Journal of the Korean Statistical Society 44 (2015) 295–311

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

Financial interpretation of herd behavior index and its statistical estimation

Woojoo Lee [a], Jae Youn Ahn [b,*]

[a] Department of Statistics, Inha University, 235 Yonghyun-Dong, Nam-Gu, Incheon 402-751, Republic of Korea

[b] Department of Statistics, Ewha Womans University, 11-1 Daehyun-Dong, Seodaemun-Gu, Seoul 120-750, Republic of Korea

ARTICLE INFO

Article history:
Received 4 February 2014
Accepted 30 September 2014
Available online 25 October 2014

AMS 2000 subject classifications:
62P05
91B30

Keywords:
Countermonotonicity
Comonotonicity
Herd behavior

ABSTRACT

Herd behavior received increasing attention as the key to understanding financial crises. Recently, Dhaene et al. (2012) proposed the herd behavior index (HIX) to measure the degree of the comonotonic movement of stock prices. Choi et al. (2013) introduced the revised version of HIX (RHIX) and illustrated why RHIX should be preferred to HIX in comparing herd behaviors based on simple toy models, but failed to offer any sufficient justification. The present paper investigates three aspects of RHIX. First, RHIX is explained as a useful tool to compare herd behaviors from different groups. For this, a new alternative representation of RHIX is provided. Second, we investigate the statistical estimation of RHIX. In particular, the realized version of RHIX is calculated using tick-by-tick stock prices and the asymptotics of bootstrap equivalence are provided to estimate the confidence interval of RHIX. Third, to extend application scope of RHIX, RHIX is employed in a clustering analysis.

1. Introduction

Herd behavior has received increasing attention as the key to understanding financial crises. Recently, Dhaene, Linders, Schoutens, and Vyncke (2012) propose the herd behavior index (HIX) to measure the degree of the comonotonic movement in a stock market. Although HIX is an intuitively appealing measure, Choi, Kim, Lee, and Ahn (2013) point out that it may not be a suitable tool for comparing the herd behavior of different sources because of a high level of sensitivity for marginal distributions. To reduce the effect of marginal distributions, Choi et al. (2013) introduce the revised version of HIX (RHIX) and illustrate why RHIX should be preferred to HIX in terms of comparing herd behaviors based on simple toy models, but their arguments offer no sufficient justification. The present paper investigates three aspects of RHIX. First RHIX is explained as a useful tool for comparing the herd behaviors of different groups. For this, a new alternative representation of RHIX is provided as an average of normalized correlations. Here the normalized correlation is defined as the ratio of the correlation of random variables to that under the comonotonic assumption. Second, we investigate the statistical estimation problem of RHIX. There are two popular estimation methods: the implied method and the realized method. As described in Dhaene et al. (2012), vanilla options can be used to calculate an implied version of HIX, and essentially, the same technique can be used to calculate an implied version of RHIX (Choi et al., 2013). Another method is to calculate a realized version of RHIX by using tick-by-tick stock prices. This paper focuses on the latter and provides asymptotics of bootstrap equivalence for

* Corresponding author.
E-mail addresses: lwj221@gmail.com (W. Lee), jy5152@gmail.com, jaeyahn@ewha.ac.kr (J.Y. Ahn).

http://dx.doi.org/10.1016/j.jkss.2014.09.005
1226-3192/

estimating the confidence interval of RHIX. Third, to extend the application scope of RHIX, RHIX is employed in a clustering analysis. This approach facilitates the grouping of several stock markets into a few clusters according to their comovement behaviors.

In the literature, herd behaviors are defined in various ways. Sushil and Sunil (2001) considered it as the comovement of group members without any reasonable justification or full appreciation of circumstances. Christie and Huang (1995) observe that returns on US equities show relatively low dispersion when there is herding, and propose the use of dispersion as a measure of herding. Chang, Cheng, and Khorana (2000), Chiang and Zheng (2010) and Prosad, Sujata, and Jhumur (2012) use essentially the same idea to demonstrate herd behavior in the market. On the other hand, Dhaene et al. (2012) adopt the concept of comonotonicity and define the level of the herd behavior based on it. As explained in Section 2, comonotonicity describes the perfect positive dependence between components of a random vector: see Cheung (2008), Dhaene et al. (2012) and Nam, Tang, and Yang (2011) for the definition and application of comonotonicity. Dhaene et al. (2012) define HIX as the ratio of the variance of the market index to that under the comonotonic assumption. Note that the variance of a market index is maximized under the comonotonic assumption. Therefore, HIX can be interpreted as an indicator of how close stock prices in a market are to perfect herd behavior. In particular, if HIX is calculated from vanilla option prices, then Dhaene et al. (2012) interpret it as the expected degree of herd behavior or comovement in the future.

Choi et al. (2013) compare herd behaviors in global stock markets and make intercontinental comparisons of this herd behavior. They observe that HIX depends heavily on the marginal variance of each stock market, and that HIX can give spurious signals of herd behaviors depending on the marginal distributions. As an alternative measure, they introduce RHIX as the ratio of the average correlation of stock prices to that of stock prices under the comonotonic assumption. Dhaene, Linders, Schoutens, and Vyncke (2014) independently propose the same herd behavior index, although from a slightly different perspective. Through some toy examples, Choi et al. (2013) explain that RHIX is less sensitive to the marginal distribution than HIX and thus conclude that RHIX is more reasonable than HIX for comparing herd behaviors. However, because the preference in Choi et al. (2013) does not have sufficient generality, there is a need for more convincing evidence. This issue will be touched in Section 3.2.

In particular, an implied version of RHIX can be calculated using vanilla option prices, as specified in Choi et al. (2013). The implied method for calculating the risk measures, including HIX, is detailed in Dhaene et al. (2012) and Linders and Schoutens (2014). However, as specified in Choi et al. (2013), use of the implied method for the calculation of RHIX can be limited in intercontinental comparisons mainly due to the absence of necessary option prices. Hence the historical stock indexes using the high frequency data will directly be used for the following analysis. However, the statistical estimation problem for the realized version of RHIX has yet to be throughly addressed, and therefore it is critical to quantify the uncertainty of RHIX estimates based on easy computing skills. This paper addresses this issue in a clear manner.

The rest of this paper is organized as follows: Section 2 provides the relevant definitions and conventions. Section 3 provides a brief overview of HIX and RHIX and explains why RHIX can be used to compare herd behaviors through its alternative representation. Section 4 considers the realized version of RHIX based on high-frequency data and shows its weak convergence. Section 5 provides some simulation results for the finite-sample property. For the practical application, Section 6 provides a clustering analysis based on RHIX and Section 7 provides the behavior of RHIX over a short interval, and concludes with some suggestions for future research.

2. Conventions and preliminaries

2.1. Conventions

A set $[a, b] \times [a, b] \cdots \times [a, b] \subseteq \mathbb{R}^d$ is denoted by $[a, b]^d$. Denote $\overrightarrow{x} = (x_1, x_2, \ldots, x_d)$ and $\overrightarrow{y} = (y_1, y_2, \ldots, y_d)$ as constant vectors in $\mathbb{R}^d$. This paper assumes that all random variables are defined in the common probability space $(\Omega, \mathcal{F}, P)$. Let $\overrightarrow{X} = (X_1, X_2, \ldots, X_d)$ and $\overrightarrow{Y} = (Y_1, Y_2, \ldots, Y_d)$ be d-dimensional random vectors and H and H^* be their cumulative distribution function defined as

$$H(\overrightarrow{x}) = P(X_1 \le x_1, \ldots, X_d \le x_d) \quad \text{and} \quad H^*(\overrightarrow{x}) = P(Y_1 \le x_1, \ldots, Y_d \le x_d) \quad \text{for } \overrightarrow{x} \in \mathbb{R}^d.$$

Marginal distributions of X_i and Y_i are $F_i(x) = P(X_i \le x)$ and $G_i(x) = P(Y_i \le x)$, respectively for $x \in \mathbb{R}$. Define $\mathcal{R}(F_1, \ldots, F_d)$ to be the Fréchet space of random vectors $\overrightarrow{X}$ with the marginal distribution $F_1, \ldots, F_d$. In addition, let $\overrightarrow{X} \in \mathcal{R}_2(F_1, \ldots, F_d)$ or equivalently $H \in \mathcal{R}_2$, when the variances of marginal distributions are strictly greater than 0 and finite. Hereafter, consider only $\overrightarrow{X} \in \mathcal{R}_2(F_1, \ldots, F_d)$, unless otherwise specified.

This paper assumes marginal distributions are continuous otherwise specified. According to Sklar (1959), given $H \in \mathcal{R}_2(F_1, \ldots, F_d)$, there exists a unique function $C : [0, 1]^d \to [0, 1]$ satisfying

$$H(\overrightarrow{x}) = C(F_1(x_1), \ldots, F_d(x_d)).$$

The function C is called a copula, which is also a distribution function on $[0, 1]^d$. Further information on copulas can be found, for example, Cherubini, Luciano, and Vecchiato (2004) and Nelsen (2006).

For any given $H \in \mathcal{R}(F_1, \ldots, F_d)$,

$$W(F_1(x_1), \ldots, F_d(x_d)) \le H(\overrightarrow{x}) \le M(F_1(x_1), \ldots, F_d(x_d)), \quad \text{for all } \overrightarrow{x} \in \mathbb{R}^d,$$