
基于宏观基本面的股市波动度量与预测

郑挺国 尚玉皇*

内容提要 宏观基本面在股市波动成分测度、长期风险识别等方面具有十分重要的作用。本文基于宏观基本面构建了多因子的广义自回归条件异方差-混频数据抽样(GARCH-MIDAS)模型,并运用该模型对中国股市日度波动率进行估计及预测比较。研究结果表明多因子 GARCH-MIDAS 模型可以改进单因子混频模型的估计效果并且能够更好地捕捉股市波动的长期成分,基于宏观经济变量的 GARCH-MIDAS 模型估计结果表明宏观经济波动对股市波动具有显著的正向影响。此外,各类混频模型尤其是多因子 GARCH-MIDAS 模型可以改进基准 GARCH 模型的预测效果,SPA 检验证实了多因子混频波动率模型在提高中国股市波动预测精度方面具有稳健性。

关键词 股市波动 宏观基本面 GARCH 模型 混频数据抽样

一 引言

现代金融理论表明股票市场波动率在金融监管、投资组合、风险管理等领域有至关重要的作用,广泛影响着金融市场参与主体的决策过程。金融监管当局密切关注股市波动以避免股市大起大落对实体经济产生的“波动溢出”效应,市场投资者更是实时跟踪股市波动以期优化投资组合策略,规避市场风险。因此,如何充分利用各类信

* 郑挺国、尚玉皇:厦门大学王亚南经济研究院 通信地址:福建省厦门市思明南路 422 号 厦门大学王亚南经济研究院 361005 电子信箱:zhengt@ gmail. com(郑挺国)。

本研究受到国家自然科学基金面上项目“状态空间混频模型及其在宏观经济中的应用(71371160)”、教育部新世纪优秀人才支持计划“混频数据建模及其在经济中的应用(NCET-13-0509)”及厦门大学基础创新基金“混频数据模型的非线性扩展及应用研究(201322G009)”的资助。作者感谢匿名审稿人的建设性意见,但文责自负。

息精确估计和预测股市波动率是实务界普遍关注的重要问题。

考虑到传统使用同频数据建模的方式将损失高频数据的有效信息以及可能引起的模型误设问题(Ghysels, 2012), 本文在 Engle 等(2009) 单因子广义自回归条件异方差混频数据抽样(GARCH-MIDAS) 模型的基础上提出了一种扩展的多因子 GARCH-MIDAS 模型。该混频波动率模型的主要特点如下: 一是可以同时测度“金融”和“宏观”因子对股市波动长期成分的影响; 二是可以避免单因子混频波动率模型因遗漏重要变量造成参数估计的不一致性问题。运用各种 GARCH-MIDAS 模型, 我们将在混频数据信息环境下对中国经济宏观基本面与股市波动率的关联性进行分析, 并对各模型的样本内拟合效果和样本外预测能力进行比较与分析。

本文其他部分安排如下: 第二部分梳理宏观经济与股市波动关联性的相关文献; 第三部分介绍各类 GARCH-MIDAS 模型; 第四部分对宏观指标和股市波动率进行数据说明; 第五部分为经验分析; 第六部分为本文的主要结论与启示。

二 文献综述

为了解宏观经济与股市波动关联性问题的研究现状, 我们梳理了相关理论假说及计量方法的演进过程, 同时对不同国家和地区的相关经验研究给予归纳总结。

(一) 经济理论假说

股市波动与宏观经济的理论假说主要基于一般均衡模型或者理性预期模型。在金融学理论中, 波动率属于资本资产定价理论的研究范畴, 其中一些经典的资本资产定价理论阐述了宏观经济与股市波动的关系,^①如基于消费的资本资产定价模型(CCAPM) (Breeden, 1979)、股利贴现模型(Campbell 和 Shiller, 1988)、长期风险模型(Bansal 和 Yaron, 2004) 等。首先, 在 Lucas(1978) 的模型框架下, Breeden(1979) 提出基于消费的资本资产定价模型, 该模型指出风险溢价的时变特征受到消费风险及风险厌恶偏好的显著影响, 当宏观经济发生变化时, 人们消费储蓄意愿及风险厌恶偏好随之改变, 进而影响风险溢价; 其次, Campbell 和 Shiller(1988) 及 Campbell(1991) 给出了一个股利贴现模型, 认为当宏观经济基本面发生变化时, 必将对企业的现金流和股利支付产生冲击, 意味着宏观经济将对股票的未预期收益产生冲击, 进而引起股市波动; 最后, Bansal 和 Yaron(2004) 提出一种长期风险模型, 假设股利支付率受长期可预期

^① Casarin 和 Trecoci(2007) 对关联宏观经济与金融风险的理论基础模型进行了详细的阐述。

成分和宏观经济波动成分的影响,该文献认为宏观经济不确定性是解释股票市场异常现象的重要源泉。

此外,还有研究还从其他角度解释宏观经济与股市波动的关联性以及金融市场中的经典悖论。例如 Barro (2006) 提出特定灾难风险 (rare disasters) 模型以解释股市异常现象; Wachter (2013) 在此框架下进一步构建出时变特定灾难风险模型来解释股市波动行为。

(二) 经济计量方法

自 Engle (1982) 提出自回归条件异方差 (ARCH) 模型以来,基于时间序列信息的波动率模型层出不穷,如 Bollerslev (1986) 的广义自回归条件异方差 (GARCH) 模型、Harvey 等 (1994) 的随机波动率 (SV) 模型、Ding 和 Granger (1996) 的分整广义自回归条件异方差 (FIGARCH) 模型以及近年来出现的基于高频数据信息包含跳跃过程的波动率模型 (Andersen 等, 2007)。

考虑到传统波动率方法无法解释和说明波动率的形成原因和经济含义, Engle 和 Lee (1999) 认为有必要对波动率进行成分分解。Engle 等 (2009) 进一步指出波动率成分分解为关联股市波动与宏观经济提供了新的线索。实际上,关于股市波动的大量经验研究表明宏观基本面是股市的主要波动源,如 Hamilton 和 Lin (1996)、Beltratti 和 Morana (2006) 及 Engle 和 Rangel (2008) 等。因此,如能将宏观经济信息引入至波动率模型中将有助于我们更好地解释股市波动成因,进而可能带来估计效果和预测精度的提高。

然而,宏观经济指标相对于高频波动率数据频率较低,关联宏观经济变量与股市波动的分析主要采用同频 (低频) VAR 模型,如 Errunza 和 Hogan (1998) 与 Wang (2011) 等。为克服同频数据建模无法直接利用高频股市波动率信息的不足,近年来一些研究考虑从波动率成分分解的角度构建混频波动率模型。Engle 等 (2009) 提出了 GARCH-MIDAS 模型,该模型区分了高频波动率的长期成分和短期成分,将低频波动率或宏观变量作为单个因子来刻画长期成分,并采用 MIDAS (混频数据抽样) 方法关联单因子与长期成分。^①

虽然 Engle 等 (2009) 提出的单因子 GARCH-MIDAS 模型可以实现高频波动率与

^① 在混频数据问题的处理方法上,由 Ghysels 等 (2004、2007) 提出的混频数据抽样模型 (MIDAS) 备受关注。该方法是处理混频数据问题的重要方法之一,并在实际经济问题分析中得到广泛应用。例如,刘汉和刘金全 (2011) 采用该方法对中国 GDP 进行混频预测分析;彭伟 (2013) 利用 MIDAS 等模型对中国上市商业银行股票日收益风险价值进行预测分析。另一种处理混频数据行之有效的方法是状态空间因子模型,如郑挺国和王霞 (2013) 采用该方法对中国经济周期进行了混频测度与实时分析。

低频宏观经济变量的关联,但其“孤立”了多个因子(宏观经济变量、低频波动率等)对高频波动率长期成分的作用。本文在现有 GARCH-MIDAS 模型的基础上,构造出同时包含低频已实现波动率和宏观经济变量的多因子 GARCH-MIDAS 模型,从而可以改善波动率的样本内拟合,并进一步提高其样本外预测精度。

(三)经验研究

更多的学者通过经验研究考察了宏观基本面对股市波动的影响,这里我们将对发达经济体和新兴市场国家的研究现状分别予以概括与归纳。

就目前经验文献来看,大部分研究关注美国等发达资本市场宏观经济对股市波动率的影响。Beltratti 和 Morana(2006)的研究发现,产出波动对于美国股市波动具有很强的因果关系,而且产出波动对股市波动既有持续性又有非持续性的影响;Engle 等(2009)指出工业生产增长率和通货膨胀率对美国股票市场波动率均有显著影响,而且工业增加值波动也对股市波动有正向影响;Kim 和 Nelson(2014)将股市波动区分为与经济周期相关成分和无关成分,并认为股市波动与经济周期之间存在潜在作用机制。Morelli(2002)发现英国宏观经济波动对股市波动具有显著影响,而 Errunza 和 Hogan(1998)指出欧洲股市波动受到货币或者实体宏观经济因子的显著影响。

近年来,对新兴市场经济体的研究越来越多。Wongbangpo 和 Sharma(2002)通过对亚洲 5 国(印度尼西亚、马来西亚、菲律宾、新加坡、泰国)股市波动率的研究发现,重要宏观经济因子(如 GDP 等)对股市波动率有决定性影响;Chiou 和 Song(2007)指出,在 1997 年金融危机后,货币供应量等宏观变量对韩国、新加坡、中国香港和台湾 4 个国家或地区股市波动率有显著影响;Engle 和 Rangel(2008)基于 Spline-GARCH 模型分析了包括大部分新兴市场 and 发达经济体近 50 个国家和地区的情况,结果表明当 GDP 或通货膨胀等宏观变量波动增加时股市波动率的低频波动成分也会随之增加。

中国作为重要新兴市场经济体之一,其股市波动与宏观经济的关联性也得到了国内学者的关注,但国内学者对该问题研究视角各不相同。首先,部分学者侧重从经济周期的角度考察宏观经济与股市波动的关联性(刘玉军,2011)。其中,赵可和丁安华(2013)探讨了经济周期与股市波动的传导机制,并认为不确定性的反周期性是研究宏观经济波动与股市波动的关键所在。其次,宏观经济变量对股市波动的直接解释能力引起部分学者的关注。赵振全和张宇(2003)采用多元回归和 VAR 模型研究发现,中国宏观经济波动对股市波动的解释能力偏弱;Wang(2011)分析指出相较于实际 GDP,股市波动与通货膨胀率波动之间具有更强的因果关系。最后,还有部分学者强调宏观经济政策在关联宏观基本面对股市波动方面的作用(唐平和刘燕,2008;陈其

安等,2010)。顾巧明(2011)基于 ARMA-FIEGARCH 模型的研究认为,货币政策的调整对股市波动存在显著的非对称效应,并且利率调整较之于准备金率调整对市场的影响更为显著。这些文献丰富了关于中国宏观经济与股市波动关联性问题的研究,但是大部分研究忽略了对波动率不同成分的分解,而且在关联宏观经济与波动率的研究中仅使用低频波动率,损失了高频波动率的有效信息。据笔者所知,利用混频数据信息分析宏观经济对中国股市波动影响的研究还十分匮乏。

三 计量模型

在资本资产定价模型中,未来现金流和折现因子是决定股票价格的两个核心要素。基于对数线性化的方法,Campbell 和 Shiller(1988)提出一种经典的股利贴现模型并以此来反映金融市场资产价格的未预期收益。其公式如下:

$$r_{i,t} - E_{i-1,t}(r_{i,t}) = (E_{i,t} - E_{i-1,t}) \sum_{j=0}^{\infty} \rho^j \Delta d_{i+j,t} - (E_{i,t} - E_{i-1,t}) \sum_{j=1}^{\infty} \rho^j r_{i+j,t} \quad (1)$$

其中, $r_{i,t}$ 表示在 t 月(季度、年)第 i 天时股票的对数收益率, $d_{i,t}$ 表示处于 t 月(季度、年)第 i 天时股票的对数股利支付, $E_{i,t}$ 表示给定对应时期信息集的条件期望, Δ 表示一阶差分算子, ρ 为贴现因子。

Campbell(1991)认为未预期收益可以由未来股利支付和预期折现因子的冲击来解释,也就是说未预期收益可以被看做是对于股票收益率的一种冲击。由式(1)可知,某一信息事件对预期股利的冲击与未预期收益正相关,而其对折现因子的冲击则与未预期收益负相关。进一步地,Engle 和 Rangel(2008)认为收益率的冲击具有时变方差,该方差会受到某一信息事件对收益率冲击的作用强度和作用乘子(impact multiplier)的影响,而该信息事件的作用乘子依赖于宏观经济环境。以状态变量 z_t 表示整个宏观经济环境,Engle 和 Rangel(2008)将收益率的冲击表示为:

$$r_{i,t} - E_{i-1,t}(r_{i,t}) = \sqrt{\tau_1(z_t)} \eta_{i,t} \quad (2)$$

其中, $\eta_{i,t}$ 表示某信息事件对收益率冲击的作用强度, $\tau_1(z_t)$ 表示该信息事件对收益率冲击的作用乘子。由于信息事件本身(强度和大小)也受到宏观经济或者其他不可观测状态变量的影响,因此将 $\eta_{i,t}$ 进一步表示为 $\eta_{i,t} = \sqrt{\tau_2(z_t)} g_{i,t} \varepsilon_{i,t}$, 其中 $\tau_2(z_t)$ 表示 $\eta_{i,t}$ 中受宏观经济或其他状态变量影响的部分, $g_{i,t}$ 视为非负的时间序列(如 GARCH 过程), $\varepsilon_{i,t}$ 为标准化的白噪声过程。在没有其他更多信息的情况下, $\tau_1(z_t)$ 和 $\tau_2(z_t)$ 两种成分一般难以识别。因此,令 $\tau(z_t) = \tau_1(z_t) \tau_2(z_t)$, 可得:

$$r_{i,t} - E_{i-1,t}(r_{i,t}) = \sqrt{\tau(z_t)g_{i,t}}\varepsilon_{i,t} \quad (3)$$

由式(3)可知,波动率可由 $g_{i,t}$ 和 $\tau(z_t)$ 两种成分构成,其中 $g_{i,t}$ 表示短期的日度高频波动, $\tau(z_t)$ 表示某种长期的低频波动。短期成分一般会受到股票市场日度流动性及其他短期因子的影响,Chordia 等(2002)的研究表明,流动性对市场波动率有重要影响,长期成分则与未来现金流和折现率密切相关。更为重要的是,宏观经济变量可以看做是股票市场的波动源(Engle 等,2009)。因此,GARCH-MIDAS 模型可以看做是一种特殊的成分波动率模型。

(一)基于低频波动率的 GARCH-MIDAS 模型

在式(3)的基础上,Engle 等(2009)提出了一种新的成分模型,称之为 GARCH-MIDAS 模型。Engle 和 Rangel(2008)指出低频波动率可以反映宏观经济行为,如果采用低频已实现波动率来刻画股票市场波动率的长期成分则会得到基于已实现波动率的 GARCH-MIDAS 模型。

$r_{i,t}$ 仍表示在 t 月(季度、年)第 i 天股票的对数收益率,即日度频率数据,并假设在时期 t 内总共有 N_t 天。于是,将式(3)略作修改整理可得,对任意 $i = 1, \dots, N_t$ 有:

$$r_{i,t} = \mu + \sqrt{\tau_t g_{i,t}} \varepsilon_{i,t} \quad (4)$$

其中,假设扰动项 $\varepsilon_{i,t}$ 服从条件标准正态分布,即 $\varepsilon_{i,t} | \psi_{i-1,t} \sim N(0,1)$, $\psi_{i-1,t}$ 表示在 t 月第 $i-1$ 天时可获得的历史信息集。进一步地,假设日度波动率短期动态成分 $g_{i,t}$ 服从 GARCH(1,1) 过程,即:

$$g_{i,t} = (1 - \alpha - \beta) + \alpha (r_{i-1,t} - \mu)^2 / \tau_t + \beta g_{i-1,t} \quad (5)$$

在 GARCH-MIDAS 模型中, τ_t 表示波动率的长期成分,其可由某个低频变量刻画。参考 Ghysels 等(2004、2007)提出的 MIDAS(混频数据抽样)回归方法,我们可以使用基于收益率的已实现波动率 $RV_t = \sum_{i=1}^{N_t} r_{i,t}^2$ 来刻画长期成分 τ_t ,则有:

$$\tau_t = m + \theta \sum_{k=1}^K \varphi_k(\omega_1, \omega_2) RV_{t-k} \quad (6)$$

式(6)中, K 表示低频变量的最大滞后阶数,而 $\varphi_k(\omega_1, \omega_2)$ 表示基于 Beta 函数构造的权重方程,即:

$$\varphi_k(\omega_1, \omega_2) = \frac{f(k/K, \omega_1, \omega_2)}{\sum_{k=1}^K f(k/K, \omega_1, \omega_2)} \quad (7)$$

其中:

$$f(x, a, b) = \frac{x^{a-1} (1-x)^{b-1} \Gamma(a+b)}{\Gamma(a) + \Gamma(b)} \quad (8)$$

上述(4)~(8)式共同构成基于收益率信息已实现波动率的 GARCH-MIDAS 模型,其中(6)式也称为 MIDAS 滤波。与基本的 GARCH 模型相比,GARCH-MIDAS 模型的主要特点是增加了成分方程的设定。另外,通过对式(6)进行对数化处理便可得到如下形式的已实现波动率 GARCH-MIDAS 模型:

$$\log \tau_t = m + \theta \sum_{k=1}^K \varphi_k(\omega_1, \omega_2) RV_{t-k} \quad (9)$$

考虑对数化的 GARCH-MIDAS 模型可以更好地匹配宏观经济变量。与波动率指标不同,宏观经济变量的符号可以为负也可以为正,采用对数形式可以更好地适用任何符号的变量,而在 $\log \tau_t$ 指数化之后,其所刻画的波动率长期成分 τ_t 始终为正值,符合混频波动率模型设定长期成分的基本要求。

(二) 基于宏观经济变量的 GARCH-MIDAS 模型

如前所述,股票价格波动的长期成分与宏观经济状态(更确切来说是宏观经济变量)密切相关,我们将分别构建基于宏观经济变量水平值和波动率的 GARCH-MIDAS 模型。类似于 Engle 等(2009)的做法,我们将宏观经济变量的变化率作为其相应的水平值,于是将式(9)替换得到:

$$\log \tau_t = m_l + \theta_l \sum_{k=1}^{K_l} \varphi_k(\omega_{1,l}, \omega_{2,l}) X_{l,t-k}^{mv} \quad (10)$$

其中, $X_{l,t-k}^{mv}$ 表示某个宏观经济变量相对当期(t 期)滞后 k 期的水平值。 K_l 表示该变量水平值的最大滞后阶数。因此,式(4)、(5)、(7)、(8)和(10)共同构成基于宏观经济变量水平值的 GARCH-MIDAS 模型。

在构建宏观经济变量波动率的 GARCH-MIDAS 模型之前,我们首先对宏观经济变量的波动率进行简单计算,常见的方法是使用 $AR(p)$ 模型得到残差项并通过残差项构造波动率的代理变量(Schwert, 1989)。本文亦考虑使用自回归模型来捕捉宏观经济的波动态势,我们将确定最大滞后阶数 p ,通过 AIC 信息准则选择最优滞后阶数,并基于相应模型的残差取平方得到宏观经济变量波动率的代理变量。我们可以构造如下方程:

$$\log \tau_t = m_v + \theta_v \sum_{k=1}^{K_v} \varphi_k(\omega_{1,v}, \omega_{2,v}) X_{v,t-k}^{mv} \quad (11)$$

其中, $X_{v,t-k}^{mv}$ 表示某个宏观经济变量相对当期(t 期)滞后 k 期的波动率。 K_v 表示该变量波动率的最大滞后阶数。因此,式(4)、(5)、(7)、(8)和(11)共同构成基于宏观经济波动率的 GARCH-MIDAS 模型。

(三)多因子 GARCH-MIDAS 模型

上述模型均是使用单变量(低频波动率或者宏观经济变量)进行混频波动率建模,本文视之为单因子 GARCH-MIDAS 模型。然而,单因子混频波动率模型无法同时捕捉多个变量的边际贡献,更为严重的是单因子模型因忽略其他因子的贡献可能存在一定的模型误设问题。为此本文构建一种多因子 GARCH-MIDAS 模型,同时捕捉不同因子对股市波动率长期成分的影响。设定多因子混频波动率模型有助于我们更好地捕捉股市波动的运行态势,深入认识宏观基本面对高频波动率的作用机制,同时还可以提高混频波动率模型的估计效果,从而改进现有波动率模型的样本外预测精度。

另外,我们需要注意的是构建多因子 GARCH-MIDAS 模型虽然重要,但如何处理变量的不同数据类型对设定多因子混频波动率模型来说不容忽视。因为股市低频波动率与宏观经济变量数据类型迥然不同,如果我们使用原始数据直接建模,无法在同一标准下衡量不同因子之间的边际贡献,甚至会导致模型的估计精度较差。本文首先对数据进行标准化处理,并在此基础上进行多因子 GARCH-MIDAS 模型的设定与估计。

多因子混频波动率模型与单因子模型的关键区别在于长期成分的设定。假设有 N 个因子共同作用于长期成分,则长期成分方程的一般形式设定如下: $\log \tau_t = \tilde{m} + \sum_{j=1}^N \tilde{\theta}_j \sum_{k=1}^K \varphi_k(\omega_{1,j}, \omega_{2,j}) X_{t-k}^j$, 其中权重方程 $\varphi_k(\omega_{1,j}, \omega_{2,j})$ 的形式与前文中式(7)和(8)的设定形式一样。例如双因子 GARCH-MIDAS 模型中刻画长期成分的方程可设定如下:

$$\log \tau_t = \tilde{m} + \tilde{\theta}_R \sum_{k=1}^K \varphi_k(\omega_{1,R}, \omega_{2,R}) RV_{n,t-k} + \tilde{\theta}_M \sum_{k=1}^K \varphi_k(\omega_{1,M}, \omega_{2,M}) X_{n,t-k}^m \quad (12)$$

其中, $RV_{n,t-k}$ 表示相对当期(t 期)滞后 k 期的标准化的对数低频波动率(这里 n 为变量的标准化标识), $X_{n,t-k}^m$ 表示相对当期(t 期)滞后 k 期标准化的宏观经济变量(水平值或波动率), $\tilde{\theta}_R$ 和 $\tilde{\theta}_M$ 分别表示低频波动率和宏观经济变量对长期成分的边际贡献。因此,式(4)、(5)、(7)、(8)和(12)共同构成双因子 GARCH-MIDAS 模型。

四 变量选择与数据说明

本文主要选取的变量为日度收益率、低频波动率、宏观经济变量,其中低频波动率指标包括已实现波动率和已实现极差,宏观经济变量为宏观经济景气一致和先行指标。其选择与构建过程为:

首先,已实现波动率和已实现极差基于股票价格日度信息构建。我们选取上证综合指数的收盘价、最高价及最低价三种价格信息作为原始数据,样本区间为 1992 年 1 月 2 日至 2012 年 12 月 31 日,每种价格信息的样本量均为 5134 个。根据收盘价信息可计算得到对数收益率,并进一步计算出基于收益率的月度已实现波动率。由于收益率指标仅考虑收盘价的信息,忽略了某一时段内价格变动信息即极差(range)信息。考虑到极差信息在构建波动率方面的明显优势(Parkinson,1980),本文还将构建如下已实现极差作为低频波动率的代理变量,即: $RV_t = (4\log 2)^{-1} \sum_{i=1}^{N_t} [\log(H_{i,t}) - \log(L_{i,t})]^2$ 。^① 其中 $H_{i,t}$ 和 $L_{i,t}$ 分别表示在 t 月(季度、年)第 i 天股票交易的最高价和最低价。以上原始数据来源于锐思(RESSET)金融研究数据库。

其次,我们选取宏观经济景气先行和一致指数作为宏观基本面的代理变量。国内生产总值(GDP)是刻画宏观经济运行状态的重要指标,但是该指标为季度数据,相对于其他月度宏观变量频率较低,往往表现出时效性差的劣势。为此许多文献考虑使用工业增加值(IP)等指标来代替 GDP(Schwert,1989;Engle 等,2009),但工业增加值不能全面反映宏观经济运行状况。

宏观经济景气指数作为月度观察指标与 GDP 相比具有更高的数据频率,而相对于工业增加值而言,可以更好地反映宏观经济周期的特征。郑挺国和王霞(2013)、李建军(2008)等国内学者认为宏观经济景气指标可以刻画中国经济周期变化趋势。与此同时,经济景气指数广泛应用于国债期限结构利差(康书隆和王志强,2010)、股指期货价格(王博和伍楠林,2012)、股票市场流动性(刘向华和柳恩普,2013)等多个金融指标的分析中。参考 Shyu 和 Hsia(2008)的做法,本文选取中国经济景气监测中心一致指数和先行指数^②水平值及其对应的波动率作为宏观基本面的代理变量。样本区间为 1992 年 1 月~2012 年 12 月。数据来源于《中国经济景气月报》、CEIC 中国经济数据库、新浪财经中国宏观经济数据。图 1 描述了上证综指日度对数收益率和月度宏观指标,他们共同构成混频数据信息集。

图 2 给出了上证综指收益率及其波动率的。如图所示,与收益率不同,波动率的自相关系数具有很强的持续性,但这种持续性并没有表现出一种平滑衰减的趋势,这与 Ding 和 Granger(1996)对美国和日本市场的描述一致。Ding 和 Granger(1996)指出当波动率的持续性表现出该行为特征时,标准的 GARCH 模型可能是不适用的,并指

① 具体构建细节可参考 Martens 和 Dijk(2007)的做法,限于篇幅及重要性这里不再赘述。

② 中国经济景气监测中心一致指数由工业企业利润、居民可支配收入、进出口、投资等宏观指标构建;先行指数则由货币供应量、工业产品产销率、消费者预期指数、恒生中国内地物流指数等指标构建。

究其原因可能是波动率受到长短期不同成分的影响。这为从波动率成分角度进行混频波动率建模提供了很好的经验支持。

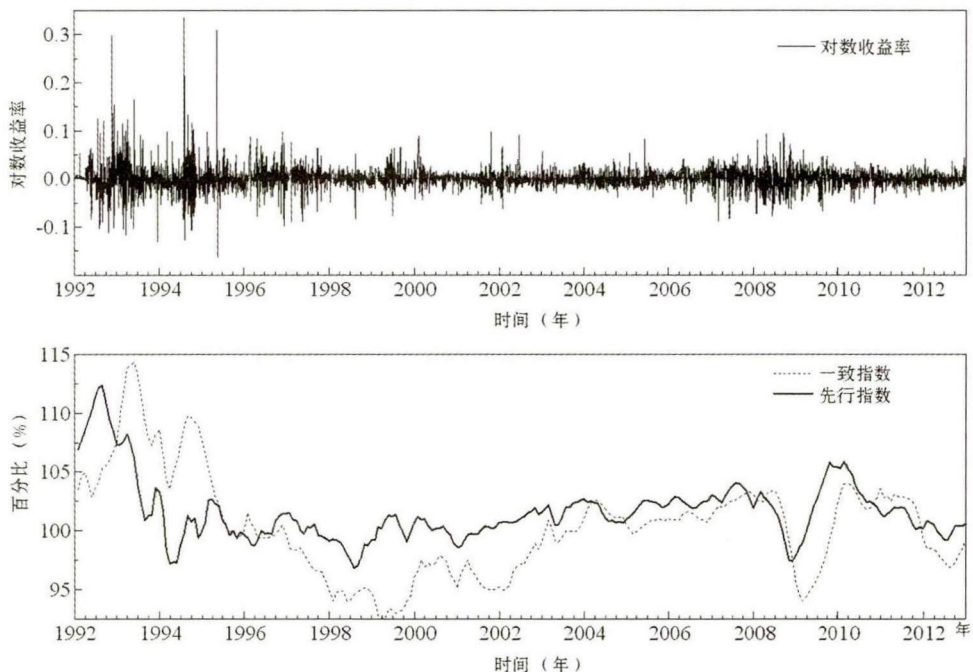


图1 上证综指对数收益率与宏观经济变量混频数据信息

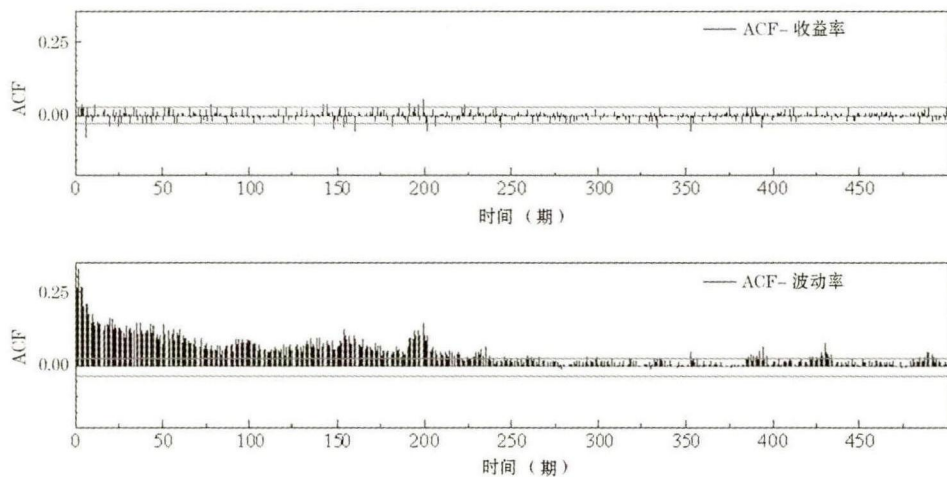


图2 上证收益率和波动率自相关关系

五 经验分析结果

(一) 单因子 GARCH-MIDAS 模型估计分析

单因子混频波动率模型中设定的低频变量为低频波动率或宏观经济变量。除估计单因子混频波动率模型外,我们还给出 GARCH(1,1)模型的估计结果。不同于现有文献的做法,本文利用已实现极差信息来构建低频波动率 GARCH-MIDAS 模型,尝试验证各类混频波动率模型在中国股市波动率研究中的适用性。文中选取一致指数和领先指数的水平值(减去基数 100)和波动率^①构建基于宏观经济变量的 GARCH-MIDAS 模型,宏观变量的水平值和波动率分别作标准化和对数标准化处理。

在 GARCH-MIDAS(P, Q, K)模型中 P 和 Q 表示 GARCH 过程的滞后阶数,本文仅考虑 P 和 Q 取值为 1 的情况; K 表示 MIDAS 滤波所选取的滞后阶数,文献中惯常以 MIDAS 滞后年来反映滞后情况,如当 K 为 1 时,在本文中表示 1 个 MIDAS 滞后年相当于低频数据包含滞后 12 个月的信息,由此得到 GARCH-MIDAS(1,1,1)模型。同理还可以构造出 GARCH-MIDAS(1,1,2)模型。表 1 报告了基于低频波动率和宏观经济变量的 GARCH-MIDAS(1,1,2)和 GARCH(1,1)模型的结果。^②

表 1 第 3 列和第 4 列分别给出基于对数标准化的已实现极差和基于相应收益率信息已实现波动率 GARCH-MIDAS 模型的估计结果。首先,模型参数的估计结果大都具有统计意义上的显著性,其中 $\tilde{\omega}$ 可以反映出基于低频波动率的单因子混频模型的最优估计权重,^③ $\tilde{\theta}$ 可以反映出中国股市低频波动率对高频波动率长期成分的影响;其次,依据 AIC 和 HQ 信息准则,可以发现各类 GARCH-MIDAS 模型优于基准的 GARCH 模型,体现出混频波动率模型在刻画中国股市波动率方面的适应性。特别的,基于已实现极差 GARCH-MIDAS 模型也优于基于收益率已实现波动率 GARCH-MIDAS 模型。

表 1 第 5~8 列给出了基于经济景气指数水平值和波动率构建的单因子混频波动率模型的估计结果。依据 AIC 和 HQ 信息准则,本文选取滞后 24 个月的经济景气指

① 我们采用 AR(p)模型得到残差项并通过残差项的平方得到宏观经济变量波动率,在分析中根据 AIC 信息准则来选取最优滞后阶数,并估计得到相应模型的残差来计算宏观经济变量的波动率。

② 本文还对基于极差和收益率的已实现波动率 GARCH-MIDAS 模型即 G-M-Range、G-M-Return 模型进行了估计,限于篇幅及重要性未给予详细报告,感兴趣的读者可向作者索取。

③ 参考 Engle 等(2009)的研究,这里对权重方程式(7)的估计,约束参数 $\omega_1 = 1$,仅仅估计参数 ω_2 。

数水平值和波动率构建的混频模型的予以报告。首先,模型参数的估计结果大都具有统计上显著性,参数 $\tilde{\theta}$ 估计值表明宏观基本面对股市波动率长期成分的边际贡献;其次,基于宏观经济波动率设定的混频波动率模型略优于相应水平值的混频波动率模型。特别的,在基于经济景气指数水平值的 GARCH-MIDAS 模型中,参数 $\tilde{\theta}$ 估计值的符号为正且在统计上显著,说明宏观经济的水平状态或者说增长率的增加可以推动股票市场波动率的上升;在基于经济景气指数波动率的 GARCH-MIDAS 模型中,参数 $\tilde{\theta}$ 估计值显著且符号为正说明宏观经济波动增加会导致股市波动增加。

表 1 单因子混频波动率模型估计结果

	GARCH ^a (1,1)	G-M-Nor-Range (1,1,2)	G-M-Nor-Return (1,1,2)	G-M-Con-L (1,1,2)	G-M-Lead-L (1,1,2)	G-M-Con-V (1,1,2)	G-M-Lead-V (1,1,1)
μ	0.0004 * (0.0002)	0.0003 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)	0.0003 (0.0002)	0.0002 (0.0002)
$\varphi \times 10^{-3}$	0.04 *** (0.01)						
α	0.0859 *** (0.0079)	0.1403 *** (0.0148)	0.1381 *** (0.0138)	0.1351 *** (0.0041)	0.1216 *** (0.0058)	0.0960 *** (0.0091)	0.0989 *** (0.0102)
β	0.9112 *** (0.0082)	0.7729 *** (0.0267)	0.7773 *** (0.0219)	0.8307 *** (0.0067)	0.8569 *** (0.0084)	0.8941 *** (0.0083)	0.8839 *** (0.0117)
m		-7.9781 *** (0.0701)	-7.9631 *** (0.0661)	-7.7917 *** (0.0973)	-7.6015 *** (0.0316)	-7.4573 *** (0.5174)	-7.7423 *** (0.1988)
θ		1.2854 *** (0.0654)	1.2029 *** (0.0636)	0.7986 *** (0.7571)	0.6313 *** (0.0075)	3.2413 *** (0.2267)	1.8841 *** (0.1697)
ω		4.0519 *** (0.5138)	3.8500 *** (0.4758)	0.8835 *** (0.0077)	0.5322 *** (0.0166)	1.1511 *** (0.0367)	1.3925 *** (0.1031)
AIC	-7.1234	-7.1644	-7.1635	-7.1728	-7.1343	-7.2213	-7.2252
HQ	-7.1205	-7.1614	-7.1606	-7.1699	-7.1314	-7.2183	-7.2221

说明:G-M-Nor-Range、G-M-Nor-Return 分别表示对数标准化的基于极差、收益率信息的已实现波动率 GARCH-MIDAS 模型;G-M-Con-L、G-M-Lead-L 分别表示基于标准化的宏观经济一致指数、先行指数水平值构造的 GARCH-MIDAS 模型;G-M-Con-V、G-M-Lead-V 分别表示基于对数标准化的宏观经济一致指数、先行指数波动率构造的 GARCH-MIDAS 模型。AIC 和 HQ 分别表示两种模型选择信息准则。括号内的值表示稳健标准误差,*、**、***分别表示在 10%、5%、1% 的显著性水平下显著。下表同。

^a本文使用 GARCH(1,1)模型作为基准模型,其均值方程为 $r_t = \mu + z_t$, 其中 $z_t = \sqrt{g_t} \varepsilon_t$ 并且 $\varepsilon_t \sim iidN(0,1)$;方差方程为 $g_t = \varphi + \alpha z_{t-1}^2 + \beta g_{t-1}$ 。

以上获得的关于中国宏观经济波动对股市波动正向影响的结果符合相关经典理论的基本观点,如 Campbell(1991)对数线性化的股利贴现模型和 Wachter(2013)时变特定灾难风险理论。同时该结果也与 Beltratti 和 Morana(2006)、Engle 等(2009)对美国等发达经济体以及 Wongbangpo 和 Sharma(2002)、Engle 和 Rangel(2008)对新兴市场经济体的研究结论相一致。正如 Engle 和 Rangel(2008)指出的那样,当宏观经济波动增加时,股市波动随之增加的现象,在新兴市场经济体和发达经济体都会出现。

我们还可以通过中国宏观基本面波动的典型事实直观分析宏观经济波动对股市波动正向影响关系。图 3 描述了中国股市波动与宏观经济基本面波动的基本情况,其中股市波动使用日度极差指标来刻画,宏观基本面波动由经济景气一致指数波动率刻画,图中阴影部分表示中国宏观经济高波动的时期。如图所示,受美国次贷危机及欧债危机影响,中国宏观经济出现大幅波动,与此同时中国股市也经历了一轮大幅波动,这一时期中国股市与宏观基本面表现出特别明显的一致高波动特征。同样追溯至上世纪末,受 1997 年东南亚金融危机的影响,中国宏观经济波动加剧而股市波动亦表现剧烈,中国股市与宏观基本面也呈现出相对一致的高波动区制。

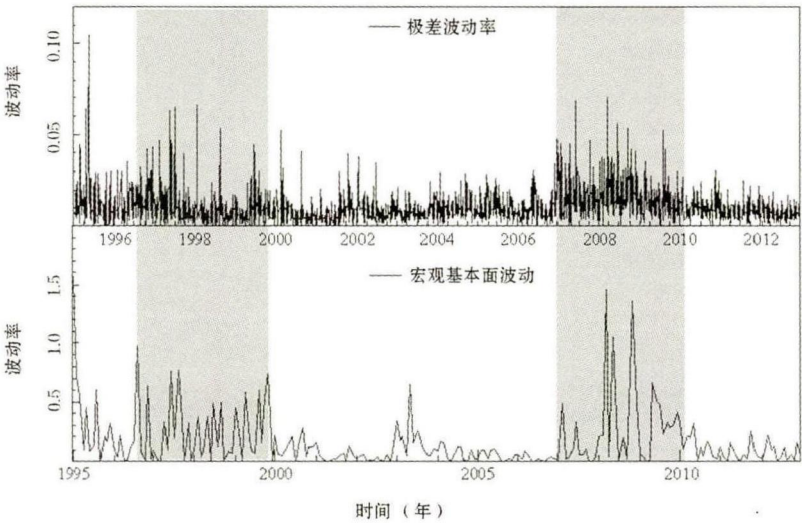


图 3 中国股市波动与宏观基本面波动

进一步的,我们可以从生产层面和宏观经济政策方面深入认识宏观经济与股市波动的关联性。首先,从生产与消费层面上来看,根据股利贴现模型的基本原理,宏观经济会从生产层面影响股市波动。当宏观经济基本面发生变化时,会影响到企业的利润

即未来现金流,企业未来现金流的波动最终导致股票价格的波动。从理论上来说股市波动与消费密切相关,陈强和叶阿忠(2009)等探讨了中国股市波动与居民消费的关系,但主要分析股市波动对居民消费的影响即股市的财富效应,多数学者认为中国股市的财富效用较弱。宏观经济从消费层面影响股市波动可能性很小,中国经济增长主要依靠投资和进出口拉动,因此宏观经济更有可能从生产层面影响股市波动。其次,有学者认为中国股市波动会受到宏观经济政策调整(如增加货币供应量)的影响(唐平和刘燕,2008)。当宏观经济基本面发生变化时,宏观经济政策的制定者会通过调整货币政策,刺激经济或者抑制通货膨胀,这同时也对股市波动产生影响。例如,受美国次贷危机及全球金融危机的影响,2008年中国宏观经济出现下行趋势,货币政策当局频繁调整法定存款准备金率,同时货币供应量增加对于股市波动产生直接影响,一定程度上说明了中国股市具有资金拉动特征(陈其安等,2010),也从侧面反映出中国股票市场是一个“政策市”。因此,短期内宏观经济政策的大幅调整会导致股市的异常波动。应该说深入认识宏观经济面与股市波动率的关联性具有很好的现实意义,监管当局需要依据宏观经济周期的变化情况制定股票市场波动的周期性监管措施,避免股票市场出现大起大落。

(二)多因子 GARCH-MIDAS 模型分析

基于混频波动率模型对中国股市波动率的估计结果可以得知,低频波动率和宏观经济变量即“金融”与“宏观”因子均能影响高频波动率的长期成分,如何同时考虑多个因子对长期成分的共同作用,是本文的主要特色。基于已实现极差和宏观经济变量(水平值或波动率)构造的多因子 GARCH-MIDAS 模型,我们实现了相应模型的估计,相关估计结果见表 2。^①

从表 2 中我们可以获得以下基本结果:首先,各类混频模型中大部分参数具有统计意义上的显著性。其中,参数 $\bar{\theta}_L$ 和 $\bar{\theta}_V$ 估计结果均显著为正,再一次说明宏观基本面对波动率长期成分的影响及中国宏观经济波动对股市波动有正向影响,这与单因子混频波动率模型的分析结论一致。其次,依据 AIC、HQ 信息准则,对比表 1 和表 2 可以发现多因子 GARCH-MIDAS 模型的设定优于相应单因子 GARCH-MIDAS 模型,三变量 GARCH-MIDAS 模型略优于双变量 GARCH-MIDAS 模型。多因子混频波动率模型具有更好的样本内拟合结果,主要是因为其包含更多影响高频波动率长期成分有

^① 这里仅报告了基于已实现极差信息和宏观经济变量构造的多因子 GARCH-MIDAS 模型的估计结果,我们还可以使用已实现波动率和宏观经济变量构造多因子 GARCH-MIDAS 模型,限于篇幅这里不做详细报告。

效信息,降低遗漏重要变量的可能性。模型信息准则结果表明“金融”因子和“宏观”因子对中国股市高频波动率均有贡献。

表 2 多因子混频波动率模型估计结果

	G-M-Ra-Con-L-(1,1,2)	G-M-Ra-Lead-L-(1,1,2)	G-M-Ra-Con-V-(1,1,2)	G-M-Ra-Lead-V-(1,1,2)	G-M-Ra-Con-L-V-(1,1,2)	G-M-Ra-Lead-L-V-(1,1,2)
μ	0.0002 (0.0002)	0.0002 (0.0002)	0.0003 (0.0491)	0.0002 (0.0009)	0.0003 (0.2071)	0.0002 (0.1981)
α	0.1595 *** (0.0147)	0.1404 *** (0.0115)	0.1172 *** (0.0137)	0.1205 *** (0.0253)	0.1202 *** (0.0123)	0.1255 *** (0.0127)
β	0.7263 *** (0.0282)	0.7702 *** (0.0225)	0.7981 *** (0.0217)	0.7952 *** (0.0459)	0.8073 *** (0.0288)	0.7943 *** (0.0213)
\tilde{m}	-7.9980 *** (0.0567)	-7.9959 *** (0.0771)	-8.0578 *** (0.0770)	-8.0495 *** (0.1985)	-7.9816 *** (0.0692)	-8.0248 *** (0.0926)
$\tilde{\theta}_R$	0.5852 *** (0.0350)	1.0943 *** (0.0431)	1.2246 *** (0.1347)	1.1493 *** (0.0264)	0.8953 *** (0.1122)	1.0282 *** (0.0388)
$\tilde{\theta}_L$	0.4376 *** (0.0585)	0.2014 *** (0.0138)			0.3955 *** (0.0916)	0.1852 *** (0.0565)
$\tilde{\theta}_V$			0.3177 *** (0.1392)	0.2688 *** (0.0271)	0.6019 *** (0.2002)	0.2483 *** (0.0829)
$\tilde{\omega}_R$	23.907 *** (5.1487)	4.3736 *** (0.1232)	3.9089 *** (0.5773)	4.1786 *** (0.1219)	4.0078 *** (1.0550)	4.0727 *** (0.4492)
$\tilde{\omega}_L$	0.8679 *** (0.0292)	0.5573 *** (0.1262)			2.6555 *** (0.6637)	0.6316 *** (0.1355)
$\tilde{\omega}_V$			1.1151 *** (0.0761)	1.3147 *** (0.0812)	1.1049 *** (0.0781)	1.2750 *** (0.0896)
AIC	-7.1797	-7.1681	-7.2275	-7.2274	-7.2348	-7.2296
HQ	-7.1758	-7.1642	-7.2235	-7.2233	-7.2297	-7.2245

说明:G-M-Ra-Con-L 表示一致指数水平值和已实现极差构造的双变量 GARCH-MIDAS 模型;G-M-Ra-Lead-L 表示先行指数水平值和已实现极差构造的双变量 GARCH-MIDAS 模型;G-M-Ra-Con-V 表示一致指数波动率和已实现极差构造的双变量 GARCH-MIDAS 模型;G-M-Ra-Lead-V 表示先行指数波动率和已实现极差构造的双变量 GARCH-MIDAS 模型;G-M-Ra-Con-L-V 一致指数水平值、波动率和已实现极差构造的多因子 GARCH-MIDAS 模型;G-M-Ra-Lead-L-V 表示先行指数水平值、波动率和已实现极差构造的多因子 GARCH-MIDAS 模型。这里仅考虑 MIDAS 滞后年取 2 即月度经济景气指数和已实现极差滞后 24 个月的情况。

进一步的,令 $\sqrt{\tau_t g_{i,t}}$ 和 $\sqrt{\tau_t}$ 分别表示日度对数收益率的条件方差和长期成分,则

可分别估计得到单因子和多因子混频波动率模型的条件方差与长期成分,见图4。如图所示,长期成分相比于条件方差的波动性较小,长期成分与条件方差的高低起伏基本吻合,表明长期成分能够反映股市波动的整体趋势;此外,与单因子混频波动率模型相比,包含宏观经济变量的多因子混频波动率模型所刻画的长期成分能够更好地识别股市波动趋势。

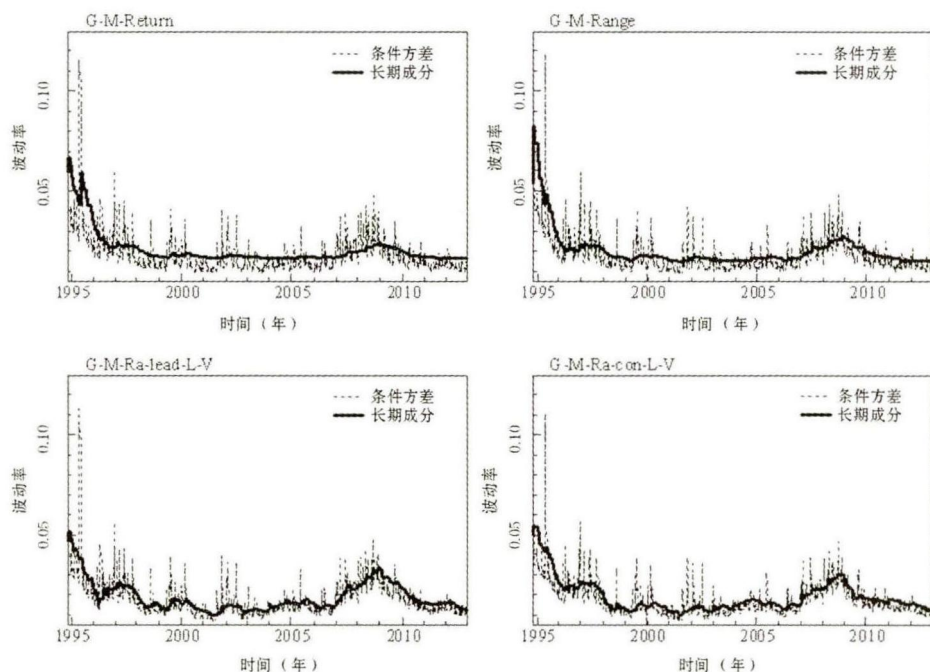


图4 单因子和多因子混频波动率模型估计的条件方差和长期成分

(三) 基于 GARCH-MIDAS 模型的波动率预测分析

本文下面将采用滚窗的样本外预测方法进一步比较多因子和单因子混频波动率模型对基本 GARCH 模型预测精度的改进效果。具体而言,我们基于滚窗样本区间进行估计,并计算出样本外(下一个月)月度波动率的预测值,进而采用已实现波动率的构造方法计算月度波动率的预测结果。本文设定的滚窗样本外预测区间为 2010 年 1 月 2 日至 2012 年 12 月 31 日共计 36 个月。样本外预测效果的评价标准为 MAE(预测绝对误差)和 RMSE(预测方根误差),具体计算如下: $MAE = T^{-1} \sum_{t=1}^T | \hat{\sigma}_t^2 - \sigma_t^2 |$, $RMSE = (T^{-1} \sum_{t=1}^T (\hat{\sigma}_t^2 - \sigma_t^2)^2)^{1/2}$, 其中 $\hat{\sigma}_t^2$ 表示模型估计的条件方差, σ_t^2 表示股市波动率的代理变量,这里使用月度已实现波动率表示。此外,我们也选取 GARCH

(1,1)模型作为基准模型,进而采用 $rMAE$ (相对预测绝对误差) 和 $rRMSE$ (相对预测绝对误差) 分析 GARCH-MIDAS 模型在改进模型预测精度方面的作用。我们基于 GARCH-MIDAS 模型得到的 MAE ($RMSE$) 除以相同样本 GARCH(1,1) 模型的 MAE ($RMSE$) 可计算出相对指标 $rMAE$ ($rRMSE$)。

表 3 汇总了各类 GARCH-MIDAS 模型计算出的 $rMAE$ 和 $rRMSE$ 值。从表中报告的结果可以发现,相对于 GARCH(1,1)模型,几乎所有的 GARCH-MIDAS 模型均可以提高预测精度;与单因子模型相比,多因子混频波动率模型对提高预测精度的贡献更大,进一步体现出多因子 GARCH-MIDAS 模型设定的合理性。^① 特别的,对于基于低频波动率的单因子 GARCH-MIDAS 模型(如 G-M-Nor-Range),当加入宏观经济变量构成多因子 GARCH-MIDAS 模型(如 G-M-Ra-Con-V)时,波动率的预测精度得到进一步提高,反映出宏观经济变量对于股市波动率的预测亦有贡献。

表 3 各类 GARCH-MIDAS 模型 $rMAE$ 和 $rRMSE$ 估计结果

单因子模型	$rRMSE$	$rMAE$	多因子模型	$rRMSE$	$rMAE$
G-M-Nor-Range	0.6806	0.9286	G-M-Ra-Con-L	0.5423	0.7545
G-M-Nor-Return	0.6206	0.8904	G-M-Ra-Lead-L	0.6886	0.7822
G-M-Con-L	0.7344	1.0298	G-M-Ra-Con-V	0.5733	0.8301
G-M-Lead-L	0.6597	0.8774	G-M-Ra-Lead-V	0.6065	0.8697
G-M-Con-V	0.5848	0.8309	G-M-Ra-Con-L-V	0.5311	0.7478
G-M-Lead-V	0.8155	1.0381	G-M-Ra-Lead-L-V	0.7145	0.7213

说明:G-M-Nor-Range、G-M-Nor-Return、G-M-CON-L、G-M-Lead-L、G-M-Con-V、G-M-Lead-V 共 6 个单因子 GARCH-MIDAS 模型与表 1 中的对应模型设定完全一致;G-M-Ra-Con-L、G-M-Ra-Lead-L、G-M-Ra-Con-V、G-M-Ra-Lead-V、G-M-Ra-Con-L-V、G-M-Ra-Lead-L-V 共 6 个多因子 GARCH-MIDAS 模型与表 2 中的对应模型设定完全一致。 $rMAE$ 和 $rRMSE$ 分别表示相对预测绝对误差和相对预测方根误差。

为说明多因子混频波动率模型样本外预测结果的稳健性,本文借鉴 Hansen (2005) 提出的模型预测能力检验方法,即 SPA (Superior Predictive Ability) 检验方法进行稳健性检验。该方法原假设为基准或目标预测模型不差于其他预测模型,并通过损失函数定义相对损失变量,基于相对损失变量构造检验统计量。该统计量在原假设下

① 需要指出的是预测效果的好坏具有一定的偶然性,这里多因子混频模型预测结果较好是一个总体概念,为了说明预测结果稳健性,一般需要进行预测检验,本文将使用 SPA 进行检验。

基于 Bootstrap 方法得到 SPA 检验相应的 p 值,具体过程可参见本文附录。SPA 检验的基准预测模型为单因子混频波动率模型,其他非目标预测模型为相关的多因子 GARCH-MIDAS 模型。^① 表 4 报告了基于 10 000 次 Bootstrap 抽样得到的 SPA 检验结果。在本文中 SPA 检验原假设的具体含义是单因子模型的预测效果不差于多因子模型,从表 4 中可以看出在一定的显著性水平下,SPA 检验的 p 值基本拒绝原假设。因此,本文认为多因子混频 GARCH-MIDAS 模型预测能力优于单因子模型,并且该预测效果具有一定的稳健性。

表 4 模型预测能力 SPA 检验结果汇总(p 值)

	G-M-Nor-Range	G-M-Con-L	G-M-Lead-L	G-M-Con-V	G-M-Lead-V
DS	0.0976 *	0.6767	0.0253 **	0.0849 *	0.3911
DA	0.0309 **	0.0419 **	0.0061 ***	0.0178 **	0.4949

说明:G-M-Nor-Range、G-M-CON-L、G-M-Lead-L、G-M-Con-V、G-M-Lead-V 共 5 个单因子 GARCH-MIDAS 模型与表 1 中的对应模型设定完全一致,这 5 个单因子模型分别作为目标预测模型。DS 和 DA 表示构造统计量的相对损失变量分别方差形式和绝对值形式。

六 结论与启示

对中国这样的新兴市场国家来说,考察宏观经济基本面对金融市场波动率的影响是一项非常重要的研究课题,它从风险波动的视角揭示了货币经济与实体经济的关联性,因此具有很强的理论参考价值。另外对该问题的深入研究还具有很好的现实指导意义,一方面广大的金融市场投资者需要基于宏观经济形势分析金融市场波动率进而完善投资组合策略;另一方面监管当局需要依据宏观经济周期的变化情况制定股票市场波动的周期性及宏观审慎性监管措施,以避免股市出现大起大落。

本文基于宏观经济变量等信息构建出一种多因子混频波动率模型,并探讨了其在中国股市波动率样本内拟合与样本外预测方面的表现。根据前面的分析,本文获得的主要结论可概述为三个方面:其一,各类 GARCH-MIDAS 模型均可识别中国股市波动率的长期成分,估计结果优于基本的 GRACH 模型,说明混频波动率模型在刻画中国

^① 目标预测模型不同,相应的非目标预测模型也不相同。如目标预测模型为 G-M-Nor-Range 时,对应的非目标预测模型为 G-M-Ra-Con-L、G-M-Ra-Lead-L、G-M-Ra-Con-V、G-M-Ra-Lead-V、G-M-Ra-Con-L-V、G-M-Ra-Lead-L-V 这 6 个多因子混频模型;当目标预测模型为 G-M-Con-L 时,非目标预测模型为 G-M-Ra-Con-L、G-M-Ra-Con-L-V 这两个多因子混频模型。其他情况可依此类推。

股市波动率方面的适用性,尤为重要的是多因子 GARCH-MIDAS 模型设定优于相应单因子混频波动率模型并且可以更好地捕捉股市波动的长期成分;其二,包含宏观变量的 GARCH-MIDAS 模型能够说明中国宏观经济波动对股市波动具有正向影响,宏观经济会从生产层面对股市波动产生影响;其三,与单因子混频波动率模型相比,多因子混频波动率模型能够更好地提高波动率的预测精度,SPA 检验结果表明该预测结果的稳健性,说明了包含宏观变量的多因子混频模型在提高中国股市波动率预测精度方面的贡献。

基于宏观经济变量等信息构建的多因子 GARCH-MIDAS 模型可以同时测度“宏观”因子和“金融”因子对高频波动率长期成分的作用。该模型不仅改进了混频波动率模型的样本内拟合效果,而且在波动率的预测方面也具有良好表现,说明宏观基本面对股市波动及其预测方面的重要影响。这些结论有助于准确把握中国股市波动率的运行特征及其变化趋势,从而为监管当局制定股票市场波动的周期性监管措施及为投资者完善投资组合策略提供了重要的经验支持。本文着重考察了宏观基本面对股市波动率及其预测方面的影响,如何利用混频数据信息研究宏观基本面与股市波动率的相互作用机制是我们需要进一步考虑的问题。

参考文献:

陈其安、张媛、刘星(2010):《宏观经济环境、政府调控政策与股票市场波动性——来自中国股票市场的经验证据》,《经济学家》第2期。

陈强、叶阿忠(2009):《股市收益、收益波动与中国城镇居民消费行为》,《经济学(季刊)》第3期。

顾巧明(2011):《我国股市波动的长记忆性与货币政策的非对称性研究》,《中国经济问题》第1期。

康书隆、王志强(2010):《中国国债利率期限结构的风险特征及其内含信息研究》,《世界经济》第7期。

李建军(2008):《中国未观测货币金融状况指数与经济景气指数——理论设计与内在关系的实证研究》,《世界经济》第7期。

刘玉军(2011):《我国证券市场与宏观经济波动关系的实证研究》,《社会科学战线》第10期。

刘汉、刘金全(2011):《中国宏观经济总量的实时预报与短期预测——基于混频数据预测模型的实证研究》,《经济研究》第3期。

刘向华、柳恩普(2013):《股票市场流动性与宏观经济的影响机制研究》,《华东经济管理》第8期。

彭伟(2013):《我国上市商业银行股票日收益风险价值研究——基于AR、HAR和MIDAS模型的分析》,《金融监管研究》第3期。

唐平、刘燕(2008):《基于宏观经济变量的中国股市波动分析》,《财经科学》第6期。

王博、伍楠林(2012):《我国股指期货收益与宏观经济关系的实证分析》,《国际贸易问题》第6期。

赵可、丁安华(2013):《经济周期与股市波动传导机制的概念模型构建》,《中南大学学报(社会科学版)》第2期。

赵振全、张宇(2003):《中国股票市场波动和宏观经济波动关系的实证分析》,《数量经济技术经济研究》第6期。

郑挺国、王霞(2013):《中国经济周期的混频数据测度及实时分析》,《经济研究》第6期。

Andersen, T. G.; Bollerslev, T. and Diebold, F. X. "Roughing It up: Including Jump Components in Measuring, Modeling and Forecasting Asset Return Volatility." *Review of Economics and Statistics*, 2007, 89(4), pp.701-720.

Bansal, R. and Yaron, A. "Risks For the Long-Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance*, 2004, 59, pp.1481-1509.

Barro, R. J. "Rare Disasters and Asset Markets in The Twentieth Century." *Quarterly Journal of Economics*, 2006, 121, pp.823-866.

Beltratti, A. and Morana, C. "Breaks and Persistency: Macroeconomic Causes of Stock Market Volatility." *Journal of Econometrics*, 2006, 131(1), pp.151-177.

Bollerslev, T. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, 1986, 31, pp.307-327.

Breeden, D. T. "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities." *Journal of Financial Economics*, 1979, 7(3), pp.265-296.

Campbell, J. and Shiller, R. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies*, 1988, 1, pp.195-228.

Campbell, J. "A Variance Decomposition for Stock Returns." *Economic Journal*, 1991, 101, pp.157-179.

Casarin, R. and Trecroci, C. "Business Cycle and Stock Market Volatilities: Are They Related?" SSRN Working paper No.888524, 2007.

Chiou, W. and Song, Z. "The Macroeconomic Determinants of Stock Price Volatility: Evidence from Taiwan, South Korea, Singapore and Hong Kong." *Journal of Environment and Management*, 2007, 8(2), pp.114-134.

Chordia, T.; Roll, R. and Subrahmanyam, A. "Order Imbalance, Liquidity, and Market Returns." *Journal of Financial economics*, 2002, 65(1), pp.111-130.

Ding, Z. and Granger, C. "Modeling Volatility Persistence of Speculative Returns: A New Approach." *Journal of Econometrics*, 1996, 73, pp.185-215.

Engle, R. F. "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U. K. Inflation." *Econometrica*, 1982, 50, pp.987-1008.

Engle, R. F.; Ghysels, E. and Sohn, B. "Stock Market Volatility and Macroeconomic Fundamentals." New York University working paper, 2009.

Engle, R. F. and Lee, G. "A Permanent and Transitory Component Model of Stock Return Volatility," in R. F. Engle and H. White ed., *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*. London: Oxford University Press, 1999, pp.475-497.

Engle, R. F. and Rangel, J. G. "The Spline-GARCH Model for Low-Frequency Volatility and Its Global Macroeconomic Causes." *The Review of Financial Studies*, 2008, 21, pp.1187-1222.

- Errunza, V. and Hogan, K. "Macroeconomic Determinants of European Stock Market Volatility." *European Financial Management*, 1998, 4(3), pp.361-377.
- Ghysels, E.; Santa-Clara, P. and Valkanov, R. "The MIDAS Touch: Mixed Data Sampling Regressions." Mimeo, Chapel Hill, N. C. 2004.
- Ghysels, E.; Arthur, S. and Rossen, V. "MIDAS Regressions: Further Results and New Directions." *Econometric Reviews*, 2007, 26(1), pp.53-90.
- Ghysels, E. "Macroeconomics and the Reality of Mixed Frequency Data." Manuscript, University of North Carolina at Chapel Hill working paper, 2012.
- Hamilton, J. D. and Lin, G. "Stock Market Volatility and the Business Cycle." *Journal of Applied Econometrics*, 1996, 11(5), pp.573-593.
- Harvey, A.; Ruiz, E. and Shephard, N. "Multivariate Stochastic Variance Models." *The Review of Economic Studies*, 1994, 61(2), pp.247-264.
- Hansen, P. R. "A Test for Superior Predictive Ability." *Journal of Business and Economic Statistics*, 2005, 23(4), pp.365-380.
- Kim, Y. and Nelson, C. R. "Pricing Stock Market Volatility: Does It Matter Whether the Volatility Is Related to the Business Cycle?" *Journal of Financial Econometrics*, 2014, 12(2), pp.307-328.
- Lucas, R. E. "Asset Prices in an Exchange Economy." *Econometrica*, 1978, 46(6), pp.1429-1445.
- Martens, M. and Dijk, V. D. "Measuring Volatility with the Realized Range." *Journal of Econometrics*, 2007, 138(1), pp.181-207.
- Morelli, D. "The Relationship between Conditional Stock Market Volatility and Conditional Macroeconomic Volatility: Empirical Evidence Based on UK Data." *International Review of Financial Analysis*, 2002, 11(1), pp.101-110.
- Parkinson, M. "The Extreme Value Method for Estimating the Variance of the Rate of Return." *Journal of Business*, 1980, 53, pp.61-65.
- Schwert, G. W. "Why Does Stock Market Volatility Change Over Time?" *The Journal of Finance*, 1989, 44(5), pp.1115-1153.
- Shyu, Y. W. and Hsia, K. "Does Stock Market Volatility with Regime Shifts Signal the Business Cycle in Taiwan?" *International Journal of Electronic Finance*, 2008, 2(4), pp.433-450.
- Wachter, J. "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *Journal of Finance*, 2013, 68(3), pp.987-1035.
- Wang, X. "The Relationship between Stock Market Volatility and Macroeconomic Volatility: Evidence from China." *Journal of Chinese Economics and Finance*, 2011, 2, pp.67-77.
- Wongbangpo, P. and Sharma, S. C. "Stock Market and Macroeconomic Fundamental Dynamic Interactions: ASEAN-5 Countries." *Journal of Asian Economics*, 2002, 13(1), pp.27-51.

附录:SPA 方法检验统计量及其 p 值的 Bootstrap 计算步骤:

本文第五部分采用 Hansen(2005)提出的模型预测能力检验方法。该方法假设存在损失函数 $L(\xi_t, \delta_{k,t-h})$, 其中 ξ_t 是随机变量, $\delta_{k,t-h}$ 代表第 k 个预测模型(预测结果), $k=0$ 表示目标预测模型, 相对损失变量则定义如下: $d_{k,t} = L(\xi_t, \delta_{0,t-h}) - L(\xi_t, \delta_{k,t-h})$ 。令 $d_t = (d_{1,t}, \dots, d_{m,t})'$ 并设定 $\mu \equiv E(d_t)$, $\mu \in R^m$ 。原假设为 $H_0: \mu \leq 0$ 。为了使得 μ 可以很好的定义并保证 Bootstrap 技术的使用, Hansen 进一步假设并得到如下表达式: $n^{1/2}(\bar{d} - \mu) \xrightarrow{d} N_m(0, \Omega)$ 其中 $\bar{d} \equiv n^{-1} \sum_{t=1}^n d_t$, Ω 为渐进协方差矩阵。在此基础上构造如下检验统计量 (studentized test statistic): $T_n^{SPA} \equiv \max_{k=1, \dots, m} [\max_{t=1, \dots, m} n^{1/2}(\hat{\omega}_k^2)^{-1/2} \bar{d}_k, 0]$, 其中, $\hat{\omega}_k^2 k^2$ 是 $\omega_k^2 \equiv \text{Var}(n^{1/2} \bar{d}_k)$ 的某个一致估计量; 同时原假设分布基于 $N_m(\hat{\mu}^c, \hat{\Omega})$, 其中, $\hat{\mu}^c = \bar{d}_k \cdot 1 \{ n^{1/2} k^{-1} \bar{d}_k \leq -\sqrt{2 \log(\log n)} \}, 1 \{ \cdot \}$ 表示指示函数。

Bootstrap 计算步骤如下:

首先, 从原始数据中产生伪时间序列 (pseudo time series) $\{d_{b,t}^*\} \equiv \{d_{\tau_{b,t}}\}$, $b=1, \dots, B$, 其中 B 表示 Bootstrap 次数, $\{\tau_{b,1}, \dots, \tau_{b,n}\}$ 表示基于 $\{1, \dots, n\}$ 的随机长度抽样。我们随机产生两个 $B \times n$ 矩阵 U 和 V , 其中 $u_{b,t}$ 和 $v_{b,t}$ 相互独立且服从 $(0, 1]$ 上的均匀分布。每次抽样第 1 个元素 $\tau_{b,1} = [nu_{b,1}]$, 其中 $[\cdot]$ 表示取整, 当 $t=2, \dots, n$ 时, 相应元素为:

$$\tau_{b,t} = \begin{cases} [nu_{b,1}] & \text{如果 } v_{b,t} < q \\ 1 \{ \tau_{b,t-1} < n \} \cdot \tau_{b,t-1} + 1 & \text{如果 } v_{b,t} \geq q \end{cases}$$

其中, 参数 $q \in (0, 1]$, Hansen(2005)在实证中设为 0.25。

其次, 通过伪时间序列计算样本均值 $\bar{d}_{\tau_{b,t}}^* \equiv n^{-1} \sum_{t=1}^n d_{\tau_{b,t}}^*$, $b=1, \dots, B$, 计算检验统计量 T_n^{SPA} , 这里对 ω_k^2 的估计量表达式为: $\hat{\omega}_k^2 = \hat{\gamma}_{0,k} + 2 \sum_{i=1}^n \kappa(n, i) \hat{\gamma}_{i,k}$ 其中 $\hat{\gamma}_{i,k} = n^{-1} \sum_{j=1}^{n-i} (d_{k,j} - \bar{d}_k)(d_{k,j+i} - \bar{d}_k)$, $\kappa(n, i) = (n-i)(1-q)^i/n + i(1-q)^{n-i}/n$ 。

最后, 为了得到统计量在原假设下的分布, 定义 $Z_{k,b,t}^* = d_{k,b,t}^* - g_c(\bar{d}_k)$, $b=1, \dots, B, t=1, \dots, n$ 。其中 $g_c(x) = x \cdot 1[x \geq -\sqrt{(k^2/n)2 \log(\log n)}]$, 我们通过 Bootstrap 抽样得到 $Z_{b,t}^*$ 的经验分布并计算出 $T_{b,n}^{SPA*} \equiv \max_{k=1, \dots, m} (n^{-1} \bar{Z}_{k,b}^* / \hat{\omega}_k)$, 最终使用 Bootstrap 方法得到 SPA 的 p 值为: $\hat{p}_{SPA} \equiv \sum_{b=1}^B 1 \{ T_{b,n}^{SPA*} > T_n^{SPA} \} / B$ 。

(截稿:2014 年 6 月 责任编辑:宋志刚)