

TI 2014-027/III
Tinbergen Institute Discussion Paper



Intraday Price Discovery in Fragmented Markets

Sait Ozturk

Michel van der Wel

Dick van Dijk

Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Intraday Price Discovery in Fragmented Markets

Sait Ozturk^(a,c) *Michel van der Wel*^(a,b,c,d) *Dick van Dijk*^(a,b,c)

February 25, 2014

(For the web appendix click [\[here\]](#).)

^(a) Econometric Institute, Erasmus School of Economics

^(b) Erasmus Research Institute of Management

^(c) Tinbergen Institute

^(d) CREATES, Aarhus

Keywords: High-frequency data, Market microstructure, Price Discovery, Kalman filter.

JEL classifications: C32, G14.

Address for correspondence: Sait Ozturk, Econometric Institute, Erasmus University Rotterdam, P.O.Box 1738, 3000 DR Rotterdam, The Netherlands.

Emails: `s.r.ozturk@ese.eur.nl` `vanderwel@ese.eur.nl` `djvandijk@ese.eur.nl`

We thank Peter Schotman for his helpful suggestions and participants at the PhD seminar in Rotterdam (May, 2013), the 6th Annual Conference of the Society for Financial Econometrics in Singapore (June, 2013), the 19th International Conference of Computing in Economics and Finance in Vancouver (July, 2013), and the Joint Congress of the European Economic Association and the Econometric Society in Gothenburg (August, 2013) for their comments. Michel van der Wel is grateful to Netherlands Organisation for Scientific Research (NWO) for a Veni grant; and acknowledges support from CREATES, funded by the Danish National Research Foundation.

Intraday Price Discovery in Fragmented Markets

Sait Ozturk Michel van der Wel Dick van Dijk

ABSTRACT

For many assets, trading is fragmented across multiple exchanges. Price discovery measures summarize the informativeness of trading on each venue for discovering the asset's true underlying value. We explore intraday variation in price discovery using a structural model with time-varying parameters that can be estimated with state space techniques. An application to the Expedia stock demonstrates intraday variation, to the extent that the overall dominant trading venue (NASDAQ) does not lead the entire day. Spreads, the number of trades and volatility can explain almost half of the intraday variation in information shares.

1 Introduction

Financial markets incorporate new information into asset prices by matching buyers and sellers. They thereby facilitate the discovery of what the price of an asset should be. Nowadays this “price discovery” role of financial markets can take place across separate exchanges and instruments, as many securities and derivatives based on the same underlying asset may trade on multiple venues. In the case of such a multiplicity, there may be variation in the share with which each market’s trades contribute to discovering the one true price of the underlying asset. Knowledge of these so-called information shares of different markets would benefit both investors concerned with price informativeness and adverse selection risk as well as policy makers investigating the determinants of price efficiency. Existing studies often assume the contributions of different markets to price discovery are constant at least over the course of the day. We analyze intraday variation in price discovery, and consider which factors may explain such variation.

The measurement of price discovery requires isolating informative price movements from noise. Observed price changes constitute the most obvious indicator of price discovery. However, they form an imperfect measure as observed prices are susceptible to transitory mispricing, caused by noise trading or temporary order imbalances, for example. In contrast, when security prices absorb new information due to informed trading, these price changes last permanently. Hasbrouck (1995) demonstrates that the above implies the existence of co-integration relationships between securities prices and develops a framework exploiting these to distinguish permanent and transitory price changes. His work initiated a booming literature on price discovery measures and information shares.

Early studies, like Hasbrouck (1995), effectively assume the contributions of different trading venues to the efficient price innovations to be constant over time, or at least for the sample period used for estimation. However, the pace of change in the characteristics of exchanges and securities - such as increases in trade volume and electronization of trading mechanisms - makes this assumption implausible. Based on these motivations, the

more recent literature mostly divides the sample into short sub-periods, and typically considers daily measurements of information shares (Chakravarty, Gulen and Mayhew, 2004; Hasbrouck, 2003; Mizrach and Neely, 2008, among others).

In spite of providing a higher level of sophistication, measurements of information shares at the daily frequency are unable to keep up with the current pace of financial markets and available data. Current information share methodologies are not able to answer questions about differences in price discovery across different parts of the day or the digestion of public news, most of which happens in a matter of minutes, if not seconds. A growing body of studies infer intraday variation in informed trading indirectly from the dynamics in other market characteristics, such as liquidity, depth and volatility in limit order markets (Ahn, Bae and Chan, 2001), asymmetric information proxies and trade volume before and after public announcements (Chae, 2005) or predictions of a model with informed and uninformed traders (Lei and Wu, 2005).

In this paper we consider the possibility of examining intraday variation in information shares directly. We propose a novel method to capture the intraday dynamics of price discovery based on the structural time series model of Hasbrouck (1995). In this structural model, the observed price series depend on a single underlying latent true price. Differences between the observed prices and the latent price consist of two components. On the one hand, these pricing errors are linked to the innovations to the latent true price capturing lagged adjustment or over-reaction to information. On the other hand, they stem from uncorrelated errors representing dynamics like noise trading. Following De Jong and Schotman (2010), information shares in this model can be expressed as a function of the structural model parameters, including the variances of the latent price innovations and the uncorrelated errors. We construct a state space model, in which the innovation and noise variances vary throughout the trading day in the form of flexible Fourier functions. This naturally leads to time-variation in the information shares, thus allowing us to capture intraday variation in the relative contributions of different trading venues to price discovery.

We examine the usefulness of our modeling approach by means of a simulation study and an empirical application. The simulation evidence demonstrates that our state space method generates accurate estimates for a wide range of settings with varying number of observations, securities and parameters of the time-varying information share model. As the range of tested parameters includes those obtained in our empirical study, the simulation results also support the validity of our approach in empirical research.

Our empirical study presents evidence for intraday variation in informed trading. We use Expedia quote data for the period July-December 2007, analyzed earlier by De Jong and Schotman (2010) in the case of constant information shares. The Expedia stock trades on a number of exchanges, including NASDAQ and the NYSE. We gather the exchanges on which the stock trades in three groups: NASDAQ, NYSE and “REST”. The NASDAQ system and the NYSE provide overall the largest contributions to price discovery with average information shares of 61.9% and 27.7%, respectively. We find that the intraday patterns of information shares display strong fluctuations. The contribution to price discovery of the NYSE reaches as high as 70% (and thus overtakes NASDAQ for part of the day), while the share of NASDAQ drops as low as 20%. The estimated latent price innovation variance follows a U-shaped intraday pattern, which is consistent with the literature (see, e.g., Admati and Pfleiderer, 1988; Foster and Viswanathan, 1993; Slezak, 1994): The average innovation variance in the first half hour is about five times larger than the rest of the day.

Lastly, we seek to explain why information shares vary so much throughout the trading day. Using a market share attraction model, we find that the number of trades, spreads and volatility have significant explanatory power for the dynamics of the relative information shares. These standard market quality measures explain 49.4% of the intraday variation. Information shares have largest elasticities for quoted spreads and the number of trades of 100 stocks, the smallest possible trade size. A 1% decrease in quoted spreads leads to 1.9% and 2.3% increases in the information shares of the NYSE and the REST groups, respectively. Similarly, a 1% decrease in the number of trades with size 100 raises the shares

of the NYSE and the NASDAQ by 0.9% and 1.5%, respectively. The amplification of a venue’s relative information share with larger trades conforms with the informed trading literature (e.g., Hasbrouck, 1991; Madhavan and Smidt, 1991; Easley et al., 1997). However, we do not observe a general monotonic relationship between trade sizes and informativeness as the most informative trades turn out to be mid-sized trades (from 200 to 500 stocks) in NASDAQ and large trades (above 500 stocks) in the NYSE.

Our work is related to a number of studies investigating price discovery via state space methods. Upper and Werner (2007) estimate a reduced-form VECM representation in the state space framework, while Frijns and Schotman (2009) and Korenok, Mizrach and Radchenko (2011) use directly the structural model of Hasbrouck in state space form, albeit not allowing for intraday variation in information shares. A closely related paper is Menkveld, Koopman and Lucas (2007), who suggest a similar structural model in state space form that allows for time-variation in parameters throughout the day. Our set-up differs in three important respects. First, in their case the comparison is for overall variation in prices throughout the day for all markets an asset trades on, and not for price discovery across markets. A result is that they study variance ratios for different parts of the trading day (a time series aspect), and not price discovery measures across the various exchanges (a cross-sectional aspect) as we do. Second, their model is designed for lower intraday frequencies such as an hour, as they assume that the innovation in the latent efficient price is fully incorporated into the observed prices at each period (which is not plausible for higher intraday frequencies). Third, we study the higher-frequency change in structural model parameters using flexible Fourier functions, while they focus on step functions to model time-variation.

The information share methodology of De Jong and Schotman (2010) that we use has several advantages over other measures in the literature. Hasbrouck (1995) estimates the contributions of each security to the variance of innovations in the latent price. Comparative studies, such as De Jong (2002) and Lehman (2002), find this focus on variance more appropriate for price discovery measurement than the common factor decomposition of Gonzalo

and Granger (1995), as Harris, McNish and Wood (1997; 2002) do. The proposal of De Jong and Schotman (2010) similarly works at the variance level and resolves two main concerns about the Hasbrouck approach. Firstly, Hasbrouck information shares are not unique but they come in the form of a range, often with a substantial difference between the upper and lower boundaries.¹ Secondly, it relies on a reduced form estimation methodology which does not provide estimates of structural parameters.

The remainder of the paper is organized as follows. Section 2 introduces the Unobserved Components Model of Hasbrouck (1995) and De Jong and Schotman (2010) information shares, followed by our extension to capture intraday variation. Section 3 provides simulation evidence for our methodology. Section 4 reports the empirical results, including the analysis of the determinants of the estimated intraday variation in information shares. Section 5 concludes.

2 Measuring price discovery

This section presents the methodology to measure the contributions of different securities (or trading venues) to price discovery. Its three parts elaborate on the structural model of Hasbrouck (1995), the information shares suggested by De Jong and Schotman (2010), and our novel implementation of intraday time-variation under the state space framework, respectively.

2.1 The Unobserved Components Model

The structural model introduced by Hasbrouck (1995) lays the foundations for much of the present price discovery literature. In this framework, all observed security prices based on the same underlying asset (such as the observed prices on multiple exchanges of the same stock) are driven by one latent efficient price process (the unknown true price of that underlying stock). This latent price is defined as the end-of-period value of the

¹Grammig and Peter (2013) provide identification restrictions using the distributional properties of financial price series to overcome the non-uniqueness problem.

asset conditional on all publicly available information at time t . Thus this price process satisfies the semi-strong form of market efficiency in line with the range of information it encompasses (Fama, 1970). Since all public information is impounded in this latent price, the best prediction for the asset price in period $t + 1$ is the price at time t and therefore it is modeled as a random walk with stationary innovations r_t . The observed asset prices deviate from this latent price with a stationary error term as long-term or unbounded deviations are ruled out by arbitrage relationships. Hence the Unobserved Components Model defines the natural logarithm of the observed asset price $p_{i,t}$ of each security i as the sum of the logarithm of the latent price p_t^* and stationary disturbance terms $u_{i,t}$. In case of N observed prices, their relation with the latent price can be represented as

$$\begin{aligned} p_t &= \iota p_t^* + u_t, \\ p_t^* &= p_{t-1}^* + r_t, \end{aligned} \tag{1}$$

where p_t is an $N \times 1$ vector of log observed prices $p_{i,t}$, u_t is an $N \times 1$ vector of stationary disturbance terms $u_{i,t}$, p_t^* is the scalar latent efficient price, r_t is the innovation in the latent price with mean zero and variance σ_r^2 and ι is an $N \times 1$ vector of ones.

The error terms in u_t capture microstructure effects in the observed prices. It comprises two components distinguished by their correlation with the efficient price innovation r_t . First, $u_{i,t}$ has an information-correlated pricing error component $\alpha_i r_t$ that captures dynamics such as adverse selection. The second error component $e_{i,t}$ is uncorrelated with information, but stems from factors such as noise trading or price discreteness. This idiosyncratic noise $e_{i,t}$ has mean zero and covariance matrix Ω , allowing for correlation in this noise component across observed prices. With these two components, the specification for the disturbance terms $u_{i,t}$ is:

$$u_t = \alpha r_t + e_t + \Psi e_{t-1} \tag{2}$$

where α is an $N \times 1$ vector of α_i 's, e_t is an $N \times 1$ vector of idiosyncratic noise $e_{i,t}$ with the

$N \times N$ covariance matrix Ω , and Ψ is an $N \times N$ coefficients matrix. De Jong and Schotman (2010) propose the inclusion of the lagged noise e_{t-1} in the observed price dynamics in order to capture serial correlation in high-frequency intraday returns. From the definition of the disturbance term $u_{i,t}$ in (2), it follows that its covariance with the innovation in the efficient price is equal to

$$Cov(u_{i,t}, r_t) = \alpha_i \sigma_r^2.$$

We provide a state space representation of the Unobserved Components Model in the Appendix. The state space system is estimated by Maximum Likelihood using the Kalman Filter. As the latent price p_t^* follows a random walk and to account for over-night price changes, we re-initialize p_t^* every day with a diffuse prior and exclude a number of initial observations from the likelihood maximization as these may be unreliable due to the initial convergence of the Kalman filter.²

2.2 De Jong-Schotman information shares

De Jong and Schotman (2010) propose a price discovery measure quantifying the explanatory power of changes in each of the observed security prices for the innovations in the latent price. The explanatory power of observed price changes is obtained by means of a regression framework. For this purpose, the total price innovation in period t is defined as

$$\nu_t = p_t - \iota p_{t-1}^* = (\iota + \alpha)r_t + e_t + \Psi e_{t-1}. \quad (3)$$

We may then consider the regression of the innovation in the latent price on the total innovations in individual prices, that is

$$r_t = \gamma' \nu_t + \eta_t, \quad (4)$$

²We exclude the first 11 observations of all days in our state space estimation from the likelihood calculation. The filter already shows signs of convergence at a smaller number of observations. However in Section 3.3 we use a step function with 10 steps as a benchmark and leaving 11 of 391 minutely observations of the trading day gives 380, which is a multiple of 10.

where η_t is the innovation in the latent price unrelated to innovations in market prices.

The regression coefficient γ is given by

$$\gamma = \frac{\text{cov}(r_t, \nu_t)}{\text{var}(\nu_t)} = \Upsilon^{-1}(\iota + \alpha)\sigma_r^2. \quad (5)$$

where $\text{cov}(r_t, \nu_t) = (\iota + \alpha)\sigma_r^2$ follows from (3) and Υ denotes the covariance matrix of ν_t .

From (3) we also have

$$\Upsilon = \sigma_r^2(\iota + \alpha)(\iota + \alpha)' + \Omega + \Psi\Omega\Psi', \quad (6)$$

Using (5), the goodness-of-fit of the regression in (4) can be expressed as

$$R^2 = 1 - \frac{\sigma_\eta^2}{\sigma_r^2} = \frac{\gamma'\Upsilon\gamma}{\sigma_r^2} = \gamma'(\iota + \alpha) = \sum_{i=1}^N \gamma_i(1 + \alpha_i).$$

This leads De Jong and Schotman (2010) to propose an information share for the i -th observed price, denoted IS_i , with a partial R^2 interpretation, namely

$$IS_i = \gamma_i(1 + \alpha_i). \quad (7)$$

The sum of these information shares, i.e. the R^2 of the regression, is not necessarily equal to one.

Computing the information shares IS_i according to (7) obviously requires estimates of the parameters in the Unobserved Components Model in Equations (1) and (2). De Jong and Schotman (2010) present a GMM approach to obtain these. The auto-covariances of the observed returns provide the following moment conditions:

$$\Gamma_0 = E[\Delta p_t \Delta p_t'] = \sigma_r^2((\iota + \alpha)(\iota + \alpha)' + \alpha\alpha') + \Omega + (\Psi - I)\Omega(\Psi - I)' + \Psi\Omega\Psi', \quad (8)$$

$$\Gamma_1 = E[\Delta p_t \Delta p_{t-1}'] = -\sigma_r^2\alpha(\iota + \alpha)' + (\Psi - I)\Omega - \Psi\Omega(\Psi - I), \quad (9)$$

$$\Gamma_2 = E[\Delta p_t \Delta p'_{t-2}] = -\Psi\Omega, \quad (10)$$

where $\Delta p_t = p_t - p_{t-1}$. These conditions identify the parameters σ_r^2 , Ω and α , but not γ . Instead γ can be computed using Equations (5) and (6).

The information shares IS_i defined in (7) improve on Hasbrouck's approach by providing unique measures of price discovery, while keeping a focus on the variance of the latent innovations. Hasbrouck (1995) estimates the reduced form of the Unobserved Components Model as a vector error correction model and uses Choleski decomposition to obtain the contribution of each security to the variance of innovations in the latent price. Because the Choleski decomposition depends on the ordering of the series, Hasbrouck's information shares come in the form of a range between certain lower and upper bounds. These bounds tend to be wide, unless the contemporaneous correlations between securities are small. In addition, the parsimony of the structural model compared to the reduced form eases statistical inference.

Following De Jong and Schotman (2010), we model Ω and Ψ as diagonal matrices for parsimony. These two diagonality assumptions are both plausible and testable. The diagonality of Ω means that the idiosyncratic noise components of the price changes in different markets are uncorrelated. A diagonal Ψ matrix implies that the mispricing in one exchange is not influenced by the previous period's noise in another exchange. In comparison, the unique identification of Hasbrouck information shares requires a far stronger assumption like the diagonality of the residual covariance matrix. This essentially states that the shocks to the prices in the reduced form system are uncorrelated, which is violated in any empirical application as De Jong and Schotman (2010) point out. The diagonality assumptions of Ω and Ψ are much weaker, and the GMM framework offers tests to evaluate their validity.

2.3 Intraday variation in information shares

Time-variation in the information shares IS_i in (7) can be introduced by considering a time-varying parameter extension of the Unobserved Components Model as given by Equations (1) and (2). This can be attained by making at least one of the parameter groups

vary over time, namely α , Ψ , σ_r^2 or Ω . The latter two variance terms have the advantage of an established literature linking intraday volatility changes to changes in informed trading. Intraday volatility is documented to follow an inverted J-shape or a U-shape pattern during trading hours (Wood, McInish and Ord, 1985; Lockwood and Linn, 1990). On the one hand, a number of asymmetric information models noted this pattern as an empirical prediction for markets with informed and uninformed traders (Admati and Pfleiderer, 1988; Foster and Viswanathan, 1993; Slezak, 1994). On the other hand, Hsieh and Kleidon (1996) document several dynamics unrelated to informed trading which aid to the formation of this intraday volatility pattern. The main area of contention lies on whether the start and the end of the day have higher levels of information absorption into prices and if this is accompanied with higher or lower amounts of noise.

Given these theoretical and empirical claims for the intraday variation of informed and noise trading, a natural way to model intraday variation in price discovery is making both the innovation and the noise variances time-varying.³ We implement time-variation using a combination of flexible Fourier trigonometric functions and a polynomial (Andersen et al., 2001; Gallant, 1981). The variance entries ζ_t^2 have the form

$$\zeta_t^2 = \exp \left(c + \sum_{p=1}^P \theta_p t^p + \sum_{q=1}^Q \left(\delta_q \cos \left(\frac{2\pi qt}{T} \right) + \phi_q \sin \left(\frac{2\pi qt}{T} \right) \right) \right), \quad (11)$$

where ζ_t^2 represents the processes of σ_r^2 and ω_i^2 's, i.e. the diagonal entries of the Ω matrix, $t = 1, \dots, T$, with T being the number of observations per day, P the order of the polynomial part, and Q the total number of flexible Fourier sets. We use an exponential specification for the variances to facilitate an unconstrained maximization procedure given that trigonometric functions can have negative values. The flexible Fourier form can model complex dynamics

³In a separate analysis, available upon request, we also model the elements of α and Ψ as a time-varying process besides the variances. The α estimates fluctuate very mildly around a constant and the resulting estimates have higher (i.e. worse) information criteria values than the constant α case. Information criteria results prefer a mild time-variation in Ψ , but this does not cause a considerable change in the information share estimates.

and smooth transitions. However using solely the flexible Fourier part would impose equality of the variances at the start and end of the day. This is avoided by complementing it with the polynomial component. We select P and Q using the Schwarz Information Criteria.

The flexible Fourier specification has several advantages over alternative specifications for capturing time-variation in parameters. A first, simpler, alternative would be to use step functions. A disadvantage of the step function approach is that it generates unlikely jumps between consecutive time periods. Moreover, it introduces the challenge of choosing the number of periods and optimizing period lengths, because assuming them to be equal in length may be too restrictive. A second alternative is to use spline functions instead of the flexible Fourier. Also here a challenge is that of finding the right number of knots for the spline and the precise knot locations.

In this time-varying setting, we can evaluate the effect of changes in the innovation and noise variances on the information share by rewriting the regression coefficient γ in (5) as

$$\gamma_t = \left((\iota + \alpha)(\iota + \alpha)' + \frac{\Omega_t + \Psi\Omega_t\Psi'}{\sigma_{r,t}^2} \right)^{-1} (\iota + \alpha). \quad (12)$$

This expression shows that the information share is shaped by a time-varying noise-to-innovation ratio. An increase in the innovation variance σ_r^2 boosts all information shares $IS_{i,t}$. Therefore both individual information shares and the total explanatory power of observed securities are amplified. By contrast, an increase in a noise variance $\omega_{i,t}^2$ for asset price i reduces the corresponding information share as well as the total explanatory power of all observed prices.

It is useful to note that the specification in (11) assumes the intraday pattern to be constant across the days under consideration. Therefore it suits best to the investigation of price discovery during recurring events like financial announcements. As such days may not be abundant, in the simulation study of Section 3.2 we reduce the number of days down to 10 to show that our methodology remains accurate even under such a data scarcity.

Finally, introducing time-variation in the error variances as in (11) obviously implies that we can no longer use the GMM approach of De Jong and Schotman (2010) to estimate the model parameters. The model, however, still keeps its state space representation, albeit with time-varying variances, and as such we can obtain parameter estimates by means of maximum likelihood combined with the Kalman filter.

3 Simulation study

In this section we provide simulation evidence for the ability of the proposed modeling framework to capture intraday variation in price discovery. Section 3.1 compares GMM and state space (i.e. maximum likelihood) results for the case without time-variation. In Section 3.2 we generate data from a model with time-varying parameters (and thus time-varying information shares) and examine to what extent our model is able to detect such time-variation. Lastly, Section 3.3 explores various parameter configurations and the case where the data generating process (DGP) differs from the model that is actually estimated.

3.1 Comparison of GMM and state space methods

We design our simulations and choose parameter values in the DGP to mimic an empirical setting in order to demonstrate the relevance of our results for empirical work. As a benchmark case, we simulate observed prices of three securities and a latent price process over 100 days with 391 intraday observations using Equations (1) and (2). This corresponds to data sampled at a 1-minute frequency for a trading day between 9:30h and 16:00h. We take the noise covariance matrix Ω and the matrix of lagged noise coefficients Ψ as diagonal. The innovation variance σ_r^2 is set to 0.816, while the noise variances in Ω take considerably smaller values of 0.016, 0.012 and 0.107. The elements of the correlation vector α have small negative magnitudes of -0.008 , -0.022 and -0.006 , such that the efficient price innovations are almost but not fully incorporated into the observed prices in each period. Lastly, the diagonal elements of the Ψ matrix are set to 0.172, 0.087, and 0.270, implying a modest

degree of autocorrelation in observed price changes. We generate 100 replications of these three observed price series and estimate them in both methods using the true parameters as initial values.

Panel A of Table 1 compares the true information shares and the estimates obtained with both GMM and the state space methods. The parameter settings of the DGP imply that the second security leads price discovery with a 53.9% information share. This is followed by the first security with a share of 39.5%, while the third security is much less important with a 5.8% information share. The results show that on average both the GMM and the state space methods provide fairly accurate estimates of the information shares although the state space method performs quite a bit better. The mean estimates are close to the true values, with a maximum difference of 0.8% for the GMM and only 0.2% for the state space case. Likewise, the estimates do not show much variation across simulations, with the maximum standard deviation at 1.5% for GMM and 1% for the state space method. The same conclusion also follows from Panel B of Table 1, showing the average and standard deviations of the root mean squared error (RMSE) for the model parameters and the three information shares. While the average RMSEs are quite small for both methods, the state space approach shows superior performance with a mean RMSE of 0.7% compared to 1.2% for GMM.

3.2 Capturing time-variation with the state space method

We now advance to testing our state space approach in the measurement of intraday variation in price discovery. Following Section 2.3, we allow for variation in the innovation and noise variances with the same pattern repeating each day. As before series for three securities are simulated for 100 days with 391 intraday observations to correspond again to data sampled at 1-minute frequency spanning from 9:30h to 16:00h. The variances now fluctuate following a flexible Fourier form complemented with a polynomial function, as given in Equation (11). In the benchmark DGP, each variance specification consists of 10

Table 1. Simulation Results of the GMM and State Space Methods in the Constant Case Benchmark

The table shows summary statistics of the simulation results for the GMM and state space methods with constant innovation and noise variances. Three stock series are generated for 100 days, each with 391 observations, using the Unobserved Components Model of equations (1) and (2). Panel A reports the summary statistics for each information share. The first column denotes the information shares for each of the simulated stocks calculated using the data generating process (DGP) parameter values. For each of the information share estimates the mean and standard deviations over all simulations are given. The results are based on 100 simulations from the corresponding data generating process. Panel B provides a more concise summary of the information share results and also provides information for the parameter estimates. The presented data consists of the means (RMSE) and of the standard deviations (SD) of root mean squared errors of the parameter estimates and the information shares. The results are based on 100 simulations from the corresponding data generating process.

Panel A: Summary Statistics for Information Share Estimates

	<i>DGP</i>	<i>GMM</i>		<i>State Space</i>	
		<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
<i>IS1</i>	39.5	38.7	1.5	39.3	0.9
<i>IS2</i>	53.9	54.6	1.5	54.1	1.0
<i>IS3</i>	5.8	5.9	0.1	5.8	0.1

Panel B: General Summary Statistics

	<i>Parameters</i>		<i>Information Shares</i>	
	<i>RMSE</i>	<i>SD</i>	<i>RMSE</i>	<i>SD</i>
<i>State Space</i>	1.7	0.4	0.7	0.5
<i>GMM</i>	2.2	0.7	1.2	0.9

flexible Fourier sets and a polynomial of order 1. For brevity, we do not report all parameter settings of the polynomials and flexible Fourier sets, which have 94 parameters in total.⁴ The mean of the innovation variance process $\sigma_{r,t}^2$ is 0.810 and the mean of noise variances in Ω_t have smaller values of 0.019, 0.009 and 0.103. We take the noise covariance matrix Ω_t and the matrix of lagged noise coefficients Ψ as diagonal like in the constant case. The diagonal elements of the Ψ matrix are 0.142, 0.122, and 0.210. Finally, the elements of the correlation vector α again have typical small magnitudes of -0.01 , -0.02 and -0.005 .⁵ As in the constant case, the parameter configuration and the aforementioned data properties of the benchmark case mirror the data and results of the empirical study in Section 4. We now consider 25 replications, due to the additional computation burden of the time-varying

⁴A full list of all parameters is available on Table 1 in our web appendix.

⁵In Table 2 of our web appendix, we also tested our method across a range of alternative parameter values.

system with a large number of parameters and the great amount of variations in settings we consider.

Figure 1 displays the true intraday information shares as implied by the parameter settings in the DGP (solid line), as well as the average estimates (dashed line), and minimum and maximum estimates (thin solid lines) across the 25 replications. The mean estimates are close to the true information shares throughout the entire day. Subtracting the mean of estimates from the true information share values at each time point and averaging the absolute values of these differences, we find a minuscule mean absolute difference of 0.2%. The mean absolute difference of the lowest and highest estimates from the true values is also modest at 2.8%.

We evaluate a number of variations in the DGP settings, with results shown in Table 2. Specifically, we consider varying the number of days in the sample, varying the number of observations per day (the observation frequency), the number of series, the number of flexible Fourier (FF) sets, and the polynomial order. We mainly focus on lowering the number of available observations in terms of the number of days and intraday observations, because this is the direction where the results tend to worsen. Also an intraday pattern can be just temporary and we would like to capture it from as little observations as possible. In terms of the variance specifications we mostly investigate cases with more flexible Fourier sets and higher polynomial degrees, since this shows if the estimation procedure can handle a large number of parameters. The number of series under consideration reflects the usual amount of asset/exchange groups used in the literature. As in Panel B of Table 1, we present means and standard deviations of RMSEs for the parameter estimates and the information share estimates.

First consider the RMSE results of the benchmark case, corresponding to the information shares of Figure 1, to provide a context to evaluate the variations in Table 2. The information shares have a mean RMSE of 1.3% with a standard deviation of 1.0%. We observe an expected but limited decline of estimation accuracy compared to the constant case of Section

Fig. 1. Simulation Results from the Benchmark Case of the Time-Varying Model

The figure shows summary statistics of the simulation results for information shares achieved by the state space method with time-varying innovation and noise variances in the flexible Fourier form. Three stock series are generated for 100 days, each with 391 intraday observations. Each figure displays for the corresponding simulated security the true values, mean estimates and the upper and lower bounds containing all the estimates of information shares. The results are based on 25 simulations.

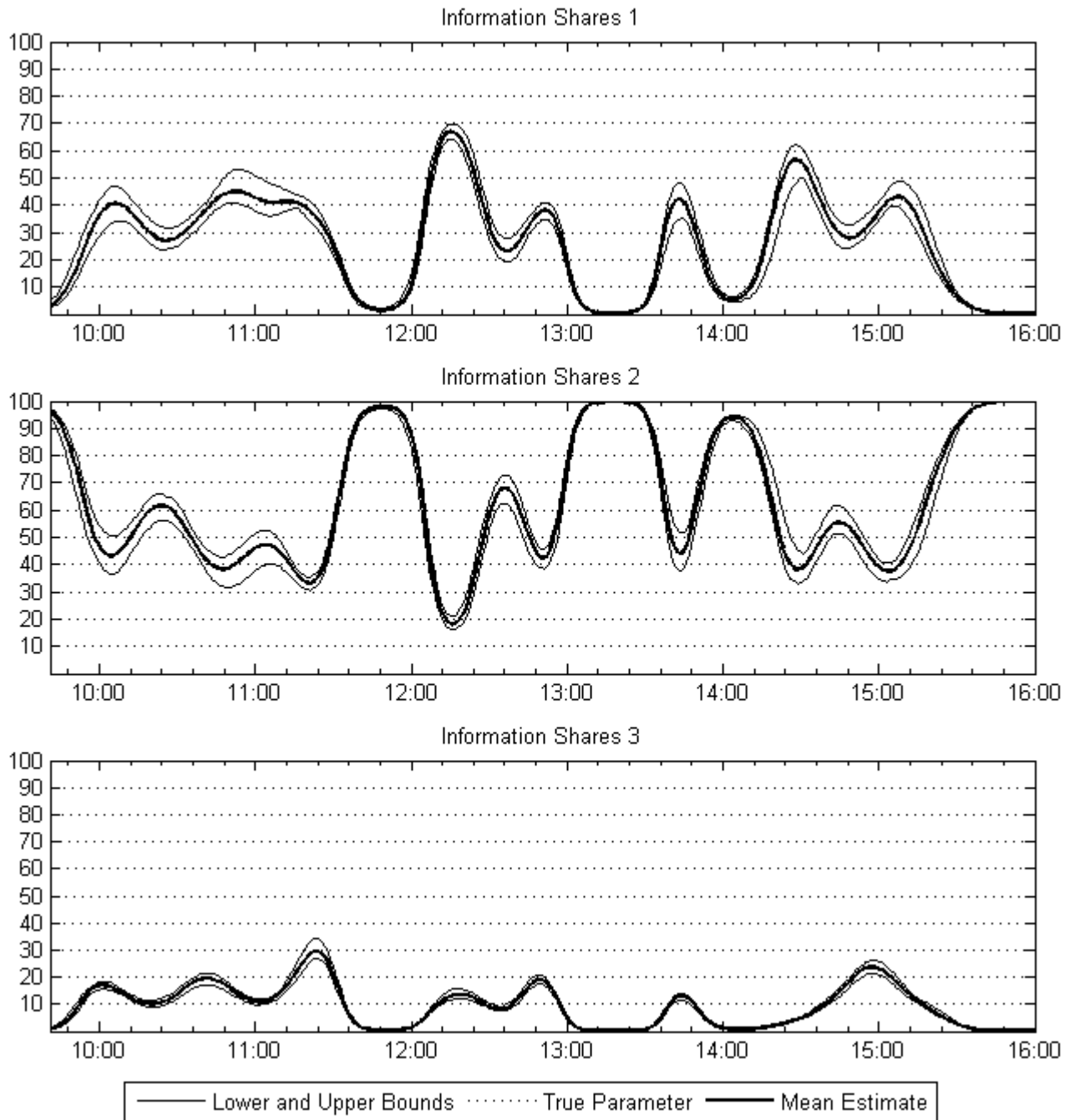


Table 2. Simulation Results of the Time-Varying Model

The table shows summary statistics of the simulation results from the state space method with time-varying innovation and noise variances. In the benchmark case, we consider a flexible Fourier type model on innovation and noise variances with 20 sets, a polynomial of order one, 3 securities, 100 days and with 391 observations. In the second column, the parameter setups corresponding to this benchmark case are emboldened. In the first column from top to bottom respectively the number of days, observations per period, securities, the flexible Fourier sets and the polynomial order are varied keeping others constant. The presented data consists of the means (RMSE) and of the standard deviations (SD) of root mean squared errors of the parameter estimates and the information shares. The results are based on 25 simulations from the corresponding data generating process.

		<i>Parameters</i>		<i>Information Shares</i>	
		<i>RMSE</i>	<i>SD</i>	<i>RMSE</i>	<i>SD</i>
<i>Days</i>	<i>10</i>	10.5	10.1	5.8	4.1
	<i>25</i>	4.6	3.7	3.1	2.2
	<i>50</i>	3.2	2.5	2.2	1.6
	100	1.5	0.6	1.3	1.0
<i>Intraday Obs.</i>	<i>71</i>	16.2	9.3	4.4	3.1
	<i>191</i>	3.0	2.0	2.4	1.7
	391	1.5	0.6	1.3	1.0
	<i>771</i>	1.2	0.8	0.9	0.6
<i>Series</i>	<i>2</i>	2.7	1.4	3.7	2.4
	3	1.5	0.6	1.3	1.0
	<i>4</i>	1.4	0.7	0.8	0.6
	<i>5</i>	1.5	0.9	0.6	0.5
<i>FF Sets</i>	<i>5</i>	1.7	0.9	0.9	0.6
	10	1.5	0.6	1.3	1.0
	<i>15</i>	2.0	1.7	1.4	1.2
	<i>20</i>	2.7	2.8	1.8	1.4
<i>Poly. Order</i>	1	1.5	0.6	1.3	1.0
	<i>2</i>	1.5	0.7	1.3	0.9
	<i>3</i>	1.9	1.0	1.6	1.5
	<i>4</i>	2.5	1.4	2.0	1.8

3.1, where the mean RMSE is 0.7% with a standard deviation of 0.5%.

Table 2 shows that a decrease in the amount of data has only a limited worsening effect on the information share estimates. Reducing the number of days from 100 to 10 increases the mean RMSE of the information shares from 1.3% to 5.8%. Likewise reducing the number of intraday observations from 771 to 71 raises the mean RMSE from 0.9% to 4.4%. These results suggest that our method can still effectively capture intraday patterns even with a limited amount of data. Similarly, estimation results improve with the number of observed price series. An increase from 2 to 5 series reduces the mean RMSE's from 3.7% to 0.6%.

Adding more flexible Fourier sets or increasing the polynomial order in the variance

specifications increases both the complexity of the pattern to be estimated and the estimation uncertainty, but this has only a weakly worsening effect on estimation accuracy. The increase of the polynomial order from 1 to 4 adds 12 more parameters, but the mean information share RMSE rises only from 1.3% to 2.0%. Similarly, the increase of the Fourier sets from 5 to 20 adds a far higher 120 parameters, yet the RMSE increase is only from 0.9% to 1.8%.

3.3 Capturing time-variation with a misspecified model

In this last part of our simulation study, we consider the effects of differences between specifications of the DGP and the model that is actually estimated. Firstly, we estimate DGP's with variances following a flexible Fourier form using models in the same form, but with correct and incorrect numbers of Fourier sets. Secondly, we introduce a state space model with a step function specification for the time-variation in variances and compare the estimation accuracy across models with and without time-varying variances.

Table 3 displays the mean RMSE's of the information shares under various cases where the DGP differs from the model. Panel A demonstrates that intraday variation in the error variances can be quite accurately captured as long as the number of Fourier sets in the model is at least as large as that of the DGP. We report nine setups with 5, 10 and 15 Fourier sets. For example, in the first row of the table we consider data generated using a flexible Fourier specification for the variances with 5 sets, and the columns represent the RMSE of the information shares when a model is estimated with 5, 10 and 15 flexible Fourier sets, respectively. The mean RMSE's are below 8% as long as the estimation model uses an equal or larger number of Fourier sets compared to the DGP. The model with 15 Fourier sets has low mean RMSE's of 2.8%, 3.0%, and 5.6% for DGP's with respectively 5, 10 and 15 sets.

Note that, in contrast to the previous sections we now use random parameter values instead of the true ones to initialize the numerical likelihood optimization, because no true initial values exist when the DGP and the model differ. In order to guard against the

Table 3. Simulation Results for Misspecified Models

The table shows summary statistics of the simulation results for cases where the data generating process (DGP) does not necessarily correspond to the estimation model. The settings of the DGP's are given in the leftmost column and those of the estimation models in the top row. In Panel A both the DGP and the estimation model are of flexible Fourier type, but have different numbers of Fourier sets. Panel B reports results for three DGP's, where variances are modeled as a constant, as a 20-period step function (SF), and as a Fourier model with 10 sets and a polynomial of order one (FF). These are estimated using the GMM method and three state space models, i.e. constant, step function with 20 periods (SF) and Fourier with 10 sets and a polynomial of order one (FF). In each case 25 data sets are generated from the DGP and for each data sets the estimation model is started for 10 trials with random initial values. All simulated data sets span 100 days with 391 intraday observations for 3 price series. The presented numbers are the means of root mean squared errors of the information shares.

Panel A: Flexible Fourier with different numbers of sets

		Estimated Model		
		<i>5</i>	<i>10</i>	<i>15</i>
DGP	<i>5</i>	1.4	3.5	2.8
	<i>10</i>	16.1	7.8	3.0
	<i>15</i>	26.9	22.0	5.6

Panel B: Estimations across models

		Estimated Model			
		<i>Constant</i>		<i>Time-Varying</i>	
		<i>GMM</i>	<i>State Space</i>	<i>SF</i>	<i>FF</i>
DGP	<i>Constant</i>	31.9	18.5	6.2	16.1
	<i>Time-Varying – SF</i>	58.7	21.5	2.7	3.2
	<i>Time-Varying – FF</i>	52.3	31.5	15.0	7.8

possibility of ending up in a local maximum of the likelihood function, we consider ten different sets of starting values for each replication.⁶ The effect of using random initial values instead of true ones can be observed from the results for the case where both the DGP and the model have 10 Fourier sets, as this corresponds with the situation considered in Section 3.2. The mean RMSE of the information shares increases from 1.3% using true initial values to a substantially higher 7.8% for random initial values.

Panel B of Table 3 considers three different DGP's, namely a constant model, a step function model and the flexible Fourier model. The estimations use the GMM method and the state space method under the constant, step function and flexible Fourier specifications for the variances. We use the step function model as a simpler functional form for time-

⁶The initial values are drawn from a uniform distribution. The support is $[-1, 1]$ for the elements of α , $[0, 1]$ for the elements of Ψ considering the positive autocorrelation of the data, and $[-2, 2]$ for the parameters of the flexible Fourier form. The log innovation variances of the step function and constant models have a support of $[0, 2]$ and the log noise variances $[-5, 0]$.

variation, where the noise and innovation variances stay at different constant levels for 20 periods per day. Like the benchmark Fourier model and constant models, the parameter configuration of the step function model comes from the empirical application in Section 4 and therefore displays an intraday variation pattern similar to the Fourier model, making the resulting true intraday patterns very comparable. The estimation models without time-variation display a low accuracy even for DGP's of their own kind. Whereas in Section 3.1 the GMM and the state space models without time-variation have similar levels of accuracy, the use of random initial values gives a lead to the state space model with a mean information share RMSE of 18.5% to 31.9%. This difference widens under time-varying DGP's, although both constant estimation models are incapable of capturing such patterns.⁷ The Fourier model can capture time-varying patterns relatively better than the step function. Both the Fourier and the step function models have low mean RMSE's at estimating the step function DGP, with respectively 3.2% to 2.7%. However under the Fourier DGP, the Fourier model's 7.8% mean RMSE is nearly the half of the 15.0% mean RMSE of the step function. Under the DGP without time-variation, the step function works comparatively well with a mean information share RMSE of 6.2%, while the Fourier model has values close to the constant state space model with 16.1%.

4 Intraday variation of price discovery in the Expedia stock

We apply our methodology in an empirical setting involving prices of the Expedia stock observed at different trading venues during the second half of 2007. This data has been analyzed earlier by De Jong and Schotman (2010) in the context of constant information shares. Section 4.1 presents some key properties of the Expedia data. Sections 4.2 and 4.3 discuss two sets of results: Firstly, results from the GMM and state space methods under the assumption of constant information shares and, secondly, results from our model with

⁷Part of the favorable performance of our methodology comes from the superiority of the maximum likelihood estimation over GMM when the estimation model is correctly specified. However the constant state space estimates remain considerably more accurate than the GMM estimates even if the model is misspecified, i.e. the DGP has time-varying variances and the estimation model has constant ones.

intraday time-variation in price discovery via the state space approach. Section 4.4 presents an investigation into the reasons for information shares varying within the day.

4.1 Data and summary statistics

We examine the prices of the Expedia stock over 127 trading days from July 2 until December 28, 2007. The stock trades from 9:30h to 16:00h (New York time) at a number of exchanges. Our high-frequency data set consists of all quotes on all these exchanges, as obtained from the Trades And Quotes (TAQ) database. We use the midquote prices to avoid the bid-ask bounce present in transaction prices. The TAQ database time-stamps the quotes to the nearest second. For each second we determine the best bid and ask prices according to the procedure outlined in Hasbrouck (2010). We sample the data at the 1-minute frequency by using the midquote at the end of each minute.

We first consider the overall price movements of the Expedia stock over our sample period. Figure 2 presents the price series without separating the quotes from different exchanges. The Expedia stock price moves around \$30. It is relatively volatile, ranging between \$25 and \$35 in the half year we consider. There are some jumps in the price series, which all correspond to large overnight returns.

Next, we arrange the data into three groups according to quote origin as in De Jong and Schotman (2010). We consider the NASDAQ group (NASD), the NYSE and NYSE Arca group (NYSE) and the remaining exchanges (REST). This grouping aids both to the model parsimony in estimation and in the interpretation of the results. In Figure 3 we present the distribution of the number of quotes across the three groups. The NASDAQ system generates a particularly large share of the activity: Quotes from the NASDAQ (TAQ codes Q, T, or D) comprise 50% of the data. The other half is divided between the NYSE and NYSE Arca (TAQ codes N and P) with 17.6% and the rest coming from mainly the NSX (TAQ code C) with 22.8% and the CBOE (TAQ codes I and W) with 9.6%. This ranking between relative quote volumes is quite stable across the considered time period (top panel

Fig. 2. The Movements of the Expedia Stock

The figure shows the value of the Expedia stock over the 127 trading days from July to December 2007. The left axis is the price in USD and the bottom axis gives the days. The data is sampled at 1-minute frequency.



of Figure 3) as well as throughout the trading day (bottom panel).

4.2 Constant information shares

We estimate the Unobserved Components Model with constant information shares as given by Equations (1) and (2) using both the GMM and state space methods. As the simulation study of Section 3.1 indicates, only small differences may be expected in results. The difference mainly stems from the different data requirements and treatment of overnight returns. The covariance matrices that serve as the input for the GMM method are computed via log price differences and exclude overnight returns. In contrast, the state space uses directly the log prices and excludes not only overnight returns but also the first 11 of 391 intraday observations when calculating the likelihood.

The results of the GMM and state space approaches are reported in panels A and B of

Fig. 3. The Distribution of Quotes

The figures depict the distribution of Expedia quotes across the time interval of trading days and throughout the hours of a trading day. We consider three groups of exchanges: the NASD group (TAQ codes T and D), the NYSE group (TAQ code P), and the REST (TAQ codes B, C, I, M, W and X). The first figure displays the total number of daily quotes coming from each of the three groups, and the second one presents the average number of quotes for each 15 minutes of a trading day.

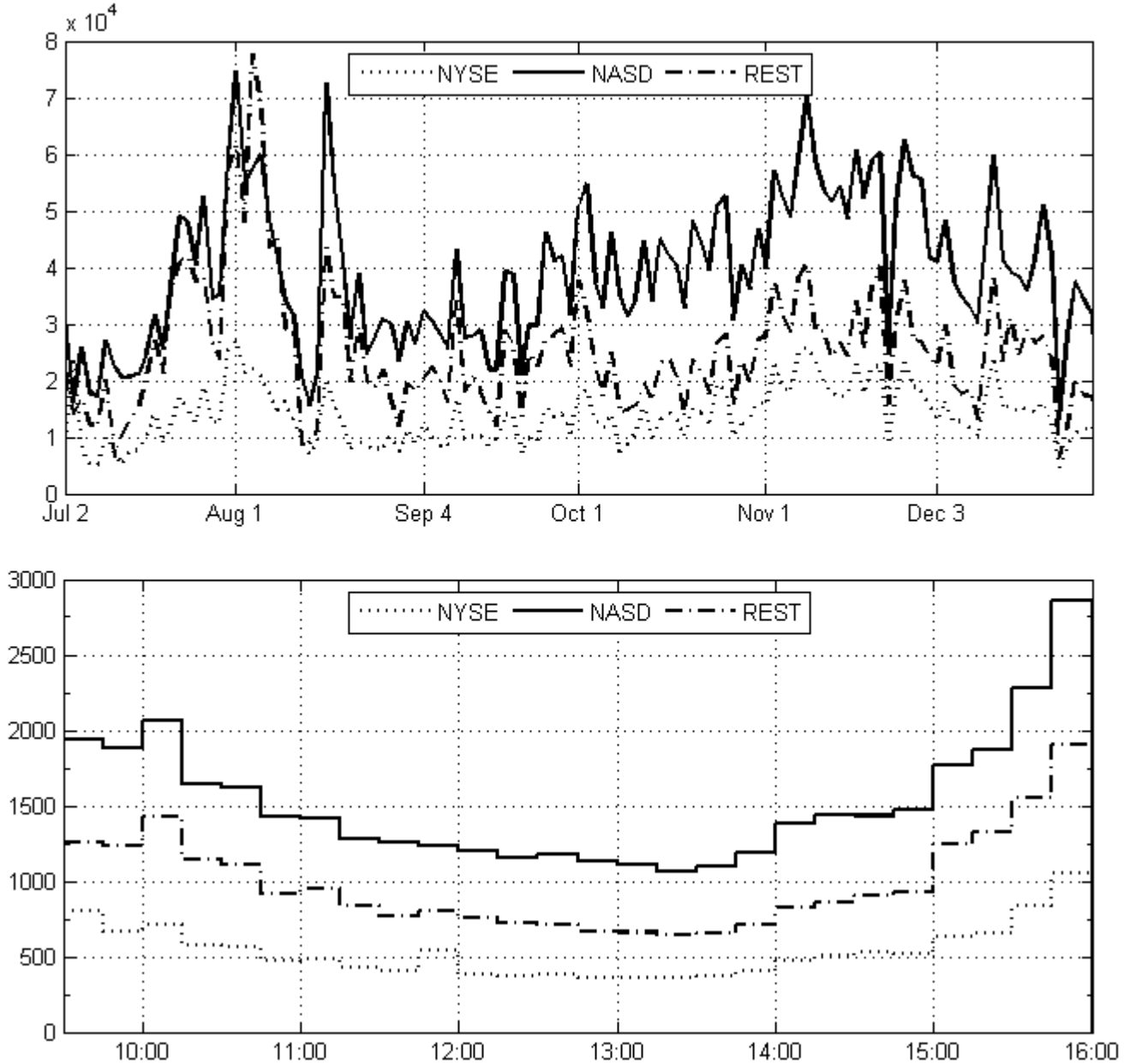


Table 4. As expected, they resemble each other quite closely.⁸ The lower innovation variance estimate of 0.82 obtained with the state space method compared to the GMM estimate of 0.94 mainly originates from the exclusion of the relatively volatile observations at the start of the trading day.⁹ The elements of α and the diagonals of the Ω and of Ψ matrices have the same qualitative ranking among each other and similar magnitudes across the two estimation methods.

Table 4. GMM and State Space Estimates

This table presents the state space and GMM estimation results for the Expedia stock. The estimates belong to the vector α , the variance of latent price innovations σ_r^2 , the idiosyncratic covariance matrix Ω and the information shares (IS). Panel A reports the GMM estimates under the diagonality restrictions on Ω and Ψ . Equations (8), (9) and (10) generate 24 moment conditions for GMM. Panel B reports the estimates from the state space model given in Equations (18) and (19). Panel C reports the daily means of time-varying model estimates for the flexible Fourier case.

Panel A: GMM Estimates				
	α	Ω	Ψ	<i>IS</i>
<i>NYSE</i>	-0.031	0.016	0.133	40.9
<i>NASD</i>	-0.043	0.012	-0.068	52.5
<i>REST</i>	-0.035	0.110	0.165	6.0
$\sigma_r^2 = 0.930$				99.4
Panel B: State Space Estimates				
	α	Ω	Ψ	<i>IS</i>
<i>NYSE</i>	-0.008	0.016	0.172	39.5
<i>NASD</i>	-0.022	0.012	0.087	53.9
<i>REST</i>	-0.006	0.107	0.270	5.8
$\sigma_r^2 = 0.816$				99.2
Panel C: Mean Flexible Fourier Estimates				
	α	Avg. Ω	Ψ	Avg. <i>IS</i>
<i>NYSE</i>	-0.010	0.019	0.142	25.9
<i>NASD</i>	-0.020	0.009	0.122	64.4
<i>REST</i>	-0.005	0.103	0.210	9.1
Avg. $\sigma_r^2 = 0.810$				99.4

The estimates point to several properties of the price discovery process. The small estimates of α suggest that price innovations are almost fully incorporated within a minute. Also

⁸Hansen's J-test for model validity cannot reject the model with diagonal Ω and Ψ matrices. The critical value of Hansen's J-statistic for the validity of the model is 23.68 for the significance level of 5% and the J-statistics of our estimates is 21.55.

⁹When we denote the inter-day observations as missing rather than initializing the Kalman filter each day, the innovation variance estimates get quite close to each other.

the noise in observed price changes is relatively small compared to genuine innovations, with the innovation variance being about 20 times higher than the mean noise variance. Lastly, the estimates of the diagonal elements of Ψ imply only a modest level of autocorrelation in the one-minute returns.

In terms of information shares both methods support that NASD dominates its main competitor, NYSE, with 52.5% against 40.9% based on the GMM estimates and with 53.9% against 39.5% based on the state space results. However the NYSE prices are very informative relative to their 17.6% share in quote volumes. By contrast, the combination of NSX and CBOE with respective volumes of 22.8% and 9.6% generates only 6% of the price discovery in the Expedia stock.

The differences in the information shares across the estimation methods stem mainly from the estimates of the idiosyncratic error variances in Ω . The negligible effect of the differences in σ_r^2 estimates can be observed from the sum of information shares, which is very similar across methods. As Equations (7) and (12) show, elements of α are summed with one in the formula for the information shares. Thus the differences of α estimates across methods have a negligible influence on the information shares, because they are very close to each other and to zero. As Equation (12) shows, the diagonal elements of the Ψ matrix are squared to compute the information shares, reducing their overall impact to almost nil. Therefore we observe that the (differences in) information shares are inversely proportional related to the (differences in) noise variance estimates.

4.3 Time-varying information shares

We model the variation in the innovation and noise variances as a deterministic time-varying function as discussed in Section 2.3. We estimate a model where these variances vary according to the flexible Fourier form complemented with a polynomial function, as given in Equation (11). To decide on the degree of the polynomial (denoted with P) and the number of Fourier sets (Q) we start with the model with the smallest number of parameters:

a model with a polynomial of order one and without any flexible Fourier sets ($P = 1$ and $Q = 0$). We then increase P and Q , while checking if the direction of increase leads to an improvement in the Schwarz Information Criteria.

A strict adherence to the information criteria does however not yield a final choice, mainly because the overlap of many intraday patterns across this half-a-year-long data set creates very complex intraday dynamics.¹⁰ As the improvements in the information criteria with each new flexible Fourier set gets smaller after 6 sets and the estimated pattern is stabilized, we decide on a specification with 10 Fourier sets and a first-order polynomial ($P = 1$, $Q = 10$), resulting in 94 parameters in total: three α_i 's, three diagonal elements in Ψ , and four variances with 22 parameters each. Further on, we also study the effect of choosing a bigger number of Fourier sets on the information shares.

Figure 4 shows the intraday patterns for the innovation variance $\sigma_{r,t}^2$ and the diagonal elements of the noise variance matrix Ω_t as obtained from the model specification with $P = 1$ and $Q = 10$. In each panel of the figure we show three lines. The middle line represents the intraday pattern based on the maximum-likelihood estimates of the parameters. Around the middle line we give 95% significance curves. As the pattern itself is not directly estimated, we use simulations to obtain these curves. We generate 50,000 draws from the asymptotic distribution of the parameter estimates and report highest density interval bounds for the variances and information shares containing 95% of these simulation results. The estimated innovation variance curve follows a pronounced U-shaped pattern. This pattern has a sound theoretical basis, as the informed trading literature documents such a U-shaped intraday innovation variance with a big peak at the start of the day (Admati and Pfleiderer, 1988; Foster and Viswanathan, 1993; Slezak, 1994). The noise variance estimates of the NYSE and REST groups also follow similar U-shaped patterns. As shown in Figure 4, the innovation and noise variance results have quite tight highest density intervals. The mean absolute

¹⁰In smaller samples, on the other hand, the improvement in the information criteria ceases at reasonable parameter numbers. In a separate analysis, available upon request, we divide our data set into six monthly periods and for four months the pattern converges under 10 flexible Fourier sets and for the other two under 20 sets.

difference between the bounds and the maximum likelihood estimates of the variances is merely 0.017 which corresponds to 1.8% of the mean estimates.

Figure 5 presents the resulting time-varying information shares computed according to Equations (5) and (6) using the constant estimates of α and the diagonal elements of Ψ , and the time-varying variance estimates. The figure displays strong variation in information shares throughout the day. On average, the NASD is the biggest contributor to price discovery followed by the NYSE and then the REST group. This result is consistent with the findings of the constant information share case in Section 4.2. In addition, we however see that the NYSE group leads over NASD for 68 minutes dispersed across the trading day. The remaining mostly regional exchanges in the REST group are never a strong alternative venue for the incorporation of new information. At the same time they are also not totally negligible with a maximum share of 29.3% and shares above 20% for 30 minutes.

In line with the variance estimates, the information share estimates in Figure 5 also have narrow highest density intervals. The mean absolute difference between the interval bounds and the estimates amounts to 6.8%. It is noteworthy that especially the nearly total dominance of the NASD group from 13:00h to 14:30h and at the end of the trading day has wider-than-average bounds.

Figure 6 presents our information share estimates in bounds of minimum and maximum estimates from models of 10 to 15 flexible Fourier sets. The model with 10 Fourier sets looks quite representative, given that the mean absolute difference between the interval bounds and the estimates is equal to 13.1% for the NYSE, 15.2% for the NASD, and 3.9% for the REST, which are considerably lower than the fluctuations of the estimates themselves. Interestingly the bounds of minimum and maximum estimates in Figure 6 show a similar pattern to the highest density intervals by increasing in the afternoon, particularly from 13:00h to 14:30h and just before the day end. The difference reaches 19.1% for the NYSE and 20.5% for the NASD in the last three trading hours, while they are 7.6% for the NYSE, 10.3% for the NASD before 13:00h.

Fig. 4. Intraday Variation in Innovation and Noise Variances

The figure displays innovation and noise variance estimates from the flexible Fourier model with their 95% highest density intervals in respectively thick and thin lines. The Fourier model consists of 10 Fourier sets and a polynomial of order one, leading to 94 parameters. The distribution of the variance estimates is approximated by 50,000 simulations using the parameter estimates from the state space framework and their covariance matrix. In cases where the distributions have more than one peak the lowest and the highest bounds are taken.

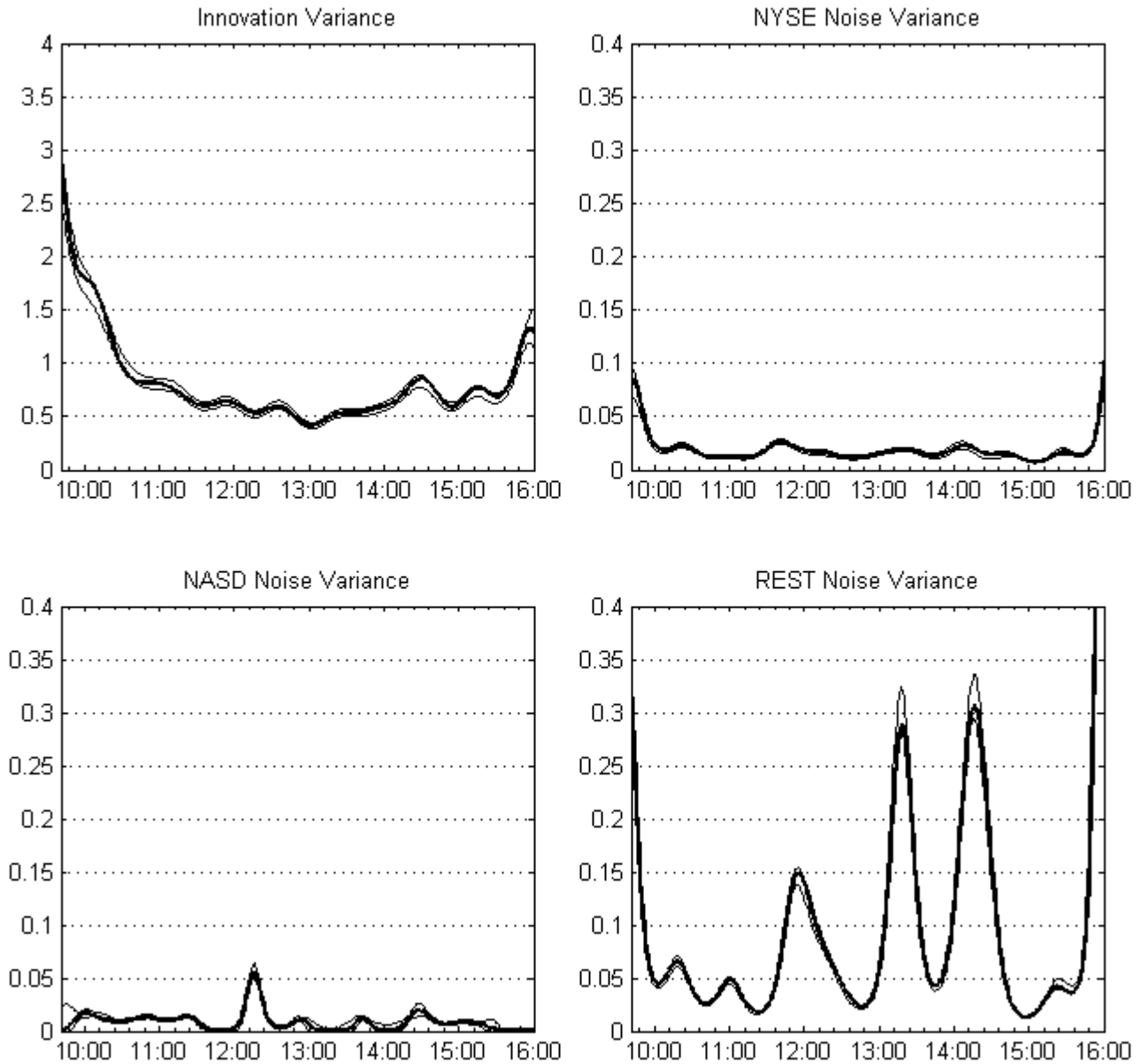


Fig. 5. Intraday Variation in Information Shares

The figure displays information share estimates with their 95% highest density intervals in respectively thick and thin lines. They shaded areas show when an exchange group has the highest information share. The Fourier model consists of 10 Fourier sets and a polynomial of order one. The distribution of the variance estimates is approximated by 50,000 simulations using the parameter estimates from the state space framework and their covariance matrix. In cases where the distributions have more than one peak the lowest and the highest bounds are taken.

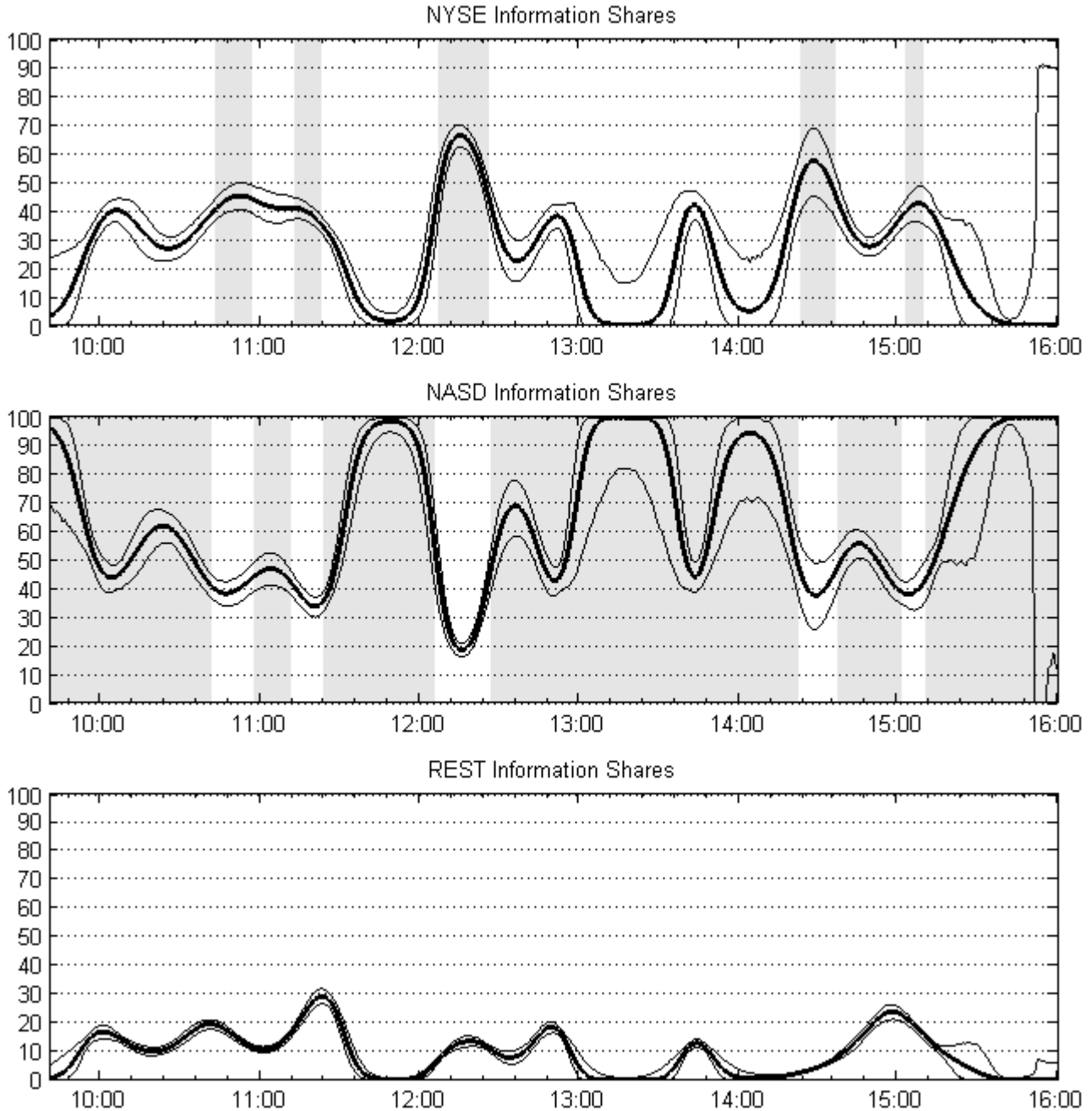
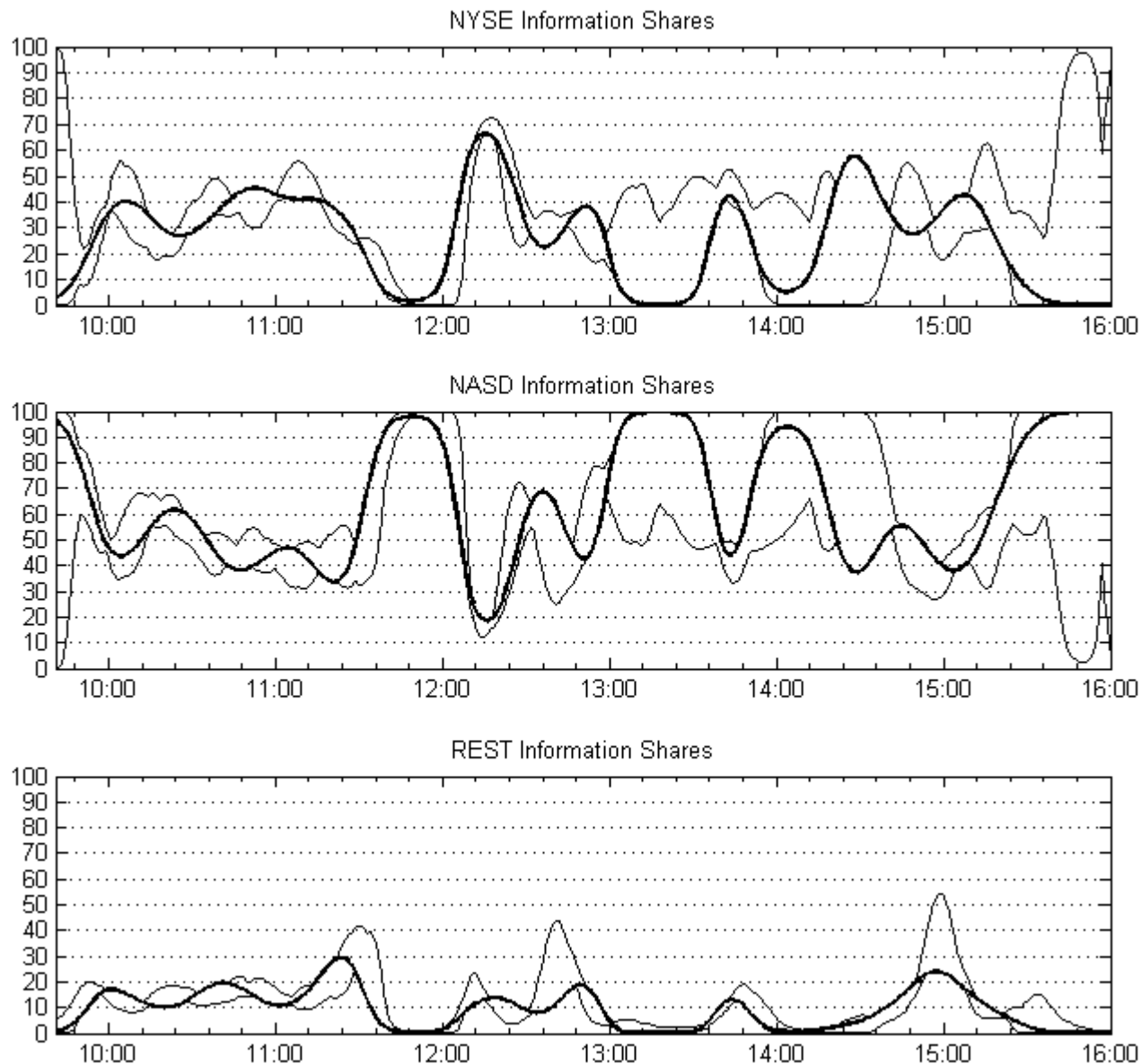


Fig. 6. Variation of Information Shares depending on the Number of Flexible Fourier Sets

The figure displays the range of information share estimates for 10 to 15 Flexible Fourier sets. The thick line gives the estimates with 10 Flexible Fourier sets and the thin lines give the minimum and the maximum of all estimates across this range.



Panel C of Table 4 allows the comparison of the results of the time-varying model with those of the constant case, reporting daily averages of the time-varying variances and information shares. The resemblance of the information share estimates stems from the similarity in the structural parameters. State space estimates have nearly identical values for both constant and time-varying models and they are at least qualitatively similar to those of the GMM method. Among the elements of the vector α , the diagonals of the noise covariance matrix Ω and of the matrix of lagged noise coefficients Ψ , the NASD group has the lowest parameter, followed by the NYSE and the REST. The conclusion is that at the daily level the same insights are learned from the constant and time-varying information share case, but that the time-varying case provides important additional insights at the intraday level.

4.4 What drives the intraday variation in information shares?

We conclude our empirical analysis by exploring which market features can explain the intraday variation of the information shares of the NASD, NYSE and REST exchanges for the Expedia stock as reported in Figure 5. We use a market share attraction model (Cooper and Nakanishi, 1988) to relate the dynamics of the estimated information share to market characteristics such as trade activity, liquidity and volatility.

Market share attraction models provide an effective framework to analyze the driving factors behind information shares. Just like the change of market shares in relation to the ability of companies to attract customers, the ability of a trading venue to attract informed trading shapes its information share. Market share attraction models also offer a gain in efficiency over separate regressions, as we can consider the determinants of the information shares for all exchanges simultaneously and in a consistent manner, as explained below.¹¹

Market share attraction models are based on the notion of a latent ‘attraction’ of a brand or company or, as in our case, a trading venue. The attraction of exchange i at time t is

¹¹The application of a simpler regression framework as in Neely and Mizrach (2008) would lead to three separate regressions for each of the three individual exchange groups. Section C of our web appendix investigates the intraday variation of the information shares using this approach.

defined as

$$A_{i,t} = \exp(\mu_i + \varepsilon_i) \prod_{j=1}^I \prod_{k=1}^K f(x_{k,j,t})^{\beta_{k,j,i}} \quad (13)$$

where $x_{k,j,t}$ is the value of the k^{th} explanatory variable of exchange j at time t , $\beta_{k,j,i}$ denotes the effect of this explanatory variable on the attraction of exchange i , I is the number of exchanges and K is the number of explanatory variables for each exchange. The function $f(\cdot)$ denotes a particular transformation of the explanatory variable $x_{k,j,t}$. We return to this below. The model is completed by setting the observed information share of exchange i at time t equal to its attraction relative to the sum of all attractions, that is,¹²

$$IS_{i,t} = \frac{A_{i,t}}{\sum_{j=1}^I A_{j,t}}. \quad (14)$$

We limit ourselves to the two most popular specifications of the function $f(\cdot)$ in (13). These are the identity function, i.e. $f(x_{k,j,t}) = x_{k,j,t}$, and the exponential transformation, i.e. $f(x_{k,j,t}) = \exp(x_{k,j,t})$. The first specification results in the so-called Multiplicative Competitive Interaction (MCI) specification, while the latter leads to the Multinomial Logit (MNL) specification. These two specifications differ in terms of the implied elasticities of the information shares with respect to the explanatory variables. Under the MCI specification the elasticity of information share i to the k^{th} explanatory variable of exchange j at the minute t is given by

$$e_{k,j,i,t}^{MCI} = \frac{\partial IS_{i,t}}{\partial x_{k,i,t}} \frac{x_{k,i,t}}{IS_{i,t}} = \beta_{k,j,i}(1 - IS_{i,t}). \quad (15)$$

while the elasticity for the MNL specification is

$$e_{k,j,i,t}^{MNL} = \frac{\partial IS_{i,t}}{\partial x_{k,i,t}} \frac{x_{k,i,t}}{IS_{i,t}} = \beta_{k,j,i}(1 - IS_{i,t})x_{k,i,t}. \quad (16)$$

¹²Note that this implies that the information shares always sum to unity at each point in time. This need not necessarily hold for the information shares of De Jong and Schotman (2010) as defined in Equation (7). Although in our empirical application the mean of the sum of information shares is very close to 1 with 99.4% (see Table 4), we therefore normalize the information share estimates for the remainder of the analysis.

In words, the MCI specification implies a monotonic decline in the elasticity with increasing values of the explanatory variable. The MNL specification on the other hand permits the increase of elasticity up to a certain level before a decline sets in (Cooper, 1993). This difference thus is mostly relevant when considering the elasticities for small values of the explanatory variables.

In our implementation of the MSA model, we consider the number of trades, quoted spreads and volatility as explanatory variables for the information shares. Considering that the information share of an exchange would be more susceptible to changes in quoted spreads and volatility when these two have small values, we model quoted spreads and volatility using the MCI specification. By contrast changes in the number of trades would not matter at small values and therefore we use the MNL specification for the number of trades. For each of these three variables we consider the average value in minute t over all days in the sample. For the quoted spread we take the average difference between the bid and ask within the minute. For volatility we use the square root of the mean squared difference between the first trade prices of consecutive minutes across trading days.

In order to estimate the model parameters, we may take one of the exchanges as the benchmark (labeled “ b ”) and rewrite the model as a set of $I - 1$ equations given by

$$\log(IS_{i,t}) - \log(IS_{b,t}) = \tilde{\mu}_i + \sum_{j=1}^I \sum_{k=1}^l \tilde{\beta}_{k,j,i} x_{k,j,t} + \sum_{j=1}^I \sum_{k=l+1}^K \tilde{\beta}_{k,j,i} \log(x_{k,j,t}) + \tilde{\varepsilon}_i \quad (17)$$

where $\tilde{\mu}_i = \mu_i - \mu_b$, $\tilde{\beta}_{k,j,i} = \beta_{k,j,i} - \beta_{k,j,b}$ and $\tilde{\varepsilon}_i = \varepsilon_i - \varepsilon_b$, and we have assumed that the first l explanatory variables enter the model with an MNL specification and the remaining $K - l$ explanatory variables are included with an MCI specification. The complete system has $(I - 1) \cdot (1 + I \cdot K)$ parameters, which can be estimated by OLS (Fok, Franses and Paap, 2002). Note that only differences of the coefficients $\beta_{k,j,i}$ can be identified. As shown by Cooper and Nakanishi (1988), this is however sufficient to completely identify the elasticities given in Equations (15) and (16).

Panel A of Table 5 shows that standard market quality variables can explain a sizable portion of the intraday variation in information shares. The adjusted R^2 's of the regression is 44.3%. We observe a negative relationship between the number of trades and information shares for all exchanges. This is somewhat surprising, as it indicates the information share becomes lower if a market has a relatively high number of trades. To examine this further, we turn to the trade size effect in more detail later. The table also shows that information shares have a negative relationship with quoted spreads. This implies that relatively narrower bid-ask spreads accelerate price discovery. The effect is particularly strong for the NYSE and the REST groups where a 1% decrease in quoted spreads increases the information shares by 1.56% and 2.47 % respectively. Volatility, on the other hand, has a mildly negative effect on the shares of the NASD and the REST, but a considerably strong positive one to the NYSE. As volatility incorporates both noise and innovations, such an ambiguous result is understandable.

Cross-elasticities imply that the NYSE and the REST act like complimentary goods, while the NASD is a substitute for both of them. All variables have elasticities with the identical signs and arguably similar magnitudes to the information shares of the NYSE and the REST, while they have the opposite signs for the NASD. The analogy of complementarity/substitution is particularly appropriate for the elasticities to quoted spreads, because these can be interpreted as a proxy for transaction costs which represent the price of choosing to trade in a particular exchange.

To examine the coefficient on the market-specific number of trades in more detail, we split the ratio of the number of trades into five groups according to their sizes. The five groups are: Number of trades with size not greater than 100 (comprising 66.7% of all trades), above 100 and less than or equal to 200 (18.5%), above 200 and less than or equal to 300 (5.8%), above 300 and less than or equal to 500 (4.8%), and above 500 (4.3%).¹³ In this manner we evaluate the effect of trade size on price informativeness.

¹³Separating the number of trades by quantiles gives similar results, because the number of trades have sharp spikes at multiples of hundred and each quantile gets one of those spikes just like the current separation.

Table 5. Elasticities of Information Shares to Market Variables

Panels A and B show the means of minute-level elasticities of the information shares to market variables. We estimate the coefficients of the market share attraction model via OLS using the set of equations (17) and compute the daily elasticity series in Equations (16) and (15). Panel A has the number of trades for each minute as the first variable of each set of explanatory variables from the three exchange groups, while in Panel B this value is split into 5 groups for trades with size 100, bigger than 100 and lesser than or equal to 200, bigger than 200 and lesser than or equal to 300, bigger than 300 and lesser than or equal to 500, and above 500. In both panels these are followed by quoted spreads and volatility. Lastly, we report the adjusted R^2 and the number of observations used in the original regression.

Panel A: Elasticities to Market Variables

<i>Info.</i>	<i>NYSE</i>			<i>NASD</i>			<i>REST</i>			<i>Adj.R²</i>	<i>N</i>
	<i>Num.ofTr.</i>	<i>Spread</i>	<i>Volatility</i>	<i>Num.ofTr.</i>	<i>Spread</i>	<i>Volatility</i>	<i>Num.ofTr.</i>	<i>Spread</i>	<i>Volatility</i>		
<i>NYSE</i>	-1.93	-1.56	0.97	1.01	0.62	0.17	-0.69	-1.05	-0.02	0.443	760
<i>NASD</i>	0.96	0.66	-0.57	-0.49	-0.33	-0.13	0.33	0.77	0.01		
<i>REST</i>	-2.20	-0.22	1.24	0.93	0.56	0.41	-0.79	-2.47	-0.04		

Panel B: Elasticities with Trades divided according to their Sizes

<i>Info.</i>	<i>NYSE</i>						<i>Q.Spread</i>	<i>Volatility</i>
	100	100 – 200	200 – 300	300 – 500	> 500			
<i>NYSE</i>	-0.87	0.00	-0.31	-0.36	0.16	-1.88	0.68	
<i>NASD</i>	0.43	0.00	0.14	0.17	-0.09	0.87	-0.38	
<i>REST</i>	-0.85	0.02	-0.31	-0.57	0.36	-0.74	0.76	

<i>Info.</i>	<i>NASD</i>						<i>Q.Spread</i>	<i>Volatility</i>
	100	100 – 200	200 – 300	300 – 500	> 500			
<i>NYSE</i>	2.67	0.38	-0.97	-1.08	0.67	0.14	0.01	
<i>NASD</i>	-1.53	-0.23	0.51	0.51	-0.31	-0.02	-0.03	
<i>REST</i>	3.85	0.78	-1.76	-1.76	1.11	-0.26	0.15	

<i>Info.</i>	<i>REST</i>						<i>Q.Spread</i>	<i>Volatility</i>	<i>Adj.R²</i>	<i>N</i>
	100	100 – 200	200 – 300	300 – 500	> 500					
<i>NYSE</i>	-0.23	-0.11	-0.15	0.07	-0.12	-0.94	-0.08	0.494	760	
<i>NASD</i>	0.09	0.06	0.06	-0.02	0.04	0.70	0.05			
<i>REST</i>	-0.09	-0.22	-0.17	0.00	-0.09	-2.27	-0.14			

Panel B of Table 5 provides the results from splitting the number of trades according to trade size. The model with disaggregated trades has a higher explanatory power with an R^2 of 49.4% and all the aforementioned results of the regression in Panel A are preserved. The negative elasticities to the number of trades seems to be mostly driven by trades that are in the smallest trade categories, as 85.2% of the trades have quantities lower than or equal to 200 which have a negative relationship with price informativeness. However the results for trades with higher sizes differs across exchanges. In the NYSE only the number of trades with sizes above 500 have a positive relationship with information shares, while for the NASD mid-sized trades ranging from 200 to 500 are more informative.

5 Conclusion

This paper proposes a novel approach to measure the contribution to price discovery made by different observed security prices, with an explicit focus on capturing intraday dynamics in information shares. We use a state space representation of the Hasbrouck (1995) Unobserved Components Model. We introduce intraday time-variation in De Jong and Schotman (2010) information shares by allowing for time-varying volatilities of the efficient price innovations and idiosyncratic noise, using flexible Fourier specifications.

Our simulation study displays the capability of our method in capturing intraday dynamics of price discovery for typical data sets used in the market microstructure literature. Across a wide range of settings and parameter configurations it consistently provides accurate estimates of the models structural parameters and the associated information shares.

In our empirical analysis we examine the Expedia stock trading in multiple exchanges. We gather these exchanges in three groups by quote origin: NASDAQ-related, NYSE-related and other exchanges. We observe that most of the new information is incorporated into prices via the first two groups, particularly by the NASDAQ system. However, exchanges related to NYSE also lead the price discovery at various parts of the trading day. We find that a number of market quality variables can explain almost half of of the intraday variation in

information shares. The number of trades, spreads and volatility have a significant effect on price discovery.

Our state space methodology advances the information shares literature to the investigation of intraday dynamics. In present-day financial markets the incorporation of news into prices takes minutes, if not seconds, and also access to high frequency data gets easier. These factors provide a fertile ground for the application of our methodology to contemporary issues in price discovery.

Appendix: The state space representation of the Unobserved Components Model

In the state space form, the Unobserved Components Model given in Equations (1) and (2) can be represented by these two equations:

$$p_t = [\iota_{N \times 1} \ \alpha \ \Psi] \begin{bmatrix} p_t^* \\ r_t \\ e_{t-1} \end{bmatrix} + G\varepsilon_t, \text{ where } G = [0_{N \times 1} \ I_N] \text{ and } \varepsilon_t = \begin{bmatrix} r_{t+1} \\ e_t \end{bmatrix}, \quad (18)$$

$$\begin{bmatrix} p_{t+1}^* \\ r_{t+1} \\ e_t \end{bmatrix} = \begin{bmatrix} 1 & 0_{1 \times N+1} \\ 0_{N+1 \times 1} & 0_{N+1 \times N+1} \end{bmatrix} \begin{bmatrix} p_t^* \\ r_t \\ e_{t-1} \end{bmatrix} + H\varepsilon_t, \text{ where } H = \begin{bmatrix} \iota_{2 \times 1} & 0_{2 \times N} \\ 0_{N \times 1} & I_N \end{bmatrix}, \quad (19)$$

with $\iota_{n \times m}$ an $n \times m$ vector of ones, $0_{n \times m}$ an $n \times m$ matrix of zeros, Ψ is an $N \times N$ matrix and I_N is an $N \times N$ identity matrix.¹⁴ The variance parameters are uniquely identified using the covariance matrix of the stacked disturbances $\begin{bmatrix} H \\ G \end{bmatrix} \varepsilon_t$, which comprises the innovation and noise variances:

$$E \left[\begin{bmatrix} H \\ G \end{bmatrix} \varepsilon_t \varepsilon_t' \begin{bmatrix} H \\ G \end{bmatrix}' \right] = \begin{bmatrix} \sigma_r^2 \iota_{2 \times 2} & 0_{2 \times N} & 0_{2 \times N} \\ 0_{N \times 2} & \Omega & \Omega \\ 0_{N \times 2} & \Omega & \Omega \end{bmatrix}.$$

¹⁴As the noise terms e_t and the innovation r_t are in different equations, we could have avoided combining them under ε_t . However this formulation is in line with the model entry requirements of the SsfPack by Koopman, Shephard and Doornik (1998) used in this paper. Oftentimes a model bears more than one equivalent state space representation.

References

- Admati, A.R. and Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies*, 1, 3-40.
- Ahn, H., Bae, H. and Chan, K. (2001). Limit orders, depth and volatility: Evidence from the stock exchange of Hong Kong. *Journal of Finance*, 56, 769-790.
- Andersen, T.G., Bollerslev, T. and Das, A. (2001). Variance-ratio statistics and high-frequency data: Testing for changes in intraday volatility patterns. *Journal of Finance*, 56, 305-327.
- Chae, J. (2005). Trading volume, information asymmetry, and timing information. *Journal of Finance*, 60, 413-442.
- Chakravarty, S., Gulen, H. and Mayhew, S. (2004). Informed trading in stock and option markets. *Journal of Finance*, 59, 1235-1256.
- Cooper, L.G., and Nakanishi, M. (1988). *Market Share Analysis*. Boston: Kluwer.
- Cooper, L.G. (1993). Market share models, in Eliasberg, J. and Lilien, G.L. (eds.), *Handbook of Operations Research and Management Science: Marketing*, Vol. 5, New York: North-Holland, pp. 259-314.
- De Jong, F. (2002). Measures of contributions to price discovery: A comparison. *Journal of Financial Markets*, 5, 323-327.
- De Jong, F. and Schotman, P.C. (2010). Price discovery in fragmented markets. *Journal of Financial Econometrics*, 8, 1-28.
- Fama, E.F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25, 383-417.
- Frijns, B. and Schotman, P.C. (2009). Price discovery in tick time. *Journal of Empirical Finance*, 16, 759-776.
- Fok, D., Franses, P.H. and Paap, R. (2002). Econometric analysis of the market share attraction model. In Franses, P.H. and Montgomery, A. L. (eds.), *Advanced in Econometrics Vol. 16: Econometric Models in Marketing*. New York: JAI/Elsevier, pp. 223-256.
- Foster, F.D. and Viswanathan, S. (1993). Variations in trading volume returns volatility and trading costs: evidence on recent price formation models. *Journal of Finance* 48, 187-211.
- Gallant, A.R. (1981). On the bias in flexible Functional forms and an essentially unbiased Form: The Fourier flexible form. *Journal of Econometrics*, 15, 211-245.
- Gonzalo, J. and Granger, C.W.J. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business and Economic Statistics*, 13, 27-36.
- Grammig, J. and Peter, F.J. (2013). Tell-tale tails: A new approach to estimating unique market information shares. *Journal of Financial and Quantitative Analysis*, 48, 459-488.
- Harris, F.H.deB., McInish, T.H. and Wood, R.A. (1997). Common long-memory components for intraday stock prices: A measure of price discovery. Working paper, Babcock School, Wake Forest University, Winston-Salam, NC.
- Harris, F.H.deB., McInish, T.H. and Wood, R.A. (2002). Security price adjustment across exchanges: an investigation of common factor components for Dow stocks. *Journal of*

- Financial Markets*, 5, 277-308.
- Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *Journal of Finance*, 50, 1175-1199.
- Hasbrouck, J. (2003). Intraday price formation in US equity index markets. *Journal of Finance*, 58, 2375-2399.
- Hasbrouck, J. (2004). Liquidity in the futures pits: Inferring market dynamics from incomplete data. *Journal of Financial and Quantitative Analysis*, 39, 305-326.
- Hasbrouck, J. (2010). The best bid and offer: A short note on programs and practices. Working Paper.
- Hsieh, D.A and Kleidon, A.W. (1996). Bid-ask spreads in foreign exchange markets: Implications for models of asymmetric information. In Frankel, J.A., Galli, G. and Giovannini, A. (eds), *The Microstructure of Foreign Exchange Markets*. Chicago: University of Chicago Press.
- Koopman, S.J., Shephard, N. and Doornik, J.A. (1998). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal*, 1, 1-5.
- Korenok, O., Mizrach, B. and Radchenko, S. (2011). A structural approach to information shares. Working paper.
- Lehmann, B. (2002). Some desiderata for the measurement of price discovery across markets. *Journal of Financial Markets*, 5, 259-276.
- Lei, Q. and Wu, G. (2005). Time-varying informed and uninformed trading activities. *Journal of Financial Markets*, 8, 153-81.
- Menkveld, A.J., Koopman, S.J. and Lucas, A. (2007). Modeling round-the-clock price discovery for cross-listed stocks using state space methods. *Journal of Business and Economic Statistics*, 25, 213-225.
- Mizrach, B. and Neely, C. (2008). Information shares in the US Treasury market. *Journal of Banking and Finance*, 32, 1221-1233.
- Slezak, S.L. (1994). A theory of the dynamics of security returns around market closures. *Journal of Finance*, 49, 1163-1211.
- Thompson, S. and Waller, M. (1988). The intraday variability of soybean futures prices: Information and trading effects. *Review of Futures Markets*, 7, 110-126.
- Upper, C. and Werner, T. (2007). The tail wags the dog: Time-varying information shares in the Bund market. BIS Working Papers No. 224.
- Wood, R.A., McInish, T.H. and Ord, J.K. (1985). An investigation of transactions data for NYSE stocks. *Journal of Finance*, 40, 723-39.