Information Shares in Stationary Time Series and Global Volatility Discovery

Rainer Baule[1], Bart Frijns[2], and Milena E. Tieves[3]

August 30, 2016

[1]University of Hagen, Universitätsstraße 41, 58084 Hagen, Germany, rainer.baule@fernuni-hagen.de

[2]Auckland University of Technology, Private Bag 92006, 1020 Auckland, New Zealand, bfrijns@aut.ac.nz

[3]University of Hagen, Universitätsstraße 41, 58084 Hagen, Germany, milena.tieves@fernuni-hagen.de.

Abstract

Standard price discovery measures, particularly information shares, rely on the concept of co-integration for non-stationary time series. For the definition of information shares, the existence of a permanent impact of innovations is crucial, as these shares measure the relative contribution of different markets to this permanent impact. For stationary time series such as interest rates, CDS prices, or volatilities, permanent impacts do not occur and thus the concept fails. In this paper, we extend the concept of information shares to the case of stationary time series. We suggest a price discovery metric based on the well-known variance decomposition for stationary time series, aiming to be equivalent to Hasbrouck's (1995) information share. This new price discovery measure converges to the standard information share in the limiting case of non-stationarity. As an application, we use the new measure to gain insight into the behavior of major implied volatility indices in Japan, Germany, and the US. We find that the US market has lost its dominant role for global volatility discovery with the emergence of the European debt crisis. In recent periods, the German volatility index exhibits the largest information shares. The Japanese market had a higher importance before the global financial crisis, which has almost vanished within recent periods.

JEL Classification: G14, C32, C10, C51

Keywords: information share, stationarity, co-integration, variance decomposition, volatility spillover.

1 Introduction

A major strand of price discovery research aims at identifying the extent to which different markets contribute to the process of determining the price of a traded asset. For this purpose, several measures have been proposed in the literature. A seminal contribution by Hasbrouck (1995) introduced the information share concept, which soon became very popular and is used in a number of influential papers such as Baillie et al. (2002), Hasbrouck (2003), Grammig et al. (2005), Tse et al. (2006), and Yan and Zivot (2010), among others. The information share measures the contribution of a market's innovation to the total variance of the efficient underlying price innovations. Some shortcomings of this approach, in particular the non-uniqueness in case of cross-correlated innovation terms, are addressed by a number of extensions of the original concept, for example the information leadership share by Putniņš (2013) and the modified and generalized information shares by Lien and Shrestha (2009) and Lien and Shrestha (2014). Furthermore, an alternative approach based on the common factor components introduced by Gonzalo and Granger (1995) was proposed by Booth et al. (1999) and Harris et al. (2002).

All of these approaches assume that the price processes on different markets are co-integrated of order 1. In particular, the single price processes are supposed to be non-stationary. This is a necessary assumption, since the measures focus on the long-term—that is, permanent—impact of price innovations on the price processes of different markets. Such a permanent impact can only occur if the processes are non-stationary.

Assuming non-stationary time series is very reasonable for asset prices such as stocks, commodities, exchange rates, etc. There are, however, financial time series which are less likely to be non-stationary. For example, interest rates or volatilities exhibit meanreverting properties, indicating stationarity.[1] Thus the concepts of Hasbrouck and others cannot be applied to this kind of financial time series without further consideration. Nonetheless, despite these theoretical shortcomings a number of papers analyze processes for interest rates, T-bill rates, CDS prices, or bond prices by means of information shares and similar measures (for example, Hendry and Juselius (2001), Zapata and Fortenbery (1996), Blanco et al. (2005), Mizrach and Neely (2008), and Fricke and Menkhoff (2011)). The reason these methods do not fail is that while time series of interest rates are in fact stationary, they are highly persistent. As a consequence, for shorter time horizons, such as a day or a month, they behave very much like non-stationary time series. Moreover, standard tests for non-stationarity very often cannot reject the null hypothesis of a unit root if time frame examined is sufficiently short.

While maybe sufficient for pragmatic reasons, such a measurement is not totally satisfying. In particular, the results lack an economic interpretation: Hasbrouck's information share measures the contributions to the variance of the permanent component of a price shock—but what are these contributions when we know that there is no permanent component? In this paper, we demonstrate how the information share concept can be extended to stationary time series. We propose an analogous measure based on the variance decomposition of the underlying time series. This new metric is closely related to the spillover indices introduced by Diebold and Yilmaz (2009) and Diebold and Yilmaz (2012). Our

[1] There are alternative approaches based on fractional co-integration, allowing for mean reversion and non-stationarity at the same time, see for example Cheung and Lai (1993), Baillie and Bollerslev (1994), Christensen and Nielsen (2006), and Dias et al. (2016). Nonetheless, volatility modelling is very often based on stationary autoregressive processes with mean reversion, as for example in the popular time-continuous model of Heston (1993).

information share for stationary time series can be seen as the directional spillover of a market, relative to the total spillover in the system. A new interpretation of Hasbrouck's original information share reveals that it is actually the limiting case of this new measure; that is, the latter converges to the former when the underlying time series tend towards non-stationary. As our metric can be used for both stationary and non-stationary time series and is identical to Hasbrouck's information share in the non-stationary case, it is a way of generalizing the classic measure.

We apply the generalized stationary information share to volatility time series, considering the CBOE volatility index VIX for the US-American S&P 500, the VDAX-New volatility index for the German DAX, and the volatility index NikkeiVI for the Japanese Nikkei on a daily basis for about 20 years. The Augmented Dickey-Fuller test rejects a unit root for these time series, thus the time series are stationary and the classical Hasbrouck information share cannot be applied. The results of the empirical analysis reveal a change in the role of the three markets for global volatility discovery in recent years: While up to the financial crisis the VIX was dominating the other markets, with the European debt crisis the German VDAX-New became increasingly important and now exhibits a higher information share than the VIX. The Nikkei volatility index has experienced a strong drop in its informational importance: Before the global financial crisis it was leading the VDAX-New, whereas in recent time its information share with respect to the VDAX-New has fallen to only 8%.

The remainder of the paper is organized as follows. In Section 2 we outline the econometric framework including the concept of information shares. In Section 3 we introduce the new measure for stationary time series and discuss its analogy to the original concept. By providing an alternative derivation of Hasbrouck's information share based on the concept of variance decomposition, we show that our measure can be seen as a generalized information share. Section 4 presents some numerical examples regarding the convergence behavior of the stationary information share to Hasbrouck's original information share when the time series move from the stationary to the non-stationary case. In Section 5, we present empirical results regarding volatility discovery for the VIX, the VDAX-New, and the NikkeiVI. Section 6 concludes.

2 Classic Information Shares

2.1 Market Microstructure Framework

In this section, we briefly sketch the standard case of non-stationary time series. The starting point is a vector autoregression that describes the behavior of the underlying variables $y_t = (y_{1,t}, \ldots, y_{n,t})'$:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} \ldots A_k y_{t-k} + \epsilon_t, \tag{1}$$

where $A_i \in \mathbb{R}^{n \times n}$, $i = 1, \ldots, k$, are fixed coefficient matrices up to the maximum lag length k, and ϵ_t denotes an $(n \times 1)$ vector representing a white noise innovation process. These innovation terms have zero mean, $E(\epsilon_t) = 0$, and a non-singular covariance matrix Ω.

The classic application deals with stock prices on different markets, but one can also think of other settings, such as common and preferred stocks, derivatives, etc. For the sake of simplicity, we will concentrate on the bivariate case in the following, referring to an asset on two markets, bearing in mind that other applications such as interest rates or volatilities exist. The multivariate case is outlined in Appendix C.

Hence we have a bivariate vector of asset prices $y_t = (y_{1,t}, y_{2,t})'$ and a two-dimensional covariance matrix $\Omega = E(\epsilon_t \cdot \epsilon_t^{'}) = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. Under the condition that the asset prices $y_{1,t}$ and $y_{2,t}$ are co-integrated of order 1, their first differences, $\Delta y_t = y_t - y_{t-1}$, can be expressed by a vector error correction model with co-integrating vector β. In the usual application with asset prices on different markets, the co-integrating vector can be forced to be $(1, -1)'$:

$$\begin{pmatrix} \Delta y_{1,t} \\ \Delta y_{2,t} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \cdot \begin{pmatrix} \beta_1, \beta_2 \end{pmatrix} \cdot \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \sum_{l=1}^{k-1} \Gamma_l \begin{pmatrix} \Delta y_{1,t-l} \\ \Delta y_{2,t-l} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix}$$

$$= \alpha \cdot (y_{1,t-1} - y_{2,t-1}) + \sum_{l=1}^{k-1} \Gamma_l \Delta y_{t-l} + \epsilon_t. \quad (2)$$

Here, α is defined as the (2×1) error correction vector containing the speed of adjustment coefficients of both markets to the error correction term $y_{1,t-1} - y_{2,t-1}$. $\Gamma_l := -\sum_{i=l+1}^{k} A_i$, $l = 1, \ldots, k-1$, are (2×2) matrices of autoregressive coefficients up to the maximum lag length $k - 1$.

The first part of the right hand side of Equation (2), $\alpha \cdot (y_{1,t-1} - y_{2,t-1})$, expresses the long-run equilibrium dynamics between both time series. Hasbrouck's information share is based on this long-run dynamics. The second part, $\sum_{l=1}^{k-1} \Gamma_l \Delta y_{t-l}$, depicts the short-run dynamics, or the transitory deviations of the system, induced by market imperfections or other short-term influences.

According to Beveridge and Nelson (1981) and Stock and Watson (1988), not only the first differences, but also the integrated levels y_t of a co-integrated I(1) time series can be decomposed into a permanent component and a transitory component. The permanent or common component corresponds to a random walk with zero drift and the transitory or cyclical component is a zero-mean covariance-stationary process. Based on the vector moving average representation of Δy_t,

$$\Delta y_t = \Psi(L)\epsilon_t, \tag{3}$$

where $\Psi(L)$ is a matrix polynomial in the lag operator, y_t can be decomposed into:

$$y_t = y_0 + \Psi(1)\sum_{s=1}^{t}\epsilon_s + \Psi^*(L)\epsilon_t \tag{4}$$

(Johansen (1991)). $\Psi(1)$ is the sum of the moving average matrices, which can be expressed as $\Psi(1) = \beta_\perp \Pi \alpha'_\perp$ with a scalar $\Pi = (\alpha'_\perp(I - \sum_{l=1}^{k-1}\Gamma_l)\beta_\perp)^{-1} \in \mathbb{R}$ and Γ_l defined as above, and $\Psi^*(L)\epsilon_t$ is a zero-mean covariance stationary process with matrices $\Psi^*_s = -\sum_{i=s+1}^{\infty}\Psi_i$. If $\beta = (1,-1)'$, we have $\beta_\perp = (1,1)'$, thus $\Psi(1)$ consists of two equal rows. The long-run impact of a single innovation ϵ_t is therefore identical for both time series, yielding the common trend representation (Stock and Watson (1988)):

$$y_t = y_0 + (1,1)' \cdot \psi \cdot \sum_{s=1}^{t}\epsilon_s + \Psi^*(L)\epsilon_t, \tag{5}$$

where $\psi := (\psi_1,\psi_2)$ is a (1×2) vector orthogonal to α' representing the common row vector for the sum of the moving average matrices $\Psi(1)$. The permanent or common component, $(1,1)'\cdot\psi\cdot\sum_{s=1}^{t}\epsilon_s$, incorporates the long-run impact of disturbances due to new innovations ϵ_t and is equal to all markets. The common trend representation of Stock and Watson (1988) is the key for the construction of Hasbroucks's information share measure presented in the following section.

2.2 The Information Share

Price discovery is driven by the question of which market incorporates new information first and thus takes the leadership. Because of arbitrage considerations, asset prices are linked by a common factor, the "efficient price" (Hasbrouck (1995)), while traded on parallel markets. As the efficient price is unobservable, market microstructure research tries to identify where exactly price discovery takes place by estimating the relative contribution of each market to the variance of this common factor.

Referring to the Stock-Watson representation (5), the term $\psi\epsilon_t$ represents the long-term, permanent impact of an innovation ϵ_t on the common efficient price. Although the long-run impact is equal in both markets, the relative contributions of the single markets to this common efficient price are not. The basic idea of Hasbrouck's information share measure is to break down the variance of this permanent impact of an innovation on the efficient price, $\psi\epsilon_t$, according to the different markets. Thus, the information share, IS, measures the contributions of the markets to the variation of the common factor, accounting for different innovation terms and also for a possible correlation between the innovations in the information set. In contrast to the price discovery measure proposed by Harris et al. (2002), the IS not only incorporates the speed of adjustment vector α, but also the structure of the covariance matrix Ω.

The variance of the long-term impact of an innovation, $\psi\,\epsilon_t$, amounts to

$$Var(\psi\,\epsilon_t) = \psi\Omega\psi' = \psi_1^2\sigma_1^2 + \psi_2^2\sigma_2^2 + 2\rho\psi_1\psi_2\sigma_1\sigma_2. \tag{6}$$

In the simplest case there is no cross correlation between the two markets, that is, $\rho = 0$. In this situation, the relative contributions of the two markets to the total variance and thus the information shares for both markets are given by

$$IS_1 = \frac{([\psi\Omega]_1)^2}{\psi\Omega\psi'} = \frac{\psi_1^2\sigma_1^2}{\psi_1^2\sigma_1^2 + \psi_2^2\sigma_2^2}, \qquad IS_2 = \frac{([\psi\Omega]_2)^2}{\psi\Omega\psi'} = \frac{\psi_2^2\sigma_2^2}{\psi_1^2\sigma_1^2 + \psi_2^2\sigma_2^2} = 1 - IS_1. \tag{7}$$

In the general case, there is contemporaneous correlation between the innovations in the different markets, resulting in a non-diagonal covariance matrix Ω. That means the innovation terms affect each other and blur the respective influence of the single markets.

Hasbrouck (1995) suggests calculating upper and lower bounds for the information share in this case. By using the Cholesky decomposition, $\Omega = MM'$, with a triangular matrix M, the innovation terms can be orthogonalized, leaving the "pure" variance of one market unaffected by the correlation to the other market. As the Cholesky decomposition always attaches a greater weight to the first market, the resulting measure depends on the order of the markets. Upper and lower bounds (expressed by $\lceil IS \rceil$ and $\lfloor IS \rfloor$, respectively) for the information share are calculated by switching the order of the markets. In the bivariate case, the lower bound $\lfloor IS \rfloor$ is simply the complement to the upper bound of the other market:

$$\lceil IS_1 \rceil = \frac{([\psi M]_1)^2}{\psi \Omega \psi'} = \frac{(\psi_1 \sigma_1 + \psi_2 \sigma_2 \rho)^2}{\psi_1^2 \sigma_1^2 + \psi_2^2 \sigma_2^2 + 2\rho \psi_1 \psi_2 \sigma_1 \sigma_2}, \qquad \lfloor IS_2 \rfloor = 1 - \lceil IS_1 \rceil. \tag{8}$$

Booth et al. (2002) suggest defining the information share of one market as the average value of the upper and lower bound.[2] Usually the information share is calculated based on the VECM representation (2), using the relation that ψ is perpendicular to α.

3 Information Shares in Stationary Time Series

3.1 Failure of the Classic Information Share

When the price time series are stationary, a long-run impact of innovations on a common fundamental price represented by the identical random walk component in the Stock-Watson decomposition (5) no longer exists. Thus the long-run influence of every single

[2] It is worth noting that the higher the cross-correlation between the two markets, the higher the distance between the upper and lower bounds of one market and the less reliable is the average of the two bounds as a price discovery measure. However, our main concern is not with problems of contemporaneous correlation between innovation terms but with stationary time series.

innovation, ϵ_t, turns to zero and, consequently, the main idea of the IS measure is pointless: It is not possible to decompose the variance of the innovation in the efficient price when there is no permanent impact.

A similar problem arises when considering the VECM representation (2). If both time series $y_t = (y_{1,t}, y_{2,t})'$ are stationary, a unique decomposition of the long-run equilibrium represented by $\alpha \cdot \beta'$ no longer exists. For co-integrated I(1) processes, there is a co-integrating vector β so that $\beta' y_t$ is stationary, which is unique up to a scalar. In contrast, when the original time series are already stationary, the combination $\beta' y_t$ remains stationary for any arbitrary choice of β. With an ambiguous choice of the co-integrating vector β, there is also no unique error correction vector α. Hence, also the vector $\psi = \alpha_\perp^{'}$ that defines the information share is arbitrary.[3]

3.2 The Information Share for Stationary Time Series

Because an explicit long-run impact does not exist in the case of stationary time series, any measure of information share necessarily focuses on transitory influences. The impact of each innovation term, ϵ_t, on the system of asset prices tends towards zero when the time horizon tends towards infinity. As the measure is intended to be closely related to the classic information share which decomposes the variance of the long-term impact of innovations, we consider the variance of an impact of innovations with finite time horizon

[3]The described non-uniqueness that emerges with stationary time series is different from the non-uniqueness in the case of cross-correlation between the innovation terms. While many researchers, including Lien and Shrestha (2009), Kehrle and Peter (2013), and Lien and Shrestha (2014), suggest solutions for the latter problem when innovation terms are correlated, we will focus on the non-uniqueness of the decomposition in the case of stationary time series, and equivalently, the absent long-run impact of innovations ϵ_t.

and take the resulting ratio of variances to the limit. In contrast to the non-stationary case, there is no single common factor which could be used for a variance decomposition. Instead, we need to consider the cross-influences of innovation terms for both markets on each other. This is done by the well-known variance decomposition for stationary time series, which we use as the basis for the stationary information share.

Under the condition of stationarity, we can transform the general VAR model (1) into a vector moving average model:

$$y_t = \Psi(L) \cdot \epsilon_t, \tag{9}$$

with $\Psi(L) := A^{-1}(L)$ representing the inverse of the original polynomial in the lag operator $A(L)$ including all fixed coefficient matrices up to the maximum lag length k. Note that in contrast to $A(L)$, the inverse operator $\Psi(L) = \Psi_0 + \Psi_1 \cdot L + \Psi_2 \cdot L^2 + \cdots$ is not necessarily finite. $\epsilon_t \sim \mathcal{N}(0,\Omega)$ depict the innovation terms with covariance structure as above.

In general, the innovation terms are cross-correlated. To cope with this issue, we use, as in the non-stationary case, the Cholesky decomposition $\Omega = M \cdot M'$:

$$y_t = \Psi(L) \cdot M \cdot M' \cdot \epsilon_t = \Phi(L) \cdot w_t = \sum_{p=0}^{\infty} \Phi_p \cdot w_{t-p}, \tag{10}$$

with $\Phi(L) = \Phi_0 + \Phi_1 \cdot L + \Phi_2 \cdot L^2 + \cdots$ representing the modified polynomial in the lag operator given by the set of $\Phi_p = \Psi_p \cdot M$ for $p = 0,1,\ldots$. Applying the Cholesky matrix to the original innovation vector forces the innovation terms, $w_t = M' \cdot \epsilon_t$, to be orthonormal; that is, all shocks are uncorrelated with unit variance, $w_t \sim \mathcal{N}(0,I)$. As with the classic information share, applying the Cholesky decomposition leads to non-unique results, as they depend on the order of markets. Hence, we get upper and lower bounds of information shares for the stationary time series as in the classic setting.

The concept of variance decomposition is to consider the forecast error of the process and to split up its variance according to the influence of different innovations on the related markets. As all innovation terms have a mean of zero, the τ-step ahead forecast at forecast origin t only depends on past realizations of innovations:

$$E_t[y_{t+\tau}] = \sum_{p=\tau}^{\infty} \Phi_p \cdot w_{t+\tau-p}. \tag{11}$$

The difference between the realized value $y_{t+\tau}$ and the forecast $E_t[y_{t+\tau}]$ is the forecast error:

$$FE(y_{t+\tau}) = y_{t+\tau} - E_t[y_{t+\tau}] = \sum_{p=0}^{\tau-1} \Phi_p \cdot w_{t+\tau-p}. \tag{12}$$

It consists solely of the realizations of the innovations between forecast origin t and forecast horizon $t+\tau$. In a bivariate system we can decompose the (2×1) forecast error vector into two parts, corresponding to the forecast error components of the respective markets:

$$FE_j(y_{t+\tau}) = y_{j,t+\tau} - E_t[y_{j,t+\tau}] = \sum_{p=0}^{\tau-1} \phi_{j1}^{(p)} \cdot w_{1,t+\tau-p} + \sum_{p=0}^{\tau-1} \phi_{j2}^{(p)} \cdot w_{2,t+\tau-p}, \tag{13}$$

where the index $j \in \{1,2\}$ represents the respective market and $\phi_{ij}^{(p)}$, $i,j \in \{1,2\}$, are the coefficients of the matrices Φ_p. The forecast error component of the first market is influenced by innovation terms of the second market and vice versa. Analogous to Hasbrouck's decomposition of the impact on the long-term variance, we consider the variance of the forecast error to be decomposed:

$$\begin{aligned} Var(FE_j(y_{t+\tau})) &= E[(y_{j,t+\tau} - E_t[y_{j,t+\tau}])^2] \\ &= E\left[\left(\sum_{p=0}^{\tau-1} \phi_{j1}^{(p)} \cdot w_{1,t+\tau-p} + \sum_{p=0}^{\tau-1} \phi_{j2}^{(p)} \cdot w_{2,t+\tau-p}\right)^2\right] \\ &= \sum_{p=0}^{\tau-1} \left(\phi_{j1}^{(p)}\right)^2 + \sum_{p=0}^{\tau-1} \left(\phi_{j2}^{(p)}\right)^2. \end{aligned} \tag{14}$$

The variance of the forecast error of one market is caused by two different kinds of innovation terms: disturbances due to a market's own innovations and disturbances due to the other markets shocks.

When the prediction length is extended to $t+\infty$, by comparing Equations (10) and (12), we notice that the forecast error converges to the process itself—obviously, as the forecast tends towards zero for a stationary process. Hence, the variance of the forecast error converges to the variance of the process. For the construction of information shares, we consider the contributions of one market to the variance of the respective other market:

$$\eta_{12} := \frac{\sum_{p=0}^{\infty}(\phi_{12}^{(p)})^2}{\sum_{p=0}^{\infty}(\phi_{11}^{(p)})^2 + \sum_{p=0}^{\infty}(\phi_{12}^{(p)})^2}, \qquad \eta_{21} := \frac{\sum_{p=0}^{\infty}(\phi_{21}^{(p)})^2}{\sum_{p=0}^{\infty}(\phi_{21}^{(p)})^2 + \sum_{p=0}^{\infty}(\phi_{22}^{(p)})^2}. \tag{15}$$

We can interpret η_{12} as the part of variance of the first market due to innovation terms of the second market; analogously, η_{21} represents the portion of variance of the second market due to the first market's innovation terms.

These two values do not yet represent relative impacts from one market to another, but relative variance impacts on the respective other markets. Based on these relative impacts we construct the stationary information share SIS by taking the ratio again:

$$\lceil SIS_1 \rceil := \frac{\eta_{21}}{\eta_{12}+\eta_{21}}, \qquad \lfloor SIS_2 \rfloor := \frac{\eta_{12}}{\eta_{12}+\eta_{21}}. \tag{16}$$

Note that these values represent upper bounds for the first market and lower bounds for the second market, according to the definition of the Cholesky matrix M. The respective other bounds are obtained by altering the order of the two markets analogously to the classic case.

The stationary information share SIS_1 measures the influence of the first market on the second market, η_{21}, relative to the aggregated cross-influences of both markets on each other, $\eta_{21}+\eta_{12}$. Vice versa, SIS_2 represents the influence of the second market on the first market, relative to the aggregated cross-influences. Comparing the interpretation of the stationary information share (16) to Hasbrouck's classic information share (8), we see one important difference: In the non-stationary case, there is a common trend component, the common efficient price. For the classic information share IS, we decompose the variance of the common efficient price, attributing the relative shares of the respective markets to their contribution in the price discovery process. In the case of stationary time series, there is no common efficient price. We therefore decompose the variance of the whole process and calculate cross-influences of the respective markets on each other to obtain the stationary information share SIS.

The stationary shares is closely related to indices used in the literature on spillover effects. Diebold and Yilmaz (2009) construct a "spillover index" $S^{(\tau)}$ which measures the total effects of cross variances in a system with a finite τ-step forecast horizon, that is,

$$S^{(\tau)} := \frac{\sum_{p=0}^{\tau-1}(\phi_{12}^{(p)})^2 + \sum_{p=0}^{\tau-1}(\phi_{21}^{(p)})^2}{\sum_{p=0}^{\tau-1}(\phi_{11}^{(p)})^2 + \sum_{p=0}^{\tau-1}(\phi_{12}^{(p)})^2 + \sum_{p=0}^{\tau-1}(\phi_{21}^{(p)})^2 + \sum_{p=0}^{\tau-1}(\phi_{22}^{(p)})^2}. \tag{17}$$

Diebold and Yilmaz (2012) extend this idea in two directions: First, they apply the generalized variance decomposition according to Koop et al. (1996) and Pesaran and Shin (1998), which does not rely on a Cholesky decomposition of the covariance matrix and is therefore independent of the order of the markets. Second, instead of a total index they construct "directional spillover" indices, which attribute the total spillover to pairwise cross-influences. In the case of a diagonal covariance matrix (when the Cholesky decomposition is also diagonal and the generalized variance decomposition equals the standard decomposition), the directional spillover indices $S_i^{(\tau)}$ are equivalent to our cross-contributions of variances when the forecast horizon τ is taken to the limit:

$$S_1^{(\infty)} = \frac{\eta_{21}}{2}, \qquad S_2^{(\infty)} = \frac{\eta_{12}}{2}. \tag{18}$$

Thus, in this special case, the stationary information shares can be expressed in terms of the directional spillover indices:

$$SIS_1 = \frac{S_1^{(\infty)}}{S_1^{(\infty)} + S_2^{(\infty)}}, \qquad SIS_2 = \frac{S_2^{(\infty)}}{S_1^{(\infty)} + S_2^{(\infty)}}. \tag{19}$$

As discussed, building these ratios leads to information shares that sum up to 1. Instead of using the generalized VAR framework, we rely on the Cholesky decomposition and calculate upper and lower bounds by changing the order of markets as with the classical information shares. The generalized variance decomposition would attribute cross-correlation effects to the different markets in a way which is not necessarily justifiable. In the following section we show that our approach is actually consistent with the Hasbrouck (1995) approach.

3.3 Convergence to the Hasbrouck (1995) Information Share

Though somehow similar, the constructions of classical and stationary information shares seem to have substantial differences. In this section, we show that the approach presented for information shares in the case of stationary time series in fact leads to the Hasbrouck information share when applied to non-stationary time series. In this sense, our measure is a generalization of the classic concept. We derive the IS by using the technique of variance decomposition. However, as the process variance is not finite for non-stationary time series, ratios in terms of Equation (15) have to be built for forecast errors with a finite time horizon and then taken to the limit.

So in the following, y_t represents a system on non-stationary, co-integrated time series. For ease of exposition we assume that Ω is diagonal, so there is no contemporaneous correlation between the innovation terms. We present the general case at the end of this section. Instead of decomposing the variance of the impact of a shock on the common efficient price—according to the original definition of the information share—we proceed analogously with the construction of the stationary information share and consider the variance of the forecast error. Recalling the common trend representation (5), we have

$$y_t = y_0 + \Psi(1) \cdot \sum_{s=1}^{t} \epsilon_s + \sum_{j=0}^{\infty} \Psi_j^* \, \epsilon_{t-j} \tag{20}$$

with $\Psi(1) = \sum_{i=0}^{\infty} \Psi_i$ representing the moving average coefficient matrices with equal rows ψ and Ψ_j^* being the transitory matrices resulting in the zero-mean covariance stationary process.

As in the stationary case, we can derive the forecast error at time t for a forecast horizon of τ periods as

$$FE(y_{t+\tau}) = \Psi(1) \cdot \sum_{s=t+1}^{t+\tau} \epsilon_s + \sum_{j=0}^{\tau-1} \Psi_j^* \, \epsilon_{t+\tau-j} = \Psi(1) \cdot \sum_{s=t+1}^{t+\tau} \epsilon_s + \sum_{s=t+1}^{t+\tau} \Psi_{t+\tau-s}^* \, \epsilon_s. \tag{21}$$

Without loss of generality, we set $t = 0$. Computing the variance of the forecast error, we obtain for the two markets $j \in \{1,2\}$:

$$\begin{aligned}
&Var\left(FE_j(y_\tau)\right) \\
&= Var\left([\Psi(1)]_j \cdot \sum_{s=1}^{\tau} \epsilon_s \right) + Var\left(\sum_{s=1}^{\tau} [\Psi_{\tau-s}^*]_j \, \epsilon_s \right) \\
&\qquad + 2\, Cov\left([\Psi(1)]_j \cdot \sum_{s=1}^{\tau} \epsilon_s; \quad \sum_{s=1}^{\tau} [\Psi_{\tau-s}^*]_j \, \epsilon_s \right) \\
&= \tau \cdot [\Psi(1)]_j \, \Omega \, [\Psi(1)]_j' + \sum_{s=1}^{\tau} [\Psi_{\tau-s}^*]_j \, \Omega \, [\Psi_{\tau-s}^{*'}]_j + 2 \sum_{s=1}^{\tau} [\Psi(1)]_j \, \Omega \, [\Psi_{\tau-s}^{*'}]_j \\
&= \tau \cdot (\psi_1^2 \sigma_1^2 + \psi_2^2 \sigma_2^2) + \sum_{s=0}^{\tau-1} \left(\psi_{j1}^{(s)*} \right)^2 \cdot \sigma_1^2 + \sum_{s=0}^{\tau-1} \left(\psi_{j2}^{(s)*} \right)^2 \cdot \sigma_2^2 \\
&\qquad + 2 \cdot \left(\psi_1 \cdot \sum_{s=0}^{\tau-1} \psi_{j1}^{(s)*} \cdot \sigma_1^2 + \psi_2 \cdot \sum_{s=0}^{\tau-1} \psi_{j2}^{(s)*} \cdot \sigma_2^2 \right) \\
&= \left(\tau \cdot \psi_1^2 + \sum_{s=0}^{\tau-1} \left(\psi_{j1}^{(s)*} \right)^2 + 2 \cdot \psi_1 \cdot \sum_{s=0}^{\tau-1} \psi_{j1}^{(s)*} \right) \cdot \sigma_1^2
\end{aligned}$$

$$+\left(\tau \cdot \psi_2^2 + \sum_{s=0}^{\tau-1} \left(\psi_{j2}^{(s)*}\right)^2 + 2 \cdot \psi_2 \cdot \sum_{s=0}^{\tau-1} \psi_{j2}^{(s)*}\right) \cdot \sigma_2^2, \tag{22}$$

with $[\Psi(1)]_j$ and $[\Psi^*_{\tau-s}]_j$ referring to the j-th-row of the matrices $\Psi(1)$ and $\Psi^*_{\tau-s}$. Considering contributions of the two markets to each other's variances, we can identify the impact of the i-th market by the terms collected with σ_i^2. For a finite time horizon τ we get:

$$\eta_{12}^{(\tau)} = \frac{\left(\tau\psi_2^2 + \sum_{s=0}^{\tau-1} \left(\psi_{12}^{(s)*}\right)^2 + 2\psi_2 \sum_{s=0}^{\tau-1} \psi_{12}^{(s)*}\right)\sigma_2^2}{\left(\tau\psi_1^2 + \sum_{s=0}^{\tau-1} \left(\psi_{11}^{(s)*}\right)^2 + 2\psi_1 \sum_{s=0}^{\tau-1} \psi_{11}^{(s)*}\right)\sigma_1^2 + \left(\tau\psi_2^2 + \sum_{s=0}^{\tau-1} \left(\psi_{12}^{(s)*}\right)^2 + 2\psi_2 \sum_{s=0}^{\tau-1} \psi_{12}^{(s)*}\right)\sigma_2^2} \tag{23}$$

and $\eta_{21}^{(\tau)}$ analogously.

When τ tends towards infinity, the variances also tend towards infinity due to the non-stationarity of the process. However, the ratios of variance contributions may remain finite. As the process $\Psi^*(L)\epsilon_t$ is covariance stationary, the sum of squared coefficients $\left(\psi_{ji}^{(s)*}\right)^2$ is bounded by a constant c_{ji}. Furthermore, using the Cauchy-Schwarz inequality, we can bound the mixed terms by[4]

$$\psi_i \cdot \sum_{s=0}^{\tau-1} \psi_{ji}^{(s)*} \leq \sqrt{\tau} \cdot \psi_i \cdot \sqrt{\sum_{s=0}^{\tau-1} \left(\psi_{ji}^{(s)*}\right)^2} \leq \sqrt{\tau} \cdot \psi_i \cdot \sqrt{c_{ji}}. \tag{24}$$

For a forecast horizon τ tending towards infinity, the ratio $\eta_{12}^{(\tau)}$ is dominated by the terms with τ, whereas terms with $\sqrt{\tau}$ and constants can be neglected.Thus we obtain a finite value for the relative contributions of the markets to the variances in the limit:

$$\eta_{12} = \lim_{\tau \to \infty} \eta_{12}^{\tau} = \frac{\tau\psi_2^2\sigma_2^2}{\tau\psi_1^2\sigma_1^2 + \tau\psi_2^2\sigma_2^2} = \frac{\psi_2^2\sigma_2^2}{\psi_1^2\sigma_1^2 + \psi_2^2\sigma_2^2} \tag{25}$$

and analogously

$$\eta_{21} = \frac{\psi_1^2\sigma_1^2}{\psi_1^2\sigma_1^2 + \psi_2^2\sigma_2^2}. \tag{26}$$

[4] In the Cauchy-Schwarz inequality $\sum_{s=0}^{\tau-1}(a_s \cdot b_s) \leq \sqrt{\sum_{s=0}^{\tau-1} a_s^2} \cdot \sqrt{\sum_{s=0}^{\tau-1} b_s^2}$ set $a_s = 1$ and $b_s = \psi_{ji}^{(s)*}$.

As in the stationary case, η_{12} is the part of the variance of the forecast error of the first market due to the second market's innovation terms, divided by the total variance of the first market's forecast error, when the forecast horizon tends towards infinity. An important point to note is that in the non-stationary case, the denominators of η_{12} and η_{21} are identical, as they represent the limit cases of a forecast horizon tending towards infinity. Hence, in the ratio of these contributions which we used to construct the stationary information share, the denominator cancels out, yielding Hasbrouck's information share (7):

$$SIS_1 = \frac{\eta_{21}}{\eta_{12} + \eta_{21}} = \frac{\sigma_1^2 \cdot \psi_1^2}{\sigma_1^2 \cdot \psi_1^2 + \sigma_2^2 \cdot \psi_2^2} = IS_1, \qquad SIS_2 = \frac{\eta_{12}}{\eta_{12} + \eta_{21}} = IS_2. \tag{27}$$

This result ensures the convergence of our new measure SIS to Hasbrouck's information share IS in the border case of non-stationarity and diagonal Ω structure. The proof of convergence for a non-diagonal covariance matrix Ω proceeds analogously, using the Cholesky decomposition M as in Section 2.2. We start with Equation (22) by setting $\Omega = M \cdot M'$:

$$\begin{aligned}
&Var\left(FE_j(y_\tau)\right) \\
&= \tau \cdot [\Psi(1)]_j \, M \cdot M' \, [\Psi(1)]_j' + \sum_{s=1}^{\tau} [\Psi^*_{\tau-s}]_j \, M \cdot M' \, [\Psi^{*'}_{\tau-s}]_j + 2 \sum_{s=1}^{\tau} [\Psi(1)]_j \, M \cdot M' \, [\Psi^{*'}_{\tau-s}]_j \\
&= \tau \cdot [\Phi(1)]_j \cdot I \cdot [\Phi(1)]_j' + \sum_{s=1}^{\tau} [\Phi^*_{\tau-s}]_j \cdot I \cdot [\Phi^{*'}_{\tau-s}]_j + 2 \sum_{s=1}^{\tau} [\Phi(1)]_j \cdot I \cdot [\Phi^{*'}_{\tau-s}]_j,
\end{aligned} \tag{28}$$

with $[\Phi(1)]_j = [\Psi(1) \cdot M]_j$ and $[\Phi^*_{\tau-s}]_j = [\Psi^*_{\tau-s} \cdot M]_j$. As in Section 3.2, we obtain orthonormal innovation terms $w_t = M'\epsilon_t$ which are uncorrelated with unit variance. Thus by following the derivation in the uncorrelated case, we get for Equation (27):

$$\lceil SIS_1 \rceil := \frac{\phi_1^2}{\phi_1^2 + \phi_2^2}, \qquad \lfloor SIS_2 \rfloor := \frac{\phi_2^2}{\phi_1^2 + \phi_2^2}, \tag{29}$$

with

$$\begin{pmatrix} \phi_1 \\ \\ \phi_2 \end{pmatrix}' = [\Psi(1) \cdot M]_j = \psi \cdot M = \begin{pmatrix} \psi_1 \cdot \sigma_1 + \psi_2 \cdot \sigma_2 \cdot \rho \\ \\ \psi_2 \cdot \sigma_2 \cdot \sqrt{1-\rho^2} \end{pmatrix}' . \tag{30}$$

So also in the general case with correlated innovation terms, we obtain the same limit result as with the original definition of the information share in Equation (8), proofing the equivalence of our new measure SIS with Hasbrouck's information share IS in the border case of non-stationarity.[5]

4 Numerical Convergence Behavior

In this section, we provide some numerical examples how the stationary information share converges to Hasbrouck's information share. For this purpose, we construct a stationary VAR(1) process with matrix

$$A_1 = \begin{pmatrix} 1 - a_{12} - \Delta & a_{12} \\ \\ a_{21} & 1 - a_{21} - \Delta \end{pmatrix}, \tag{31}$$

with minor diagonal elements a_{12} and a_{21}, and a convergence parameter $\Delta > 0$. One of the roots of the characteristic polynomial of A_1 is $1-\Delta$. By letting $\Delta \to 0$, the root approaches unity, thus the process converges from the stationary case to the non-stationary case. As an additional example, we furthermore consider a VAR(5) process with matrices A_i for $i = 1, \dots 5$ by setting

$$A_i = \begin{pmatrix} \frac{2^{5-i}}{2^5-1} - \frac{a_{12}+\Delta}{5} & \frac{a_{12}}{5} \\ \\ \frac{a_{21}}{5} & \frac{2^{5-i}}{2^5-1} - \frac{a_{21}+\Delta}{5} \end{pmatrix}. \tag{32}$$

[5]In Appendix A we additionally provide a more technical proof for the convergence of the stationary information share (16) to the classic information share (8) in case of a VAR(1)-process.

The idea of this construction is to have an impact that decreases with lag order, achieved by the term $\frac{2^{5-i}}{2^5-1}$, while the sum of the A_i equals the matrix A_1 in the VAR(1) process. Hence, we have the same convergence to a unit root of the characteristic polynomial and thus from the stationary to the non-stationary case when Δ tends towards zero. Appendix B demonstrates the structure of the A_i and the calculation of the SIS by an example.

In the following, we analyze the convergence from SIS to IS. Note that depending on the convergence parameter Δ, by construction of the A_i matrices the setting refers to substantially different processes, not to equivalent processes sampled at different frequencies. Without loss of generality, we demonstrate numerical convergence by referring to the first market, so strictly speaking we do not compute SIS and IS, but the upper bounds for the first market, $\lceil SIS_1 \rceil$ and $\lceil IS_1 \rceil$.

As mentioned, we consider lag orders of one and five, and we choose different minor diagonal elements $a_{12}, a_{21} \in \{0.1, 0.3\}$. Moreover, we consider potential contemporaneous correlation ($\rho = 0.7$) in the innovation terms ϵ_t and assume different variances ($\sigma_1 = 0.2$ and $\sigma_2 \in \{0.1, 0.2\}$). We obtain 24 different cases listed in Table 1 representing all potential variations of the underlying VAR matrices A_i. The results of numerical convergence of the 24 cases are presented in Figures 1 and 2.

Insert Table 1 about here.

Insert Figure 1 and Figure 2 about here.

The left columns of each subfigure (1(a), 1(b), 2(a), 2(b)) refer to the VAR(1) processes, the right columns to the VAR(5) processes. We exemplarily take a closer look at the situation with identical minor diagonal elements $a_{12} = a_{21} = 0.1$ (Figure 1(a)). In the first row, also the variance terms σ_1 and σ_2 are identical and there is no cross-correlation. It is therefore not surprising that both measures, IS and SIS, are identical at 0.5 in all cases. In the second row, the innovation term variance of the second market is lower, resulting in a noticeable rise in both information share measures of the first market. Moreover, we perceive an approach of the dotted line, representing our new measure in the stationary case, starting at approximately $SIS = 0.9$, to the bold line, representing Hasbrouck's reference value in the non-stationary case ($IS = 0.8$). Similar results can be found in the third row, where error terms exhibit both different variance and cross-correlation. Because the Cholesky decomposition attaches greater weight to the first market, the correlation term $\rho = 0.7$ increases both information share measures further: The dashed line starts at about $SIS = 0.98$ and decreases to the reference value of $IS = 0.93$. Finally, we remark that the results are fairly consistent through the two columns, representing the different maximum lag orders $k = 1$ and $k = 5$.

The other situations of minor diagonal elements (Subfigures 1(b), 2(a), and 2(b)) exhibit a qualitatively similar convergence behavior and are therefore not discussed in detail. Instead, we consider the log-log graphs of the differences $|IS - SIS|$ with respect to the convergence parameter Δ in Figure 3.

Insert Figure 3 about here.

The four graphs of Figure 3 correspond with the four subfigures 1(a)–2(b) and thus differ by the the minor diagonal elements. The six graphs of each subfigure 1(a)–2(b) are each represented by a convergence line in the respective graph of Figure 3.[6] VAR(1) processes are shown with solid lines, VAR(5) processes with dashed lines.

[6]For the first row of Figure 1(a) and Figure 2(b) the differences are zero and thus not meaningful in a log-log scale, hence the corresponding graphs only show four lines.

Together with Figure 1(b), the second graph of Figure 3 shows that with asymmetric minor diagonal elements the difference $|IS - SIS|$ is about a factor ten smaller than in the symmetric case of Figure 1(a). This can be explained by the high reference value of the IS, particularly when also the variances are asymmetric ($IS = 0.989$), which leaves little leeway for deviations. The same can be observed with switched minor diagonal elements and identical variances (first row of Figure 2(a), corresponding to the respective line in the third graph of Figure 3). In contrast, with asymmetries in the minor diagonal elements and the variances in opposite directions (second and third row of Figure 2(a)), the deviations are the largest. Finally, for larger symmetric minor diagonal elements (Figure 2(b) and fourth graph of Figure 3), the deviations are similar to the first setting, albeit a little bit smaller. Regarding the VAR(1) and VAR(5) processes, the convergence behavior is quite similar in all situations.

The log-log plots furthermore allow a determination of the order of convergence. The stationary information share $SIS(\Delta)$ converges at order a to Hasbrouck's information share $IS = SIS(0)$ if there exists a constant $c > 0$ such that

$$|SIS(\Delta) - IS| \leq c \cdot \Delta^a \tag{33}$$

for all sufficiently small $\Delta \in \mathbb{R}_+$. If $a = 1$, SIS is said to converge linearly to IS. Taking logarithms at each side of Equation (33) yields

$$\log(|SIS(\Delta) - IS|) \leq \log(c) + a \cdot \log(\Delta), \tag{34}$$

hence, the order of convergence is graphically given by the slopes of the convergence line on a log-log scale. We see in Figure 3 that the slope a of the lines are close to one. That means we obtain a linear order of convergence of SIS to IS.

5 Application: Volatility Discovery Around the World

5.1 Review of Empirical Literature

In this section, we discuss the results of an empirical application, analyzing information shares of volatility indices of different countries: the US, German, and Japanese market. Effects of volatility spillovers betweens different countries or regions have been analyzed in the literature by numerous papers. Most of them, however, focus on realized volatilities instead of implied volatilities as condensed in volatility indices such as the VIX. Seminal contributions are those of Engle et al. (1990), who find a "meteor shower effect"—a phenomenon of intra-daily volatility spillovers from one market to the next—for the influence from news in the yen/dollar exchange rates to volatility in the Japanese market, and Hamao et al. (1990), who utilize ARCH models, finding strong volatility spillover effects from the US and the UK stock markets to the Japanese market and vice versa.

The predominant methods for analyzing spillover effects are Granger causality tests and insights from (structural) VAR or GARCH models. While these methods allow an identification of a direction of spillover or a quantification of cross-effects, respectively, they do not provide a measure of relative informational leadership such as the information share. As discussed, most closely related to our metric are the spillover indices used by Diebold and Yilmaz (2009) and Diebold and Yilmaz (2012). In contrast to these indices, the stationary information share allows for attributing relative shares of informational leadership.

Especially regarding our three considered markets—the US, Germany, and Japan—, previous literature mostly finds a dominance of the US market. Diebold and Yilmaz (2009) show that innovations in US volatility are responsible for 26.9 percent of the error variance in forecasting German volatility and for 2.7 percent of the error variance in forecasting

Japanese volatility. On the other side, innovations in German volatility are responsible for only 1.9 percent of the error variance in forecasting US volatility and even less in forecasting Japanese volatility (0.7 percent). The Japanese influence on German and US volatility seems to be negligible (below 1 percent). Dimpfl and Jung (2012) use structural VAR and GARCH models to show that there are spillovers from the US market to the Japanese market and from the Japanese market to the European market in the pre-crisis period 2002–2006. Also Ehrmann et al. (2011) confirm the dominant role of the US market. Studying transmission between money, bond, and equity markets, they find that US markets explain, on average, more than 25% of European market movements in the period 1989–2004, whereas European markets account for 8% of the variance of US asset prices. Slightly different evidence is provided by Jung and Maderitsch (2014), who investigate volatility transmission between stock markets in Hong Kong, Europe, and the United States from 2000 to 2011 using a Heterogeneous Autoregressive Distributed Lag Model. Concerning the US market, they find evidence for a strong positive statistically significant volatility spillover from Europe. The European market, on the other hand, is also influenced by the US market, but not as much as vice versa. Also Clements et al. (2015) report a higher importance of Europe for global volatility discovery for more recent data (2005–2013). They investigate the effects between foreign exchange, equity, and bond future markets in the US, Japan, and Europe by a multivariate GARCH model, measuring bivariate volatility spillovers from one country to the proceeding zone. While US news seems to have only small but significant explanatory power on Japanese volatility, there is a huge significant spillover from Europe to the US market.

While the previously mentioned studies focus on realized volatilities, only a few papers analyze implied volatility indices. The two concepts are closely related, however, there is an important difference with respect to the interpretation of the results: Spillover of realized volatilities refers to short-term reactions of the stock markets, whereas implied volatility indices refer to mid-term expectations of volatility. Aboura (2003) analyzes implied volatility indices for the US, the French, and the German market in the years 1994–1999. In line with other results for realized volatilities, she finds a predominant role of the VIX compared to the VDAX and the French volatility index VXI. Konstantinidi et al. (2008), considering European and US implied volatility indices in the period 2001–2007, validate this result by showing that the information contained in all implied volatility indices can be used to predict each European index separately but not the US index.

More timely data covering the recent crises underline the importance of the German market and the decreasing influence of US market volatility. Narwal et al. (2012) use a BEKK-GARCH model to examine the bivariate spillover and transmission of the newly developed implied Indian volatility index on the US volatility index VIX, on the German volatility index VDAX, and other indices in the period 2007–2011. They find traces of unidirectional shock spillover from Indian volatility to the VIX, while spillovers are running unidirectional from Germany to India. On the contrary, Gupta and Kamilla (2015) report that the VIX also in the most recent period 2011–2013 has the highest causality effects over its other peer indices and explains approximately 11% or 5% of the fluctuation of German and Japanese volatility, respectively. However, these findings might be flawed by the method of multivariate variance decomposition, which heavily depends on the ordering of the markets.

5.2 Empirical Findings

We consider the CBOE volatility index VIX for the US S&P 500, the VDAX-New for the German DAX, and the NikkeiVI for the Japanese Nikkei 225. Data is taken from Thomson Reuters on a daily basis of closing prices, starting in June 1996 (VIX and VDAX-New), respectively in January 1998 (NikkeiVI), and ending in June 2016. To investigate potential changes in volatility discovery over time, we divide the whole period in 4 subperiods, referring to crisis periods as defined by Ballester et al. (2016):

- The pre-crisis period starts in 1996/1998 and ends at 2007/07/17, the date when Bear Stearns disclosed that its major subprime hedge fund had lost nearly all of its value.
- The global financial crisis starts with the Bear Stearns distress and results in the European sovereign debt crisis, which starts with the EU summit in October 2009.
- The European debt crisis ends with a package for the Bank of Cyprus at 2013/03/19.
- The post-crisis period runs from 2013/03/20 to the end of the data sample in June 2016.

An overview of the three different volatility time series over the entire period is presented by Figure 4. All volatility indices reach their maxima of 80% or more within the global financial crisis, while the minima are reached in the pre-crisis period.

Insert Figure 4 about here.

Descriptive statistics, correlations, and results of Augmented Dickey-Fuller tests for stationarity are reported in Table 2. Correlations between the daily index returns tend to increase in crisis periods. The correlation between VIX and VDAX-New is the highest, followed by the correlation between VDAX-New and NikkeiVI, and the lowest correlation is observed between VIX and NikkeiVI.

Insert Table 2 about here.

For all indices, the null hypothesis of a unit root is rejected at a significance level of 1% for the entire period and at least at the 10% level for the non-crisis subperiods. In the second subperiod, during the global financial crisis, non-stationarity is not rejected for all indices, while in the third subperiod, the European sovereign debt crisis, it is not rejected for the VIX and the VDAX-New. This could be explained by the high increase in volatilities during the crisis periods and by the solicited fact that the tested time series could be too short to deliver reasonable results. Concluding, we can assume stationarity for all three volatility time series.

We compute bivariate stationary information shares.[7] Results are presented in Table 3. The table reports upper and lower bounds of the SIS for each of the three pairs of markets for the entire period and the four subperiods. Following Baillie et al. (2002), we consider the mean of the upper and lower bound as the actual information share of a market. By this, correlation effects are allocated evenly to the respective markets. However, as we have no further information of the direction of correlation effects, this allocation is only based on a simple rule of thumb. We therefore label a market as significantly dominating another markets in terms of volatility discovery if the lower bound of the market is still larger than the upper bound of the other market.

Insert Table 3 about here.

[7] We refrain from considering the three-markets system as a whole because of too high cross-correlations.

Considering the entire period, the VIX has the largest information shares with a mean of 70% relative to the VDAX-New and 73% relative to the NikkeiVI, however, only the latter being significant. Comparing the VDAX-New and the NikkeiVI, with mean information shares of 58% and 42% there is no clear leader.

This opaque picture becomes clearer when we look at the different subperiods. Within the pre-crisis period until 2007, we observe a strong significant leadership of the VIX with respect to the VDAX-New with an SIS of 89%. The VIX also dominates the NikkeiVI with an SIS of 71%, which is slightly smaller than in the entire period. Regarding the VDAX-New and the NikkeiVI, in contrast to the entire period we observe a larger SIS for the NikkeiVI with 61%, which is significant. In interpreting these results one should keep in mind that the "pre-crisis" period actually covers Japan's lost decade in the aftermath of the Japanese crisis in the early 1990s.

During the global financial crisis, the dominance of the VIX becomes even more pronounced with SIS values of 90% with respect to the VDAX-New and 95% with respect to the NikkeiVI. The relation between the VDAX-New and the NikkeiVI is switched to a leadership of the German volatility index with an SIS of 69%, albeit not significant.

Within the following European debt crisis, the importance of the German market in terms of volatility discovery becomes stronger. The VDAX-New is now dominating the NikkeiVI significantly with an SIS of 78%. With respect to the VIX, the leadership of the US market is no longer existent; the respective SIS values of 55% and 45% indicate equable influences in both directions. Further, the VIX still dominates the NikkeiVI with an SIS of 84%.

The trend of an increasing importance of the German market continues in recent years in the "post-crisis" period. The VDAX-New now exhibits an SIS of 70% with respect to the VIX. The SIS with respect to the NikkeiVI rises to 92%. Again, we further see a leadership of the VIX over the NikkeiVI with an SIS of 81%.

The core findings can be summarized as follows:

- Regarding the entire period, the VIX is the most important index in terms of volatility discovery with SIS values of 70% and 73% with respect to the VDAX-New and the NikkeiVI, respectively.
- The relevance of the VDAX-New has continuously increased, in particular with the European debt crisis and its aftermath. While previously being dominated by the other indices, the VDAX-New now has a higher information share compared to the NikkeiVI and also the VIX.
- The NikkeiVI had a higher information share compared to the VDAX-New in the pre-crisis period until 2007. It is now dominated by both the VIX and the VDAX-New.

These results are in line with previous literature, which has reported a dominance of US volatility in previous periods, which has decreased in favor of European markets, in particular the German market, in more recent periods. Our findings provide a clearer picture of this change in global volatility discovery and allow an interpretation in terms of the origin of a crisis: Those regions which are most influenced by crises tend to send shock waves around the world. In the first subperiod, the Japanese crisis was still prevalent in the country, and obviously changes in market expectations about future volatility in Japan were transmitted into similar changes in Germany. The global financial crisis had its origin in the United States, and most shock events spread from the US around the globe, including Germany and Japan. As a consequence, the SIS values of the VIX reached a maximum during this period. The following European debt crisis shifted the focus to

Europe, represented by Germany in our sample. Thus, SIS values of the VDAX-New increased and still tend to increase although the peak of the European debt crisis seems to have passed. Obviously, global market participants still consider Europe as the most important origin for global shock waves. This is not surprising, since for instance the Greek crisis is far from being solved with potential impacts on Germany and markets around the world.

6 Conclusion

Information share measures are based on the contribution of a market's innovation to the total variance of the efficient underlying price innovations. All modifications of the seminal information share concept proposed by Hasbrouck (1995) rely on the assumption that the price processes on the different markets are co-integrated of order 1, and that single price processes are non-stationary. This necessary assumption seems to have been subtly ignored in the recent past by several authors applying information share measures to stationary time series. While such an application might work for stationary, though highly persistent time series, the results lack an economic interpretation: Information shares measure the relative contribution to the permanent impact of shocks in the time series systems—when time series are stationary, there is however no permanent shock.

In this paper, we present a new measure, a stationary information share, based on the idea of the classic variance decomposition for stationary time series. This measure is related to the spillover indices introduced by Diebold and Yilmaz (2009) and Diebold and Yilmaz (2012). We show that the stationary information share converges to Hasbrouck's information share in the border case of non-stationarity. Thus, we establish a link between the spillover literature and the information share literature.

We apply the stationary information share to gain insight into the interdependence between implied volatilities for countries around the globe: the United States, Germany, and Japan. By considering different subperiods, defined by recent financial crises, we demonstrate that the origin of a crisis plays a decisive role for volatility discovery: In the aftermath of the Japan crisis, the Japanese volatility index has a significantly larger information share for the German volatility index than vice versa. This relationship turns into its opposite with the emergence of the European debt crisis. The US volatility index VIX exhibits the largest information shares over the entire period of 20 years with a peak during the global financial crisis which spread from the US across the world. In recent periods, the VIX has lost its dominance over the German market, as global shocks tend to originate from Europe and its sovereign debt crisis.

References

Aboura, S. (2003). International transmission of volatility: A study on the volatility indexes VX1, VDAX and VIX. Working Paper, ESSEC Business School.

Anthony, M. and M. Harvey (2012). *Linear algebra: concepts and methods*. Cambridge University Press.

Baillie, R. T. and T. Bollerslev (1994). Cointegration, fractional cointegration, and exchange rate dynamics. *Journal of Finance 49*, 737–745.

Baillie, R. T., G. G. Booth, Y. Tse, and T. Zabotina (2002). Price discovery and common factor models. *Journal of Financial Markets 5*, 309–321.

Ballester, L., B. Casu, and A. González-Urteaga (2016). Bank fragility and contagion: Evidence from the CDS market. *Journal of Empirical Finance forthcoming*.

Beveridge, S. and C. R. Nelson (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary Economics 7*, 151–174.

Blanco, R., S. Brennan, and I. W. Marsh (2005). An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps. *Journal of Finance 60*, 2255–2281.

Booth, G. G., J.-C. Lin, T. Martikainen, and Y. Tse (2002). Trading and pricing in upstairs and downstairs stock markets. *Review of Financial Studies 15*, 1111–1135.

Booth, G. G., R. W. So, and Y. Tse (1999). Price discovery in the German equity index derivatives markets. *Journal of Futures Markets 19*, 619–643.

Cheung, Y.-W. and K. S. Lai (1993). A fractional cointegration analysis of purchasing power parity. *Journal of Business & Economic Statistics 11*, 103–112.

Christensen, B. J. and M. Ø. Nielsen (2006). Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting. *Journal of Econometrics 133*, 343–371.

Clements, A. E., S. A. Hurn, and V. V. Volkov (2015). Volatility transmission in global fnancial markets. *Journal of Empirical Finance 32*, 3–18.

Dias, G. F., C. M. Scherrer, and F. Papailias (2016). Volatility discovery. Working Paper, Aarhus University and Queen's University.

Diebold, F. X. and K. Yilmaz (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal 119*, 158–171.

Diebold, F. X. and K. Yilmaz (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting 28*, 57–66.

Dimpfl, T. and R. C. Jung (2012). Financial market spillovers around the globe. *Applied Financial Economics 22*, 45–57.

Ehrmann, M., M. Fratzscher, and R. Rigobon (2011). Stocks, bonds, money markets and exchange rates: Measuring international financial transmission. *Journal of Applied Econometrics 26*, 948–974.

Engle, R. F., T. Ito, and W.-L. Lin (1990). Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market. *Econometrica 58*, 525–542.

Fricke, C. and L. Menkhoff (2011). Does the "Bund" dominate price discovery in euro bond futures? Examining information shares. *Journal of Banking & Finance 35*, 1057–1072.

Gonzalo, J. and C. Granger (1995). Estimation of common long-memory components on cointegrated systems. *Journal of Business & Economic Statistics 13*, 27–35.

Grammig, J., M. Melvin, and C. Schlag (2005). Internationally cross-listed stock prices during overlapping trading hours: Price discovery and exchange rate effects. *Journal of Empirical Finance 12*, 139–164.

Gupta, D. and U. Kamilla (2015). Dynamic linkages between implied volatility indices of developed and emerging financial markets: An econometric approach. *Global Business Review 16*, 456–575.

Hamao, Y., R. W. Masulis, and V. Ng (1990). Correlations in price changes and volatility across international stock markets. *Review of Financial Studies 3*, 281–307.

Harris, F. H., T. H. McInish, and R. A. Wood (2002). Security price adjustment across exchanges: An investigation of common factor components for dow stocks. *Journal of Financial Markets 5*, 277–308.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *Journal of Finance 50*, 1175–1199.

Hasbrouck, J. (2002). Stalking the 'efficient price' in market microstructure specifications: An overview. *Journal of Financial Markets 5*, 329–339.

Hasbrouck, J. (2003). Intraday price formation in U.S. equity index markets. *Journal of Finance 58*, 2375–2399.

Hendry, D. F. and K. Juselius (2001). Explaining Cointegration Analysis: Part II. *The Energy Journal 22*, 75–120.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*, 327–343.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica 59*, 1551–1580.

Jung, R. C. and R. Maderitsch (2014). Structural breaks in volatility spillovers between international financial markets: Contagion or mere interdependence. *Journal of Banking & Finance 47*, 331–342.

Kehrle, K. and F. J. Peter (2013). Who moves first? An intensity-based measure for information flows across stock exchanges. *Journal of Banking & Finance 37*, 1629–1642.

Konstantinidi, E., G. Skiadopoulos, and E. Tzagkaraki (2008). Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices. *Journal of Banking & Finance 32*, 2401–2411.

Koop, G., H. M. Pesaran, and S. M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics 74*, 119–147.

Lien, D. and K. Shrestha (2009). A new information share measure. *Journal of Futures Markets 29*, 377–395.

Lien, D. and K. Shrestha (2014). Price discovery in interrelated markets. *Journal of Futures Markets 34*, 203–219.

Mizrach, B. and C. J. Neely (2008). Information shares in the U.S. treasury market. *Journal of Banking & Finance 32*, 1221–1233.

Narwal, K. P., V. P. Sheera, and R. Mittal (2012). Spillovers and transmission in emerging and mature markets implied volatility indices. *International Journal of Financial Management 2*, 47–59.

Pesaran, H. H. and Y. Shin (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters 58*, 17–29.

Putniņš, T. J. (2013). What do price discovery metrics really measure? *Journal of Empirical Finance 23*, 68–83.

Schay, G. (2012). *A concise introduction to linear algebra.* Birkhäuser.

Stock, J. H. and M. W. Watson (1988). Testing for common trends. *Journal of American Statistical Association 83*, 1097–1107.

Tse, Y., P. Bandyopadhyay, and Y.-P. Shen (2006). Intraday price discovery in the DJIA index markets. *Journal of Business Finance & Accounting 33*, 1572–1585.

Yan, B. and E. Zivot (2010). A structural analysis of price discovery measures. *Journal of Financial Markets 13*, 1–19.

Zapata, H. O. and R. T. Fortenbery (1996). Stochastic interest rates and price discovery in selected commodity markets. *Review of Agricultural Economics 18*, 643–654.

A Technical Proof of Convergence

In addition to the general proof of convergence given in Section 3.3, we present a more direct, but technical proof of convergence of the SIS to the IS in the case of a VAR(1)-process when the root of its characteristic polynomial tends towards one.

Theorem 1 (Convergence of stationary information share to Hasbrouck's information share)**.** *Assume a VAR(1) process $y_t = A_1 \cdot y_{t-1} + \epsilon_t$. In the border case of non-stationarity the stationary information share converges to Hasbrouck's information share.*

Proof. While developing the non-stationary reference value IS, we make the assumption of a VAR(1) process with $A_1 := A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}_+^{2\times 2}$, where each of the coefficients are less than one and the sum over rows equals one:

$$a_{ij} \leq 1 \; \forall i,j \in \{1,2\}, \qquad a_{11} + a_{12} = 1, \qquad a_{21} + a_{22} = 1. \tag{35}$$

First we develop Hasbrouck's information share IS under the presented restrictions as the reference value. In the VAR(1) case, the cointegrating vector α is given by (see also Equation (55) below)

$$\alpha = (-a_{12}, a_{21})'. \tag{36}$$

Without loss of generality we calculate the upper bound $\lceil IS_1 \rceil$ (in the following simply referred to as IS_1). Computing ψ' as an orthogonal complement of α:

$$\psi' = \alpha_\perp := \left(\frac{1}{a_{12}}, \frac{1}{a_{21}}\right)', \tag{37}$$

and setting this result into Equation (8), we obtain Hasbrouck's information share IS_1 as the reference value:

$$
\begin{aligned}
IS_1 &= \frac{(\psi_1\sigma_1 + \psi_2\sigma_2\rho)^2}{(\psi_1\sigma_1 + \psi_2\sigma_2\rho)^2 + \left(\psi_2\sigma_2\sqrt{1-\rho^2}\right)^2} = \frac{1}{1 + \dfrac{\left(\psi_2\sigma_2\sqrt{1-\rho^2}\right)^2}{(\psi_1\sigma_1 + \psi_2\sigma_2\rho)^2}} \\
&= \frac{1}{1 + \dfrac{\left(\dfrac{1}{a_{21}}\sigma_2\sqrt{1-\rho^2}\right)^2}{\left(\dfrac{1}{a_{12}}\sigma_1 + \dfrac{1}{a_{21}}\sigma_2\rho\right)^2}} = \frac{1}{1 + \dfrac{(\sigma_2)^2 \cdot (1-\rho^2)}{\left(\dfrac{a_{21}}{a_{12}}\sigma_1 + \sigma_2\rho\right)^2}}.
\end{aligned} \tag{38}
$$

After developing the price discovery measure in the non-stationary case, we will now focus on the information share concept when time series are stationary. Each of the coefficients of A is thus less than one and the sums over the rows tend toward one:

$$
a_{ij} \leq 1 \; \forall i,j \in \{1,2\}, \qquad a_{11} + a_{12} \lesssim 1, \qquad a_{21} + a_{22} \lesssim 1. \tag{39}
$$

In Section 3, we defined SIS based on an orthonormalized VAR system $y_t = \sum_{p=0}^{\infty} \Phi_p \cdot w_{t-p}$, see Equation (10). To characterize the moving average matrices Φ_p, we develop the moving average representation for a general VAR(1) process:

$$
\begin{aligned}
y_t &= A \cdot y_{t-1} + \epsilon_t \\
&= \sum_{p=0}^{\infty} A^p \epsilon_{t-p} \\
&= \sum_{p=0}^{\infty} A^p \cdot M \cdot M' \cdot \epsilon_{t-p} \\
&= \sum_{p=0}^{\infty} \Phi_p \cdot w_{t-p},
\end{aligned}
$$

where M represents a triangular matrix given by the Cholesky decomposition of the matrix A, $\Phi_p := A^p \cdot M$ and $w_{t-p} := M' \cdot \epsilon_{t-p}$.

How do the matrices A^p and consequently Φ_p look like? We use the theory of diagonalization and eigendecomposition to get an analytical solution for the shape of A^p. Applying basic linear algebra, if a matrix A is diagonalizable we can decompose it into matrices containing its eigenvectors and eigenvalues as follows (e.g., Schay (2012) or Anthony and Harvey (2012)):

$$A = S \cdot D \cdot S^{-1},$$

where S contains the eigenvectors of A and D contains its eigenvalues. Furthermore we know that, if A is diagonalizable,

$$A^p = S \cdot D^p \cdot S^{-1}. \tag{40}$$

In Lemma 1 below, we calculate eigenvalues $\lambda_{1,2}$ and eigenvectors $v_{1,2}$, finally concluding:

$$\lambda_{1/2} = \frac{a_{11} + a_{22}}{2} \pm \sqrt{\left(\frac{a_{11} + a_{22}}{2}\right)^2 - a_{11}a_{22} + a_{12}a_{21}}, \tag{41}$$

$$v_1 = \left(\frac{1}{1+B^2}, \frac{B}{1+B^2}\right)', \tag{42}$$

$$v_2 = \left(\frac{C}{1+C^2}, \frac{1}{1+C^2}\right)', \tag{43}$$

where

$$B := \frac{\lambda_1 - a_{11}}{a_{12}}, \qquad C := \frac{\lambda_2 - a_{22}}{a_{21}}. \tag{44}$$

By normalization of the eigenvalues, the matrix S arises:

$$S = \begin{pmatrix} \frac{1}{1+B^2} & \frac{C}{1+C^2} \\ \frac{B}{1+B^2} & \frac{1}{1+C^2} \end{pmatrix},$$

$$S^{-1} = \left(\frac{(1+B^2)(1+C^2)}{1-BC}\right) \begin{pmatrix} \frac{1}{1+C^2} & -\frac{C}{1+C^2} \\ -\frac{B}{1+B^2} & \frac{1}{1+B^2} \end{pmatrix}.$$

Putting our results into Equation (40), we obtain:

$$A^p = S \cdot D^p \cdot S^{-1}$$

$$= \frac{(1+B^2)(1+C^2)}{1-BC} \cdot \begin{pmatrix} \frac{1}{1+B^2} & \frac{C}{1+C^2} \\ \frac{B}{1+B^2} & \frac{1}{1+C^2} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1^p & 0 \\ 0 & \lambda_2^p \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{1+C^2} & -\frac{C}{1+C^2} \\ -\frac{B}{1+B^2} & \frac{1}{1+B^2} \end{pmatrix}$$

$$= \frac{1}{1-BC} \cdot \begin{pmatrix} \lambda_1^p - BC\lambda_2^p & C(\lambda_2^p - \lambda_1^p) \\ B(\lambda_1^p - \lambda_2^p) & \lambda_2^p - BC\lambda_1^p. \end{pmatrix}.$$

So, we finally get an analytical closed formula for the moving average matrices Φ_p:

$$\Phi_p = A^p \cdot M$$

$$= \frac{1}{1-BC} \cdot \begin{pmatrix} \lambda_1^p - BC\lambda_2^p & C(\lambda_2^p - \lambda_1^p) \\ B(\lambda_1^p - \lambda_2^p) & \lambda_2^p - BC\lambda_1^p \end{pmatrix} \cdot \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2(1-\rho^2)^{\frac{1}{2}} \end{pmatrix}$$

$$= \frac{1}{1-BC} \cdot \begin{pmatrix} \sigma_1(\lambda_1^p - BC\lambda_2^p) + \rho\sigma_2 C(\lambda_2^p - \lambda_1^p) & \sigma_2 C(1-\rho^2)^{\frac{1}{2}}(\lambda_2^p - \lambda_1^p) \\ \sigma_1 B(\lambda_1^p - \lambda_2^p) + \rho\sigma_2(\lambda_2^p - BC\lambda_1^p) & \sigma_2(1-\rho^2)^{\frac{1}{2}}(\lambda_2^p - BC\lambda_1^p) \end{pmatrix}. \quad (45)$$

To obtain SIS_1 in Equation (16), we have to calculate η_{12} and η_{21} by Equation (15) from the entries of the moving average matrices Φ_p. Neglecting the factor $\frac{1}{1-BC}$ in Equation (45) (which is possible because we build ratios in the following), we obtain for the first entry of Φ_p:

$$\left(\phi_{11}^{(p)}\right)^2 = \left(\sigma_1(\lambda_1^p - BC\lambda_2^p) + \rho\sigma_2 C(\lambda_2^p - \lambda_1^p)\right)^2 \quad (46)$$

$$= \sigma_1^2 \cdot (\lambda_1^p - BC\lambda_2^p)^2 + 2 \cdot \sigma_1\sigma_2\rho C(\lambda_1^p - BC\lambda_2^p) \cdot (\lambda_2^p - \lambda_1^p) + (\rho\sigma_2)^2 C^2(\lambda_2^p - \lambda_1^p)^2.$$

Calculating the η's requires summing up $\phi_{11}^{(p)}$ (and additionally $\phi_{12}^{(p)}$, $\phi_{21}^{(p)}$ and $\phi_{22}^{(p)}$) from $p = 0, \dots \infty$. For the first expression on the right side of Equation (46), we can use the infinite geometric series:

$$\sum_{p=0}^{\infty} \sigma_1^2 \cdot (\lambda_1^p - BC\lambda_2^p)^2$$

$$= \sum_{p=0}^{\infty} \sigma_1^2 \cdot \left(\lambda_1^{2p} - 2 \cdot BC(\lambda_1\lambda_2)^p + (BC)^2\lambda_2^{2p}\right)$$

$$=\sigma_1^2 \cdot \left(\frac{1}{1-\lambda_1^2} - 2 \cdot \frac{BC}{1-\lambda_1\lambda_2} + \frac{(BC)^2}{1-\lambda_2^2} \right).$$

Analogously, the other two terms of Equation (46) are computed. Finally we get:

$$\sum_{p=0}^{\infty} \left(\phi_{11}^{(p)}\right)^2 = \sigma_1^2 \cdot K_{\lambda_2} + 2\sigma_1\sigma_2\rho C \cdot X_{\lambda_2} + \sigma_2^2\rho^2 C^2 \cdot K,$$

where

$$\begin{aligned}
K_{\lambda_2} &:= \left(\frac{1}{1-\lambda_1^2} - 2 \cdot \frac{BC}{1-\lambda_1\lambda_2} + \frac{(BC)^2}{1-\lambda_2^2} \right), \\
X_{\lambda_2} &:= \left(-\frac{1}{1-\lambda_1^2} + \frac{1+BC}{1-\lambda_1\lambda_2} - \frac{BC}{1-\lambda_2^2} \right), \\
K &:= \left(\frac{1}{1-\lambda_1^2} - 2 \cdot \frac{1}{1-\lambda_1\lambda_2} + \frac{1}{1-\lambda_2^2} \right).
\end{aligned}$$

The missing terms are obtained analogously:

$$\begin{aligned}
\sum_{p=0}^{\infty} \left(\phi_{12}^{(p)}\right)^2 &= \sigma_2^2 C^2 \cdot (1-\rho^2) \cdot K, \\
\sum_{p=0}^{\infty} \left(\phi_{21}^{(p)}\right)^2 &= \sigma_1^2 B^2 \cdot K + 2\sigma_1\sigma_2\rho B \cdot X_{\lambda_1} + \sigma_2^2\rho^2 \cdot K_{\lambda_1}, \\
\sum_{p=0}^{\infty} \left(\phi_{22}^{(p)}\right)^2 &= \sigma_2^2(1-\rho^2) \cdot K_{\lambda_1},
\end{aligned}$$

where

$$\begin{aligned}
K_{\lambda_1} &:= \left(\frac{(BC)^2}{1-\lambda_1^2} - 2 \cdot \frac{BC}{1-\lambda_1\lambda_2} + \frac{1}{1-\lambda_2^2} \right), \\
X_{\lambda_1} &:= \left(-\frac{BC}{1-\lambda_1^2} + \frac{1+BC}{1-\lambda_1\lambda_2} - \frac{1}{1-\lambda_2^2} \right).
\end{aligned}$$

Setting the results into Equation (15), we obtain η_{12} as follows:

$$\begin{aligned}
\eta_{12} &= \frac{\sigma_2^2 C^2 \cdot (1-\rho^2) \cdot K}{\sigma_2^2 C^2 \cdot (1-\rho^2) \cdot K + \sigma_1^2 \cdot K_{\lambda_2} + 2\sigma_1\sigma_2\rho C \cdot X_{\lambda_2} + \sigma_2^2\rho^2 C^2 \cdot K} \\
&= \frac{\sigma_2^2 C^2 \cdot (1-\rho^2) \cdot K}{\sigma_2^2 C^2 \cdot K + \sigma_1^2 \cdot K_{\lambda_2} + 2\sigma_1\sigma_2\rho C \cdot X_{\lambda_2}}
\end{aligned}$$

$$= \frac{\sigma_2^2 C^2 \cdot (1-\rho^2) \cdot \left[\frac{1}{1-\lambda_1^2} - 2 \cdot \frac{1}{1-\lambda_1\lambda_2} + \frac{1}{1-\lambda_2^2}\right]}{\sigma_2^2 C^2 \left[\frac{1}{1-\lambda_1^2} - 2 \cdot \frac{1}{1-\lambda_1\lambda_2} + \frac{1}{1-\lambda_2^2}\right] + \sigma_1^2 \left[\frac{1}{1-\lambda_1^2} - 2\frac{BC}{1-\lambda_1\lambda_2} + \frac{(BC)^2}{1-\lambda_2^2}\right] + 2\sigma_1\sigma_2\rho C \left[-\frac{1}{1-\lambda_1^2} + \frac{1+BC}{1-\lambda_1\lambda_2} - \frac{BC}{1-\lambda_2^2}\right]}.$$

Since we are interested in the border case when time series switch from stationarity to non-stationarity, we develop the eigenvalues $\lambda_{1,2}$ under this restriction in Lemma 1 below. Building the common denominator and letting $\lambda_1 \to 1$, η_{12} tends to:

$$\eta_{12} \to \frac{\sigma_2^2 C^2 \cdot (1-\rho^2) \cdot \left[(1-\lambda_1\lambda_2)(1-\lambda_2^2)\right]}{\sigma_2^2 C^2 \cdot \left[(1-\lambda_1\lambda_2)(1-\lambda_2^2)\right] + \sigma_1^2 \cdot \left[(1-\lambda_1\lambda_2)(1-\lambda_2^2)\right] - 2\sigma_1\sigma_2\rho C \cdot \left[(1-\lambda_1\lambda_2)(1-\lambda_2^2)\right]}$$
$$= \frac{\sigma_2^2 C^2 \cdot (1-\rho^2)}{\sigma_2^2 C^2 + \sigma_1^2 - 2\sigma_1\sigma_2\rho C}. \tag{47}$$

Analogously:

$$\eta_{21} \to \frac{\sigma_1^2 B^2 - 2\sigma_1\sigma_2\rho B^2 C + \sigma_2^2\rho^2 B^2 C^2}{\sigma_1^2 B^2 - 2\sigma_1\sigma_2\rho B^2 C + \sigma_2^2 B^2 C^2} = \frac{\sigma_1^2 - 2\sigma_1\sigma_2\rho C + \sigma_2^2\rho^2 C^2}{\sigma_1^2 - 2\sigma_1\sigma_2\rho C + \sigma_2^2 C^2}. \tag{48}$$

Putting these values into Equation (16), we get SIS_1 in the border case of non-stationarity:

$$SIS_1 = \frac{1}{1+\frac{\eta_{12}}{\eta_{21}}} \to \frac{1}{1+\frac{\frac{\sigma_2^2 C^2 \cdot (1-\rho^2)}{\sigma_2^2 C^2 + \sigma_1^2 - 2\sigma_1\sigma_2\rho C}}{\frac{\sigma_1^2 - 2\sigma_1\sigma_2\rho C + \sigma_2^2\rho^2 C^2}{\sigma_1^2 - 2\sigma_1\sigma_2\rho C + \sigma_2^2 C^2}}} = \frac{1}{1+\frac{\sigma_2^2 \cdot (1-\rho^2)}{\left(\frac{\sigma_1}{C} - \sigma_2\rho\right)^2}}.$$

In the non-stationary border case we moreover know from Lemma 1 below that $C \to -\frac{a_{12}}{a_{21}}$, so it finally follows

$$SIS_1 \to \frac{1}{1+\frac{\sigma_2^2 \cdot (1-\rho^2)}{\left(\sigma_1 \cdot (-\frac{a_{21}}{a_{12}}) - \sigma_2\rho\right)^2}} = \frac{1}{1+\frac{\sigma_2^2 \cdot (1-\rho^2)}{\left(\sigma_1 \cdot \frac{a_{21}}{a_{12}} + \sigma_2\rho\right)^2}} = IS_1,$$

by comparison with the reference value in Equation (38). □

Eigenvalues of A We calculate the eigenvalues $\lambda_{1/2}$ of A by setting:

$$det(A - \lambda I) \stackrel{!}{=} 0$$
$$\Rightarrow \lambda_{1/2} = \frac{a_{11}+a_{22}}{2} \pm \sqrt{\left(\frac{a_{11}+a_{22}}{2}\right)^2 - a_{11}a_{22} + a_{12}a_{21}}$$

$$= \frac{a_{11} + a_{22}}{2} \pm \sqrt{\left(\frac{a_{11} - a_{22}}{2}\right)^2 + a_{12} \cdot a_{21}}. \tag{49}$$

To obtain the associated eigenvectors $v_{1/2}$, from $A \cdot v_{1/2} = \lambda_{1/2} \cdot v_{1/2}$ we get

$$v_1 = \left(\frac{1}{1 + B^2}, \frac{B}{1 + B^2}\right)',$$
$$v_2 = \left(\frac{C}{1 + C^2}, \frac{1}{1 + C^2}\right)',$$

where $B = \frac{\lambda_1 - a_{11}}{a_{12}}$ and $C = \frac{\lambda_2 - a_{22}}{a_{21}}$.

To ensure that A is diagonalizable, λ_1 and λ_2 must be different from each other. This is the case if the sum under the root is greater than zero:

$$\left(\frac{a_{11} - a_{22}}{2}\right)^2 + a_{12} \cdot a_{21} \overset{!}{>} 0$$
$$\Leftrightarrow \left(\frac{a_{11} - a_{22}}{2}\right)^2 \overset{!}{>} -a_{12} \cdot a_{21}. \tag{50}$$

This inequality is satisfied if A has strictly positive entries, meaning $A \in \mathbb{R}_+^{n \times n}$, which is a reasonable assumption.

Furthermore, we need to discuss the behavior of the eigenvalues and eigenvectors under the assumption of the border case of non-stationarity. Under the assumed relations of Equation (39), we now consider the border case $a_{11} + a_{12} \to 1$ and $a_{21} + a_{22} \to 1$.

Lemma 1 (Eigenvalues and eigenvectors in border case of non-stationarity)**.** *The eigenvalues for $A \in \mathbb{R}_+^{2 \times 2}$, in the border case of non-stationarity converge to*

$$\lambda_1 \to 1,$$

$$\lambda_2 \to a_{11} + a_{22} - 1.$$

The auxiliary variables for the eigenvectors tend to

$$B \to \frac{1 - a_{11}}{a_{12}} = \frac{a_{12}}{a_{12}} = 1,$$

$$C \to -\frac{a_{12}}{a_{21}}.$$

Proof. Putting $a_{11} + a_{12} \to 1$ and $a_{21} + a_{22} \to 1$ in Equation (49) immediately yields:

$$\begin{aligned}
\lambda_{1/2} &= \frac{a_{11} + a_{22}}{2} \pm \sqrt{\left(\frac{a_{11} - a_{22}}{2}\right)^2 + a_{12} \cdot a_{21}} \\
&\to \frac{a_{11} + a_{22}}{2} \pm \sqrt{\left(\frac{a_{11} - a_{22}}{2}\right)^2 + (1 - a_{11}) \cdot (1 - a_{22})} \\
&= \frac{a_{11} + a_{22}}{2} \pm \sqrt{\left[1 - \left(\frac{a_{11} + a_{22}}{2}\right)\right]^2} \\
\Rightarrow \lambda_1 &\to \frac{a_{11} + a_{22}}{2} + \left[1 - \left(\frac{a_{11} + a_{22}}{2}\right)\right] = 1; \\
\lambda_2 &\to \frac{a_{11} + a_{22}}{2} - \left[1 - \left(\frac{a_{11} + a_{22}}{2}\right)\right] = a_{11} + a_{22} - 1.
\end{aligned}$$

Setting these results in Equation (44) proves the equations for B and C. □

B Exemplary Calculation of Stationary Information Shares

We refer to the setting in Section 4 and first demonstrate the structure of the A_i, $i = 1, \ldots, 5$ by an example: The coefficients $\frac{2^{5-i}}{2^5-1}$ are $\frac{16}{31}$, $\frac{8}{31}$, $\frac{4}{31}$, $\frac{2}{31}$, and $\frac{1}{31}$, and add up to 1. Using minor diagonal elements $a_{12} = a_{21} = 0.1$ and a convergence parameter $\Delta = 0.05$, the matrices read

$$\begin{aligned}
A_1 &= \begin{pmatrix} 0.486 & 0.02 \\ 0.02 & 0.486 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0.228 & 0.02 \\ 0.02 & 0.228 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0.099 & 0.02 \\ 0.02 & 0.099 \end{pmatrix}, \\
A_4 &= \begin{pmatrix} 0.035 & 0.02 \\ 0.02 & 0.035 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 0.002 & 0.02 \\ 0.02 & 0.002 \end{pmatrix}.
\end{aligned} \tag{51}$$

We assume a covariance matrix with $\sigma_1 = 0.2$, $\sigma_2 = 0.1$, and no cross-correlation. To calculate the stationary information share depicted in Section 3.2, we have to transform the constructed VAR(5) process into a moving average process according to Equation (9), which can be done by the method of recursion, obtaining the moving average matrices Ψ_p.[8] In the case of no cross-correlation, the Cholesky matrix is simply the root of the covariance matrix, hence the Φ_p matrices are obtained by multiplying the columns of Ψ_p with the respective standard deviations σ_j.[9] Following further the procedure in Section 3.2, we determine the contributions of the respective market's variances by Equation (15), which requires calculating the sum of the squared entries of the Φ_p matrices:

$$\eta_{12} = \frac{0.008}{0.128 + 0.008} = 0.059, \qquad \eta_{21} = \frac{0.032}{0.032 + 0.032} = 0.500. \tag{52}$$

So, for the stationary information share, we finally obtain:

$$SIS_1 = \frac{0.500}{0.059 + 0.500} = 0.894, \qquad SIS_2 = \frac{0.059}{0.059 + 0.500} = 0.106. \tag{53}$$

As we expected, information share of the first market is larger than of the second as the variance in the error terms is positively related to the information share.

To calculate the information share in the non-stationary case as a reference value, we note

---

[8]The recursion is simply defined as $\Psi_p = \sum_{i=1}^{k} A_i \cdot \Psi_{p-i}$ for $p \geq 1$ and Ψ_0 is set as unit matrix. Note that even for a VAR(1) process, the number of Ψ_p matrices is infinite. The first matrices read:

$$\Psi_0 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Psi_1 = \begin{pmatrix} 0.486 & 0.020 \\ 0.020 & 0.486 \end{pmatrix}, \quad \Psi_2 = \begin{pmatrix} 0.237 & 0.019 \\ 0.019 & 0.237 \end{pmatrix}, \quad \Psi_3 = \begin{pmatrix} 0.465 & 0.039 \\ 0.039 & 0.465 \end{pmatrix}.$$

[9]For the first Φ_p matrices we obtain

$$\Phi_0 = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.1 \end{pmatrix}, \quad \Phi_1 = \begin{pmatrix} 0.097 & 0.002 \\ 0.004 & 0.049 \end{pmatrix}, \quad \Phi_2 = \begin{pmatrix} 0.047 & 0.002 \\ 0.004 & 0.024 \end{pmatrix}, \quad \Phi_3 = \begin{pmatrix} 0.093 & 0.004 \\ 0.008 & 0.047 \end{pmatrix}.$$

that the VECM representation (2) can be derived from the VAR(k) process as[10]

$$\Delta y_t = \left(\sum_{i=1}^{k}(A_i) - I\right) \cdot y_{t-1} - \sum_{i=2}^{k-1}\sum_{j=i}^{k} A_j \cdot \Delta y_{t-i+1} + \epsilon_t. \tag{54}$$

A comparison with (2), where the long-run impact is contained in the vectors α and β, yields:

$$\sum_{i=1}^{k}(A_i) - I = \alpha\beta' = \begin{pmatrix} \alpha_1 & -\alpha_1 \\ \alpha_2 & -\alpha_2 \end{pmatrix}, \tag{55}$$

since $\beta = (1; -1)'$. The common row vector in the Stock-Watson decomposition (5) is defined as $\psi = \alpha_\perp^{'} = (\alpha_2, -\alpha_1)$, and is thus given by the sum over the minor diagonal elements of the matrices $A_i^{(k)}$. Hence, by construction, $\psi = (a_{21}, a_{12})$ throughout our numerical examples (with $\Delta = 0$).

With these results we obtain by Equation (7):

$$IS_1 = \frac{\psi_1^2\sigma_1^2}{\psi_1^2\sigma_1^2 + \psi_2^2\sigma_2^2} = \frac{(0.1 \cdot 0.2)^2}{(0.1 \cdot 0.2)^2 + (0.1 \cdot 0.1)^2} = 0.8, \qquad IS_2 = 1 - IS_1 = 0.2. \tag{56}$$

[10]To see this, note that:

$$\begin{aligned}
\Delta y_t &= (A_1 - I) \cdot y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} \ldots + A_k y_{t-k} + \epsilon_t \\
&= (A_1 - I) \cdot y_{t-1} + A_2 y_{t-1} - A_2 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} \ldots + A_k y_{t-k} + \epsilon_t \\
&= (A_1 + A_2 - I) \cdot y_{t-1} - A_2(y_{t-1} - y_{t-2}) + A_3 y_{t-3} \ldots + A_k y_{t-k} + \epsilon_t \\
&= (A_1 + A_2 - I) \cdot y_{t-1} - A_2(y_{t-1} - y_{t-2}) + A_3 y_{t-2} - A_3 y_{t-2} + A_3 y_{t-3} \ldots + A_k y_{t-k} + \epsilon_t \\
&= (A_1 + A_2 + A_3 - I) \cdot y_{t-1} - (A_2 + A_3)(y_{t-1} - y_{t-2}) - A_3(y_{t-2} - y_{t-3}) \ldots + A_k y_{t-k} + \epsilon_t \\
&= (A_1 + A_2 + A_3 - I) \cdot y_{t-1} - (A_2 + A_3) \cdot \Delta y_{t-1} - A_3 \cdot \Delta y_{t-2} \ldots + A_k y_{t-k} + \epsilon_t \\
&\ldots \\
&= \left(\sum_{i=1}^{k}(A_i) - I\right) \cdot y_{t-1} - (A_2 + A_3 + \ldots + A_k) \cdot \Delta y_{t-1} \\
&\quad - (A_3 + A_4 + \ldots + A_k) \cdot \Delta y_{t-2} - \cdots - (A_{k-1} + A_k) \cdot \Delta y_{t-k+1} + \epsilon_t.
\end{aligned}$$

C Stationary Information Shares in the Multivariate Case

Hasbrouck's Information share concept can easily be extended to a multivariate n-market setting. In the simplest case without cross correlation between any markets, the relative contribution of market k with $k \in \{1 \dots n\}$ to the total variance and thus the k'th information shares is given by:

$$IS_k = \frac{([\psi\Omega]_k)^2}{\psi\Omega\psi'} = \frac{\psi_k^2\sigma_k^2}{\sum_{j=1}^n(\psi_j^2\sigma_j^2)}. \tag{57}$$

with the covariance matrix $\Omega \in \mathbb{R}^{n\times n}$ and $\psi \in \mathbb{R}^{n\times 1}$ representing the common row vector for the sum of the moving average matrices.[11]

In the general case with contemporaneous correlation between the innovations in the different markets, again the Cholesky decomposition $\Omega = MM'$ with a triangular matrix $M \in \mathbb{R}^{n\times n}$ is applied. Following Hasbrouck (2002) and Yan and Zivot (2010), all permutations of the n markets have to be considered to get the lower and upper bounds of the information share. With a given order and thus a given Cholesky matrix, the k'th information share is defined as:

$$IS_k = \frac{([\psi M]_k)^2}{\psi\Omega\psi'}. \tag{58}$$

The bounds are calculated by switching the order of the markets resulting in $n!$ values for the information share for one market k, of which the largest and the smallest define the upper and lower bound, respectively.

Regarding the generalization of the information share for stationary time series, the idea of variance decomposition can also be applied for a multivariate n-market system. In this n-market system we can break down the forecast error vector into n parts, corresponding

[11]Note that ψ is constructed as the $(n \times 1)$-dimensional subspace orthogonal to the error correction vector $\alpha \in \mathbb{R}^{n\times(n-1)}$ in Equation 2.

to the forecast error components of the respective markets. Analogous to Equation (13) in Section 3.2 we get:

$$\begin{aligned} FE_j(y_{t+\tau}) &= y_{j,t+\tau} - E_t[y_{j,t+\tau}] \\ &= \sum_{p=0}^{\tau-1}\sum_{k=1}^{n} \phi_{jk}^{(p)} \cdot w_{k,t+\tau-p}, \end{aligned}$$

where the index $j \in \{1, 2, \ldots, n\}$ represents the respective market and $\phi_{jk}^{(p)}$, $j,\ k \in \{1, 2, \ldots, n\}$, are the coefficients of the matrices $\Phi_p \in \mathbb{R}^{n\times n}$. τ remains the forecast horizon as in the two-market system. The forecast error component of the first market ($j = 1$) is thus influenced by itself ($k = 1$) and moreover by innovation terms of the second market, third market and so on ($k = 2, 3, \ldots, n$). Analogous to the bivariate market case, we consider the variance of the forecast error to be decomposed:

$$\begin{aligned} Var(FE_j(y_{t+\tau})) &= E[(y_{j,t+\tau} - E_t[y_{j,t+\tau}])^2] \\ &= \sum_{p=0}^{\tau-1}\sum_{k=1}^{n} \left(\phi_{jk}^{(p)}\right)^2. \end{aligned} \tag{59}$$

The relative contribution of market k to the variance of market j is given by

$$\eta_{jk} = \frac{\sum\limits_{p=0}^{\infty} \left(\phi_{jk}^{(p)}\right)^2}{\sum\limits_{p=0}^{\infty}\sum\limits_{l=1}^{n} \left(\phi_{jl}^{(p)}\right)^2}. \tag{60}$$

Based on these relative impacts we construct the stationary information share SIS_k of market k in a multivariate market system by considering all cross-influences of market k on the other $n-1$ markets relative to the aggregated cross-influences of all markets on each other:

$$SIS_k = \frac{\sum\limits_{j=1,j\neq k}^{n} \eta_{jk}}{\sum\limits_{l=1}^{n}\sum\limits_{j=1,j\neq k}^{n} \eta_{jl}}. \tag{61}$$

Note that this value depends on the Cholesky matrix M and thus on the order of the markets. To obtain the upper and lower bounds, as in the non-stationary case, all permutations have to be evaluated.

Convergence to Hasbrouck's information share can be shown analogously to the bivariate case.

(a) $a_{12} = a_{21} = 0.1$

(b) $a_{12} = 0.1$ *and* $a_{21} = 0.3$

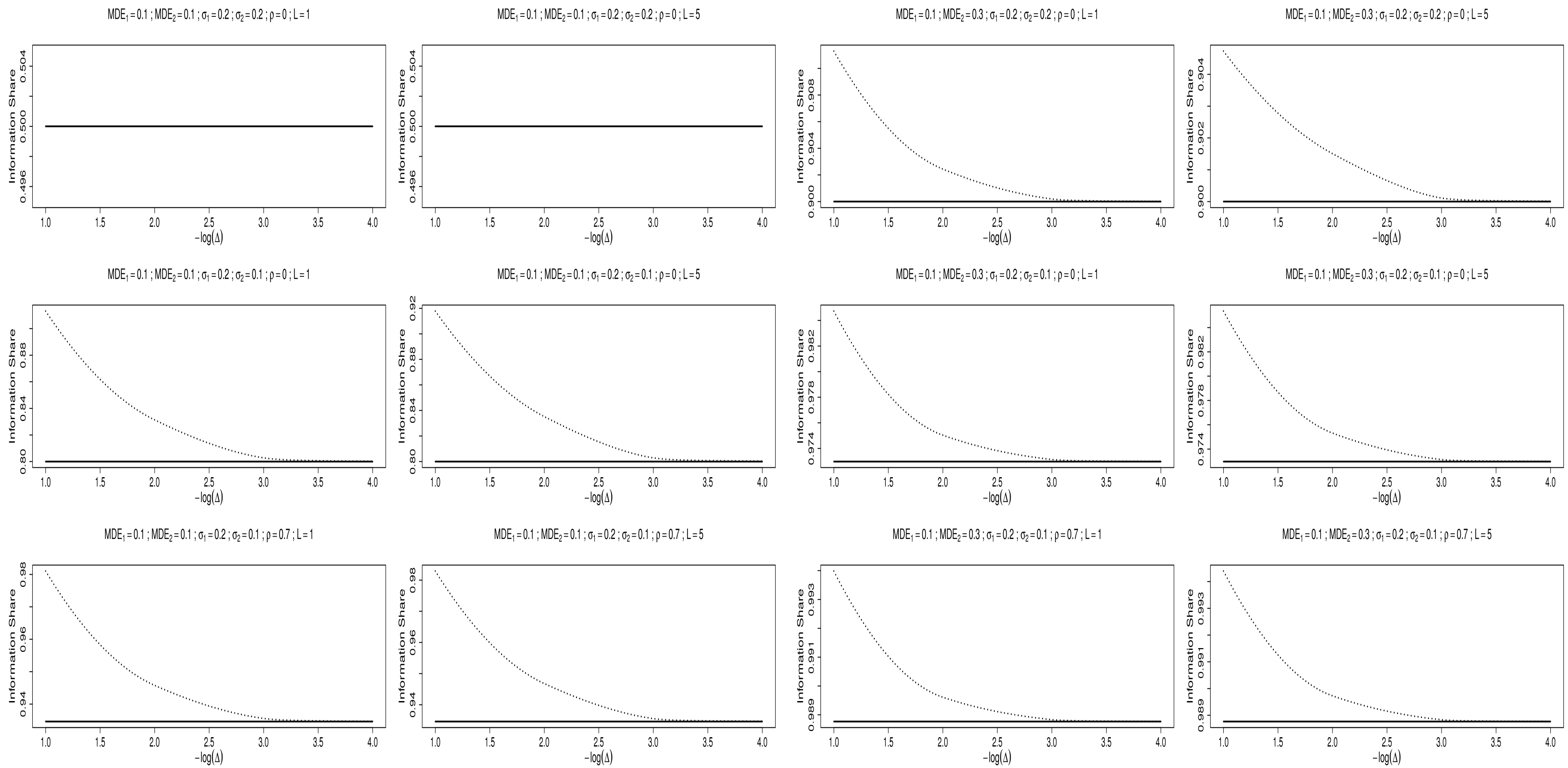

Figure 1. *Numerical convergence of SIS_1 to IS_1. By decreasing Δ, the dashed line shows the behavior of SIS_1 when the process converges to the non-stationary case. The underlying VAR process is constructed according to Equation (32) with minor diagonal elements (MDE) as in the header of the two subfigures. The left column of both subfigures shows the VAR(1) process, the right the VAR(5) process. The covariance matrix is changing by rows: In the first row, identical variances and no cross-correlation, $\Omega = \begin{pmatrix} 0.04 & 0 \\ 0 & 0.04 \end{pmatrix}$; in the second row different variances but still no cross-correlation, $\Omega = \begin{pmatrix} 0.04 & 0 \\ 0 & 0.01 \end{pmatrix}$; finally in the third row different variances and cross-correlation, $\Omega = \begin{pmatrix} 0.040 & 0.014 \\ 0.014 & 0.010 \end{pmatrix}$.*

(a) $a_{12} = 0.3$ *and* $a_{21} = 0.1$ **(b)** $a_{12} = a_{21} = 0.3$

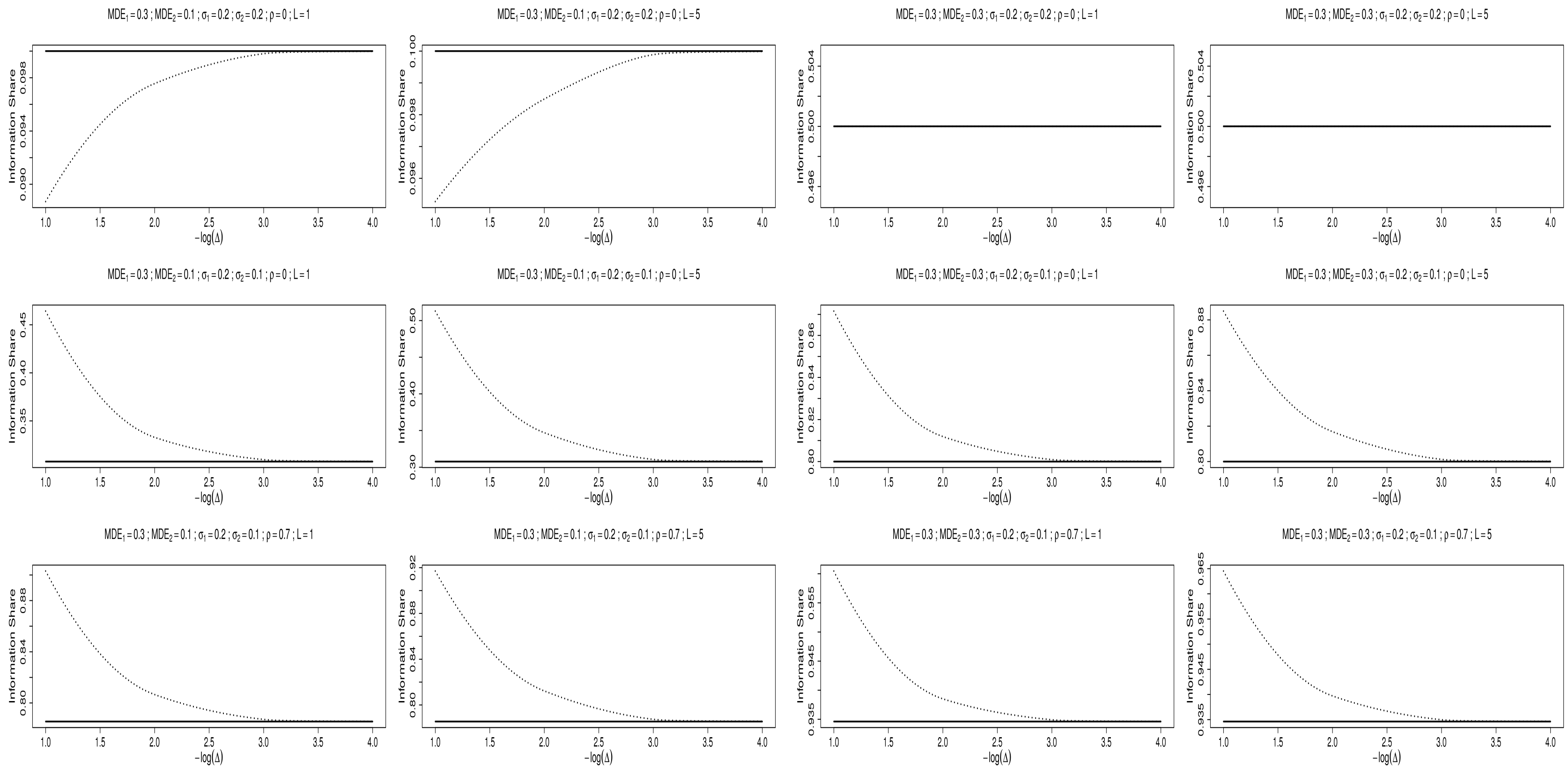

Figure 2. *Numerical convergence of SIS_1 to IS_1. By decreasing Δ, the dashed line shows the behavior of SIS_1 when the process converges to the non-stationary case. The underlying VAR process is constructed according to Equation (32) with minor diagonal elements (MDE) as in the header of the two subfigures. The left column of both subfigures shows the VAR(1) process, the right the VAR(5) process. The covariance matrix is changing by rows: In the first row, identical variances and no cross-correlation, $\Omega = \begin{pmatrix} 0.04 & 0 \\ 0 & 0.04 \end{pmatrix}$; in the second row different variances but still no cross-correlation, $\Omega = \begin{pmatrix} 0.04 & 0 \\ 0 & 0.01 \end{pmatrix}$; finally in the third row different variances and cross-correlation, $\Omega = \begin{pmatrix} 0.040 & 0.014 \\ 0.014 & 0.010 \end{pmatrix}$.*

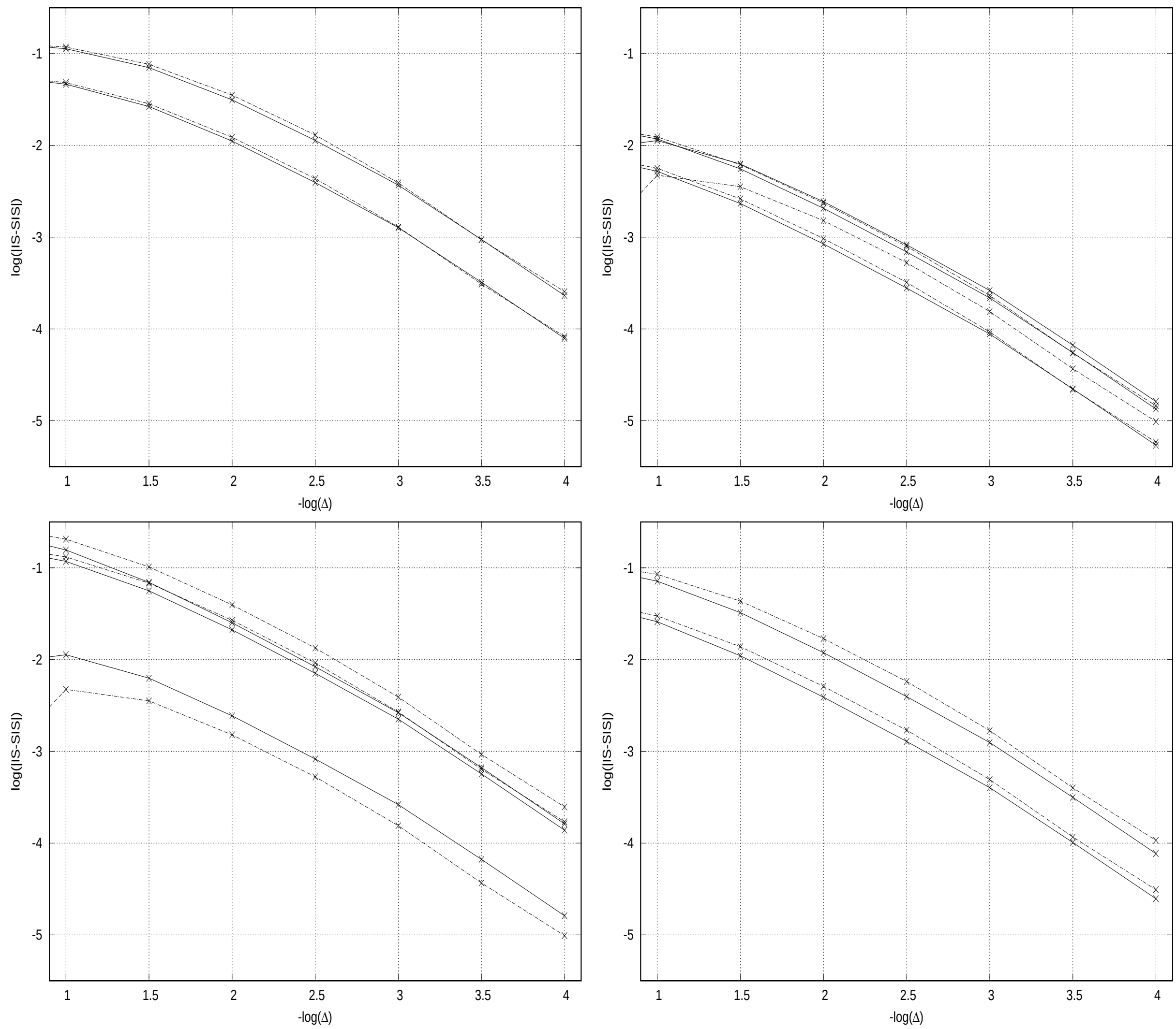

Figure 3. *Numerical convergence of SIS_1 to IS_1 on a log-log scale. The four graphs correspond to the four subfigures 1(a)–2(b): upper left graph with minor diagonal elements $a_{12} = a_{21} = 0.1$ (1a); upper right graph with minor diagonal elements $a_{12} = 0.1, a_{21} = 0.3$ (1b); lower left graph with minor diagonal elements $a_{12} = 0.3, a_{21} = 0.1$ (2a); lower right graph with minor diagonal elements $a_{12} = a_{21} = 0.3$ (2b). Solid lines represent $VAR(1)$ processes, dashed lines $VAR(5)$ processes. The slopes of the lines indicate at least linear convergence.*

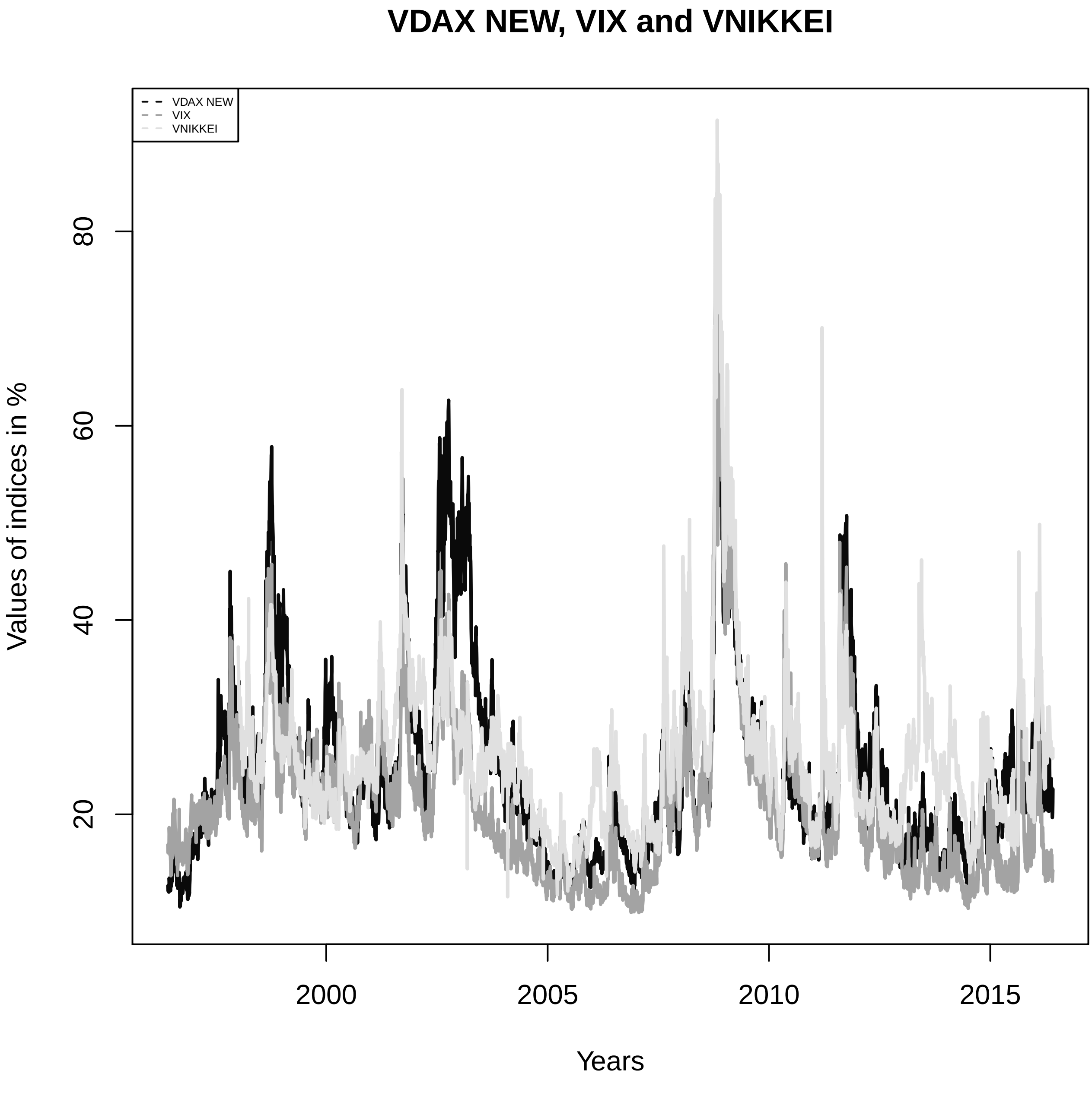

Figure 4. *VIX, VDAX-New, and NikkeiVI over the entire period 1996–2016.*

k	a_{12}	a_{21}	σ_1	σ_2	ρ
1	0.1	0.1	0.2	0.2	0.0
			0.2	0.1	0.0
			0.2	0.1	0.7
1	0.1	0.3	0.2	0.2	0.0
			0.2	0.1	0.0
			0.2	0.1	0.7
1	0.3	0.1	0.2	0.2	0.0
			0.2	0.1	0.0
			0.2	0.1	0.7
1	0.3	0.3	0.2	0.2	0.0
			0.2	0.1	0.0
			0.2	0.1	0.7

k	a_{12}	a_{21}	σ_1	σ_2	ρ
5	0.1	0.1	0.2	0.2	0.0
			0.2	0.1	0.0
			0.2	0.1	0.7
5	0.1	0.3	0.2	0.2	0.0
			0.2	0.1	0.0
			0.2	0.1	0.7
5	0.3	0.1	0.2	0.2	0.0
			0.2	0.1	0.0
			0.2	0.1	0.7
5	0.3	0.3	0.2	0.2	0.0
			0.2	0.1	0.0
			0.2	0.1	0.7

Table 1. *Description of the 24 different cases used to demonstrate the speed of convergence of stationary information shares SIS to Hasbrouck's information share IS. The left table considers a fixed lag length $k = 1$, and the four blocks represent different minor diagonal element combinations $a_{12}, a_{21} \in \{0.1, 0.3\}$. Within each block, the first row uses equal variances in the innovation terms, $\sigma_1 = \sigma_2 = 0.2$; the second row a lower variance of the second market ($\sigma_1 = 0.2$ and $\sigma_2 = 0.1$); and the third row additionally cross-correlation between the innovation terms ($\rho = 0.7$). The right table is built up analogously to the left one, except the lag length $k = 5$.*

	Distribution						Correlation			Dickey-Fuller		
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	VIX	VDAX	NikkeiVI	ADF	Lags	p-value
Entire period 1996/1998–2016												
VDAX-New	10.47	18.16	22.42	24.71	27.99	83.23	1	0.50	0.23	−4.657	23	0.01***
VIX	9.89	15.12	19.56	21.03	24.42	80.86	0.50	1	0.11	−4.509	23	0.01***
NikkeiVI	11.52	20.71	24.73	26.07	28.77	91.45	0.23	0.11	1	−4.781	22	0.01***
First subperiod 1996/1998–2007												
VDAX-New	10.47	17.60	22.72	24.92	28.69	62.63	1	0.46	0.13	−3.380	15	0.06*
VIX	9.89	15.22	20.06	20.57	24.40	45.74	0.46	1	0.07	−4.219	15	0.01***
NikkeiVI	11.52	20.04	24.37	24.74	27.89	63.73	0.13	0.07	1	−4.154	15	0.01***
Second subperiod 2007–2009												
VDAX-New	15.85	22.48	27.54	31.21	36.13	83.23	1	0.45	0.38	−1.890	9	0.62
VIX	15.23	22.72	25.71	31.11	37.15	80.86	0.45	1	0.08	−1.769	9	0.68
NikkeiVI	17.33	27.27	31.52	36.87	42.11	91.45	0.38	0.08	1	−1.816	9	0.66
Third subperiod 2009–2013												
VDAX-New	13.31	18.92	21.9	23.76	26.20	50.74	1	0.58	0.31	−2.549	10	0.35
VIX	11.3	16.81	19.22	21.16	23.70	48.00	0.58	1	0.17	−2.854	10	0.22
NikkeiVI	16.37	20.16	22.98	24.07	26.73	70.09	0.31	0.17	1	−4.270	10	0.01***
Fourth subperiod 2013–2016												
VDAX-New	12.17	16.77	19.57	20.53	23.17	38.29	1	0.52	0.26	−4.350	10	0.01***
VIX	10.32	13.18	14.38	15.51	16.61	40.74	0.52	1	0.15	−5.003	10	0.01***
NikkeiVI	14.00	20.49	24.24	24.79	27.74	49.84	0.26	0.15	1	−3.354	10	0.06*

Table 2. *Descriptive statistic of the closing prices of the volatility time series VDAX-New, VIX and NikkeiVI over the entire time period and over four subperiods. Minimum, Maximum, first-third quartile and mean of the respective indices are reported; furthermore bivariate correlations between the daily index returns. The final columns show results of Augmented Dickey-Fuller tests for stationarity (with constant and linear trend): the ADF test statistics, the number of estimated lags according to the suggested upper bound on the rate at which the number of lags should be made to grow with the sample size for the general ARMA(p,q) setup, and the corresponding p-values. Significance at the 10% level is denoted by *, at the 1% level by ***.*

	VDAX	VIX	VDAX	NikkeiVI	VIX	NikkeiVI
Entire period 1996/1998–2016						
SIS Mean	0.30	0.70	0.58	0.42	0.73*	0.27
SIS Upper Bound	0.59	0.98	0.69	0.53	0.81	0.35
SIS Lower Bound	0.02	0.41	0.47	0.31	0.65	0.19
First subperiod 1996/1998–2007						
SIS Mean	0.11	0.89*	0.39	0.61*	0.71*	0.29
SIS Upper Bound	0.19	0.97	0.45	0.67	0.77	0.34
SIS Lower Bound	0.03	0.81	0.33	0.55	0.66	0.23
Second subperiod 2007–2009						
SIS Mean	0.10	0.90*	0.69	0.31	0.95*	0.05
SIS Upper Bound	0.18	0.98	0.91	0.53	0.98	0.08
SIS Lower Bound	0.02	0.82	0.47	0.09	0.92	0.02
Third subperiod 2009–2013						
SIS Mean	0.55	0.45	0.78*	0.22	0.84*	0.16
SIS Upper Bound	0.89	0.79	0.92	0.36	0.93	0.25
SIS Lower Bound	0.21	0.11	0.64	0.08	0.75	0.07
Fourth subperiod 2013–2016						
SIS Mean	0.70	0.30	0.92*	0.08	0.81*	0.19
SIS Upper Bound	0.95	0.56	0.99	0.14	0.88	0.26
SIS Lower Bound	0.44	0.05	0.86	0.01	0.74	0.12

Table 3. *Bivariate stationary information shares over the entire period and the four subperiods for the combinations of VDAX-New & VIX, VDAX-New & NikkeiVI and VIX & NikkeiVI. The rows represent upper and lower bounds and mean values of the respective SIS. Information shares are declared to be significant (*) if the upper-lower bound range of one volatility index is not touched by the respective upper-lower bound range of the second volatility index. For the entire period and for the four subperiods, VIX is leading NikkeiVI significantly. For the first and second subperiods, VIX also has a significant higher information share compared with VDAX-New. VDAX-New is gaining information share after the global financial crisis and tends to lead the VIX. The relation between VDAX-New and NikkeiVI is also changing: While for the first period NikkeiVI is leading the VDAX-New significantly, VDAX-New is taking leadership in and after the European sovereign debt crisis in the third and fourth subperiod.*