



Munich Personal RePEc Archive

Autocorrelation in an unobservable global trend: Does it help to forecast market returns?

Peresetsky, Anatoly and Yakubov, Ruslan

Higher School of Economics, Moscow

2015

Online at <https://mpra.ub.uni-muenchen.de/64579/>

MPRA Paper No. 64579, posted 26 May 2015 13:50 UTC

Autocorrelation in an unobservable global trend: Does it help to forecast market returns?

In this paper a Kalman-filter type model is used to extract a global stochastic trend from discrete non-synchronous data on daily stock market index returns from different markets[‡]. The model allows for the autocorrelation in the global stochastic trend, which means that its increments are predictable. It does not necessarily mean the predictability of market returns, since the global trend is unobservable. The performance of the model for the forecast of market returns is explored for three markets: Japan, UK, US.

Keywords: financial market integration; stock market returns; state space model; Kalman filter; non-synchronous data; market returns forecast.

JEL classification: C49; C58; G10; G15; F36; F65.

1. Introduction

The plot of daily closing values of stock indices of different exchanges reveals a common trend in their behavior, which is referred to as the “global stochastic trend” in Korhonen, Peresetsky (2013b). Stock exchanges are placed in different parts of the globe. Such a plot of daily closing values is misleading, because exchanges close at different times. Korhonen, Peresetsky (2013b) suggested a Kalman-filter type model, which splits the daily returns of each index into two independent components — local and global. Both components are unobservable. Additionally, they suppose a martingale property of the global component. They designed the model, which takes into account the asynchrony of the stock indices, and estimated it with a state-space model. This very simple model produces results in good accordance with the observed properties of stock market returns.

However, estimates of the global trend reveal a correlation between the increments of the global trend in neighboring time intervals. This could be the result of a violation of the model assumption of a martingale property of the global trend, or the result of the estimation method. In our paper we suggest a modified model which allows for the autocorrelation in the global trend increments.

Historically, the existence of a “common stochastic trend” in synchronous data was discussed in the framework of error-correction VAR models. The description of this approach, based on Engle, Granger, (1987), Johansen (1991) can be found in Kasa (1992).

This approach for finding a common stochastic trend is widely used in the literature. To avoid the problem of asynchronous trading returns, long time intervals are considered, so that this problem can be neglected. To mention some of them: Chung and Lin (1994), Jeon and Chiang (1991) use weekly data; Siclos and Ng (2001), Phengpis, Apilado (2004) use monthly data; Rangvid (2001) use quarterly data. Some papers, for example Choudhry, Peng (2007), Lucey, Muckley (2011), Bentes (2015) use daily market data. They

* International Laboratory of Quantitative Finance of the Higher School of Economics, Moscow. The research was funded by the Russian Federation government grant 14.A12.31.0007.

† International Laboratory of Quantitative Finance of the Higher School of Economics, Moscow. The research was funded by the Russian Federation government grant 14.A12.31.0007.

‡ The description of this model was included in the previous paper of authors (Durdyev, Peresetsky, 2014), which is in Russian. To make this paper understandable for English language reader, we include here the translation of the theoretical description of the model in the section 2. With an exception: in previous paper different set of exchanges was discussed.

consider markets from different geographical regions. In this case it is not clear to which point in world time the common trend is attributed. Flad, Jung (2008) use high-frequency intra-day one-minute returns data. To avoid the problem of non-synchronous data they use data only for the overlapping hours of the US and German stock market trading sessions.

The results of the search for a common stochastic trend could depend on the frequency of data. Loh (2013) stressed that “The stock market is a complex system consisting of investors with different investment horizons, and they collectively form aggregate market co-movements. Time scale (investment horizon) is an important aspect that affects trading behaviour”.

Error correction VAR models could be formulated in the state-space approach (see Aoki, 1987, 1988). This approach is more flexible. Kasa (1992) points out that in this approach a common trend can contain transitory elements, rather than being simple random walks.

There are some papers which apply Kalman filter type state-space models to extract a common stochastic trend from synchronous data. Dungey, Martin, and Pagan (2000) use weekly data to study the spreads on long-term bonds for Australia, Japan, Germany, Canada and the UK, all relative to the USA, over the period 1991 to 1999. They found that the world factor is the dominant influence for Australian and Canadian spreads, while both the UK and Germany display country-specific effects, albeit not as strong as does Japan.

Chang, Miller and Park (2009) use a Kalman filter model to derive a latent common stochastic trend from daily observations on the 30 price series of the stocks that comprise the Dow Jones Industrial Average (DJIA). They found that the extracted common stochastic trend resembles the DJIA quite closely up to an affine transformation.

Cartea and Karyampas (2011) use a Kalman filter-based methodology to decompose price series into the true efficient price and the microstructure noise. They found that the decomposition allows one to estimate the variance covariance matrix of asset returns in a more efficient way than the usual methods, which is important for portfolio optimization.

Bae and Kim (2011) use a two-step state space model and monthly data to examine the existence of global and regional factors and to analyze the effects of both factors on four Asian country yields curves.

Felices and Wieladek (2012) use monthly observations (1995–2007) on vulnerability indicators of financial crises, real exchange rate appreciation and international reserve growth, for 31 emerging market economies (EMEs) and 10 industrialized countries. They employ a Kalman filter-based model with time variation (Bayesian dynamic common factor model) to estimate the common component and find that a single common factor contributes on average up to 60% of the total variation in the vulnerability indicators.

Note that when monthly or weekly data are used the asynchrony of the data could be neglected.

There are two papers which explicitly use models close to the concept of unobserved global stochastic trend, which we use in our model.

King and Wadhvani (1990) study the transmission of volatility between stock markets. They construct a model in which “contagion” between markets occurs as a result of attempts of rational agents to infer information from price changes in other markets. They consider a model for the two markets which operates in different time zones. This model exploits the concept of a “shadow index”: “When a market is closed there is no explicit index of prices. But we may define the shadow index as the price that would clear the market if trading

were to take place conditional upon the information that is available when it is closed. Although the shadow index is unobservable, the concept plays an important role in our model” (page 12).

Lin, Engle, and Ito (1994) most clearly state the idea of splitting stock returns in global and local parts. They state that they “designed an econometric framework which is designed to separate the global factor that affects stock returns globally from a local factor that affects stock returns locally”. They improved the approach of King, Wadhvani (1990) decomposing close-to-close returns into daytime and overnight returns and by allowing time-varying volatility in performing the signal extraction. In empirical application the study relation between Tokyo and New York markets was considered using daily data on open and close prices for the period 1985–1989.

Following Korhonen, Peresetsky (2013b), we exploit the linear Kalman filter model to extract common stochastic trend from non-synchronous data on daily returns of the stock indices. In addition we allow for the correlation of the increments of the global stochastic trend in neighboring time intervals. The model with correlation in common stochastic trend was included in the previous paper of authors (Durdyev, Peresetsky, 2014), which is in Russian. To make this paper understandable for English language reader, we include here the translation of the theoretical description of the model in the section 2. In the empirical part of this paper we consider three market indices (Nikkei 225, FTSE, S&P 500). It is demonstrated that there is a statistically significant negative autocorrelation in the global stochastic trend which increases during the crises. However that autocorrelation does not help to predict daily stock market returns.

2. The model with autocorrelation

The model is an extension of the model in Korhonen, Peresetsky (2013b). We consider three stock index daily returns: Japan (Tokyo, Nikkei 225), UK (London, FTSE), and US (New-York, S&P 500). Closing times of the stock exchanges are respectively 6:00, 16:30, and 21:00 UTC. Figure 1 illustrates the closing times for the two days: $t-1$ and t .

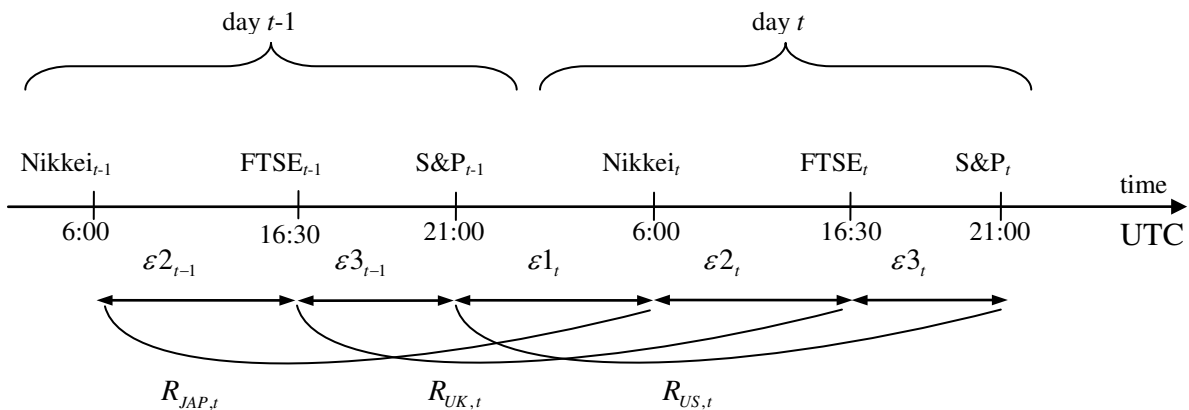


Figure 1. Daily returns and state variables.

The daily index return is defined as the difference of the logs of closing values of the indices, e.g. $R_{US,t} = \ln(S\&P_t) - \ln(S\&P_{t-1})$. Obviously daily returns of two stock indices are correlated since the corresponding 24-hour time intervals are common.

Following Korhonen, Peresetsky (2013a, 2013b), the daily return of market index I is split into two components — global and local:

$$R_{I,t} = R_{IG,t} + R_{IL,t}. \quad (1)$$

The global component $R_{IG,t}$ absorbs all news from the world economy and world financial markets that is available at time t and important for all markets. Local component $R_{IL,t}$ absorbs local news that is important only to investors in the local market. Such splitting corresponds to the market globalization: more and more traders/financial firms operate, not in one but in many markets, and there is continuous 24 hours trading for many financial instruments.

Note that both components are unobservable, only their sum $R_{I,t}$ is observable. The global component (global stochastic trend) exists only in the brains of traders.

Denote as $\varepsilon 1_t, \varepsilon 2_t, \varepsilon 3_t$ the log-returns of the global stochastic trend from 21:00 (day $t-1$) to 6:00 (day t), from 6:00 (day t) to 16:30 (day t), and from 16:30 (day t) to 21:00 (day t). For example, the global part of the US index return is proportional to $\varepsilon 1_t + \varepsilon 2_t + \varepsilon 3_t$ (see Figure 1).

The linear Kalman filter model is given in (2):

$$\begin{aligned} y_t &= Ax_t + Hs_t + u_t & - \text{observation equation,} & \quad t = 1, \dots, T, \\ s_t &= Fs_{t-1} + v_t & - \text{state equation,} \end{aligned} \quad (2)$$

where $y_t - (n \times 1)$, observables; $s_t - (m \times 1)$, state variables; $A - (n \times k)$, $H - (n \times m)$, $F - (m \times m)$, matrices; $u_t \sim (0, \Sigma_u) - (n \times 1)$ random vector; $v_t \sim (0, \Sigma_v) - (m \times 1)$ random vector.

In our case the observables are $y_t = [R_{JAP,t}, R_{UK,t}, R_{US,t}]'$, and the epsilons are state variables. Since the indices have different scales, their returns are affine transformations of the common stochastic trend; hence the observation equations are (3):

$$\begin{aligned} R_{JAP,t} &= \alpha_1 + \beta_1(\varepsilon 2_{t-1} + \varepsilon 3_{t-1} + \varepsilon 1_t) + u1_t, \\ R_{UK,t} &= \alpha_2 + \beta_2(\varepsilon 3_{t-1} + \varepsilon 1_t + \varepsilon 2_t) + u2_t, \\ R_{US,t} &= \alpha_3 + \beta_3(\varepsilon 1_t + \varepsilon 2_t + \varepsilon 3_t) + u3_t, \end{aligned} \quad (3)$$

Since standard linear Kalman filter model (2) does not allow for lagged state variables in the observation equations, it is necessary to increase the dimension of the state vector by introducing two additional components, $\varepsilon 2L_t, \varepsilon 3L_t$.

Thus, the observation equations (3) in new state variables are (4)

$$\begin{aligned} R_{JAP,t} &= \alpha_1 + \beta_1(\varepsilon 2L_t + \varepsilon 3L_t + \varepsilon 1_t) + u1_t, \\ R_{UK,t} &= \alpha_2 + \beta_2(\varepsilon 3L_t + \varepsilon 1_t + \varepsilon 2_t) + u2_t, \\ R_{US,t} &= \alpha_3 + \beta_3(\varepsilon 1_t + \varepsilon 2_t + \varepsilon 3_t) + u3_t. \end{aligned} \quad (4)$$

If we introduce autocorrelation in the global trend in the state equation we get (5):

$$\begin{aligned} \varepsilon 1_t &= e1_t + \rho_1 \varepsilon 3_{t-1}, \\ \varepsilon 2_t &= e2_t + \rho_2 \varepsilon 1_t, \\ \varepsilon 3_t &= e3_t + \rho_3 \varepsilon 2_t, \\ \varepsilon 2L_t &= \varepsilon 2_{t-1}, \end{aligned} \quad (5)$$

$$\varepsilon 3L_t = \varepsilon 3_{t-1},$$

where $e_t = (e1_t, e2_t, e3_t)$ and $u_t = (u1_t, u2_t, u3_t)$ are random vectors both with zero mean. It is supposed that the components of the vector e_t are random shocks i.e. are uncorrelated with corresponding state variables: $\text{Cov}(e1_t, \varepsilon 3_{t-1}) = \text{Cov}(e2_t, \varepsilon 1_t) = \text{Cov}(e3_t, \varepsilon 2_t) = 0$. Note that (5) imply a predictability of the returns of the global trend, e.g. $E(\varepsilon 1_t | \varepsilon 3_{t-1}) = \rho_1 \varepsilon 3_{t-1}$.

Equations (5) can be presented in the standard form of the state equations (2) by using the substitutions (6):

$$\begin{aligned} \varepsilon 1_t &= e1_t + \rho_1 \varepsilon 3_{t-1}, \\ \varepsilon 2_t &= e2_t + \rho_2 e1_t + \rho_2 \rho_1 \varepsilon 3_{t-1}, \\ \varepsilon 3_t &= e3_t + \rho_3 e2_t + \rho_3 \rho_2 e1_t + \rho_3 \rho_2 \rho_1 \varepsilon 3_{t-1}, \\ \varepsilon 2L_t &= \varepsilon 2_{t-1}, \\ \varepsilon 3L_t &= \varepsilon 3_{t-1}. \end{aligned} \quad (6)$$

Thus model (4), (6) in vector form (2) looks like (4'), (6') :

$$\begin{pmatrix} R_{JAP,t} \\ R_{UK,t} \\ R_{US,t} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \cdot 1 + \begin{pmatrix} \beta_1 & 0 & 0 & \beta_1 & \beta_1 \\ \beta_2 & \beta_2 & 0 & 0 & \beta_2 \\ \beta_3 & \beta_3 & \beta_3 & 0 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon 1_t \\ \varepsilon 2_t \\ \varepsilon 3_t \\ \varepsilon 2L_t \\ \varepsilon 3L_t \end{pmatrix} + \begin{pmatrix} u1_t \\ u2_t \\ u3_t \end{pmatrix}, \quad (4')$$

$$\begin{pmatrix} \varepsilon 1_t \\ \varepsilon 2_t \\ \varepsilon 3_t \\ \varepsilon 2L_t \\ \varepsilon 3L_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & \rho_1 & 0 & 0 \\ 0 & 0 & \rho_2 \rho_1 & 0 & 0 \\ 0 & 0 & \rho_3 \rho_2 \rho_1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon 1_{t-1} \\ \varepsilon 2_{t-1} \\ \varepsilon 3_{t-1} \\ \varepsilon 2L_{t-1} \\ \varepsilon 3L_{t-1} \end{pmatrix} + \begin{pmatrix} e1_t \\ e2_t + \rho_2 e1_t \\ e3_t + \rho_3 e2_t + \rho_3 \rho_2 e1_t \\ 0 \\ 0 \end{pmatrix}. \quad (6')$$

With covariance matrix Σ_v of the vector of errors in the state equation:

$$\text{Var}(v_t) = \Sigma_v = \begin{pmatrix} \sigma_1^2 & \rho_2 \sigma_1^2 & \rho_3 \rho_2 \sigma_1^2 & 0 & 0 \\ \rho_2 \sigma_1^2 & \sigma_2^2 + \rho_2^2 \sigma_1^2 & \rho_3 \sigma_2^2 + \rho_3 \rho_2^2 \sigma_1^2 & 0 & 0 \\ \rho_3 \rho_2 \sigma_1^2 & \rho_3 \sigma_2^2 + \rho_3 \rho_2^2 \sigma_1^2 & \sigma_3^2 + \rho_3^2 \sigma_2^2 + \rho_3^2 \rho_2^2 \sigma_1^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

where $\sigma_1^2 = \text{Var}(e1_t)$, $\sigma_2^2 = \text{Var}(e2_t)$, $\sigma_3^2 = \text{Var}(e3_t)$.

We assume the simplest diagonal specification for the covariance matrix of the errors in the observation equation (4): $\Sigma_u = \text{diag}(\sigma_{u1}^2, \sigma_{u2}^2, \sigma_{u3}^2)$.

Model (4'), (6') is not identifiable. One can multiply σ_1 , σ_2 , σ_3 and ε by some scale factor λ and multiply all β by λ^{-1} and this procedure does not change the model. The global trend can be identified only up to an affine transformation. To make the model identifiable we impose the restriction on the parameters: $\sigma_1^2 + \sigma_2^2 + \sigma_3^2 = 3$.

Model (4'), (6') can be estimated with a standard Kalman filter procedure realized in EViews using actual observations of the market index returns. Thus we obtain parameter estimates $\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2, \hat{\alpha}_3, \hat{\beta}_3, \hat{\sigma}_{u1}^2, \hat{\sigma}_{u2}^2, \hat{\sigma}_{u3}^2, \hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\sigma}_3^2, \hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3$ and estimates for the state variables $\hat{\varepsilon}1_t, \hat{\varepsilon}2_t, \hat{\varepsilon}3_t$ using all available information ($\hat{\varepsilon}1_t = E(\varepsilon 1_t | I_T)$, $\hat{\varepsilon}2_t = E(\varepsilon 2_t | I_T)$, $\hat{\varepsilon}3_t = E(\varepsilon 3_t | I_T)$). By summing the state

variables one can estimate the evolution of the global trend at the closing times of the three stock exchanges. It is also possible to decompose stock market returns into global and local components (7):

$$\begin{aligned}
R_{JAP,t} &= \hat{\alpha}_1 + \hat{\beta}_1(\hat{\varepsilon}2_{t-1} + \hat{\varepsilon}3_{t-1} + \hat{\varepsilon}1_t) + \hat{u}1_t = \hat{\alpha}_1 + \widehat{R_{JAP,G,t}} + \widehat{R_{JAP,L,t}}, \\
R_{UK,t} &= \hat{\alpha}_2 + \hat{\beta}_2(\hat{\varepsilon}3_{t-1} + \hat{\varepsilon}1_t + \hat{\varepsilon}2_t) + \hat{u}2_t = \hat{\alpha}_2 + \widehat{R_{UK,G,t}} + \widehat{R_{RUS,L,t}}, \\
R_{US,t} &= \hat{\alpha}_3 + \hat{\beta}_3(\hat{\varepsilon}1_t + \hat{\varepsilon}2_t + \hat{\varepsilon}3_t) + \hat{u}3_t = \hat{\alpha}_3 + \widehat{R_{US,G,t}} + \widehat{R_{US,L,t}}.
\end{aligned} \tag{7}$$

It is assumed that all α_i in (7) are zeros, but we include them into (7) by analogy between (7) and CAPM.

3. The estimation of autocorrelation in global stochastic trend

3.1. Data

Data for the closing quotes for the three stock indices Nikkei 225, FTSE, S&P 500 for the period 22.09.1997–06.03.2014 were obtained from Bloomberg. All weekends and holidays were deleted from the sample. Descriptive statistics are presented in the table 1.

Table 1. Descriptive statistics

	R_{UK}	R_{JAP}	R_{US}
Mean	-7.14E-05	-4.83E-05	0.00017
Median	7.97E-05	0.000000	0.00033
Maximum	0.094	0.132	0.102
Minimum	-0.082	-0.121	-0.095
Standard deviation	0.0125	0.0158	0.0130
Skewness	-0.165	-0.291	-0.274
Kurtosis	7.45	9.038	10.47
Jarque-Bera test	3376.1	6231.2	9503.6
p-value	0.00000	0.00000	0.00000
Serial correlation LM test*	3.70	11.4	36.7
p-value	0.055	0.001	0.000
ARCH-LM test*	144.5	308.5	215.2
p-value	0.000	0.000	0.000
Observations	4069	4069	4069

* Both Breush-Godfrey serial correlation LM test and ARCH-LM tests were evaluated for lag = 1.

The results of the estimation of the model (4), (6) are presented in Table 2. As expected, alphas are not significantly different from 0. All three coefficients ρ_i , responsible for the autocorrelation in the global trend are statistically different from 0 and negative. The negative value of ρ_3 means that if the return of the global trend in the time interval 6:00–16:30 UTC (closing time Tokyo exchange – closing time London exchange) is positive then expected value of return of the global trend in the next interval 16:30–21:00 UTC (closing time London exchange – closing time New-York exchange) is negative. A similar interpretation of the global trend “overshooting” is valid for other intervals.

Estimates of the slope coefficients β_1 (*JAP*), β_2 (*UK*), β_3 (*US*) are close to each other, β_2 and β_3 are almost identical. Same behavior of betas in rolling widows is presented in Figure 5 below. This could be explained by the fact that US, UK and Japan markets are largest markets in their time zones.

Table 2. Model with autocorrelation

Parameter	Estimate	Standard error
α_1 (<i>JAP</i>)	-4.85E-05	(0.000242)
β_1 (<i>JAP</i>)	0.00745***	(0.00012)
α_2 (<i>UK</i>)	7.09E-5	(0.00019)
β_2 (<i>UK</i>)	0.00783***	(0.00009)
α_3 (<i>US</i>)	0.00017	(0.00019)
β_3 (<i>US</i>)	0.00783***	(0.00010)
$\ln \sigma_{u1}^2$	-9.053***	(0.028)
$\ln \sigma_{u2}^2$	-10.703***	(0.095)
$\ln \sigma_{u3}^2$	-11.114***	(0.132)
$\ln \sigma_1^2$	-0.2413	(0.0350)
$\ln \sigma_2^2$	0.2115***	(0.0191)
ρ_1	-0.353***	(0.018)
ρ_2	-0.298***	(0.033)
ρ_3	-0.070***	(0.018)

Note. *** — significant at 1% level.

3.2. Correlation in the global stochastic trend

Given coefficients ρ_1 , ρ_2 , ρ_3 it is possible to calculate the correlations of the returns of the global trend in adjacent intervals. From (5) we have (8):

$$\begin{aligned}\varepsilon 1_t &= e1_t + \rho_1 \varepsilon 3_{t-1}, \\ \varepsilon 2_t &= e2_t + \rho_2 e1_t + \rho_2 \rho_1 \varepsilon 3_{t-1}, \\ \varepsilon 3_t &= e3_t + \rho_3 e2_t + \rho_3 \rho_2 e1_t + \rho_3 \rho_2 \rho_1 \varepsilon 3_{t-1}.\end{aligned}\tag{8}$$

Under condition $|\rho_3 \rho_2 \rho_1| < 1$, series $\varepsilon 3_t$ is a weak stationary first order autocorrelation process[§]. Thus $E(\varepsilon 3_t) = 0$ and

$$\text{Var}(\varepsilon 3_t) = \frac{\sigma_3^2 + \rho_3^2 \sigma_2^2 + \rho_3^2 \rho_2^2 \sigma_1^2}{1 - (\rho_3 \rho_2 \rho_1)^2}.\tag{9}$$

From (8) we have

$$\text{Corr}(\varepsilon 1_t, \varepsilon 3_{t-1}) = \rho_1 \sqrt{\frac{\text{Var}(\varepsilon 3_{t-1})}{\text{Var}(\varepsilon 1_t)}} = \rho_1 \sqrt{\frac{\text{Var}(\varepsilon 3_{t-1})}{\sigma_1^2 + \rho_1^2 \text{Var}(\varepsilon 3_{t-1})}}.\tag{10}$$

$$\text{Corr}(\varepsilon 2_t, \varepsilon 1_t) = \rho_2 \frac{\sigma_1^2 + \rho_1^2 \text{Var}(\varepsilon 3_{t-1})}{\sqrt{(\sigma_2^2 + \rho_2^2 \sigma_1^2 + \rho_2^2 \rho_1^2 \text{Var}(\varepsilon 3_{t-1}))(\sigma_1^2 + \rho_1^2 \text{Var}(\varepsilon 3_{t-1}))}},\tag{11}$$

$$\text{Corr}(\varepsilon 3_t, \varepsilon 2_t) = \rho_3 \frac{\sigma_2^2 + \rho_2^2 \sigma_1^2 + \rho_2^2 \rho_1^2 \text{Var}(\varepsilon 3_{t-1})}{\sqrt{\text{Var}(\varepsilon 3_{t-1})(\sigma_2^2 + \rho_2^2 \sigma_1^2 + \rho_2^2 \rho_1^2 \text{Var}(\varepsilon 3_{t-1}))}}.\tag{12}$$

Using (9)–(11), estimates of corresponding parameters (Table 2) and the delta-method we obtain estimates of the correlations and their standard errors (see Table 3).

[§] The autocorrelation of this process is $\rho_1 \rho_2 \rho_3$, its estimate is -0.0073 with s.e. 0.0017.

Table 3. The correlation of global trend returns in adjacent intervals and sample correlations of epsilon estimates.

Coefficients	Estimates	Sample correlation
$\text{Corr}(\varepsilon_{1_t}, \varepsilon_{3_{t-1}})$	-0.368 (0.017)	-0.382
$\text{Corr}(\varepsilon_{2_t}, \varepsilon_{1_t})$	-0.248 (0.028)	0.130
$\text{Corr}(\varepsilon_{3_t}, \varepsilon_{2_t})$	-0.081 (0.020)	0.066

Note. Standard errors are in parentheses.

All three correlation coefficients are statistically significant and their signs coincide with signs of ρ_i . Sample correlation between corresponding estimates $\widehat{\varepsilon}_3$, $\widehat{\varepsilon}_3$, $\widehat{\varepsilon}_3$ differ from the estimates, calculated by (10)–(12) since they include more noise. Both estimates show the largest negative correlation between global trend returns at the interval (closing London, closing NY) and the adjacent interval (closing NY, closing Tokyo).

3.3. Comparison with S&P 500 e-mini

The concept of the (unobserved) global stochastic trend suggests that traders have an idea of its value at each moment, 24 hours a day. There are assets which are traded continuously. The most well-known asset of this type is the contract S&P 500 e-mini, which is traded 24 hours with a break of 45 minutes. It was introduced by the Chicago Exchange (CME) on September 9, 1997, after the value of the existing S&P contract became too large for many small traders. The E-Mini quickly became the most popular equity index futures contract in the world. The current average daily implied volume for the E-mini is over \$100 billion, far exceeding the combined traded dollar volume of the underlying 500 stocks. Thus, it is interesting to compare the estimated global stochastic trend with the S&P 500 e-mini.

For comparison we take the period June 2009 – March 2014, three points each day (at 6:00, 16:30, and 21:00 UTC).

The plots of the dynamics of the S&P 500 e-mini** and estimated global trend are presented at the Figure 2. Since the global trend is determined up to an affine transformation, the scale of the left axis was chosen in order to make two graphs comparable.

** Data on S&P e-mini are taken from www.finam.ru

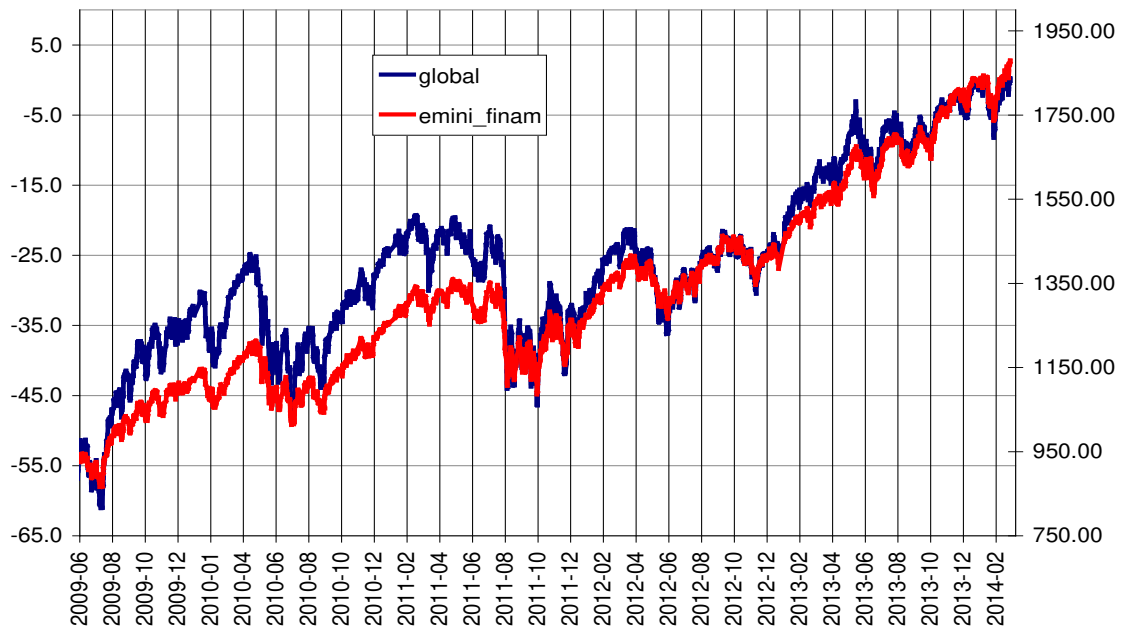


Figure 2. Global stochastic trend (“global”, left axis) and S&P 500 e-mini (“emini_finam”, right axis)

Note that S&P 500 e-mini mainly reflects information important for the US stock market and global stochastic trend accumulates information important for all markets. US is the largest economy, hence the US market is the most important component of the global trend (Korhonen, Peresetsky, 2013b). This is why the plots are so close to each other. On the other hand, other markets (in our example – Japan, UK) also impact the global trend, so the plots are different. This is in agreement with the conclusion in Lin et al. (1994) that not only does the US market affect Japan, but the Japan market also affects the US.

For a more detailed visual comparison of the global trend and S&P 500 e-mini Figure 3 presents scatter plots of 24-hour returns of the global trend (r_{global}) and contract S&P 500 e-mini (r_{emini}) for closing of exchanges in Tokyo (JAP), London (UK), and New York (US). Correlation between r_{global} and r_{emini} is 0.837 for Tokyo, 0.883 for London, and 0.968 for New York. Most similarity is observed for US, as expected.

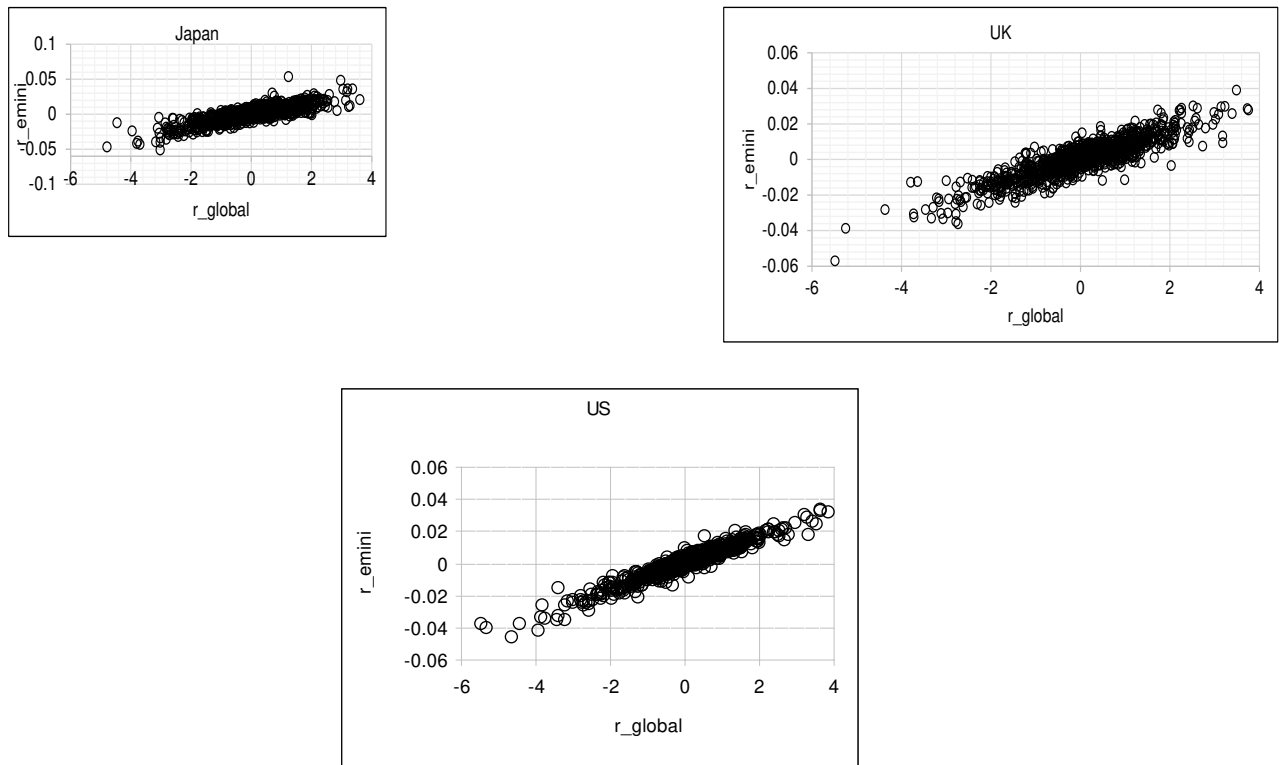


Figure 3. Scatter plots of 24-hour returns of the global trend (r_{global}) and contract S&P 500 e-mini (r_{emini}) for the closing of exchanges in Tokyo (JAP), London (UK), and New York (US)

3.4. The dynamics of correlations

14 parameters of the model (4), (6) were estimated on the whole sample (see Table 2, 4096 obs.). It was assumed that the parameters are constant in time. To study the evolution of coefficients ρ_1, ρ_2, ρ_3 in time we estimate them in a rolling window of length τ . Is it true that during the crisis periods the correlations increase?

In order to estimate a Kalman filter model with a large number of parameters we need many observations for the convergence of the likelihood function maximization algorithm. In our case a stable convergence of the algorithm is achieved only for windows of around 500 observations.

The evolution of the coefficients ρ_1, ρ_2, ρ_3 in rolling windows of length $\tau = 500$ is presented in Figure 4. Points correspond to the right end of the rolling window. Markers highlight estimates which are significant at 10% level. Almost all significant values are negative. Contrary to ρ_1, ρ_2 which are significant in most of windows, ρ_3 after 2008 is significant only between 11.2010 and 08.2011. The evolution of respective correlations (10)–(12) is presented in Figure 6 in Appendix.

The evolution of betas is presented in Figure 5. Plot of the log-likelihood $\log L$ — a measure of goodness-of-fit — is also presented. During the crisis of 2008 one can observe a drop of log-likelihood. Thus, the model fits data less than in quiet periods of the markets. However, during the crisis all betas increase which means there is a synchronization of the markets.

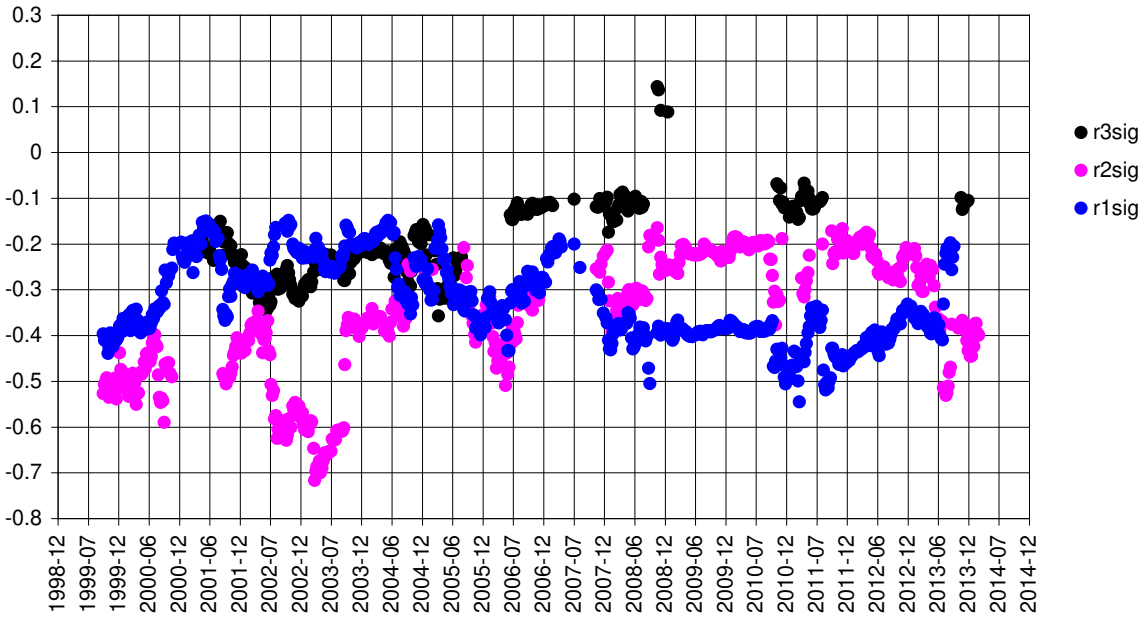


Figure 4. Evolution of ρ_1, ρ_2, ρ_3 (r1sig, r2sig, r3sig in the plot)

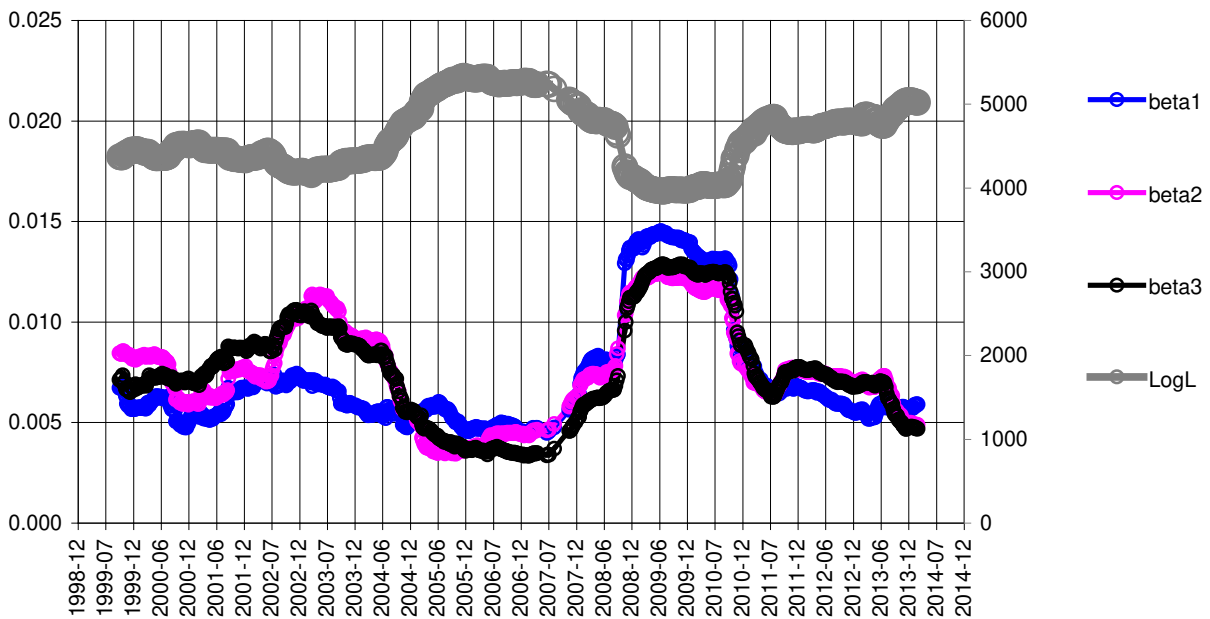


Figure 5. Evolution of the beta coefficients $\beta_1, \beta_2, \beta_3$ in rolling windows (beta1, beta2, beta3 at the plot). LogL — log-likelihood.

4. Does the model of the global trend help to forecast market returns?

The existence of a non-zero autocorrelation in the global trend allows us to make a non-trivial one step ahead forecast of its increment. Since the increment of the global trend is a part of the daily market return it also provides the possibility to forecast market returns. The success of such a forecast depends upon the balance between the global and local parts of market return and on the precision of the global market return forecast.

Given the estimates of the coefficients of the model (4), (6), in the interval $[1, T]$ we can estimate the state variables $\widehat{\varepsilon}1_t, \widehat{\varepsilon}2_t, \widehat{\varepsilon}3_t, t=1, \dots, T$ and using (8) to forecast the next step increment of the global trend $\widehat{\varepsilon}1_{T+1} = \hat{\rho}_1 \widehat{\varepsilon}3_T$. Then by (3) we forecast $R_{JAP, T+1}$ (assuming $E(u_{T+1} | I_T) = 0$) as

$$\hat{R}_{JAP, T+1} = E(R_{JAP, T+1} | I_T) = \hat{\alpha}_1 + \hat{\beta}_1(\widehat{\varepsilon}2_T + \widehat{\varepsilon}3_T + \widehat{\varepsilon}1_{T+1}) = \hat{\alpha}_1 + \hat{\beta}_1(\widehat{\varepsilon}2_T + \widehat{\varepsilon}3_T + \hat{\rho}_1 \widehat{\varepsilon}3_T). \quad (13)$$

Similarly by (6) we can forecast the next two intervals

$$\widehat{\varepsilon}2_{T+1} = \hat{\rho}_1 \hat{\rho}_2 \widehat{\varepsilon}3_T \quad \text{and} \quad \widehat{\varepsilon}3_{T+1} = \hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_3 \widehat{\varepsilon}3_T,$$

and then using (3) forecast two other market returns as

$$\hat{R}_{UK, T+1} = E(R_{UK, T+1} | I_T) = \hat{\alpha}_2 + \hat{\beta}_2(\widehat{\varepsilon}3_T + \widehat{\varepsilon}1_{T+1} + \widehat{\varepsilon}2_{T+1}) = \hat{\alpha}_2 + \hat{\beta}_2(\widehat{\varepsilon}3_T + \hat{\rho}_1 \widehat{\varepsilon}3_T + \hat{\rho}_1 \hat{\rho}_2 \widehat{\varepsilon}3_T), \quad (14)$$

$$\hat{R}_{US, T+1} = E(R_{US, T+1} | I_T) = \hat{\alpha}_3 + \hat{\beta}_3(\widehat{\varepsilon}1_{T+1} + \widehat{\varepsilon}2_{T+1} + \widehat{\varepsilon}3_{T+1}) = \hat{\alpha}_3 + \hat{\beta}_3(\hat{\rho}_1 \widehat{\varepsilon}3_T + \hat{\rho}_1 \hat{\rho}_2 \widehat{\varepsilon}3_T + \hat{\rho}_1 \hat{\rho}_2 \hat{\rho}_3 \widehat{\varepsilon}3_T). \quad (15)$$

However, the distance between information set $I_T = \{y_t = (R_{JAP, t}, R_{UK, t}, R_{US, t}); t=1, \dots, T\}$ and events forecasted in (14) and (15) is larger than for the event forecasted in (13). To forecast $R_{UK, T+1}$ it is possible to use information from $R_{JAP, T+1}$ and for $R_{US, T+1}$ information from $R_{JAP, T+1}$ and $R_{UK, T+1}$.

We construct that ‘‘improved’’ forecast with the following trick. We redefine ‘‘24 hours day’’ and use $y_t = (R_{UK, t}, R_{US, t}, R_{JAP, t+1})$ as the vector of observables^{††}. Using equations (4), (6) with that vector of observables we get forecast $\hat{R}_{UK, T+1}$ as $E(R_{UK, T+1} | I_T, R_{JAP, T+1})$ with a formula corresponding to (13). Similarly for the ‘‘improved’’ forecast of $R_{US, T+1}$ we use model (4), (6) with vector of observables $y_t = (R_{US, t}, R_{JAP, t+1}, R_{UK, t+1})$ ^{‡‡} and get the forecast by (13) as $E(R_{US, T+1} | I_T, R_{JAP, T+1}, R_{UK, T+1})$ (see Figure 1).

In order to estimate the quality of these forecasts we use the following procedure. Model (4), (6) was estimated in rolling windows of size 500 observations with step 21. For each window a one day ahead forecast was made. Thus $N = 165$ forecasts were obtained for the days from 30.09.1997 to 06.03.2014. Two measures of the precision of the forecast were used mean absolute error (MAE) and root mean squared error (RMSE):

$$MAE = \frac{1}{N} \sum_{t=1}^N |R_t^F - R_t|, \quad RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (R_t^F - R_t)^2},$$

here R_t^F is the forecast for day t , and R_t is the observed return. Column ‘‘model’’ in Table 4 presents the precision measures for the ‘‘improved’’ forecasts for the three market returns Nikkei 225, FTSE and S&P 500.

The next four columns in the table 4 present ratios of the measures of the model precision to the similar measures for the four benchmark models. Column ‘‘martingale’’ corresponds to the forecast $R_t^F \equiv 0$, corresponding to the perfect market hypothesis; column ‘‘trend’’ corresponds to the naïve forecast $R_t^F = R_{t-1}$. The next two columns ‘‘regression (16)’’ and ‘‘regression (17)’’ correspond to the forecast made by two regression models which regress market return on the preceding market returns. For example, for the Nikkei 225 index return the regressions are:

$$R_{JAP, t} = c_1 + c_2 R_{US, t-1} + c_3 R_{UK, t-1} + w_t, \quad t = \tau - 500, \tau - 499, \dots, \tau, \quad (16)$$

^{††} It corresponds to the choice of the beginning of the day not 0:00 UTC but for example 10:00 UTC.

^{‡‡} It corresponds to the choice of the beginning of the day for example 17:00 UTC.

$$\text{forecast: } \hat{R}_{JAP,\tau+1} = \hat{c}_1 + \hat{c}_2 R_{US,\tau} + \hat{c}_3 R_{UK,\tau}.$$

and

$$R_{JAP,t} = c_1 + c_2 R_{US,t-1} + c_3 R_{UK,t-1} + c_4 R_{JAP,t-1} + w_t, \quad t = \tau - 500, \tau - 499, \dots, \tau, \quad (17)$$

$$\text{forecast: } \hat{R}_{JAP,\tau+1} = \hat{c}_1 + \hat{c}_2 R_{US,\tau} + \hat{c}_3 R_{UK,\tau} + c_4 R_{JAP,\tau}.$$

Table 4. Precision of the forecast by 500 observations

		Error	Relative Error: Model/(Benchmark Model)			
			model	martingale	trend	regression (16)
Japan	MAE	0.0099	0.8810	0.5810	0.9792	0.9842
	RMSE	0.0140	0.8393	0.5774	0.9578	0.9619
UK	MAE	0.0078	0.8742	0.6058	0.9821	1.0140
	RMSE	0.0105	0.8349	0.5899	0.9561	1.0071
US	MAE	0.0065	0.8331	0.6101	0.9513	1.0277
	RMSE	0.0086	0.7843	0.5923	0.9362	1.0015

RMSE of the forecasts (column “model” in table 4) are approximately equal to the standard errors of corresponding market returns (Table 1). The model forecast is more precise than the zero (“martingale”) forecast and much better than naïve (trend) forecast. However, the model forecast has approximately same precision as these two simple benchmark regression models forecasts.

Diebold, Mariano (1995) suggested the test to compare the relative predictive accuracy of the two forecasts. Results of the tests for the same 165 forecasts are presented in the table 5 in Appendix for two loss functions: MAE and MSE. Our model is not significantly different from the “long” regression (17) forecasts, but at 10% significant level it beats “short” regression (16) for the MSE criteria and Japan and US forecasts.

Results for 3569 forecasts (rolling windows of size 500 with step 1) are presented in the table 6 in Appendix^{§§}. Now our model beats benchmark models for the forecasts of the UK and US indices for MSE loss function. Most likely this result is statistically significant only because of large number of observations although this does not imply that we can construct an arbitrage strategy out of these forecasts.

6. Conclusion

The model based on a linear Kalman filter which extracts a common stochastic trend from nonsynchronous observations was suggested in Korhonen, Peresetsky (2013b). We extend this model allowing for the autocorrelation in the global stochastic trend, and estimate this model using daily observations for the period 22.09.1997–06.03.2014 for three markets indices Nikkei 225, FTSE, S&P 500.

We found a statistically significant autocorrelation in the global trend. However, this observation does not help to make market return forecasts statistically reliable for practical applications. Still it might be that

^{§§} Table 4 for 3569 forecasts does not differ significantly from the table for 165 forecasts (with step 21 rolling windows).

estimating a model for more markets uniformly distributed on the UTC time scale will show a better performance.

Comparing the extracted global stochastic trend with the S&P 500 e-mini, which is traded continuously 23 hours per day, we found that their 24-hour returns have the largest correlations for the closing time of the NY exchange. This is not unexpected since the US market has the largest impact on the global trend and that S&P 500 e-mini is mostly based on the US company stocks.

References

- Aoki M. (1988). On alternative state space representations of time series models. *Journal of Economic Dynamics and Control*, 12, 595–607.
- Aoki, M. (1987). *State space modeling of time series*, Springer Verlag, Heidelberg and New York.
- Bae B. Y., Kim D. H. (2011). Global and regional yield curve dynamics and interactions: The case of some Asian countries. *International Economic Journal*, 25 (4), 717–738.
- Bentes S. R. (2015). On the integration of financial markets: How strong is the evidence from five international stock markets? *Physica A: Statistical Mechanics and its Applications*, 429, 205–214.
- Cartea A., Karyampas D. (2011). Volatility and covariation of financial assets: A high-frequency analysis. *Journal of Banking and Finance*, 35 (12), 3319–3334.
- Chang Y., Miller J. I., Park J. Y. (2009). Extracting a common stochastic trend: Theory with some applications. *Journal of Econometrics*, 150 (2), 231–247.
- Choudhry T., Lu L., Peng K. (2007). Common stochastic trends among far east stock prices: effects of the Asian financial crisis. *International Review of Financial Analysis*, 16 (3), 242–261.
- Chung P. J., Liu D. J. (1994). Common stochastic trends in pacific rim stock markets. *The Quarterly Review of Economics and Finance*, 34 (3), 241–259.
- Diebold X. D., Mariano R. M. (1995). Comparing predictive ability. *Journal of Business and Economic Statistics*, 13 (3), 253–263.
- Dungey M., Martin V. L., Pagan A. R. (2000). A multivariate latent factor decomposition of international bond yield spreads. *Journal of Applied Econometrics*, 15 (6), 697–715.
- Durdyev R. I., Peresetsky A. A. (2014). Autocorrelation in the global stochastic trend. *Applied Econometrics*, 35 (3), 81–100. (in Russian).
- Engle R., Granger C. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55 (2), 251–276.
- Felices G., Wieladek T. (2012). Are emerging market indicators of vulnerability to financial crises decoupling from global factors? *Journal of Banking and Finance*, 36 (2), 321–331.
- Flad M., Jung R. C. (2008). A common factor analysis for the US and the German stock markets during overlapping trading hours. *Journal of International Financial Markets, Institutions and Money*, 18 (5), 498–512.
- Hamilton J. (1994). *Time series analysis*. Princeton University Press.
- Jeon B. N., Chiang T. C. (1991). A system of stock prices in world stock exchanges: Common stochastic trends for 1975-1990? *Journal of Economics and Business*, 43 (4), 329–338.
- Johansen S. (1991). Estimation and hypothesis testing of co-integration vectors in Gaussian vector autoregressive models. *Econometrica*, 59 (6), 1551–1580.
- Kasa K. (1992). Common stochastic trends in international stock markets. *Journal of Monetary Economics*, 29 (1), 95-124.
- King M., Wadhvani S. (1990). Transmission of volatility between stock markets. *Review of Financial Studies*, 3, 5–33.

Korhonen I., Peresetsky A. (2013a). What determines stock market behavior in Russia and other emerging countries? *Bank of Finland, BOFIT Discussion Papers*, No 4/2013.

Korhonen I., Peresetsky A. (2013b). Extracting global stochastic trend from non-synchronous data. *Bank of Finland, BOFIT Discussion Papers*, No 15/2013.

Lin W., Engle R., Ito T. (1994). Do bulls and bears move across borders? Transmission of international stock returns and volatility, *Review of Financial Studies*, 7, 507–538.

Loh L. (2013). Co-movement of Asia-Pacific with European and US stock market returns: A cross-time-frequency analysis. *Research in International Business and Finance*, 29(C), 1–13.

Lucey B. M., Muckley C. (2011). Robust global stock market interdependencies. *International Review of Financial Analysis*, 20 (4), 215–224.

Phengpis C., Apilado V. P. (2004). Economic interdependence and common stochastic trends: A comparative analysis between EMU and non-EMU stock markets. *International Review of Financial Analysis*, 13 (3), 245–263.

Rangvid J. (2001). Increasing convergence among European stock markets? A recursive common stochastic trends analysis. *Economics Letters*, 71 (3), 383–389.

Siclos P. L., Ng P. (2001). Integration among Asia-Pacific and international stock markets: Common stochastic trend and regime shifts. *Pacific Economic Review*, 6 (1), 89–100.

Appendix

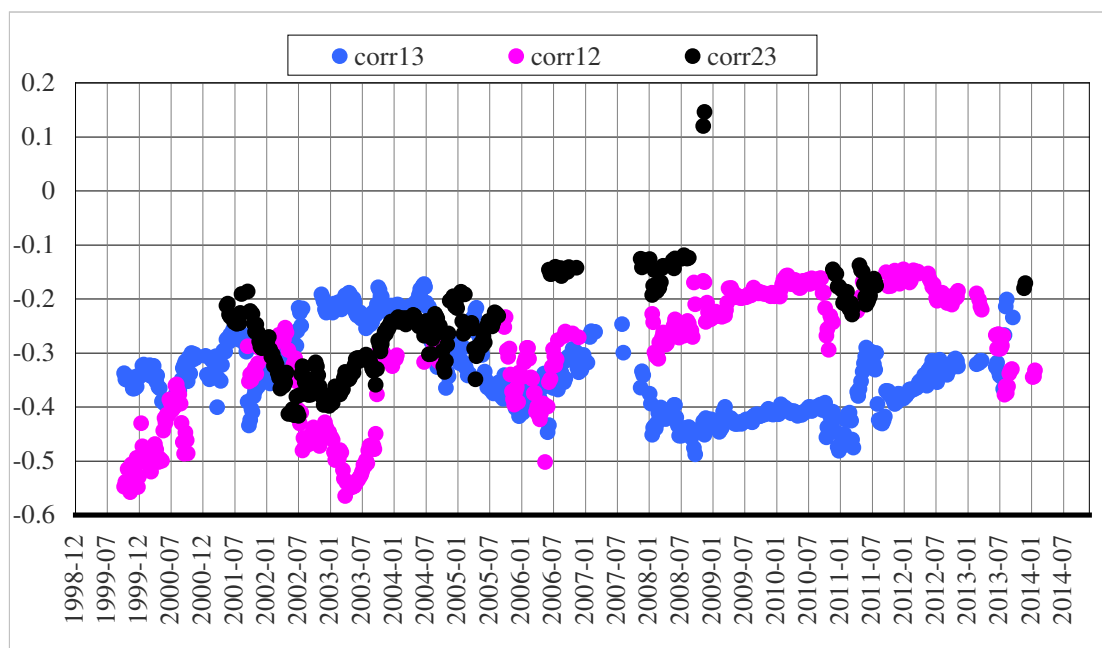


Figure 6. Evolution of the correlations of the returns of the global trend in adjacent intervals, calculated by (10)–(12). $\text{corr13} = \text{Corr}(\varepsilon_{1,t}, \varepsilon_{3,t-1})$, $\text{corr12} = \text{Corr}(\varepsilon_{2,t}, \varepsilon_{1,t})$, $\text{corr23} = \text{Corr}(\varepsilon_{3,t}, \varepsilon_{2,t})$.

Table 5. Diebold-Mariano test for comparing predictive ability of the model (4)–(6) vs. benchmark models, 165 forecasts*

		martingale	trend	regression (16)	regression (17)
Japan	MAE	–2.42 [0.015]	–6.44 [0.000]	–0.96 [0.337]	–0.92 [0.358]
	MSE	–2.18 [0.030]	–4.1 [0.000]	–1.84 [0.066]	–1.37 [0.170]
UK	MAE	–2.38 [0.017]	–5.89 [0.000]	–0.54 [0.589]	0.87 [0.386]
	MSE	–2.20 [0.028]	–3.34 [0.001]	–1.39 [0.165]	0.50 [0.618]
US	MAE	–2.82 [0.005]	–5.8 [0.000]	–1.32 [0.188]	0.97 [0.334]
	MSE	–3.07 [0.002]	–4.21 [0.000]	–1.78 [0.076]	0.064 [0.949]

* P-values in brackets, null hypothesis is H_0 : equal predictive accuracy. Negative value of the test statistic corresponds to better performance of the model in comparison with a benchmark model.

Table 6. Diebold-Mariano test for comparing predictive ability of the model (4)–(6) vs. benchmark models, 3569 forecasts*

		martingale	trend	regression (16)	regression (17)
Japan	MAE	–6.17 [0.000]	–13.66 [0.000]	–2.94 [0.003]	–1.30 [0.192]
	MSE	–3.14 [0.002]	–4.76 [0.000]	–1.47 [0.143]	–1.04 [0.299]
UK	MAE	–4.64 [0.000]	–12.08 [0.000]	–4.50 [0.000]	–1.58 [0.114]
	MSE	–3.31 [0.001]	–5.81 [0.000]	–3.81 [0.000]	–2.66 [0.001]
US	MAE	–7.52 [0.000]	–10.95 [0.000]	–3.82 [0.000]	–3.54 [0.000]
	MSE	–4.08 [0.000]	–4.18 [0.000]	–2.51 [0.012]	–3.31 [0.001]

* P-values in brackets, null hypothesis is H_0 : equal predictive accuracy. Negative value of the test statistic corresponds to better performance of the model in comparison with a benchmark model.