Journal of Banking & Finance 37 (2013) 2733–2749

Contents lists available at SciVerse ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Identifying the interaction between stock market returns and trading flows of investor types: Looking into the day using daily data

Numan Ülkü [a,*], Enzo Weber [b]

[a] American University in Bulgaria, Bulgaria
[b] University of Regensburg, IAB Nuremberg and IOS Regensburg, Germany

ARTICLE INFO

Article history:
Received 6 December 2012
Accepted 23 March 2013
Available online 15 April 2013

JEL classification:
G11
G14
G15
C32

Keywords:
The interaction between trading flows and returns
Identification
Structural conditional correlation
Investor types
Feedback trading behavior

ABSTRACT

This paper introduces a new method for identifying the simultaneity between returns and trading flows. The proposed method enables us to identify the intraday interaction using daily data, and provides measures of the information content of trading flows, and their instantaneous response to public information and information revealed by market prices. Applying this method to daily data on investor types from the Korea Stock Exchange, we find significant intraday bi-directional interaction between flows and returns and their latent common drivers, altering some of the results of the previous literature based on Cholesky assumptions. Thus, we obtain a number of new insights concerning the behavior of investor types.

1. Introduction

The interpretation of the contemporaneous association between trading flows and returns at the daily frequency has been a notorious problem in the microstructure literature. Use of tick data can be of partial help, however tick data on stock market trading flows are scarce and in most cases only privately available for only short sample periods that risk not being representative of the full data.[1] On the other hand, some stock exchanges regularly publish long periods of trading flows data at the daily frequency.

The contemporaneous correlation between net trading flows and returns may reflect three possibilities: price impact (contemporaneous, and in the form of intra-period forecast ability), intra-period feedback trading, and common factor influence driving both flows and returns simultaneously. Classical models of price formation in microstructure theory (see Hasbrouck, 1991) assume that public information arrivals are fully and instantaneously incorporated only by return innovations, excluding the possibility of accompanying flow (or trade) innovations. They further assume that feedback trading can only appear at lags. Thus, the contemporaneous positive correlation between flows and returns has been named *price impact*. It is usually interpreted as private information content of trading if permanent or attributed to price pressure if reversed subsequently. However, regarding empirical work, two facts render this decomposition and its interpretation questionable: First, both flow and return innovations may be contemporaneously driven by common latent effects such as public information arrivals (see Green, 2004; Love and Payne, 2008; Riordan et al., 2013). This is especially relevant, even under tick data, for most emerging stock exchanges operating electronic order book systems with irreversible limit orders, where prices may only change by trading and not by dealer's quote revision.[2] Thus, the positive correlation between flows and returns may result from common factor influence, such as public information arrivals. Whenever a time series of all public information arrivals is not available to the researcher, public information needs to be treated as a latent common factor. Second, intra-period feedback

* Corresponding author. Tel.: +359 73 888493.
E-mail address: nulku@aubg.bg (N. Ülkü).

[1] Further, even when tick data are available, in many cases they need to be aggregated into time intervals for econometric analysis (e.g., Cerrato et al., 2011).

[2] North American and European stock exchanges have also been migrating towards electronic limit order book systems or a hybrid, hence this issue is becoming pervasive.

0378-4266/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jbankfin.2013.03.021

trading may occur when investors observe price changes and react within the same period. In a model using daily data (more generally any data which are less frequent than tick data), this would induce a contemporaneous spillover from returns to flows. The standard treatment in the microstructure literature by ordering flows (net trading) before returns to enable the contemporaneous identification in a recursive (Cholesky-type) structural vector autoregression (SVAR) system, following Hasbrouck (1991), imposes causality to run from flows to returns. For the above reasons, the standard approach is questionable at best and may potentially lead to misspecification.

Most of the extant empirical literature has followed the standard assumption that only flows can cause returns despite its shortcomings (e.g., Froot et al., 2001; Griffin et al., 2004; Richards, 2005; Green, 2004), usually verbally arguing about the reasonableness of the assumption. The same presumption underlies a large literature on foreign exchange market microstructure: the order flow approach, pioneered by Evans and Lyons (2002), uses order flow as a regressor in the exchange rate change equation [see Cerrato et al. (2011) for a recent example]. Some papers present the results under the two opposite extreme ordering assumptions (e.g., Gradojevic and Neely, 2008; Kim et al., 2009). We are aware of only two papers which explicitly deal with the issue of simultaneity between trades (investor net flows) and returns. Danielsson and Love (2006) offer a solution based on instrumental variables approach on a foreign exchange market example. Sias et al. (2006) obtain a term structure of correlations in their study of US institutional investors. Some shortcomings of these approaches are discussed in the next section. Two other related papers, Love and Payne (2008) and Evans and Lyons (2008), employ the idea of identification by heteroscedasticity, which also underlies the approach introduced in the current paper, focusing on public information arrivals as a common driver of flows and returns.

In this paper, we propose a frontier methodology that exploits time variation in the volatility of shocks to achieve identification (see e.g. Sentana and Fiorentini, 2001 or Rigobon, 2003). Specifically, we employ the structural conditional correlation (SCC) model of Weber (2010) to identify the contemporaneous return-flow interaction at the daily frequency. Importantly, the contributions of all three possible sources of the contemporaneous correlation can be estimated without zero-restrictions. SCC method permits to identify from data the information content of a particular investor group's trading and its instantaneous response to public information and to information conveyed by market prices.

We implement this approach on Korea Stock Exchange (KSE) which offers (publicly available) daily data on trading flows of a rich breakdown of investor types for a long sample period. Employing this methodology, we show that the standard assumption, that flows cause returns but not vice versa, is not justified, as there exist significant bi-directional intraday interactions between flows and returns and their common latent drivers. In some cases, freely-estimated contemporaneous interaction parameters alter the impulse responses based on (inaccurate) Cholesky assumptions. Thus, we obtain new insights concerning the trading of various investor types. In particular, despite the strong negative daily contemporaneous correlation between individuals' net trading and market returns, our results suggest that individuals are intraday positive feedback traders. On the other hand, our results also confirm many of the findings in the previous literature/under Cholesky assumption, such as significant information content of institutional and foreign flows and positive feedback trading by foreign investors. We also obtain new results from the more-detailed break-down of the broad group of institutions. For example, merchants' trading has both significant information content and forecast ability, and they are intraday negative feedback traders, possibly reflecting their liquidity supplier role. Private funds are positive feedback traders, however their trading has little information content, which seems to be an indication of technical trading strategies they employ. Many institutional investor types (pension funds, banks) display negative feedback trading, but unlike individuals, their trading is not negatively associated with future returns.

In Section 2, we introduce the SCC methodology, describe the data and review the literature on the interaction of trading by investor types with stock returns, which provides a link between the methodological contribution of the current study and issues discussed in the previous literature. Results are presented in Section 3. Section 4 concludes.

2. Methodology, related literature and data

2.1. Standard methodology

As the relation between returns and trading flows involves a bi-directional interaction with lagged responses, vector autoregression (VAR) methodology has been a standard in this line of the literature, following Hasbrouck (1991). The fact that the bulk of the relationship is observed within the contemporaneous period requires a SVAR specification with contemporaneous identification assumptions that need to be justifiable. This leaves us with the following SVAR model where the net trading flows (purchases minus sales) of a particular investor type i during period t is denoted as $F_{i,t}$ and the returns of the market as R_t:

$$\begin{bmatrix} 1 & a_{12} \\ a_{21} & 1 \end{bmatrix} \begin{bmatrix} F_{it} \\ R_t \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} + \begin{bmatrix} b_{11}(L) & b_{12}(L) \\ b_{21}(L) & b_{22}(L) \end{bmatrix} \begin{bmatrix} F_{it} \\ R_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \quad (1)$$

k's are constants, $b(L)$'s are polynomials of order p in lag coefficients (beginning with exponent one, i.e. excluding the contemporaneous value), and ε's are SVAR innovations. The left-hand-side matrix of contemporaneous impacts (A) has elements normalized to one on its main diagonal. We set the lag order $p = 5$ which ensures elimination of serial correlation in residuals across all investor types in our sample.

The contemporaneous relation among the variables in an estimated standard VAR equation system is hidden in the covariance matrix of reduced-form disturbances (i.e., VAR residuals) and not uniquely identifiable by standard methods. The typical approach to tackle this problem is to impose identifying restrictions in the form of a Cholesky ordering. Such restrictions, however, amount to allocating the contemporaneous correlation to causation from one variable to another while excluding the reverse causality and any latent common driver. Unless the imposed restrictions are based on sound theory and accurate for the data under consideration, this approach may lead to misleading interpretations.

In the case of return-flow interaction, there is no theoretical consensus on the appropriate restrictions. Using tick data, Hasbrouck's (1991) model, which allows F to affect R contemporaneously but restricts the reverse coefficient to zero, can be legitimate under a dealer system without frictions. However, with daily or less frequent data, feedback trading models that attribute the contemporaneous correlation between flows and returns to intra-period feedback trading (e.g., Brennan and Cao, 1997) imply the opposite restriction. Alternating the two extreme assumptions will not solve the problem, as allowing both effects simultaneously, along with latent common drivers, may sharply change the conclusion.

Hasbrouck's model assumes that public information surprises are incorporated by return innovations that are not accompanied by flows, thus ignores common drivers of returns and flows (i.e., assumes that shocks from the structural return equation and reduced-form flow equation are uncorrelated). Under a