

农产品期货市场波动率的动态特征及其预测模型

杨 科 田风平*

摘要: 本文以 8 种农产品期货的高频数据为样本,实证考察了我国农产品期货市场已实现波动率的动态特征,发现农产品期货已实现波动率同时具有长记忆性和区制转换性。在此基础上构建了长记忆马尔科夫区制转换模型来预测农产品期货的已实现波动率,并比较和评价了该模型与其他嵌套模型的预测性能。结果发现,我国农产品期货的已实现波动率具有高波动和低波动两种不同的状态,状态之间的转换概率较小,低波动状态的稳定性比高波动状态强;同时引入长记忆性和区制转换能进一步提高模型的预测性能,长记忆马尔科夫区制转换模型是预测性能最好的模型。

关键词: 农产品期货 已实现波动率 长记忆性 区制转换 预测

一、引言及文献综述

农产品期货价格是市场交易机制的核心,准确认识农产品期货价格波动的动态特征,有利于投资者和监管者做出科学合理的决策。然而,农产品期货同时具有衍生品固有的高风险特性,易受外来宏观因素以及季节因素的影响,使用不当易诱发金融市场的极端风险。因此,农产品期货市场波动率的准确预测对于防范投资风险、促进经济金融安全和粮食生产安全的实现以及促进高效期货市场体系的建立等方面都有重要的理论价值和现实意义。

国内对农产品期货市场的研究大致可以分为理论研究和实证研究两个方面。理论研究在农产品期货市场发展初期占据主要地位,主要集中于从制度方面探讨我国农产品期货市场如何建设等问题。国内关于农产品期货市场的实证研究在 2000 年以后逐步增多,主要侧重于借鉴西方金融工程理论,对我国农产品期货市场的有效性、市场功能等方面进行数量分析,例如,王辉等(2012)基于修正 ADCC 模型和 DADCC 模型研究了我国农产品期货套期保值中基差和“消息”的非对称效应及其对套期保值效果的影响。国内学者对我国农产品期货市场波动率

* 杨科,华南农业大学经济管理学院,邮政编码:510642,电子信箱:yangkedc@foxmail.com;田风平(通讯作者),中山大学国际商学院,邮政编码:510275,电子信箱:tianfpdc@gmail.com。

本文得到国家自然科学基金项目“农产品期货市场波动率的预测以及预测精度评价研究”(项目编号:71203067)、广东省高校优秀青年创新人才培养项目(育苗项目)“我国农产品期货市场波动率的预测研究”(项目编号:2012WYM_0033)、广东省哲学社会科学规划项目“广东省居民消费平滑研究”(项目编号:GD11YJ01)以及中央高校基本科研业务费专项资金中山大学青年教师培育项目“中国城乡居民消费行为差异与消费金融产品创新研究”(项目编号:13wkpy21)的资助。感谢匿名评审人和编辑部富有建设性的修改意见,文责自负。

的研究文献非常少,并且绝大部分都是采用基于低频收益率数据的 GARCH 族模型来研究我国农产品期货市场波动率的特征。例如,王金媛(2009)基于 GARCH 族模型发现我国农产品期货市场波动率具有长记忆性和不对称性;魏宇(2009a)基于能够刻画非对称杠杆效应的 EGARCH、GJR 和 APARCH 等非线性 GARCH 模型发现我国商品期货市场的波动率的非对称杠杆效应不明显;刘向丽等(2012)通过构建 ACD - GARCH - M 模型研究了我国期货市场价格波动特征,发现久期、交易量等微观结构变量能很好地吸收波动率的聚集性。然而,国内鲜有文献在高频数据环境下同时探讨农产品期货市场波动率的长记忆性和区制转换特征及其预测等问题。因此,本文在高频数据环境下研究我国农产品期货市场波动率的区制转换和长记忆性特性,并在此基础上进一步研究我国农产品期货市场波动率的预测问题。

在基于高频数据的已实现波动率的预测建模方面,Andersen 等(2003)的 ARFIMA 模型和 Corsi(2009)的 HAR 模型是最具代表性的两类模型。近年来,这两类模型得到了一些重要的扩展和改进,陈浪南和杨科(2013)对这方面的研究进行了综述。例如,Giot 和 Laurent(2004)考虑波动率的不对称性的 ARFIMAX 模型以及 Corsi 和 Reno(2009)的不对称性的 AHAR 模型,Beltratti 和 Morana(2005)、Corsi 等(2008)、Degiannakis(2008)以及 Yang 和 Chen(2014)通过考虑扰动项的条件异方差性,构建的 ARFIMA - FIGARCH 模型、HAR - GARCH 模型、ARFIMA - TARCH 模型以及不对称性的 HAR - D - FIGARCH 模型,魏宇(2009b)基于多分形波动率的 ARFIMA 模型,杨科和陈浪南(2012)的半参数预测模型等都较大程度提高 ARFIMA 模型和 HAR 模型的样本外预测性能。

然而,金融市场周期性会受到重要的突发金融事件的影响,可能会导致金融市场波动的大幅变动,如最近的次贷危机引发的全球金融危机等。但上述模型均没有考虑金融市场波动率可能存在的区制转换性。此外,国外学者发现波动率的区制转换可能造成“伪”长记忆性(Diebold and Inoue, 2001),Liu 和 Maheu(2008)以及 Choi 等(2010)的研究表明已实现波动率的结构突变只能部分解释波动率的长记忆性。最近,外国一些学者建议在已实现波动率模型中同时考虑长记忆性和结构突变特征,例如,McAleer 和 Medeiros(2008)的具有区制平滑转换的 HAR 模型以及 Lux 和 Morales - Arias(2010)的具有区制转换的多重分形模型。鉴于此,本文首先检验农产品期货的已实现波动率序列长记忆性特征的真伪性;然后构建具有马尔科夫区制转换性质的 ARFIMA 模型来刻画农产品期货已实现波动率的动态特征,该模型不但同时考虑了波动率的长记忆性和区制转换特征,还考察了长记忆性参数的时变性,模型采用 MCMC 方法进行估计;最后采用 MCS 检验和加权似然比(WLR)检验,比较和评价新构建的马尔科夫区制转换 ARFIMA 模型及其 4 种嵌套模型对我国农产品期货市场波动率的预测性能。

二、模型与方法

(一)波动率的高频估计

近年来,随着高频环境逐步建立,基于高频数据的已实现波动率(Realized Volatility, RV)的方法成为金融计量领域的研究热点。已实现波动率是日内高频收益率的平方和,可作为日波动率的代理变量。假设 $P_{i,t}$ 表示农产品期货在交易日 t 的第 i 个价格观测值,则交易日 t 的第 i 个日内高频收益率为 $r_{i,t} = 100 \times (\log P_{i,t} - \log P_{i-1,t})$ 。假定交易日 t 可以获得的高频收益率数目为 M ,则交易日 t 的已实现波动率可以表述为:

$$RV_t = \sum_{i=1}^M r_{i,t}^2 \quad (1)$$

由于有记录的高频数据仅包含有交易时段的市场波动信息,而无法反映无交易时段的市场波动状况,因此,本文采用 Martens(2002)的尺度参数 δ 对式(1)进行偏差修正,以考虑农产品期货市场休市期间的波动状况,修正后的已实现波动率估计量为:

$$RV_t^\delta = \sum_{i=1}^M r_{i,t}^2, \delta = \frac{\hat{\sigma}_{oc}^2 + \hat{\sigma}_{co}^2}{\hat{\sigma}_{oc}^2} \quad (2)$$

其中, $\hat{\sigma}_{oc}^2 = (1/T) \sum_{t=1}^T (\log P_{M,t} - \log P_{0,t})$ 为当天开盘至当天收盘这段时间的样本方差, $\hat{\sigma}_{co}^2 = (1/T) \sum_{t=1}^T (\log P_{0,t} - \log P_{M,t-1})$ 为收盘至第二天开盘这段时间的样本方差, $P_{0,t}$ 和 $P_{M,t}$ 为在交易日 t 的农产品期货价格的第一个观测值和最后一个观测值。

(二) 长记忆性检验

由于金融资产波动率的长记忆性的存在使得金融市场的行为变得更为复杂,对金融市场的长记忆性进行深入研究是非常重要的。Andersen 等(2003)通过 ARFIMA 模型较好地刻画了已实现波动率的长记忆性,并取得良好的样本外预测效果。已实现波动率的 ARFIMA(p, d, q) 模型可以表述为:

$$\Phi(L)(1-L)^d(y_t - \mu) = \Theta(L)\varepsilon_t \quad (3)$$

其中 $y_t (t=1, \dots, T)$ 表示已实现波动率, μ 为已实现波动率的均值, $\Phi(L)$ 和 $\Theta(L)$ 分别为自回归滞后 p 阶算子以及移动平均滞后 q 阶算子, $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ 。 $(1-L)^d$ 为分数阶积分算子, d 表示分数阶差分参数或记忆参数。

记忆参数 d 除了运用 ARFIMA 模型估计外,文献中运用最为广泛还有半参数对数周期图(log-periodogram, 后文记为 GPH)法。记忆参数 d 可以通过如下的回归方程估计得到:

$$\log(I(\lambda_j)) = \alpha_0 - d \log[4 \sin^2(\lambda_j/2)] + \varepsilon_j \quad j = l+1, \dots, m \quad (4)$$

其中 $I(\lambda_j)$ 表示在频率 $\lambda_j = 2\pi j/T$ 下的样本谱密度, T 为样本观测值数目。GPH 方法的一个优点是关于记忆参数 d 的假设可以直接采用标准的 t 统计量进行检验。但 Choi 和 Zivot(2007)的研究发现,短记忆的强自相关性可能导致 GPH 方法的有限样本偏差。出于比较目的,本文参照 Coakley 等(2011)的研究,还进一步考虑参数 ARFIMA 模型中参数 d 的基于极大似然估计的误差配准算法估计值。

(三) 伪长记忆性检验

Diebold 和 Inoue(2001)以及 Perron 和 Qu(2010)的研究表明,波动率序列的结构突变能产生“伪长记忆性”(spurious long memory)。Choi 等(2010)的研究发现结构突变能部分解释已实现波动率的长记忆性。由此可见,波动率序列所表现出的长记忆性可能与区制转换或结构突变混淆。为了区分农产品期货已实现波动率的长记忆性和区制转换特性以及区制转换对长记忆性估计值的影响,本文采用最新发展的 4 类检验方法对农产品期货已实现波动率长记忆性的真伪进行检验。第一种检验方法基于这样的事实:如果一个时间序列是分数阶积分过程,即 $I(d)$ 过程,则该序列的每个子样本也应该具有相同的 d 值。因此,该检验的零假设是 $H_0: d_0 = d_{0,1} = \dots = d_{0,b}$, 其中 d_0 表示记忆参数 d 在全样本下的真实值, $d_{0,a}$ 表示 d 的第 a 个子样本真实值。对于伪 $I(d)$ 过程, d 在各个子样本下的估计值的平均值将与全样本估计值存在差别,

其差异程度随取样频率的增加而增加。基于这一事实, Shimotsu (2006) 建议采用如下的统计量来检验 $I(d)$ 过程的真伪性:

$$W_c = 4m(c_{m/b}/(m/b)) Ad_b (A \Omega A')^+ (Ad_b)' \quad (5)$$

其中 $c_m = \sum_{j=1}^m \nu_j^2$, $\nu_j = \log j - \frac{1}{m} \sum_{j=1}^m \log j$ 。Shimotsu (2006) 已经证明当 $n \rightarrow \infty$ 时, $W_c \xrightarrow{d} \chi^2(b-1)$ 。

第二种检验方法基于 $I(d)$ 过程的时域特征: 若一个时间序列遵循 $I(d)$ 过程, 则该序列的 d 阶差分应遵循 $I(0)$ 过程。该检验方法的检验步骤如下: 首先, 对时间序列的 d 值进行估计, 然后, 运用 d 的估计值对原序列进行 d 阶差分, 最后, 运用 KPSS 检验以及 PP 检验对差分后的数据进行单位根检验。

第三种检验方法基于 Perron 和 Qu (2010) 的研究, 该检验的零假设是时间序列遵循平稳的长记忆过程, 备择假设是时间序列遵循具有均值突变的短记忆过程, 该检验的检验统计量为:

$$t_d(a, c_1, b, c_2) = \sqrt{24c_1 [T^a] / 2\pi^2} (\hat{d}_{a, c_1} - \hat{d}_{b, c_2}) \quad (6)$$

其中 $\hat{d}_{a, c}$ 为记忆参数 d 在频率 $m_{a, c} = c[T^a]$ 时的对数周期图估计值, $0 < a < b < 1$, $a < 4/5$ 。在零假设条件下该检验统计量渐进遵循标准正态分布。此外, 本文还考虑另外两个检验统计量: 一是当固定 b 等于 $1/2$ 和 $1 < c_1 < 2$ 时, 统计量 $t_d(1/3, c_1, b, 1)$ 的最大值, 即 $\text{Sup} - t_d = \sup_{c_1 \in [1, 2]} t_d(1/3, c_1; b, 1)$; 二是当固定 b 等于 $1/2$ 和 $1 < c_1 < 2$ 时, 统计量 $t_d(1/3, c_1, b, 1)$ 的均值, 即 $\text{mean} - t_d = \text{mean}_{c_1 \in [1, 2]} t_d(1/3, c_1; b, 1)$ 。由于这两类统计量的极限分布无法直接得到, 本文采用 Perron 和 Qu (2010) 的建议, 运用“自助法”(Bootstrap) 来计算相关的临界值。

第四种检验方法基于 Qu (2011) 的研究, 该检验的零假设是时间序列遵循平稳的长记忆过程, 备择假设是长记忆性受马尔科夫区制转换或平滑可变的趋势的影响。Qu (2011) 构建的检验统计量如下:

$$W = \sup_{r \in [\varepsilon, 1]} \left(\sum_{j=1}^m \nu_j^2 \right)^{-1/2} \left| \sum_{j=1}^{[mr]} \nu_j \left(\frac{I_j}{G(\hat{d}) \lambda_j^{-2\hat{d}}} - 1 \right) \right| \quad (7)$$

其中 $\hat{d}$ 为记忆参数 d 在频率 m 下的局部消减估计值 (local Whittle estimate), ε 为消减参数, I_j 表示待研究时间序列在频率 λ_j 下的周期图。此外 $G(\hat{d}) = m^{-1} \sum_{j=1}^m \lambda_j^{2\hat{d}} I_j$, $\nu_j = \log \lambda_j - \frac{1}{m} \sum_{j=1}^m \log \lambda_j$ 。

(四) 马尔科夫区制转换 ARFIMA 模型 (MS-ARFIMA 模型)

参照 Ray 和 Tsay (2002) 的研究, 本文构建一个基于两个 $\text{ARFIMA}(1, d_i, 0)$ ($i = 0, 1$) 模型动态混合的长记忆马尔科夫区制转换模型, 用来同时刻画农产品期货已实现波动率的长记忆性特征和区制转换特征。该模型将均值漂移和时变长记忆结合在一起, 进一步扩展了 Ray 和 Tsay (2002) 的模型。该模型记为 $\text{MS-ARFIMA}(1, d_i, 0)$ 模型, 表述如下:

$$y_t = \mu_{S_t} + \beta_{S_t} y_{t-1} + (1-L)^{-d_{S_t}} \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma_{S_t}^2) \quad (8)$$

其中, y_t ($t = 1, \dots, T$) 表示已实现波动率, S_t ($t = 1, \dots, T$) 表示无法观察的两种状态 $\{0, 1\}$,

遵循 1 阶马尔科夫链,其转换概率矩阵 $\Pi = \begin{bmatrix} p_{00} & p_{10} \\ p_{01} & p_{11} \end{bmatrix}$, $p_{ij} = P(s_t = j | s_{t-1} = i)$, $\sum_{j=0}^1 p_{ij} = 1$ 。此外,本文采用如下的先验分布设定: $\mu_i \sim N(1, 0.5)$, $\beta_i \sim N(1, 0.5)$, $d_i \sim U(0, 0.5)$, $\sigma_i^2 \sim IG(2.5, 0.75)$ 和 $p_{ij} \sim \text{Beta}(5, 95)$ ($i, j = 0, 1$)。

MS-ARFIMA(1, d_i , 0) 模型同时考虑了波动率的长记忆性和区制转换特征,潜在的区制转换过程 $S = (S_1, \dots, S_T)$ 无法直接观测到,需要考虑所有时点的状态,并且关于长记忆性参数的推断也不明显,似然方程的计算较复杂。因此,本文采用 MCMC 方法来估计该模型。为使 MCMC 估计可行,本文首先通过 M 阶截断将 ARFIMA 过程重新表述为线性 MA 过程;然后基于区制转换向量 S 将该模型转化为如下的时变参数的状态空间模型:

$$\begin{aligned} \underbrace{\begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon_t \\ \vdots \\ \varepsilon_{t-M+1} \end{bmatrix}}_{\alpha_{t+1(M+1)} \times 1} &= \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}}_{T_{(M+1) \times (M+1)}} \underbrace{\begin{bmatrix} \varepsilon_t \\ \varepsilon_{t-1} \\ \vdots \\ \varepsilon_{t-M} \end{bmatrix}}_{\alpha_t} + \underbrace{\begin{bmatrix} \varepsilon_{t+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{u_{t(M+1)} \times 1} \\ y_t &= \mu_{s_t} + \beta_{s_t} y_{t-1} + \underbrace{\begin{bmatrix} 1 & \pi_1(d_{s_t}) & \pi_2(d_{s_t}) & \cdots & \pi_M(d_{s_t}) \end{bmatrix}}_{Z_t \ 1 \times (M+1)} \alpha_t \end{aligned} \quad (9)$$

其中 $\pi_j(d_{s_t}) = \frac{\Gamma(j+d_{s_t})}{\Gamma(j+1)\Gamma(d_{s_t})}$ 。基于式(9)表述形式,我们可以运用 ARMA 模型的

MCMC 估计的标准结果来对 MS-ARFIMA(1, d_i , 0) 模型的参数进行推断。此外,本文采用 MCMC 算法来模拟 (θ, S) , 其中 $\theta = (\mu_i, \beta_i, \sigma_i, d_i, p_{01}, p_{10})$ ($i = 0, 1$), 表示模型的参数向量。MCMC 算法如下:

(1) 设定初始值 $(M^{(0)}, \theta^{(0)}, S^{(0)})$;

(2) 从 $j=1$ 至 $j=n$:

其一,运用 RJMCMC 算法更新截断阶数 M ;

其二,运用 Metropolis-Hastings 算法从概率分布 $p(\theta_i | M^{(j)}, S^{(j-1)}, \theta_{-i}^{(j)}, \theta_{-i+}^{(j)}, Y)$ 中对 θ 进行抽样,其中 θ_{-i-} 表示参数向量 θ 的前 $i-1$ 个元素, θ_{-i+} 表示从 $i+1$ 至最后一个元素;

其三,参照 Chib(1996)的研究,从概率分布 $p(\theta_i | M^{(j)}, \theta^{(j)}, Y)$ 中按块抽样得到 $S^{(j)}$ 。

在实践中, y_{T+h} 的预测均值和预测密度都无法通过分析解方法直接计算得到。因此,本文采用预测模拟的方法获得。为了简便,我们定义 $X_h = (X_{T+1}, \dots, X_{T+h})$, $S_h = (S_{T+1}, \dots, S_{T+h})$ 和 $Y_h = (y_{T+1}, \dots, y_{T+h})$ 。预测密度的抽样可以通过如下步骤模拟得到:(1) 在第 j 步 Gibbs 抽样迭代中,设定 $\theta = \theta^{(j)}$ 和 $M = M^{(j)}$;(2) 对于 $i=1, \dots, h$, 从概率分布 $p(S_{T+i} | S_{T+i-1}^{(j)}, \theta^{(j)})$ 抽样得到 $S_{T+i}^{(j)}$;(3) 对于 $i=1, \dots, h$, 从概率分布 $p(y_{T+i} | Y, Y_{T+i-1}^{(j)}, M, X_{T+i}, S, S_{T+i-1}^{(j)}, \theta)$ 中抽样得到 $y_{T+i}^{(j)}$ 。上述模拟过程得到的序列 $y_{T+i}^{(j)}$ ($j=1, \dots, n, i=1, \dots, h$) 可以用于估计预测均值 $E(y_{T+i} | Y)$ 得到点预测值以及估计预测密度 $p(y_{T+i} | Y)$ 得到密度预测值。

(五) 样本外预测精度评价

不同预测时域的点预测和密度预测通过滚动时间窗的样本外预测方法获得,即将估计窗口固定为 N ,运用 $(y_1, \dots, y_N)$ 估计模型并产生第 1 个向前 h 步预测值,第 2 个向前 h 步预测值

基于 $(y_2, \dots, y_{N+1})$ 得到, 而最后 1 个向前 h 步预测值由 $(y_{T-N-h+1}, \dots, y_{T-h})$ 得到。最终, 我们可以分别获得 $T_h = T - N - h + 1$ 个点预测值和密度预测值。本文采用 Hansen 等 (2011) 的 MCS (Model Confidence Set) 检验来评价和比较各种模型的点预测性能, 采用如下的两种常见的损失函数:

$$MSE^i = \frac{1}{T_h} \sum_{t=N}^{T-h} (y_t - \hat{y}_{i,t})^2 \quad (10)$$

$$QLIKE^i = \frac{1}{T_h} \sum_{t=N}^{T-h} \left[\frac{y_t}{\hat{y}_{i,t}} + \log(\hat{y}_{i,t}) \right] \quad (11)$$

其中, $\hat{y}_{i,t}$ 表示由模型 i 得到的点预测值, y_t 表示真实的潜在波动率。给定一组候选预测模型集合 M_0 , MCS 方法的目标是确定的预测模型置信集 $\hat{M}_{1-\alpha}^* \subset M_0$, 该集合以 $1 - \alpha$ 的置信水平包含了最优的预测模型。MCS 检验程序开始于总体模型集 $M = M_0 = \{1, \dots, m_0\}$, 然后反复检验预测精度相等 (Equal Forecasting Accuracy, EFA) 这一零假设:

$$H_{0,M}: E(d_{ij,t}) = 0, \quad \forall i, j \in M \quad (12)$$

其中, $d_{ij,t} = L_{i,t} - L_{j,t}$ 表示模型 i 与模型 j 之间的损失函数值之差。只要零假设 EFA 在置信水平 α 下被拒绝, MCS 检验程序会序贯的从模型集合 M 中剔除预测精度最差的预测模型, 该过程一直持续到零假设 EFA 在置信水平 α 下不再被拒绝, 最后幸存下来的模型就是预测模型置信集 $\hat{M}_{1-\alpha}^*$ 。若在序贯剔除的每一步中都把置信水平固定为 α , 则集合 $\hat{M}_{1-\alpha}^*$ 以 $1 - \alpha$ 的置信水平包含了最优预测模型。MCS 检验过程给出了每个模型在初始集合中的 p 值。对于一个给定的模型 $i \in M$, MCS 检验过程估计的 $\hat{p}_i$ 是确定该模型是否属于模型置信集 MCS 的门阈置信水平, 当且仅当 $\hat{p}_i \geq \alpha$ 时, 模型 $i \in \hat{M}_{1-\alpha}^*$ 。在实证研究中, 报告了置信水平为 10% 的检验结果。

两种波动率模型间密度预测性能的比较基于损失函数的加权差:

$$WLR_{t+h} = \omega(\tilde{y}_{t+h}) [\log \hat{f}_i(y_{t+h}) - \log \hat{f}_j(y_{t+h})] \quad (13)$$

其中 $\tilde{y}_{t+h}$ 表示标准化的观测值, $\hat{f}_i$ 和 $\hat{f}_j$ 分别表示基于模型 i 和模型 j 由 $(y_{t-N+1}, \dots, y_t)$ 估计得到的向前 h 步预测密度, $\omega(\cdot)$ 表示加权系数。若式 (13) 为正, 意味着模型 i 的预测密度性能比模型 j 高。参照 Amisano 和 Giacomini (2007) 的研究, 检验两种模型具有相同的密度预测性能等价于检验如下的假设系统:

$$H_0: E[WLR_{t+h}] = 0, \quad t > N \quad \text{vs.} \quad H_1: E\left[\sum_{t=N}^{T-h} \frac{WLR_{t+h}}{T_h}\right] \neq 0 \quad (14)$$

该检验的统计量为:

$$t_h = \frac{\sum_{t=N}^{T-h} \frac{WLR_{t+h}}{T_h}}{\widehat{avar}(WLR_{t+h})} \quad (15)$$

其中 $\widehat{avar}(WLR_{t+h})$ 表示渐进方差, 本文采用对异方差和自相关稳健的估计量。

三、数据与统计描述

本文研究的数据样本为橡胶期货、黄豆期货、豆粕期货、豆油期货、玉米期货、白砂糖期货、强筋小麦期货以及棉花期货的日内高频交易数据, 取样区间为 2006 年 1 月 9 日至 2012 年 4

月 27 日,数据来源于 Wind 资讯金融数据库。由于在同一时间,对于每一个期货品种都有不同的合约在进行交易,本文按照交易最活跃的合约来构造主力连续期货合约。按照 Corsi 等(2008)的建议,对日内高频交易数据进行每 5 分钟的抽样,以平衡测量误差和微观市场结构噪声对已实现波动率估计值的影响,通过剔除周末、公共节假日以及一些数据不全的交易日,最后得到每种期货合约总共 $T=1\,497$ 个交易日的 5 分钟高频数据。表 1 报告了 8 种期货合约已实现波动率序列的描述性统计量以及由 ARFIMA(1, d ,0)模型估计得到的长记忆性参数 d 。由表 1 可知,8 种期货合约已实现波动率序列都表现出尖峰厚尾的特征和较强的序列相关性。长记忆性参数 d 的估计值均大于 0.25,表明我国农产品期货市场波动率具有较强的长记忆性。

表 1 已实现波动率序列的描述性统计量

RV	极小值	均值	极大值	标准差	偏度	峰度	JB	d
橡胶	0.0543	2.4614	37.7120	2.8915	4.9622	43.3830	1.07E+05	0.5154 (0.1186)
黄豆	0.0264	0.6366	17.378	0.9910	7.0523	84.2620	4.2430	0.3420 (0.0486)
豆粕	0.0485	0.7790	23.874	1.1564	8.2247	126.4000	9.67E+05	0.3458 (0.0459)
豆油	0.0301	0.8927	23.08	1.5208	6.9674	76.7540	3.51E+05	0.4038 (0.0663)
玉米	0.0460	0.3973	7.1231	0.4684	5.9028	57.2720	1.92E+05	0.3187 (0.0582)
白砂糖	0.0286	1.0819	15.8930	1.3147	4.4746	34.4040	6.65E+04	0.3615 (0.0441)
强筋小麦	0.0269	0.4494	18.7310	0.7269	12.8850	282.0600	4.9E+06	0.2579 (0.0434)
棉花	0.0182	0.7393	18.0410	1.4818	5.9468	52.2190	1.60E+05	0.4357 (0.0696)

注:RV 表示已实现波动率,下同。

四、实证结果与分析

(一)长记忆性检验结果与分析

表 2 报告了 ADF 检验、PP 检验和 KPSS 检验三种单位根检验结果。ADF 和 PP 检验结果显示所有已实现波动序列在 1% 的显着性水平都拒绝单位根的零假设,而 KPSS 检验结果显示,除了橡胶期货外,所有已实现波动序列在 1% 的显着性水平也拒绝平稳性的原假设。

为了探讨农产品期货已实现波动率是否具有分数阶积分性质,本文采用式(4)所述的半参数的 GPH 法得到长记忆参数 d 的估计值,并运用对应的 t 统计量来检验无长记忆性假设 $H_0:d=0$ 和单位根假设 $H_0:d=1$,检验结果如表 3 所示。表 3 的结果表明:第一,所有农产品期货的已实现波动率序列在 1% 的显着性水平下既拒绝无长记忆性的原假设,也拒绝单位根的原假设。这一结论说明我国农产品期货市场的已实现波动率序列既不是 $I(1)$ 过程也不是 $I(0)$ 过程,而更加遵循 $0 < d < 1$ 的分数阶积分过程。第二,基于半参数的 GPH 法得到长记忆参数 d 的估计值在 0.2637 到 0.6407 之间,说明已实现波动率序列都表现出很强的长记忆性,

并且绝大多数都落在稳定区域内($0 < d < 0.5$)。表 4 给出了基于 ARFIMA(1, d , 0) 和 ARFIMA(1, d , 1) 两类参数模型的检验结果。由表 4 可知, 所有农产品期货的已实现波动率序列在 1% 的显着性水平上再次拒绝了无长记忆性和单位根的原假设。这一发现进一步说明, 我国农产品期货的已实现波动率序列应遵循 $0 < d < 1$ 的分数阶积分过程。

表 2 单位根检验结果

RV	ADF		PP		KPSS
	τ	τ_{μ}	$Z_{t_{\alpha}}$	$Z_{t_{\alpha}^*}$	η_{μ}
橡胶	-7.6184 **	-11.4160 **	-14.51 **	-27.50 **	0.4986 *
黄豆	-10.2761 **	-13.7901 **	-19.20 **	-28.88 **	2.8207 **
豆粕	-9.7216 **	-13.1537 **	-17.80 **	-27.78 **	3.9744 **
豆油	-9.3997 **	-12.0543 **	-21.95 **	-31.26 **	3.0114 **
玉米	-9.2823 **	-14.2506 **	-14.19 **	-26.12 **	4.1360 **
白砂糖	-8.8035 **	-13.1649 **	-14.51 **	-26.39 **	1.7354 **
强筋小麦	-11.6154 **	-15.9122 **	-21.57 **	-31.18 **	4.1100 **
棉花	-8.1281 **	-9.5260 **	-16.61 **	-21.76 **	5.6797 **

注: (1) $\tau(Z_{t_{\alpha}})$ 表示无常数项模型的 ADF(PP) 检验统计量, $\tau_{\mu}(Z_{t_{\alpha}^*})$ 表示带有常数项模型的 ADF(PP) 检验统计量; (2) η_{μ} 表示有常数项的 KPSS 检验统计量; (3) ** 和 * 分别表示 1% 和 5% 显著水平。

表 3 GPH 检验结果

RV	$d_{GPH}(m = T^{0.5})$	$\tau_d = 1$	$\tau_d = 0$	$d_{GPH}(m = T^{0.7})$	$\tau_d = 1$	$\tau_d = 0$
橡胶	0.4276(0.1213)	-4.7196	3.5257	0.4441(0.0529)	-10.5094	8.3942
黄豆	0.4099(0.1213)	-4.8657	3.3796	0.4069(0.0529)	-11.2109	7.6927
豆粕	0.4907(0.1213)	-4.1995	4.0458	0.5162(0.0529)	-9.1457	9.7579
豆油	0.5431(0.1213)	-3.7669	4.4784	0.4973(0.0529)	-9.5030	9.4001
玉米	0.2637(0.1213)	-6.0711	2.1742	0.3270(0.0529)	-12.7218	6.1818
白砂糖	0.3367(0.1213)	-5.4688	2.7765	0.4338(0.0529)	-10.7041	8.1995
强筋小麦	0.5073(0.1212)	-4.0622	4.1831	0.2878(0.0529)	-13.4628	5.4408
棉花	0.6407(0.1213)	-2.9624	5.2829	0.5446(0.0529)	-8.6096	10.2940

注: (1) 括号中为 d_{GPH} 估计值的标准误(σ_d); (2) $\tau_d = 1$ 列表示检验统计量 $(d_{GPH} - 1)/\sigma_d$; (3) $\tau_d = 0$ 列表示检验统计量 $(d_{GPH} - 0)/\sigma_d$; (4) 5% 和 10% 的显著水平的临界值分别为 ± 1.96 和 ± 1.64 。

表 4 基于 ARFIMA 模型的检验结果

RV	$d_{ARFIMA}(1, d, 0)$ with constant	$\tau_d = 1$	$\tau_d = 0$	$d_{ARFIMA}(1, d, 1)$ with constant	$\tau_d = 1$	$\tau_d = 0$
橡胶	0.5154(0.1186)	-4.0860	4.3457	0.4828(0.1175)	-4.4017	4.1089
黄豆	0.3420(0.0486)	-13.5391	7.0370	0.3138(0.0559)	-12.2755	5.2387
豆粕	0.3458(0.0459)	-14.2527	7.5338	0.5177(0.1350)	-3.5726	3.8348
豆油	0.4038(0.0663)	-8.9925	6.0905	0.4893(0.1513)	-3.3754	3.2340
玉米	0.3187(0.0582)	-11.7062	5.4759	0.2730(0.0804)	-9.0423	3.3955
白砂糖	0.3615(0.0441)	-14.4785	8.1973	0.3234(0.0340)	-19.9000	9.5118
强筋小麦	0.2579(0.0434)	-17.0991	5.9424	0.3284(0.0875)	-7.6754	3.7531
棉花	0.4357(0.0696)	-8.1078	6.2601	0.5444(0.1514)	-3.0092	3.5958

注: (1) 括号中为 d_{ARFIMA} 估计值的标准误(σ_d); (2) $\tau_d = 1$ 列表示检验统计量 $(d_{ARFIMA} - 1)/\sigma_d$; (3) $\tau_d = 0$ 列表示检验统计量 $(d_{ARFIMA} - 0)/\sigma_d$; (4) 5% 和 10% 的显著水平的临界值分别为 ± 1.96 和 ± 1.64 。

(二) 伪长记忆性检验结果与分析

表 5 给出了前文所述的前两种伪长记忆性检验的检验结果, 其中基于全样本的长记忆性

参数估计值 $\hat{d}_{fetu}$ 与基于子样本的长记忆性参数估计值的均值 $\bar{d}$ 都采用 Shimotsu (2006) 的两步可行的正合局部消减法得到, 统计量 $W_c, Z_t, \hat{\eta}_\mu$ 分别基于 $m = \{90, 120\}$ 和 $b = \{2, 3\}$ 计算得到。从表 5 可知, 虽然 8 种农产品期货的已实现波动率的 $\hat{d}_{fetu}$ 和 $\bar{d}$ 都比较接近, 但从检验统计量 W_c 可知, 除棉花期货外, 其他 7 种农产品期货的已实现波动率序列在所有的 b 值和 m 值下都拒绝了长记忆性参数为常数的原假设, 并且由检验统计量 Z_t 和 $\hat{\eta}_\mu$ 可知, 所有农产品期货的已实现波动率序列同时也不能拒绝遵循 $I(d)$ 过程的原假设。此外, 除橡胶期货外其他 7 种农产品期货的 $\hat{d}_{fetu}$ 和 $\bar{d}$ 都随 m 的增加而减小, 说明我国农产品期货的已实现波动率序列可能具有区制转换或结构突变特征。表 6 报告了前文所述的后两种伪长记忆性检验的检验结果。由表 6 中的检验统计量可以发现, 8 种农产品期货的已实现波动率序列在 1% 的显著性水平下都不能拒绝遵循具有均值漂移的短记忆过程的原假设, 但由检验统计量 $\text{Sup} - t_d$ 和 $\text{Mean} - t_d$ 可知, 只有黄豆期货和豆油期货的已实现波动率序列拒绝遵循长记忆过程的原假设。进一步由第四种检验方法的检验统计量 W 可知, 只有强筋小麦的已实现波动率序列在显著水平 2.5% 和 $\varepsilon = 0.02$ 以及显著水平 5% 和 $\varepsilon = 0.05$ 的条件下才拒绝遵循平稳长记忆过程的原假设。

因此, 尽管我国农产品期货的已实现波动率序列的长记忆性可能存在结构突变或局部变化, 但我们也找不到拒绝其遵循真实的 $I(d)$ 过程的有力证据, 说明纯粹的 $I(d)$ 过程或纯粹的区制转换或结构突变模型都不能全部解释我国农产品期货的已实现波动率序列所表现出的强长记忆性。所以, 在对我国农产品期货的已实现波动率建模时, 必须同时考虑已实现波动率的真实长记忆性和区制转换(或结构突变)特征。

表 5 第一种和第二种检验方法的检验结果

RV	$\hat{d}_{fetu}$	$\bar{d}$		W_c		Z_i	$\hat{\eta}_{\mu}$
		$b = 2$	$b = 3$	$b = 2$	$b = 3$		
$m = 90$							
橡胶	0.3763	0.4191	0.4030	2.8646	3.0099	-2.3588	0.0531
黄豆	0.4450	0.4134	0.4273	0.8748	0.00379	-1.5949	0.1304
豆粕	0.4919	0.4570	0.4433	0.6049	1.0762	-1.6524	0.1229
豆油	0.4875	0.4871	0.4711	0.0239	1.3819	-1.8215	0.1075
玉米	0.2906	0.2422	0.2986	0.6319	0.2071	-0.9548	0.3428
白砂糖	0.4526	0.4810	0.4881	3.1895	0.9968	-2.6152	0.0857
强筋小麦	0.3254	0.3177	0.3449	3.5280	3.3999	-1.1267	0.2491
棉花	0.5150	0.4561	0.4803	8.5146 *	7.2435 *	-1.8733	0.0795
$m = 120$							
橡胶	0.3917	0.4060	0.4032	0.3571	1.0840	-2.4660	0.0493
黄豆	0.4105	0.3855	0.3969	1.7539	1.1904	-1.3678	0.1676
豆粕	0.4876	0.4510	0.4256	2.4373	2.2539	-1.6176	0.1273
豆油	0.4371	0.4462	0.4141	0.0097	5.5291	-1.4371	0.1560
玉米	0.2881	0.2502	0.2916	0.4248	0.2601	-0.9390	0.3508
白砂糖	0.4478	0.4676	0.4494	2.5010	0.4392	-2.5639	0.0885
强筋小麦	0.3139	0.2968	0.3292	1.3931	3.6114	-1.0576	0.2752
棉花	0.4734	0.4190	0.4629	6.0965 *	7.1804 *	-1.6110	0.1090

注: (1) Z_t 表示 PP 检验统计量; (2) $\hat{\eta}_\mu$ 表示 KPSS 检验统计量; (3) * 表示 5% 的显著水平; (4) $\chi^2_{0.95}(1) = 3.84$, $\chi^2_{0.95}(2) = 5.99$ 。

表 6 第三种和第四种检验方法的检验结果

RV	Perron 和 Qu(2010) tests					Qu(2011) tests	
	检验统计量			模拟的临界值		W	
	t_d	Sup - t_d	Mean - t_d	Sup - t_d	Mean - t_d	$\varepsilon = 0.02$	$\varepsilon = 0.05$
橡胶	3.5986	-0.0872	-0.6240	1.27, 1.70, 2.59	0.35, 0.62, 1.33	0.7064	0.5557
黄豆	3.7996	1.5710	0.8018	1.29, 1.70, 2.48	0.35, 0.70, 1.38	0.5515	0.5515
豆粕	4.4339	1.7004	0.7549	0.94, 1.38, 2.25	0.03, 0.40, 1.09	0.8633	0.8249
豆油	5.4933	0.3664	-0.3204	2.21, 2.65, 3.54	1.23, 1.61, 2.28	0.6130	0.6130
玉米	3.1748	0.6023	0.0087	2.21, 2.62, 3.51	1.22, 1.58, 2.19	0.3923	0.3923
白砂糖	3.2688	-0.8907	-1.1721	1.51, 1.94, 2.83	0.59, 0.92, 1.58	0.4052	0.2746
强筋小麦	5.0854	1.7578	0.9982	2.22, 2.66, 3.53	1.23, 1.58, 2.32	1.4618	1.1867
棉花	6.2045	1.2381	-0.3255	1.27, 1.66, 2.65	0.33, 0.73, 1.40	0.4247	0.4032

注: (1) 模拟的临界值分别为 10%, 5% 和 1% 的显著水平; (2) 当消减参数 ε 为 0.02 时, Qu(2011) 的 W 统计量在 10%, 5%, 2.5% 和 1% 的显著水平下的渐进临界值分别为 1.118, 1.252, 1.374 和 1.517, 当消减参数 ε 为 0.05 时, 分别为 1.022, 1.155, 1.277 和 1.426。

(三) MS - ARFIMA(1, d_i , 0) 模型估计结果与分析

本文采用 MCMC 方法估计 MS - ARFIMA(1, d_i , 0) 模型, 由于 MCMC 算法的计算量很大, 因此本文的估计程序使用了类似于 C 语言的 Ox©6.0 语言以及 ssfpack 程序包编程实现。本文进行了 10 000 次 Gibbs 抽样, 前 2 500 次作为预烧过程, 以剔除对初始条件的依赖性, 后 7500 次抽样用来产生实证结果。表 7 是用 MCMC 方法估计 MS - ARFIMA(1, d_i , 0) 模型参数的后验均值以及 95% 的置信区间。

表 7 MS - ARFIMA(1, d_i , 0) 模型参数的估计结果

	橡胶	黄豆	豆粕	豆油	玉米	白砂糖	强筋小麦	棉花
μ_0	1.5873 [1.00, 2.07]	0.3867 [0.24, 0.53]	0.5428 [0.34, 0.76]	0.4781 [0.29, 0.66]	0.2536 [0.15, 0.36]	0.6169 [0.41, 0.86]	0.2585 [0.14, 0.41]	0.2228 [0.09, 0.44]
μ_1	5.5041 [4.28, 6.43]	2.5260 [1.93, 3.08]	3.2330 [2.50, 3.96]	3.5913 [2.94, 4.12]	1.8069 [1.48, 2.12]	2.8102 [2.18, 3.35]	2.9096 [2.06, 3.56]	3.7514 [2.97, 4.48]
β_0	-0.0721 [-0.17, 0.04]	0.1042 [0.00, 0.24]	0.0710 [-0.08, 0.21]	0.1501 [-0.02, 0.32]	0.2164 [-0.02, 0.50]	0.1151 [-0.06, 0.29]	0.2804 [0.02, 0.53]	0.4433 [0.08, 0.69]
β_1	-0.2405 [-0.41, -0.04]	-0.3979 [-0.52, -0.24]	-0.2855 [-0.43, -0.13]	-0.3977 [-0.57, -0.23]	-0.2379 [-0.41, -0.06]	-0.3544 [-0.50, -0.21]	-0.3746 [-0.55, 0.00]	-0.1605 [-0.37, 0.02]
d_0	0.3546 [0.27, 0.43]	0.2025 [0.11, 0.29]	0.2805 [0.20, 0.37]	0.2537 [0.15, 0.36]	0.1842 [0.02, 0.35]	0.2625 [0.16, 0.37]	0.1215 [0.00, 0.26]	0.1198 [0.00, 0.32]
d_1	0.4180 [0.29, 0.50]	0.4525 [0.31, 0.50]	0.3545 [0.19, 0.49]	0.3606 [0.19, 0.49]	0.2134 [0.05, 0.39]	0.4704 [0.37, 0.50]	0.4269 [0.22, 0.50]	0.3527 [0.18, 0.48]
σ_0^2	1.5896 [1.32, 1.88]	0.2472 [0.13, 0.37]	0.4593 [0.20, 0.63]	0.3385 [0.24, 0.61]	0.0706 [0.05, 0.08]	0.4223 [0.28, 0.59]	0.2284 [0.09, 0.30]	0.2998 [0.15, 0.41]
σ_1^2	30.7605 [23.68, 40.75]	4.6307 [2.83, 7.29]	5.0963 [2.65, 10.17]	10.6806 [7.44, 15.05]	0.9610 [0.65, 1.49]	4.9392 [3.46, 6.85]	3.3642 [1.34, 8.66]	8.7096 [6.02, 12.65]
ρ_1	0.0306 [0.02, 0.04]	0.0240 [0.01, 0.04]	0.0175 [0.01, 0.03]	0.0244 [0.02, 0.03]	0.0178 [0.01, 0.03]	0.0305 [0.02, 0.04]	0.0138 [0.01, 0.02]	0.0149 [0.01, 0.02]
ρ_0	0.1629 [0.11, 0.22]	0.1608 [0.10, 0.22]	0.1177 [0.07, 0.17]	0.1428 [0.10, 0.20]	0.1399 [0.09, 0.20]	0.1489 [0.10, 0.21]	0.1274 [0.07, 0.20]	0.0942 [0.05, 0.14]
M	28.7079 [19.0, 38.0]	31.3550 [14.0, 46.0]	25.4644 [15.0, 36.0]	29.1147 [14.0, 41.0]	22.2423 [16.0, 33.0]	25.1231 [16.0, 41.0]	23.9366 [15.0, 37.0]	23.4141 [13.0, 32.0]

从表 7 可以看出, $MS-ARFIMA(1, d_i, 0)$ 模型很明显的从 8 种农产品期货的已实现波动率序列中分离出了两种不同的波动状态: 所有农产品期货市场的已实现波动率序列在状态 1 中 σ_1^2 和 d_1 的估计值均高于在状态 0 中 σ_0^2 和 d_0 的估计值, 并且在状态 1 中自回归参数 β_1 的估计值都为负, 而在状态 0 中自回归参数 β_0 的估计值都为正 (除橡胶期货外)。根据 σ_i^2 和 d_i 的大小, 可以将状态 0 称为稳定波动状态, 状态 1 称为高波动状态。此外, 从表 7 中还可以看出, 所有农产品期货市场的已实现波动率序列的 ρ_0 ($\rho_0 = 1 - \rho_1$) 估计值均高于 ρ_1 ($\rho_1 = 1 - \rho_0$), 但都很接近 1, 这说明从 $t-1$ 期到 t 期两种波动状态之间的转换概率较小, 都趋向于更会保持原有的波动状态, 而稳定波动状态的稳定性比高波动状态的强。

图 1 是 $MS-ARFIMA(1, d_i, 0)$ 模型的高波动率状态概率 ($\Pr(S_t = 1 | Y)$) 图。从图 1 中我们可以清楚的看到, 绝大部分的农产品期货的已实现波动率序列都呈现出比较有序的波动状态转换。在 2008 年初至 2009 年期间以及 2010 年 9 月至 2011 年 1 月期间, 农产品期货市场较多的表现为高波动率状态, 而在其他时间区间上, 农产品期货市场较多的表现为稳定波动率状态。这两次结构性转变与 2007-2009 年期间的美国次贷危机以及 2011-2012 年期间的欧债危机的时间较为吻合, 说明美国次贷危机和欧债危机对我国农产品期货市场有较大影响, 在此期间, 我国农产品期货市场进入高波动性主导时期, 本文构建的 $MS-ARFIMA(1, d_i, 0)$ 模型较好地提取了我国农产品期货市场的区制转换性。

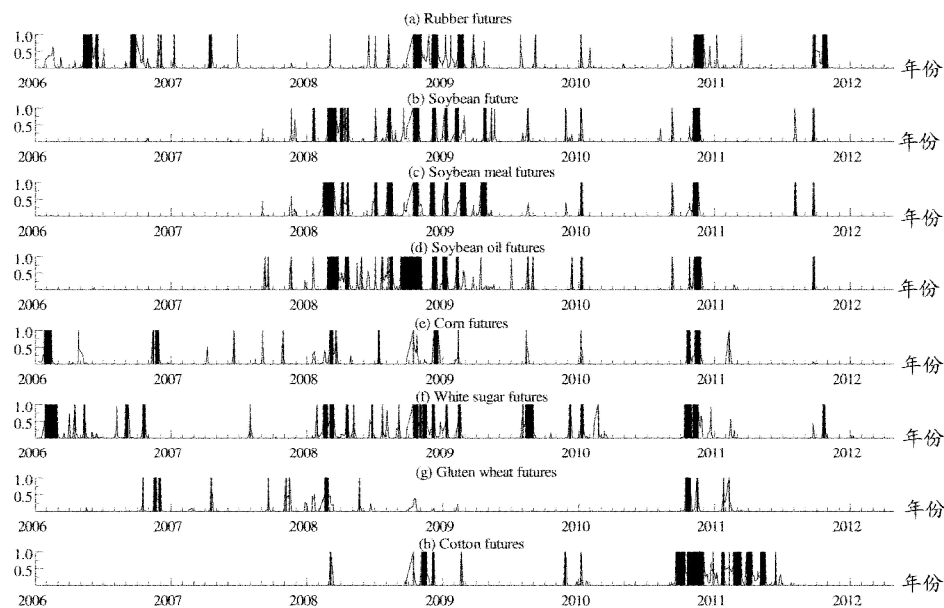


图 1 高波动率状态概率图

图 2 给出了所有农产品期货的已实现波动率长记忆性的路径图。从图 2 可知, 所有农产品期货的已实现波动率序列较为明显的呈现出两个层次的长记忆性, 高波动率状态下的长记忆性比稳定状态下的长记忆性强 ($d_0 < d_1$), 但都落在稳定区域 ($0 < d < 0.5$) 内, 说明我国农产品期货市场的已实现波动率具有很强的长记忆性, 这种动态特征由不同波动状态下的长记忆成分主导, 意味着对市场波动的冲击效应将会持续较长一段时间, 之后才会缓慢转换到稳定状态。从图 2 中长记忆性的状态转换时点可以看出, 在 2008 年初和 2010 年 9 月, 我国农产品期货市场进入高波动性主导时期, 同时具有较高的长记忆性参数估计值, 意味着在全球市场动荡

时期,我国农产品期货市场的波动性的记忆性比稳定时期强,进一步说明金融危机对我国农产品期货市场波动性的影响主要是结构性和长期性的。

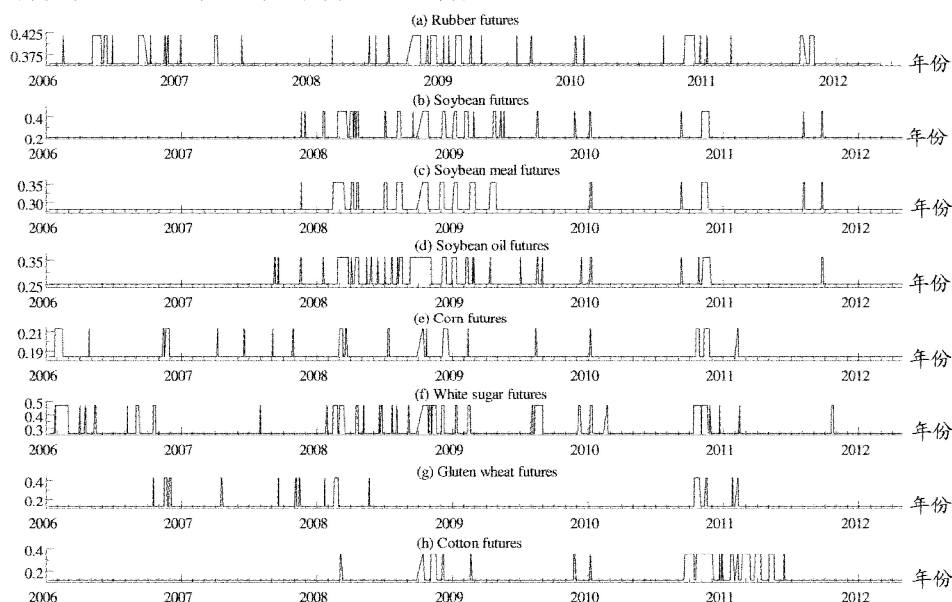


图2 长记忆性参数的估计值

(四) 样本外预测分析

本文还评价和比较了4类常见的预测模型和MS-ARFIMA(1, d_i , 0)模型(记为Mod 1)的向前1天、5天和20天的点预测和密度预测。样本外的点预测值和密度预测值通过滚动固定时间窗口(固定为 N)的序贯估计得到。本文使用的总体样本数据从2006年1月9日至2012年4月27日,将估计窗口固定为 $N=1\,000$,不同预测模型可以分别产生497、493和478个向前1天、5天(1周)以及20天(1月)的样本外点预测值和密度预测值。

我们考察的第1个嵌套模型具有固定的 β 和长记忆参数值,但均值和条件方差具有状态依赖性,记为Mod 2:

$$y_t = \mu_{s_t} + \beta y_{t-1} + (1-L)^{-d} \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma_{s_t}^2) \quad (16)$$

第2个嵌套模型为ARFIMA(1, d , 0)模型,记为Mod 3:

$$y_t = \mu + \beta y_{t-1} + (1-L)^{-d} \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma^2) \quad (17)$$

第3个嵌套模型为短记忆的马尔科夫区制转换模型,记为Mod 4:

$$y_t = \mu_{s_t} + \beta_{s_t} y_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma_{s_t}^2) \quad (18)$$

第4个嵌套模型为AR(1)模型,记为Mod 5:

$$y_t = \mu + \beta y_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma^2) \quad (19)$$

表8给出了所有预测模型在损失函数MSE和QLIKE评价标准下的经过10 000次Bootstrap模拟过程后得到的MCS检验结果。由表8可知:(1)当预测时域为1天时,模型置信集MCS在10%的置信水平和损失函数MSE的评价标准下都包含了Mod 1至Mod 4,只有Mod 5排除在外,说明在损失函数MSE的评价标准下,Mod 1至Mod 4对农产品期货的已实现波动率的预测性能在统计上是无差异的,但在损失函数QLIKE的评价标准下,模型置信集MCS在

10% 的置信水平下仅包含 MS - ARFIMA(1, d_i , 0) 模型 (除强筋小麦期货外), 两种损失函数标准下检验结果的不一致的原因是: QLIKE 损失函数对带有噪声的波动率代理变量更加稳健, 在评价和比较各种波动率模型的预测精度的差异性方面更有说服力; (2) 当预测时域为 5 天时, 模型置信集 MCS 在 10% 的置信水平和损失函数 MSE 的评价标准下仅包含 MS - ARFIMA(1, d_i , 0) 模型和 Mod 2, 而 Mod 3 和 Mod 4 被排除在外, 而在 QLIKE 损失函数评价标准下, MS - ARFIMA(1, d_i , 0) 模型是唯一被置信集 MCS 所包含的模型; (3) 当预测时域为 20 天时, 模型置信集 MCS 在 10% 的置信水平和两种损失函数评价标准下都只包含 MS - ARFIMA(1, d_i , 0) 模型。这些发现进一步表明引入马尔科夫区制转换动态和长记忆性动态可以提高预测模型对农产品期货已实现波动率的预测性能, 并且 MS - ARFIMA(1, d_i , 0) 模型在所有预测时域上的预测性能都明显的好于其他嵌套模型, 尤其是在长期预测时域上。

表 8 MCS 检验结果

	向前 1 天				向前 5 天				向前 20 天			
	MSE		QLIKE		MSE		QLIKE		MSE		QLIKE	
	T_R	T_{SQ}	T_R	T_{SQ}	T_R	T_{SQ}	T_R	T_{SQ}	T_R	T_{SQ}	T_R	T_{SQ}
橡胶												
Mod 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Mod 2	0.744	0.827	0.077	0.082	0.376	0.388	0.016	0.054	0.064	0.026	0.051	0.038
Mod 3	0.283	0.221	0.035	0.036	0.014	0.008	0.003	0.002	0.064	0.027	0.051	0.038
Mod 4	0.283	0.221	0.035	0.036	0.014	0.008	0.003	0.002	0.064	0.027	0.050	0.038
Mod 5	0.040	0.050	0.000	0.001	0.001	0.000	0.001	0.000	0.006	0.027	0.042	0.032
黄豆												
Mod 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Mod 2	0.711	0.644	0.064	0.082	0.324	0.359	0.009	0.012	0.074	0.055	0.042	0.031
Mod 3	0.381	0.188	0.002	0.000	0.026	0.004	0.001	0.002	0.067	0.020	0.042	0.031
Mod 4	0.381	0.188	0.002	0.000	0.026	0.004	0.001	0.002	0.067	0.020	0.042	0.027
Mod 5	0.002	0.092	0.000	0.000	0.000	0.000	0.000	0.000	0.010	0.014	0.012	0.027
豆粕												
Mod 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Mod 2	0.687	0.565	0.053	0.064	0.248	0.250	0.010	0.016	0.024	0.036	0.029	0.022
Mod 3	0.201	0.472	0.032	0.037	0.023	0.011	0.002	0.000	0.024	0.032	0.029	0.022
Mod 4	0.201	0.472	0.032	0.037	0.023	0.011	0.002	0.000	0.024	0.028	0.016	0.020
Mod 5	0.010	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.010	0.014
豆油												
Mod 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Mod 2	0.774	0.810	0.094	0.097	0.442	0.320	0.016	0.015	0.088	0.042	0.054	0.030
Mod 3	0.263	0.312	0.001	0.003	0.041	0.056	0.018	0.017	0.043	0.030	0.054	0.030
Mod 4	0.263	0.312	0.001	0.002	0.041	0.056	0.008	0.006	0.043	0.030	0.052	0.025
Mod 5	0.005	0.003	0.000	0.000	0.000	0.000	0.001	0.001	0.011	0.014	0.021	0.009
玉米												
Mod 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Mod 2	0.691	0.753	0.072	0.079	0.184	0.220	0.026	0.025	0.048	0.022	0.040	0.031
Mod 3	0.528	0.344	0.024	0.038	0.092	0.083	0.014	0.008	0.016	0.018	0.012	0.030
Mod 4	0.528	0.406	0.024	0.038	0.092	0.083	0.014	0.008	0.014	0.018	0.012	0.030
Mod 5	0.000	0.001	0.002	0.002	0.000	0.000	0.001	0.002	0.014	0.009	0.008	0.007

续表 8 MCS 检验结果

	向前 1 天				向前 5 天				向前 20 天			
	MSE		QLIKE		MSE		QLIKE		MSE		QLIKE	
	T_R	T_{SQ}	T_R	T_{SQ}	T_R	T_{SQ}	T_R	T_{SQ}	T_R	T_{SQ}	T_R	T_{SQ}
白砂糖												
Mod 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Mod 2	0.288	0.412	0.044	0.026	0.110	0.119	0.009	0.012	0.018	0.016	0.020	0.014
Mod 3	0.174	0.165	0.016	0.026	0.002	0.004	0.006	0.009	0.017	0.012	0.017	0.012
Mod 4	0.174	0.165	0.016	0.011	0.002	0.004	0.006	0.009	0.017	0.012	0.016	0.012
Mod 5	0.082	0.002	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.002	0.002
强筋小麦												
Mod 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Mod 2	0.911	0.948	0.112	0.128	0.407	0.430	0.022	0.038	0.090	0.073	0.044	0.040
Mod 3	0.247	0.297	0.001	0.000	0.088	0.092	0.009	0.014	0.067	0.060	0.031	0.026
Mod 4	0.247	0.315	0.001	0.000	0.088	0.092	0.009	0.014	0.067	0.042	0.030	0.026
Mod 5	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.024	0.029	0.030	0.020
棉花												
Mod 1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Mod 2	0.483	0.527	0.038	0.045	0.214	0.307	0.044	0.068	0.051	0.048	0.013	0.025
Mod 3	0.483	0.126	0.001	0.001	0.065	0.070	0.010	0.009	0.027	0.020	0.012	0.025
Mod 4	0.483	0.126	0.001	0.002	0.065	0.070	0.010	0.009	0.027	0.020	0.012	0.025
Mod 5	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.001	0.000

为了比较各种预测模型的样本外密度预测性能,表 9 报告了 WLR 检验的 t 统计量及对应的 p 值。表中 t 统计量若为负值同时对应的 p 值非常小,则表明可选模型的样本外密度预测精度比基础模型的高。由表 9 可知:(1)在所有预测时域上,AR(1)模型(Mod 5)是所有预测模型中的样本外密度预测精度最低的模型,这一结论与基于 MSE、QLIKE 以及 MCS 检验得到的点预测的评价和比较结论一致;(2)MS-ARFIMA(1, d_i ,0)模型和 Mod 2 的样本外密度预测精度都要比 Mod 3 和 Mod 4 高,说明分别引入马尔科夫区制转换动态和长记忆性动态同样可以提高预测模型对我国农产品期货已实现波动率的样本外密度预测精度;(3)当预测时域为 1 天和 5 天时,MS-ARFIMA(1, d_i ,0)模型与 Mod 2 的样本外密度预测精度十分接近,其差异在统计上不显著,但当预测时域为 20 天时,MS-ARFIMA(1, d_i ,0)模型的样本外密度预测精度显著地比 Mod 2 高,这一发现表明考虑分数阶积分过程的区制转换特征可以在长期预测时域上进一步提高模型的密度预测精度;(4)与表 8 和表 9 得到的点预测的评价和比较结论不同的是,短记忆的马尔科夫区制转换模型(Mod 4)在所有预测时域上对我国农产品期货已实现波动率的样本外密度预测精度都比 ARFIMA(1, d ,0)模型(Mod 3)高,表明相比于长记忆性,区制转换是影响预测模型的样本外密度预测精度更为重要的因素。

综上所述,同时引入已实现波动率的长记忆性和马尔科夫区制转换可以进一步提高预测模型对我国农产品期货已实现波动率的预测精度(包括点预测和密度预测),进一步考虑分数阶积分过程的时变性的 MS-ARFIMA(1, d_i ,0)模型是所有预测模型中预测性能最好的模型。

表 9 WLR 检验结果

基础模型	向前 1 天				向前 5 天				向前 20 天			
	可选模型				可选模型				可选模型			
	Mod2	Mod3	Mod4	Mod5	Mod2	Mod3	Mod4	Mod5	Mod2	Mod3	Mod4	Mod5
橡胶												
Mod1	-0.092 (0.931)	9.548 (<0.000)	5.477 (<0.000)	97.870 (<0.000)	0.514 (0.562)	10.925 (<0.000)	5.382 (<0.000)	28.924 (<0.000)	4.853 (<0.000)	22.636 (<0.000)	6.458 (<0.000)	16.044 (<0.000)
Mod2		8.430 (<0.000)	4.652 (<0.000)	47.247 (<0.000)		16.514 (<0.000)	4.674 (<0.000)	23.724 (<0.000)		19.547 (<0.000)	5.389 (<0.000)	12.641 (<0.000)
Mod3			-72.043 (<0.000)	32.382 (<0.000)			-36.533 (<0.000)	9.045 (<0.000)			-9.205 (<0.000)	12.537 (<0.000)
Mod4				40.529 (<0.000)				18.630 (<0.000)				15.384 (<0.000)
黄豆												
Mod1	0.174 (0.842)	7.439 (<0.000)	4.977 (<0.000)	125.033 (<0.000)	-1.037 (0.163)	16.447 (<0.000)	6.592 (<0.000)	15.206 (<0.000)	4.312 (<0.000)	18.077 (<0.000)	5.472 (<0.000)	12.385 (<0.000)
Mod2		6.504 (<0.000)	5.026 (<0.000)	58.430 (<0.000)		18.904 (<0.000)	5.308 (<0.000)	16.488 (<0.000)		14.392 (<0.000)	4.350 (<0.000)	12.679 (<0.000)
Mod3			-64.835 (<0.000)	28.947 (<0.000)			-28.472 (<0.000)	12.470 (<0.000)			-2.042 (0.039)	13.067 (<0.000)
Mod4				39.605 (<0.000)				21.393 (<0.000)				14.952 (<0.000)
豆粕												
Mod1	-0.437 (0.284)	9.403 (<0.000)	6.715 (<0.000)	105.371 (<0.000)	0.158 (0.874)	15.473 (<0.000)	7.476 (<0.000)	32.100 (<0.000)	4.320 (<0.000)	19.174 (<0.000)	6.185 (<0.000)	14.126 (<0.000)
Mod2		7.548 (<0.000)	5.022 (<0.000)	53.620 (<0.000)		17.194 (<0.000)	5.038 (<0.000)	25.644 (<0.000)		15.648 (<0.000)	5.884 (<0.000)	11.592 (<0.000)
Mod3			-43.174 (<0.000)	38.336 (<0.000)			-41.506 (<0.000)	12.148 (<0.000)			-8.372 (<0.000)	12.394 (<0.000)
Mod4				42.510 (<0.000)				19.740 (<0.000)				16.445 (<0.000)
豆油												
Mod1	0.188 (0.821)	6.089 (<0.000)	6.128 (<0.000)	68.302 (<0.000)	0.505 (0.576)	16.230 (<0.000)	7.464 (<0.000)	20.558 (<0.000)	4.368 (<0.000)	12.167 (<0.000)	4.994 (<0.000)	11.945 (<0.000)
Mod2		6.594 (<0.000)	5.485 (<0.000)	25.107 (<0.000)		18.175 (<0.000)	6.107 (<0.000)	21.196 (<0.000)		11.382 (<0.000)	5.179 (<0.000)	12.698 (<0.000)
Mod3			-57.643 (<0.000)	36.171 (<0.000)			-38.625 (<0.000)	3.487 (0.016)			-8.841 (<0.000)	12.022 (<0.000)
Mod4				32.612 (<0.000)				8.953 (<0.000)				15.841 (<0.000)
玉米												
Mod1	0.244 (0.792)	8.346 (<0.000)	5.826 (<0.000)	92.001 (<0.000)	0.166 (0.824)	14.306 (<0.000)	5.473 (<0.000)	24.730 (<0.000)	3.924 (0.008)	21.374 (<0.000)	7.182 (<0.000)	9.472 (<0.000)
Mod2		9.077 (<0.000)	5.063 (<0.000)	42.255 (<0.000)		15.527 (<0.000)	4.548 (<0.000)	25.157 (<0.000)		16.470 (<0.000)	5.322 (<0.000)	9.182 (<0.000)
Mod3			-70.521 (<0.000)	39.383 (<0.000)			-32.162 (<0.000)	9.275 (<0.000)			-8.746 (<0.000)	9.496 (<0.000)
Mod4				36.412 (<0.000)				16.544 (<0.000)				12.646 (<0.000)

续表 9 WLR 检验结果

基础模型	向前 1 天				向前 5 天				向前 20 天			
	可选模型				可选模型				可选模型			
	Mod2	Mod3	Mod4	Mod5	Mod2	Mod3	Mod4	Mod5	Mod2	Mod3	Mod4	Mod5
白砂糖												
Mod1	-1.110 (0.264)	8.621 (<0.000)	7.338 (<0.000)	86.305 (<0.000)	0.162 (0.822)	12.384 (<0.000)	5.142 (<0.000)	27.672 (<0.000)	4.758 (<0.000)	14.567 (<0.000)	8.150 (<0.000)	11.832 (<0.000)
Mod2		6.863 (<0.000)	4.927 (<0.000)	28.454 (<0.000)		16.219 (<0.000)	5.463 (<0.000)	21.447 (<0.000)		18.644 (<0.000)	6.809 (<0.000)	16.685 (<0.000)
Mod3			-38.269 (<0.000)	24.526 (<0.000)			-35.044 (<0.000)	9.928 (<0.000)			-7.541 (<0.000)	16.364 (<0.000)
Mod4				28.395 (<0.000)				12.306 (<0.000)				15.435 (<0.000)
强筋小麦												
Mod1	-1.412 (0.163)	7.894 (<0.000)	8.306 (<0.000)	76.691 (<0.000)	0.249 (0.790)	9.668 (<0.000)	6.494 (<0.000)	23.000 (<0.000)	4.106 (<0.000)	16.711 (<0.000)	9.605 (<0.000)	12.659 (<0.000)
Mod2		8.434 (<0.000)	6.593 (<0.000)	21.363 (<0.000)		12.450 (<0.000)	7.921 (<0.000)	19.185 (<0.000)		12.547 (<0.000)	8.864 (<0.000)	12.361 (<0.000)
Mod3			-46.037 (<0.000)	20.055 (<0.000)			-30.184 (<0.000)	9.363 (<0.000)			-8.117 (<0.000)	15.004 (<0.000)
Mod4				18.484 (<0.000)				13.852 (<0.000)				14.725 (<0.000)
棉花												
Mod1	-0.937 (0.348)	12.446 (<0.000)	6.408 (<0.000)	64.473 (<0.000)	0.169 (0.825)	6.790 (<0.000)	6.100 (<0.000)	23.036 (<0.000)	3.944 (0.007)	20.107 (<0.000)	5.644 (<0.000)	14.159 (<0.000)
Mod2		9.507 (<0.000)	5.385 (<0.000)	28.407 (<0.000)		12.438 (<0.000)	5.366 (<0.000)	23.284 (<0.000)		14.371 (<0.000)	5.903 (<0.000)	17.394 (<0.000)
Mod3			-65.144 (<0.000)	22.015 (<0.000)			-42.139 (<0.000)	6.160 (<0.000)			-8.747 (<0.000)	11.065 (<0.000)
Mod4				28.144 (<0.000)				15.372 (<0.000)				13.392 (<0.000)

五、研究结论

本文首先采用最新发展起来的四种伪长记忆性检验方法检验橡胶、黄豆、豆粕、豆油、玉米等 8 类农产品期货的已实现波动率序列在 2006 年 1 月 9 日至 2012 年 4 月 27 日间的长记忆性特征的真伪性;然后,运用同时考虑波动率的长记忆性和区制转换性的长记忆马尔科夫区制转换模型($MS - ARFIMA(1, d_i, 0)$)来考察我国农产品期货已实现波动率的长记忆性和区制转换;最后,采用 MCS 检验和加权似然比(WLR)检验,评价和比较 $MS - ARFIMA(1, d_i, 0)$ 模型及其 4 种嵌套模型对我国农产品期货市场波动率的预测性能。

研究结论表明:(1)纯粹的分数阶积分过程或纯粹的结构突变模型都不能全部解释我国农产品期货的已实现波动率序列所表现出的强长记忆性;(2)我国农产品期货的已实现波动率明显地表现出两种不同的波动状态,一种是高波动率状态,另一种是稳定状态,市场都趋向于更会保持原有的波动状态,而由高波动率状态转换到低波动率状态的概率较大,稳定波动状态的稳定性比高波动状态的强,其原因是在高波动率状态的持续时间上会有大量的政策性冲击;(3)绝大部分农产品期货的已实现波动率序列都呈现出比较有序的波动状态转换,结构突

变的时点与美国次贷危机以及欧债危机的时间较为吻合;(4)同时引入已实现波动率的长记忆性和马尔科夫区制转换可以进一步提高预测模型对我国农产品期货已实现波动率的预测精度,而进一步考虑分数阶积分过程的时变性的 $MS-ARFIMA(1, d_i, 0)$ 模型是所有预测模型中预测性能最好的模型。本文的研究对于投资者有效规避投资风险以及管理层科学合理发挥农产品期货的市场功能,最终促进金融市场平稳较快发展具有重要的理论价值和现实意义。

参考文献:

1. 刘向丽,成思危、汪寿阳、洪永森,2012:《基于 ACD 模型的中国期货市场波动性》,《系统工程理论与实践》第 2 期,第 268 – 273 页。
2. 王辉、孙志凌、谢幽篁,2012:《中国农产品期货套期保值非对称效应研究》,《统计研究》第 7 期,第 68 – 74 页。
3. 王金媛,2009:《我国农产品期货价格波动率分析》,《东北农业大学学报》(社会科学版)第 3 期,第 30 – 31 页。
4. 魏宇,2009a:《中国商品期货市场的风险价值模型及其后验分析》,《财贸经济》第 2 期,第 63 – 68 页。
5. 魏宇,2009b:《金融市场的多分形波动率测度、模型及其 SPA 检验》,《管理科学学报》第 5 期,第 88 – 99 页。
6. 杨科、陈浪南,2012:《股市波动率的短期预测模型和预测精度评价》,《管理科学学报》第 5 期,第 19 – 31 页。
7. 陈浪南,杨科,2013:《中国股市高频波动率的特征、预测模型以及预测精度比较》,《系统工程理论与实践》第 2 期,第 296 – 307 页。
8. Amisano, G., and R. Giacomini. 2007. “Comparing Density Forecasts via Weighted Likelihood Ratio Tests.” *Journal of Business & Economic Statistics*, 25(2):177 – 190.
9. Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens. 2003. “Modeling and Forecasting Realized Volatility.” *Econometrica*, 71(3):579 – 625.
10. Beltratti, A., and C. Morana. 2005. “Statistical Benefits of Value – at – Risk with Long Memory.” *Journal of Risk*, 7(4):21 – 45.
11. Chib, S. 1996. “Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models.” *Journal of Econometrics*, 75(1):79 – 97.
12. Choi, K., W. C. Yu, and E. Zivot. 2010. “Long Memory versus Structural Breaks in Modeling and Forecasting Realized Volatility.” *Journal of International Money and Finance*, 29(5):857 – 875.
13. Choi, K., and E. Zivot. 2007. “Long Memory and Structural Changes in the Forward Discount: An Empirical Investigation.” *Journal of International Money and Finance*, 26(3):342 – 363.
14. Coakley, J., J. Dollery, and N. Kellard. 2011. “Long Memory and Structural Breaks in Commodity Futures Markets.” *Journal of Futures Markets*, 31(11):1076 – 1113.
15. Corsi, F., S. Mittnik, and U. Pigorsch. 2008. “The Volatility of Realized Volatility.” *Econometric Reviews*, 27(2):46 – 78.
16. Corsi, F. 2009. “A Simple Approximate Long Memory Model of Realized Volatility.” *Journal of Financial Econometrics*, 7(4):174 – 196.
17. Corsi, F., and R. Reno. 2009. “HAR Volatility Modeling With Heterogeneous Leverage and Jumps.” Working Paper, University of Siena. Available at <http://www.econometrics.org/ecaresearch/seminars0910/corsi2.pdf>.
18. Degiannakis, S. 2008. “ARFIMAX and ARFIMAX – TARCH Realized Volatility Modeling.” *Journal of Applied Statistics*, 35(1):1169 – 1180.
19. Diebold, F. X., and A. Inoue. 2001. “Long Memory and Regime Switching.” *Journal of Econometrics*, 105(1):131 – 159.
20. Giot, P., and S. Laurent. 2004. “Modeling Daily Value – at – Risk Using Realized Volatility and ARCH Type Models.” *Journal of Empirical Finance*, 11(3):379 – 398.
21. Hansen, P. R., and A. Lunde, and M. Nason. 2011. “The Model Confidence Set.” *Econometrica*, 79(2):453 – 497.
22. Liu, C., and J. M. Maheu. 2008. “Are There Structural Breaks in Realized Volatility?” *Journal of Financial Econometrics*, 6(3):326 – 360.

23. Lux, T. , and L. Morales – Arias. 2010. “Forecasting Volatility under Fractality, Regime – switching, Long Memory and Student – t Innovations.” *Computational Statistics & Data Analysis*, 54(11) :2676 – 2692.
24. Martens, M. 2002. “Measuring and Forecasting S&P 500 Index – futures Volatility Using High – frequency Data.” *Journal of Futures Markets*, 22(6) : 497 – 518.
25. McAleer, M. , and M. C. Medeiros. 2008. “A Multiple Regime Smooth Transition Heterogeneous Autoregressive Model for Long Memory and Asymmetries.” *Journal of Econometrics*, 147(1) :104 – 119.
26. Perron, P. , and Z. Qu. 2010. “Long – memory and Level Shifts in the Volatility of Stock Market Return Indices.” *Journal of Business & Economic Statistics*, 28(2) :275 – 290.
27. Qu, Z. 2011. “A Test against Spurious Long Memory.” *Journal of Business & Economic Statistics*, 29(3) :423 – 438.
28. Ray, B. K. , and R. S. Tsay. 2002. “Bayesian Methods for Change – point Detection in Long – range Dependent Processes.” *Journal of Time Series Analysis*, 23(6) :687 – 705.
29. Shimotsu, K. 2006. “Simple (but effective) Tests of Long Memory versus Structural Breaks.” Queen’ s Economics Dept. Working Paper, 1101 :1 – 46.
30. Yang, K. , and L. Chen. 2014. “Realized Volatility Forecast: Structural Breaks, Long Memory, Asymmetry, and Day – of – the – Week Effect.” *International Review of Finance*, Forthcoming. Available at <http://onlinelibrary.wiley.com/doi/10.1111/irfi.12030/abstract;jsessionid=902CB5239D597699474EE1867945EBF4.f01t02?deniedAccessCustomisedMessage=&userIsAuthenticated=false>.

Realized Volatility Dynamics and Forecasting Model for Agricultural Commodity Futures

Yang Ke¹ and Tian Fengping²

(1 : South China Agricultural University, College of Economics & Management;

2 : Sun Yat – sen University, International Business School)

Abstract: This paper empirically analyzes the dynamic properties of the realized volatility in China’ s agricultural commodity future markets by employing the high – frequency data of eight agricultural commodity futures, and finds that the realized volatility exhibits both long memory and regime switching features. To capture these properties simultaneously, we modeled the realized volatility of agricultural commodity futures in China’ s future markets by a MS – ARFIMA model, and compared its forecasting performance with that of a variety of competing models which have been encompassed by the proposed model. The estimated results show that the dynamics of realized volatility for agricultural commodity futures are characterized by two levels of long memory: one associated with the low – volatility regime and the other one associated with the high – volatility regime, and the probability to stay in the low – volatility regime is higher than that of the high – volatility regime. The out – sample volatility forecast results show that the combination of long memory with switching regimes can improve the out – of – sample realized volatility forecast performance, and the proposed MS – ARFIMA model represents the superior out – of – sample realized volatility forecasts over those obtained using nested models in China’ s agricultural commodity future markets.

Key Words: Agricultural Commodity Futures; Realized Volatility; Long Memory; Regime Switching; Forecast

JEL Classification: G10, C22, C53

(责任编辑:彭爽)