Journal of Financial Econometrics, 2021, Vol. 19, No. 5, 934–959
doi: 10.1093/jjfinec/nbz035
Advance Access Publication Date: 12 November 2019
Article

Bootstrap Confidence Intervals and Hypothesis Testing for Market Information Shares*

Karsten Schweikert
University of Hohenheim

Address correspondence to Karsten Schweikert, Core Facility Hohenheim & Institute of Economics, University of Hohenheim, Schloss Hohenheim 1 C, 70593 Stuttgart, Germany, or e-mail: karsten.schweikert@uni-hohenheim.de.

Received September 7, 2018; revised October 4, 2019; editorial decision 11 October 2019; accepted October 16, 2019

Abstract

Market information shares are widely used in empirical finance to measure one market's contributions to price discovery. In contrast to common factor components, the literature on market information shares only provides rudimentary tools to test general hypotheses. Using Monte Carlo simulations, we show that bootstrap confidence bands proposed by Sapp (2002) perform well if markets have similar information shares but are too narrow if one market dominates price discovery. We design a new bootstrap-based method to test the "one-central-market" hypothesis and show that our tests have correct size and substantial power against the null hypothesis. Empirical results in the context of CDS and bonds markets complement the theoretical analysis.

Key words: bootstrap inference, cointegration, Monte Carlo simulation, price discovery

JEL classification: C12, C15, G14

Price discovery is one of the main functions of financial markets. It is of particular interest to determine which market impounds information first if an asset is traded on several exchanges. Multiple competing measures to estimate the contribution of a market to price discovery have been proposed in the literature. Booth, So, and Tse (1999) and Harris, McInish, and Wood (2002a) use common factor components estimated from the Gonzalo

* I thank Thomas Dimpfl, Robert Jung, Niraj KC, Konstantin Kuck, Alexander Schmidt, Markus Mößler and the editor, Allan Timmermann, an assistant editor as well as two anonymous referees for valuable comments and suggestions. Further, I thank Yang Liu for excellent research assistance. Access to Thomson Reuters Datastream, provided by the Hohenheim Datalab (DALAHO), is gratefully acknowledged.

© The Author(s) 2019. Published by Oxford University Press. All rights reserved.
For permissions, please email: journals.permissions@oup.com

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

and Granger (1995) permanent-transitory decomposition (PT/GG). In contrast, Hasbrouck (1995) constructs information shares based on the common trend representation by Stock and Watson (1988) incorporating both the system dynamics and the innovation variances. While the methodology is intuitive and easy to compute, it employs a Cholesky decomposition of the reduced-form innovation covariance matrix and therefore depends on the ordering of price series in the model. In practice, the midpoint between minimum and maximum information share estimates obtained from all possible permutations of orderings is computed to provide a representative information share value (Baillie et al., 2002; Carbera, Wang, and Yang, 2009; Fricke and Menkhoff, 2011; Dimpfl, Flad, and Jung, 2017). In the remainder of this article, we refer to this midpoint as a Hasbrouck information share (HIS). Lien and Shrestha (2009) propose a modified information share (MIS) which improves over the HISs by being order independent and Lien and Shrestha (2014) extend their methodology to interrelated markets where prices are cointegrated but the cointegrating vector does not have to be one-to-one. Putniņš (2013) provides a comprehensive discussion about the differences between common factor components and market information shares. Further important contributions are found in De Jong and Schotman (2010) where information share measures are based on the covariance between transitory components and the efficient price and in Grammig and Peter (2013) who identify tail-dependent information shares from a mixture distribution.

While it is straightforward to conduct inference on price discovery measures based on the PT/GG decomposition, where hypotheses can directly be stated as restrictions on the long-run multiplier, it is more difficult to test hypotheses in the context of market information shares like the HIS or MIS. Paruolo (1997) provides the necessary tools to conduct asymptotic inference on the long-run impact matrix and its row and column spaces in cointegrated VAR processes. These tools could in principle be used to design hypothesis tests for market information shares. However, these methods are quite complicated and are rarely applied in practice. Sapp (2002) proposes a stationary bootstrap algorithm to construct confidence intervals for the HIS. The algorithm is also used in Grammig, Melvin, and Schlag (2005) to assess the estimation precision of information shares in the context of internationally cross-listed stocks and in Grammig and Peter (2013) to compare the performance of HIS against tail-dependent information shares. Since the Sapp algorithm has not been evaluated in a Monte Carlo study, practitioners might be reluctant to use it. Still, the need for a reliable assessment of precision of information share estimates has been clearly formulated in the literature. For example, Harris, McInish, and Wood (2002b) base their critique of market information shares on the fact that hypothesis tests can easily be conducted under the PT/GG framework but not for market information shares. They argue that findings of numerical differences between common factor components and information shares are not meaningful if the estimation uncertainty is not properly incorporated. A simulation study is conducted by the authors to show that numerical evidence obtained for common factor components against the "one-central-market" hypothesis does not hold up if hypothesis tests are considered. Figuerola-Ferretti and Gonzalo (2010, p. 17) comment on bootstrap inference and argue that "it is always possible to use some bootstrap methods as in Sapp (2002) for testing single hypotheses (for instance $S_1 = 0$), but it is unclear how to proceed for testing joint hypotheses on different IS (for example $S_1 = S_2$)." While it is certainly difficult to test the "equal shares" null hypothesis in large systems, it amounts to a single hypothesis in two-variable systems the authors refer to. Furthermore, it is impossible

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

for the percentile method employed by Sapp (2002) to construct valid confidence bands for the hypothesis that one information share is zero because the algorithm produces strictly positive information shares. Neither of the authors' claims is supported by simulation evidence. Lien and Shrestha (2009) argue that the null hypothesis that one market does not contribute to price discovery can be tested if the innovations are uncorrelated. Because this situation is unlikely to occur in practice, hypothesis testing has to be based on more general assumptions.

This article addresses the lack of inferential tools for market information shares and proposes a bootstrap framework for specific hypotheses. Although it is difficult to test general hypotheses in settings with more than two markets, information shares are commonly applied to two variable systems (e.g., analysis of spot and futures markets or cross-listed stocks). From a practitioner's perspective, two types of null hypotheses are of interest where it is important to design valid statistical tests. First, we consider the null hypothesis of equal information shares for both markets, such as spot and futures for a particular underlying. It might then be of interest to determine which market contributes significantly more to price discovery than the other one. Formally, we test the "equal shares" null hypothesis $IS_{p_1} = IS_{p_2} = 0.5$. Second, we test the null hypothesis of "one-central-market" assuming that one market contributes 100% to price discovery whereas the other market has a zero information share. Formally, we test $IS_{p_1} = 0$ or $IS_{p_2} = 0$, depending on the hypothesized role of both prices. In principle, it is always possible to make use of the duality between confidence bands and statistical tests to evaluate hypotheses. One of these approaches should be chosen depending on which one has better statistical properties in a given situation. We evaluate both approaches in various situations and derive recommendations based on our simulation study.

In the following, we design a Monte Carlo study to investigate the properties of the Sapp algorithm for the "equal shares" and "one-central-market" hypotheses using different data generating processes (DGPs). To the author's knowledge, the statistical properties of bootstrap confidence intervals for information shares have not been systematically investigated yet. It is therefore unclear whether this algorithm yields the intended coverage and how the length of the confidence bands evolves with increasing sample sizes, that is, how powerful the implied hypothesis tests are when they are based on these confidence intervals. Particularly, it remains an open question how the procedure can be used to test the "one-central-market" null hypothesis as the percentile method employed by Sapp (2002) produces confidence bands with strictly positive information shares so that zero information shares always lie outside of the estimated confidence band. We therefore propose a studentized version of the Sapp algorithm to properly evaluate the "one-central-market" null hypothesis. We can show that the enhanced algorithm maintains good properties in case of "equal shares" but confidence bands fail to achieve their nominal coverage for zero information shares. Hence, we additionally design a resampling algorithm to draw data under the null restriction and construct a valid bootstrap test with data-specific critical values. We show that this test has the correct empirical size under different data-generating processes and good power in small samples. Finally, we apply our new bootstrap tests to an empirical illustration presented in Lien and Shrestha (2014) where the contribution to price discovery of CDS and bond markets is investigated for a panel of actively traded U.S. companies using generalized information shares (GISs). We show that our methodology allows us to test the

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

hypotheses that price discovery takes place only in one market which yields different results from corresponding tests in the PT/GG framework.

The article is organized as follows. Section 1 introduces the reduced-form vector error correction model, both information share measures and our bootstrap methods. Section 2 is devoted to the Monte Carlo simulation study. Section 3 reports the results of an empirical application in the context of CDS and bond markets, and Section 4 concludes.

1 Methodology

In this section, we first introduce both information share measures and the underlying vector error correction model. Second, the procedure to compute bootstrap confidence intervals proposed by Sapp (2002) is described. Finally, we present a new bootstrap algorithm to evaluate the "one-central-market" hypothesis.

1.1 Vector Error Correction Model and Market Information Shares

Following Johansen (1988, 1991)'s notation, the linear VECM for an $N \times 1$ vector of $I(1)$ price variables is given as

$$\Delta p_t = \Pi p_{t-1} + \sum_{i=1}^{K-1} \Gamma_i \Delta p_{t-i} + u_t, \tag{1}$$

where u_t is a vector of i.i.d. Gaussian error terms. The $N \times N$ parameter matrix $\Pi = \alpha\beta'$ captures both the long-run equilibrium relations and the adjustment behavior. The matrix β contains r cointegrating vectors and α carries the loadings on each cointegrating vector. The covariance matrix of the error terms is given by $E(u_t u_t') = \Omega$ and expresses the contemporaneous linear dependencies for each variable with the other variables.

As shown in Stock and Watson (1988), the cointegrated system in Equation (1) can equivalently be written in a common trend representation as follows

$$p_t = p_0 + \Psi(1) \sum_{i=1}^{t} u_i + \Psi^*(L) u_t, \tag{2}$$

where $\Psi(L)$ and $\Psi^*(L)$ are matrix polynomials in the lag operator L derived from the moving average representation of the process and p_0 are initial values. The long-run impact matrix $\Psi(1)$ denotes the sum of the moving average coefficients and $\Psi^*(L) = (\Psi(L) - \Psi(1))/(1 - L))$ denotes transitory components. For each period, we can express the impact of an innovation on each of the prices using $\Psi(1)u_t$. Since the price series are cointegrated, the conditions $\beta'\Psi(1) = 0$ and $\Psi(1)\alpha = 0$ hold and we can write

$$p_t = p_0 + \beta_\perp \left(\alpha_\perp' \Gamma \beta_\perp\right)^{-1} \alpha_\perp' \sum_{i=1}^{t} u_i + \Psi^*(L) u_t, \tag{3}$$

where $\Gamma = I_N - \sum_{i=1}^{K-1} \Gamma_i$, I_N is the $N \times N$ identity matrix and $\perp$ denotes the orthogonal complement. The common stochastic trend shared by all price variables is then given by $\alpha_\perp' \sum_{i=1}^{t} u_i$ which has serially uncorrelated innovations $\alpha_\perp' u_t$ by construction.

We restrict our analysis to N-variable cointegrated systems with $r = N - 1$ cointegrating vectors and one common stochastic trend. This means that the cointegrated systems are driven by one common source of information. Following the empirical price discovery

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

literature, we perceive the common stochastic trend as the latent efficient price of one asset which is traded on different exchanges. In these cases, the matrix $\Psi(1)$ has rank one and identical rows $\psi = (\psi_1, \psi_2, \ldots, \psi_n)$. The variance of the common stochastic trend is $\psi\Omega\psi'$. The method proposed by Hasbrouck (1995) first allocates the contemporaneous correlation between innovations to one market using the Cholesky factorization $\Omega = FF'$. Then, the HIS is computed as

$$S_j = \frac{[\psi F]_j^2}{\psi\Omega\psi'}, \tag{4}$$

for each market j. Since the ordering of markets has a substantial effect on the outcome (except for uncorrelated innovations), all possible permutations of the ordering had to be considered. Baillie et al. (2002) show that the upper bound in the sequence of HIS estimates incorporates the series' own contribution and its correlation with the other series. The corresponding lower bound only considers the series' contribution that is uncorrelated with the other series. In empirical studies, the midpoint between lower and upper bound is usually taken as the HIS estimate to solve the uniqueness problem. Although lower and upper bounds of the HIS have a clear economic interpretation, the discrepancies between orderings become large when the contemporaneous correlation of disturbances across markets increases.

An alternative measure proposed by Lien and Shrestha (2009) yields an order-independent information share. Their factorization is based on the innovation correlation matrix $\Phi = V^{-1}\Omega V^{-1}$ where $V = \mathrm{diag}\left(\sqrt{\Omega_{11}}, \sqrt{\Omega_{22}}, \ldots, \sqrt{\Omega_{NN}}\right)$. Furthermore, the diagonal matrix Λ is defined with diagonal elements equal to the eigenvalues of the correlation matrix Φ and the regular matrix G where the corresponding eigenvectors are stacked column wise. The factorization for this information share measure is $\Omega = F^*(F^*)'$ where $F^* = [G\Lambda^{-1/2}G'V^{-1}]^{-1}$ and the MIS is then given by

$$S_j^* = \frac{[\psi F^*]_j^2}{\psi\Omega\psi'}. \tag{5}$$

The methodology is extended by Lien and Shrestha (2014) to include cointegrated systems where the relationship is not necessarily one-to-one. The GIS incorporates the estimated cointegrating vector for the computation of the respective measures. In contrast to the lower and upper bounds of the HIS and similar to the HIS midpoint, the MIS (GIS) is lacking economic rationale and mainly has a statistical interpretation. Lien and Wang (2016) show, using a Monte Carlo study, that the MIS performs similarly to the HIS measure based upon the upper/lower bound midpoint.

Finally, we briefly outline the PT/GG framework under which common factor components can be obtained. Gonzalo and Granger (1995) propose an alternative decomposition of p_t into permanent and transitory components, where

$$p_t = A_1 f_t + A_2 z_t, \tag{6}$$

$f_t = \alpha_\perp' p_t$ is the permanent component and $z_t = \beta' p_t$ is the transitory component. The loading matrices are given by $A_1 = \beta_\perp \left(\alpha_\perp' \beta_\perp\right)^{-1}$ and $A_2 = \alpha\left(\beta'\alpha\right)^{-1}$. This definition of the permanent component is different from the Stock–Watson definition because the changes in f_t can be serially correlated (De Jong, 2002). Harris, McInish, and Wood (2002a) suggest to

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

compute common factor components from elements of $\alpha_\perp$, that is, from the normalized orthogonal to the adjustment coefficient matrix. Since the computation of common factor components is only based on the adjustment dynamics of the cointegrated system, it does not take the innovation variances into account. Whereas market information shares measure the contribution of an innovation in market i to the total variance of the innovation in the permanent component, common factor components measure the impact of an innovation in market i on the innovation in the permanent component. One advantage of common factor components over the HIS is its unique determination of a market's contribution to price discovery.

1.2 Bootstrap Confidence Intervals

The bootstrap procedure proposed in Sapp (2002) aims to closely mimic the data so that confidence intervals can be centered around the estimated HIS midpoints through successive reestimation of the model. A stationary bootstrap according to Politis and Romano (1994) is used for this purpose which allows to maintain some degree of autocorrelation in the residuals. This is of particular importance for empirical applications where the underlying theoretical model cannot be accurately captured by reduced-from models. The author resamples blocks of data where the length of each block is determined from a geometric distribution and proposes to use the percentile method to construct confidence bands around the estimated information shares. Conceptually, the algorithm employed by Sapp (2002) has some drawbacks as it does not account for the fact that the range of information shares is bounded on the interval (0, 1). The "one-central-market" hypothesis leads to information shares of $S_1 = 0$ and $S_2 = 1$ or vice versa. Hence, the true information shares could only be placed within the confidence band if the lower bound of the confidence band takes on negative values which is theoretically impossible. It is, however, possible to obtain negative lower bounds of the confidence bands after studentizing the information shares. Therefore, we introduce an updated studentized bootstrap confidence band which in principle allows to evaluate the performance under the "one-central-market" hypothesis. The algorithm is outlined in the following for a $N = 2$ variable cointegrated system[1]:

1. Estimate the unrestricted VECM.
2. Compute the information share $\hat{S}_j$ according to Equation (4).
3. Using the estimated sets of residuals $\hat{u}_t$, create a pseudo time series set of bootstrap residuals, u_t^b, by performing a stationary bootstrap as follows:
   i. Randomly, sample one set of residuals $\{\hat{u}_{1j}, \hat{u}_{2j}\}$, $j \in \{1, \ldots, T\}$.
   ii. With probability q concatenate the set of residuals immediately following those obtained in Step i to the bootstrap residual sample, so that $\{\hat{u}_{1j+1}, \hat{u}_{2j+1}\}$ is drawn. With probability 1 – q, Step i is repeated. Hence, a new set of residuals is drawn, $\{\hat{u}_{1k}, \hat{u}_{2k}\}$, $k \in \{1, \ldots, T\}$.
   iii. Repeat until T sets of residuals have been drawn.
4. Construct a bootstrap series, p_t^b, by recursively inserting the bootstrapped residuals into the estimated VECM where the initial p differences are given by $\Delta p_j^b = \left\{u_{1j}^b, u_{2j}^b\right\}$, $j \in \{1, \ldots, p\}$ and the initial p values of p_t^b are given by $p_j^b = p_j$, $j \in \{1, \ldots, p\}$.

1 Sapp (2002) describes the algorithm for $N = 5$ variable system which does not pose any additional challenges.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

5. Use the bootstrap series, p_t^b, to re-estimate the VECM and compute the corresponding bootstrap information share S_j^b.
6. Repeat Steps 3–5 sufficiently often to obtain a bootstrap sample of information shares.
7. Determine the α-th quantile (c_α) and the $1-\alpha$-th quantile ($c_{1-\alpha}$) of the distribution of the studentized bootstrap statistic $(S_j^b - \hat{S}_j)$ to construct the confidence band $[CI_{lb}; CI_{ub}] = [\hat{S}_j + c_\alpha; \hat{S}_j + c_{1-\alpha}]$.

While the algorithm was originally designed for the HIS, we exploit the fact that the stationary bootstrap is a parametric algorithm based on the reduced-form VECM and analogously use it for the MIS. Since our simulation study in Section 2 reveals that bootstrap confidence bands should not be applied to test the "one-central-market" hypothesis, we develop a specific bootstrap test for these cases.

1.3 Bootstrap Tests of the "One-Central-Market" Hypothesis

The following discussion of a new bootstrap test for the "one-central-market" hypothesis is limited to $N=2$ variable systems but can straightforwardly be extended to multiple variable systems. For two variables, the covariance matrix Ω takes the form of

$$\Omega = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}, \tag{7}$$

where σ_1 (σ_2) is the standard deviation of u_{1t} (u_{2t}) and ρ is the correlation between u_{1t} and u_{1t}. It is easily shown, for example, in Baillie et al. (2002), that the Cholesky decomposition of Ω used to compute the HIS yields the factor

$$F = \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\left(1-\rho^2\right)^{1/2} \end{bmatrix}. \tag{8}$$

The first series is the central contributor of new information and the second series' information share is zero if the numerator of S_2 is zero. Hence, it has to hold that

$$[\psi F]_j = \begin{bmatrix} \psi_1 & \psi_2 \end{bmatrix} \begin{bmatrix} 0 \\ \sigma_2\left(1-\rho^2\right)^{1/2} \end{bmatrix} = \psi_2\sigma_2\left(1-\rho^2\right)^{1/2} = 0. \tag{9}$$

Because σ_2 and $\left(1-\rho^2\right)^{1/2}$ cannot be zero, we have to set $\psi_2 = 0$ to generate data under the null hypothesis. This can be achieved if the first series is strongly exogenous, that is, does not adjust to the long-run equilibrium and does not react to the second series in the short run. If, however, the first market has zero information share, the condition

$$\begin{bmatrix} \psi_1 & \psi_2 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \rho\sigma_2 \end{bmatrix} = \psi_1\sigma_1 + \psi_2\rho\sigma_2 = 0 \tag{10}$$

has to be satisfied. In this case, we have to draw data under the assumption that the second series is strongly exogenous, $\psi_1 = 0$, which directly implies $\psi_2 \neq 0$ so that we have to exclude the possibility of contemporaneous correlation, that is, $\rho = 0$.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

In comparison, the MIS is based on a decomposition of Ω which employs the factor,

$$F^* = \begin{bmatrix} 0.5(\sqrt{1+\rho}+\sqrt{1-\rho})\sigma_1 & 0.5(\sqrt{1+\rho}-\sqrt{1-\rho})\sigma_1 \\ 0.5(\sqrt{1+\rho}-\sqrt{1-\rho})\sigma_2 & 0.5(\sqrt{1+\rho}+\sqrt{1-\rho})\sigma_2 \end{bmatrix}. \tag{11}$$

The first series has full information share and the second series' information share is zero if the condition

$$\begin{aligned} &\begin{bmatrix} \psi_1 & \psi_2 \end{bmatrix} \begin{bmatrix} 0.5(\sqrt{1+\rho}-\sqrt{1-\rho})\sigma_1 \\ 0.5(\sqrt{1+\rho}+\sqrt{1-\rho})\sigma_2 \end{bmatrix} \\ &= 0.5\psi_1(\sqrt{1+\rho}-\sqrt{1-\rho})\sigma_1 + 0.5\psi_2(\sqrt{1+\rho}+\sqrt{1-\rho})\sigma_2 = 0 \end{aligned} \tag{12}$$

holds. To achieve this, we have to draw data under the restrictions $\psi_2 = 0$ and $\rho = 0$. Since the MIS is order invariant, the complementary condition of zero information share for the first series is straightforwardly given by $\psi_1 = 0$ and $\rho = 0$. The "one-central-market" hypothesis for MIS always requires to destroy any contemporaneous correlation between u_{1t} and u_{2t}.

Instead of constructing skewed confidence bands around the lower or upper boundary of the information shares, we draw data under the "one-central-market" null hypothesis and obtain critical values from the bootstrap distribution of our test statistics. Since we now use the derivations above to impose restrictions on the reduced-form VECM coefficient matrices, we do not encounter the problem that the "one-central-market" hypothesis is on the boundary of the parameter space. The restricted bootstrap algorithm is given in the following for the HIS and the null hypothesis that the second market has full information share:

1. Estimate the unrestricted two-variable VECM.
2. Compute the information share $\hat{S}_j$ according to Equation (4).
3. Reestimate the VECM under the restriction of p_2 being strongly exogenous,

$$\begin{bmatrix} \Delta p_{1t} \\ \Delta p_{2t} \end{bmatrix} = \begin{bmatrix} \tilde{\alpha}_1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & \beta \end{bmatrix} \begin{bmatrix} p_{1t-1} \\ p_{2t-1} \end{bmatrix} + \begin{bmatrix} \tilde{\gamma}_{11} & \tilde{\gamma}_{12} \\ 0 & \tilde{\gamma}_{22} \end{bmatrix} \begin{bmatrix} \Delta p_{1t-1} \\ \Delta p_{2t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix},$$

   and save the residual vector $\hat{e}_t$.
4. Generate a bootstrap residual series e_t^b using residuals from the restricted model. Resample each component of the residual vector individually to destroy any contemporaneous correlation:
   i. Randomly sample one residual element $\hat{e}_{jk}$, $j \in \{1,2\}$, $k \in \{1,\dots,T\}$.
   ii. With probability q, the adjacent residual to the one obtained in Step i is concatenated to the bootstrap residual sample, so that $\hat{e}_{jk+1}$ is drawn. With probability $1-q$, Step i is repeated. Hence, a new element of the residual vector is drawn, $\hat{e}_{jl}$, $j \in \{1,2\}$, $l \in \{1,\dots,T\}$.
   iii. Repeat until T observations of each residual component have been drawn.
5. Generate a bootstrap series p_t^b, by recursively inserting the bootstrapped residuals e_t^b into the estimated constrained model of Step 3. The initial p differences are given by

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

$\Delta p_j^b = \{\hat{e}_{1j}, \hat{e}_{2j}\}, j \in \{1, \ldots, p\}$ and the initial p-values of p_t^b are given by $p_j^b = p_j, j \in \{1, \ldots, p\}$.

6. Reestimate the unconstrained VECM for the bootstrap series p_t^b and compute the corresponding bootstrap information share S_j^b.
7. Repeating Steps 1–6 sufficiently often yields a distribution of bootstrap information shares. Compute the $1 - \alpha$-th quantile of the bootstrap distribution.
8. If the information share is greater than the bootstrap critical value, reject the "one-central-market" hypothesis at the α% significance level.

Since our algorithm only involves OLS estimations, we are confronted with low computational costs and can draw many bootstrap replications for an arbitrary precision of the p-value.

Note that the strong exogeneity constraints we put on p_2 might create problems if the test is applied to situation where p_1 has an information share close to one. Values close to one could be driven by the strong exogeneity of the respective variables. Setting the adjustment dynamics of p_2 to zero might result in very little adjustment after equilibrium errors and could, therefore, destroy the cointegrating relationship. Consequently, we recommend that the test should be applied with caution if the hypothesized dominant market has empirical information share of less than 50%.

A simple alternative to the bootstrap algorithm above is a test for strong exogeneity of one variable corresponding to the hypothesized dominant market, say p_1, under the assumption of contemporaneously uncorrelated reduced-form innovations. This test can be designed as a variable omission test in the equation of p_1, that is, omission of the error correction term and all lags of p_2. Since Wald tests of restrictions on the coefficients of cointegrated systems are known to have nonstandard asymptotic properties, it is advisable to use the test proposed in Dolado and Lütkepohl (1996) which adds an extra lag for inference.[2] However, as we show in the following section, this test can be oversized for small samples and should only be employed in large sample applications. The proposed bootstrap variant has the advantage that it evaluates the market information share measure directly and, hence, takes the nonlinear computation and compound error into account.

It is important to emphasize that the "one-central-market" hypothesis for market information shares is distinctly different from the null hypothesis of no contribution to price discovery in the PT/GG framework. Whereas the restrictions of strict exogeneity and uncorrelated innovations are imposed under the null hypothesis for market information shares, the likelihood ratio test outlined by Gonzalo and Granger (1995) requires only weak exogeneity under the null hypothesis. Consequently, using the likelihood ratio test for market information shares or the restricted bootstrap test for common factor components potentially leads to wrong conclusions.

2 Simulation Results

In our simulation experiments, we consider three DGPs proposed by De Jong (2002) which have a parametric design based on the reduced-form VECM. Hence, we know the exact lag length, true parameters, and structural information shares. Additionally, we also consider

2 We thank an anonymous referee for this suggestion.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

two economically motivated models discussed in Hasbrouck (2002), Lien and Shrestha (2009), and Lien and Wang (2016). These processes are based on classical microstructure models by Roll (1984) and do not have a straightforward reduced-form representation. In fact, the lag length is of infinite order and our reduced-form VECMs are necessarily misspecified. We still use them to evaluate how robust our parametric bootstrap procedure is to a misspecification of the lag structure. This could be relevant for practical applications as the true lag length is usually unknown and the lag structure has to be truncated using, for example, information criteria.

Our Parametric DGPs Are Based on the VECM

$$\begin{bmatrix} \Delta p_{1t} \\ \Delta p_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} (p_{1t-1} - p_{2t-1}) + \begin{bmatrix} -0.1 & 0 \\ 0 & -0.1 \end{bmatrix} \begin{bmatrix} \Delta p_{1t-1} \\ \Delta p_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}, \tag{13}$$

where u_{1t} and u_{2t} are i.i.d. processes with variances σ_1^2 and σ_2^2, respectively. The first configuration (DGP I.a) is given by $\alpha = (-0.5, 0.5)'$ and equal variances ($\sigma_1^2 = \sigma_2^2 = 1$) and yields a system where both markets have equal information shares. The second specification (DGP II) shifts more adjustment to the first variable, $\alpha = (-0.75, 0.25)'$, but keeps the variances ($\sigma_1^2 = \sigma_2^2 = 1$) unchanged. This results in a 90% information share of the second market. Third, we set $\alpha = (-1, 0)'$ to give the second market a 100% information share (DGP III.a). In principle, we could choose the innovation variances to be of equal size, since the ratio of innovation variances does not influence the measure if one price is strongly exogenous and innovations are uncorrelated. This is in stark contrast to DGP I.a where the ratio of variances has substantial influence on the value of the measure. Still, our data are generated with a larger variance for u_{1t} balancing the different roles of both innovation terms. Whereas the innovation u_{2t} represents the increment of a random walk process, the innovation u_{1t} drives a stationary process and needs to have a larger variance to make this relationship distinguishable. Otherwise, the generated time series p_{1t} and p_{2t} are almost perfectly correlated and it becomes difficult to estimate parameters of the second equation. Note that a parametric bootstrap algorithm can be distorted by noisy parameter estimates although the deviation from zero might not be statistically significant.

The second part of our simulation experiment focusses on microstructural models. We consider an asset traded in two venues. In the first microstructural model, the efficient price is driven by nontrade public information (u_t). The prices of both markets are determined by the efficient price and bid-ask bounce. The observed price in market $i = 1, 2$ is denoted by p_{it} and the unobserved common stochastic trend (efficient price) is denoted by m_t. DGP I.b is generating data according to

$$\begin{aligned} m_t &= m_{t-1} + u_t \qquad u_t \sim N(0, \sigma_u^2) \\ q_{it} &= \begin{cases} -1 \text{ with prob. } 1/2 \\ +1 \text{ with prob. } 1/2 \end{cases} \qquad i = 1, 2 \\ p_{it} &= m_t + c_i q_{it}, \end{aligned}$$

where the stochastic elements q_{1t}, q_{2t}, and u_t are uncorrelated. Both markets have equal information share if the half-spreads are specified such that $c_1 = c_2 = 1$. Furthermore, we set $\sigma_u^2 = 1$ for our simulations. Although the structural information shares are identical to

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

DGP I.a, estimation of both information share measures is complicated by the fact that the process can only be approximated by a reduced-form VECM. Since the process has a vector moving average representation (Hasbrouck, 2002), we have to take into account that the approximation by fixed lag length VECMs leads to underspecified models affecting the properties of HIS and MIS.

The second microstructural model considers an efficient price that incorporates private information. The first price series follows the efficient prices lagged by one period, while the second price series is equal to the efficient price plus a half-spread. Therefore, only the second market impounds new information in the efficient price and the first market has zero information share. The process is denoted by DGP III.b and is given by

$$m_t = m_{t-1} + \lambda q_{2t}$$

$$q_{it} = \begin{cases} -1 \text{ with prob. } 1/2 \\ +1 \text{ with prob. } 1/2 \end{cases} \qquad i = 1, 2$$

$$p_{1t} = m_{t-1} + c_1 q_{1t}$$

$$p_{2t} = m_t + c_2 q_{2t}.$$

A Monte Carlo simulation experiment concerned with bootstrap procedures has to fulfill $B, R \to \infty$, where R is the number of replications and B is the number of bootstrap draws. Assuming that the number of bootstrap replications is fixed, every added Monte Carlo iteration contributes multiplicatively to the overall computational cost. To avoid this inefficiency, we refer to the "Warp-speed" bootstrap described in Giacomini, Politis, and White (2013). The authors provide a formal proof that it is sufficient to draw only one bootstrap replication in each Monte Carlo replication and to evaluate the statistic of interest against the resulting bootstrap distribution of size R.

The results for the bootstrap confidence bands are reported in Tables 1–4. We find that the coverage rates for HIS and MIS under the "equal shares" hypothesis converge to their nominal levels with increasing sample size. Since the lag order for DGP I.a is correctly specified, we observe that the coverage closest to the nominal level is reached for $q = 0$, that is, i.i.d. draws from the residuals. Accounting for nonexistent autocorrelation in the residuals leads to confidence bands that are slightly too narrow. In contrast, the lag order for DGP I.b is mis-specified by construction so that some autocorrelation is left in the residuals. Accounting for this fact in our stationary bootstrap procedure, that is, increasing the probability q to select the adjacent residual set, guarantees that our bootstrap draws better approximate the observed data. For empirical applications, it is therefore recommended to evaluate the robustness of test decisions on the choice of q if lag selection is not straightforward (e.g., in situations where information criteria and diagnostic tests do not agree unanimously on a specific lag length).

Panels 4 and 5 of Tables 1 and 3 show that the Sapp algorithm does not yield appropriate coverage rates if the true information shares are at the boundaries of the interval. The confidence bands for DGPs under the "one-central-market" hypothesis are much too narrow and do not converge to the nominal levels. Turning to the average length $(1/R \sum_{i=1}^{R} c_\alpha^i - c_{1-\alpha}^i)$ and average shape $(1/R \sum_{i=1}^{R} c_\alpha^i / c_{1-\alpha}^i)$ of bootstrap confidence intervals (Tables 2 and 4), we first observe that the average shape of confidence bands for DGP

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Table 1 Empirical coverage rates for bootstrap confidence intervals (HIS)

		DGP I.a			DGP I.b			DGP II			DGP III.a			DGP III.b		
	Parameters	$\alpha_1 = -0.5, \alpha_2 = 0.5$			$c_1 = c_2 = 1, \sigma_u = 1$			$\alpha_1 = -0.75, \alpha_2 = 0.25$			$\alpha_1 = -1, \alpha_2 = 0$			$c_1 = c_2 = 1, \lambda = 1$		
	Structural IS	50			50			90			100			100		
	q	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
$T = 200$	0.00	89.1	94.6	98.9	92.2	96.4	99.6	86.4	90.7	96.1	77.0	85.8	94.7	72.9	82.5	94.1
	0.05	89.0	94.4	98.8	92.0	96.1	99.3	86.5	90.9	96.1	76.9	85.9	95.1	73.3	81.8	94.0
	0.10	88.9	94.4	99.0	91.4	96.0	99.3	86.6	90.5	96.2	77.8	86.4	94.8	73.9	83.2	94.7
$T = 400$	0.00	89.5	94.8	98.9	91.7	96.2	99.4	87.9	92.0	96.5	77.3	86.1	95.1	74.0	83.6	94.1
	0.05	89.3	94.6	98.8	91.7	96.1	99.3	88.3	92.1	96.7	76.9	85.7	94.9	75.3	83.5	93.9
	0.10	89.3	94.5	99.0	91.6	96.0	99.2	88.2	91.9	96.7	76.8	85.9	95.2	75.2	84.1	94.5
$T = 800$	0.00	89.9	95.1	98.9	91.4	96.0	99.2	89.3	93.5	97.2	77.5	86.2	95.5	74.3	83.1	94.3
	0.05	89.9	95.1	99.1	91.4	95.8	99.3	89.5	93.6	97.6	77.4	85.9	95.6	75.6	84.1	94.2
	0.10	90.0	95.2	98.9	91.1	95.8	99.2	89.5	93.9	97.4	77.5	86.4	95.4	75.4	84.1	93.9
$T = 1600$	0.00	90.1	95.0	99.0	91.3	95.9	99.2	89.4	94.1	98.0	78.3	86.0	95.3	76.0	84.7	95.0
	0.05	89.8	95.1	99.1	91.2	95.9	99.2	89.5	94.3	98.1	77.4	86.4	95.5	75.4	84.3	94.8
	0.10	89.6	94.4	99.0	91.1	95.6	99.2	89.7	94.3	98.3	77.4	86.4	95.5	74.4	83.8	94.1

Notes: We draw $R = 25{,}000$ replications from the DGPs described in De Jong (2002) and Hasbrouck (2002) and apply the "Warp-speed" bootstrap algorithm described in Giacomini, Politis, and White (2013) to obtain empirical coverage rates. The structural information shares are computed for p_2 and are expressed in percent. The lag length for DGP I.a, DGP II, and DGP III.a is fixed at $K = 2$. The lag length for DGP I.b and DGP III.b is determined by Akaike information criterion (AIC) and maximum lag length $K = 8$.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Table 2 Average length (shape) of bootstrap confidence intervals (HIS)

		DGP I.a			DGP I.b			DGP II			DGP III.a			DGP III.b		
	Parameters	$\alpha_1=-0.5, \alpha_2=0.5$			$c_1=c_2=1, \sigma_u=1$			$\alpha_1=-0.75, \alpha_2=0.25$			$\alpha_1=-1, \alpha_2=0$			$c_1=c_2=1, \lambda=1$		
	Structural IS	50			50			90			100			100		
	q	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
$T=200$	0.00	0.349 (1.008)	0.415 (1.007)	0.537 (1.000)	0.312 (1.003)	0.383 (0.984)	0.545 (0.989)	0.213 (1.551)	0.265 (1.574)	0.375 (1.615)	0.082 (3.612)	0.110 (3.302)	0.177 (3.035)	0.103 (2.692)	0.140 (2.571)	0.248 (2.631)
	0.05	0.347 (0.994)	0.413 (1.000)	0.532 (0.996)	0.309 (1.016)	0.373 (1.010)	0.518 (1.023)	0.213 (1.570)	0.263 (1.586)	0.365 (1.609)	0.083 (3.726)	0.111 (3.356)	0.174 (2.841)	0.105 (2.720)	0.141 (2.637)	0.241 (2.584)
	0.10	0.349 (1.004)	0.417 (0.985)	0.540 (1.006)	0.307 (1.011)	0.375 (0.997)	0.527 (0.987)	0.214 (1.552)	0.264 (1.577)	0.373 (1.607)	0.083 (3.501)	0.113 (3.156)	0.181 (3.022)	0.105 (2.663)	0.144 (2.568)	0.257 (2.661)
$T=400$	0.00	0.252 (1.000)	0.299 (1.012)	0.395 (1.037)	0.224 (1.014)	0.273 (0.996)	0.387 (0.983)	0.147 (1.406)	0.181 (1.466)	0.248 (1.502)	0.042 (3.702)	0.057 (3.352)	0.090 (2.972)	0.058 (3.006)	0.079 (2.776)	0.134 (2.732)
	0.05	0.250 (0.989)	0.297 (0.993)	0.387 (0.987)	0.225 (0.987)	0.275 (0.985)	0.383 (0.966)	0.148 (1.404)	0.180 (1.467)	0.251 (1.562)	0.042 (3.688)	0.056 (3.366)	0.089 (3.011)	0.057 (2.778)	0.079 (2.776)	0.130 (2.703)
	0.10	0.248 (1.014)	0.296 (0.995)	0.392 (1.010)	0.223 (0.994)	0.270 (0.990)	0.382 (0.989)	0.146 (1.413)	0.177 (1.466)	0.245 (1.507)	0.041 (3.754)	0.055 (3.310)	0.089 (2.943)	0.057 (2.804)	0.078 (2.680)	0.139 (2.714)
$T=800$	0.00	0.181 (1.018)	0.216 (1.013)	0.279 (1.022)	0.162 (1.017)	0.197 (0.993)	0.282 (0.964)	0.104 (1.263)	0.125 (1.311)	0.169 (1.439)	0.021 (3.600)	0.027 (3.252)	0.045 (2.874)	0.031 (2.886)	0.043 (2.828)	0.072 (2.679)
	0.05	0.179 (0.995)	0.214 (0.995)	0.278 (0.991)	0.161 (1.003)	0.197 (1.006)	0.277 (0.999)	0.104 (1.264)	0.125 (1.317)	0.166 (1.351)	0.021 (3.673)	0.028 (3.358)	0.046 (2.929)	0.032 (2.818)	0.044 (2.778)	0.072 (2.663)
	0.10	0.180 (0.998)	0.216 (0.987)	0.279 (0.968)	0.164 (0.986)	0.201 (0.990)	0.278 (0.975)	0.104 (1.274)	0.125 (1.308)	0.168 (1.419)	0.020 (3.650)	0.028 (3.315)	0.047 (3.062)	0.032 (2.834)	0.043 (2.642)	0.073 (2.776)
$T=1600$	0.00	0.127 (0.998)	0.151 (0.992)	0.202 (1.023)	0.122 (1.010)	0.147 (1.019)	0.203 (1.039)	0.073 (1.180)	0.088 (1.226)	0.118 (1.305)	0.010 (3.544)	0.014 (3.287)	0.023 (3.044)	0.018 (2.882)	0.024 (2.788)	0.041 (2.565)
	0.05	0.127 (0.991)	0.153 (0.990)	0.202 (1.000)	0.122 (1.016)	0.149 (0.998)	0.206 (1.016)	0.073 (1.182)	0.088 (1.216)	0.118 (1.283)	0.010 (3.604)	0.014 (3.210)	0.022 (2.785)	0.018 (3.072)	0.024 (2.903)	0.041 (2.729)
	0.10	0.128 (0.990)	0.152 (0.989)	0.199 (0.992)	0.120 (1.016)	0.147 (1.014)	0.207 (0.964)	0.073 (1.175)	0.087 (1.202)	0.116 (1.230)	0.010 (3.595)	0.014 (3.265)	0.023 (2.912)	0.018 (3.109)	0.024 (2.899)	0.041 (2.873)

Notes: We draw $R=25{,}000$ replications from the DGPs described in De Jong (2002) and Hasbrouck (2002) and apply the "Warp-speed" bootstrap algorithm described in Giacomini, Politis, and White (2013). The structural information shares are computed for p_2 and are expressed in percent. The lag length for DGP I.a, DGP II, and DGP III.a is fixed at $K=2$. The lag length for DGP I.b and DGP III.b is determined by Akaike information criterion (AIC) and maximum lag length $K=8$. The average length is reported first, the average shape is given in parentheses.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Table 3 Empirical coverage rates for bootstrap confidence intervals (MIS)

		DGP I.a			DGP I.b			DGP II			DGP III.a			DGP III.b		
	Parameters	$\alpha_1=-0.5, \alpha_2=0.5$			$c_1=c_2=1, \sigma_u=1$			$\alpha_1=-0.75, \alpha_2=0.25$			$\alpha_1=-1, \alpha_2=0$			$c_1=c_2=1, \lambda=1$		
	Structural IS	50			50			90			100			100		
	q	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
$T=200$	0.00	89.2	94.7	99.0	92.1	96.4	99.4	86.9	90.9	96.2	79.6	87.2	95.2	75.0	83.5	94.7
	0.05	89.1	94.4	98.8	92.0	96.2	99.5	87.1	91.2	96.3	79.9	87.6	95.7	75.0	83.2	94.3
	0.10	89.0	93.6	98.9	91.4	96.0	99.3	87.1	91.0	96.3	80.3	87.6	95.3	75.7	84.4	95.2
$T=400$	0.00	89.6	94.7	99.0	91.9	96.1	99.3	88.2	92.4	96.7	79.9	87.5	95.9	76.2	84.8	94.6
	0.05	89.4	94.7	98.8	91.8	96.1	99.3	88.5	92.3	96.8	79.7	87.5	95.3	77.1	84.5	94.2
	0.10	89.4	94.6	98.9	91.7	95.9	99.3	88.5	92.1	96.8	79.7	87.5	95.7	76.7	85.3	94.9
$T=800$	0.00	89.9	95.1	98.9	91.4	95.9	99.2	89.5	93.6	97.3	80.2	87.9	96.0	76.0	84.2	94.9
	0.05	90.0	95.1	99.1	91.4	96.0	99.3	89.6	93.8	97.7	80.1	87.5	95.9	77.2	85.2	94.6
	0.10	90.1	95.2	98.9	91.1	95.7	99.2	89.7	94.0	97.5	80.5	87.7	95.9	77.1	85.4	94.4
$T=1600$	0.00	90.0	94.9	99.0	91.3	95.6	99.2	89.5	94.2	98.1	80.5	87.9	95.6	77.6	85.6	95.3
	0.05	90.1	95.3	99.1	91.2	95.9	99.3	89.5	94.3	98.2	80.2	87.9	95.9	77.4	85.2	95.0
	0.10	90.1	95.4	99.1	91.1	95.8	99.3	89.7	94.3	98.4	80.1	87.8	95.9	76.2	84.5	94.5

Notes: We draw $R=25{,}000$ replications from the DGPs described in De Jong (2002) and Hasbrouck (2002) and apply the "Warp-speed" bootstrap algorithm described in Giacomini, Politis, and White (2013) to obtain empirical coverage rates. The structural information shares are computed for p_2 and are expressed in percent. The lag length for DGP I.a, DGP II, and DGP III.a is fixed at $K=2$. The lag length for DGP I.b and DGP II.b is determined by Akaike information criterion (AIC) and maximum lag length $K=8$.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Table 4 Average length (shape) of bootstrap confidence intervals (MIS)

		DGP I.a			DGP I.b			DGP II			DGP III.a			DGP III.b		
	Parameters		$\alpha_1 = -0.5, \alpha_2 = 0.5$			$c_1 = c_2 = 1, \sigma_u = 1$			$\alpha_1 = -0.75, \alpha_2 = 0.25$			$\alpha_1 = -1, \alpha_2 = 0$			$c_1 = c_2 = 1, \lambda = 1$	
	Structural IS		50			50			90			100			100	
	q	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
$T = 200$	0.00	0.351 (1.006)	0.417 (1.008)	0.540 (1.003)	0.382 (1.003)	0.467 (0.981)	0.659 (0.972)	0.214 (1.521)	0.266 (1.565)	0.376 (1.612)	0.080 (3.395)	0.109 (3.200)	0.175 (2.949)	0.102 (2.558)	0.139 (2.511)	0.248 (2.605)
	0.05	0.349 (0.996)	0.414 (1.001)	0.536 (0.997)	0.374 (1.001)	0.457 (0.997)	0.646 (1.001)	0.214 (1.542)	0.264 (1.576)	0.366 (1.603)	0.082 (3.514)	0.110 (3.223)	0.172 (2.776)	0.104 (2.603)	0.141 (2.570)	0.241 (2.578)
	0.10	0.351 (1.004)	0.419 (0.985)	0.542 (1.009)	0.374 (1.017)	0.457 (1.009)	0.631 (1.002)	0.215 (1.521)	0.265 (1.553)	0.375 (1.596)	0.083 (3.318)	0.111 (3.073)	0.179 (2.941)	0.104 (2.555)	0.143 (2.483)	0.258 (2.608)
$T = 400$	0.00	0.253 (1.002)	0.300 (1.012)	0.396 (1.036)	0.276 (1.015)	0.333 (0.995)	0.470 (0.996)	0.147 (1.386)	0.181 (1.447)	0.249 (1.508)	0.041 (3.486)	0.055 (3.277)	0.088 (2.858)	0.058 (2.833)	0.078 (2.675)	0.132 (2.647)
	0.05	0.250 (0.989)	0.298 (0.992)	0.388 (0.985)	0.275 (0.980)	0.336 (0.966)	0.470 (0.968)	0.148 (1.381)	0.181 (1.454)	0.252 (1.554)	0.041 (3.455)	0.056 (3.180)	0.088 (2.979)	0.057 (2.650)	0.079 (2.721)	0.130 (2.649)
	0.10	0.249 (1.011)	0.297 (0.996)	0.394 (1.006)	0.273 (0.993)	0.329 (0.990)	0.461 (0.989)	0.147 (1.389)	0.178 (1.457)	0.245 (1.492)	0.040 (3.495)	0.054 (3.165)	0.089 (2.914)	0.057 (2.714)	0.078 (2.592)	0.139 (2.672)
$T = 800$	0.00	0.181 (1.018)	0.216 (1.012)	0.279 (1.021)	0.199 (1.014)	0.242 (0.989)	0.345 (0.976)	0.104 (1.251)	0.126 (1.304)	0.169 (1.433)	0.020 (3.379)	0.027 (3.107)	0.044 (2.779)	0.031 (2.764)	0.042 (2.726)	0.072 (2.603)
	0.05	0.179 (0.997)	0.214 (0.995)	0.278 (0.990)	0.198 (1.010)	0.241 (1.001)	0.339 (0.986)	0.104 (1.252)	0.125 (1.304)	0.166 (1.337)	0.021 (3.464)	0.028 (3.181)	0.045 (2.864)	0.032 (2.716)	0.043 (2.682)	0.072 (2.604)
	0.10	0.180 (0.998)	0.216 (0.986)	0.279 (0.969)	0.200 (0.986)	0.245 (0.996)	0.340 (0.979)	0.104 (1.260)	0.125 (1.298)	0.168 (1.415)	0.020 (3.374)	0.027 (3.211)	0.046 (3.016)	0.032 (2.705)	0.043 (2.537)	0.072 (2.710)
$T = 1600$	0.00	0.127 (0.998)	0.152 (0.992)	0.202 (1.023)	0.149 (1.010)	0.180 (1.024)	0.248 (1.024)	0.073 (1.172)	0.088 (1.219)	0.118 (1.291)	0.010 (3.360)	0.014 (3.110)	0.023 (3.003)	0.017 (2.783)	0.024 (2.717)	0.041 (2.518)
	0.05	0.127 (0.990)	0.153 (0.990)	0.203 (1.003)	0.149 (1.014)	0.182 (0.999)	0.252 (1.015)	0.073 (1.173)	0.088 (1.209)	0.118 (1.275)	0.010 (3.414)	0.014 (3.082)	0.022 (2.720)	0.018 (2.900)	0.024 (2.834)	0.041 (2.721)
	0.10	0.128 (0.989)	0.152 (0.989)	0.199 (0.992)	0.147 (1.016)	0.179 (1.017)	0.254 (0.968)	0.073 (1.169)	0.087 (1.195)	0.116 (1.220)	0.010 (3.382)	0.014 (3.113)	0.023 (2.864)	0.017 (2.961)	0.024 (2.844)	0.041 (2.811)

Notes: We draw $R = 25{,}000$ replications from the DGPs described in De Jong (2002) and Hasbrouck (2002) and apply the "Warp-speed" bootstrap algorithm described in Giacomini, Politis, and White (2013). The structural information shares are computed for p_2 and are expressed in percent. The lag length for DGP I.a, DGP II, and DGP III.a is fixed at $K = 2$. The lag length for DGP I.b and DGP II.b is determined by Akaike information criterion (AIC) and maximum lag length $K = 8$. The average length is reported first, the average shape is given in parentheses.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

I.a and DGP I.b are close to one and therefore symmetric. In contrast, we find that the average shape of confidence bands for DGP III.a and DGP III.b are strongly skewed which leads to wrong coverage rates. To investigate this further, we draw from a model where one market dominates prices discovery but does not have 100% information share. The results for DGP II show that bootstrap confidence bands, in general, are too narrow if the average length is large enough to include boundary values. In these cases, large samples are needed to reach nominal coverage rates. We observe that the rate of convergence becomes slower, the closer the true information shares are to the boundaries.

While it seems that the average length of bootstrap confidence intervals for DGP III is substantially smaller than for DGP I, the average length depends, inter alia, on the speed of adjustment and maximum lag length which is particularly relevant for processes with infinite lag order. Hence, the average lengths cannot be directly compared across DGPs. Our results suggest that the performance of the Sapp algorithm depends heavily on the underlying process. Bootstrap confidence bands should be used cautiously if the initial information share estimates are too close to their boundaries and, in particular, should not be used to evaluate the "one-central-market" hypothesis.[3]

Following up on the shortcomings of the Sapp algorithm, we use the same "one-central-market" processes to evaluate our new bootstrap test under zero information share restrictions. The results are reported in Tables 5–8. The tests for HIS seem to be slightly undersized but converge slowly to their nominal significance levels. Similar to the confidence bands, we find that accounting for additional autocorrelation in the residuals can improve the statistical properties of the test for microstructural models with infinite lag order (truncation effects) but not for the parametric models with fixed lag order. Again, the tests for MIS seem to converge faster to their nominal significance levels. The power against the null hypotheses is evaluated under different parametrizations of the DGP in Equation (13). We find that the power converges to unity for all specifications. As expected, the differences between HIS and MIS are marginal in large samples. Additionally, we consider alternative strong exogeneity tests and report our results in Table 9. The DGPs feature slow adjustment generating noisy cointegrated systems. The Dolado–Lütkepohl test is evaluated against the strong exogeneity test in a VECM without imposing contemporaneously uncorrelated reduced-form innovations and the bootstrap test for the MIS. We find that the performances of the Dolado–Lütkepohl test and the mis-specified VECM test depend on the adjustment coefficient, that is, on the signal-to-noise ratio, for small to medium sample sizes whereas the bootstrap variant maintains its nominal size throughout. Size distortions for all tests vanish in large samples.[4] Overall, we find that our proposed restricted bootstrap procedure has statistical properties superior to the Sapp algorithm if one market dominates

3 Because the length and shape of bootstrap confidence bands depend on, inter alia, the adjustment dynamics, innovation variances, and the sample size, it is difficult to decide whether information shares are too close to the boundaries to obtain correct coverage rates. As a general guideline, it is advisable to discard confidence intervals which include zero or one. These confidence intervals are usually skewed and do not hold their nominal size.

4 Further simulation experiments revealed that the bootstrap test has the highest size-adjusted power in small to medium sample size out of all tests considered. These results are not reported but can be obtained from the author upon request.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Table 5 Empirical confidence levels for the "one-central-market" hypothesis test (HIS)

		DGP III.a			DGP III.b		
	Parameters	$\alpha_1 = -1, \alpha_2 = 0$			$c_1 = c_2 = 1, \lambda = 1$		
	Structural IS	100			100		
	q	90%	95%	99%	90%	95%	99%
$T = 200$	0.00	91.4	96.1	99.4	92.3	97.0	99.8
	0.05	91.4	96.1	99.5	92.2	96.8	99.8
	0.10	91.2	96.5	99.6	93.2	97.7	99.8
$T = 400$	0.00	90.7	95.3	99.0	90.2	95.5	99.4
	0.05	90.7	95.4	99.3	90.0	95.3	99.4
	0.10	91.8	96.2	99.3	90.7	96.0	99.6
$T = 800$	0.00	90.4	95.1	99.1	88.8	94.4	98.9
	0.05	91.0	95.6	99.3	89.6	94.9	99.2
	0.10	91.2	95.8	99.3	90.7	95.5	99.3
$T = 1600$	0.00	90.5	95.3	99.1	87.6	93.4	98.7
	0.05	90.8	95.5	99.2	89.0	94.4	99.0
	0.10	91.2	95.9	99.2	90.4	95.3	99.3

Notes: We draw $R = 25{,}000$ replications from the DGPs described in De Jong (2002) and Hasbrouck (2002) and apply the "Warp-speed" bootstrap algorithm described in Giacomini, Politis, and White (2013) to obtain bootstrap distributions. The structural information shares are computed for p_2 and are expressed in percent. The lag length for DGP III.a is fixed at $K = 2$, while the lag length for DGP III.b is determined by Akaike information criterion (AIC) and maximum lag length $K = 8$.

Table 6 Empirical confidence levels for the "one-central-market" hypothesis test (MIS)

		DGP III.a			DGP III.b		
	Parameters	$\alpha_1 = -1, \alpha_2 = 0$			$c_1 = c_2 = 1, \lambda = 1$		
	Structural IS	100			100		
	q	90%	95%	99%	90%	95%	99%
$T = 200$	0.00	90.5	95.8	99.3	88.1	94.5	99.4
	0.05	90.1	95.1	99.0	87.9	94.0	99.5
	0.10	90.4	95.1	99.0	89.2	95.3	99.6
$T = 400$	0.00	91.2	95.7	99.2	87.9	94.0	99.1
	0.05	90.4	95.0	99.1	87.8	93.9	99.2
	0.10	90.1	95.0	98.9	88.6	94.7	99.5
$T = 800$	0.00	90.9	95.8	99.2	87.5	93.5	98.7
	0.05	90.7	95.3	99.2	88.3	94.1	99.0
	0.10	90.2	95.0	99.1	89.5	94.9	99.2
$T = 1600$	0.00	91.0	95.9	99.3	86.9	92.8	98.6
	0.05	90.8	95.4	99.3	88.4	94.0	98.8
	0.10	90.5	95.2	99.2	89.7	94.8	99.1

Notes: We draw $R = 25{,}000$ replications from the DGPs described in De Jong (2002) and Hasbrouck (2002) and apply the "Warp-speed" bootstrap algorithm described in Giacomini, Politis, and White (2013) to obtain bootstrap distributions. The structural information shares are computed for p_2 and are expressed in percent. The lag length for DGP III.a is fixed at $K = 2$, while the lag length for DGP III.b is determined by Akaike information criterion (AIC) and maximum lag length $K = 8$.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Table 7 Size-adjusted power of bootstrap test (HIS)

Hypothesis	"One-central-market"								
Parameters		$\alpha_1 = -0.9, \alpha_2 = 0.1$ $\sigma_1 = 1, \sigma_2 = 1, \rho = 0$			$\alpha_1 = -0.75, \alpha_2 = 0.25$ $\sigma_1 = 1, \sigma_2 = 1, \rho = 0$			$\alpha_1 = -0.5, \alpha_2 = 0.5$ $\sigma_1 = 1, \sigma_2 = 1, \rho = 0$	
Structural IS		$\approx$99			90			50	
T	10%	5%	1%	10%	5%	1%	10%	5%	1%
200	21.4	12.3	3.3	91.3	83.2	57.5	100	100	100
400	39.4	26.0	8.3	99.7	99.0	92.9	100	100	100
800	66.8	54.1	24.9	100	100	100	100	100	100
1600	92.1	85.1	63.3	100	100	100	100	100	100
Parameters		$\alpha_1 = -0.2, \alpha_2 = 0.2$ $\sigma_1 = 1, \sigma_2 = 16, \rho = 0.5$			$\alpha_1 = -0.2, \alpha_2 = 0.2$ $\sigma_1 = 1, \sigma_2 = 8, \rho = 0.5$			$\alpha_1 = -0.2, \alpha_2 = 0.2$ $\sigma_1 = 1, \sigma_2 = 4, \rho = 0.5$	
Structural IS		$\approx$85			$\approx$82			$\approx$77	
T	10%	5%	1%	10%	5%	1%	10%	5%	1%
200	75.1	56.1	16.8	94.8	85.6	48.7	98.8	96.2	76.4
400	97.2	91.7	62.6	99.9	99.6	95.4	100	100	99.6
800	100	100	99.0	100	100	100	100	100	100
1600	100	100	100	100	100	100	100	100	100

Notes: We draw $R = 10{,}000$ replications from the DGPs described in De Jong (2002) and apply the "Warp-speed" bootstrap algorithm described in Giacomini, Politis, and White (2013) to obtain bootstrap distributions. The structural information shares are computed for p_2 and are expressed in percent. The lag length is fixed at $K = 2$. The probability to draw adjacent residuals is set to $q = 0.1$.

Table 8 Size-adjusted power of bootstrap test (MIS)

Hypothesis	"one-central-market"								
Parameters		$\alpha_1 = -0.9, \alpha_2 = 0.1$ $\sigma_1 = 1, \sigma_2 = 1, \rho = 0$			$\alpha_1 = -0.75, \alpha_2 = 0.25$ $\sigma_1 = 1, \sigma_2 = 1, \rho = 0$			$\alpha_1 = -0.5, \alpha_2 = 0.5$ $\sigma_1 = 1, \sigma_2 = 1, \rho = 0$	
Structural IS		$\approx$99			90			50	
T	10%	5%	1%	10%	5%	1%	10%	5%	1%
200	24.4	15.1	4.9	88.4	80.1	56.8	100	100	100
400	40.5	27.8	10.0	99.3	98.3	91.7	100	100	100
800	66.6	54.7	26.1	100	100	100	100	100	100
1600	91.6	84.5	63.5	100	100	100	100	100	100
Parameters		$\alpha_1 = -0.2, \alpha_2 = 0.2$ $\sigma_1 = 1, \sigma_2 = 16, \rho = 0.5$			$\alpha_1 = -0.2, \alpha_2 = 0.2$ $\sigma_1 = 1, \sigma_2 = 8, \rho = 0.5$			$\alpha_1 = -0.2, \alpha_2 = 0.2$ $\sigma_1 = 1, \sigma_2 = 4, \rho = 0.5$	
Structural IS		$\approx$91			$\approx$87			$\approx$80	
T	10%	5%	1%	10%	5%	1%	10%	5%	1%
200	75.1	56.1	16.8	91.8	81.8	47.5	97.4	93.1	71.1
400	97.2	91.7	62.6	99.7	99.2	93.2	100	99.9	99.0
800	100	100	99.0	100	100	100	100	100	100
1600	100	100	100	100	100	100	100	100	100

Notes: We draw $R = 10{,}000$ replications from the DGPs described in De Jong (2002) and apply the "Warp-speed" bootstrap algorithm described in Giacomini, Politis, and White (2013) to obtain bootstrap distributions. The structural information shares are computed for p_2 and are expressed in percent. The lag length is fixed at $K = 2$. The probability to draw adjacent residuals is set to $q = 0.1$.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Table 9 Empirical confidence levels for strict exogeneity tests

DGP III.a	VECM			Dolado–Lütkepohl			Bootstrap (MIS)		
Parameters					$\alpha_1 = -0.05, \alpha_2 = 0$				
Structural IS					100				
T	90%	95%	99%	90%	95%	99%	90%	95%	99%
200	83.0	90.4	97.2	84.8	91.8	98.7	90.7	94.8	98.7
400	87.1	92.3	98.6	87.3	93.4	98.8	90.5	94.7	98.8
800	89.4	94.9	99.0	89.4	94.8	98.8	90.2	95.1	98.9
Parameters:					$\alpha_1 = -0.1, \alpha_2 = 0$				
Structural IS					100				
T	90%	95%	99%	90%	95%	99%	90%	95%	99%
200	86.1	92.4	98.2	86.8	92.4	98.8	90.4	95.2	98.7
400	88.2	93.6	98.8	88.6	94.1	98.9	90.5	95.1	98.9
800	90.0	95.0	98.8	89.7	94.6	99.0	90.5	95.0	98.9
Parameters					$\alpha_1 = -0.2, \alpha_2 = 0$				
Structural IS					100				
T	90%	95%	99%	90%	95%	99%	90%	95%	99%
200	87.8	93.4	98.8	87.3	93.7	99.0	91.1	95.2	98.6
400	88.4	94.2	98.8	88.9	94.2	98.9	90.6	95.1	98.7
800	89.6	95.1	99.0	89.7	95.1	99.1	90.1	95.0	98.9

Notes: We draw $R = 10{,}000$ replications from the DGP III.a and apply the "Warp-speed" bootstrap algorithm described in Giacomini, Politis, and White (2013) to obtain bootstrap distributions. To obtain noisy cointegration systems, we choose $\sigma_1^2 = 1$, $\sigma_2^2 = 5$, $\rho = 0$ and slow speeds of adjustment. The structural information shares are computed for p_2 and are expressed in percent. The lag length is fixed at $K = 2$. The probability to draw adjacent residuals is set to $q = 0.1$.

price discovery. Since it also has slightly better properties than the Dolado–Lütkepohl strong exogeneity test, it should be the method of choice to test the "one-central-market" hypothesis.

Finally, we consider two specifications of five-variable cointegrated systems to investigate whether our results for two-variable systems can be extended to larger systems. We specify both VECMs according to Equation (1), where $K = 2$, $\Gamma_1 = -0.1 I_5$ and

$$\beta' = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}. \tag{14}$$

Using this structure guarantees that both models are driven by one common stochastic trend. We generate data from two models with different adjustment behavior. The first

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

model (DGP IV) imposes equal information shares for all variables. Here, the adjustment coefficient matrix is set to

$$\alpha = \begin{bmatrix} -0.2 & 0 & 0 & 0 \\ 0 & -0.2 & 0 & 0 \\ 0 & 0 & -0.2 & 0 \\ 0 & 0 & 0 & -0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 \end{bmatrix}. \tag{15}$$

In the second model (DGP V), we generate data under the assumption that the first price variable does not contribute any adjustment to the long-run equilibrium. This results in a 100% information share of the first price series. The corresponding adjustment coefficient matrix is given by

$$\alpha = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -0.2 & 0 & 0 \\ 0 & 0 & -0.2 & 0 \\ 0 & 0 & 0 & -0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 \end{bmatrix}. \tag{16}$$

The empirical coverage rates for the MIS are reported in Table 10. Since the MIS is order-invariant, it is straightforward to extend it to N-variable systems. In contrast, the HIS midpoint would have to be computed over all possible ordering combinations and therefore does not give accurate estimates for structural information share close to the boundaries of the interval. We can infer from Table 10 that bootstrap confidence bands yield accurate coverage rates for truly equal information shares.[5] However, they are again too narrow if one market dominates price discovery. While coverage rates for dominated markets are only slightly too narrow, they strongly deviate from their nominal levels for the dominant market which invalidates tests based on this bootstrap algorithm. Our restricted bootstrap algorithm specifically designed for the "one-central-market" hypothesis yields accurate coverage rates.

3 Empirical Application: Price Discovery in CDS and Bond Markets

In this section, we revisit the empirical illustration presented in Lien and Shrestha (2014). The authors investigate the hypothesis that price discovery takes place mainly in the CDS market which is based on the fact that the CDS market is more liquid than the bond market. Lower short-selling costs and higher liquidity might attract traders with private information about the credit risk of an entity which could lead to faster impounding of new information in the CDS market. They analyze a panel of sixteen actively traded CDS with underlying bonds issued by large U.S. corporations over a 1 year span reaching from July 1, 2009 to June 30, 2010. We rebuild their dataset from Thomson Reuters Datastream and

5 It is important to note that bootstrap confidence intervals cannot be used straightforwardly to test the equal information shares hypothesis for $N > 2$. Confidence intervals may overlap, yet there may be a statistically significant difference between information shares.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Table 10 Empirical coverage rates for five-variable systems (MIS)

	DGP IV					DGP V					DGP V				
Method	Sapp					Sapp					Restricted				
Structural IS	20	20	20	20	20	100	0	0	0	0	100	0	0	0	0
Nominal Conf.								90 %							
T															
200	89.5	89.8	89.1	90.0	90.1	57.2	83.4	83.3	82.7	83.4	85.8	86.1	86.7	87.8	85.5
400	88.3	88.5	89.6	89.0	89.9	56.2	80.2	82.4	81.1	80.0	88.6	87.8	87.9	89.1	88.7
800	89.2	88.2	88.6	88.0	88.3	54.8	82.0	81.8	80.7	81.4	89.1	89.7	89.0	89.3	88.8
1600	89.1	88.8	89.2	89.2	89.2	54.9	81.9	81.0	81.6	80.7	90.9	90.4	91.6	90.9	89.6
Nominal Conf.								95 %							
T															
200	93.2	93.5	92.1	94.0	93.5	72.9	89.9	90.0	89.8	89.4	94.0	93.4	93.7	93.2	92.0
400	93.5	93.5	94.1	94.1	94.3	73.3	87.7	89.7	88.6	88.0	94.3	93.7	94.5	94.4	94.3
800	94.5	94.2	93.8	93.8	94.0	73.3	88.6	88.2	88.7	89.4	94.5	94.7	94.1	94.5	94.2
1600	94.4	94.1	94.9	94.4	93.9	72.5	88.9	87.9	88.1	88.9	95.6	94.7	95.6	95.0	94.2
Nominal Conf.								99 %							
T															
200	97.6	97.6	97.2	97.8	97.7	86.7	96.8	97.5	97.3	97.1	98.7	98.8	99.0	98.3	97.9
400	97.3	97.7	97.4	97.3	97.8	92.0	96.0	96.9	96.0	96.4	99.1	98.4	99.0	99.0	98.4
800	97.7	98.2	98.5	97.5	98.0	91.3	96.8	95.6	96.3	96.3	98.7	98.9	98.6	98.5	98.8
1600	98.4	98.0	98.4	98.2	98.3	91.1	96.1	96.7	96.8	96.4	98.8	99.0	99.0	99.0	98.6

Notes: We draw $R = 25{,}000$ replications from a five-variable VECM and apply the "Warp-speed" bootstrap algorithm described in Giacomini, Politis, and White (2013) to obtain bootstrap distributions. The structural information shares are computed for $(p_1, p_2, p_3, p_4, p_5)$ and are expressed in percent. The lag length is fixed at $K = 2$. The probability to draw adjacent residuals is set to $q = 0.1$.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Table 11 Summary statistics

No.	Firm name	MSCI Industry group	Average Basis	Coupon	Maturity	S&P Ratings
1.	American Express Co.	Diversified financial	–87	5.500	September 12, 2016	BBB+
2.	Avnet, Inc.	Technology hardware and equipment	–145	6.625	September 15, 2016	BBB−
3.	CMS Energy Corp.	Utilities	–120	6.875	December 15, 2015	BBB−
4.	First Data Corp.	Software and services	–212	9.875	September 24, 2015	B
5.	Halliburton Co.	Energy	–73	7.530	May 12, 2017	A
6.	Lennar Corp.	Consumer durables and apparel	–205	5.600	May 31, 2015	B+
7.	Limited Brands	Retailing	–208	6.900	July 15, 2017	BB+
8.	Macy's, Inc.	Retailing	–173	7.450	July 15, 2017	BBB−
9.	Motorola, Inc.	Technology hardware and equipment	–173	6.000	November 15, 2017	BB+
10.	Pioneer Natural Resources Co.	Energy	–192	6.650	March 15, 2017	BB+
11.	Standard Pacific Corp.	Consumer durables and apparel	–115	7.000	August 15, 2015	B+
12.	Allstate Corp.	Insurance	–110	6.750	May 15, 2018	A+
13.	Hertz Corp.	Transportation	–213	7.875	January 1, 2014	B+
14.	Neiman-Marcus Gr Incorp.	Retailing	–178	10.375	October 15, 2015	B+
15.	New York Times Co.	Media	–251	5.000	March 15, 2015	B+
16.	Toys R US, Inc.	Retailing	100	7.375	October 15, 2018	B

Notes: This table provides details of bonds issued by all firms considered in the sample. The average basis is computed as the difference between the CDS spread and the underlying bond yield spread. Note that the average basis slightly deviates from those values reported in Table I in Lien and Shrestha (2014).

report summary statistics in Table 11. We observe that the average basis which is computed as the average difference between the CDS spread and bond yield spread slightly deviates from those values reported in Table I in Lien and Shrestha (2014). Since the bond contracts are identical, we suspect that the data for some bond contracts have been revised by Thomson Reuters Datastream. However, the values deviate at most by 10 basis points and should not influence our results. Since the relationship between CDS spreads and bond yield spreads is not necessarily one-to-one, the MIS (or HIS) cannot be directly applied. Consequently, Lien and Shrestha (2014) propose the GIS which is based on the estimated cointegrating vector. The bootstrap algorithms described in Section 1.2 and Section 1.3 account for the fact that the cointegrating vector can be different from $(1, -1)$.

Lien and Shrestha (2014) first show that all sixteen pairs are cointegrated. Then, they report the GIS estimates with additional PT/GG common factor components. The hypotheses that price discovery takes place only in one market is evaluated exclusively for the PT/GG common factor components because the test can be easily applied as a likelihood ratio test on the vector of adjustment coefficients. In contrast, the corresponding tests of the "one-central-market" hypotheses for the GIS are not conducted. Since information shares and common factor components measure different aspects of the price discovery process,

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

rejection of a hypothesis in the context of common factor components does not contain any meaning for hypotheses on the GIS. Particularly, the test in the PT/GG framework amounts to a test for weak exogeneity whereas a test for strict exogeneity under the assumption of uncorrelated innovations is needed for market information shares as demonstrated in Section 1. Also, the reported GIS and PT/GG estimates do not agree with the ordering of information roles for all cases in the original study. For those reasons, we apply the newly proposed test of the "one-central-market" hypothesis to all pairs and compare our results to the PT/GG results. GIS estimates and p-values are reported in the first part of Table 12, while common factor components and the p-values of likelihood ratio tests are reported in the second part of Table 12.

Overall, we can closely replicate the original results reported in Lien and Shrestha (2014). In most cases, we find small numerical differences between the GIS estimates in Table 12 and those in the original article which might be attributed to different lag lengths and slight deviations from the average basis (Table 11).[6] The ordering of information roles according to the GIS and common factor components is identical for all CDS/bond pairs except for the New York Times Co., where the common factor component of CDS is 0.574 instead of 0.4613 in the original article. Similar to Lien and Shrestha (2014), we find that the GIS and common factor components do not agree on the ordering of information roles for Limited Brands. The results of the bootstrap test are only in line with reported PT/GG hypothesis tests for eleven out of sixteen pairs (American Express Co., Halliburton Co., Lennar Corp., Limited Brands, Motorola, Inc., Pioneer Natural Resources Co., Standard Pacific Corp., Hertz Corp., Neiman-Marcus Gr Corp., New York Times Co., Toys R US, Inc.). In the remaining five cases (Avnet, Inc., CMS Energy Corp., First Data Corp., Macy's, Inc. Allstate Corp.), we find significant contribution of the bond market to price discovery using the bootstrap test, but would not reject the null hypothesis in the PT/GG framework. This clearly highlights the need for additional testing in empirical applications concerned with market information share measures. Moreover, finding that the "one-central-market" hypothesis is rejected for eleven out of sixteen pairs at the 5% significance level, provides sufficient evidence to reconsider the null hypothesis that the bond market does not contribute substantially to price discovery. Taking into consideration that the bootstrap test almost exactly holds its nominal size at $T = 200$ and shows considerable power against the null hypothesis, we recommend using the newly proposed test instead of tests under the PT/GG framework if hypotheses on market information shares are of interest.

4 Conclusion

This article proposes new inferential tools for the analysis of contributions to price discovery. Although it has proven difficult to test general hypotheses about market information shares directly, we identify two important hypotheses and extensively discuss bootstrap methods to conduct hypothesis tests in these cases. We can show that bootstrap confidence bands based on the Sapp algorithm have sufficient coverage in situations where the

6 It seems that the estimated cointegrated systems are quite noisy. The null hypothesis that the bond market of Hertz Corp. or Neiman-Marcus Gr Corp. has zero information share is not rejected at the 10% level although the estimated GIS is 0.280 and 0.169, respectively.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Table 12 GISs and hypothesis tests

No.	Firm name	Generalized Information Share: CDS	Generalized Information Share: Bond	H_0: (0, 1)	H_0: (1, 0)	Common Factor Components: CDS	Common Factor Components: Bond	H_0: (0, 1)	H_0: (1, 0)
1.	American Express Co.	0.988	0.012	0.011	0.593	0.982	0.018	0.009	0.940
2.	Avnet, Inc.	0.841	0.159	0.000	0.013	1.081	–0.081	0.020	0.080
3.	CMS Energy Corp.	0.756	0.244	0.014	0.034	0.757	0.243	0.000	0.100
4.	First Data Corp.	0.763	0.237	0.024	0.041	0.826	0.174	0.040	0.640
5.	Halliburton Co.	0.070	0.930	0.018	0.001	–0.008	1.008	0.980	0.000
6.	Lennar Corp.	0.996	0.004	0.001	0.570	1.155	–0.155	0.000	0.380
7.	Limited Brands	0.477	0.523	0.023	0.016	0.682	0.318	0.200	0.170
8.	Macy's, Inc.	0.547	0.453	0.018	0.036	0.645	0.355	0.040	0.070
9.	Motorola, Inc.	0.977	0.023	0.016	0.529	1.009	–0.009	0.030	0.970
10.	Pioneer Natural Resources Co.	0.378	0.622	0.043	0.010	0.469	0.531	0.150	0.050
11.	Standard Pacific Corp.	0.063	0.937	0.221	0.014	0.098	0.902	0.830	0.010
12.	Allstate Corp.	0.657	0.343	0.022	0.016	0.755	0.245	0.010	0.080
13.	Hertz Corp.	0.720	0.280	0.016	0.146	0.785	0.215	0.040	0.230
14.	Neiman-Marcus Gr Corp.	0.831	0.169	0.210	0.105	0.796	0.204	0.050	0.650
15.	New York Times Co.	0.333	0.667	0.005	0.048	0.574	0.426	0.030	0.000
16.	Toys R US, Inc.	0.264	0.736	0.063	0.014	0.252	0.748	0.160	0.010

Notes: The hypothesis that the bond market is the central market and the bond market has zero information share is denoted by H_0: (0, 1). The opposing hypothesis that the CDS market is the central market and the bond market has zero information share is denoted by H_0: (1, 0). Note that both null hypotheses have a different meaning for market information shares and in the PT/GG framework. The probability of drawing the adjacent residual element is $q = 0.05$. Bootstrap p-values are computed from 800 bootstrap replications.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

contributions of all markets are relatively similar. By contrast, evaluating the "one-central-market" hypothesis at the bounds of the interval requires a different statistical approach. In these situations, we recommend to use a bootstrap test which generates data under the implied parameter restrictions and has improved statistical properties. Since both approaches are based on parametric bootstrap algorithms, their performance depends on carefully specified reduced-form VECMs. Our proposed bootstrap methods can be applied in all price discovery studies where inferential statements on market information shares are needed.

References

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.

Booth, G. G., R. W. So, and Y. Tse. 1999. Price Discovery in the German Equity Index Derivatives Markets. *The Journal of Futures Markets* 19: 619–643.

Carbera, J., T. Wang, and J. Yang. 2009. Do Futures Lead Price Discovery in Electronic Foreign Exchange Markets? *The Journal of Futures Markets* 29: 137–156.

De Jong, F. 2002. Measures of Contributions to Price Discovery: A Comparison. *Journal of Financial Markets* 5: 323–327.

De Jong, F., and P. C. Schotman. 2010. Price Discovery in Fragmented Markets. *Journal of Financial Econometrics* 8: 1–28.

Dimpfl, T., M. Flad, and R. C. Jung. 2017. Price Discovery in Agricultural Commodity Markets in the Presence of Futures Speculation. *Journal of Commodity Markets* 5: 50–62.

Dolado, J. J., and H. Lütkepohl. 1996. Making Wald Tests Work for Cointegrated VAR Systems. *Econometric Reviews* 15: 369–386.

Figuerola-Ferretti, I., and J. Gonzalo. 2010. Modelling and Measuring Price Discovery in Commodity Markets. *Journal of Econometrics* 158: 95–107.

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.

Giacomini, R., D. N. Politis, and H. White. 2013. A Warp-Speed Method for Conducting Monte Carlo Experiments Involving Bootstrap Estimators. *Econometric Theory* 29: 567–589.

Gonzalo, J., and C. W. J. Granger. 1995. Estimation of Common Long-Memory Components in Cointegrated Systems. *Journal of Business & Economic Statistics* 13: 27–35.

Grammig, J., and F. J. Peter. 2013. Telltale Tails: A New Approach to Estimating Unique Market Information Shares. *Journal of Financial and Quantitative Analysis* 48: 459–488.

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.

Harris, F. H. d., T. H. McInish, and R. A. Wood. 2002a. Security Price Adjustment across Exchanges: An Investigation of Common Factor Components for Dow Stocks. *Journal of Financial Markets* 5: 277–308.

Harris, F. H. d., T. H. McInish, and R. A. Wood. 2002b. Common Factor Components versus Information Shares: A Reply. *Journal of Financial Markets* 5: 341–348.

Hasbrouck, J. 1995. One Security, Many Markets: Determining the Contributions to Price Discovery. *The 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.

Johansen, S. 1988. Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control* 12: 231–254.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022

Johansen, S. 1991. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* 59: 1551–1580.

Lien, D., and K. Shrestha. 2009. A New Information Share. *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.

Lien, D., and Z. Wang. 2016. Estimation of Market Information Shares: A Comparison. *Journal of Futures Markets* 36: 1108–1124.

Paruolo, P. 1997. Asymptotic Inference on the Moving Average Impact Matrix in Cointegrated I(1) VAR Systems. *Econometric Theory* 13: 79–118.

Politis, D. N., and J. P. Romano. 1994. The Stationary Bootstrap. *Journal of the American Statistical Association* 89: 1303–1313.

Putniņš, T. J. 2013. What Do Price Discovery Metrics Really Measure? *Journal of Empirical Finance* 23: 68–83.

Roll, R. 1984. A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *The Journal of Finance* 39: 1127–1139.

Sapp, S. G. 2002. Price Leadership in the Spot Foreign Exchange Market. *The Journal of Financial and Quantitative Analysis* 37: 425–448.

Stock, J. H., and M. W. Watson. 1988. Testing for Common Trends. *Journal of the American Statistical Association* 83: 1097–1107.

Downloaded from https://academic.oup.com/jfec/article/19/5/934/5622754 by Erasmus University Rotterdam user on 11 January 2022