

中国股市波动率预测研究：基于实时已实现 EGARCH-MIDAS 模型

吴鑫育¹, 赵安¹, 谢海滨², 马超群³

(1. 安徽财经大学金融学院, 蚌埠 233030; 2. 对外经济贸易大学金融学院, 北京 100029; 3. 湖南大学工商管理学院, 长沙 410082)

摘要 本文构建了一个能够充分捕获高频数据信息、当前收益率信息以及波动率长记忆性的实时已实现 EGARCH-MIDAS (RT-REGARCH-MIDAS) 模型对中国股市波动率进行建模和预测。采用上证综合指数 (SSEC) 和深证成份指数 (SZSEC) 5 分钟高频数据进行实证研究, 结果表明: RT-REGARCH-MIDAS 模型相比其它众多竞争模型具有更好的收益率数据拟合效果, 能够更好地描述股市波动性。利用稳健的损失函数以及模型置信集 (MCS) 检验作为判断准则, 实证比较了该模型与其它竞争模型对中国股市波动率的样本外预测能力。实证结果表明: 捕获高频数据信息、当前收益率信息和波动率长记忆性对于股市波动率预测具有重要作用; 在众多竞争模型中, RT-REGARCH-MIDAS 模型具有最为优越的波动率预测能力。进一步, 采用不同的已实现测度、不同的预测窗口、不同的 MIDAS 滞后阶数、不同的预测期以及样本外 R^2 检验进行稳健性检验, 证实了该模型优越的波动率预测能力具有稳健性。最后, 通过考察模型波动择时策略发现, 该模型能够获得相比其它模型显著更高的投资组合经济价值。

关键词 波动率预测; 高频数据信息; 当前收益率信息; 波动率长记忆性; 波动择时

收稿日期: 2023-07-09

基金项目: 国家自然科学基金 (71971001); 安徽省自然科学基金 (2208085Y21); 安徽省高校杰出青年科研项目 (2022AH020047); 安徽省高校学科 (专业) 拔尖人才学术资助项目 (gxbjZD2022019); 安徽省高校优秀科研创新团队 (2022AH010045); 安徽高校协同创新项目 (GXXT-2021-078)

Supported by National Natural Science Foundation of China (71971001); Natural Science Foundation of Anhui Province (2208085Y21); Outstanding Youth Research Project for Anhui Universities (2022AH020047); Academic Funding Project for Top Academic Talents in Anhui Universities (gxbjZD2022019); Excellent Scientific Research and Innovation Team for Anhui Universities (2022AH010045); Anhui University Collaborative Innovation Project (GXXT-2021-078)

作者简介: 通信作者: 吴鑫育, 教授, 博士生导师, 研究方向: 金融工程与风险管理, E-mail: xywu@aufe.edu.cn; 赵安, 硕士研究生, 研究方向: 金融工程与风险管理, E-mail: anzhao97@126.com; 谢海滨, 教授, 博士生导师, 研究方向: 金融工程与风险管理, E-mail: hbxi@amss.ac.cn; 马超群, 教授, 博士生导师, 研究方向: 金融工程与风险管理, E-mail: cqma1998@126.com.

Forecasting Chinese Stock Market Volatility: A Real-Time Realized EGARCH-MIDAS Model

WU Xinyu¹, ZHAO An¹, XIE Haibin², MA Chaoqun³

(1. School of Finance, Anhui University of Finance and Economics, Bengbu 233030, China; 2. School of Banking and Finance, University of International Business and Economics, Beijing 100029, China; 3. Business School, Hunan University, Changsha 410082, China)

Abstract This paper proposes the real-time Realized EGARCH-MIDAS (RT-REGARCH-MIDAS) model which adequately captures the information content of high-frequency data, the current return information and the long memory of volatility to model and forecast Chinese stock market volatility. An empirical analysis based on the 5-minute high-frequency data of the Shanghai Stock Exchange Composite Index (SSEC) and the Shenzhen Stock Exchange Component Index (SZSEC) shows that the RT-REGARCH-MIDAS model outperforms a variety of competitor models in fitting the return data and can describe the stock market volatility better. Using robust loss functions and the model confidence set (MCS) test, the paper compares the out-of-sample forecasting ability of the model and other competitor models for Chinese stock market volatility. Our empirical results show that accounting for the information content of high-frequency data, the current return information and the long memory of volatility plays an important role in forecasting stock market volatility. As a consequence, the proposed RT-REGARCH-MIDAS model performs the best in forecasting Chinese stock market volatility. Further, according to the robustness checks, the superior volatility forecasting ability of the model is robust to alternative realized measure, alternative forecast windows, alternative MIDAS lags, alternative forecasting horizons and out-of-sample R^2 test. Finally, a volatility timing strategy shows that the proposed model yields more significant economic value of portfolio compared to the other models.

Keywords volatility forecasting; information content of high-frequency data; current return information; long memory volatility; volatility timing

1 引言

波动性是股市的重要特征之一。对股市波动率进行准确的估计和预测一直以来是金融计量学领域研究的核心内容, 其对于资产组合配置、金融风险管理以及期权定价等都具有十分重要的理论价值和现实意义。随着全球市场一体化程度的提高和经济的不断发展, 中国已经成为全球第二大经济体。与此同时, 中国股市已经成为全球最大的股市之一, 近年来其产生的剧烈波动引起了广泛关注。准确的建模和预测中国股市波动率对于投资者、风险管理者和监管者等都具有重要意义。本文的研究目的是提出一种新的波动率模型来对中国股市波动率进行合理建模和预测。

目前, 最为流行的波动率模型主要有 Bollerslev (1986) 提出的 GARCH 模型和 Taylor (1986) 提出的随机波动率 (SV) 模型。这两类模型在文献中和实际应用中均获得了广泛的

关注. SV 模型在拟合金融时间序列数据上比 GARCH 模型更为灵活, 但其波动率的动态变化由一阶自回归过程描述, 波动率是一个不可观测变量, 模型实现起来相对较为困难. 而在 GARCH 模型中, 波动率由可观测到的历史信息 (例如资产收益率) 驱动, 具有结构简单、直观和易于实现等优点.

随着信息技术的快速发展, 日内高频数据的获得变得越来越容易. 采用日内高频数据构建已实现测度 (realized measure) 对波动率进行估计成为金融计量学研究的热点问题. 已实现测度的优点是构造简单且充分利用了交易日内的信息, 其相比日度收益率数据包含更多的关于当前波动率水平的信息. 但是已实现测度本质还是反映历史信息. 已有研究表明, 仅利用历史信息来建模和预测波动率, 而忽略当前的收益率信息, 并不能适应金融市场的快速变化 (Politis, 2007; Breitung and Hafner, 2016). 引入随机新息 (实时信息) 的 SV 模型由于属于参数驱动模型, 其参数估计和统计推断相比于 GARCH 模型更为复杂. 鉴于此, Smetanina (2017) 同时结合 GARCH 和 SV 模型的优点, 引入当前收益率信息, 提出实时 GARCH (RT-GARCH) 模型来对波动率进行建模和预测.

相较于传统的单成分波动率模型, 将波动率乘性分解为一个短期成分和一个长期成分的双成分波动率模型在刻画波动率的长记忆性方面具有明显优势 (Maheu, 2005; Ding, 2016). 近年来最为典型的是 Engle et al. (2013) 提出的基于乘性成分分解的 GARCH-MIDAS 模型. 该模型由于其优越的实证表现受到了学者们的广泛关注.

在现有关于波动率建模和预测的研究中, 大多数学者仅从考虑日内高频数据信息、当前收益率信息或波动率长记忆性的单一视角出发, 鲜有学者同时考虑日内高频数据信息、当前收益率信息和波动率长记忆性的综合影响. 鉴于此, 本文试图拓展已有研究, 同时引入高频数据信息、当前收益率信息和波动率长记忆性, 提出基于乘性成分分解的实时已实现 EGARCH-MIDAS (RT-REGARCH-MIDAS) 模型, 基于多视角融合来对中国股市波动率进行建模和预测.

2 文献综述

越来越多的研究表明, 充分考虑日内高频数据信息对波动率建模和预测至关重要 (龚旭等, 2020; 苑慧玲等, 2020; 鲁万波和亢晶浩, 2022; 刘小军等, 2022; 周辰月和崔文昊, 2023; Hansen and Lunde, 2005; Dobrev and Szerszen, 2010; Christoffersen et al., 2014). 特别地, Hansen et al. (2012) 在 GARCH 模型框架下, 提出对收益率和已实现测度联合建模的已实现 GARCH (RGARCH) 模型, 将日内与日间信息完美结合起来对波动率建模. RGARCH 模型结构简单、易于实现, 灵活的二次型杠杆函数的引入使得模型能够很好地描述波动率动态性 (非对称性), 度量方程的引入使得模型能够自动修正由于非交易时间和微观结构噪声导致的已实现测度的偏差, 日内高频信息的引入使得模型能够对股市波动率的急剧变化进行快速反应. 随后, 为了更好地捕获杠杆效应, Hansen and Huang (2016) 在 RGARCH 模型的基础上进一步提出了已实现 EGARCH (REGARCH) 模型. 近年来, 众多学者实证考察了 R(E)GARCH 模型在波动率预测、市场风险度量以及期权定价中的表现, 例如王天一和黄卓 (2015), 黄友珀等 (2015), 王天一等 (2018), 于孝建和王秀花 (2018), 蒋伟和顾研 (2019), 吴鑫育和侯信盟 (2020), 白娟娟和师荣蓉 (2021), 蔡光辉等 (2021), 高雷阜和李伟梅 (2022), 郭

宝才和项琳 (2022), 邢艳春和廖晗 (2023), Huang et al. (2017, 2020), Banulescu-Radu et al. (2018), Hansen et al. (2021).

最近, Smetanina (2017) 在传统 GARCH 模型中引入当前收益率信息, 提出 RT-GARCH 模型对波动率进行建模. 随后, Ding (2023) 拓展了 RT-GARCH 模型, 提出能够捕获当前收益率信息以及波动率的波动率的随机异方差自回归波动率 (stochastic heteroskedastic autoregressive volatility, SHARV) 模型. 实证结果表明, 与传统 GARCH 模型相比, 该模型具有更好的样本内拟合、更有效的参数估计以及更优越的样本外波动率和 VaR 预测表现.

传统 GARCH 模型属于单成分波动率模型, 对于刻画波动率的长记忆性存在不足. 为了捕获波动率长记忆性, 一些学者通过将条件方差分解为短期成分和长期成分提出了双成分波动率模型 (Ding and Granger, 1996; Engle and Lee, 1999). 最近, Engle et al. (2013) 提出了基于乘性成分分解的 GARCH-MIDAS 模型. 该模型在 GARCH 框架下引入 Ghysels et al. (2004, 2007) 提出的混频数据抽样 (mixed data sampling, MIDAS) 结构, 将条件方差乘性分解为一个短期成分和一个长期成分. GARCH-MIDAS 模型可以将具有不同频率的变量 (高频日度收益率与低频月度或季度宏观经济变量) 结合起来对波动率进行建模, 具有较大的建模灵活性. Wang and Ghysels (2015), Conrad and Kleen (2020) 证明 GARCH-MIDAS 模型能够有效捕捉到波动率的高持续性、结构突变和非平稳性等复杂动态特征. GARCH-MIDAS 模型一经提出, 在金融计量学文献中获得了广泛关注, 例如郑挺国和尚玉皇 (2014), 雷立坤等 (2018), 夏婷和闻岳春 (2018), 张一锋等 (2020), 梁超等 (2022), 柏建成等 (2022), Asgharian et al. (2013), Conrad and Loch (2015), Wang and Ghysels (2015), Dorion (2016), Pan et al. (2017), Conrad and Kleen (2020), Wang et al. (2022). 但值得指出的是, GARCH-MIDAS 模型仅利用了股市日度收益率 (收盘价) 信息对波动率建模, 忽略了股市日内价格变动的信息. 当处在金融市场高波动时期, 这类模型对于波动率的急剧变化往往无法及时反应. 因此, Borup and Jakobsen (2019) 在 GARCH-MIDAS 模型的基础上, 借鉴 REGARCH 模型的建模思路, 提出了 REGARCH-MIDAS 模型. 该模型进一步引入了已实现测度, 充分利用了交易日内的高频数据信息, 相比 GARCH-MIDAS 模型具有更为优越的波动率预测表现.

基于以上分析, 本文同时引入高频数据信息和当前收益率信息, 提出基于乘性成分分解的 RT-REGARCH-MIDAS 模型. 为了验证该模型的合理性与优越性, 选取上证综合指数 (SSEC) 和深证成份指数 (SZSEC) 5 分钟高频数据进行实证分析, 比较该模型与其它众多竞争模型 (包括 GARCH 模型、EGARCH 模型、GJR-GARCH 模型、REGARCH 模型、REGARCH-MIDAS 模型和 RT-GARCH 模型) 的收益率数据拟合效果. 进一步, 运用稳健的损失函数以及模型置信集 (model confidence set, MCS) 检验方法, 实证比较该模型与其它竞争模型对中国股市波动率的样本外预测能力. 通过采用不同的已实现测度、不同的预测窗口、不同的 MIDAS 滞后阶数、不同的预测期以及样本外 R^2 检验进行稳健性检验, 分析该模型波动率预测能力的稳健性. 最后, 通过检验波动择时策略, 考察该模型优越的波动率预测结果的经济价值.

3 模型与方法

3.1 RT-REGARCH-MIDAS 模型

设资产 (对数) 收益率定义为 $r_t = \log(P_t) - \log(P_{t-1})$, 其中 P_t 是第 t 日的资产收盘价. RT-REGARCH-MIDAS 模型假设 r_t 服从如下动态过程:

$$r_t = \mu + \sigma_t \varepsilon_t, \quad (1)$$

$$\sigma_t^2 = \tau_t g_t, \quad (2)$$

$$\varepsilon_t | \Phi_{t-1} \sim N(0, 1), \quad (3)$$

其中 Φ_{t-1} 是截止第 $t-1$ 日的信息集, $\mu = E[r_t | \Phi_{t-1}]$ 是资产收益率的条件均值¹, $\sigma_t^2 = \text{Var}[r_t | \Phi_{t-1}]$ 是资产收益率的条件方差, 被乘性分解为两个成分: 长期成分 τ_t 和短期成分 g_t . 本文假设短期成分 g_t 服从如下过程:

$$\log(g_t) = \beta \log(g_{t-1}) + \nu(\varepsilon_{t-1}) + \gamma(\log(\varepsilon_t^2) - \log(2) - \psi(1/2)) + \alpha u_{t-1}, \quad (4)$$

$$\log(x_t) = \xi + \phi \log(\sigma_t^2) + \delta(\varepsilon_t) + u_t, \quad (5)$$

$$u_t | \Phi_{t-1} \sim N(0, \sigma_u^2), \quad (6)$$

其中 x_t 是第 t 日的已实现测度 (例如已实现波动率), $\psi(\cdot)$ 是 digamma 函数, γ 描述当前收益率信息对短期成分的影响, $\nu(\varepsilon_t)$ 和 $\delta(\varepsilon_t)$ 是两个二次型杠杆函数, 用于捕获波动率非对称性, 可以写为:

$$\nu(\varepsilon_t) = \nu_1 \varepsilon_t + \nu_2 (\varepsilon_t^2 - 1), \quad (7)$$

$$\delta(\varepsilon_t) = \delta_1 \varepsilon_t + \delta_2 (\varepsilon_t^2 - 1). \quad (8)$$

上述杠杆设定具有灵活的形式, 满足 $E[\nu(\varepsilon_t)] = E[\delta(\varepsilon_t)] = 0$. 在平稳性条件 ($|\beta| < 1$) 下, $\log(g_t)$ 的无条件均值为 $E[\log(g_t)] = 0$.

注意到, 式 (4) 中引入了已实现测度的新息项 αu_{t-1} , 系数 α 反映已实现测度包含关于未来波动率信息的情形. 式 (5) 将事后已实现测度与事前波动率联系起来, u_{t-1} 是测量误差. 考虑到实际金融市场受到微观结构噪声和非交易时间的影响, 已实现测度并非真实波动率的完美测度, 为此式 (5) 中引入了偏差修正系数: ξ (比例修正系数) 和 ϕ (指数修正系数).

值得指出的是, 式 (4) 可以等价地写为:

$$\log(g_t) = \frac{\beta}{1+\gamma} \log(g_{t-1}) + \frac{1}{1+\gamma} (\nu(\varepsilon_{t-1}) + \gamma(\log(r_t - \mu)^2 - \log(\tau_t) - \log(2) - \psi(1/2)) + \alpha u_{t-1}). \quad (9)$$

由上式可以看出, 短期成分 g_t 由当前收益率信息 r_t 驱动, 并且可以基于 g_{t-1} 和 τ_t 更新.

¹我们可以将条件均值建模为一个更复杂的动态结构, 例如自回归模型. 但是对于股市日度收益率数据而言, 假设一个常数的条件均值通常是足够的.

对于长期成分 τ_t , 本文采用灵活的 MIDAS 回归方法进行建模. 长期成分 τ_t 可以表示为:

$$\log(\tau_t) = m + \theta \sum_{k=1}^K \varphi_k(\omega_1, \omega_2) y_{t-1,k}^{(N)}, \quad (10)$$

其中 $y_{t,k}^{(N)} = (1/N) \sum_{i=1}^N \log(x_{t-N(k-1)-i})$ 是对数已实现测度的 N 天均值. 本文与 Borup and Jakobsen (2019) 一致, 选取 $N = 22$, 即选取月度均值. $\varphi_k(\omega_1, \omega_2) \geq 0$ 是权重函数, 满足 $\sum_{k=1}^K \varphi_k(\omega_1, \omega_2) = 1$, K 是最大 MIDAS 滞后阶数. 文献中通常选择加权方式为灵活的 Beta 多项式函数 (Ghysels et al., 2007), 即设定:

$$\varphi_k(\omega_1, \omega_2) = \frac{(k/K)^{\omega_1-1} (1-k/K)^{\omega_2-1}}{\sum_{j=1}^K (j/K)^{\omega_1-1} (1-j/K)^{\omega_2-1}}. \quad (11)$$

根据 Engle et al. (2013), Asgharian et al. (2016) 的研究, 本文设定 $\omega_1 = 1$, $\omega_2 > 1$. 这可以保证权重函数是单调递减的, 即越近的观测值对当期的影响越大, 被赋予更大的权重. 由此, 式 (11) 可以改写为:

$$\varphi_k(\omega) = \frac{(1-k/K)^{\omega-1}}{\sum_{j=1}^K (1-j/K)^{\omega-1}}, \quad (12)$$

其中 ω 相应于式 (11) 中的 ω_2 , 它决定了权重函数的衰减速度.

在上述 RT-REGARCH-MIDAS 模型中, 式 (1)、(4)、(5) 和 (10) 分别称为收益方程、短期方程、度量方程和长期方程. 该模型能够充分捕获高频数据信息、当前收益率信息以及波动率长记忆性. 值得指出的是, 该模型包含了 REGARCH-MIDAS 模型和 REGARCH 模型作为特殊形式, 具有较强的一般性. 事实上, 当 $\gamma = 0$ 时, 该模型退化为 REGARCH-MIDAS 模型; 进一步, 约束 $\theta = 0$ (长期成分为常数), REGARCH-MIDAS 模型退化为 REGARCH 模型. 此外, 根据式 (2) 和 (4), RT-REGARCH-MIDAS 模型可以重新表述为一个实时时变系数自回归波动率模型:

$$\log(\sigma_t^2) = \omega_t + \beta \log(\sigma_{t-1}^2) + v_t, \quad (13)$$

其中系数 $\omega_t = (1-\beta) \log(\tau_t)$ 是时变的, 这允许模型可以自动及时适应当前的市场环境, 捕获更为复杂、丰富的波动率动态特征, 例如波动率结构性突变. 大量文献证实了波动率结构性突变会导致对波动率持续性的高估, 因而对波动率长记忆性具有重要的解释效力 (Diebold, 1986; Lamoureux and Lastrapes, 1990; 李木易和方颖, 2020). $v_t = \nu(\varepsilon_{t-1}) + \gamma(\log(\varepsilon_t^2) - \log(2) - \psi(1/2)) + \alpha v_{t-1}$ 是波动率更新项, 包含了当前收益率和已实现测度的新息项.

3.2 极大似然估计

RT-REGARCH-MIDAS 模型可以采用极大似然方法进行估计. 该模型的对数似然函数可以写为:

$$\begin{aligned}\ell(r, x; \Theta) &= \sum_{t=1}^T \log p(r_t, x_t | \Phi_{t-1}) \\ &= \underbrace{\sum_{t=1}^T \log p(r_t | \Phi_{t-1})}_{\ell(r; \Theta)} + \underbrace{\sum_{t=1}^T \log p(x_t | r_t, \Phi_{t-1})}_{\ell(x|r; \Theta)},\end{aligned}\quad (14)$$

其中 $\ell(r; \Theta)$ 是关于收益率的局部对数似然函数, $\ell(x|r; \Theta)$ 是关于已实现测度的局部对数似然函数. $\ell(r; \Theta)$ 和 $\ell(x|r; \Theta)$ 分别可以写为:

$$\ell(r; \Theta) = -\frac{1}{2} \sum_{t=1}^T [\log(2\pi) - \log(d_{2,t}^2) + d_{1,t}^2], \quad (15)$$

$$\ell(x|r; \Theta) = -\frac{1}{2} \sum_{t=1}^T [\log(2\pi) + \log(\sigma_u^2) + u_t^2/\sigma_u^2], \quad (16)$$

其中 $\Theta = (\mu, m, \theta, \omega, \alpha, \beta, \gamma, \nu_1, \nu_2, \xi, \phi, \sigma_u^2, \delta_1, \delta_2)'$ 是模型参数向量, $d_{1,t}$ 和 $d_{2,t}$ 分别可以写为:

$$d_{1,t} = \frac{(r_t - \mu)^{\frac{1}{1+\gamma}}}{(\tau_t \exp(-\gamma(\log(2) + \psi(1/2)) + \beta \log(g_{t-1}) + \nu(\varepsilon_{t-1}) + \alpha u_{t-1}))^{\frac{1}{2+2\gamma}}}, \quad (17)$$

$$d_{2,t} = \frac{1}{1 + \gamma} \frac{(r_t - \mu)^{-\frac{\gamma}{1+\gamma}}}{(\tau_t \exp(-\gamma(\log(2) + \psi(1/2)) + \beta \log(g_{t-1}) + \nu(\varepsilon_{t-1}) + \alpha u_{t-1}))^{\frac{1}{2+2\gamma}}}. \quad (18)$$

通过最大化对数似然函数 $\ell(r, x; \Theta)$, 即得到模型参数的极大似然估计为:

$$\hat{\Theta} = \arg \max_{\Theta} \ell(r, x; \Theta). \quad (19)$$

3.3 波动率预测能力评价

3.3.1 损失函数

本文采用损失函数来评价 RT-REGARCH-MIDAS 模型的波动率预测精确性. 具体地, 采用文献中常见的两种不同的损失函数来对波动率预测误差进行衡量, 这两种损失函数分别为均方误差 (MSE) 和拟似然误差 (QLIKE). 值得指出的是, MSE 和 QLIKE 是两个稳健的损失函数, 它们针对不完美的 (存在噪声的) 波动率代理变量也能给出一致的波动率预测评价结果 (Patton, 2011). 这两种损失函数分别定义为:

$$\text{MSE} : L_t(\sigma_t^2, \hat{\sigma}_t^2) = (\sigma_t^2 - \hat{\sigma}_t^2)^2, \quad (20)$$

$$\text{QLIKE} : L_t(\sigma_t^2, \hat{\sigma}_t^2) = \frac{\sigma_t^2}{\hat{\sigma}_t^2} + \log(\hat{\sigma}_t^2), \quad (21)$$

其中 σ_t^2 是真实的波动率, $\hat{\sigma}_t^2$ 是预测的波动率.

3.3.2 MCS 检验

为了考察模型预测结果之间的差异是否在统计上显著, 本文进一步采用 Hansen et al. (2011) 提出的模型置信集 (model confidence set, MCS) 方法检验模型预测结果的统计精度. MCS 程序检验一组给定的竞争模型, 并识别出一组具有一定可信度的最优预测模型或 MCS. 具体地, MCS 检验过程是基于一个等价检验 (equivalence test) $\delta_{\mathcal{M}}$ 和一个剔除规则 (elimination rule) $e_{\mathcal{M}}$ 来实现的. 设 \mathcal{M}^0 是所有竞争模型的初始集合. 设定 $\mathcal{M} = \mathcal{M}^0$, 利用等价检验 $\delta_{\mathcal{M}}$ 检验原假设: 竞争模型具有相同的期望损失 (预测能力), 表示为:

$$H_{0,\mathcal{M}} : E[d_{uv,t}] = 0, \quad \forall u, v \in \mathcal{M}, \quad (22)$$

其中 $d_{uv,t} = L_t(\sigma_t^2, \hat{\sigma}_t^2(u)) - L_t(\sigma_t^2, \hat{\sigma}_t^2(v))$ 是模型 u 和 v 波动率预测值的相对损失函数值. 为了对原假设 $H_{0,\mathcal{M}}$ 进行检验, 构建如下统计量:

$$T_{\mathcal{M}} = \max_{u,v \in \mathcal{M}} |t_{uv}|, \quad t_{uv} = \frac{\bar{d}_{uv}}{\sqrt{\widehat{\text{var}}(\bar{d}_{uv})}}, \quad (23)$$

其中 \bar{d}_{uv} 是模型 u 和 v 波动率预测值的相对损失函数的平均值, $\widehat{\text{var}}(\bar{d}_{uv})$ 是采用“自举法” (bootstrapping) 对 $\text{var}(\bar{d}_{uv})$ 的估计. 在显著性水平 α 下, 如果原假设 $H_{0,\mathcal{M}}$ 被接受, 则定义 $\widehat{\mathcal{M}}_{1-\alpha}^* = \mathcal{M}$, 否则使用剔除规则 $e_{\mathcal{M}} = \arg \max_{m \in \mathcal{M}} \sup_{n \in \mathcal{M}} t_{mn}$, 将拒绝原假设的模型从集合 \mathcal{M} 中剔除. 这一过程不断重复, 一直持续到不再出现拒绝原假设的情况. 最后, 得到 MCS 检验下的幸存模型集 (surviving objects) $\widehat{\mathcal{M}}_{1-\alpha}^* = \mathcal{M}$, 即在给定的置信水平 $1 - \alpha$ 下的最优预测模型集合或 MCS.

由于检验统计量 $T_{\mathcal{M}}$ 的渐近分布是非标准的, 本文采用基于 10000 次抽样的分块自举法 (block-bootstrapping) 进行近似计算. 在 MCS 检验中, 设定显著性水平为 $\alpha = 10\%$ ².

4 实证研究

4.1 数据

本文基于 RT-REGARCH-MIDAS 模型实证考察高频数据信息、当前收益率信息和波动率长记忆性对中国股市波动率预测的作用. 实证研究选取的数据为上证综合指数 (SSEC) 和深证成份指数 (SZSEC) 5 分钟高频数据. SSEC 和 SZSEC 指数数据的抽样阶段选取为 2005 年 1 月 4 日至 2022 年 12 月 30 日, 包括 4374 个交易日的数据. 数据来源于 Wind 数据库. 本文采用指数日度收盘价计算日度 (对数) 收益率, 并采用 5 分钟日内高频价格数据计算日度 RV, 以此作为本文模型中的已实现测度³.

表 1 给出了 SSEC 和 SZSEC 指数日度收益率和 RV 的描述性统计量. 从表 1 可以看到, 两指数收益率的均值均大于零, 但在统计上不显著; 偏度均小于 0, 峰度均大于 3, 表明两指数收益率具有负偏、尖峰厚尾分布特征; Jarque-Bera 统计量明显拒绝两指数收益率服从正态分布的假定. 两指数 RV 的偏度均大于 0, 峰度均明显大于 3, 表明其具有正偏、尖峰厚尾分

²选取显著性水平为 10% 是文献中较为常见的做法, 例如 Conrad and Kleen (2020).

³本文所使用的数据和代码请参见科学数据银行 (ScienceDB) 期刊社区, DOI: 10.57760/sciencedb.j00214.00012 和 CSTR: 31253.11.sciencedb.j00214.00012. 若使用文中数据信息, 请注明引文和数据出处.

布特征; Jarque-Bera 统计量明显拒绝其服从正态分布的假定. 比较两指数 RV 和 log-RV 的偏度、峰度和 Jarque-Bera 统计量可以看出, log-RV 的偏度更接近于 0, 峰度更接近于 3, 因此相较于 RV 而言, log-RV 的偏度和峰度更接近于正态分布的值, Jarque-Bera 统计量也明显更小, 说明 log-RV 的分布更接近于正态分布. Ljung-Box Q 统计量表明, SSEC 和 SZSEC 指数的波动率均表现出高的持续性 (显著的长期记忆特征).

表 1 SSEC 和 SZSEC 指数日度收益率和 RV 的描述性统计量

	SSEC			SZSEC		
	收益率	RV	log-RV	收益率	RV	log-RV
均值	0.0002	0.0002	-9.3737	0.0003	0.0002	-8.9776
最小值	-0.0926	0.0000	-12.0260	-0.0975	0.0000	-11.8412
最大值	0.0903	0.0039	-5.5450	0.0916	0.0053	-5.2414
标准差	0.0155	0.0003	1.0262	0.0180	0.0003	0.9506
偏度	-0.5808	6.0055	0.5530	-0.5045	5.7511	0.3732
峰度	7.9211	58.6474	3.0746	6.1537	56.7239	3.1590
Jarque-Bera	4658.4755	590517.3214	223.8753	1997.2329	549880.2496	106.0759
$Q(10)$	37.6847	10586.6970	23189.0082	33.1335	8780.2091	20829.5662
$Q(20)$	68.9416	17118.2262	40987.1194	49.4201	14130.7314	36288.0783

注: $Q(\cdot)$ 表示对应滞后阶数的 Ljung-Box Q 统计量.

4.2 参数估计结果

为了说明本文提出的 RT-REGARCH-MIDAS 模型的优越性, 本文考虑 6 个竞争模型, 包括 GARCH 模型、GJR-GARCH 模型、EGARCH 模型、REGARCH 模型、REGARCH-MIDAS 模型和 RT-GARCH 模型, 将其与 RT-REGARCH-MIDAS 模型进行实证比较分析.

为了实现包含 MIDAS 结构的模型, 需要确定最大 MIDAS 滞后阶数 K . 本文在实证中选取 $K = 22$, 即采用滞后一个月的 RV 来估计长期波动率. 此外, 与 Banulescu-Radu et al. (2018) 一致, 本文在 REGARCH 及其扩展模型 (REGARCH-MIDAS 和 RT-REGARCH-MIDAS 模型) 中设定 $\phi = 1$, 以简化模型结构, 改进模型样本外预测表现.

运用极大似然方法得到 RT-REGARCH-MIDAS 模型以及 6 个竞争模型的参数估计结果如表 2 和表 3 所示. 从表中 GARCH 模型、GJR-GARCH 模型、EGARCH 模型、REGARCH 模型和 RT-GARCH 模型的估计结果可以看到, SSEC 和 SZSEC 指数条件方差的持续性系数: $\alpha + \beta$ (GARCH/RT-GARCH)、 $\alpha + \beta + \zeta/2$ (GJR-GARCH) 和 β (EGARCH/REGARCH) 的估计值均接近于 1, 表明两个指数均展现非常高的波动率持续性.

在引入 MIDAS 结构的两个模型 (REGARCH-MIDAS 和 RT-REGARCH-MIDAS) 中, 条件方差被乘性分解为长期成分 (τ_t) 和短期成分 (g_t). 可以看到, 这些模型短期成分的持续性系数 β 的估计值相比未引入 MIDAS 结构的模型的条件方差的持续性系数的估计值明显要低, 表明 MIDAS 结构能够捕获波动率长记忆性. 在 REGARCH-MIDAS 模型和 RT-REGARCH-MIDAS 模型的长期成分中, θ 系数的估计值均显著为正, 表明 RV 对长期波动

表 2 参数估计结果: SSEC

	GARCH	GJR-GARCH	EGARCH	REGARCH	REGARCH-MIDAS	RT-GARCH	RT-REGARCH-MIDAS
μ	0.0002 (0.0002)	0.0001 (0.0002)	0.0001 (0.0002)	0.0003 (0.0002)	0.0001 (0.0002)	0.0007 (0.0002)	2.1006×10^{-5} (6.2843×10^{-6})
m	1.1710×10^{-6} (1.8237×10^{-7})	1.2005×10^{-6} (1.6821×10^{-7})	-0.0705 (0.0009)	-0.2756 (0.0017)	-0.7325 (0.0027)	-4.3509×10^{-6} (8.2172×10^{-7})	-2.8769 (0.0084)
θ				0.8503 (0.0017)	0.8503 (0.0017)		0.6764 (0.0025)
ω				12.6342 (0.0320)	12.6342 (0.0320)		23.0512 (0.0685)
α	0.0675 (0.0031)	0.0654 (0.0048)	0.1584 (0.0033)	0.3443 (0.0041)	0.3551 (0.0052)	0.0500 (0.0087)	0.4274 (0.0052)
β	0.9304 (0.0028)	0.9296 (0.0032)	0.9907 (0.0004)	0.9684 (0.0007)	0.7992 (0.0049)	0.9124 (0.0104)	0.8897 (0.0045)
γ						9.9326×10^{-6} (9.1063×10^{-7})	0.0829 (0.0025)
ζ		0.0052 (0.0052)	-0.0108 (0.0024)				
ν_1				-0.0534 (0.0024)	-0.0772 (0.0038)		-0.0671 (0.0029)
ν_2				0.0680 (0.0019)	0.0760 (0.0026)		0.0313 (0.0016)
ξ				-0.6819 (0.0051)	-0.6705 (0.0028)		-0.1710 (0.0073)
σ_u^2				0.2270 (0.0041)	0.2214 (0.0037)		0.2661 (0.0046)
δ_1				-0.0876 (0.0038)	-0.0921 (0.0042)		-0.1032 (0.0039)
δ_2				0.0891 (0.0020)	0.0886 (0.0022)		-0.0113 (0.0020)
$\ell(r)$	12721.5178	12721.7284	12726.0665	12805.1984	12822.2532	12811.3589	13507.8648
$\ell(r, x)$				9842.0027	9913.7527		10195.6116

注: () 中是极大似然估计的渐近标准误差, $\ell(r)$ 是关于收益率的局部极大对数似然值, $\ell(r, x)$ 是全极大对数似然值。

表3 参数估计结果: SZSEC

	GARCH	GJR-GARCH	EGARCH	REGARCH	REGARCH-MIDAS	RT-GARCH	RT-REGARCH-MIDAS
μ	0.0003 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)	-0.0001 (0.0002)	-0.0001 (0.0002)	0.0003 (0.0002)	0.0002 (8.9252×10^{-6})
m	2.8492×10^{-6} (3.6780×10^{-7})	3.1089×10^{-6} (3.9428×10^{-7})	-0.0950 (0.0014)	-0.3075 (0.0025)	-0.6753 (0.0081)	-7.4664×10^{-6} (1.2426×10^{-7})	-2.9385 (0.0093)
θ				0.8490 (0.0016)	0.8490 (0.0016)		0.8905 (0.0045)
ω				21.3230 (0.0535)	21.3230 (0.0535)		10.4424 (0.0360)
α	0.0596 (0.0032)	0.0540 (0.0052)	0.1405 (0.0037)	0.3400 (0.0041)	0.3421 (0.0053)	0.0254 (0.0070)	0.4150 (0.0046)
β	0.9322 (0.0033)	0.9294 (0.0041)	0.9877 (0.0001)	0.9628 (0.0004)	0.7732 (0.0051)	0.9549 (0.0082)	0.8905 (0.0045)
γ						1.2292×10^{-5} (1.3261×10^{-6})	0.0743 (0.0026)
ζ		0.0146 (0.0062)	-0.0150 (0.0027)				
ν_1				-0.0697 (0.0028)	-0.0936 (0.0040)		-0.0858 (0.0033)
ν_2				0.0642 (0.0019)	0.0736 (0.0022)		0.0315 (0.0033)
ξ				-0.6807 (0.0085)	-0.6724 (0.0058)		-0.2556 (0.0060)
σ_u^2				0.2212 (0.0041)	0.2134 (0.0041)		0.2548 (0.0063)
δ_1				-0.1137 (0.0038)	-0.1121 (0.0045)		-0.0128 (0.0020)
δ_2				0.0830 (0.0119)	0.0833 (0.0022)		-0.0858 (0.0033)
$\ell(r)$	11846.0918	11847.6597	11853.9636	11935.4139	11946.4255	11924.3818	12530.3081
$\ell(r, x)$				9029.6284	9118.9081		9320.1459

注: () 中是极大似然估计的渐近标准误差, $\ell(r)$ 是关于收益率的局部极大对数似然值, $\ell(r, x)$ 是全局极大对数似然值。

率具有正向影响, 即 RV 水平增加, 预期长期波动率会上升. 此外, ω 的估计值大于 1, 说明 RV 对长期波动率的影响随着滞后时间的增加呈现递减趋势. 图 1 和图 2 分别给出 SSEC 指数和 SZSEC 指数基于 REGARCH-MIDAS 模型和 RT-REGARCH-MIDAS 模型得到的条件方差及其长期成分的时间序列图. 从图 1 和图 2 可以看到, 两个模型的条件方差的长期成分都能够较好地捕捉条件方差的长期趋势.

在引入 MIDAS 结构的两个模型 (REGARCH-MIDAS 和 RT-REGARCH-MIDAS) 中, 条件方差被乘性分解为长期成分 (τ_t) 和短期成分 (g_t). 可以看到, 这些模型短期成分的持续性系数 β 的估计值相比未引入 MIDAS 结构的模型的条件方差的持续性系数的估计值明显要低, 表明 MIDAS 结构能够捕获波动率长记忆性. 在 REGARCH-MIDAS 模型和 RT-REGARCH-MIDAS 模型的长期成分中, θ 系数的估计值均显著为正, 表明 RV 对长期波动率具有正向影响, 即 RV 水平增加, 预期长期波动率会上升. 此外, ω 的估计值大于 1, 说明 RV 对长期波动率的影响随着滞后时间的增加呈现递减趋势. 图 1 和图 2 分别给出 SSEC 指数和 SZSEC 指数基于 REGARCH-MIDAS 模型和 RT-REGARCH-MIDAS 模型得到的条件方差及其长期成分的时间序列图. 从图 1 和图 2 可以看到, 两个模型的条件方差的长期成分都能够较好地捕捉条件方差的长期趋势.

REGARCH 模型及其拓展模型引入灵活的二次型杠杆函数能够较为充分地捕获中国股市的杠杆效应. 事实上, 在这些模型中, 杠杆系数 ν_1 、 ν_2 、 δ_1 和 δ_2 的估计值在统计上都是显著的, 且 ν_1 和 δ_1 估计值均小于零, 表明中国股市存在显著的杠杆效应. REGARCH 模型及其两个拓展模型的度量方程中 RV 偏差修正系数 ξ 的估计值显著小于零, 说明市场的非交易时间导致 RV 低估了真实的日度波动率.

最后注意到, 在引入了已实现测度的 REGARCH 模型、REGARCH-MIDAS 模型和 RT-REGARCH-MIDAS 模型中, α 的估计值显著为正, 这表明已实现测度包含关于未来波动率的重要信息; 在引入了当前收益率信息的 RT-GARCH 模型和 RT-REGARCH-MIDAS 模型中, γ 的估计值显著为正, 这表明当前收益率信息对波动率建模具有重要影响.

比较各模型关于收益率的局部对数似然 $\ell(r)$ 值可以看到, REGARCH 模型及其拓展模型相比 GARCH 模型、GJR-GARCH 模型和 EGARCH 模型都具有明显更高的 $\ell(r)$ 值, 表明引入高频数据信息能够显著改进模型对收益率数据的拟合效果; RT-GARCH 模型相比 GARCH 模型、GJR-GARCH 模型和 EGARCH 模型都具有明显更高的 $\ell(r)$ 值, 表明引入当前收益率信息能够显著改进模型对收益率数据的拟合效果. 总之, 同时引入日内高频数据信息和当前收益率信息的 RT-REGARCH-MIDAS 模型具有最优的数据拟合效果.

4.3 样本外预测表现

与样本内表现相比, 模型样本外预测表现对于市场参与者来说具有更为重要的意义. 本文采用滚动时间窗方法对波动率进行预测. 将数据样本总体分为两部分: “估计样本”和“预测样本”, 其中估计样本包含前 3500 个交易日的数据 (收益率和 RV), 预测样本包含最后 873 个交易日的数据 (即预测窗口为 873 天).

由于真实的波动率是不可观测的, 在评价和比较波动率预测模型时需要借助真实波动率的“代理值” (proxy). Andersen et al. (2001) 提出已实现波动率 (realized volatility, RV) 作

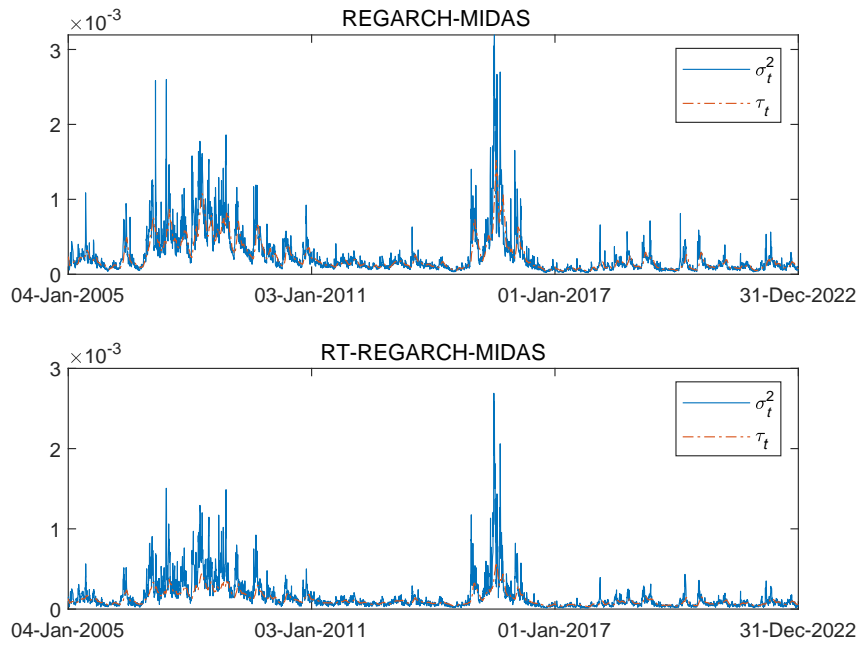


图 1 SSEC 指数条件方差及其长期成分的时间序列图

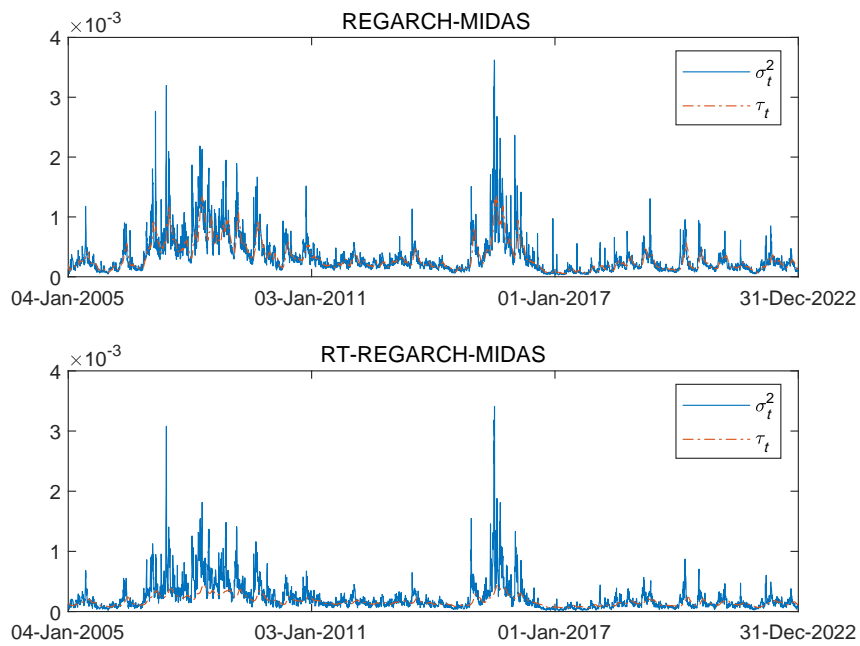


图 2 SZSEC 指数条件方差及其长期成分的时间序列图

为真实波动率的代理:

$$RV_t = \sum_{j=1}^M r_{t,j}^2, \quad (24)$$

其中 $r_{t,j}$ 是第 t 日的第 j 个日内收益率, M 是日内收益率数目, 取决于抽样频率. Andersen et al. (2001) 研究表明, 当不存在市场微观结构噪声且资产可连续交易时, 随着抽样频率的增加 ($J \rightarrow \infty$), RV 收敛于真实的积分波动率 (integrated volatility, IV), 也即 RV 是 IV 的一致估计量. 然而, 现实市场由于微观结构噪声的存在, 当抽样频率过高, 会导致 RV 较为严重的偏差, 而过低的抽样频率则会损失数据信息, 导致估计方差的增加. 权衡市场微观结构噪声 (偏差) 与信息损失 (方差), 本文与 Andersen et al. (2001) 中的做法一致, 采用 5 分钟高频数据来计算 RV .

采用 RV 作为真实波动率代理的另一个问题就是非交易时间. 如果仅仅采用开盘期间的日内收益率来计算 RV , 可能会低估真实的波动率. 为了克服这个问题, 根据 Hansen and Lunde (2005) 的研究, 本文采用如下调整的 RV 作为真实波动率的代理:

$$RV_t^{\text{scale}} = cRV_t, \quad c = \frac{\sum_{t=1}^T (r_t - \bar{r})^2}{\sum_{t=1}^T RV_t}, \quad (25)$$

其中 $\bar{r} = \sum_{t=1}^T r_t / T$ 是资产收益率的样本均值.

4.3.1 波动率预测结果

图 3 和图 4 分别给出了两个指数基于 REGARCH 模型、REGARCH-MIDAS 模型、RT-GARCH 模型和 RT-REGARCH-MIDAS 模型的波动率预测结果⁴. 从图 3 和图 4 可以看到, 这四个模型都能够较好地预测我国股市波动率, 表现出较好的预测精确性. 表 4 给出了基于两个损失函数对各模型对于 SSEC 指数和 SZSEC 指数波动率预测的评价结果. 从表 4 可以看到, 引入日内高频数据信息的 REGARCH 模型显著改进了 GARCH 模型、GJR-GARCH 模型和 EGARCH 模型的波动率预测能力, 说明日内高频数据信息对于波动率预测具有重要作用. RT-GARCH 模型相比于 GARCH 模型、GJR-GARCH 模型和 EGARCH 模型具有更为优越的波动率预测表现, 说明引入当前收益率信息能够改善波动率预测精确性. 引入 MIDAS 结构的 REGARCH-MIDAS 模型和 RT-REGARCH-MIDAS 模型波动率预测效果要优于相应的原始单成分 REGARCH 模型和 RT-GARCH 模型, 说明捕获波动率长记忆性对于波动率预测具有重要作用. 总体上, 充分捕获了日内高频信息、当前收益率信息和波动率长记忆性的 RT-REGARCH-MIDAS 模型获得最低的损失值, 具有最为优越的波动率预测能力.

进一步, 表 5 给出了 MCS 检验结果. 从表 5 中可以看到, RT-REGARCH-MIDAS 模型在所有情形下均被包含在 MCS 中. 特别值得指出的是, 该模型在所有情形下均获得了最高的 MCS 检验 p 值 ($p = 1$), 充分表明了该模型在中国股市波动率预测方面显著优于其它模型.

⁴其它模型的波动率预测结果如有需要可向作者索取.

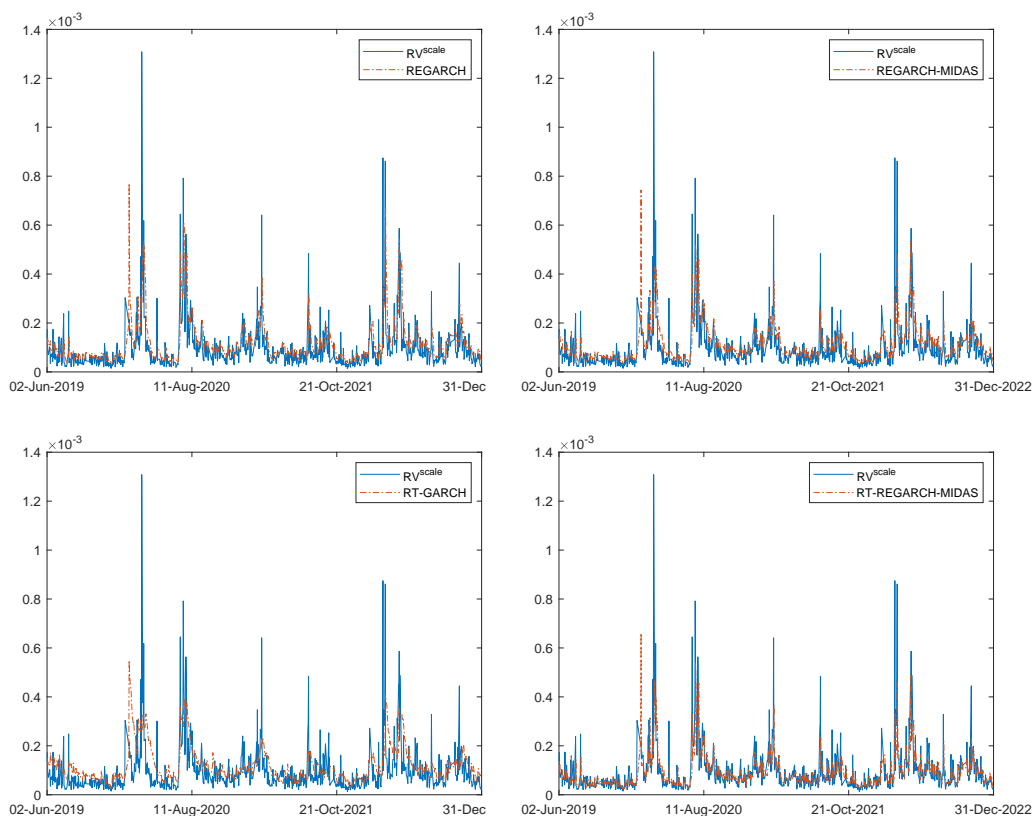


图 3 SSEC 指数波动率预测结果

表 4 波动率预测评价结果

	SSEC		SZSEC	
	MSE	QLIKE	MSE	QLIKE
GARCH	1.0852×10^{-8}	-8.2994	2.8992×10^{-8}	-7.7115
GJR-GARCH	1.0681×10^{-8}	-8.2995	2.9228×10^{-8}	-7.7118
EGARCH	9.9715×10^{-9}	-8.2992	2.8881×10^{-8}	-7.7063
REGARCH	7.8829×10^{-9}	-8.3777	2.1432×10^{-8}	-7.7852
REGARCH-MIDAS	6.8990×10^{-9}	-8.3840	1.9718×10^{-8}	-7.7886
RT-GARCH	8.7940×10^{-9}	-8.3316	2.5157×10^{-8}	-7.7544
RT-REGARCH-MIDAS	6.5041×10^{-9}	-8.3893	1.9315×10^{-8}	-7.7889

注: 表中加粗的数字表示最低损失值. MSE 是均方误差, QLIKE 是拟似然误差.

4.3.2 稳健性检验

为了说明本文提出的 RT-REGARCH-MIDAS 模型波动率预测能力的稳健性, 本文考虑不同的已实现测度、不同的预测窗口、不同的 MIDAS 滞后阶数和不同的预测期下模型的预测表现. 首先, 本文在引入了已实现测度的模型 (REGARCH 模型、REGARCH-MIDAS

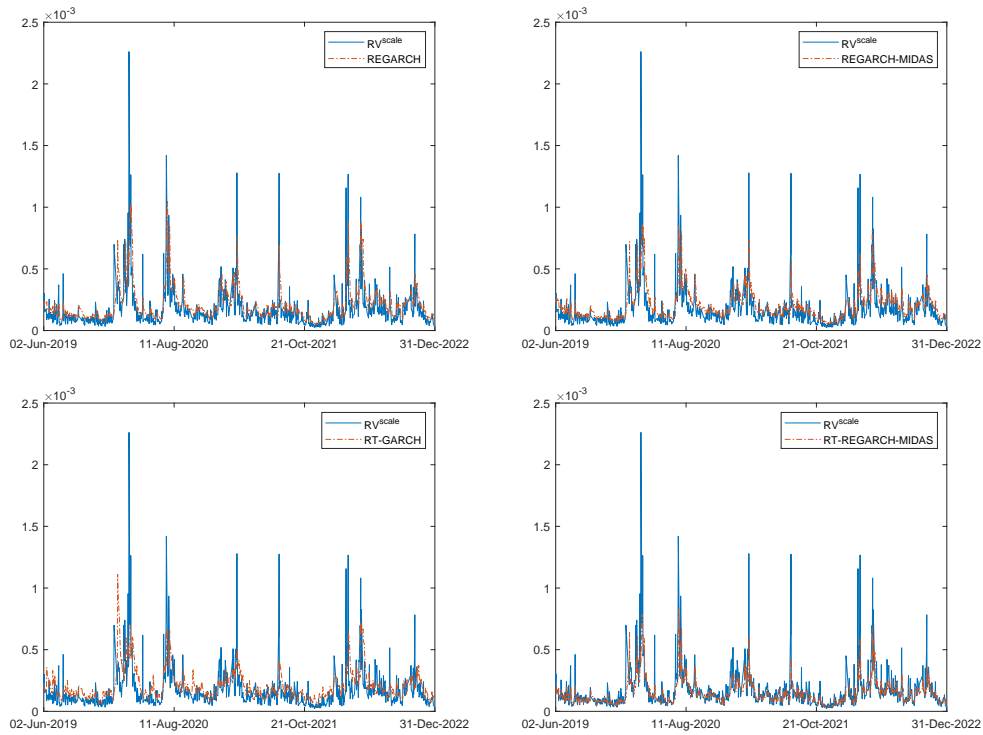


图 4 SZSEC 指数波动率预测结果

表 5 MCS 检验结果

	SSEC		SZSEC	
	MSE	QLIKE	MSE	QLIKE
GARCH	0.0000	0.0000	0.0000	0.0000
GJR-GARCH	0.0000	0.0000	0.0000	0.0000
EGARCH	0.0000	0.0000	0.0000	0.0000
REGARCH	0.0207	0.0167	0.1303	0.2529
REGARCH-MIDAS	0.0479	0.2410	0.6024	0.9588
RT-GARCH	0.0000	0.0000	0.0025	0.0000
RT-REGARCH-MIDAS	1.0000	1.0000	1.0000	1.0000

注: 表中数字是 MCS 检验 p 值. p 值大于 0.1 (加粗的数字) 表示模型包含在 MCS 中, 即预测能力较好的模型. MSE 是均方误差, QLIKE 是拟似然误差.

模型和 RT-REGARCH-MIDAS 模型) 中, 选取另外一种已实现测度, 即已实现极差波动率

(realized range volatility, RRV)⁵, 对模型预测能力进行检验, 波动率预测评价结果如表 6 所示. 另外, 本文考虑预测窗口为 120 天、240 天和 480 天 (即半年、一年和两年) 下模型的预测表现, 波动率预测评价结果如表 7 所示. 进一步, 本文考虑 MIDAS 滞后阶数为 $K = 60$ 和 120 (即季度 RV 和半年度 RV) 下模型的预测表现, 波动率预测评价结果如表 8 所示. 最后, 本文考虑预测期为 5 天和 10 天 (即一周和两周) 模型的预测表现, 波动率预测评价结果如表 9 所示.

从表 6~9 中可以看到, 与前面表 4 和表 5 中得到的结果一致, RT-REGARCH-MIDAS 模型获得了最低的损失函数值, 并且在所有情形下均被包含在 MCS 中, 说明该模型在中国股市波动率预测上显著优于其它模型.

本文还进一步采用样本外 R^2 检验来评估模型的波动率预测能力. Campbell and Thompson (2008) 提出了样本外预测 R^2 检验, 定义为:

$$R^2 = 1 - \frac{\sum_{t=1}^T (\sigma_t^2 - \hat{\sigma}_t^2(m))^2}{\sum_{t=1}^T (\sigma_t^2 - \hat{\sigma}_t^2(bm))^2}, \quad (26)$$

其中 σ_t^2 是真实的波动率, $\hat{\sigma}_t^2(m)$ 是模型预测的波动率, $\hat{\sigma}_t^2(bm)$ 是基准模型预测的波动率. 本文选取标准的 GARCH 模型作为基准模型. 当 $R^2 > 0$ 时, 说明预测模型的预测能力优于基准模型. 本文选取 Clark and West (2007) 提出的调整后的 MSPE 统计量来检验 R^2 在统计上的显著性. 样本外 R^2 检验结果如表 10 所示. 从表 10 可以看到, RT-REGARCH-MIDAS 模型的 R^2 均为正且在统计上显著, 而且高于其他竞争模型的 R^2 , 说明该模型拥有最佳的样本外波动率预测表现.

表 6 波动率预测评价结果: 不同已实现测度 (RRV)

	SSEC		SZSEC	
	MSE	QLIKE	MSE	QLIKE
GARCH	1.0852×10^{-8}	-8.2994	2.8992×10^{-8}	-7.7115
GJR-GARCH	1.0681×10^{-8}	-8.2995	2.9228×10^{-8}	-7.7118
EGARCH	9.9715×10^{-9}	-8.2992	2.8881×10^{-8}	-7.7063
REGARCH	9.1853×10^{-9}	-8.3369	2.5259×10^{-8}	-7.7517
REGARCH-MIDAS	7.9333×10^{-9}	-8.3431	2.2914×10^{-8}	-7.7530
RT-GARCH	8.7940×10^{-9}	-8.3316	2.5157×10^{-8}	-7.7544
RT-REGARCH-MIDAS	7.0140×10^{-9}	-8.3703	2.1275×10^{-8}	-7.7725

注: MSE 是均方误差, QLIKE 是拟似然误差. MCS 检验 p 值大于 0.1 (加粗的数字) 表示模型包含在 MCS 中, 即预测能力较好的模型.

⁵RRV 定义为:

$$RRV_t = \frac{1}{4 \log(2)} \sum_{j=1}^M (\log(H_{t,j}) - \log(L_{t,j}))^2,$$

其中 $H_{t,j}$ 和 $L_{t,j}$ 分别是第 t 日的第 j 个区间上的最高价和最低价. Christensen and Podolskij (2007) 研究表明, 理论上 RRV 的精度是 RV 的 5 倍. 本文采用 5 分钟高频数据来计算 RRV.

表 7 波动率预测评价结果: 不同预测窗口

	SSEC		SZSEC	
	MSE	QLIKE	MSE	QLIKE
预测窗口: 120 天				
GARCH	4.2739×10^{-9}	-8.3390	1.0924×10^{-8}	-7.7865
GJR-GARCH	4.2473×10^{-9}	-8.3401	1.0668×10^{-8}	-7.7902
EGARCH	4.3321×10^{-9}	-8.3412	1.1356×10^{-8}	-7.7802
REGARCH	3.6882×10^{-9}	-8.3871	9.6648×10^{-9}	-7.8294
REGARCH-MIDAS	3.2960×10^{-9}	-8.3967	8.6530×10^{-9}	-7.8356
RT-GARCH	4.1521×10^{-9}	-8.3481	9.7807×10^{-9}	-7.8107
RT-REGARCH-MIDAS	3.0339×10^{-9}	-8.4142	7.0412×10^{-9}	-7.8598
预测窗口: 240 天				
GARCH	1.2373×10^{-8}	-8.1306	2.6001×10^{-8}	-7.5939
GJR-GARCH	1.2358×10^{-8}	-8.1306	2.5949×10^{-8}	-7.5952
EGARCH	1.1713×10^{-8}	-8.1352	2.5513×10^{-8}	-7.5905
REGARCH	9.1924×10^{-9}	-8.1951	2.0720×10^{-8}	-7.6403
REGARCH-MIDAS	7.6334×10^{-9}	-8.2066	1.8021×10^{-8}	-7.6479
RT-GARCH	9.8656×10^{-9}	-8.1567	2.3415×10^{-8}	-7.6142
RT-REGARCH-MIDAS	7.4869×10^{-9}	-8.2073	1.6808×10^{-8}	-7.6606
预测窗口: 480 天				
GARCH	8.3299×10^{-9}	-8.2844	2.1939×10^{-8}	-7.7185
GJR-GARCH	8.3128×10^{-9}	-8.2845	2.2005×10^{-8}	-7.7185
EGARCH	8.3048×10^{-9}	-8.2774	2.2585×10^{-8}	-7.7120
REGARCH	6.7252×10^{-9}	-8.3285	1.8248×10^{-8}	-7.7687
REGARCH-MIDAS	5.6935×10^{-9}	-8.3363	1.5793×10^{-8}	-7.7762
RT-GARCH	4.1521×10^{-9}	-8.3481	1.9775×10^{-8}	-7.7449
RT-REGARCH-MIDAS	5.5053×10^{-9}	-8.3381	1.5423×10^{-8}	-7.7861

注: MSE 是均方误差, QLIKE 是拟似然误差. MCS 检验 p 值大于 0.1 (加粗的数字) 表示模型包含在 MCS 中, 即预测能力较好的模型.

综上, 本文提出的 RT-REGARCH-MIDAS 模型优越的波动率预测能力具有关于不同的已实现测度、不同的预测窗口、不同的 MIDAS 滞后阶数、不同的预测期和样本外 R^2 检验的稳健性.

4.3.3 波动择时的经济价值

为了评价模型样本外预测能力的提升是否能带来经济价值, 本文基于各模型的波动率预测值构建最优资产组合, 检验其波动择时策略. 具体地, 本文基于 Fleming et al. (2001, 2003) 提出的方法, 通过采用均值-方差策略来考察各个模型波动率预测结果的经济价值.

考虑具有均值-方差偏好的风险厌恶型投资者, 他们考虑两种金融资产的投资, 一种是无风险资产, 另一种是风险资产 (股票). 投资者使用条件均值-方差分析来对未来投资做出决

表 8 波动率预测评价结果: 不同 MIDAS 滞后阶数

	SSEC		SZSEC	
	MSE	QLIKE	MSE	QLIKE
MIDAS 滞后阶数: $K = 60$				
GARCH	1.0852×10^{-8}	-8.2994	2.8992×10^{-8}	-7.7115
GJR-GARCH	1.0681×10^{-8}	-8.2995	2.9228×10^{-8}	-7.7118
EGARCH	9.9715×10^{-9}	-8.2992	2.8881×10^{-8}	-7.7063
REGARCH	7.8829×10^{-9}	-8.3777	2.1432×10^{-8}	-7.7852
REGARCH-MIDAS	6.8983×10^{-9}	-8.3840	1.9718×10^{-8}	-7.7886
RT-GARCH	8.7940×10^{-9}	-8.3316	2.5157×10^{-8}	-7.7544
RT-REGARCH-MIDAS	6.6634×10^{-9}	-8.3851	1.9195×10^{-8}	-7.8001
MIDAS 滞后阶数: $K = 120$				
GARCH	1.0852×10^{-8}	-8.2994	2.8992×10^{-8}	-7.7115
GJR-GARCH	1.0681×10^{-8}	-8.2995	2.9228×10^{-8}	-7.7118
EGARCH	9.9715×10^{-9}	-8.2992	2.8881×10^{-8}	-7.7063
REGARCH	7.8829×10^{-9}	-8.3777	2.1432×10^{-8}	-7.7852
REGARCH-MIDAS	6.9109×10^{-9}	-8.3843	1.9718×10^{-8}	-7.7886
RT-GARCH	8.7940×10^{-9}	-8.3316	2.5157×10^{-8}	-7.7544
RT-REGARCH-MIDAS	6.7212×10^{-9}	-8.3851	1.9363×10^{-8}	-7.7890

注: MSE 是均方误差, QLIKE 是拟似然误差. MCS 检验 p 值大于 0.1 (加粗的数字) 表示模型包含在 MCS 中, 即预测能力较好的模型.

策, 投资组合的效用为:

$$U(r_{t+1}^p) = E_t[r_{t+1}^p] - \frac{A}{2} \text{Var}_t[r_{t+1}^p], \quad (27)$$

其中 $E_t[r_{t+1}^p]$ 和 $\text{Var}_t[r_{t+1}^p]$ 分别是投资组合收益率的条件均值和条件方差, A 是投资者的风险厌恶系数. 最大化投资组合的效用 $U(r_{t+1}^p)$ 得到风险资产的最优权重为:

$$w_t^* = \frac{E_t[r_{t+1}] - r_{t+1}^f}{A \text{Var}_t[r_{t+1}]}, \quad (28)$$

其中 $E_t[r_{t+1}]$ 和 $\text{Var}_t[r_{t+1}]$ 分别是风险资产收益率和方差的预测值, r_{t+1}^f 是无风险利率. 根据 Fleming et al. (2001, 2003), $E_t[r_{t+1}]$ 基于滚动样本内均值估计得到. 显然, 在给定的期望收益率水平下, 当预期波动率水平较高, 投资者将分配更多的资金到无风险资产, 反之亦然. 考虑卖空约束以及不超过 50% 的杠杆率, 本文设定 $0 \leq w_t^* \leq 1.5$. 投资组合的实际收益率为:

$$r_{t+1}^p = r_{t+1}^f + w_t^*(r_{t+1} - r_{t+1}^f), \quad (29)$$

其中无风险利率 r_{t+1}^f 采用 3 个月的上海银行间同业拆借利率 (SHIBOR) 作为代理.

为了衡量投资组合的表现, 根据 Fleming et al. (2003) 的研究, 计算如下平均实际效用:

$$\bar{U}(r_{t+1}^p) = \frac{1}{I} \sum_{t=1}^I \left(r_{t+1}^p - \frac{A}{2} \left(r_{t+1}^p - \frac{1}{I} \sum_{t=1}^I r_{t+1}^p \right)^2 \right). \quad (30)$$

表 9 波动率预测评价结果: 不同预测期

	SSEC		SZSEC	
	MSE	QLIKE	MSE	QLIKE
预测期: 5 天				
GARCH	1.3302×10^{-8}	-8.2463	3.4610×10^{-8}	-7.6695
GJR-GARCH	1.3071×10^{-8}	-8.2468	3.5069×10^{-8}	-7.6689
EGARCH	1.2508×10^{-8}	-8.2449	3.5500×10^{-8}	-7.6603
REGARCH	1.2282×10^{-8}	-8.2742	3.3653×10^{-8}	-7.6972
REGARCH-MIDAS	1.0991×10^{-8}	-8.2690	3.3406×10^{-8}	-7.6802
RT-GARCH	1.1129×10^{-8}	-8.2624	3.1903×10^{-8}	-7.6859
RT-REGARCH-MIDAS	1.0047×10^{-8}	-8.2859	2.8652×10^{-8}	-7.7081
预测期: 10 天				
GARCH	1.4954×10^{-8}	-8.2119	3.7923×10^{-8}	-7.6396
GJR-GARCH	1.4745×10^{-8}	-8.2124	3.8348×10^{-8}	-7.6389
EGARCH	1.4687×10^{-8}	-8.2038	4.0105×10^{-8}	-7.6257
REGARCH	1.4515×10^{-8}	-8.2251	3.7771×10^{-8}	-7.6542
REGARCH-MIDAS	1.3140×10^{-8}	-8.2109	3.9908×10^{-8}	-7.6280
RT-GARCH	1.2434×10^{-8}	-8.2218	3.3635×10^{-8}	-7.6511
RT-REGARCH-MIDAS	1.1855×10^{-8}	-8.2289	3.3554×10^{-8}	-7.6587

注: MSE 是均方误差, QLIKE 是拟似然误差. MCS 检验 p 值大于 0.1 (加粗的数字) 表示模型包含在 MCS 中, 即预测能力较好的模型.

表 10 波动率预测评价结果: 样本外 R^2 检验

	SSEC			SZSEC		
	R^2	MSPE-adjusted	p 值	R^2	MSPE-adjusted	p 值
GJR-GARCH	0.0157	2.0196	0.0217	-0.0081	-1.3214	0.9068
EGARCH	0.0811	1.6205	0.0526	0.0038	1.0345	0.1504
REGARCH	0.2736	5.5924	0.0000	0.2607	6.7811	0.0000
REGARCH-MIDAS	0.3642	4.7743	0.0000	0.3199	7.9552	0.0000
RT-GARCH	0.1896	4.9895	0.0000	0.1323	5.0029	0.0000
RT-REGARCH-MIDAS	0.4006	4.7548	0.0000	0.3338	5.8865	0.0000

注: 表中 R^2 是样本外 R^2 , MSPE-adjusted 是调整后的 MSPE, $R^2 > 0$ 说明对应模型的样本外预测能力优于基准模型.

为了量化相对于基准模型的经济效益, 计算如下投资组合的业绩表现费 (performance fee):

$$\Delta = \bar{U}(r_{t+1}^p) - \bar{U}(r_{t+1}^{p,bm}), \quad (31)$$

其中 $\bar{U}(r_{t+1}^{p,bm})$ 是基准模型的平均实际效用. 业绩表现费 Δ 测度投资者从基准策略切换到竞争策略所愿意支付的费用 (年化基点), Δ 越大, 表明模型的经济价值越高.

同时, 本文采用确定性等价收益 (certainty equivalent return, CER) 来衡量投资组合的

表现. CER 可以表示为:

$$\text{CER} = E[r_{t+1}^p] - \frac{A}{2} \text{Var}[r_{t+1}^p]. \quad (32)$$

CER 越大, 说明模型的经济价值越高.

本文选取标准的 GARCH 模型作为基准模型, 并选取投资者的风险厌恶系数为 $A = 3$ 、6 和 9, 得到不同模型下投资组合的业绩表现 (年化基点费用) 如表 11 所示. 从表 11 可以看到, 在所有模型中, RT-REGARCH-MIDAS 模型的业绩表现费最高, 达到 74~596 个基点, REGARCH-MIDAS 模型次之, 达到 43~343 个基点. 值得指出的是, 虽然 GJR-GARCH 和 RT-GARCH 能够改进 GARCH 模型的波动率预测能力, 但并不能带来经济价值的提升 (业绩表现费在某些情形下为负). 此外, 从表中还可以看到, RT-REGARCH-MIDAS 模型也获得了最高的 CER (年化百分比值), 说明该模型的经济价值最高, REGARCH-MIDAS 模型次之. 综上, 引入日内高频信息和当前收益率信息的 RT-REGARCH-MIDAS 模型相比其它所有模型获得了显著更高的投资组合经济价值的提升, 且这种提升对于不同的投资者风险厌恶水平 ($A = 3$ 、6 和 9) 具有稳健性.

表 11 投资组合的业绩表现

	A = 3		A = 6		A = 9	
	Δ	CER	Δ	CER	Δ	CER
SSEC						
GARCH	-	0.1504	-	0.1165	-	0.1044
GJR-GARCH	1.6125	0.1509	1.0384	0.1168	0.6922	0.1046
EGARCH	21.3046	0.1575	8.0734	0.1192	5.3821	0.1062
REGARCH	21.6988	0.1576	34.4875	0.1280	22.9901	0.1120
REGARCH-MIDAS	91.4051	0.1808	65.5094	0.1383	43.6716	0.1189
RT-GARCH	-82.9584	0.1227	10.7807	0.1201	8.1875	0.1071
RT-REGARCH-MIDAS	188.0060	0.2130	111.7531	0.1537	74.5012	0.1292
SZSEC						
GARCH	-	0.3544	-	0.2173	-	0.1716
GJR-GARCH	-0.0061	0.3544	-0.0028	0.2173	-0.0016	0.1716
EGARCH	52.1986	0.3718	26.0999	0.2260	17.4003	0.1774
REGARCH	160.0133	0.4078	80.0091	0.2440	53.3410	0.1894
REGARCH-MIDAS	343.6988	0.4690	171.8509	0.2746	114.5682	0.2098
RT-GARCH	52.3828	0.3719	28.1653	0.2267	18.7788	0.1778
RT-REGARCH-MIDAS	596.5541	0.5532	272.2472	0.3080	181.5024	0.2321

注: 表中加粗且带下划线的数字表示最大数值, 加粗但不带下划线的数字表示次大数值. A 是风险厌恶系数, Δ 是业绩表现费 (年化基点费用), CER 是确定性等价收益 (年化百分比值).

5 结论与展望

本文综合考虑日内高频数据信息、当前收益率信息和波动率长记忆性, 构建了 RT-REGARCH-MIDAS 模型对中国股市波动率进行建模和预测. 采用 SSEC 指数和 SZSEC 指

数 5 分钟高频数据进行实证研究, 结果表明: 中国股市波动率存在高持续性以及显著的杠杆效应; 本文提出的模型相比其它众多竞争模型 (包括 GARCH 模型、GJR-GARCH 模型、EGARCH 模型、REGARCH 模型、REGARCH-MIDAS 模型和 RT-GARCH 模型) 具有更好的收益率数据拟合效果, 能够更好地描述股市波动性. 运用稳健的损失函数以及 MCS 检验, 实证检验了本文模型与其它竞争模型对中国股市波动率的样本外预测能力, 结果表明: 在波动率模型中引入高频数据信息、当前收益率信息和波动率长记忆性对于波动率预测具有重要作用; 本文构建的模型相比其它竞争模型具有更为优越的波动率预测能力, 而且这种优越的波动率预测能力具有关于不同的已实现测度、不同的预测窗口、不同的 MIDAS 滞后阶数、不同的预测期以及样本外 R^2 检验的稳健性. 进一步, 通过考察基于不同波动率模型的波动择时策略, 验证了本文提出的模型优越的波动率预测能力能够带来更高的投资组合经济价值.

本文的研究工作为中国股市波动率建模与预测提供了有益的方法借鉴, 无论对于学术研究人员还是实际从业人员都具有重要的意义. 值得指出的是, 本文研究工作还可以进一步的拓展, 例如可以考虑在现有模型基础上引入宏观经济变量, 考察宏观信息对波动率的影响. 此外, 考虑资产收益率分布的偏斜、尖峰厚尾及其时变性 (时变高阶矩) 特征, 对本文模型进行扩展并将其应用于市场风险度量也值得进一步深入研究. 当然, 将本文模型应用于期权定价中, 也是未来一个重要的研究方向.

参 考 文 献

- 柏建成, 黄云飞, 高增安, 何田, (2022). 经济政策不确定性与数字货币市场波动影响研究——基于比特币市场的实证分析 [J]. 运筹与管理, 31(5): 183-189.
- Bai J C, Huang Y F, Gao Z A, He T, (2022). Research on the Impact of Economic Policy Uncertainty and Cryptocurrency Market Volatility — Empirical Analysis Based on GARCH-MIDAS Model[J]. Operations Research and Management Science, 31(5): 183-189.
- 白娟娟, 师荣蓉, (2021). 基于广义已实现测度的中国股市波动预测与 VaR 度量 [J]. 系统科学与数学, 41(3): 653-666.
- Bai J J, Shi R R, (2021). The Volatility Forecasting and VaR Measurement of Chinese Stock Market Based on Generalized Realized Measures[J]. Journal of Systems Science and Mathematical Sciences, 41(3): 653-666.
- 蔡光辉, 徐君, 应雪海, (2021). 基于混频抽样和杠杆效应的高频波动率模型预测 [J]. 系统科学与数学, 41(7): 1985-2005.
- Cai G H, Xu J, Ying X H, (2021). High-Frequency Volatility Prediction Based on Mixed Frequency Sampling and Leverage Effect[J]. Journal of Systems Science and Mathematical Sciences, 41(7): 1985-2005.
- 高雷阜, 李伟梅, (2022). 基于 Expectile 和 Realized GARCH 模型的波动率预测 [J]. 运筹与管理, 31(2): 99-103.
- Gao L F, Li W M, (2022). Volatility Prediction Based on Expectile and Realized GARCH model[J]. Operations Research and Management Science, 31(2): 99-103.
- 龚旭, 曹杰, 文凤华, 杨晓光, (2020). 基于杠杆效应和结构突变的 HAR 族模型及其对股市波动率的预测研究 [J]. 系统工程理论与实践, 40(5): 1113-1133.
- Gong X, Cao J, Wen F H, Yang X G, (2020). The HAR-type Models with Leverage and Structural

- Breaks and Their Applications to the Volatility Forecasting of Stock Market[J]. *Systems Engineering — Theory & Practice*, 40(5): 1113–1133.
- 郭宝才, 项琳, (2022). 基于跳跃、好坏波动率的混频已实现 EGARCH 模型的波动率预测与风险度量 [J]. *商业经济与管理*, 5: 79–97.
- Guo B C, Xiang L, (2022). Realized EGARCH-MIDAS Model Based on Jump and Good-bad Volatility with Its Volatility and VaR Forecasting[J]. *Journal of Business Economics*, 5: 79–97.
- 黄友珀, 唐振鹏, 周熙雯, (2015). 基于偏 t 分布 realized GARCH 模型的尾部风险估计 [J]. *系统工程理论与实践*, 35(9): 2200–2208.
- Huang Y P, Tang Z P, Zhou X W, (2015). Estimation of Tail Risk Based on Realized GARCH Model with Skew- t Distribution[J]. *Systems Engineering — Theory & Practice*, 35(9): 2200–2208.
- 蒋伟, 顾研, (2019). 基于广义已实现测度的 Realized GARCH 模型改进及应用 [J]. *数量经济技术经济研究*, 36(7): 156–173.
- Jiang W, Gu Y, (2019). Forecasting VaR Realized GARCH Incorporating Generalized Realized Measures[J]. *Journal of Quantitative & Technological Economics*, 36(7): 156–173.
- 雷立坤, 余江, 魏宇, 赖晓东, (2018). 经济政策不确定性与我国股市波动率预测研究 [J]. *管理科学学报*, 21(6): 88–98.
- Lei L K, Yu J, Wei Y, Lai X D, (2018). Forecasting Volatility of Chinese Stock Market with Economic Policy Uncertainty[J]. *Journal of Management Sciences in China*, 21(6): 88–98.
- 李木易, 方颖, (2020). 动态混合 HGARCH 模型的估计和预测 [J]. *管理科学学报*, 23(5): 1–12.
- Li M Y, Fang Y, (2020). Estimation and Forecasting of Mixture HGARCH Model[J]. *Journal of Management Sciences in China*, 23(5): 1–12.
- 梁超, 魏宇, 马锋, 李薇, (2022). 投资者关注对中国黄金价格波动率的影响研究 [J]. *系统工程理论与实践*, 42(2): 320–332.
- Liang C, Wei Y, Ma F, Li W, (2022). A Study on the Impact of Investor Attention on Chinese Gold Volatility[J]. *Systems Engineering — Theory & Practice*, 42(2): 320–332.
- 刘小军, 汪寿阳, 谢海滨, (2022). 已实现波动率预测: 非对称二次滑动平均模型 [J]. *计量经济学报*, 2(4): 930–945.
- Liu X J, Wang S Y, Xie H B, (2022). Realized Volatility Forecasting: An Asymmetric Double Moving Average Model[J]. *China Journal of Econometrics*, 2(4): 930–945.
- 鲁万波, 亢晶浩, (2022). GAS-SKST-F 模型及其在高频多元波动率预测中的应用 [J]. *中国管理科学*, 30(1): 77–87.
- Lu W B, Kang J H, (2022). GAS-SKST-F Model and Its Application in High Frequency Multivariate Volatility Forecast[J]. *Chinese Journal of Management Science*, 30(1): 77–87.
- 王天一, 黄卓, (2015). Realized GAS-GARCH 及其在 VaR 预测中的应用 [J]. *管理科学学报*, 18(5): 79–86.
- Wang T Y, Huang Z, (2015). Realized GAS-GARCH Model and Its Application in Value-at-Risk Forecast[J]. *Journal of Management Sciences in China*, 18(5): 79–86.
- 王天一, 刘浩, 黄卓, (2018). 基于混频数据抽样的已实现波动率长记忆模型 [J]. *系统工程学报*, 33(6): 812–822.
- Wang T Y, Liu H, Huang Z, (2018). Model for the Long Memory of Realized Volatility Based on Mixed Data Sampling[J]. *Journal of Systems Engineering*, 33(6): 812–822.
- 吴鑫彦, 侯信盟, (2020). 基于双因子已实现 GARCH 模型的波动率预测研究 [J]. *运筹与管理*, 29(12): 207–214.
- Wu X Y, Hou X M, (2020). Volatility Forecasting Based on Two-factor Realized GARCH Model[J].

- Operations Research and Management Science, 29(12): 207–214.
- 夏婷, 闻岳春, (2018). 经济不确定性是股市波动的因子吗? ——基于 GARCH-MIDAS 模型的分析 [J]. 中国管理科学, 26(12): 1–11.
- Xia T, Wen Y C, (2018). Does Economic Uncertainty Matter for Stock Market Volatility? An Analysis Based on GARCH-MIDAS[J]. Chinese Journal of Management Science, 26(12): 1–11.
- 邢艳春, 廖晗, (2023). 经济政策不确定性对中国股市波动率的影响 [J]. 统计与信息论坛, 38(1): 71–80.
- Xing Y C, Liao H, (2023). The Impact of Economic Policy Uncertainty on China's Stock Market Volatility[J]. Journal of Statistics and Information, 38(1): 71–80.
- 于孝建, 王秀花, (2018). 基于混频已实现 GARCH 模型的波动预测与 VaR 度量 [J]. 统计研究, 35(1): 104–116.
- Yu X J, Wang X H, (2018). Mix Frequency Realized GARCH Models: The Forecasting of Volatility and Measure of VaR[J]. Statistical Research, 35(1): 104–116.
- 苑慧玲, 徐路, 周勇, (2020). 带有市场交易信息和随机微观噪声下的杠杆效应研究 [J]. 中国管理科学, 28(9): 12–22.
- Yuan H L, Xu L, Zhou Y, (2020). Leverage Effect Combining Trading Information with Stochastic Microstructure Noise[J]. Chinese Journal of Management Science, 28(9): 12–22.
- 张一锋, 雷立坤, 魏宇, (2020). 羊群效应的新测度指数及其对我国股市波动的预测作用研究 [J]. 系统工程理论与实践, 40(11): 2810–2824.
- Zhang Y F, Lei L K, Wei Y, (2020). A New Herd Index and Volatility Forecasting of China's Stock Market[J]. Systems Engineering — Theory & Practice, 40(11): 2810–2824.
- 郑挺国, 尚玉皇, (2014). 基于宏观基本面的股市波动度量与预测 [J]. 世界经济, 37(12): 118–139.
- Zheng T G, Shang Y H, (2014). Measuring and Forecasting Stock Market Volatility with Macro Fundamentals[J]. The Journal of World Economy, 37(12): 118–139.
- 周辰月, 崔文昊, (2023). 金融资产波动率估计的最优内生抽样方案的设计与应用 [J]. 计量经济学报, 3(1): 238–258.
- Zhou C Y, Cui W H, (2023). The Optimal Endogenous Sampling Scheme for Financial Asset Volatility Estimation[J]. China Journal of Econometrics, 3(1): 238–258.
- Andersen T G, Bollerslev T, Diebold F X, Labys P, (2001). The Distribution of Realized Exchange Rate Volatility[J]. Journal of the American Statistical Association, 96(453): 42–55.
- Asgharian H, Christiansen C, Hou A J, (2016). Macro-finance Determinants of the Long-run Stock-bond Correlation: The DCC-MIDAS Specification[J]. Journal of Financial Econometrics, 14: 617–642.
- Asgharian H, Hou A J, Javed F, (2013). The Importance of the Macroeconomic Variables in Forecasting Stock Return Variance: A GARCH-MIDAS Approach[J]. Journal of Forecasting, 32: 600–612.
- Banulescu-Radu D, Hansen P R, Huang Z, Matei M, (2018). Volatility During the Financial Crisis Through the lens of High Frequency Data: A Realized GARCH Approach[R]. Working Paper.
- Bollerslev T, (1986). Generalized Autoregressive Conditional Heteroskedasticity[J]. Journal of Econometrics, 31: 307–327.
- Borup D, Jakobsen J S, (2019). Capturing Volatility Persistence: A Dynamically Complete Realized EGARCH-MIDAS Model[J]. Quantitative Finance, 19(11): 1839–1855.
- Breitung J, Hafner C M, (2016). A Simple Model for Now-casting Volatility Series[J]. International Journal of Forecasting, 32(4): 1247–1255.
- Campbell J Y, Thompson S B, (2008). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?[J]. Review of Financial Studies, 21(4): 1509–1531.

- Christensen K, Podolskij M, (2007). Realized Range-based Estimation of Integrated Variance[J]. *Journal of Econometrics*, 141(2): 323–349.
- Christoffersen P, Feunou B, Jacobs K, Meddahi N, (2014). The Economic Value of Realized Volatility: Using high-frequency Returns for Option Valuation[J]. *Journal of Financial and Quantitative Analysis*, 49(3): 663–697.
- Clark T E, West K D, (2007). Approximately Normal Tests for Equal Predictive Accuracy in Nested Models[J]. *Journal of Econometrics*, 138(1): 291–311.
- Conrad C, Kleen O, (2020). Two are Better Than One: Volatility Forecasting Using Multiplicative Component GARCH-MIDAS Models[J]. *Journal of Applied Econometrics*, 35(1): 19–45.
- Conrad C, Loch K, (2015). Anticipating Long-term Stock Market Volatility[J]. *Journal of Applied Econometrics*, 30: 1090–1114.
- Diebold F X, (1986). Modelling the Persistence of Conditional Variance: A Comment[J]. *Econometric Reviews*, 5: 51–56.
- Ding Y D, (2023). A Simple Joint Model for Returns, Volatility and Volatility of Volatility[J]. *Journal of Econometrics*, 232(2): 521–543.
- Ding Z, (2016). Volatility Modeling Using GARCH: Theory and Practice[R]. Working Paper.
- Ding Z, Granger C W J, (1996). Modeling Volatility Persistence of Speculative Returns: A New Approach[J]. *Journal of Econometrics*, 73(1): 185–215.
- Dobrev D, Szerszen P, (2010). The Information Content of High-frequency Data for Estimating Equity Return Models and Forecasting Risk[R]. Working Paper.
- Dorion C, (2016). Option Valuation with Macro-finance Variables[J]. *Journal of Financial and Quantitative Analysis*, 51: 1359–1389.
- Engle R F, Ghysels E, Sohn B, (2013). Stock Market Volatility and Macroeconomic Fundamentals[J]. *Review of Economics and Statistics*, 95(3): 776–797.
- Engle R F, Lee G G L, (1999). A Long-run and Short-run Component Model of Stock Return Volatility[M]. *Cointegration, Causality, and Forecasting*. Oxford: Oxford University Press.
- Fleming J, Kirby C, Ostdiek B, (2001). The Economic Value of Volatility Timing[J]. *Journal of Finance*, 56(1): 329–352.
- Fleming J, Kirby C, Ostdiek B, (2003). The economic Value of Volatility Timing Using “Realized” Volatility[J]. *Journal of Financial Economics*, 67(3): 473–509.
- Ghysels E, Santa-Clara P, Valkanov R, (2004). The MIDAS Touch: Mixed Data Sampling Regression Models[R]. Working Paper.
- Ghysels E, Sinko A, Valkanov R, (2007). MIDAS Regressions: Further Results and New Directions[J]. *Econometric Reviews*, 26: 53–90.
- Hansen P R, Huang Z, (2016). Exponential GARCH Modeling with Realized Measures of Volatility[J]. *Journal of Business & Economic Statistics*, 34(2): 269–287.
- Hansen P R, Huang Z, Shek H H, (2012). Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility[J]. *Journal of Applied Econometrics*, 27: 877–906.
- Hansen P R, Huang Z, Tong C, Wang T Y, (2021). Realized GARCH, CBOE VIX, and the Volatility Risk Premium[R]. Working Paper.
- Hansen P R, Lunde A, (2005). A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?[J]. *Journal of Applied Econometrics*, 20(7): 873–889.
- Hansen P R, Lunde A, Nason J M, (2011). The Model Confidence Set[J]. *Econometrica*, 79(2): 453–497.

- Huang Z, Tong C, Wang T Y, (2020). Which Volatility Model for Option Valuation in China? Empirical Evidence from SSE 50 ETF Options[J]. *Applied Economics*, 52(17): 1866–1880.
- Huang Z, Wang T Y, Hansen P, (2017). Option Pricing with the Realized Garch Model: An Analytical Approximation Approach[J]. *Journal of Futures Markets*, 37(4): 328–358.
- Lamoureux C G, Lastrapes W D, (1990). Persistence in Variance, Structural Change and the GARCH Model[J]. *Journal of Business & Economic Statistics*, 8: 225–234.
- Maheu J, (2005). Can GARCH Models Capture Long-range Dependence?[J]. *Studies in Nonlinear Dynamics & Econometrics*, 9(4), Article 1.
- Pan Z Y, Wang Y D, Wu C F, Yin L B, (2017). Oil Price Volatility and Macroeconomic Fundamentals: A Regime Switching GARCH-MIDAS Model[J]. *Journal of Empirical Finance*, 43: 130–142.
- Patton A, (2011). Volatility Forecast Comparison Using Imperfect Volatility Proxies[J]. *Journal of Econometrics*, 160(1): 246–256.
- Politis D N, (2007). Model-free Versus Model-based Volatility Prediction[J]. *Journal of Financial Econometrics*, 5(3): 358–359.
- Smetanina E, (2017). Real-time GARCH[J]. *Journal of Financial Econometrics*, 15(4): 561–601.
- Taylor S J, (1986). *Modelling Financial Time Series*[M]. Chichester: John Wiley and Sons.
- Wang F F, Ghysels E, (2015). Econometric Analysis of Volatility Component Models[J]. *Econometric Theory*, 31: 362–393.
- Wang L, Wu J B, Cao Y, Hong Y R, (2022). Forecasting Renewable Energy Stock Volatility Using Short and Long-term Markov Switching GARCH-MIDAS Models: Either, Neither or Both?[J]. *Energy Economics*, 111, Article 106056.