{\kappa_t^2} \to_p \rho^2 \Rightarrow \frac{\kappa_{t+1}}{\kappa_t^2} \to_d \rho^2. \]  (2.9)

Since the square root function is measurable and continuous over the domain \([0, \infty)\), we also have \( \frac{\kappa_{t+1}}{\kappa_t} \to_d \rho \). Since \( \frac{\kappa_{t+1}}{\kappa_t} \) converges to a constant in distribution, we may conclude that

\[ \lim_{t \to \infty} P \left( \frac{\kappa_{t+1}}{\kappa_t} > \rho \right) = 1 \Rightarrow \lim_{t \to \infty} P \left( \frac{\kappa_{t+1}}{\kappa} > t\rho \right) = 1. \]  \( \square \)

**Proposition 2.2** In a circular social belief system with updating rules (??) and (??), in determining agents’ actions, the probability that the weight of public belief (relative to private signals) tending to infinity becomes equal to one as \( t \to \infty \).

**Proof.** This is a direct indication of Lemma ??.

**Remark 2.1** Notice that proposition ?? is for a belief distributed with some circular distribution updated by (??) and (??). It is valid for, but not limited to beliefs distributed with a von Mises distribution.

**Remark 2.2** Mardia and Jupp (2000) also show that the conditional limiting distribution of \( \mu_{t+1} \) when the belief has a von Mises distribution is also von Mises: \( \mu_{t+1} | \kappa_{t+1} \sim M(\mu, \kappa_{t+1}) \) so that \( E[\mu_{t+1} | \kappa_{t+1}] = \mu \) as \( t \to \infty \). The expected value of the public belief, given the value of precision of the belief, converges to the true value \( \mu \) as time goes to infinity, and with \( \kappa_{t+1} \) tending to infinity in the sense of Lemma ??.

**2.3 Learning a leading indicator**

In this section, a two dimensional circular learning system will be investigated. For the sake of simplicity, two stock markets will be considered. The intrinsic values of markets

\[ \text{The converge concepts here are} \to_{L_1}: \text{convergence in the first moment,} \to_p: \text{convergence in probability, and} \to_d: \text{convergence in distribution.} \]
are given by the firm profits, which fluctuate with frequencies $\mu_1$ and $\mu_2$ for market 1 and market 2 respectively. These frequencies are not known by traders. Traders in these two markets play a guessing game: their returns from trade are maximized if they can forecast the movement of the market value with minimum squared deviation from the actual frequency $\mu_j$, $j = 1, 2$. In other words, beating the market maximizes profits. As explained in the previous section, the decision rule implies perfect revelation of beliefs, and the forecast of each trader becomes public belief one period later.

The support of the distribution in this case, similar to a circle, will be a manifold, namely a torus. It will be assumed that the expected cosine of the difference between signal $x_{1t}$ and $x_{2t}$ is the cosine of a certain angle $\alpha$. For this case, Mardia and Jupp (2000) propose a distribution on a torus with density proportional to:

$$
\exp\{\kappa_1 \cos(x_{1t} - \mu_1) + \kappa_2 \cos(x_{2t} - \mu_2) + (\cos x_{1t}, \sin x_{1t})D(\cos x_{2t}, \sin x_{2t})' \} \quad (2.10)
$$

where $D = a \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$ which is a multiple of a rotation matrix. Also, we used the notation $(\cdot, \cdot)'$ to denote transpose of a vector.

This setting is similar to the one dimensional setting discussed before. Nature picks two variables $\mu_1$ and $\mu_2$, the constant change (frequency) in firm profits in two stock markets at each time interval, from a distribution proportional to

$$
\exp\{\kappa_{\mu_1} \cos(\mu_1 - \delta_{\mu_1}) + \kappa_{\mu_2} \cos(\mu_2 - \delta_{\mu_2}) + (\cos \mu_1, \sin \mu_1)D(\cos \mu_2, \sin \mu_2)' \} \quad (2.11)
$$

where $D$ is as defined above.
The parameters of the distribution, $\kappa_{\mu_1}, \delta_{\mu_1}, \kappa_{\mu_2}, \delta_{\mu_2}$ and $D$ are known to all investors in both markets. Each time interval has two sub-intervals, $t_-$ and $t_+$. Signal $x_{1t}$ arrives at the beginning of $t_-$ (or $t$) and signal $x_{2t}$ arrives at the beginning of $t_+$. Since the signals are correlated through matrix $D$, agents will use $x_{1t}$ as a “leading indicator” to form their expectations on $x_{2t}$. The Bayesian updating rules for this case will be

\begin{align}
\mu_{jt+1} &= \tan^{-1} \frac{\kappa_{jt} \cos \mu_{jt} + \kappa_j \cos x_{jt}}{\kappa_{jt} \sin \mu_{jt} + \kappa_j \sin x_{jt}}, \quad (2.12) \\
\kappa_{jt+1}^2 &= \kappa_{jt}^2 + \kappa_j^2 + 2\kappa_j \kappa_{jt} \cos(x_{jt} - \mu_{jt})
\end{align}

for $j = 1, 2$. Association between the signals will be learned according to the updating rules below:

\begin{align}
\triangle \psi_{t+1} &= \tan^{-1} \frac{R_t \cos \triangle \psi_t + a \cos(x_{1t} - x_{2t})}{R_t \sin \triangle \psi_t + a \sin(x_{1t} - x_{2t})}, \quad (2.13) \\
R_{t+1}^2 &= R_t^2 + a^2 + 2aR_t \cos(x_{1t} - x_{2t} - \triangle \psi_t)
\end{align}

where the belief on $\cos \alpha$ has mean $\cos \triangle \psi_t$ and precision $R_t$ at the end of period $t$. Notice that by Remark ??, Proposition ?? applies to the belief on $\alpha$. The agents can use the identities $E_{t_-} \cos(x_{1t} - x_{2t}) = \rho_t \cos \triangle \psi_t$ and $E_{t_-} \sin(x_{1t} - x_{2t}) = \rho_t \sin \triangle \psi_t$ to derive

\begin{align}
E_{t_-} \cos(x_{2t} - \mu_{2t}) &= \rho_t \cos(x_{1t} - \mu_{2t} - \triangle \psi_t). \quad (2.14)
\end{align}

The unknown parameter $\rho_t$ could be approximated as $\rho_t \simeq A(R_t + A^{-1}(A(\kappa_{1t})A(\kappa_{2t})))$ under the assumption that $\alpha = \mu_1 - \mu_2$. Note that $\rho_t$ is increasing in $R_t$, $\kappa_{1t}$, and $\kappa_{2t}$. 
The crash scenario is as follows: at time $t$ the index of market 1 makes a big jump down as a result of an outlier local shock higher in magnitude. The investor $t$ in market 2 forecasts their market index taking market 1 index as a leading indicator using (??). Their forecast indicates a fall in their index as there is a learned association between two indices. Investor $t$ in market 2 starts to sell in order not to get hit by the forecasted downfall. However, since it is known to everyone that actions reflect signals perfectly, investors in markets 1 and 2 regard this as a private signal to sell received by investor $t$ independent from the downturn in market 1. The investor $t + 1$ in market 1 take this as a leading indicator, and the downturn in market 2, initiated by market 1 will bounce back, further pulling the index of 1 down.

The accuracy of the association $R_t$ may fall at the beginning, but as both indices start falling fast, $R_t$ increase with repetitive observations arriving confirming similar changes (assertive signals) in the phases of the indices. This is a direct consequence of proposition ??$. With increasing $R_t$ investors start to ignore the new signals about the profitability of the firms with stocks in their own markets, but go with their belief that the indices have a strong association between them. The strong association between the indices turns into a self-fulfilling prophecy.

Even when small local shocks start to take effect again, the learned strong association between the indices prevails. It takes a long time before the investors de-learn the strong association between the indices through negative signals, and for the correlation between the indices to go down to the level before the shock.

3 Dynamical Systems Approach

In this section, an alternative approach will be used to explain the downward stickiness of
the correlation coefficient after crashes. The setup is a continuous dynamical system where $X_t$ and $Y_t$ are the indices of the stock markets 1 and 2 respectively. The system presented below is not based on learning theories, so it is discussed only heuristically as a supporting argument for the learning approach from a different point of view.

### 3.1 Synchronization of indices

It is known that two systems linked with “coupling parameters” $\delta_1$ and $\delta_2$, and “natural frequencies” $\omega_1$ and $\omega_2$ will synchronize, and there will be a steady-state equilibrium where the phase difference between the series is constant over time. An unstable state is also present where the phase difference will be fluctuating, generating a quasiperiodic state in time domain (Stratonovich, 1967, chapter 9, sec.2).

**Definition 3.1** Two coupled systems are defined, in general, as

\[
\begin{align*}
\dot{X}_t & = f_1(X_t) - \delta_1 P_1(X_t, Y_t), \\
\dot{Y}_t & = f_2(Y_t) + \delta_2 P_2(X_t, Y_t), \quad \delta_1 > 0, \delta_2 > 0.
\end{align*}
\]

where $P_1$ and $P_2$ are deterministic functions. When the functions $f_1$ and $f_2$ are assumed to be generating series with constant amplitudes, and with noisy frequencies averaging to $\omega_1$ and $\omega_2$ respectively (i.e., the processes $\{X_t\}$ and $\{Y_t\}$ go through a steady state cycle when independent) the analysis could be carried out in terms of phase series only.

\[
\begin{align*}
\dot{\phi}_{1t} & = \omega_1 - \delta_1 Q_1(\phi_{1t}, \phi_{2t}) + \zeta_{1t}, \\
\dot{\phi}_{2t} & = \omega_2 + \delta_2 Q_2(\phi_{1t}, \phi_{2t}) + \zeta_{2t},
\end{align*}
\]
where $\zeta_1t$ and $\zeta_2t$ are additive noise terms on observed frequencies of the series. The functions $Q_1$ and $Q_2$ are the phase space counterparts of functions $P_1$ and $P_2$ respectively, derived under the assumption of constant amplitudes. For details on phase space representation and approximation of functions, see Aronson, Ermentrout, Kopell (1990) and Pikovsky et. al. (2001, pp.222-229).

Assuming that synchronization occurs in only one frequency of the series, and omitting all the resonance terms, the system can be represented in terms of the phase difference $\psi_t \equiv \phi_1t - \phi_2t$ between the series.

$$\dot{\psi}_t = \nu - (\delta_2 + \delta_1) \sin \psi_t + \zeta_{1t} - \zeta_{2t},$$

(3.3)

with a first degree Fourier expansion of functions $Q_1$ and $Q_2$ that are assumed to be identical. Also $\nu \equiv \omega_1 - \omega_2$.

Stratonovich (1967) gives the solution for the stationary distribution of $\psi$ under bounded noise⁴, and for $\delta_2 + \delta_1 > 0$. When $\nu = 0$, this distribution is von Mises, with $\psi \sim M(0, \kappa)$, $\kappa \equiv (\delta_2 + \delta_1)/\sigma^2$ where $\sigma^2 \equiv Var(\zeta_1 - \zeta_2) \equiv \sigma_1^2 + \sigma_2^2 - 2\sigma_{12}$, $\sigma_j^2 \equiv Var \zeta_j$, $j = 1, 2$, $\sigma_{12} \equiv Cov(\zeta_1, \zeta_2)$. The correlation coefficient $r$ in this case will be an increasing function of $\kappa$ as derived by Koopmans (1974):

$$r \equiv A(\kappa)$$

(3.4)

where $A(\kappa) \equiv I_1(\kappa)/I_0(\kappa)$ and $I_p$ is the modified Bessel function of the first kind of order $p$. Since $A(\kappa)$ is monotonically increasing, this approach implies increasing correlation between the series when covariance between the noise terms is increasing. Also when $\triangle \delta_2 + \triangle \delta_1 > 0$,

⁴Namely, if $|\zeta_1 - \zeta_2| < \pi$. 
Revisiting stock market index correlations

the correlation between the series will be increasing. This feature is similar to features of the
models introduced by Barnett and Dalkir (2007, equation 3.34) and Dalkir (2004, equation
8).

The case where \( \nu \neq 0 \) is slightly more complicated. In this situation the mean of the
stationary distribution is not zero. Under bounded noise the phase difference will be stable
around the steady state solution \( \psi^* \neq 0 \).

If we define \( \gamma_\psi \triangleq \dot{\psi} \), then the time average of \( \gamma_\psi \) will be

\[
\overline{\gamma_\psi} \equiv \pi^{-1} \sigma^2 |I_{\nu}(\kappa)|^{-2} \sinh(\pi \bar{\nu})
\]

(3.5)

where \( \bar{\nu} \equiv \nu/\sigma^2 \) and \( \kappa \) is as defined above (Stratonovich, 1967). Also \( I_{\nu}(\cdot) \) is the modified
Bessel function of the first kind of complex order.

The dynamical system approach is also in parallel with the Bayesian learning approach
of section 2. When the index of market 1 starts falling as a result of a local shock, and
if the investors in market 2 are aware that \( \sigma_{12} \) is nonzero, they would know that the local
shock in market 1 will affect the index of market 2 as well. During times of turmoil, and
with uncertainty in the model (??) investors in market 2 may believe that it is not only
the shock being transmitted by the correlated shocks, but the parameter \( \delta_2 \) is increasing
as well. That belief will further decrease the average change in the phase difference \( \overline{\gamma_\psi} \) as
indicated in Stratonovich (1967, pg.242) (or increase the correlation coefficient \( r \)). Likewise,
if the markets have comparable sizes, investors in market 1 will observe that both indices
are moving together, so they may as well believe that \( \delta_1 \) is increasing. Following their belief
that two markets are now structurally close with a higher \( \delta_1 \) and a higher \( \delta_2 \), investors in

\[5\text{If } |\zeta_1 - \zeta_2| < \pi - 2\psi^* \text{ where } \psi^* = \sin^{-1} \frac{\nu}{\sigma_1 + \sigma_2}, -\frac{\pi}{2} < \psi^* < \frac{\pi}{2} \text{ is the steady state phase difference. Also } \nu \neq 0 \Rightarrow \psi^* \neq 0.\]
both markets will make their trading decisions in a synchronized manner. That will further improve the comovement, and the structural change in model (??) will turn into a self-fulfilling prophecy. In the case where investors believe that the effect of the shock vanished, the persisting higher level of $\delta_1 + \delta_2$ will keep $\gamma_{\psi}$ at its lower level ($r$ at its higher level).

3.2 Reduction in volatility

An immediate question follows after the previous sections is whether the stronger association between stock market indices have any advantages. One advantage that could be cited under the dynamical systems framework is the reduction in the volatility of the systems when there is an association between them.

Malakhov (1968) provides the formal treatment of the situation. If the stock market indices are closely associated so that we can neglect the volatility in their phase difference relative to the volatility in the individual phase series, the condition for reduced volatility compared to the case where the indices have no association between them becomes as the following.

Proposition 3.1 If two stock markets are close to a synchronized state, that is the volatility in their phase difference $\psi$ is negligible relative to the volatility in the individual phase series, $\sigma_1^2$ and $\sigma_2^2$, the variance of both index series will be at the same level $(\delta_2^2 \sigma_1^2 + \delta_1^2 \sigma_2^2)/((\delta_1 + \delta_2)^2)$ and further, if

$$\frac{\delta_2}{2\delta_1 + \delta_2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{\delta_1 + 2\delta_2}{\delta_1}$$

(3.6)

holds then this common variance will be smaller than the variances of the independent indices, $\sigma_1^2$ and $\sigma_2^2$. 
Proof.

The proof closely follows Malakhov (1968) as translated in Pikovsky et al. (2001.) First define $\psi = \phi_1 - \phi_2$, $\theta = \delta_2 \phi_1 + \delta_1 \phi_2$, $\nu = \omega_1 - \omega_2$. Then

\[
\dot{\psi} = \nu - (\delta_1 + \delta_2) \sin \psi + \zeta_1 - \zeta_2, \quad (3.7)
\]

\[
\dot{\theta} = \delta_2 \omega_1 + \delta_1 \omega_2 + \delta_2 \zeta_1 + \delta_1 \zeta_2,
\]

\[
Var(\dot{\theta}) = \delta_2^2 \sigma_1^2 + \delta_1^2 \sigma_2^2.
\]

Write the phases in the form

\[
\phi_1 = \frac{\delta_1 \psi + \theta}{\delta_1 + \delta_2} \quad \text{and} \quad \phi_2 = \frac{-\delta_2 \psi + \theta}{\delta_1 + \delta_2}.
\]

When variance of $\dot{\psi}$ is negligible we find

\[
Var(\dot{\phi}_1) = \frac{\delta_2^2 \sigma_1^2 + \delta_1^2 \sigma_2^2}{(\delta_1 + \delta_2)^2} = Var(\dot{\phi}_2).
\]

\[
(3.8)
\]

For the markets with no shock transmission in between them, variance of $\dot{\phi}_j$ is $\sigma_j^2$, $j = 1, 2$. It is easy to verify that

\[
\frac{\delta_2^2 \sigma_1^2 + \delta_1^2 \sigma_2^2}{(\delta_1 + \delta_2)^2} < \sigma_j^2, \quad j = 1, 2,
\]

\[
(3.9)
\]

when the proposition is satisfied. □

The proposition ?? could specifically be considered for equally strong association between the markets: $\delta_1 = \delta_2$. In this case, the condition in proposition ?? will be reduced to

\[
1/3 < \sigma_2^2/\sigma_1^2 < 3.
\]

Both stock markets will benefit from a reduction in volatility in this case, and their common volatility will be $(\sigma_1^2 + \sigma_2^2)/4$. Additionally, if these two markets have equal volatility when they are not associated ($\sigma_1^2 = \sigma_2^2$), then their volatility will reduce by half once they become synchronized.
4 Conclusion

It is an established fact in physical sciences that systems with quite a weak connection between them display high levels of comovement (for example, in Ermentrout (1981)). Friedman, Johnson, and Landsberg (2003) specifically state that there is an event strong enough to synchronize two economic systems no matter how weak is the coupling in between them so that they feel the shock to one of the economies in a similar way. This paper states that the individuals in the stock market may not be aware of this fact: the high correlation between the indices is high because of a strong event, and the effect is only temporary. However, if the individuals believe that the stronger correlation is there to stay, that belief will turn into a self-fulfilling prophecy, and returning to the level of correlation before the strong event may not be momentarily.

Our main conclusion is that the correlation between stock market indices does not go down after a crash because it is learned during the crash. This learned level of correlation has high precision, so there is little doubt that it is at a higher level because of a numerical discrepancy. The belief that market movements are loyal to each other turns into a self-fulfilling prophecy after the crisis. Traders follow other markets closely before making trading decisions. It is the belief that interdependence between markets are high during the crash that avoids correlation to fall to its previous level after the crash.

Further synchronization of stock markets during crashes has two effects: (i) the association between the markets sticks to a higher level so that crashes are shared afterwards at a larger extent, and (ii) the high level of association brings about lower volatility to the prices. The question of which of these effect are dominating the other, and the desirability
of compromising immunity to outside shocks for the benefits of stability calls for further research in the area.

One policy implication to stop a contagious crash is to send frequent and accurate signals to the stock market investors on the profitability of the firms with stocks traded in a stock market. If the investors’ trading policy depends on the intrinsic value of the stocks that are traded rather than the short run price fluctuations, price signals from other markets will be less effective.
References


