# Lead-lag analysis of non-synchronously observed time series with R

Tokyo Metropolitan University Yuta Koike

A lead-lag relationship is a time-lagged correlation structure of two time series wherein one is correlated to the other with a delay. Investigation of such relationships has a long history in financial econometrics in connection with studies of market efficiency and price discovery (cf. [1, 3, 7]). However, it has also been well-recognized since long ago that the non-synchronicity of observation times of two time series causes a serious bias in estimating lead-lag relationships (cf. [8]). This is problematic especially in high-frequency financial data because their observation times are generically stochastic and irregular and thus non-synchronous observations naturally arise. To overcome this issue, the recent work of [5] has introduced a novel approach to compute the cross-covariances of the returns of two non-synchronously observed time series, which is based on the method of [4] to deal with non-synchronous observations in covariance estimation. It has also proposed a simple statistic to measure their lead-lag relationship. The methodology is not only applicable to high-frequency financial data (cf. [6]) but also applicable to SNS data (cf. [2]), so it could provide a useful tool for lead-lag analysis of time series to empirical researchers in any areas.

R package **yuima** provides systematic functions to conveniently apply the methodology of [5] to real time series. The aim of this talk is to present them to empirical researchers as well as to show what we can really do in **yuima**. As an illustration we will demonstrate its application to *Google Trends* data.

## References

[1] Agmon, T. (1972). The relations among equity markets: A study of share price co-movements in the United States, United Kingdom, Germany and Japan. *Journal of Finance* **27**, 839–855.

[2] Ceron, A., Curini, L. & Iacus, S. M. (2016). First- and second-level agenda setting in the Twittersphere: An application to the Italian political debate. *Journal of Information Technology & Politics* **13**, 159–174.

[3] Garbade, K. D. & Silber, W. L. (1979). Dominant and satellite markets: A study of dually-traded securities. *The Review of Economics and Statistics* **61**, 455–460.

[4] Hayashi, T. & Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli* **11**, 359–379.

[5] Hoffmann, M., Rosenbaum, M. & Yoshida, N. (2013). Estimation of the lead-lag parameter from non-synchronous data. *Bernoulli* **19**, 426–461.

[6] Huth, N. & Abergel, F. (2014). High frequency lead/lag relationships — empirical facts. *Journal of Empirical Finance* **26**, 41–58.

[7] Kawaller, I. G., Koch, P. D. & Koch, T. W. (1987). The temporal price relationship between S&P 500 futures and the S&P 500 index. *Journal of Finance* **42**, 1309–1329.

[8] Lo, A. W. & MacKinlay, A. C. (1990). An econometric analysis of nonsynchronous trading. *J. Econometrics* **45**, 181–211.