Physica A 392 (2013) 6004–6014

Contents lists available at ScienceDirect

Physica A

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

Distribution characteristics of stock market liquidity

Jiawen Luo [a,b], Langnan Chen [a,b,*], Hao Liu [c]

[a] *Lingnan College (University), Sun Yat-Sen University, Guangzhou 510275, China*
[b] *Institute for Economics, Sun Yat-Sen University, Guangzhou 510275, China*
[c] *Guangfa Bank, Guangzhou 510000, China*

HIGHLIGHTS

- This paper investigates the distribution characteristics of stock market liquidity.
- This paper utilizes the GAMLSS model and high-frequency data from Chinese stock markets.
- The mean and variance of stock market liquidity move in the opposite directions.
- The stock market liquidity normally exhibits a positive skewness.
- The stock market liquidity exhibits a normal distribution at a low level of stock market index.

ARTICLE INFO

Article history:
Received 27 March 2013
Received in revised form 15 July 2013
Available online 29 July 2013

Keywords:
Liquidity
GAMLSS model
BCPE distribution
Non-parameter cubic splines regression
GAIC

ABSTRACT

We examine the distribution characteristics of stock market liquidity by employing the generalized additive models for location, scale and shape (GAMLSS) model and three-minute frequency data from Chinese stock markets. We find that the BCPE distribution within the GAMLSS framework fits the distributions of stock market liquidity well with the diagnosis test. We also find that the stock market index exhibits a significant impact on the distributions of stock market liquidity. The stock market liquidity usually exhibits a positive skewness, but a normal distribution at a low level of stock market index and a high-peak and fat-tail shape at a high level of stock market index.

1. Introduction

Liquidity can be statistically defined as market width, depth and immediacy, and dynamically defined as market power and resiliency. Among all the liquidity proxies, market depth and width are commonly used in the existing literature. Brockman and Chung [1] investigate the liquidity distribution by employing the high frequency data from the Hong Kong Stock Exchange and find that the intraweek depth follows a reverse *U* pattern. Similarly, Plerou et al. [2] reveal that the distribution of the stock price spread has a fat tailed feature and follows a power law. Yan et al. [3] find that the probability distribution of relative changes in returns satisfies a power law form. Ponzi et al. [4] examine the relaxation dynamics of the bid–ask spread and of the midprice after an abrupt change of the spread in a double auction financial market. Mu et al. [5] observe the distributions of trade sizes and trading volumes for 22 Chinese stocks and find that the size distributions for individual stocks has jumps and the probability density functions exhibit power-law tails for large volumes. Kagraoka [6] examines the liquidity processes and reveals the leptokurtic properties of the return distributions. Toth et al. [7] derive a dynamical model of market liquidity and find the average supply/demand profile is V-shaped and disappears close to the current price. Feng

* Corresponding author at: Lingnan College (University), Sun Yat-Sen University, Guangzhou 510275, China. Tel.: +86 13570115839.
E-mail address: lnscln@mail.sysu.edu.cn (L. Chen).

0378-4371/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.physa.2013.07.046

et al. [8] develop an agent-based model with stochastic process to show the "fat" tail property in price return distributions. The above studies confirm that the liquidity does not follow the traditional exponential distribution families.

In this paper we employ a new approach or a class of GAMLSS models to investigate the distribution of market liquidity, which proves to be a convenient option when the response variable does not follow the exponential family distributions or when the shape of the response variable's distribution is clearly determined by covariates. Within the GAMLSS framework, we can expand distribution parameters up to four dimensions at a time, which improves the flexibility of analysis. The application of a non-parametric model within the GAMLSS framework avoids modeling with parameters subjectively. In this paper we also show that the GAMLSS framework can compete with traditional linear regression methods for this high-frequency market liquidity data set in terms of forecast accuracy.

The remainder of this paper is organized as follows: Section 2 presents the GAMLSS model proposed by Rigby & Stasinopoulos [9]. Section 3 describes the statistical characteristics of the data. Section 4 discusses the empirical results. Section 5 concludes the paper.

2. Model and estimation

2.1. The GAMLSS model

The "additive models for location, scale and shape" (GAMLSS) was originally developed by Rigby and Stasinopoulos [9]. For a GAMLSS model, the conditional distribution of the response variable is not required provided that a set of covariates is a member of the exponential family. Nevertheless, wide ranges of discrete, continuous and mixed discrete–continuous distribution are possible, for example, the distributions based on Box–Cox transformations such as the Box–Cox-t (BCT) distribution, the Box–Cox power exponential (BCPE) distribution, the Student t-family (TF) distribution, the generalized gamma family including generalized gamma (GG) distribution, generalized inverse gamma (GIG) distribution, etc. A detailed list of optional distributions for GAMLSS is presented in [10].

As a methodology for financial economics, the GAMLSS framework allows the distribution parameters of the response variable to be modeled as linear/nonlinear or smooth functions of the explanatory variables. It also allows the simultaneous modeling of several parameters that document the response distribution using the parametric and/or non-parametric functions. In addition, the GAMLSS model has the advantage of extra flexibility as different additive terms such as smoothing splines and random effects can be included in the regression function for the parameters that document the distribution.

2.2. Assumptions

A typical GAMLSS model assumes an independent observations set $y = (y_1, y_2, y_3 \ldots\ldots y_n)^T$, each y_i $(i = 1, 2 \ldots n)$ having conditional probability (density) function $f(y_i|\theta^i)$, where $\theta^i = \left(\theta_{i1}, \theta_{i2} \ldots \theta_{ip}\right)^T$ is a vector of p parameters associated with the explanatory variables related to random effects. When the covariates are stochastic, $f(y_i|\theta^i)$ is considered to be conditional on their values. The parameters can be derived by regressing the explanatory variables through the known monotonic link functions $g_k()$.

An addicted model of GAMLSS is as follows:

$$g_k(\theta_k) = \eta_k = X_k\beta_k + \sum_{j=1}^{J_k} Z_{jk}\gamma_{jk} \tag{1}$$

where θ_k and η_k are $n \times 1$-dimension vectors, $\beta_k = \left(\beta_{1k}, \beta_{2k}, \ldots, \beta_{J'_k k}\right)^T$ is a vector of length J'_k, X_k and Z_{jk} are fixed design covariance matrices of length $n \times J'_k$, and $n \times q_{jk}$, and γ_{jk} is a q_{jk}-dimension vector of random variables.

In most cases, $\theta^i = \left(\theta_{i1}, \theta_{i2}, \ldots, \theta_{ip}\right)^T$ is a four-dimension vector of distribution parameters, i.e. $\theta^i = (\mu_i, \sigma_i, \nu_i, \tau_i)^T$ which are the distribution parameters, where μ_i is the location indicating the mean of distribution, σ_i is the scale indicating the variance of distribution, and ν_i and τ_i are shape parameters representing the skewness and kurtosis of distribution.

As proposed by Rigby and Stasinopoulos [11], the addicted model $g_k(\theta_k)$ can be expanded to the following semi-parametric form (Eq. (2)) and nonlinear semi-parametric additive form (Eq. (3)):

$$g_k(\theta_k) = \eta_k = X_k\beta_k + \sum_{j=1}^{J_k} h_{jk}(\gamma_{jk}) \tag{2}$$

$$g_k(\theta_k) = \eta_k = h(X_k, \beta_k) + \sum_{j=1}^{J_k} h_{jk}(\gamma_{jk}) \tag{3}$$

where $h()$ is the nonlinear model, and h_{jk} is the non-parameter fitting model (cubic smoothing function, etc.).