Appendix: Proxy for the Loadings of Information Content on Unexpected Day Trading Volume

Not only the empirical analysis relied on econometric methods, by the VAR analysis and EGARCH model, we also create an information proxy, loadings of information-based trading ratio (LIN), to test if unexpected day trading volume is indeed related to news coming to the market. LIN is similar to the spirit of the probability of informed trading (PIN), which is used to estimate the probability that a given security is subject to informed trading over a certain period of time.

Based on previous theoretical work from Glosten and Milgrom (1985), Easley and O'Hara (1987, 1992) and Easley, Kiefer and O' Hara (1997), Easley, Kiefer, O'Hara and Paperman (1996) and Easley, Hvidkjaer and O'Hara (2002) impose an assumed microstructure model on the data and then estimate the model's parameters, estimated via maximum likelihood, in order to construct an estimate of the PIN. These parameters then become the building blocks of the PIN variable, which is the ratio of the expected number of informed trading over the expected total number of trades. Thus, the PIN variable capture the probability of informed trading.[1]

The major assumption of PIN is assumed to that information events occur (with a given probability) independently on a daily basis and that the numbers of buy and sell orders are assumed to be independently distributed by the Poisson distribution. However, contrast to the PIN, the LIN, is more concisely used to identify the information content of unexpected day trading volume in this paper, since we release the limitations both on the independence and distribution in the PIN measurement to avoid the independence of estimated parameter implicit in Poisson procedure. Moreover, corresponding with the decomposing procedure, the LIN provides more clear measurement which aggregates the positive and negative values, which could violate the assumption of the probability in

[1] For example, at the beginning of each period, an informational event could occur with probability α. These events are assumed to be independent. Given an informational event has occurred, this could be positive with probability 1-δ, or negative with probability δ. Informed traders anticipate that at the end of the day, the value of the stock, for example, will be $\tilde{v}_i$ if news is bad or ν_i if news is good, where $\tilde{v}_i > \nu_i$ and i is the trading period.

PIN. Finally, contrast to the PIN which uses the antithesis of buy/sell sides trading, the LIN earns more information among the decomposing expected (unexpected) volume variables.[2] In short, the LIN ration is more suitable and intuitive than the PIN in the paper's major issue, if unexpected day trading volume is indeed related to news coming to the market. Similar to the concept of Easley, Hvidkjaer and O'Hara (2002), we formally define the LIN variable as,

$$LIN_t = \frac{DT_t^{un}}{DT_t^{un} + (V_t^e + OI_t^e + DT_t^e)}, \tag{A1}$$

where, V_t^e, OI_t^e, and DT_t^e is the expected total trading volume, open interest and day trading volume variables, respectively. DT_t^{un} proxies for unexpected day trading volume. The new information token by unexpected day trading volume in a given day is ratio of DT_t^{un} over the sum of (the expected total number of trades is $(V_t^e + OI_t^e + DT_t^e)$ and DT_t^{un}. Including the DT_t^{un} in the denominator is used to avoid multicollinearity in OLS regression on DT_t^{un}. Thus, the fraction of informed trades to the total expected number of trades is a way to estimate the loadings of informed trading. The LIN takes on high values if there are relatively big and infrequent jumps in the number of day trades. Intuitively, such jumps are more likely due to the arrival of good or bad news into the market.

Before we formally test the information content by the LIN ratio, we preliminary illustrate the descriptive statistics of LIN. The mean of LIN is -0.014 with a standard deviation, -0.013, and follows a normal distribution at 1% statistically significant level. According to the test of equality, we find that the futures' return series and LIN series are no difference based on the value of t-statistics, 0.298. Next, similar to the procedure of Mohanram and Rajgopal (2009) who find that PIN is a priced risk factor to capture a stock's liquidity and that PIN factor loadings predict returns, we use LIN as an information proxy to investigate that whether the new shocks are absorbed by the (lag) unexpected day trading volume and then if reflect this shock to forecast futures' returns.

[2] It is important to mention that PIN does not generally take high values when there are frequent jumps.

The models to test the information transferring channel of unexpected day trading volume are specified as follows,

$$\Delta OI_t = a_0 + a_1 R_t + a_2 LIN_t * DT_t^{un} + \varepsilon_t, \tag{A2}$$

where ΔOI_t is the daily chance of level on open interest at day t. The results are provided in Table A1.

$$\begin{aligned} R_t = a_0 + a_1 V_t^e + a_2 V_t^{un} + a_3 OI_t^e + a_4 OI_t^{un} + \\ a_5 DT_t^e + a_6 DT_t^{un} + a_7 DT_{t-1}^{un} + a_8 LIN_t \sigma_{t-1} + \varepsilon_t, \end{aligned} \tag{A3}$$

where σ_{t-1} is the lagged conditional volatility estimated from EGARCHM model. In contrast to open interest, equation A2 is adopted to test whether day traders are typically sophisticated investors who take risks and seek to exploit their information advantages through the use of day trading orders in stock markets (Odean, 1998, and Büyükşahin and Harris, 2009). Specifically, if the coefficient of the cross term of information loading factors, LIN, and unexpected day trading volume is significant, then we infer that the new information is first reflected by unexpected day trading volume and then changes the open interest corresponding with the changes in the price-volume relationship to a new equilibrium. Equation (A3) is used to investigate whether information loading factor, LIN, the ratio of unexpected day trading volume to all expected volume variables, is indeed related to news coming to the returns under controlling the other volume variables. If the unexpected change of day trading volume, corresponding with LIN ratio, significantly reflect the more shock of the new information to the change of price behaviors than the evidence of open interest of Hong and Yogo (2012).

Table A1 and A2 provide the results. The cross term is statistically significant to the change of the level on open interest in Table A1, suggesting that new shocks of the information in unexpected day trading volume provides information to the change of open interest. Moreover the cross term of lagged conditional volatility (shocks) and information loadings ratio is, for example, 0.240 in Table A2 Model 1, at 5% statistical significance to futures return process. Similar results can be found in Model 2 – 4. This result indicates that this ratio, LIN, indeed provides information coming to the market, and the ratio $LIN_t = \frac{DT_t^{un}}{DT_t^{un} + (V_t^e + OI_t^e + DT_t^e)}$ is a valuable and simple index for counting the news into the market. Finally, the LIN ratio is useful to calculate and forecast the changes of the open interest.

Reference

Easley D., S., Hvidkjaer and M., O'Hara, 2002, Is Information Risk a Determinant of Asset Returns?, *Journal of Finance* 47, 2185–2221.

Easley D., N., Kiefer and M., O'Hara, and J., Paperman, 1996, Liquidity, Information, and Infrequently Traded Stocks, *Journal of Finance* 51, 1405–1436.

Easley D., and M., O'Hara, 1987, Price, Trade Size and Information in Securities Markets, *Journal of Financial Economics* 19, 69–90.

Easley D., and M., O'Hara, 1992, Time and the Process of Security Price Adjustment, *Journal of Finance* 47, 577–605.

Glosten, L., and P. Milgrom, 1985, Bid, Ask and Transaction Prices in A Specialist Market with Heterogeneously Informed Traders. *Journal of Financial Economics* 14, 71–100.

Mohanram, P., and S., Rajgopal, 2009, Is PIN Priced Risk?, *Journal of Accounting and Economics* 47, 226–243.

Table A1. Information Transferring Channel of Unexpected Day Trading Volume to the Change of Open Interest

$$\Delta OI_t = a_0 + a_1 R_t + a_2 LIN_t * DT_t^{un} + \varepsilon_t,$$

Variable	Coefficient	t-Statistic	Prob.
C	-1.022⁂	-36.325	0.000
R_t	-0.046⁂	-3.897	0.000
$LIN_t * DT_t^{un}$	0.356‡	2.280	0.023

Adjusted R^2	3.40%
F-statistics	9.585
Prob(F-statistic)	0.000

Note: †, ‡, and ⁂ represent that the null hypothesis is rejected at 1%, 5% and 10% significance, respectively.

Table A2. Results: Forecasting of Unexpected Day Trading Volume

$$R_t = a_0 + a_1 V_t^{e} + a_2 V_t^{un} + a_3 OI_t^{\ e} + a_4 OI_t^{\ un} + a_5 DT_t^{e} + a_6 DT_t^{un} + a_7 DT_{t-1}^{\ un} + a_8 LIN_t \sigma_{t-1} + \varepsilon_t,$$

	Model 1.		Model 2.		Model 3.		Model 4.	
Variable	Coef.	t-value (p-sta.)	Coef.	t-value (p-sta.)	Coef.	t-value (p-sta.)	Coef.	t-value (p-sta.)
C	6.283	0.149	-1.891⁂	-3.044	21.132	0.496	16.211	0.385
		(0.882)		(0.003)		(0.620)		(0.701)
V^{e}_{t}	-6.603	-0.194			-1.768	-0.514	-1.346	-0.301
		(0.846)				(0.607)		(0.487)
V^{un}_{t}	-1.049	-0.141			-2.808	-0.514	-2.136	-0.396
		(0.816)				(0.607)		(0.692)
OI^{e}_{t}	2.872	1.279	0.237‡	2.093			0.126†	1.668
		(0.202)		(0.037)				(0.096)
OI^{un}_{t}	-0.577⁂	-3.534	-0.580⁂	-3.569			-0.628⁂	-3.793
		(0.000)		(0.000)				(0.000)
DT^{e}_{t}	-0.124	-0.808	-0.136	-1.271	-0.200	-1.299		
		(0.420)		(0.204)		(0.195)		
DT^{un}_{t}	-0.270‡	-2.119	-0.270‡	-2.126	-0.288‡	-2.229		
		(0.035)		(0.034)		(0.026)		
DT^{un}_{t-1}	0.634	1.028	0.646	1.054	0.648	1.036		
		(0.305)		(0.293)		(0.301)		
$LIN_t\ \sigma_{t-1}$	0.204‡	2.266	0.202‡	2.300	0.215‡	2.361	0.239‡	2.060
		(0.024)		(0.022)		(0.019)		(0.040)

Model Diagnosis	Model 1.	Model 2.	Model 3.	Model 4.
Adjusted R^2	3.964%	3.401%	1.858%	3.570%
F-statistic	4.427	3.313	1.584	3.732
Prob(F-sta.)	0.000	0.001	0.150	0.002

Note: The measurement of expected (unexpected) volume and open interest is in terms of 10,000 contracts, and that of day trading is in terms of 1,000 contracts. †, ‡, and ⁂ represent that the null hypothesis is rejected at 1%, 5% and 10% significance, respectively.