Han, Q. and Zhao, C. and Chen, J. and Guo, Christine (2025) Does asynchronous market update matter? Re-examining the price discovery of stock index and futures in China. Emerging Markets Review 67 (101307), ISSN 1566-0141.

Emerging Markets Review 67 (2025) 101307

Contents lists available at ScienceDirect

Emerging Markets Review

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

Does asynchronous market update matter? Re-examining the price discovery of stock index and futures in China

Qian Han [a], Chengzhi Zhao [b], Jing Chen [c,*], Qian Guo [d]

[a] Lingnan College, Sun Yat-sen University, Guangzhou 510275, China
[b] School of Mathematics and Statistics, Guangxi Normal University, Guilin 541004, China
[c] Cardiff University, Senghennydd Road, Cardiff CF24 4AG, UK
[d] Birkbeck College, University of London, London WC1E 7HU, UK

A R T I C L E I N F O

JEL:
G10
G13
G14

Keywords:
Price discovery
Frequency of market update
Chinese stock index futures
Lead-lag relationship
Microstructure noise

A B S T R A C T

Uniquely addressing asynchronous informational update between index and futures, we find that reduction in data frequency depicts a dual effect of "noise reduction" and "speed reduction" on Hasbrouck's (1995) information share (IS) and Gonzalo-Granger's (1995) component share (CS) indicators. Furthermore, the "noise reduction" effect does not exist significantly on CS, thereby preventing Putniņš's (2013) information leading share (ILS) indicator from eliminating noise under low-frequency data. Our novel leading time (LT) indicator suggests that the Shanghai-Shenzhen Stock Exchange 300 (CSI 300) and China Stock Exchange 500 (CSI 500) futures dominate price discovery. An asynchronous informational update overestimates the price discovery ability of futures.

1. Introduction

Price discovery is the process by which information is incorporated into asset prices (Hasbrouck, 1995). Prior literature confirmed that, due to lower transaction costs and more flexible trading mechanisms, stock index futures generally dominate price discovery, and reflect changes in fundamental information faster than spot (Stoll and Whaley, 1990). Primarily comprising the Shanghai Stock Exchange 50 (SSE 50), China Stock Exchange 500 (CSI 500), and Shanghai-Shenzhen Stock Exchange 300 (CSI 300), the Chinese stock index futures market performed well in terms of price discovery since its launch (Guo et al., 2013). However, there exist certain misunderstandings about the futures price discovery function among retail investors in China, especially under extreme market conditions. During the Chinese stock market crash in 2015, the stock index futures became one of the scapegoats. Many retail investors believe that the decline in futures caused the decline of the spot index, and the futures price discovery led to the stock market crash (Hao et al., 2019). Apparently, these investors misunderstood price discovery by incorrectly interpreting the lead-lag relationship between markets as a causal relationship.

* Corresponding author.
E-mail address: ChenJ60@cardiff.ac.uk (J. Chen).

https://doi.org/10.1016/j.ememar.2025.101307
Received 28 December 2023; Received in revised form 23 January 2025; Accepted 12 May 2025
Available online 13 May 2025

It is worth noting that, due to technical reasons, the Chinese stock (futures) exchange provides real-time data feeds for the index (futures) every 5 (0.5) seconds.[1] This implies that the futures appear ten times faster than the spot in providing a timely market update. Inevitably, such a difference may lead traders to believe that futures lead spot, and the academic literature depicts plenty of instances where futures usually reflect changes in fundamental information faster than spot (see, for example, Stoll and Whaley, 1990; Brooks et al., 2001; Tse and Chan, 2010; Lien and Shrestha, 2009; Gong et al., 2016; Da Fonseca and Zaatour, 2017; Jiang et al., 2019). Nonetheless, in China, the root of this phenomenon is twofold. First, the futures themselves respond to information more quickly and sensitively than the spot (Chu et al., 1999). The second aspect emerges from the technical aspect of futures providing a timely market update much faster than the spot. During a market crash where both futures and spot fall rapidly at the same time, if one fails to distinguish between the two fundamental causes, the asynchrony in market price updates may exaggerate the lead-lag relationship between futures and spot. In other words, the asynchrony in market price updates may reinforce investors' sense that futures are falling faster than spot, which in turn generates the erroneous belief that price discovery of stock index futures causes the decline in the spot market. Thus, the futures price discovery may be overestimated, especially with higher-frequency data. Additionally, the relatively slow frequency of spot market updates reduces the information efficiency of the spot market.

To what extent could such an asynchrony in the speed of market updates exaggerates the lead-lag relationship between futures and spot, or overestimate the information share (IS) of futures? Stoll and Whaley (1990) studied the lead-lag relationship between S&P 500 index futures, MMI index futures and the corresponding stock market index. They noted that some of the delays in the computation and reporting of the stock index were due to purely technical problems, and such delays would tend to show futures market returns leading stock index returns. Also, existing literature focusing on the market update on the spot suggests that an increase (decrease) in tick sizes is associated with decreases (increases) in the price efficiency of stocks, indicating that price efficiency decreases when the smallest price increment by which the stock prices are quoted increases (i.e., the market update on the spot decelerates) (Chung and Chuwonganant, 2023). Whilst the study shed some light on the change of frequency in reporting of the spot market, no prior studies have carried out any systematic analyses linking the relative time in reporting spot and futures index values and its impact on the price discovery of underlying financial assets.

Since both futures and spot prices are updated within 5 s, we must use data with a frequency higher than 5 s (e.g. 1 s) to investigate the effect of market asynchrony on price discovery between the Chinese stock index and index futures. When using such high-frequency data, the market noise implied by the data needs to be carefully considered, as noise can have a significant impact on price discovery measures. The most widely used price discovery metrics include Hasbrouck's (1995) information share (IS) and Gonzalo-Granger's (1995) component share (CS). It is imperative to emphasize that both IS and CS can be influenced by data noise. Prior studies confirmed the existence of two components for both IS and CS, a noise avoidance component (i.e. higher if less noisy) and a velocity/speed component (i.e. higher if faster information reflection). The presence of greater noise leads to smaller values for these two indicators (Yan and Zivot, 2010; Putniņš, 2013; Wu and Ma, 2014). However, the overall price discovery function of the underlying asset may not be necessarily weak.

One way to mitigate the effects of noise is to use relatively low-frequency data (Brandvold et al., 2015; Hao et al., 2019), e.g. 5-min data. However, it can be difficult to capture the subtle price lead-lag relationship between markets using such low-frequency data. Hasbrouck (2021) confirmed that the price discovery differentials among markets that are indistinguishable at 1-s resolution can be revealed at a resolution of 10 microseconds, thus recommending the use of higher frequency data. There also exists robustness and reliability problems when calculating IS with low-frequency data. Due to price correlation, the upper and lower bounds of IS calculated by low-frequency data will be large. After averaging the upper and lower bounds (Baillie et al., 2002), the final value of IS will be close to 50%, thus tends to underestimate the price discovery differentials between markets. As pointed out by Hasbrouck (2002), the large gap between the upper and lower bounds of IS reflects insufficient identification. This gap is largely a function of the time resolution.

Another way to mitigate the effects of noise is to construct noise-robust price discovery indicators. Yan and Zivot (2010) confirmed that the noise effects can be effectively removed by combining IS and CS. Based on this idea, Putniņš (2013) proposed an information leading share (ILS) indicator, and confirmed through Monte Carlo simulation that ILS is effective in removing the effect of noise, making it a more reliable price discovery indicator than IS and CS indicators. Nevertheless, the ILS relies on the calculation of IS and only demonstrates optimal performance under certain conditions. Moreover, previous literature lacks sufficient analysis of the performance of ILS under different data frequencies. Gathering some low-frequency data analyses on price discovery (Chen and Gau, 2010; Hsieh et al., 2008; Rittler, 2012), Putniņš (2013) concluded that ILS is less than 50% where IS and CS are greater than 50%, thus the conclusions from IS and CS differ from those from ILS. Putniņš (2013) further explains that this is the advantage of ILS, since the effect of noise is removed. However, the study does not consider the insufficient identification problem in IS. Due to the low frequency of the data, the IS, when averaged, is lower than the CS value, leading to the ILS exceeding 50%, indicating a potential bias of ILS. Therefore, a systematic analysis of the performance of these price discovery metrics by different data frequencies is necessary. Our Monte Carlo simulations in section 3.1 demonstrate that the ILS tends to generate incorrect inferences at low-frequency data.

[1] An interesting contrast is documented in Stoll and Whaley (1990): "Prior to June 13, 1986, the [US] stock index was computed approximately once a minute; but, since that time, it has been computed and reported approximately four times per minute." It appears that the reporting and frequency of calculating spot prices are different between China and the US. More recent evidence from spglobal.com in 2023 reveals that the real-time index calculation (for example, S&P DJI) involves real-time exchange-traded prices and computes the intraday index on a pre-determined fixed interval (for example, every 5 s). At each fixed interval, the index calculates the latest real-time pricing for each underlying security included in the index. If a new price is not available since the last real-time calculation, the calculation leverages the last available traded price provided by the exchange (https://www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-equity-indices-policies-practices.pdf).

Based on the above discussion, this study employs 1-s data in conjunction with the ILS to analyze price discovery. However, the investors are more concerned about the lead-lag time difference between markets, not the ILS (or IS / CS) measurement, which is expressed as a percentage. It is vital that the price discovery indicators could directly measure the speed at which futures incorporate new information. Therefore, we propose a new concept of leading time (LT) and it is set out as follows. If the price discovery indicator of futures contracts is around 50%, futures and index incorporate new information almost at the same speed. In the case where futures lead the index, the calculated price discovery indicator of the futures will be greater than 50%. If the index time series are brought forward by a few seconds, the price discovery indicator will generally decrease. When the indicator drops below 50% for the first time, the time brought forward is defined as the LT of the futures. Using this concept, we further test the validity of LT indicators in analyzing the price discovery of futures and spot through a series of Monte Carlo simulations in section 3.2.

In the empirical part, we first construct real-time index quotes with a 1-s update frequency using the data underlying CSI 300 constituents and estimate the IS and LT of futures. This is followed by their comparison against the real index released every five seconds to estimate the extent to which the asynchrony in the speed of market updates exaggerates the lead-lag relationship between futures and spot. We also discuss how the regulatory policies on futures affect the lead-lag relationship between futures and spot.

We find that the CSI 300 stock index futures lead their underlying stock indices across the majority, 97.14% (92.14%, 71.03%) of trading days according to ILS (IS, CS); the average value of the futures price discovery indicator, ILS (IS, CS) is 91.38% (81.79%, 59.2%); and futures lead the index by an average of 32.23 (19.78, 9.76) seconds according to the ILS (IS, CS) calculation. On most trading days, the ILS (IS, CS) of futures is overvalued by an average of 2.47% (2.08%, 0.16%) due to asynchronous informational updates between the futures and spot markets. The average LT of stock index futures is overestimated by 3.26 (1.55, 0.22) seconds as based on ILS (IS, CS) calculations, which is not small considering that the index is updated every 5 s. Meanwhile, our results show that the restrictive policies in futures markets reduce the price discovery ability of futures, while deregulation has no significant impact on the price discovery of futures. The analyses using CSI 500 stock index futures provide similar conclusions.

This article serves as a valuable addition to the existing literature. Although previous literature has noted the issue of data frequency (Brandvold et al., 2015; Entrop et al., 2020), there is still a lack of systematic analysis of how data frequency affects price discovery indicators such as IS, CS, and especially ILS. Our theoretical and empirical findings confirm that price discovery indicators under different data frequencies can be very different and even contradictory, thus confirming that data frequency has a significant impact on these price discovery indicators. Our results help to explain the conflicting results of prior research, especially when IS, CS and ILS give different conclusions. It is also common in some studies that high-frequency data are not available. In such cases, understanding the sensitivity of different price discovery indicators to the frequency of data is needed, and the empirical results given by price discovery indicators need to be discussed more carefully.

Prior studies on price discovery of China's stock index futures employ either one- or five-minute data to calculate IS (Lin and Wang, 2018; Xu and Liu, 2019; Hao et al., 2019; He et al., 2020; Huang et al., 2021). This exhibits a large gap between the upper and lower bounds of IS. Miao et al. (2017) employ 5-s data, but use IS instead of ILS to measure price discovery. Evidenced by our Monte Carlo simulations results, we find that the same price discovery indicator, especially calculated from low-frequency data, is affected by data frequency and noise, producing misleading results. Hence, we incorporate the use of 1-s data combined with ILS indicators, encompassing an 11-year timeframe since the introduction of stock index futures in China. This enables us to accurately portray the time-varying characteristics of China's stock index futures.

In addition, we propose the LT indicator, which presents a useful tool for practitioners, offering a quantitative measure of the time gap between futures and spot prices. Moreover, this study is the first to address asynchronous informational updates and their impact on the price discovery of stock index futures. This provides meaningful insights for regulatory authorities when adjusting the frequency of index market updates.

The remainder of the paper is organized as follows. Section 2 presents a literature review, particularly about the technical aspects of the price discovery metrics. Section 3 presents the Monte Carlo simulations and discusses how data frequency impacts price discovery indicators using different parameter settings. Subsequently, Section 4 empirically analyzes the price discovery between Chinese stock index futures and spot. It further addresses the impact of asynchronous informational updates between the futures and spot markets and describes the impact of regulatory policies on price discovery indicators of futures. The study ends with a conclusion in Section 5.

2. Literature review and price discovery indicators

We begin with a literature review of price discovery indicators, IS, CS, and ILS, particularly about their technical aspects, followed by a detailed construction of our novel indicator of LT.

2.1. IS, CS, and ILS indicators

Most of the current price discovery measures are based on cointegration theory. Stock and Watson (1988) stated that if a cointegration relationship exists between two price series, x_t and y_t, there must be a common permanent component f_t such that

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} A \\ 1 \end{bmatrix} f_t + \begin{bmatrix} \widetilde{x}_t \\ \widetilde{y}_t \end{bmatrix}.$$

Stock and Watson (1988) and Gonzalo and Granger (1995) also depicted that the coefficients of f_t, which are $(A, 1)'$, are orthogonal to $\alpha = (\alpha_1, \alpha_2)'$, the coefficient of the error correction term in VECM. Thus, when taking the orthogonal vector $(\alpha_2, -\alpha_1)'$ of α and normalizing it, one can measure the contributions of different prices to the common component as:

$$CS_1 = \gamma_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1}, CS_2 = \gamma_2 = \frac{-\alpha_1}{\alpha_2 - \alpha_1},$$

where CS is the component share of the two prices. It is imperative to note that CS relates to α only.

Hasbrouck (1995) proposed the IS indicator. Suppose there are n I(1) price series, denoted as $p_t = \left[p_{1,t}, p_{2,t}, \cdots, p_{n,t}\right]'$. Since their first-differences Δp_t are stationary, it follows that:

$$\Delta p_t = \Psi(\mathrm{L}) e_t$$

which is the vector of moving average representations $VMA(\infty)$ of Δp_t, $\Psi(\mathrm{L})$ is a polynomial of the lag operator L, and e_t is a zero-mean independently distributed innovation whose covariance matrix is denoted as Ω. A model satisfying a particular cointegration relation such that $\beta' p_t$ remains stationary is assumed, with the cointegration vector defined as follows:

$$\beta' = \begin{bmatrix} 1 & -1 & & & \\ 1 & & -1 & & \\ \vdots & & & \ddots & \\ 1 & & & & -1 \end{bmatrix}.$$

As $\beta' p_t$ is stationary, $\beta' \Psi(1) = 0$. This implies that all rows of $\Psi(1)$ are identical. Denoting that common row in $\Psi(1)$ as ψ, some further analogues to the Beveridge-Nelson Decomposition (Beveridge and Nelson, 1981) render the following:

$$p_t = p_0 + \psi\left(\sum_{s=1}^{t} e_s\right)\mathbf{1} + \Psi^*(L)e_t,$$

where $\mathbf{1}$ is a column vector with elements of 1. Hasbrouck (1995) defined ψe_t as part of the novel information permanently incorporated into asset prices. Through the variance ($\psi\Omega\psi'$) of ψe_t, one can define i$^{\text{th}}$ price's share of information as follows:

$$IS_i = \frac{\psi_i^2 \Omega_{ii}}{\psi\Omega\psi'},$$

where Ω must be diagonal. To reflect a more realistic setting where Ω is a non-diagonal (i.e., the existence of correlations between price changes), Hasbrouck (1995) provided the following extension:

$$IS_i = \frac{\left[(\psi F)_i\right]^2}{\psi\Omega\psi'},$$

where $(\psi F)_i$ is the i$^{\text{th}}$ element of ψF and F is the Cholesky decomposition of Ω, i.e., $\Omega = FF'$. Since the Cholesky decomposition links with the order of the variables, one common practice is to calculate the IS in accordance with varying orders of the variables and treat their average as the final IS value (Baillie et al., 2002).

We follow the aforementioned approach, considering that a VECM is estimated as follows (Hasbrouck, 1995):

$$\Delta p_t = \alpha(\beta' p_{t-1} - E\beta' p_{t-1}) + \sum_{i=1}^{p-1} \Gamma_i \Delta p_{t-i} + e_t,$$

where the data is de-meaned. Thus, the estimated VECM, together with the corresponding nonstationary VAR, generates the VMA coefficient and value for IS.

Furthermore, Baillie et al. (2002) depicted that IS and CS rely solely on α and Ω. In particular, the IS takes the following form when $n = 2$:

$$IS_1 = \frac{(\gamma_1 m_{11} + \gamma_2 m_{21})^2}{(\gamma_1 m_{11} + \gamma_2 m_{21})^2 + (\gamma_2 m_{22})^2},$$
$$IS_2 = \frac{(\gamma_2 m_{22})^2}{(\gamma_1 m_{11} + \gamma_2 m_{21})^2 + (\gamma_2 m_{22})^2},$$

where m_{ij} is the Cholesky decomposition coefficient of Ω, i.e., $\Omega = FF^T$,

$$F = \begin{bmatrix} m_{11} & 0 \\ m_{21} & m_{22} \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix}, \Omega = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

To maintain simplicity, this study employs such an approach to calculate the value of IS.

Both CS and IS are dependent on the coefficients of VECM. CS is particularly associated with α and IS with α and Ω. CS focuses on the common components of prices and the "effectiveness" of price discovery, whereas IS focuses on the contributions of various markets to changes in the common components. Therefore, when compared with CS, IS implies moving to the common trend of prices, and thus the "speed" of price discovery (De Jong, 2002).

IS and CS are very widely used to analyze price discovery between spot, futures, options and ETF markets. However, IS and CS often give contradictory conclusions. It is worth noting that VECM is a reduced form and lacks structural modelling between prices. Some researchers have analyzed IS and CS from the perspective of structural cointegration, and thus proposed new price discovery indicators. Accordingly, Yan and Zivot (2010) propose the "speed" and "noise avoidance" components for IS and CS through structural cointegration, which represents a new scope for price discovery analyses. Putniņš (2013) further defines information leading (IL) to measure the IL of the asset price:

$$IL_1 = \left|\frac{IS_1}{IS_2}\Big/\frac{CS_1}{CS_2}\right|, IL_2 = \left|\frac{IS_2}{IS_1}\Big/\frac{CS_2}{CS_1}\right|.$$

Putniņš (2013) also presents ILS to measure the share of the leading information by the corresponding IL indicator:

$$ILS_1 = \frac{IL_1}{IL_1 + IL_2}, ILS_2 = \frac{IL_2}{IL_1 + IL_2}.$$

Through Monte Carlo simulations, Putniņš (2013) illustrates that the ILS indicator has the ability to eliminate the noise present in IS and CS measures. As a result, ILS proves itself as a suitable measure of the speed of price discovery.

In this paper, the difference between the upper and lower bounds of IS is calculated from the varying orders of the variables and denoted as information share difference (ISD). When the data frequency is high, the correlation between prices is low, which is manifested as the informational correlation ρ that is closer to 1. In certain extreme scenarios where $\rho = 1$ and $IS_1 = 1, IS_2 = 0$, if we alter the variable order and calculate the average of the resultant IS values, the IS values of the final two prices are both 50%. Therefore, faced with low-frequency data, the IS is unable to reveal the lead-lag relationship of prices. Hasbrouck (2002) called it the insufficient identification problem of IS. However, if $\rho = 0$, no correlation is observed between the prices. The upper and lower bounds of IS remain unchanged, and ISD is 0. We see that ISD actually reflects the correlations between prices. On the other hand, CS is independent of Ω; therefore, data frequency does not directly affect CS.

Getting the final IS by averaging the upper and lower bounds is problematic when data frequency is low. Lien and Shrestha (2009) propose the MIS indicator to address the upper and lower bounds that arise when computing IS. However, the MIS indicator suffers from an insufficient identification problem because it is close to the average of the upper and lower bounds of IS.

Although confirming ILS can remove the noise, Putniņš (2013) does not take into account the effect of data frequency. ILS relies on the calculation of IS, and thus also suffers from an insufficient identification problem. In other words, the performance of ILS under different data frequencies needs to be examined more carefully.

2.2. Leading time indicators

Although IS, CS, and ILS reveal the price discovery differentials among markets, they are expressed as percentages and so cannot reflect the lead-lag relation concretely. In fact, market traders are concerned regarding how faster futures incorporate new information compared to stock indices. Accordingly, we have proposed a new indicator named IS/CS/ILS Leading Time, denoted as IS-LT/CS-LT/ILS-LT, to measure the speed of price discovery. Here, we have used IS as an example to illustrate the mechanism of the IS-LT. The remaining indicators follow the same fashion.

If futures IS appears to be less than 50%, meaning that the futures are not ahead of the index, we mark the IS-LT of the futures as 0. If futures IS is greater than 50%, the index series is then gradually brought forward by a few seconds to calculate futures IS successively. In general, futures IS will gradually decline. When futures IS starts to fall below 50%, we define the time by which the index is brought forward as the IS-LT of futures. In other words, if the IS-LT of futures is m seconds, the futures lead spot by roughly m-1 to m seconds.

Historically, the LT differentials among asset markets are estimated through regressions among returns and their lagged values (Kawaller et al., 1987; Stoll and Whaley, 1990; Chan, 1992; De Jong and Nijman, 1997). Although it is reasonable to use regression to study lead-lag relationships, it is important to note that regression estimates correlations between returns rather than the precise lead-lag relationship between prices. The LT indicator we introduce provides a more explicit tool as it quantifies the average time difference between price movements, explicitly capturing the lead-lag relationship between prices.

3. Monte Carlo simulations

To examine the merits of the various methods in measuring price discovery, we utilize the Monte Carlo simulations, which first generate a price series with the given price discovery differential through simulations (constructed by structural models), then estimate the price discovery ability using econometrics methods. This approach allows us to see whether IS, CS, ILS, and LT can correctly capture true price discovery differences under various data frequencies.

3.1. Performance of price discovery indicators under diverse data frequency

Following Putniņš (2013), we assume two logarithmic price series of futures and index, $p_{1,t}$ and $p_{2,t}$, share a common trend m_t, assumed to be a random walk:

$$m_t = m_{t-1} + e_t, e_t \sim i.i.d.N(0, \sigma_m^2),$$
$$p_{1,t} = m_t + s_{1,t}, s_{1,t} \sim i.i.d.N(0, \sigma_{s1}^2),$$
$$p_{2,t} = m_{t-\delta} + s_{2,t}, s_{2,t} \sim i.i.d.N(0, \sigma_{s2}^2).$$

Here, δ, as lagged periods of $p_{2,t}$, is regarded as the time period for the index to fall behind futures. σ_{s1} and σ_{s2} denote the noise of the two prices around the common trend. By selecting different values for δ, σ_{s1} and σ_{s2}, we examine the sensitivity of IS, CS, and ILS to these parameters. In particular, we fix the value of $\sigma_m = 1 \times 10^{-5}$ and $(\sigma_{s1}, \sigma_{s2}) \in \{(2,4),(3,6),(6,6),(6,3),(4,2)\} \times 10^{-5}$.

In the first two noise combinations, futures are relatively less noisy, while the relative noise ratio of futures and index $\frac{\sigma_{s1}}{\sigma_{s2}} = \frac{1}{2}$ is fixed. In the third combination, the noises of both prices are identical. In the last two combinations, we swap their parameters with those in the first two noise combinations. Such a noise setting allows us to study the noise effect on the price discovery indicator, keeping the relative noise ratio constant. To explore the speed effect on the price discovery indicator, we set three speeds on δ, $\delta \in \{0, 2, 10\}$.

We simulate two price series $p_{1,t}$ and $p_{2,t}$ for each set of parameters for 1000 times, each contains two price series of length 14,400, resembling a trading time of seconds (i.e. 4 h a day). For each set of parameters, we calculate the mean and standard deviations of 1000 samples for all price discovery indicators.[2]

3.1.1. Noise reduction effect from low-frequency data on price discovery indicators

The noise retained in high-frequency data is a common problem in empirical analysis. Intuitively, low-frequency data helps preserve more fundamental information regarding the asset price with less market microstructure noise and may optimize the measure of price discovery indicator. In other words, Reducing the data frequency can have a "noise reduction effect" on price discovery indicators. However, such intuition must be formally testified. Table 1 presents the average values of the IS, CS, and ILS of the simulated futures prices at $\delta = 0$.[3] As the lead-lag relationship between prices is non-existent, we can examine whether the "noise reduction effect" exists on the price discovery indicators by excluding the "speed" factor.

We first examine the CS (Panel C) presented in Table 1. We see that CS possesses the ability to avoid noise: when the relative noise is high, the CS is low. Additionally, when the relative noise ratio remains constant, the CS remains unchanged. This shows that the noise component of CS actually measures the relative noise difference between prices. Furthermore, CS does not change significantly if data frequency decreases, which indicates that CS is not sensitive to data frequency. This implies that the noise avoidance component in CS does not change if data frequency changes. Therefore, the "noise reduction effect" due to frequency reduction is not evident on the CS.[4]

We then assess the results of ISD (Panel A) in Table 1. We can see that if the data frequency is low, the price correlations are strong and ISD is high. The first two rows illustrate that when the relative noise of the two prices is constant, the greater the $(\sigma_{s1}, \sigma_{s2})$ (i.e., the greater the noise relative to the common trend), the smaller the price correlations, and thus, the smaller the ISD. Conversely, when the values of $(\sigma_{s1}, \sigma_{s2})$ are smaller, it becomes less likely for prices to deviate from the common trend. This leads to higher correlations between prices, and consequently, an increase in the ISD. Thus, the magnitudes of σ_{s1} and σ_{s2} relative to σ_m play a role in determining the value of ISD.

Panel B shows that the IS is affected by noise. Under the same data frequency, the smaller the noise, the higher the IS. Thus, both IS and CS possess noise avoidance mechanisms. As the data frequency decreases, the "noise reduction" from frequency reduction impacts IS as it gradually approaches 50%. This is consistent with the results of ISD. As the data frequency gradually decreases, the ISD increases and the IS value approaches 50% after averaging. This phenomenon indicates a "noise reduction effect," where the relative difference in noise between prices decreases at low data frequencies. When the frequency is very low to make the correlation between prices strong enough, the ISD can be very large, resulting in the IS being very close to 50% after the average. At this point IS has completely failed to capture the noise relationship between prices, reflecting the insufficient identification problem.

It's worth noting that the "noise reduction effect" impacts differently on the two prices under different noise levels. When $\sigma_{s1} < \sigma_{s2}$, the futures noise is relatively small and the "noise reduction effect" weakens the low-noise advantage of futures under low-frequency data. IS decreases with a decrease in data frequency. On the contrary, when $\sigma_{s1} > \sigma_{s2}$, the disadvantages of high noise of futures at high data frequencies gradually weaken under low-frequency data, and IS increases as frequency decreases.

Finally, it can be seen in Panel D that ILS is closer to 50% under high-frequency data, indicating that it can remove the influence of noise more effectively.[5] This finding is consistent with Putniņš (2013). By definition, ILS has a positive/negative relationship with IS/CS. ILS is greater/smaller than 50% if IS is greater/smaller than CS. When $\sigma_{s1} < \sigma_{s2}$ ($\sigma_{s1} > \sigma_{s2}$), futures IS gradually decreases (increases) but CS remains largely unchanged following the frequency reduction. Hence, ILS gradually decreases (increases) and falls (rises) below (above) 50%. Therefore, at low data frequencies, ILS may be greater or less than 50%, resulting in completely inconsistent results with the speed set by the model. As the frequency decreases, the noise components present in IS and CS no longer align consistently. Consequently, the ILS becomes incapable of eliminating the impact of noise by combining IS and CS. This limitation could lead to biased statistical inferences.

2 Sometimes, the α coefficients of VECM generate the same sign in the simulated sample. To minimize the influence of such a problem, we remove these extreme samples from our simulations.

3 The results of standard deviations are presented in Appendix A.

4 From Table 1, it can be seen that CS slightly approaches 50% at the lowest data frequency, which means the noise reduction effect on CS may be rather slow.

5 From the results of the 1-s frequency data, we can see that ILS does not remove noise perfectly.

Table 1
Price discovery indicator of futures under diverse data frequency $\delta = 0$.

Panel A: ISD

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	17.74	37.9	43.76	48.29	50.31	58.33	63.7	76.37	82.87
(3, 6)	11.95	26.14	30.62	34.43	36.05	42.86	47.86	61.32	69.27
(6, 6)	11.95	25.81	30.25	33.86	35.22	42.01	46.97	59.29	66.41
(6, 3)	11.97	26.09	30.7	34.58	35.98	42.56	47.99	61.06	68.89
(4, 2)	17.78	37.83	43.7	48.44	50.23	58	63.75	76.08	82.57

Panel B: IS

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	78.6	73.85	72.08	70.68	69.85	66.9	64.89	59.46	56.2
(3, 6)	79.38	76.85	75.87	74.94	74.33	72.2	70.65	65.49	61.89
(6, 6)	50.01	49.84	50.04	49.99	49.97	49.79	50.08	49.81	49.76
(6, 3)	20.69	23.01	24.23	25.23	25.57	27.5	29.45	34.22	37.61
(4, 2)	21.5	26.04	27.84	29.46	30.05	32.81	35.11	40.26	43.43

Panel C: CS

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	80.04	79.96	79.97	80.13	79.92	79.79	79.98	78.74	75.15
(3, 6)	80.04	79.97	80.03	80.12	79.94	79.86	80.04	79.66	78.27
(6, 6)	50.01	49.9	50.02	49.99	49.99	49.85	50.08	49.77	49.64
(6, 3)	20	19.95	20.03	20.02	19.97	19.89	20.05	19.86	20.63
(4, 2)	20.02	19.96	19.93	20.01	19.98	19.83	19.99	20.47	23.32

Panel D: ILS

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	45.69	33.33	29.38	26.15	25.1	20.57	17.57	14.82	17.84
(3, 6)	48.06	40.95	38.1	35.46	34.47	29.81	26.09	18.84	17.49
(6, 6)	50.01	49.87	50.06	49.99	49.96	49.87	50	50.09	50.25
(6, 3)	52.03	58.92	61.98	64.57	65.54	70.25	73.87	81.65	83.69
(4, 2)	54.42	66.64	70.72	73.79	74.97	79.78	82.49	85.97	83.62

Note: 1) Table 1 reports the average values of the IS, CS, and ILS of the simulated futures prices at δ = 0. The price discovery indicator is the average of the simulated values. The unit of measurement is in percentage. 2) IS and CS have the ability to avoid noise. 3) Noise reduction effect due to frequency reduction is not evident on CS. 4) Noise reduction effect is significant on IS. 5) When frequency decreases, ILS is unable to completely remove the noise.

3.1.2. *Speed reduction effect from low-frequency data on price discovery indicators*

Intuitively, the lower the data frequency, the more “blurred” the price series and the less prominent the difference in the information transmission speed among prices. Here, by taking a non-zero value of δ, one can investigate how frequency reduction influences the speed component of the price discovery indicator, or the “slowdown effect.” Tables 2 and 3 report the average values of the IS, CS, and ILS indicators for futures prices simulated at δ = 2 (a small speed component) and δ = 10 (a large speed component).

Panel C in Tables 2 and 3 depicts the CS. When there is no relative noise difference between prices (see the third row of Panel C), CS depicts a downward trend after frequency reduction. This suggests that the “slowdown effect” exists in CS. Moreover, as the frequency decreases, the CS at different noise ratios gradually decreases and approaches the value without the speed dimension when δ = 0. This shows that the frequency reduction reduces the speed component of CS. When comparing the results based on different values of δ, it was found that the larger the δ, the greater the speed component of CS.

ISD and IS are then examined. Consistent with its counterpart results in Table 1, ISD gradually increases when frequency decreases, suggesting an increased correlation between prices. However, the changes in IS are more complex. From the third row of Panel B in Tables 2 and 3, it can be seen that IS decreases as frequency decreases in the absence of a relative noise difference. Hence, the “slowdown effect” on IS exists. From the last two rows of Panel B in Table 2 ($\delta = 2$), the futures noise is larger ($\sigma_{s1} > \sigma_{s2}$), and it was observed that IS first rises and then falls as the frequency decreases. If the “slowdown effect” does not exist, IS should always rise and not fall due to the “noise reduction effect.” This again suggests the existence of the “slowdown effect.” Moreover, this also suggests the “noise reduction effect” and “speed reduction effect” on the IS operate in contrasting directions. When δ is small, the “slowdown effect” dominates and the speed component of IS decreases rapidly, resulting in decreasing values of IS. When the speed component remains stable, the “noise reduction effect” begins to dominate, and IS starts to rise. When δ is large, the speed component of IS is large as well. As frequency decreases, the “slowdown effect” may persist for a considerable amount of time, causing IS to constantly exhibit a decreasing trend. From

Table 2
Price discovery indicator of futures under diverse data frequency $\delta = 2$.

Panel A: ISD									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	3.59	24.54	31.73	37.47	40.01	50.08	56.84	72.5	80.26
(3, 6)	4.26	18.63	23.63	28.04	29.88	37.47	43.06	58.07	66.95
(6, 6)	8.79	23.23	27.82	31.56	33	40.1	45.17	58	65.37
(6, 3)	12.19	27.36	32	35.76	37.18	43.82	49.11	61.92	69.27
(4, 2)	16.19	39.15	45.24	49.95	51.79	59.52	64.98	76.62	82.21

Panel B: IS									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	95.4	83.18	80	77.51	76.32	71.9	68.91	61.63	57.65
(3, 6)	92.3	83.56	81.49	79.59	78.74	75.85	73.73	67.45	63.24
(6, 6)	66.14	56.95	55.85	54.96	54.63	53.2	52.88	51.69	51.14
(6, 3)	42	31.49	31.06	30.82	30.84	31.69	32.92	36.29	38.9
(4, 2)	53.55	37.71	37.29	37.4	37.46	38.31	39.42	42.43	44.44

Panel C: CS									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	90.4	84.5	83.98	83.73	83.45	82.97	82.83	81.07	77.39
(3, 6)	87.58	83.35	82.85	82.35	82.08	81.77	81.75	81.13	79.42
(6, 6)	58.64	54.09	53.61	53.22	53.07	52.34	52.3	51.82	51.55
(6, 3)	30.94	25.88	25.34	24.82	24.68	24.35	24.33	23.81	23.95
(4, 2)	37.31	29.93	29.4	29.13	28.97	28.46	28.31	27.84	28.61

Panel D: ILS									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	83.07	45.08	36.65	30.68	28.68	21.35	17.32	13.49	15.94
(3, 6)	74.52	50.94	45.46	41.1	39.5	32.66	27.72	18.62	17.03
(6, 6)	65.53	55.81	54.55	53.53	53.17	51.76	51.18	49.71	49.12
(6, 3)	72.28	63.33	63.76	64.54	64.94	67.64	70.25	77.26	80.11
(4, 2)	78.95	66.77	67.12	67.95	68.41	71.1	73.35	77.95	77.64

Note: 1) Table 2 reports the average values of the IS, CS, and ILS indicators for futures prices simulated at $\delta = 2$ (a small speed component). 2) The price discovery indicator is the average of the simulated values. The unit of measurement is in percentage. 3) Frequency reduction reduces the speed component of CS and IS. 4) ILS may be biased under low-frequency data; the indicator is more suitable with high-frequency data as it removes the noise influences.

the last two rows of Panel B in Table 3 ($\delta = 10$), it can be seen that IS decreases without rising as frequency decreases.

In contrast, the first two rows of Panel B in Tables 2 and 3 suggest that the "noise reduction" and "speed reduction" work in the same direction for the IS indicator when futures noise is smaller ($\sigma_{s1} < \sigma_{s2}$). Along with frequency reduction, both the speed and noise avoidance components of IS decrease, thus causing IS to show a downward trend. When the frequency is sufficiently low to bring the IS very near to 50% after the average, the IS reflects neither a noise difference nor a speed difference between the two prices.

Finally, from Panel D in Tables 2 and 3, we can see that ILS at high frequencies can remove the effects of noise and can more accurately reflect the true speed difference between two prices. The performance of ILS at low frequencies is somewhat similar to Panel D in Table 1. As the frequency decreases, the speed component contained in IS and CS decreases, and ILS approaches the value when there is no speed difference between the two prices. At low frequencies, IS contains small speed and noise components, whereas CS contains a small speed component but a large noise component since the noise component on CS does not change under different frequencies. Thus, ILS at low frequencies reflects only the relative noise difference between prices and does not include the speed component. Especially when the futures noise is relatively small, CS retains more noise components at low-frequency data and its value is significantly larger than 50%. Although IS is also greater than 50%, it is closer to 50% and smaller than CS, resulting in ILS being less than 50%. This scenario is very similar to the three articles discussed by Putniņš (2013) mentioned before. For example, Hsieh et al. (2008) confirm that the index futures dominate the price discovery compared to the spot prices implied by index options using the put-call-parity method since the IS and CS are 66% and 83%, respectively. If ILS is taken into account, then spot prices should be considered to dominate price discovery since futures ILS is 14%. However, the difference between the upper and lower bound of IS is up to 62%, indicating the data frequency is very low, and the result of ILS is likely to be misleading. The difference between the upper and lower bounds of IS is even 99% when the spot prices implied by the index options are calculated by the Black-Scholes method. In this case, IS has almost no information about price discovery and ILS only reflects noise differences between prices.

Table 3
Price discovery indicator of futures under diverse data frequency $\delta = 10$.

Panel A: ISD									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	0.03	1.16	3.59	4.9	5.42	18.49	29.66	56.51	69.77
(3, 6)	0.07	2.88	5.13	7.53	8.3	17.01	24.28	45.05	57.76
(6, 6)	1.31	11.57	16.14	19.78	21.11	30.03	36.14	51.47	60.4
(6, 3)	2.11	18.84	24.81	29.9	31.76	40.56	46.93	61.42	68.96
(4, 2)	0.6	15.94	25.08	30.92	33.41	47.1	55.69	72.13	78.99

Panel B: IS									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	99.94	99	97.35	96.83	96.61	89.72	83.98	70.26	63.38
(3, 6)	99.79	96.74	95.5	93.88	93.51	88.92	85.31	74.94	68.43
(6, 6)	93.93	79.16	75.51	73.43	72.88	66.8	64.31	59.26	56.69
(6, 3)	91.93	64.32	59.81	55.23	54.33	49.96	47.89	44.97	44.76
(4, 2)	98.56	81.38	72.93	70.57	69.53	61.59	57.59	51.99	50.53

Panel C: CS									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	98.9	95.49	92.99	93.17	93.15	90.98	89.85	86.46	82.87
(3, 6)	97.91	91.63	90.99	89.57	89.25	87.91	87.37	85.31	82.81
(6, 6)	80.03	67.22	65.25	64.25	64.02	60.99	60.17	58.9	58.17
(6, 3)	65.01	44.42	42.84	40.57	40.15	38.89	38.29	36.78	36.05
(4, 2)	83.03	58.98	54.06	53.69	53.25	51.22	50.25	48.67	47.93

Panel D: ILS									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	99.7	94.44	87.76	81.37	78.31	40.6	25.03	12.83	12.77
(3, 6)	99.02	88.34	81.81	75.96	75.21	54.34	40.17	20.47	16.93
(6, 6)	93.82	77.62	73.15	70.46	69.7	62.5	58.82	50.65	46.6
(6, 3)	97.44	83.61	79.81	76.59	75.91	71.16	68.8	66.69	67.77
(4, 2)	99.51	90.29	84.05	81.12	80.12	70.07	64.44	56.51	54.99

Note: 1) Table 3 reports the average values of the IS, CS, and ILS indicators for futures prices simulated at δ = 10 (a large speed component). 2) The price discovery indicator is the average of the simulated values. The unit of measurement is in percentage. 3) Frequency reduction reduces the speed component of CS and IS. 4) ILS may be biased under low-frequency data; the indicator is more suitable with high-frequency data as it removes the noise influences.

Table 4 summarizes the "noise reduction" and "speed reduction" effects along with frequency reduction upon different price discovery indicators. In general, the frequency reduction demonstrates a dual effect of "noise reduction" and "speed reduction" on IS and CS, where the former does not significantly impact CS. As a result, the ILS indicator becomes ineffective in eliminating noise through the combination of IS and CS when the data frequency is low and generates incorrect inferences. However, when the data frequency is high, ILS can remove the noises. To ensure that the data frequency is high enough, a relatively reliable approach is to focus on the difference between the upper and lower bounds of IS, which is ISD. If the value of ISD is small (for example, at 20%), it means that the data frequency is high enough and ILS is reliable. Thus, it is recommended to use high-frequency data to calculate the ILS indicator and present results of ISD, IS and CS as supporting evidence.

Appendix A presents the standard deviation of all price discovery indicators. The standard deviations of all indicators are smaller with high-frequency data and larger with low-frequency data, especially for CS and ILS. This also suggests that ILS under high-frequency data is more reliable.

3.2. Performance of leading time indicators under high-frequency data

We test the reliability of the LT indicator through Monte Carlo simulations under the following scenarios:

Scenario 1: $\delta = 10$, $\sigma_m = 2 \times 10^{-5}$, $\sigma_{s1} = 4 \times 10^{-5}$, $\sigma_{s2} = 3 \times 10^{-5}$,
Scenario 2: $\delta = 10$, $\sigma_m = 2 \times 10^{-5}$, $\sigma_{s1} = 3 \times 10^{-5}$, $\sigma_{s2} = 3 \times 10^{-5}$,

Table 4
Performance of price discovery indicators under frequency reduction.

		Noise Reduction Effect	Speed Reduction Effect	Noise Reduction + Speed Reduction Effects
IS	Smaller noise	↓	↓	↓
	Larger noise	↑		↓ or first↓ then ↑
CS		→	↓	↓
ILS		Effective for high-frequency data but unreliable for low-frequency data.		

Note: 1) Table 4 summarizes the "noise reduction" and "speed reduction" effects along with frequency reduction upon different price discovery indicators. 2) Frequency reduction has a dual effect of "noise reduction" and "speed reduction" on IS and CS. 3) Noise reduction does not exist significantly on CS. 4) ILS can remove the noises when data frequency is high but unreliable when data frequency is low.

where $\delta = 10$ means the true leading time of futures is 10 s.

Table 5 reports the statistics of the futures IS, CS, and ILS with the index brought forward by various time periods under scenario 1. The grey highlighted parts in Table 5 signify that the value of the indicator is below 50%. The price discovery indicators decline along with the index being forwarded. The LT indicator is the time at which the index is advanced when the price discovery indicator first drops below 50%. From Table 5, we can see that the LT for IS is 10 s, the LT for CS is 9 or 10 s, while the LT for ILS is 11 s, which is consistent with the true leading time of 10 s. Table 6 reports the results under scenario 2. We can see that LT indicators for all price discovery indicators are 10 s or 11 s, which are also consistent with the leading time set by the model. By definition, if we calculate LT to be m seconds, then futures actually lead spot by roughly between m-1 and m seconds. This implies that LT will slightly overestimate futures' leading times, but such bias is relatively small. Overall, the LT indicators perform well at capturing the actual lead-lag relationship between prices.

In Scenario 1, When the index is advanced 10 s forward, the means of IS and CS are 37.94% and 35.99%, respectively, while the mean of ILS is 54.16%, much closer to 50%. In Scenario 2, IS, CS and ILS are all around 50% when the index is advanced 10 s forward. This also suggests that the ILS indicator can reduce the data noise. Given that the level of noise in the price series of futures and spot indices is time-varying and extremely difficult to measure, we argue that the LT indicator based on ILS would be more reliable, while the LT indicators based on IS and CS should be taken as references.

It is worth noting that the true prices of futures and indexes may not satisfy the random walk assumption of the Monte Carlo simulations. Therefore, our analyses may not be entirely compatible with real-world situations. Nevertheless, our Monte Carlo simulations highlight the necessity of using high-frequency data in analyzing futures price discovery. Furthermore, the leading time indicator based on the ILS emerges as a sensible and valuable tool in this analysis.

Table 5
IS, CS, and ILS of futures after indices brought forward by different time periods (Scenario 1).

Panel A: IS								
	1 s	7 s	9 s	10 s	11 s	15 s	20 s	40 s
Mean	99.95	93.44	66.47	37.94	14.19	0.15	0.03	0.24
Max	100	96.11	72.82	43.25	19.43	1.09	0.37	5.19
Min	99.37	89.56	60.13	31.52	10.14	0	0	0
Panel B: CS								
	1 s	7 s	9 s	10 s	11 s	15 s	20 s	40 s
Mean	97.86	77.12	54.53	35.99	21.32	2.44	1.14	3.01
Max	100	82.14	60.01	40.6	26.53	8.6	5.86	16.01
Min	91.33	71.1	49.53	31	16.78	0	0	0
Panel C: ILS								
	1 s	7 s	9 s	10 s	11 s	15 s	20 s	40 s
Mean	99.67	94.73	73.22	54.16	27.09	0.73	0.36	0.48
Max	100	96.78	76.61	56.96	31.36	99.79	99.89	99.79
Min	0.02	91.88	69.61	50.96	22.11	0	0	0

Note: 1) Table 5 reports the statistics of the futures IS, CS, and ILS with the index trading time brought forward by different time periods under different noise conditions when the true lead time of futures is 10 s. 2) The grey areas show a price discovery indicator lower than 50%. 3) The lead time calculated based on IS and CS is 10 s on average. The lead time calculated based on ILS is 11 s. 4) When the index data is 10 s ahead, the IS and CS values are further away from 50%. The value of ILS (54.16%) is the closest of the three indicators to 50%.

Table 6
IS, CS, and ILS of futures after indices brought forward by different time periods (Scenario 2).

Panel A: IS

	1 s	7 s	9 s	10 s	11 s	15 s	20 s	40 s
Mean	99.97	97.96	79.85	50.00	20.14	0.19	0.02	0.34
Max	100	99.29	85.25	55.26	25.17	1.01	0.3	6.17
Min	99.57	95.8	74.4	43.56	15.13	0	0	0

Panel B: CS

	1 s	7 s	9 s	10 s	11 s	15 s	20 s	40 s
Mean	98.63	89.14	69.89	49.99	30.09	3.36	1.35	4.41
Max	100	94.56	75.85	54.82	34.73	10.1	6.79	20.45
Min	93.36	83.51	64.36	44.57	24.67	0	0	0

Panel C: ILS

	1 s	7 s	9 s	10 s	11 s	15 s	20 s	40 s
Mean	99.71	97.2	74.49	50.01	25.52	0.47	0.31	0.4
Max	100	98.82	78.19	52.14	29.08	99.92	99.86	99.9
Min	0.04	94.64	70.72	47.31	21.14	0	0	0

Note: 1) Table 6 reports the statistics of the futures IS, CS, and ILS with the index trading time brought forward by different time periods under the same noise conditions when the true lead time of futures is 10 s. 2) The grey areas show that a price discovery indicator is lower than 50%. 3) On average, the IS, CS, and ILS lead times are similar (10 s in some samples and 11 s in other samples).

4. Empirical analyses

4.1. Data

We conduct an empirical analysis of the price discovery between futures and their underlying stock index, specifically focusing on CSI 300 and CSI 500 data. Since SSE 50 constituents are generally also CSI 300 constituents, and CSI 300 and CSI 500 stocks include the most important part of the stocks in Chinese stock markets, this paper selects these two stock index futures for analysis.

The data for the empirical analysis are obtained from the CSMAR high-frequency database. The CSI 300 futures were launched in 2010, whilst SSE 50 and CSI 500 futures were launched in 2015. Due to data availability, May 31, 2021 was chosen as the sample endpoint for this paper. Therefore, the sample range of CSI 300 (CSI 500) stock index futures is from April 16, 2010 (April 16, 2015), to May 31, 2021. On January 4th and 7th, 2016, stock market circuit breakers led to the early closure of the Chinese stock exchanges. From July 8th to 10th 2015, and on January 5th, 2016 and February 3rd, 2020, stock index futures hit the price limit, which caused suspension of the futures prices for a long time. To avoid the impact of these samples on our results, we excluded these trading days from the sample. The sample ultimately contains a total of 2696 (1485) trading days for the CSI 300 and CSI 500 respectively. Before 2016, trading hours of futures and stocks did not coincide. Therefore, we select the overlapping trading hours of both futures and stock on each trading day, that is, 09:30:00 to 14:57:00 after excluding opening and closing call auction sessions. This includes the price series at the 1-s frequency with each series containing 14,222 consecutive trades.

Consistent with previous literature (Yang et al., 2021), we select the prices of the most active futures contracts (contracts with the largest trading volumes of the day) as the representative futures prices of the day. The original stock index futures data was updated every 0.5 s and we reduced the data frequency to 1 s. The update frequency of the CSI 300 and CSI 500 indices is roughly every 5 s, so it is not possible to obtain the index data at a 1-s frequency directly. To this end, we first generate 1-s data for each component stock through historical tick data of CSI 300 and CSI 500 constituents and then construct the index at a 1-s frequency according to the CSMAR daily weights of constituent stocks of the stock index.[6]

[6] We do not replace the index with the price of an index ETFs. The reason is that the update frequency of ETFs is the same as that of the index, and so ETFs cannot be used to study the impact of asynchronous market update on the price discovery. Moreover, existing literature (see, for example, Buckle et al., 2018) reveals the differences in price discovery on ETFs and indices, which means they cannot replace with one another. In addition, limited to data availability, we only employ individual stock tick data from 2017 to 2021 and snapshot data from 2010 to 2017. The latter is updated approximately every 3 or 5 s. Since the tick data and snapshot data overlap in 2017, we verify that these two types of data are highly close and the resultant price discovery indicators are consistent. The difference in the 2017 price discovery indicators calculated separately from the two types of data is illustrated in Appendix B.

Table 7
Statistics of daily price discovery indicators of futures.

Panel A: CSI 300 Stock Index Futures							
	No. of obs.	Periods of Information Dominance	Percentage of Information Dominant Periods	Mean	Std. Dev	Max	Min
IS	2696	2484	92.14%	81.79	20.21	100	0
CS	2696	1915	71.03%	59.2	19.5	100	0
ILS	2696	2619	97.14%	91.38	15.02	100	0

Panel B: CSI 500 Stock Index Futures							
	No. of obs.	Periods of Information Dominance	Percentage of Information Dominant Periods	Mean	Std. Dev	Max	Min
IS	1485	1318	88.75%	76.88	21.12	100	0
CS	1485	517	34.81%	43.32	20.15	100	0
ILS	1485	1459	98.25%	94.12	12.36	100	0

Note: 1) Table 7 provides the preliminary statistics of the IS, CS, and ILS price discovery indicators for CSI 300 and CSI 500 futures. 2) Futures price discovery indicators are predominantly higher than the corresponding indicators of the stock index, with the percentage of information dominant periods by the three indicators being 92.14%, 71.03%, and 97.14%, respectively. 3) Based on ILS, the price discovery ability of China's stock index futures is much stronger than that of their underlying indices. 4) The CS of futures is not so high (especially in the case of CSI 500).

4.2. Empirical results on price discovery indicators

Before VECM can be estimated, it is necessary to examine the cointegration between prices. Since the IS is based on a specific cointegrating vector $\beta' = (1, -1)$, ADF is used as the unit root test on the two prices. During the majority of trading days, the price is integrated of order one, whilst the first difference is stationary. On some days, the first difference in the price series did not pass the unit root test, exhibiting a non-stationary property. However, a visual inspection of the graphs suggests that the two price movements are close. Further, non-cointegrated trading days account for a small proportion of the sample with a minimal impact on the results. Under the strong linkage between futures and index themselves, we assume that futures and index prices on all trading days are cointegrated. Finally, to determine the order of the VAR while estimating the VECM, we use the BIC criterion with a maximum lag order of 20. We then estimate the VECM according to the order selected by the minimum BIC, followed by calculations of the price discovery indicator.

We utilize intraday trading data to calculate futures IS, CS,[7] and ILS for each trading day (the corresponding price discovery indicator of the index is 100% minus the futures counterpart). Table 7 provides the preliminary statistics on these indicators. We find that the average values of IS, CS, and ILS of CSI 300 stock index futures are 81.79%, 59.2%, and 91.38%, respectively. On most of the trading days, futures price discovery indicators are higher than the corresponding indicators of the stock index (i.e., futures dominate the price discovery), with the percentage of information dominant periods by the three indicators being 92.14%, 71.03%, and 97.14%, respectively. The average values of IS, CS, and ILS of CSI 500 stock index futures are 76.88%, 43.32%, and 94.12%, with the percentage of information dominant periods being 88.75%, 34.81%, and 98.25%, respectively. Based on ILS, the price discovery ability of China's stock index futures is significantly stronger than that of their underlying indices. The CS of futures is relatively low, indicating that the noise within the stock index futures is higher compared to that within the index. This disparity is particularly prominent in the case of CSI 500.

Fig. 1 plots the daily price discovery indicators of stock index futures, which appear to be time-varying. The ILS of stock index futures is relatively large during most of the trading days; however, some days depict stock index futures lagging behind the index. Comparing CSI 500 with CSI 300, the price discovery ability of the CSI 500 futures is significantly stronger relative to its underlying indices. In addition, ISD gradually increases from 2019 with a rapid upward trend after 2020. This suggests that the correlation between stock index futures and the index has increased in recent years, implying that the asset market linkages (in terms of information transmission) have risen rapidly, reflecting an enhanced synergy between China's stock index futures and spot markets. From the results, we can see that ISD can provide reliable information regarding the IS indicator and reflect correlations between prices. Nevertheless, previous literature has paid relatively little attention to the ISD indicator, with the value of ISD rarely reported.

Table 8 presents preliminary statistics of the price discovery indicators of CSI 300 futures under diverse data frequencies.[8] According to Panel A of the table, there is a positive relationship between lower data frequency and an increase in ISD, indicating that the gap between the prices narrows. Consequently, the IS indicator approaches 50% as data frequency decreases. When data frequency reaches every 20 s, the ISD generates an average of 31.79% whereas the IS appears to be only 65.12% after taking the average of the upper and lower bounds. As a result, ILS exhibits significant variation at low frequencies compared to high-frequency cases. Data frequencies lower than 20 s appear to be unsuitable for measuring price discovery with the given sample. Panel B reports statistics of the percentages of dominant price discovery indicators[9] under different frequencies, revealing a significant difference in these values

[7] When the estimated α coefficients of the VECM have the same sign, the CS is not between 0 and 1. We use a truncation method to calculate the CS (see Appendix C for more details).

[8] The results of CSI 500 stock index futures are similar and are available upon request.

[9] Dominant price discovery indicator equals 1 if futures price discovery indicator is higher than 50% and 0 otherwise.

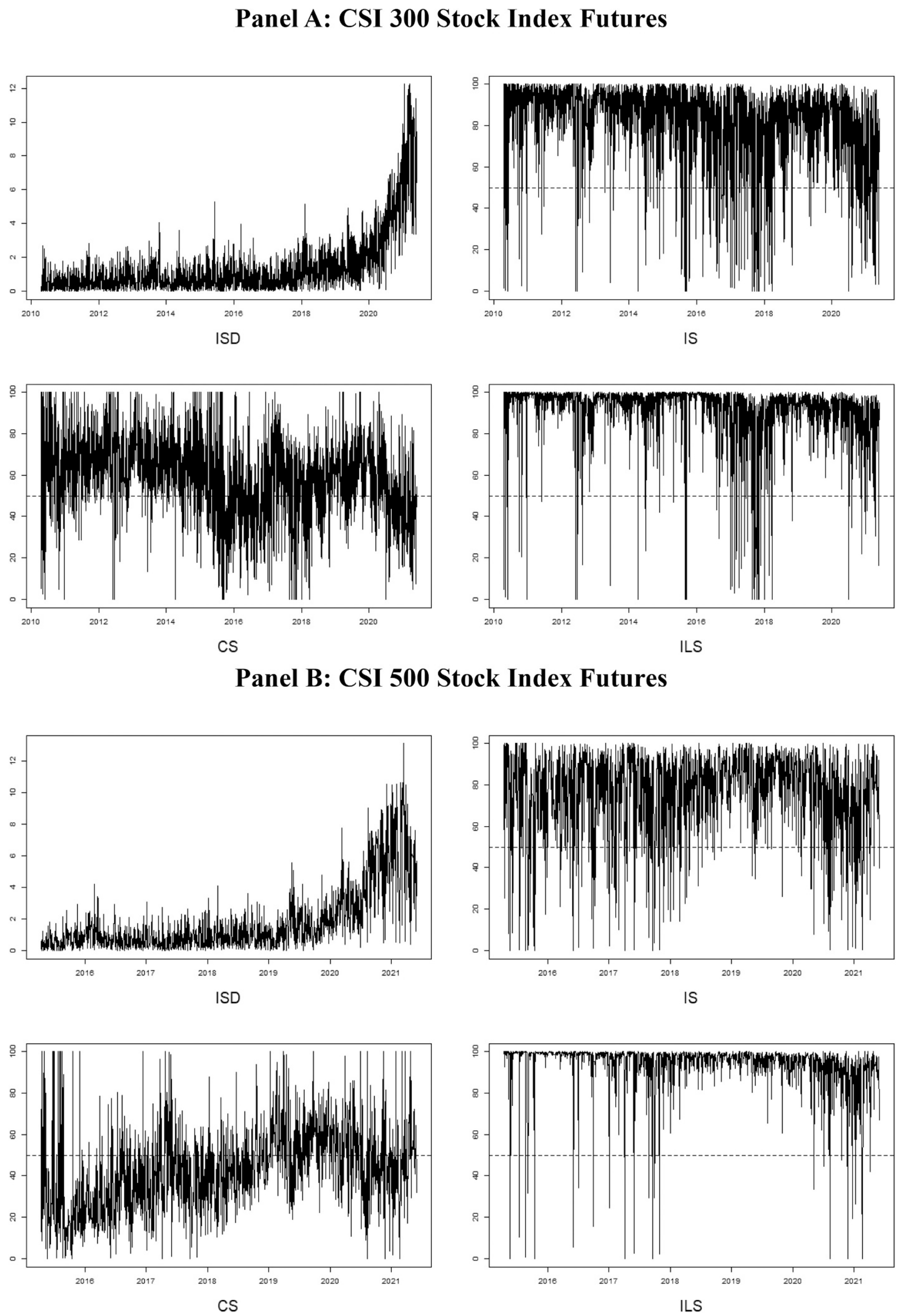


Fig. 1. Daily price discovery indicators of futures.
Note: 1) Fig. 1 plots the price discovery indicators of daily stock index futures. 2) Comparing CSI 500 with CSI 300, the price discovery ability of the CSI 500 futures is significantly stronger relative to its underlying indices. 3) ISD suggests that the correlation between stock index futures and the index has increased in recent years.

across a wide range of high to low-frequency data. Therefore, we conclude that the price discovery estimated from low-frequency data may appear different from the true price discovery. Overall, Table 8 and the previous Monte Carlo simulations reflect that the ILS indicator is biased under low-frequency data and is more robust under high-frequency data.

4.3. Empirical results on leading time indicators

To further quantify the lead-lag relationship between prices, we bring forward the index series by a certain amount of time and then

Table 8
Statistics of price discovery indicators of CSI 300 futures under diverse data frequencies.

Panel A: Price Discovery Indicators								
	1 s	5 s	10 s	20 s	30 s	40 s	50 s	60 s
ISD	1.37 (1.83)	8.6 (8.84)	17.53 (13.66)	31.79 (16.63)	41.74 (17.15)	46.54 (16.49)	50.61 (16.19)	58.22 (14.94)
IS	81.79 (20.21)	72.58 (24.79)	69.37 (25.43)	65.12 (24.47)	62.2 (22.8)	60.62 (21.95)	59.72 (20.83)	57.8 (18.52)
CS	59.2 (19.5)	53.93 (24.05)	53.23 (27.41)	54.34 (31.36)	55.46 (33.53)	56.32 (34.51)	57.37 (35.21)	58.48 (36.91)
ILS	91.38 (15.02)	86.38 (17.81)	83.88 (19.16)	78.03 (25.1)	73.87 (29.32)	70.51 (31.71)	67.88 (33.56)	65.72 (36.44)

Panel B: Dominant Price Discovery Indicators								
	1 s	5 s	10 s	20 s	30 s	40 s	50 s	60 s
IS	92.14 (26.92)	82.64 (37.88)	78.52 (41.07)	74.7 (43.48)	71.88 (44.96)	70.4 (45.66)	68.84 (46.32)	68.4 (46.5)
CS	71.03 (45.37)	57.08 (49.5)	52.71 (49.94)	52.6 (49.94)	54.01 (49.85)	54.86 (49.77)	56.31 (49.61)	57.83 (49.39)
ILS	97.14 (16.66)	95.62 (20.46)	94.47 (22.85)	88.65 (31.73)	82.97 (37.59)	78.89 (40.81)	75.07 (43.27)	70.55 (45.59)

Note: 1) Table 8 presents preliminary statistics of the price discovery indicators of CSI 300 futures under diverse data frequencies. 2) The value is in averages. The standard deviation is in parentheses. The unit of measurement is in percentage. 3) ILS at low frequencies changes a lot relative to high-frequency cases. 4) There is a significant difference in percentages of dominant price discovery indicators across a wide range of high to low-frequency data.

Table 9
Futures IS after the index trading time brought forward by a few seconds in the first 10 trading days of 2019.

Trading day	0 s	3 s	5 s	7 s	10 s	15 s	20 s	30 s	40 s	50 s
2019/1/2	80.07	64.23	54.61	45.67	36.02	10.25	3.77	0.67	5.63	12.59
2019/1/3	96.1	92.13	87.08	79.46	67.78	61.04	43.47	12.78	2.22	2.42
2019/1/4	89.08	79.79	70.57	60.07	46.68	25.21	10.5	0.59	8.84	13.85
2019/1/7	91.89	85.88	82.44	73.56	51.88	26.63	14.02	0.47	4.24	7.94
2019/1/8	87.12	81.44	74.08	65.63	52.05	28.1	14.31	0.71	0.38	2.34
2019/1/9	99.14	95.98	94.64	95.61	82.54	71.06	57.53	35.18	9.1	0.64
2019/1/10	86.05	82.21	73.47	67.63	46.16	20.07	10.8	0.63	1.67	3.59
2019/1/11	88.55	77.46	70.21	56.53	43.55	19.29	8.99	0.35	2.38	5.68
2019/1/14	87.19	79.37	75.11	62.29	49.79	28.74	13.79	1.98	0.17	2.92
2019/1/15	92.31	83.11	72.96	60.14	43.29	19.8	7.7	0.69	6.07	12.44

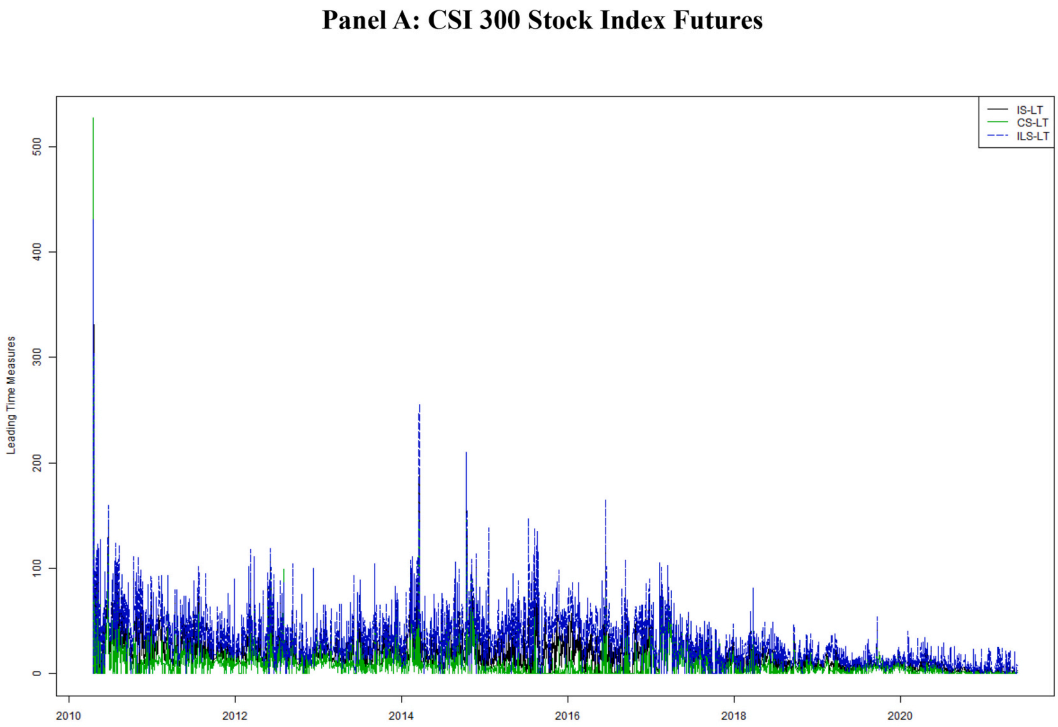
Note: 1) Table 9 shows the futures IS indicator after the trading time is brought forward by a few seconds during the first 10 trading days of 2019. 2) The grey areas imply that the values are below 50%. Results suggest that on 80% of the trading days, the lead time of futures is between 7 s and 15 s.

calculate the LT of futures IS, CS, and ILS, respectively. Generally, the price discovery indicator of futures decreases as the time the index is advanced increases. Table 9 shows the futures IS indicator after the index trading time is brought forward by a few seconds during the first 10 trading days of 2019, which is similar to the Monte Carlo simulations.

Fig. 2 plots the daily LT indicators of futures for the whole sample, and Table 10 provides the corresponding statistics.[10] In general, the LT is the longest for ILS and the shortest for CS, consistent with the results based on the price discovery indicator. The average LT of CSI 300 futures is 19.78 s (IS), 9.76 s (CS), and 32.23 s (ILS). In addition, the LT of CSI 300 futures generally depicts a downward trend, especially after 2019, when the LT of these futures was found to be relatively short (usually within 30 s). This is consistent with the results of ISD above. This is also in line with the experience of market participants and simultaneously reflects the gradual improvement of information efficiency in the spot market. The results suggest that stock index futures dominate price discovery. Futures and index are highly correlated, and both of them can quickly incorporate and reflect new information with futures performing at a faster speed.

A similar result is detected in CSI 500 stock index futures. The average LT of CSI 500 futures is 18.26 s (IS), 2.43 s (CS), and 46.53 s (ILS). Compared to CSI 300, a shorter lead time is detected by the IS and CS while a longer lead time is detected by the ILS of CSI 500 stock index futures. This implies that CSI 500 stock index futures contain large noise, while their price discovery speed is rapid.

[10] CSI 500 stock index futures have ILS leading time longer than 15 min for 5 trading days. We find that the ILS lead time for these trading days is on the extreme side of the calculated sample, thus these five trading days are removed from the statistics and graphs.



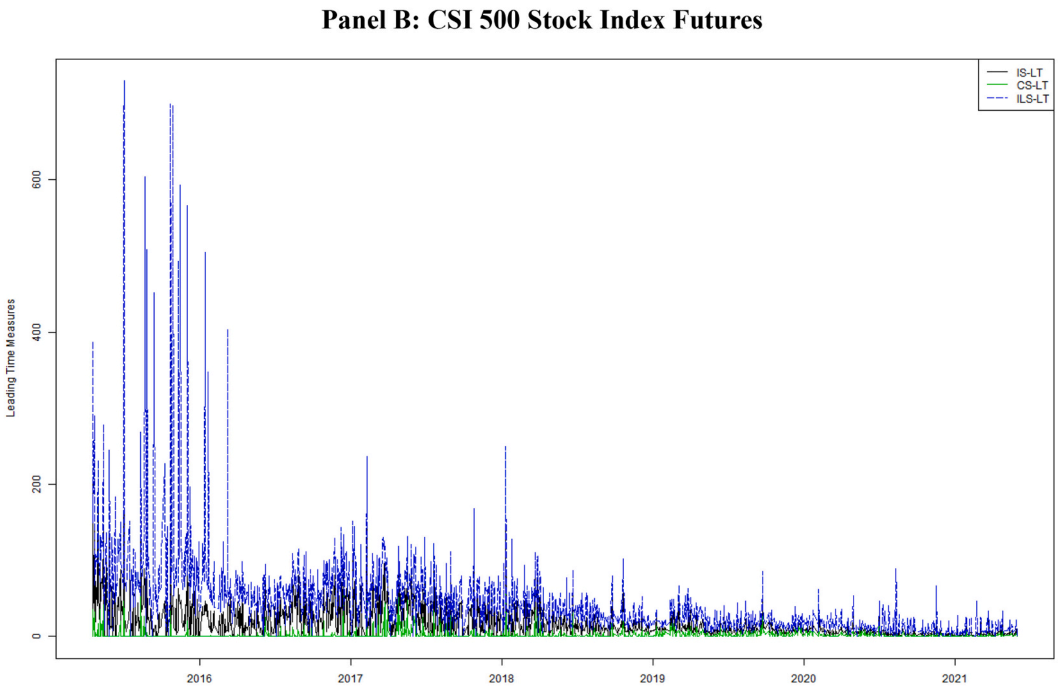


Fig. 2. Daily leading time indicators of futures.
Note: 1) Fig. 2 plots the daily lead time of futures. The lead time is the longest under ILS and the shortest under CS, which coincides with the results based on the price discovery indicators. 2) The leading time of CSI 300 futures generally depicts a downward trend, especially following 2019 where the leading time of these futures is relatively short.

Table 10
Statistics of the daily leading time of futures.

	CSI 300 Stock Index Futures				CSI 500 Stock Index Futures			
	Mean	Std. Dev	Max	Min	Mean	Std. Dev	Max	Min
IS-LT	19.78	21.54	527	0	18.26	22.17	202	0
CS-LT	9.76	16.99	527	0	2.43	6.18	63	0
ILS-LT	32.23	26.36	431	0	46.53	63.87	730	0

Note: 1) Table 10 provides statistics on the daily leading time of stock index futures. 2) The unit of measurement is in seconds. 3) The lead time is the longest under ILS and the shortest under CS. The average lead time of CSI 300 futures is 19.78 s (IS), 9.76 s (CS), and 32.23 s (ILS). 4) A similar result is detected in CSI 500 stock index futures. Moreover, compared to CSI 300, a shorter lead time is detected by the IS and CS, whereas a longer lead time is detected by the ILS of CSI 500 stock index futures.

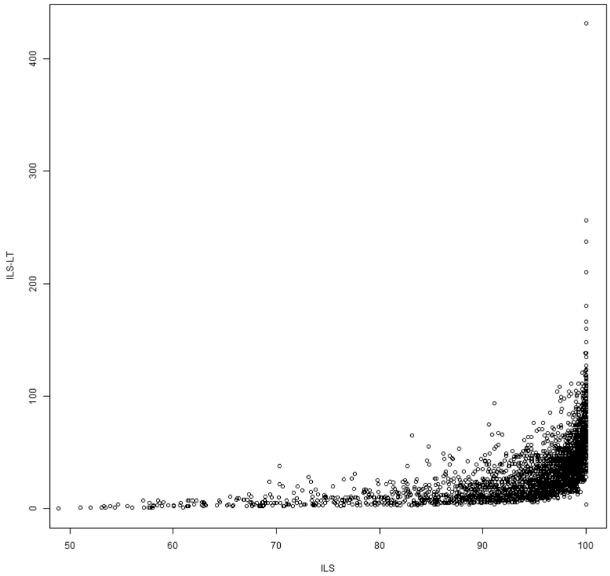


Fig. 3. Scatter plot of ILS and ILS leading time for CSI 300 stock index futures.
Note: 1) Fig. 3 shows a scatter plot of ILS and ILS Leading Time for CSI 300 futures. 2) There is a clear positive correlation between ILS and ILS lead time, depicting a non-linear relationship.

Fig. 3 shows a scatter plot of ILS and ILS Leading Time for CSI 300 futures.[11] There is, on average, a clear positive correlation between ILS and ILS lead time. Such a pattern suggests that our ILS Leading Time is a reliable indicator as it implies consistency with the traditional ILS. Fig. 3 also illustrates a nonlinear relationship between ILS and ILS-LT, indicating that ILS-LT captures additional information that the price discovery indicator ILS alone does not encompass and can indeed reflect different lead times under the same price discovery indicator. For example, when ILS is high (above 90%), the ILS LTs can vary. This aspect is not discernible through conventional indicators such as IS, CS, and ILS.

4.4. China's stock market crash in 2015 and regulatory policy on stock index futures

In 2015, China's stock market witnessed a significant downturn. Many investors hold the belief that the price discovery of futures played a role in this decline (Hao et al., 2019), and the speculations in futures contributed to the accumulation and transmission of risks to the stock market, ultimately leading to market turmoil (Miao et al., 2017). Faced with intense regulatory pressure, the regulators limited China's stock index futures trading by imposing severe restrictions (Han and Liang, 2017; Lin and Wang, 2018). There are three main rounds of restrictive policies on stock index futures, which came into effect on July 8th, 2015, August 26th, 2015, and September 7th, 2015 respectively. When the stock market gradually recovered, regulators began to deregulate the futures market. There are four main rounds of deregulation policies for stock index futures, which came into effect on February 17th, 2017, September 18th, 2017, December 3rd, 2018, and April 22nd, 2019, respectively. Appendix D outlines some of the main regulatory policies for stock index futures in the past.

To examine the nature of changes to the futures price discovery during the 2015 stock market crash and how the restrictive policy

[11] The results for IS and CS are similar and available upon request.

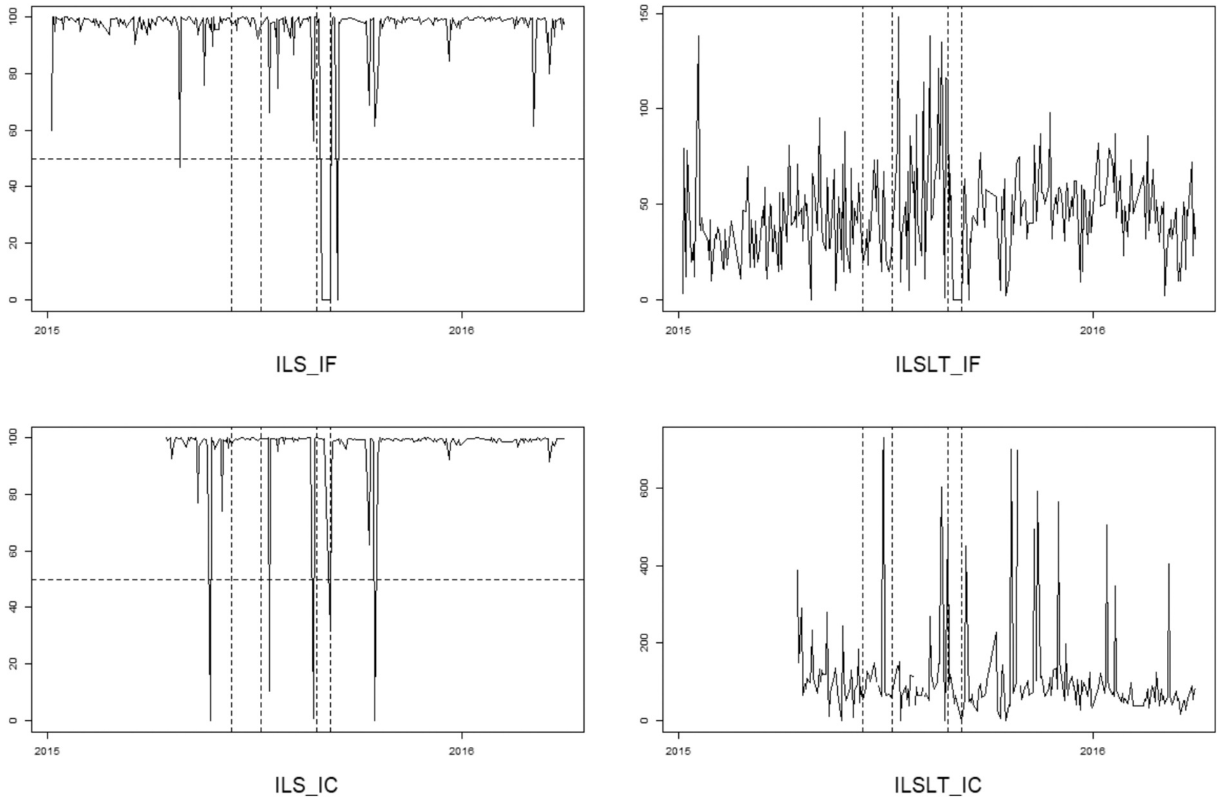


Fig. 4. ILS and ILS-LT of futures (January 1, 2015–March 31, 2016).
Note: 1) IF and IC represent CSI 300 and CSI 500 index futures respectively. The first vertical line corresponds to June 12th, 2015, when the index is the highest in 2015. 2) The second, third, and fourth vertical lines correspond to July 8th, 2015; August 26th, 2015; and September 7th, 2015, respectively. These are the dates when the restrictive policy on the stock index futures takes place. 3) Through the plot of the futures ILS and ILS-LT, the price discovery of CSI 500 futures appears stable with no obvious changes during this period. 4) The ILS of CSI 300 futures has shown an upward trend since 2015. Thus, the price discovery of CSI 300 futures enhances during the dramatic up and down of the market. 5) The restrictive policy causes a significant decline in the price discovery of CSI 300 stock index futures; however, such an effect does not last long.

impacts futures price discovery, we limit the sample to January 1 st, 2015, until March 31 st, 2016. Through the plot of the futures ILS and ILS-LT (see Fig. 4), the price discovery of CSI 500 futures appears stable with no obvious changes during this period. Despite the belief that futures price discovery contributes to the market turmoil in 2015, the upward trend in the ILS LT of CSI 300 futures suggests a potential positive role in providing timely and accurate price information. In particular, it faces a rapid increase when the market reaches its peak. It then falls sharply during the second and third rounds of the restrictive policy, slowly recovering to the level before the crash. Therefore, the price discovery of CSI 300 futures enhances during the dramatic ebbs and flows of the market, indicating that under the turmoil informed traders form their expectations towards the outlook of the market where they are more engaging in trading stock index futures.

The restrictive policy appears to impact the price discovery of stock index futures, causing a significant decline in the price discovery of CSI 300 stock index futures. Nevertheless, it is important to note that this effect is transient and does not persist for an extended period. When the market recovers from the crash, we see that the severely restricted Chinese stock index futures fare a good price discovery function.

Prior literature has examined the impact of China's stock index futures regulatory policies on futures price discovery with mixed conclusions. Some confirm that the restrictive policies weaken the futures price discovery (Miao et al., 2017; Xu et al., 2020; Xu et al., 2021) which is consistent with our results, while others find the opposite (Lin and Wang, 2018; Hao et al., 2019). However, none of these studies uses ILS indicators, and the data frequency is quite low (e.g. 1 min or 5 min). For example, Hao et al. (2019) divide the sample into two subsamples before and after the policy. The differences between the upper and lower bound of CSI 300 futures IS for these two subsamples are 45% and 63%, respectively, indicating possible bias in their results.

We then restrict the sample to April 1 st, 2016, to May 31 st, 2021, to examine the impact of deregulation on stock index futures on its price discovery (see Fig. 5). During the sample period, futures price discovery does not exhibit an upward trend. Even if stock index futures are severely restricted, they still function well in terms of price discovery ability. Although stock index futures were deregulated, the price discovery function of stock index futures was already strong before that, so the policy did not significantly improve the price discovery function of stock index futures. Therefore, deregulation has no real impact on the price discovery of stock

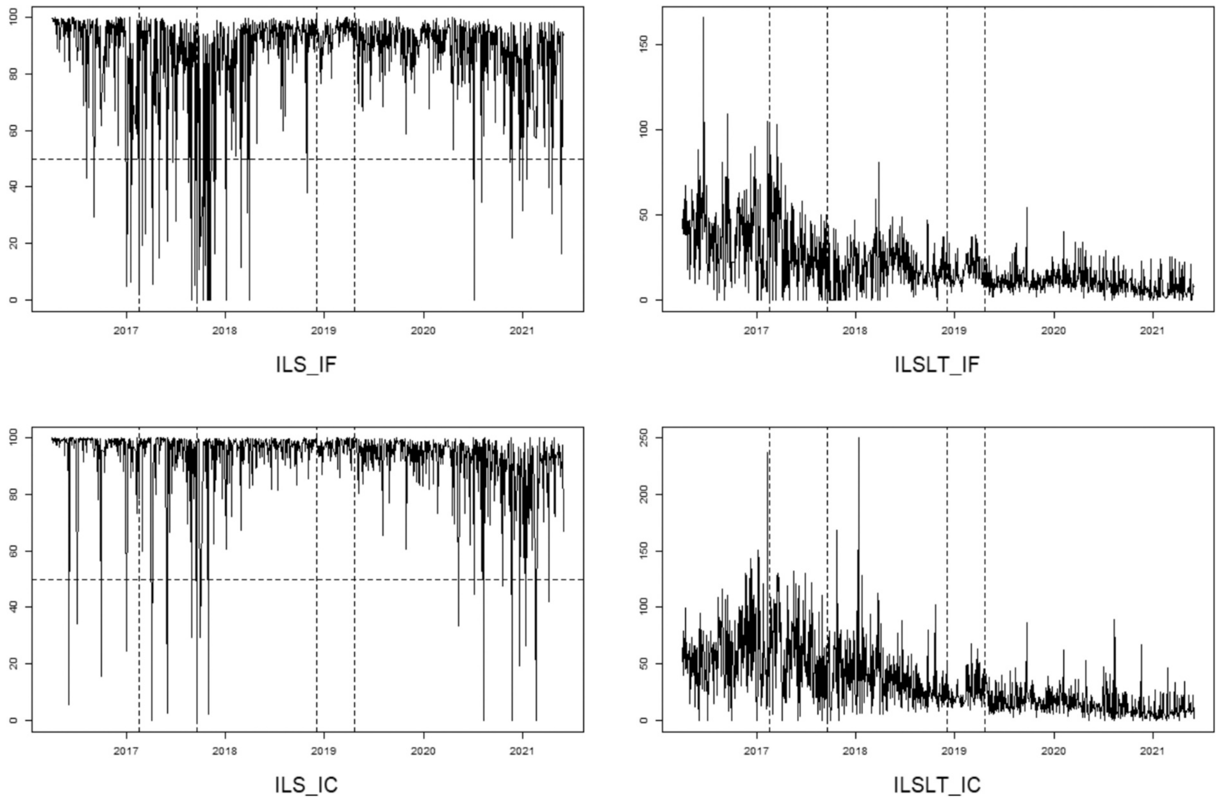


Fig. 5. ILS and ILS-LT of futures (April 1, 2016–May 31, 2021).
Note: 1) IF and IC represent CSI 300 and CSI 500 index futures respectively. The four vertical lines correspond to February 17th, 2017, September 18th, 2017, December 3rd, 2018, and April 22nd, 2019, respectively. These are the dates for the stock index futures deregulation policy to take effect. 2) When the market is stable, even if stock index futures are severely restricted, they still function well in terms of price discovery ability. Moreover, deregulation has no real impact on the price discovery of stock index futures.

Table 11
Statistics of price discovery and leading time indicators of CSI 300 stock index futures under different frequencies of updates on index information.

—	Panel A: Price Discovery Indicators				Panel B: Leading Time Indicators			
		Mean	Max	Min		Mean	Max	Min
	1 s	81.79 (20.21)	100	0	1 s	19.59 (19.2)	191	0
IS	5 s	83.87 (18.84)	100	0	5 s	21.14 (19.26)	193	0
	Difference	2.08 (5.27)	24.58	−24.23	Difference	1.55 (2.36)	18	−16
	1 s	59.2 (19.5)	100	0	1 s	9.57 (13.76)	152	0
CS	5 s	59.36 (20.11)	100	0	5 s	9.79 (13.69)	154	0
	Difference	0.16 (6.65)	29.51	−37.76	Difference	0.22 (3.7)	17	−39
	1 s	91.38 (15.02)	100	0	1 s	32.08 (25.22)	256	0
ILS	5 s	93.84 (12.71)	100	0	5 s	35.34 (24.8)	239	0
	Difference	2.47 (5.36)	53.24	−35.98	Difference	3.26 (4.05)	66	−31

Note: 1) Table 11 reports the statistics of futures price discovery and leading time indicators under various index update frequencies. 2) The unit of measurement of the price discovery indicator is in percentage. The unit of measurement of the leading time indicator is in seconds. The standard deviation is in parentheses. 3) All three indicators are overestimated under the 5-s information-update frequency when compared to the 1-s case. 4) The minimal overestimated values are negative. The leading time ahead of the index is overvalued by 1.55 s (IS-LT), 0.22 s (CS-LT), and 3.26 s (ILS-LT).

index futures. By discussing the coefficients of the VECM model, Huang et al. (2021) find that deregulation strengthens the lead role of futures, which is inconsistent with our results. However, their approach is similar to CS and also fails to take into account the key information about linkages between markets implied by the covariance matrix of the residual terms in the VECM model.

4.5. Impact of asynchronous informational updates on price discovery between futures and spot markets

In reality, the index/futures update frequency a trader can observe is 5/0.5 s. When futures update their prices, the index may not necessarily change. Therefore, the asynchronous informational updates between futures and spot will likely cause the traders to believe that futures lead the index, reducing the efficiency of information discovery in the spot market. To examine the impact of market update frequency on price discovery, we reconstruct 1-s index data that is updated every five seconds to mimic the index observed and calculate price discovery indicators. Then we compare the results with an index update frequency of 1 and 5 s. Panel A of Table 11 reports the statistics of futures price discovery indicators and their difference under two different index update frequencies. The results illustrate that all three indicators are overestimated under the 5-s information-update frequency when compared to the 1-s case.

It is worth noting that the minimal overestimated values are negative across the three cases. This implies that over some trading days, the price discovery indicator based on data updated every five seconds is higher than those under the 1-s specification. We argue that when the index update frequency is reduced from one to five seconds, the index remains unchanged between updates and index noise is much smaller than the noise of futures price, resulting in a decline of the futures price discovery indicator.

Panel B of Table 11 reports the statistics on futures lead times under different index-information update frequencies. When the index update frequency is reduced from one second to five seconds, the LT of CSI 300 stock index futures increases, indicating that the LT ahead of the index is overvalued by 1.55 s (IS-LT), 0.22 s (CS-LT), and 3.26 s (ILS-LT), respectively. In general, due to the asynchronous informational updates between futures and index markets, the futures price discovery ability is overestimated. Such overvaluation is more evident when the LT of stock index futures is relatively short.[12]

5. Discussions

5.1. Policy implications

This paper confirms that during the stock market crisis of 2015, price discovery in stock index futures was indeed stronger than in the stock market. On the surface, stock index futures did play a "leading" role in the market crisis. However, it is worth noting that this so-called stock index futures "lead" the stock market crash is actually a misunderstanding of the essence of price discovery. Stock index futures reflect market prices earlier simply because informed traders have a sense of what the market expects in the future and thus react earlier in the stock index futures market. The price discovery function of stock index futures is based on the trend of the stock market itself and cannot really determine the stock market price. The decline in the stock market is attributed to the valuation of the stock market itself. The Chinese stock market is a market where retail investors are in the majority. Misunderstanding of price discovery of stock index futures creates public opinion pressure, which is not conducive to the supervision of stock index futures by the regulators, and to a certain extent hinders the healthy development of the stock index futures market. Therefore, regulators should guide and promote investors to correctly understand the price discovery function of stock index futures.

Our results show that the price discovery ability of stock index futures is indeed reduced under extremely severe restrictions during extremely volatile market conditions. However, such temporary regulatory policies can only reduce speculative behavior in the futures market in a short period of time. On the other hand, the weakened price discovery of stock index futures reflects the relative separation of the futures and the spot markets. When the market falls, the risk-hedging function of stock index futures will become more important. Separation of the spot and futures markets is detrimental to investors' risk management. Therefore, regulators should take measures to restore the good functioning of the stock index futures market as soon as possible after the market has gradually recovered.

This paper confirms that a low market update overestimates the stock index futures price discovery and reduces the efficiency of the spot market. Therefore, it is recommended that the regulators adjust the update frequency of the stock index to be in line with the stock index futures. This can promote the speed of information dissemination in the stock market and also reduce arbitrage caused by the asynchrony. For investors, market update asynchrony exaggerates the price discovery function of futures. Investors can construct quantitative strategies to trade based on this fact.

5.2. Practical implications

In this paper, we are able to quantify the lead time of futures with respect to the corresponding indices and identify the lead time differences during the market crash, trading restriction and deregulation periods. These results are of great value for high-frequency traders in particular as they can now optimize their trading strategies based on the time series of lead time of futures. Currently, traders only have a rough idea from their daily trading experience about how much faster the futures market responds to information shock relative to their corresponding spot markets. And they do not differentiate between the effect of information update asynchrony and that of the index futures themselves. Our results can certainly help these futures traders to better understand the futures information edge and accordingly refine their arbitrage strategies across the spot and futures markets, leveraging on the fact that they now know the time lapse before information gets equally reflected in the spot market.

[12] CSI 500 stock index futures achieve similar results, but by a relatively small percentage of overvalued leading times. The possible reason may be that CSI 500 stock index futures lead CSI 500 index much longer; therefore, the frequency of market updates has a small impact on price discovery. The results of CSI 500 futures are available upon request.

It will be important to estimate the significance of the potential economic benefits generated from our findings by testing a simple arbitrage strategy. However, in our sample, the price discovery of futures comes from two sources: one due to the futures trading itself, the other due to the information update asynchrony. It would be very difficult at this point to differentiate between and to separate out these two effects. But if later on the Chinese regulators change the policy and the information update frequency of futures and spot becomes the same, then we may be able to obtain an accurate estimate of the economic significance from information update asynchrony by comparing the arbitrage benefits before and after the policy change.

5.3. Limitations and future research

This paper has also some limitations. First, the LT metrics proposed in this paper are mainly used to discuss price discovery between two markets. Constructing metrics to measure the lead-lag time difference across multiple markets is a future research direction.

Second, our results show that ISD can be used to analyze whether the data frequency is adequate or not. The ILS indicator is reliable only when the ISD is small, or the data frequency is high. However, we don't know what range of ISD would guarantee the reliability of ISD. It is therefore necessary to find clear guidelines for determining whether the frequency of the data selected is sufficiently high.

In many studies, high-frequency data are not available. The price discovery indicators used in this paper are problematic with low-frequency data. How to construct noise-robust price discovery indicators that are also robust to low-frequency data is thus an important direction for future research.

Lastly, the seasonality of trading during a trading day may have a differential impact on the price discovery of futures and it will be very interesting to see how the effect of information update asynchrony may vary across different parts of the trading.[13]

6. Conclusion

Through Monte Carlo simulations, we find that the price discovery indicator ILS may be biased under low-frequency data; the indicator is more suitable with high-frequency data. We propose an LT indicator to quantify the lead-lag relationship between futures and index. Based on 1-s data, we observe that the Chinese stock index futures possess a fair price discovery function, with an average LT of 32.23 s (CSI 300 stock index futures) and 46.53 s (CSI 500 stock index futures) ahead of the underlying indices. In the later period of the sample, we find evidence of a decrease in the LT of futures, indicating a shorter time gap between futures and spot prices. Moreover, during the same period, an enhanced synergy between the two markets is also observed. This implies that the information transmission within the two financial assets is more engaging. Past literature using low-frequency data (i.e., one- or five-minute data) seems unsuitable for analyzing the current financial market. In other words, the interpretations of IS, CS, and ILS must be handled with caution, and ILS may no longer be reliable. A few studies advocate the superiority of IS over CS (Baillie et al., 2002; Yan and Zivot, 2010); however, our results show that CS is not susceptible to data frequency. Therefore, in cases where researchers do not have access to higher-frequency data or when the research problem does not necessitate its use, the CS can serve as a dependable metric for measuring price discovery.

The asynchrony in the market update on futures and spot causes the overestimation of the price discovery ability of futures and the deceleration of the information transmission speed in the stock market. Under the existing condition where information update on futures and spot is not synchronized, we find that the LT of futures is prolonged by about three seconds, purely caused by the technical issues of the stock exchange. This delays the process of price discovery and jeopardizes the informational efficiency of the market. In particular, price discovery spills over to the index futures and index-tracking ETFs,[14] which may lead to speed arbitrage opportunities between the two markets. To this end, the regulatory bodies ought to urgently alter the current frequency of informational updates on the stock index to ensure that China's securities market can further align with international standards.

CRediT authorship contribution statement

Qian Han: Visualization, Investigation, Conceptualization, Supervision. **Chengzhi Zhao:** Conceptualization, Methodology, Software, Data curation, Writing – original draft, Validation. **Jing Chen:** Visualization, Investigation, Supervision, Conceptualization. **Qian Guo:** Investigation, Conceptualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that there is no conflict of interest.

Acknowledgements

Chengzhi Zhao expresses gratitude for the valuable feedback received from the discussants at the SKKU-ZJU International Conference on Economics and Finance.

[13] We thank the anonymous reviewer for raising this possibility.

[14] There are 25 index tracking ETFs (from 13 indices) in the Chinese stock market. (https://www.justetf.com/en/how-to/invest-in-china.html)

Appendix A. Appendix

Table A1
Standard deviations of futures price discovery indicators under diverse data frequencies $\delta = 0$.

Panel A: ISD									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	0.85	2.02	2.34	2.69	2.83	3.27	3.9	4.22	3.77
(3, 6)	0.77	1.94	2.37	2.57	2.71	3.26	4	5.31	5.67
(6, 6)	0.82	1.74	2.07	2.16	2.35	2.66	2.96	3.75	3.81
(6, 3)	0.8	2.01	2.39	2.5	2.71	3.37	3.97	5.24	5.88
(4, 2)	0.86	2.07	2.39	2.65	2.84	3.39	3.87	4.22	3.91

Panel B: IS									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	1.53	2.27	2.37	2.63	2.78	3.03	3.57	3.81	3.8
(3, 6)	1.71	2.71	2.88	2.85	2.92	3.29	3.98	4.87	5.38
(6, 6)	2.4	3.66	3.82	3.99	4.24	4.23	4.65	5.75	6.63
(6, 3)	1.74	2.6	2.82	2.79	2.96	3.46	3.95	4.76	5.27
(4, 2)	1.52	2.21	2.38	2.64	2.76	3.16	3.52	3.75	3.75

Panel C: CS									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	1.01	2.12	2.48	3.05	3.35	4.43	6.2	10.41	13.98
(3, 6)	1.03	2.06	2.35	2.51	2.63	3.35	4.59	7.93	11.1
(6, 6)	1.36	2.47	2.74	3.04	3.29	3.67	4.46	7.29	10.44
(6, 3)	1.04	1.94	2.29	2.47	2.69	3.56	4.56	7.78	10.81
(4, 2)	0.99	2.02	2.45	3.07	3.34	4.68	6.11	10.26	13.95

Panel D: ILS									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	1.6	1.78	2.39	3.15	3.49	4.87	6.68	10.6	15.49
(3, 6)	2.16	2.11	2.13	2.27	2.4	3.23	4.79	8.44	11.29
(6, 6)	2.16	2.65	2.53	2.39	2.43	1.96	1.85	3.95	8.49
(6, 3)	2.2	2.12	2.12	2.31	2.44	3.52	4.71	8.35	10.94
(4, 2)	1.6	1.69	2.32	3.17	3.51	5.22	6.6	10.35	14.99

Table A2
Standard deviations of futures price discovery indicators under diverse data frequencies $\delta = 2$.

Panel A: ISD									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	0.51	2.06	2.41	2.92	3.15	3.7	4.41	4.73	4.31
(3, 6)	0.57	1.89	2.36	2.62	2.8	3.43	4.22	5.66	5.99
(6, 6)	0.8	1.76	2.1	2.2	2.43	2.74	3.03	3.83	3.96
(6, 3)	0.82	1.85	2.19	2.3	2.46	3.04	3.58	4.89	5.55
(4, 2)	0.81	1.69	1.96	2.11	2.23	2.63	3.05	3.76	3.81

Panel B: IS									
$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	0.82	2.02	2.15	2.55	2.74	3.05	3.64	4.02	4.04
(3, 6)	1.19	2.4	2.63	2.71	2.84	3.24	3.96	4.99	5.46
(6, 6)	2.39	3.63	3.84	4	4.27	4.26	4.74	5.88	6.79
(6, 3)	2.23	2.89	3.09	3.02	3.16	3.66	4.18	5.01	5.58
(4, 2)	2.05	2.54	2.63	2.92	3.04	3.39	3.8	4.18	4.34

(continued on next page)

Table A2 (*continued*)

Panel B: IS

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s

Panel C: CS

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	1.06	2	2.26	2.91	3.2	4.16	5.87	9.94	13.29
(3, 6)	1.09	1.98	2.25	2.44	2.58	3.25	4.48	7.86	10.67
(6, 6)	1.44	2.41	2.7	2.98	3.22	3.58	4.42	7.22	10.39
(6, 3)	1.11	1.89	2.23	2.41	2.59	3.47	4.47	7.68	11.04
(4, 2)	1.14	1.94	2.29	2.91	3.16	4.4	5.84	10.33	14.87

Panel D: ILS

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	2.16	2.72	3.25	4.17	4.42	5.42	7.02	10.04	14.21
(3, 6)	2.79	2.84	2.81	2.85	3.01	3.67	5.31	8.71	10.95
(6, 6)	2.22	2.78	2.71	2.56	2.64	2.15	1.97	3.67	8.2
(6, 3)	1.65	2.07	1.96	1.85	1.88	2.42	3.41	7.71	11.31
(4, 2)	1.18	1.19	1.15	1.52	1.71	3.48	5.37	11.3	17.08

Table A3

Standard deviations of futures price discovery indicators under diverse data frequencies $\delta = 10$.

Panel A: ISD

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	0.04	0.57	1.1	1.45	1.61	3.25	4.49	5.79	5.45
(3, 6)	0.07	0.96	1.46	1.92	1.96	2.98	4.18	6.05	6.65
(6, 6)	0.44	1.73	2.19	2.39	2.62	3.07	3.43	4.35	4.62
(6, 3)	0.51	1.77	2.12	2.25	2.41	2.69	2.94	3.78	4.15
(4, 2)	0.24	1.68	2.13	2.36	2.5	2.75	2.82	2.85	2.94

Panel B: IS

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	0.08	0.59	0.94	1.12	1.2	2.23	3.09	4.14	4.24
(3, 6)	0.21	1.24	1.65	1.91	1.83	2.5	3.41	4.78	5.38
(6, 6)	1.48	3.28	3.9	3.72	4.08	4.29	4.7	6.05	7.03
(6, 3)	1.45	3.49	3.95	3.97	3.55	4.02	4.63	5.59	6.32
(4, 2)	0.48	2.39	2.79	2.77	2.95	3.64	4.19	4.89	5.32

Panel C: CS

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	0.79	2.05	2.23	2.53	2.78	3.68	4.74	7.93	10.34
(3, 6)	1.16	2.14	2.53	2.76	2.49	3.08	4.16	6.96	9.09
(6, 6)	2.19	2.65	3.09	2.97	3.26	3.5	4.13	6.9	9.92
(6, 3)	2.36	2.35	2.67	2.76	2.5	3.17	4.12	7.2	10.76
(4, 2)	2.66	2.52	2.63	2.62	2.87	3.9	5.07	9.27	13.97

Panel D: ILS

$(\sigma_{s1}, \sigma_{s2})$	1 s	5 s	7 s	9 s	10 s	15 s	20 s	40 s	60 s
(2, 4)	1.18	8.19	5.98	10.37	13.25	14.12	12.05	9.65	10.82
(3, 6)	4.02	4.3	4.53	6.75	4.98	7.46	9.61	10.28	10.45
(6, 6)	1.46	3.03	3.36	3.11	3.28	3.03	2.83	3.6	7.56
(6, 3)	0.47	1.73	2.02	1.98	1.81	1.74	1.58	4.43	9.82
(4, 2)	0.16	1.12	1.39	1.42	1.44	1.54	2.23	9.1	17.04

Appendix B. Appendix

Table B1
Statistics of the differences in price discovery indicators calculated from two types of high-frequency data for 2017.

	CSI 300 Stock Index Futures				CSI 500 Stock Index Futures			
	$\lvert\Delta ISD\rvert$	$\lvert\Delta IS\rvert$	$\lvert\Delta CS\rvert$	$\lvert\Delta ILS\rvert$	$\lvert\Delta ISD\rvert$	$\lvert\Delta IS\rvert$	$\lvert\Delta CS\rvert$	$\lvert\Delta ILS\rvert$
Mean	0.53	1.89	6.23	5.72	0.51	2.36	4.65	1.26
Std. Dev	0.54	2.71	3.32	5.24	0.49	3.62	5.62	2.46
Max	3.39	20.4	32.35	24.32	2.25	26.04	58.63	23.19
Min	0	0.02	0.02	0.01	0	0.01	0.01	0

Note: $\lvert\Delta ISD\rvert$, $\lvert\Delta IS\rvert$, $\lvert\Delta CS\rvert$, and $\lvert\Delta ILS\rvert$ represent the absolute values of difference of the corresponding price discovery indicators calculated from the two types of data on 2017 stocks.

Appendix C. Appendix

When CS is calculated, the coefficients of the error correction term in the estimated VECM model, $\alpha = (\alpha_1, \alpha_2)$, generally have the opposite signs, i.e., $\alpha_1 < 0$ and $\alpha_2 > 0$. However, in empirical research, it is not unusual for the α coefficients to have the same sign. This does not imply that the cointegration relation is invalid.[15] According to Bohl et al. (2011), when $\alpha_1 < 0, \alpha_2 < 0$, one can guarantee a long-term equilibrium between the two prices as long as $\mid \alpha_2 \mid < \mid \alpha_1 \mid$. Similarly, when $\alpha_1 > 0, \alpha_2 > 0$, one can guarantee a long-term equilibrium between the two prices if $\mid \alpha_2 \mid > \mid \alpha_1 \mid$. Therefore, it is possible that the coefficients α have the same sign. Under such a situation, although the sum of the CS of the two prices is still 1, the CS value may lie outside $[0, 1]$, which somewhat deviates from the original meaning of CS of *share* of information. Nevertheless, although the CS value does not lie within $[0, 1]$, it can be seen from the formula that IS and ILS still fall in between $[0, 1]$; hence, IS and ILS still preserve their nature of "information share." When the α coefficients of the estimated VECM have the same sign, there are generally two ways to deal with it to obtain CS. The first method is to take the absolute value of α

$$CS_1 = \frac{|\alpha_2|}{|\alpha_2| + |\alpha_1|}, CS_2 = \frac{|\alpha_1|}{|\alpha_2| + |\alpha_1|},$$

(see, for example, Bohl et al., 2011; Cabrera et al., 2009; Entrop et al., 2020)

Another method truncates the elements of α that do not meet expectations to 0 (see, for example, Alexander et al., 2020), in which case the CS is 100%. This is because when the α coefficients have the same sign, the elements that do not conform to expectations are usually much smaller and less significant than the other element. Truncating them to 0 seems a reasonable way to handle the problem. Therefore, CS is truncated in this paper. Since we use 1-s high-frequency data, the ISD is small. When CS is truncated to 100%, IS is close to 100% but less than 100%. In this case, the ILS is undefined since the ILS formula uses the CS as a denominator. Since both IS and CS are very high, the market's price discovery function can be considered quite strong. Thus, we define such an ILS as 100%.

Appendix D. Appendix

Table D1
Major regulatory policies on stock index futures.

Notification Date	Effective Date	Adjustment of Daily Closing Premium	Adjustment of Daily Excessive Trades	Margin Adjustment			
				Non-Hedging Margin		Hedging Margin	
		IF, IC, IH	IF, IC, IH	IF, IH	IC	IF, IH	IC
2015-07-08	2015-07-08				10%→ 20%		
2015-07-08	2015-07-09				20%→ 30%		
2015-07-31	2015-08-03	0.23/10,000					
2015-08-25	2015-08-26	1.15/10,000		10%→ 12%	10%→ 12%		
2015-08-25	2015-08-27			12%→ 15%	12%→ 15%		

(continued on next page)

[15] Bohl et al. (2011); Kapar and Olmo (2019),; Xu and Liu (2019); and Xu et al. (2021) provide the α estimates of the same sign.

Table D1 (*continued*)

Notification Date	Effective Date	Adjustment of Daily Closing Premium	Adjustment of Daily Excessive Trades	Margin Adjustment			
				Non-Hedging Margin		Hedging Margin	
		IF, IC, IH	IF, IC, IH	IF, IH	IC	IF, IH	IC
2015-08-25	2015-08-28			15%→ 20%	15%→ 20%		
2015-08-28	2015-08-31		100 contracts	20%→ 30%	20%→ 30%		
2015-09-02	2015-09-07	23/10,000	10 contracts	30%→ 40%	30%→ 40%	10%→ 20%	10%→ 20%
2017-02-16	2017-02-17	9.2/10,000	20 contracts	40%→ 20%	40%→ 30%		
2017-09-15	2017-09-18	6.9/10,000		20%→ 15%		20%→ 15%	
2018-12-02	2018-12-03	4.6/10,000	50 contracts	15%→ 10%	30%→ 15%	15%→ 10%	20%→ 15%
2019-04-19	2019-04-22	3.45/10,000	500 contracts		15%→ 12%		15%→ 12%

Note: The data is compiled from the official website of the China Financial Futures Exchange.

Appendix E. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ememar.2025.101307.

Data availability

Data will be made available on request.

References

Alexander, C., Choi, J., Park, H., Sohn, S., 2020. BitMEX bitcoin derivatives: Price discovery, informational efficiency, and hedging effectiveness. J. Futur. Mark. 40 (1), 23–43. https://doi.org/10.1002/fut.22050.

Baillie, R.T., Booth, G.G., Tse, Y., Zabotina, T., 2002. Price discovery and common factor models. J. Financ. Mark. 5 (3), 309–321. https://doi.org/10.1016/S1386-4181(02)00027-7.

Beveridge, S., Nelson, C.R., 1981. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. J. Monet. Econ. 7 (2), 151–174. https://doi.org/10.1016/0304-3932(81)90040-4.

Bohl, M.T., Salm, C.A., Schuppli, M., 2011. Price discovery and investor structure in stock index futures. J. Futur. Mark. 31 (3), 282–306. https://doi.org/10.1002/fut.20469.

Brandvold, M., Molnár, P., Vagstad, K., Andreas Valstad, O.C., 2015. Price discovery on bitcoin exchanges. J. Int. Financ. Mark. Inst. Money 36, 18–35. https://doi.org/10.1016/j.intfin.2015.02.010.

Brooks, C., Rew, A.G., Ritson, S., 2001. A trading strategy based on the lead-lag relationship between the spot index and futures contract for the FTSE 100. Int. J. Forecast. 17, 31–44. https://doi.org/10.1016/S0169-2070(00)00062-5.

Buckle, M., Chen, J., Guo, Q., Tong, C., 2018. Do ETFs lead the price moves? Evidence from the major US markets. Int. Rev. Financ. Anal. 58, 91–103. https://doi.org/10.1016/j.irfa.2017.12.005.

Cabrera, J., Wang, T., Yang, J., 2009. Do futures lead price discovery in electronic foreign exchange markets? J. Futur. Mark. 29 (2), 137–156. https://doi.org/10.1002/fut.20352.

Chan, K., 1992. A further analysis of the lead-lag relationship between the cash market and stock index futures market. Rev. Financ. Stud. 5 (1), 123–152. https://doi.org/10.1093/rfs/5.1.123.

Chen, Y.L., Gau, Y.F., 2010. News announcements and price discovery in foreign exchange spot and futures markets. J. Bank. Financ. 34 (7), 1628–1636. https://doi.org/10.1016/j.jbankfin.2010.03.009.

Chu, Q.C., Hsieh, W.-L.G., Tse, Y., 1999. Price discovery on the S&P 500 index markets: an analysis of spot index, index futures, and SPDRs. Int. Rev. Financ. Anal. 8 (1), 21–34. https://doi.org/10.1016/S1057-5219(99)00003-4.

Chung, K.H., Chuwonganant, C., 2023. Tick size and price efficiency: further evidence from the tick size pilot program. Financ. Manag. 52, 483–511. https://doi.org/10.1111/fima.12419.

Da Fonseca, J., Zaatour, R., 2017. Correlation and lead-lag relationships in a Hawkes microstructure model. J. Futur. Mark. 37 (3), 260–285. https://doi.org/10.1002/fut.21800.

De Jong, F., 2002. Measures of contributions to price discovery: a comparison. J. Financ. Mark. 5 (3), 323–327. https://doi.org/10.1016/S1386-4181(02)00028-9.

De Jong, F., Nijman, T., 1997. High frequency analysis of lead-lag relationships between financial markets. J. Empir. Financ. 4 (2), 259–277. https://doi.org/10.1016/S0927-5398(97)00009-1.

Entrop, O., Frijns, B., Seruset, M., 2020. The determinants of price discovery on bitcoin markets. J. Futur. Mark. 40 (5), 816–837. https://doi.org/10.1002/fut.22101.

Gong, C., Ji, S., Su, L., Li, S., Ren, F., 2016. The lead-lag relationship between stock index and stock index futures: a thermal optimal path method. Phys. A 444, 63–72. https://doi.org/10.1016/j.physa.2015.10.028.

Gonzalo, J., Granger, C., 1995. Estimation of common long-memory components in cointegrated systems. J. Bus. Econ. Stat. 13 (1), 27–35. https://doi.org/10.1080/07350015.1995.10524576.

Guo, B., Han, Q., Liu, M., Ryu, D., 2013. A tale of two index futures: the intraday price discovery and volatility transmission processes between the China financial futures exchange and the Singapore exchange. Emerg. Mark. Financ. Trade 49 (sup4), 197–212. https://doi.org/10.2753/REE1540-496X4905S414.

Han, Q., Liang, J., 2017. Index futures trading restrictions and spot market quality: evidence from the recent Chinese stock market crash. J. Futur. Mark. 37 (4), 411–428. https://doi.org/10.1002/fut.21825.

Hao, J., Xiong, X., He, F., Ma, F., 2019. Price discovery in the Chinese Stock index futures market. Emerg. Mark. Financ. Trade 55 (13), 2982–2996. https://doi.org/10.1080/1540496X.2019.1598368.

Hasbrouck, J., 1995. One security, many markets: determining the contributions to price discovery. J. Financ. 50 (4), 1175–1199. https://doi.org/10.1111/j.1540-6261.1995.tb04054.x.

Hasbrouck, J., 2002. Stalking the "efficient price" in market microstructure specifications: an overview. J. Financ. Mark. 5 (3), 329–339. https://doi.org/10.1016/S1386-4181(02)00029-0.

Hasbrouck, J., 2021. Price discovery in high resolution. J. Financ. Econ. 19 (3), 395–430. https://doi.org/10.1093/jjfinec/nbz027.

He, F., Liu-Chen, B., Meng, X., Xiong, X., Zhang, W., 2020. Price discovery and spillover dynamics in the Chinese stock index futures market: a natural experiment on trading volume restriction. Quant. Financ. 20 (12), 2067–2083. https://doi.org/10.1080/14697688.2020.1814037.

Hsieh, W.L.G., Lee, C.S., Yuan, S.F., 2008. Price discovery in the options markets: an application of put-call parity. J. Futur. Mark. 28 (4), 354–375. https://doi.org/10.1002/fut.20302.

Huang, W., Luo, J., Qian, Y., Zheng, Y., 2021. The impact of decreased margin requirements on futures markets: Evidence from CSI 300 index futures. Emerg. Mark. Financ. Trade 57 (7), 2052–2064. https://doi.org/10.1080/1540496X.2020.1852925.

Jiang, T., Bao, S., Li, L., 2019. The linear and nonlinear lead-lag relationship among three SSE 50 index markets: the index futures, 50ETF spot and options markets. Phys. A 525, 878–893. https://doi.org/10.1016/j.physa.2019.04.056.

Kapar, B., Olmo, J., 2019. An analysis of price discovery between bitcoin futures and spot markets. Econ. Lett. 174, 62–64. https://doi.org/10.1016/j.econlet.2018.10.031.

Kawaller, I.G., Koch, P.D., Koch, T.W., 1987. The temporal price relationship between S&P 500 futures and the S&P 500 index. J. Financ. 42 (5), 1309–1329. https://doi.org/10.1111/j.1540-6261.1987.tb04368.x.

Lien, D., Shrestha, K., 2009. A new information share measure. J. Futur. Mark. 29 (4), 377–395. https://doi.org/10.1002/fut.20356.

Lin, H., Wang, Y., 2018. Are tightened trading rules always bad? Evidence from the Chinese index futures market. Quant. Financ. 18 (9), 1453–1470. https://doi.org/10.1080/14697688.2018.1445586.

Miao, H., Ramchander, S., Wang, T., Yang, D., 2017. Role of index futures on China's stock markets: evidence from price discovery and volatility spillover. Pac. Basin Financ. J. 44, 13–26. https://doi.org/10.1016/j.pacfin.2017.05.003.

Putniņš, T.J., 2013. What do price discovery metrics really measure? J. Empir. Financ. 23, 68–83. https://doi.org/10.1016/j.jempfin.2013.05.004.

Rittler, D., 2012. Price discovery and volatility spillovers in the European Union emissions trading scheme: a high-frequency analysis. J. Bank. Financ. 36 (3), 774–785. https://doi.org/10.1016/j.jbankfin.2011.09.009.

Stock, J.H., Watson, M.W., 1988. Testing for common trends. J. Am. Stat. Assoc. 83 (404), 1097–1107. https://doi.org/10.1080/01621459.1988.10478707.

Stoll, H.R., Whaley, R.E., 1990. The dynamics of stock index and stock index futures returns. J. Financ. Quant. Anal. 25 (4), 441–468. https://doi.org/10.2307/2331010.

Tse, Y., Chan, W., 2010. The lead-lag relation between the S&P500 spot and futures markets: an intraday-data analysis using a threshold regression model. Jpn. Econ. Rev. 61 (1), 133–144. https://doi.org/10.1111/j.1468-5876.2009.00481.x.

Wu, L., Ma, J., 2014. A review on the measurement of price discovery. China Econ. Quart. 13 (1), 399–424. https://doi.org/10.13821/j.cnki.ceq.2014.01.015.

Xu, R., Liu, C., 2019. How restrictive policy on trade affects the relationship between futures and spot?——an empirical test on the price discovery of Stock index futures. J. Financ. Res. 464 (2), 154–168.

Xu, K., Xiong, X., Shen, D., 2021. The impact of transaction cost on the price discovery of Stock index futures: evidence from the stock market crash. Econ. J. 1 (3), 637–654. https://doi.org/10.12012/CJoE2020-0017.

Yan, B., Zivot, E., 2010. A structural analysis of price discovery measures. J. Financ. Mark. 13, 1–19. https://doi.org/10.1016/j.finmar.2009.09.003.

Yang, J., Li, Z., Wang, T., 2021. Price discovery in Chinese agricultural futures markets: a comprehensive look. J. Futur. Mark. 41 (4), 536–555. https://doi.org/10.1002/fut.22179.