Bayesian Inference with Applications to Macroeconomics and Financial Market Price Discovery

Wenqiu Ma

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2026

Reading Committee:
Eric Zivot, Chair
Thomas Gilbert
Chang-Jin Kim

Program Authorized to Offer Degree:
Economics

University of Washington

Abstract

Bayesian Inference with Applications to Macroeconomics and Financial Market Price Discovery

Wenqiu Ma

Chair of the Supervisory Committee:
Eric Zivot
Department of Economics

This dissertation develops a unified Bayesian framework for analyzing and forecasting macroeconomic time series and for measuring price discovery in fragmented financial markets. Across three papers, it proposes flexible yet disciplined methodologies that exploit Bayesian shrinkage, stochastic volatility, and structural identification to handle large systems, noisy data, and time-varying dynamics while delivering finite-sample-relevant inference.

The first chapter introduces a large time-varying parameter vector autoregression (TVP-VAR) with stochastic volatility and informative priors that shrink the richly parameterized model toward a parsimonious benchmark. This Bayesian design stabilizes estimation in high dimensions and mitigates overfitting. Using U.S. macroeconomic data of varying dimensions and levels of aggregation, the chapter provides evidence that the proposed TVP-VAR improves forecast accuracy—both point and density—relative to standard benchmarks.

The second chapter develops a Bayesian vector error correction (BVECM) framework for structural analysis in cointegrated markets. By decomposing shocks into permanent and

transitory components, the framework yields structural price discovery measures that depend only on permanent innovations, abstracting from microstructure noise and temporary liquidity imbalances. Simulation studies and an empirical application to S&P 500 ETFs demonstrate robust, interpretable price discovery rankings with coherent uncertainty quantification.

The third chapter studies time-varying price discovery by embedding structural measures in a Bayesian order-invariant VAR with stochastic volatility. It shows that measures grounded in the permanent component are robust to time-varying volatility and noise, while conventional metrics can be distorted. Collectively, the three chapters demonstrate how Bayesian inference supports macroeconomic forecasting and the measurement of evolving information flows in modern financial markets.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Prof. Eric Zivot, for his unwavering support, guidance, and patience throughout my PhD journey. His insight, encouragement, and high standards have shaped not only this dissertation but also the way I think as a researcher. I am especially grateful for his help and advice during my job search, which has been invaluable for the next stage of my career.

I am also indebted to my committee members, Prof. Domenico Giannone, Prof. Chang-Jin Kim, Prof. Thomas Gilbert, and Prof. David F. Layton, for their thoughtful feedback, constructive suggestions, and support at every stage of this work.

I am deeply thankful to my cohort, as well as all the faculty and staff in the department, particularly Michelle Foshee, for creating a stimulating, supportive, and collegial environment in which to learn, teach, and grow. Their dedication and generosity have had a profound impact on my development as an economist and as a person.

My deepest thanks go to my parents and my family, whose unconditional love, understanding, and encouragement have sustained me through all the ups and downs. Their belief in me has been a constant source of strength. I am also profoundly grateful to my grandpa for his care, concern, and constant encouragement, and to my forever beloved grandma, whose passing has never diminished her presence in my life. She remains with me spiritually, and her love and example continue to inspire and comfort me throughout this journey.

I am truly grateful to my friends, especially Ruixuan Wan, Gusangyu Yin, Minhao Jiang, Liu Cao, Zihao Chen, Hao Yin, Mingfei Chen, Yanxiao Sun, Yuhua Nie, Xinrui Zheng, Guanping Li, Julian Lai, Shaoyu Liu, Xingbang Weng, Yi Cui, Aidan Beggs, and Yoon Choi, for walking with me through both the happiest and the most difficult times. Their companionship, humor, and emotional support made this journey not only possible, but also meaningful and memorable.

During my professional work, I have been fortunate to receive support and mentorship from many generous colleagues. I would like to especially thank my managers, mentors, and coworkers, including Domenico Giannone, Michiel De Pooter, Andrea Carriero, Leilei Zhang, Ashish Rajbhandari, Collin McCormack, Matthew Cocci, and Bartley Tablante, for their guidance, collaboration, and encouragement. Their insights and support have greatly enriched my research and helped me grow as an applied economist.

Finally, I would like to thank everyone—teachers, colleagues, collaborators, and friends—who has, in one way or another, contributed to this dissertation and to my growth during these years.

Contents

Chapter 1

Macroeconomic Forecasting with Large Time-Varying Parameter VARs

Wenqiu Ma

Abstract

Time-varying parameter VARs with stochastic volatility are flexible time series models for structural analysis and forecasting macroeconomic variables due to their ability to capture complex dynamic interrelationships. However, their dense parameterization can lead to huge computational burden, unstable inference and inaccurate out-of-sample forecasts, especially for large VAR models. To address this challenge, we propose a TVP-VAR model with informative priors to shrink the richly parameterized unrestricted model towards a parsimonious naïve benchmark, reducing estimation uncertainty. Through empirical analysis using U.S. datasets of various dimensions, we find compelling evidence that the proposed TVP-VAR model with informative priors outperforms many standard benchmarks in terms of forecast accuracy. Moreover, it demonstrates remarkable performance for level data.

1.1 Introduction

Vector autoregressive models (VARs) are standard benchmark for the analysis of dynamic economic problems, routinely used for forecasting and policy analysis today. "The vector autoregression specification is very general. It is capable of modeling arbitrarily well any covariance stationary stochastic process (Litterman et al., 1979)." VARs have provided many valuable insights in applied macroeconometrics, among which time-varying parameter vector autoregressions (TVP-VARs) developed by Cogley and Sargent (2001, 2005) and Primiceri (2005) have become the workhorse models in empirical macroeconomics.

TVP-VARs, particularly with heteroscedasticity, are flexible to allow for the time-varying macroeconomic processes, to better capture the dynamics of the underlying variables, including different forms of structural instabilities and the non-linear relationships between the dependent variables. Macroeconometricians have come to understand the importance, for a range of reasons, of allowing model parameters to evolve over time. For example, the US economy has undergone many structural changes in the past decades. Stock and Watson (2002) investigated the Great Moderation, in which period the volatility of economic activity, particularly real GDP growth, has declined sharply. Persistence of inflation increases during the 1970s, then falls in the 1980s and 1990s (Cogley and Sargent, 2001, 2005). In addition to these series-specific changes, many important shifts in the relationships between macroeconomic variables have been documented. For example, the performance Phillips curve forecasts is ahown to be episodic (Stock and Watson, 2008). Primiceri (2005) shows that the last 40 years have seen evident changes in both systematic and non-systematic monetary policy. Many researchers provided the evidence that TVP-VARs forecast substantially better than their homoscedastic or constant-coefficient counterparts (Clark, 2011; D'Agostino et al., 2013; Clark and Ravazzolo, 2015; Cross and Poon, 2016). These models have provided important ways to addressing a range of questions such as the transmission of monetary policy (Cogley and Sargent, 2001, 2005; Primiceri, 2005).

The literature has pointed out empirical and theoretical issues arising from using small VARs (Koop and Korobilis, 2013; Giannone et al., 2014; Carriero et al., 2015). In empirical work, however, the applications are mostly limited to modeling small systems involving only a few variables because of the computational burden and over-parameterization concerns. This feature

is maybe most evident in VARs that feature time-varying parameters and stochastic volatility, which are the main focus of this article. The number of free parameters increases quadratically with the number of variables in a system, and for even moderately-sized systems the model becomes highly overparameterized relative to the number of available observations.

There has been considerable interest in estimating higher-dimensional state space models, and more recently researchers have been developing methods to estimate larger systems of variables in VARs. Bańbura et al. (2010) demonstrated that large VARs with richer information have better forecast performance and more sensible impulse-response analysis. Subsequently, large VARs have been used in many empirical applications in both macroeconomics and finance (Bańbura et al., 2015; Carriero et al., 2018, 2019, 2012; Chan et al., 2020; Gefang, 2014; Giannone et al., 2014; Koop and Korobilis, 2016, 2019). However, the literature on large VARs with time-varying parameters remains relatively scarce (Chan, 2022).

These two streams of literature have naturally paved the way for the development of higher-dimensional and time-varying models. As the number of variables included in the TVP-VAR grows, such models quickly become overparameterized or overfitting which lead to suboptimal even poor estimation and inference. To mitigate the overparameterization issues, researchers have been working on reducing the dimensions to improve estimation.

Bayesian inference, by using priors, allows researchers to avoid overfitting the observed sample. A common practice is to focus on prior distributions that themselves depend on a substantially smaller number of parameters (which are commonly called hyperparameters). One prominent example that uses this approach is the "Minnesota" prior for VARs (Doan et al., 1984) which is especially useful in applications with many observable variables (Bańbura et al., 2010).

This paper investigates whether and how to employ informative priors will improve the performance of the time-varying coefficients VAR model with stochastic volatility in Primiceri (2005). The objective is to provide a flexible framework for the accurate estimation and interpretation of TVP-VARs. Specifically, Primiceri (2005) established the model on a small system with stationary data. In this paper, we adapt informative priors —Minnesota prior (Litterman et al., 1979; Litterman, 1980), Sum-of-coefficients prior (Doan et al., 1984), and Dummy-initial-observation prior (Sims, 1993)—instead of normal priors in Primiceri (2005) on the pre-sample to mitigate overparameterization issues. Besides, these priors help to fit the nonstationary pattern. The estimation

is done using Markov chain Monte Carlo (MCMC) methods.

This paper contributes to the literature along three dimensions. First, we use informative priors for TVP-VARs to reduce estimation uncertainty, namely the overparameterization or overfitting problem. Second, our estimation approach is exact and fully Bayesian, and the modeling framework is more flexible than many of those in earlier papers based on the approximations. Third, the new TVP-BVARs can be applied to variables in levels instead of stationary data, particularly with unit roots and cointegration issues.

1.2 Related Literature

This section reviews the efforts spent over the past several years to address the overfitting problem by adopting a Bayesian approach to specification and estimation of VARs.

Variable selection allows the data to select parameters to be set to zero. This gives a prior distribution for the parameter that is a discrete-continuous mixture: the prior involves equating the value of a binary indicator, $\gamma \in \{0, 1\}$, to the the restriction that a parameter is set to has a spike at zero with a continuous distribution elsewhere and for this reason is commonly called the spike-and-slab prior. The aim in using this prior is to empirically determine whether a parameter is exactly zero, or is drawn from a continuous distribution. In the context of a VAR, this variable selection technique can be used to automatically select lag length, to identify the set of variables to be included in each equation or to identify which variables have time-varying coefficients and which do not (Bitto and Frühwirth-Schnatter, 2019). Korobilis (2013) develops methods for automatic selection of variables in Bayesian vector autoregressions using the Gibbs sampler.

Frühwirth-Schnatter and Wagner (2010) use related approaches to let the data decide which parameters are time varying. The more general and earlier approach is that of Gerlach et al. (2000) in which the time-variation is allowed to switch on and off where necessary. To do this, they developed a dynamic mixture approach that could allow time variation of the states at some points in time but to turn off time variation at other times. This approach uses an indicator variable, γ_t, that could turn off the shocks to the states. A value of $\gamma_t = 0$ implies the state does not vary from time $t - 1$ to t.

Though these two approaches provide feasible solutions to the overparameterization issue of

large VARs, they are not consistent with macroeconometricians' perspective. For example, Lucas critique (Lucas Jr, 1976) states that the parameters of traditional macroeconometric models cannot be invariant with respect to changes in government policy variables. Therefore, considering DSGE models for policy analysis, it is not reasonable to rule out some variables or treat some variables as time-invariant over the time.

Shrinkage priors, on the other hand, do not set parameters to zero but place mass around zero and are continuous everywhere. The Bayesian approach to estimation regards the true population structure as uncertain and does not assign too much "weight" on any particular value of the model parameters (e.g., zero-restrictions on certain coefficients). Instead, it takes this uncertainty into account in the form of a prior probability distribution over the model parameters. The degree of uncertainty represented by this prior distribution can then be altered by the information contained in the data. Thus, it reduces the risk of overfitting and is computationally more efficient.

The choice of a prior distribution summarizing the researcher's uncertainty over the model parameters is a crucial step in specifying a Bayesian VAR. The Bayesian approach to estimation provides the way to combining sample information with prior information, while the range of alternative priors depends on the particular economic problem at hand. In this paper, we mainly discuss three priors in Section 1.4.1: Minnesota prior (Litterman et al., 1979; Litterman, 1980), Sum-of-coefficients prior (Doan et al., 1984), and Dummy-initial-observation prior (Sims, 1993).

1.3 Model

We choose the TVP-VAR in Primiceri (2005) as the baseline and reparameterize it for equation-by-equation representation following Carriero et al. (2019) and Carriero et al. (2022).

We first start with standard VAR with three variables and two lags:

$$\boldsymbol{y_t} = \boldsymbol{c} + \boldsymbol{B_1}\ \boldsymbol{y_{t-1}} + \boldsymbol{B_2}\ \boldsymbol{y_{t-2}} + \boldsymbol{u_t}, \qquad \boldsymbol{u_t} \sim \mathcal{N}(0, \boldsymbol{\Omega})$$

$\boldsymbol{y_t}$ is an 3×1 vector of observed endogenous variables; $\boldsymbol{c}$ is an 3×1 vector of coefficients that multiply constant terms; $\boldsymbol{B_1}$ and $\boldsymbol{B_2}$ are 3×3 matrices of time varying coefficients; $\boldsymbol{u_t}$ are heteroscedastic unobservable shocks with variance covariance matrix $\boldsymbol{\Omega}$. When both the coefficients $\boldsymbol{B_1}$, $\boldsymbol{B_2}$ and

variance covariance matrix $\boldsymbol{\Omega}$ are time-varying, we rewrite as

$$\boldsymbol{y_t} = \boldsymbol{c_t} + \boldsymbol{B_{1,t}}\, \boldsymbol{y_{t-1}} + \boldsymbol{B_{2,t}}\, \boldsymbol{y_{t-2}} + \boldsymbol{u_t}, \qquad \boldsymbol{u_t} \sim \mathcal{N}(0, \boldsymbol{\Omega_t})$$

Without loss of generality, consider the triangular reduction of $\boldsymbol{\Omega_t}$, defined by

$$\boldsymbol{A_t \Omega_t A_t}' = \boldsymbol{\Sigma_t \Sigma_t}'$$

where $\boldsymbol{A_t}$ is the lower triangular matrix

$$\boldsymbol{A_t} = \begin{bmatrix} 1 & 0 & 0 \\ \alpha_{21,t} & 1 & 0 \\ \alpha_{31,t} & \alpha_{32,t} & 1 \end{bmatrix},$$

and $\boldsymbol{\Sigma_t}$ is the diagonal matrix

$$\boldsymbol{\Sigma_t} = \begin{bmatrix} \sigma_{1,t} & 0 & 0 \\ 0 & \sigma_{2,t} & 0 \\ 0 & 0 & \sigma_{3,t} \end{bmatrix}.$$

It follows that

$$\boldsymbol{y_t} = \boldsymbol{c_t} + \boldsymbol{B_{1,t}}\, \boldsymbol{y_{t-1}} + \boldsymbol{B_{2,t}}\, \boldsymbol{y_{t-2}} + \boldsymbol{A_t^{-1}} \boldsymbol{\Sigma_t} \boldsymbol{\varepsilon_t},$$
$$V\left(\boldsymbol{\varepsilon_t}\right) = \boldsymbol{I_n}.$$

Reparameterize in the following structural form:

$$\boldsymbol{A_t y_t} = \dot{\boldsymbol{b}}_{\boldsymbol{t}} + \dot{\boldsymbol{B}}_{\boldsymbol{1,t}}\, \boldsymbol{y_{t-1}} + \dot{\boldsymbol{B}}_{\boldsymbol{2,t}}\, \boldsymbol{y_{t-2}} + \boldsymbol{v_t}, \qquad \boldsymbol{v_t} \sim \mathcal{N}(0, \boldsymbol{\Sigma_t \Sigma_t}')$$

where

$$\dot{\boldsymbol{b}}_{\boldsymbol{t}} = \boldsymbol{A_t c_t}, \quad \dot{\boldsymbol{B}}_{\boldsymbol{p,t}} = \boldsymbol{A_t B_{j,t}}, \quad j = 1, 2.$$

Then, the ith ($i = 1, 2, 3$) equation of the system can be rewritten as

$$
\begin{aligned}
y_{1,t} &= \dot{b}_{1,t} + \dot{\boldsymbol{B}}_{1,1,t}\, \boldsymbol{y}_{t-1} + \dot{\boldsymbol{B}}_{1,2,t}\, \boldsymbol{y}_{t-2} + v_{1,t}, \qquad v_{1,t} \sim \mathcal{N}(0, \sigma_{1,t}^2) \\
y_{2,t} &= \dot{b}_{2,t} + \dot{\boldsymbol{B}}_{2,1,t}\, \boldsymbol{y}_{t-1} + \dot{\boldsymbol{B}}_{2,2,t}\, \boldsymbol{y}_{t-2} - \alpha_{21,t} y_{1,t} + v_{2,t}, \qquad v_{2,t} \sim \mathcal{N}(0, \sigma_{2,t}^2) \\
y_{1,t} &= \dot{b}_{3,t} + \dot{\boldsymbol{B}}_{3,1,t}\, \boldsymbol{y}_{t-1} + \dot{\boldsymbol{B}}_{3,2,t}\, \boldsymbol{y}_{t-2} - \alpha_{31,t} y_{1,t} - \alpha_{32,t} y_{2,t} + v_{3,t}, \qquad v_{3,t} \sim \mathcal{N}(0, \sigma_{3,t}^2)
\end{aligned}
$$

where $\dot{\boldsymbol{B}}_{i,j,t}$ represent the ith row of $\dot{\boldsymbol{B}}_{j,t}$.

The main drawback of this representation, however, is that the implied reduced-form estimates depend on how the variables are ordered in the system. Carriero et al. (2019) and Chan (2022) investigated the extent to which these estimates depend on the ordering, and showed that the variability of the estimates is mild and comparable to that of the model in Primiceri (2005). There are a few recent papers that aim to develop order-invariant VARs with multivariate stochastic volatility, such as Shin and Zhong (2020); Arias et al. (2022); Chan et al. (2021). These models, however, are either designed for small TVP-VARs or they do not feature time-varying VAR coefficients. Therefore, we choose the TVP-VAR of Primiceri (2005) to be our benchmark, even though it is not order invariant, because it is generally viewed as the state-of-the-art and it is widely used as a reduced-form VAR for both forecasting (D'Agostino et al., 2013) and structural analysis using non-recursive identification schemes (Baumeister and Peersman, 2013).

Now consider the general TVP-VAR model with n observed endogenous variables

$$
\boldsymbol{y_t} = \boldsymbol{c_t} + \boldsymbol{B_{1,t}}\, \boldsymbol{y_{t-1}} + \cdots + \boldsymbol{B_{p,t}}\, \boldsymbol{y_{t-p}} + \boldsymbol{u_t}, \quad \boldsymbol{u_t} \sim \mathcal{N}(0, \boldsymbol{\Omega_t}) \tag{1.1}
$$

$\boldsymbol{y_t}$ is an $n \times 1$ vector of observed endogenous variables; $\boldsymbol{c}$ is an $n \times 1$ vector of coefficients that multiply constant terms; $\boldsymbol{B_j}, j = 1 \cdots p$ are $n \times n$ matrices of time varying coefficients; $\boldsymbol{u_t}$ are heteroscedastic unobservable shocks with variance covariance matrix $\boldsymbol{\Omega}$.

Similarly, we assume the time-varying covariance-matrix $\boldsymbol{\Omega_t}$ is decomposed into two pieces: time-varying covariance states in $\boldsymbol{A_t}$ and time-varying variance states in $\boldsymbol{\Sigma_t}$

$$
\boldsymbol{A_t \Omega_t A_t'} = \boldsymbol{\Sigma_t \Sigma_t'} \tag{1.2}
$$

$$\begin{aligned} \boldsymbol{y_t} &= \boldsymbol{c_t} + \boldsymbol{B_{1,t}}\ \boldsymbol{y_{t-1}} + \cdots + \boldsymbol{B_{p,t}}\ \boldsymbol{y_{t-p}} + \boldsymbol{A_t^{-1}}\boldsymbol{\Sigma_t}\boldsymbol{\varepsilon_t}, \\ V(\boldsymbol{\varepsilon_t}) &= \boldsymbol{I_n}. \end{aligned} \tag{1.3}$$

where similarly $\boldsymbol{A_t}$ is the lower triangular matrix

$$\boldsymbol{A_t} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \alpha_{21,t} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \alpha_{n1,t} & \cdots & \alpha_{nn-1,t} & 1 \end{bmatrix},$$

and $\boldsymbol{\Sigma_t}$ is the diagonal matrix

$$\boldsymbol{\Sigma_t} = \begin{bmatrix} \sigma_{1,t} & 0 & \cdots & 0 \\ 0 & \sigma_{2,t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{n,t} \end{bmatrix}.$$

Reparameterize in the following structural form:

$$\boldsymbol{A_t y_t} = \boldsymbol{\dot{b}_t} + \boldsymbol{\dot{B}_{1,t}}\ \boldsymbol{y_{t-1}} + \cdots + \boldsymbol{\dot{B}_{p,t}}\ \boldsymbol{y_{t-p}} + \boldsymbol{v_t}, \quad \boldsymbol{v_t} \sim \mathcal{N}(0, \boldsymbol{\Sigma_t}\boldsymbol{\Sigma_t}')$$

where

$$\boldsymbol{\dot{b}_t} = \boldsymbol{A_t c_t}, \quad \boldsymbol{\dot{B}_{p,t}} = \boldsymbol{A_t B_{j,t}}, \quad j = 1, \cdots, p.$$

Then, the ith ($i = 1, \cdots, n$) equation of the system can be rewritten as

$$\begin{aligned} y_{i,t} &= \dot{b}_{i,t} + \boldsymbol{\dot{B}}_{i,1,t}\ \boldsymbol{y}_{t-1} + \cdots + \boldsymbol{\dot{B}}_{i,p,t}\ \boldsymbol{y}_{t-p} - \alpha_{i1,t}y_{1,t} - \cdots - \alpha_{i(i-1),t}y_{2,t} + v_{i,t} \\ y_{i,t} &= \widetilde{\boldsymbol{x}}_t\boldsymbol{\beta}_{i,t} + \widetilde{\boldsymbol{w}}_{i,t}\boldsymbol{\alpha}_{i,t} + v_{i,t}, \quad v_{i,t} \sim \mathcal{N}(0, \sigma_{i,t}^2) \end{aligned} \tag{1.4}$$

where $\boldsymbol{\dot{B}}_{i,j,t}$ represent the ith row of $\boldsymbol{\dot{B}}_{j,t}$, $\boldsymbol{\beta}_{i,t} = \left(\dot{b}_{i,t}, \boldsymbol{\dot{B}}_{i,1,t}, \ldots, \boldsymbol{\dot{B}}_{i,p,t}\right)'$ is the intercept and VAR coefficients of the ith equation, $\boldsymbol{\alpha}_{i,t} = \left(\alpha_{i1,t}, \cdots, \alpha_{i(i-1),t}\right)'$ denotes the free elements in the ith row of the contemporaneous impact matrix A_t, $\widetilde{\boldsymbol{x}}_t = \left(1, \boldsymbol{y}'_{t-1}, \ldots, \boldsymbol{y}'_{t-p}\right)$, and $\widetilde{\boldsymbol{w}}_{i,t} = (-y_{1,t}, \ldots, -y_{i-1,t})$.

Let $\boldsymbol{x}_{i,t} = (\widetilde{\boldsymbol{x}}_t, \widetilde{\boldsymbol{w}}_{i,t})$, $\boldsymbol{\theta}_{i,t} = \left(\boldsymbol{\beta}'_{i,t}, \boldsymbol{\alpha}'_{i,t}\right)'$, we can further simplify the ith equation as

$$y_{i,t} = \boldsymbol{x}_{i,t}\boldsymbol{\theta}_{i,t} + v_{i,t}, \quad v_{i,t} \sim \mathcal{N}\left(0, \sigma^2_{i,t}\right) \tag{1.5}$$

Thus, we have rewritten the TVP-VAR in (1.1) as n unrelated regressions.

The model specifies a conditional-on-parameters (and lags) normal distribution for outcomes:

$$p\big(y_{i,t} \,|\, \boldsymbol{\theta}_{i,t}, \sigma_{i,t}; x_{i,t}\big) \sim \mathcal{N}\big(x_{i,t}\boldsymbol{\theta}_{i,t}, \sigma^2_{i,t}\big)$$

Next, we specify the dynamics of time-varying coefficients and stochastic volatility. Both the coefficients on lags $\boldsymbol{\beta}_{i,t}$ and the non-zero and non-one elements (free elements) of the matrix A_t, $\boldsymbol{\alpha}_{i,t}$, are assumed to evolve according to independent random walks:

$$\boldsymbol{\beta}_{i,t} = \boldsymbol{\beta}_{i,t-1} + \boldsymbol{v}_{i,t} \qquad \boldsymbol{v_{i,t}} \sim \mathcal{N}(0, \boldsymbol{Q}_i) \tag{1.6}$$

$$\boldsymbol{\alpha}_{i,t} = \boldsymbol{\alpha}_{i,t-1} + \boldsymbol{\zeta}_{i,t} \qquad \boldsymbol{\zeta}_{i,t} \sim \mathcal{N}(0, \boldsymbol{S}_i) \tag{1.7}$$

Therefore

$$\boldsymbol{\theta}_{i,t} = \boldsymbol{\theta}_{i,t-1} + \boldsymbol{\eta}_{i,t} \qquad \boldsymbol{\eta_{i,t}} \sim \mathcal{N}(0, \boldsymbol{R}_i) \tag{1.8}$$

where $\boldsymbol{R_i} = \text{diag}(\boldsymbol{Q_i}, \boldsymbol{S_i})$.

Since $\boldsymbol{u_t} = \boldsymbol{A_t^{-1}}\boldsymbol{\Sigma_t}\boldsymbol{\varepsilon_t}$, the diagonality of the matrix $\boldsymbol{\Sigma_t}$ implies that the generic jth element of the rescaled VAR disturbances $\tilde{\boldsymbol{u}}_t = \boldsymbol{A}_t\boldsymbol{u}_t$ is $\tilde{u}_{j,t} = \sigma_{j,t}\varepsilon_{j,t}$. Taking logs of $\tilde{u}_{j,t}$ yields the following set of observation equations:

$$\log \tilde{u}_{j,t} = \log \sigma_{j,t} + \log \varepsilon_{j,t}, \quad j = 1, \ldots, n$$

The model is completed by specifying laws of motion for the unobserved states:

$$\log \sigma_{j,t} = \log \sigma_{j,t-1} + \omega_{j,t}, \quad j = 1, \ldots, n$$

Thus, the standard deviations $\sigma_{i,t}$ are assumed to evolve as geometric random walks:

$$\log \sigma_{i,t} = \log \sigma_{i,t-1} + \omega_{i,t} \qquad \eta_t \sim \mathcal{N}(0, W_i) \tag{1.9}$$

As mentioned by Primiceri (2005), one could adopt more general processes rather than a random walk in (C.3)-(C.5), but this comes at the huge cost of additional model parameters.

Note that the innovations in (C.3)-(C.5) are commonly assumed to be independent across equations in literature. This assumption is partly motivated by the concern of proliferation of correlation parameters, especially when n is large. Carriero et al. (2019) show that the computational cost of adding correlations across coefficients belonging to different equations is negligible as they need to be computed only once outside the main MCMC sampler. Therefore, the following section 1.4.1 introduces the generalized priors allowing correlations between equations.

1.4 Priors and Bayesian Estimation

1.4.1 Priors

We estimate the model using the Bayesian VAR (BVAR) approach which helps to overcome the curse of dimensionality via the imposition of prior beliefs on the parameters.

The first T_0 observations are defined as the pre-sample, which are used to calibrate the prior distributions over the initial states and the innovation variances for $\boldsymbol{\theta}_{i,t}$ and the standard errors $\sigma_{i,t}$.

We modify the normal (flat) priors used in Primiceri (2005) to calibrate the prior distributions. First, we use a prior often referred to as the *Minnesota prior* developed in Litterman et al. (1979); Litterman (1986) with modifications proposed by Kadiyala and Karlsson (1997) and Sims and Zha (1998). The basic principle behind it is that all the equations are centered around the random walk, possibly with drift, which is a parsimonious yet "reasonable approximation of the behavior of an economic variable" (Litterman et al., 1979); i.e., the prior mean can be associated with the following representation for y_t:

$$y_t = c + y_{t-1} + u_t$$

This amounts to shrinking the diagonal elements of B_1 toward one and the remaining coefficients in $B_2, \ldots, B_p$ toward zero. In addition, the prior specification incorporates the belief that the more recent lags should provide more reliable information than the more distant ones and that own lags should explain more of the variation of a given variable than the lags of other variables in the equation. The coefficients associated with the same variable and lag in different equations can be correlated. These prior beliefs are imposed by setting the following moments for the prior distribution of the coefficients:

$$\beta_0 \sim \mathcal{N}(a_\beta, V_\beta)$$

where

$$a_\beta := E[(\dot{B}_s)_{ij}] = \begin{cases} 1 & i = j \text{ and } s = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$V_\beta := Cov\big((\dot{B}_s)_{ij}, (\dot{B}_r)_{hm}\big) = \begin{cases} \lambda^2 \frac{1}{s^2} \frac{\Sigma_{ih}}{\Psi_{jj}/(d-n-1)} & m = j \text{ and } r = s \\ 0 & \text{otherwise} \end{cases}$$

For the prior on the covariance matrix of the errors, Σ, we set the degrees of freedom $d = n + 2$, which is the minimum value that guarantees the existence of the prior mean Ψ, which is diagonal. Σ denotes the sample variance of the residuals from regressing (1.1) in presample, and Ψ can be treated as a hyperparameter as Giannone et al. (2015).

The key hyperparameter λ controls the scale of all of the variances and covariances, and effectively determines the overall tightness of this prior. For $\lambda \to 0$ the posterior equals the prior and the data do not influence the estimates. If $\lambda \to \infty$, on the other hand, posterior expectations coincide with the ordinary least squares (OLS) estimates. $\frac{1}{s^2}$ is the rate at which prior variance decreases with increasing lag length, and $\frac{\Sigma_{ih}}{\Psi_{jj}/(d-n-1)}$ accounts for the different scale and variability of the data.

The literature following Litterman's work introduced refinements of the Minnesota prior to further "favor unit roots and cointegration, which fits the beliefs reflected in the practices of many

applied macroeconomists" (Sims and Zha, 1998). We introduce two more priors to improve the forecasting performance in the case of level data.

Rewrite the VAR (1.1) equation in an error correction form:

$$\Delta y_t = c + \Pi y_{t-1} + \tilde{B}_1 \Delta y_{t-1} + \cdots + \tilde{B}_{\tilde{p}} \Delta y_{t-\tilde{p}} + \varepsilon_t$$

where $\tilde{p} = p - 1, B_s = -A_{s+1} - \cdots - A_p, s = 1, \ldots, \tilde{p}$ and $\Pi = A_1 + \cdots + A_p - I_n$.

A VAR in first differences implies the restriction $\Pi = 0$ (or $B_1 + \cdots + B_p = I_n$). We follow Doan et al. (1984) and set a prior known as *sum-of-coefficient prior* that shrinks Π to zero. Specifically, we set a prior that is centered at 1 for the sum of coefficients on the own lags for each variable, and at 0 for the sum of coefficients on other variables' lags. This prior introduces correlations among the coefficients on each variable in each equation. It is implemented using Theil mixed estimation, with a set of n artificial observation. More precisely, we construct the following set of dummy observations:

$$\begin{aligned} y^+_{n \times n} &= \operatorname{diag}\left(\frac{\bar{y}_0}{\mu}\right) \\ \underset{n \times (1+np)}{x^+} &= [\underset{n \times 1}{0}, y^+, \ldots, y^+] \end{aligned}$$

where $\bar{y}_0$ is an $n \times 1$ vector containing the average of the first p observations for each variable. These artificial observations are added on top of the data matrices $y \equiv [y_{p+1}, \ldots, y_T]'$ and $x \equiv [x_{p+1}, \ldots, x_T]'$, which are then used for inference. The hyperparameter μ controls the variance of these prior beliefs: as $\mu \to \infty$, the prior becomes uninformative, while $\mu \to 0$ means the case of exact differencing, which implies the presence of a unit root in each equation and rules out cointegration.

Following Sims (1993) and Sims and Zha (1998), we complement such "inexact differencing" with an additional prior, known as the *dummy-initial-observation prior*. This involves prior correlations among coefficients of the same equations, but not across equations. It is implemented

using the following dummy observation:

$$y^{++}_{1\times n} = \frac{y_0'}{\delta}$$
$$x^{++}_{1\times(1+np)} = \left[\frac{1}{\delta}, y^{++}, \ldots, y^{++}\right]$$

which states that a no-change forecast for all variables is a good forecast at the beginning of the sample. The hyperparameter δ controls the tightness of the prior implied by this artificial observation. As $\delta \to \infty$, the prior becomes uninformative. And, as $\delta \to 0$, all the variables of the VAR are forced to be at their unconditional mean, or the system is characterized by the presence of an unspecified number of unit roots without drift, which is consistent with cointegration.

In the similar way as β_0, the coefficients associated with the same variable in different equations can be correlated. A reasonable prior for α_0 can be obtained:

$$\alpha_0 \sim \mathcal{N}(a_\alpha, V_\alpha)$$

where

$$a_\alpha = 0$$

$$V_\alpha := Cov\big((A)_{ij}, (A)_{hm}\big) = \begin{cases} \lambda^2 \frac{\Sigma_{ih}}{\Psi_{jj}/(d-n-1)} & m = j \\ 0 & \text{otherwise} \end{cases}$$

For $log\sigma_0$, the mean of the distribution is chosen to be the logarithm of the point estimates of the standard errors using $\boldsymbol{\theta}_0$, while the variance covariance matrix is arbitrarily assumed to be the identity matrix. Finally, degrees of freedom and scale matrices are needed for the inverse-Wishart prior distributions of the hyperparameters. The degrees of freedom for $\boldsymbol{R}_i$ and W_i are set to one plus the dimension of each matrix.

To summarize, for the ith equation, $i = 1, \cdots, n$, the priors take the forms:

$$\beta_{i,0} \sim \mathcal{N}(a_\beta, V_\beta) \tag{1.10}$$

$$\alpha_{i,0} \sim \mathcal{N}(a_\alpha, V_\alpha) \tag{1.11}$$

$$\log \sigma_{i,0} \sim \mathcal{N}(a_{\log \sigma_i}, V_{\log \sigma_i}) \tag{1.12}$$

$$\boldsymbol{Q}_i \sim \mathcal{IW}\left(k_Q^2 \cdot V_\beta, d_{Q,i}\right) \tag{1.13}$$

$$\boldsymbol{S}_i \sim \mathcal{IW}\left(k_S^2 \cdot V_\alpha, d_{S,i}\right) \tag{1.14}$$

$$W_i \sim \mathcal{IW}\left(k_W^2, d_{W_i}\right) \tag{1.15}$$

The setting of the prior $\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0$ depends on the hyperparameters $\boldsymbol{\kappa} = (\lambda, \mu, \delta, \Psi)$. As hyperpriors for λ, μ and δ, we choose gamma densities with mode equal to 0.2, 1, and 1, the values recommended by Sims and Zha (1998), and standard deviations equal to 0.4, 1, and 1, respectively. For the hyperprior for each element of Ψ, We pick an inverse-Gamma with scale and shape equal to $(0.02)^2$ as in Giannone et al. (2015). The tuning parameters are k_R, k_W, which determine the scales of $\mathcal{IW}$ distribution. We choose $k_W = 0.01$, $k_S = 0.1$. $k_Q = 0.01$ if it is associated with a VAR coefficient and 0.1 for an intercept.

1.4.2 Estimation Algorithm

One can simulate from the joint posterior distribution using the following posterior sampler. After reparameterizing the TVP-VAR, we can adopt the efficient simulation algorithm for state space models in Chan and Jeliazkov (2009). For each step, we leave the details to Appendix B.

1. $p\left(\boldsymbol{\beta}_i, \boldsymbol{\alpha}_i \lfloor \boldsymbol{y}, \boldsymbol{\sigma}, \boldsymbol{Q}, \boldsymbol{S}, \boldsymbol{W}, \boldsymbol{\kappa}\right), i = 1, \ldots, n;$

2. $p\left(\boldsymbol{\sigma}_i \mid \boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{Q}, \boldsymbol{S}, \boldsymbol{W}, \boldsymbol{\kappa}\right), i = 1, \ldots, n;$

3. $p\left(\boldsymbol{Q}_i, \boldsymbol{S}_i, W_i \mid \boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\sigma}, \boldsymbol{\kappa}\right), i = 1, \ldots, n;$

4. $p\left(\boldsymbol{\beta}_{i,0}, \boldsymbol{\alpha}_{i,0} \mid \boldsymbol{y}, \boldsymbol{\sigma}, \boldsymbol{Q}, \boldsymbol{S}, \boldsymbol{W}, \boldsymbol{\kappa}\right), i = 1, \ldots, n;$

5. $p\left(\boldsymbol{\kappa} \mid \boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\sigma}, \boldsymbol{Q}, \boldsymbol{S}, \boldsymbol{W}\right).$

Note that if we treat $\boldsymbol{\kappa} = (\lambda, \mu, \delta, \Psi)$ as predetermined values (density means above) instead of hyperparameters, then step 4 and 5 are not necessary.

1.5 Forecasting Evaluation

The U.S. dataset for our empirical application consists of 21 quarterly variables with a sample period from 1959Q1 to 2018Q4. It is sourced from the FRED-QD database at the Federal Reserve Bank of St. Louis as described in McCracken and Ng (2020) [1]. Our dataset contains a variety of standard macroeconomic and financial variables, such as Real GDP, industrial production, inflation rates, labor market variables, money supply and interest rates. The variables enter the models in annualized log levels (i.e., we take logs and multiply by 4), except those already defined in terms of annualized rates, such as interest rates, which are taken in levels. The number of lags (p) in all the VARs is set to two, and T_0 is four. Further details on the database are reported in Table 3.1. We work with three different VAR models, including progressively larger sets of variables:

1. A Small-scale model—the prototypical monetary VAR—with three variables, i.e. GDP, the GDP deflator, and the federal funds rate.

2. A Medium-scale model, which includes the variables used for the estimation of the DSGE model of Smets and Wouters (2007) for the U.S. economy. In other words, we add consumption, investment, hours worked, and wages to the small model.

3. A Large-scale model, with 21 variables, using a data set that nests the previous two specifications and also includes a number of important additional labor market, financial, and monetary variables.

Note that federal funds rates are ordered last in the VAR. This is not an ordering issue but an identification condition, which is completely standard in most of the existing literature.

The time series plots in Figure 1.1 convey well-known features of the medium-scale series of interest, and Figure 1.2 provides standard deviations computed on rolling windows of 60 quarters

[1] https://research.stlouisfed.org/econ/mccracken/fred-databases/

Table 1.1: Description of the Database

Variables	Mnemonic	Transformations
Real GDP	GDPC1	$4 \cdot \log s$
GDP deflator	GDPCTPI	$4 \cdot \log s$
Federal funds rate	FEDFUNDS	raw
Consumer price index	CPIAUCSL	$4 \cdot \log s$
Commodity price	PPIACO	$4 \cdot \log s$
Industrial production	INDPRO	$4 \cdot \log s$
Employment	CE16OV	$4 \cdot \log s$
Employment in the services sector	SRVPRD	$4 \cdot \log s$
Real consumption	PCECC96	$4 \cdot \log s$
Real investment	GPDIC1	$4 \cdot \log s$
Personal consumption Expenditures	PCECTPI	$4 \cdot \log s$
Gross private domestic investment	GPDICTPI	$4 \cdot \log s$
Capacity utilization	TCU	raw
Consumer expectations	UMCSENTx	raw
Hours worked	HOABS	$4 \cdot \log s$
Real compensation per hours	RCPHBS	$4 \cdot \log s$
One-year bond rate	GS1	raw
Ten-years bond rate	5GS10	raw
SP500	S&P500	$4 \cdot \log s$
Effective exchange rate	TWEXAFEGSMTHx	$4 \cdot \log s$
M2	M2REAL	$4 \cdot \log s$

and it shows that the movements in volatility have been gradual and substantial. In particular, these series seem to be affected by secular forces, especially the Great Moderation (Stock and Watson, 2002) taking place from the mid 1980s onwards and the Great Financial Crisis from 2007Q4.

Using each of these three datasets, we produce the VAR forecasts recursively for two horizons (one and four quarters), starting with the estimation sample that ranges from 1959Q1 to 1990Q4. The procedure consists of generating the forecasts using the predictive density with the same information available to the econometrician at the time the forecasts are made. More precisely, using data from 1959Q1 to 1990Q4, we generate draws from the posterior predictive density of the model for 1991Q1 (one quarter ahead) and 1991Q4 (one year ahead). This procedure is then repeated by updating the estimation sample, one quarter at a time, using all the available vintages. Predictions are compared with ex post realized data vintages.

This section begins from the evolving patterns using estimation of the time-varying parameters.

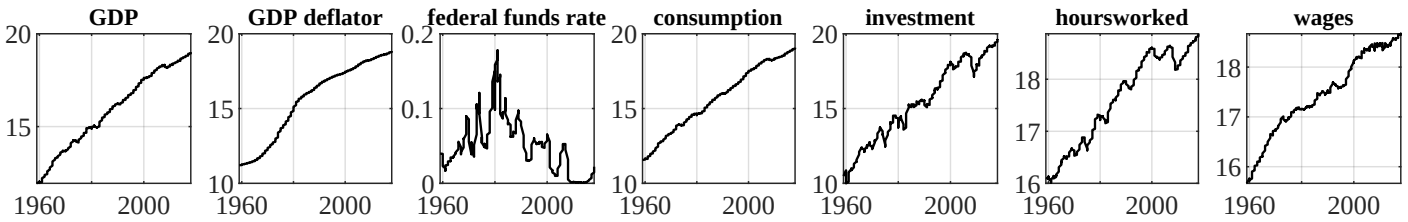

Figure 1.1: Series being predicted

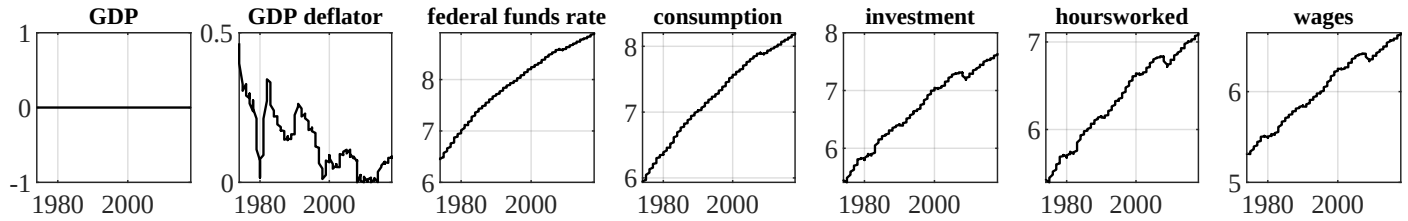

Figure 1.2: Standard deviations of the series being predicted. Computations based on rolling window of 60 observations

We start to assess the accuracy of our models in terms of point forecasts, defined as the mean squared forecast error (MSFE) at each point in time. We then turn to the evaluation of the density forecasts to assess how accurately different models capture the uncertainty around the point forecasts. The basis of comparison used is the predictive log scoring rule—the model's probability density for the return at the relevant horizon before it is observed, evaluated at the actual value of the return after it is observed. The basis of evaluation in this study is the probability integral transformation (PIT), which is the inverse of the sequence of *ex ante* predictive cumulative distribution functions (CDFs) evaluated on the sequence of actual returns *ex post*.

We compare the forecasts obtained with the TVP-VAR with those obtained using different standard forecasting models. We consider nine models in total:

1. A time-invariant VAR with Normal prior, estimated by VAR(p) (we refer to this model as VAR or VAR-OLS).

2. A time-invariant VAR with Minnesota prior (referred to as VAR-MN), and one by adding sum-of-coefficient prior and dummy-initial-observation prior (referred to as VAR-MSD). This is the same model as in Giannone et al. (2015).

3. A time-varying coefficients VAR (TV-VAR) with time-invariant volatility. We apply either Normal prior (referred to as TV-VAR-OLS), or Minnesota prior (referred to as TV-VAR-

MN), or along with sum-of-coefficient prior and dummy-initial-observation prior (referred to as TV-VAR-MSD).

4. A time-varying coefficients VAR with stochastic volatility (TV-SV-VAR). When we apply Normal prior to coefficients (referred to as TV-SV-VAR-OLS), it is consistant with the TVP-VAR in Primiceri (2005). All the above models are compared with the time-varying coefficients VAR with stochastic volatility, applying Minnesota prior (referred to as TV-SV-VAR-MN), and along with sum-of-coefficient prior and dummy-initial-observation prior (referred to as TV-SV-VAR-MSD).

1.5.1 How Much Time Variation?

As a preliminary step of our analysis, in order to understand whether time variation is an important characteristic of the dataset, we estimate the TVP-VAR model over the whole sample in the case of medium-scale model.

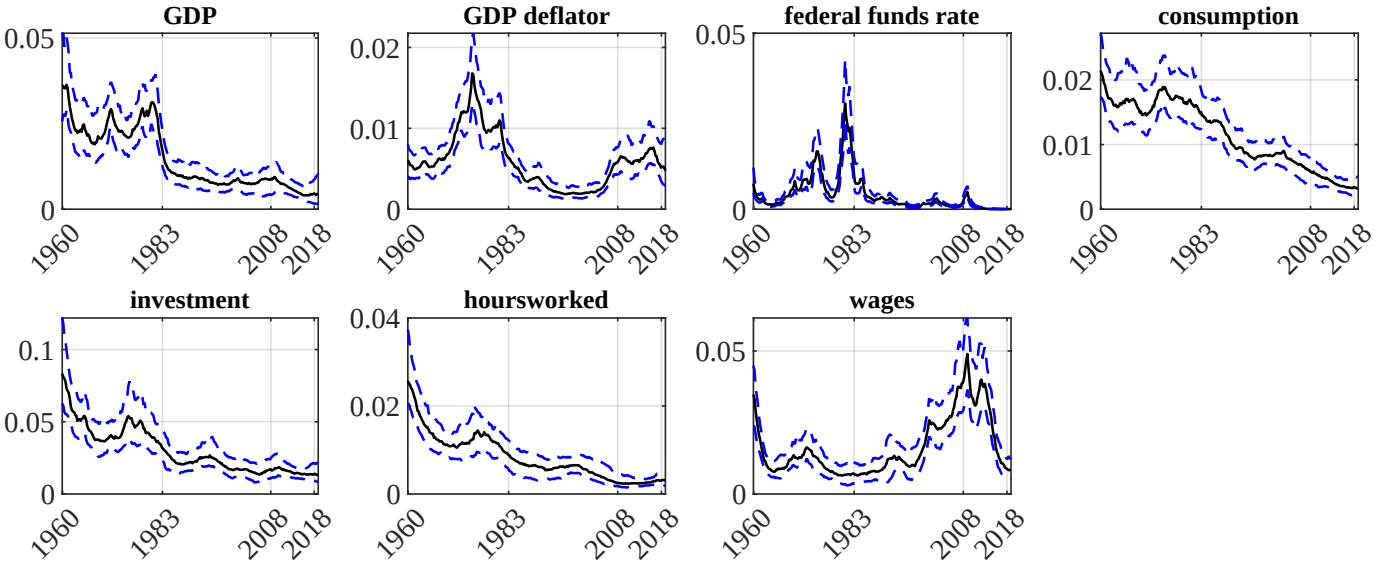

Figure 1.3: Estimation of volatility $\sigma_{i,t}$. Solid lines are the posterior median; dashed lines are the 84% confidence bands.

Figure 1.3 shows the evolution of the standard deviations of the residuals. The solid lines are the posterior medians of the parameters and the dashed lines are the 84% confidence bands. All the volatilities exhibit substantial time variation over the sample. The figure also shows that,

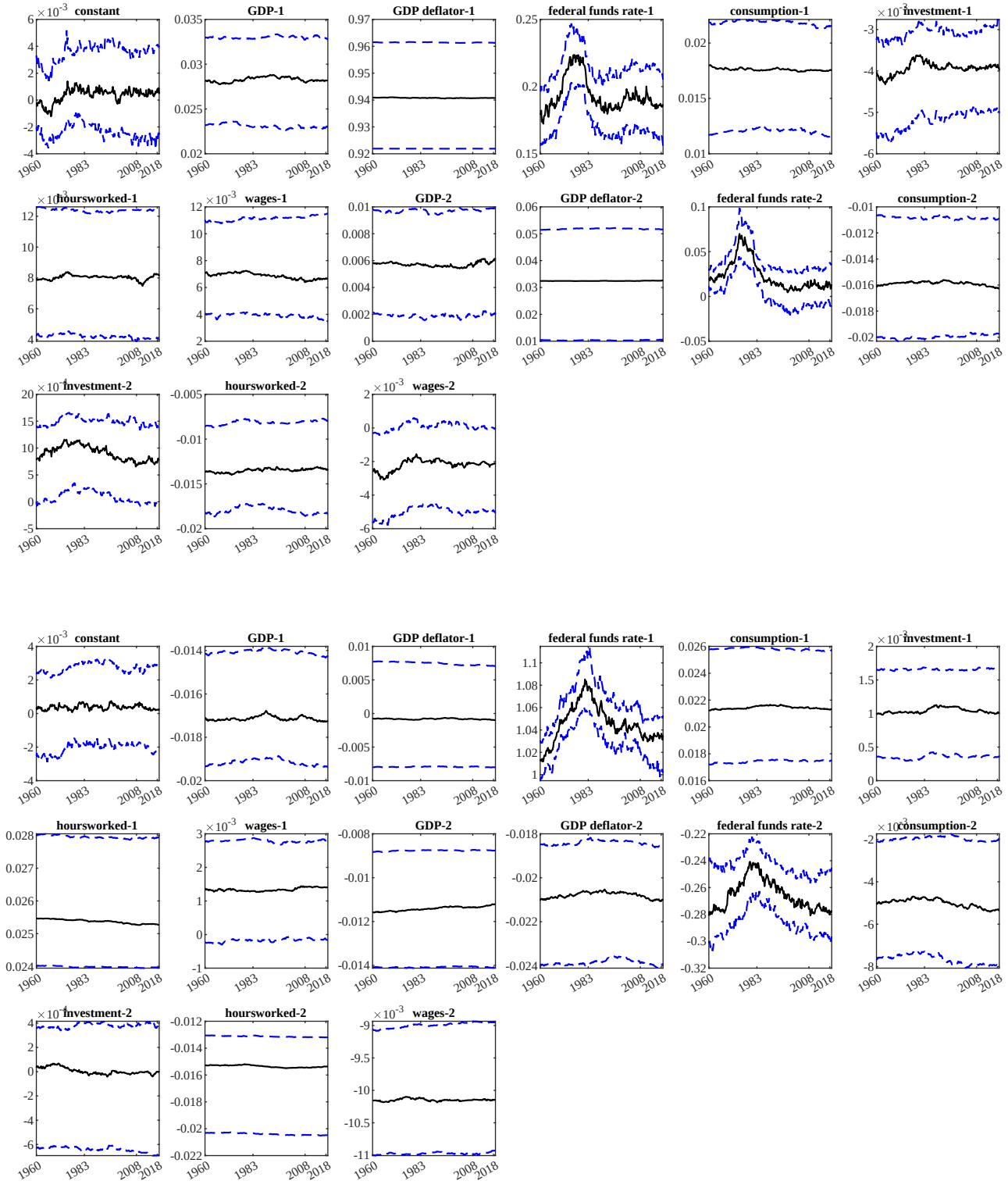

Figure 1.4: Time-varying parameters $B_{i,t}$ of GDP deflator (top) and federal funds rate (bottom).Solid lines are the posterior median; dashed lines are the 68% confidence bands.

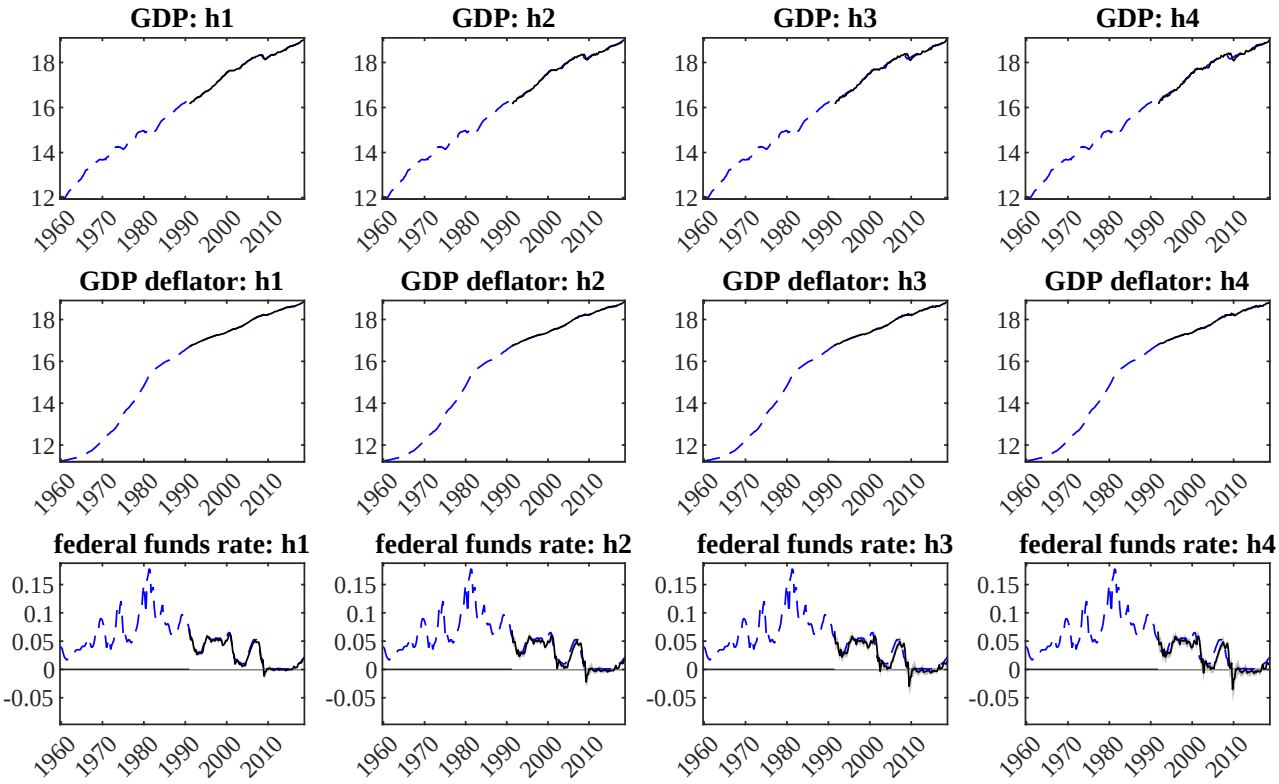

Figure 1.5: Forecasting Distributions. The figure shows the time series evolution of the predicted distribution from one-quarter-ahead (left) to four-quarters-ahead (right). The dashed lines are realized values; solid lines are median predicted values.

concomitant with the great moderation period around the mid 1980s, there is a sharp drop in the volatilities, while some variables have a peak of their volatilaties along with the financial crisis around 2008. Figure 1.4 shows the evolution of the coefficients over time. The solid lines are the posterior medians of the coefficients and the dashed lines are the 68% confidence bands. Many coefficients display constant patterns, while two parameters—federal funds rate and investment—are characterized by remarkable fluctuations over time. All in all, the results show that time variation is an important feature of the parameters.

1.5.2 Forecast Accuracy: Point Forecast

To preliminarily evaluate the forecast accuracy, Figure 1.5 shows one- and four-quarter-ahead out-of-sample forecasting distributions with the conditional medians and the conditional 5, 25, 75, and 95 percent quantiles.

We start by assessing the accuracy of our models in terms of point forecasts, defined as the

median of the predictive density density at each point in time. For each variable, the target of our evaluation is defined in terms of the h-period annualized average growth rates, $z_{i,t+h}^{h} = \frac{1}{h}\left[y_{i,t+h} - y_{i,t}\right]$.

The metric used to evaluate the point forecasts for variable i with horizon h is the mean squared forecast error (MSFE) defined as

$$\text{MSFE}_{i,h} = \frac{\sum_{t=t_0}^{T-m}\left(z_{i,t+h}^{0} - \mathbb{E}\left(z_{i,t+h} \mid \boldsymbol{y}_{1:t}\right)\right)^2}{T-h-T_0+1},$$

where $z_{i,t+h}^{0}$ is the actual observed value of $z_{i,t+h}$. For MSFE, a smaller value indicates better forecast performance.

Table A.3 analyzes the accuracy of point forecasts by reporting the MSFE of real GDP, the GDP deflator, and the federal funds rate. To facilitate the comparison between various models, the results are reported in terms of relative MSFE (RMSFE), i.e. the ratio between the MSFE of a particular model and the MSFE of the TV-SV-VAR-MN model, used as benchmark. When the relative MSFE is *larger* than one, the forecasts of the benchmark model are, on average, more accurate that those produced with a given model. The VAR forecasts with informative priors are systematically more accurate than the flat-prior VAR forecasts for almost all the variables and horizons that we consider. Exceptionally, the comparison to the VAR model with no time variation is not always favorable to the BVARs especially in small scale, which means we don't benefit as much as in large VARs.The results from several alternative informative priors, however, are mixing, which means we have to carefully select priors case by case. In addition, the BVAR models with only time varying coefficients have similar performance with those adding stochastic volatility, which means that the coefficients provide more time variation than volatilities. Overall, the TVP-VAR produces very accurate forecasts for all the variables and, on average, performs better than any other model considered. In particular it outperforms the VARs with Normal priors for all the variables at all horizons, except for federal funds rates at the horizon of one quarter.

Table 1.2: Point Forecasts: Relative MSFE

Horizons	Variables	VAR-OLS	VAR-MN	VAR-MSD	TV-VAR-OLS	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-OLS	TV-SV-VAR-MN	TV-SV-VAR-MSD
A. Small-scale model										
One quarter	Real GDP	0.9544	0.7442	1.0670	1.3107	1.0842	1.8932	0.9259	1.8979	1.0220
	GDP deflator	1.0178	0.9443	1.0466	1.3091	1.0087	1.4871	1.3306	0.2377	1.0979
	Federal funds rates	1.5465	0.9275	1.1766	1.2855	1.0538	0.7874	0.8878	0.0322	1.4586
One year	Real GDP	1.0336	0.7682	1.3495	2.2314	1.0653	1.7242	2.4824	1.1434	1.4426
	GDP deflator	2.5745	1.1742	1.6108	5.0673	1.0067	2.7533	2.5400	0.1535	1.1467
	Federal funds rates	2.7456	2.2170	2.1306	1.1345	1.0212	2.0048	1.0695	0.0173	2.7485
B. Medium-scale model										
Horizons	Variables	VAR-OLS	VAR-MN	VAR-MSD	TV-VAR-OLS	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-OLS	TV-SV-VAR-MN	TV-SV-VAR-MSD
One quarter	Real GDP	1.1864	1.1942	1.1286	1.1639	1.2765	1.1770	1.1712	1.4799	1.1856
	GDP deflator	1.6449	0.9177	0.8747	0.9500	1.1270	0.9168	1.1943	0.3043	0.8025
	Federal funds rates	1.6341	1.1965	0.6619	1.0464	1.0262	0.8987	1.3905	0.0532	0.8006
One year	Real GDP	1.4164	1.0440	0.8660	1.1506	1.0947	0.8840	0.9735	1.2409	0.8970
	GDP deflator	1.7737	1.0712	1.3283	1.3409	0.9694	1.2583	1.0707	0.2276	1.1533
	Federal funds rates	1.4867	1.1582	0.8976	1.2548	1.3326	0.9312	1.2833	0.0478	0.9528
C. Large-scale model										
Horizons	Variables	VAR-OLS	VAR-MN	VAR-MSD	TV-VAR-OLS	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-OLS	TV-SV-VAR-MN	TV-SV-VAR-MSD
One quarter	Real GDP	2.3688	1.0836	0.9187	1.0949	0.9987	0.7034	1.0931	1.6591	0.8022
	GDP deflator	1.1821	1.0042	1.0630	1.0814	0.8811	0.6233	1.0761	0.4021	0.6703
	Federal funds rates	1.4087	1.0740	1.1182	1.1403	1.1247	1.2857	1.1234	0.0609	1.2611
One year	Real GDP	1.1956	1.0429	0.4708	0.9615	0.8146	0.5556	0.9727	1.5093	0.5141
	GDP deflator	1.4167	1.0634	1.0629	1.0863	1.0093	0.5751	1.0841	0.5646	0.5979
	Federal funds rates	1.6387	1.0934	1.0324	1.0774	1.0691	0.3817	1.0312	0.0986	0.3667

Note: The table reports the results relative to the forecasting accuracy using point forecasts. The variables we forecast are real GDP, GDP deflator and federal funds rates. For the TV-SV-VAR-MN model we report the mean square forecast error (MSFE). For the other models we report the relative mean square forecast error (RMSFE), i.e. the ratio of the MSFE of a particular model to the MSFE of the baseline model.

1.5.3 Forecast Accuracy: Density Forecast

The point forecast evaluation is a useful tool to discriminate among models, but it disregards the uncertainty assigned by each model to its point prediction. For this reason, we evaluate the density forecasts using two metrics. The first we use is the log-predictive score, which is the logarithm of the value of the predictive density evaluated at the true realized value of the time series. A high value is obtained when high probability is assigned to the actual outcome. Specifically, the log-predictive score for variable i with horizon h is defined as

$$\text{Logscore}_{i,h} = \frac{1}{T-h-T_0+1} \sum_{t=T_0}^{T-h} \log p\left(z_{i,t+m} = z_{i,t+m}^0 \mid \boldsymbol{y}_{1:t}\right),$$

where we measure the log-predictive score using a gaussian approximation of the predictive density for all models.

Table A.4 reports the average difference between the log predictive scores of the TV-SV-VAR-MN model and the competing models, for each variable and horizon. A *positive* number indicates that the density forecasts produced by our proposed procedure are more accurate than those of the alternative models. In addition, the HAC estimate of its standard deviation (in parentheses) gives a rough idea of the statistical significance and the volatility of this difference. Table A.4 makes clear that the BVAR generally forecasts outperform those with a Normal prior at all scales also when evaluating the whole density. Again, no particular one informative prior outperforms the other priors. In general, the TV-SV-VAR with informative priors tends to significantly dominate all the other models, for all the horizons and all the variables.

Figure 1.6 compares the log predictive scores of the TVP-VAR model with flat or informative priors, i.e., TV-SV-VAR-OLS and TV-SV-VAR-MSD model over time in terms of medium-scale data. In particular, the model with informative priors (solid lines) outperforms the one with a Normal prior (dashed lines) for GDP deflator and federal funds rate during the financial crisis around 2008.

Next, to gauge the quality of the density intervals shown in Figure 1.5 we perform probability integral transforms (PITs) tests. The PITs are the value of the predictive cumulative distribution evaluated at the true realized values of the time series. PITs are widely used to assess the calibration of density forecasts (Jore et al., 2010; Geweke and Amisano, 2010; Clark, 2011; Amisano

Table 1.3: Density Forecasts: log scores

Horizons	Variables	VAR-OLS	VAR-MN	VAR-MSD	TV-VAR-OLS	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-OLS	TV-SV-VAR-MSD
A. Small-scale model									
One quarter	Real GDP	0.5322	0.0006	0.0160	0.0044	0.0055	0.0105	0.0083	0.0159
		(0.9844)	(1.3972)	(1.3949)	(1.4043)	(1.4145)	(1.3983)	(1.3926)	(1.4019)
	GDP deflator	-0.5242	0.0131	0.0050	0.0008	0.0212	-0.0030	0.0098	-0.0072
		(1.0410)	(1.4878)	(1.4726)	(1.4695)	(1.5099)	(1.4671)	(1.4861)	(1.4691)
	Federal funds rates	0.5559	-0.0099	-0.0272	0.0044	-0.0068	-0.0291	0.0013	-0.0305
		(0.8471)	(1.1895)	(1.1806)	(1.1965)	(1.2022)	(1.1821)	(1.1970)	(1.1851)
One year	Real GDP	-0.2330	0.0250	0.9936	0.0072	0.0633	0.9651	0.0440	0.9576
		(5.7027)	(8.0575)	(8.8203)	(7.9636)	(8.1474)	(8.7490)	(8.0048)	(8.7676)
	GDP deflator	-0.1120	0.2691	-1.0024	0.1624	0.3083	-0.7050	0.1566	-0.9755
		(22.2384)	(31.8175)	(31.6246)	(31.4293)	(31.8783)	(31.5999)	(31.5677)	(31.8332)
	Federal funds rates	-0.2575	0.0325	-0.3079	0.0137	0.1031	-0.2030	0.0100	-0.1419
		(4.9521)	(7.0694)	(6.8838)	(7.0802)	(7.1458)	(6.9662)	(7.0626)	(7.0576)
B. Medium-scale model									
One quarter	Real GDP	0.5842	0.0073	-0.0560	0.0120	0.0132	-0.0468	0.0036	-0.0589
		(1.0152)	(1.4349)	(1.4260)	(1.4366)	(1.4642)	(1.4198)	(1.4454)	(1.4073)
	GDP deflator	0.6060	0.0057	-0.0509	0.0401	-0.0276	-0.0526	-0.0224	-0.0391
		(1.1068)	(1.5631)	(1.5006)	(1.5297)	(1.5464)	(1.5153)	(1.5350)	(1.5303)
	Federal funds rates	0.0534	0.0005	0.0787	0.0022	0.0112	-0.0847	0.0153	-0.0691
		(0.9029)	(1.2852)	(1.2214)	(1.2832)	(1.2862)	(1.2180)	(1.2860)	(1.2204)
One year	Real GDP	0.2439	0.0775	0.1281	0.1206	0.2140	0.0399	0.0130	0.0063
		(7.2287)	(10.2792)	(10.6678)	(10.3078)	(10.5021)	(10.5345)	(10.3627)	(10.6221)
	GDP deflator	0.8429	0.2575	2.3283	0.1698	0.1132	0.2320	0.0599	2.1444
		(18.0077)	(25.4739)	(29.3281)	(25.2951)	(25.4412)	(29.1229)	(25.2993)	(29.2058)
	Federal funds rates	0.2376	0.0669	-0.8071	0.0556	0.0675	-0.7383	0.0251	-0.7094
		(5.5009)	(7.8021)	(7.3775)	(7.8658)	(7.9428)	(7.5868)	(7.7909)	(7.5327)
C. Large-scale model									
One quarter	Real GDP	0.1741	0.0068	-0.0804	0.0063	0.0026	-0.0814	0.0008	-0.0812
		(1.2089)	(1.7358)	(1.6541)	(1.7135)	(1.7100)	(1.6364)	(1.7170)	(1.6404)
	GDP deflator	0.1891	0.0047	0.0898	-0.0029	-0.0104	-0.0665	0.0091	-0.0642
		(1.3904)	(1.9699)	(1.8947)	(1.9754)	(1.9947)	(1.8713)	(1.9652)	(1.9078)
	Federal funds rates	0.4716	0.0088	-0.1439	0.0016	-0.0032	-0.1305	0.0029	-0.1225
		(1.0140)	(1.4488)	(1.4039)	(1.4243)	(1.4537)	(1.4081)	(1.4619)	(1.4099)
One year	Real GDP	0.2213	-0.0611	-0.2313	0.0364	-0.0326	-0.3076	0.0066	-0.3440
		(8.9188)	(12.7650)	(12.9247)	(12.6792)	(12.8603)	(12.9173)	(12.9304)	(12.8769)
	GDP deflator	0.5905	0.0152	0.7184	0.0119	-0.0091	0.8252	0.0223	1.2011
		(20.5918)	(28.9367)	(34.8658)	(29.3078)	(29.4492)	(35.0813)	(29.3800)	(35.5785)
	Federal funds rates	0.6030	0.0111	-0.3176	0.0092	0.0054	-0.3953	0.0196	-0.3415
		(7.2958)	(10.3524)	(9.4985)	(10.3156)	(10.4595)	(9.4699)	(10.5727)	(9.5890)

Note: The table reports the results relative to the forecasting accuracy using predictive densities. The variables we forecast are real GDP, GDP deflator and federal funds rates. For each model we report the sample average of the difference between the log score of TV-SV-VAR-MN and the log score of that model. Standard deviation (HAC) are reported in parentheses.

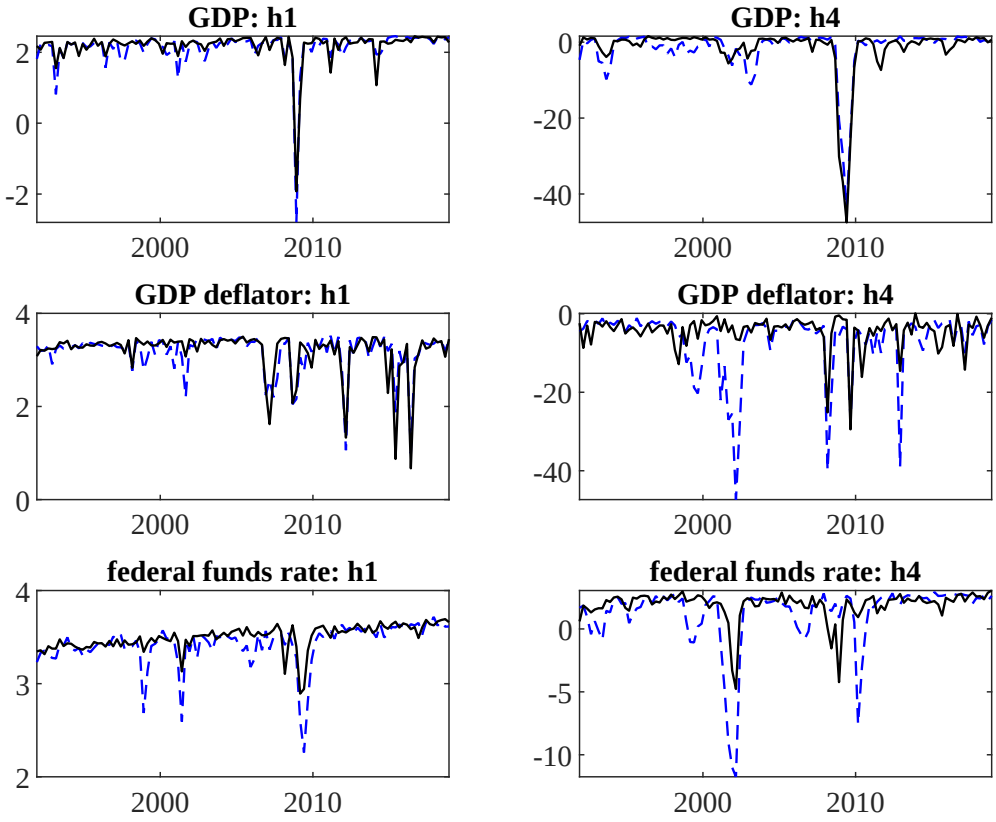

Figure 1.6: Log predictive scores. The figure compares the out-of-sample log predictive scores of TV-SV-VAR-OLS (dashed lines) and TV-SV-VAR-MSD (solid lines).

and Geweke, 2017).

Suppose that a model A assigns one-step-ahead predictive densities $p\left(y_t \mid \boldsymbol{y}_{t-1}, A\right)$. Denote the corresponding sequence of cumulative distribution functions

$$F\left(c \mid \boldsymbol{y}_{t-1}, A\right) = P\left(y_t \leq c \mid \boldsymbol{y}_{t-1}, A\right).$$

The PIT corresponding to model A and the sequence $\{y_t\}$ is $p_1\left(t; \boldsymbol{y}_T, A\right) = F\left(y_t \mid \boldsymbol{y}_{t-1}, A\right)$. If $A = D$, the true data generating process for $\{y_t\}$, then the sequence $\{p_1\left(t; \boldsymbol{y}_T, A\right)\}$ is distributed i.i.d. uniform $(0, 1)$ *ex ante*. This result dates back at least to Rosenblatt (1952). Following Smith (1985) and Berkowitz (2001), if $A = D$ and $\Phi(z)$ is the $N(0, 1)$ pdf computed at z then

$$f_1\left(t; \boldsymbol{y}_T, A\right) = \Phi^{-1}\left[p_1\left(t; \boldsymbol{y}_T, A\right)\right] \overset{\text{iid}}{\sim} N(0, 1),$$

Table 1.4: Density Forecasts: PITs tests

A. Small-scale model

Horizons	Variables	VAR-OLS	VAR-MN	VAR-MSD	TV-VAR-OLS	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-OLS	TV-SV-VAR-MN	TV-SV-VAR-MSD
One quarter	Real GDP	9274.0154	8872.1068	30.8204	20.6793	373.8509	2428.4852	7172.2850	19.19834	5399.5463
		(0.0000)	(0.0000)	(0.0000)	(0.0004)	(0.0000)	(0.0000)	(0.0000)	(0.0007)	(0.0000)
	GDP deflator	8516.017	27.3619	8.3448	9.3937	10.2019	8.6522	10.9311	9.3135	9.5188
		(0.0000)	(0.0000)	(0.0797)	(0.0520)	(0.0372)	(0.0704)	(0.0273)	(0.0537)	(0.0494)
	Federal funds rates	9891.5667	36.9894	34.71888	35.7784	35.5597	35.2467	36.0294	35.0585	33.3910
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
One year	Real GDP	8221.3467	7180.9186	4917.9310	8439.1195	4268.4217	5218.8476	6270.4617	24.6616	61.3362
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
	GDP deflator	8618.7900	20.3463	22.0893	23.5770	22.7860	2725.6065	22.7511	21.6819	21.2510
		(0.0000)	(0.0004)	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0002)	(0.0002)
	Federal funds rates	4431.5249	35.8160	31.8133	8103.6109	34.8100	29.0092	35.2163	34.5799	30.0986
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

B. Medium-scale model

Horizons	Variables	VAR-OLS	VAR-MN	VAR-MSD	TV-VAR-OLS	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-OLS	TV-SV-VAR-MSD	
One quarter	Real GDP	8350.9447	1332.6690	23.5712076886	3465.2394	4104544.0322	2954.9737	1884.6209	14.4362	0403.2827
		(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0060)	(0.0000)
	GDP deflator	9492.3699	6.9154	8.0626	6.7441	5.8297	9.3926	6.4776	7.6769	9.5964
		(0.0000)	(0.1404)	(0.0893)	(0.1500)	(0.2122)	(0.0520)	(0.1662)	(0.1042)	(0.0478)
	Federal funds rates	6161.1165	24.4832	31.6079	23.9001	24.5185	35.0840	25.1169	24.6555	33.1340
		(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
One year	Real GDP	9212.1281	88.4700	42.3808	6.0013	8830.0394	2470.7264	3157.4666	9037.9375	44.5488
		(0.0000)	(0.0000)	(0.0000)	(0.1990)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
	GDP deflator	6275.0602	8.1029	20.9937	7.5379	7.3585	22.2568	7.2060	7.4847	20.6274
		(0.0000)	(0.0879)	(0.0003)	(0.1100)	(0.1181)	(0.0002)	(0.1254)	(0.1124)	(0.0004)
	Federal funds rates	2038.4059	7.2464	26.8559	1954.3504	8.3700	27.3118	7.9091	8.1303	26.5152
		(0.0000)	(0.1234)	(0.0000)	(0.0000)	(0.0789)	(0.0000)	(0.0950)	(0.0869)	(0.0000)

C. Large-scale model

Horizons	Variables	VAR-OLS	VAR-MN	VAR-MSD	TV-VAR-OLS	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-OLS	TV-SV-VAR-MSD	
One quarter	Real GDP	9970.3741	14.7887	8670.5729	7798.5269	5734.2418	2890.2129	6148.2757	14.7315	3863.5464
		(0.0000)	(0.0052)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0053)	(0.0000)
	GDP deflator	6271.6513	6.83527	954.07253	5.8826	4898.9424	2.8956	5.4104	6.9875	3.2689
		(0.0000)	(0.1449)	(0.0000)	(0.2081)	(0.0000)	(0.5754)	(0.2477)	(0.1365)	(0.5139)
	Federal funds rates	7744.7786	34.1017	29.9967	34.0210	32.0204	29.9997	32.6111	31.8814	29.3626
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)
One year	Real GDP	6200.5903	18.2341	13.3035	21.3793	19.5193	22.6566	18.8196	20.7063	13.0767
		(0.0000)	(0.0011)	(0.0099)	(0.0003)	(0.0006)	(0.0001)	(0.0009)	(0.0004)	(0.0109)
	GDP deflator	8277.2890	16.9914	8.5880	12.0646	22.7629	5182.6749	19.3429	23.4467	10.0402
		(0.0000)	(0.0019)	(0.0723)	(0.0169)	(0.0001)	(0.0000)	(0.0007)	(0.0001)	(0.0398)
	Federal funds rates	9366.2786	20.7035	19.6452	18.4228	18.8256	19.6224	21.2588	21.4581	19.17883
		(0.0000)	(0.0004)	(0.0006)	(0.0010)	(0.0009)	(0.0006)	(0.0003)	(0.0003)	(0.0007)

Note: The table reports the test statistics of PIT tests. The p-values are reported in parentheses based on the simulation sample of size 10^5 described in the Appendix C. For reported values above 0.02 these values are close to those of the asymptotic distributions.

and for analytical purposes it is often more convenient to work with $\{f_1(t; \boldsymbol{y}_T, A)\}$ than with $\{p_1(t; \boldsymbol{y}_T, A)\}$. For h-step-ahead predictive distributions, let $p_h(t; \boldsymbol{y}_T, A) = F\left(y_t \mid \boldsymbol{y}_{t-h}, A\right)$. If $A = D$, then the distribution of $p_h(t; \boldsymbol{y}_T, A)$ is uniform on $(0, 1)$ *ex ante*. The goodness-of-fit tests employed in this paper is chi-squared test in Geweke and Amisano (2010). The details are presented in Appendix C.

The variant of the portmanteau test in this work uses the first four moments ($q = 1, 2, 3, 4$) of the normalized PIT z. For each constituent the asymptotic distribution of the moment test statistics are $\chi^2(4)$. In Table 1.4 we report the test statistics, along with p-values in parentheses. The reported values above 0.02 are close to those of the asymptotic distributions. Results suggest that the density forecasts produced by the TVP-VAR model with informative priors are well calibrated for GDP deflator of medium- and large-scale data. Evidence is more mixed for real GDP and federal funds rates, which is misspecified at some horizons and scales. In general, the calibration of the density forecasts produced by the TVP-VAR model with informative priors compares well with the ones obtained using the other models, particularly with a Normal prior.

1.5.4 Structural analysis: Impulse Response Functions

VARs are used in the literature not only for forecasting but also as a tool to identify structural shocks and assess their transmission mechanism. In this section, we estimate the impulse responses to monetary policy shocks using our medium-scale BVAR (DSGE) with seven variables, estimated over the entire available sample. However, we do not have an observable counterpart of these impulse responses in the data that can be used to directly check their accuracy. This motivates the future work to perform a controlled Monte Carlo experiment.

The monetary policy shock is identified using a relatively standard recursive identification scheme, assuming that only federal fund rates react contemporaneously to the monetary policy shock. Figures 1.7 and 1.8 report the median and the 16th and 84th percentiles of the posterior distribution of the impulse responses to a monetary policy shock using TVP-VAR with informative priors using the full sample (solid lines). The dates chosen for the comparison are 1974Q4 and 2007Q4. They are somewhat representative of the typical economic conditions of the Great Moderation and the financial crisis.

A one standard deviation exogenous increase in the federal funds rate generates a substantial

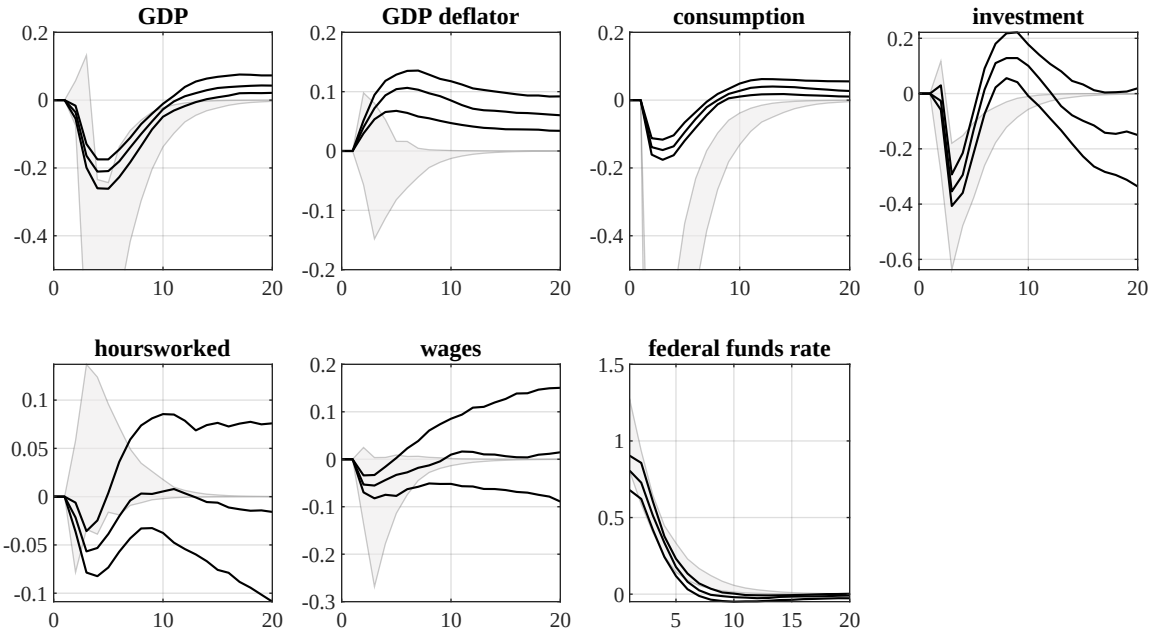

Figure 1.7: Impulse responses. The figure reports the 16th, 50th, and 84th percentiles (solid lines) of the distribution of the impulse response functions of the TV-SV-VAR-MSD to a 1 standard deviation monetary policy shock in 1984Q4. The gray area refers to the corresponding quantiles of the distribution of impulse responses of TV-SV-VAR-OLS.

contraction in GDP, consumption, investment and all the other variables related to economic activity. The responses show the similar patterns of two dates, while the magnitude of the effects is larger in 1984Q4. For comparison, figures 1.7 and 1.8 also report the corresponding quantiles of the distribution of impulse responses of TVP-VAR with a Normal prior (gray shaded areas). It is evident that these error bands reflect a considerable amount of estimation uncertainty, and even when initially significant, these impulse responses tend to revert to zero at a very fast pace. In a nutshell, the results are all in line with intuition, and hence supportive to our proposed model.

1.6 Conclusion

In this paper, we have studied the problem of whether and how to choose the informativeness of a variety of commonly used prior distributions for VAR models. Using U.S. data, we demonstrated the superior forecast performance of the proposed TVP-VAR with informative priors compared to the Normal prior. We have shown that this approach is theoretically grounded, easy to implement

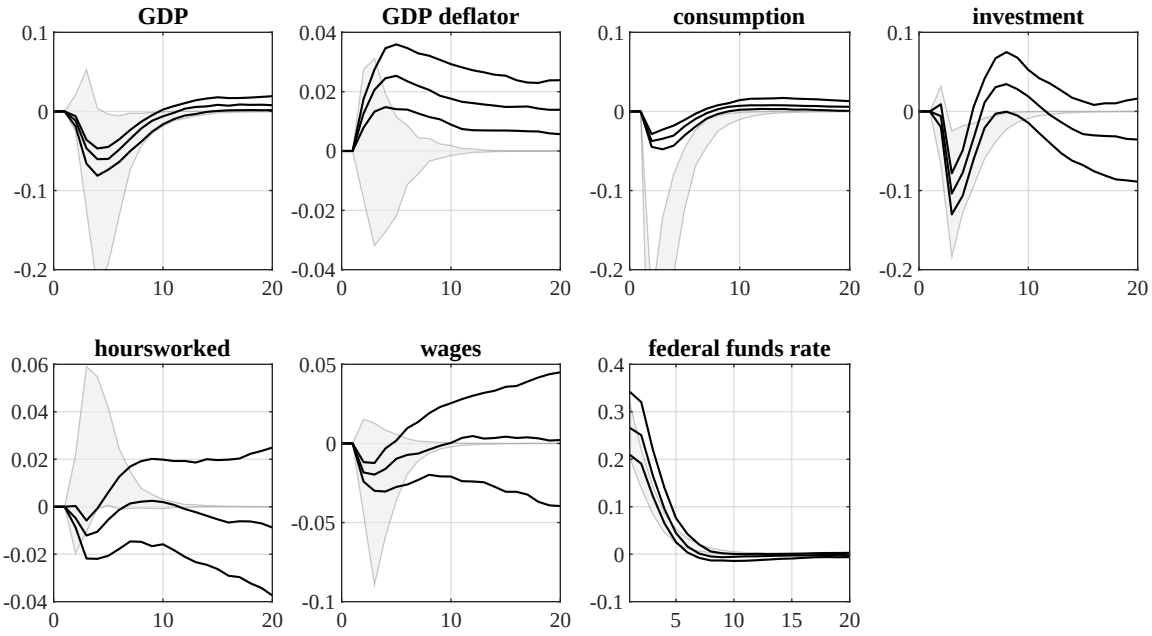

Figure 1.8: Impulse responses. The figure reports the 16th, 50th, and 84th percentiles (solid lines) of the distribution of the impulse response functions of the TV-SV-VAR-MSD to a 1 standard deviation monetary policy shock in 2007Q4. The gray area refers to the corresponding quantiles of the distribution of impulse responses of TV-SV-VAR-OLS.

by adopting equation-by-equation estimation procedure, and performs very well in terms of out-of-sample forecasting and accuracy. Moreover, this approach proves particularly useful for the increasingly literature on large VAR models. The models are also validated for level data.

In future work, developing an order-invariant version of these TVP-VARs would be an interesting and important extension. The state-of-the-art paper is Arias et al. (2022). In addition, the model can be implemented on price discovery, which is one of the central functions of financial markets. The price discovery is dynamic in nature, and this fast adjustment process performs more volatility during the crisis. This leaves us a vast potential in applying time varying models to capture the dynamics.

Chapter 2

Bayesian Analysis of Price Discovery

Wenqiu Ma, Eric Zivot

Abstract

We study price discovery in fragmented markets through the lens of a structural cointegration model with identified permanent and transitory shocks using Bayesian methods. Utilizing recent results in the Bayesian analysis of vector autoregression models, we propose a Bayesian inference procedure to derive posterior distributions of commonly used measures of price discovery.

Through simulations, we show that the proposed structural measures are robust to realistic degrees of noise, asynchronous trading, and misspecification, and that Bayesian estimation substantially improves the precision and stability of inference relative to classical methods. An empirical application to fragmented trading in major U.S. equity index ETFs—SPY and IVV—illustrates how the framework delivers economically interpretable rankings of venues and instruments in terms of their contribution to the permanent price, and how posterior uncertainty can be used to assess the statistical significance of those rankings. Overall, the paper demonstrates that combining structural identification of permanent shocks with Bayesian inference yields a coherent, noise-robust toolkit for price discovery analysis in modern financial markets.

2.1 Introduction

Financial markets are increasingly fragmented. The same or closely related claims trade contemporaneously across multiple venues—primary exchanges, alternative trading systems, dark pools, futures and options markets, and exchange-traded funds. In such an environment, understanding price discovery—the process by which new information about fundamentals is impounded into transaction prices—is essential for evaluating market quality and the allocation of capital in a competitive economy. A central empirical question is which venue or instrument first reflects news about fundamental value and to what extent observed leadership reflects genuine information acquisition rather than differences in trading frictions or microstructure noise.

The modern empirical literature formalizes price discovery for cointegrated prices by positing a common efficient price that follows a martingale and by allowing observed quotes and trades to contain transitory components. The main empirical model used is a bivariate vector error correction model (VECM) with a known cointegrating vector $\boldsymbol{\beta} = (1, -1)'$. Two traditions anchor the literature. In the first, variance-decomposition methods apportion efficient-price innovations to venue-specific shocks, delivering information shares (IS) as defined in Hasbrouck (1995). In the second, common-factor representations extract the latent permanent component and attribute component shares (CS) to each venue (see (Baillie et al., 2002; Harris et al., 2002; De Jong, 2002). These methods, however, have several drawbacks (e.g., non-uniqueness due to order-dependence, ambiguous interpretation, impacted by transient dynamics) which makes their empirical use problematic. Recent research led by Yan and Zivot (2007), Yan and Zivot (2010), Putniņš (2013), Shen et al. (2025), and Lautier et al. (2024) propose alternatives to the IS and CS measures based on structural cointegration models that are order invariant/unique and not impacted by transient dynamics. Among these alternatives, the preferred measures include the improved information leadership share measure of Shen et al. (2025), denoted PILS, and the covariance information share measures of Lautier et al. (2024), denoted CovIS and CovISQ.[1].

The classical frequentist statistical methodology is used in almost all price discovery studies and the emphasis is on point estimation rather than inference. This is due to the fact that price discovery measures like IS, CS, PILS, and CovIS are ratios of highly nonlinear functions

[1]Shen et al. (2025) shows that CovISQ is equivalent to PILS.

of reduced-form VECM parameters and no general asymptotic distribution theory results are available for these measures.[2] If standard errors are reported, then these are typically computed using bootstrapping techniques. While there have been a few investigations of the accuracy of bootstrap standard errors for some price discovery measures (e.g., Sapp (2002) and Schweikert (2021)), the bootstrap is known to be unreliable for ratio-type estimates. This places doubt on the reliability of confidence intervals and inference for price discovery measures based on the bootstrap. An additional problem with the classical estimation of the empirical VECM used to compute price discovery measures is the specification of lag length. If very high frequency data are used (e.g. 1-second) then many lags may be needed in the estimation of the VECM which can increase the variances of price discovery measures. Also, many empirical studies use standard information criteria such as AIC or BIC and these methods may not work well for model specification as they focus on common fixed lag length models.

In this paper, we advance a Bayesian approach for inference on price discovery measures that addresses the above mentioned concerns of the classical approach. Our Bayesian approach offers three advantages that are especially salient for discovery metrics. The first advantage is regularization and shrinkage of high dimensional parameters. We use Minnesota priors and their hierarchical variants to shrink autoregressive lag coefficients toward zero mitigating overfitting and improving finite-sample behavior (Doan et al., 1984; Litterman, 1986; Kadiyala and Karlsson, 1997; De Mol et al., 2008; Bańbura et al., 2010; Giannone et al., 2015). The second advantage is finite sample inference for price discovery measures and price discovery impulse response functions (IRFs) that are based on simulated posterior distributions of structural parameters related to identified permanent shocks. The third advantage is formal inference about price discovery between competing assets and markets. Using marginal likelihoods and posterior odds we can formally evaluate hypotheses concerning price discovery such as equal price discovery shares.

The outline of our paper is as follows. In Section 2, we provide a short review of the price discovery literature focusing on recent price discovery measures derived from structural cointegration models. In Section 3, we detail our Bayesian modeling approach for price discovery (given details of the subsections). In Section 4 we present our Bayesian approach using simulated data

[2]Two notable exceptions are the recent papers by Lien et al. (2025), who report asymptotic distribution results for PILS based on the delta method, and Schweikert (2026) who derive the asymptotic distribution of IS and provide a large sample approximation to conduct standard hypothesis tests.

from a stylized price discovery model commonly used in the literature. In Section 5, we apply our methodology to the analysis of price discovery among exchange traded funds that track the S&P 500 index. Finally, our conclusions are presented in Section 6.

2.2 Literature Review

Research on price discovery in fragmented markets and among arbitrage-linked assets has developed along two intertwined tracks: (i) price discovery measurement based on reduced–form cointegrated systems represented by VECMs with known cointegrating vectors; and (ii) structural interpretations of price discovery measures that aim to separate permanent information shocks from transitory frictions such as inventory risk and microstructure noise.

The seminal contributions by Hasbrouck (1995) and Gonzalo and Granger (1995) formalize price discovery when the same asset trades in multiple markets or when multiple assets are arbitrage-linked. The information share (IS) of Hasbrouck (1995) attributes the variance of innovations to the common efficient price to individual venues (markets or assets), while the component share (CS) of Gonzalo and Granger (1995) traces the venue impact of the permanent component through common factor loadings. Hasbrouck (1995) defines IS as the proportional contribution of each market's innovation variance to the variance of the latent efficient price. Practically, a VECM for the change in cointegrated prices is estimated and the variance matrix of the reduced-form residuals is decomposed via a Cholesky factorization to construct IS. Because contemporaneous venue shocks are typically correlated, IS is not unique and depends on the ordering of prices in the VECM. As a result, IS is only defined within lower and upper bounds which can be uninformative if the VECM residuals are highly correlated. Nonetheless, IS has become the workhorse measure of price discovery. Using the Gonzalo–Granger framework, CS measures each market's weight in the latent permanent component (Baillie et al., 2002) and so it is thought to be less about "who moves first" and more about long-run contribution. Harris et al. (2002) and De Jong (2002) emphasize that CS and IS answer different questions and should not be conflated. As noted in Baillie et al. (2002) and De Jong (2002), IS is a variance weighted version of CS, and is the most widely used measure of price discovery in empirical studies of price discovery. Empirically, there is a large literature using IS and CS which shows how price discovery varies with liquid-

ity, trading intensity, and market design (e.g., Hasbrouck (2003); Theissen (2016); Ozturk et al. (2017)). Derivatives can meaningfully contribute to price discovery in underlyings (Chakravarty et al., 2004), and high–frequency trading tends to increase quote informativeness (Hendershott et al., 2011; Brogaard et al., 2014).

Despite the popularity of IS and CS for measuring price discovery, they can mislead when venues have a high degree of transitory noise. The IS order dependence and innovation correlation issues, which complicate IS estimation, and the focus of CS on composition rather than informativeness, are analytically laid out by Yan and Zivot (2007, 2010) and discussed in applied contexts (Caporale and Girardi, 2013; Hauptfleisch et al., 2016). High-frequency sampling amplifies transitory frictions, making IS particularly sensitive. Yan and Zivot (2010) show that both IS and CS are functions of two primitives—informativeness and noise/liquidity shocks—and only IS contains information about relative informativeness. They also emphasize that IS is a measure of contemporary price discovery and when computed using ultra-high-frequency data it can be distorted by transitory frictions.

Motivated by identification concerns, numerous studies have proposed different solutions to address the order dependence of IS. For example, Hasbrouck (1995) suggested the use of high-frequency data to reduce residual correlation, Baillie et al. (2002) advocated using the mean or midpoint of IS measures, and Grammig and Peter (2013) proposed using a measure of tail price discovery. These solutions, however, have been found to be either only effective in a particular context or have their own identification issues. Some new order invariant measures of price discovery have been proposed. Lien and Shrestha (2009) proposed the Modified Information Share (MIS) measure based on the spectral decomposition of the correlation matrix of the reduced-form residuals, and Sultan and Zivot (2015) and Shen et al. (2025) proposed the Price Discovery Share (PDS) based on a simple additive decomposition of the variance of the underlying efficient market innovations. Through simulation studies and analytic examination, these new measures have been shown to be generally superior to IS.

A general problem with the traditional price discovery measures IS and CS and the order-invariant measures MIS and PDS is that they are functions of parameters from a reduced-form VECM and lack a clear structural interpretation. Using a general structural cointegration model that features two types of structural shocks (a permanent news innovation to the common funda-

mental value and a transitory liquidity/noise trading shock), Yan and Zivot (2010) derive structural representations of CS and IS and show two striking results. First, the CS of a venue is a function of only the contemporaneous impacts of transitory shocks and accounts for the avoidance of noise trading and liquidity shocks to that venue. Second, IS is a function of the contemporaneous impacts of both permanent and transitory shocks, and so is not a pure measure of price discovery. Shen et al. (2025) show that MIS and PDS have similar structural representations as IS.

To sort out the confounding effects of transitory shocks on IS in the special case of uncorrelated reduced-form residuals, Yan and Zivot (2010) combine IS and CS and form the information leadership (IL) measure, which is the ratio of the contemporaneous impacts of the permanent shock for each venue. Putniņš (2013) transformed IL into the information leadership share (ILS) measure taking a value between 0 and 1 and showed the superiority of ILS over CS and IS in an extensive simulation study. To reduce bias, ILS was transformed into a binary information leadership indicator (ILI) by Patel et al. (2020). A limitation of the IL, ILS and ILI measures is that they are derived under the special case of uncorrelated reduced-form errors. Empirically, reduced-form errors can be correlated even with high frequency data and this puts doubt on the general usefulness of these measures.

To improve the IL, ILS and ILI measures in the general case of correlated reduced-firm residuals, Shen et al. (2024a) combine PDS with CS into a new correlation-robust price information leadership (PIL) measure that is free of the structural impacts of transitory frictions and has the same structural interpretation as IL in the case of uncorrelated reduced-form residuals. In the same spirit as Putniņš (2013) and Patel et al. (2020), they define the correlation-robust price information leadership share (PILS) measure and the binary price information leadership share indicator (PILI). Using simulation designs similar to those used in Yan and Zivot (2010) and Putniņš (2013) allowing for correlated reduced-form residuals, Shen et al. (2025) demonstrated the superior performance of PILS and PILI over CS, IS, MIS, PDS, ILS, and ILI.

Independently of Shen et al. (2025), Lautier et al. (2024) derive the Covariance Information Share (CovIS) of a venue as the covariance between the reduced-form venue specific shock and the latent permanent shock normalized by the sum of venue specific covariances, and show that CovIS is the ratio of contemporaneous venue impact of the permanent shock divided by the sum of the contemporaneous venue impacts which is free of the impact of transitory shocks. They also

define a measure called CovISQ which is equivalent to the PILS measures of Shen et al. (2025).

All of the price discovery measures discussed so far are measures of contemporaneous price discovery. That is, they are only functions of the contemporaneous impacts of shocks. They do not say anything about the path or speed of price discovery. Yan and Zivot (2007) were the first to define and identify venue specific permanent shock structural impulse response functions (SIRFs) to measure the path and speed of price discovery over time. They also defined statistics based on the permanent shock SIRFs to measure the speed of price discovery in a single number. Using a different identification scheme than Yan and Zivot (2007), Lautier et al. (2024) derived the same permanent shock SIRFs as Yan and Zivot (2007).

Most of the econometric analysis of price discovery measures has focused on estimation and very little analysis has centered on inference. Price discovery measures like IS, CS, MIS, IL, PILS, and CovIS are ratios of highly nonlinear functions of reduced-form VECM parameters and asymptotic distributions for these parameters can be derived using the delta method based on results presented in **?**. However, to our knowledge, only a few asymptotic distribution results based on the delta method are available for these measures. Lien et al. (2025) derived the asymptotic distribution for PILS/CovISQ and provided analytical standard errors in the bivariate case. Schweikert (2026) derived the asymptotic distribution of Hasbrouck's information shares and proposed large-sample approximations that enable testing of certain hypotheses. The lack of asymptotic distribution results for price discovery measures means that there are no general statistical tests for hypotheses concerning price discovery. Because of the lack of asymptotic distribution results, inference on price discovery measures is largely based on the bootstrap. However, not much is known about the accuracy of the bootstrap for price discovery measures beyond the results reported in Schweikert (2021). Schweikert (2021) showed that the bootstrap confidence bands for IS proposed by Sapp (2002) perform well if markets have similar information shares but are too narrow if one market dominates price discovery. Using this result, Schweikert (2021) developed bootstrap tests for equal information shares and for zero information shares and evaluated these tests using Monte Carlo simulations.

Only a few papers have applied Bayesian methods to study price discovery. Madhavan and Smidt (1991) developed and evaluated a Bayesian model of intraday security price movements and devised a natural measure of the degree of information asymmetry. Kim and Linn (2022)

developed an economic model of supply and demand in a commodity spot and futures market in which the parameters of the equilibrium price function are subject to regime changes. They used Bayesian methods to produce posterior distributions of key structural parameters as well as a new measure of price discovery. Mohamad and Inani (2023) examined price discovery between bitcoin spot and futures markets during the Covid-19 pandemic using time-varying parameter Bayesian vector autoregressive model (TVP-VAR) with stochastic volatility from Primiceri (2005). Although Mohamad and Inani (2023) used BVARs to describe the dynamics of bitcoin spot and futures data and to compute point estimates of CS, IS, MIS, and IL, they did not study the posterior distributions of these price discovery measures.

We contribute to the price discovery literature by using recent research on Bayesian VECMs to study the posterior distributions of the static price discovery measures CS, IS, MIS, ILS, PDS, PILS, PIES, and CovIS as well as the dynamic permanent/transitory shock SCIRFs. These posterior distributions allow us to perform finite sample inference on price discovery measures and to evaluate hypotheses related to price discovery without the need for delta-method approximations.

2.3 Modeling Price Discovery

In this section, we present our methodology for modeling price discovery using cointegration. We first review the structural cointegration model of Yan and Zivot (2007) and Yan and Zivot (2010) and the associated reduced-form VECM. We then discuss identification of the parameters of the structural cointegration model. This is followed by a review of price discovery measures derived from both the reduced-form VECM and structural model.

2.3.1 Structural Cointegration Model

Let $\mathbf{p}_t = (p_{1t}, p_{2t})'$ denote a vector of log prices for two assets that are closely related by arbitrage. For example, these can be the same asset traded in different markets or two nearly identical assets traded in the same market. In the price discovery literature, it is assumed that these two prices series are intergated of order 1, or I(1), the price changes, $\Delta\mathbf{p}_t$, are stationary, or I(0), and that $\mathbf{p}_t$ is cointegrated with cointegrating vector $\beta = (1, -1)'$.

Yan and Zivot (2010) start with the following structural moving average (SMA) representation of $\Delta \mathbf{p}_t$

$$\Delta \mathbf{p}_t = \mathbf{D}(L)\boldsymbol{\eta}_t = \mathbf{D}_0\boldsymbol{\eta}_t + \mathbf{D}_1\boldsymbol{\eta}_{t-1} + \mathbf{D}_2\boldsymbol{\eta}_{t-2} + \dots \tag{2.1}$$

$$\mathbf{D}(L) = \sum_{k=0}^{\infty} \mathbf{D}_k L^k, \quad \mathbf{D}_0 \neq \mathbf{I}, \tag{2.2}$$

where the elements of $\{\mathbf{D}_k\}_{k=0}^{\infty}$ are 1-summable and $\mathbf{D}_0$ is invertible. They assume that the number of structural shocks is equal to the number of observed prices, so that $\mathbf{D}(L)$ is invertible. The innovation to the common efficient price of the asset, η_t^P, is labeled the permanent shock and the noise innovation, η_t^T, is labeled the transitory shock so that $\boldsymbol{\eta}_t = \left(\eta_t^P, \eta_t^T\right)'$. These structural shocks are assumed to be serially and mutually uncorrelated with diagonal covariance matrix $\mathbf{C} = \text{diag}\left(\sigma_P^2, \sigma_T^2\right)$. The matrix $\mathbf{D}_0$ contains the initial impacts of the structural shocks on $\Delta \mathbf{p}_t$, and defines the contemporaneous correlation structure of $\Delta \mathbf{p}_t$:

$$\mathbf{D}_0 = \begin{pmatrix} d_{0,1}^P & d_{0,1}^T \\ d_{0,2}^P & d_{0,2}^T \end{pmatrix} = (\mathbf{d}_0^P \vdots \mathbf{d}_0^T). \tag{2.3}$$

The permanent innovation η_t^P carries new information on the fundamental value of the asset, and permanently moves the market prices. Given that the cointegrating vector is $\boldsymbol{\beta} = (1, -1)'$, the defining characteristic of η_t^P is that it has a one-to-one long-run effect on the price levels for each market. The transitory innovation η_t^T summarizes non-information related shocks, such as trading by uninformed or liquidity traders. The defining characteristic of η_t^T is that it is uncorrelated with the informational innovation η_t^P, and has no long-run effect on price levels. Hence, the long-run impact matrix $\mathbf{D}(1)$ of the structural innovations $\boldsymbol{\eta}_t$ takes the form

$$\mathbf{D}(1) = \begin{pmatrix} d_1^P(1) & d_1^T(1) \\ d_2^P(1) & d_2^T(1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = (\mathbf{d}^P(1) \vdots \mathbf{d}^T(1)). \tag{2.4}$$

The Beveridge-Nelson decomposition applied to Eq. (2.1) yields the level relationship:

$$\mathbf{p}_t = \mathbf{p}_0 + \mathbf{D}(1)\sum_{j=1}^{t}\boldsymbol{\eta}_j + \mathbf{s}_t \tag{2.5}$$

$$= \mathbf{p}_0 + \mathbf{1}\sum_{j=1}^{t}\eta_t^P + \mathbf{s}_t \tag{2.6}$$

$$= \mathbf{p}_0 + \mathbf{1}m_t + \mathbf{s}_t, \tag{2.7}$$

where $\mathbf{D}(1) = \sum_{k=0}^{\infty}\mathbf{D}_k$, $\mathbf{s}_t = (s_{1t}, s_{2t})' = \mathbf{D}^*(L)\boldsymbol{\eta}_t \sim I(0)$, and $\mathbf{D}^*(L) = -\sum_{j=k+1}^{\infty}\mathbf{D}_j$ $(k = 0, \ldots, \infty)$, and $m_t = m_{t-1} + \eta_t^P$. Multiplying (5) by $\boldsymbol{\beta} = (1, -1)'$ shows that $\boldsymbol{\beta}'\mathbf{p}_t \sim I(0)$ since $\boldsymbol{\beta}'\mathbf{D}(1) = \mathbf{0}$.

2.3.2 Reduced-form VECM

The reduced-form VECM associated with the structural model has the form:

$$\Delta\mathbf{p}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{p}_{t-1} + \sum_{j=1}^{k}\boldsymbol{\Gamma}_j\Delta\mathbf{p}_{t-j} + \boldsymbol{\varepsilon}_t, \tag{2.8}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$ is the error correction vector, $\boldsymbol{\Gamma}_j$ $(j = 1, \ldots, k)$ are the short-run coefficient matrices, and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is the vector of reduced-form VECM residuals with $E[\boldsymbol{\varepsilon}_t] = \mathbf{0}$ and

$$E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_s'] = \begin{cases} \mathbf{0}, & \text{if } t \neq s, \\ \boldsymbol{\Omega}, & \text{otherwise,} \end{cases}$$

and the residual covariance matrix $\boldsymbol{\Omega}$ is

$$\boldsymbol{\Omega} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

where σ_i^2 denotes the variance of each market's idiosyncratic error and ρ denotes the correlation coefficient between these two errors.

From the Granger Representation Theorem, the VECM (2.8) can be inverted to give the

reduced-form vector moving average (VMA) representation:

$$\Delta \mathbf{p}_t = \mathbf{\Psi}(L)\varepsilon_t = \varepsilon_t + \mathbf{\Psi}_1\varepsilon_{t-1} + \mathbf{\Psi}_2\varepsilon_{t-2} + \cdots , \tag{2.9}$$

and its integrated form (or Beveridge-Nelson decomposition):

$$\begin{aligned} \mathbf{p}_t &= \mathbf{p}_0 + \mathbf{\Psi}(1)\sum_{s=1}^{t}\boldsymbol{\varepsilon}_s + \mathbf{\Psi}^*(L)\boldsymbol{\varepsilon}_t \\ &= \mathbf{p}_0 + \mathbf{\Psi}(1)\sum_{s=1}^{t}\boldsymbol{\varepsilon}_s + \mathbf{s}_t, \end{aligned} \tag{2.10}$$

where $\mathbf{\Psi}(1) = \sum_{k=0}^{\infty}\mathbf{\Psi}_k$ with $\mathbf{\Psi}(L)$ and $\mathbf{\Psi}^*(L)$ being matrix polynomials in the lag operator, L, $\mathbf{\Psi}_t^*(k) = -\sum_{j=k+1}^{\infty}\mathbf{\Psi}_j$, and $\mathbf{s}_t = \mathbf{\Psi}^*(L)\boldsymbol{\varepsilon}_t \sim I(0)$. The matrix $\mathbf{\Psi}(1)$ contains the cumulative impacts of the innovation ε_t on all future price movements, and acts as a measure of the long-run impact of ε_t on prices.

Johansen (1991) shows that the reduced-form VECM model and the integrated VMA model are linked through the factorization:

$$\mathbf{\Psi}(1) = \beta_\perp \mathbf{\Pi}\alpha_\perp',$$

where $\boldsymbol{\beta}_\perp$ is orthogonal to $\boldsymbol{\beta}$, and $\mathbf{\Pi} = \left(\boldsymbol{\alpha}_\perp'\left(\mathbf{I} - \sum_{j=1}^{k}\mathbf{\Gamma}_j\right)\boldsymbol{\beta}_\perp\right)^{-1}$ where $\mathbf{I}$ is the identity matrix and $\boldsymbol{\alpha}_\perp$ is orthogonal to $\boldsymbol{\alpha}$. We can see that $\mathbf{\Pi}$ for the bivariate VECM model is a scalar.

Since the cointegrating vector is $\boldsymbol{\beta} = (1, -1)'$ and $\boldsymbol{\beta}'\mathbf{\Psi}(1) = \mathbf{0}$ then the rows of $\mathbf{\Psi}(1)$ are the same and we can represent $\mathbf{\Psi}(1)$ as:

$$\mathbf{\Psi}(1) = \begin{pmatrix} \psi_1 & \psi_2 \\ \psi_1 & \psi_2 \end{pmatrix} = \mathbf{1}\boldsymbol{\psi}',$$

where $\boldsymbol{\psi} = (\psi_1, \psi_2)'$ is the common row of $\mathbf{\Psi}(1)$ and $\mathbf{1}$ is a 2×1 vector of ones. Then (2.10) can

be re-written as

$$\begin{aligned}\mathbf{p}_t &= \mathbf{p}_0 + \mathbf{1}\sum_{s=1}^{t}\boldsymbol{\psi}'\boldsymbol{\varepsilon}_s + \mathbf{s}_t \\ &= \mathbf{p}_0 + \mathbf{1}\sum_{s=1}^{t}\eta_t^P + \mathbf{s}_t \\ &= \mathbf{p}_0 + \mathbf{1}m_t + s_t \end{aligned} \tag{2.11}$$

where $\eta_t^P = \boldsymbol{\psi}'\boldsymbol{\varepsilon}_t$ and $m_t = m_{t-1} + \eta_t^P$ is the efficient price.

2.3.3 Identification of the Structural Model

Comparing the SMA representation (2.1) and the reduced-form VMA representation (2.9) shows that the structural shocks $\boldsymbol{\eta}_t$ are related to the reduced-form shocks $\boldsymbol{\varepsilon}_t$ via

$$\boldsymbol{\varepsilon}_{t-j} = \mathbf{D}_0\boldsymbol{\eta}_{t-j},\ j = 0, 1, 2, \ldots \tag{2.12}$$

It follows that SMA coefficients $\mathbf{D}_j$ are related to the reduced-form VMA coefficients $\boldsymbol{\Psi}_j$ via

$$\mathbf{D}_j = \mathbf{D}_0\boldsymbol{\Psi}_j. \tag{2.13}$$

In addition, the structural shock covariance matrix $\mathbf{C} = \text{diag}(\sigma_P^2, \sigma_T^2)$ is related to the reduced-form shock covariance matrix $\boldsymbol{\Omega}$ via

$$\mathbf{C} = \mathbf{D}_0^{-1}\boldsymbol{\Omega}\mathbf{D}_0^{-1'} \tag{2.14}$$

From these results, it is clear that identification of all the parameters of the SMA model requires the identification of all of the elements in $\mathbf{D}_0$.

Identification of the elements in $\mathbf{D}_0$ was first done by Yan and Zivot (2007) using the Gonzalo-Ng decomposition. Using direct calculations, Lautier et al. (2024) derived the same representation as Yan and Zivot (2007) for $\mathbf{d}_0^P$. The derivation is instructive. Using $\boldsymbol{\varepsilon}_t = \mathbf{D}_0\boldsymbol{\eta}_t$ we have

$$\varepsilon_{it} = d_{0,i}^P\eta_t^P + d_{0,i}^T\eta_t^T\ (i = 1, 2). \tag{2.15}$$

Then, using $\eta_t^P = \boldsymbol{\psi}'\boldsymbol{\varepsilon}_t = \psi_1\varepsilon_{1t} + \psi_2\varepsilon_{2t}$, we have

$$\text{cov}(\varepsilon_{1t}, \eta_t^P) = \psi_1\sigma_1^2 + \psi_2\sigma_{12} = (\boldsymbol{\Omega}\boldsymbol{\psi})_1 = d_{0,1}^P\sigma_P^2, \tag{2.16}$$

$$\text{cov}(\varepsilon_{1t}, \eta_t^P) = \psi_2\sigma_2^2 + \psi_1\sigma_{12} = (\boldsymbol{\Omega}\boldsymbol{\psi})_2 = d_{0,2}^P\sigma_P^2, \tag{2.17}$$

where $(\boldsymbol{\Omega}\boldsymbol{\psi})_i$ denotes the i^{th} row of $\boldsymbol{\Omega}\boldsymbol{\psi}$. Using $\sigma_P^2 = \boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi}$, it follows that

$$\mathbf{d}_0^P = \frac{\boldsymbol{\Omega}\boldsymbol{\psi}}{\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi}} = \boldsymbol{\beta}^P, \tag{2.18}$$

where $\boldsymbol{\beta}^P$ is the linear projection coefficient of $\boldsymbol{\varepsilon}_t$ on $\eta_t^P = \boldsymbol{\psi}'\boldsymbol{\varepsilon}_t$.

Identification of $\mathbf{d}_0^T$ requires additional assumptions. Using the Gonzalo-Nq decomposition Yan and Zivot (2007) showed that

$$\mathbf{d}_0^T = \begin{pmatrix} \psi_2/(\psi_1 + \psi_2) \\ -\psi_1/(\psi_1 + \psi_2) \end{pmatrix}. \tag{2.19}$$

Lautier et al. (2024) assumed that $\sigma_P^2 = \sigma_T^2$ and showed that this implies

$$\mathbf{d}_0^T = \begin{pmatrix} \psi_2\sigma_1\sigma_2\sqrt{1-\rho^2} \\ \psi_1\sigma_1\sigma_2\sqrt{1-\rho^2} \end{pmatrix}. \tag{2.20}$$

2.3.4 Static Price Discovery Measures

The empirical measures most widely used to quantify price discovery are the information share (IS) measure of Hasbrouck (1995) and the component share measure (CS) of Booth et al. (1999), Chu et al. (1999) and Harris et al. (2002). The IS measures are defined as:

$$\text{IS}_i = \frac{\left([\boldsymbol{\psi}'\mathbf{F}]_i\right)^2}{\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi}} \ (i = 1, 2), \tag{2.21}$$

where $\mathbf{F}$ is the lower triangular Cholesky factor of $\boldsymbol{\Omega}$, $[\boldsymbol{\psi}'\mathbf{F}]_i$ is the i^{th} element of the row matrix $\boldsymbol{\psi}'\mathbf{F}$. IS measures the contribution of a venue to the variance of the permanent shock. When $\boldsymbol{\Omega}$ is non-diagonal, IS is not unique and depends on the ordering of the variables in $\mathbf{p}_t$. The CS

measures are defined as

$$\text{CS}_i = \frac{\psi_i}{\psi_1 + \psi_2} \ (i = 1, 2), \tag{2.22}$$

and measure a venue's permanent component contribution to the sum of the permanent component weights.

To address the order-dependence of IS, numerous studies have proposed different solutions and new measures. Most notably are the Modified Information Share (MIS) measure of Lien and Shrestha (2009), and the Price Discovery Share (PDS) measure based of Sultan and Zivot (2015) and Shen et al. (2025). The MIS measures are based on a spectral decomposition of the correlation matrix of the reduced-form residuals and Shen et al. (2025) shows that they have the form

$$\text{MIS}_i = \frac{\psi_i^2\sigma_i^2(1 + \sqrt{1-\rho^2})/2 + \psi_j^2\sigma_j^2(1 - \sqrt{1-\rho^2})/2 + \psi_i\psi_j\sigma_{i,j}}{\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi}} \ (i = 1, 2). \tag{2.23}$$

The PDS measures are based on the additive Euler decomposition of σ_P and are given by

$$\text{PDS}_i = \psi_i\beta_i^P = \frac{\psi_i^2\sigma_i^2 + \sum_{j=1}^n \psi_i\psi_{j\neq i}\sigma_{i,j\neq i}}{\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi}}, \ (i = 1, 2) \tag{2.24}$$

where β_i^P is the projection coefficient from the projection of ε_{it} on η_t^P.

The price discovery measures IS, CS, MIS and PDS are based on residuals from a reduced-form vector error correction model (VECM) and do not have a clear structural interpretation. Using the structural cointegration model (2.1), Yan and Zivot (2010) and Shen et al. (2024a) derive the following structural representations of IS, CS, MIS and PDS[3]:

$$\text{IS}_1 = \frac{d_{0,1}^P d_{0,2}^T}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}, \quad \text{IS}_2 = \frac{-d_{0,1}^T d_{0,2}^P}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}, \tag{2.25}$$

$$\text{CS}_1 = \frac{d_{0,2}^T}{d_{0,2}^T - d_{0,1}^T}, \quad \text{CS}_2 = -\frac{d_{0,1}^T}{d_{0,2}^T - d_{0,1}^T}, \tag{2.26}$$

$$\text{MIS}_1 = \frac{1}{2} + \frac{1}{2}\frac{d_{0,1}^P d_{0,2}^T + d_{0,1}^T d_{0,2}^P}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}\sqrt{1-\rho^2}, \quad \text{MIS}_2 = \frac{1}{2} - \frac{1}{2}\frac{d_{0,1}^P d_{0,2}^T + d_{0,1}^T d_{0,2}^P}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}\sqrt{1-\rho^2}, \tag{2.27}$$

$$\text{PDS}_1 = \frac{d_{0,1}^P d_{0,2}^T}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}, \quad \text{PDS}_2 = \frac{-d_{0,1}^T d_{0,2}^P}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}. \tag{2.28}$$

[3]The representation for IS is for the special case of uncorrelated reduced-form residuals.

The structural representations show some surprising results. First, CS only depends on the contemporaneous impacts of the transitory shock and does not reflect price discovery at all. Second, IS, MIS and PDS respond to be permanent and transitory shocks and are not pure measures of price discovery.

To help sort out the confounding effects of responses to transitory shocks, Yan and Zivot (2010) proposed the information leadership measure (IL hereafter), a combination of the IS and CS measures, which in the restricted case with uncorrelated reduced-form residuals is defined as

$$\mathrm{IL}_1 = \left|\frac{\mathrm{IS}_1/\mathrm{CS}_1}{\mathrm{IS}_2/\mathrm{CS}_2}\right| = \left|\frac{\beta_1^P}{\beta_2^P}\right| = \left|\frac{d_{0,1}^P}{d_{0,2}^P}\right|, \quad \mathrm{IL}_2 = \left|\frac{\mathrm{IS}_2/\mathrm{CS}_2}{\mathrm{IS}_1/\mathrm{CS}_1}\right| = \left|\frac{\beta_2^P}{\beta_1^P}\right| = \left|\frac{d_{0,2}^P}{d_{0,1}^P}\right|. \tag{2.29}$$

For the more general case with possibly correlated reduced-form errors, Shen et al. (2024a) proposed the Price Information Leadership (PIL) measure, a combination of the PDS and CS measures, to unravel the confounding effects of transitory noise:

$$\mathrm{PIL}_1 = \left|\frac{\mathrm{PDS}_1/\mathrm{CS}_1}{\mathrm{PDS}_2/\mathrm{CS}_2}\right| = \left|\frac{\beta_1^P}{\beta_2^P}\right| = \left|\frac{d_{0,1}^P}{d_{0,2}^P}\right|, \quad \mathrm{PIL}_2 = \left|\frac{\mathrm{PDS}_2/\mathrm{CS}_2}{\mathrm{PDS}_1/\mathrm{CS}_1}\right| = \left|\frac{\beta_2^P}{\beta_1^P}\right| = \left|\frac{d_{0,2}^P}{d_{0,1}^P}\right|. \tag{2.30}$$

Following Putniņš (2013), Shen et al. (2025) converted PIL into the Price Information Leadership Share (PILS) measure

$$\begin{aligned}
\mathrm{PILS}_1 &= \frac{\mathrm{PIL}_1}{\mathrm{PIL}_1 + \mathrm{PIL}_2} = \frac{(\beta_1^P)^2}{(\beta_1^P)^2 + (\beta_2^P)^2} = \frac{(d_{0,1}^P)^2}{(d_{0,1}^P)^2 + (d_{0,2}^P)^2}, \\
\mathrm{PILS}_2 &= \frac{\mathrm{PIL}_2}{\mathrm{PIL}_1 + \mathrm{PIL}_2} = \frac{(\beta_2^P)^2}{(\beta_1^P)^2 + (\beta_2^P)^2} = \frac{(d_{0,2}^P)^2}{(d_{0,1}^P)^2 + (d_{0,2}^P)^2}.
\end{aligned} \tag{2.31}$$

PILS is a pure measure of price discovery as it depends only on the contemporaneous impacts of the permanent shock. A venue has a high PILS if it has a stronger contemporaneous impact from the permanent shock than the other venue.

Independently of Shen et al. (2025), Lautier et al. (2024) proposed the Covariance Information Share (CovIS) and the Squared Covariance Information Share (CovISQ) measures of price

discovery. CovISQ is equivalent to PILS and CovIS is given by

$$\begin{aligned}\text{CovIS}_1 &= \frac{\beta_1^P}{\beta_1^P + \beta_2^P} = \frac{d_{0,1}^P}{d_{0,1}^P + d_{0,2}^P},\\ \text{CovIS}_2 &= \frac{\beta_2^P}{\beta_1^P + \beta_2^P} = \frac{d_{0,2}^P}{d_{0,1}^P + d_{0,2}^P}.\end{aligned} \tag{2.32}$$

The interpretation of PILS and CovIS can be misleading in the unusual case when $d_{0,i}^P$ is large and negative. Following Shen and Zivot (2024) we define the the instantaneous pricing error rule

$$\text{PIES}_1 = \frac{(\beta_1^P - 1)^2}{(\beta_1^P - 1)^2 + (\beta_2^P - 1)^2} = \frac{(d_{0,1}^P - 1)^2}{(d_{0,1}^P - 1)^2 + (d_{0,2}^P - 1)^2}, \tag{2.33}$$

$$\text{PIES}_2 = \frac{(\beta_2^P - 1)^2}{(\beta_1^P - 1)^2 + (\beta_2^P - 1)^2} = \frac{(d_{0,2}^P - 1)^2}{(d_{0,1}^P - 1)^2 + (d_{0,2}^P - 1)^2}. \tag{2.34}$$

to define the leading market as the market with the smaller instantaneous absolute pricing error $d_{0,i}^P - 1$, instead of the one with a larger instantaneous absolute response $d_{0,i}^P$.

Table 2.1 summarizes the different measures of price discovery in terms of both reduced-form and structural parameters. Note that the structural representations depend only on $\mathbf{D_0}$, σ_P^2 and σ_T^2. The reduced-form representations depend only on $\boldsymbol{\Omega}$ and $\boldsymbol{\psi}$.

2.3.5 Price Discovery Impulse Response Functions

The structural cointegration model in (2.1) offers a convenient tool to directly characterize how multiple market prices for the same underlying asset discover the new equilibrium price following the arrival of new information. The tool is based on the impulse response function (IRF) of a market's price to the permanent innovation of the common efficient price implied by the cointegration model. Based on the work in Yan and Zivot (2007) and the extension in Lautier et al. (2024), the expected cumulative price response h periods after a one unit increase to the permanent shock η_t^P, is given by

$$\text{SCIRF}_i^P(h) = \frac{\partial E_t\left[p_{i,t+h}\right]}{\partial \eta_t^P} = \sum_{l=0}^{h} \frac{\partial E_t\left[\Delta p_{i,t+l}\right]}{\partial \eta_t^P} = \sum_{l=0}^{h} d_{l,i}^P, \quad k = 0, 1, \cdots \tag{2.35}$$

Table 2.1: Price Discovery Measures

Structural Form	**Reduced Form**
Information Share: Hasbrouck (1995)	
Market 1 ranked First (F) in Cholesky decomposition:	
$\mathrm{IS}_1^F = \frac{d_{0,1}^{P2}}{d_{0,1}^{P2}+d_{0,1}^{T2}}$	$\mathrm{IS}_1^F = \frac{(\psi_1\sigma_1+\rho\psi_2\sigma_2)^2}{\psi_1^2\sigma_1^2+2\psi_1\psi_2\sigma_{12}+\psi_2^2\sigma_2^2}$
$\mathrm{IS}_2^F = \frac{d_{0,1}^{T2}}{d_{0,1}^{P2}+d_{0,1}^{T2}}$	$\mathrm{IS}_2^F = \frac{\psi_2^2\sigma_2^2\left(1-\rho^2\right)}{\psi_1^2\sigma_1^2+2\psi_1\psi_2\sigma_{12}+\psi_2^2\sigma_2^2}$
Market 1 ranked Second (S) in Cholesky decomposition:	
$\mathrm{IS}_1^S = \frac{d_{0,2}^{T2}}{d_{0,2}^{P2}+d_{0,2}^{T2}}$	$\mathrm{IS}_1^S = \frac{\psi_1^2\sigma_1^2\left(1-\rho^2\right)}{\psi_1^2\sigma_1^2+2\psi_1\psi_2\sigma_{12}+\psi_2^2\sigma_2^2}$
$\mathrm{IS}_2^S = \frac{d_{0,2}^{P2}}{d_{0,2}^{P2}+d_{0,2}^{T2}}$	$\mathrm{IS}_2^S = \frac{(\rho\psi_1\sigma_1+\psi_2\sigma_2)^2}{\psi_1^2\sigma_1^2+2\psi_1\psi_2\sigma_{12}+\psi_2^2\sigma_2^2}$
Hasbrouck's IS Arithmetic Averages:	
$\overline{\mathrm{IS}}_1 = \frac{\mathrm{IS}_1^F+\mathrm{IS}_1^S}{2}; \overline{\mathrm{IS}}_2 = \frac{\mathrm{IS}_2^F+\mathrm{IS}_2^S}{2}$	$\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}$ is a correlation coefficient
Covariance Information Share: Lautier et al. (2024)	
$\mathrm{CovIS}_1 = \frac{d_{0,1}^{P}}{d_{0,1}^{P}+d_{0,2}^{P}}$	$\mathrm{CovIS}_1 = \frac{\psi_1\sigma_1^2+\psi_2\sigma_{12}}{\psi_1\sigma_1^2+(\psi_1+\psi_2)\sigma_{12}+\psi_2\sigma_2^2}$
$\mathrm{CovIS}_2 = \frac{d_{0,2}^{P}}{d_{0,1}^{P}+d_{0,2}^{P}}$	$\mathrm{CovIS}_2 = \frac{\psi_1\sigma_{12}+\psi_2\sigma_2^2}{\psi_1\sigma_1^2+(\psi_1+\psi_2)\sigma_{12}+\psi_2\sigma_2^2}$
Covariance Information Share: Quadiratic variation	
$\mathrm{CovISQ}_1 = \frac{d_{0,1}^{P2}}{d_{0,1}^{P2}+d_{0,2}^{P2}}$	$\mathrm{CovISQ}_1 = \frac{(\psi_1\sigma_1^2+\psi_2\sigma_{12})^2}{(\psi_1\sigma_1^2+\psi_2\sigma_{12})^2+(\psi_1\sigma_{12}+\psi_2\sigma_2^2)^2}$
$\mathrm{CovISQ}_2 = \frac{d_{0,2}^{P2}}{d_{0,1}^{P2}+d_{0,2}^{P2}}$	$\mathrm{CovISQ}_2 = \frac{(\psi_1\sigma_{12}+\psi_2\sigma_2^2)^2}{(\psi_1\sigma_1^2+\psi_2\sigma_{12})^2+(\psi_1\sigma_{12}+\psi_2\sigma_2^2)^2}$
Component Shares/ Weights in common component	

Structural Form	**Reduced Form**
$\mathrm{CS}_1 = \dfrac{d_{0,2}^T}{d_{0,2}^T - d_{0,1}^T}$	$\mathrm{CS}_1 = \dfrac{\psi_1}{\psi_1 + \psi_2}$
$\mathrm{CS}_2 = -\dfrac{d_{0,1}^T}{d_{0,2}^T - d_{0,1}^T}$	$\mathrm{CS}_2 = \dfrac{\psi_2}{\psi_1 + \psi_2}$
Modified Information Share: Lien and Shrestha (2009)	
	$\mathrm{MIS}_i = \dfrac{\left([\boldsymbol{\psi}'\mathbf{F}^*]_i\right)^2}{\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi}} = \dfrac{\left(\boldsymbol{\psi}_i^*\right)^2}{\sum_{i=1}^n \left(\boldsymbol{\psi}_i^*\right)^2}$ where ψ_i^* represents the i^{th} element of the row matrix $\boldsymbol{\psi}'\mathbf{F}^*$
Price Discovery Share: Sultan and Zivot (2015)	
$\mathrm{PDS}_1 = \dfrac{d_{0,1}^P d_{0,2}^T}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}$	$\mathrm{PDS}_i = \dfrac{\psi_i \frac{\partial \sigma_\eta(\boldsymbol{\psi})}{\partial \psi_i}}{\sigma_\eta(\boldsymbol{\psi})}$
$\mathrm{PDS}_2 = -\dfrac{d_{0,1}^T d_{0,2}^P}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}$	where $\sigma_\eta(\boldsymbol{\psi}) = (\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi})^{1/2}$
Information Leadership Share: Yan and Zivot (2010); Putniņš (2013)	
$\mathrm{IL}_1 = -\frac{d_{0,1}^T\left(d_{0,1}^{P2} d_{0,2}^{P2}\sigma_P^4 + 2d_{0,1}^{P2} d_{0,2}^{T2}\sigma_P^2\sigma_T^2 + d_{0,1}^{T2} d_{0,2}^{T2}\sigma_T^4\right)}{d_{0,2}^T\left(d_{0,1}^{P2} d_{0,2}^{P2}\sigma_P^4 + 2d_{0,2}^{P2} d_{0,1}^{T2}\sigma_P^2\sigma_T^2 + d_{0,1}^{T2} d_{0,2}^{T2}\sigma_T^4\right)}, \quad \mathrm{IL}_2 = \frac{1}{\mathrm{IL}_1}$	
$\mathrm{ILS}_1 = \dfrac{\mathrm{IL}_1}{\mathrm{IL}_1 + \mathrm{IL}_2}, \quad \mathrm{ILS}_2 = \dfrac{\mathrm{IL}_2}{\mathrm{IL}_1 + \mathrm{IL}_2}$	
Modified Information Leadership Share: Shen et al. (2025)	
$\mathrm{MIL}_1 = \left\lvert\dfrac{\mathrm{MIS}_1/\mathrm{CS}_1}{\mathrm{MIS}_2/\mathrm{CS}_2}\right\rvert, \quad \mathrm{MIL}_2 = \left\lvert\dfrac{\mathrm{MIS}_2/\mathrm{CS}_2}{\mathrm{MIS}_1/\mathrm{CS}_1}\right\rvert$	
$\mathrm{MILS}_1 = \dfrac{\mathrm{MIL}_1}{\mathrm{MIL}_1 + \mathrm{MIL}_2}, \quad \mathrm{MILS}_2 = \dfrac{\mathrm{MIL}_2}{\mathrm{MIL}_1 + \mathrm{MIL}_2}$	
Price Information Leadership Share: Shen et al. (2025)	

Structural Form	Reduced Form
$\text{PIL}_1 = \left\lvert \frac{\text{PDS}_1/\text{CS}_1}{\text{PDS}_2/\text{CS}_2} \right\rvert = \left\lvert \frac{d^P_{0,1}}{d^P_{0,2}} \right\rvert, \quad \text{PIL}_2 = \left\lvert \frac{\text{PDS}_2/\text{CS}_2}{\text{PDS}_1/\text{CS}_1} \right\rvert = \left\lvert \frac{d^P_{0,2}}{d^P_{0,1}} \right\rvert$	
$\text{PILS}_1 = \frac{\text{PIL}_1}{\text{PIL}_1 + \text{PIL}_2} = \frac{(d^P_{0,1})^2}{(d^P_{0,1})^2 + (d^P_{0,2})^2},$	
$\text{PILS}_2 = \frac{\text{PIL}_2}{\text{PIL}_1 + \text{PIL}_2} = \frac{(d^P_{0,2})^2}{(d^P_{0,1})^2 + (d^P_{0,2})^2}$	
Price Instantaneous Efficiency Share: Shen and Zivot (2024)	
$\text{PIES}_1 = \frac{(d^P_{0,2} - 1)^2}{(d^P_{0,1} - 1)^2 + (d^P_{0,2} - 1)^2},$	
$\text{PIES}_2 = \frac{(d^P_{0,1} - 1)^2}{(d^P_{0,1} - 1)^2 + (d^P_{0,2} - 1)^2}$	

where $d_{i,l}^P$ is the coefficient on the lth lag of $d_i^P(L)$. This sequence begins with $d_{0,i}^P$ and ends with the long-run impact of 1. The impulse response dynamics between these two points describe the price discovery process. Questions such as how quickly the lagging market catches up, whether it can overtake the leading market, and the time it takes for markets to converge, are answered by examining the time path $\text{SCIRF}_i^P(h)$.

The SCIRF with respect to the permanent shock η_t^P is identified from the reduced-form VECM without restrictions. However, the SCIRF with respect to the transitory shock η_t^T

$$\text{SCIRF}_i^T(h) = \frac{\partial E_t\left[p_{i,t+h}\right]}{\partial \eta_t^T} = \sum_{l=0}^{h} \frac{\partial E_t\left[\Delta p_{i,t+l}\right]}{\partial \eta_t^T} = \sum_{l=0}^{h} d_{l,i}^T, \quad k = 0, 1, \cdots \tag{2.36}$$

is not identified without additional restrictions. Yan and Zivot (2007) identify $\text{SCIRF}_i^T(h)$ using the Gonzalo-Ng decomposition, and Lautier et al. (2024) identify $\text{SCIRF}_i^T(h)$ using the assumption that $\sigma_P^2 = \sigma_T^2$.

2.4 Bayesian Inference

Consider the bivariate reduced-form VECM from Equation (2.8):

$$\Delta \mathbf{p}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{p}_{t-1} + \sum_{j=1}^{k} \boldsymbol{\Gamma}_j \Delta \mathbf{p}_{t-j} + \boldsymbol{\varepsilon}_t, \tag{2.37}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$ is the error correction vector, $\boldsymbol{\Gamma}_j (j = 1, \ldots, k)$ are the short-run coefficient matrices, and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is the vector of reduced-form VECM residuals with $E\left[\boldsymbol{\varepsilon}_t\right] = \mathbf{0}$ and

$$E\left[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s'\right] = \begin{cases} 0, & \text{if } t \neq s, \\ \boldsymbol{\Omega}, & \text{otherwise.} \end{cases}$$

We assume only one cointegrating vector $\boldsymbol{\beta} = (1, -1)'$, removing cointegration-space identification concerns.[4]

[4] If $\boldsymbol{\beta}$ were treated as unknown, priors defined directly on the cointegration space (e.g., MACG/Grassmann-manifold priors) and collapsed/parameter-augmented Gibbs schemes are recommended; see Koop et al. (2009) and Strachan and Inder (2004).

We use the following notation. Let $\mathbf{\Gamma} \equiv [\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_k]$ and collect regressors as $\mathbf{X}_t = [\,\boldsymbol{\beta}'\mathbf{p}_{t-1}, \Delta\mathbf{p}_{t-1}, \ldots, \Delta\mathbf{p}_{t-k}]$, so that $\mathbf{Y} = \mathbf{XB} + \mathbf{U}$ with $\mathbf{U} \sim \mathcal{MN}(0, \mathbf{I}_T \otimes \mathbf{\Omega})$ and $\mathbf{B} = [\,\boldsymbol{\alpha}',\ \mathbf{\Gamma}']'$.

2.4.1 Priors

We implement sensible informative and noninformative priors over the VECM parameters as in Koop et al. (2009). The priors are summarized below.

1. **Adjustment (loading) matrix $\boldsymbol{\alpha}$.** A shrinkage matrix-normal prior (see, e.g., Bauwens et al. (2000)) centered at zero:

$$\boldsymbol{\alpha} \mid \mathbf{\Omega}, \nu \ \sim\ \mathcal{MN}\Big(\mathbf{0},\ \nu(\boldsymbol{\beta}'\boldsymbol{\beta})^{-1} \otimes \mathbf{\Omega}\Big), \tag{2.38}$$

   where $\nu > 0$ controls tightness, i.e., the degree of shrinkage. The noninformative limit is obtained by letting $1/\nu \to 0$ (Geweke, 1996; Strachan, 2003; Villani, 2005). The scalar v can either be selected by the researcher subjectively or can be treated as a hierarchical prior with v being an unknown parameter.

2. **Short-run dynamics $\mathbf{\Gamma}$.** A conjugate Gaussian prior

$$\mathrm{vec}(\mathbf{\Gamma}) \ \sim\ \mathcal{N}\big(\underline{\mathbf{\Gamma}},\ \underline{\mathbf{V}}_\Gamma\big), \tag{2.39}$$

   with $\underline{\mathbf{V}}_\Gamma$ chosen as diffuse (or Minnesota-style, if desired).

3. **Innovation covariance $\mathbf{\Omega}$.** standard noninformative prior

$$p(\mathbf{\Omega}) \ \propto\ |\mathbf{\Omega}|^{-(n+1)/2}, \tag{2.40}$$

   and an inverse-Wishart prior is also straightforward to accommodate.

2.4.2 Posterior Simulation

With $\boldsymbol{\beta}$ fixed, the model is linear in $(\boldsymbol{\alpha}, \mathbf{\Gamma})$ and conjugate. Let $\mathbf{Z}_t = \boldsymbol{\beta}'\mathbf{p}_{t-1}$ and define residuals $\boldsymbol{\varepsilon}_t = \Delta\mathbf{p}_t - \big(\boldsymbol{\alpha}\mathbf{Z}_t + \sum_{j=1}^{k} \mathbf{\Gamma}_j \Delta\mathbf{p}_{t-j}\big)$. Stacking over t yields $\mathbf{U} = [\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_T]'$. Thus, we can derive

efficient and simple posterior computation through use of Gibbs sampler:

Full conditionals:

$$\boldsymbol{\Omega} \mid \boldsymbol{\alpha}, \boldsymbol{\Gamma}, \text{data} \;\sim\; IW\Big(\mathbf{U}'\mathbf{U} + \nu^{-1}\boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{\beta})\boldsymbol{\alpha}',\; T + r\Big), \tag{2.41}$$

$$\operatorname{vec}(\boldsymbol{\alpha}) \mid \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \text{data} \;\sim\; \mathcal{N}\big(\,\overline{\boldsymbol{\alpha}},\; \overline{\mathbf{V}}_{\alpha}\,\big) \tag{2.42}$$

$$\text{where} \qquad \overline{\mathbf{V}}_{\alpha} = \Big[\nu^{-1}(\boldsymbol{\beta}'\boldsymbol{\beta}) \otimes \boldsymbol{\Omega}^{-1} + (\mathbf{Z}\mathbf{Z}' \otimes \boldsymbol{\Omega}^{-1})\Big]^{-1} \tag{2.43}$$

$$\text{and} \qquad \overline{\boldsymbol{\alpha}} \;=\; \overline{\mathbf{V}}_{\alpha}\, \operatorname{vec}\big(\boldsymbol{\Omega}^{-1}\Delta\mathbf{p}\, Z'\big), \qquad \mathbf{Z} = [\mathbf{Z}_1, \ldots, \mathbf{Z}_T], \tag{2.44}$$

$$\operatorname{vec}(\boldsymbol{\Gamma}_j) \mid \boldsymbol{\Omega}, \boldsymbol{\alpha}, \text{data} \;\sim\; \mathcal{N}\big(\,\overline{\boldsymbol{\Gamma}}_j,\; \overline{\mathbf{V}}_{\Gamma_j}\,\big) \tag{2.45}$$

$$\text{where} \qquad \overline{\mathbf{V}}_{\Gamma_j} \;=\; \Big[\underline{\mathbf{V}}_{\Gamma_j}^{-1} + \big(\Delta\mathbf{p}_{t-j}\Delta\mathbf{p}_{t-j}' \otimes \boldsymbol{\Omega}^{-1}\big)\Big]^{-1} \tag{2.46}$$

$$\text{and} \qquad \overline{\boldsymbol{\Gamma}}_j \;=\; \overline{\mathbf{V}}_{\Gamma_j}\Big[\underline{\mathbf{V}}_{\Gamma_j}^{-1}\underline{\boldsymbol{\Gamma}}_j + \operatorname{vec}\big(\boldsymbol{\Omega}^{-1}\Delta\mathbf{p}\,\Delta\mathbf{p}_{t-j}'\big)\Big]. \tag{2.47}$$

Gibbs sampler: Initialize $\boldsymbol{\Omega}^{(0)}$ (e.g., diagonal) and fix $\boldsymbol{\beta} = (1, -1)'$. For $s = 1, \ldots, S$:

1. Draw $\boldsymbol{\alpha}^{(s)} \sim \mathcal{MN}\big(\overline{\boldsymbol{\alpha}}, \overline{\mathbf{V}}_{\alpha}\big)$.
2. Draw $\boldsymbol{\Gamma}^{(s)} \sim \mathcal{N}\big(\overline{\boldsymbol{\Gamma}}, \overline{\mathbf{V}}_{\Gamma}\big)$.
3. Draw $\boldsymbol{\Omega}^{(s)} \sim IW(\cdot)$ using (2.41).

Discard burn-in to obtain draws $\{\boldsymbol{\alpha}^{(s)}, \boldsymbol{\Gamma}^{(s)}, \boldsymbol{\Omega}^{(s)}\}_{s=1}^{S^\star}$. To implement the MCMC algorithm, we fix $\boldsymbol{\beta} = (1, -1)'$ and run the three-block Gibbs sampler above with priors (2.38)–(2.40). Unless noted, we use weakly informative $\underline{\mathbf{V}}_{\Gamma}$ and assess sensitivity to ν (noninformative limit $1/\nu \to 0$). For each posterior draw, we impose stability of the implied VECM and discard any draw that violates this condition. Price discovery measures are then computed from the remaining draws, with point estimates reported as posterior means and uncertainty summarized by 15/85% credible intervals.

2.4.3 Price-discovery Computation

Classical price–discovery measures are typically computed at point estimates of the VECM and a single orthogonalization, which obscures uncertainty and identification sensitivity. In contrast, our

Bayesian procedure averages each measure over the posterior of $(\boldsymbol{\alpha}, \boldsymbol{\Gamma}, \boldsymbol{\Omega})$, yielding credible sets that reflect parameter and covariance uncertainty. This subsection lays out (i) the reduced-form VECM estimation and its mapping to the reduced-form VMA used for order-invariant measures, and (ii) the structural analysis used to recover contemporaneous impact matrices and permanent/transitory shocks for structurally interpretable measures.

Reduced-form VECM and VMA. Let the bivariate VECM for log prices be

$$\Delta \mathbf{p}_t = \boldsymbol{\alpha}\boldsymbol{\beta}' \mathbf{p}_{t-1} + \sum_{i=1}^{k} \boldsymbol{\Gamma}_i \Delta \mathbf{p}_{t-i} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}), \tag{2.48}$$

with fixed $\boldsymbol{\beta} = (1, -1)'$. For a given posterior draw $\boldsymbol{\theta}^{(s)} = \{\boldsymbol{\alpha}^{(s)}, \boldsymbol{\Gamma}^{(s)}, \boldsymbol{\Omega}^{(s)}\}$, the associated VAR is

$$\mathbf{p}_t = \sum_{j=1}^{k} \mathbf{A}_j^{(s)} \mathbf{p}_{t-j} + \boldsymbol{\varepsilon}_t^{(s)}, \qquad \boldsymbol{\varepsilon}_t^{(s)} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}^{(s)}), \tag{2.49}$$

obtained by the standard mapping from (2.48). Let $\mathbf{A}^{(s)}(L) = I - \sum_{j=1}^{k} \mathbf{A}_j^{(s)} L^j$ and $\mathbf{A}^{(s)}(1) = \mathbf{I} - \sum_{j=1}^{k} \mathbf{A}_j^{(s)}$. The reduced-form VMA for differences is

$$\Delta \mathbf{p}_t = \boldsymbol{\Psi}^{(s)}(L)\, \boldsymbol{\varepsilon}_t^{(s)} = \sum_{\ell=0}^{\infty} \boldsymbol{\Psi}_\ell^{(s)}\, \boldsymbol{\varepsilon}_{t-\ell}^{(s)}, \qquad \boldsymbol{\Psi}^{(s)}(L) = I + \sum_{\ell=1}^{\infty} \boldsymbol{\Psi}_\ell^{(s)} L^\ell, \tag{2.50}$$

and the corresponding level VMA is obtained by summing the differenced coefficients. The long-run (permanent) reduced-form impact on levels is the $n \times n$ matrix

$$\boldsymbol{\Psi}^{(s)}(1) \equiv \sum_{\ell=0}^{\infty} \left(\sum_{j=0}^{\ell} \boldsymbol{\Psi}_j^{(s)} \right) = \boldsymbol{\beta}_\perp \Big(\boldsymbol{\alpha}_\perp^{(s)\prime} \mathbf{A}^{(s)}(1)\, \boldsymbol{\beta}_\perp \Big)^{-1} \boldsymbol{\alpha}_\perp^{(s)\prime}, \tag{2.51}$$

where $\boldsymbol{\beta}_\perp$ and $\boldsymbol{\alpha}_\perp^{(s)}$ denote orthogonal complements.[5] Two classes of measures follow from the reduced form:

- *Order-invariant, reduced-form measures* such as CS, MIS and PDS use only $\boldsymbol{\psi}^{(s)}$ and $\boldsymbol{\Omega}^{(s)}$

[5] For $r = 1$, $\boldsymbol{\Psi}^{(s)}(1)$ has identical rows (the Beveridge–Nelson common-trend loadings). We denote that common row by $\boldsymbol{\psi}^{(s)\prime} = (\psi_1^{(s)}, \ldots, \psi_n^{(s)})$.

(and, for MIS, the correlation structure). Example:

$$\mathrm{CS}_i^{(s)} = \frac{\psi_i^{(s)}}{\mathbf{1}'\boldsymbol{\psi}^{(s)}}, \qquad \mathrm{PDS}_i^{(s)} = \frac{\psi_i^{(s)}\,[\boldsymbol{\Omega}^{(s)}\boldsymbol{\psi}^{(s)}]_i}{\boldsymbol{\psi}^{(s)\prime}\boldsymbol{\Omega}^{(s)}\boldsymbol{\psi}^{(s)}}.$$

- *Orthogonalization-based measures* such as Hasbrouck's IS compute the long-run price impact of orthogonal shocks. With Cholesky factor $B_\ell^{(s)}$ of $\boldsymbol{\Omega}^{(s)}$ under ordering ℓ,

$$\boldsymbol{\Theta}_\ell^{(s)} \equiv \boldsymbol{\Psi}^{(s)}(1)B_\ell^{(s)}, \qquad \mathrm{IS}_{i,\ell}^{(s)} = \frac{\Theta_{\ell,\,i1}^{(s)2}}{\sum_{j=1}^{n}\Theta_{\ell,\,j1}^{(s)2}},$$

  and bounds are obtained by varying ℓ (e.g., forward and reverse orderings).

Structural analysis and contemporaneous impacts. To disentangle permanent news from transitory noise, we adopt a structural moving-average (SMA) representation for $\Delta\mathbf{p}_t$,

$$\Delta\mathbf{p}_t = \mathbf{D}(L)\,\boldsymbol{\eta}_t = \mathbf{D}_0\boldsymbol{\eta}_t + \mathbf{D}_1\boldsymbol{\eta}_{t-1} + \cdots, \qquad \boldsymbol{\eta}_t = (\eta_t^P, \eta_t^T)', \tag{2.52}$$

where $\boldsymbol{\eta}_t$ are serially and mutually uncorrelated with diagonal covariance $\mathbf{C} = \mathrm{diag}(\sigma_P^2, \sigma_T^2)$, and $\mathbf{D}_0$ is invertible; see Yan and Zivot (2010). The permanent innovation η_t^P moves levels one-for-one in the long run while η_t^T has no long-run effect:

$$\mathbf{D}(1) = \sum_{j=0}^{\infty}\mathbf{D}_j = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}. \tag{2.53}$$

Reduced-form shocks and structural shocks are linked by $\boldsymbol{\varepsilon}_t = \mathbf{D}_0\boldsymbol{\eta}_t$, so the $n \times 2$ matrix $\mathbf{D}_0$ collects the contemporaneous responses of $\Delta\mathbf{p}_t$ to (η_t^P, η_t^T).

Let $\boldsymbol{\Psi}^{(s)}(1)$ be as in (2.51) and denote its common row by $\boldsymbol{\psi}^{(s)\prime}$. Following Yan and Zivot (2007) and Lautier et al. (2024), the permanent shock is the projection of reduced-form innovations onto the common trend,

$$\eta_t^{P,(s)} = \boldsymbol{\psi}^{(s)\prime}\boldsymbol{\varepsilon}_t^{(s)}, \qquad \mathrm{var}\big(\eta_t^{P,(s)}\big) = \sigma_P^{2,(s)} = \boldsymbol{\psi}^{(s)\prime}\boldsymbol{\Omega}^{(s)}\boldsymbol{\psi}^{(s)}. \tag{2.54}$$

The contemporaneous impact matrix $\mathbf{D}_0^{(s)}$ is then pinned down by long-run restrictions and the covariance of reduced-form shocks. In the bivariate ($n = 2$, $r = 1$) case, Lautier et al. (2024) show that

$$\mathbf{D}_0^{(s)} = \frac{1}{\sigma_P^{2,(s)}} \begin{bmatrix} \psi_1^{(s)}\sigma_1^{2,(s)} + \psi_2^{(s)}\sigma_{12}^{(s)} & \psi_2^{(s)}\,\sigma_1^{(s)}\sigma_2^{(s)}\sqrt{1-\rho^{(s)2}} \\ \psi_1^{(s)}\sigma_{12}^{(s)} + \psi_2^{(s)}\sigma_2^{2,(s)} & -\psi_1^{(s)}\,\sigma_1^{(s)}\sigma_2^{(s)}\sqrt{1-\rho^{(s)2}} \end{bmatrix}, \qquad \rho^{(s)} = \frac{\sigma_{12}^{(s)}}{\sigma_1^{(s)}\sigma_2^{(s)}}, \tag{2.55}$$

which is observationally equivalent to the long-run identification in Yan and Zivot (2007) for the permanent shock (we abstract from identification of the transitory shock because our IRFs and shares are driven by the first column of $\mathbf{D}_0^{(s)}$).

Measures from reduced and structural forms. Given $\{\boldsymbol{\psi}^{(s)}, \boldsymbol{\Omega}^{(s)}\}$ (reduced form) and $\mathbf{D}_0^{(s)}$ (structural), we compute the following for each draw:

(a) **Reduced-form measures.**

- *Hasbrouck IS*: $\mathrm{IS}_{i,\ell}^{(s)}$ from $\boldsymbol{\Theta}_\ell^{(s)} = \boldsymbol{\Psi}^{(s)}(1)B_\ell^{(s)}$.
- *Component shares (CS)*: $\mathrm{CS}_i^{(s)} = \psi_i^{(s)} / \sum_j \psi_j^{(s)}$.
- *Modified information share (MIS)*: spectral variant based on the correlation structure (order-invariant).
- *Price discovery share (PDS)*: variance attribution of the efficient-price innovation, $\mathrm{PDS}_i^{(s)} = \psi_i^{(s)}\,[\boldsymbol{\Omega}^{(s)}\boldsymbol{\psi}^{(s)}]_i / (\boldsymbol{\psi}^{(s)\prime}\boldsymbol{\Omega}^{(s)}\boldsymbol{\psi}^{(s)})$.

(b) **Structural measures.**

- *Covariance information share (CovIS)*: uses the first column of $\mathbf{D}_0^{(s)}$,

$$\mathrm{CovIS}_i^{(s)} = \frac{d_{0,i}^{P,(s)}}{\sum_{j=1}^n d_{0,j}^{P,(s)}}, \quad \text{where} \quad \mathbf{D}_0^{(s)} = \left[d^{P,(s)},\, d^{T,(s)}\right].$$

  A quadratic-variation analogue (CovISQ) is proportional to $d_{0,i}^{P,(s)2}$.

- *Information-leadership families (ILS/MILS/PILS).* Combine speed (IS/MIS/PDS) with precision (CS) to mitigate noise confounding.

Posterior summarization and uncertainty. Within each MCMC iteration we:

1. Draw $(\boldsymbol{\alpha}^{(s)}, \boldsymbol{\Gamma}^{(s)}, \boldsymbol{\Omega}^{(s)})$ from their full conditionals;
2. Construct $\boldsymbol{\psi}^{(s)}$ via (2.51) and, if needed, IS bounds via orthogonalizations;
3. Recover $\mathbf{D}_0^{(s)}$ using (2.55) (bivariate case) and compute structural measures (CovIS, PIES, PILS/CovISQ, ILS variants) and SCIRFs.

We report posterior means and 15/85% credible intervals for all measures. This presentation integrates parameter and covariance uncertainty.[6]

2.4.4 Marginal Likelihoods and Posterior Odds

Most empirical studies of price discovery report point estimates and, at best, classical confidence intervals or bootstrap bands (see, e.g., Hasbrouck, 1995; Baillie et al., 2002; Patel et al., 2020). As emphasized by Kass and Raftery (1995) and Berger and Pericchi (1996), however, questions such as "do the two markets contribute equally to price discovery?" are naturally formulated as competing models rather than as large-sample approximations around a single estimate. Within a Bayesian VECM framework, we can exploit marginal likelihoods to conduct formal model comparison and derive posterior odds for economically meaningful hypotheses on price discovery, while fully accounting for parameter and covariance uncertainty.

Hypotheses and Bayes factors. Let $\vartheta(\theta)$ denote a scalar price-discovery functional of the reduced-form and structural parameters $\theta = (\boldsymbol{\alpha}, \boldsymbol{\Gamma}, \boldsymbol{\Omega}, \mathbf{D}_0)$, such as PILS_1, CovIS_1 or IS_1 (for a fixed orthogonalization). A prototypical hypothesis is

$$H_0 : \ \vartheta(\theta) = \tfrac{1}{2} \qquad \text{vs.} \qquad H_1 : \ \vartheta(\theta) \neq \tfrac{1}{2},$$

which corresponds to equal versus unequal price-discovery contributions of the two markets. In a Bayesian setting, this is implemented as a comparison between a *constrained* model M_0 and an

[6] If $\boldsymbol{\beta}$ is estimated rather than fixed, one can sample on the cointegration *space* and employ a collapsed (or parameter-augmented) Gibbs sampler to avoid local non-identification and improve mixing; see Koop et al. (2009). Our two-ETF application fixes $\boldsymbol{\beta} = (1, -1)'$, so the simple three-block Gibbs suffices.

unconstrained model M_1:

$$M_0 : \ \theta \in \Theta_0 = \{\theta : \vartheta(\theta) = \tfrac{1}{2}\}, \qquad M_1 : \ \theta \in \Theta_1 \supset \Theta_0.$$

The posterior odds in favor of M_0 are

$$\frac{p(M_0 \mid y)}{p(M_1 \mid y)} \;=\; \underbrace{\frac{m(y \mid M_0)}{m(y \mid M_1)}}_{\text{Bayes factor } BF_{01}} \;\times\; \underbrace{\frac{p(M_0)}{p(M_1)}}_{\text{prior odds}}, \tag{2.56}$$

where $m(y \mid M) = \int p(y \mid \theta, M) p(\theta \mid M)\, d\theta$ is the marginal likelihood. Following Kass and Raftery (1995), we interpret BF_{01} using standard evidence scales (e.g., "substantial," "strong," "decisive" support for H_0 or H_1).

Computing marginal likelihoods. For each model $M \in \{M_0, M_1\}$ we obtain posterior draws of the reduced-form VECM parameters by the Gibbs sampler described in Section 4.2. The corresponding marginal likelihood $m(y \mid M)$ is then computed using established methods from the Bayesian time series literature:

- *Chib's method* (Chib, 1995):

  $$\log m(y \mid M) = \log p(y \mid \theta^\star, M) + \log p(\theta^\star \mid M) - \log p(\theta^\star \mid y, M),$$

  where $\theta^\star$ is a high-posterior-density point, the likelihood and prior ordinates are analytic for our VECM, and the posterior ordinate is evaluated via Rao–Blackwellized Gibbs conditionals.

- *Savage–Dickey density ratio* (Dickey, 1971; Verdinelli and Wasserman, 1995): when M_0 is nested in M_1 by a sharp restriction such as $\delta_1 - \delta_2 = 0$, and the prior under M_1 admits a proper density at that point, the Bayes factor can be written as

  $$BF_{01} = \frac{p(\delta_1 - \delta_2 = 0 \mid y, M_1)}{p(\delta_1 - \delta_2 = 0 \mid M_1)},$$

  where the posterior ordinate is estimated from the MCMC draws of $\delta_1 - \delta_2$ induced by the structural mapping from $(\boldsymbol{\alpha}, \boldsymbol{\Gamma}, \boldsymbol{\Omega})$ to $\mathbf{D}_0$ (see also Koop, 2017).

Posterior probabilities and regions of practical equivalence. Beyond Bayes factors, posterior draws of any functional $\vartheta(\theta)$ yield direct probability statements. For instance, for dominance hypotheses we compute

$$\Pr\left(\vartheta(\theta) > \tfrac{1}{2} \mid y\right) = \Pr\left(\text{market 1 leads according to } \vartheta \mid y\right),$$

which translates into posterior odds and, under equal prior odds, into an implied Bayes factor. Following Kruschke (2018), we define "regions of practical equivalence" (ROPE), e.g.

$$\Pr\left(|\vartheta(\theta) - \tfrac{1}{2}| \leq \epsilon \mid y\right),$$

for small tolerances ϵ, to assess whether the two markets are practically indistinguishable in terms of price discovery.

2.5 Illustration and Simulation

2.5.1 Illustration: A Partial Price Adjustment Model

To illustrate the empirical performance of the discussed price discovery measures, we consider the stylized partial adjustment micro-structure model utilized by Amihud and Mendelson (1987), Hasbrouck and Ho (1987) and Yan and Zivot (2010):

$$p_{1t} = p_{1,t-1} + \delta_1 \left(m_t - p_{1,t-1}\right) + b_{0,1}^T \eta_t^T \tag{2.57}$$

$$p_{2t} = p_{2,t-1} + \delta_2 \left(m_t - p_{2,t-1}\right) + b_{0,2}^T \eta_t^T \tag{2.58}$$

$$m_t = m_{t-1} + \eta_t^P, \boldsymbol{\eta}_t = \left(\eta_t^P, \eta_t^T\right)' \sim i.i.d.N\left(\mathbf{0}, \begin{bmatrix} \sigma_P^2 & 0 \\ 0 & \sigma_T^2 \end{bmatrix}\right). \tag{2.59}$$

Solving for Δp_{it} gives

$$\Delta p_{it} = d_i^P(L)\eta_t^P + d_i^T(L)\eta_t^T,$$

where

$$d_1^P(L) = [1-(1-\delta_1)\,L]^{-1}\,\delta_1, \quad d_2^P(L) = [1-(1-\delta_2)\,L]^{-1}\,\delta_2,$$
$$d_1^T(L) = [1-(1-\delta_1)\,L]^{-1}\,(1-L)b_{0,1}^T, \quad d_2^T(L) = [1-(1-\delta_2)\,L]^{-1}\,(1-L)b_{0,2}^T.$$

The SMA representation of the partial adjustment model is determined from the appropriate elements of the lag polynomials $d_i^P(L)$ and $d_i^T(L)$. In particular, the initial impact and the long-run impact matrices are given by:

$$\mathbf{D}_0 = \begin{pmatrix} d_{0,1}^P & d_{0,1}^T \\ d_{0,2}^P & d_{0,2}^T \end{pmatrix} = \begin{pmatrix} \delta_1 & b_{0,1}^T \\ \delta_2 & b_{0,2}^T \end{pmatrix}, \quad \mathbf{D}(1) = \begin{pmatrix} d_1^P(1) & d_1^T(1) \\ d_2^P(1) & d_2^T(1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}. \tag{2.60}$$

The initial responses of the two assets to a one-unit permanent shock are equal to δ_i $(i = 1, 2)$ and the long-run responses to one-unit permanent shock are equal one. The asset with a larger value of the initial response is identified as the leader in the process of the price discovery. Rewriting Δp_{it} in the reduced form as Eq. (2.8), we have

$$\alpha_1 = \frac{-b_{0,1}^T\delta_1\delta_2}{b_{0,1}^T\delta_2 - b_{0,2}^T\delta_1}, \quad \alpha_2 = \frac{-b_{0,2}^T\delta_1\delta_2}{b_{0,2}^T\delta_1 - b_{0,1}^T\delta_2},$$
$$\varepsilon_1 = \delta_1\eta_t^P + b_{0,1}^T\eta_t^T, \quad \varepsilon_2 = \delta_2\eta_t^P + b_{0,2}^T\eta_t^T.$$

The CS, PDS, CovIS, PILS/CovISQ and PIES measures are independent of any reduced form

correlation and have structural representations:

$$\begin{aligned}
\text{CS}_1 &= \frac{b_{0,2}^T}{b_{0,2}^T - b_{0,1}^T}, \quad \text{CS}_2 = \frac{-b_{0,1}^T}{b_{0,2}^T - b_{0,1}^T} \\
\text{PDS}_1 &= \frac{\delta_1 b_{0,2}^T}{\delta_1 b_{0,2}^T - b_{0,1}^T \delta_2}, \quad \text{PDS}_2 = \frac{-b_{0,1}^T \delta_2}{\delta_1 b_{0,2}^T - b_{0,1}^T \delta_2} \\
\text{PIL}_1 &= \left|\frac{\delta_1}{\delta_2}\right|, \quad \text{PIL}_2 = \left|\frac{\delta_2}{\delta_1}\right| \\
\text{PILS}_1 &= \text{CovISQ}_1 = \frac{\delta_1^2}{\delta_1^2 + \delta_2^2}, \quad \text{PILS}_2 = \text{CovISQ}_2 = \frac{\delta_2^2}{\delta_1^2 + \delta_2^2} \\
\text{CovIS}_1 &= \frac{\delta_1}{\delta_1 + \delta_2}, \quad \text{CovIS}_2 = \frac{\delta_2}{\delta_1 + \delta_2} \\
\text{PIES}_1 &= \frac{(\delta_1 - 1)^2}{(\delta_1 - 1)^2 + (\delta_2 - 1)^2}, \quad \text{PIES}_2 = \frac{(\delta_2 - 1)^2}{(\delta_1 - 1)^2 + (\delta_2 - 1)^2}
\end{aligned}$$

provided $b_{0,2}^T \neq b_{0,1}^T$. Note that when $\delta_1, \delta_2 \in [0, 1]$, then PIES is equivalent to PILS.

When the reduced-form innovations, ε_t, are uncorrelated, the IS metric is unique. It can be shown that $\text{cov}(\varepsilon_{1t}, \varepsilon_{2t}) = 0$ when

$$\frac{\sigma_T^2}{\sigma_P^2} = \frac{d_{0,1}^P d_{0,2}^P}{-d_{0,1}^T d_{0,2}^T} = \frac{\delta_1 \delta_2}{-b_{0,1}^T b_{0,2}^T}$$

and all elements of $\mathbf{D_0}$ are non-zero (hence $|\mathbf{D_0}| \neq 0$). Then the IS measures are uniquely defined as follows:

$$\text{IS}_1 = \frac{\delta_1 b_{0,2}^T}{\delta_1 b_{0,2}^T - b_{0,1}^T \delta_2}, \quad \text{IS}_2 = \frac{-b_{0,1}^T \delta_2}{\delta_1 b_{0,2}^T - b_{0,1}^T \delta_2}.$$

The whole price discovery process shows the progressive absorption of permanent and transitory shocks. We compute and draw the SCIRFs for both the permanent (P) and the transitory (T) shocks:

$$\text{SCIRF}^P(h) = \sum_{i=0}^{h} d_i^P \quad \text{and} \quad \text{SCIRF}^T(h) = \sum_{i=0}^{h} d_i^T.$$

2.5.2 Simulation Evidence from Partial Price Adjustment Model

In this subsection, we evaluate the empirical performance of the discussed price discovery measures using simulated data from different parameterizations of the stylized partial adjustment model. Specifically, we evaluate Bayesian posterior mean estimates like they are classical estimates. The

Bayesian toolkit produces sensible results with good frequentist sampling properties[7].

Equal price discovery share

To give hypotheses in Section 4.4 a structural interpretation, we exploit the partial price-adjustment microstructure model, where the contemporaneous impact matrix and long-run impact matrix are

$$\mathbf{D}_0 = \begin{pmatrix} \delta_1 & b_{0,1}^T \\ \delta_2 & b_{0,2}^T \end{pmatrix}, \qquad \mathbf{D}(1) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix},$$

so that the speed-of-adjustment parameters (δ_1, δ_2) fully characterize the initial responses of the two markets to permanent news. As shown in Section 2.4.3 and Table 2.1, several order-invariant measures admit simple structural forms:

$$\text{PILS}_1 = \frac{\delta_1^2}{\delta_1^2 + \delta_2^2}, \quad \text{PIES}_1 = \frac{(\delta_1 - 1)^2}{(\delta_1 - 1)^2 + (\delta_2 - 1)^2}, \quad \text{CovIS}_1 = \frac{\delta_1}{\delta_1 + \delta_2},$$

and analogously for market 2. In particular, under this structural model,

$$H_0 : \ \delta_1 = \delta_2 \quad \Longleftrightarrow \quad \text{PILS}_1 = \text{PIES}_1 = \text{CovIS}_1 = \tfrac{1}{2},$$

so testing equal price-discovery contributions is equivalent to testing equal speeds of adjustment to the permanent shock.[8]

For the symmetric case $\delta_1 = \delta_2 = 0.5$ (Figure 2.1), the posterior distributions of these measures are tightly centered at 0.5, leading to Bayes factors that support H_0 and large ROPE probabilities around 0.5. Specifically, the posterior distributions of PILS, PIES, CovIS and CovISQ look normally distributed with posterior mean equal to 0.5. The results provide evidence that PILS, PIES, CovIS, and CovISQ are only dependent on $(d_{0,1}^P, d_{0,2}^P) = (\delta_1, \delta_2)$ as shown in Table 2.1, which correctly represent the ratio of initial responses to the permanent shock. On the other hand, the other measures depend on $(d_{0,1}^T, d_{0,2}^T) = (b_{0,1}^T, b_{0,2}^T) = (0.8, -0.2)$, which distorts the results. When $d_{0,1}^P = d_{0,2}^P$, IS = CS = PDS and these measures only reflect the response of one market to transi-

[7] Please see more figures in Appendix B.1

[8] This contrasts with IS/CS-type measures whose structural expressions also depend on the noise parameters $b_{0,i}^T$ and on the reduced-form covariance $\mathbf{\Omega}$.

tory frictions relative to the other market. When $\delta_1 > \delta_2$, the posterior mass shifts toward one for market 1, and the associated Bayes factors provide "strong" or "decisive" evidence against equal discovery, in line with the structural interpretation of δ_i as the speed of adjustment to the permanent shock. The results here confirm that posterior distributions of the measures PILS/CovISQ, PIES, and CovIS provide clear and accurate inference regarding price discovery.

Full simulation results

For results in this section, we set $\delta_2 = 1-\delta_1$ and let δ_1 takes values from 0.9 to 0.1 with a reduction of 0.1. When $\delta_1 > \delta_2$ (i.e., $\delta_1 > 0.5$), Market 1 has a greater speed of price discovery than Market 2. We set the variance of the permanent shock to unity ($\sigma_P^2 = 1$). We simulate 1000 samples of 21600 observations of two price series for each given value of δ_1. For each simulated sample, we estimate the Bayesian VECM and calculate the corresponding price discovery measures. We take 10,000 interations and 5,000 burn-in using the Gibbs sampler.

Consider first the results in Figures 2.2 and 2.3, when $(b_0^{T,1}, b_0^{T,2}) = (0.8, -0.2)$ and $\sigma_T^2 = 10$, which display the full posterior distributions of the price discovery measures for two representative parameter settings with strong asymmetry in the adjustment speeds. Figure 2.2 corresponds to $\boldsymbol{\delta} = (0.8, 0.2)$, where market 1 adjusts much faster to permanent news than market 2. The posterior densities of PILS/CovISQ, PIES, and CovIS for market 1 are sharply concentrated near one (around 0.8 or above), with virtually no overlap with the corresponding densities for market 2, which are concentrated near zero. The posteriors for these measures thus deliver an unambiguous probabilistic statement that market 1 dominates price discovery. In contrast, the posteriors for IS, MIS and PDS are more dispersed and are centered close to 0.5. The ILS and MILS panels show very spiky densities close to 0 and 1, reflecting the fact that these measures are essentially indicator functions of leadership: they almost always pick market 1 as the leader when $\delta_1 = 0.8$, but their discrete nature makes them less informative about the magnitude of the leadership gap.

Figure 2.3 flips the speeds to $\boldsymbol{\delta} = (0.2, 0.8)$. As expected from the structural formulas, the posterior distributions of PILS/CovISQ, PIES and CovIS mirror those in Figure 2.2: the mass now lies near zero for market 1 and near one for market 2, clearly indicating that market 2 leads the price discovery process. The densities of ILS and MILS become more diffuse and cluster around values slightly above and below 0.5, showing that, in the presence of strongly correlated residuals,

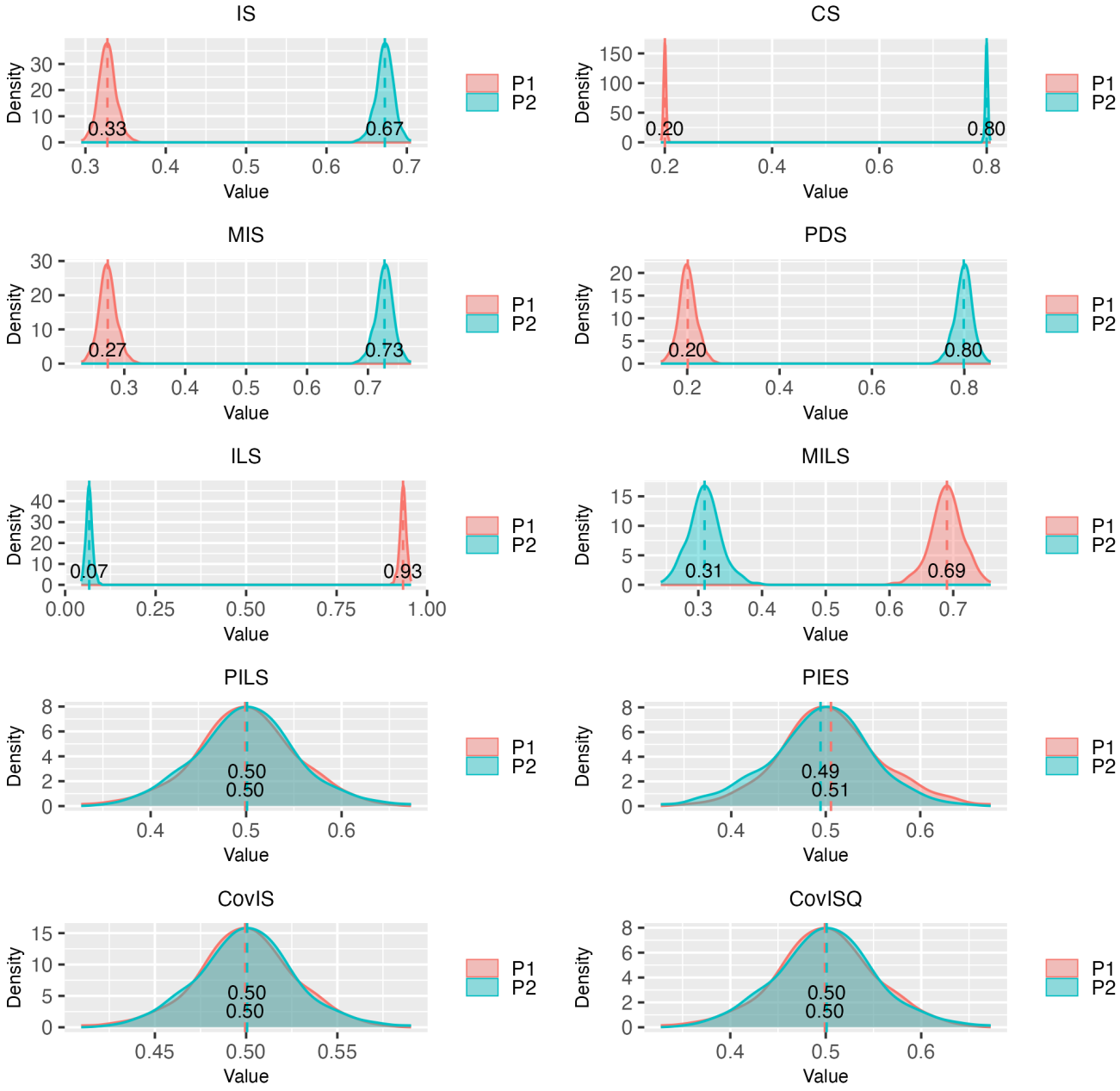

Figure 2.1: Distribution of price discovery measurements for $\delta = (0.5, 0.5)$ in panel C.

these combined measures can struggle to separate the two markets as cleanly as PILS, PIES and CovIS do.

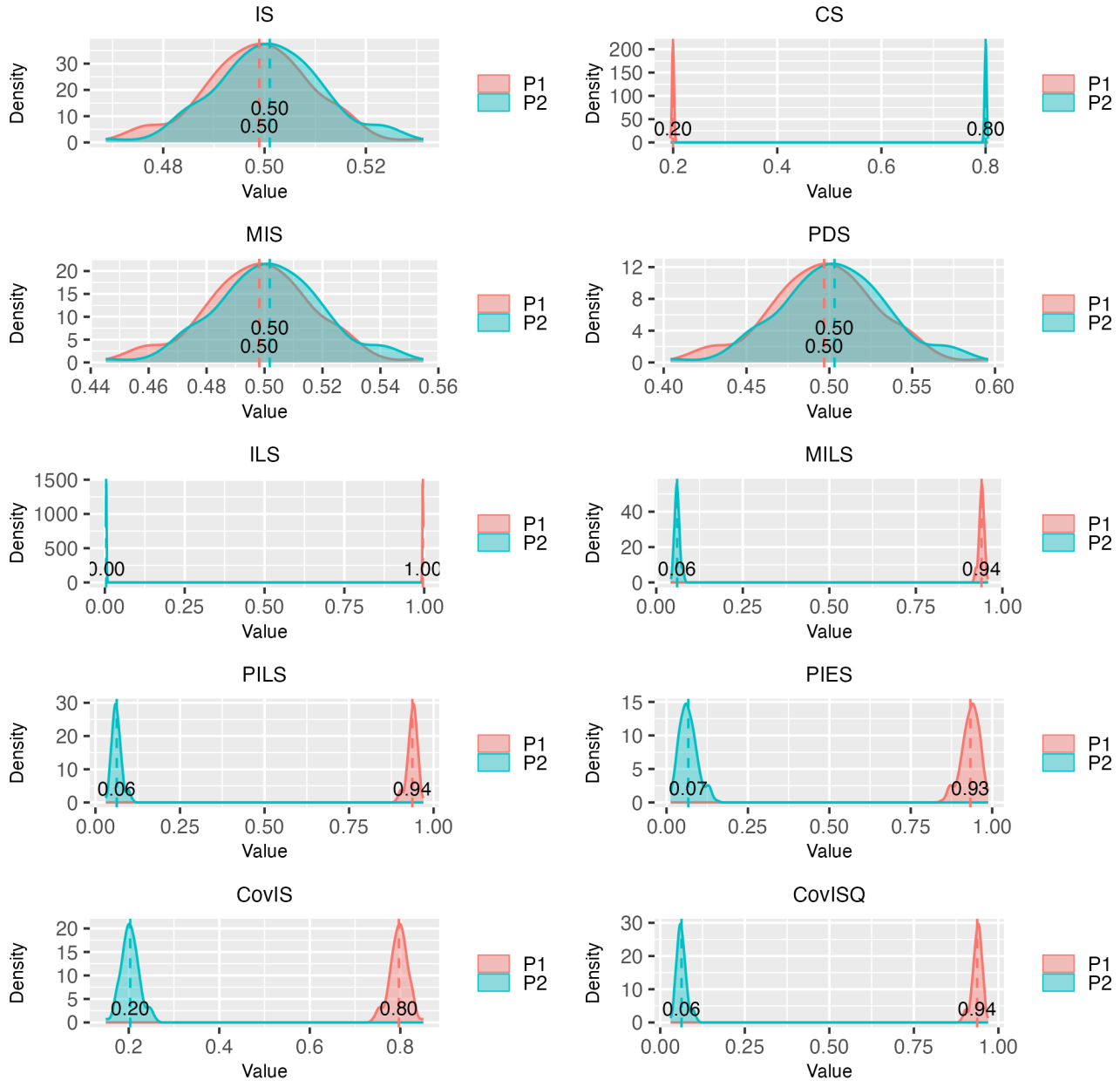

Figure 2.2: Distribution of price discovery measurements for $\delta = (0.8, 0.2)$ in panel C.

Under the above setting, we firstly consider the sub-case when both markets' responses to transitory shocks are equal and set as $\left(b_{0,1}^T, b_{0,2}^T\right) = (0.5, -0.5)$. We further set the variance of the transitory shock to $\sigma_T^2 = \frac{\delta_1 \delta_2}{-b_{0,1}^T b_{0,2}^T}$ which makes the reduced-form residuals uncorrelated for these two markets. We summarize the averages of each price discovery measure over the 1000 samples

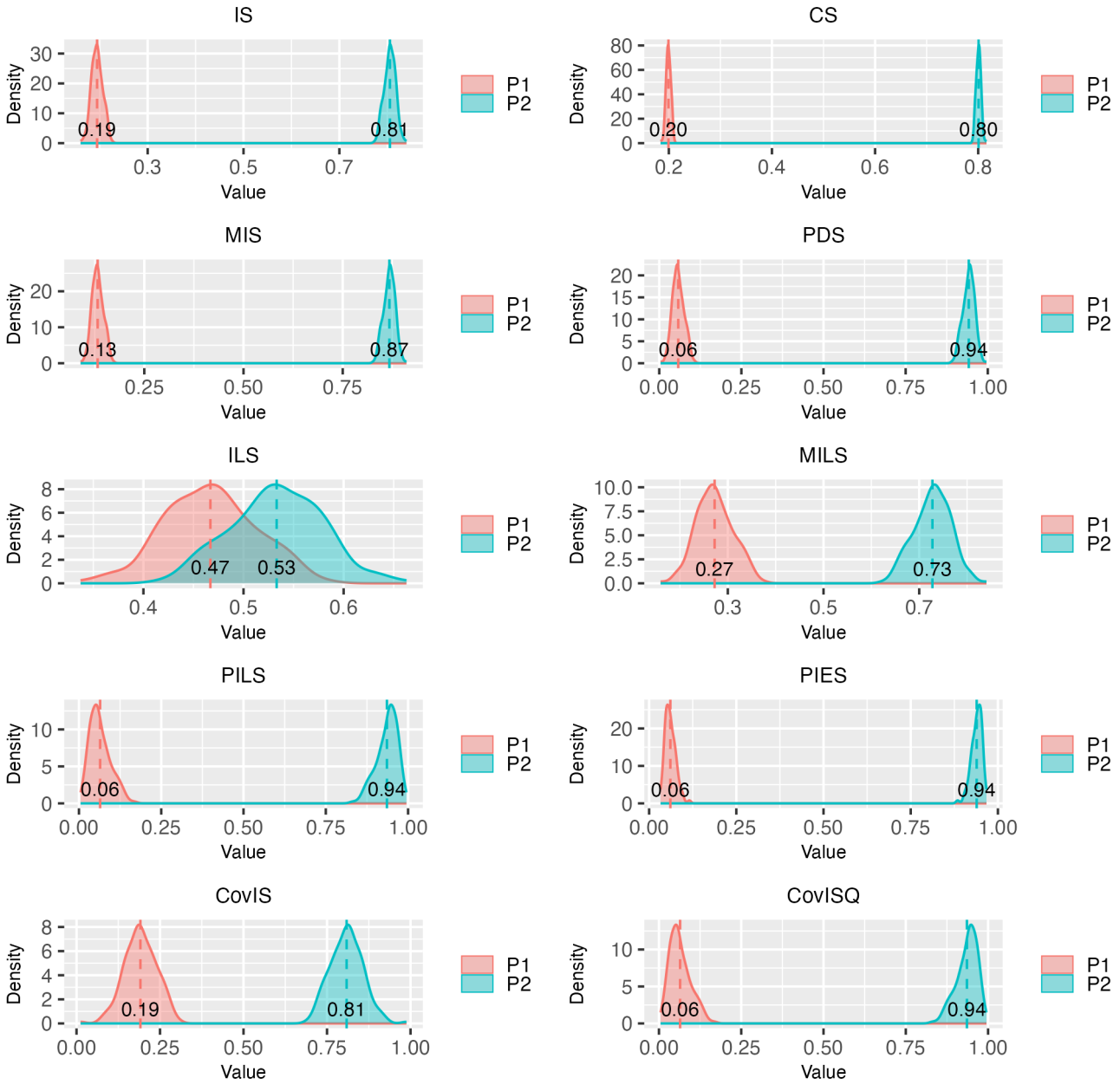

Figure 2.3: Distribution of price discovery measurements for $\delta = (0.2, 0.8)$ in panel C.

and report results in Panel A of Table 2.2.

We now examine the finite-sample (sampling) behavior of the Bayesian estimators. As we can see from Panel A of Table 2.2, when the two markets have the same response to transitory shocks, $(b_{0,1}^T, b_{0,2}^T) = (0.5, -0.5)$, and the reduced–form innovations are uncorrelated by construction, all IS–type measures line up very closely with the structural speeds of adjustment (δ_1, δ_2). For each value of δ_1 the posterior means of IS, MIS and PDS for market 1 are almost identical and essentially equal to δ_1 (and likewise for market 2). The combined measures ILS, MILS, CovIS, PILS/CovISQ and PIES also collapse to the same numbers, reflecting the fact that, under diagonal $\mathbf{\Omega}$ and symmetric transitory responses, these measures are algebraically equivalent to IS an PDS. By contrast, the CS measure is anchored by the transitory loadings and therefore remains close to 0.5 for both markets in this symmetric case. For $\delta_1 > 0.5$ the IS, MIS and PDS shares for market 1 all exceed 0.5, and the combined measures correctly identify market 1 as the price leader; for $\delta_1 < 0.5$ the roles are reversed and all measures correctly select market 2 as the leader. Thus, in the benchmark case with uncorrelated innovations and symmetric noise, classical measures and their "information leadership" combinations all behave as intended and recover the true ordering implied by (δ_1, δ_2).

Panel B of Table 2.2 breaks this symmetry by increasing the transitory loading of market 1 to $(b_{0,1}^T, b_{0,2}^T) = (0.8, -0.2)$, while still choosing $\sigma_T^2 = \delta_1\delta_2/(-b_{0,1}^T b_{0,2}^T)$ so that the reduced–form residuals remain uncorrelated. In this case market 1 is both faster (for $\delta_1 > 0.5$) and noisier. The table shows that IS, MIS, and PDS are now systematically pulled toward the less noisy market 2. For instance, when $\delta_1 = 0.7$ or $\delta_1 = 0.6$ the true leader is market 1, yet the mean IS/MIS/PDS shares for market 1 are only around 0.37 and 0.27, respectively, and would lead an applied researcher to incorrectly conclude that market 2 dominates price discovery. This reproduces the structural critique in Yan and Zivot (2010) and Shen et al. (2024a): variance–decomposition measures based on reduced–form residuals confound informational leadership with the avoidance of transitory noise. In contrast, the combined measures ILS, MILS, CovIS, PILS/CovISQ and PIES strip out this noise component and continue to assign the larger share to the market with the higher δ_i; they always pick market 1 as the leader for $\delta_1 > 0.5$, market 2 for $\delta_1 < 0.5$, and split leadership roughly evenly when $\delta_1 = \delta_2 = 0.5$.

Panel C of Table 2.2 holds the asymmetric transitory responses $(b_{0,1}^T, b_{0,2}^T) = (0.8, -0.2)$ fixed

Table 2.2: Price Discovery Measures from Partial Price Adjustment Model

This table reports price discovery measure estimates from the price data simulated from the following 2-market model:

$$p_{1t} = p_{1,t-1} + \delta_1 (m_t - p_{1,t-1}) + b_{0,1}^T \eta_t^T,$$
$$p_{2t} = p_{2,t-1} + \delta_2 (m_t - p_{2,t-1}) + b_{0,2}^T \eta_t^T$$

where $m_t = m_{t-1} + \eta_t^P, \boldsymbol{\eta_t} = \left(\eta_t^P, \eta_t^T\right)'$ are Gaussian white noise with diagonal covariance matrix $\text{diag}\,(\sigma_P^2, \sigma_T^2)$. The simulation parameterization is set as $\delta_2 = 1 - \delta_1, \sigma_P^2 = 1$. We simulate 1000 samples of 21600 observations. For each sample, we take 5000 burn-in and save 10000 iterations.

Panel A: $(b_0^{\top,1}, b_0^{\top,2}) = (0.5, -0.5), \quad \sigma_T^2 = \frac{\delta_1 \delta_2}{-b_{0,1}^T b_{0,2}^T}$

	IS		CS		MIS		PDS		ILS		MILS		PILS/CovISQ		PIES		CovLS	
δ_1	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
0.90	0.90	0.10	0.50	0.50	0.90	0.10	0.90	0.10	0.99	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.90	0.10
0.80	0.80	0.20	0.50	0.50	0.80	0.20	0.80	0.20	0.94	0.06	0.94	0.06	0.94	0.06	0.94	0.06	0.80	0.20
0.70	0.70	0.30	0.50	0.50	0.70	0.30	0.70	0.30	0.69	0.31	0.69	0.31	0.85	0.15	0.85	0.15	0.70	0.30
0.60	0.60	0.40	0.50	0.50	0.60	0.40	0.60	0.40	0.50	0.50	0.50	0.50	0.69	0.31	0.69	0.31	0.60	0.40
0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.31	0.69	0.31	0.69	0.50	0.50	0.50	0.50	0.50	0.50
0.40	0.40	0.60	0.50	0.50	0.40	0.60	0.40	0.60	0.16	0.84	0.16	0.84	0.31	0.69	0.31	0.69	0.40	0.60
0.30	0.30	0.70	0.50	0.50	0.30	0.70	0.30	0.70	0.06	0.94	0.06	0.94	0.16	0.84	0.16	0.84	0.30	0.70
0.20	0.20	0.80	0.50	0.50	0.20	0.80	0.20	0.80	0.06	0.94	0.06	0.94	0.06	0.94	0.06	0.94	0.20	0.80
0.10	0.10	0.90	0.50	0.50	0.10	0.90	0.10	0.90	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.99	0.10	0.90

Panel B: $(b_0^{\top,1}, b_0^{\top,2}) = (0.8, -0.2), \quad \sigma_T^2 = \frac{\delta_1 \delta_2}{-b_{0,1}^T b_{0,2}^T}$

	IS		CS		MIS		PDS		ILS		MILS		PILS/CovISQ		PIES		CovLS	
δ_1	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
0.90	0.69	0.31	0.20	0.80	0.69	0.31	0.69	0.31	0.99	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.90	0.10
0.80	0.50	0.50	0.20	0.80	0.50	0.50	0.50	0.50	0.94	0.06	0.94	0.06	0.94	0.06	0.94	0.06	0.80	0.20
0.70	0.37	0.63	0.20	0.80	0.37	0.63	0.37	0.63	0.84	0.16	0.84	0.16	0.84	0.16	0.84	0.16	0.70	0.30
0.60	0.27	0.73	0.20	0.80	0.27	0.73	0.27	0.73	0.69	0.31	0.69	0.31	0.69	0.31	0.69	0.31	0.60	0.40
0.50	0.20	0.80	0.20	0.80	0.20	0.80	0.20	0.80	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
0.40	0.14	0.86	0.20	0.80	0.14	0.86	0.14	0.86	0.31	0.69	0.31	0.69	0.31	0.69	0.31	0.69	0.40	0.60
0.30	0.10	0.90	0.20	0.80	0.10	0.90	0.10	0.90	0.15	0.85	0.15	0.85	0.15	0.85	0.15	0.85	0.30	0.70
0.20	0.06	0.94	0.20	0.80	0.06	0.94	0.06	0.94	0.06	0.94	0.06	0.94	0.06	0.94	0.06	0.94	0.20	0.80
0.10	0.03	0.97	0.20	0.80	0.03	0.97	0.03	0.97	0.01	0.99	0.01	0.99	0.01	0.99	0.01	0.99	0.10	0.90

Panel C: $(b_0^{\top,1}, b_0^{\top,2}) = (0.8, -0.2), \quad \sigma^2 = 10$

	IS		CS		MIS		PDS		ILS		MILS		PILS/CovISQ		PIES		CovLS	
δ_1	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
0.90	0.54	0.46	0.20	0.80	0.59	0.41	0.69	0.31	0.96	0.04	0.97	0.03	0.99	0.01	0.99	0.01	0.90	0.10
0.80	0.50	0.50	0.20	0.80	0.50	0.50	0.50	0.50	0.94	0.06	0.94	0.06	0.94	0.06	0.94	0.06	0.80	0.20
0.70	0.44	0.56	0.20	0.80	0.41	0.59	0.37	0.63	0.91	0.09	0.89	0.11	0.84	0.16	0.94	0.06	0.70	0.30
0.60	0.38	0.62	0.20	0.80	0.34	0.66	0.27	0.73	0.86	0.14	0.80	0.20	0.69	0.31	0.69	0.31	0.60	0.40
0.50	0.33	0.67	0.20	0.80	0.27	0.73	0.20	0.80	0.79	0.21	0.69	0.31	0.50	0.50	0.50	0.50	0.50	0.50
0.40	0.28	0.72	0.20	0.80	0.22	0.78	0.14	0.86	0.70	0.30	0.55	0.45	0.31	0.69	0.31	0.69	0.40	0.60
0.30	0.23	0.77	0.20	0.80	0.17	0.83	0.10	0.90	0.59	0.41	0.40	0.60	0.16	0.84	0.16	0.84	0.30	0.70
0.20	0.20	0.80	0.20	0.80	0.13	0.87	0.06	0.94	0.49	0.51	0.28	0.72	0.06	0.94	0.06	0.94	0.20	0.80
0.10	0.17	0.83	0.20	0.80	0.10	0.90	0.03	0.97	0.39	0.61	0.17	0.83	0.02	0.98	0.06	0.94	0.10	0.90

Notes: 1. Numbers shown are the averages of price discovery measure estimates of 1000 samples.
2. In the case of estimating VECM using MLE method (Shen et al., 2025), ILS, MILS and PILS for P_1 are 0.16 when $\delta_1 = 0.30$ in panel B; MILS is equal to 0.27 when $\delta_1 = 0.20$ in panel C.

but inflates the variance of the transitory shock to $\sigma_T^2 = 10$, thereby inducing substantial correlation in the reduced–form innovations. In this more realistic setting the simple IS, MIS and PDS measures diverge from one another, reflecting their different ways of allocating variance when shocks are correlated, and they provide little guidance on which market leads. The combined measures ILS and MILS (constructed from IS/MIS and CS) perform better than the raw IS/MIS shares, but they now tend to over-select market 1 as the leader even when $\delta_1 \leq 0.4$. By contrast, the structurally motivated CovIS, PILS and PIES, which depend only on the permanent-impact parameters $(d_{0,1}^P, d_{0,2}^P) = (\delta_1, \delta_2)$, continue to track the true leadership pattern perfectly: they assign higher shares to market 1 whenever $\delta_1 > \delta_2$, higher shares to market 2 whenever $\delta_1 < \delta_2$, and approximately equal shares when $\delta_1 = \delta_2 = 0.5$. In all three panels CovIS, PILS and PIES coincide because the estimated values of $d_{0,1}^P$ and $d_{0,2}^P$ lie in $[0, 1]$, making the two definitions equivalent in this parameterization.

Based on

$$\varepsilon_1 = \delta_1 \eta_t^P + b_{0,1}^T \eta_t^T, \quad \varepsilon_2 = \delta_2 \eta_t^P + b_{0,2}^T \eta_t^T$$

we have

$$Var(\varepsilon_i) = \delta_i^2 \sigma_P^2 + (b_{0,i}^T)^2 \sigma_T^2, \quad i = 1, 2$$
$$Cov(\varepsilon_1, \varepsilon_2) = \delta_1 \delta_2 \sigma_P^2 + b_{0,1}^T b_{0,2}^T \sigma_T^2$$

and the correlation coefficient between these two errors is

$$\rho = \frac{Cov(\varepsilon_1, \varepsilon_2)}{\sqrt{Var(\varepsilon_1) Var(\varepsilon_2)}}$$

Table A.4 and Figure 2.4 shows that the estimated values ρ are the same as true values.

Figure 2.5 plots the cumulative price responses (SCIRFs) implied by the partial price adjustment model subsequent to one unit innovation in the efficient price. The values of initial permanent SCIRF of each price are equal to $\delta_i, (i = 1, 2)$, moreover, the permanent and transitory SCIRFs converge to 1 and 0 respectively as shown in Equation (2.60).

To sum up, the simulation results highlight that structurally grounded measures based on the permanent component (PILS/CovISQ, PIES, CovIS) are robust to both asymmetries in noise and correlation in reduced–form shocks, and provide a much clearer Bayesian characterization of price discovery leadership than other measures. More importantly, only the measure CovIS, PILS and

Table 2.3: Comparison of correlation coefficient from Partial Price Adjustment Model

	Panel A		Panel B		Panel C	
δ_1	ρ	$\hat{\rho}$	ρ	$\hat{\rho}$	ρ	$\hat{\rho}$
0.9	0	0	0	0	-0.88	-0.88
0.8	0	0	0	0	-0.82	-0.82
0.7	0	0	0	0	-0.76	-0.76
0.6	0	0	0	0	-0.70	-0.70
0.5	0	0	0	0	-0.65	-0.65
0.4	0	0	0	0	-0.61	-0.61
0.3	0	0	0	0	-0.58	-0.58
0.2	0	0	0	0	-0.56	-0.56
0.1	0	0	0	0	-0.54	-0.54

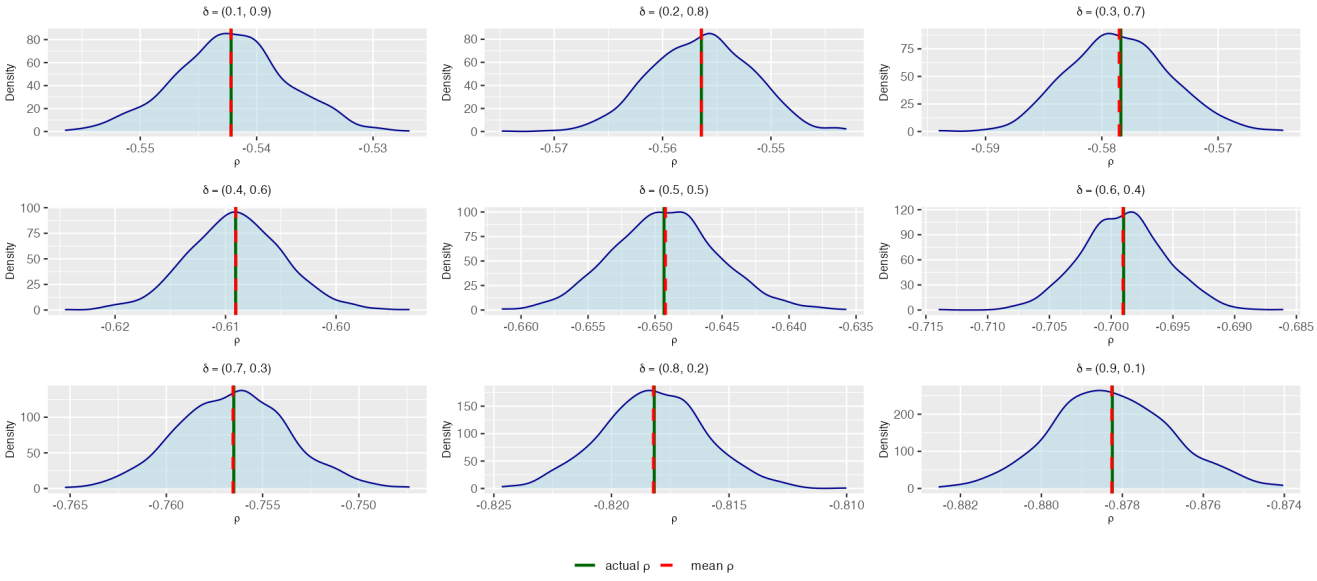

Figure 2.4: Comparison of correlation coefficients in panel C.

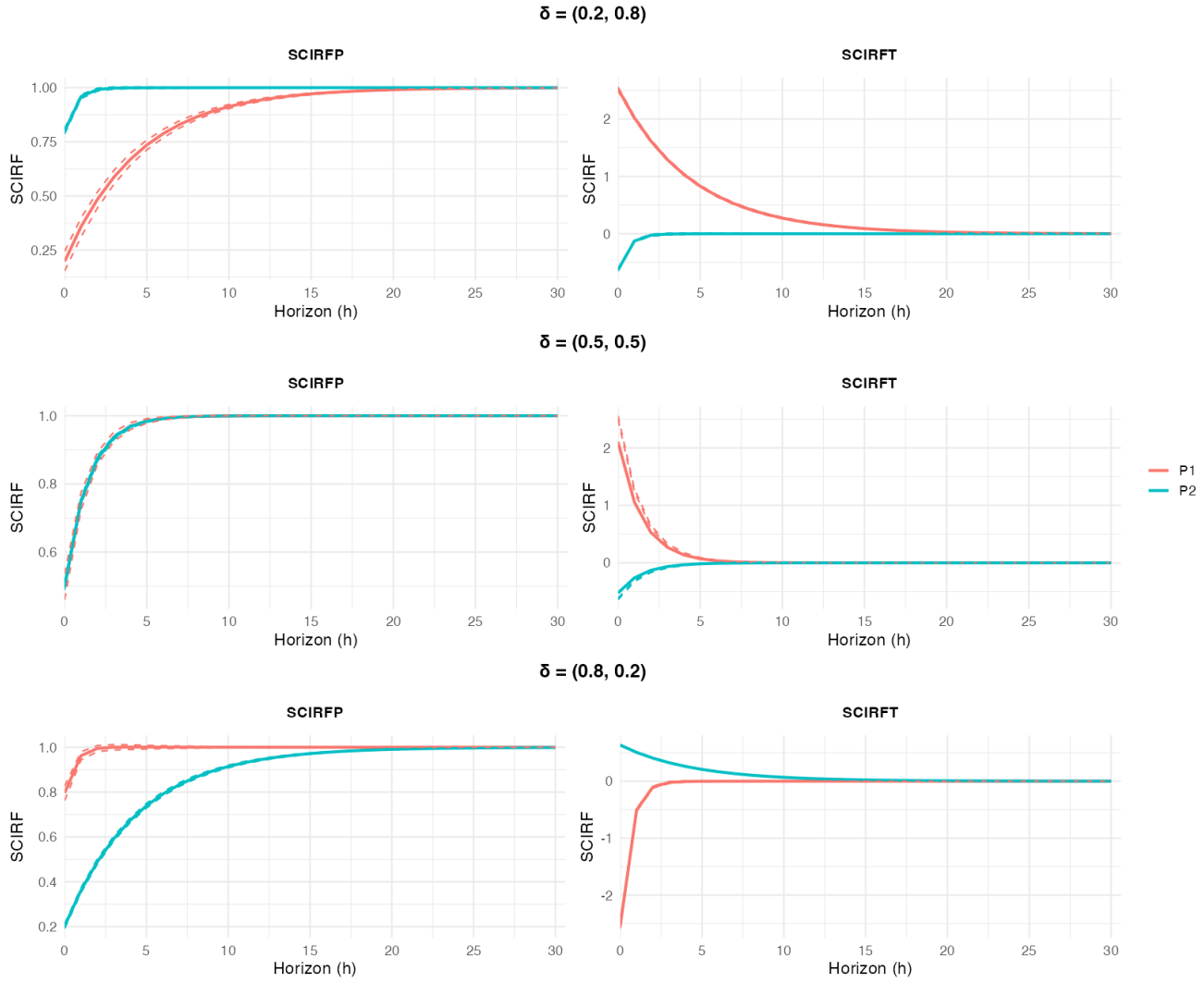

Figure 2.5: Structural cumulative impulse response functions (SCIRFs) of simulated prices in panel C. Each rows presents different pair of δ.

PIES provide correct results when reduced-form residuals are correlated. A market's response to transitory shocks (or just noise) may misguide price leadership identification using a single metric like IS, MIS or PDS and hence should be taken into account.

2.6 Empirical Application: Price Discovery in ETF Market

In this section, we apply our Bayesian price discovery methodology to analyze price discovery in the market for ETFs that track the S&P 500 index. In particular, we examine two competing S&P 500 ETFs (SPY and IVV) and discuss the effect of duplication of ETFs on price discovery. We calculate the price discovery measures across these two ETFs in three different stock exchanges– BATS, NASDAQ and ARCA. We first give an overview of the market for S&P 500 ETFs, then describe the data and present our results and analysis[9].

2.6.1 Market for S&P 500 ETFs

We focus on the spot market for exchange-traded funds (ETFs) tracking the S&P 500 index. Our setting is closely related to Hasbrouck (2003), who studied price discovery between the E-mini S&P 500 futures and spot instruments (including SPY mid-quotes and trades). Since that study, trading in S&P 500 ETFs has expanded markedly. In addition to SPDR S&P 500 (SPY), iShares Core S&P 500 (IVV) has long been available, and other vehicles (e.g., Vanguard's VOO) now attract substantial flows.

In what follows we restrict attention to SPY and IVV for three reasons. First, SPY and IVV account for the bulk of capitalization and turnover among S&P 500 ETFs, so adding further tickers contributes little incremental information for our questions about spot-market leadership. Second, SPY and IVV are constructed to track the same index with highly similar portfolios and operational designs, yielding very tight co-movement at intraday and daily horizons; differences in share price levels across ETFs (due to share scaling, fees, and operational conventions) are immaterial for cointegration. Third, investors treat SPY and IVV as close substitutes: Marshall

[9]For this analysis, we use the ETF data from Sultan and Zivot (2015).

Table 2.4: Comparison of SPY and IVV

	SPY	**IVV**
Overview		
Issuer	State Street	BlackRock, Inc.
Inception	1993-01-22	2000-05-15
Assets Under Management	$680B	$ 705B
Shares Outstanding	1.01B	1.04B
Expense Ratio	0.09%	0.03%
Performance Comparison		
1 Week Return	-1.57%	-1.56%
1 Month Return	0.24%	0.27%
26 Week Return	26.22%	26.37%
1 Year Return	14.83%	14.94%
3 Year Return	24.39%	24.51%
Beta	1.00	1.00
P/E ratio	17.86	19.66
Annual Dividend Rate	$7.25	$7.76
Annual Dividend Yield	1.10%	1.17%
5 day Volatility	101.32%	93.96%
200 day Volatility	10.39%	10.36%
Standard Devuation	36.72%	37.02%

Source: ETF Database. Web link: `http://etfdb.com/tool/etf-comparison/IVV-SPY/`. All results are reported on October 17st, 2025.

et al. (2013) show that when relative mispricing arises, arbitrage activity links the two funds, implying that price deviations are corrected over time.

These institutional features imply a cointegration relation between SPY and IVV. Let $\mathrm{P}_{SPY,t}$ and $\mathrm{P}_{IVV,t}$ denote log prices. Creation–redemption arbitrage and near-identical index exposure ensure that the spread $\mathrm{P}_{SPY,t} - \mathrm{P}_{IVV,t}$ is stationary: transitory deviations reflect microstructure noise, small fee/expense differences, and tracking-error realizations, but they mean-revert as authorized participants and liquidity providers exploit violations of the law of one price. We therefore model $(\mathrm{P}_{SPY,t}, \mathrm{P}_{IVV,t})'$ as an I(1) vector with cointegration rank one, and estimate a reduced-form VECM in the spirit of Hasbrouck (1995). This specification accommodates short-run lead–lag adjustments while imposing the long-run parity implied by the ETF arbitrage mechanism, providing a natural platform for our price-discovery analysis.

Interpreting price-discovery shares for two near-identical S&P 500 ETFs requires clarity about

the information that drives trading in these instruments. For a single stock, an "information shock" is naturally tied to firm-specific fundamentals. By contrast, for broad-market ETFs such as SPY and IVV, the relevant innovation is to investors' expectations about the aggregate market—i.e., news about the S&P 500's fundamental value. In a cointegrated two-ETF setting, a permanent shock therefore represents common information about the index, whereas transitory movements primarily reflect microstructure frictions and venue-specific noise.

Following the microstructure perspective presented in Hasbrouck (2003), we model the price vector P_t as an information set summarizing the prices readily observed and used by intraday traders. In Hasbrouck (2003), P_t included ETF quotes and trades alongside futures (pit and E-mini) prices, a deliberately broad proxy for public information available to market participants. Fang and Sanger (2011) similarly construct a rich intraday price set—mid-quotes for SPY and IVV and a second-by-second proxy for the S&P 500 index—to study spot-market discovery.

Our application is deliberately narrower for two reasons. First, we restrict P_t at intraday frequency to SPY and IVV bid–ask mid-quotes. Creation/redemption links between ETF prices and the underlying index operate at daily frequency via authorized participants; the index level (or end-of-day NAV) disciplines ETF pricing primarily at the close (Ben-David et al., 2014). Including a high-frequency index proxy in intraday discovery may therefore overweight a signal that is not directly arbitrageable within the day. In contrast, cross-ETF arbitrage (SPY vs. IVV) is actionable throughout the trading day when relative mispricing arises (Marshall et al., 2013), making contemporaneous SPY/IVV quotes the most relevant inputs for intraday decisions. Second, mid-quotes attenuate classic microstructure distortions (e.g., bid–ask bounce), improving the signal-to-noise ratio for estimating short-run leadership while preserving the long-run cointegration implied by the ETF arbitrage mechanism.

2.6.2 Data Description and Descriptive Statistics

In the empirical analysis high-frequency trade and quote data from TAQ database have been used, the access to which was possible via Wharton Research Data Service (WRDS). Tick-by-tick raw trade and quote data contains numerous data errors and requires thorough data-cleaning before analysis. For this study, the step-by-step data cleaning procedure proposed by Barndorff-Nielsen

et al. (2008) is followed[10]:

1. Restrict data to exchange hours (9:30 am to 4:30 pm);
2. Delete entries with zero quotes;
3. Delete entries with negative spreads;
4. Delete entries if spread > maxi*median daily spread;
5. Delete entries for which the mid-quote is outlying with respect to surrounding entries;
6. Restrict Data to specific exchange (these step is done three times for the three stock exchanges that the paper analyzes-BATS, NASDAQ, ARCA);
7. Delete entries with same time stamp and use median quotes;
8. When there is a time-stamp with no trade or ask/bid price reported for it, use the last observed trade or ask/bid price to replace the missing values. Therefore, each ETF (SPY and IVV) contains 25201 observations for a given stock exchange and for a given day.

We use the NYSE TAQ database as our source of high-frequency quotes for SPY and IVV (1 second). We choose two snapshots of data in two distinct trading environment. For a normal trading period, we choose intra-day quotes data from Dec 3rd to Dec 7th in 2012. For an extremely volatile trading period, we choose the quotes data of May 6th in 2010 (the day of Flash Crash) and Aug 8th, 2011 (the day of the worst stock market fall in US since 2008 till 2015).

On May 6th, 2010 the US stock market experienced an abrupt crash. The abnormal plunge in the market index was first seen at 2:42 pm and the fall in prices continued for next 20 minutes. The Dow Jones Industrial Average experienced the biggest one-day point decline during that day. August 11th, 2011 is considered to be the worst day in Wall Street since the crisis of 2008. All three major stock market indexes (S&P 500, Dow Jones Industrial Average, NASDAQ composite) fell sharply (between 5% to 7%) during that day. The day was also known for the wide-spread panic among the investors regarding the US losing its AAA credit rating on Aug 6th, 2011.

Table 2.5 shows descriptive statistics of bid-ask mid-quotes of SPY and IVV. The mean values of SPY and IVV are very close in all three stock exchanges at all dates, while the standard deviation is much higher for abnormal trading days.

Figure 2.6 provides evidence that both ETF prices move in tandem throughout the day (i.e., highly cointegrated). The only exception was the short period in the afternoon during the Flash

[10]The data cleaning procedure was carried out using the R package highfrequency.

Table 2.5: Descriptive Statistics of bid-ask mid-quotes of SPY and IVV in different Stock Exchanges

Stock Exchange	Price	Mean			Standard Deviation		
		Average of Dec 3rd-Dec 7th, 2012	Aug 8th, 2011	May 6th, 2010 (flashcrash)	Average of Dec 3rd-Dec 7th, 2012	Aug 8th, 2011	May 6th, 2010 (flash-crash)
NASDAQ	SPY	141.740	115.104	114.751	0.255	1.627	1.652
	IVV	142.334	115.500	115.026	0.256	1.631	2.010
BATS	SPY	141.740	115.106	114.751	0.255	1.624	1.650
	IVV	142.334	115.503	115.024	0.256	1.629	1.955
ARCA	SPY	141.740	115.106	114.749	0.255	1.624	1.652
	IVV	142.334	115.503	114.993	0.257	1.629	2.295

Crash when the IVV price plummeted by a significant amount compared to SPY.

2.6.3 Estimation and Results

Our assumption that SPY and IVV are cointegrated with cointegrating vector $\boldsymbol{\beta} = (1, -1)'$ in the three stock exchanges we consider is confirmed from the Johansen test of co-integration. To compute the price discovery measures, we put the price of SPY as the first element of the price vector and IVV as the second element. We also calculate the measures reversing the order, that is, putting IVV quotes as the first and SPY quotes as the second element of the vector. The results provide the same interpretation of the price discovery process[11]. For each trading day and each exchange we simulate posteriors from a bivariate Bayesian VECM for the log bid-ask mid-quotes of SPY and IVV at the one–second frequency and compute the full set of reduced–form and structural price discovery measures introduced in Section 3.3.

In table 2.6, we look into log bid-ask mid-quotes of SPY and IVV of one volatile trading day (flash-crash). The table reports the price discovery measures across SPY and IVV in three different stock exchanges. All variance–based reduced–form measures (IS, MIS and PDS), as well as PIES, assign a dominant role to SPY in price discovery on each of NASDAQ, BATS and ARCA. The posterior means of these measures for SPY lie between 0.78 and 0.83, whereas the corresponding values for IVV range between 0.17 and 0.22. CS also allocate the majority of the common efficient–price component to SPY. The posterior densities in Figure 2.7 are tightly concentrated around these values, implying that posterior uncertainty about the identity of the leading market is negligible on this day. By contrast, the information–leadership–type measures ILS, MILS and

[11] We focus on NASDAQ in this section, please see more figures in Appendix B.2

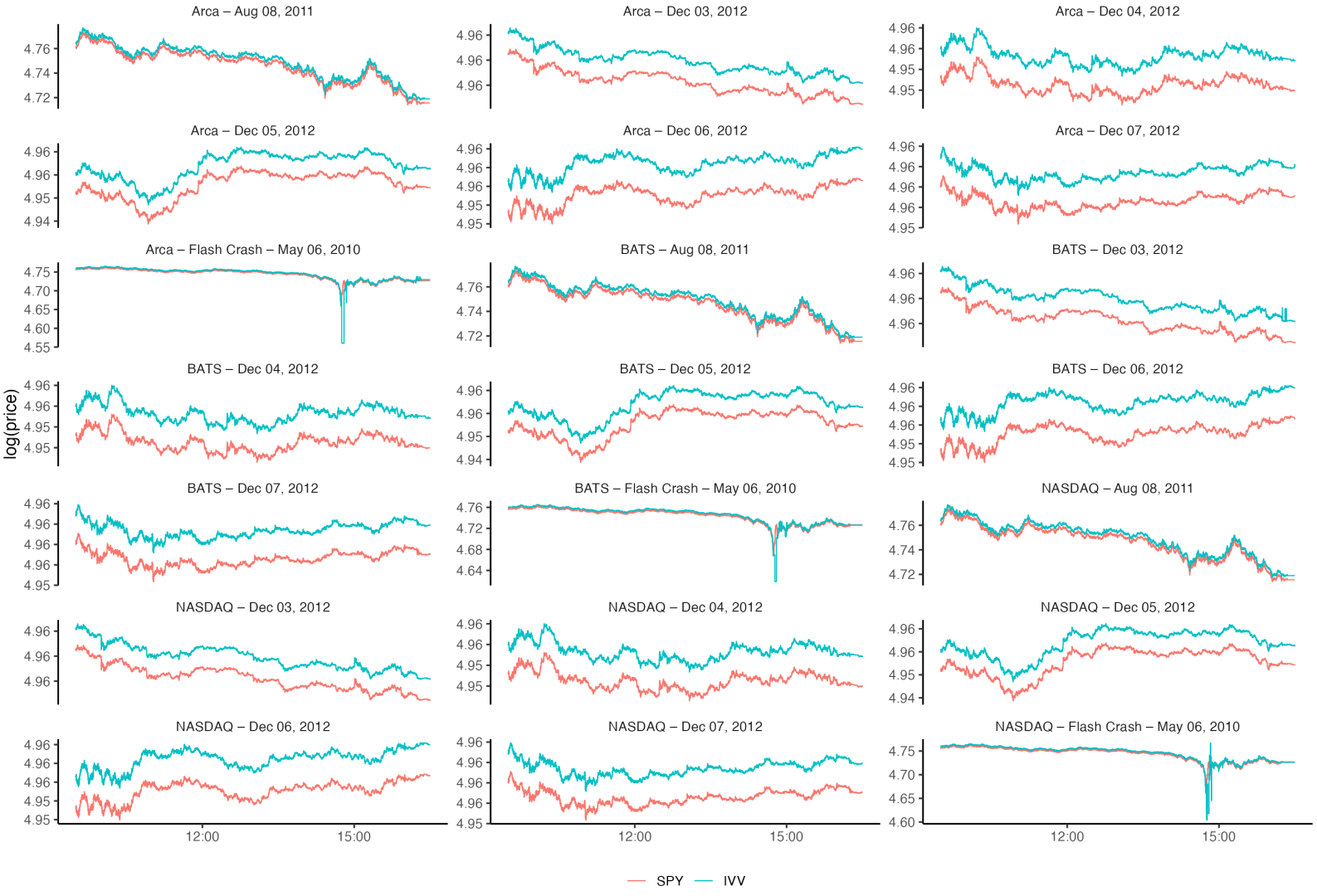

Figure 2.6: Time series of SPY & IVV

PILS attribute most of the leadership to IVV, with posterior means around 0.8–0.9 for IVV and 0.1–0.2 for SPY. Moreover, CovIS delivers negative shares for SPY and values greater than one for IVV. As stressed by Lautier et al. (2024), "a negative share corresponds to a contrarian move, corrected later by convergence. Interpretations depend on the specific markets and structures". Such outcomes complicate the interpretation of CovIS as a straightforward leadership measure.

The structural estimates reported in Table 2.9 help reconcile these apparently conflicting findings. On 6 May 2010, 2010 the estimate of d_0^P is positive for SPY across all exchanges, while for IVV it is large in magnitude and negative. This configuration implies that IVV tends to move initially in the opposite direction to the permanent news shock and overshoots the new fundamental value. At the same time, the estimates of d_0^T are strongly negative for IVV, and substantially less negative (or even positive on BATS) for SPY, indicating that IVV is more severely affected by transitory liquidity–related disturbances. The SCIRFs in Figure 2.10 confirm these patterns: the permanent SCIRFs for both ETFs converge to one, whereas the transitory SCIRFs mean–revert to zero; however, the initial permanent and transitory response of IVV is much larger in absolute value than that of SPY. In such an environment, measures that are sensitive to the transitory column D_0^T (e.g., ILS and MILS) may misrepresent leadership, while measures primarily driven by D_0^T or D_0^P (e.g., CS and PIES) correctly identify SPY as the informational leader.

Table 2.7 and Figure 2.8 report results for 8 August 2011, a highly volatile day associated with the U.S. sovereign downgrade episode. The posterior distributions of all price discovery measures place most of their mass above 0.5 for SPY and below 0.5 for IVV, so the posterior odds strongly favor SPY as the price leader on this day.

Across measures, the posterior means differ in how strongly they quantify SPY's lead. IS, CS and PIES yield posterior means only moderately above 0.5 (around 0.55–0.60), with credible intervals that still leave some posterior mass near equal shares. In contrast, PDS, ILS, MILS, PILS and CovIS have posterior means in the 0.6–0.9 range and credible intervals that lie much further away from 0.5. For these measures, the posterior probability that SPY's share exceeds one half is effectively one, and the implied posterior odds against "equal market share" are very large.

The corresponding structural estimates in Table 2.9 show that both ETFs load positively on the permanent shock, but SPY has a slightly larger D_0^P than IVV on all exchanges. The transitory impacts D_0^T are small in absolute value for both assets and display less asymmetry across SPY and

IVV than on the flash–crash day. Consequently, all measures convey a broadly consistent message of SPY leadership, with those relying more heavily on D_0^P (PILS, PIES, CovIS) emphasizing the leadership more sharply than IS or MIS. The posterior densities in Figure 2.8 are noticeably more dispersed than on 6 May 2010, reflecting that the leadership differential is economically meaningful but less extreme in this "normal high–volatility" environment.

For the relatively tranquil trading day 4 December 2012, Table 2.8 and Figure 2.9 show a very different Bayesian picture. The posterior means of almost all measures (IS, CS, MIS, PDS, ILS, MILS, PILS and PIES) are close to 0.5 for both SPY and IVV, and the corresponding posterior densities are wide and centered around 0.5. As a result, the posterior probability that each measure lies in a neighborhood of equal market share is large, and the posterior odds in favor of one market having a strictly higher share than the other are close to one (i.e. essentially indecisive). Inference on this date is therefore consistent with no clearly dominant leader in terms of price discovery: SPY and IVV appear to contribute roughly symmetrically to the formation of the common S&P 500 efficient price.

CovIS again exhibits different behavior, assigning a larger share to IVV on NASDAQ and ARCA and a negative share to SPY on BATS. According to Table 2.9, the estimated D_0^P coefficients for SPY and IVV are both small and of similar magnitude on 4 December 2012, while the estimates of D_0^T are moderate. In such situations, small sampling variation in D_0^P can change the sign of CovIS, rendering this measure difficult to interpret as a ranking device when leadership differences are minor. The remaining measures, which do not rely on the sign of D_0^P, provide a more robust assessment by indicating that there is no distinct leader between SPY and IVV on this calm trading day.

Overall, the empirical evidence shows that the Bayesian framework provides full posterior distributions for all price discovery measures, so uncertainty about market leadership can be assessed explicitly. When one asset clearly dominates—as on the flash–crash day and on 8 August 2011—the posterior distributions are highly concentrated (especially during the flash–crash), whereas on the calm day of 4 December 2012 they are diffuse and centered around 0.5, indicating no clear leader. Structural measures that depend only on the permanent–impact matrix D_0^P (PILS/CovISQ and PIES) emerge as relatively robust indicators of informational leadership, particularly in stressed market conditions where temporary overshooting and contrarian moves are common. Finally, ex-

amining the estimated (D_0^P, D_0^T) jointly with the associated SCIRFs is crucial both for diagnosing why different measures may deliver conflicting signals and for linking reduced–form price discovery statistics to the underlying microstructure dynamics of the SPY–IVV ETF market.

Table 2.6: Price discovery measurements of SPY and IVV on May 6th, 2010 (flash-crash)

Stock Exchange	Price	IS	CS	MIS	PDS	ILS	MILS	PILS/CovISQ	PIES	CovIS
NASDAQ	SPY	0.83	0.92	0.83	0.83	0.04	0.15	0.16	0.97	-0.86
	IVV	0.17	0.08	0.17	0.17	0.96	0.85	0.84	0.03	1.86
BATS	SPY	0.78	0.83	0.79	0.79	0.31	0.37	0.40	0.92	-0.49
	IVV	0.22	0.17	0.21	0.21	0.69	0.63	0.60	0.08	1.49
ARCA	SPY	0.78	0.90	0.78	0.78	0.04	0.14	0.16	0.93	-0.82
	IVV	0.22	0.10	0.22	0.22	0.96	0.86	0.84	0.07	1.82

Table 2.7: Price discovery measurements of SPY and IVV on August 8th, 2011

Stock Exchange	Price	IS	CS	MIS	PDS	ILS	MILS	PILS/CovISQ	PIES	CovIS
NASDAQ	SPY	0.55	0.52	0.64	0.86	0.60	0.81	0.61	0.51	0.82
	IVV	0.45	0.48	0.36	0.14	0.40	0.19	0.39	0.49	0.18
BATS	SPY	0.59	0.55	0.67	0.84	0.63	0.83	0.65	0.53	0.62
	IVV	0.41	0.45	0.33	0.16	0.37	0.17	0.35	0.47	0.38
ARCA	SPY	0.57	0.54	0.67	0.91	0.61	0.85	0.63	0.52	0.70
	IVV	0.43	0.46	0.33	0.09	0.39	0.15	0.37	0.48	0.30

Table 2.8: Price discovery measurements of SPY and IVV on Dec 4th, 2012

Stock Exchange	Price	IS	CS	MIS	PDS	ILS	MILS	PILS/CovISQ	PIES	CovIS
NASDAQ	SPY	0.49	0.50	0.48	0.47	0.48	0.52	0.47	0.46	0.15
	IVV	0.51	0.50	0.52	0.53	0.52	0.48	0.53	0.54	0.85
BATS	SPY	0.49	0.49	0.49	0.48	0.51	0.54	0.48	0.48	-0.54
	IVV	0.51	0.51	0.51	0.52	0.49	0.46	0.52	0.52	1.54
ARCA	SPY	0.49	0.48	0.48	0.47	0.53	0.57	0.50	0.51	0.46
	IVV	0.51	0.52	0.52	0.53	0.47	0.43	0.50	0.50	0.54

2.6.4 Sensitivity Analysis: Data Frequency

The choice of sampling frequency is pivotal in price discovery measurement because it governs the trade-off between informativeness and market microstructure noise. At very high frequencies, quotes and trades arrive asynchronously across venues and are contaminated by bid–ask

Table 2.9: Estimation of D_0^P and D_0^T

		Dec 04, 2012		May 06, 2010 (flash-crash)		Aug 08, 2011	
Stock Exchange	Price	D_0^P	D_0^T	D_0^P	D_0^T	D_0^P	D_0^T
NASDAQ	SPY	0.39	0.14	0.58	-0.25	0.22	0.06
	IVV	0.43	-0.13	-1.34	-3.14	0.21	-0.03
BATS	SPY	-0.29	0.13	0.52	0.79	0.25	0.04
	IVV	-1.46	-0.16	-0.63	0.21	0.21	-0.09
ARCA	SPY	0.42	0.13	0.45	-0.25	0.23	0.04
	IVV	0.41	-0.14	-1.07	-2.27	0.21	-0.05

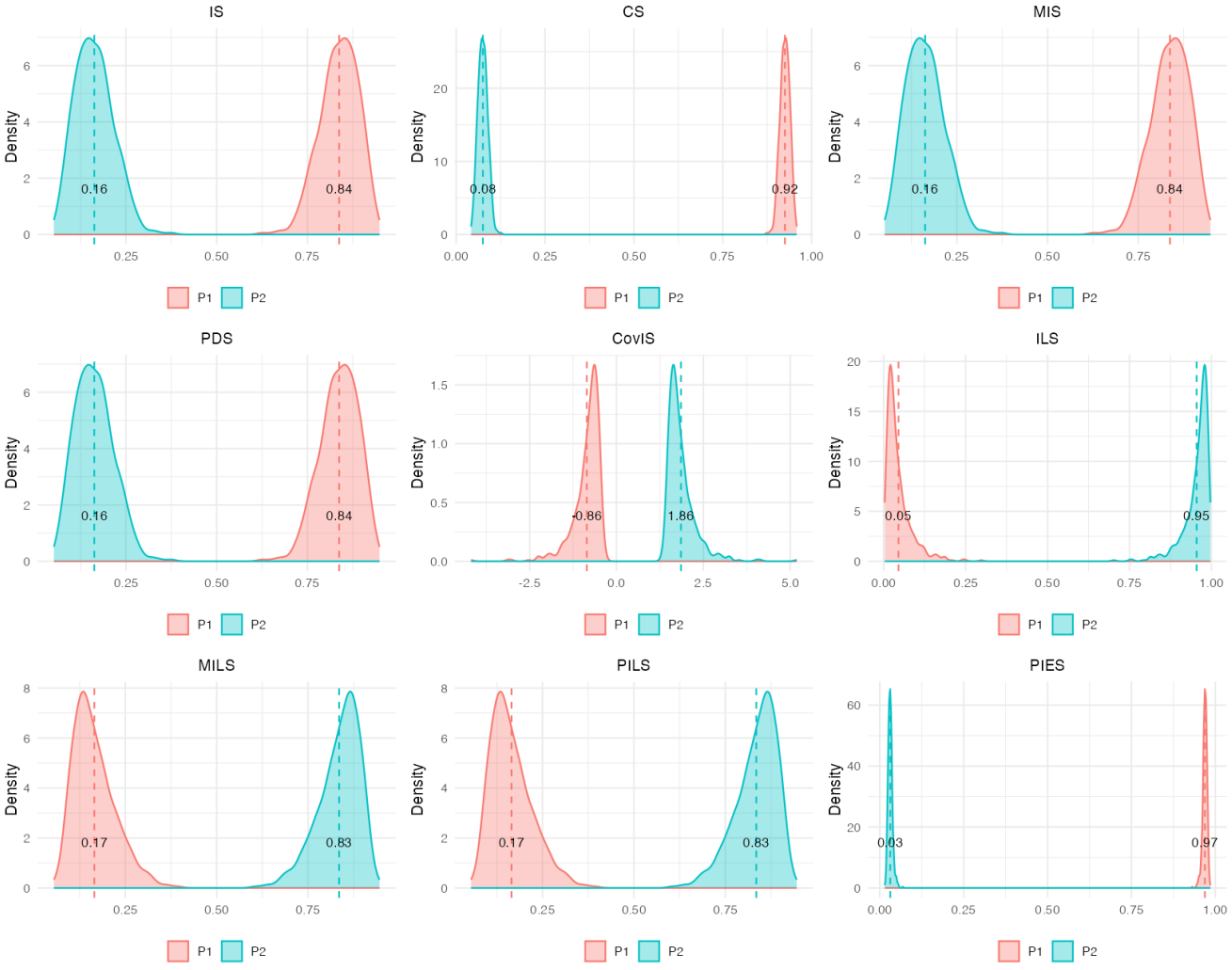

Figure 2.7: Distribution of price discovery measures at NASDAQ on May 6th, 2010 (flash-crash)

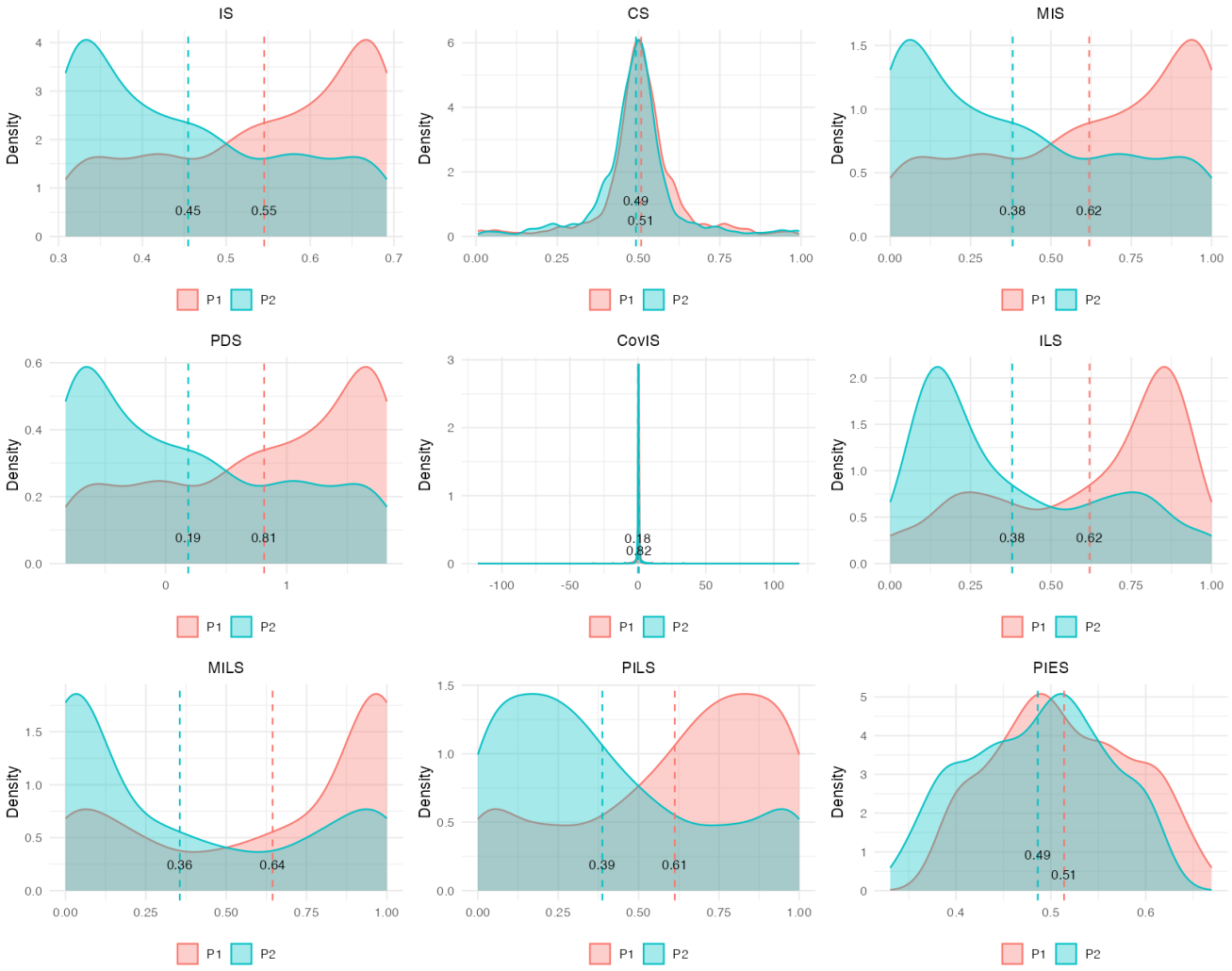

Figure 2.8: Distribution of price discovery measures at NASDAQ on Aug 8th, 2011

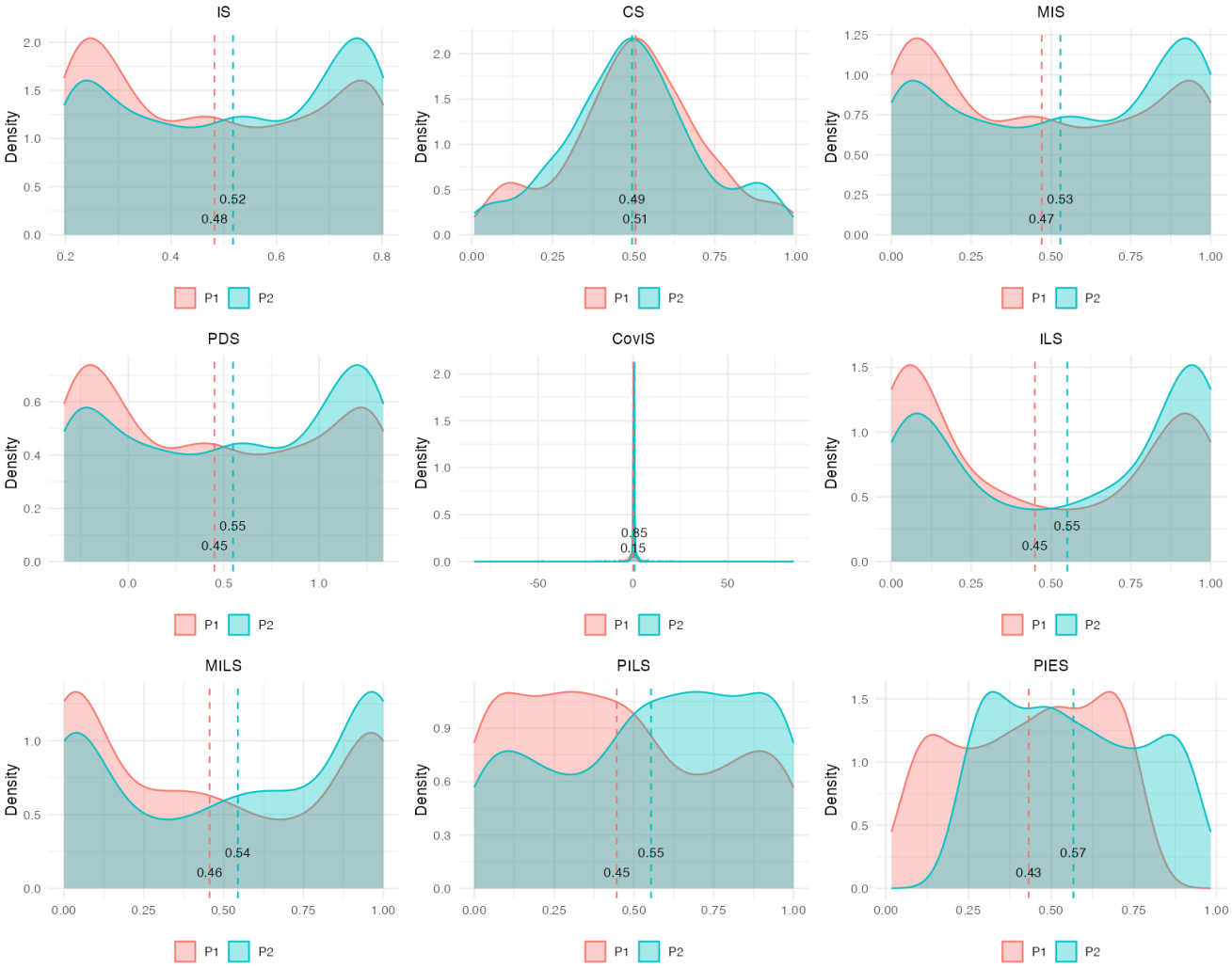

Figure 2.9: Distribution of price discovery measures at NASDAQ on Dec 4th, 2012

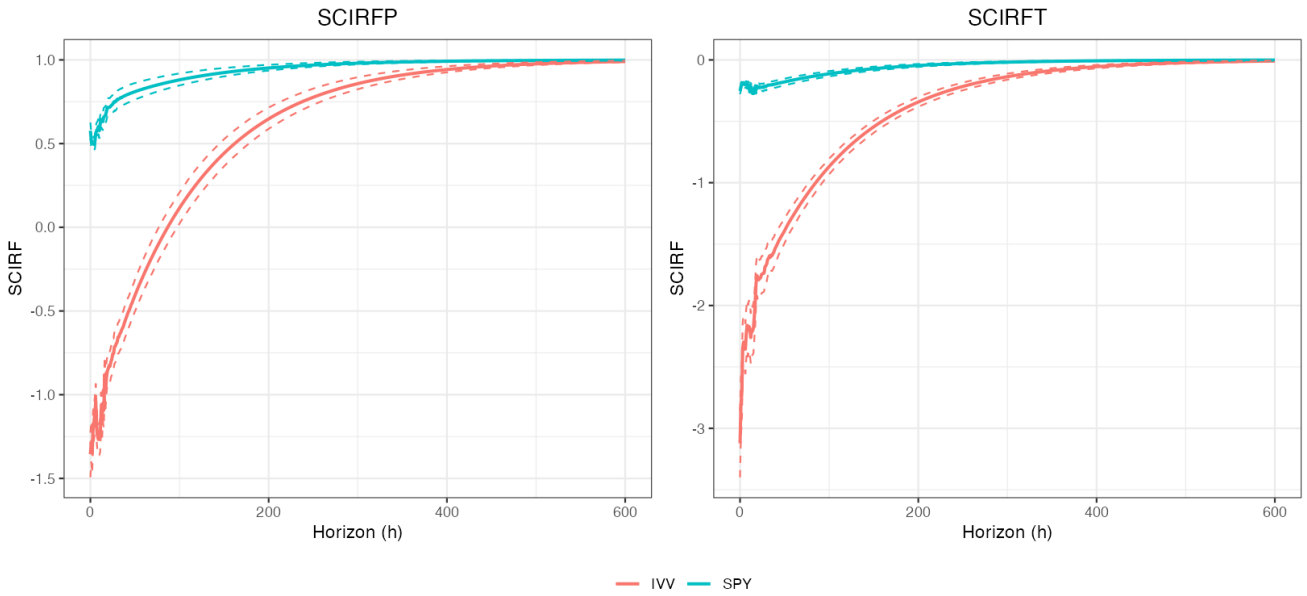

Figure 2.10: SCIRF of SPY & IVV at NASDAQ on May 6th, 2010 (flash crash)

bounce, price discreteness and latency; as a result, information–share (IS) bounds widen and component–share (CS) estimates become unstable (Hasbrouck, 1995; Putniņš, 2013). Lead–lag in message arrival also induces the Epps effect—dampened short-horizon comovement due to non-synchronous trading—which can bias discovery toward the venue sampled more densely unless timestamps are carefully synchronized (Epps, 1979; Hayashi and Yoshida, 2005). The microstructure literature therefore advocates the use of mid–quotes rather than transaction prices, refresh–time or event–time alignment across venues, and various noise-robust filters (e.g. subsampling, pre-averaging) prior to VAR/VECM estimation (Hasbrouck, 2003; Theissen, 2016). Moving to coarser calendar-time intervals reduces noise and mitigates asynchrony, but excessive aggregation may wash out genuine intraday leadership (e.g. futures-to-spot or venue-to-venue precedence measured in seconds) and attenuate the separation between permanent and transitory shocks (Hasbrouck, 2003; Ozturk et al., 2017).

Against this background, Table 2.10 and Figures 2.11-2.12 examine how the estimated price discovery shares for SPY vary with the sampling interval when comparing NASDAQ and BATS on August 8th, 2011. We focus on log mid–quotes of SPY, the most actively traded S&P 500 ETF in these venues. The VECM is estimated at six calendar-time frequencies—1s, 5s, 10s, 20s, 30s, 1 minute and 2 minutes—and the full suite of reduced-form and structural discovery measures is

computed for each frequency.

Table 2.10 shows that, across all frequencies, NASDAQ is identified as the informationally dominant venue. For the highest frequencies (1–10 seconds), IS, MIS and PDS attribute roughly 90% or more of the discovery to NASDAQ, with PILS and PIES delivering similarly extreme shares (around 0.9) in favor of NASDAQ. CS and CovIS are slightly less extreme but still allocate substantially larger weights to NASDAQ than to BATS. As the sampling interval is lengthened from 20s to 1 minute, the NASDAQ shares decline gradually: IS falls from 0.75 to 0.59, PDS from roughly unity to around 1.1 (reflecting its unconstrained nature), and CovIS decreases from 0.66 to 0.55. At the coarsest 2-minute interval, the evidence remains tilted toward NASDAQ (e.g. IS = 0.62, CovIS = 0.59), but the leadership gap narrows considerably and some measures (notably ILS and MILS) become more erratic due to the reduced effective sample size.

Taken together, the table indicates that the direction of leadership (NASDAQ vs. BATS) is remarkably robust to the choice of frequency, whereas the magnitude of leadership is frequency–dependent. Very fine sampling yields near-degenerate shares that suggest a single dominant venue; moderate aggregation (30–60s) still favors NASDAQ strongly but produces less extreme values; and heavy aggregation (2 minutes) pushes all measures toward more balanced contributions.

Figure 2.11 and Figure 2.12 report the full posterior distributions of PILS and PIES for NASDAQ and BATS at each sampling frequency. At the highest frequencies (1s, 5s, and 10s), the posterior for NASDAQ is sharply concentrated near one (posterior means around 0.86–0.99), while that for BATS is concentrated near zero, implying an overwhelming posterior probability that NASDAQ is the price leader at these horizons. At 30s, the PIES measure in particular tilts even more toward the extremes, with NASDAQ's distribution massed close to one and BATS' close to zero, reinforcing the notion of very strong leadership by NASDAQ. At 60s, the two posteriors remain clearly separated but become noticeably more dispersed, indicating that some of NASDAQ's advantage is attenuated once the most transient microstructure noise has been averaged out. By 120s, the distributions overlap substantially: NASDAQ's PILS and PIES center around 0.67, BATS around 0.33, and posterior mass is much more spread out, signaling weaker evidence in favor of a unique leading venue.

Overall, the qualitative ranking of venues in terms of informational leadership is robust: NAS-

DAQ leads BATS at essentially all sampling intervals, and this conclusion is particularly strong at the highest frequencies where the permanent shocks are identified most sharply. As the sampling interval lengthens, episodes in which BATS briefly adjusts first are averaged in, gradually eroding the apparent information advantage of NASDAQ.

Table 2.10: Price discovery in log of bid-ask mid-quotes of SPY in NASDAQ and BATS (from high frequency to low frequency data) on Aug 8th, 2011

Stock Exchange	Stock Exchange	IS	CS	MIS	PDS	ILS	MILS	PILS	PIES	CovIS
1 sec	NASDAQ	0.93	0.91	0.95	0.96	0.70	0.74	0.89	0.86	0.74
	BATS	0.07	0.09	0.05	0.04	0.30	0.26	0.11	0.14	0.26
5 sec	NASDAQ	0.93	0.96	0.97	1.01	0.21	0.49	0.90	0.98	0.75
	BATS	0.07	0.04	0.03	-0.01	0.19	0.51	0.10	0.02	0.25
10 sec	NASDAQ	0.88	0.94	0.95	1.02	0.15	0.48	0.88	0.99	0.73
	BATS	0.12	0.06	0.05	-0.02	0.85	0.52	0.12	0.01	0.27
20 sec	NASDAQ	0.75	0.93	0.87	1.03	0.03	0.26	0.79	0.99	0.66
	BATS	0.25	0.07	0.13	-0.03	0.97	0.74	0.21	0.01	0.34
30 sec	NASDAQ	0.69	0.90	0.82	1.06	0.03	0.29	0.71	0.98	0.61
	BATS	0.31	0.10	0.18	-0.06	0.97	0.71	0.29	0.02	0.39
1 min	NASDAQ	0.59	0.85	0.73	1.09	0.04	0.27	0.60	0.89	0.55
	BATS	0.41	0.15	0.27	-0.09	0.96	0.73	0.40	0.11	0.45
2 min	NASDAQ	0.62	0.57	0.93	2.04	0.68	0.94	0.67	0.66	0.59
	BATS	0.38	0.43	0.07	-1.04	0.32	0.06	0.33	0.34	0.41

2.7 Conclusion

This paper develops a Bayesian framework for measuring price discovery in fragmented markets that unifies recent structural approaches to discovery with modern Bayesian VECM methods. We embed a wide range of reduced-form and structural price discovery measures in a bivariate Bayesian VECM with fixed cointegrating vector $\beta = (1, -1)'$, and show how posterior simulations of the VECM parameters can be mapped into posterior distributions for discovery shares, structural impulse responses, and formally testable leadership hypotheses. By doing so, we address two long-standing challenges in the literature: the lack of reliable finite-sample inference for highly nonlinear discovery functionals and the difficulty of separating true informational leadership from transitory microstructure frictions.

A central contribution of the paper is to emphasize structurally grounded price discovery mea-

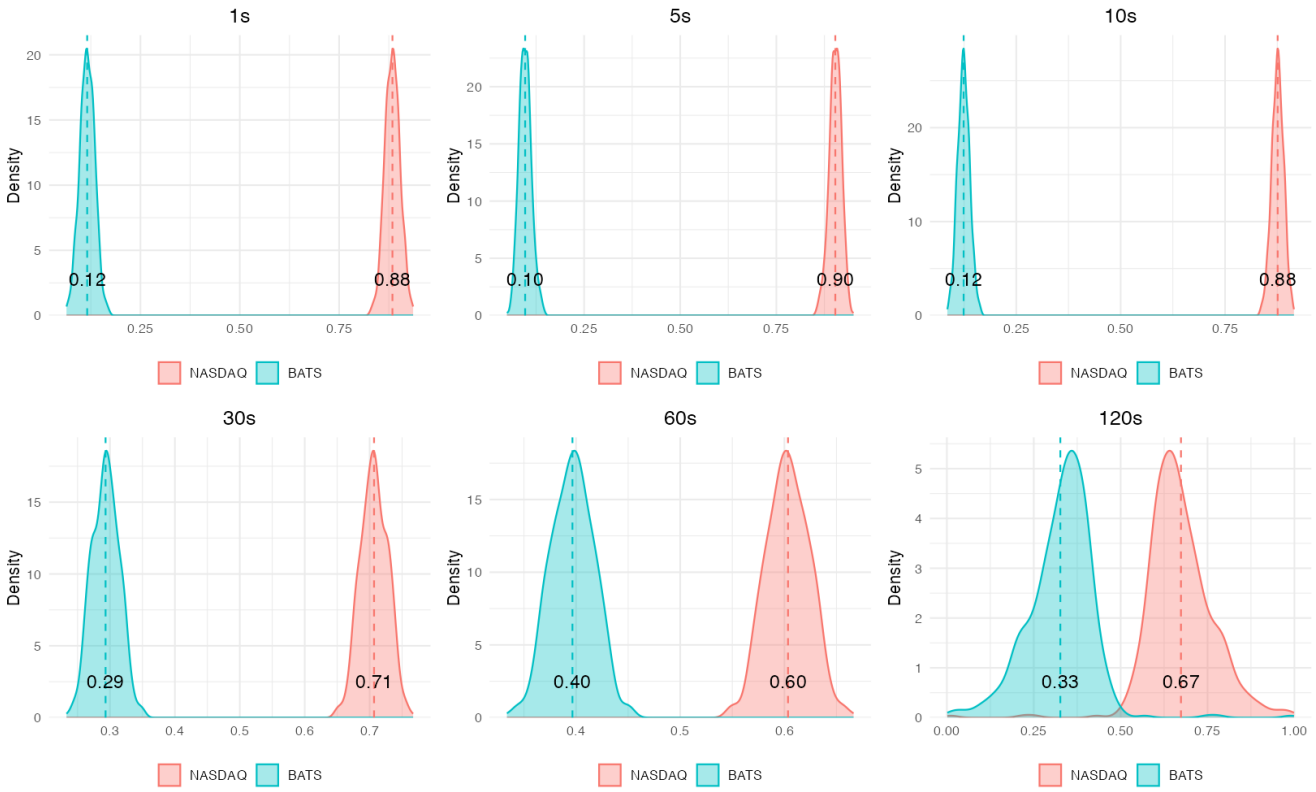

Figure 2.11: PILS comparison of data frequency for SPY on Aug 8th,2011

sures that depend only on the permanent component of prices. Following the structural cointegration model of Yan and Zivot (2010) and the identification results of Lautier, Ling and Villeneuve (2024), we recover the contemporaneous impact matrix D_0 and, in particular, its permanent column D_0^P, which captures the initial response of each market to the common efficient-price shock. Measures such as PILS, PIES, CovIS and CovISQ can all be expressed as simple functions of $(d_{0,1}^P, d_{0,2}^P)$ and the variance of the permanent shock. As a result, these measures are invariant to the covariance structure of reduced-form residuals and to the transitory impact parameters D_0^T. They rank markets purely on how strongly and how quickly they impound permanent news into prices, rather than on how effectively they avoid or absorb noise. This structural dependence on permanent shocks gives these measures a clear economic interpretation and makes them robust benchmarks for price discovery analysis.

The Bayesian VECM provides the natural environment for inference on these structural measures. Shrinkage priors regularize high-dimensional short-run dynamics, improving finite-sample performance when many lags are needed to capture high-frequency price adjustments. Given con-

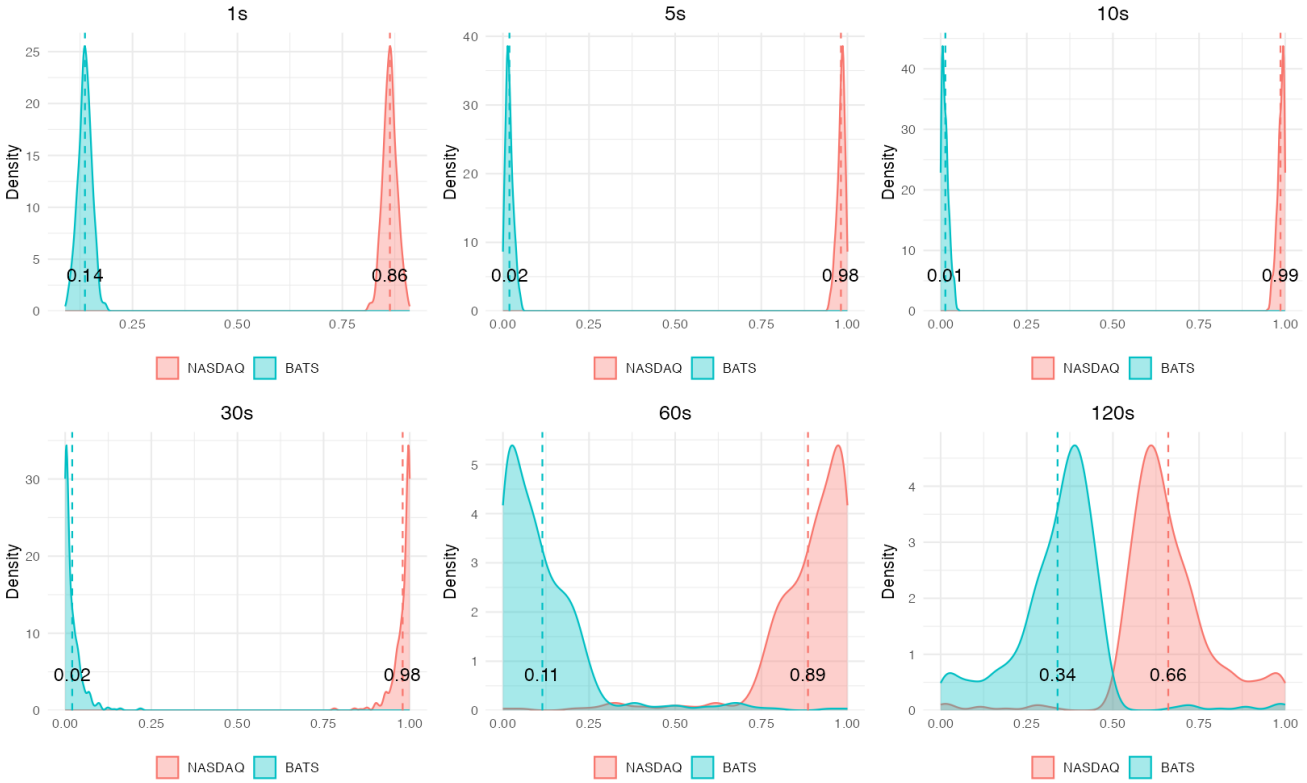

Figure 2.12: PIES comparison of data frequency for SPY on Aug 8th,2011

jugate or semi-conjugate priors for (α, Γ, Ω), posterior simulation by Gibbs sampling is straightforward. For each posterior draw we compute reduced-form objects, recover the structural impact matrix D_0, and then transform these into the full suite of price discovery measures and structural cumulative impulse responses (SCIRFs). The resulting posterior distributions yield credible intervals and probability statements for any functional of interest without relying on fragile delta-method approximations or bootstrap schemes that are known to perform poorly for ratio-type statistics.

Simulation evidence from a stylized partial price-adjustment microstructure model highlights the practical advantages of focusing on permanent-shock-based measures within a Bayesian framework. When both markets have similar noise and uncorrelated reduced-form residuals, all discovery measures—including classical IS, MIS and PDS—track the true leadership implied by the structural adjustment speeds. Once we introduce asymmetric noise or strong correlation in reduced-form innovations, however, the variance-decomposition measures often misclassify the leader, favoring the cleaner but slower market. In contrast, the structurally grounded measures

PILS/CovISQ, PIES and CovIS remain tightly aligned with the true speeds of adjustment across all parameterizations. Bayesian posterior distributions for these measures exhibit the expected shape—centered at 0.5 under symmetry and shifting decisively toward the true leader when asymmetries are present—providing a coherent probabilistic characterization of leadership that is robust to transitory frictions.

The empirical application to SPY and IVV ETFs, across exchanges and trading environments, further illustrates the value of combining structural price discovery measures with Bayesian inference. On stressed days such as the Flash Crash and the U.S. sovereign downgrade episode, the structural estimates of D_0^P and the permanent SCIRFs reveal that SPY more cleanly and more quickly absorbs permanent news about the S&P 500 index, while IVV exhibits overshooting and stronger exposure to transitory shocks. Permanent-shock-based measures (PILS/CovISQ, PIES) consistently identify SPY as the informational leader, with posterior mass concentrated near one for SPY and near zero for IVV. On normal days, posterior distributions for these measures become diffuse and center around 0.5, indicating no clear leader and correctly reflecting the economic reality that SPY and IVV share discovery in normal conditions. The analysis of sampling frequency shows that while the qualitative ranking of venues (e.g., NASDAQ vs. BATS for SPY) is robust, the magnitude of leadership is horizon-dependent—something our Bayesian approach makes explicit through posterior distributions of discovery shares at each frequency.

Overall, our results suggest that future empirical work on price discovery should (i) explicitly distinguish permanent information shocks from transitory noise; (ii) prioritize structural measures that depend only on permanent shocks when ranking markets or assets; and (iii) adopt Bayesian VECM methods to quantify uncertainty and conduct formal hypothesis tests on discovery shares. Extensions of the framework to higher-dimensional systems, time-varying parameters and stochastic volatility are natural next steps. Incorporating additional assets (e.g., futures, options, or index proxies) and exploring cross-asset discovery while retaining a structural, permanent-shock perspective within a Bayesian setting promises a more nuanced and reliable understanding of how information flows through today's fragmented financial markets.

Chapter 3

Dynamics of Time-varying Price Discovery

Wenqiu Ma, Eric Zivot

Abstract

We study the dynamics of price discovery in fragmented markets when the informational roles of trading venues vary over time. Building on the structural cointegration literature, we define price discovery in terms of each market's contribution to the permanent component of a common efficient price and explicitly separate permanent shocks from transitory shocks. To model time variation in a coherent and order-invariant way, we adopt the Bayesian order-invariant VAR with stochastic volatility (OI-VAR-SV) of Chan et al. (2024) and map its reduced-form representation into structural price discovery measures, including Hasbrouck's information shares (IS), component shares (CS), modified information shares (MIS), price discovery shares (PDS), information leadership measures (ILS/MILS), the price information leadership share (PILS), covariance information shares (CovIS/CovISQ), and the price instantaneous efficiency share (PIES).

Using a stylized partial-adjustment model with time-varying volatility and adjustment speeds, we show that measures grounded purely in the permanent component—PILS and CovISQ—track the true informational leadership accurately and are robust to both permanent and transitory heteroskedasticity. In contrast, the other measures can be severely distorted by time-varying noise and reduced-form correlation, and may generate spurious leadership even in symmetric environments. The proposed Bayesian framework delivers full posterior distributions for time-varying price discovery shares and structural impulse responses, providing a flexible and disciplined toolbox for empirical work on dynamic information flows in modern financial markets.

3.1 Introduction

Market fragmentation is pervasive in modern financial markets. The same underlying risk is often traded simultaneously in multiple venues and through closely related instruments such as spot, futures, options, and exchange-traded funds (ETFs) tracking the same index. In such an environment, a central question for traders, regulators, and researchers is which asset or market best incorporates new fundamental information into prices. Equivalently, we want to know which of the observed prices is most indicative of the asset's fundamental value and reflects relevant information first.

Price discovery, the process by which new information is impounded into prices through trading, plays a key role in the allocative efficiency of a free-market economy. In well-functioning and informationally efficient markets, news that affects the fundamental value of a security should be rapidly reflected in its price (Hasbrouck, 1995). In practice, however, imperfections such as transaction costs, information asymmetries, regulatory frictions, and market design differences prevent information from being instantaneously reflected in all related prices. When homogeneous or tightly linked securities are traded in multiple markets, it becomes especially important to understand which venues lead in processing information and how these informational roles evolve over time.

Over the past three decades, the empirical literature has developed a rich toolbox to measure price discovery in interdependent and cointegrated markets. The modern approach posits a common efficient price that follows a martingale and allows observed quotes and transaction prices to contain transitory components. Empirically, this is typically implemented in a bivariate vector error correction model (VECM) with a known cointegrating vector $\boldsymbol{\beta} = (1, -1)'$, within which price discovery is defined relative to the common stochastic trend. Two traditions anchor this literature. In the first, variance–decomposition methods apportion innovations in the efficient price to venue-specific shocks, yielding information shares (IS) in the sense of Hasbrouck (1995). In the second, common-factor representations identify the permanent component of the VECM and attribute component shares (CS) to each venue through the adjustment loadings, as in Gonzalo and Granger (1995) and subsequent applications (e.g., Baillie et al., 2002; Harris et al., 2002; De Jong, 2002). While IS and CS have become workhorses in applications to equities, futures,

options, commodities, credit, and foreign exchange markets, they suffer from several well-known drawbacks, including non-uniqueness due to order dependence and VECM residual error correlation, ambiguous interpretation when transient dynamics are strong, and sensitivity to short-run noise.

Building on these foundations, subsequent research has connected price discovery to bid–ask spreads, volatility, transaction costs, and liquidity provision, documenting how market microstructure characteristics shape the informational content of prices (e.g., Madhavan et al., 1997; Hasbrouck and Seppi, 2001). The availability of high-frequency data has enabled detailed studies of intraday price discovery in equity and ETF markets (Hasbrouck, 2003; Theissen, 2016; Ozturk et al., 2017; Hasbrouck, 2021), showing, for example, that high-frequency and algorithmic trading can improve liquidity and quote informativeness (Hendershott et al., 2011), that options markets contribute nontrivially to discovery of the underlying price (Chakravarty et al., 2004), and that certain high-frequency traders help predict price changes over very short horizons (Brogaard et al., 2014; Fang et al., 2022). More recent contributions such as Yan and Zivot (2007, 2010); Putniņš (2013); Shen et al. (2024b); Lautier et al. (2024) develop structural cointegration-based measures that explicitly separate permanent and transitory shocks and are invariant to the ordering of variables. Among these alternatives, the improved price information leadership share of Shen et al. (2024b) (PILS) and the covariance information shares of Lautier et al. (2024) (CovIS and CovISQ)[1] have emerged as preferred benchmarks, as they deliver unique, structurally interpretable rankings of venues that are robust to VECM residual error correlation and not contaminated by transient dynamics.

A central limitation of most conventional price discovery measures is that they are defined in static environments: VECM parameters are assumed constant over the sample, and informational roles are treated as time-invariant averages. This is at odds with what we know about the time-varying nature of information flows in financial markets (Andersen and Bollerslev, 1997). Markets are subject to recurrent shocks to fundamentals, liquidity, regulation, and trading technology; the moments of both fundamental values and noise components evolve; and the long-run relationships linking related assets may themselves shift. As emphasized by Hasbrouck (1995, 2003, 2007), static VECMs are not adequate for modeling price discovery over extended periods that span multiple

[1] Shen et al. (2024b) shows that CovISQ is equivalent to PILS.

regimes. A growing empirical literature documents that the relative contribution of different markets to price discovery is highly time-varying.

Existing approaches incorporate time variation in several ways, including rolling-window estimation, smoothly time-varying coefficients, and regime-switching models. Here we emphasize a common gap: many time-varying methods either (i) treat price discovery measures as reduced-form functions of changing VECM parameters without explicitly separating permanent from transitory shocks; (ii) rely on ordering-sensitive identification (e.g., Cholesky schemes) in settings where no natural ordering exists; or (iii) provide limited uncertainty quantification for dynamic leadership patterns.

A natural idea is to use Bayesian time-varying parameter VARs with stochastic volatility (TVP-VAR-SV), following Primiceri (2005) and Del Negro and Primiceri (2015), and then compute time-varying price discovery measures from the implied time-varying VMA representation. However, Primiceri's framework achieves identification through a triangular factorization of the time-varying covariance matrix Ω_t (equivalently, a lower-triangular contemporaneous impact matrix). In fragmented markets this creates two problems. First, the implied structural shocks are ordering dependent: the "first" shock is mechanically tied to the first equation's innovation, so leadership conclusions can change when the series are relabelled. Second, in a cointegrated price system where price discovery should be defined by innovations to the common permanent component, a triangular identification can impose strong and economically unintended restrictions on the permanent/transitory decomposition. In particular, interpreting the first triangular shock as the permanent shock implies that the long-run impact vector and contemporaneous permanent impacts become tightly linked to the triangular coefficient that governs reduced-form correlation, so measured "leadership" can largely reflect covariance normalization rather than true differences in how venues absorb permanent information. Therefore, this can yield mechanically time-varying discovery shares even in symmetric environments and makes the resulting measures difficult to interpret as structural contributions to the efficient price.

A full TVP-VECM with stochastic volatility allows both short-run dynamics (adjustment speeds and lag coefficients) and shock variances/covariances to drift over time. While conceptually appealing, this level of flexibility is often unnecessary—and can be empirically fragile—in high-frequency price discovery settings. The most pervasive and empirically dominant form of

time variation in high-frequency returns is time-varying volatility and time-varying correlations—including volatility clustering and strong intraday periodicity in second moments (e.g., Andersen and Bollerslev, 1997; Longin and Solnik, 2001; Zhang et al., 2005). In contrast, rapid drift in the short-run adjustment dynamics of a cointegrated price system is typically difficult to identify reliably at very high frequencies and can be confounded with market microstructure effects (e.g., bid–ask bounce, discreteness, asynchronous trading), which induce transient autocorrelation and time-varying dependence in observed returns (e.g., Hasbrouck, 1995; Ait-Sahalia et al., 2005; Hasbrouck, 2007; Aït-Sahalia and Jacod, 2014).

Motivated by this, we adopt a parsimonious specification in which the VECM coefficients are constant over the sample while the reduced-form innovation covariance matrix Ω_t evolves through stochastic volatility. This restriction delivers two benefits. First, it yields an analytically tractable mapping from the Bayesian model to structural price discovery objects at each time t, while preserving the economically meaningful separation between permanent and transitory shocks. Second, it isolates the empirically important channel through which high-frequency environments change—the volatility/correlation structure—without overparameterizing the short-run dynamics. Our simulation designs confirm that this class of models can reproduce a wide range of time variation implied by stylized structural price discovery models (including time-varying permanent and transitory heteroskedasticity and time-varying reduced-form correlation), while delivering stable and interpretable time-varying posterior inference for price discovery shares.

These considerations point to the need for a framework that (i) models the evolving joint distribution of innovations in related prices, (ii) defines price discovery structurally in terms of permanent shocks to the common trend, (iii) is invariant to arbitrary ordering of variables, and (iv) provides coherent uncertainty quantification for dynamic leadership.

This paper investigates how order-invariant Bayesian VAR models with stochastic volatility, which is developed by Chan et al. (2021), can be used to study the dynamics of time-varying price discovery in fragmented markets. The objective is to provide a flexible yet disciplined framework for the accurate estimation and interpretation of structural price discovery measures over time. We model the joint evolution of a vector of related prices in an order-invariant VAR with stochastic volatility (OI-VAR-SV) and then define price discovery in terms of each market's contribution to the permanent component of the common stochastic trend. Crucially, our structural decom-

position depends only on permanent shocks, explicitly separating information-bearing innovations from transitory noise and market microstructure disturbances. The estimation is carried out via Markov chain Monte Carlo (MCMC), delivering a full joint posterior for the stochastic volatilities, structural shocks, and resulting price discovery measures.

The paper makes three main contributions to the literature on price discovery in fragmented markets. First, we develop measures of time-varying price discovery that are grounded in the structural decomposition of permanent and transitory shocks. By focusing on each market's contribution to the permanent component of the common stochastic trend, we bring empirical measures closer to the economic notion of information about fundamentals and distinguish informational leadership from transitory noise. Second, we introduce an order-invariant Bayesian VAR with stochastic volatility framework that can be applied to systems of multiple assets and venues. Time variation in the shock variances, together with informative priors, allows the model to capture rich dynamics while avoiding overparameterization, and the identification scheme ensures that the resulting structural price discovery measures do not depend on arbitrary variable orderings. Third, by working in a fully Bayesian setting, we provide a comprehensive characterization of the dynamics and uncertainty of price discovery. Posterior distributions for dynamic price discovery measures at each point in time support probabilistic statements about which market leads in information incorporation and by how much, and they enable formal assessment of the statistical significance of changes in information leadership. In our simulation study, we demonstrate how different sources of time variation—in permanent or transitory shocks and in market efficiency—jointly shape the evolution of price discovery.

The remainder of the paper is structured as follows. Section 2 reviews the price discovery literature, emphasizing time-varying and regime-dependent approaches. Section 3 presents the structural cointegration-based framework and summarizes key price discovery measures. Section 4 develops our Bayesian order-invariant stochastic-volatility specification and its mapping to time-varying structural price measures. Section 5 presents simulation evidence highlighting how different sources of time variation affect competing measures. Section 6 concludes.

3.2 Related Literature

The literature on price discovery in fragmented markets is vast. Here we organize it around three themes: (i) classical cointegration-based measures of price discovery, (ii) structural refinements that separate permanent and transitory shocks, and (iii) time-varying and regime-dependent approaches, with a particular emphasis on Bayesian methods.

3.2.1 Classical Cointegration-Based Measures

The dominant empirical framework models arbitrage related prices as cointegrated and extracts price discovery measures from reduced-form VECMs. With a known cointegrating vector $\boldsymbol{\beta} = (1, -1)'$, the long-run common stochastic trend is interpreted as the efficient price, and the short-run dynamics capture temporary deviations due to noise and microstructure frictions. Within this setup, Hasbrouck (1995) introduces the information share (IS), which attributes variance in innovations to the efficient price to venue-specific shocks via a variance decomposition of the vector moving average (VMA) representation. Gonzalo and Granger (1995) propose the component share (CS), based on permanent–transitory (PT) decompositions of the price vector, and interpret the weights in the permanent component as each market's contribution to price discovery. These measures have been applied to equities, futures, options, commodities, credit, and foreign exchange markets (e.g., Baillie et al., 2002; Harris et al., 2002; De Jong, 2002).

Subsequent work connected IS and CS to microstructure variables such as bid–ask spreads, order flow, and liquidity provision, showing that spreads and quote sizes help explain cross-sectional and time-series variation in price impacts and price discovery shares (Madhavan et al., 1997; Hasbrouck and Seppi, 2001). With the advent of high-frequency data, Hasbrouck (2003) and others (e.g., Theissen, 2016; Ozturk et al., 2017; Hasbrouck, 2021; Shen et al., 2024b) study intraday price discovery and show that high-frequency and algorithmic trading can improve liquidity and enhance quote informativeness (Hendershott et al., 2011), that options markets can contribute nontrivially to discovery of the underlying price (Chakravarty et al., 2004), and that some high-frequency traders predict short-horizon returns (Brogaard et al., 2014; Fang et al., 2022).

A well-known limitation of IS is its dependence on the ordering of variables in the Cholesky factorization of the VECM residual covariance matrix, which yields only bounds for each market's

contribution. CS, while order-invariant, can be difficult to interpret structurally.

3.2.2 Order-Invariant Measures and Structural Refinements

To overcome the order dependence problem of IS, Lien and Shrestha (2009) proposed the modified information share (MIS) measure of price discovery based on the spectral decomposition of the correlation matrix of VECM residuals. Additionally, Sultan and Zivot (2015) introduced the price discovery share (PDS), which uniquely decomposes the variance of the efficient price into additive contributions of each market. Both MIS and PDS are order-invariant. However, they are reduced-form measures of price discovery and have ambiguous structural interpretations.

To improve structural interpretation, several authors recast the VECM in terms of permanent and transitory structural shocks and derive price discovery measures that depend only on the permanent component. Yan and Zivot (2007) and Yan and Zivot (2010) develop a structural cointegration model in which a permanent "news" shock moves the common efficient price, while a transitory "noise" shock represents uninformed trading and liquidity effects. They derive structural expressions for IS and CS and show that both are affected by the response of prices to transitory shocks. To mitigate this, Yan and Zivot (2010) propose an information leadership (IL) measure that combines IS and CS and captures relative responses to news in stylized settings with uncorrelated reduced-form innovations. Building on the structural decomposition of Yan and Zivot (2010), Shen et al. (2024b) combined PDS with CS to define the price information leadership share (PILS), which depends only on the impact of the permanent shock and is robust to correlation in reduced-form errors. Lautier et al. (2024) proposed covariance information shares (CovIS and CovISQ), which select as leader the price whose innovations have the highest covariance with those of the efficient price and show that CovISQ is equivalent to PILS. A related line of work emphasizes instantaneous efficiency. Shen and Zivot (2024) argue that, since both prices eventually move one-for-one with the efficient price following a permanent shock, leadership should be defined by the size of the instantaneous pricing error rather than the size of the instantaneous response. They proposed a price instantaneous efficiency share (PIES) based on the squared deviation of the initial impact to the permanent shock from the long-run response of unity. These structural measures motivate our focus on permanent shocks and order-invariant decompositions in a dynamic Bayesian environment.

3.2.3 Time-Varying and Regime-Dependent Approaches

A large empirical literature documents that the price discovery process is time-varying. A first and widely used approach estimates "static" VECMs on rolling or subsample windows and compares IS/CS across periods (Hasbrouck, 2003; Frijns et al., 2015; Hauptfleisch et al., 2016; Wallace et al., 2019; Narayan and Sharma, 2018; Conlon et al., 2022; Mohamad and Inani, 2023). This strategy is simple and intuitive, but it entails a trade-off between window length and temporal resolution: windows must be long enough for reliable estimation, yet long windows smooth out potentially interesting short-run dynamics and make results sensitive to window choice.

A second strand allows VECM parameters to vary smoothly over time. Dufour and Nguyen (2008) allow adjustment coefficients to depend on observed state variables; Bierens and Martins (2010) specify cointegrating vectors that evolve smoothly via orthogonal Chebyshev polynomials; and Taylor (2011), Avino et al. (2015), Qiao et al. (2019) and others estimate time-varying VECMs using kernel, spline, or multivariate GARCH methods. These models capture gradual evolution in adjustment speeds and cointegration relationships, but can be computationally demanding and do not always map directly into structural permanent–transitory decompositions.

A third line of research models discrete regime changes using Markov-switching frameworks. Giudici and Abu Hashish (2020), Segnon and Bekiros (2020), Kim and Linn (2022), Kuck and Schweikert (2023) and others show that price discovery in Bitcoin and commodity markets is regime-dependent, with parameters and leadership patterns changing across volatility and liquidity regimes. Conlon et al. (2022) emphasize regime shifts in Bitcoin spot–futures price discovery around the COVID-19 period. These models highlight the importance of allowing for time-varying dynamics but often treat price discovery measures as reduced-form functions of regime-specific VECM parameters and rely on Cholesky identification.

Finally, a growing literature applies Bayesian time-varying parameter VARs with stochastic volatility (TVP-VAR-SV) to financial markets, following Primiceri (2005) and the corrections in Del Negro and Primiceri (2015). These models use structured priors to regularize high-dimensional parameter spaces and allow both coefficients and volatilities to evolve over time. Mohamad and Inani (2023), for example, apply a TVP-VAR-SV to study Bitcoin spot–futures price discovery around the COVID-19 pandemic. Chan et al. (2021) develop order-invariant VARs with stochastic

volatility (OI-VAR-SV), which separate identification from variable ordering and are particularly suited for large systems.

Our paper is closely related to this third strand. We differ from most existing work by (i) explicitly linking time-varying price discovery measures to a structural permanent–transitory decomposition, (ii) using an order-invariant Bayesian VAR with stochastic volatility to avoid dependence on arbitrary variable orderings, and (iii) providing a fully Bayesian quantification of the time variation and uncertainty in structural price discovery measures, including PILS, CovISQ, and related benchmarks.

3.3 Modeling Static Price Discovery

In this section, we present our methodology for modeling static price discovery using cointegration. We first review the structural cointegration model of Yan and Zivot (2007) and Yan and Zivot (2010) and the associated reduced-form VECM. We then discuss identification of the parameters of the structural cointegration model. This is followed by a review of price discovery measures derived from both the reduced-form VECM and structural model.

3.3.1 Structural Cointegration Model

Let $\mathbf{p}_t = (p_{1t}, p_{2t})'$ denote a vector of log prices for two assets that are closely related by arbitrage. For example, these can be the same asset traded in different markets or two nearly identical assets traded in the same market. In the price discovery literature, it is assumed that these two prices series are intergated of order 1, or I(1), the price changes, $\Delta\mathbf{p}_t$, are stationary, or I(0), and that $\mathbf{p}_t$ is cointegrated with cointegrating vector $\boldsymbol{\beta} = (1, -1)'$.

Yan and Zivot (2010) start with the following structural moving average (SMA) representation of $\Delta\mathbf{p}_t$

$$\Delta\mathbf{p}_t = \mathbf{D}(L)\boldsymbol{\eta}_t = \mathbf{D}_0\boldsymbol{\eta}_t + \mathbf{D}_1\boldsymbol{\eta}_{t-1} + \mathbf{D}_2\boldsymbol{\eta}_{t-2} + \dots \tag{3.1}$$

$$\mathbf{D}(L) = \sum_{k=0}^{\infty} \mathbf{D}_k L^k, \quad \mathbf{D}_0 \neq \mathbf{I}, \tag{3.2}$$

where the elements of $\{\mathbf{D}_k\}_{k=0}^{\infty}$ are 1-summable and $\mathbf{D}_0$ is invertible. They assume that the number of structural shocks is equal to the number of observed prices, so that $\mathbf{D}(L)$ is invertible. The innovation to the common efficient price of the asset, η_t^P, is labeled the permanent shock and the noise innovation, η_t^T, is labeled the transitory shock so that $\boldsymbol{\eta}_t = \left(\eta_t^P, \eta_t^T\right)'$. These structural shocks are assumed to be serially and mutually uncorrelated with diagonal covariance matrix $\mathbf{C} = \text{diag}\left(\sigma_P^2, \sigma_T^2\right)$. The matrix $\mathbf{D}_0$ contains the initial impacts of the structural shocks on $\Delta\mathbf{p}_t$, and defines the contemporaneous correlation structure of $\Delta\mathbf{p}_t$:

$$\mathbf{D}_0 = \begin{pmatrix} d_{0,1}^P & d_{0,1}^T \\ d_{0,2}^P & d_{0,2}^T \end{pmatrix} = (\mathbf{d}_0^P \vdots \mathbf{d}_0^T). \tag{3.3}$$

The permanent innovation η_t^P carries new information on the fundamental value of the asset, and permanently moves the market prices. Given that the cointegrating vector is $\boldsymbol{\beta} = (1, -1)'$, the defining characteristic of η_t^P is that it has a one-to-one long-run effect on the price levels for each market. The transitory innovation η_t^T summarizes non-information related shocks, such as trading by uninformed or liquidity traders. The defining characteristic of η_t^T is that it is uncorrelated with the informational innovation η_t^P, and has no long-run effect on price levels. Hence, the long-run impact matrix $\mathbf{D}(1)$ of the structural innovations $\boldsymbol{\eta}_t$ takes the form

$$\mathbf{D}(1) = \begin{pmatrix} d_1^P(1) & d_1^T(1) \\ d_2^P(1) & d_2^T(1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = (\mathbf{d}^P(1) \vdots \mathbf{d}^T(1)). \tag{3.4}$$

The Beveridge-Nelson decomposition applied to Eq. (3.1) yields the level relationship:

$$\mathbf{p}_t = \mathbf{p}_0 + \mathbf{D}(1)\sum_{j=1}^{t} \boldsymbol{\eta}_j + \mathbf{s}_t \tag{3.5}$$

$$= \mathbf{p}_0 + \mathbf{1}\sum_{j=1}^{t} \eta_t^P + \mathbf{s}_t \tag{3.6}$$

$$= \mathbf{p}_0 + \mathbf{1}m_t + \mathbf{s}_t, \tag{3.7}$$

where $\mathbf{D}(1) = \sum_{k=0}^{\infty}\mathbf{D}_k$, $\mathbf{s}_t = (s_{1t}, s_{2t})' = \mathbf{D}^*(L)\boldsymbol{\eta}_t \sim I(0)$, and $\mathbf{D}^*(L) = -\sum_{j=k+1}^{\infty}\mathbf{D}_j$ $(k = 0, \ldots, \infty)$, and $m_t = m_{t-1} + \eta_t^P$. Multiplying (5) by $\boldsymbol{\beta} = (1, -1)'$ shows that $\boldsymbol{\beta}'\mathbf{p}_t \sim I(0)$ since $\boldsymbol{\beta}'\mathbf{D}(1) = \mathbf{0}$.

3.3.2 Reduced-form VECM

The reduced-form VECM associated with the structural model has the form:

$$\Delta \mathbf{p}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{p}_{t-1} + \sum_{j=1}^{k} \boldsymbol{\Gamma}_j \Delta \mathbf{p}_{t-j} + \boldsymbol{\varepsilon}_t, \tag{3.8}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$ is the error correction vector, $\boldsymbol{\Gamma}_j$ $(j = 1, \ldots, k)$ are the short-run coefficient matrices, and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is the vector of reduced-form VECM residuals with $E[\boldsymbol{\varepsilon}_t] = \mathbf{0}$ and

$$E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s'] = \begin{cases} \mathbf{0}, & \text{if } t \neq s, \\ \boldsymbol{\Omega}, & \text{otherwise,} \end{cases}$$

and the residual covariance matrix $\boldsymbol{\Omega}$ is

$$\boldsymbol{\Omega} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

where σ_i^2 denotes the variance of each market's idiosyncratic error and ρ denotes the correlation coefficient between these two errors.

From the Granger Representation Theorem, the VECM (3.8) can be inverted to give the reduced-form vector moving average (VMA) representation:

$$\Delta \mathbf{p}_t = \boldsymbol{\Psi}(L)\varepsilon_t = \varepsilon_t + \boldsymbol{\Psi}_1\varepsilon_{t-1} + \boldsymbol{\Psi}_2\varepsilon_{t-2} + \cdots, \tag{3.9}$$

and its integrated form (or Beveridge-Nelson decomposition):

$$\begin{aligned} \mathbf{p}_t &= \mathbf{p}_0 + \boldsymbol{\Psi}(1)\sum_{s=1}^{t} \boldsymbol{\varepsilon}_s + \boldsymbol{\Psi}^*(L)\boldsymbol{\varepsilon}_t \\ &= \mathbf{p}_0 + \boldsymbol{\Psi}(1)\sum_{s=1}^{t} \boldsymbol{\varepsilon}_s + \mathbf{s}_t, \end{aligned} \tag{3.10}$$

where $\boldsymbol{\Psi}(1) = \sum_{k=0}^{\infty} \boldsymbol{\Psi}_k$ with $\boldsymbol{\Psi}(L)$ and $\boldsymbol{\Psi}^*(L)$ being matrix polynomials in the lag operator, L, $\boldsymbol{\Psi}_t^*(k) = -\sum_{j=k+1}^{\infty} \boldsymbol{\Psi}_j$, and $\mathbf{s}_t = \boldsymbol{\Psi}^*(L)\boldsymbol{\varepsilon}_t \sim I(0)$. The matrix $\boldsymbol{\Psi}(1)$ contains the cumulative

impacts of the innovation ε_t on all future price movements, and acts as a measure of the long-run impact of ε_t on prices.

Johansen (1991) shows that the reduced-form VECM model and the integrated VMA model are linked through the factorization:

$$\boldsymbol{\Psi}(1) = \beta_{\perp}\boldsymbol{\Pi}\alpha'_{\perp},$$

where $\boldsymbol{\beta}_{\perp}$ is orthogonal to $\boldsymbol{\beta}$, and $\boldsymbol{\Pi} = \left(\boldsymbol{\alpha}'_{\perp}\left(\mathbf{I} - \sum_{j=1}^{k}\boldsymbol{\Gamma}_j\right)\boldsymbol{\beta}_{\perp}\right)^{-1}$ where $\mathbf{I}$ is the identity matrix and $\boldsymbol{\alpha}_{\perp}$ is orthogonal to $\boldsymbol{\alpha}$. We can see that $\boldsymbol{\Pi}$ for the bivariate VECM model is a scalar.

Since the cointegrating vector is $\boldsymbol{\beta} = (1, -1)'$ and $\boldsymbol{\beta}'\boldsymbol{\Psi}(1) = \mathbf{0}$ then the rows of $\boldsymbol{\Psi}(1)$ are the same and we can represent $\boldsymbol{\Psi}(1)$ as:

$$\boldsymbol{\Psi}(1) = \begin{pmatrix} \psi_1 & \psi_2 \\ \psi_1 & \psi_2 \end{pmatrix} = \mathbf{1}\boldsymbol{\psi}',$$

where $\boldsymbol{\psi} = (\psi_1, \psi_2)'$ is the common row of $\boldsymbol{\Psi}(1)$ and $\mathbf{1}$ is a 2×1 vector of ones. Then (3.10) can be re-written as

$$\begin{aligned} \mathbf{p}_t &= \mathbf{p}_0 + \mathbf{1}\sum_{s=1}^{t}\boldsymbol{\psi}'\boldsymbol{\varepsilon}_s + \mathbf{s}_t \\ &= \mathbf{p}_0 + \mathbf{1}\sum_{s=1}^{t}\eta_t^P + \mathbf{s}_t \\ &= \mathbf{p}_0 + \mathbf{1}m_t + s_t \end{aligned} \tag{3.11}$$

where $\eta_t^P = \boldsymbol{\psi}'\boldsymbol{\varepsilon}_t$ and $m_t = m_{t-1} + \eta_t^P$ is the efficient price.

3.3.3 Identification of the Structural Model

Comparing the SMA representation (3.1) and the reduced-form VMA representation (3.9) shows that the structural shocks $\boldsymbol{\eta}_t$ are related to the reduced-form shocks $\boldsymbol{\varepsilon}_t$ via

$$\boldsymbol{\varepsilon}_{t-j} = \mathbf{D}_0\boldsymbol{\eta}_{t-j},\ j = 0, 1, 2, \ldots \tag{3.12}$$

It follows that SMA coefficients $\mathbf{D}_j$ are related to the reduced-form VMA coefficients $\mathbf{\Psi}_j$ via

$$\mathbf{D}_j = \mathbf{D}_0 \mathbf{\Psi}_j. \tag{3.13}$$

In addition, the structural shock covariance matrix $\mathbf{C} = \text{diag}(\sigma_P^2, \sigma_T^2)$ is related to the reduced-form shock covariance matrix $\mathbf{\Omega}$ via

$$\mathbf{C} = \mathbf{D}_0^{-1} \mathbf{\Omega} \mathbf{D}_0^{-1'} \tag{3.14}$$

From these results, it is clear that identification of all the parameters of the SMA model requires the identification of all of the elements in $\mathbf{D}_0$.

Identification of the elements in $\mathbf{D}_0$ was first done by Yan and Zivot (2007) using the Gonzalo-Ng decomposition. Using direct calculations, Lautier et al. (2024) derived the same representation as Yan and Zivot (2007) for $\mathbf{d}_0^P$. The derivation is instructive. Using $\boldsymbol{\varepsilon}_t = \mathbf{D}_0 \boldsymbol{\eta}_t$ we have

$$\varepsilon_{it} = d_{0,i}^P \eta_t^P + d_{0,i}^T \eta_t^T \ (i = 1, 2). \tag{3.15}$$

Then, using $\eta_t^P = \boldsymbol{\psi}' \boldsymbol{\varepsilon}_t = \psi_1 \varepsilon_{1t} + \psi_2 \varepsilon_{2t}$, we have

$$\text{cov}(\varepsilon_{1t}, \eta_t^P) = \psi_1 \sigma_1^2 + \psi_2 \sigma_{12} = (\mathbf{\Omega}\boldsymbol{\psi})_1 = d_{0,1}^P \sigma_P^2, \tag{3.16}$$

$$\text{cov}(\varepsilon_{1t}, \eta_t^P) = \psi_2 \sigma_2^2 + \psi_1 \sigma_{12} = (\mathbf{\Omega}\boldsymbol{\psi})_2 = d_{0,2}^P \sigma_P^2, \tag{3.17}$$

where $(\mathbf{\Omega}\boldsymbol{\psi})_i$ denotes the i^{th} row of $\mathbf{\Omega}\boldsymbol{\psi}$. Using $\sigma_P^2 = \boldsymbol{\psi}' \mathbf{\Omega} \boldsymbol{\psi}$, it follows that

$$\mathbf{d}_0^P = \frac{\mathbf{\Omega}\boldsymbol{\psi}}{\boldsymbol{\psi}' \mathbf{\Omega} \boldsymbol{\psi}} = \boldsymbol{\beta}^P, \tag{3.18}$$

where $\boldsymbol{\beta}^P$ is the linear projection coefficient of $\boldsymbol{\varepsilon}_t$ on $\eta_t^P = \boldsymbol{\psi}' \boldsymbol{\varepsilon}_t$.

Identification of $\mathbf{d}_0^T$ requires additional assumptions. Using the Gonzalo-Nq decomposition Yan and Zivot (2007) showed that

$$\mathbf{d}_0^T = \begin{pmatrix} \psi_2/(\psi_1 + \psi_2) \\ -\psi_1/(\psi_1 + \psi_2) \end{pmatrix}. \tag{3.19}$$

Lautier et al. (2024) assumed that $\sigma_P^2 = \sigma_T^2$ and showed that this implies

$$\mathbf{d}_0^T = \begin{pmatrix} \psi_2 \sigma_1 \sigma_2 \sqrt{1-\rho^2} \\ \psi_1 \sigma_1 \sigma_2 \sqrt{1-\rho^2} \end{pmatrix}. \tag{3.20}$$

3.3.4 Static Price Discovery Measures

The empirical measures most widely used to quantify price discovery are the information share (IS) measure of Hasbrouck (1995) and the component share measure (CS) of Booth et al. (1999), Chu et al. (1999) and Harris et al. (2002). The IS measures are defined as:

$$\text{IS}_i = \frac{\left([\boldsymbol{\psi}'\mathbf{F}]_i\right)^2}{\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi}} \ (i = 1, 2), \tag{3.21}$$

where $\mathbf{F}$ is the lower triangular Cholesky factor of $\boldsymbol{\Omega}$, $[\boldsymbol{\psi}'\mathbf{F}]_i$ is the i^{th} element of the row matrix $\boldsymbol{\psi}'\mathbf{F}$. IS measures the contribution of a venue to the variance of the permanent shock. When $\boldsymbol{\Omega}$ is non-diagonal, IS is not unique and depends on the ordering of the variables in $\mathbf{p}_t$. The CS measures are defined as

$$\text{CS}_i = \frac{\psi_i}{\psi_1 + \psi_2} \ (i = 1, 2), \tag{3.22}$$

and measure a venue's permanent component contribution to the sum of the permanent component weights.

To address the order-dependence of IS, numerous studies have proposed different solutions and new measures. Most notably are the Modified Information Share (MIS) measure of Lien and Shrestha (2009), and the Price Discovery Share (PDS) measure based of Sultan and Zivot (2015) and Shen et al. (2025). The MIS measures are based on a spectral decomposition of the correlation matrix of the reduced-form residuals and Shen et al. (2025) shows that they have the form

$$\text{MIS}_i = \frac{\psi_i^2 \sigma_i^2 (1 + \sqrt{1-\rho^2})/2 + \psi_j^2 \sigma_j^2 (1 - \sqrt{1-\rho^2})/2 + \psi_i \psi_j \sigma_{i,j}}{\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi}} \ (i = 1, 2). \tag{3.23}$$

The PDS measures are based on the additive Euler decomposition of σ_P and are given by

$$\text{PDS}_i = \psi_i \beta_i^P = \frac{\psi_i^2 \sigma_i^2 + \sum_{j=1}^{n} \psi_i \psi_{j \neq i} \sigma_{i, j \neq i}}{\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi}}, \ (i = 1, 2) \tag{3.24}$$

where β_i^P is the projection coefficient from the projection of ε_{it} on η_t^P.

The price discovery measures IS, CS, MIS and PDS are based on residuals from a reduced-form vector error correction model (VECM) and do not have a clear structural interpretation. Using the structural cointegration model (3.1), Yan and Zivot (2010) and Shen et al. (2024a) derive the following structural representations of IS, CS, MIS and PDS[2]:

$$\mathrm{IS}_1 = \frac{d_{0,1}^P d_{0,2}^T}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}, \quad \mathrm{IS}_2 = \frac{-d_{0,1}^T d_{0,2}^P}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}, \tag{3.25}$$

$$\mathrm{CS}_1 = \frac{d_{0,2}^T}{d_{0,2}^T - d_{0,1}^T}, \quad \mathrm{CS}_2 = -\frac{d_{0,1}^T}{d_{0,2}^T - d_{0,1}^T}, \tag{3.26}$$

$$\mathrm{MIS}_1 = \frac{1}{2} + \frac{1}{2}\frac{d_{0,1}^P d_{0,2}^T + d_{0,1}^T d_{0,2}^P}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}\sqrt{1-\rho^2}, \quad \mathrm{MIS}_2 = \frac{1}{2} - \frac{1}{2}\frac{d_{0,1}^P d_{0,2}^T + d_{0,1}^T d_{0,2}^P}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}\sqrt{1-\rho^2}, \tag{3.27}$$

$$\mathrm{PDS}_1 = \frac{d_{0,1}^P d_{0,2}^T}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}, \quad \mathrm{PDS}_2 = \frac{-d_{0,1}^T d_{0,2}^P}{d_{0,1}^P d_{0,2}^T - d_{0,1}^T d_{0,2}^P}. \tag{3.28}$$

The structural representations show some surprising results. First, CS only depends on the contemporaneous impacts of the transitory shock and does not reflect price discovery at all. Second, IS, MIS and PDS respond to be permanent and transitory shocks and are not pure measures of price discovery.

To help sort out the confounding effects of responses to transitory shocks, Yan and Zivot (2010) proposed the information leadership measure (IL hereafter), a combination of the IS and CS measures, which in the restricted case with uncorrelated reduced-form residuals is defined as

$$\mathrm{IL}_1 = \left|\frac{\mathrm{IS}_1/\mathrm{CS}_1}{\mathrm{IS}_2/\mathrm{CS}_2}\right| = \left|\frac{\beta_1^P}{\beta_2^P}\right| = \left|\frac{d_{0,1}^P}{d_{0,2}^P}\right|, \quad \mathrm{IL}_2 = \left|\frac{\mathrm{IS}_2/\mathrm{CS}_2}{\mathrm{IS}_1/\mathrm{CS}_1}\right| = \left|\frac{\beta_2^P}{\beta_1^P}\right| = \left|\frac{d_{0,2}^P}{d_{0,1}^P}\right|. \tag{3.29}$$

For the more general case with possibly correlated reduced-form errors, Shen et al. (2024a) proposed the Price Information Leadership (PIL) measure, a combination of the PDS and CS measures, to unravel the confounding effects of transitory noise:

$$\mathrm{PIL}_1 = \left|\frac{\mathrm{PDS}_1/\mathrm{CS}_1}{\mathrm{PDS}_2/\mathrm{CS}_2}\right| = \left|\frac{\beta_1^P}{\beta_2^P}\right| = \left|\frac{d_{0,1}^P}{d_{0,2}^P}\right|, \quad \mathrm{PIL}_2 = \left|\frac{\mathrm{PDS}_2/\mathrm{CS}_2}{\mathrm{PDS}_1/\mathrm{CS}_1}\right| = \left|\frac{\beta_2^P}{\beta_1^P}\right| = \left|\frac{d_{0,2}^P}{d_{0,1}^P}\right|. \tag{3.30}$$

[2]The representation for IS is for the special case of uncorrelated reduced-form residuals.

Following Putniņš (2013), Shen et al. (2025) converted PIL into the Price Information Leadership Share (PILS) measure

$$\begin{aligned}\text{PILS}_1 &= \frac{\text{PIL}_1}{\text{PIL}_1 + \text{PIL}_2} = \frac{(\beta_1^P)^2}{(\beta_1^P)^2 + (\beta_2^P)^2} = \frac{(d_{0,1}^P)^2}{(d_{0,1}^P)^2 + (d_{0,2}^P)^2},\\ \text{PILS}_2 &= \frac{\text{PIL}_2}{\text{PIL}_1 + \text{PIL}_2} = \frac{(\beta_2^P)^2}{(\beta_1^P)^2 + (\beta_2^P)^2} = \frac{(d_{0,2}^P)^2}{(d_{0,1}^P)^2 + (d_{0,2}^P)^2}.\end{aligned} \tag{3.31}$$

PILS is a pure measure of price discovery as it depends only on the contemporaneous impacts of the permanent shock. A venue has a high PILS if it has a stronger contemporaneous impact from the permanent shock than the other venue.

Independently of Shen et al. (2025), Lautier et al. (2024) proposed the Covariance Information Share (CovIS) and the Squared Covariance Information Share (CovISQ) measures of price discovery. CovISQ is equivalent to PILS and CovIS is given by

$$\begin{aligned}\text{CovIS}_1 &= \frac{\beta_1^P}{\beta_1^P + \beta_2^P} = \frac{d_{0,1}^P}{d_{0,1}^P + d_{0,2}^P},\\ \text{CovIS}_2 &= \frac{\beta_2^P}{\beta_1^P + \beta_2^P} = \frac{d_{0,2}^P}{d_{0,1}^P + d_{0,2}^P}.\end{aligned} \tag{3.32}$$

The interpretation of PILS and CovIS can be misleading in the unusual case when $d_{0,i}^P$ is large and negative. Following Shen and Zivot (2024) we define the the instantaneous pricing error rule

$$\text{PIES}_1 = \frac{(\beta_1^P - 1)^2}{(\beta_1^P - 1)^2 + (\beta_2^P - 1)^2} = \frac{(d_{0,1}^P - 1)^2}{(d_{0,1}^P - 1)^2 + (d_{0,2}^P - 1)^2}, \tag{3.33}$$

$$\text{PIES}_2 = \frac{(\beta_2^P - 1)^2}{(\beta_1^P - 1)^2 + (\beta_2^P - 1)^2} = \frac{(d_{0,2}^P - 1)^2}{(d_{0,1}^P - 1)^2 + (d_{0,2}^P - 1)^2}. \tag{3.34}$$

to define the leading market as the market with the smaller instantaneous absolute pricing error $d_{0,i}^P - 1$, instead of the one with a larger instantaneous absolute response $d_{0,i}^P$.

Table 3.1 summarizes the different measures of price discovery in terms of both reduced-form and structural parameters. Note that the structural representations depend only on $\mathbf{D_0}$, σ_P^2 and σ_T^2. The reduced-form representations depend only on $\boldsymbol{\Omega}$ and $\boldsymbol{\psi}$.

Table 3.1: Price Discovery Measures

Structural Form	**Reduced Form**
Information Share: Hasbrouck (1995)	
Market 1 ranked First (F) in Cholesky decomposition:	
$\text{IS}_1^F = \frac{d_{0,1}^{P2}}{d_{0,1}^{P2}+d_{0,1}^{T2}}$	$\text{IS}_1^F = \frac{(\psi_1\sigma_1+\rho\psi_2\sigma_2)^2}{\psi_1^2\sigma_1^2+2\psi_1\psi_2\sigma_{12}+\psi_2^2\sigma_2^2}$
$\text{IS}_2^F = \frac{d_{0,1}^{T2}}{d_{0,1}^{P2}+d_{0,1}^{T2}}$	$\text{IS}_2^F = \frac{\psi_2^2\sigma_2^2\left(1-\rho^2\right)}{\psi_1^2\sigma_1^2+2\psi_1\psi_2\sigma_{12}+\psi_2^2\sigma_2^2}$
Market 1 ranked Second (S) in Cholesky decomposition:	
$\text{IS}_1^S = \frac{d_{0,2}^{T2}}{d_{0,2}^{P2}+d_{0,2}^{T2}}$	$\text{IS}_1^S = \frac{\psi_1^2\sigma_1^2\left(1-\rho^2\right)}{\psi_1^2\sigma_1^2+2\psi_1\psi_2\sigma_{12}+\psi_2^2\sigma_2^2}$
$\text{IS}_2^S = \frac{d_{0,2}^{P2}}{d_{0,2}^{P2}+d_{0,2}^{T2}}$	$\text{IS}_2^S = \frac{(\rho\psi_1\sigma_1+\psi_2\sigma_2)^2}{\psi_1^2\sigma_1^2+2\psi_1\psi_2\sigma_{12}+\psi_2^2\sigma_2^2}$
Hasbrouck's IS Arithmetic Averages:	
$\overline{\text{IS}}_1 = \frac{\text{IS}_1^F+\text{IS}_1^S}{2}; \overline{\text{IS}}_2 = \frac{\text{IS}_2^F+\text{IS}_2^S}{2}$	$\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}$ is a correlation coefficient
Covariance Information Share: Lautier et al. (2024)	
$\text{CovIS}_1 = \frac{d_{0,1}^P}{d_{0,1}^P + d_{0,2}^P}$	$\text{CovIS}_1 = \frac{\psi_1\sigma_1^2 + \psi_2\sigma_{12}}{\psi_1\sigma_1^2 + (\psi_1+\psi_2)\sigma_{12} + \psi_2\sigma_2^2}$
$\text{CovIS}_2 = \frac{d_{0,2}^P}{d_{0,1}^P + d_{0,2}^P}$	$\text{CovIS}_2 = \frac{\psi_1\sigma_{12} + \psi_2\sigma_2^2}{\psi_1\sigma_1^2 + (\psi_1+\psi_2)\sigma_{12} + \psi_2\sigma_2^2}$
Covariance Information Share: Quadiratic variation	
$\text{CovISQ}_1 = \frac{d_{0,1}^{P2}}{d_{0,1}^{P2} + d_{0,2}^{P2}}$	$\text{CovISQ}_1 = \frac{(\psi_1\sigma_1^2 + \psi_2\sigma_{12})^2}{(\psi_1\sigma_1^2 + \psi_2\sigma_{12})^2 + (\psi_1\sigma_{12} + \psi_2\sigma_2^2)^2}$
$\text{CovISQ}_2 = \frac{d_{0,2}^{P2}}{d_{0,1}^{P2} + d_{0,2}^{P2}}$	$\text{CovISQ}_2 = \frac{(\psi_1\sigma_{12} + \psi_2\sigma_2^2)^2}{(\psi_1\sigma_1^2 + \psi_2\sigma_{12})^2 + (\psi_1\sigma_{12} + \psi_2\sigma_2^2)^2}$
Component Shares/ Weights in common component	

Structural Form	**Reduced Form**
$\mathrm{CS}_1 = \dfrac{d^T_{0,2}}{d^T_{0,2} - d^T_{0,1}}$	$\mathrm{CS}_1 = \dfrac{\psi_1}{\psi_1 + \psi_2}$
$\mathrm{CS}_2 = -\dfrac{d^T_{0,1}}{d^T_{0,2} - d^T_{0,1}}$	$\mathrm{CS}_2 = \dfrac{\psi_2}{\psi_1 + \psi_2}$
Modified Information Share: Lien and Shrestha (2009)	
	$\mathrm{MIS}_i = \dfrac{\left([\boldsymbol{\psi}'\mathbf{F}^*]_i\right)^2}{\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi}} = \dfrac{\left(\boldsymbol{\psi}^*_i\right)^2}{\sum_{i=1}^n \left(\boldsymbol{\psi}^*_i\right)^2}$ where ψ^*_i represents the i^{th} element of the row matrix $\boldsymbol{\psi}'\mathbf{F}^*$
Price Discovery Share: Sultan and Zivot (2015)	
$\mathrm{PDS}_1 = \dfrac{d^P_{0,1} d^T_{0,2}}{d^P_{0,1} d^T_{0,2} - d^T_{0,1} d^P_{0,2}}$	$\mathrm{PDS}_i = \dfrac{\psi_i \frac{\partial \sigma_\eta(\boldsymbol{\psi})}{\partial \psi_i}}{\sigma_\eta(\boldsymbol{\psi})}$
$\mathrm{PDS}_2 = -\dfrac{d^T_{0,1} d^P_{0,2}}{d^P_{0,1} d^T_{0,2} - d^T_{0,1} d^P_{0,2}}$	where $\sigma_\eta(\boldsymbol{\psi}) = (\boldsymbol{\psi}'\boldsymbol{\Omega}\boldsymbol{\psi})^{1/2}$
Information Leadership Share: Yan and Zivot (2010); Putniņš (2013)	
$\mathrm{IL}_1 = -\frac{d^T_{0,1}\left(d^{P2}_{0,1} d^{P2}_{0,2} \sigma^4_P + 2 d^{P2}_{0,1} d^{T2}_{0,2} \sigma^2_P \sigma^2_T + d^{T2}_{0,1} d^{T2}_{0,2} \sigma^4_T\right)}{d^T_{0,2}\left(d^{P2}_{0,1} d^{P2}_{0,2} \sigma^4_P + 2 d^{P2}_{0,2} d^{T2}_{0,1} \sigma^2_P \sigma^2_T + d^{T2}_{0,1} d^{T2}_{0,2} \sigma^4_T\right)}, \quad \mathrm{IL}_2 = \frac{1}{\mathrm{IL}_1}$	
$\mathrm{ILS}_1 = \dfrac{\mathrm{IL}_1}{\mathrm{IL}_1 + \mathrm{IL}_2}, \quad \mathrm{ILS}_2 = \dfrac{\mathrm{IL}_2}{\mathrm{IL}_1 + \mathrm{IL}_2}$	
Modified Information Leadership Share: Shen et al. (2025)	
$\mathrm{MIL}_1 = \left\lvert \dfrac{\mathrm{MIS}_1/\mathrm{CS}_1}{\mathrm{MIS}_2/\mathrm{CS}_2} \right\rvert, \quad \mathrm{MIL}_2 = \left\lvert \dfrac{\mathrm{MIS}_2/\mathrm{CS}_2}{\mathrm{MIS}_1/\mathrm{CS}_1} \right\rvert$	
$\mathrm{MILS}_1 = \dfrac{\mathrm{MIL}_1}{\mathrm{MIL}_1 + \mathrm{MIL}_2}, \quad \mathrm{MILS}_2 = \dfrac{\mathrm{MIL}_2}{\mathrm{MIL}_1 + \mathrm{MIL}_2}$	
Price Information Leadership Share: Shen et al. (2025)	

Structural Form	Reduced Form
$\text{PIL}_1 = \left\lvert \frac{\text{PDS}_1/\text{CS}_1}{\text{PDS}_2/\text{CS}_2} \right\rvert = \left\lvert \frac{d^P_{0,1}}{d^P_{0,2}} \right\rvert, \quad \text{PIL}_2 = \left\lvert \frac{\text{PDS}_2/\text{CS}_2}{\text{PDS}_1/\text{CS}_1} \right\rvert = \left\lvert \frac{d^P_{0,2}}{d^P_{0,1}} \right\rvert$	
$\text{PILS}_1 = \frac{\text{PIL}_1}{\text{PIL}_1 + \text{PIL}_2} = \frac{(d^P_{0,1})^2}{(d^P_{0,1})^2 + (d^P_{0,2})^2},$	
$\text{PILS}_2 = \frac{\text{PIL}_2}{\text{PIL}_1 + \text{PIL}_2} = \frac{(d^P_{0,2})^2}{(d^P_{0,1})^2 + (d^P_{0,2})^2}$	
Price Instantaneous Efficiency Share: Shen and Zivot (2024)	
$\text{PIES}_1 = \frac{(d^P_{0,2} - 1)^2}{(d^P_{0,1} - 1)^2 + (d^P_{0,2} - 1)^2},$	
$\text{PIES}_2 = \frac{(d^P_{0,1} - 1)^2}{(d^P_{0,1} - 1)^2 + (d^P_{0,2} - 1)^2}$	

3.4 Modeling Time-varying Price Discovery

In this section, we describe several models for capturing time-varying price discovery. We first present a general time-varying parameter structural cointegration model and describe how time-varying structural price discovery measures can be computed. Next, we discuss the time-varying parameter VAR model with stochastic volatility (TVP-VAR-SV) of Primiceri (2005) and show that it is not well suited for modeling price discovery. Then we introduce the constant parameter VAR model with order invariant stochastic volatility of Chan et al. (2021) and show that it can be used to construct time-varying price discovery measures.

3.4.1 General Time-varying Parameter Structural Cointegration Model

A general time-varying parameter structural moving average model with stochastic volatility (TVP-SMA-SV) has the form

$$\Delta \mathbf{p}_t = \mathbf{D}_{0,t}\boldsymbol{\eta}_t + \mathbf{D}_{1,t}\boldsymbol{\eta}_{t-1} + \mathbf{D}_{2,t}\boldsymbol{\eta}_{t-2} + \cdots, \tag{3.35}$$

$$E[\boldsymbol{\eta}_t\boldsymbol{\eta}_t'] = \mathbf{C}_t, \tag{3.36}$$

where

$$\mathbf{D}_{j,t} = \begin{pmatrix} d^P_{j,1,t} & d^T_{j,1,t} \\ d^P_{j,2,t} & d^T_{j,2,t} \end{pmatrix}, \ j = 0, 1, 2, \ldots \tag{3.37}$$

$$\mathbf{C}_t = \text{diag}(\sigma^2_{P,t}, \sigma^2_{T,t}), \tag{3.38}$$

and $\mathbf{D}_{0,t}$ is invertible.

The levels representation (BN decomposition) is

$$\mathbf{p}_t = \mathbf{p}_0 + \mathbf{D}_t(1)\sum_{k=0}^{t}\boldsymbol{\eta}_t + \mathbf{D}^*_t(L)\boldsymbol{\eta}_t \tag{3.39}$$

$$= \mathbf{p}_0 + \mathbf{1}\sum_{k=0}^{t}\eta^P_t + \mathbf{s}_t \tag{3.40}$$

$$= \mathbf{p}_0 + \mathbf{1}m_t + \mathbf{s}_t \tag{3.41}$$

where $m_t = m_{t-1} + \eta_t^P \sim I(1)$ is the efficient price, and $\mathbf{s}_t = \mathbf{D}_t^*(L)\boldsymbol{\eta}_t \sim I(0)$ is the transitory shock.

The empirical time-varying parameter reduced-form VECM with stochastic volatility (TVP-VECM-SV) is

$$\Delta \mathbf{p}_t = \boldsymbol{\alpha}_t \boldsymbol{\beta}' \mathbf{p}_{t-1} + \sum_{j=1}^{k} \boldsymbol{\Gamma}_{j,t} \Delta \mathbf{p}_{t-j} + \boldsymbol{\varepsilon}_t, \tag{3.42}$$

$$E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] = \boldsymbol{\Omega}_t = \begin{pmatrix} \sigma_{1,t}^2 & \sigma_{12,t} \\ \sigma_{12,t} & \sigma_{2,t}^2 \end{pmatrix}. \tag{3.43}$$

The TVP-VECM-SV can be inverted to give the time-varying parameter reduced form VMA model with stochastic volatility (TVP-VMA-SV):

$$\Delta \mathbf{p}_t = \boldsymbol{\varepsilon}_t + \boldsymbol{\Psi}_{1,t} \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\Psi}_{2,t} \boldsymbol{\varepsilon}_{t-2} + \cdots \tag{3.44}$$

The levels representation (BN decomposition) is

$$\mathbf{p}_t = \mathbf{p}_0 + \boldsymbol{\Psi}_t(1) \sum_{k=0}^{t} \boldsymbol{\varepsilon}_t + \boldsymbol{\Psi}_t^*(L)\boldsymbol{\varepsilon}_t \tag{3.45}$$

$$= \mathbf{p}_0 + \mathbf{1} \sum_{k=0}^{t} \boldsymbol{\psi}_t' \boldsymbol{\varepsilon}_t + \boldsymbol{\Psi}_t^*(L)\boldsymbol{\varepsilon}_t \tag{3.46}$$

$$= \mathbf{p}_0 + \mathbf{1} m_t + \mathbf{s}_t \tag{3.47}$$

where $m_t = m_{t-1} + \eta_t^P \sim I(1)$ is the efficient price, $\eta_t^P = \boldsymbol{\psi}_t' \boldsymbol{\varepsilon}_t$ is the permanent shock, and $\mathbf{s}_t = \boldsymbol{\Psi}_t^*(L)\boldsymbol{\varepsilon}_t \sim I(0)$ is the transitory shock.

Equating $\boldsymbol{\varepsilon}_t$ in (3.44) with $\mathbf{D}_{0,t}\boldsymbol{\eta}_t$ in (3.35) allows us to link the reduced-form parameters with the structural parameters:

$$\boldsymbol{\Psi}_{j,t} = \mathbf{D}_{j,t}\mathbf{D}_{0,t-j}^{-1} \Rightarrow \mathbf{D}_{j,t} = \boldsymbol{\Psi}_{j,t}\mathbf{D}_{0,t-j} \tag{3.48}$$

$$\boldsymbol{\Omega}_t = \mathbf{D}_{0,t}\mathbf{C}_t\mathbf{D}_{0,t}' \Rightarrow \mathbf{C}_t = \mathbf{D}_{0,t}^{-1}\boldsymbol{\Omega}_t\mathbf{D}_{0,t}^{-1'} \tag{3.49}$$

Now, the elements of $\mathbf{D}_{0,t}$ associated with the permanent shock can be identified. The elements

of the first column of $\mathbf{D}_{0,t}$ have the form

$$d^P_{0,1,t} = \frac{\text{cov}\left(e_{1,t}, \eta^P_t\right)}{\text{var}\left(\eta^P_t\right)} = \frac{(\boldsymbol{\Omega}_t \boldsymbol{\psi}_t)_1}{\boldsymbol{\psi}_t \boldsymbol{\Omega}_t \boldsymbol{\psi}'_t}, \tag{3.50}$$

$$d^P_{0,2,t} = \frac{\text{cov}\left(e_{2,t}, \eta^P_t\right)}{\text{var}\left(\eta^P_t\right)} = \frac{(\boldsymbol{\Omega}_t \boldsymbol{\psi}_t)_2}{\boldsymbol{\psi}_t \boldsymbol{\Omega}_t \boldsymbol{\psi}'_t}. \tag{3.51}$$

The time-varying structural price discovery measures are then identified.

3.4.2 Primiceri's TVP-VAR for Price discovery

In this section, we consider constructing time-varying price discovery share measures computed from the TVP-BVAR model of Primiceri (2005) utilized recently by Mohamad and Inani (2023). Following Primiceri (2005) we can allow for time variation in the VECM parameters $\boldsymbol{\alpha}, \boldsymbol{\Gamma}_j$ and $\boldsymbol{\Omega}$ in (3.8) as follows. Rewrite the VECM as:

$$\Delta \mathbf{p}_t = \boldsymbol{\alpha}_t \boldsymbol{\beta}' \mathbf{p}_{t-1} + \sum_{j=1}^{k} \boldsymbol{\Gamma}_{j,t} \Delta \mathbf{p}_{t-j} + \boldsymbol{\varepsilon}_t, \tag{3.52}$$

$$E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}'_t] = \boldsymbol{\Omega}_t = \begin{pmatrix} \sigma^2_{1t} & \rho_t \sigma_{1t} \sigma_{2t} \\ \rho_t \sigma_{1t} \sigma_{2t} & \sigma^2_{2t} \end{pmatrix}, \tag{3.53}$$

where $\boldsymbol{\alpha_t} = (\alpha_{1,t}, \alpha_{2,t})'$ is the time-varying error correction speed of adjustment vector, $\boldsymbol{\Gamma}_{j,t} (i = 1, \ldots, k)$ are the time-varying short-run coefficient matrices, and $\boldsymbol{\Omega}_t$ is the time-varying residual covariance matrix. Primiceri (2005) considered the triangular reduction of $\boldsymbol{\Omega}_t$, defined by

$$\mathbf{A}_t \boldsymbol{\Omega}_t \mathbf{A}'_t = \boldsymbol{\Sigma}_t \boldsymbol{\Sigma}'_t \tag{3.54}$$

where $\mathbf{A}_t$ is the lower triangular matrix

$$\mathbf{A}_t = \begin{pmatrix} 1 & 0 \\ a_t & 1 \end{pmatrix}, \tag{3.55}$$

and $\mathbf{\Sigma}_t$ is the diagonal matrix

$$\mathbf{\Sigma}_t = \begin{pmatrix} \omega_{1,t} & 0 \\ 0 & \omega_{2,t} \end{pmatrix}. \tag{3.56}$$

Then (3.52) can be expressed as

$$\Delta \mathbf{p}_t = \boldsymbol{\alpha}_t \boldsymbol{\beta}' \mathbf{p}_{t-1} + \sum_{j=1}^{k} \mathbf{\Gamma}_{j,t} \Delta \mathbf{p}_{t-j} + \mathbf{A}_t^{-1} \mathbf{\Sigma}_t \mathbf{e}_t, \tag{3.57}$$

$$E[\mathbf{e}_t \mathbf{e}_t'] = \mathbf{I}_2. \tag{3.58}$$

The time-varying VECM (3.52) can be inverted into the time-varying reduced-form vector moving average model:

$$\Delta \mathbf{p}_t = \mathbf{\Psi}_t(L)\varepsilon_t = \varepsilon_t + \mathbf{\Psi}_{1,t}\varepsilon_{t-1} + \mathbf{\Psi}_{2,t}\varepsilon_{t-2} + \cdots, \tag{3.59}$$

with its time-varying integrated form (Beveridge-Nelson decomposition):

$$\mathbf{p}_t = \mathbf{p}_0 + \mathbf{\Psi}_t(1) \sum_{s=1}^{t} \boldsymbol{\varepsilon}_s + \Psi_t^*(L)\boldsymbol{\varepsilon}_t. \tag{3.60}$$

Here, $\mathbf{\Psi}_t(1) = \sum_{k=0}^{\infty} \mathbf{\Psi}_{k,t}$ is the time-varying long-run impact matrix and $\mathbf{\Psi}_t^*(k) = -\sum_{j=k+1}^{\infty} \mathbf{\Psi}_{j,t}$. The matrix $\mathbf{\Psi}_t(1)$ contains the time-t cumulative impacts of the innovations $\boldsymbol{\varepsilon}_t$ on all future price movements, and is a time-t measure of the long-run impact of $\boldsymbol{\varepsilon}_t$ on prices.

Using the Johansen factorization, $\mathbf{\Psi}_t(1)$ can be rewritten as

$$\mathbf{\Psi}_t(1) = \boldsymbol{\beta}_\perp \Pi_t \boldsymbol{\alpha}_{t\perp}', \tag{3.61}$$

where $\boldsymbol{\beta}_\perp$ is orthogonal to $\boldsymbol{\beta}$, $\boldsymbol{\alpha}_{t\perp}$ is orthogonal to $\boldsymbol{\alpha}_t$, and $\mathbf{\Pi}_t = \left(\boldsymbol{\alpha}_{t\perp}' \left(\mathbf{I}_2 - \sum_{j=1}^{k} \mathbf{\Gamma}_{j,t}\right) \boldsymbol{\beta}_\perp\right)^{-1}$. We can see that $\mathbf{\Pi}_t$ for the bivariate VECM model is a scalar. Since the cointegrating vector is $\boldsymbol{\beta} = (1, -1)'$, one choice for $\boldsymbol{\beta}_\perp$ is $\boldsymbol{\beta}_\perp = (1, 1)'$. If we denote $\boldsymbol{\alpha}_{t\perp} = (\gamma_{1,t}, \gamma_{2,t})'$, then we can represent $\mathbf{\Psi}_t(1)$ as:

$$\mathbf{\Psi}_t(1) = \mathbf{\Pi}_t \beta_\perp \alpha_{t\perp}' = \mathbf{\Pi} \begin{pmatrix} \gamma_{1,t} & \gamma_{2,t} \\ \gamma_{1,t} & \gamma_{2,t} \end{pmatrix}. \tag{3.62}$$

The rows of $\boldsymbol{\Psi}_t(1)$ are identical, and the long-run impacts of an innovation $\boldsymbol{\varepsilon}_t$ on each of the prices are identical. The increment $\boldsymbol{\psi}_t'\boldsymbol{\varepsilon}_t$ (with $\psi_{\mathbf{t}} = (\psi_{1,t}, \psi_{2,t})'$ the common row of $\boldsymbol{\Psi}_t(1)$) is the component of the price change that is permanently impounded into the price (presumably due to new information) which we denote as:

$$\eta_t^P = \boldsymbol{\psi}_t'\boldsymbol{\varepsilon}_t = \psi_{1,t}\varepsilon_{1,t} + \psi_{2,t}\varepsilon_{2,t}. \tag{3.63}$$

This common efficient price evolves as a random walk driven by the permanent shock η_t^P.

3.4.3 Structural Interpretation of Primiceri's TVP-VECM

The TVP-VAR model from Primiceri (2005) imposes undesirable restrictions due to the assumption of a triangular factorization $\boldsymbol{\Omega}_t$. To see this, The TVP-VECM (3.52) can be thought of as a particular structural TVP-VECM

$$\Delta\mathbf{p}_t = \boldsymbol{\alpha}_t\boldsymbol{\beta}'\mathbf{p}_{t-1} + \sum_{j=1}^{k}\boldsymbol{\Gamma}_{j,t}\Delta\mathbf{p}_{t-j} + \mathbf{A}_t^{-1}\boldsymbol{\eta}_t, \tag{3.64}$$

where $\mathbf{A}_t$ is the structural initial impact matrix, and $\boldsymbol{\eta}_t = \boldsymbol{\Sigma}_t\mathbf{e}_t$ is a vector of independent structural shocks. Equating $\boldsymbol{\varepsilon}_t$ from (3.44) with $\mathbf{A}_t^{-1}\eta_t$ from (3.35) gives $\eta_t = \mathbf{A}_t\varepsilon_t$. Therefore, we may represent the structural shocks as

$$\begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ a_t & 1 \end{pmatrix}\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{3.65}$$

which implies that

$$\eta_{1t} = \varepsilon_{1t}, \tag{3.66}$$

$$\eta_{2t} = a_t\varepsilon_{1t} + \varepsilon_{2t}. \tag{3.67}$$

If we interpret η_{1t} as the permanent shock η_t^P then we know that

$$\eta_t^P = \psi_{1,t}\varepsilon_{1,t} + \psi_{2,t}\varepsilon_{2,t} = \varepsilon_{1,t}$$

However, (3.4) implies that $\psi_{1,t} = 1$ and $\psi_{2,t} = 0$ and the long-run impact matrix $\mathbf{\Psi}(1)$ reduces to the structural long-run impact matrix

$$\mathbf{\Psi_t} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \mathbf{D}(1).$$

The underlying SMA model has the form

$$\Delta \mathbf{p}_t = \boldsymbol{\varepsilon}_t + \mathbf{\Psi}_{1,t}\boldsymbol{\varepsilon}_{t-1} + \mathbf{\Psi}_{2,t}\boldsymbol{\varepsilon}_{t-2} + \cdots \tag{3.68}$$

$$= \mathbf{D}_{0,t}\boldsymbol{\eta}_t + \mathbf{D}_{1,t}\boldsymbol{\eta}_{t-1} + \mathbf{D}_{2,t}\boldsymbol{\eta}_{t-2} + \cdots \tag{3.69}$$

where $\boldsymbol{\eta}_t = (\eta_t^P, \eta_t^T)'$, and $E[\boldsymbol{\eta}_t\boldsymbol{\eta}_t'] = \mathbf{\Sigma_t}$. Since $\boldsymbol{\eta}_t = \mathbf{A}_t\boldsymbol{\varepsilon}_t$, it follows that

$$\mathbf{D}_{0,t} = \begin{pmatrix} d_{0,1,t}^P & d_{0,1,t}^T \\ d_{0,2,t}^P & d_{0,2,t}^T \end{pmatrix} = \mathbf{A}_t^{-1} = \begin{pmatrix} 1 & 0 \\ -a_t & 1 \end{pmatrix}$$

So that $d_{0,1,t}^P = 1$ which is not time-varying, and $d_{0,2,t}^P = -a_t$ which is time-varying.

The structural price discovery meansures CovIS and PILS/CovISQ are then

$$\text{CovIS}_{1,t} = \frac{d_{0,1,t}^P}{d_{0,1,t}^P + d_{0,2,t}^P} = \frac{1}{1-a_t}, \quad \text{CovIS}_{2,t} = \frac{d_{0,2,t}^P}{d_{0,1,t}^P + d_{0,2,t}^P} = \frac{-a_t}{1-a_t}$$

$$\text{PILS}_{1,t} = \frac{(d_{0,1,t}^P)^2}{(d_{0,1,t}^P)^2 + (d_{0,2,t}^P)^2} = \frac{1}{1+a_t^2}, \quad \text{PILS}_{2,t} = \frac{(d_{0,2,t}^P)^2}{(d_{0,1,t}^P)^2 + (d_{0,2,t}^P)^2} = \frac{a_t^2}{1+a_t^2}.$$

Since a_t captures the correlation between p_{1t} and p_{2t}, if $a_t = 0$, then the two prices are uncorrelated and

$$\text{PILS}_{1,t} = 1, \quad \text{PILS}_{2,t} = 0$$

So Primiceri's triangular model places rather strong restrictions on the price discovery structural model.

3.4.4 Constant Parameter VECM with Order–invariant SV

Due to the limitations of Primiceri's TVP-VAR for modeling price discovery, we instead consider an order-invariant constant parameter VECM with time-varying stochastic volatility (abbreviated OI–VECM–SV):

$$\Delta \mathbf{p}_t = \boldsymbol{\alpha}\,\boldsymbol{\beta}'\mathbf{p}_{t-1} + \sum_{j=1}^{k} \boldsymbol{\Gamma}_j \Delta \mathbf{p}_{t-j} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_t), \tag{3.70}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$ is the error–correction vector and $\{\boldsymbol{\Gamma}_j\}_{j=1}^{k}$ are short–run matrices. Order–invariant identification is achieved by writing the shocks as

$$\boldsymbol{\varepsilon}_t = \mathbf{B}_0^{-1}\mathbf{e}_t, \qquad \mathbf{e}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_t), \qquad \boldsymbol{\Sigma}_t = \operatorname{diag}\big(e^{h_{1,t}}, e^{h_{2,t}}\big), \tag{3.71}$$

with unrestricted, nonsingular $\mathbf{B}_0$ and diagonal, time–varying $\boldsymbol{\Sigma}_t$. Hence $\boldsymbol{\Omega}_t = (\mathbf{B}_0'\boldsymbol{\Sigma}_t^{-1}\mathbf{B}_0)^{-1}$. Each log–variance follows a stationary AR(1):

$$h_{i,t} = \phi_i h_{i,t-1} + u_{i,t}^h, \qquad u_{i,t}^h \sim \mathcal{N}(0, \omega_i^2), \quad |\phi_i| < 1, \quad h_{i,1} \sim \mathcal{N}\Big(0, \frac{\omega_i^2}{1-\phi_i^2}\Big). \tag{3.72}$$

The representation (3.71) delivers identification that does not depend on the ordering of $\Delta\mathbf{p}_t$, while allowing the unconditional correlation of reduced–form innovations $\boldsymbol{\varepsilon}_t$ to evolve entirely through $\boldsymbol{\Sigma}_t$ and the static mixing matrix $\mathbf{B}_0$. In the empirical section, we keep $\boldsymbol{\alpha}$ and $\{\boldsymbol{\Gamma}_j\}$ constant over the sample, while all time variation enters via $\boldsymbol{\Omega}_t$. Extensions with drifting coefficients are straightforward.

Priors

Let $\boldsymbol{\gamma}_i$ denote the stacked coefficient vector in equation i ($i = 1, 2$), collecting the elements of $\{\boldsymbol{\Gamma}_j\}_{j=1}^{k}$ as well as deterministic terms (if any). We adopt a Minnesota–type horseshoe prior that shrinks more aggressively across lags and across "other" variables:

$$\begin{aligned}(\gamma_{i,j} \mid \kappa_1, \kappa_2, \psi_{i,j}) &\sim \mathcal{N}\big(m_{i,j},\ \kappa_{i,j}\psi_{i,j}C_{i,j}\big), \\ \sqrt{\psi_{i,j}} \sim \mathcal{C}^+(0,1), &\qquad \sqrt{\kappa_1}, \sqrt{\kappa_2} \sim \mathcal{C}^+(0,1),\end{aligned} \tag{3.73}$$

where $\mathcal{C}^+$ is the standard half–Cauchy, $\kappa_{i,j} = \kappa_1$ for own lags and $\kappa_{i,j} = \kappa_2$ for other-variable lags, and the Minnesota scaling constants are

$$C_{i,j} = \begin{cases} \frac{1}{l^2}, & \text{for the coefficient on the } l\text{th lag of variable } i, \\ \frac{s_i^2}{l^2 s_j^2}, & \text{for the coefficient on the } l\text{th lag of variable } j \neq i, \end{cases}$$

with s_r^2 the residual variance from an AR(4) fit to series r used for initialization. For levels data we set $m_{i,j} = 0$ (including intercepts unless otherwise noted). When all $\psi_{i,j} \equiv 1$, this reduces to a standard Minnesota prior.

For the order–invariant mixing matrix, write $\mathbf{B}_0' = (\mathbf{b}_1, \mathbf{b}_2)$. We use independent Gaussian priors

$$\mathbf{b}_i \sim \mathcal{N}(\mathbf{b}_{0,i}, \mathbf{V}_{b_i}), \qquad i = 1, 2, \tag{3.74}$$

with diffuse centers (e.g., $\mathbf{b}_{0,i} = \mathbf{0}$) and diagonal $\mathbf{V}_{b_i}$.

For the ARSV parameters in (3.72), we use independent priors

$$\phi_i \sim \mathcal{N}(\phi_{0,i}, V_{\phi_i}) \cdot \mathbb{I}(|\phi_i| < 1), \qquad \omega_i^2 \sim \mathcal{IG}(\nu_i, S_i), \qquad i = 1, 2, \tag{3.75}$$

where $\mathcal{IG}$ denotes the inverse–Gamma distribution. Hyperparameters are chosen to be weakly informative so that most regularization comes via the horseshoe components.

Posterior and Sampling

Conditional on $\{\mathbf{\Omega}_t\}_{t=1}^T$ and the cointegration restriction $(1, -1)'\mathbf{p}_t$, the likelihood implied by (3.70) is Gaussian:

$$\mathcal{L}(\boldsymbol{\alpha}, \{\mathbf{\Gamma}_j\}, \mathbf{B}_0, \{h_{i,t}\}) \;\propto\; \prod_{t=1}^{T} |\mathbf{\Omega}_t|^{-1/2} \exp\Big(-\tfrac{1}{2} \boldsymbol{\varepsilon}_t' \mathbf{\Omega}_t^{-1} \boldsymbol{\varepsilon}_t \Big).$$

Combining with the priors yields a conditionally conjugate/block–conditionally conjugate posterior. A standard and efficient Gibbs/Metropolis-within-Gibbs sampler cycles through:

1. **VAR coefficients $\{\mathbf{\Gamma}_j\}$ and $\boldsymbol{\alpha}$**: stack the regression for each equation with time–varying weights from $\mathbf{\Omega}_t^{-1}$; conditional on the local/global horseshoe scales, the posterior is Gaussian.

Update the local scales $\{\psi_{i,j}\}$ and globals (κ_1, κ_2) via scale–mixture representations of the half–Cauchy.

2. **Order–invariant mixing $\mathbf{B}_0$**: given $\{\boldsymbol{\varepsilon}_t\}$ and $\{h_{i,t}\}$, the full conditional for each row $\mathbf{b}_i$ is Gaussian (a weighted regression of reduced–form $\boldsymbol{\varepsilon}_t$ on structural shocks), updated independently across i.[3]

3. **Log–volatilities $\{h_{i,t}\}$ and ARSV parameters** (ϕ_i, ω_i^2): sample $\{h_{i,t}\}$ using state–space samplers for Gaussian ARSV (e.g., ancillarity–sufficiency interweaving, auxiliary mixture samplers), then draw (ϕ_i, ω_i^2) from their linear–Gaussian posteriors subject to $|\phi_i| < 1$.

The above blocks are iterated to obtain draws from the joint posterior of all unknowns. If desired, one can treat $\boldsymbol{\alpha}$ as known (from a preliminary cointegration step) to focus purely on time–varying covariance effects.

3.4.5 Time-varying Price Discovery Measures

The OI–VECM–SV in (3.70) features constant reduced-form dynamics $(\boldsymbol{\alpha}, \{\boldsymbol{\Gamma}_j\})$ and a constant order-invariant mixing matrix $\boldsymbol{B}_0$, and the reduced-form innovation covariance $\boldsymbol{\Omega}_t$ is time varying. The key is to distinguish (i) the statistical order-invariant factorization used for estimation from (ii) the economic permanent–transitory decomposition that defines price discovery.

Because $(\boldsymbol{\alpha}, \{\boldsymbol{\Gamma}_j\})$ are constant, the VECM admits the usual reduced-form VMA representation

$$\Delta \boldsymbol{p}_t = \boldsymbol{\Psi}(L)\boldsymbol{\varepsilon}_t = \sum_{\ell=0}^{\infty} \boldsymbol{\Psi}_\ell \boldsymbol{\varepsilon}_{t-\ell}, \qquad \boldsymbol{\Psi}_0 = \boldsymbol{I}, \tag{3.76}$$

where $\{\boldsymbol{\Psi}_\ell\}$ depends only on $(\boldsymbol{\alpha}, \{\boldsymbol{\Gamma}_j\})$ and therefore is time-invariant. All time variation enters through the distribution of $\boldsymbol{\varepsilon}_t$ via $\boldsymbol{\Omega}_t$.

Under the OI–VECM–SV, we write

$$\boldsymbol{\varepsilon}_t = \boldsymbol{B}_0^{-1}\boldsymbol{e}_t, \qquad \boldsymbol{e}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_t), \qquad \boldsymbol{\Sigma}_t = \operatorname{diag}(e^{h_{1,t}}, e^{h_{2,t}}), \tag{3.77}$$

[3]Equivalently, draw $\mathbf{B}_0$ by stacking moment conditions $\mathbf{B}_0\boldsymbol{\varepsilon}_t = \mathbf{e}_t$ with diagonal $\boldsymbol{\Sigma}_t$; see Chan et al. (2024) for details.

so that

$$\boldsymbol{\Omega}_t = (\boldsymbol{\varepsilon}_t) = \boldsymbol{B}_0^{-1}\boldsymbol{\Sigma}_t\boldsymbol{B}_0^{-1\prime}. \tag{3.78}$$

Here $\boldsymbol{B}_0$ is a static matrix that delivers order-invariant identification for the SV component; time variation in $\boldsymbol{\Omega}_t$ is driven by $\boldsymbol{\Sigma}_t$ (and its interaction with $\boldsymbol{B}_0$). Importantly, $\boldsymbol{B}_0$ is not the structural impact matrix that defines permanent and transitory shocks in the structural cointegration model.

Price discovery is defined by the innovation to the common efficient price in Section 3.3:

$$\eta_t^P = \boldsymbol{\psi}'\,\boldsymbol{\varepsilon}_t.$$

Even though $\boldsymbol{\psi}$ is constant, the projection of reduced-form innovations onto the permanent shock is time varying because it depends on $\boldsymbol{\Omega}_t$:

$$\boldsymbol{d}_{0,t}^P \;\equiv\; \begin{pmatrix} d_{0,1,t}^P \\ d_{0,2,t}^P \end{pmatrix} \;=\; \frac{(\boldsymbol{\varepsilon}_t, \eta_t^P)}{(\eta_t^P)} \;=\; \frac{\boldsymbol{\Omega}_t\boldsymbol{\psi}}{\boldsymbol{\psi}'\boldsymbol{\Omega}_t\boldsymbol{\psi}}. \tag{3.79}$$

Thus, although $\boldsymbol{B}_0$ is constant, the economically meaningful contemporaneous permanent-impact vector $\boldsymbol{d}_{0,t}^P$ (and therefore the price-discovery measures built from it) *changes over time* through $\boldsymbol{\Omega}_t$.

Let $\boldsymbol{\eta}_t = (\eta_t^P, \eta_t^T)'$ denote economic shocks with η_t^P a permanent shock and η_t^T a transitory shock normalized to have no long-run effect on prices. As in the static structural cointegration model, we write

$$\boldsymbol{\varepsilon}_t = \boldsymbol{D}_{0,t}\boldsymbol{\eta}_t, \qquad \boldsymbol{D}_{0,t} = \big(\boldsymbol{d}_{0,t}^P \;\; \boldsymbol{d}_0^T\big), \tag{3.80}$$

where $\boldsymbol{d}_0^T$ can be fixed using the same normalization adopted in the static model (e.g., the Gonzalo–Ng/Yan–Zivot choice based on $\boldsymbol{\psi}$), so that time variation is concentrated in the permanent-impact column $\boldsymbol{d}_{0,t}^P$. Under this construction, $\boldsymbol{D}_{0,t}$ is an impact matrix implied by $(\boldsymbol{\psi}, \boldsymbol{\Omega}_t)$, and it should not be conflated with the statistical mixing matrix $\boldsymbol{B}_0$ in (3.77).

Combining (3.76) and (3.80) yields a SMA representation with time-varying impacts:

$$\Delta\boldsymbol{p}_t = \boldsymbol{\Psi}(L)\boldsymbol{D}_{0,\cdot}\boldsymbol{\eta}_\cdot = \sum_{\ell=0}^{\infty}\boldsymbol{\Psi}_\ell\boldsymbol{D}_{0,t-\ell}\,\boldsymbol{\eta}_{t-\ell} \;\equiv\; \sum_{\ell=0}^{\infty}\boldsymbol{D}_{\ell,t}\,\boldsymbol{\eta}_{t-\ell}, \qquad \boldsymbol{D}_{\ell,t} = \boldsymbol{\Psi}_\ell\boldsymbol{D}_{0,t-\ell}, \tag{3.81}$$

so in particular

$$\boldsymbol{D}_{0,t} = \boldsymbol{\Psi}_0 \boldsymbol{D}_{0,t} = \boldsymbol{D}_{0,t}. \tag{3.82}$$

Therefore, even with constant VECM coefficients, the SMA coefficients are time varying whenever $\boldsymbol{D}_{0,t}$ is time varying, which occurs here because $\boldsymbol{d}^P_{0,t}$ depends on $\boldsymbol{\Omega}_t$. For horizon h, we report stochastic cumulative impulse responses (SCIRFs) computed draw-by-draw as

$$\mathrm{SCIRF}^P_t(h) = \sum_{\ell=0}^{h} \boldsymbol{D}^P_{\ell,t}, \qquad \mathrm{SCIRF}^T_t(h) = \sum_{\ell=0}^{h} \boldsymbol{D}^T_{\ell,t}, \tag{3.83}$$

where $\boldsymbol{D}^P_{\ell,t}$ and $\boldsymbol{D}^T_{\ell,t}$ denote the columns of $\boldsymbol{D}_{\ell,t}$ associated with η^P_t and η^T_t.

At posterior draw s and time t, form

$$\boldsymbol{\Omega}^{(s)}_t = \boldsymbol{B}^{(s)-1}_0 \boldsymbol{\Sigma}^{(s)}_t \boldsymbol{B}^{(s)-1\prime}_0, \qquad \boldsymbol{d}^{P,(s)}_{0,t} = \frac{\boldsymbol{\Omega}^{(s)}_t \boldsymbol{\psi}}{\boldsymbol{\psi}' \boldsymbol{\Omega}^{(s)}_t \boldsymbol{\psi}}. \tag{3.84}$$

Then the time-varying structural measures that depend only on the contemporaneous permanent impacts are computed pointwise:

$$\mathrm{PILS}^{(s)}_{i,t} = \mathrm{CovISQ}^{(s)}_{i,t} = \frac{\left(d^{P,(s)}_{0,i,t}\right)^2}{\left(d^{P,(s)}_{0,1,t}\right)^2 + \left(d^{P,(s)}_{0,2,t}\right)^2}, \tag{3.85}$$

$$\mathrm{CovIS}^{(s)}_{i,t} = \frac{d^{P,(s)}_{0,i,t}}{d^{P,(s)}_{0,1,t} + d^{P,(s)}_{0,2,t}}, \tag{3.86}$$

$$\mathrm{PIES}^{(s)}_{i,t} = \frac{\left(d^{P,(s)}_{0,i,t} - 1\right)^2}{\left(d^{P,(s)}_{0,1,t} - 1\right)^2 + \left(d^{P,(s)}_{0,2,t} - 1\right)^2}, \tag{3.87}$$

and analogously for other measures using their reduced-form representations with $\boldsymbol{\Omega}^{(s)}_t$. We summarize the time-varying price discovery measures by posterior means and credible bands at each t.

The OI–VECM–SV separates three objects: (i) constant propagation dynamics $\{\boldsymbol{\Psi}_\ell\}$ determined by $(\boldsymbol{\alpha}, \{\boldsymbol{\Gamma}_j\})$; (ii) time-varying second moments $\boldsymbol{\Omega}_t$ driven by stochastic volatility $\boldsymbol{\Sigma}_t$ (and the static mixing $\boldsymbol{B}_0$); and (iii) time-varying permanent impacts $\boldsymbol{d}^P_{0,t}$ induced by (3.79), which in turn generate time variation in structural shares and in the SMA impacts $\boldsymbol{D}_{\ell,t}$ through (3.81).

3.5 A Partial Price Adjustment Model

3.5.1 Model Illustration

To illustrate the empirical performance of the discussed price discovery measures, we consider the stylized partial adjustment micro-structure model utilized by Amihud and Mendelson (1987), Hasbrouck and Ho (1987) and Yan and Zivot (2010):

$$p_{1t} = p_{1,t-1} + \delta_1 (m_t - p_{1,t-1}) + b_{0,1}^T \eta_t^T \tag{3.88}$$

$$p_{2t} = p_{2,t-1} + \delta_2 (m_t - p_{2,t-1}) + b_{0,2}^T \eta_t^T \tag{3.89}$$

$$m_t = m_{t-1} + \eta_t^P, \boldsymbol{\eta}_t = \left(\eta_t^P, \eta_t^T\right)' \sim i.i.d.N\left(\mathbf{0}, \begin{bmatrix} \sigma_P^2 & 0 \\ & \sigma_T^2 \end{bmatrix}\right). \tag{3.90}$$

Solving for Δp_{it} gives

$$\Delta p_{it} = d_i^P(L)\eta_t^P + d_i^T(L)\eta_t^T$$

where

$$d_1^P(L) = [1 - (1-\delta_1)L]^{-1}\delta_1, \quad d_2^P(L) = [1 - (1-\delta_2)L]^{-1}\delta_2,$$
$$d_1^T(L) = [1 - (1-\delta_1)L]^{-1}(1-L)b_{0,1}^T, \quad d_2^T(L) = [1 - (1-\delta_2)L]^{-1}(1-L)b_{0,2}^T$$

The SMA representation of the partial adjustment model is determined from the appropriate elements of the lag polynomials $d_i^P(L)$ and $d_i^T(L)$. In particular, the initial impact and the long-run impact matrices are given by:

$$\mathbf{D}_0 = \begin{pmatrix} d_{0,1}^P & d_{0,1}^T \\ d_{0,2}^P & d_{0,2}^T \end{pmatrix} = \begin{pmatrix} \delta_1 & b_{0,1}^T \\ \delta_2 & b_{0,2}^T \end{pmatrix}, \quad \mathbf{D}(1) = \begin{pmatrix} d_1^P(1) & d_1^T(1) \\ d_2^P(1) & d_2^T(1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \tag{3.91}$$

The initial responses of the two assets to a one-unit permanent shock are equal to δ_i and the long-run responses to one-unit permanent shock are equal one. The asset with a larger value of the initial response is identified as the leader in the process of the price discovery.

Rewrite Δp_{it} in the reduced form as (1), we have

$$\alpha_1 = \frac{-b_{0,1}^T \delta_1 \delta_2}{b_{0,1}^T \delta_2 - b_{0,2}^T \delta_1}, \quad \alpha_2 = \frac{-b_{0,2}^T \delta_1 \delta_2}{b_{0,2}^T \delta_1 - b_{0,1}^T \delta_2}$$

$$\varepsilon_1 = \delta_1 \eta_t^P + b_{0,1}^T \eta_t^T, \quad \varepsilon_2 = \delta_2 \eta_t^P + b_{0,2}^T \eta_t^T$$

The CS, PDS, CovIS, PILS and PIES measures are independent of any reduced form correlation and have structural representations:

$$\begin{aligned}
\mathrm{CS}_1 &= \frac{b_{0,2}^T}{b_{0,2}^T - b_{0,1}^T}, \quad \mathrm{CS}_2 = \frac{-b_{0,1}^T}{b_{0,2}^T - b_{0,1}^T} \\
\mathrm{PDS}_1 &= \frac{\delta_1 b_{0,2}^T}{\delta_1 b_{0,2}^T - b_{0,1}^T \delta_2}, \quad \mathrm{PDS}_2 = \frac{-b_{0,1}^T \delta_2}{\delta_1 b_{0,2}^T - b_{0,1}^T \delta_2} \\
\mathrm{PILS}_1 = \mathrm{CovISQ}_1 &= \frac{\delta_1^2}{\delta_1^2 + \delta_2^2}, \quad \mathrm{PILS}_2 = \mathrm{CovISQ}_2 = \frac{\delta_2^2}{\delta_1^2 + \delta_2^2} \\
\mathrm{CovIS}_1 &= \frac{\delta_1}{\delta_1 + \delta_2}, \quad \mathrm{PILS}_2 = \frac{\delta_2}{\delta_1 + \delta_2} \\
\mathrm{PIES}_1 &= \frac{(\delta_2 - 1)^2}{(\delta_1 - 1)^2 + (\delta_2 - 1)^2}, \quad \mathrm{PIES}_2 = \frac{(\delta_1 - 1)^2}{(\delta_1 - 1)^2 + (\delta_2 - 1)^2}
\end{aligned}$$

provided $b_{0,2}^T \neq b_{0,1}^T$. Note that when $\delta_1, \delta_2 \in [0, 1]$, then PIES is equivalent to PILS.

The structural representations of IS, MIS, IL, and ILS for these two products is complicated by the covariance structure of the reduced-form forecasting errors. When the reduced-form innovations, ε_t, are uncorrelated, the IS metric is unique. It can be shown that $\mathrm{cov}\,(\varepsilon_{1t}, \varepsilon_{2t}) = 0$ when

$$\frac{\sigma_T^2}{\sigma_P^2} = \frac{d_{0,1}^P d_{0,2}^P}{-d_{0,1}^T d_{0,2}^T} = \frac{\delta_1 \delta_2}{-b_{0,1}^T b_{0,2}^T}$$

and all elements of $\mathbf{D_0}$ are non-zero (hence $|\mathbf{D_0}| \neq 0$). Then the IS measures are uniquely defined as follows:

$$\mathrm{IS}_1 = \frac{\delta_1 b_{0,2}^T}{\delta_1 b_{0,2}^T - b_{0,1}^T \delta_2}, \quad \mathrm{IS}_2 = \frac{-b_{0,1}^T \delta_2}{\delta_1 b_{0,2}^T - b_{0,1}^T \delta_2}.$$

The whole price discovery process shows the progressive absorption of permanent and transitory shocks. We compute and draw the SCIRFs for both the permanent (P) and the transitory (T) shocks:

$$\mathrm{SCIRF}^P(h) = \sum_{i=0}^{i=h} d_i^P \quad \text{and} \quad \mathrm{SCIRF}^T(h) = \sum_{i=0}^{i=h} d_i^T.$$

3.5.2 Time-varying Dynamics

To study how price discovery evolves when the data-generating process changes over time, we allow three building blocks of the partial adjustment model to vary with t:

1. the volatility of the *permanent* innovation, $\sigma^2_{P,t}$;
2. the volatility of the *transitory* innovations, $\sigma^2_{T,t}$;
3. the *adjustment speeds*, $\delta_{i,t} \in (0,1)$, which govern how fast venue i absorbs the permanent shock.

In the simulations, time variation is introduced parsimoniously:

$$\eta^P_t \sim \mathcal{N}\big(0, \sigma^2_{P,t}\big), \qquad \eta^T_t \sim \mathcal{N}\big(0, \sigma^2_{T,t}\big),$$

and $\delta_{i,t}$ follows one of: constant, AR(1) around a mean, random walk on the probability or logit scale, or a smooth sinusoidal/logit–RW path. Volatility paths $\{\sigma^2_{P,t}\}, \{\sigma^2_{Ti,t}\}$ can be i.i.d. or GARCH(1,1).

Let the reduced-form one-step errors be

$$\varepsilon_{it} = \delta_{i,t}\, \eta^P_t + b^T_{0,i}\, \eta^T_t, \qquad i \in \{1,2\}. \tag{3.92}$$

Then, period by period,

$$\mathrm{Var}(\varepsilon_{it}) = \delta^2_{i,t}\, \sigma^2_{P,t} + (b^T_{0,i})^2\, \sigma^2_{T,t}, \tag{3.93}$$

$$\mathrm{Cov}(\varepsilon_{1t}, \varepsilon_{2t}) = \delta_{1,t}\delta_{2,t}\, \sigma^2_{P,t} \;+\; b^T_{0,1} b^T_{0,2}\, \sigma^2_{T,t}, \tag{3.94}$$

$$\mathrm{Corr}(\varepsilon_{1t}, \varepsilon_{2t}) \equiv \rho^{(\varepsilon)}_t = \frac{\delta_{1,t}\delta_{2,t}\sigma^2_{P,t} + b^T_{0,1} b^T_{0,2}\, \sigma^2_{T,t}}{\sqrt{\delta^2_{1,t}\sigma^2_{P,t} + (b^T_{0,1})^2\sigma^2_{T,t}}\, \sqrt{\delta^2_{2,t}\sigma^2_{P,t} + (b^T_{0,2})^2\sigma^2_{T,t}}}. \tag{3.95}$$

Hence, even though the structural shocks (η^P_t, η^T_t) are orthogonal, the reduced-form errors ε_t generally display time-varying correlation driven by the time variation in $\{\delta_{i,t}\}$ and $\{\sigma^2_{P,t}, \sigma^2_{T,t}\}$. In the "no-covariance" design, we choose $\{\sigma^2_{T,t}\}$ (given $\{\delta_{i,t}\}$ and $\sigma^2_{P,t}$) so that $Cov(\varepsilon_{1t}, \varepsilon_{2t}) \equiv 0$ for

each t, i.e.,

$$\sigma^2_{T,t} \;=\; -\,\frac{\delta_{1,t}\delta_{2,t}}{b^T_{0,1}b^T_{0,2}}\,\sigma^2_{P,t}, \tag{3.96}$$

whenever $b^T_{0,1}b^T_{0,2} < 0$ so that the right-hand side is nonnegative.

Price-discovery measures split into two classes as below:

(i) Structural, Σ-invariant (period by period). CS_i, PDS_i, PIL_i, PILS_i, CovIS_i, PIES_i depend only on $\{\delta_{i,t}\}$ and $b^T_{0,i}$. With time-varying $\delta_{i,t}$ they evolve deterministically with $\delta_{i,t}$. In particular, for each t,

$$\mathrm{PILS}_{i,t} = \frac{\delta^2_{i,t}}{\delta^2_{1,t}+\delta^2_{2,t}}, \qquad \mathrm{CovIS}_{i,t} = \frac{\delta_{i,t}}{\delta_{1,t}+\delta_{2,t}}, \tag{3.97}$$

$$\mathrm{PIES}_{1,t} = \frac{(1-\delta_{2,t})^2}{(1-\delta_{1,t})^2+(1-\delta_{2,t})^2}, \qquad \mathrm{PIES}_{2,t} = \frac{(1-\delta_{1,t})^2}{(1-\delta_{1,t})^2+(1-\delta_{2,t})^2}. \tag{3.98}$$

When $\delta_{i,t} \in [0,1]$, $\mathrm{PIES}_{i,t} \equiv \mathrm{PILS}_{i,t}$.

(ii) Correlation–sensitive. IS_i, MIS_i, IL_i, ILS_i depend on $\{\delta_{i,t}\}$ *and* on the reduced-form covariance in (3.93)–(3.94). If $\rho^{(\varepsilon)}_t = 0$, Hasbrouck IS_i is point-identified (the interval collapses); otherwise IS_i is interval-identified and its midpoint (MIS_i) and leadership ratio (ILS_i) co-move with volatility and correlation regimes.

Five empirical scenarios. We implement five cases to separate channels of time variation:

1. $\eta^P_t \sim i.i.d.$, $\sigma^2_P = 1$, $\eta^T_t \sim i.i.d.$, $\sigma^2_T = 10$, δ constant: Homoskedastic benchmark.
2. $\eta^P_t \sim$ GARCH(1,1), $\eta^T_t \sim i.i.d.$, $\sigma^2_T = 10$, δ constant: Permanent volatility clusters.
3. $\eta^P_t \sim i.i.d.$, $\sigma^2_T = 10$, $\eta^T_t \sim$ GARCH(1,1), δ constant: Transitory heteroskedasticity.
4. η^P_t $i.i.d.$, $\sigma^2_P = 1$, η^T_t $i.i.d.$, $\sigma^2_T = 10$, δ_t AR(1): Slowly evolving adjustment speeds.
5. η^P_t $i.i.d.$, $\sigma^2_P = 1$, η^T_t $i.i.d.$, $\sigma^2_T = 10$, δ_t smooth RW: Smooth leadership cycles.

3.5.3 Time-invariant δ

In this and the following subsection, we evaluate the empirical performance of the discussed price discovery measures using simulated data from different parameterizations of the stylized partial

adjustment model.

we firstly consider the sub-case when the initial responses to a one-unit permanent shock δ_i are time constant. We set the markets' responses to the transitory shock as $\left(b_{0,1}^T, b_{0,2}^T\right) = (0.8, -0.2)$, and $\delta_2 = 1 - \delta_1$, where δ_1 takes values from 0.9 to 0.1 with a reduction of 0.1. When $\delta_1 > \delta_2$ (i.e., $\delta_1 > 0.5$), Market 1 has a greater speed of price discovery than Market 2. We simulate 1000 samples of 21600 observations of two price series for each given value of δ_1. For each simulated sample, we estimate the Bayesian OI-VAR-SV model and calculate the corresponding price discovery measures. We take 10,000 interations and 5,000 burn-in using the Gibbs sampler. We summarize the averages of each price discovery measure over the 1000 samples and report results in Panel A of Table 3.2.

We consider three cases:

1. $\eta_t^P \sim i.i.d.$, $\sigma_P^2 = 1$, $\eta_t^T \sim i.i.d.$, $\sigma_T^2 = 10$, δ constant.

2. $\eta_t^P \sim$ GARCH(1,1), $\eta_t^T \sim i.i.d.$, $\sigma_T^2 = 10$, δ constant.

3. $\eta_t^P \sim i.i.d.$, $\sigma_T^2 = 10$, $\eta_t^T \sim$ GARCH(1,1), δ constant.

We report accuracy of time-varing measures in three panels of Table 3.2. The accuracy means the percentage of time that the shares identify the leadership correctly. When $\delta_1 = \delta_2 = 0.5$, we specify the accuracy as the value of measure between 0.45 to 0.55. The figures of price dicovery measures are in Appendix B.

When both permanent and transitory innovations are *i.i.d.* (Panel A), the measures that are structurally tied to the adjustment speeds—in particular PILS, CovIS, and its quadratic variant CovISQ—perform exceptionally well across the entire range of δ_1. For sizeable leadership gaps (e.g. $\delta_1 \geq 0.7$ or $\delta_1 \leq 0.3$), PILS and CovISQ attain essentially perfect accuracy (values very close to 1.00), and CovIS achieves 100% accuracy for all $\delta_1 \neq 0.5$. Even in the knife-edge symmetric case $\delta_1 = 0.5$, these measures remain informative: CovIS correctly signals approximate equality in 90% of simulations, while PILS and CovISQ fall into the $[0.45, 0.55]$ band in about 74% of samples.

By contrast, the correlation–sensitive measures that combine δ_i with reduced-form covariance information—Hasbrouck's IS and MIS, PDS, and the leadership ratios ILS and MILS—are markedly less reliable. For moderate differences in adjustment speed (e.g. $\delta_1 = 0.7$ or $\delta_1 = 0.6$), IS,

Table 3.2: Accuracy of Price Discovery Measures from Partial Price Adjustment Model

This table reports price discovery measure accuracy from the price data simulated from the following 2-market model:

$$p_{1t} = p_{1,t-1} + \delta_1 (m_t - p_{1,t-1}) + b_{0,1}^T \eta_t^T,$$
$$p_{2t} = p_{2,t-1} + \delta_2 (m_t - p_{2,t-1}) + b_{0,2}^T \eta_t^T$$

where $m_t = m_{t-1} + \eta_t^P, \boldsymbol{\eta_t} = \left(\eta_t^P, \eta_t^T\right)'$ are Gaussian white noise with diagonal covariance matrix $\text{diag}\left(\sigma_P^2, \sigma_T^2\right)$. The simulation parameterization is set as $\delta_2 = 1 - \delta_1, \sigma_P^2 = 1$, $\left(b_{0,1}^T, b_{0,2}^T\right) = (0.8, -0.2)$. We simulate 1000 samples of 21600 observations. For each sample, we take 5000 burn-in and save 10000 iterations.

Panel A: $\eta_t^P \sim i.i.d.$, $\sigma_P^2 = 1$, $\eta_t^T \sim i.i.d.$, $\sigma_T^2 = 10$										
δ_1	IS	CS	MIS	PDS	ILS	MILS	PILS	PIES	CovIS	CovISQ
0.9	1.00	0.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00	1.00
0.8	0.39	0.00	0.39	0.39	1.00	1.00	1.00	0.00	1.00	1.00
0.7	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.04	1.00	1.00
0.6	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.30	1.00	1.00
0.5	0.00	0.00	0.00	0.00	0.15	0.43	0.74	0.07	0.90	0.74
0.4	1.00	1.00	1.00	1.00	0.51	0.62	0.80	0.98	0.80	0.80
0.3	1.00	1.00	1.00	1.00	0.57	0.77	0.99	0.99	0.99	0.99
0.2	1.00	1.00	1.00	1.00	0.87	0.95	1.00	1.00	1.00	1.00
0.1	1.00	1.00	1.00	1.00	0.90	0.95	1.00	1.00	1.00	1.00
Panel B: $\eta_t^P \sim GARCH(1,1)$, $\eta_t^T \sim i.i.d.$, $\sigma_T^2 = 10$										
δ_1	IS	CS	MIS	PDS	ILS	MILS	PILS	PIES	CovIS	CovISQ
0.9	1.00	0.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00	1.00
0.8	0.25	0.00	0.25	0.25	1.00	1.00	1.00	0.01	1.00	1.00
0.7	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.01	1.00	1.00
0.6	0.00	0.00	0.00	0.00	0.84	0.82	0.78	0.25	0.78	0.78
0.5	0.00	0.00	0.00	0.00	0.18	0.59	0.93	0.12	0.99	0.93
0.4	1.00	1.00	1.00	1.00	0.52	0.69	0.97	0.99	0.97	0.97
0.3	1.00	1.00	1.00	1.00	0.97	0.99	1.00	1.00	1.00	1.00
0.2	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00
0.1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00
Panel C: $\eta_t^P \sim i.i.d.$, $\sigma_P^2 = 1$, $\eta_t^T \sim GARCH(1,1)$										
δ_1	IS	CS	MIS	PDS	ILS	MILS	PILS	PIES	CovIS	CovISQ
0.9	1.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.8	0.70	0.00	0.70	0.70	1.00	1.00	1.00	1.00	1.00	1.00
0.7	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00
0.6	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.92	0.87	0.98	0.92
0.4	1.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00
0.3	1.00	1.00	1.00	1.00	0.01	0.10	1.00	1.00	1.00	1.00
0.2	1.00	1.00	1.00	1.00	0.13	0.35	1.00	1.00	1.00	1.00
0.1	1.00	1.00	1.00	1.00	0.86	0.95	1.00	1.00	1.00	1.00

MIS, and PDS display zero accuracy, i.e. they systematically assign leadership to the wrong venue. ILS and MILS also exhibit substantial misclassification around the symmetric case ($\delta_1 = 0.5$), with accuracies of only 0.15 and 0.43, respectively. Only when the leadership gap becomes very pronounced ($\delta_1 \in \{0.9, 0.1\}$ or $\delta_1 \in \{0.8, 0.2\}$) do IS/MIS/PDS recover high accuracy. The Gonzalo–Granger CS share, which depends solely on $(b^T_{0,1}, b^T_{0,2})$, is completely misaligned in this design: it always assigns full leadership to Market 2, yielding 0% accuracy when $\delta_1 > 0.5$ and 100% accuracy when $\delta_1 < 0.5$. Overall, in the simplest homoskedastic environment, measures that are purely δ-based (PILS, CovIS, CovISQ) track the true structural leadership almost perfectly, whereas correlation–sensitive measures frequently misclassify the leader for empirically relevant ranges of δ_1.

Panel B introduces clustering in the permanent innovation via a GARCH(1,1) process, while transitory noise remains homoskedastic. The accuracy patterns confirm that the Σ–invariant measures remain robust to permanent volatility dynamics. PILS, CovIS, and CovISQ display almost the same performance as in Panel A: for $|\delta_1 - 0.5| \geq 0.2$, their accuracy is essentially 1.00, and even at $\delta_1 = 0.5$ CovIS is accurate in 99% of cases, with PILS and CovISQ around 0.93. Thus, heteroskedasticity in $\sigma^2_{P,t}$ has virtually no effect on the ability of these measures to rank the venues, consistent with their structural dependence on δ_i only.

In sharp contrast, IS, MIS, and PDS again perform poorly in the empirically relevant region with moderate leadership differences. For $\delta_1 \in \{0.7, 0.6\}$, their accuracy remains at 0.00, and even at $\delta_1 = 0.8$ they correctly classify the leader in only about 25% of simulations. The leadership ratios ILS and MILS improve somewhat relative to Panel A but still exhibit substantial underperformance near $\delta_1 = 0.5$, with accuracies of 0.18 for ILS and 0.59 for MILS. Thus, allowing for clustered permanent volatility amplifies the contrast between the structurally robust δ-based measures and the volatility-sensitive IS-type measures: the former are essentially invariant to the GARCH regime, while the latter remain fragile.

Panel C instead introduces GARCH(1,1) dynamics in the transitory microstructure noise while keeping the permanent shock homoskedastic. This design is particularly demanding for measures that rely on reduced-form correlations, because both the noise-to-signal ratio and the covariance of ε_t fluctuate substantially over time. Nevertheless, PILS, CovIS, and CovISQ again deliver near-perfect performance. For all $\delta_1 \neq 0.5$, their reported accuracies are 1.00 (or extremely close), and

at $\delta_1 = 0.5$ they still classify "approximate equality" correctly in the vast majority of replications (0.92 for PILS and CovISQ and 0.98 for CovIS). The robustness of PILS and CovISQ is not accidental: by construction they depend only on the adjustment speeds and are designed to be invariant to the variance–covariance matrix of $\boldsymbol{\eta}_t$.

As before, IS, MIS, and PDS struggle whenever leadership is not extreme. For $\delta_1 = 0.8$, they are correct in only about 70% of simulations, and for $\delta_1 = 0.7$ or 0.6 their accuracy again collapses to zero. ILS and MILS are particularly unstable under transitory GARCH: their accuracy is essentially zero for $\delta_1 \in \{0.6, 0.4\}$ and remains very low around the symmetric case. These patterns highlight that correlation–sensitive measures can be severely distorted by time-varying microstructure volatility, even when the underlying structural speeds of adjustment are fixed.

Taken together, Panels A–C of Table 3.2 deliver a clear message. When informational leadership is structurally determined by the adjustment speeds δ_i, measures that are explicitly built from these coefficients—PILS, CovIS, and CovISQ—recover the correct ranking of venues with very high frequency, virtually irrespective of whether volatility is homoskedastic, driven by GARCH in the permanent component, or driven by GARCH in the transitory component. Their only deviations from perfect accuracy occur in the near-symmetric case $\delta_1 = 0.5$, where one should not expect sharp identification of the leader in finite samples.

In contrast, measures that combine δ_i with reduced-form covariances (IS, MIS, PDS, ILS, MILS) and the CS measure based solely on $\boldsymbol{b}_0^T$ can be systematically misleading, especially when the leader's advantage is moderate rather than extreme. From the perspective of structural interpretation, these findings underscore the advantages of PILS, CovISQ, and CovIS as measures of price discovery in environments consistent with the partial adjustment model: they are tightly linked to the economically meaningful parameter δ_i, robust to changes in the volatility and correlation structure of the shocks, and therefore provide much more reliable guidance on which market leads in the incorporation of permanent information.

3.5.4 Time-varying δ

Next, we let the initial responses to a one–unit permanent shock, $\delta_{i,t}$, evolve over time. We keep the markets' responses to the transitory shock fixed at $\left(b_{0,1}^T, b_{0,2}^T\right) = (0.8, -0.2)$ and impose $\delta_{2,t} = 1 - \delta_{1,t}$. We simulate 1,000 samples of 21,600 observations from the partial adjustment

model, estimate the Bayesian order–invariant VAR with stochastic volatility for each sample, and compute posterior draws of all price discovery measures. For each horizon t we work with posterior means of the measures, and we then average across simulations.

To isolate the role of time variation in $\delta_{i,t}$, we maintain homoskedastic structural shocks, $\eta_t^P \sim \mathcal{N}(0, \sigma_P^2)$ and $\eta_t^T \sim \mathcal{N}(\mathbf{0}, \sigma_T^2 I_2)$ with $\sigma_P^2 = 1$ and $\sigma_T^2 = 10$. We consider two specifications for the adjustment speeds. In the first, $\delta_{1,t}$ follows a mean–reverting AR(1) process around a baseline value $\bar{\delta}_1$, and $\delta_{2,t} = 1 - \delta_{1,t}$. In the second, which is our main focus in this subsection, $\delta_{1,t}$ is generated as a random walk on the logit scale and then smoothed with an EWMA filter (the "RW–smooth" design). This construction produces smooth, persistent leadership cycles in which $\delta_{1,t}$ and $1 - \delta_{1,t}$ drift between 0 and 1 while remaining well inside the unit interval.

Figures 3.1 and 3.2 plot the time dimension of the RW–smooth experiments for two simulations. In each panel, the dashed line shows the true path of $\delta_{1,t}$ and the solid line reports the estimated mean of the corresponding price discovery measure for market 1, with a light band indicating the 15–85% posterior quantile range.

Several robust patterns emerge. First, the structural measures that depend only on the adjustment speeds and on the permanent/transitory decomposition behave as theory would suggest. In both calibrations, PDS, PIL, PILS, CovISQ, PIES, and the structural leadership measures (MILS, PILS, SILS) track the smooth movements in $\delta_{1,t}$ very closely. Whenever the true $\delta_{1,t}$ rises above 0.5, these measures for market 1 move smoothly above 0.5, and when $\delta_{1,t}$ falls below 0.5 they decline symmetrically. Because the shocks are homoskedastic and uncorrelated, the posterior variability of these measures is modest, and their paths are smooth functions of time. In particular, CovISQ and PILS are numerically indistinguishable up to Monte Carlo noise, reflecting their common structural dependence on the squared long–run permanent impacts.

Second, the Gonzalo–Granger component share (CS) behaves in a way that may appear "odd" at first glance. In both figures the CS line is essentially flat over time, and its level does not coincide with the simple benchmark $(\mathrm{CS}_1, \mathrm{CS}_2) = (0.2, 0.8)$ that arises in the static partial–adjustment model with $(b_{0,1}^T, b_{0,2}^T) = (0.8, -0.2)$. This is not a failure of the Bayesian estimation but a structural property of CS in our empirical implementation. Within the OI–VAR–SV, CS is obtained from the long–run impact vector ψ, which is a function of the estimated cointegration and adjustment matrices $(\alpha, \beta, \Gamma(1))$ only and does not depend on either the time–varying covariance matrices Σ_t

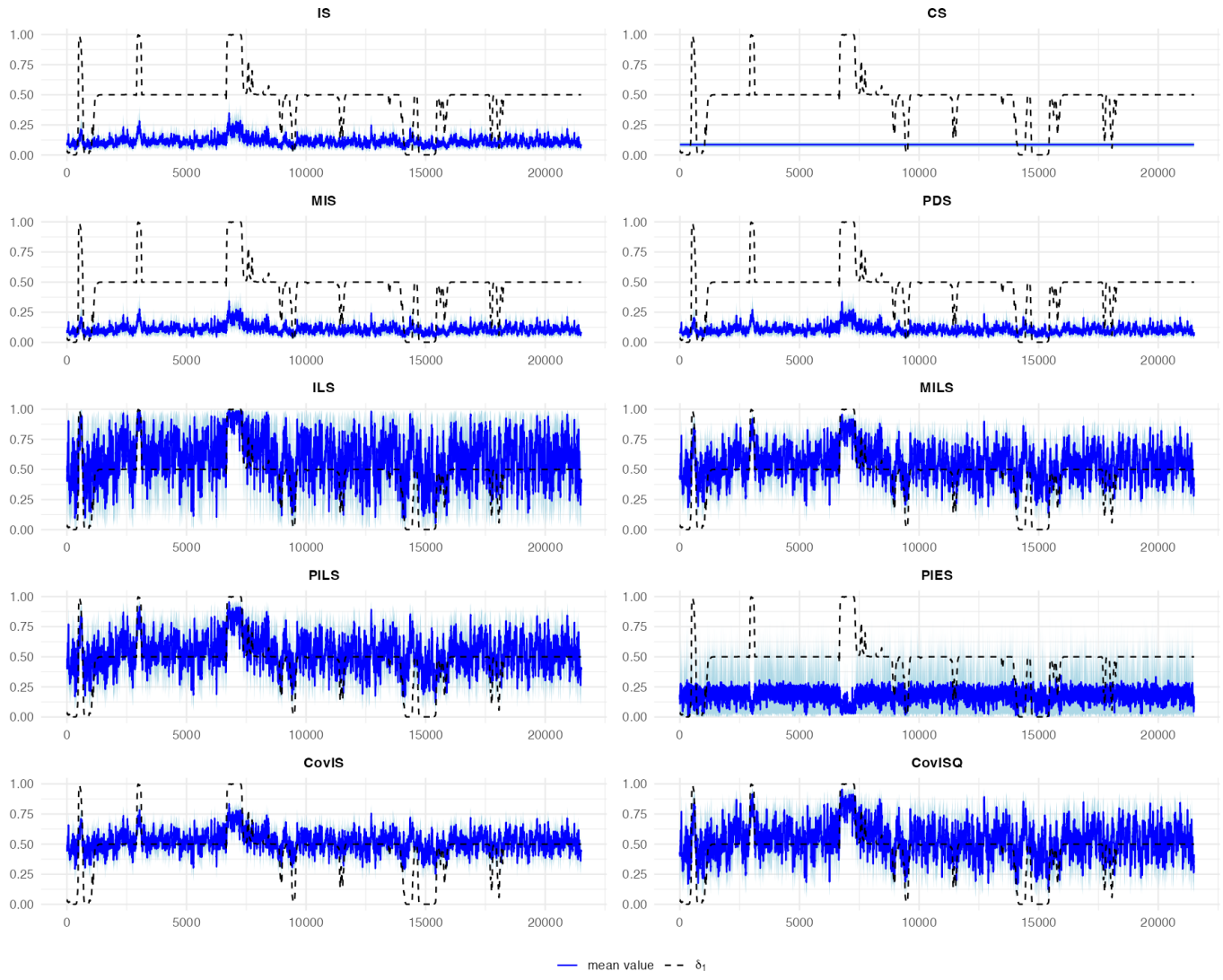

Figure 3.1: price discovery measures for smoothing RW δ

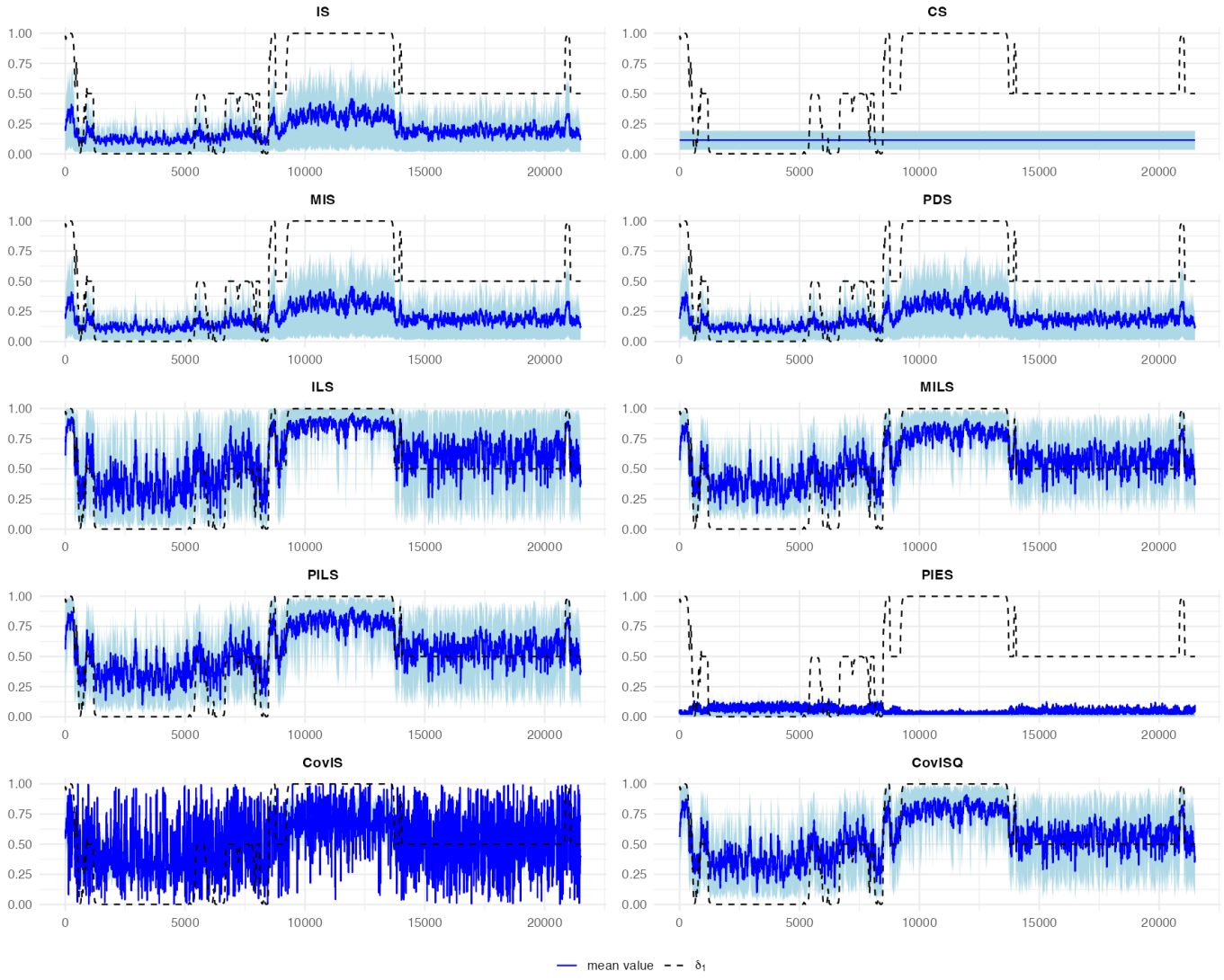

Figure 3.2: price discovery measures for smoothing RW δ

or the period–by–period adjustment speeds $\delta_{i,t}$. Because we fit a constant–coefficient VAR as a reduced–form approximation to a time–varying data–generating process, $(\alpha, \beta, \Gamma(1))$, and hence ψ and CS, are time invariant and represent a pseudo–true sample–average long–run structure rather than the closed–form expression from the static partial–adjustment model. The horizontal CS lines in Figures 3.1–3.2 therefore highlight a fundamental limitation of component shares in dynamic environments: CS summarizes how prices load on the permanent component in a time–invariant VECM, but it is intrinsically unable to track shifts in leadership driven by changes in the adjustment speeds $\delta_{i,t}$.

Third, the covariance–based measure CovIS is noticeably more erratic than its squared counterpart CovISQ and the other structural measures. By construction, CovIS is proportional to the signed long–run impact $D_{0P,i}$ normalized by the sum $\sum_j D_{0P,j}$. In finite samples, $D_{0P,i}$ is inferred from estimated VAR coefficients and the reduced–form covariance matrix, and it can occasionally become negative or very close to zero, especially when $\delta_{1,t}$ is near the boundary or when the posterior for $\delta_{1,t}$ places non–negligible mass near zero. When one $D_{0P,i}$ is negative while the other is positive, the normalization by $\sum_j D_{0P,j}$ amplifies small estimation errors and can generate sharp spikes or values outside $[0, 1]$. This is exactly what we observe in the bottom–left panels of Figures 3.1–3.2, where CovIS for market 1 exhibits large excursions even though the underlying $\delta_{1,t}$ path is very smooth. Squaring $D_{0P,i}$ in CovISQ eliminates the sign, stabilizes the normalization, and yields trajectories that coincide with PILS and faithfully reflect the evolving leadership.

Correlation–sensitive measures such as IS, MIS and ILS also respond to the changing $\delta_{i,t}$, but their paths are visibly noisier than those of the purely structural measures. Even under homoskedastic structural shocks, the reduced–form errors inherit time variation from the evolving $\delta_{i,t}$, so small estimation errors in the reduced–form covariance matrix translate into fluctuations in IS and MIS. The problem is most acute when $\delta_{1,t}$ is close to 0.5, where leadership is only weakly identified. Overall, the information–leadership ratios (ILS, MILS, PILS) are relatively robust: because they rescale IS/MIS by CS, they damp some of the noise and still convey the correct direction of leadership most of the time.

In Table 3.3 we report the accuracy of each measure, defined as the fraction of times it correctly identifies the leading market, conditional on the contemporaneous value of $\delta_{1,t}$ being close to one of the grid points $\{0.9, 0.8, \ldots, 0.1\}$. When $\delta_{1,t} \approx 0.5$, we treat the outcome as "correct" if the

Table 3.3: Accuracy of Price Discovery Measures from Partial Price Adjustment Model

This table reports price discovery measure accuracy from the price data simulated from the following 2-market model:

$$p_{1t} = p_{1,t-1} + \delta_{1t}\left(m_t - p_{1,t-1}\right) + b_{0,1}^T \eta_t^T,$$
$$p_{2t} = p_{2,t-1} + \delta_{2t}\left(m_t - p_{2,t-1}\right) + b_{0,2}^T \eta_t^T$$

where $m_t = m_{t-1} + \eta_t^P, \boldsymbol{\eta_t} = \left(\eta_t^P, \eta_t^T\right)'$ are Gaussian white noise with diagonal covariance matrix $\text{diag}\left(\sigma_P^2, \sigma_T^2\right)$. $\delta_{2t} = 1 - \delta_{1t}$, $\left(b_{0,1}^T, b_{0,2}^T\right) = (0.8, -0.2)$. η_t^P *i.i.d.*, $\sigma_P^2 = 1$, η_t^T *i.i.d.*, $\sigma_T^2 = 10$.

Panel D: $\delta_t \sim smoothRW$

$\bar{\delta}_1$	IS	CS	MIS	PDS	ILS	MILS	PILS	PIES	CovIS	CovISQ
0.9	0.98	0.98	0.98	0.98	0.74	0.74	0.74	0.68	0.75	0.74
0.8	0.74	0.05	0.74	0.74	0.73	0.73	0.74	0.02	0.74	0.74
0.7	0.56	0.22	0.56	0.56	0.32	0.34	0.34	0.20	0.36	0.34
0.6	0.39	0.35	0.39	0.38	0.23	0.29	0.29	0.34	0.39	0.29
0.5	0.84	0.11	0.84	0.84	0.84	0.84	0.84	0.04	0.84	0.84
0.4	0.06	0.06	0.06	0.06	0.19	0.30	0.42	0.10	0.60	0.42
0.3	0.74	0.57	0.74	0.74	0.46	0.46	0.46	0.47	0.48	0.46
0.2	0.62	0.38	0.62	0.62	0.44	0.44	0.44	0.31	0.45	0.44
0.1	0.88	0.88	0.88	0.88	0.75	0.76	0.77	0.77	0.77	0.77
Mean	0.65	0.40	0.65	0.64	0.52	0.54	0.56	0.33	0.60	0.56

Panel E: $\delta_t \sim AR(1)$

$\bar{\delta}_1$	IS	CS	MIS	PDS	ILS	MILS	PILS	PIES	CovIS	CovISQ
0.9	0.34	0.13	0.34	0.34	0.84	0.84	0.83	0.13	0.84	0.83
0.8	0.23	0.20	0.23	0.23	0.75	0.75	0.75	0.20	0.75	0.75
0.7	0.28	0.27	0.28	0.28	0.66	0.66	0.65	0.28	0.66	0.65
0.6	0.37	0.37	0.37	0.37	0.57	0.57	0.56	0.37	0.57	0.56
0.5	0.47	0.47	0.47	0.47	0.51	0.51	0.51	0.47	0.53	0.51
0.4	0.58	0.58	0.58	0.58	0.51	0.53	0.53	0.56	0.55	0.53
0.3	0.68	0.68	0.68	0.68	0.64	0.65	0.65	0.64	0.66	0.65
0.2	0.76	0.76	0.76	0.76	0.73	0.76	0.76	0.76	0.76	0.76
0.1	0.83	0.83	0.83	0.83	0.82	0.83	0.83	0.83	0.83	0.83
Mean	0.50	0.48	0.50	0.50	0.67	0.68	0.67	0.47	0.68	0.67

Notes: The first column labeled $\bar{\delta}_1$ represents a baseline parameter used to generate a time-varying adjustment path $\{\delta_{1t}\}$. For each value of $\bar{\delta}_1$, we simulate $\{\delta_{1t}\}$ either from a smoothed random walk (Panel D) or an AR(1) process (Panel E), rescale together with δ_{2t} so that $\delta_{1t} + \delta_{2t} \approx 1$ at each t, and report accuracy measures averaged over the resulting time-varying paths and Monte Carlo replications. Hence the δ_1 entries should be interpreted as baseline price discovery shares, not constant values of δ_{1t}.

measure lies between 0.45 and 0.55.

Panel D corresponds to the smooth random–walk design. Several features emerge clearly. First, no measure achieves the near–perfect accuracy seen in the time–invariant δ case. This is natural: with smoothly evolving $\delta_{1,t}$, the process spends a nontrivial fraction of time in "ambiguous" regions where the leadership advantage is modest, making correct classification inherently harder in finite samples. Even so, some measures perform systematically better than others.

For extreme leadership ($\delta_1 \approx 0.9$ or 0.1), the main measures (IS, MIS, PDS, PILS, CovIS, CovISQ) all attain reasonably high accuracy (roughly 0.74–0.98 in the first row and 0.75–0.88 in the last row). In these regions the signal is very strong, and most measures, including the correlation–sensitive ones, correctly attribute leadership a large fraction of the time. By contrast, the CS and PIES measures are notably weaker away from the boundaries: at $\delta_1 = 0.8$ and 0.7 their accuracies fall to 0.05 and 0.22 (for CS) and to 0.02 and 0.20 (for PIES), indicating that they often point to the wrong venue or fail to reflect the underlying leadership cycles.

A more stringent test is provided by intermediate values of δ_1. Around $\delta_1 = 0.6$–0.4, the accuracies of IS/MIS/PDS drop markedly (e.g. IS accuracy of about 0.39 at $\delta_1 = 0.6$ and 0.06 at $\delta_1 = 0.4$), indicating substantial misclassification when the leadership gap is moderate. Here the δ–based measures, and in particular CovIS and CovISQ, retain a comparatively stronger signal: at $\delta_1 = 0.4$ CovIS attains an accuracy of 0.60, substantially above IS/MIS/PDS (0.06) and above PIES (0.10), while PILS and CovISQ still achieve accuracy around 0.42. Thus, although no measure is perfect in the presence of smooth leadership cycles, CovIS and CovISQ tend to provide more reliable guidance than CS and PIES and often dominate IS/MIS/PDS in the "hard" region where leadership is present but not extreme.

The symmetric case $\delta_1 \approx 0.5$ is particularly revealing. Here we ask whether the measures correctly return a value near 1/2. Most of the structurally δ–based shares—PILS, CovIS, CovISQ, as well as the leadership ratios ILS and MILS—achieve an accuracy of 0.84, indicating that in roughly 84% of the simulation time they correctly reflect near–equality of contributions. CS and PIES, in contrast, achieve accuracies of only 0.11 and 0.04 respectively, underscoring their lack of symmetry under this design. Overall, in the smooth random–walk case, PILS/CovIS/CovISQ are not uniformly dominant, but they deliver comparatively stable and economically sensible signals, especially in regions where correlation–based measures and CS/PIES struggle.

Panel E reports results for the AR(1) specification of δ_t. In this case, the adjustment speeds fluctuate around a mean rather than drifting smoothly over the entire unit interval. The resulting leadership cycles are less extreme but still persistent. Here the contrast between different classes of measures is sharper. For large leadership advantages (e.g. $\delta_1 = 0.9$ or 0.8), the accuracies of IS, MIS, and PDS remain quite low (around 0.23–0.34), whereas PILS, CovIS, CovISQ and the leadership ratios ILS and MILS attain accuracies in the 0.75–0.84 range. Thus, when the structural leader is clearly dominant, the purely δ–based shares and their leadership counterparts capture this dominance much more reliably than the Hasbrouck–type measures.

As δ_1 moves towards 0.5, the same pattern persists. For $\delta_1 = 0.7$ or 0.6, IS/MIS/PDS show accuracies around 0.28–0.37, whereas PILS, CovIS, and CovISQ yield accuracies around 0.56–0.65, and ILS/MILS are close behind. At the symmetric point $\delta_1 = 0.5$, all δ–based measures and leadership ratios have accuracy near 0.5 (e.g. PILS and CovISQ at 0.51, CovIS at 0.53), which is the natural benchmark given that the true leadership is perfectly balanced and small sample noise is unavoidable. IS/MIS/PDS also reach about 0.47 in this special case, but they lag behind PILS/CovIS/CovISQ across most of the off–symmetric grid.

For values $\delta_1 < 0.5$, where Market 2 is the structural leader, the pattern continues to favour the δ–based shares. At $\delta_1 = 0.3$, IS/MIS/PDS accuracy is about 0.68, while PILS, CovIS, and CovISQ are in the 0.65–0.66 range; at $\delta_1 = 0.2$ and 0.1, all of these measures converge to high accuracy (around 0.76–0.83). In other words, when leadership differences become extreme, the information in δ_t is so strong that even the correlation–sensitive measures perform well, but for moderate and realistic gaps the outperformance of PILS, CovIS, and CovISQ is substantial.

In sum, the time–varying δ experiments highlight the dynamic robustness of the δ–based price discovery measures. In both the smooth random–walk and AR(1) designs, PILS, CovIS, and CovISQ track the evolving leadership cycles in a stable and interpretable way. In the more realistic AR(1) case, they outperform IS, MIS, and PDS by a wide margin whenever leadership is moderate rather than extreme, and they dominate CS and PIES across virtually all configurations. Even in the more challenging smooth random–walk design, they remain competitive with IS/MIS/PDS and clearly superior to CS and PIES in ambiguous regions and around symmetry. Together with the time–invariant results, these findings suggest that PILS, CovIS, and CovISQ provide structurally meaningful and empirically reliable summaries of price discovery when the underlying adjustment

speeds evolve over time.

3.5.5 Discussion of Equal Price Discovery Share

We examine the symmetric benchmark $\delta_1 = \delta_2 = 0.5$ with time–invariant adjustment speeds. In the partial–adjustment model this configuration implies that the two markets contribute equally to the permanent price at every point in time, so any sensible price discovery measure should be tightly centered at 0.5. The three panels in Figures 3.3–3.5 correspond to (i) homoskedastic permanent and transitory shocks, (ii) GARCH(1,1) volatility in the permanent shock only, and (iii) GARCH(1,1) volatility in the transitory shock only. In all cases the true leadership is exactly balanced.

The structural measures that depend only on the permanent component and the adjustment speeds behave as expected. PILS, CovIS and CovISQ are essentially flat and hover very close to 0.5 across all three volatility scenarios, with only mild sampling variability. This reflects the fact that, in population, these measures collapse to $\text{PILS}_1 = \text{CovIS}_1 = \text{CovISQ}_1 = 0.5$ when $\delta_1 = \delta_2$, and that the Bayesian OI-VAR-SV estimator is able to recover this symmetry with high precision even in the presence of stochastic volatility. The same is true, to a slightly lesser extent, for the leadership ratios that are constructed from these shares: when $\delta_1 = \delta_2$ the underlying structural leadership is one, so the corresponding leadership measures fluctuate around their symmetric benchmark rather than indicating a persistent leader.

By contrast, correlation–sensitive measures display systematic distortions away from 0.5. In all three figures, Hasbrouck's IS and MIS settle around values far below 0.5 (roughly 0.1–0.2 in the first two panels and closer to 0.25 when transitory volatility is heteroskedastic), despite the fact that the two markets are structurally identical in their response to the permanent shock. PDS shows a similar downward bias. The Gonzalo–Granger component share (CS) is nearly constant but also far from 0.5, reflecting the fact that it is pinned down by the estimated long–run cointegration structure rather than by the true symmetric adjustment speeds. Moreover, the information leadership measures ILS and MILS indicate strong and persistent "leadership" for one venue, with values above 0.7–0.8, even though no such leadership exists in the data-generating process. These spurious leadership signals arise because IS and MIS are functions of the reduced-form covariance matrix; small estimation errors and volatility dynamics are propagated into the

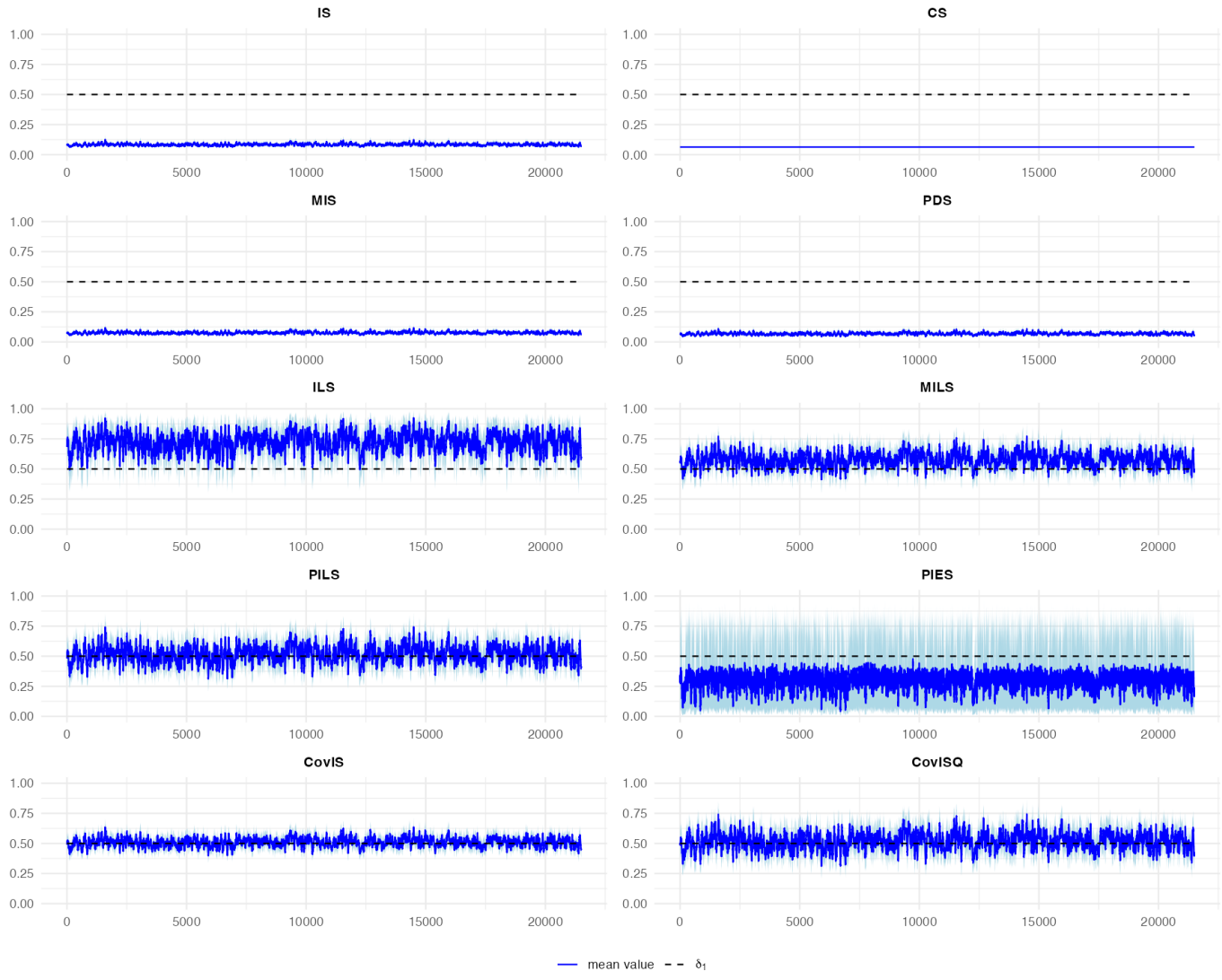

Figure 3.3: price discovery measures of Panel A in Table 3.2. $\delta = (0.5, 0.5)$.

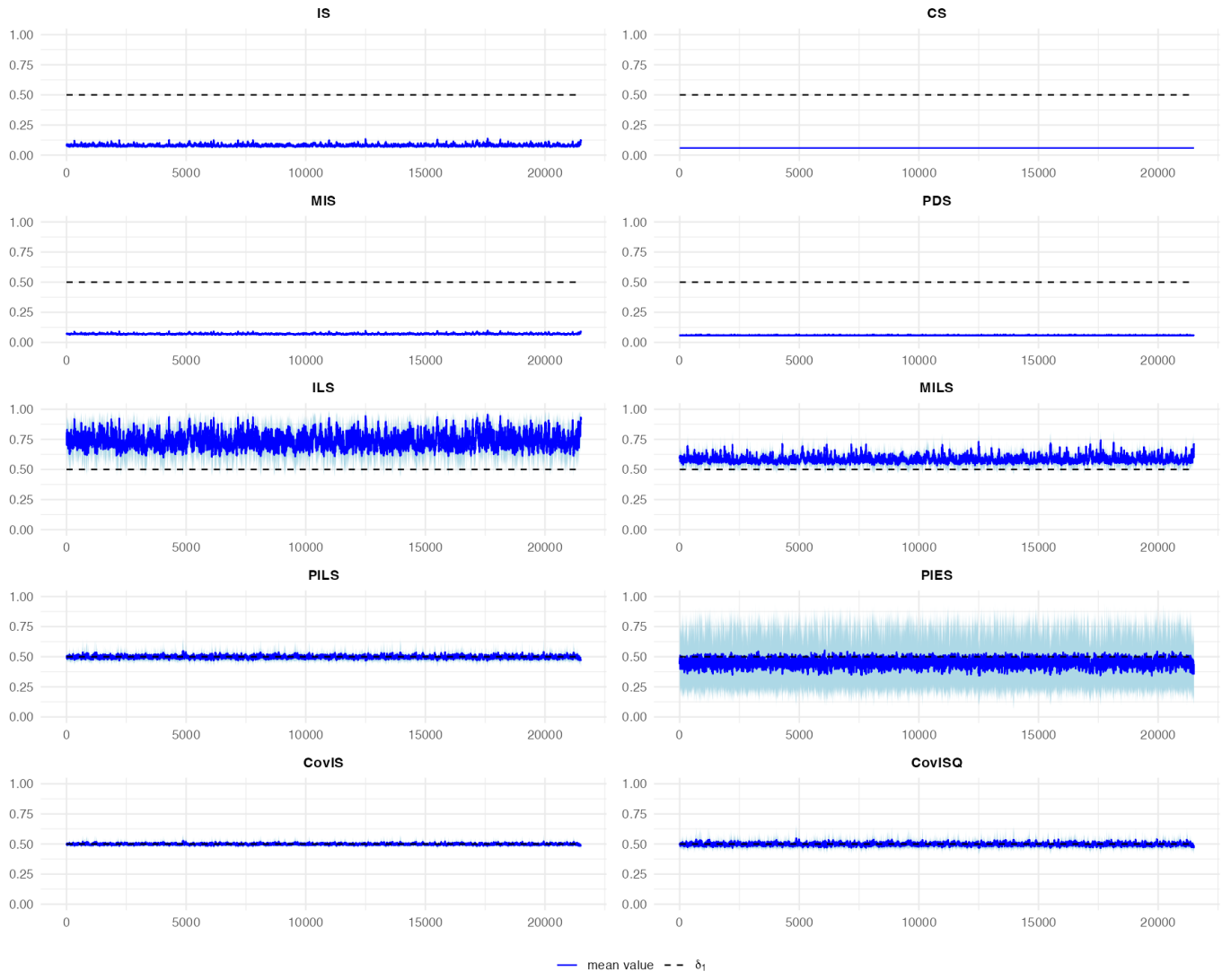

Figure 3.4: price discovery measures of Panel B in Table 3.2. $\delta = (0.5, 0.5)$.

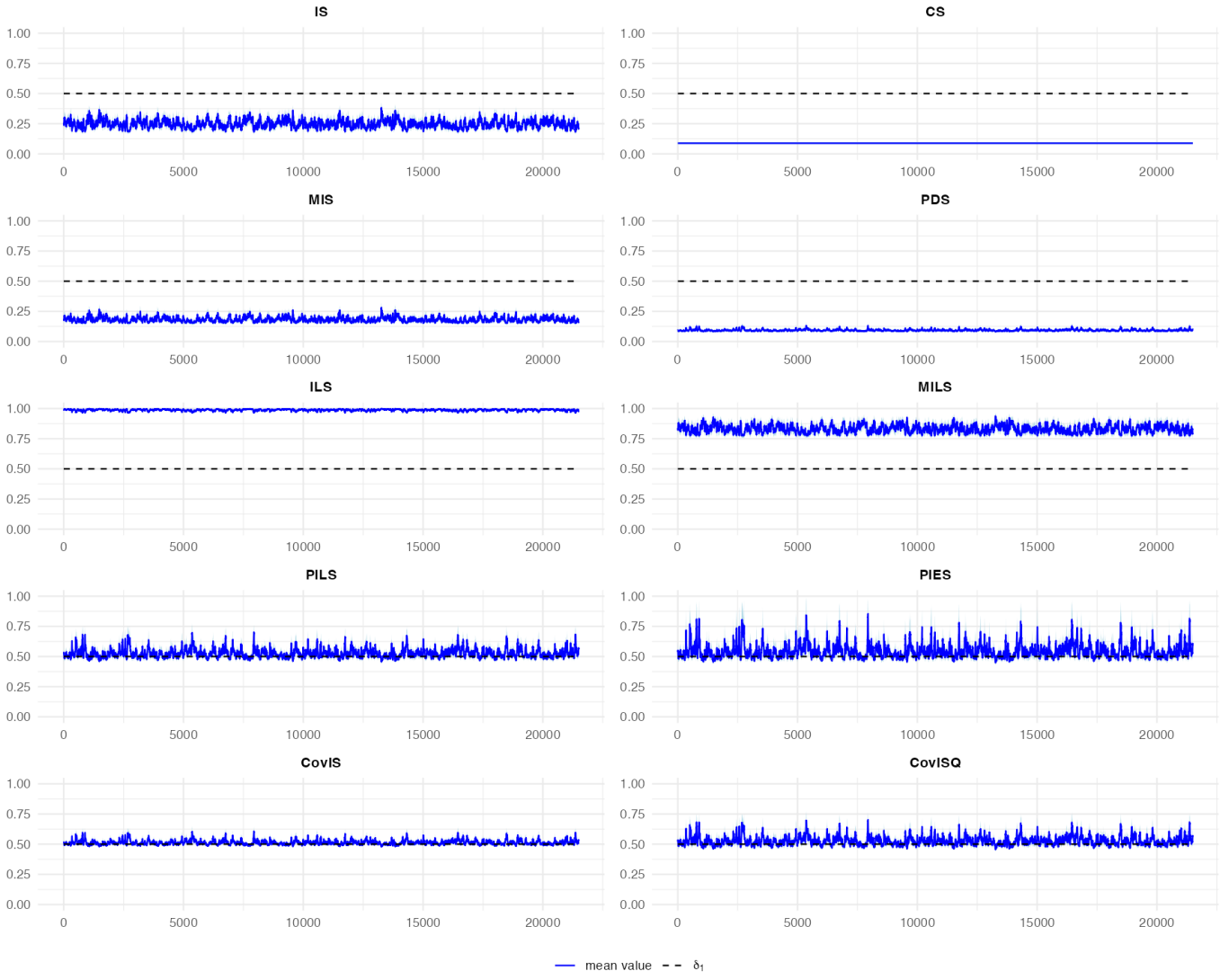

Figure 3.5: price discovery measures of Panel C in Table 3.2. $\delta = (0.5, 0.5)$.

ratio that defines ILS and MILS, creating the illusion of dominance when the true contributions are identical.

Overall, the symmetric design underscores the main message of our simulation study: measures that depend only on the permanent component (PILS, CovIS, CovISQ) recover the correct equal–share benchmark even under substantial heteroskedasticity, while correlation–based measures and component shares can generate large and persistent deviations from 0.5 and spurious market leadership.

3.6 Conclusion

This paper develops a structural, order-invariant Bayesian framework for analyzing the dynamics of time-varying price discovery in fragmented markets. Starting from a cointegration-based representation of related prices, we define price discovery in terms of each venue's contribution to the permanent component of the common stochastic trend, explicitly separating information-bearing permanent shocks from transitory microstructure noise. Embedding this decomposition in an order-invariant VAR with stochastic volatility allows us to model rich time variation in the volatility and covariance of shocks, while avoiding the arbitrary ordering assumptions that plague traditional Cholesky-based approaches. The framework yields a family of time-varying price discovery measures and structural impulse responses, together with full Bayesian uncertainty quantification.

Our simulation study, based on a stylized partial-adjustment microstructure model, highlights sharp differences in the performance of competing measures. When informational leadership is governed by markets' adjustment speeds to permanent shocks, measures that depend only on this structural channel—most notably PILS, CovIS, and its quadratic variant CovISQ—recover the true ranking of venues with very high accuracy across a wide range of designs. This robustness holds when shocks are homoskedastic, when the permanent innovation follows a GARCH process, and when transitory microstructure noise exhibits strong heteroskedasticity. Even in challenging scenarios with smoothly time-varying adjustment speeds, these structural measures track the evolution of leadership in a stable and interpretable way.

In contrast, measures that combine adjustment speeds with reduced-form covariance informa-

tion—IS, MIS, PDS, and the associated information-leadership ratios ILS and MILS—are much more fragile. In empirically relevant regions where leadership is present but not extreme, these measures often misclassify the leading venue, and they can produce highly distorted signals when the noise-to-signal ratio or reduced-form correlation varies over time. The Gonzalo–Granger component share (CS) is essentially time-invariant in our dynamic environment and reflects a pseudo-true average long-run structure rather than the underlying time-varying leadership. The PIES measure, which emphasizes instantaneous pricing errors, can also behave erratically when adjustment speeds wander near the boundaries.

The symmetric benchmark with equal adjustment speeds further underscores these differences. When both markets contribute equally to the permanent price, structurally grounded measures such as PILS, CovIS, and CovISQ remain tightly centered around one-half, even in the presence of substantial stochastic volatility. By contrast, IS, MIS, PDS, CS, ILS, and MILS can deviate persistently from 0.5 and suggest spurious leadership in markets that are, by construction, informationally identical. This experiment clarifies that sensitivity to reduced-form covariance and identification schemes can generate misleading conclusions about informational dominance, particularly in long samples spanning multiple volatility regimes.

Taken together, our results argue strongly in favor of structurally interpretable, order-invariant measures of price discovery that depend only on permanent shocks and adjustment speeds. Within the Bayesian OI-VAR-SV framework, PILS, CovIS, and CovISQ provide such measures: they deliver coherent rankings of venues, are robust to changes in volatility and noise, and naturally accommodate full posterior inference on the dynamics of price discovery shares and structural impulse responses. For empirical applications—such as equity and ETF markets, spot–futures pairs, or crypto-asset venues—these tools offer a disciplined way to study how informational leadership evolves over time, how it responds to changes in volatility, liquidity, and regulation, and how uncertainty around leadership should be quantified.

Several avenues for future research remain. On the methodological front, allowing for time-varying short-run coefficients and cointegration vectors within the order-invariant Bayesian framework would broaden the set of environments in which dynamic price discovery can be studied. Extending the analysis to higher-dimensional systems with more than two venues, and integrating microstructure covariates such as spreads, depth, or order-flow imbalances into the dynamics of

stochastic volatility and adjustment speeds, are natural next steps. On the empirical side, applying the proposed framework to high-frequency data for equities, ETFs, options, and crypto-assets would shed light on how technological change, market design, and regulatory interventions reshape informational leadership over time.

Appendix A

Supplementary Material for Chapter 1

A.1 Kalman Filter and Smoother Algorithm

This section provides details for MCMC algorithm in Primiceri (2005).

The first T_0 observations are used to calibrate the prior distributions over the initial states and the innovation variances for the time varying coefficients β_t, simultaneous relations α_t, and the standard errors σ_t:

$$p\big(\beta_0, \alpha_0, \sigma_0, Q, S, W\big)$$

To match the independence of their innovations, we also assume the prior for the time-varying parameter states are independent. This implies the prior can be written as

$$p\big(\beta_0, \alpha_0, \sigma_0, Q, S, W\big) = p(\beta_0, Q)\; p(\alpha_0, S)\; p(\sigma_0, W)$$

We further assume that

$$p(\beta_0, Q) = p(\beta_0|Q)\; p(Q)$$

$$p(\alpha_0, S) = p(\alpha_0|S)\; p(S)$$

$$p(\log \sigma_0, W) = p(\log \sigma_0|W)\; p(W)$$

The priors take the forms:

$$\alpha_0 \sim \mathcal{N}(a_\alpha, V_\alpha) \tag{A.1}$$

$$\log \sigma_0 \sim \mathcal{N}(a_{\log\sigma}, V_{\log\sigma}) \tag{A.2}$$

$$Q \sim \mathcal{IW}\left(k_Q^2 \cdot d_Q \cdot V_\beta, d_Q\right) \tag{A.3}$$

$$W \sim \mathcal{IW}\left(k_W^2 \cdot d_W \cdot I_n, d_W\right) \tag{A.4}$$

$$S_j \sim \mathcal{IW}\left(k_S^2 \cdot d_{S,j} \cdot V_\alpha, d_{S,j}\right) \tag{A.5}$$

Specifically, $d_W = n+1$, $\Sigma_j(d_{S,j}-1) = d_W - 1$, $d_{S,j} = \text{dimension}(S_j)+1$, S_j is the j^{th} block of S, whose dimension is the number of free elements in each row of A. Default d_Q is the presample size, since "a slightly tighter prior could avoid implausible behaviours of the time varying coefficients" (Primiceri, 2005).

The tuning parameters are $\kappa_Q, \kappa_S, \kappa_W$, which determine the scales of $\mathcal{IW}$ distribution. We get to constant parameter BVAR by setting $\kappa_Q, \kappa_S, \kappa_W$ all as zero. Primiceri (2005) uses $k_Q = 0.01, k_S = 0.1, k_W = 0.01$.

We can construct the full posterior for endogenous variables and time-varying coefficient states from the priors and the joint specification. Importantly, because of the model structure and the use of conjugate priors, sampling from this posterior is easily done in pieces given the data. We follow the corrected algorithm in Del Negro and Primiceri (2015):

0. Initialize $\alpha_{0:T}, \sigma_{0:T}, Q, S, W$ based on priors.

1. Sample $\beta_{0:T} \sim p\left(\beta_{0:T} \mid \alpha_{0:T}, \sigma_{0:T}, Q, S, W, y_{1:T}\right)$.

   Conditional on the data $y_{1:T}$, the volatility states $\alpha_{0:T}$ and $\sigma_{0:T}$, and the innovation variance Q, the model for y_t and the dynamics of β_t imply, respectively, a measurement equation and state transition equation, reproduced below:

$$\begin{aligned} y_t &= X_t'\beta_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}\big(0, A_t^{-1}\,\text{diag}(\sigma_t)\,\text{diag}(\sigma_t)'\,[A_t^{-1}]'\big) \\ \beta_t &= \beta_{t-1} + v_t, \qquad v_t \sim \mathcal{N}(0, Q) \end{aligned}$$

   Together with the Gaussian prior over β_0, this fully describes a Gaussian state space model

for $\beta_{0:T}$, and we can draw $\beta_{0:T}$ efficiently using the simulation smoother.

2. Sample $\alpha_{1:T} \sim p\left(\alpha_{1:T} \mid \beta_{1:T}, \sigma_{1:T}, Q, S, W, y_{1:T}\right)$.

   We will have the following state transition equation:

$$\alpha_t = \alpha_{t-1} + \zeta_t, \qquad \zeta_t \sim \mathcal{N}(0, S)$$

   For example, in the case of three variables

$$\begin{bmatrix} \begin{pmatrix} \alpha_{21,t} \\ \alpha_{31,t} \end{pmatrix} \\ \begin{pmatrix} \alpha_{32,t} \end{pmatrix} \end{bmatrix} = \begin{bmatrix} \begin{pmatrix} \alpha_{21,t} \\ \alpha_{31,t} \end{pmatrix} \\ \begin{pmatrix} \alpha_{32,t} \end{pmatrix} \end{bmatrix}$$

   The VAR model can equivalently be written to make A_t and σ_t explicit:

$$y_t = \underbrace{X_t' \beta_t}_{\widehat{y}_t} + A_t^{-1}\, \tilde{u}_t, \qquad \tilde{u}_t \sim \mathcal{N}(0, \mathrm{diag}(\sigma_t)\mathrm{diag}(\sigma_t)')$$

   To transform this into a measurement equation in A_t (or α_t), premultiply by A_t and rearrange

$$\tilde{u}_t = A_t \underbrace{\left(y_t - \widehat{y}_t\right)}_{\widetilde{y}_t} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \alpha_{21,t} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \alpha_{n1,t} & \cdots & \alpha_{nn-1,t} & 1 \end{bmatrix} \widetilde{y}_t = \begin{bmatrix} \widetilde{y}_{t,1} \\ \alpha_{21,t}\, \widetilde{y}_{t,1} + \widetilde{y}_{t,2} \\ \alpha_{31,t}\, \widetilde{y}_{t,1} + \alpha_{32,t}\, \widetilde{y}_{t,2} + \widetilde{y}_{t,3} \\ \vdots \end{bmatrix}$$

   Given that α_t is a stacking of A_t by rows, can rewrite again as

$$\widetilde{y}_t = Z_t \alpha_t + \tilde{u}_t, \qquad \tilde{u}_t \sim \mathcal{N}(0, \mathrm{diag}(\sigma_t)\mathrm{diag}(\sigma_t)')$$

where

$$\underset{(n\times\frac{n(n-1)}{2})}{Z_t} = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ -\widetilde{y}_{1,t} & 0 & \cdots & 0 \\ 0 & -\widetilde{y}_{[1,2],t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\widetilde{y}_{[1,\ldots,n-1],t} \end{bmatrix},$$

$$\widetilde{y}_{[1,\ldots,i],t} = \left[\widetilde{y}_{1,t}, \widetilde{y}_{2,t}, \ldots, \widetilde{y}_{i,t}\right].$$

Notice that Z_t has block structure.

3. Sample $\sigma_{1:T} \sim p\left(\sigma_{1:T} \mid \beta_{1:T}, \alpha_{1:T}, Q, S, W, y_{1:T}\right)$.

   The VAR model can equivalently be written to make A_t and σ_t explicit:

$$y_t = \underbrace{X_t'\beta_t}_{\widehat{y}_t} + A_t^{-1}\,\mathrm{diag}(\sigma_t)\,\varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, I_n)$$

   To transform this into a measurement equation in σ_t, again start by premultiplying by A_t and rearranging

$$y_t^* := A_t(y_t - \widehat{y}_t) = \mathrm{diag}(\sigma_t)\,\varepsilon_t$$

   This is a nonlinear measurement equation (because σ_t and ε_t are both random in the system), but can turn it into a linear measurement equation by squaring and taking logs, element-wise:

$$\log\left[\left(y_{t,i}^*\right)^2\right] = \log\left[\left(\left[\mathrm{diag}(\sigma_t)\,\varepsilon_t\right]_i\right)^2\right] = \log\left(\sigma_{t,i}^2\right) + \log\left(\varepsilon_{t,i}^2\right)$$

   Proceeding and joining this with the assumed dynamics for the log standard deviations, we get to the following measurement and state transition equations:

$$y_t^{**} = 2h_t + e_t$$
$$h_t = h_{t-1} + \eta_t$$

   where $y_{i,t}^{**} = \log\left[\left(y_{i,t}^*\right)^2 + \bar{c}\right]$; $\bar{c}$ is the offset constant (set to 0.001); $e_{i,t} = \log\left(w_{i,t}^2\right)$; $h_{i,t} =$

$\log \sigma_{i,t}$, the e's an the η 's are not correlated.

The above system, while linear, is non-Gaussian because of $e_{i,t}$, which is distributed as $\chi^2(1)$. Therefore, approximate this system with a linear Gaussian system using a mixture of normals approximation (Kim et al., 1998). The result is a linear Gaussian state space system, which we can use to draw $h_{0:T}$ $(\log \sigma_{0:T})$.

4. Sample $s_{1:T} \sim p\left(s_{1:T} \mid \alpha_{1:T}, \sigma_{1:T}, Q, S, W, y_{1:T}\right)$.

   $s^T = [s_1, \ldots, s_T]'$ is the matrix of indicator variables selecting at every point in time which member of the mixture of the normal approximation has to be used for each element of e.

   Conditional on $y^{**1:T}$ and the new $h_{1:T}$, it is possible to sample the new $s_{1:T}$ matrix, to be used in the next iteration. As in Kim et al. (1998), this is easily done by independently sampling each $s_{i,t}$ from the discrete density defined by

$$\Pr\left(s_{i,t} = j \mid y^{**}_{i,t}, h_{i,t}\right) \propto q_j f_N\left(y^{**}_{i,t} \mid 2h_{i,t} + m_j - 1 \cdot 2704, v_j^2\right), \quad j = 1, \ldots, 7, \quad i = 1, \ldots, n.$$

5. Lastly, sample Q, W and S from appropriate inverse Wishart distributions. The posterior is inverse Wishart because of conjugacy.

6. Return to 1.

A.2 Equation-by-Equation Algorithm

One can simulate from the joint posterior distribution using the following posterior sampler.

1. $p\left(\boldsymbol{\beta}_i, \boldsymbol{\alpha}_i \lfloor \boldsymbol{y}, \boldsymbol{\sigma}, \boldsymbol{Q}, \boldsymbol{S}, \boldsymbol{W}, \boldsymbol{\kappa}\right), i = 1, \ldots, n$;

2. $p\left(\boldsymbol{\sigma}_i \mid \boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{Q}, \boldsymbol{S}, \boldsymbol{W}, \boldsymbol{\kappa}\right), i = 1, \ldots, n$;

3. $p\left(\boldsymbol{Q}_i, \boldsymbol{S}_i, W_i \mid \boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\sigma}, \boldsymbol{\kappa}\right), i = 1, \ldots, n$;

4. $p\left(\boldsymbol{\beta}_{i,0}, \boldsymbol{\alpha}_{i,0} \mid \boldsymbol{y}, \boldsymbol{\sigma}, \boldsymbol{Q}, \boldsymbol{S}, \boldsymbol{W}, \boldsymbol{\kappa}\right), i = 1, \ldots, n$;

5. $p\left(\boldsymbol{\kappa} \mid \boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\sigma}, \boldsymbol{Q}, \boldsymbol{S}, \boldsymbol{W}\right)$.

Note that if we treat $\boldsymbol{\kappa} = (\lambda, \mu, \delta, \Psi)$ as predetermined values (density means above) instead of hyperparameters, then step 4 and 5 are not necessary.

Basically, the steps 1-3 aim to sample the same parameters as those in Appendix A, while the main difference is estimate the time-variant parameters at the same time, equation by equation, instead of filtering and smoothing. Here we show how we operate step 1, because this where we gain the huge computational benefit. For more details, please see Chan and Jeliazkov (2009) and Chan (2022).

The ith $(i = 1, \cdots, n)$ equation of the system can be rewritten as

$$y_{i,t} = \widetilde{\boldsymbol{x}}_t \boldsymbol{\beta}_{i,t} + \widetilde{\boldsymbol{w}}_{i,t} \boldsymbol{\alpha}_{i,t} + v_{i,t}, \quad v_{i,t} \sim \mathcal{N}(0, \sigma_{i,t}^2)$$

where $\dot{\boldsymbol{B}}_{i,j,t}$ represent the ith row of $\dot{\boldsymbol{B}}_{j,t}$, $\boldsymbol{\beta}_{i,t} = \left(\dot{b}_{i,t}, \dot{\boldsymbol{B}}_{i,1,t}, \ldots, \dot{\boldsymbol{B}}_{i,p,t}\right)'$ is the intercept and VAR coefficients of the ith equation, $\boldsymbol{\alpha}_{i,t} = \left(\alpha_{i1,t}, \cdots, \alpha_{i(i-1),t}\right)'$ denotes the free elements in the ith row of the contemporaneous impact matrix A_t, $\widetilde{\boldsymbol{x}}_t = \left(1, \boldsymbol{y}'_{t-1}, \ldots, \boldsymbol{y}'_{t-p}\right)$, and $\widetilde{\boldsymbol{w}}_{i,t} = (-y_{1,t}, \ldots, -y_{i-1,t})$.

Let $\boldsymbol{x}_{i,t} = (\widetilde{\boldsymbol{x}}_t, \widetilde{\boldsymbol{w}}_{i,t})$, $\boldsymbol{\theta}_{i,t} = \left(\boldsymbol{\beta}'_{i,t}, \boldsymbol{\alpha}'_{i,t}\right)'$, we can further simplify the ith equation as

$$y_{i,t} = \boldsymbol{x}_{i,t} \boldsymbol{\theta}_{i,t} + v_{i,t}, \quad v_{i,t} \sim \mathcal{N}\left(0, \sigma_{i,t}^2\right)$$

Next, stack the state equations (C.3)-(C.4) over $t = 1, \ldots, T$:

$$\boldsymbol{H}_{\theta_i} \boldsymbol{\theta}_i = \boldsymbol{\varepsilon}_i^{\theta}, \quad \boldsymbol{\varepsilon}_i^{\theta} \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{I}\right)$$

where $\boldsymbol{H}_{\theta_i}$ is the first difference matrix. Since $\boldsymbol{H}_{\theta_i}$ is a square matrix with unit determinant, it is invertible. It then follows that $\boldsymbol{\theta}_i \sim \mathcal{N}\left(\boldsymbol{0}, \left(\boldsymbol{H}'_{\theta_i} \boldsymbol{H}_{\theta_i}\right)^{-1}\right)$.

Finally, using standard linear regression results, we have

$$\left(\boldsymbol{\theta}_i \mid \boldsymbol{y}_i, \boldsymbol{\sigma}_i, \boldsymbol{R}_i, \boldsymbol{\theta}_{i,0}\right) \sim \mathcal{N}\left(\widehat{\boldsymbol{\theta}}_i, \boldsymbol{K}_{\boldsymbol{\theta}_i}^{-1}\right),$$

where

$$\boldsymbol{K}_{\boldsymbol{\theta}_i} = \boldsymbol{H}'_{\theta_i'} \boldsymbol{H}_{\theta_i} + \boldsymbol{x_i}' \boldsymbol{R_i}^{-1} \boldsymbol{x}_i,$$

$$\widehat{\boldsymbol{\theta}}_i = \boldsymbol{K}_{\boldsymbol{\theta}_i}^{-1} \left(\boldsymbol{x_i}' \boldsymbol{R_i}^{-1} \left(\boldsymbol{y}_i - \boldsymbol{x}_i \boldsymbol{\theta}_{i,0}\right)\right).$$

Since the precision matrix $\boldsymbol{K}_{\theta_i}$ is a band matrix, one sample $(\boldsymbol{\theta}_i \mid \boldsymbol{y}_i, \boldsymbol{\sigma}_i, \boldsymbol{R}_i, \boldsymbol{\theta}_{i,0})$ efficiently using the algorithm in Chan and Jeliazkov (2009).

The assumption of independence across equations covers most of the practical implementations of VARs. It could be relaxed while maintaining the same triangular structure. This would entail using the properties of the multivariate normal to derive a triangular factorization for the prior in which β_i appears only in equation $i, \cdots, n$. While this requires the inversion of large matrices, such inversions need to be performed only once, outside the main MCMC algorithm, and therefore with not much computational cost.

A.3 PIT tests

This paper follows the PIT test in Amisano and Geweke (2017).

For the stated distributions of $\{\pi_{jt}\}$ and $\{z_{jt}\}$ to be literally true, a model specification would have to be dogmatic and correct for θ_i. The usual criterion of correct specification of an econometric model is weaker: that for some value of θ_i, the distribution of $Y_{1:T}$ coincides with D. The actual size of test statistics for the properties of $\{\pi_{jt}\}$ or $\{z_{jt}\}$ is likely to be larger than the nominal size when the model is correctly specified up to unknown parameters. More relevant is the degree to which different models depart from the ideal of the PIT, and the particular ways in which this happens for different models; Geweke and Amisano (2010) illustrates this use of PIT tests.

The i.i.d. normal distribution of $\{z_{jt}\}$ under the hypothesis of correct specification is analytically more tractable than that of $\{\pi_{iT}\}$ and the tests here proceed from $\{z_{iT}\}$. These series are by-products of the computation of variance decompositions described in the previous section. Corresponding to each parameter vector $\theta^{(m)}_{t-1,i}$, compute

$$\pi^{(m)}_{jt,i} = \Phi\left[\left(y_{jt} - \mu^{(m)}_{j,t-1,i}\right) \cdot \left(v^{(m)}_{jj,t-1,i}\right)^{-1/2}\right] \times \left(v^{(m)}_{jj,t-1,i}\right)^{-1/2} (m = 1, \ldots, M)$$

and then $M^{-1}\sum_{m=1}^{M} \pi^{(m)}_{j,t-1,i} \approx \pi_{jt,i}$ and $\Phi\left(\pi_{jt,i}\right) \approx z_{jti}$. For each model A_i and each constituent time series j the ideal of correct model specification implies

$$z_{jti}(t = 1, \ldots, T) \overset{iid}{\sim} N(0, 1).$$

This hypothesis can be tested in a great many ways, each with its own power against alternatives. Here we use novel PIT tests, based on the distribution under the null hypothesis (5) of a set of raw moments, each element of the form $T^{-1}\sum z_{jt,i}^{q}$, where q is a particular positive integer unique to that element, and of a set of of cross products, each element of the form $(T-\ell)^{-1}\sum_{t=1}^{T-\ell} z_{jt,i}\cdot z_{j,t-\ell,i}$, where ℓ is a particular positive integer unique to that element. In particular, the following theorem can be used to establish the asymptotic distribution of the proposed statistics.

Theorem 1. Given $z_t, t = 1, 2, \ldots, T$, a sequence of i.i.d. draws from a N$(0, 1)$ distribution, then it holds that

$$T\left(\bar{m}_T - m_{q+L}\right)'\Omega^{-1}\left(\bar{m}_T - m_{q+L}\right) \xrightarrow{d} \chi^2_{Q+L}$$

In the paper we compute the tests (6) for the transformed PIT sequences of each of the seven series common to all models. In particular, we choose $q = 4$ and $L = 4$ for all series, i.e. we consider the first four moments and the first cross moments of the transformed PITs. In this particular case we have

$$\bar{m}_{i,T} = \begin{bmatrix} \bar{\mu}_{1,iT} \\ \bar{\mu}_{2,iT} \\ \bar{\mu}_{3,iT} \\ \bar{\mu}_{4,iT} \\ \bar{\rho}_{1,iT} \\ \bar{\rho}_{2,iT} \\ \bar{\rho}_{3,iT} \\ \bar{\rho}_{4,iT} \end{bmatrix}, m_{q+L} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 3 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \Omega = \begin{bmatrix} 1 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 12 & 0 & 0 & 0 & 0 \\ 3 & 0 & 15 & 0 & 0 & 0 & 0 & 0 \\ 0 & 12 & 0 & 96 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Under the hypothesis of correct model specification the exact distribution of the test statistics depends only on the sample size T, and it is easy to access this distribution by simulating $z_t \overset{iid}{\sim} N(0, 1)(t = 1, \ldots, T)$. The work here uses 10^5 simulations, which reliably establishes p-values of PIT test statistics in the first three decimal places. Moreover, except for very small p values, it turns out that the asymptotic approximations are quite good. Note that, given the asymptotic distribution (6), it is also possible to construct a separate test based on only the raw moments up

to order q and another test based only on the correlations up to order L, in order to separately assess which properties of the PITs hold.

A.4 Additional empirical results

Table A.1: Point Forecasts: Relative MSFE

B. Medium-scale model										
Horizons	Variables	VAR-flat	VAR-MN	VAR-MSD	TV-VAR-flat	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-flat	TV-SV-VAR-MN	TV-SV-VAR-MSD
One quarter	Real GDP	1.1864	1.1942	1.1286	1.1639	1.2765	1.1770	1.1712	1.4799	1.1856
	GDP deflator	1.6449	0.9177	0.8747	0.9500	1.1270	0.9168	1.1943	0.3043	0.8025
	Federal funds rates	1.6341	1.1965	0.6619	1.0464	1.0262	0.8987	1.3905	0.0532	0.8006
	consumption	0.9791	0.9775	0.7420	0.9582	1.0990	0.7790	1.1788	1.7613	0.6155
	investment	1.0200	0.9381	1.1895	0.9074	1.0186	1.0647	1.1592	34.3437	1.1762
	hoursworked	1.0701	1.0242	1.0105	1.0051	0.8741	1.0083	1.0437	1.5637	0.9033
	wages	1.1272	1.0358	1.0849	1.0397	1.0753	0.9720	0.9892	5.5513	1.1350
One year	Real GDP	1.4164	1.0440	0.8660	1.1506	1.0947	0.8840	0.9735	1.2409	0.8970
	GDP deflator	1.7737	1.0712	1.3283	1.3409	0.9694	1.2583	1.0707	0.2276	1.1533
	Federal funds rates	1.4867	1.1582	0.8976	1.2548	1.3326	0.9312	1.2833	0.0478	0.9528
	consumption	1.0122	1.0471	0.6890	1.1992	1.0620	0.5295	0.9596	1.4565	0.6690
	investment	1.4954	1.2079	0.5359	1.2418	1.2265	0.5112	1.2698	23.6708	0.5564
	hoursworked	0.8031	0.7558	0.5331	0.9371	0.7590	0.4262	1.0298	1.8869	0.5030
	wages	1.0096	1.0122	1.0296	0.9111	1.0397	0.8578	1.0870	1.4985	0.9387

Note: The table reports the results relative to the forecasting accuracy using point forecasts for medium-scale model. For the TV-SV-VAR-MN model we report the mean square forecast error (MSFE). For the other models we report the relative mean square forecast error (RMSFE), i.e. the ratio of the MSFE of a particular model to the MSFE of the baseline model.

Table A.2: Point Forecasts: Relative MSFE one-quater horizon

C. Large-scale model									
Variables	VAR-flat	VAR-MN	VAR-MSD	TV-VAR-flat	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-flat	TV-SV-VAR-MN	TV-SV-VAR-MSD
Real GDP	2.3688	1.0836	0.9187	1.0949	0.9987	0.7034	1.0931	1.6591	0.8022
GDP deflator	1.1821	1.0042	1.0630	1.0814	0.8811	0.6233	1.0761	0.4021	0.6703
Federal funds rates	1.4087	1.0740	1.1182	1.1403	1.1247	1.2857	1.1234	0.0609	1.2611
CPI	1.7764	1.1103	0.7433	1.0557	0.9777	0.8633	1.1024	1.3816	0.8109
commodity price	1.9205	1.0365	0.8178	0.9514	1.0969	1.1221	1.0399	14.3606	0.9930
industrial production	2.1629	0.8831	0.7108	0.8896	0.9140	0.7373	0.8550	4.2049	0.7790
employment	2.2304	0.9652	0.8663	0.9571	0.9565	0.8920	0.9712	0.6038	0.7623
employment services	1.7594	0.7872	0.7993	0.8803	0.8934	0.8547	0.9255	0.1534	0.7722
consumption	2.0414	1.5025	1.2775	1.1851	1.1225	1.1765	1.1124	1.1386	1.3551
investment	1.2617	1.0679	0.6322	0.8775	1.0725	0.5211	0.9399	36.3496	0.5882
PCE	1.4270	0.9314	0.6605	0.8618	0.9449	0.8205	0.8575	0.7789	0.8710
domestic investment	0.6839	0.9870	0.5434	0.9950	0.8368	0.5824	0.8176	0.8569	0.5736
capacity utilization	2.4201	0.7540	0.9557	0.8773	0.8200	0.9374	0.8645	149.9861	0.7157
consumer expectations	2.1434	1.0407	1.2995	1.0426	0.9807	1.1053	0.9847	190.9373	1.1732
hours worked	1.3622	0.9235	0.9977	0.8138	0.7795	1.0088	0.9890	1.4676	0.8995
wages	1.6926	1.0363	1.1648	0.9777	1.0401	1.1115	1.1826	5.5381	1.1195
one-year bond rate	4.3007	0.8599	0.9545	0.9099	1.1991	1.0451	0.8096	750.6073	1.0792
ten-years bond rate	2.3206	0.8651	0.9931	1.0268	0.9831	0.9879	1.2181	618.8149	0.9168
SP500	3.0510	1.4814	1.0330	1.3200	1.0133	1.0609	1.2089	191.6677	0.8810
exchange rate	1.9453	1.0735	1.0219	1.1431	1.0897	0.9961	0.9123	64.4798	1.1314
M2	1.8130	0.8482	1.1489	0.9572	1.1226	1.2138	0.8569	3.3681	1.2032

Note: The table reports the results relative to the forecasting accuracy using point forecasts for large-scale model at one quarter horizon. For the TV-SV-VAR-MN model we report the mean square forecast error (MSFE). For the other models we report the relative mean square forecast error (RMSFE), i.e. the ratio of the MSFE of a particular model to the MSFE of the baseline model.

Table A.3: Point Forecasts: Relative MSFE one-year horizon

C. Large-scale model									
Variables	VAR-flat	VAR-MN	VAR-MSD	TV-VAR-flat	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-flat	TV-SV-VAR-MN	TV-SV-VAR-MSD
Real GDP	1.1956	1.0429	0.4708	0.9615	0.8146	0.5556	0.9727	1.5093	0.5141
GDP deflator	1.4167	1.0634	1.0629	1.0863	1.0093	0.5751	1.0841	0.5646	0.5979
Federal funds rates	1.6387	1.0934	1.0324	1.0774	1.0691	0.3817	1.0312	0.0986	0.3667
CPI	1.7256	0.9279	0.6565	1.0889	0.9944	0.5783	0.9892	1.2059	0.5747
commodity price	1.7181	1.0844	0.5585	1.0790	0.9818	0.6842	0.9718	13.5162	0.6640
industrial production	0.9024	0.9667	0.2635	1.0289	1.0887	0.3041	0.9276	9.7797	0.2312
employment	0.9681	0.8734	0.6030	0.9821	0.9630	0.7326	0.9668	0.4884	0.6797
employment services	1.6473	1.0475	0.6626	1.1489	0.9780	0.7160	1.0962	0.2509	0.7593
consumption	1.0936	1.0015	0.4147	1.0539	0.9289	0.3748	1.0079	1.5649	0.4440
investment	0.9897	0.8486	0.3095	0.9638	0.8510	0.3102	0.9661	42.2301	0.2887
PCE	2.3071	1.1301	0.9333	0.9408	1.0693	0.8317	0.9891	0.6218	0.9529
domestic investment	0.9859	1.1859	0.5673	1.2083	1.1483	0.6318	1.2213	1.0574	0.5556
capacity utilization	1.3937	1.1328	0.3390	1.0259	0.8796	0.3168	0.8631	313.0338	0.3566
consumer expectations	2.2046	1.0443	0.7402	0.8879	0.8741	0.9066	1.0234	414.7756	0.8694
hours worked	0.4393	0.8205	0.3478	0.8335	0.9282	0.3983	0.8433	3.1172	0.4351
wages	1.4085	1.1060	0.8123	1.2579	1.1167	0.7391	1.3393	1.6580	0.7729
one-year bond rate	1.1372	0.9329	0.2777	0.8594	0.7659	0.3266	0.8520	197.0586	0.3197
ten-years bond rate	2.0298	1.2289	0.5156	1.3163	1.2023	0.5835	1.0090	414.7117	0.5764
SP500	2.1531	1.0969	0.9380	1.0574	1.0315	0.9427	1.0009	146.0654	0.9614
exchange rate	1.8385	1.1196	1.0707	1.4372	0.9654	1.1128	0.9883	25.5724	0.9749
M2	1.3601	1.2127	1.2301	1.2046	1.0840	1.3574	1.1299	3.5881	1.1666

Note: The table reports the results relative to the forecasting accuracy using point forecasts for large-scale model at one year horizon. For the TV-SV-VAR-MN model we report the mean square forecast error (MSFE). For the other models we report the relative mean square forecast error (RMSFE), i.e. the ratio of the MSFE of a particular model to the MSFE of the baseline model.

Table A.4: Density Forecasts: log scores

B. Medium-scale model									
Horizons	Variables	VAR-flat	VAR-MN	VAR-MSD	TV-VAR-flat	TV-VAR-MN	TV-VAR-MSD	TV-SV-VAR-flat	TV-SV-VAR-MSD
One quarter	Real GDP	0.5842	0.0073	-0.0560	0.0120	0.0132	-0.0468	0.0036	-0.0589
		(1.0152)	(1.4349)	(1.4260)	(1.4366)	(1.4642)	(1.4198)	(1.4454)	(1.4073)
	GDP deflator	0.6060	0.0057	-0.0509	0.0401	-0.0276	-0.0526	-0.0224	-0.0391
		(1.1068)	(1.5631)	(1.5006)	(1.5297)	(1.5464)	(1.5153)	(1.5350)	(1.5303)
	Federal funds rates	0.0534	0.0005	0.0787	0.0022	0.0112	-0.0847	0.0153	-0.0691
		(0.9029)	(1.2852)	(1.2214)	(1.2832)	(1.2862)	(1.2180)	(1.2860)	(1.2204)
	consumption	0.2305	0.0024	0.0156	0.0013	-0.0024	0.0226	0.0048	0.0058
		(1.0914)	(1.5493)	(1.4782)	(1.5298)	(1.5681)	(1.4856)	(1.5509)	(1.4761)
	investment	0.3132	-0.0052	-0.0013	0.0030	0.0022	-0.0128	0.0058	0.0016
		(1.0571)	(1.4737)	(1.4661)	(1.4836)	(1.4889)	(1.4462)	(1.4901)	(1.4390)
	hoursworked	0.2896	0.0027	-0.0365	0.0038	0.0087	-0.0322	0.0027	-0.0324
		(1.1182)	(1.6043)	(1.5396)	(1.6031)	(1.6073)	(1.5210)	(1.5906)	(1.5173)
	wages	0.0977	0.0031	0.0093	0.0102	0.0139	0.0070	0.0042	0.0126
		(1.4649)	(2.1129)	(2.0483)	(2.0997)	(2.1091)	(2.0303)	(2.1007)	(2.0379)
One year	Real GDP	0.2439	0.0775	0.1281	0.1206	0.2140	0.0399	0.0130	0.0063
		(7.2287)	(10.2792)	(10.6678)	(10.3078)	(10.5021)	(10.5345)	(10.3627)	(10.6221)
	GDP deflator	0.8429	0.2575	2.3283	0.1698	0.1132	0.2320	0.0599	2.1444
		(18.0077)	(25.4739)	(29.3281)	(25.2951)	(25.4412)	(29.1229)	(25.2993)	(29.2058)
	Federal funds rates	0.2376	0.0669	-0.8071	0.0556	0.0675	-0.7383	0.0251	-0.7094
		(5.5009)	(7.8021)	(7.3775)	(7.8658)	(7.9428)	(7.5868)	(7.7909)	(7.5327)
	consumption	0.8795	0.0306	-0.0344	0.0305	0.0422	-0.0693	0.0111	-0.0369
		(8.0424)	(11.4445)	(11.0417)	(11.3603)	(11.5313)	(11.0261)	(11.4541)	(11.0436)
	investment	0.3659	0.0342	0.0480	0.0354	0.0536	0.0002	0.0054	0.0470
		(6.0436)	(8.6177)	(8.9589)	(8.5949)	(8.6724)	(8.8524)	(8.5312)	(8.8856)
	hoursworked	1.2417	-0.0100	0.0791	0.0369	0.0100	0.1415	0.0168	0.0740
		(11.0395)	(15.8128)	(16.9942)	(15.8025)	(15.9508)	(16.7394)	(15.7924)	(16.6976)
	wages	0.5451	0.0502	0.0289	0.0160	0.0009	0.0076	0.0092	-0.0442
		(6.3849)	(9.0846)	(9.1744)	(9.1120)	(9.1232)	(9.1541)	(9.0318)	(9.0972)

Note: The table reports the results relative to the forecasting accuracy using predictive densities for medium-scale model. For each model we report the sample average of the difference between the log score of TV-SV-VAR-MN and the log score of that model. Standard deviation (HAC) are reported in parentheses.

Appendix B

Supplementary Material for Chapter 2

B.1 More Plots in Section 5.2

B.1.1 Panel A: $(b_0^{\top,1}, b_0^{\top,2}) = (0.5, -0.5)$, $\sigma_T^2 = \frac{\delta_1\delta_2}{-b_{0,1}^T b_{0,2}^T}$

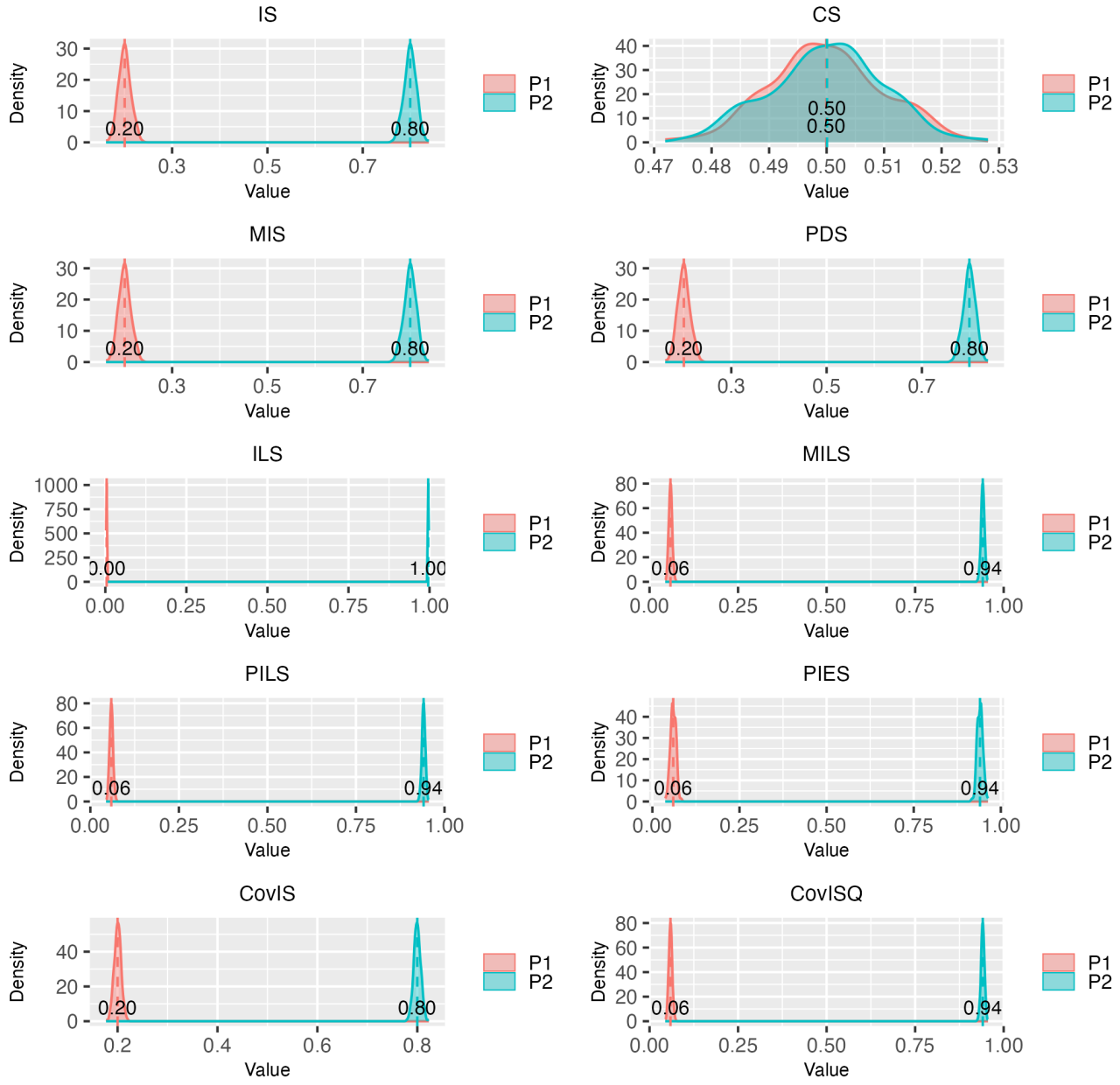

Figure B.1: Distribution of price discovery measurements for $\delta = (0.2, 0.8)$ in panel A.

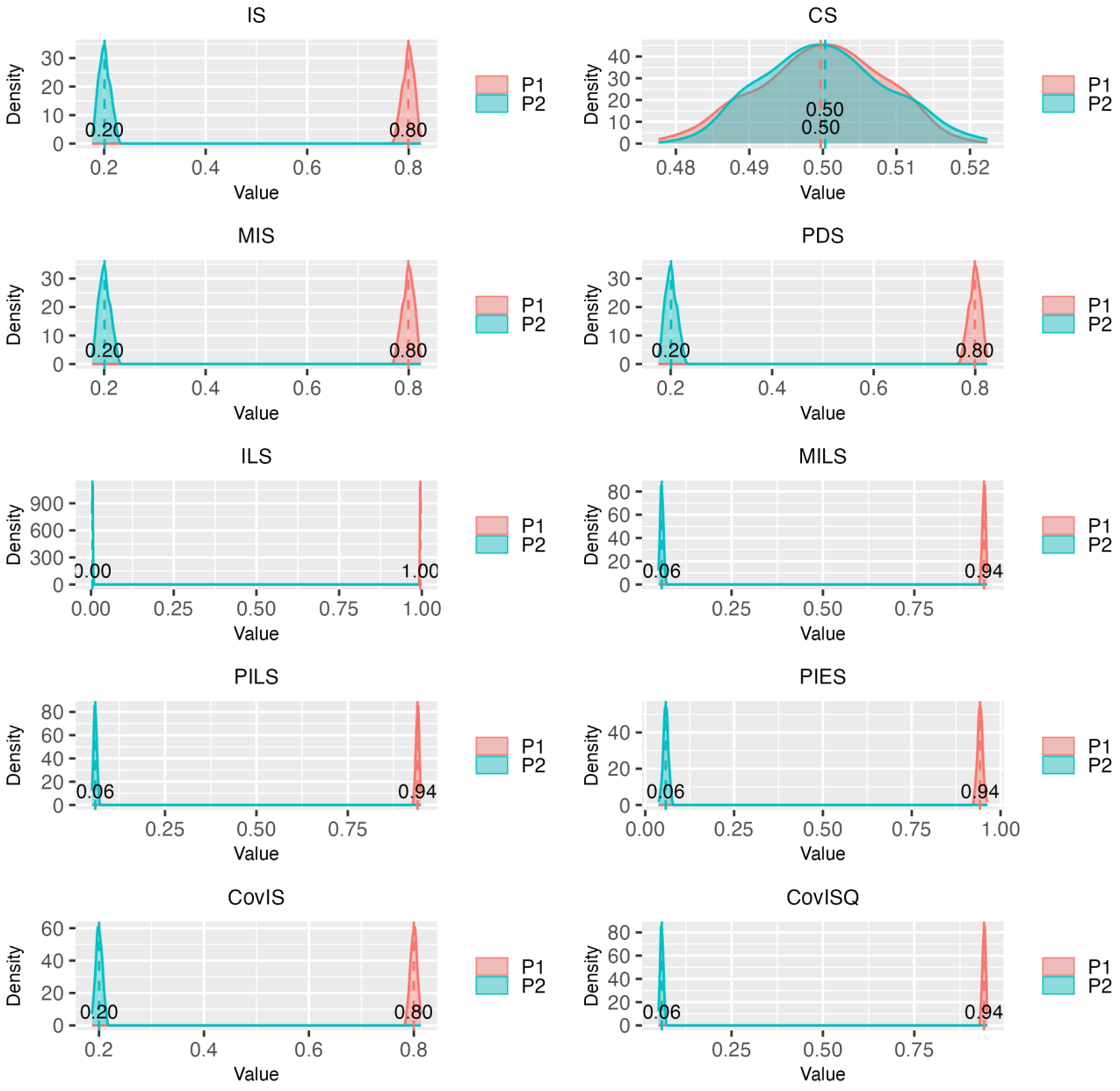

Figure B.2: Distribution of price discovery measurements for $\delta = (0.8, 0.2)$ in panel A.

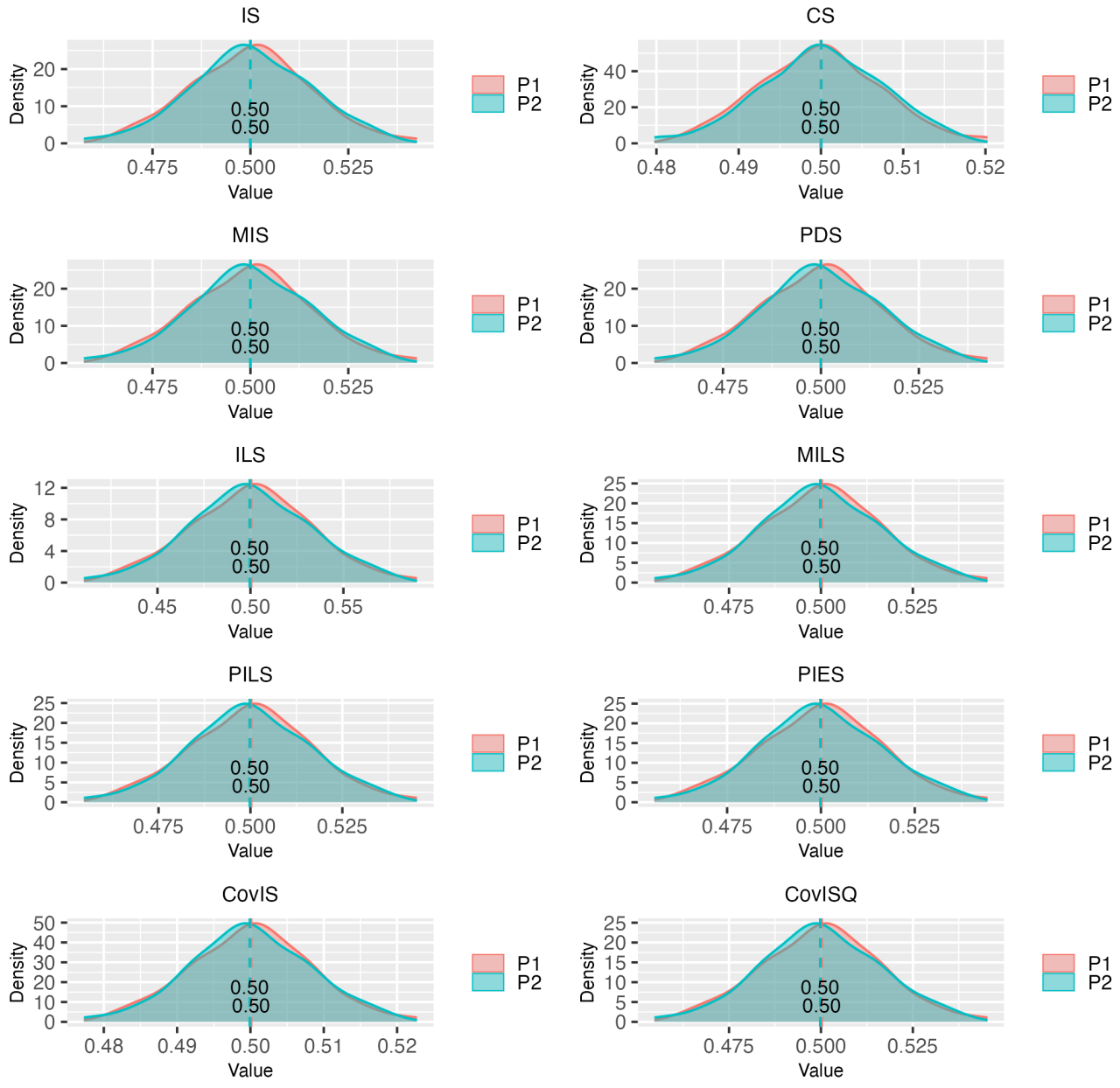

Figure B.3: Distribution of price discovery measurements for $\delta = (0.5, 0.5)$ in panel A.

B.1.2 Panel B: $(b_0^{\top,1}, b_0^{\top,2}) = (0.8, -0.2)$, $\sigma_T^2 = \frac{\delta_1 \delta_2}{-b_{0,1}^T b_{0,2}^T}$

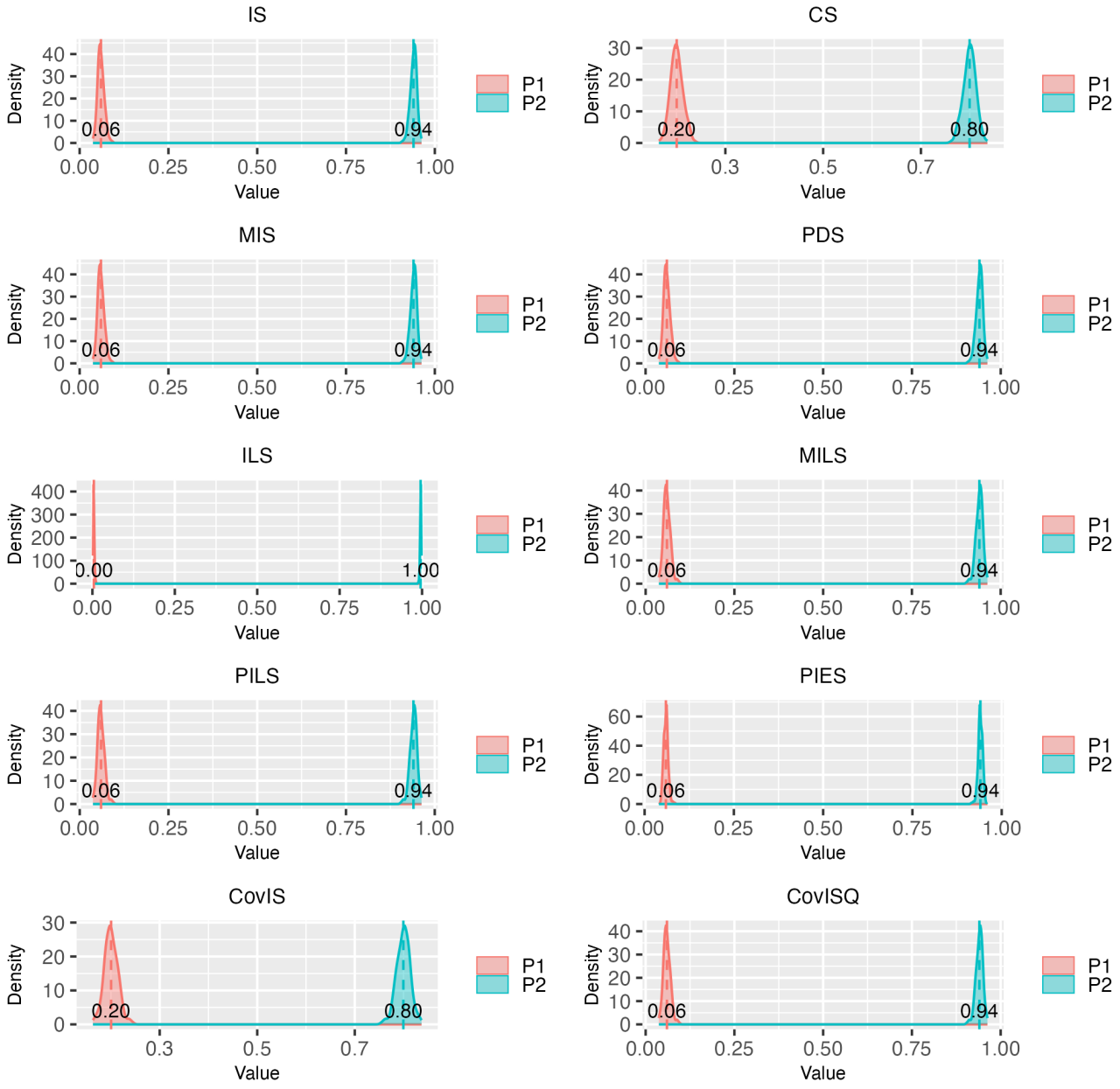

Figure B.4: Distribution of price discovery measurements for $\delta = (0.2, 0.8)$ in panel B.

clearpage

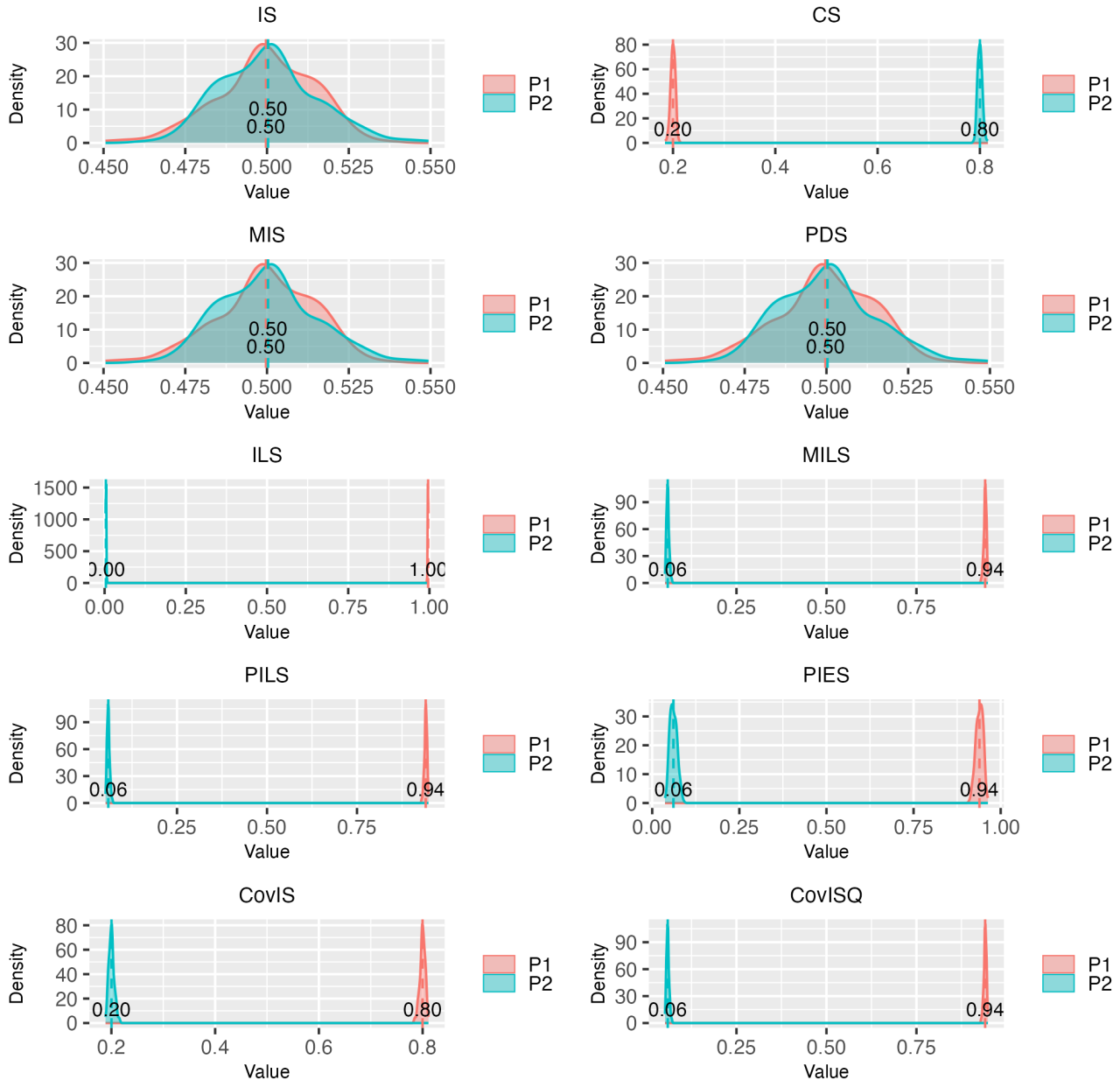

Figure B.5: Distribution of price discovery measurements for $\delta = (0.8, 0.2)$ in panel B.

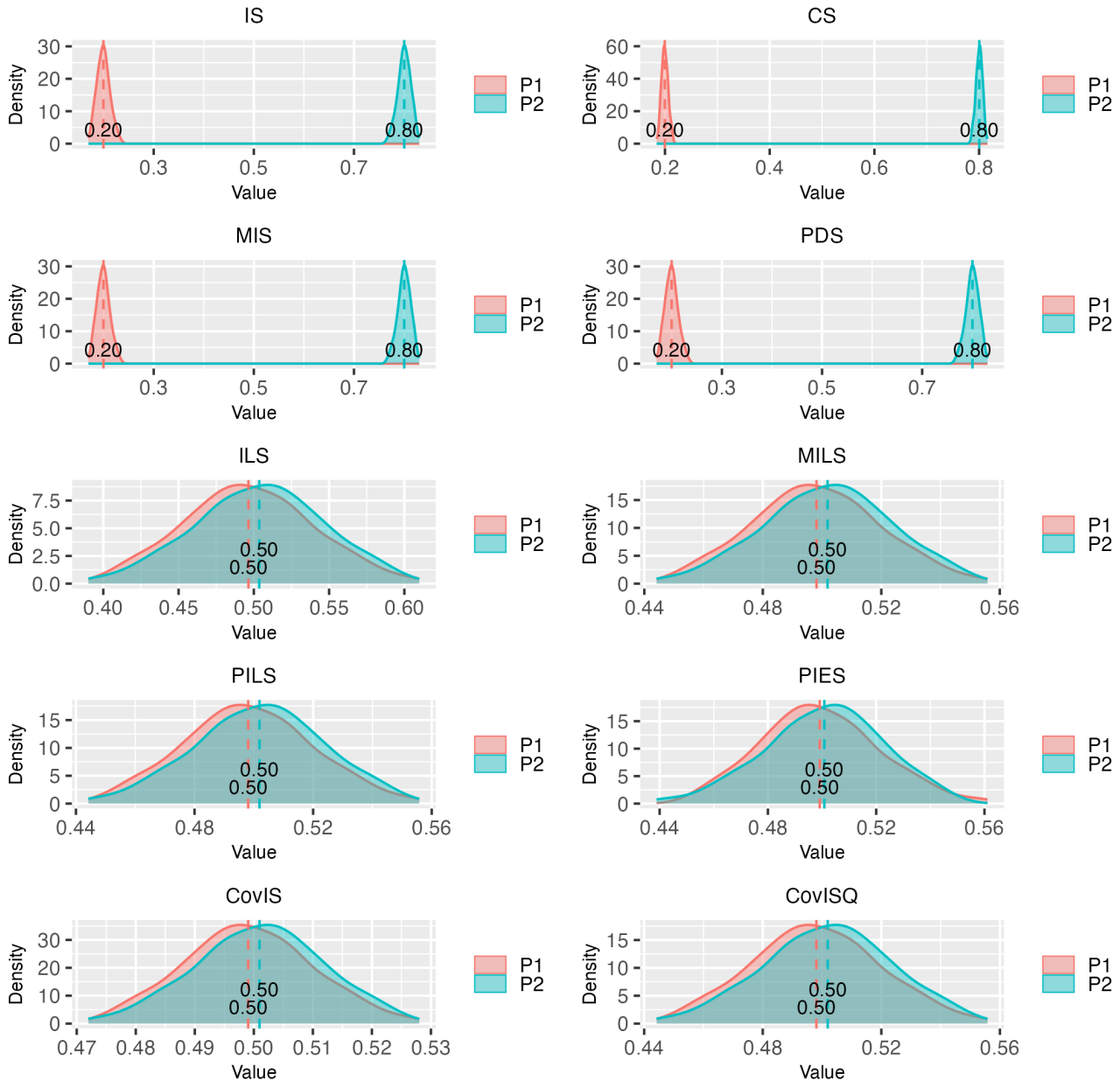

Figure B.6: Distribution of price discovery measurements for $\delta = (0.5, 0.5)$ in panel B.

B.2 More Plots in Section 6.3

B.2.1 Distribution of price discovery measures at Arca

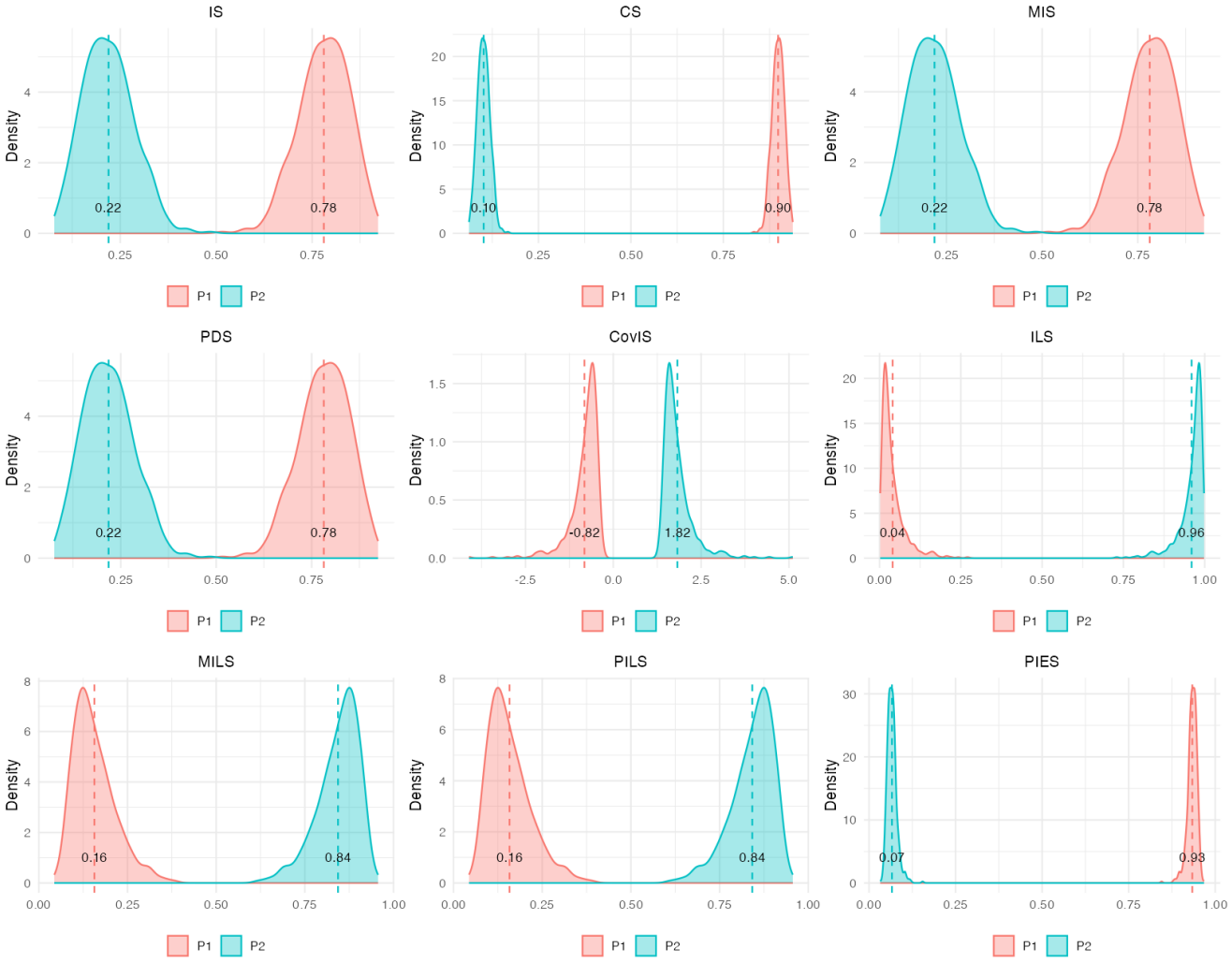

Figure B.7: Distribution of price discovery measures at Arca on May 6th, 2010 (flash-crash)

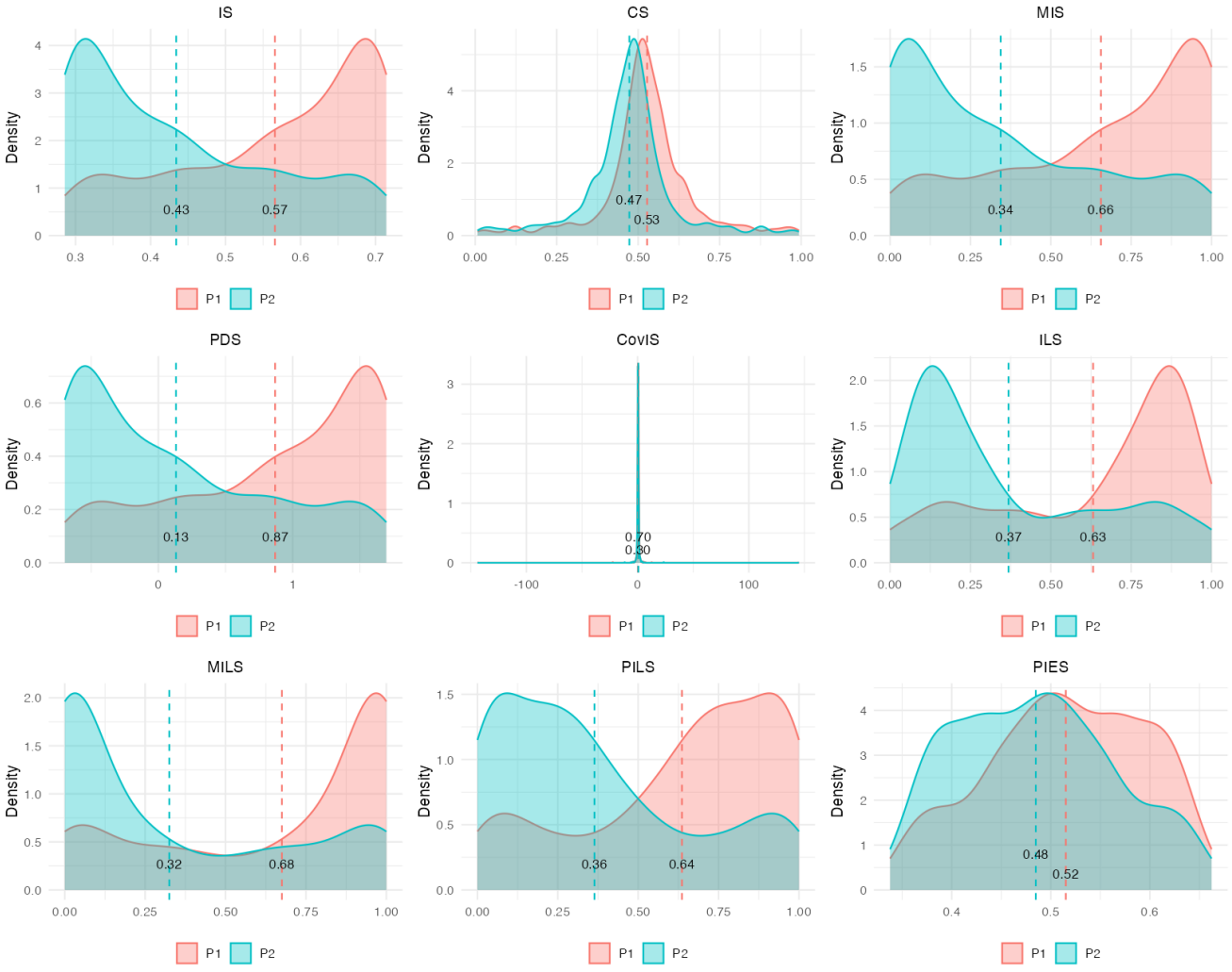

Figure B.8: Distribution of price discovery measures at Arca on Aug 8th, 2011

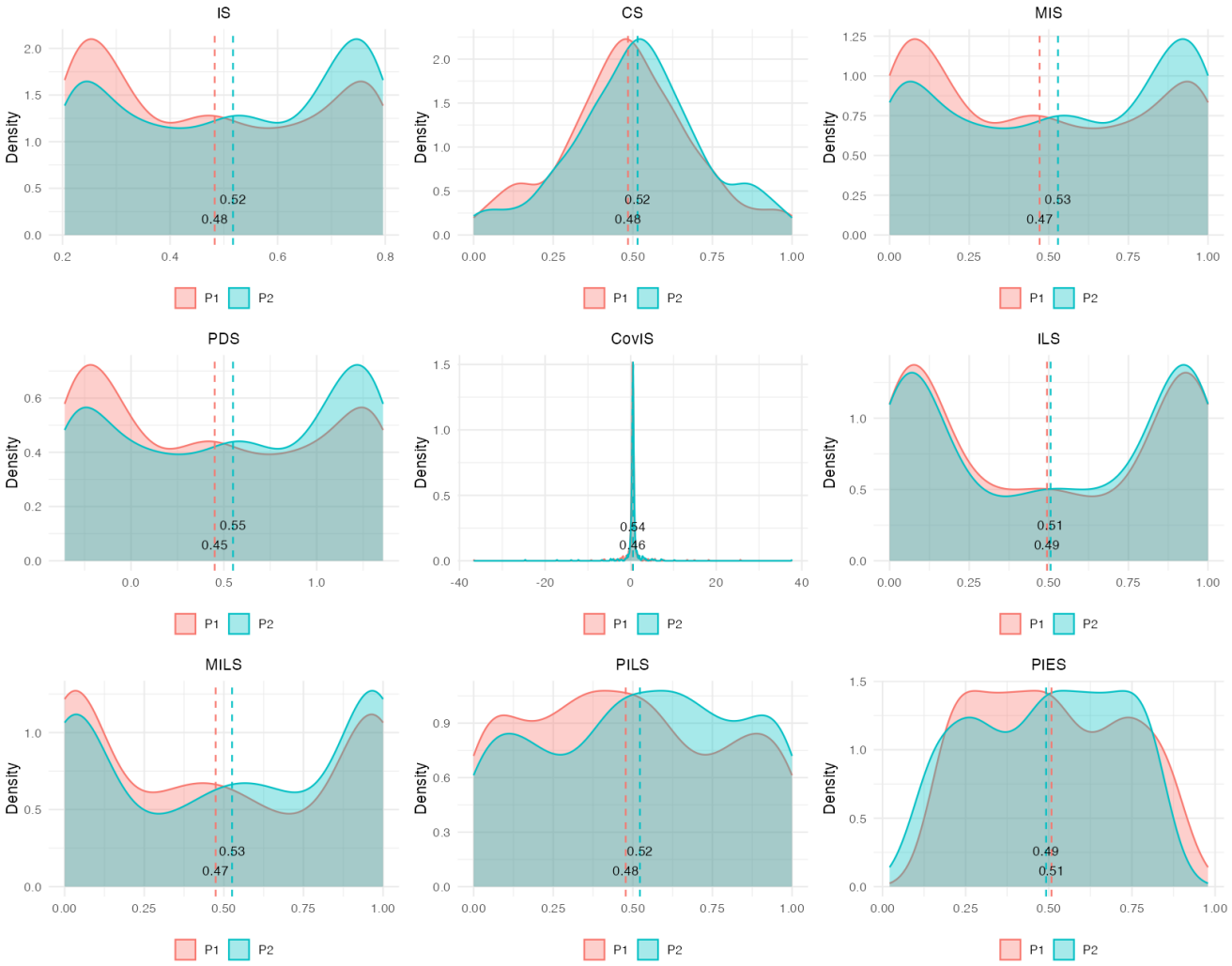

Figure B.9: Distribution of price discovery measures at Arca on Dec 4th, 2012

B.2.2 Distribution of price discovery measures at BATS

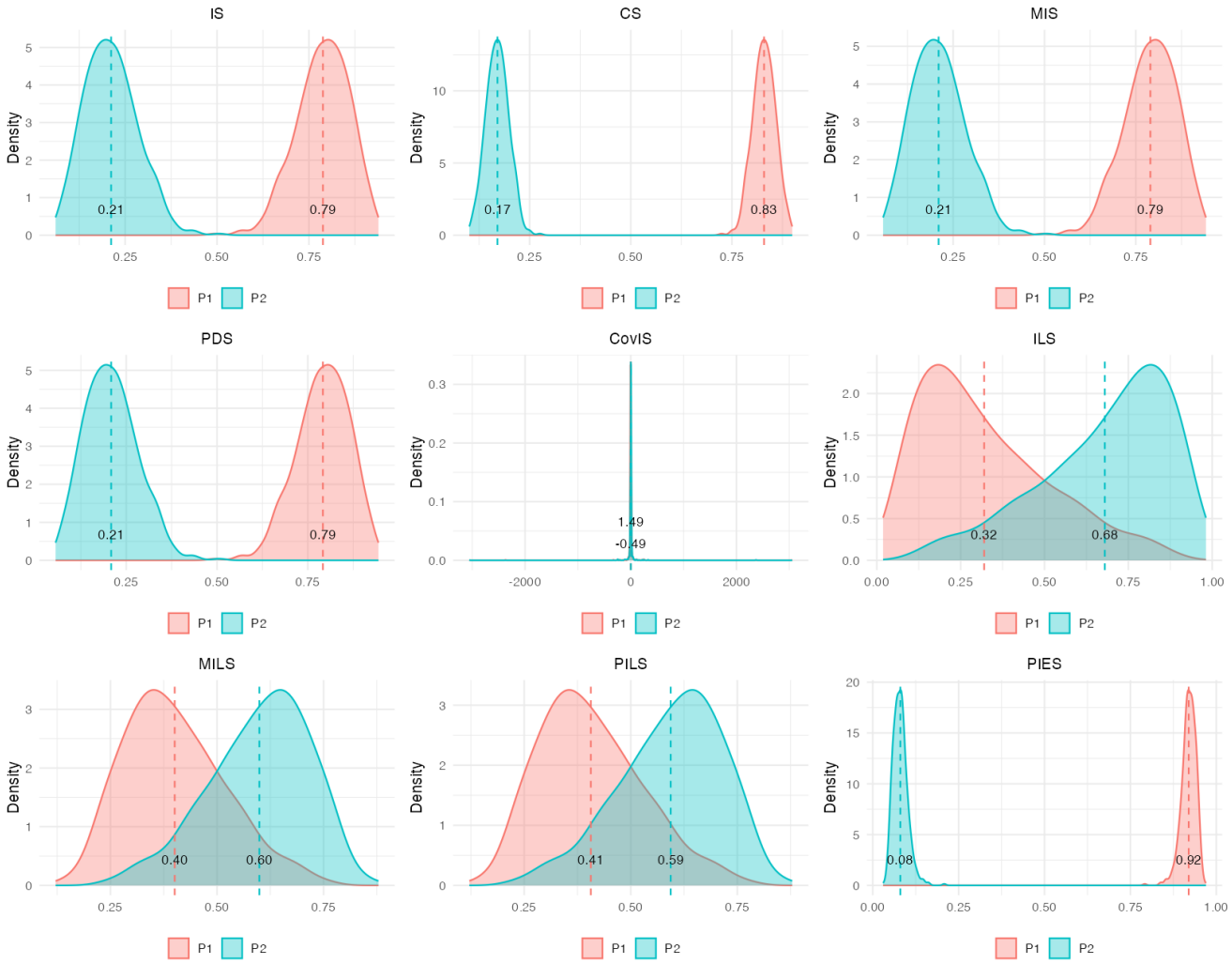

Figure B.10: Distribution of price discovery measures at BATS on May 6th, 2010 (flash-crash)

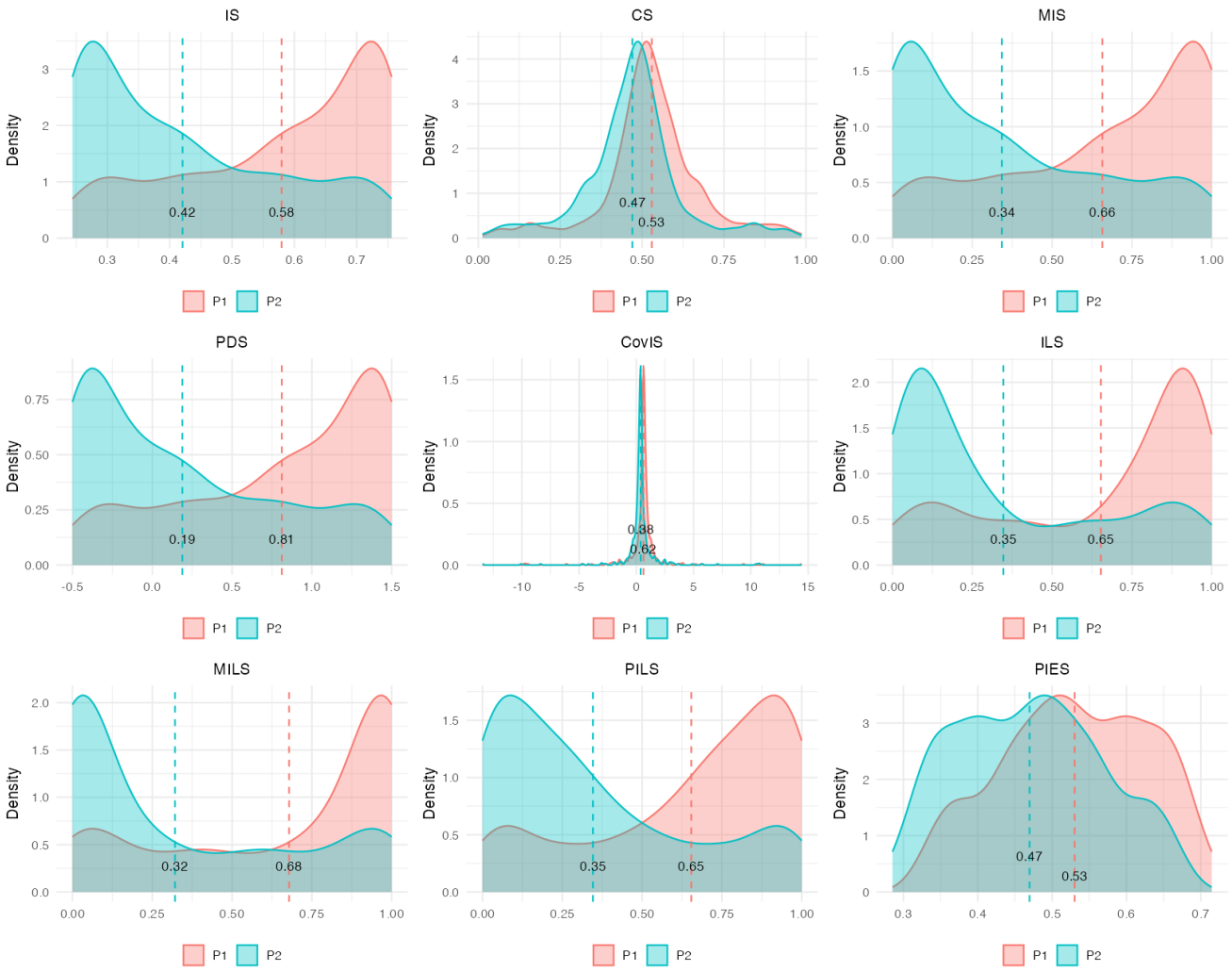

Figure B.11: Distribution of price discovery measures at BATS on Aug 8th, 2011

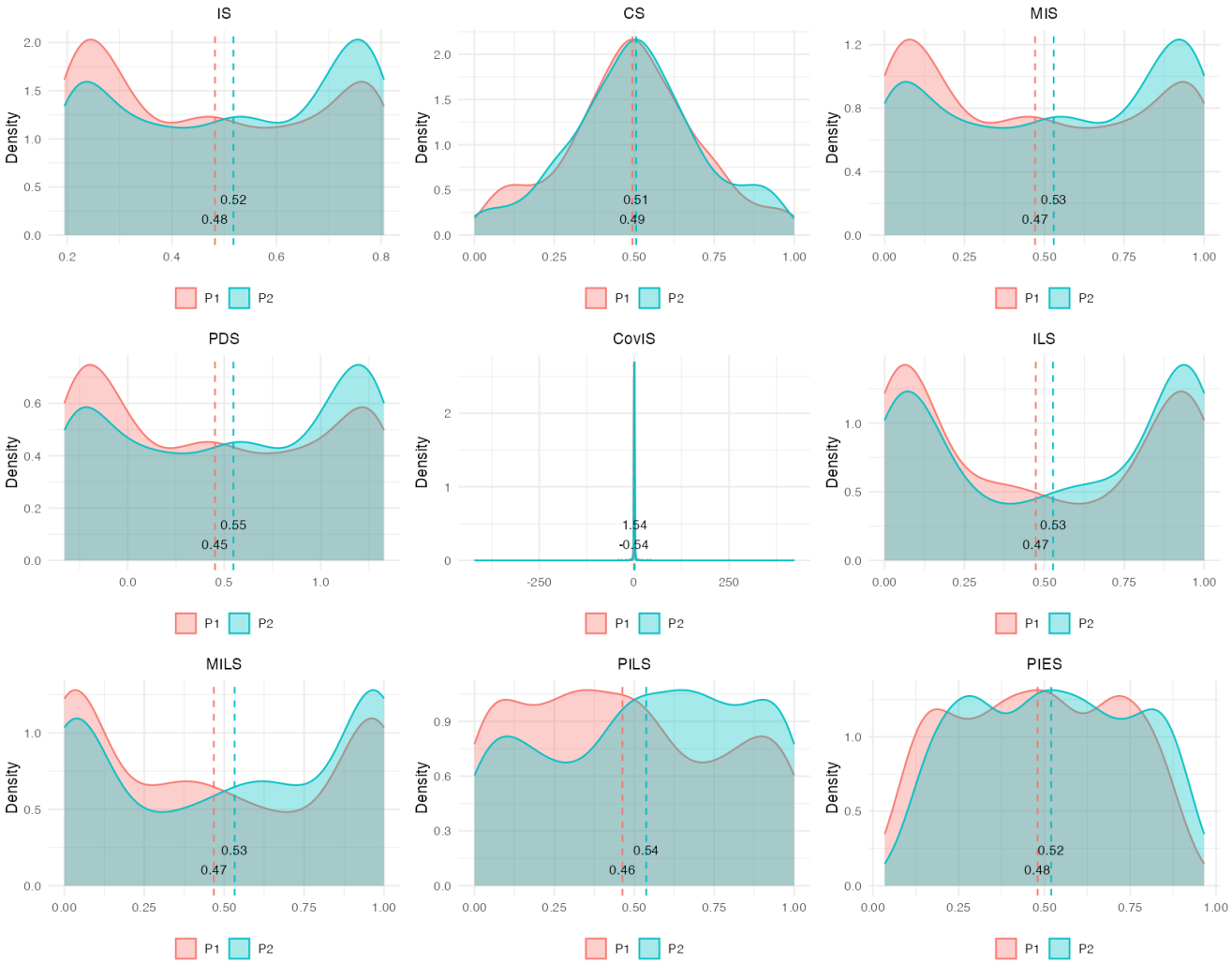

Figure B.12: Distribution of price discovery measures at BATS on Dec 4th, 2012

Appendix C

Supplementary Material for Chapter 3

C.1 Ordering Issue for TVP-VAR Price discovery

In this section, we introduce the time-varying price discovery share measures based on the TVP-BVAR model in Primiceri (2005). In order to give clear interpretations of these price discovery measures, we follow the structural analysis in Yan and Zivot (2010) to derive expressions for these price discovery measures in terms of interpretable structural parameters.

Let $\mathbf{p}_t = (p_{1t}, p_{2t})$ denote a vector of log prices for two assets that are closely related by arbitrage. For example, these can be the same asset traded in different markets or two nearly identical assets traded in the same market. In the price discovery literature, it is assumed that these two prices series are intergated of order 1, or I(1), the price changes, $\Delta\mathbf{p}_t$, are stationary, or I(0), and that $\mathbf{p}_t$ is cointegrated with cointegrating vector $\beta = (1, -1)'$. Price discovery measures are typically derived from a reduced-form VECM of the form:

$$\Delta\mathbf{p}_t = \alpha\beta'\mathbf{p}_{t-1} + \sum_{j=1}^{k} \mathbf{\Gamma}_j \Delta\mathbf{p}_{t-j} + \varepsilon_t, \tag{C.1}$$

where $\alpha = (\alpha_1, \alpha_2)'$ is the error correction vector, $\mathbf{\Gamma}_j (j = 1, \ldots, k)$ are the short-run coefficient matrices, and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ is the vector of reduced-form VECM residuals with $E[\varepsilon_t] = \mathbf{0}$ and

$$E[\varepsilon_t \varepsilon_s'] = \begin{cases} 0, & \text{if } t \neq s, \\ \mathbf{\Omega}, & \text{otherwise.} \end{cases}$$

and the residual covariance matrix Ω is

$$\mathbf{\Omega} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

where σ_i^2 denotes the variance of each market's idiosyncratic error and ρ denotes the correlation coefficient between these two errors.

In the spirit of time variation, the VECM can be rewritten as

$$\Delta \mathbf{p}_t = \alpha_t \beta' \mathbf{p}_{t-1} + \sum_{j=1}^{k} \mathbf{\Gamma}_{j,t} \Delta \mathbf{p}_{t-j} + \varepsilon_t, \tag{C.2}$$

now $\alpha_{\mathbf{t}} = (\alpha_{1,t}, \alpha_{2,t})'$ is the error correction vector, $\mathbf{\Gamma}_{j,t} (i = 1, \ldots, k)$ are the short-run coefficient matrices, and

$$E\left[\varepsilon_t \varepsilon_s'\right] = \begin{cases} 0, & \text{if } t \neq s, \\ \mathbf{\Omega}_t, & \text{otherwise.} \end{cases}$$

and the residual covariance matrix $\mathbf{\Omega}_t$ is

$$\mathbf{\Omega}_t = \begin{pmatrix} \sigma_{1t}^2 & \rho_t\sigma_{1t}\sigma_{2t} \\ \rho_t\sigma_{1t}\sigma_{2t} & \sigma_{2t}^2 \end{pmatrix}.$$

Without loss of generality, consider the triangular reduction of $\mathbf{\Omega}_t$, defined by

$$\mathbf{A}_t \mathbf{\Omega}_t \mathbf{A}_t' = \mathbf{\Sigma}_t \mathbf{\Sigma}_t'$$

where $\mathbf{A}_t$ is the lower triangular matrix

$$\mathbf{A}_t = \begin{pmatrix} 1 & 0 \\ a_t & 1 \end{pmatrix},$$

and $\mathbf{\Sigma}_t$ is the diagonal matrix

$$\mathbf{\Sigma}_t = \begin{pmatrix} \omega_{1,t} & 0 \\ 0 & \omega_{2,t} \end{pmatrix}.$$

It follows that

$$\Delta \mathbf{p}_t = \alpha_t \beta' \mathbf{p}_{t-1} + \sum_{j=1}^{k} \mathbf{\Gamma}_{j,t} \Delta \mathbf{p}_{t-j} + \mathbf{A}_t^{-1} \mathbf{\Sigma}_t \mathbf{e}_t,$$

$$V(\mathbf{e}_t) = \mathbf{I_n}.$$

Next, we specify the dynamics of time-varying coefficients and stochastic volatility. Stack in a vector $\mathbf{B}_t$ all the R.H.S. coefficient α_t and $\mathbf{\Gamma}_{j,t}, j = 1, ..., k$.

The error correction α_t, the short-run coefficients $\mathbf{\Gamma}_{j,t}$ and the non-zero and non-one elements (free elements) of the matrix $\mathbf{A}_t$, a_t, are all assumed to evolve according to independent random walks:

$$\mathbf{B}_t = \mathbf{B}_{t-1} + \mathbf{v}_t \qquad \mathbf{v}_t \sim \mathcal{N}(0, \mathbf{Q}) \tag{C.3}$$

$$a_t = a_{t-1} + \zeta_t \qquad \zeta_t \sim \mathcal{N}(0, S) \tag{C.4}$$

The standard deviation $\omega_t = (\omega_{1,t}, \omega_{2,t})'$ is assumed to evolve as geometric random walks:

$$\log \omega_t = \log \omega_{t-1} + \eta_t \qquad \eta_t \sim \mathcal{N}(0, \mathbf{W}) \tag{C.5}$$

After estimating $\alpha_t, \mathbf{\Gamma}_{j,t}$, and $\mathbf{\Omega}_t$, we transform the above time-varying VECM model into a reduced-form time-varying vector moving average (VMA) model following Hasbrouck (1995):

$$\Delta \mathbf{p}_t = \mathbf{\Psi}_t(L)\varepsilon_t = \varepsilon_t + \mathbf{\Psi}_{1,t}\varepsilon_{t-1} + \mathbf{\Psi}_{2,t}\varepsilon_{t-2} + \cdots, \tag{C.6}$$

and its integrated form (or Beveridge-Nelson decomposition):

$$\mathbf{p}_t = \mathbf{p}_0 + \mathbf{\Psi}_t(1) \sum_{s=1}^{t} \varepsilon_s + \Psi_t^*(L)\varepsilon_t, \tag{C.7}$$

where $\mathbf{\Psi}_t(1) = \sum_{k=0}^{\infty} \mathbf{\Psi}_{k,t}$ with $\mathbf{\Psi}_t(L)$ and $\mathbf{\Psi}_t^*(L)$ being matrix polynomials in the lag operator, L, and $\mathbf{\Psi}_t^*(k) = -\sum_{j=k+1}^{\infty} \mathbf{\Psi}_{j,t}$. The matrix $\mathbf{\Psi}_t(1)$ contains the cumulative impacts of the innovation ε_t on all future price movements, and acts as a measure of the long-run impact of ε_t on prices.

Johansen (1991) shows that the reduced-form VECM model and the integrated VMA model

are linked through the factorization:

$$\mathbf{\Psi}_t(1) = \beta_\perp \mathbf{\Pi}_t \alpha'_{t\perp},$$

where $\beta_\perp$ is orthogonal to β, and $\mathbf{\Pi}_t = \left(\alpha'_{t\perp}\left(\mathbf{I} - \sum_{j=1}^{k}\mathbf{\Gamma}_{j,t}\right)\beta_\perp\right)^{-1}$ where $\mathbf{I}$ is the identity matrix and $\alpha_{t\perp}$ is orthogonal to α_t. We can see that $\mathbf{\Pi}_t$ for the bivariate VECM model is a scalar.

Since the cointegrating vector is $\beta = (1, -1)'$, one choice for $\beta_\perp$ is $\beta_\perp = (1, 1)'$. If we denote $\alpha_{t\perp} = (\gamma_{1,t}, \gamma_{2,t})'$, then we can represent $\mathbf{\Psi}_t(1)$ as:

$$\mathbf{\Psi}_t(1) = \mathbf{\Pi}_t \beta_\perp \alpha'_{t\perp} = \mathbf{\Pi} \begin{pmatrix} \gamma_{1,t} & \gamma_{2,t} \\ \gamma_{1,t} & \gamma_{2,t} \end{pmatrix}.$$

The rows of $\mathbf{\Psi}_t(1)$ are identical, and the long-run impacts of an innovation ε_t on each of the prices are identical. The increment $\psi'_t\varepsilon_t$ (with $\psi_\mathbf{t} = (\psi_{1,t}, \psi_{2,t})'$ the common row of $\mathbf{\Psi}_t(1)$) is the component of the price change that is permanently impounded into the price (presumably due to new information) which we denote as:

$$\eta_t^P = \mathbf{\Psi}'\varepsilon_t = \psi_{1,t}\varepsilon_{1,t} + \psi_{2,t}\varepsilon_{2,t}.$$

This common efficient price evolves as a random walk driven by the permanent shock η_t^P.

C.1.1 Discussion for Time-varying Price Discovery

The TVP-VAR model from Primiceri (2005) is likely to cause problems due to the assumption of a triangular factorization of the residual covariance matrix in the VECM model. To see this, consider the TVP-VECM

$$\Delta\mathbf{p}_t = \alpha_t\beta'\mathbf{p}_{t-1} + \sum_{j=1}^{k}\mathbf{\Gamma}_{j,t}\Delta\mathbf{p}_{t-j} + \varepsilon_t, \tag{C.8}$$

$$E\left[\varepsilon_t\varepsilon'_s\right] = \begin{cases} 0, & \text{if } t \neq s, \\ \mathbf{\Omega}_t, & \text{otherwise.} \end{cases}$$

and the residual covariance matrix $\mathbf{\Omega}_t$ is

$$\mathbf{\Omega}_t = \begin{pmatrix} \omega_{1t}^2 & \omega_{12,t} \\ \omega_{12,t} & \omega_{2t}^2 \end{pmatrix}, \omega_{12,t} = \rho_{12,t}\omega_{1t}\omega_{2t}.$$

Primiceri employs the triangular factorization of $\mathbf{\Omega}_t$

$$\mathbf{A}_t\mathbf{\Omega}_t\mathbf{A}_t' = \mathbf{\Sigma}_t\mathbf{\Sigma}_t'$$

where $\mathbf{A}_t$ is the lower triangular matrix

$$\mathbf{A}_t = \begin{pmatrix} 1 & 0 \\ a_t & 1 \end{pmatrix},$$

and $\mathbf{\Sigma}_t$ is the diagonal matrix

$$\mathbf{\Sigma}_t = \begin{pmatrix} \omega_{1,t} & 0 \\ 0 & \omega_{2,t} \end{pmatrix}.$$

Then the VECM model can be rewritten as

$$\Delta\mathbf{p}_t = \alpha_t\beta'\mathbf{p}_{t-1} + \sum_{j=1}^{k}\mathbf{\Gamma}_{j,t}\Delta\mathbf{p}_{t-j} + \mathbf{A}_t^{-1}\mathbf{\Sigma}_t\mathbf{e}_t,$$

$$E\left(\mathbf{e}_t\mathbf{e}_t'\right) = \mathbf{I_n}$$

Remarks:

(1) The rewritten model can be thought of as a particular SVECM

$$\Delta\mathbf{p}_t = \alpha_t\beta'\mathbf{p}_{t-1} + \sum_{j=1}^{k}\mathbf{\Gamma}_{j,t}\Delta\mathbf{p}_{t-j} + \mathbf{A}_t^{-1}\eta_t,$$

$$\eta_t = \mathbf{\Sigma}_t\mathbf{e}_t$$

where $\mathbf{A}_t$ is the structural initial impact matrix and η_t is a vector of independent structural shocks.

$$\varepsilon_t = \mathbf{A}_t^{-1}\eta_t \Rightarrow \eta_t = \mathbf{A}_t\varepsilon_t$$

that is

$$\begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ a_t & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$
$$\Rightarrow \eta_{1t} = \varepsilon_{1t}, \quad \varepsilon_{2t} = a_t \varepsilon_{1t} + \eta_{2t}.$$

(2) If we interpret η_{1t} as the permanent shock η_t^P and η_{2t} as the transitory shock η_t^T, then we know that

$$\eta_t^P = \psi_{1,t}\varepsilon_{1,t} + \psi_{2,t}\varepsilon_{2,t} = \varepsilon_{1,t}$$

where $\mathbf{\Psi_t}$ is long-run impact matrix

$$\mathbf{\Psi_t} = \begin{pmatrix} \psi_{1,t} & \psi_{2,t} \\ \psi_{1,t} & \psi_{2,t} \end{pmatrix}$$

Hence $\psi_{1,t} = 1$ and $\psi_{2,t} = 0$, which means that ψ's are not time-varying. $\mathbf{\Psi_t}$ is the structural long-run impact matrix

$$\mathbf{\Psi_t} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = D(1).$$

(3) The underlying SMA model has the form

$$\begin{aligned} \Delta \mathbf{p}_t &= \varepsilon_t + \mathbf{\Psi}_{1,t}\varepsilon_{t-1} + \mathbf{\Psi}_{2,t}\varepsilon_{t-2} + \cdots \\ &= \mathbf{D}_{0,t}\eta_t + \mathbf{D}_{1,t}\eta_{t-1} + \cdots \end{aligned}$$

where $\eta_t = (\eta_t^P, \eta_t^T)$, $\eta_t \sim \mathcal{N}(0, \mathbf{\Sigma_t})$.

Since $\eta_t = \mathbf{A}_t\varepsilon_t$,

$$\mathbf{D}_{0,t} = \begin{pmatrix} d_{0,1,t}^P & d_{0,1,t}^T \\ d_{0,2,t}^P & d_{0,2,t}^T \end{pmatrix} = \mathbf{A}_t^{-1} = \begin{pmatrix} 1 & 0 \\ \tilde{a}_t & 1 \end{pmatrix}$$

So that $d_{0,1,t}^P = d_{0,1}^P = 1$ which is not time-varying, and $d_{0,2,t}^P = -a_t$ which is time-varying.

(4) The price discovery meansure are then

$$\text{PILS}_1 = \frac{(d^P_{0,1})^2}{(d^P_{0,1})^2 + (d^P_{0,2})^2} = \frac{1}{1+\tilde{a}_t},$$
$$\text{PILS}_2 = \frac{(d^P_{0,2})^2}{(d^P_{0,1})^2 + (d^P_{0,2})^2} = \frac{(\tilde{a}_t)^2}{1+\tilde{a}_t}.$$

Since $\tilde{a}_t$ captures the correlation between p_{1t} and p_{2t}, if $\tilde{a}_t = 0$, then the two prices are uncorrelated and

$$\text{PILS}_1 = 1, \quad \text{PILS}_2 = 0$$

So Primiceri's triangular model places rather strong restrictions on the price discovery structural model.

C.1.2 General SVAR Framework

The general SMA framework is

$$\Delta \mathbf{p}_t = \mathbf{D}_{0,t}\eta_t + \mathbf{D}_{1,t}\eta_{t-1} + \cdots$$

where

$$\mathbf{D}_{0,t} = \begin{pmatrix} d^P_{0,1,t} & d^T_{0,1,t} \\ d^P_{0,2,t} & d^T_{0,2,t} \end{pmatrix}$$

$\mathbf{D}_{0,t}$ is unrestricted.

The reduced-form VMA is

$$\Delta \mathbf{p}_t = \varepsilon_t + \mathbf{\Psi}_{1,t}\varepsilon_{t-1} + \mathbf{\Psi}_{2,t}\varepsilon_{t-2} + \cdots$$

where $\varepsilon_t = \mathbf{D}_{0,t}\eta_t$, i.e., $\eta_t = \mathbf{D}^{-1}_{0,t}\varepsilon_t$. Then

$$\begin{aligned}\Delta \mathbf{p}_t &= \mathbf{D}_{0,t}\mathbf{D}^{-1}_{0,t}\varepsilon_t + \mathbf{\Psi}_{1,t}\mathbf{D}_{0,t-1}\mathbf{D}^{-1}_{0,t-1}\varepsilon_{t-1} + \cdots \\ &= \mathbf{D}_{0,t}\eta_t + \mathbf{D}_{1,t}\eta_{t-1} + \cdots\end{aligned}$$

Then the VECM can be expressed as

$$\Delta \mathbf{p}_t = \alpha_t \beta' \mathbf{p}_{t-1} + \sum_{j=1}^{k} \mathbf{\Gamma}_{j,t} \Delta \mathbf{p}_{t-j} + \mathbf{D}_{0,t} \eta_t, \tag{C.9}$$

$$Var(\varepsilon_t) = \mathbf{\Omega_t} = \mathbf{D}_{0,t} Var(\eta_t) \mathbf{D}'_{0,t} = \mathbf{D}_{0,t} \Sigma_t \mathbf{D}'_{0,t}.$$

Now, the elements of $\mathbf{D}_{0,t}$ can be identified. The first column of $\mathbf{D}_{0,t}$ has the form

$$d^P_{0,1,t} = \frac{\text{cov}\left(e_{1,t}, \eta^P_t\right)}{\text{var}\left(\eta^P_t\right)} = \frac{\left(\mathbf{\Omega_t \Psi}_t\right)_1}{\mathbf{\Psi}_t \mathbf{\Omega_t} \mathbf{\Psi}'_t},$$
$$d^P_{0,2,t} = \frac{\text{cov}\left(e_{2,t}, \eta^P_t\right)}{\text{var}\left(\eta^P_t\right)} = \frac{\left(\mathbf{\Omega_t \Psi}_t\right)_2}{\mathbf{\Psi}_t \mathbf{\Omega_t} \mathbf{\Psi}'_t}$$

where $\mathbf{\Omega_t} = Var(\varepsilon_t)$ is the reduced-form covariance matrix.

Johansen (1991) shows that the reduced-form VECM model and the integrated VMA model are linked through the factorization:

$$\mathbf{\Psi}_t(1) = \begin{pmatrix} \psi_{1,t} & \psi_{2,t} \\ \psi_{1,t} & \psi_{2,t} \end{pmatrix} = \beta_\perp \mathbf{\Pi}_t \alpha'_{t\perp},$$

where $\beta_\perp$ is orthogonal to β, and $\mathbf{\Pi}_t = \left(\alpha'_{t\perp}\left(\mathbf{I} - \sum_{j=1}^{k} \mathbf{\Gamma}_{j,t}\right)\beta_\perp\right)^{-1}$ where $\mathbf{I}$ is the identity matrix and $\alpha_{t\perp}$ is orthogonal to α_t. Note that the time variation in ψ_{1t} and ψ_{2t} depends on the time variation in $\mathbf{\Pi}_t$ and $\alpha_{t\perp}$.

C.2 More Plots for Section 5.3

C.2.1 η^P_t *i.i.d.*, $\sigma^2_P = 1$, η^T_t *i.i.d.*, $\sigma^2_T = 10$, δ **constant**

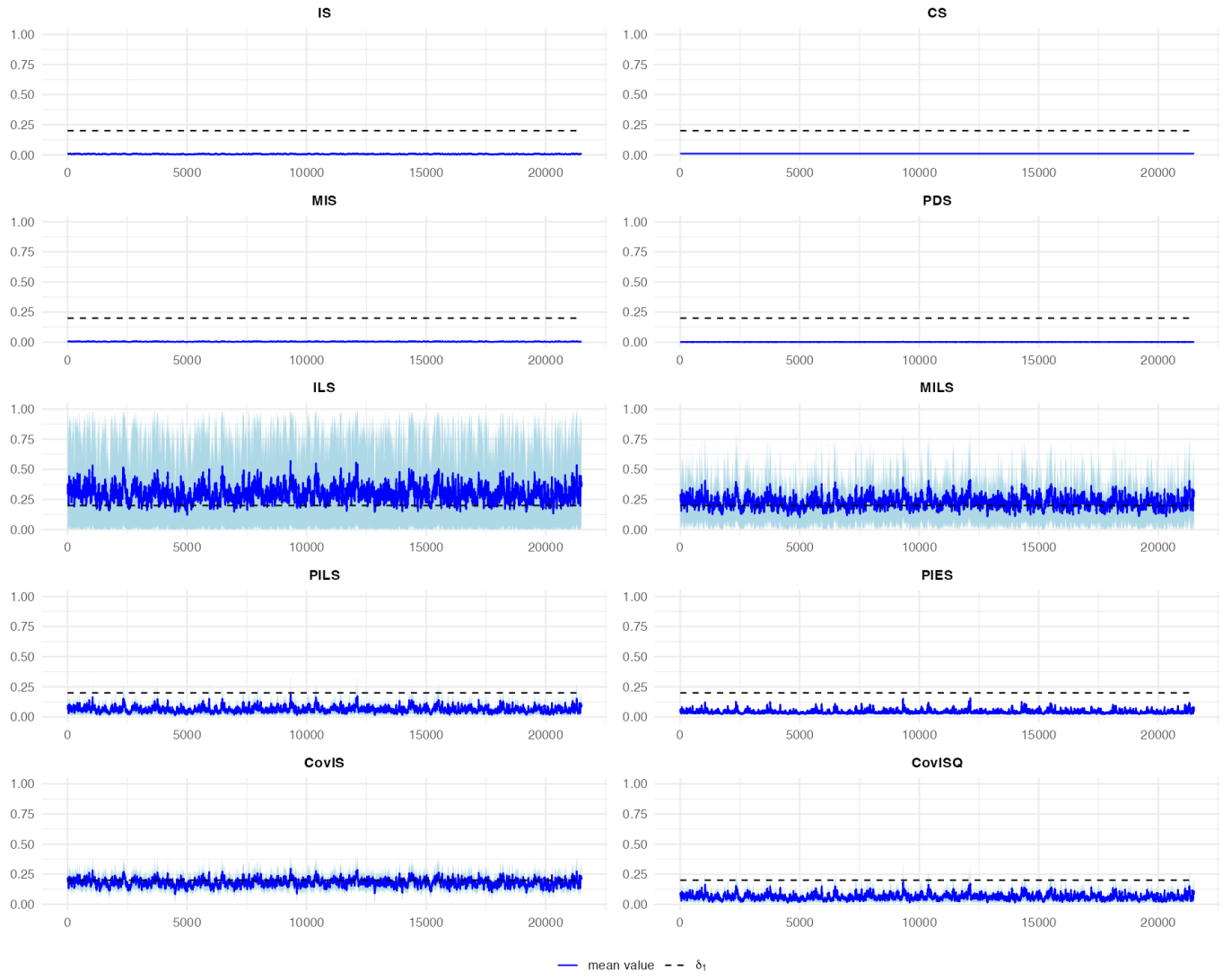

Figure C.1: price discovery measures of Panel A in Table A.3. $\delta = (0.2, 0.8)$.

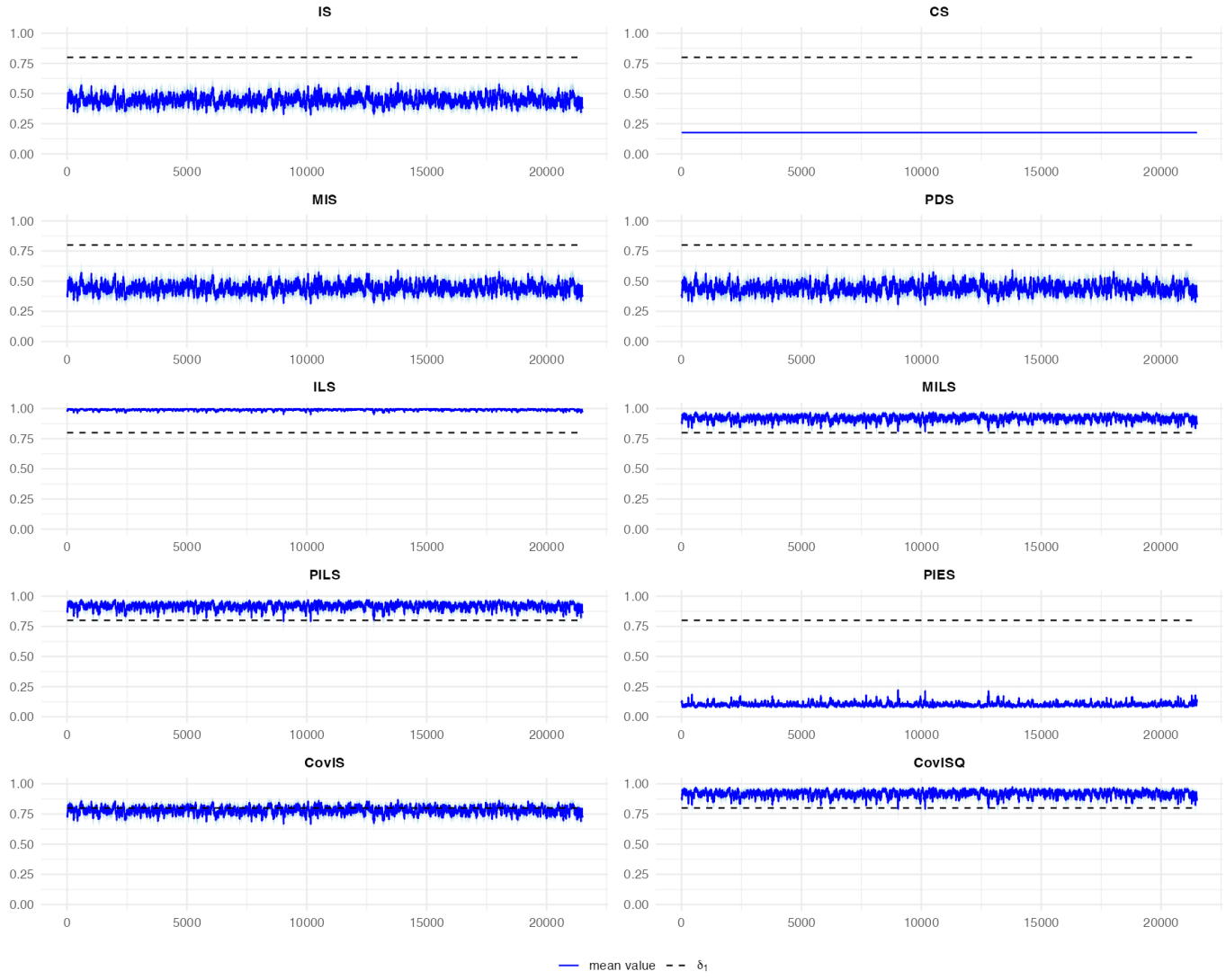

Figure C.2: price discovery measures of Panel A in Table A.3. $\delta = (0.8, 0.2)$.

C.2.2 η_t^P GARCH(1,1), η_t^T *i.i.d.*, $\sigma_T^2 = 10$, δ constant

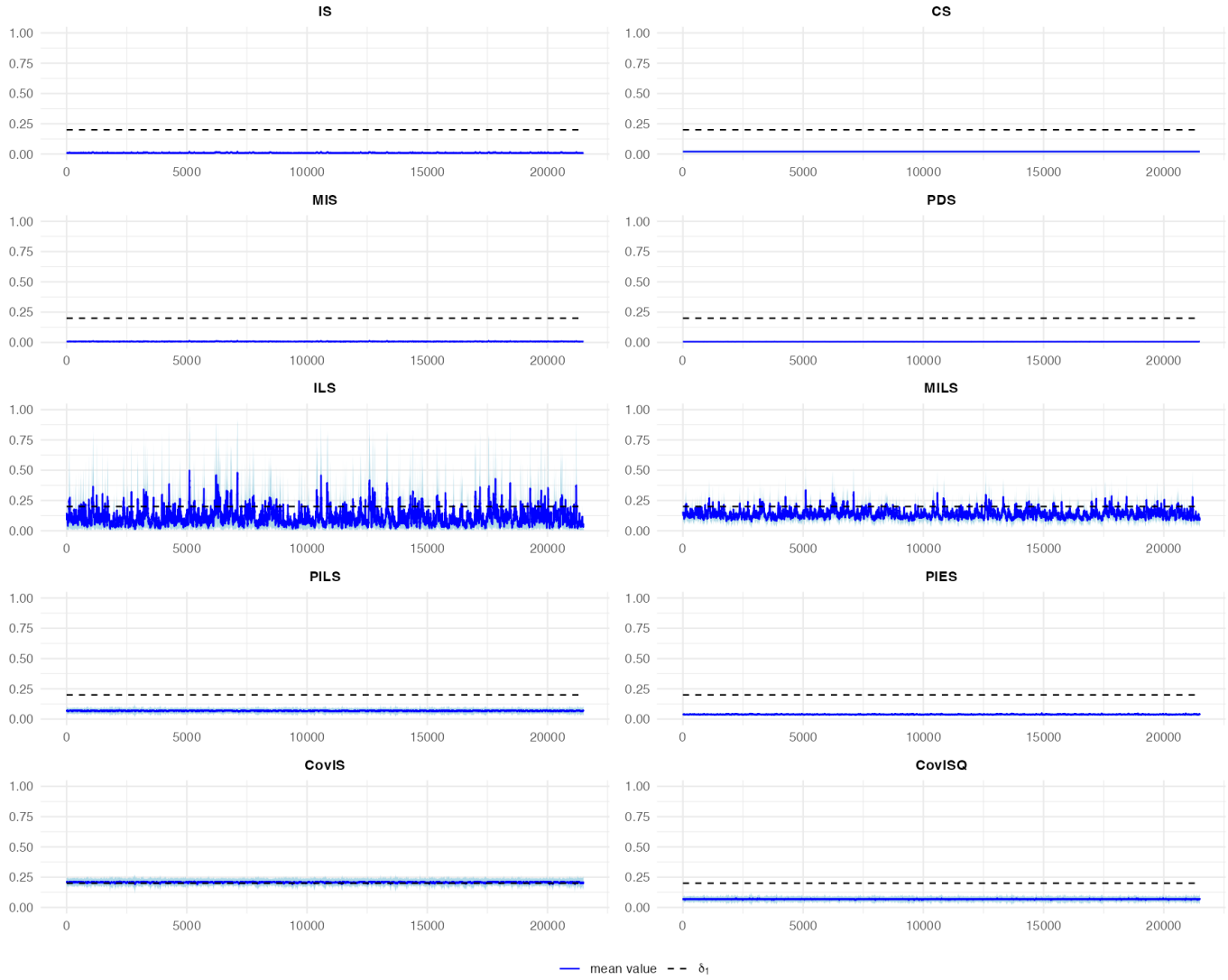

Figure C.3: price discovery measures of Panel B in Table A.3. $\delta = (0.8, 0.2)$.

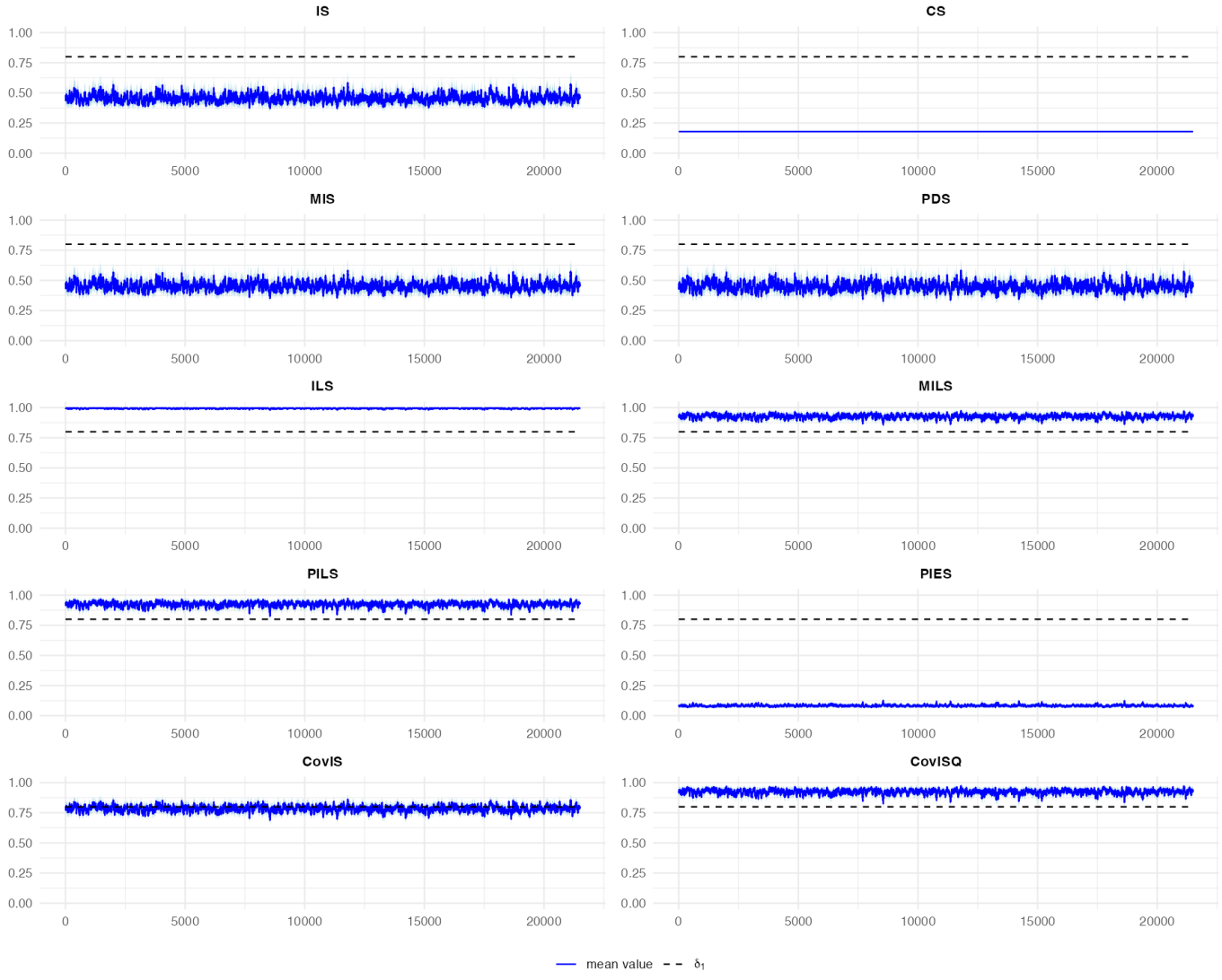

Figure C.4: price discovery measures of Panel B in Table A.3. $\delta = (0.8, 0.2)$.

C.2.3 η_t^P *i.i.d.*, $\sigma_T^2 = 10$, η_t^T GARCH(1,1), δ constant

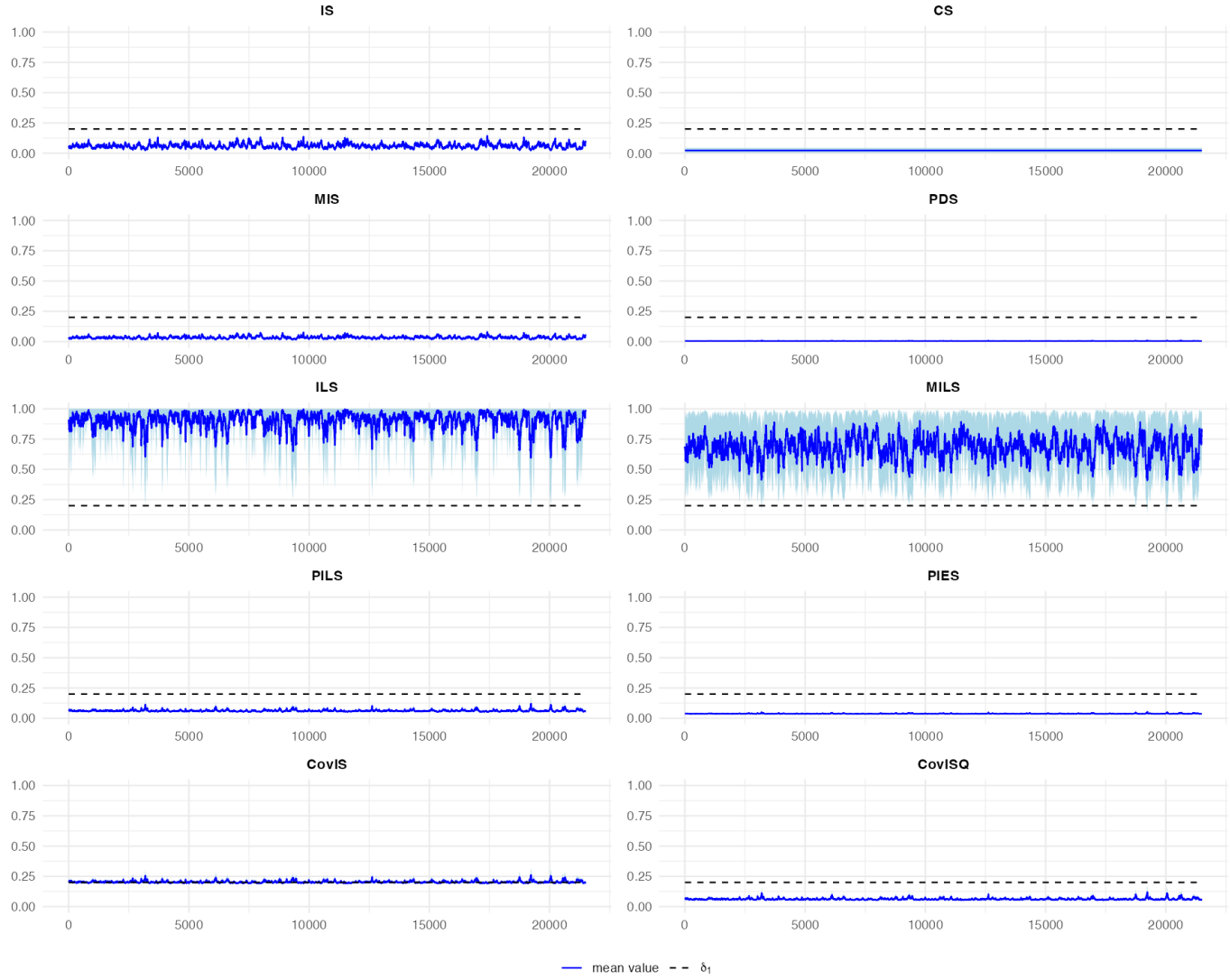

Figure C.5: price discovery measures of Panel C in Table 3.2. $\delta = (0.8, 0.2)$.

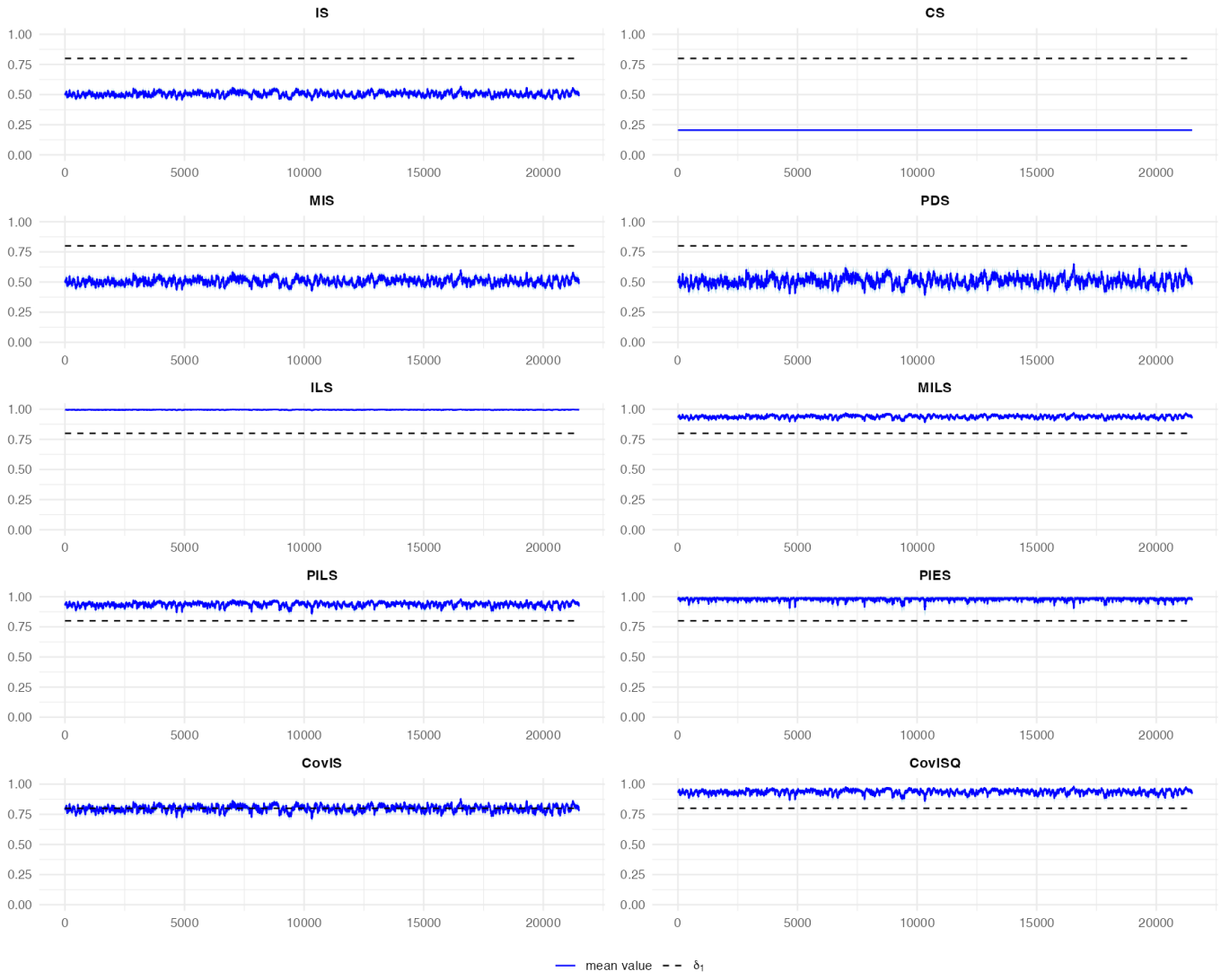

Figure C.6: price discovery measures of Panel C in Table 3.2. $\delta = (0.8, 0.2)$.

References

Aït-Sahalia, Y. and Jacod, J. (2014). *High-frequency financial econometrics*. Princeton University Press.

Ait-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *The review of financial studies*, 18(2):351–416.

Amihud, Y. and Mendelson, H. (1987). Trading mechanisms and stock returns: An empirical investigation. *The Journal of Finance*, 42(3):533–553.

Amisano, G. and Geweke, J. (2017). Prediction using several macroeconomic models. *Review of Economics and Statistics*, 99(5):912–925.

Andersen, T. G. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of empirical finance*, 4(2-3):115–158.

Arias, J. E., Rubio-Ramirez, J. F., and Shin, M. (2022). Macroeconomic forecasting and variable ordering in multivariate stochastic volatility models. *Journal of Econometrics*.

Avino, D., Lazar, E., and Varotto, S. (2015). Time varying price discovery. *Economics Letters*, 126:18–21.

Baillie, R. T., Booth, G. G., Tse, Y., and Zabotina, T. (2002). Price discovery and common factor models. *Journal of financial markets*, 5(3):309–321.

Bańbura, M., Giannone, D., and Lenza, M. (2015). Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections. *International Journal of forecasting*, 31(3):739–756.

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of applied Econometrics*, 25(1):71–92.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536.

Baumeister, C. and Peersman, G. (2013). Time-varying effects of oil supply shocks on the us economy. *American Economic Journal: Macroeconomics*, 5(4):1–28.

Bauwens, L., Lubrano, M., and Richard, J.-F. (2000). *Bayesian inference in dynamic econometric models*. OuP Oxford.

Ben-David, I., Franzoni, F., and Moussawi, R. (2014). Do etfs increase volatility? Technical report, National Bureau of Economic Research.

Berger, J. O. and Pericchi, L. R. (1996). The intrinsic bayes factor for model selection and prediction. *Journal of the American Statistical Association*, 91(433):109–122.

Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Journal of Business & Economic Statistics*, 19(4):465–474.

Bierens, H. J. and Martins, L. F. (2010). Time-varying cointegration. *Econometric Theory*, 26(5):1453–1490.

Bitto, A. and Frühwirth-Schnatter, S. (2019). Achieving shrinkage in a time-varying parameter model framework. *Journal of Econometrics*, 210(1):75–97.

Booth, G. G., So, R. W., and Tse, Y. (1999). Price discovery in the german equity index derivatives markets. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 19(6):619–643.

Brogaard, J., Hendershott, T., and Riordan, R. (2014). High-frequency trading and price discovery. *The Review of Financial Studies*, 27(8):2267–2306.

Caporale, G. M. and Girardi, A. (2013). Price discovery and trade fragmentation in a multi-market environment: Evidence from the mts system. *Journal of Banking & Finance*, 37(2):227–240.

Carriero, A., Chan, J., Clark, T. E., and Marcellino, M. (2022). Corrigendum to "large bayesian vector autoregressions with stochastic volatility and non-conjugate priors"[j. econometrics 212 (1)(2019) 137–154]. *Journal of Econometrics*, 227(2):506–512.

Carriero, A., Clark, T. E., and Marcellino, M. (2015). Bayesian vars: specification choices and forecast accuracy. *Journal of Applied Econometrics*, 30(1):46–73.

Carriero, A., Clark, T. E., and Marcellino, M. (2018). Measuring uncertainty and its impact on the economy. *Review of Economics and Statistics*, 100(5):799–815.

Carriero, A., Clark, T. E., and Marcellino, M. (2019). Large bayesian vector autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics*, 212(1):137–154.

Carriero, A., Kapetanios, G., and Marcellino, M. (2012). Forecasting government bond yields with large bayesian vector autoregressions. *Journal of Banking & Finance*, 36(7):2026–2047.

Chakravarty, S., Gulen, H., and Mayhew, S. (2004). Informed trading in stock and option markets. *The Journal of Finance*, 59(3):1235–1257.

Chan, J. C. (2022). Large hybrid time-varying parameter vars. *Journal of Business & Economic Statistics*, pages 1–16.

Chan, J. C., Eisenstat, E., and Strachan, R. W. (2020). Reducing the state space dimension in a large tvp-var. *Journal of Econometrics*, 218(1):105–118.

Chan, J. C. and Jeliazkov, I. (2009). Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(1-2):101–120.

Chan, J. C., Koop, G., and Yu, X. (2021). Large order-invariant bayesian vars with stochastic volatility. *arXiv preprint arXiv:2111.07225*.

Chan, J. C., Koop, G., and Yu, X. (2024). Large order-invariant bayesian vars with stochastic volatility. *Journal of Business & Economic Statistics*, 42(2):825–837.

Chib, S. (1995). Marginal likelihood from the gibbs output. *Journal of the american statistical association*, 90(432):1313–1321.

Chu, Q. C., Hsieh, W.-l. G., and Tse, Y. (1999). Price discovery on the s&p 500 index markets: An analysis of spot index, index futures, and spdrs. *International Review of Financial Analysis*, 8(1):21–34.

Clark, T. E. (2011). Real-time density forecasts from bayesian vector autoregressions with stochastic volatility. *Journal of Business & Economic Statistics*, 29(3):327–341.

Clark, T. E. and Ravazzolo, F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, 30(4):551–575.

Cogley, T. and Sargent, T. J. (2001). Evolving post-world war ii us inflation dynamics. *NBER macroeconomics annual*, 16:331–373.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post wwii us. *Review of Economic dynamics*, 8(2):262–302.

Conlon, T., Corbet, S., Hou, G., Hu, Y., and Oxley, L. (2022). Beyond the noise: Information discovery in bitcoin revisited. *Available at SSRN 4649154*.

Cross, J. and Poon, A. (2016). Forecasting structural change and fat-tailed events in australian macroeconomic variables. *Economic Modelling*, 58:34–51.

D'Agostino, A., Gambetti, L., and Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of applied econometrics*, 28(1):82–101.

De Jong, F. (2002). Measures of contributions to price discovery: A comparison. *Journal of Financial markets*, 5(3):323–327.

De Mol, C., Giannone, D., and Reichlin, L. (2008). Forecasting using a large number of predictors: Is bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, 146(2):318–328.

Del Negro, M. and Primiceri, G. E. (2015). Time varying structural vector autoregressions and monetary policy: a corrigendum. *The review of economic studies*, 82(4):1342–1345.

Dickey, J. M. (1971). The weighted likelihood ratio, linear hypotheses on normal location parameters. *The Annals of Mathematical Statistics*, pages 204–223.

Doan, T., Litterman, R., and Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric reviews*, 3(1):1–100.

Dufour, A. and Nguyen, M. (2008). Time-varying price discovery in the european treasury markets.

Epps, T. W. (1979). Comovements in stock prices in the very short run. *Journal of the American Statistical Association*, 74(366a):291–298.

Fang, F., Ventre, C., Basios, M., Kanthan, L., Martinez-Rego, D., Wu, F., and Li, L. (2022). Cryptocurrency trading: a comprehensive survey. *Financial Innovation*, 8(1):13.

Fang, Y. and Sanger, G. C. (2011). Index price discovery in the cash market. In *Midwest Finance Association 2012 Annual Meetings Paper*.

Frijns, B., Gilbert, A., and Tourani-Rad, A. (2015). The determinants of price discovery: Evidence from us-canadian cross-listed shares. *Journal of Banking & Finance*, 59:457–468.

Frühwirth-Schnatter, S. and Wagner, H. (2010). Stochastic model specification search for gaussian and partial non-gaussian state space models. *Journal of Econometrics*, 154(1):85–100.

Gefang, D. (2014). Bayesian doubly adaptive elastic-net lasso for var shrinkage. *International Journal of Forecasting*, 30(1):1–11.

Gerlach, R., Carter, C., and Kohn, R. (2000). Efficient bayesian inference for dynamic mixture models. *Journal of the American Statistical Association*, 95(451):819–828.

Geweke, J. (1996). Bayesian reduced rank regression in econometrics. *Journal of econometrics*, 75(1):121–146.

Geweke, J. and Amisano, G. (2010). Comparing and evaluating bayesian predictive distributions of asset returns. *International Journal of Forecasting*, 26(2):216–230.

Giannone, D., Lenza, M., Momferatou, D., and Onorante, L. (2014). Short-term inflation projections: A bayesian vector autoregressive approach. *International journal of forecasting*, 30(3):635–644.

Giannone, D., Lenza, M., and Primiceri, G. E. (2015). Prior selection for vector autoregressions. *Review of Economics and Statistics*, 97(2):436–451.

Giudici, P. and Abu Hashish, I. (2020). A hidden markov model to detect regime changes in cryptoasset markets. *Quality and Reliability Engineering International*, 36(6):2057–2065.

Gonzalo, J. and Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics*, 13(1):27–35.

Grammig, J. and Peter, F. J. (2013). Telltale tails: A new approach to estimating unique market information shares. *Journal of Financial and Quantitative Analysis*, 48(2):459–488.

Harris, F. H. d., McInish, T. H., and Wood, R. A. (2002). Security price adjustment across exchanges: an investigation of common factor components for dow stocks. *Journal of financial markets*, 5(3):277–308.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The journal of Finance*, 50(4):1175–1199.

Hasbrouck, J. (2003). Intraday price formation in us equity index markets. *The Journal of Finance*, 58(6):2375–2400.

Hasbrouck, J. (2007). *Empirical market microstructure: The institutions, economics, and econometrics of securities trading*. Oxford University Press.

Hasbrouck, J. (2021). Price discovery in high resolution. *Journal of Financial Econometrics*, 19(3):395–430.

Hasbrouck, J. and Ho, T. S. (1987). Order arrival, quote behavior, and the return-generating process. *The Journal of Finance*, 42(4):1035–1048.

Hasbrouck, J. and Seppi, D. J. (2001). Common factors in prices, order flows, and liquidity. *Journal of financial Economics*, 59(3):383–411.

Hauptfleisch, M., Putniņš, T. J., and Lucey, B. (2016). Who sets the price of gold? london or new york. *Journal of Futures Markets*, 36(6):564–586.

Hayashi, T. and Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, 11(2):359–379.

Hendershott, T., Jones, C. M., and Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *The Journal of finance*, 66(1):1–33.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: journal of the Econometric Society*, pages 1551–1580.

Jore, A. S., Mitchell, J., and Vahey, S. P. (2010). Combining forecast densities from vars with uncertain instabilities. *Journal of Applied Econometrics*, 25(4):621–634.

Kadiyala, K. R. and Karlsson, S. (1997). Numerical methods for estimation and inference in bayesian var-models. *Journal of Applied Econometrics*, 12(2):99–132.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the american statistical association*, 90(430):773–795.

Kim, J. and Linn, S. C. (2022). Price discovery under model uncertainty. *Energy Economics*, 107:105833.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The review of economic studies*, 65(3):361–393.

Koop, G. (2017). Bayesian methods for empirical macroeconomics with big data. *Review of Economic Analysis*, 9(1):33–56.

Koop, G. and Korobilis, D. (2013). Large time-varying parameter vars. *Journal of Econometrics*, 177(2):185–198.

Koop, G. and Korobilis, D. (2016). Model uncertainty in panel vector autoregressive models. *European Economic Review*, 81:115–131.

Koop, G. and Korobilis, D. (2019). Forecasting with high-dimensional panel vars. *Oxford Bulletin of Economics and Statistics*, 81(5):937–959.

Koop, G., León-González, R., and Strachan, R. W. (2009). Efficient posterior simulation for cointegrated models with priors on the cointegration space. *Econometric Reviews*, 29(2):224–242.

Korobilis, D. (2013). Var forecasting using bayesian variable selection. *Journal of Applied Econometrics*, 28(2):204–230.

Kruschke, J. K. (2018). Rejecting or accepting parameter values in bayesian estimation. *Advances in methods and practices in psychological science*, 1(2):270–280.

Kuck, K. and Schweikert, K. (2023). Price discovery in equity markets: A state-dependent analysis of spot and futures markets. *Journal of Banking & Finance*, 149:106808.

Lautier, D., Ling, J., and Villeneuve, B. (2024). Rediscovering price discovery. *Available at SSRN 4972062*.

Lien, D., Roseman, B., and Shi, Y. (2025). A new leadership share measure for price discovery. *Journal of Banking & Finance*, page 107527.

Lien, D. and Shrestha, K. (2009). A new information share measure. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 29(4):377–395.

Litterman, R. B. (1980). *Bayesian procedure for forecasting with vector autoregressions*. Massachusetts Institute of Technology.

Litterman, R. B. (1986). Forecasting with bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics*, 4(1):25–38.

Litterman, R. B. et al. (1979). Techniques of forecasting using vector autoregressions. Technical report.

Longin, F. and Solnik, B. (2001). Extreme correlation of international equity markets. *The journal of finance*, 56(2):649–676.

Lucas Jr, R. E. (1976). Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy*, volume 1, pages 19–46. North-Holland.

Madhavan, A., Richardson, M., and Roomans, M. (1997). Why do security prices change? a transaction-level analysis of nyse stocks. *The Review of Financial Studies*, 10(4):1035–1064.

Madhavan, A. and Smidt, S. (1991). A bayesian model of intraday specialist pricing. *Journal of Financial Economics*, 30(1):99–134.

Marshall, B. R., Nguyen, N. H., and Visaltanachoti, N. (2013). Etf arbitrage: Intraday evidence. *Journal of Banking & Finance*, 37(9):3486–3498.

McCracken, M. and Ng, S. (2020). Fred-qd: A quarterly database for macroeconomic research. Technical report, National Bureau of Economic Research.

Mohamad, A. and Inani, S. K. (2023). Price discovery in bitcoin spot or futures during the covid-19 pandemic? evidence from the time-varying parameter vector autoregressive model with stochastic volatility. *Applied Economics Letters*, 30(19):2749–2757.

Narayan, P. K. and Sharma, S. S. (2018). An analysis of time-varying commodity market price discovery. *International Review of Financial Analysis*, 57:122–133.

Ozturk, S. R., Van der Wel, M., and van Dijk, D. (2017). Intraday price discovery in fragmented markets. *Journal of Financial Markets*, 32:28–48.

Patel, V., Putniņš, T. J., Michayluk, D., and Foley, S. (2020). Price discovery in stock and options markets. *Journal of Financial Markets*, 47:100524.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, 72(3):821–852.

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

Qiao, G., Zhao, P., and Li, W. (2019). Time varying price discovery of the new third board market in china: does the market-making system help? *Applied Economics*, 51(45):4902–4919.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The annals of mathematical statistics*, 23(3):470–472.

Sapp, S. G. (2002). Price leadership in the spot foreign exchange market. *Journal of Financial and Quantitative Analysis*, 37(3):425–448.

Schweikert, K. (2021). Bootstrap confidence intervals and hypothesis testing for market information shares. *Journal of Financial Econometrics*, 19(5):934–959.

Schweikert, K. (2026). Asymptotic inference for hasbrouck information shares. *Economics Letters*, 258:112756.

Segnon, M. and Bekiros, S. (2020). Forecasting volatility in bitcoin market. *Annals of Finance*, 16(3):435–462.

Shen, S., Sultan, S. G., and Zivot, E. (2024a). Price discovery share: An order invariant measure of price discovery. *Finance Research Letters*, 67:105734.

Shen, S., Zhang, Y., and Zivot, E. (2024b). Improving price leadership share for measuring price discovery. *Available at SSRN 4759845*.

Shen, S., Zhang, Y., and Zivot, E. (2025). Improving information leadership share for measuring price discovery. *Journal of Empirical Finance*, page 101638.

Shen, S. and Zivot, E. (2024). In defense of information leadership share: A response to shrestha and lee (2023). *Available at SSRN 4752194*.

Shin, M. and Zhong, M. (2020). A new approach to identifying the real effects of uncertainty shocks. *Journal of Business & Economic Statistics*, 38(2):367–379.

Sims, C. A. (1993). A nine-variable probabilistic macroeconomic forecasting model. In *Business cycles, indicators, and forecasting*, pages 179–212. University of Chicago press.

Sims, C. A. and Zha, T. (1998). Bayesian methods for dynamic multivariate models. *International Economic Review*, pages 949–968.

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American economic review*, 97(3):586–606.

Smith, J. (1985). Diagnostic checks of non-standard time series models. *Journal of Forecasting*, 4(3):283–291.

Stock, J. H. and Watson, M. W. (2002). Has the business cycle changed and why? *NBER macroeconomics annual*, 17:159–218.

Stock, J. H. and Watson, M. W. (2008). Phillips curve inflation forecasts.

Strachan, R. W. (2003). Valid bayesian estimation of the cointegrating error correction model. *Journal of Business & Economic Statistics*, 21(1):185–195.

Strachan, R. W. and Inder, B. (2004). Bayesian analysis of the error correction model. *Journal of Econometrics*, 123(2):307–325.

Sultan, S. G. and Zivot, E. (2015). Price discovery share-an order invariant measure of price discovery with application to exchange-traded funds. *University of Washington*.

Taylor, N. (2011). Time-varying price discovery in fragmented markets. *Applied Financial Economics*, 21(10):717–734.

Theissen, E. (2016). Price discovery in spot and futures markets: A reconsideration. In *High Frequency Trading and Limit Order Book Dynamics*, pages 249–268. Routledge.

Verdinelli, I. and Wasserman, L. (1995). Computing bayes factors using a generalization of the savage-dickey density ratio. *Journal of the American Statistical Association*, 90(430):614–618.

Villani, M. (2005). Bayesian reference analysis of cointegration. *Econometric Theory*, 21(2):326–357.

Wallace, D., Kalev, P. S., and Lian, G. (2019). The evolution of price discovery in us equity and derivatives markets. *Journal of Futures Markets*, 39(9):1122–1136.

Yan, B. and Zivot, E. (2007). The dynamics of price discovery. In *Afa 2005 philadelphia meetings*.

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

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411.