Free Temporal Causality from Market Microstructure Noise in Foreign Exchange Markets: A Brief Proposal

Luofeng Zhou

1. Causal Inference in Financial Markets

Price discovery in financial markets is in fact a causal inference process. When investors apply a model or an algorithm in forming portfolio or doing algorithmic trading, the underlying can always be described as follow: if certain factors exhibit certain characteristics, it is believed to have a short-run or long-run increase (decrease) that enables profitability. This cause could either be statistical or actual since algorithmic trading concerns only the profitability. Causal inference in financial markets is also important for regulators to determine whether it is the existing or lacking regulations that is to blame.

A large part of empirical research in financial markets, corporate finance and accounting relies on observational (non-experimental) data. Gow et al. (2016) survey the causal inference methods in accounting research and these methods are also the most common ones. Ordinary least-square (OLS) regression after controlling other influential, instrumental variables (IVs), natural experiments (exogenous shocks), regression discontinuity (RD) and so on.

Natural experiments aim at revealing the effect of the exogenous shocks. Informed traders may form portfolio before a certain exogenous shock takes place. For other investors, it is usually not a good investment opportunity. So scholars apply natural experiments in financial markets to check whether investors are informed (e.g. Christophe et al., 2004, 2010). Regression discontinuity is designed to identify main treatment impacts (Hahn et al., 2001).

Instrumental variables are described as "the most powerful weapon in the arsenal" (Angrist and Pischke, 2008) . However, good instruments are always hard to find since "good instruments come from a combination of institutional knowledge and ideas about the process determining the variable of interest" (Angrist and Pischke, 2008). For example, when measuring market sentiment, indicators like put-call ratio (Dennis and Mayhew, 2002), Barron's confidence index (Lashgari, 2000) and TED spread (Lashgari, 2000) are calculated through market data and regressions are suspicious through this way. Bollen et al. (2011) use twitter mood drawn by text mining technique to predict stock market return. Same measure could be used to determine causality between investors' sentiment and returns.

Most of extant research about financial markets are based on vector autoregression (VAR) model which is the direct extension of OLS. The causal relation is given by the transmission matrix (coefficient matrix). If a non-diagonal term equals to zero, it indicates no one-way causal relation corresponding to its footnotes. But if a non-diagonal term is not zero, significance of one-way causality is tested through Granger causality test (Block Exogeneity Wald test) using the estimated variance of the error terms. However, there are some main demerits of this method. First, the Granger causality test determines the so-called Granger causality which is statistical rather than actual or theoretical. Thus, only prediction models can utilize the Granger causality. Fortunately, price discovery only cares the statistical causality from which algorithmic trading can benefit. Second, it relies heavily on the sampling rate (Gong et al., 2015, 2017). Chances are that there exist no causal relations between A and B when sampled every 30 minutes while A is the Granger cause of B when sampled every 15 minutes.

2. High Frequency Causality in Financial Markets

2.1. Problem: A Dilemma

Even though the high frequency data is available, high frequency causality is almost unable to detect because high frequency data are contaminated by market microstructure effect, such as bid-ask spread, liquidity ratios, turnover and asymmetric information, leading extremely high observed volatility and low signal-to-noise ratio (Hasbrouck and Seppi, 2001; O'Hara, 2003; Aït-Sahalia et al., 2005). Aït-Sahalia et al. (2005) also modeled it directly: since the log-return (difference between two adjacent log-price) is believed to obey Brownian motion, its volatility should be linear to sampling rate while from observed data it exhibits a U-shape caused by market microstructure noise whose volatility is set to be time independent. The authors also point out for empirical researchers in high frequency financial markets, most of them use lower resolution data whose frequency ranges from 5-min (e.g. Andersen et al., 2001) to 1-hour (e.g. Chen et al., 2012). While the aggregation or subsampling process does resolve the market microstructure noise, as pointed out before, this process may mask high frequency causality. For high frequency trading algorithm, low frequency causality is likely to be useless. Plus, although O'Hara (2003) urges that information-based microstructure models should include microstructure effects, it is not practical to allow all kinds of influentials to the regression model. Some of the sources are not able to assess directly, either. Through more delicate modeling is not a good solution.

2.2. A Possible Solution: EM algorithm

Fortunately, Gong et al. (2015, 2017) offer an EM algorithm focusing on recovering temporal causality from subsampled or aggregated data respectively. Although the original problems in these two paper is the inaccessibility of high frequency data, it can also be applied to dissolve the market microstructure noise from high frequency observations. Since the low frequency data is irrelevant to the high frequency data, thus microstructure noise, the EM algorithm acts as a latent estimation from artificially subsampled or aggregated time series to high frequency data generating process. The basic idea behind EM algorithm is to make use of the information contained in residue terms (noise terms). It assumes non-Gaussian noise terms while in financial markets, indicators like log-price always exhibit fat-tail (Geweke, 1994; Gallant et al., 1997). In VAR model, econometrics often use t-distribution embedded in a GARCH specification rather than Gaussian distribution to describe the noise terms (Glosten et al., 1993).

If we aggregate or subsample the high frequency data to an alleged microstructure noise-free level (e.g. hourly) and recover higher frequency data generating process from it, the noise term will no longer preserve microstructure noise and Granger causality test (Block Exogeneity Wald test) could be done to reveal causality.

3. Modeling Foreign Exchange Markets

Price discovery in foreign exchange (FX) markets is a quite controversial issue. Some authors argue that information transmission explains most of short run asset price (e.g. Evans and Lyons, 2002; Andersen et al., 2003) while others believe price discovery relies more on liquidity (e.g. Roll, 1984; Grossman and Miller, 1988). When it comes to measure each factors, different variables are defined by different research.

Order flow, also called order imbalance, is defined as the difference between buyer-initiated trade volume and seller-initiated trade volume within a given period. Lee and Ready (1991) propose an algorithm to determine whether a transaction is buyer-initiated or seller-initiated in stock market (i.e. the transaction direction). In traditional order flow model, order imbalance is calculated as the amount traded at ask price minus the amount traded at bid price. However, Cont et al. (2014) propose another formulation referred to the mechanism of FX markets. Tick data in FX markets, which offer investors a snapshot of the situation of the market with a short disclosure interval, contain four primitive variables: ask price, bid price, ask volume and bid volume. By using these four variables and there lags, an alternative measurement of order flow is drawn (this measurement is also called volume order imbalance).

Liquidity, from another perspective, indicates the matching of buyers and sellers (Chen et al., 2012). Volume is believed to be associated inextricably to liquidity (Benston and Hagerman, 1974; Stoll, 1978) and quoted or realized bid-ask spread are another type of liquidity measurement (Amihud and Mendelson, 1986; Stoll, 1989).

O'Hara (2003) introduces the mediation process to illustrate that these two aspects are correlated but not the same, indicating that liquidity measures should also be added to order flow model in FX markets. So in this paper, a VAR model is formed combining order flow model and liquidity as part of the microstructure effect measurement.

One thing worth mentioning is that variables within one regression model should either all be subsampled data (i.e. snapshots) or all be aggregated data (i.e. flows). When sampling rate changes, a regression model with both snapshots and flows in it might tell a rather different story and EM algorithm is also infeasible. So we form two regression models corresponding to these two types of data as follow:

$$\begin{bmatrix} logP_t \\ OI_t^S \\ RS_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} logP_{t-1} \\ OI_{t-1}^S \\ RS_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} \Delta logP_t \\ OI_t^A \\ Vol_t \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \Delta logP_{t-1} \\ OI_{t-1}^A \\ Vol_{t-1} \end{bmatrix} + \begin{bmatrix} \varphi_{1t} \\ \varphi_{2t} \\ \varphi_{3t} \end{bmatrix} \quad (2)$$

In regression (1), P_t denotes the midpoint of bid and ask price.RS_t is defined as realized spread calculated as $(P_t^A - P_T^B)/(P_t^A + P_T^B)$. OI_t^S is the subsampled order imbalance calculated through formula in Cont et al. (2014):

$$OI_t^S = \delta V_t^B - \delta V_t^A$$

where

$$\delta V_t^B = \begin{cases} 0 & P_t^B < P_{t-1}^B \\ V_t^B - V_{t-1}^B & P_t^B = P_{t-1}^B \\ V_t^B & P_t^B > P_{t-1}^B \end{cases}, \delta V_t^A = \begin{cases} V_t^A & P_t^A < P_{t-1}^A \\ V_t^A - V_{t-1}^A & P_t^A = P_{t-1}^A \\ 0 & P_t^A > P_{t-1}^A \end{cases}$$

In regression (2), $\Delta logP_t$ is the log-return from t-1 to t. OI_t^A is gap between buyer-initiated transactions minus seller-initiated ones from t-1 to t whose classification is gained by algorithm in Lee and Ready (1991). Vol_t is the trade volume from t-1 to t.

It is easy to see that regression (1) is estimated through subsampled data while regression (2) is estimated through aggregated data. Robustness would be tested through the consistency of the two regression models. Before that, the consistency between different measurements are ascertained by correlation analysis. To show that EM algorithm does eliminate the microstructure noise, the recovered standard error of noise terms are evaluated. In the meantime, mean-square error (MSE) between high frequency coefficient matrix and recovered one is expected to be small.

4. Miscellaneous

4.1. FX Markets as a Whole or Respectively

Easley et al. (1996) point out the risk of information-based trading is lower for active securities than for infrequently traded securities. In FX markets, those with a high trading density act very different from those with low trading density (Chen et al., 2012). The bilateral political factors are also different among

FX markets. It is believed that there are many reasons to treat them respectively.

With the development of transfer learning, it makes the combination of different FX markets to be possible even if they have not much in common. For instance, Rashid et al. (2017) propose a method for localizing facial keypoints on animals by transferring knowledge gained from human faces. Differences between animal faces and human faces are salient while some keypoints are the same (if there is nothing in common, other knowledge is useless). In financial markets, the fundamental incentives of investors are basically alike. The knowledge from one financial markets could be applied to trading in another one. Through transfer learning model such as convolutional neural network (CNN), the underlying causality across different FX markets could be revealed better.

4.2. Mulitvariate GRACH: VAR-MGARCH Specification

GARCH effect is common in economic world. The original EM algorithm does not allow a GARCH effect in its VAR specification. When applied to financial markets, the strict assumption on noise terms should be loosen to include a GARCH effect in VAR model. Since EM estimators and GARCH estimators are all MLEs, it is possible to allow a modification on original algorithm. Before evaluating empirical data, tests for GARCH effect should be done prior to the selection of a VAR or a VAR-MGARCH specification.

4.3. Limitation of EM Algorithm

Some practical issues are listed below when EM algorithm is evaluated through empirical data:

- **Time Complexity**. Although Gong et al. (2015, 2017) propose a mean field approximation to speed up, the time complexity is also high in high frequency trading data.

- **Spacial Complexity**. Since the estimations of the distribution characteristics are through Gaussian mixture model, when data grows larger and sampling rate gets lower, the spacial complexity increases exponentially. That is to say, it is hardly practical to recover 1-minute causality from daily data.

References

Aït-Sahalia, Y., Mykland, P.A., Zhang, L., 2005. How often to sample a continuous-time process in the presence of market microstructure noise. The Review of Financial Studies 18, 351–416.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics 17, 223–249.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001. The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96, 42–55.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Vega, C., 2003. Micro effects of macro announcements: Real-time price discovery in foreign exchange. American Economic Review 93, 38–62.

Angrist, J.D., Pischke, J.S., 2008. Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Benston, G.J., Hagerman, R.L., 1974. Determinants of bid-asked spreads in the over-the-counter market. Journal of Financial Economics 1, 353–364.

Bollen, J., Mao, H., Zeng, X., 2011. Twitter mood predicts the stock market. Journal of Computational Science 2, 1–8.

Chen, S., Chien, C.C., Chang, M.J., 2012. Order flow, bid-ask spread and trading density in foreign exchange markets. Journal of Banking & Finance 36, 597–612.

Christophe, S.E., Ferri, M.G., Angel, J.J., 2004. Short-selling prior to earnings announcements. The Journal of Finance 59, 1845–1876.

Christophe, S.E., Ferri, M.G., Hsieh, J., 2010. Informed trading before analyst downgrades: Evidence from short sellers. Journal of Financial Economics 95, 85–106.

Cont, R., Kukanov, A., Stoikov, S., 2014. The price impact of order book events. Journal of Financial Econometrics 12, 47–88.

Dennis, P., Mayhew, S., 2002. Risk-neutral skewness: Evidence from stock options. Journal of Financial and Quantitative Analysis 37, 471–493.

Easley, D., Kiefer, N.M., O'hara, M., Paperman, J.B., 1996. Liquidity, information, and infrequently traded stocks. The Journal of Finance 51, 1405–1436.

Evans, M.D., Lyons, R.K., 2002. Order flow and exchange rate dynamics. Journal of Political Economy 110, 170–180.

Gallant, A.R., Hsieh, D., Tauchen, G., 1997. Estimation of stochastic volatility models with diagnostics. Journal of Econometrics 81, 159–192.

Geweke, J., 1994. Bayesian analysis of stochastic volatility models: Comment. Journal of Business & Economic Statistics 12, 397–99.

Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance 48, 1779–1801.

Gong, M., Zhang, K., Schoelkopf, B., Tao, D., Geiger, P., 2015. Discovering temporal causal relations from subsampled data, in: International Conference on Machine Learning, pp. 1898–1906.

Gong, M., Zhang, K., Schölkopf, B., Glymour, C., Tao, D., 2017. Causal discovery from temporally aggregated time series, in: Uncertainty in artificial intelligence: proceedings of the... conference. Conference on Uncertainty in Artificial Intelligence, NIH Public Access.

Gow, I.D., Larcker, D.F., Reiss, P.C., 2016. Causal inference in accounting research. Journal of Accounting Research 54, 477–523.

Grossman, S.J., Miller, M.H., 1988. Liquidity and market structure. The Journal of Finance 43, 617–633.

Hahn, J., Todd, P., Van der Klaauw, W., 2001. Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica 69, 201–209.

Hasbrouck, J., Seppi, D.J., 2001. Common factors in prices, order flows, and liquidity. Journal of Financial Economics 59, 383–411.

Lashgari, M., 2000. The role of ted spread and confidence index in explaining the behavior of stock prices. American Business Review 18, 9.

Lee, C.M., Ready, M.J., 1991. Inferring trade direction from intraday data. The Journal of Finance 46, 733–746.

O'Hara, M., 2003. Presidential address: Liquidity and price discovery. The Journal of Finance 58, 1335–1354.

Rashid, M., Gu, X., Lee, Y.J., 2017. Interspecies knowledge transfer for facial keypoint detection, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. The Journal of Finance 39, 1127–1139.

Stoll, H.R., 1978. The pricing of security dealer services: An empirical study of nasdaq stocks. The Journal of Finance 33, 1153–1172.

Stoll, H.R., 1989. Inferring the components of the bid-ask spread: Theory and empirical tests. The Journal of Finance 44, 115–134.