Journal of Empirical Finance 23 (2013) 68–83

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

What do price discovery metrics really measure?

Tālis J. Putniņš *,1

University of Technology Sydney, PO Box 123 Broadway, NSW 2007, Australia
Stockholm School of Economics in Riga, Strelnieku iela 4a, Riga LV-1010, Latvia

ARTICLE INFO

Article history:
Received 24 October 2012
Received in revised form 6 May 2013
Accepted 15 May 2013
Available online 24 May 2013

JEL classification:
G14

Keywords:
Price discovery
Information share
Component share
Information leadership share

ABSTRACT

A market is typically considered to dominate price discovery if it is the first to reflect new information about the fundamental value. Our simulations indicate that common price discovery metrics – Hasbrouck information share and Harris–McInish–Wood component share – are only consistent with this view of price discovery if the price series have equal levels of noise, including microstructure frictions and liquidity. If the noise in the price series differs, the information and component shares measure a combination of leadership in impounding new information and relative avoidance of noise, to varying degrees. A third price discovery metric, the 'information leadership share' uses the information share and the component share together to identify the price series that is first to impound new information. This third metric is robust to differences in noise levels and therefore correctly attributes price discovery in a wider range of settings. Using four recent empirical studies of price discovery we show that the choice and interpretation of price discovery metrics can have a substantial impact on conclusions about price discovery.

1. Introduction

Price discovery, a fundamentally important role of secondary markets, is the "efficient and timely incorporation of the information implicit in investor trading into market prices" (Lehmann, 2002, p. 259). When multiple price series are related via a common asset (e.g., a stock trading in multiple venues, order flow for one security from different market participants, and different derivative securities linked to the same underlying asset) the contribution of a price series to price discovery is typically considered to be the extent to which it is the first to reflect new information about the 'true' underlying asset value. There are two main (occasionally competing) empirical measures of the contribution of different price series to price discovery: Hasbrouck information share (*IS*) and Harris–McInish–Wood component share (*CS*).[2] Market microstructure scholars have made substantial progress in reconciling and understanding the two measures (e.g., Baillie et al., 2002; Lehmann, 2002). In a recent and notable contribution to this effort, Yan and Zivot (2010) analytically show how to interpret *IS* and *CS* in the context of a structural cointegration model.

Despite the substantial progress made to date, there is still not a consensus in practice on what each metric really measures. Consequently, the approach to applying and interpreting the metrics in different applications is far from consistent.[3] Furthermore, Yan and Zivot (2010) propose combining *IS* and *CS* to specifically measure impounding of new information, and the usefulness of this technique is yet to be thoroughly tested.

* University of Technology Sydney, PO Box 123 Broadway, NSW 2007, Australia. Tel.: +61 2 95144022.
E-mail address: talis.putnins@uts.edu.au.

[1] The author is grateful to three anonymous referees for their helpful comments.

[2] For a good review of *IS* and *CS* and how they are related see Baillie et al. (2002), Lehmann (2002) and Yan and Zivot (2010). The next section provides a brief overview of these metrics. We refer to *CS* as the Harris–McInish–Wood component share due to their role in popularizing this measure of price discovery, but we note that others were involved in pioneering the use of this metric to measure price discovery (e.g., Booth et al., 1999; Chu et al., 1999).

[3] For example, see recent studies such as Chen and Gau (2010), Frijns et al. (2010), Korczak and Phylaktis (2010), Anand et al. (2011), Fricke and Menkhoff (2011), Liu and An (2011), Chen and Chung (2012), Chen and Sub Choi (2012), Rittler (2012), and Chen et al. (2013).

0927-5398/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jempfin.2013.05.004

This article aims to illustrate what each of the price discovery metrics really measures, in the hope of promoting a more consistent approach to their use and interpretation. We do this by estimating each price discovery metric on simulated data for a simple structural model, choosing parameter values and sample size to mimic an empirical research setting. A key feature of the model is that it allows price series to differ in two important ways: (i) noise (e.g., microstructure frictions such as tick discreteness and bid–ask bounce); and (ii) speed of adjustment to new information. The interaction of these two dimensions is the source of much of the confusion and debate about the different price discovery metrics. By varying each of these two characteristics independently we obtain a two-dimensional grid of parameters describing the relations between two price series, i.e., relative to the reference price series, a price series of interest can be: (i) noisier and slower; (ii) noisier and faster; (iii) less noisy and slower; or (iv) less noisy and faster. By estimating all of the price discovery metrics for each of these cases (and many values in between) we demonstrate what each metric actually measures and thus how the price discovery metrics should be interpreted with respect to the relative amount of noise and speed of incorporation of information.

We find that *IS* and *CS* only accurately measure contribution to price discovery, in the conventional sense of being the first to reflect innovations in the fundamental value, when the price series being compared have a similar level of noise. When the levels of noise differ both *IS* and *CS* measure to varying extents: (i) the relative speed at which a price impounds new information; and (ii) the relative avoidance of noise. That is, price series that reflect new information faster will tend to have higher *IS* and *CS*, but so too will price series that are less noisy, holding constant the speed with which information is impounded. When the difference in noise levels is sufficiently large relative to the differences in the speed at which information is impounded, conventional interpretations of *IS* and *CS* can lead to incorrect conclusions about which price series incorporates information first. For example, in a two-market setting *IS* and *CS* could be larger (greater than 50%) for the market that is *slower* to incorporate new information as long as it is sufficiently less noisy. Our simulations indicate that *IS* places greater weight on the speed at which a price series impounds information compared to the *CS* metric.

In contrast, an adaptation of an expression derived by Yan and Zivot (2010), which we call the 'informational leadership share' (*ILS*), is able to correctly attribute contributions to price discovery in the presence of different levels of noise. This metric, which is only applicable in bivariate systems, uses *IS* and *CS* together to identify the price series that leads the process of impounding new information.

We illustrate the practical importance of the choice and interpretation of the price discovery metric using four recent empirical price discovery studies. For the four studies we calculate *ILS* and compare the conclusions from this metric with those made by the studies. The results are striking: in many cases *ILS* leads to opposite conclusions to those made by the studies. For example, Rittler (2012) concludes that futures contracts dominate price discovery in the EU emission trading market and that their dominance has increased through time, whereas *ILS* indicates the opposite: the spot market leads price discovery overall, and the contribution of the futures market to price discovery is in fact declining through time. The differences in conclusions stem from the failure of many empirical price discovery studies to account for the way *IS* and *CS* are affected by differences in noise levels across the price series of interest. We discuss several implications of our findings for existing and future research.

The present paper complements Yan and Zivot (2010) by using Monte Carlo simulations to illustrate the insights that Yan and Zivot provide analytically, and also provides some extensions and novel insights. We design our simulations and choose parameter values to mimic an empirical setting in order to demonstrate the relevance of the analytic results for empirical work. Another reason for turning to simulations is that the structural cointegration specifications implied by our models of price formation are not invertible and therefore we cannot perform the analytic exercise that is done by Yan and Zivot. An advantage of our structural models, in addition to being simple and intuitive, is that each of the two price series has an independent source of noise (trading frictions) rather than sharing one transitory shock, which we argue is a realistic feature in many price discovery applications. A further advantage is that we are able to provide evidence on how Yan and Zivot's results and measure of informational leadership hold up in a model of price formation that does not satisfy the assumptions under which Yan and Zivot derived their results. Interestingly, the measure of informational leadership reliably attributes price discovery in our setting, suggesting that it is also useful in at least some models of price formation that do not satisfy Yan and Zivot's assumptions. We do not know whether this holds for other models of price formation. We also document some differences from Yan and Zivot's analytic results. For example, our simulations indicate that *CS* measures to some extent the relative speed at which a price series impounds new information, not just the relative avoidance of noise as documented by Yan and Zivot.

This paper also provides a modification of the price discovery metric proposed by Yan and Zivot to make it less susceptible to extreme values and make it easily comparable to the well-established *IS* and *CS* metrics. The result is the 'information leadership share' (*ILS*). Importantly, we provide the first tests of the performance of this new metric. Given that our simulations highlight some very attractive features of *ILS*, particularly compared to existing metrics, we hope it will be used in future empirical work. Finally, we delve deeper into the implications for existing and future research, of both Yan and Zivot and the present paper's findings about what price discovery metrics really measure.

2. What do researchers mean by "price discovery"?

2.1. The economic process of price discovery

Part of the confusion and debate about alternative price discovery metrics stems from differences and lack of precision in definitions of price discovery. Here, we refer to definitions of the economic process rather than the empirical metrics, which we review in the next subsection. In order to correctly interpret the price discovery metrics it is essential to distinguish between