Time-varying price discovery in fragmented markets.*

Nicholas Taylor†
Cardiff University

June 5, 2008

Abstract

This paper examines temporal aspects of the price discover process in the (fragmented) S&P 500 market. This is achieved by augmenting the coefficients in the model upon which the price discovery measures are based, by a set of time-varying (theoretically-implied) scaling factors. The factors considered can be characterized as those that measure market liquidity and those that measure the degree of information asymmetry that exists at a particular time. Regarding the latter measures, this feature of financial markets is assessed by considering, inter alia, price discovery around the release of key macroeconomic information. Using high frequency data from five constituent S&P 500 index markets, the results provide two main insights into price discovery in this fragmented market. First, the majority of price discovery appears to occur in the market for the individual stocks making up the index and in the (electronically traded) E-mini futures market. And second, the E-mini futures market becomes the dominant price discovery market only during periods of extreme information asymmetry and when this market is liquid – a finding that supports theoretical arguments proposed in related literature.

Key Words: Price discovery, information asymmetry, macroeconomic announcements.

JEL Classification Code: C32; G14.

*I thank Cardiff University for funding received while undertaking this paper, John Huang of *Tick Data Inc.* for supplying the financial data used in this paper, and Kevin Evans for help received while collecting the macroeconomic announcement data.

†Cardiff Business School, Cardiff University, Cardiff CF10 3EU, United Kingdom. `Email:` `taylorn@cardiff.ac.uk`.

In addition to providing a (liquid) trading mechanism, financial markets enable price discovery; that is, '*the impounding of new information into the security price*' (Hasbrouck, 1995). While the liquidity provision role of financial markets has been extensively studied in the microstructure literature, the latter role has only recently come under scrutiny. The primary reason for this is the institutional trend toward multiple trading of securities on competing markets, with market diversity usually facilitated by the introduction of electronic trading systems.[1] Therefore, as markets (and order flow) become more fragmented, the relative performance of the constituent markets (in terms of their respective contribution to the price discovery process) becomes increasingly relevant to investors and trading system designers alike.[2]

The market fragmentation trend has also fueled a debate in the finance literature concerning the relative merits of the two trading systems prevalent in fragmented markets, viz., open outcry (quote-driven dealer markets) and electronic trading systems (order-driven auction markets). In turn, given the preferences of informed (as opposed to uninformed) traders for, inter alia, anonymity in the trading process, the relative performance of these two systems is also informative regarding the nature of the trading activities of these types of traders and, ultimately, regarding price determination in financial markets.[3] Therefore, by examining measures of price discovery in fragmented markets one can shed light on the relative benefits (and costs) of the above trading system designs, and on issues pertaining to trading in the presence of asymmetric information.

Though fragmented markets have long been the subject of academic inquiry (see Garbade and Silber, 1979, 1982, for early examples), it is only since the work of Hasbrouck (1995) that richer characterizations of price discovery have been possible; see Harris et al (2002), de Jong and Schotman (2005), and Yan and Zivot (2006a). Each of these methodological approaches share the assumption that a common implicit efficient price exists, which, in turn, is implied by the presence of a cointegrated system of constituent market prices. Moreover, each measure of price discovery is based on innovations to this efficient price. For instance, Hasbrouck (1995) defines a market's contribution to price discovery as the proportion of the efficient price innovation variance attributed to that market. However, while these common assumptions are well motivated (both theoretically and empirically), the existing methodologies also share an assumption that is clearly violated on both theoretical and empirical terms; specifically, that price discovery does not vary over time. It is this particular assumption that is relaxed in this paper.

Use of the time-invariant price discovery assumption is somewhat surprising given the fact that likely covariates to the relative amount of price discovery vary considerably over time. For instance, Admati and Pfleiderer (1988) demonstrate (theoretically) that the clustering of information acquisition and informed trading are the result of liquidity-motivated trading.[4] Moreover, Lei and Wu (2005) and Ting (2006) both find (empirically) that measures of information asymmetry vary over time, with the latter finding variation over the trading day and around the release of public information. Given this strong evidence, we introduce a model that allows price discovery to vary over time. This is achieved by allowing the coefficients in the model upon which measures of price discovery are based, to covary with relevant theoretically motivated market-based measures. Indeed, to anticipate some of the results, we find that using data from the S&P 500 market, this covariation is both statistically and economically significant, and results in measures of price discovery that vary over the trading day, around the release of macroeconomic information, and with respect to measures of information asymmetry.

The rest of the paper is organized as follows. The next section describes the framework within which the above issues are examined, while Section II describes the hypotheses tested using this framework. Section III contains results pertaining to various aspects of the estimated models, with Section IV focusing on hypothesis tests relating to these models and their products, viz., the price discovery measures. The final section concludes. In addition, a technical appendix is also included that contains derivations of the price discovery measures used in this paper.

I. Return dynamics and price discovery

In this section a description of the model used to assess the extent of price discovery in a fragmented market is given, and measures of price discovery are described in the context of this model (see also Appendix A). In both cases, it is assumed that this market consists of M constituent markets, each of which competes for the order flow associated with identical or closely related securities.

A. The econometric model

In this subsection a generalized model of returns is described that represents an augmented version of an existing model of returns used in the context of price discovery. This augmentation is achieved by relaxing the assumption that the coefficient and variance-covariance matrices in such a model do not vary over time. In addition to being able to deliver measures of price discovery, the model also possesses the desirable property of being able to incorporate the complex dependencies (including periodicities) inherent in high frequency financial data.

The above requirements are satisfied by the following *linearly scaled* vector equilibrium correction model,

$$\Delta \mathbf{p}_t = \mathbf{\Psi}_0 \mathbf{h}_t + \mathbf{\Psi}(L)(\Delta \mathbf{p}_{t-1} \otimes \mathbf{h}_t) + \boldsymbol{\alpha}(\mathbf{z}_{t-1} \otimes \mathbf{h}_t) + \boldsymbol{\epsilon}_t, \tag{1}$$

where $\mathbf{p}_t$ is an $(M \times 1)$ vector of log prices in the constituent markets, that is, $\mathbf{p}_t' := [p_{1t}, \ldots, p_{Mt}]$; $\mathbf{h}_t$ is an $(R \times 1)$ vector of scaling factors (including unity); $\mathbf{\Psi}_0$ is an $(M \times R)$ matrix of coefficients; $\mathbf{\Psi}(L) = \mathbf{\Psi}_1 + \mathbf{\Psi}_2 L + \ldots + \mathbf{\Psi}_{K-1} L^{K-2} + \mathbf{\Psi}_K L^{K-1}$, is a polynomial matrix with $(M \times MR)$ coefficient matrix elements $\{\mathbf{\Psi}_k\}_{k=1}^{K-1}$; $\boldsymbol{\alpha}$ is an $(M \times NR)$ matrix of adjustment coefficients; $\mathbf{z}_{t-1} := \boldsymbol{\beta}' \mathbf{p}_{t-1} - \boldsymbol{\mu}$, such that $\mathbf{z}_{t-1}' := [z_{1t-1}, \ldots, z_{Nt-1}]$, has bilateral mispricing elements $\{z_{nt-1}\}$ defined as the (demeaned) difference between the (log) price in the 1st market and the (log) price in the uth market, that is, $z_{nt-1} := p_{1t-1} - p_{ut-1} - \mu_n, \; \forall \, u \in \{2, \ldots, M\}$, $n \in \{1, \ldots, N\}$, $n = u - 1$, $N = M - 1$; $\boldsymbol{\beta}' := (\mathbf{1}_{N \times 1} \vdots - \mathbf{I}_N)$; $\mathbf{1}_{N \times 1}$ is an $(N \times 1)$ vector of ones; $\mathbf{I}_N$ is an N-dimension identity matrix; $\boldsymbol{\mu}$ is a $(N \times 1)$ vector containing the means of $\boldsymbol{\beta}' \mathbf{p}_{t-1}$; $\boldsymbol{\epsilon}_t' := [\epsilon_{1t}, \ldots, \epsilon_{Mt}]$, with $\boldsymbol{\epsilon}_t | \mathcal{I}_{t-1} \sim \mathcal{D}(\mathbf{0}_{M \times 1}, \mathbf{\Sigma}_t)$ such that $\mathcal{I}_{t-1}$ represents the information set available at time $t-1$, $\mathcal{D}$ is a continuous distribution with support over $(-\infty, \infty)$, mean equal to zero in all constituent markets, and variance-covariance matrix $\mathbf{\Sigma}_t$; $\mathbf{0}_{M \times 1}$ is an $(M \times 1)$ vector of zeros; and $\otimes$ is the Kronecker product operator.

Rearranging (1) gives the following equivalent time-varying coefficient representation of the above equilibrium correction model:

$$\Delta \mathbf{p}_t = \mathbf{\Psi}_{0t} + \mathbf{\Psi}_t(L) \Delta \mathbf{p}_{t-1} + \boldsymbol{\alpha}_t \mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t, \tag{2}$$

where the $(M \times 1)$ vector $\mathbf{\Psi}_{0t} \equiv \mathbf{\Psi}_0 \mathbf{h}_t$; the $(M \times M)$ matrix $\mathbf{\Psi}_t(L) \equiv \mathbf{\Psi}(L)(\mathbf{I}_M \otimes \mathbf{h}_t)$; and the $(M \times N)$ matrix $\boldsymbol{\alpha}_t \equiv \boldsymbol{\alpha}(\mathbf{I}_N \otimes \mathbf{h}_t)$. This equation states that the prices in each of the constituent markets are cointegrated, with equilibrium correction occurring via the actions of arbitrageurs seeking riskless (or very low risk) profit opportunities. That is, if a difference occurs between prices in the constituent markets, then arbitrageurs will take offsetting positions in the relevant securities. In doing this, prices will eventually revert back (correct) to their equilibrium levels and deliver profits to the arbitrageurs. However, in contrast to that assumed in the time-invariant parameter (i.e., *unscaled*) version of the equilibrium correction model, the rate at which prices correct to their equilibrium levels is functionally dependent upon market conditions that are, in turn, measured by a set of time-varying scaling factors.[5]

B. Measuring price discovery

The above models form the foundations of the augmented price discovery measures used in this paper. These measures are derived in Appendix A. As is clear from this appendix, the derivations are divided into three parts. The first part demonstrates the equivalence of an augmented version of Hasbrouck's (1993) unobservable efficient price model (see (A.1a), (A.1b), and (A.1c), in Appendix A), with the equilibrium correction model in (1), under a variety of different assumptions. Maintaining these assumptions, explicit expressions for the parameters in the former of these models are derived that are based on the parameters obtained from the latter model (i.e., on observable data). In turn, these are then used to derive a number of price discovery measures, each of which is based on the different assumptions made during the first stage of the derivations. This process eventually leads to the following price discovery measures: the augmented Harris et al (2002) component share measure (denoted $\mathbf{f}_{1t}$); the augmented Hasbrouck (1995) information share measure (denoted $\mathbf{f}_{2t}$); the augmented de Jong and Schotman (2005) information share measure (denoted $\mathbf{f}_{3t}$); and the augmented Yan and Zivot (2006a) non-standardized, standardized (relative), and aggregated price discovery efficiency loss measures (denoted $\mathbf{f}_{4K,t}$, $s\mathbf{f}_{4K,t}$, and $a \ln \mathrm{f}_{4K,t}$, respectively).[6]

Though extensive descriptions of all the above price discovery measures is beyond the scope of the current paper, for illustration purposes it is useful to demonstrate the intuition lying behind one of the measures.[7] Consider, for example, the augmented Harris et al (2002) component share measure. Under the assumptions necessary to derive this measure, the permanent innovation, η_t^P in (A.1b), is a function of the return shocks to each constituent market, as measured by $\boldsymbol{\epsilon}_t$ in (1). Moreover, this function is linear, with the marginal contribution of each constituent market's return shock to the permanent innovation given by the time-varying (observable) 'long-run impact' column vector, denoted $\boldsymbol{\theta}_t$. It is this measure, albeit standardized so that the elements sum to unity, that is referred to as the augmented Harris et al (2002) component share measure of price discovery. It differs from its traditional (non-augmented) counterpart in that the augmented version allows the marginal contributions to be functionally dependent on a series of time-varying scaling factors.[8] Consequently, it represents a measure of price discovery that is free from the assumption of time-invariance – an assumption undoubtedly violated given the time-varying nature of the trading process.

II. Hypotheses

In this section a number of hypotheses are proposed regarding the appropriateness of the models underlying the price discovery measures, and the nature of the measures themselves.

A. Preliminary hypotheses

The validity of the results presented in the subsequent section depend heavily on the appropriateness of the model of returns upon which the price discovery measures are based. As the primary objective of this paper is to introduce time-varying price discovery measures, then the appropriateness of the linearly scaled equilibrium correction model given by (1), versus the traditional (unscaled) version of this model, must be demonstrated. Therefore, the first null hypothesis considered in this paper can be formally stated as

H_0^1: *The linearly scaled and unscaled versions of the equilibrium correction model provide equitable representations of the data*,

where representativeness is measured in terms of (penalized) model fit, and via a series of Wald tests applied to the coefficients in the unrestricted (linearly scaled) model in order to yield the restricted (unscaled) functional form.

If the linearly scaled equilibrium correction model does provide a superior representation of the data then it is also of interest to determine which particular scaling factors drive the results. Formally, we test the null hypothesis

H_0^2: *The superiority of the linearly scaled version of the equilibrium correction model is not driven by the ith scaling factor*,

where the significance of the ith scaling factor is assessed via a Wald test applied to the ith scaling factor coefficients in the unrestricted (linearly scaled) model in order to yield a model containing no ith scaling factor coefficients.

B. Main hypotheses

A number of recent studies have focused on the role of trading system design in determining price discovery. Most notably, Hasbrouck (2003) examines relative price discovery in various US equity index markets, each composed of constituent markets that employ both electronic and non-electronic trading systems. Concerning the S&P 500 market, he finds that approximately 90% of the price discovery occurs in the electronically traded E-mini futures market – a result confirmed by the subsequent studies of Ates and Wang (2004), and Kurov and Lasser (2004).[9] Therefore, in order to replicate theses studies, the following null hypothesis is tested:

H_0^3: *Relative price discovery is not a function of trading system design*, where trading system design is primarily defined according to whether a market trades electronically or otherwise. Moreover, if one is willing to accept the conjecture that informed traders prefer to trade via (anonymous) electronic trading systems then testing this hypothesis also casts light on the role of asymmetric information in determining relative price discovery.

While previous studies provide important insight into a crucial determinant of price discovery, they are necessarily limited in that only cross-sectional (constituent market) variation in price discovery is permitted. However, given the theoretical and empirical evidence concerning the time-varying nature of asymmetric information (see Admati and Pfleiderer, 1988, and Ting, 2006, respectively), it would appear appropriate to test for (time-series) relations amongst relative price discovery and the degree of information asymmetry. Indeed, Barclay and Hendershott (2003) have already established that such relations do exist. However, their methodology relies on sub-dividing the sample into discrete blocks of data across which the degree of information asymmetry is expected to vary, i.e., the pre-open, daytime, and post-close periods. To overcome the discrete nature of the sample sub-division process, we examine the relations using the continuously changing price discovery measures outlined in the previous section. Formally, we test the null hypothesis

H_0^4: *Relative price discovery does not co-vary (over time) with the degree of information asymmetry*,

where the degree of information asymmetry is measured by various time-varying proxies, including, inter alia, a set of intraday periodicity variables, a set of liquidity variables, and a set of variables that measure the impact of macroeconomic announcements.[10]

It is quite possible that price discovery in *all* markets contemporaneously increases (decreases). However, the standardized nature of the relative price discovery measures described in the previous section will fail to detect such events. This is a serious flaw if one believes that the total amount of informed trading in a market can vary over time. To overcome this issue we make use of the measure of aggregated price discovery given by (A.33) in Appendix A. This measure of price discovery allows an examination of the total amount of price discovery present in a market at a particular time. Thus, it is possible to examine whether this amount varies over time or, more importantly, co-varies with the degree of information asymmetry. Formally, we test

H_0^5: *Aggregate price discovery does not co-vary (over time) with the degree of information asymmetry*,

where aggregate price discovery is measured by (A.33) in Appendix A.

III. Empirical results

Having defined the models and hypotheses considered in this paper, we proceed with an empirical analysis of price discovery in the largest and most fragmented of all the US index markets, viz., the S&P 500 market. This begins with a description of the data used, and is followed by an evaluation of the models used to generate the price discovery measures, and an examination of the hypothesis tests described in the previous section.

A. Data

We make use of various pieces of information concerning all trades in various securities relating to the S&P 500 index, over the period, January 2, 2002 to December 31, 2002. In particular, returns, defined as the log change in transaction prices; and trading volume, defined as the number of trades that occur within an interval, were obtained for S&P 500 SPDR exchange traded fund (ETF) shares (with ticker symbol SPY), which trade on the American Stock Exchange (AME); S&P 500 iShares ETF shares (with ticker symbol IVV), which trade on the AME; S&P 500 regular futures contracts (with ticker symbol SP), which (mainly) trade in the Chicago Mercantile Exchange (CME) trading pit; and S&P 500 E-mini futures contracts (with ticker symbol ES), which trade (exclusively) on the (electronic) CME GLOBEX trading system.[11] In addition, the levels of the S&P 500 index were obtained for each minute of the sample period. All of these data were obtained from *Tick Data, Inc.*

Both of the S&P 500 futures contracts specify cash settlement of the contract at 8:30 a.m. on the 3rd Friday of March, June, September, and December. To obtain a single continuous series for each type of futures contract, we adopt the common practise that futures contracts with the nearest maturity are replaced (through trading) by contracts with the next nearest maturity when the next contract's daily tick count exceeds the current contract tick count.

The above transaction data are then converted to one-minute frequency data. This frequency is deemed to be sufficiently low enough to avoid stale data, and high enough to avoid loss of information. Furthermore, as the trading hours of each security differ, these data are necessarily truncated to produce a fully synchronized intraday dataset. In particular, the (Chicago time) trading hours for the S&P 500 index, S&P 500 SPDR ETF shares, S&P 500 iShares ETF shares, S&P 500 regular futures contracts, and S&P 500 E-mini futures contracts are 8:30 a.m. to 3:05 p.m., 8:30 a.m. to 3:15 p.m., 8:30 a.m. to 3:15 p.m., 8:30 a.m. to 3:15 p.m., and 12:00 (midnight) to 12:00 (midnight), respectively.[12] Therefore, to generate an overlapping dataset, to avoid use of duplicate closing and opening prices, and to avoid use of outlying and/or stale prices at the close of trading, these series are truncated to contain data over the intraday period, 8:31 a.m. to 3:00 p.m.

The resulting (three) spot and (two) futures series are then used to construct various bilateral mispricing series. One possible method of constructing mispricing series involving spot and futures prices is to construct the theoretical futures prices using government-issued interest rates with a maturity closest to that of the futures contract, and expected (or realized) dividend payments for each stock in the index over the life of the futures contract. However, the complex nature of the timing of the dividend income associated with ownership of the ETF shares makes it very difficult to construct a realistic mispricing series. Consequently, a more simplistic approach is adopted. Specifically, as the interest rate and dividend series can only change on a daily basis, we use the daily demeaned difference between the (log) spot and futures prices as the measure of mispricing. Similarly, for mispricing series involving only spot (or only futures) prices, the differential effects of dividends and interest rates are controlled for by using the daily demeaned difference between the series.

In addition to the above data, survey data associated with all US macroeconomic announcements made during 2002 were obtained from *Money Market Services International*. This dataset is based on a survey of 40 leading economists who are asked at the end of each week to forecast various macroeconomic series with realizations due within the next seven days. The difference between the median of these survey forecasts and the actual value of the series at the exact time of the announcement is taken as the unexpected value of the series. This procedure results in data pertaining to 408 macroeconomic announcements, the majority of which occur at 7.30 a.m. and 9.00 a.m. (Chicago time).[13]

B. Summary statistics

A selection of descriptive statistics pertaining to one-minute frequency returns for each constituent market, one-minute frequency bilateral mispricings for the index market against the other four constituent markets, and one-minute frequency trading intensities for the regular and E-mini futures markets are given in Table I.

INSERT TABLE I HERE

The results highlight five main characteristics of the data. First, index returns are less volatile than single security (i.e., ETF shares and futures contracts) returns. This ordering almost certainly reflects the use of imperfectly cross-correlated constituent share transaction prices when constructing index levels (and index returns). Second, index returns are positively autocorrelated, while all single security returns are negatively autocorrelated – a result mostly likely due to the non-synchronous trading and bid-ask bounce microstructure effects, respectively. Third, the results indicate that the data appear to be highly non-normal, with excess kurtosis being the likely cause of this non-normality. When considering the return series, it is possible that this is due to time-variation in the conditional volatility of returns, as indicated by the tests for (no) autoregressive conditional heteroscedasticity (ARCH) effects. Nevertheless, some attention must be paid to the validity of the normality assumption when estimating the models described previously. Fourth, regarding the bilateral mispricing series, all autocorrelation coefficients are less than unity. This observation most likely indicates the presence of arbitrageurs who take offsetting spot and/or futures positions when these series take large absolute values. In turn, this behavior results in stationary mispricing series (as opposed to the non-stationary prices upon which the mispricing series are based). Finally, the E-mini futures market is more actively traded that the regular futures market. This is evinced by the fact that for every trade in the latter market, more than twenty trades occur in the former market despite the fact that the size of an E-mini contract is only one-fifth the size of a regular contract.[14]

C. Model specifics

Two versions of the vector equilibrium correction model given by (1) are estimated in this paper. The first is a restricted unscaled version of this model where no scaling factors are applied to the parameters in (1). By contrast, the second model assumes the following set of scaling factors: an intraday periodic scaling factor set (henceforth denoted $\mathcal{S}_1$); an interday periodic scaling factor set based on the time-to-maturity of the futures contracts (henceforth denoted $\mathcal{S}_2$); a scaling factor set based on a measure of liquidity (henceforth denoted $\mathcal{S}_3$); and two scaling factor sets based on the time period before and after the announcement of key macroeconomic data (henceforth denoted $\mathcal{S}_{4a}$ and $\mathcal{S}_{4b}$, respectively); with unions of these sets defined (and denoted) such that $\mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4$ and $\mathcal{S}_4 := \mathcal{S}_{4a} \cup \mathcal{S}_{4b}$. By allowing all of these scaling factors to affect each of the parameters in (1), one is explicitly adopting a highly flexible transmission mechanism that assumes that time-variation in price discovery is determined by variation in (any of) the parameters in (1). In turn, it is assumed that this parameter variation (and hence price discovery variation) is fully captured by the above scaling factors.

With regards to the periodic scaling factor sets, one possible model is one that allows the parameters in (1) to take different values during each t_s (e.g., intraday) time period within the periodic cycle (e.g., the trading day) of length T_s, where $t_s \in \{1, 2, \ldots, T_s - 1, T_s\}$, and t_s and t are related by the function $f_s(t)$.[15] Under this assumption (and ignoring other scaling factors for now), $\mathbf{h}'_t = [1, D_{2,t}, \ldots, D_{t_s,t}, \ldots, D_{T_s,t}]$, with the dummy variable, $D_{t_s,t}$, taking a value of unity if the current observation is in the t_sth stage of the periodic cycle, and a value of zero otherwise. A problem with this model is that a large number of coefficients are required if there are many time periods within each periodic cycle (i.e., T_s is large). It is possible to overcome this problem by selecting dummy variables that span more than one time period. However, this assumes that all parameters in (1) are constant within the time period covered by the dummy variable and then change abruptly whenever a new time period is entered.

To overcome the above problems, a cubic spline-version of the model is used. This model permits a parsimonious representation of parameter variation characterized by smooth variation over the scaling factor space (e.g., the trading day). Specifically, different cubic spline functions are estimated between selected points (referred to as knots) within the scaling factor space. In particular, letting k_r denote the (equally spaced) rth knot, with $k_r = \{k_r \in \mathbb{Z}^+ : 1 \leq k_r \leq T_s\}$, $r \in \{1, \ldots, R-1\}$, and $k_1 = 0$, the spline-based periodic scaling factors are given by

$$\mathbf{h}'_t := [1, \mathbf{h}'_{1,t}, \ldots, \mathbf{h}'_{r,t}, \ldots, \mathbf{h}'_{R-1,t}], \tag{3a}$$

$$\mathbf{h}'_{r,t} := D_r[(t_s - k_r), (t_s - k_r)^2, (t_s - k_r)^3], \tag{3b}$$

and D_r equals unity if $t_s \geq k_r$, and zero otherwise.[16]

While the relationship between intraday periodicity and price discovery can be easily motivated with respect to the information asymmetry-based arguments in Section II, the motivation for including other scaling factor sets requires further reasoning based on an additional key variable. Specifically, the validity of the arguments proposed in Section II regarding the positive relationship between price discovery and the degree of information asymmetry depend crucially on the assumption that the electronic trading system is relatively liquid. If, however, market makers trading on this system increase their effective spreads in the presence of a large number of informed traders then this may result in a relatively illiquid market – conditions that will adversely affect the price discovery process.[17] To examine this issue we consider interday periodic and liquidity-based scaling factor sets. The former set measure liquidity via the time-to-maturity of the nearest futures contract, while the latter are a function of the one-period lagged value of trade intensity in the E-mini market, with selection of this market motivated by virtue of the fact that it is the only market in the sample based on an electronic trading system (i.e., the GLOBEX system).[18] To avoid the potentially restricted assumption of *linear* dependence between the scaling factor and liquidity, we also perform a spline-based transformation of the data. This is achieved by transforming the data according to (3b), with t_s being replaced by the time-to-maturity of the nearest futures contract and trading intensity.

Finally, given the expectation that informed traders are more likely to be trade around the time of public announcements, we consider the thirty minute period before and after the release of all major US macroeconomic announcements.[19] In this particular case, the time from the start of the pre-announcement window and the time to the end of the post-announcement window are used as the inputs into (3b). By defining time in this way, and applying the spline-based transformation given by (3b), the parameter values are assured to converge to their non-announcement period values as the outer edges of the announcement window are approached.

D. Model estimation

The ordinary least squares (OLS) estimates of the parameters (and measures of model performance) associated with the unscaled and scaled versions of the equilibrium correction model are given in Table II.[20] The results of each are based on first-order lag dependence structure and single knot versions of the four scaling factors described in the previous subsection. Selection of this particular parameter space is based on the superior system Akaike Information Criterion (AIC) value associated with this space, against those obtained when up to a fifth-order lag dependence structure, and up to three knots are considered.[21]

INSERT TABLE II HERE

The first thing to note about the estimated models is that they differ from those given in

(1) in that the constant is unscaled. This is because we have no a priori reason to believe that the scaling factors will be related to the returns in each constituent market. Indeed, when the scaling factors are included in the constant term, the resulting model provides an inferior representation of the data.[22] Also, space limitations prevent presentation of all the estimated parameters of the scaled version of the model. Consequently, we present a summary of these estimated parameters in the form of the mean estimated parameters calculated over the sample period.[23]

The results in Table II indicate that the vast majority of the coefficients are significantly different from zero. Thus, all constituent markets appear to be characterized by high levels of interdependence in terms both the mispricing series and the lagged returns series observed in the 'other' constituent markets. This interdependence appears to be strongest in the iShares ETF shares market, and weakest in the E-mini futures market – a result also evinced by the differences in the measures of model fit in these markets. Given the association between this form of interdependence and information flow, it would appear that returns in the former market are heavily dependent on the information flow generated in other markets, and *not* on new information. By contrast, returns in the E-mini market appear to be the product of trading new information. Whether this trading is fundamental to the efficient price is formally tested in the subsequent analysis.

Regarding the performance of the two competing models, the results indicate that the scaled version of the model provides a superior representation of the data. This can be seen by the marginally superior residual diagnostics, and the superior system $\overline{R}^2$, AIC, and Schwarz Information Criterion (SIC) measures, associated with the scaled version of the model. For example, the unscaled version of the model produces a system $\overline{R}^2$ of 15.02%, while the scaled version has an associated system $\overline{R}^2$ of 17.21%. Therefore, this collective evidence points to a rejection of the (preliminary) hypothesis given by H_0^1.

IV. Testing the hypotheses

In this section a formal examination of the validity of the null hypotheses, H_0^2, H_0^3, H_0^4, and H_0^5, is carried out. The analysis is divided into those relating to model performance (H_0^2), and those relating to the model-based measures of price discovery (H_0^3 to H_0^5).

A. Model performance

To formally examine the relative contributions of the parameters within the scaling factor sets to the variation in the parameters in (1), we carry out a series of Wald tests, each of which is based on the estimated coefficients (and White's, 1982, robust 'sandwich' estimator of the asymptotic variance-covariance matrix) from the estimated unscaled and scaled versions of the model. These tests are designed to examine the significance (or otherwise) of the scaling factor sets and subsets therein. Specifically, for each scaling factor set (and the union of all sets), we test the significance of the following sets of coefficients in each of the constituent market equations: (*i*). all coefficients, (*ii*). the coefficients on lagged returns observed in the same constituent market (henceforth referred to as the *own lagged return coefficients*), (*iii*). the coefficients on lagged returns observed in the other constituent markets (henceforth referred to as the *other lagged return coefficients*), (*iv*). all mispricing coefficients, and (*v*). the coefficients in (*iii*). and (*iv*) above. By applying the tests in this way one is able to examine the specific nature of the driving forces behind the variation in the parameters in (1). So, for instance, to examine the significance of the intraday variation in the own lagged return coefficient, one restricts all parameters associated with this coefficient within the $\mathcal{S}_1$ scaling factor set to equal zero. The resulting Wald statistics for all variants of the scaling factor sets and selected coefficient groups are given in Table III.

Insert Table III here

The first thing to note from the results is that all coefficients in the unscaled version of the model are significant (at the 1% level), except for the own lagged return coefficient. Similarly, all coefficients in the scaled version of the model are also significant (at the 1% level), with the exception of the own lagged return coefficients in the iShares ETF share and E-mini futures markets. The results in the table also permit an evaluation of the marginal contribution of the scaled version of the model over the unscaled version. This is achieved by testing the significance of the restriction that all the parameters in the scaling factor sets equal zero; equivalently, by testing the significance of all the coefficients in the scaled version that do not appear in the unscaled version. In doing this, we find that the parameters in the scaling factor sets are highly significant. Indeed, only the scaled own lagged return coefficients in the iShares ETF share and E-mini futures markets are insignificant.

The results also highlight a number of features underlying the dynamics of the parameters in (1). With regards to the individual scaling factor sets, the results indicate that, overall, the coefficients are significantly related to the space spanned by each of the scaling factor sets. However, analysis of particular subsets of the coefficients within each scaling factor set reveals interesting performance differences. Most notably, the vast majority of the coefficients within the liquidity-based scaling factor set appear to be significant, while variation in the majority of coefficients within the macroeconomic announcement period appear to be insignificant. Thus, with respect to the null hypothesis, H_0^2, the evidence points to a rejection of this hypothesis for all scaling factor sets, with the possible exception of the macroeconomic announcement scaling factor sets. The impact of these observations (if any) upon price discovery is examined in the subsequent analysis.

B. Price discovery and trading system design

The next stage of the analysis involves examining the relative amount of price discovery in each of the constituent markets. To this end, the four main augmented measures of relative price discovery described in Section I (and Appendix A) are calculated, viz., the augmented Harris et al (2002) component share measure ($\mathbf{f}_{1t}$); the augmented Hasbrouck (1995) information share measure ($\mathbf{f}_{2t}$); the augmented de Jong and Schotman (2005) information share measure ($\mathbf{f}_{3t}$); and the augmented Yan and Zivot (2006a) standardized price discovery measure ($-s\mathbf{f}_{4K,t}$) – all of which are based on the scaled version of the equilibrium correction model given by (1). In addition, non-augmented (unscaled) versions of these measures are calculated, with the non-augmented (unscaled) version of the equilibrium correction model used to obtain the relevant parameter values.[24]

Further enhancements to the estimation of price discovery are necessary as the residuals from (1) exhibit significant conditional heteroscedasticity; see the ARCH test statistics in Table II. Specifically, we model the residuals from (1) using the dynamic conditional correlation (DCC) multivariate generalized ARCH (GARCH) model of Engle (2002).[25] The simplistic implementation and relative parsimony of this model is particularly useful in the current context as a very large number of observations are considered in the context of a complex multivariate model.[26]

The estimated measures of relative price discovery (and intra-market differences therein) based on the above assumptions are presented in Table IV. To give an indication of the magnitude of the augmented versions of relative price discovery, the time-series averages of these measures are presented. Moreover, as the ordering of the five equations in (1) affects certain measures of non-augmented and augmented price discovery, all measures are estimated for each of the 120 possible orderings (i.e., permutations) of the five constituent markets, with the mean relative price discovery (over all orderings) presented.[27]

Insert Table IV here

Inference regarding the price discovery measures is obtained via a simple bootstrap procedure. Specifically, for each permutation of the equation orderings, returns are generated according to (1), with parameters given by the original estimated parameters, and errors given by a random sample of the original residuals. This procedure is repeated 1,000 times with the estimated parameters (including time-varying estimates of the conditional variance-covariance matrix) obtained during each repetition used to generate bootstrap estimates of the measures of relative price discovery. These estimates are then averaged over all permutations of the equation orderings and also, in the case of the scaled versions of price discovery, averaged over the sample period to given the time-series mean relative price discovery measures. Finally, these estimates are used to form the confidence intervals that are used to draw inference regarding whether relative price discovery is significantly different from some benchmark value within each constituent market, and whether there are significant differences amongst relative price discovery across the constituent markets.

The results in Table IV indicate that the index and the E-mini futures markets are the two dominant constituent markets, with all unscaled, and most scaled relative price discovery measures significantly greater than those expected if relative price discovery were uniformly distributed across all constituent markets (i.e., greater than $1/M$), and significantly different from relative price discovery in the other constituent markets. Therefore, with respect to the hypothesis H_0^3 in Section II, it appears that trading system design does have an effect on price discovery. Specifically, the superiority of the electronic E-mini market over all other single security (i.e., not the index) constitutional markets points to a rejection of this null hypothesis – a result that is similar to that obtained by previous investigators (Hasbrouck, 2003, Ates and Wang, 2004, and Kurov and Lasser, 2004).

C. Price discovery and information asymmetry

The nature of the augmented (scaled) price discovery measures considered in this paper make it possible to analyze explicitly the relationship between price discovery and the degree of information asymmetry. This is achieved by examining these measures over the space spanned by the scaling factor sets, and over pre-defined pairs of events based on this space. The pairs of events considered in the paper are designed to capture major differences in the degree of information asymmetry that exists in a particular constituent market. Specifically, we consider the intraday opening and closing of each market, high versus low relative liquidity, and the period before and after a macroeconomic announcement.

The plots in Figure 1 show the conditional mean of the augmented version of Hasbrouck's (1995) measure of relative price discovery over the trading day, over the maturity cycle of the futures contracts, as a function of E-mini futures market liquidity, and thirty minutes before and after a pre-10.00 a.m. macroeconomic announcement.[28] These plots are obtained by taking the conditional mean of the measures at each point within the space spanned by the scaling factor sets. For instance, the intraday plot in Figure 1 gives the conditional mean of each of the relative price discovery measures during each minute of the trading day.

INSERT FIGURE 1 HERE

It is evident from these plots that there exists a considerable amount of variation within the spaces spanned by the scaling factor sets. Most notably in this regard, price discovery appears to vary positively (and non-linearly) with the degree of information asymmetry. For instance, during periods of low liquidity in the E-mini futures market, most prices are discovered in the market for constituent shares (as revealed by the index). By contrast, when the E-mini futures market is highly liquid, this market dominates all other markets in terms of price discovery.[29] A similar conclusion can be drawn when considering the period around macroeconomic announcements. During the period prior to the announcement, the market for constituent shares dominates all other markets. However, at the precise point in time when the announcement is made, the E-mini futures market takes over as the dominant market. Thus, if one assumes that the vast majority of informed trades take place immediately after (and not before) the release of macroeconomic information, then this result supports the notion that price discovery in markets based on electronic trading systems is positively related to the degree of information asymmetry.[30]

To formally test the above conjecture regarding price discovery and the degree of information asymmetry, a series of tests are conducted that are designed to examine the differences between the conditional means of the price discovery measures across events characterized by extreme trading behavior. In all cases, inference is based on the bootstrap samples obtained previously. The results of these procedures are presented in Table V (market open-close variation), Table VI (low-high liquidity variation), and Table VII (before-after announcement variation), where the market open period is defined as the first fifteen minutes of trading after the start of trading in the major S&P 500 markets, the market close period is defined as the last fifteen minutes of trading prior to the end of trading in the major S&P 500 markets, low (high) liquidity is defined as all liquidity amounts below (above) the median level of liquidity in the E-mini futures market, and the before (after) announcement period is defined as the five minute period prior to (after) a macroeconomic announcement.

Insert Table V here

The results in Table V indicate that for the majority of the price discovery measures, the E-mini futures market is the most informative of all the constituent markets during the market open period, while the second most informative is the market for constituent shares. By contrast, this ranking is reversed at the close of trading, with the latter market being the dominant market in terms of price discovery. Moreover, during the market open period, most of the E-mini futures price discovery measures are significantly greater than the measures observed in the market for constituent shares, with the exact reverse holding during the market close period. Combining this finding with the empirical evidence of Barclay and Hendershott (2003) and Ting (2006) who both find that the degree of information asymmetry decreases over the trading day, one can interprete this as evidence in favor of the argument that price discovery in operationally efficient markets (e.g., the E-mini futures market) is positively related to the degree of information asymmetry.

INSERT TABLES VI & VII HERE

Similar evidence is found in Table VI and Table VII. Specifically, the results in Table VI indicate that price discovery in the E-mini futures market is significantly higher during periods of high liquidity than during periods of low liquidity. Moreover, this market becomes the dominant price discovery market (according to two of the four price discovery measures considered) during periods of high liquidity. These results are consistent with the argument that markets based on electronic trading systems (i.e., the E-mini futures market) are the preferred choice of informed traders providing that liquidity is sufficiently high (equivalently, trading costs are sufficiently low). Similarly, the results in Table VII show that during periods of high information asymmetry (i.e., immediately after the announcement of price sensitive news), the E-mini futures market becomes the dominant price discovery market (according to three of the four price discovery measures considered). Therefore, the collective evidence in Tables V, VI, and VII, points to a rejection of the null hypothesis given by H_0^4 that relative price discovery does not co-vary with the degree of information asymmetry. Rather, relative price discovery appears to significantly co-vary with variables that measure the degree of information asymmetry.

D. Price discovery, information asymmetry, and liquidity

It could be argued that the results in Tables V, VI, and VII are flawed in that the results may be driven by contemporaneous variable correlation across the event pairs. For example, the results regarding the difference in relative price discovery over the market open and close periods may be driven by variation in liquidity over these periods. To remedy this potential shortcoming, we increase the dimension of the conditionality by further subdividing the event pairs. Specifically, each time period within the market open and close periods is subdivided according to whether the time period was characterized by low or high levels of liquidity. The conditional means of the price discovery measures are then calculated in each of these four new time periods. The resulting conditional mean measures of relative price discovery are given in Table VIII.

Insert Tables VIII & IX here

The results in Table VIII indicate that during periods when the degree of information asymmetry is likely to be at its highest (i.e., during the market open period), the E-mini futures market is the dominant constituent price discovery market *only* when the market is highly liquid (i.e., trading costs are low). Thus, this result provides further support for the argument that markets based on electronic trading systems are the preferred choice of informed traders providing that liquidity is sufficiently high. Similar evidence is provided in Table IX, where the before and after macroeconomic announcement periods are subdivided into periods of low and high liquidity. Specifically, only immediately after an announcement *and* when the market is relatively liquid do informed traders appear to prefer to trade on electronic trading systems.

E. Aggregate price discovery

Thus far the analysis has focused on relative price discovery in markets than compete for order flow. In doing this we have provided empirical evidence to show that particular markets are favored by informed traders under certain conditions. Most notably, electronic markets are preferred during periods of extreme information asymmetry and high liquidity. However, this analysis has ignored the possibility that price discovery may uniformly increase (or decrease) across all constituent markets. To remedy this potential shortcoming we consider an aggregate measure of price discovery; specifically, the sum of the log of the augmented Yan and Zivot (2006a) price discovery measure over all constituent markets.[31]

Insert Figure 2 here

The conditional means of the above measure are calculated using the space spanned by selected scaling factor sets, and the three event pairs considered previously. The results associated with aggregate price discovery around pre-10.00a.m. macroeconomic announcements are represented diagrammatically in Figure 2. It is clear from this particular plot that aggregate price discovery falls dramatically before the announcement and increases immediately after the release of the information. Likewise, when the conditional means are calculated for each event pair considered previously (see Tables V, VI, VII, VIII, and IX, for these aggregated price discovery measures), the results lead to a similar conclusion. Indeed, the results in Tables VIII and IX confirm the previous result that informed trading appears to be concentrated in the market open and post-announcement periods *only* when markets are relatively liquid. However, the results are not significant even at the 10% level. Thus, though the evidence is suggestive, it does not enable a clear rejection of the null hypothesis, H_0^5, concerning the relationship between aggregate price discovery and the degree of information asymmetry.

V. Concluding remarks

The multivariate model of fragmented market returns introduced in this paper differentiates itself from competing models in its treatment of time in the arbitrage process. Existing models assume that the propensity-to-arbitrage (as given by the coefficients on past mispricings) is constant. However, when these coefficients are scaled, they are found to covary with, inter alia, measures of the degree of information asymmetry and liquidity. Moreover, this model provides a superior representation of the S&P 500 data considered in this paper.

Having established that the model introduced in this paper most likely represents the dynamics of returns in fragmented markets, the paper proceeds by using this model to obtain measures of relative and aggregate price discovery in the S&P 500 market. As these measures are (significantly) functionally dependent upon the coefficients of the underlying model via a set of scaling factors, then the measures themselves are also dependent upon these factors. Indeed, this dependence is shown to cast light on important aspects of trading on competing systems. Specifically, most price discovery appears to occur in the market for the stocks making up the index and in the (electronically traded) E-mini futures market. However, the operational efficiencies inherent in the E-mini futures market result in it becoming the dominant price discovery market under two conditions: extreme information asymmetry and high liquidity – a finding that supports theoretical arguments proposed in related literature, particularly those recently advocated by Ates and Wang (2004, 2005).

Appendix A: Derivations of augmented price discovery measures

This appendix contains derivations of the augmented price discovery measures used in this paper. These derivations are divided into three parts. The first part establishes the equivalence between an augmented version of Hasbrouck's (1993) unobservable efficient price model, and the equilibrium correction model given by (1). Then, explicit expressions for the coefficients in these models are derived and subsequently used to define the price discovery measures given in the final part of the appendix.

A.1 Derivation of the efficient price model

The econometric model used in this paper can be expressed in terms of the following multivariate (time-varying parameter) version of Hasbrouck's (1993) univariate (time-invariant) unobservable components efficient price model,

$$\mathbf{p}_t = \mathbf{D}_t + m_t + \boldsymbol{\upsilon}_t, \tag{A.1a}$$

$$m_t = m_{t-1} + \eta_t^P, \tag{A.1b}$$

$$\boldsymbol{\upsilon}_t = \boldsymbol{\gamma}_t \eta_t^P + \boldsymbol{\zeta}_t, \tag{A.1c}$$

where $\mathbf{D}_t$ represents all deterministic terms; m_t is the unobservable common fundamental full-information asset price, which Hasbrouck (1995) refers to as the *efficient price*; $\boldsymbol{\upsilon}_t$ is an $(M \times 1)$ vector of stationary pricing error terms; η_t^P is the permanent innovation; $\boldsymbol{\gamma}_t$ is an $(M \times 1)$ parameter vector; $\boldsymbol{\zeta}_t | \mathcal{I}_{t-1} \sim \mathcal{D}(\mathbf{0}_{M\times 1}, \boldsymbol{\Omega}_t)$; and observed price innovations (obtained by substituting (A.1b) and (A.1c) into (A.1a)) are given by $\boldsymbol{\upsilon}_t^* = (\boldsymbol{\gamma}_t + \mathbf{1}_{M\times 1})\eta_t^P + \boldsymbol{\zeta}_t$, with $\boldsymbol{\upsilon}_t^* | \mathcal{I}_{t-1} \sim \mathcal{D}(\mathbf{0}_{M\times 1}, \boldsymbol{\Upsilon}_t)$. This model states that, excluding deterministic terms, the observed price in each of the M constituent markets equals the efficient price plus a market-specific stationary pricing error, with the efficient price following a random walk subject to a permanent innovation that, in turn, has a time-varying relationship with the stationary pricing error.

To see the equivalence between the above model and (2), firstly consider the following Wold representation of $\Delta \mathbf{p}_t$:

$$\Delta \mathbf{p}_t = \boldsymbol{\theta}_{0t} + \boldsymbol{\theta}_t(L)\boldsymbol{\epsilon}_t = \boldsymbol{\theta}_{0t} + \boldsymbol{\theta}_{1t}\boldsymbol{\epsilon}_t + \boldsymbol{\theta}_{2t}\boldsymbol{\epsilon}_{t-1} + \cdots \tag{A.2}$$

where the matrix polynomial $\boldsymbol{\theta}_t(L)$ has full rank, with $\boldsymbol{\theta}_{1t} = \mathbf{I}_M$, and $(M \times M)$ coefficient matrix elements $\{\boldsymbol{\theta}_{kt}\}_{k=1}^{\infty}$ that are 1-summable. The importance of this representation in the current context can be demonstrated by applying the Beveridge-Nelson decomposition (Beveridge and Nelson, 1981) to $\boldsymbol{\theta}_t(L)$ and iterating through time,

$$\mathbf{p}_t = \mathbf{D}(\boldsymbol{\theta}_{0t}) + \boldsymbol{\theta}_t(1)\sum_{j=1}^{t}\boldsymbol{\epsilon}_j + \boldsymbol{\nu}_t, \tag{A.3}$$

where $\mathbf{D}(\boldsymbol{\theta}_{0t})$ represents all deterministic terms with functional dependence, inter alia, upon $\boldsymbol{\theta}_{0t}$; $\boldsymbol{\nu}_t$ is an $(M \times 1)$ vector of stationary pricing error terms given by $\big((\boldsymbol{\theta}_t(L) - \boldsymbol{\theta}_t(1))(1-L)^{-1}\big)\boldsymbol{\epsilon}_t$; and $\boldsymbol{\theta}_t(1)$ measures the cumulative impact of all past and present $\boldsymbol{\epsilon}_t$ on the current price with rows that are identical to each other. Given this definition of $\boldsymbol{\theta}_t(1)$, and without loss of generality, the subsequent analysis is based on the 1st row of $\boldsymbol{\theta}_t(1)$ transposed to yield an $(M \times 1)$ 'long-run impact' column vector defined as $\boldsymbol{\theta}_t := [\boldsymbol{\theta}_t(1)]'_{(1,1:M)}$, where superscripted square parentheses indicates the selection of specific rows and columns of the matrix therein.

Furthermore, it is also possible to express (2) as a function of a common stochastic trend via the decomposition proposed by Stock and Watson (1988); specifically,

$$\mathbf{p}_t = \mathbf{D}(\boldsymbol{\theta}_{0t}) + m_t + \boldsymbol{\nu}_t, \tag{A.4a}$$

$$m_t = m_{t-1} + \eta_t^P, \tag{A.4b}$$

where the permanent innovation, η_t^P, is given by $\boldsymbol{\theta}_t'\boldsymbol{\epsilon}_t$; and $\mathsf{E}(\boldsymbol{\nu}_t\eta_t^P|\mathcal{I}_{t-1}) = 0$, hence $\boldsymbol{\gamma}_t = 0\ \forall\, t \in \{1,\ldots,T\}$. Thus, we obtain a representation analogous to that given in (A.1a), (A.1b), and (A.1c), whereby prices in each of the M constituent markets is composed of, inter alia, a common efficient price, and a set of N orthogonal transitory pricing errors. Moreover, this efficient price evolves as a random walk driven by the arrival of new information concerning the asset's future value (i.e., the permanent innovation, η_t^P).

An alternative way of deriving the representation given by (A.1a), (A.1b) and (A.1c) is to directly imposed the Beveridge-Nelson restriction on the reduced form version of this representation. In doing this we obtain $\eta_t^P = \boldsymbol{\theta}_t'\boldsymbol{\epsilon}_t$, $\mathsf{Var}(\eta_t^P|\mathcal{I}_{t-1}) = \boldsymbol{\theta}_t'\boldsymbol{\Sigma}_t\boldsymbol{\theta}_t$; with parameter expressions,

$$\boldsymbol{\gamma}_t = \frac{\boldsymbol{\Sigma}_t\boldsymbol{\theta}_t}{\boldsymbol{\theta}_t'\boldsymbol{\Sigma}_t\boldsymbol{\theta}_t} - \mathbf{1}_{M\times 1}, \tag{A.5a}$$

$$\boldsymbol{\Omega}_t = \boldsymbol{\Sigma}_t - \frac{\boldsymbol{\Sigma}_t\boldsymbol{\theta}_t\boldsymbol{\theta}_t'\boldsymbol{\Sigma}_t}{\boldsymbol{\theta}_t'\boldsymbol{\Sigma}_t\boldsymbol{\theta}_t}, \tag{A.5b}$$

where $\boldsymbol{\gamma}_t$ takes the maximal permissable value for each t; see Appendix A of de Jong and Schotman (2005) for proof of this result in the case of the time-invariant parameter (i.e., unscaled) unobserved components model. As will shown in subsection A.3, this particular representation forms the basis of an alternative measure of price discovery considered in this paper.

Finally, a third approach is possible in order to derive the representation given by (A.1a), (A.1b), and (A.1c).

Specifically, Yan and Zivot (2006a) propose use of an (unscaled) structural moving average model with innovations separated into those that are have a permanent impact on prices, and those that have a transitory impact on prices. As with previous derivations, this model is also augmented in the current application to allow for time variation in the coefficients. This is achieved via use of the following linearly scaled structural vector moving average model:

$$\Delta \mathbf{p}_t = \boldsymbol{\theta}^*_{0t} + \boldsymbol{\theta}^*_t(L)\boldsymbol{\eta}_t = \boldsymbol{\theta}^*_{0t} + \boldsymbol{\theta}^*_{1t}\boldsymbol{\eta}_t + \boldsymbol{\theta}^*_{2t}\boldsymbol{\eta}_{t-1} + \cdots \tag{A.6}$$

where the matrix polynomial $\boldsymbol{\theta}^*_t(L)$ has full rank, with $\{\boldsymbol{\theta}^*_{1t}\} \neq \mathbf{I}_M$, and $(M \times M)$ coefficient matrix elements $\{\boldsymbol{\theta}^*_{kt}\}_{k=1}^{\infty}$ that are 1-summable; and $\boldsymbol{\eta}_t := (\eta^P_t, \boldsymbol{\eta}^T_t)'$, with $\boldsymbol{\eta}_t|\mathcal{I}_{t-1} \sim \mathcal{D}(\mathbf{0}, \boldsymbol{\Sigma}^*_t)$, such that $\boldsymbol{\Sigma}^*_t$ is a variance-covariance matrix with the variances of $\boldsymbol{\eta_t}$ on the diagonal, and zeros elsewhere. It is a trivial matter to show that η^P_t has a one-to-one long-run effect on the price of each asset, while $\boldsymbol{\eta}^T_t$ has no long-run effect on the price of each asset – both of which imply that $\boldsymbol{\theta}^*_t(1) = (\mathbf{1}_{M\times 1} \vdots \mathbf{0}_{M\times 1})$.

By applying the Beveridge-Nelson decomposition to (A.6), one can demonstrate the similarities (and differences) between the linearly scaled structural moving average and equilibrium corrections models; specifically, (A.6) becomes

$$\mathbf{p}_t = \mathbf{D}(\boldsymbol{\theta}^*_{0t}) + \boldsymbol{\theta}^*_t(1)\sum_{j=1}^{t}\boldsymbol{\eta}_t + \boldsymbol{\nu}_t = \mathbf{D}(\boldsymbol{\theta}^*_{0t}) + m_t + \boldsymbol{\nu}^*_t, \tag{A.7a}$$

$$m_t = m_{t-1} + \eta^P_t, \tag{A.7b}$$

where the $(M \times 1)$ vector of stationary pricing error terms $\boldsymbol{\nu}^*_t$ is given by $\big((\boldsymbol{\theta}^*_t(L) - \boldsymbol{\theta}^*_t(1))(1-L)^{-1}\big)\boldsymbol{\eta}_t$. As is clear from (A.7a) and (A.7b), prices are determined by a common efficient price, with the same permanent innovation, η^P_t, as in the model given by (A.4a) and (A.4b). However, the two models differ in that transitory pricing errors are now a function of $\boldsymbol{\eta}_t$, implying that $\boldsymbol{\gamma}_t$ may differ from zero and thus pricing errors may be a function of the arrival of new information.

A.2 Deriving coefficient expressions

Having established the relation between the unobservable efficient price model given by (A.1a), (A.1b), and (A.1c), and the equilibrium model given by (1), this subsection proceeds with derivations of the expressions for the individual components in the subsequently defined price discovery measures. Most notably, the coefficients in the directly unobservable Wold representation of the equilibrium correction and structural moving average models, given by (A.2) and (A.6), respectively. In each case, these coefficients are necessarily functions of the directly observable coefficients in the linearly scaled equilibrium correction model given by (2).

In the spirit of Yan and Zivot's (2006a) derivation of their time-invariant price discovery measure, the first step toward deriving explicit expressions for the coefficients in (A.2), and also $\boldsymbol{\theta}_t(1)$, is to rewrite (2) as a first-order process;

specifically,

$$\begin{bmatrix} \Delta\mathbf{p}_t \\ \Delta\mathbf{p}_{t-1} \\ \Delta\mathbf{p}_{t-2} \\ \vdots \\ \Delta\mathbf{p}_{t-(K-1)} \\ \mathbf{z}_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}_{0t} \\ \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1} \\ \vdots \\ \mathbf{0}_{M\times 1} \\ \boldsymbol{\beta}'\boldsymbol{\Psi}_{0t} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Psi}_{1t} & \boldsymbol{\Psi}_{2t} & \cdots & \boldsymbol{\Psi}_{K-1t} & \boldsymbol{\Psi}_{Kt} & \boldsymbol{\alpha}_t \\ \mathbf{I}_M & \mathbf{0}_{M\times M} & \cdots & \mathbf{0}_{M\times M} & \mathbf{0}_{M\times M} & \mathbf{0}_{M\times N} \\ \mathbf{0}_{M\times M} & \mathbf{I}_M & \cdots & \mathbf{0}_{M\times M} & \mathbf{0}_{M\times M} & \mathbf{0}_{M\times N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{0}_{M\times M} & \mathbf{0}_{M\times M} & \cdots & \mathbf{I}_M & \mathbf{0}_{M\times M} & \mathbf{0}_{M\times N} \\ \boldsymbol{\beta}'\boldsymbol{\Psi}_{1t} & \boldsymbol{\beta}'\boldsymbol{\Psi}_{2t} & \cdots & \boldsymbol{\beta}'\boldsymbol{\Psi}_{K-1t} & \boldsymbol{\beta}'\boldsymbol{\Psi}_{Kt} & \boldsymbol{\beta}'\boldsymbol{\alpha}_t + \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \Delta\mathbf{p}_{t-1} \\ \Delta\mathbf{p}_{t-2} \\ \Delta\mathbf{p}_{t-3} \\ \vdots \\ \Delta\mathbf{p}_{t-K} \\ \mathbf{z}_{t-1} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_t \\ \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1} \\ \vdots \\ \mathbf{0}_{M\times 1} \\ \boldsymbol{\beta}'\boldsymbol{\epsilon}_t \end{bmatrix}, \tag{A.8}$$

or in compact form:

$$\Delta\widetilde{\mathbf{p}}_t = \widetilde{\boldsymbol{\Psi}}_{0t} + \widetilde{\boldsymbol{\Psi}}_t \Delta\widetilde{\mathbf{p}}_{t-1} + \widetilde{\boldsymbol{\epsilon}}_t, \tag{A.9}$$

where covariance-stationarity at time t is assured if all the eigenvalues of the Q-dimension square matrix $\widetilde{\boldsymbol{\Psi}}_t$ lie inside the unit circle; equivalently, in terms of spectral radius, we require that $\varrho(\widetilde{\boldsymbol{\Psi}}_t) < 1$. By repeated substitution in (A.9) we obtain,

$$\begin{aligned} \Delta\widetilde{\mathbf{p}}_t = \widetilde{\boldsymbol{\Psi}}_{0t} + \sum_{h=1}^{k-1}\left(\prod_{h^*=1}^{h}\widetilde{\boldsymbol{\Psi}}_{t-h^*+1}\right)\widetilde{\boldsymbol{\Psi}}_{0t-h} + \left(\prod_{h=1}^{k}\widetilde{\boldsymbol{\Psi}}_{t-h+1}\right)\Delta\widetilde{\mathbf{p}}_{t-k} \\ + \left(\prod_{h=1}^{k-1}\widetilde{\boldsymbol{\Psi}}_{t-h+1}\right)\widetilde{\boldsymbol{\epsilon}}_{t-k+1} + \left(\prod_{h=1}^{H-1}\widetilde{\boldsymbol{\Psi}}_{t-h+1}\right)\widetilde{\boldsymbol{\epsilon}}_{t-k+2} + \ldots + \widetilde{\boldsymbol{\Psi}}_t\widetilde{\boldsymbol{\epsilon}}_{t-1} + \widetilde{\boldsymbol{\epsilon}}_t, \end{aligned} \tag{A.10}$$

with coefficient matrix elements of $\boldsymbol{\theta}_t(L)$ given by $\boldsymbol{\theta}_{1t} = \mathbf{I}_M$,

$$\{\boldsymbol{\theta}_{kt}\}_{k=2}^{\infty} := \frac{\partial\Delta\mathbf{p}_t}{\partial\boldsymbol{\epsilon}_{t-k+1}} = \left[\prod_{h=1}^{k-1}\widetilde{\boldsymbol{\Psi}}_{t-h+1}\right]_{(1:M,1:M)} + \left[\prod_{h=1}^{k-1}\widetilde{\boldsymbol{\Psi}}_{t-h+1}\right]_{(1:M,Q-N+1:Q)}\boldsymbol{\beta}', \tag{A.11}$$

and long-run impact matrix $\boldsymbol{\theta}_t(1)$ given by

$$\boldsymbol{\theta}_t(1) := \sum_{k=1}^{\infty}\boldsymbol{\theta}_{kt} = \mathbf{I}_M + \left[\sum_{k=2}^{\infty}\prod_{h=1}^{k-1}\widetilde{\boldsymbol{\Psi}}_{t-h+1}\right]_{(1:M,1:M)} + \left[\sum_{k=2}^{\infty}\prod_{h=1}^{k-1}\widetilde{\boldsymbol{\Psi}}_{t-h+1}\right]_{(1:M,Q-N+1:Q)}\boldsymbol{\beta}', \tag{A.12}$$

where, without loss of generality, the long-run impact vector can be redefined as the $(M\times 1)$ vector $\boldsymbol{\theta}_t := [\boldsymbol{\theta}_t(1)]'_{(1,1:M)}$.

To obtain an expression for the coefficients in (A.6), we first apply the Gonzalo and Ng (2001) decomposition to calculate the permanent and transitory innovations from the errors in the linearly scaled equilibrium correction model,

$$\boldsymbol{\xi}_t := \mathbf{B}_t\boldsymbol{\epsilon}_t, \tag{A.13}$$

where $\boldsymbol{\xi}_t := (\xi_t^P, \boldsymbol{\xi}_t^T)'$; and $\mathbf{B}_t := (\boldsymbol{\theta}_t \vdots \boldsymbol{\beta})'$, with the elements of $\mathbf{B}_t$, specifically $\boldsymbol{\theta}_t$, such that the permanent innovation, ξ_t^P, has a unit long-run impact on prices. To orthogonalize these errors, we again follow Gonzalo and Ng (2001) and define the errors in the structural moving average models as

$$\boldsymbol{\eta}_t := \mathbf{C}_t^{-1}\boldsymbol{\xi}_t, \tag{A.14}$$

where $\mathbf{C}_t$ is a unique lower triangular matrix with ones along the principal diagonal such that $\mathbf{C}_t\mathbf{D}_t\mathbf{C}_t' = \boldsymbol{\Sigma}_t$, and $\mathbf{D}_t$ is a unique diagonal matrix with positive entries along the principal diagonal. Combining the above results with the Wold representation of the linear scaled equilibrium correction model we see that

$$\begin{aligned}\Delta\mathbf{p}_t &= \boldsymbol{\theta}_{0t} + \boldsymbol{\theta}_t(L)\boldsymbol{\epsilon}_t \\ &= \boldsymbol{\theta}_{0t} + \boldsymbol{\theta}_t(L)\mathbf{B}_t^{-1}\mathbf{C}_t\mathbf{C}_t^{-1}\mathbf{B}_t\boldsymbol{\epsilon}_t \\ &= \boldsymbol{\theta}_{0t}^* + \boldsymbol{\theta}_t^*(L)\boldsymbol{\eta}_t, \end{aligned} \tag{A.15}$$

where $\boldsymbol{\theta}_{0t} \equiv \boldsymbol{\theta}_{0t}^*$; $\boldsymbol{\eta}_t := \mathbf{C}_t^{-1}\mathbf{B}_t\boldsymbol{\epsilon}_t$; and thus

$$\boldsymbol{\theta}_t^*(L) = \boldsymbol{\theta}_t(L)\mathbf{B}_t^{-1}\mathbf{C}_t, \tag{A.16}$$

with coefficient matrix elements in the matrix polynomial $\boldsymbol{\theta}_t^*(L)$ given by

$$\{\boldsymbol{\theta}_{kt}^*\}_{k=1}^{\infty} = \{\boldsymbol{\theta}_{kt}\mathbf{B}_{t-k+1}^{-1}\mathbf{C}_{t-k+1}\}_{k=1}^{\infty}. \tag{A.17}$$

A.3 Defining price discovery

The first two price discovery measures considered in this paper are based explicitly on the linearly scaled long-run impact matrix associated with the equilibrium correction model given in (2). The relevance of this matrix in the current context can by seen in (A.4a) and (A.4b). These equations show that the permanent innovation, η_t^P, to the efficient price impounds new information regarding the fundamental value of the asset. Moreover, as this innovation is given by $\boldsymbol{\theta}_t'\boldsymbol{\epsilon}_t$ then the importance of the long-run impact matrix becomes evident. This relevance motivates the *component share* measure of relative price discovery proposed by Harris et al (2002). However, their measure is augmented in the current paper to take into account the time-varying nature of the long-run impact matrix; specifically, we consider the following linearly scaled version of their measure based on the long-run *future* impact matrix:

$$\mathbf{f}_{1t} := \frac{\boldsymbol{\theta}_t^f}{\mathbf{1}_{1\times M}\boldsymbol{\theta}_t^f}, \tag{A.18}$$

where $\mathbf{f}_{1t}$ is an $(M \times 1)$ vector of relative price discovery measures that has elements that contain each constituent market's share in the composition of the permanent innovation to the efficient price; $\mathbf{1}_{1\times M}\mathbf{f}_{1t} = 1$; and $\boldsymbol{\theta}_t^f$ is defined as

$$\boldsymbol{\theta}_t^{f\prime} := \left[\sum_{k=1}^{\infty} \frac{\partial \Delta \boldsymbol{p}_{t+k-1}}{\partial \boldsymbol{\epsilon}_t}\right]_{(1,1:M)} = \left[\sum_{k=1}^{\infty} \boldsymbol{\theta}_{kt+k-1}\right]_{(1,1:M)}. \tag{A.19}$$

The first stage toward obtaining an implementable definition of $\boldsymbol{\theta}_t^f$ involves making use of the expression for the long-run impact matrix given by (A.12). Using this expression leads to the following expression for the second summation term in (A.19):

$$\sum_{k=1}^{\infty} \boldsymbol{\theta}_{kt+k-1} = \mathbf{I}_M + \left[\sum_{k=2}^{\infty}\prod_{h=1}^{k-1} \widetilde{\boldsymbol{\Psi}}_{t+k-h}\right]_{(1:M,1:M)} + \left[\sum_{k=2}^{\infty}\prod_{h=1}^{k-1} \widetilde{\boldsymbol{\Psi}}_{t+k-h}\right]_{(1:M,Q-N+1:Q)} \boldsymbol{\beta}', \tag{A.20}$$

where $\widetilde{\boldsymbol{\Psi}}_t$ is a Q-dimension compact matrix of time-dependent coefficients given by (A.8) and (A.9). To obtain an analytical solution to the above expression, we assume that for each t,

$$\widetilde{\boldsymbol{\Psi}}_{t+k-h} = \widetilde{\boldsymbol{\Psi}}_t, \qquad \forall\, k-h \in \mathbb{N}. \tag{A.21}$$

Though this assumption is somewhat restrictive, it does yield an implementable estimate of the long-run future impact matrix at time t based on the ceteris paribus assumption that the currently observed coefficient values will hold for all future time periods. Consequently, it can be interpreted as an instantaneous measure of the long-run future impact matrix at time t. Substituting (A.21) into (A.20) eventually leads to the following analytical definition for the long-run future impact matrix at time t:

$$\boldsymbol{\theta}_t^{f\prime} := \left[\left[\mathbf{G}(\widetilde{\boldsymbol{\Psi}}_t)\right]_{(1:M,1:M)} + \left[\mathbf{G}(\widetilde{\boldsymbol{\Psi}}_t)\right]_{(1:M,Q-N+1:Q)} \boldsymbol{\beta}'\right]_{(1,1:M)}, \tag{A.22}$$

where $\mathbf{G}(\widetilde{\boldsymbol{\Psi}}_t) = (\mathbf{I}_Q - \widetilde{\boldsymbol{\Psi}}_t)^{-1}$. Thus, we have a useable measure of price discovery, where a high (low) component share value at time t implies that an unexpected price change in a particular constituent market at time t has, on average, a high (low) impact upon the unexpected change in the efficient price at time t; consequently, prices are generally (not) discovered in this market at time t.

As with the component share measure, the *information share* measure of relative price discovery originally proposed by Hasbrouck (1995) is based on the long-run impact matrix associated with the linearly scaled equilibrium correction model given by (2). However, it differs in that it measures the contribution of an unexpected price change in a particular market to the total variance of the permanent innovation; thus, the linearly scaled version of Hasbrouck's (1995) measure is given by

$$\mathbf{f}_{2t} := \frac{\mathbf{A}_t' \boldsymbol{\theta}_t^f}{\boldsymbol{\theta}_t^{f\prime} \boldsymbol{\Sigma}_t \boldsymbol{\theta}_t^f}, \tag{A.23}$$

where $\mathbf{f}_{2t}$ is an $(M \times 1)$ vector of relative price discovery measures; $\mathbf{A}_t$ is the lower triangular matrix such that $\mathbf{A}_t \mathbf{A}_t' = \boldsymbol{\Sigma}_t$; $\boldsymbol{\theta}_t^f$ is functionally dependent on $\widetilde{\boldsymbol{\Psi}}_t$ as described by the expression in (A.22); and $\mathbf{1}_{1\times M}\mathbf{f}_{2t} = 1$. In this instance, a high (low) information share value at time t implies that an unexpected price change in a particular constituent market at time t is likely to have a high (low) impact upon the variance of the unexpected change in the efficient price at time t; consequently, prices are generally (not) discovered in this market at time t.[32]

As with Hasbrouck's (1995) measure, de Jong and Schotman (2005) define the information share of a market in terms of the proportion of the efficient price variance explained by observed price innovations. However, rather than using the somewhat arbitrary Cholesky decomposition of $\boldsymbol{\Sigma}_t$ to attribute covariance terms to each market as in (A.23), they define price discovery directly within the unobserved components model given by (A.1a), (A.1b), and (A.1c). To obtain a scaled version of their measure, one takes conditional variances of both sides of

$$\eta_t^p = \boldsymbol{\gamma}_t^{*\prime} \boldsymbol{\upsilon}_t^* + \varsigma_t, \tag{A.24}$$

to give

$$\mathsf{Var}(\eta_t^p | \mathcal{I}_{t-1}) = \mathsf{Var}(\boldsymbol{\gamma}_t^{*\prime} \boldsymbol{\upsilon}_t^* | \mathcal{I}_{t-1}) + \mathsf{Var}(\varsigma_t | \mathcal{I}_{t-1}), \tag{A.25}$$

where $\varsigma_t | \mathcal{I}_{t-1} \sim \mathcal{D}(0, \sigma_{\varsigma_t}^2)$ and $\mathsf{E}(\boldsymbol{\upsilon}_t^* \varsigma_t | \mathcal{I}_{t-1}) = 0,\ \forall\, t \in \{1, \ldots, T\}$. It follows that the vector of scaled de Jong-Schotman information shares is given by

$$\mathbf{f}(\boldsymbol{\gamma}_t^*, \boldsymbol{\Upsilon}_t, \sigma_{\varsigma_t}^2) := \frac{\mathsf{Var}(\boldsymbol{\gamma}_t^{*\prime} \boldsymbol{\upsilon}_t^* | \mathcal{I}_{t-1})}{\mathsf{Var}(\eta_t^P | \mathcal{I}_{t-1})} = \frac{\boldsymbol{\gamma}_t^{*\prime} \odot \boldsymbol{\Upsilon}_t \boldsymbol{\gamma}_t^*}{\boldsymbol{\gamma}_t^{*\prime} \boldsymbol{\Upsilon}_t \boldsymbol{\gamma}_t^* + \sigma_{\varsigma_t}^2}, \tag{A.26}$$

where $\mathsf{Var}(\eta_t^P | \mathcal{I}_{t-1}) \equiv \sigma_{\varsigma_t}^2$. Under the Beveridge-Nelson implied parameter values associated with the unobserved components model given in (A.5a) and (A.5b), we find that $\mathsf{Var}(\eta_t^p | \mathcal{I}_{t-1}) = \mathsf{Var}(\boldsymbol{\gamma}_t^{*\prime} \boldsymbol{\upsilon}_t^* | \mathcal{I}_{t-1}) = \boldsymbol{\theta}_t' \boldsymbol{\Sigma}_t \boldsymbol{\theta}_t$, hence we obtain the vector of scaled relative price discovery measures

$$\mathbf{f}_{3t} := \frac{\boldsymbol{\theta}_t^f \odot \boldsymbol{\Sigma}_t \boldsymbol{\theta}_t^f}{\boldsymbol{\theta}_t^{f\prime} \boldsymbol{\Sigma}_t \boldsymbol{\theta}_t^f}, \tag{A.27}$$

where $\mathbf{f}_{3t}$ is an $(M \times 1)$ vector of relative price discovery measures; $\boldsymbol{\theta}_t^f$ is again given by the expression in (A.22); and $\mathbf{1}_{1\times M}\mathbf{f}_{3t} = 1$ as $\mathsf{Var}(\varsigma_t|\mathcal{I}_{t-1}) = 0$. It is noticeable that this particular definition of price discovery relies on the same parameter space as the vector of information shares defined in (A.23). However, they differ in that the latter avoids use of an arbitrary Cholesky decomposition in favor of the covariance structure implied by the unobserved components model.

The fourth of the price discovery measures considered in this paper is based on the linearly scaled structural moving average model given by (A.6). Specifically, Yan and Zivot (2006a) use the coefficients in an unscaled version of this equation to derive a relationship between future price changes in a particular constituent market and changes in the permanent innovation to the common efficient price. In the current paper, this is achieved by constructing an impulse response function, which, in the presence of linear scaling is given by

$$\mathbf{f}(\boldsymbol{\theta}^*_{kt}; \widetilde{K}) := \sum_{k=1}^{\widetilde{K}} \frac{\partial \mathsf{E}[\Delta \mathbf{p}_{t+k-1}|\mathcal{I}_t]}{\partial \eta_t^P} = \left[\sum_{k=1}^{\widetilde{K}} \boldsymbol{\theta}^*_{kt+k-1}\right]_{(1:M,1)}, \qquad \widetilde{K} = 1, 2, \ldots \tag{A.28}$$

where $\mathbf{f}(\boldsymbol{\theta}^*_{kt}; \widetilde{K}) \in \mathbb{R}^M$, $\forall \widetilde{K} \in \mathbb{N}$. By noting the result given in (A.16), the second summation term in (A.28) can be expressed as

$$\sum_{k=1}^{\widetilde{K}} \boldsymbol{\theta}^*_{kt+k-1} = \sum_{k=1}^{\widetilde{K}} \boldsymbol{\theta}_{kt+k-1}\mathbf{B}_t^{-1}\mathbf{C}_t; \tag{A.29}$$

moreover, by imposing the assumption given in (A.21), $\mathbf{f}(\boldsymbol{\theta}^*_{kt}; \widetilde{K})$ can eventually be expressed in the following analytical form:

$$\mathbf{f}(\widetilde{\boldsymbol{\Psi}}_t, \boldsymbol{\Sigma}_t; \widetilde{K}) = \left[\left(\left[\mathbf{G}(\widetilde{\boldsymbol{\Psi}}_t; \widetilde{K})\right]_{(1:M,1:M)} + \left[\mathbf{G}(\widetilde{\boldsymbol{\Psi}}_t; \widetilde{K})\right]_{(1:M,Q-N+1:Q)}\boldsymbol{\beta}'\right)\mathbf{B}_t^{-1}\mathbf{C}_t\right]_{(1:M,1)}, \tag{A.30}$$

where $\mathbf{G}(\widetilde{\boldsymbol{\Psi}}_t; \widetilde{K}) = (\mathbf{I}_Q - \widetilde{\boldsymbol{\Psi}}_t^{\widetilde{K}})(\mathbf{I}_Q - \widetilde{\boldsymbol{\Psi}}_t)^{-1}$; $\mathbf{B}_t := (\boldsymbol{\theta}_t^f \vdots \boldsymbol{\beta})'$; $\mathbf{C}_t$ is a unique lower triangular matrix with ones along the principal diagonal such that $\mathbf{C}_t\mathbf{D}_t\mathbf{C}_t' = \boldsymbol{\Sigma}_t$, and $\mathbf{D}_t$ is a unique diagonal matrix with positive entries along the principal diagonal. This definition of price discovery efficiency loss provides a measure of the deviation of prices in each constituent market from equilibrium when a shock to the permanent innovation to the common efficient price occurs – all of which is measured as a function of the time since the shock occurred. Moreover, as η_t^P has a one-to-one long-run effect on the prices in each constituent market, then these deviations can be compared with each other across all markets. Consequently, for markets in which a large amount price discovery occurs, deviations from unity are likely to be short-lived. By contrast, if a constituent market does not generally discover the efficient price then deviations from unity are likely to be persistent.

To summarize the information in the above expression, the function can be summed over a range of values of

$\widetilde{K}$ and a suitable symmetric loss function can be applied to the resulting vector to give a summary measure of price discovery efficiency loss, that is,

$$\mathbf{f}_{4K,t} := \sum_{\widetilde{K}=1}^{K} \mathbf{L}\big(\mathbf{f}(\widetilde{\mathbf{\Psi}}_t, \mathbf{\Sigma}_t; \widetilde{K}) - 1\big), \tag{A.31}$$

where $\mathbf{f}_{4K,t}$ is an $(M \times 1)$ vector of price discovery efficiency loss measures; K is sufficiently large such that $\mathbf{f}(\widetilde{\mathbf{\Psi}}_t, \mathbf{\Sigma}_t; K) \approx 1$; $\mathbf{L}(.)$ is a symmetric loss function to be selected by the investigator; and $\mathbf{1}_{1\times M}\mathbf{f}_{4K,t} \in \mathbb{R}$, $\forall K \in \mathbb{N}$. Unlike the first three measures of price discovery given in (A.18), (A.23), and (A.26), the dynamic nature of the design of the measures defined in (A.30) and (A.31) is such that it takes into consideration the "who moves first" dimension of the price discovery measurement problem. Thus, these measures appear to provide more information regarding the nature of the price discovery process despite relying on the same parameter space as the price discovery measures given by (A.23) and (A.26).

In addition to the above measures, we also consider two variants on these measures. Specifically, the $\mathbf{f}_{4K,t}$ measure of price discovery efficiency loss can be used to provide a standardized (relative) measure of price discovery efficiency loss that sums to unity over all constituent markets, and an aggregate measure of price discovery efficiency loss for the asset under consideration. These measures are obtained via the expressions,

$$s\mathbf{f}_{4K,t} := \frac{\mathbf{f}_{4K,t}}{\mathbf{1}_{1\times M}\mathbf{f}_{4K,t}}, \tag{A.32}$$

$$a\mathrm{f}_{4K,t} := \mathbf{1}_{1\times M}\mathbf{f}_{4K,t}, \tag{A.33}$$

where $s\mathbf{f}_{4K,t}$ is an $(M \times 1)$ vector of standardized (relative) price discovery efficiency loss measures, and $a\mathrm{f}_{4K,t}$ is an aggregated measure of price discovery efficiency loss.

Notes

[1]This is exemplified by the recent acquisition of the *Archipelago* electronic trading network by the New York Stock Exchange (NYSE).

[2]Hasbrouck (1995) defines market fragmentation as the '*dispersal of trading in a security* [or securities that may be technically distinct, but are closely linked by arbitrage or short-term equilibrium considerations] *to multiple sites*'.

[3]See Kyle (1985) and Foster and Viswanathan (1996) for theoretical motivation for the relationship between asymmetric information and asset prices.

[4]A similar theoretical model is introduced by Back and Pedersen (1998); however, they relax the somewhat implausible assumption adopted in the Admati-Pfleiderer model that the *number* of agents collecting information varies systematically over the trading day. This is achieved by allowing the *intensity* with which agents trade on information to vary systematically – an assumption made possible by allowing private information to be long-lived.

[5]The functional form of (2) means that coefficients on past returns from each of the constituent markets are also assumed to be dependent on the set of scaling factors. In addition to the usual microstructure motivation (e.g., bid-ask bounce over the trading day), this dependence (at least for the coefficient on lagged returns observed in the same constituent market) can also be motivated with reference to the rational equilibrium model of Campbell et al (1993). Specifically, this model predicts that non-informed trading causes price movements that, when absorbed by liquidity suppliers, causes prices to revert. Moreover, non-informed (informed) trading will be accompanied by high (low) trading volume. Therefore, price changes accompanied by high (low) trading volume should (should not) revert. Empirical support for this argument is provided by Conrad et al (1999), and Avramov et al (2006). Indeed, the latter study demonstrates that the argument is valid when it is extended to include measures of liquidity instead of trading volume.

[6]The component share and information share measures are *relative* in that they measure the proportional allocation of price discovery across the constituent markets, i.e., they sum to unity.

[7]This paper makes use of a wide range of price discovery measures and does not attempt to differentiate between the quality of each. Rather, the subsequent analysis treats each with uniform emphasis and, as such, avoids criticisms of favoritism for one measure over another. See Baillie et al (2002), de Jong (2002), and Yan and Zivot (2006b), for explicit discussions regarding the similarities amongst, and differences between,

non-augmented versions of these measures.

[8]A similar comparison can be made between the other augmented versions of the price discovery measures and their non-augmented counterparts.

[9]Related results found by Huang (2002) and Barclay and Hendershott (2003) show that electronic communications network (ECN) quotes in NASDAQ stocks are more informative that the quotes posted by dealers.

[10]Though the high frequency impact of macroeconomic announcements upon financial markets has been previously studied (see, e.g., Andersen et al, 2005), this is the first time their impact has been examined with respect to price discovery.

[11]The iShares ETF shares have recently switched to the NYSE. However, during the sample period used in this paper these shares traded on the AME.

[12]The S&P 500 E-mini futures market is not open 24 hours a day. Rather, it closes over the periods: 3:15 p.m. to 3:45 p.m. (Monday to Thursday), and 3:15 p.m. (Friday) to 5:30 p.m. (Sunday).

[13]Further details of the macroeconomic series included in the sample are available upon request.

[14]The E-mini market trade unit is $50 times the index level, while the regular market trade unit is $250 times the index level.

[15]The function $f_s(t)$ is defined such that

$$(t_s, t) := (1,1), (2,2), \ldots, (T_s - 1, T_s - 1), (T_s, T_s), (1, T_s + 1), (2, T_s + 2), \ldots, (T_s, T).$$

[16]Though few studies have used this functional form in the context of equilibrium correction models, it has been used extensively in trade duration models. Most notably, cubic splines have been successfully incorporated into a variety of autoregressive duration (ACD) models (see, e.g., Engle and Russell, 1995, 1998, Zhang, Russell and Tsay, 2001, and Taylor, 2004). It is also possible to use the flexible Fourier form (FFF) to model the periodicity of the parameter space. This functional form has been used to model intraday return volatility dynamics (see, e.g., Andersen and Bollerslev, 1997, 1998). Though the FFF specification is parsimonious and allows for smooth dynamics, it is somewhat rigid in functional form and, therefore, may not be able to capture complex dynamics. For this reason the FFF specification is not used in the current application.

[17]See Ates and Wang (2004, 2005) for empirical evidence regarding the relationship between price discovery and various measures of liquidity.

[18]Trade intensity in the subsequent analysis is defined as the logarithm of one plus one tenth the number of trades in the E-mini futures market. Use of this transformation yields a higher degree of model fit than when no transformation is applied or when other transformation of these data are considered. Further details of these results are available upon request.

[19]We also considered the fifteen minute and one hour periods before and after the announcements. However, the models based on these assumptions exhibited a less satisfactory level of model fit. Further details of these results are available upon request.

[20]Though the model is actually a system of equations, each equation contains the same RHS variables. Therefore, the parameters of the model can be efficiently estimated by single equation OLS.

[21]Details associated with the estimated models based on all parameter values within the space considered are available upon request.

[22]A scaled mean version based on the unexpected change in each of the macroeconomic variables (at the time of the announcement) was also considered. Despite its intuitive appeal, the results proved to be insignificant. For this reason we do not present the results pertaining to this particular scaling factor. Further details are available upon request.

[23]Details pertaining to all estimated parameters (including plots of the parameters over the scaling factor spaces) are available upon request.

[24]The estimates based on the Yan and Zivot (2006a) efficiency loss measure in (A.31) assume that $K = 30$, and $\mathbf{L}(.)$ is given by a quadratic loss function. Selection of this value for K is such that it is sufficiently large to capture the full dynamic responses of shocks to each of the constituent market equations, while the use of quadratic loss function yields similar results to those obtained when an absolute loss function is assumed. The results pertaining to the latter loss function are available upon request.

[25]In modeling the residuals in this way, we are allowing for functional dependence between relative price discovery and relative volatility in (and between) each constituent market. This is because relative volatility is used as an input in most of the measures of price discovery considered in this paper; see, e.g., (A.23) in Appendix A. This represents an improvement in the modeling of price discovery given the previous evidence of a difference in relative price discovery during periods of high and low volatility (Ates and Wang, 2004).

[26]The simplicity of this approach lies in the fact that the estimation process in broken down into three separate procedures. First, the residuals from (1) are (consistently) obtained using OLS. Then, the heteroscedasticity in these residuals is modeled via the application of individual GARCH models. And finally,

a parsimonious parametric model for the correlations is then applied to deliver estimates of the conditional variance-covariance matrix – estimates of which are used as inputs in the price discovery formulae. Further details of the DCC-GARCH model estimated in this paper are available upon request.

[27]The use of all permutations of the equation orderings was originally proposed by Hasbrouck (1995).

[28]Plots associated with the other augmented measures of relative price discovery give similar results and are available upon request.

[29]Such a result is consistent with the hypothesis that E-mini futures liquidity and the degree of information asymmetry are positively related, and inconsistent with the counterargument that market markers in the E-mini futures market increase (decrease) their relative spreads during periods of concentrated informed (uniformed) trading.

[30]These results are not highly dependent on the use of DCC-GARCH conditional volatility estimates. Indeed, similar results are obtained when the (time-invariant) unconditional variances of the residuals in (1) are used to construct the variance-covariance matrix of residuals.

[31]Taking the log of these measures delivers an aggregated measure that is robust to outlying price discovery measures associated with individual constituent markets.

[32]The difference between the component and information share measures of price discovery is analogous to the difference between a coefficient and a partial R^2 in a standard regression framework (de Jong, 2002).

References

Admati, A., and P. Pfleiderer, 1988, A theory of intraday patterns: Volume and price variability, *Review of Financial Studies* 1, 3–40.

Andersen, T., and T. Bollerslev, 1997, Intraday periodicity and volatility persistence in financial markets, *Journal of Empirical Finance* 4, 115–158.

———, 1998, DM-Dollar volatility: Intraday activity patterns, macroeconomic announcements and longer run dependencies, *Journal of Finance* 53, 219–265.

Andersen, T., Bollerslev, T., Diebold, F., and C. Vega, 2005, Real-time price discovery in stock, bond, and foreign exchange markets, *AEA 2005 Philadelphia Meetings*.

Ates, A., and G. Wang, 2004, Liquidity and evolution of price discovery in E-mini versus regular index futures markets, CFTC Working Paper.

———, 2005, Liquidity and the evolution of price discovery floor versus screen-based trading systems: An analysis of the foreign exchange futures markets, CFTC Working Paper.

Avramov, D., Chordia, T., and A. Goyal, 2006, Liquidity and autocorrelations in individual stock returns, *Journal of Finance* 61, 2365–2394.

Back, K., and H. Pedersen, 1998, Long-lived information and intraday patterns, *Journal of Financial Markets* 1, 385–402.

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

Barclay, M., and T. Hendershott, 2003, Price discovery and trading after hours, *Review of Financial Studies* 16, 1041–1073.

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

Campbell, J., Grossman, S., and J. Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905–939.

Conrad, J., Hameed, A., and C. Niden, 1994, Volume and auto-covariances in short-horizon individual security returns, *Journal of Finance* 49, 1305–1329.

Engle, R., 2002, Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Business and Economic Statistics* 20, 339–350.

Engle, R., and J. Russell, 1995, Autoregressive conditional duration: A new model for irregularly spaced transaction data, UCSD Working Paper.

———, 1998, Autoregressive conditional duration: A new model for irregularly spaced transaction data, *Econometrica* 66, 1127–1162.

Foster, F., and S. Viswanathan, 1996, Strategic trading when agents forecast the forecast of others, *Journal of Finance* 51, 1437–1478.

Garbade, K., and W. Silber, 1979, Dominant and satellite markets: A study of dually-traded securities, *Review of Economics and Statistics* 61, 455–460.

———, 1982, Price movements and price discovery in futures and cash markets, *Review of Economics and Statistics* 64, 455–460.

Gonzalo, J., and S. Ng, 2001, A systematic framework for analyzing the dynamic effects of permanent and transitory shocks, *Journal of Economic Dynamics and Control* 25, 1527–1546.

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

Hasbrouck, J., 1993, Assessing the quality of a security market: A new approach to transaction-cost measurement, *Review of Financial Studies* 6, 191–212.

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

———, 2002, Stalking the "efficient price" in market microstructure specifications: An overview, *Journal of Financial Markets* 5, 329–339.

———, 2003, Intraday price formation in US equity index markets, *Journal of Finance* 58, 2375–2399.

Huang, R., 2002, The quality of ECN and Nasdaq market quotes, *Journal of Finance* 57, 1285–1319.

De Jong, F., 2002, Measures of contributions to price discovery: A comparison, *Journal of Financial Markets* 5, 323–327.

De Jong, F., and P. Schotman, 2005, Price discovery in fragmented markets, Maastricht University Working Paper.

Kurov, A., and D. Lasser, 2004, Price dynamics in the regular and E-mini futures markets, *Journal of Financial and Quantitative Analysis* 39, 365–384.

Kyle, A., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1335.

Lei, Q., and G. Wu, 2005, Time-varying informed and uninformed trading activities, *Journal of Financial Markets* 8, 153–181.

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

Taylor, N., Trading intensity, volatility, and arbitrage activity, *Journal of Banking and Finance* 28, 2004, 1137–1162.

Ting, C., 2006, Determinants of intraday stock price changes and asymmetric information, Lee Kong Chian School of Business, Singapore Management University.

White, H., 1982, Maximum likelihood estimation of misspecified models, *Econometrica* 50, 1–25.

Yan, B., and E. Zivot, 2006a, The dynamics of price discovery, unpublished manuscript, Department of Economics, University of Washington.

________, 2006b, A structural analysis of price discovery measures, unpublished manuscript, Department of Economics, University of Washington.

Zhang, M., Russell, J., and R. Tsay, 2001, A nonlinear autoregressive conditional duration model with applications to financial transaction data, *Journal of Econometrics* 61, 2001, 179–207.

Table I: Summary statistics

This table contains summary statistics pertaining to one-minute frequency returns, bilateral mispricings, and trading intensities observed in five S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. Returns are defined as the log of the first difference in prices, multiplied by 100; bilateral mispricing is defined as the demeaned difference between the index level and the price of each of four other securities, multiplied by 100; and trading intensity is defined as the number of trades carried out per minute. The significance of the mean, minimum, maximum, and standard deviation are assessed using a standard asymptotic test obtained under the null hypothesis that each of the measures equal zero. Normality is tested using the Jarque-Bera test, while skewness and (excess) kurtosis are tested under the (asymptotic) assumption that these moments have, respectively, $\mathsf{N}(0, 6/T)$ and $\mathsf{N}(0, 24/T)$ distributions under the appropriate null hypotheses. Autocorrelation is measured as the first-order autoregressive coefficient, with the associated t-test being used to assess its significance. The ARCH test statistic is given by $T \times R^2$, where R^2 is obtained from an OLS regression of squared returns on up to ten lagged values of these squared returns. The significance of each moment or test is denoted by *** (1% significance), ** (5% significance), and * (10% significance).

	Constituent Market				
	(1)	(2)	(3)	(4)	(5)
Panel A: Returns					
Mean	−0.0002	−0.0002	−0.0002	−0.0002	−0.0002
Minimum	−0.6385***	−2.3430***	−0.6367***	−0.7418***	−1.3534***
Maximum	0.7346***	1.8635***	0.6164***	0.8554***	0.6365***
Standard Deviation	0.0521***	0.0789***	0.0704***	0.0717***	0.0705***
Skewness	0.0746***	0.0423***	0.0250***	0.0548***	−0.0179**
Kurtosis	7.3498***	43.2673***	6.1480***	7.0553***	7.8118***
Normality	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Autocorrelation	0.3250***	−0.0444***	−0.0058*	−0.0037	−0.0086***
ARCH	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Panel B: Bilateral Mispricings					
Mean		−0.0018	−0.0008	−0.0008	−0.0007
Minimum		−1.9300***	−1.9100***	−1.7300***	−1.7600***
Maximum		1.8700***	1.3700***	1.4100***	1.5000***
Standard Deviation		0.1219***	0.0598***	0.0663***	0.0611***
Skewness		−0.7507***	−2.9436***	−3.0109***	−2.4373***
Kurtosis		20.5112***	98.8296***	98.4476***	88.0485***
Normality		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Autocorrelation		0.7484***	0.6710***	0.7281***	0.6505***
ARCH		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Panel C: Trading Intensities					
Mean				6.9417***	145.8719***
Minimum				0.0000	0.0000
Maximum				41.0000***	1200.0000***
Standard Deviation				7.9253***	185.6195***
Skewness				1.3220***	1.7592***
Kurtosis				1.9567***	4.0036***
Normality				> 99.9999***	> 99.9999***
Autocorrelation				0.4684***	0.7308***
ARCH				> 99.9999***	> 99.9999***

Table II: Model estimates

This table contains the estimated parameters (and model performance measures) associated with the unscaled (Panel A) and scaled (Panel B) versions of the equilibrium correction model given by equation (1), applied to five S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. In Panel A, numbers in parentheses are the Newey-West standard errors associated with the estimated parameters; while in Panel B, the estimated parameters are represented by their time-series means, with numbers in parentheses the associated time-series Newey-West standard errors. Residual quality is assessed via residual skewness (SK) and kurtosis (KT), the Ljung-Box (LB) test statistic for autocorrelation, and the LM test for ARCH errors. Model fit is measured by the adjusted R^2, and the Akaike and Schwarz Information Criteria (denoted AIC and SIC, respectively). Significance is denoted in the usual way.

	Constituent Market				
Coefficient	(1)	(2)	(3)	(4)	(5)
Panel A: Unscaled model					
$\{\widehat{\mathbf{\Psi}}_0\}_{i1}$	−0.0002	0.0001	−0.0000	−0.0001	−0.0001***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
$\{\widehat{\mathbf{\Psi}}_1\}_{i1}$	0.0900***	−0.0182***	−0.1893***	−0.1731***	0.0163
	(0.0081)	(0.0068)	(0.0116)	(0.0116)	(0.0184)
$\{\widehat{\mathbf{\Psi}}_1\}_{i2}$	−0.0030	0.1341***	0.2654***	0.2524***	0.2366***
	(0.0027)	(0.0113)	(0.0107)	(0.0115)	(0.0112)
$\{\widehat{\mathbf{\Psi}}_1\}_{i3}$	−0.0228***	−0.0544***	−0.0213***	−0.0229***	−0.0241***
	(0.0076)	(0.0098)	(0.0038)	(0.0041)	(0.0041)
$\{\widehat{\mathbf{\Psi}}_1\}_{i4}$	0.0469***	−0.0315***	−0.0417***	−0.0672***	−0.0615***
	(0.0074)	(0.0105)	(0.0110)	(0.0114)	(0.0118)
$\{\widehat{\mathbf{\Psi}}_1\}_{i5}$	0.1770***	0.0445***	0.1889***	0.1874***	−0.0157
	(0.0107)	(0.0117)	(0.0027)	(0.0167)	(0.0115)
$\{\widehat{\boldsymbol{\alpha}}\}_{i1}$	0.0014	0.3281***	−0.0259***	−0.0308***	−0.0288***
	(0.0020)	(0.0067)	(0.0182)	(0.0029)	(0.0029)
$\{\widehat{\boldsymbol{\alpha}}\}_{i2}$	−0.0153	0.1039***	0.5972***	−0.1134***	−0.0991***
	(0.0111)	(0.0125)	(0.0182)	(0.0091)	(0.0178)
$\{\widehat{\boldsymbol{\alpha}}\}_{i3}$	0.1025***	−0.1440***	−0.1327***	0.5103***	−0.0597***
	(0.0103)	(0.0125)	(0.0155)	(0.0084)	(0.0165)
$\{\widehat{\boldsymbol{\alpha}}\}_{i4}$	−0.2038***	−0.1140***	−0.3386***	−0.3453***	0.2885***
	(0.0182)	(0.0163)	(0.0300)	(0.0089)	(0.0328)
SK	0.0501***	0.2807***	0.0389***	0.0415***	0.0324***
KT	6.4580***	37.3709***	6.6730***	6.7468***	6.7141***
LB	12.8042***	1.1730	4.4043**	1.8304	0.0250
ARCH	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
$\overline{R}^2$	0.1901	0.2708	0.1434	0.1112	0.0248
System $\overline{R}^2$			0.1502		
AIC	−3.2833	−2.5583	−2.6248	−2.5498	−2.4928
System AIC			−17.9044		
SIC	−3.2823	−2.5573	−2.6238	−2.5489	−2.4918
System SIC			−17.9000		

Table II: Model estimates (continued)

This table contains the estimated parameters (and model performance measures) associated with the unscaled (Panel A) and scaled (Panel B) versions of the equilibrium correction model given by equation (1), applied to five S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. In Panel A, numbers in parentheses are the Newey-West standard errors associated with the estimated parameters; while in Panel B, the estimated parameters are represented by their time-series means, with numbers in parentheses the associated time-series Newey-West standard errors. Residual quality is assessed via residual skewness (SK) and kurtosis (KT), the Ljung-Box (LB) test statistic for autocorrelation, and the LM test for ARCH errors. Model fit is measured by the adjusted R^2, and the Akaike and Schwarz Information Criteria (denoted AIC and SIC, respectively). Significance is denoted in the usual way.

	Constituent Market				
Coefficient	(1)	(2)	(3)	(4)	(5)
Panel B: Scaled model					
$\{\widehat{\mathbf{\Psi}}_0\}_{i1}$	−0.0002	0.0003	0.0001	0.0000	0.0000
	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
$\{\widehat{\mathbf{\Psi}}_{1t}\}_{i1}$	0.0994***	−0.0115***	−0.1026***	−0.0628***	0.0218***
	(0.0019)	(0.0006)	(0.0008)	0.0011)	(0.0028)
$\{\widehat{\mathbf{\Psi}}_{1t}\}_{i2}$	−0.0036***	0.0500***	0.2474***	0.2422***	0.2288***
	(0.0003)	(0.0022)	(0.0031)	(0.0027)	(0.0028)
$\{\widehat{\mathbf{\Psi}}_{1t}\}_{i3}$	−0.0046**	−0.0266***	−0.0103***	−0.0119***	−0.0180***
	(0.0019)	(0.0011)	(0.0003)	(0.0006)	(0.0005)
$\{\widehat{\mathbf{\Psi}}_{1t}\}_{i4}$	0.0553***	−0.0025	0.0085***	−0.0354***	−0.0402***
	(0.0012)	(0.0009)	(0.0011)	(0.0011)	(0.0018)
$\{\widehat{\mathbf{\Psi}}_{1t}\}_{i5}$	0.1351***	0.0301***	0.1022***	0.1069***	0.0141***
	(0.0018)	(0.0009)	(0.0018)	(0.0017)	(0.0007)
$\{\widehat{\boldsymbol{\alpha}}_t\}_{i1}$	0.0009***	0.3101***	−0.0099***	−0.0127***	−0.0158***
	(0.0002)	(0.0018)	(0.0002)	(0.0003)	(0.0004)
$\{\widehat{\boldsymbol{\alpha}}_t\}_{i2}$	0.0201***	−0.0546***	0.7814***	−0.0445***	−0.0529***
	(0.0017)	(0.0017)	(0.0076)	(0.0024)	(0.0024)
$\{\widehat{\boldsymbol{\alpha}}_t\}_{i3}$	0.0890***	−0.0550***	−0.0557***	0.7484***	−0.0379***
	(0.0012)	(0.0009)	(0.0023)	(0.0048)	(0.0019)
$\{\widehat{\boldsymbol{\alpha}}_t\}_{i4}$	−0.2547***	−0.1256***	−0.4756***	−0.4759***	0.3418***
	(0.0034)	(0.0018)	(0.0057)	(0.0047)	(0.0026)
SK	0.0656***	0.4403***	0.0554***	0.0761***	0.0314***
KT	6.2148***	36.0296***	6.4142***	6.6072***	6.6122***
LB	6.2122**	1.7093	1.5663	0.5099	0.0067
ARCH	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
$\overline{R}^2$	0.2142	0.2891	0.1692	0.1419	0.0370
System $\overline{R}^2$			0.1720		
AIC	−3.3115	−2.5823	−2.6543	−2.5838	−2.5034
System AIC			−18.2246		
SIC	−3.2973	−2.5682	−2.6402	−2.5697	−2.4893
System SIC			−18.1544		

Table III: Testing coefficient restrictions

This table contains the (robust) Wald test statistics associated with various coefficient restrictions imposed on the unscaled and scaled versions of the equilibrium correction model given by equation (1), applied to five S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. Panel A of the table contains Wald test statistics associated with restrictions regarding the overall significance of various groups of coefficients, while Panel B contains Wald test statistics associated with restrictions imposed upon various groups of coefficients within various scaling factor sets. The scaling factor sets are defined as follows: an intraday periodic scaling factor set ($\mathcal{S}_1$), an interday periodic scaling factor set based on the time-to-maturity of the futures contracts ($\mathcal{S}_2$), a scaling factor set based on a measure of liquidity ($\mathcal{S}_3$), and two scaling factor sets based on the time period before and after the announcement of key macroeconomic data ($\mathcal{S}_{4a}$ and $\mathcal{S}_{4b}$, respectively), with unions of these sets defined (and denoted) such that $\mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4$ and $\mathcal{S}_4 := \mathcal{S}_{4a} \cup \mathcal{S}_{4b}$. Significance is denoted in the usual way.

		Constituent Market				
Variable Restriction	Factor(s)	(1)	(2)	(3)	(4)	(5)
Panel A: Unscaled model						
All	Unity	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Own lagged		> 99.9999***	> 99.9999***	> 99.9999***	34.9260***	1.8498
Other Lagged		> 99.9999***	61.6873***	> 99.9999***	> 99.9999***	> 99.9999***
Mispricing		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Other Lagged plus Mispricing		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Panel B: Scaled model						
All	$\mathcal{S}$ (inc. unity)	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Own lagged		> 99.9999***	63.6800***	14.9605	38.0400***	17.8297
Other Lagged		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Mispricing		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Other Lagged plus Mispricing		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
All	$\mathcal{S}$ (ex. unity)	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Own lagged		> 99.9999***	30.0009**	6.1395	25.1910**	17.1307
Other Lagged		> 99.9999***	78.5429*	> 99.9999***	> 99.9999***	> 99.9999***
Mispricing		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Other Lagged plus Mispricing		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
All	$\mathcal{S}_1$ (ex. unity)	> 99.9999***	99.2854***	> 99.9999***	> 99.9999***	> 99.9999***
Own Lagged		47.9649***	6.6194*	0.9865	6.0994	1.2241
Other Lagged		22.3836**	13.9475	54.3221***	33.3101***	34.7689***
Mispricing		38.8455***	55.9264***	82.6989***	84.3354***	> 99.9999***
Other Lagged plus Mispricing		57.5550***	64.9721***	> 99.9999***	> 99.9999***	> 99.9999***
All	$\mathcal{S}_2$ (ex. unity)	> 99.9999***	47.4034***	> 99.9999***	> 99.9999***	90.1758***
Own Lagged		21.9916***	7.8567**	1.8136	1.2783	1.0117
Other Lagged		21.5526**	15.3107	58.3592***	40.4438***	44.6238***
Mispricing		54.5560***	7.6534	15.9407	64.1538***	26.4574***
Other Lagged plus Mispricing		78.2147***	27.2261	95.8773***	> 99.9999***	67.4363***
All	$\mathcal{S}_3$ (ex. unity)	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Own Lagged		86.4056***	14.8912***	1.0192	8.5171**	1.8978
Other Lagged		92.6038***	25.8265**	> 99.9999***	71.3226***	> 99.9999***
Mispricing		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***
Other Lagged plus Mispricing		> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***	> 99.9999***

Table III: Testing coefficient restrictions (continued)

This table contains the (robust) Wald test statistics associated with various coefficient restrictions imposed on the unscaled and scaled versions of the equilibrium correction model given by equation (1), applied to five S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. Panel A of the table contains Wald test statistics associated with restrictions regarding the overall significance of various groups of coefficients, while Panel B contains Wald test statistics associated with restrictions imposed upon various groups of coefficients within various scaling factor sets. The scaling factor sets are defined as follows: an intraday periodic scaling factor set ($\mathcal{S}_1$), an interday periodic scaling factor set based on the time-to-maturity of the futures contracts ($\mathcal{S}_2$), a scaling factor set based on a measure of liquidity ($\mathcal{S}_3$), and two scaling factor sets based on the time period before and after the announcement of key macroeconomic data ($\mathcal{S}_{4a}$ and $\mathcal{S}_{4b}$, respectively), with unions of these sets defined (and denoted) such that $\mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 \cup \mathcal{S}_4$ and $\mathcal{S}_4 := \mathcal{S}_{4a} \cup \mathcal{S}_{4b}$. Significance is denoted in the usual way.

		Constituent Market				
Variable Restriction	Factor(s)	(1)	(2)	(3)	(4)	(5)
Panel B: Scaled model (continued)						
All	$\mathcal{S}_{4a}$ (ex. unity)	31.6949	26.6906	33.0253	34.2065	35.0701
Own Lagged		5.1151	3.1250	1.0427	0.6679	4.9261
Other Lagged		13.6388	13.6040	10.5016	7.2182	9.5631
Mispricing		14.2009	14.2978	23.1663**	20.7860*	19.4763*
Other Lagged plus Mispricing		25.4043	24.8564	31.5229	31.0738	29.9009
All	$\mathcal{S}_{4b}$ (ex. unity)	32.2093	19.7389	31.6679	34.2519	34.6724
Own Lagged		1.7545	0.6914	0.9972	0.6288	6.5830*
Other Lagged		18.5827*	8.0193	18.6623*	17.4175	6.7915
Mispricing		10.6581	7.0062	16.9962	13.4591	18.6823*
Other Lagged plus Mispricing		30.0571	18.9389	30.7439	30.5287	26.5234

Table IV: Estimated price discovery

This table contains estimates of the following measures of relative price discovery: the non-augmented and augmented Harris et al (2002) component share measures ($\widehat{\mathbf{f}}_1$ and $\widehat{\mathbf{f}}_{1t}$, respectively); the non-augmented and augmented Hasbrouck (1995) information share measures ($\widehat{\mathbf{f}}_2$ and $\widehat{\mathbf{f}}_{2t}$, respectively); the non-augmented and augmented de Jong and Schotman (2005) information share measures ($\widehat{\mathbf{f}}_3$ and $\widehat{\mathbf{f}}_{3t}$, respectively); and the non-augmented and augmented Yan and Zivot (2006a) standardized price discovery measures ($-s\widehat{\mathbf{f}}_{4K}$ and $s\widehat{\mathbf{f}}_{4K,t}$, respectively) – all of which are based on the unscaled version (Panel A) or scaled version (Panel B) of the equilibrium correction model given by equation (1). In the case of the augmented measures of price discovery, the entries correspond to the time-series averages of each measure over the sample period. All of these measures are calculated for all five (M) S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. Inference is based on a 1,000 repetition bootstrap procedure with significance denoted in the usual way.

	Constituent Market				
Discovery measure	(1)	(2)	(3)	(4)	(5)
Panel A: Unscaled model					
$\{\widehat{\mathbf{f}}_1\}_i - 1/M$	0.2410***	−0.1573	−0.1099	−0.2034	0.2297***
$\{\widehat{\mathbf{f}}_2\}_i - 1/M$	0.0645***	−0.1904	0.0190***	0.0078***	0.0991***
$\{\widehat{\mathbf{f}}_3\}_i - 1/M$	0.1586***	−0.1901	−0.1015	−0.2038	0.3368***
$-(\{s\widehat{\mathbf{f}}_{4K}\}_i - 1/M)$	0.1049***	0.0057	−0.0253	−0.1085	0.0232***
$\{\widehat{\mathbf{f}}_1\}_1 - \{\widehat{\mathbf{f}}_1\}_i$	0.0000	0.3983***	0.3509***	0.4444***	0.0113
$\{\widehat{\mathbf{f}}_1\}_2 - \{\widehat{\mathbf{f}}_1\}_i$		0.0000	−0.0474***	0.0461	−0.3870***
$\{\widehat{\mathbf{f}}_1\}_3 - \{\widehat{\mathbf{f}}_1\}_i$			0.0000	0.0935***	−0.3396***
$\{\widehat{\mathbf{f}}_1\}_4 - \{\widehat{\mathbf{f}}_1\}_i$				0.0000	−0.4331***
$\{\widehat{\mathbf{f}}_2\}_1 - \{\widehat{\mathbf{f}}_2\}_i$	0.0000	0.2549***	0.0455***	0.0567***	−0.0346***
$\{\widehat{\mathbf{f}}_2\}_2 - \{\widehat{\mathbf{f}}_2\}_i$		0.0000	−0.2094***	−0.1982***	−0.2895***
$\{\widehat{\mathbf{f}}_2\}_3 - \{\widehat{\mathbf{f}}_2\}_i$			0.0000	0.0122***	−0.0801***
$\{\widehat{\mathbf{f}}_2\}_4 - \{\widehat{\mathbf{f}}_2\}_i$				0.0000	−0.0913***
$\{\widehat{\mathbf{f}}_3\}_1 - \{\widehat{\mathbf{f}}_3\}_i$	0.0000	0.3487***	0.2601***	0.3624***	−0.1782***
$\{\widehat{\mathbf{f}}_3\}_2 - \{\widehat{\mathbf{f}}_3\}_i$		0.0000	−0.0886***	0.0137	−0.5269***
$\{\widehat{\mathbf{f}}_3\}_3 - \{\widehat{\mathbf{f}}_3\}_i$			0.0000	0.1023***	−0.4383***
$\{\widehat{\mathbf{f}}_3\}_4 - \{\widehat{\mathbf{f}}_3\}_i$				0.0000	−0.5406***
$-(\{s\widehat{\mathbf{f}}_{4K}\}_1 - \{s\widehat{\mathbf{f}}_{4K}\}_i)$	0.0000	0.0992***	0.1302***	0.2134***	0.0817***
$-(\{s\widehat{\mathbf{f}}_{4K}\}_2 - \{s\widehat{\mathbf{f}}_{4K}\}_i)$		0.0000	0.0501***	0.1142***	−0.0175**
$-(\{s\widehat{\mathbf{f}}_{4K}\}_3 - \{s\widehat{\mathbf{f}}_{4K}\}_i)$			0.0000	0.0832***	0.0485***
$-(\{s\widehat{\mathbf{f}}_{4K}\}_4 - \{s\widehat{\mathbf{f}}_{4K}\}_i)$				0.0000	−0.1317***
Panel B: Scaled model					
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	0.3706***	−0.1772	−0.1386	−0.2375	0.1826***
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$	0.1137***	−0.1930	−0.0050	−0.0066	0.0907***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$	0.3268***	−0.1943	−0.1725	−0.2472	0.2872***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$	0.1074***	0.0364***	−0.0452	−0.0819	−0.0167
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$	0.0000	0.5478***	0.5092***	0.6080***	0.1880***
$\{\widehat{\mathbf{f}}_{1t}\}_2 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	−0.0386	0.0603***	−0.3598***
$\{\widehat{\mathbf{f}}_{1t}\}_3 - \{\widehat{\mathbf{f}}_{1t}\}_i$			0.0000	0.0989***	−0.3212***
$\{\widehat{\mathbf{f}}_{1t}\}_4 - \{\widehat{\mathbf{f}}_{1t}\}_i$				0.0000	−0.4201***

Table IV: Estimated price discovery (continued)

This table contains estimates of the following measures of relative price discovery: the non-augmented and augmented Harris et al (2002) component share measures ($\widehat{\mathbf{f}}_1$ and $\widehat{\mathbf{f}}_{1t}$, respectively); the non-augmented and augmented Hasbrouck (1995) information share measures ($\widehat{\mathbf{f}}_2$ and $\widehat{\mathbf{f}}_{2t}$, respectively); the non-augmented and augmented de Jong and Schotman (2005) information share measures ($\widehat{\mathbf{f}}_3$ and $\widehat{\mathbf{f}}_{3t}$, respectively); and the non-augmented and augmented Yan and Zivot (2006a) standardized price discovery measures ($-s\widehat{\mathbf{f}}_{4K}$ and $s\widehat{\mathbf{f}}_{4K,t}$, respectively) – all of which are based on the unscaled version (Panel A) or scaled version (Panel B) of the equilibrium correction model given by equation (1). In the case of the augmented measures of price discovery, the entries correspond to the time-series averages of each measure over the sample period. All of these measures are calculated for all five (M) S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. Inference is based on a 1,000 repetition bootstrap procedure with significance denoted in the usual way.

	Constituent Market				
Discovery measure	(1)	(2)	(3)	(4)	(5)
Panel B: Scaled model (continued)					
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$	0.0000	0.3067***	0.1187***	0.1203***	0.0230***
$\{\widehat{\mathbf{f}}_{2t}\}_2 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	−0.1880***	−0.1864***	−0.2837***
$\{\widehat{\mathbf{f}}_{2t}\}_3 - \{\widehat{\mathbf{f}}_{2t}\}_i$			0.0000	0.0016***	−0.0957***
$\{\widehat{\mathbf{f}}_{2t}\}_4 - \{\widehat{\mathbf{f}}_{2t}\}_i$				0.0000	−0.0973***
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$	0.0000	0.5212***	0.4993***	0.5740***	0.0396***
$\{\widehat{\mathbf{f}}_{3t}\}_2 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	−0.0218***	0.0528***	−0.4993***
$\{\widehat{\mathbf{f}}_{3t}\}_3 - \{\widehat{\mathbf{f}}_{3t}\}_i$			0.0000	0.0747***	−0.4598***
$\{\widehat{\mathbf{f}}_{3t}\}_4 - \{\widehat{\mathbf{f}}_{3t}\}_i$				0.0000	−0.5344***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$	0.0000	0.0710***	0.1525***	0.1893***	0.1241***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_2 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	0.0815***	0.1182***	0.0530***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_3 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$			0.0000	0.0367***	−0.0285***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_4 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$				0.0000	−0.0652***

Table V: Estimated price discovery and the trading day

This table contains estimates of the following event-specific conditional mean measures of price discovery: the augmented Harris et al (2002) component share measure ($\widehat{\mathbf{f}}_{1t}$); the augmented Hasbrouck (1995) information share measure ($\widehat{\mathbf{f}}_{2t}$); the augmented de Jong and Schotman (2005) information share measure ($\widehat{\mathbf{f}}_{3t}$); the augmented Yan and Zivot (2006a) standardized price discovery measure ($-s\widehat{\mathbf{f}}_{4K,t}$); and the augmented Yan and Zivot (2006a) aggregated price discovery measure ($-a\ln\widehat{\mathrm{f}}_{4K,t}$) – all of which are based on the scaled version of the equilibrium correction model given by equation (1). The pair of events considered in this table are the market open and market close periods, where the market open period is defined as the first fifteen minutes of trading after the start of trading in the major S&P 500 markets, and the market close period is defined as the last fifteen minutes of trading prior to the end of trading in the major S&P 500 markets. Differences between the market open and close conditional mean measures of price discovery over these events are denoted by $\Delta(\widehat{\mathbf{f}}_{1t})$, $\Delta(\widehat{\mathbf{f}}_{2t})$, $\Delta(\widehat{\mathbf{f}}_{3t})$, $-\Delta(s\widehat{\mathbf{f}}_{4K,t})$, and $-\Delta(a\ln\widehat{\mathrm{f}}_{4K,t})$. All of these measures are calculated for all five (M) S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. Inference is based on a 1,000 repetition bootstrap procedure with significance denoted in the usual way.

		Constituent Market				
Discovery measure	Event(s)	(1)	(2)	(3)	(4)	(5)
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	Open (O)	−0.0887	−0.0094	−0.6603	−0.5937	1.3521***
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$		0.0524***	−0.1741	0.0150***	−0.0100	0.1167***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$		0.1059***	−0.1690	−0.0392	−0.2798	0.3821***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$		0.0089***	0.0435***	−0.0038	−0.0605	0.0119***
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	−0.0793	0.5716	0.5050***	−1.4408
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	0.2265***	0.0374***	0.0625***	−0.0643***
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	0.2748***	0.1450***	0.3857***	−0.2762***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	−0.0346***	0.0127***	0.0694***	−0.0030
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	Close (C)	0.6399***	−0.2438	−0.3345	−0.1773	0.1157***
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$		0.1388***	−0.1960	−0.0296	−0.0004	0.0872***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$		0.4070***	−0.2020	−0.3042	−0.1690	0.2681***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$		0.1123***	0.0380***	−0.0631	−0.0524	−0.0348
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	0.8836***	0.9743***	0.8171***	0.5242***
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	0.3348***	0.1684***	0.1392***	0.0516***
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	0.6090***	0.7112***	0.5760***	0.1389***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	0.0743***	0.1755***	0.1647***	0.1471***
$\Delta(\{\widehat{\mathbf{f}}_{1t}\}_i)$	O *v.* C	−0.7285***	0.2344***	−0.3259	−0.4164	1.2364
$\Delta(\{\widehat{\mathbf{f}}_{2t}\}_i)$		−0.0864***	0.0220***	0.0446***	−0.0096***	0.0295***
$\Delta(\{\widehat{\mathbf{f}}_{3t}\}_i)$		−0.3012***	0.0330***	0.2651***	−0.1108***	0.1140***
$-\Delta(\{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		−0.1034***	0.0055	0.0593***	0.0081**	0.0467***
$-\Delta(a\ln\mathrm{f}_{4K,t})$				1.1497		

Table VI: Estimated price discovery and liquidity

This table contains estimates of the following event-specific conditional mean measures of price discovery: the augmented Harris et al (2002) component share measure ($\widehat{\mathbf{f}}_{1t}$); the augmented Hasbrouck (1995) information share measure ($\widehat{\mathbf{f}}_{2t}$); the augmented de Jong and Schotman (2005) information share measure ($\widehat{\mathbf{f}}_{3t}$); the augmented Yan and Zivot (2006a) standardized price discovery measure ($-s\widehat{\mathbf{f}}_{4K,t}$); and the augmented Yan and Zivot (2006a) aggregated price discovery measure ($-a \ln \widehat{\mathrm{f}}_{4K,t}$) – all of which are based on the scaled version of the equilibrium correction model given by equation (1). The pair of events considered in this table are low liquidity and high liquidity, where low (high) liquidity is defined as all liquidity amounts below (above) the median level of liquidity in the E-mini futures market. Differences between the low and high liquidity conditional mean measures of price discovery over these events are denoted by $\Delta(\widehat{\mathbf{f}}_{1t})$, $\Delta(\widehat{\mathbf{f}}_{2t})$, $\Delta(\widehat{\mathbf{f}}_{3t})$, $-\Delta(s\widehat{\mathbf{f}}_{4K,t})$, and $-\Delta(a \ln \widehat{\mathrm{f}}_{4K,t})$. All of these measures are calculated for all five (M) S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. Inference is based on a 1,000 repetition bootstrap procedure with significance denoted in the usual way.

		Constituent Market				
Discovery measure	Event(s)	(1)	(2)	(3)	(4)	(5)
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	Low Liq. (L)	0.3504***	−0.1739	−0.0837	−0.1940	0.1012***
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$		0.1248***	−0.1930	−0.0001	−0.0044	0.0726***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$		0.3561***	−0.1945	−0.1530	−0.2084	0.1997***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$		0.1067***	0.0531***	−0.0430	−0.0703	−0.0464
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	0.5243***	0.4342***	0.5444***	0.2493***
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	0.3178***	0.1249***	0.1292***	0.0523***
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	0.5506***	0.5092***	0.5645***	0.1564***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	0.0536***	0.1497***	0.1769***	0.1531***
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	High Liq. (H)	0.3910***	−0.1805	−0.1942	−0.2815	0.2652***
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$		0.1025***	−0.1929	−0.0099	−0.0088	0.1091***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$		0.2971***	−0.1942	−0.1923	−0.2865	0.3759***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$		0.1081***	0.0194***	−0.0473	−0.0937	0.0135***
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	0.5715***	0.5851***	0.6725***	0.1258***
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	0.2954***	0.1124***	0.1113***	−0.0066***
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	0.4913***	0.4894***	0.5836***	−0.0788***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	0.0887***	0.1554***	0.2018***	0.0946***
$\Delta(\{\widehat{\mathbf{f}}_{1t}\}_i)$	L v. H	−0.0405	0.0066	0.1104	0.0875***	−0.1640***
$\Delta(\{\widehat{\mathbf{f}}_{2t}\}_i)$		0.0223***	−0.0001	0.0099***	0.0044***	−0.0366***
$\Delta(\{\widehat{\mathbf{f}}_{3t}\}_i)$		0.0590***	−0.0003	0.0393***	0.0782***	−0.1762***
$-\Delta(\{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		−0.0015	0.0336***	0.0043***	0.0234***	−0.0599***
$-\Delta(a \ln \mathrm{f}_{4K,t})$				−3.0992***		

Table VI: Estimated price discovery around macroeconomic announcements

This table contains estimates of the following event-specific conditional mean measures of price discovery: the augmented Harris et al (2002) component share measure ($\widehat{\mathbf{f}}_{1t}$); the augmented Hasbrouck (1995) information share measure ($\widehat{\mathbf{f}}_{2t}$); the augmented de Jong and Schotman (2005) information share measure ($\widehat{\mathbf{f}}_{3t}$); the augmented Yan and Zivot (2006a) standardized price discovery measure ($-s\widehat{\mathbf{f}}_{4K,t}$); and the augmented Yan and Zivot (2006a) aggregated price discovery measure ($-a\ln\widehat{\mathrm{f}}_{4K,t}$) – all of which are based on the scaled version of the equilibrium correction model given by equation (1). The pair of events considered in this table are the before and after macroeconomic announcement periods, where the before (after) announcement period is defined as the fifteen minute period prior to (after) the announcement. Differences between the before and after announcement conditional mean measures of price discovery over these events are denoted by $\Delta(\widehat{\mathbf{f}}_{1t})$, $\Delta(\widehat{\mathbf{f}}_{2t})$, $\Delta(\widehat{\mathbf{f}}_{3t})$, $-\Delta(s\widehat{\mathbf{f}}_{4K,t})$, and $-\Delta(a\ln\widehat{\mathrm{f}}_{4K,t})$. All of these measures are calculated for all five (M) S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. Inference is based on a 1,000 repetition bootstrap procedure with significance denoted in the usual way.

		Constituent Market				
Discovery measure	Event(s)	(1)	(2)	(3)	(4)	(5)
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	Before Ann. (B)	0.2955*	−0.1615	−0.2323	0.0949	0.0034
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$		0.1220***	−0.1920	−0.0140	0.0122	0.0718***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$		0.3556***	−0.1935	−0.1861	−0.1776	0.2017***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$		0.0936***	0.0616***	−0.0579	−0.0873	−0.0100
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	0.4570***	0.5278***	0.2006	0.2921***
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	0.3139***	0.1360***	0.1097***	0.0502***
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	0.5491***	0.5417***	0.5332***	0.1539***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	0.0319***	0.1514***	0.1808***	0.1035***
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	After Ann. (A)	0.2792***	−0.2087	−0.1695	−0.1890	0.2881***
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$		0.0782***	−0.1921	0.0020	0.0077***	0.1043***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$		0.1938***	−0.1957	−0.1546	−0.1833	0.3398***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$		0.0749***	0.0047***	−0.0374	−0.0678	0.0256***
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	0.4880***	0.4488***	0.4682***	−0.0088
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	0.2704***	0.0763***	0.0706***	−0.0260***
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	0.3894***	0.3483***	0.3771***	−0.1460***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	0.0702***	0.1122***	0.1427***	0.0493***
$\Delta(\{\widehat{\mathbf{f}}_{1t}\}_i)$	B *v.* A	0.0163	0.0473**	−0.0628***	0.2839	−0.2847***
$\Delta(\{\widehat{\mathbf{f}}_{2t}\}_i)$		0.0437***	0.0002	−0.0160***	0.0046***	−0.0325***
$\Delta(\{\widehat{\mathbf{f}}_{3t}\}_i)$		0.1618***	0.0022**	−0.0316	0.0057	−0.1381***
$-\Delta(\{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0187***	0.0570***	−0.0205***	−0.0195***	−0.0356***
$-\Delta(a\ln\mathrm{f}_{4K,t})$				−1.2936		

Table VIII: Estimated price discovery, the trading day, and liquidity

This table contains estimates of the following event-specific conditional mean measures of price discovery: the augmented Harris et al (2002) component share measure ($\widehat{\mathbf{f}}_{1t}$); the augmented Hasbrouck (1995) information share measure ($\widehat{\mathbf{f}}_{2t}$); the augmented de Jong and Schotman (2005) information share measure ($\widehat{\mathbf{f}}_{3t}$); the augmented Yan and Zivot (2006a) standardized price discovery measure ($-s\widehat{\mathbf{f}}_{4K,t}$); and the augmented Yan and Zivot (2006a) aggregated price discovery measure ($-a\ln\widehat{\mathrm{f}}_{4K,t}$) – all of which are based on the scaled version of the equilibrium correction model given by equation (1). The two pair of events considered in this table are the market open and close, and low and high liquidity, where the market open period is defined as the first fifteen minutes of trading after the start of trading in the major S&P 500 markets, the market close period is defined as the last fifteen minutes of trading prior to the end of trading in the major S&P 500 markets, and low (high) liquidity is defined as all liquidity amounts below (above) the median level of liquidity in the E-mini futures market. Differences between the market open and close conditional mean measures of price discovery for each level of liquidity are denoted by $\Delta(\widehat{\mathbf{f}}_{1t})$, $\Delta(\widehat{\mathbf{f}}_{2t})$, $\Delta(\widehat{\mathbf{f}}_{3t})$, $-\Delta(s\widehat{\mathbf{f}}_{4K,t})$, and $-\Delta(a\ln\widehat{\mathrm{f}}_{4K,t})$. All of these measures are calculated for all five (M) S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. Inference is based on a 1,000 repetition bootstrap procedure with significance denoted in the usual way.

Discovery measure	Event(s)	Constituent Market (1)	(2)	(3)	(4)	(5)	Constituent Market (1)	(2)	(3)	(4)	(5)
		Event(s)									
		Low Liquidity ($n = 33$)					High Liquidity ($n = 218$)				
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	Open (O)	-1.4802	0.4201	-4.6463	-2.5858	8.2922	0.1220***	-0.0744	-0.0569	-0.2921	0.3015***
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$		0.0710***	-0.1764	0.0123	-0.0171	0.1103***	0.0496***	-0.1737	0.0154***	-0.0090	0.1177***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$		0.1391*	-0.1726	0.0249	-0.1505	0.1590**	0.1008***	-0.1684	-0.0489	-0.2994	0.4159***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$		-0.0092	0.0669***	0.0065*	-0.0364	-0.0148	0.0116***	0.0400***	-0.0034	-0.0642	0.0160***
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	-1.9003	3.1662	1.1056*	-9.7724	0.0000	0.1964***	0.1789***	0.4140***	-0.1795^{***}
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	0.2474***	0.0587	0.0881***	-0.0393	0.0000	0.2233***	0.0342***	0.0586	-0.0681^{***}
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	0.3117***	0.1142	0.2896***	-0.0199	0.0000	0.2693***	0.1497***	0.4002***	-0.3150^{***}
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	-0.0761^{***}	-0.0026	0.0272***	0.0056	0.0000	-0.0284^{***}	0.0151***	0.0758***	0.0044
		Low Liquidity ($n = 100$)					High Liquidity ($n = 151$)				
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	Close (C)	0.8445***	-0.2536	-0.3568	-0.1568	-0.0773	0.5044***	-0.2373	-0.3197	-0.1908	0.2435***
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$		0.1687***	-0.1952	-0.0324	-0.0035	0.0624***	0.1190***	-0.1966	-0.0277	0.0016	0.1037***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$		0.1030***	-0.2009	-0.2920	-0.1367	0.1486***	0.3580***	-0.2026	-0.3123	-0.1904	0.3473***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$		0.0945***	0.0475***	-0.0581	-0.0370	-0.0554	0.1185***	0.0318***	-0.0665	-0.0626	-0.0212^{***}
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	1.0980***	1.2012***	1.0013***	0.9218***	0.0000	0.7417***	0.8241***	0.6952***	0.2609***
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	0.3639***	0.2011***	0.1722***	0.1063***	0.0000	0.3156***	0.1467***	0.1173***	0.0153***
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	0.6820***	0.7730***	0.6177***	0.3325***	0.0000	0.5606***	0.6703***	0.5483***	0.0106
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	0.0555***	0.1611***	0.1399***	0.1584***	0.0000	0.0867***	0.1850***	0.1811***	0.1397***
		Low Liquidity ($n = 133$)					High Liquidity ($n = 369$)				
$\Delta(\{\widehat{\mathbf{f}}_{1t}\}_i)$	O *v.* C	-2.3247	0.6737	-4.2896	-2.4289	8.3695	-0.3824^{***}	0.1629***	-0.3197^{***}	-0.1908^{***}	0.0580***
$\Delta(\{\widehat{\mathbf{f}}_{2t}\}_i)$		-0.0977^{***}	0.0188***	0.0447***	-0.0137	0.0479**	-0.0694^{***}	0.0229***	-0.0277^{***}	0.0016	0.0140***
$\Delta(\{\widehat{\mathbf{f}}_{3t}\}_i)$		-0.3419^{***}	0.0283***	0.3169***	-0.0138^{***}	0.0105	-0.2572^{***}	0.0342***	-0.3123^{***}	-0.1904^{***}	0.0685***
$-\Delta(\{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		-0.1121^{***}	0.0194	0.0516***	0.0006	0.0406***	-0.1069^{***}	0.0082***	-0.0665^{***}	-0.0626^{***}	0.0372***
$-\Delta(a\ln\mathrm{f}_{4K,t})$				-5.1715					0.6278		

Table IX: Estimated price discovery, macroeconomic announcements, and liquidity

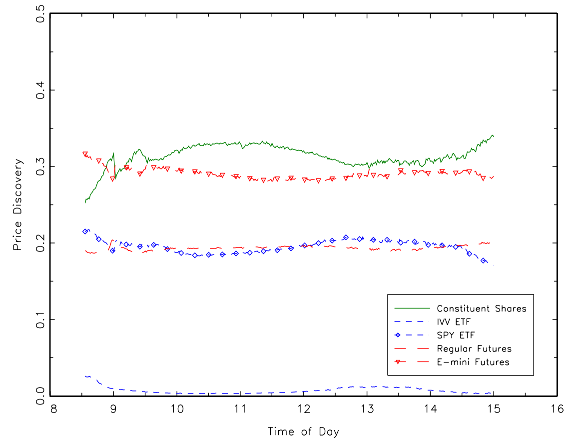
This table contains estimates of the following event-specific conditional mean measures of price discovery: the augmented Harris et al (2002) component share measure ($\widehat{\mathbf{f}}_{1t}$); the augmented Hasbrouck (1995) information share measure ($\widehat{\mathbf{f}}_{2t}$); the augmented de Jong and Schotman (2005) information share measure ($\widehat{\mathbf{f}}_{3t}$); the augmented Yan and Zivot (2006a) standardized price discovery measure ($-s\widehat{\mathbf{f}}_{4K,t}$); and the augmented Yan and Zivot (2006a) aggregated price discovery measure ($-a\ln\widehat{\mathrm{f}}_{4K,t}$) – all of which are based on the scaled version of the equilibrium correction model given by equation (1). The two pair of events considered in this table are the before and after macroeconomic announcement periods, and low and high liquidity, where the before (after) announcement period is defined as the fifteen minute period prior to (after) the announcement, and low (high) liquidity is defined as all liquidity amounts below (above) the median level of liquidity in the E-mini futures market. Differences between the before and after announcement conditional mean measures of price discovery for each level of liquidity are denoted by $\Delta(\widehat{\mathbf{f}}_{1t})$, $\Delta(\widehat{\mathbf{f}}_{2t})$, $\Delta(\widehat{\mathbf{f}}_{3t})$, $-\Delta(s\widehat{\mathbf{f}}_{4K,t})$, and $-\Delta(a\ln\widehat{\mathrm{f}}_{4K,t})$. All of these measures are calculated for all five (M) S&P 500 constituent markets; viz., (1) the index market, (2) the SPDR ETF share market, (3) the iShares ETF share market, (4) the regular futures market, and (5) the E-mini futures market. Inference is based on a 1,000 repetition bootstrap procedure with significance denoted in the usual way.

		Constituent Market					Constituent Market				
Discovery measure	Event(s)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
		Event(s)									
		Low Liquidity ($n = 271$)					High Liquidity ($n = 374$)				
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	Before (B)	0.1672	-0.1402	-0.2458	0.3568	0.1380	0.3884***	-0.1769	-0.2225	-0.0949	0.1058
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$		0.1254***	-0.1922	-0.0158	0.0179	0.0646***	0.1195***	-0.1918	-0.0127	0.0081	0.0769***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$		0.3701***	-0.1941	-0.1870	-0.1442	0.1552***	0.3450***	-0.1930	-0.1855	-0.2019	0.2354***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$		0.0912***	0.0752***	-0.0575	-0.0767	0.0322***	0.0952***	0.0518***	-0.0582	-0.0949	0.0061***
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	0.3074	0.4130	-0.1897	0.3052**	0.0000	0.5653***	0.6109***	0.4834	0.2826***
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	0.3175***	0.1412***	0.1074***	0.0608***	0.0000	0.3113***	0.1323***	0.1114***	0.0426***
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	0.5643***	0.5571***	0.5143***	0.2149***	0.0000	0.5381***	0.5306***	0.5469***	0.1097*
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	0.0160***	0.1487***	0.1680***	0.1234***	0.0000	0.0434***	0.1534***	0.1902***	0.0892***
		Low Liquidity ($n = 110$)					High Liquidity ($n = 535$)				
$\{\widehat{\mathbf{f}}_{1t}\}_i - 1/M$	After (A)	0.2762***	-0.1853	-0.1551	-0.1140	0.1782***	0.2798***	-0.2136	-0.1725	-0.2044	0.3107***
$\{\widehat{\mathbf{f}}_{2t}\}_i - 1/M$		0.0836***	-0.1930	0.0071	0.0223***	0.0800***	0.0771***	-0.1919	0.0009	0.0046***	0.1092***
$\{\widehat{\mathbf{f}}_{3t}\}_i - 1/M$		0.1964***	-0.1948	-0.1378	-0.0890	0.2252***	0.1933***	-0.1958	-0.1580	-0.2027	0.3633***
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_i - 1/M)$		0.0540***	0.0434***	-0.0395	-0.0416	-0.0164	0.0792***	-0.0033	-0.0369	-0.0732	0.0342***
$\{\widehat{\mathbf{f}}_{1t}\}_1 - \{\widehat{\mathbf{f}}_{1t}\}_i$		0.0000	0.4615***	0.4312***	0.3902***	0.0980**	0.0000	0.4934***	0.4524***	0.4843***	-0.0308
$\{\widehat{\mathbf{f}}_{2t}\}_1 - \{\widehat{\mathbf{f}}_{2t}\}_i$		0.0000	0.2766***	0.0764***	0.0613***	0.0035	0.0000	0.2691***	0.0762***	0.0725***	-0.0321^{***}
$\{\widehat{\mathbf{f}}_{3t}\}_1 - \{\widehat{\mathbf{f}}_{3t}\}_i$		0.0000	0.3911***	0.3341***	0.2854***	-0.0289	0.0000	0.3891***	0.3513***	0.3960***	-0.1700^{***}
$-(\{s\widehat{\mathbf{f}}_{4K,t}\}_1 - \{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0000	0.0106	0.0934***	0.0956***	0.0704***	0.0000	0.0825	0.1161***	0.1524***	0.0450***
		Low Liquidity ($n = 381$)					High Liquidity ($n = 909$)				
$\Delta(\{\widehat{\mathbf{f}}_{1t}\}_i)$	B $v.$ A	-0.1090	0.0451	-0.0908^{*}	0.4709	-0.3161	0.1086	0.0367**	-0.0500^{*}	0.1095	-0.2048^{**}
$\Delta(\{\widehat{\mathbf{f}}_{2t}\}_i)$		0.0418***	0.0009	-0.0229^{**}	-0.0043	-0.0154^{*}	0.0424***	0.0001	-0.0136^{***}	0.0035	-0.0323^{***}
$\Delta(\{\widehat{\mathbf{f}}_{3t}\}_i)$		0.1738***	0.0006	-0.0492	-0.0551	-0.0700	0.1518***	0.0028***	-0.0275	0.0008	-0.1279^{***}
$-\Delta(\{s\widehat{\mathbf{f}}_{4K,t}\}_i)$		0.0081	-0.0448^{***}	0.0097	0.0082	0.0188*	-0.0159^{***}	0.0075	0.0069***	0.0316***	-0.0151^{***}
$-\Delta(a\ln\mathrm{f}_{4K,t})$				-0.6215					-1.1188		

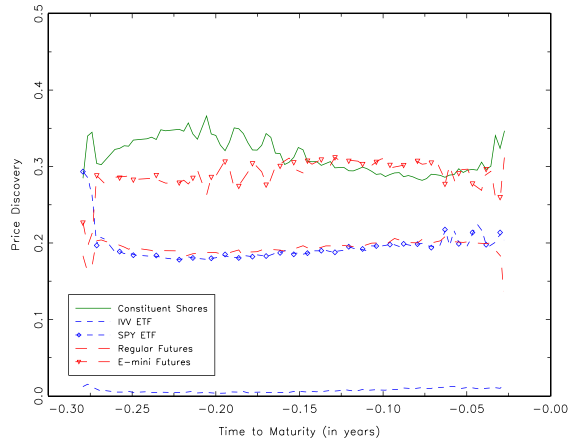
Figure 1: Relative price discovery in factor space.

The figure contains the augmented Hasbrouck (1995) information share measure of relative price discovery ($\widehat{\mathbf{f}}_{2t}$) plotted against the intraday, interday, liquidity and pre-10.00 a.m. macroeconomic announcement scaling factor spaces.

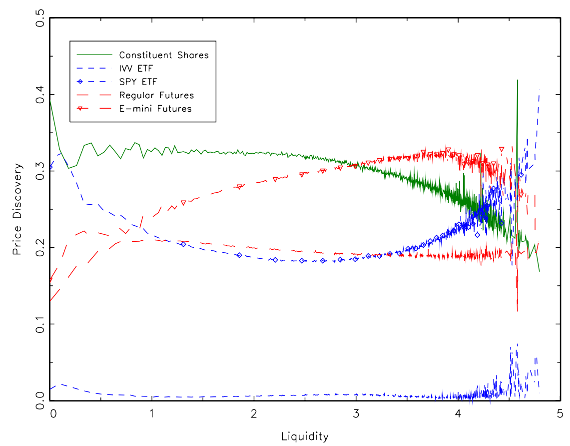
(a) Intraday variation



(b) Interday variation



(c) Liquidity variation



(d) Announcement variation

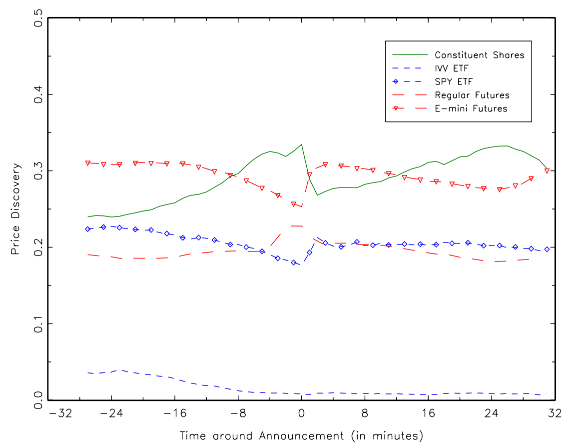


Figure 2: Aggregate price discovery around macroeconomic announcements.
The figure contains the augmented Yan and Zivot (2006a) aggregated price discovery measure ($-a\ln\widehat{\mathrm{f}}_{4K,t}$) plotted against the pre-10.00 a.m. macroeconomic announcement scaling factor space.