

# 基于 MIDAS-Expectile 回归模型的加密货币风险测度

张志远<sup>1</sup>, 叶五一<sup>1,2</sup>

(1. 中国科学技术大学管理学院, 安徽合肥 230026; 2. 中国科学技术大学国际金融研究院, 安徽合肥 230601)

**摘要:** 风险测度 EVaR(以 Expectile 模型为基础)作为 QVaR(以分位数为基础)的替代技术,其计算更加简便,且能够更加准确地反映极端值的影响.为了充分综合利用不同频率数据所包含的信息,构建了 MIDAS-Expectile 回归模型,并基于非线性非对称最小二乘方法得到参数及条件 EVaR 的估计,同时给出了估计的渐近正态性以及条件 Expectile 的 coverage 检验.此外,还从极大似然估计的角度给出了 Expectile 回归模型的似然函数及信息准则,以完成不同 Expectile 回归模型的比较与检验.为了对加密货币的金融风险进行研究,在实证部分,将 MIDAS-Expectile 回归模型应用于加密货币收益风险的度量,同时探讨了其他传统金融市场对这一新兴金融资产的风险传染现象.加密货币月度数据的风险实证结果表明其他金融市场的信号将对加密货币市场风险有显著的或正向或负向的影响,加密货币市场不是孤立于传统金融市场.

**关键词:** MIDAS-Expectile 回归模型; EVaR; 加密货币; 非线性非对称最小二乘; 极大似然

**中图分类号:** F821; F830.9 **文献标识码:** A **doi:** 10.3969/j.issn.0253-2778.2020.06.020

**引用格式:** 张志远,叶五一. 基于 MIDAS-Expectile 回归模型的加密货币风险测度[J]. 中国科学技术大学学报, 2020, 50(6): 860-872.

ZHANG Zhiyuan, YE Wuyi. Cryptocurrency risk measurement based on MIDAS-Expectile regression model[J]. Journal of University of Science and Technology of China, 2020, 50(6): 860-872.

## Cryptocurrency risk measurement based on MIDAS-Expectile regression model

ZHANG Zhiyuan<sup>1</sup>, YE Wuyi<sup>1,2</sup>

(1. School of Management, University of Science and Technology of China, Hefei 230026, China;

2. International Institute of Finance, University of Science and Technology of China, Hefei 230601, China)

**Abstract:** As an alternative to the quantile-based QVaR, the risk measure EVaR based on the Expectile model is simpler to calculate and can more accurately reflect the effects of extreme values. In order to make full use of the information contained in mixed frequency data, a MIDAS-Expectile regression model was constructed, and the estimation of the parameters and conditional EVaR were obtained based on the nonlinear asymmetric least squares method. The asymptotic normality of the estimates and coverage test for conditional Expectile were also given. In addition, the likelihood function and information criterion of the Expectile regression model were given from the perspective of maximum likelihood estimation, which could compare and test different models. In order to study the financial risks of cryptocurrencies, in the empirical part, the MIDAS-Expectile regression model was applied to the measurement of cryptocurrency returns risk, and the risk contagion of other tradition financial markets to this emerging financial asset was discussed. The empirical results of the risk of cryptocurrency monthly data indicate that signals from other financial markets will have a significant or positive or negative impact on the risks of the cryptocurrency market, and that the cryptocurrency market is not isolated from traditional financial markets.

**Key words:** MIDAS-Expectile regression model; EVaR; cryptocurrency; nonlinear asymmetric least squares; maximum likelihood

## 0 引言

良好的风险度量工具对于风险度量与管理极

为重要,是风险度量与管理的基础.众多学者都在这个研究领域做出了很多贡献.风险度量工具各有特点,对金融数据的某一特征具有针对性,但很难

收稿日期: 2020-04-09; 修回日期: 2020-05-20

基金项目: 国家自然科学基金面上项目(71973133, 71671171), 国家自然科学基金重点项目(71631006)资助.

作者简介: 张志远,男,1994年生,硕士.研究方向:风险管理. E-mail: zhzhy1@mail.ustc.edu.cn

通讯作者: 叶五一,博士/副教授. E-mail: wyye@ustc.edu.cn

做到尽善尽美。尽管常用的风险度量方法在险价值 (value at risk, VaR) 已经经历了几十年的研究发展,但是其仍然存在不足,因而一些研究者将目光转向了新的风险度量工具。Newey 和 Powell<sup>[1]</sup>提出的非对称最小二乘 (asymmetric least square, ALS) 方法,将最小二乘方法中的损失函数修改为非对称的二次形式,即对正负两个部分分别以  $\tau$  ( $0 < \tau < 1$ ) 和  $1 - \tau$  的权重进行加权,可以通过回归得到被解释变量的 Expectile 的估计。非对称最小二乘方法自提出以后得到了越来越多研究者的关注,原因在于 Expectile 计算更简单,对极端值的反映也更加敏感,并且可以和分位数相互转化,是一种很好的风险测度。

度量风险时经常使用的 VaR 是尾部分位数, Kuan 等<sup>[2]</sup>将它称作 QVaR,同时把尾部 Expectile 作为一种风险度量称为 EVaR,构建了条件自回归 Expectile 模型 (CARE) 来研究在平稳以及弱相依金融时间序列中的风险度量。为了处理数据中存在的异方差性,谢尚宇等<sup>[3]</sup>拓展 Kuan 等<sup>[2]</sup>的模型从而构建了 ARCH-Expectile 模型。Xie 等<sup>[4]</sup>将非参数变系数方法的统计技术引入到 Expectile 模型中以此发展出了变系数 Expectile 模型,并基于局部线性方法得到了系数的非参数估计。Kim 等<sup>[5]</sup>将非对称最小二乘应用到了非线性回归模型中,提出了非线性 Expectile 回归并给出了回归估计的统计渐近性质。变点检测的统计方法常被用于金融危机前后的机制变化研究,谭常春等<sup>[6]</sup>将这一思想和 Expectile 模型结合,使用 Expectile 变点检测方法研究了金融危机前后的金融风险变化。Daouia 等<sup>[7]</sup>指出分位数和 Expectile 可以包含在更加广泛的 M-quantiles 中并且 Expectile 是其中唯一的相合风险测度, M-quantiles 可以通过  $L^p$  ( $1 < p < 2$ ) 最优化来得到估计结果,该研究表明上述估计也具有渐近正态性。M-quantiles 这一更具包容性的框架是分位数和 Expectile 的折中,并且不会损失二者的优点。许启发等<sup>[8]</sup>通过 Expectile 回归来简化均值-ES 组合投资决策中的计算问题。

金融市场中,存在很多包含高、低频数据的混合频率数据,如何充分利用这些包含不同频率信息的数据值得研究。当在同一个框架下使用混合频率数据时常用的方法就是将高频数据等权重加总成为低频数据,这样数据的频率都变成相同。然而,诸多研究已经证实,金融时间序列数据的权重不再像传统统计中是等权重的。高频数据的结构特征隐含着更多信息,这些需要研究者去构建模型进行挖掘。

Ghysels 等<sup>[9]</sup>将日度数据加总成月度数据的权重设为某一权重函数,以此来体现每个高频数据的不同重要性,从而提出了 MIDAS (mixed-data sampling) 方法。很多研究者使用 MIDAS 方法来综合混合频率数据的信息,发现 MIDAS 方法可以取得良好的结果。Ghysels 等<sup>[10]</sup>基于 MIDAS 方法研究了股票市场收益率和收益率波动之间的关系,验证了 Merton<sup>[11]</sup>的 ICAPM 模型中市场收益与波动的正相关性。Ghysels 等<sup>[12]</sup>则基于 MIDAS 方法研

究了对波动的最佳预测。Ghysels 等<sup>[13]</sup>将 MIDAS 方法与分位数回归模型结合来得到条件分位数的估计,同时发现 MIDAS 方法能够过滤掉高频数据的噪声。Pettenuzzo 等<sup>[14]</sup>使用 MIDAS 方法对随机波动进行建模并且基于贝叶斯方法得到了参数的估计。Andreou<sup>[15]</sup>研究了高频数据等权重加总成为低频数据的标准最小二乘回归模型和使用 MIDAS 方法的非线性最小二乘回归模型,比较了两个模型估计结果的偏差和相对有效性,研究结果表明 MIDAS 方法更好。夏婷等<sup>[16]</sup>使用 GARCH-MIDAS 探究经济不确定性对中国股票市场波动率的作用。尚玉皇等<sup>[17]</sup>构建了一种混频利率模型 (BHK-MIDAS 模型),更好地对短期利率波动性进行了预测。Xu 等<sup>[18]</sup>结合 UMIDAS 和 SVQR 提出了一个新的 UMIDAS-SVQR 模型来实现分位数回归。

综合考虑 Expectile 回归模型和 MIDAS 方法的优势,本文试图结合这两种方法,从而构建 MIDAS-Expectile 回归模型。本文构建 MIDAS-Expectile 回归的目的是,基于混合频率数据得到更精确的 Expectile 回归参数的估计以及 EVaR 的估计,从而得到更优良的金融风险度量结果,以此帮助风险管理的使用者获得更准确的参考信息。本文提出的 MIDAS-Expectile 回归模型既有非对称最小二乘作为估计方法的优点又有 MIDAS 充分挖掘混合频率数据信息的优势,因此成为传统 Expectile 回归的进一步发展。MIDAS 方法中权重函数往往选取非线性参数的形式使得对它的估计需要使用非线性估计的数值计算技术,即本文构建的 MIDAS-Expectile 回归模型要用非线性非对称最小二乘 (nonlinear asymmetric least squares, NALS) 来进行估计。本文将使用高斯-牛顿迭代数值方法来求解损失函数中非线性最优化的问题,并且给出 NALS 估计的渐近正态性来方便构建假设检验。进一步本文完成了对条件 Expectile 估计效果的检验,基于似然比检验方法构造了一种 coverage 检验,该检验的思路与分位数的 coverage 检验类似。此外,参照 Aigner 等<sup>[19]</sup>的做法,本文给出了 Expectile 回归模型的似然函数及极大似然估计。为了方便比较不同的 Expectile 回归模型,本文也给出了 Expectile 回归模型信息准则:赤池信息准则 (Akaike information criterion, AIC) 和贝叶斯信息准则 (Bayesian information criterion, BIC)。

基于本文提出的 MIDAS-Expectile 回归模型的优势,将该模型应用于相关金融风险问题的研究可以得到更优良的结果。因此,本文将该模型用于加密货币市场的金融风险的研究。近期人们高度关注区块链技术的发展和应用。2016 年 12 月我国发布的《“十三五”国家信息化规划》将区块链技术列为“十三五”期间的“重大任务和重点工程”。区块链技术已然上升成为国家战略,它将是全球争夺的关键技术,有着服务实体经济、改善社会民生的巨大能量。故而对加密货币市场相应的金融风险测度问题进行深入探究很有实际意义,可以为政策制定者提供参考。

加密货币作为一种货币去中心化的技术工具,一开始影响范围很小.随着科技技术的快速发展与普及,加密货币作为新奇的技术产物成为了热门投资产品.这种新鲜事物引发的投机热潮往往会形成金融泡沫,带来金融风险.加密货币的价格经历几次大涨大跌,对它的质疑声也一直不断,这也引起了监管部门的注意,一些监管政策相继推出.加密货币已经成为金融市场上的一大类资产,越来越多的金融研究者开始关注于这方面的研究,谢平等<sup>[20]</sup>对加密货币的相关文献进行了综述.其中一个主要的研究领域是加密货币的定价问题,如 Buchholz 等<sup>[21]</sup>使用黄金标准模型分析加密货币价格剧烈波动原因;van Wijk<sup>[22]</sup>发现股票指数、汇率和石油价格对加密货币价格有显著性影响;而 Kristoufek<sup>[23]</sup>则认为常用的货币价格供需理论不再适用于加密货币,投机是加密货币价格波动的主要原因.另外一个方面便是加密货币的金融风险研究,Walther 和 Klein<sup>[24]</sup>借助 GARCH-MIDAS 模型研究了外部变量对加密货币波动率预测的效果,Fry 等<sup>[25]</sup>建立了经济泡沫与崩溃的模型并将其应用于加密货币市场.此外,Urquhart 等<sup>[26]</sup>的研究发现在市场动荡时加密货币具有避险作用,Li 等<sup>[27]</sup>研究了技术和经济因素对比特币兑换率的影响.由于加密货币价格的剧烈波动及其泡沫破灭可能带来的巨大灾难,对于研究者和加密货币投资者来说其金融风险的测度是一个非常重要的关注点,然而这方面的研究还相对较少.本文将应用提出的 MIDAS-Expectile 回归模型研究外部变量对加密货币收益 EVaR 的影响,其中选取的外部变量是一些传统金融资产市场,这将有助于探索传统金融市场对加密货币市场风险的传染和影响.

## 1 模型原理

### 1.1 Expectile 回归模型

假设从如下的线性回归模型中产生一组样本观测值  $\{(y_t, x_t), t=1, 2, \dots, T\}$ :

$$y_t = x_t^T \beta + u_t \quad (1)$$

式中,  $x_t = (1, x_{t,1}, \dots, x_{t,p})^T$ , 含截距项的系数向量  $\beta = (\beta_0, \beta_1, \dots, \beta_p)^T$ ,  $u_t$  是误差项.当取不同的损失函数时,式(1)将变为不同的回归模型,如最常见的线性均值回归模型以及分位数回归模型<sup>[28]</sup>,回归系数  $\beta$  的估计能够由最小化特定的损失函数获得.可以给出线性均值回归即最小二乘法的损失函数为

$$L(\beta) = \sum_{t=1}^T (y_t - x_t^T \beta)^2$$

此外可以将线性分位数回归的损失函数表示为

$$L(\beta; \alpha) = \sum_{t=1}^T \rho_\alpha(y_t - x_t^T \beta) \quad (2)$$

式中,  $\alpha$  取自  $(0, 1)$ ,  $\rho_\alpha(\cdot)$  是凸函数,形式为

$$\rho_\alpha(\lambda) = |\alpha - I(\lambda < 0)| \cdot |\lambda|,$$

其中,  $I(A)$  是事件  $A$  的示性函数.

值得注意的是,分位数回归的损失函数在零处不可导,最小化式(2)来进行参数估计时存在困难.

文献[1]介绍了 Expectile 回归模型以此来得到一种跟分位数回归相似但是更为简便计算的度量风险的方法,并基于 ALS 对参数进行估计,具体为最小化如下的损失函数:

$$L(\beta; \tau) = \sum_{t=1}^T \eta_\tau(y_t - x_t^T \beta) \quad (3)$$

式中,  $0 < \tau < 1$ ,不同于线性均值和分位数回归的是  $\eta_\tau(\cdot)$  为非对称平方损失函数,具体表达式为  $\eta_\tau(\lambda) = |\tau - I(\lambda < 0)| \cdot \lambda^2$ .那么可以容易得到最小化式(3)的一阶条件如下所示:

$$\sum_{t=1}^T |\tau - I(y_t - x_t^T \beta < 0)| (y_t - x_t^T \beta) x_t = 0,$$

进而,基于 ALS 方法的参数  $\beta$  的估计式为

$$\hat{\beta}(\tau) = \left( \sum_{t=1}^T |\tau - I(y_t - x_t^T \hat{\beta}(\tau) < 0)| x_t x_t^T \right)^{-1} \cdot \left( \sum_{t=1}^T |\tau - I(y_t - x_t^T \hat{\beta}(\tau) < 0)| x_t y_t \right) \quad (4)$$

对于从某一分布产生的数据,如果满足文献[1]设定的某些条件,Expectile 回归的参数  $\beta(\tau)$  的 ALS 估计是唯一的,且满足相合性以及渐近正态性等良好性质.文献[2]把 ALS 拓展到了平稳和弱相依的金融数据的同时也给出了参数估计的统计渐近性.对于这些 ALS 估计的统计性质的更为详细的介绍和证明请参考文献[1,2].从形式上看,ALS 相当于权重为  $|\tau - I(y_t - x_t^T \beta < 0)|$  的加权最小二乘,那自然地文献[1,2]的结论可以被视为加权最小二乘估计的推广性质.基于 Expectile 回归的 ALS 估计的渐近正态性,可以容易地进行参数的假设检验和置信区间估计.

与分位数回归相似的,得到参数估计后自然有

$$\hat{\mu}(\tau) = x^T \hat{\beta}(\tau).$$

文献[1]将  $\mu(\tau)$  称作  $Y$  的  $\tau$  水平 Expectile.  $\mu(\tau)$  是  $\tau$  的严格单调增函数,且对  $\tilde{Y} = aY + b$ ,  $a > 0$ , 有  $\mu_{\tilde{Y}}(\tau) = a\mu_Y(\tau) + b$ .上述讨论表明 Expectile 是一种良好的风险测度工具,因而文献[2]依据 Expectile 给出了新的风险测度:  $EVaR = |\mu(\tau)|$ ,那么基于 Expectile 回归模型就便可以得到 EVaR 的估计.

### 1.2 MIDAS-Expectile 回归模型

常用的回归模型中因变量  $y_t$  和自变量  $x_t$  往往是以同一时间频率抽取得到的数据,例如季度或是月度数据,但一些情况下某些变量的数据可以更频繁地获取,比如日数据.假设对于取样时间点  $t = 1, 2, \dots, T$ , 因变量  $y_t$  来自低频数据而自变量  $z_t$  来自高频数据,且在  $t-1$  到  $t$  期间  $y_t$  只有一个观测数据,  $z_t = (z_{t-1/n}^H, z_{t-2/n}^H, \dots, z_{t-n/n}^H)^T$  可以获取  $n$  个观测数据.例如因变量  $y_t$  来自季度数据而自变量  $z_t$  来自月度数据,那么  $n=3$ .为了构建这种不同频率数据混合的回归模型,文献[9]引入了新的时间序列统计技术: MIDAS.

传统的 Expectile 回归模型中解释变量不具有 MIDAS 结构,这会损失高频数据中的信息.为此本文考虑当解释变量具有 MIDAS 结构时,如何得到

更好的条件 Expectile 估计. 先考虑解释变量是一维的情况, 而高维情况的处理是相似的. 那么, 构建 MIDAS-Expectile 回归模型为

$$y_t = \beta_0 + \beta_1 x_t + u_t \quad (5)$$

式中, 解释变量  $x_t$  由高频数据  $z_t$  以某种权重结构加权得到,

$$x_t = \sum_{d=1}^D B(d; \theta) z_t^{\frac{H-d}{n}}$$

$n$  是混合数据的频率倍数,  $D$  是滞后阶数. 权重函数  $B(d; \theta)$  是由  $q$  维参数向量  $\theta = (\theta_1, \theta_2, \dots, \theta_q)^T$  生成的特定形式函数. 为了能识别参数  $\beta_1$ , 通常有约束条件使权重的和为 1:

$$\sum_{d=1}^D B(d; \theta) = 1 \quad (6)$$

当然, 在 MIDAS 中  $\beta_1$  的识别不是必需的, 但实际应用中为了能得到高频数据的整体对低频数据的影响, 即  $x_t$  和  $y_t$  的关系,  $\beta_1$  的识别很重要.

对于 MIDAS 方法, 权重函数  $B(d; \theta)$  是关键, 好的权重函数能灵活便捷地刻画  $x_t$  内高频数据  $z_t$  的权重结构. 关于权重函数的选取已经有很多文献做过研究. 文献[10, 29]使用了指数 Almon 多项式函数, 文献[30]则使用了 Beta 函数作为权函数.

本文将使用指数 Almon 多项式函数, 即

$$B(d; \theta) = \frac{\exp(\sum_{i=1}^q \theta_i d^i)}{\sum_{d=1}^D \exp(\sum_{i=1}^q \theta_i d^i)} \quad (7)$$

设定该权重函数的参数是二维的,  $\theta = (\theta_1, \theta_2)^T$ . 选用指数 Almon 多项式函数的原因在于它是最常用的且具备诸多优势. 指数 Almon 多项式函数能保证权重始终为正, 并且总能满足识别条件式(6). 另外这是一种灵活的权重函数, 控制参数可以得到各种形状的权重. 特别是当参数都为 0 时, 我们得到的权重是相等的, 即  $B(d; \theta) = 1/D$ . 因此可以看出, MIDAS 方法包含了在处理混合频率数据时一般常用的将高频数据等权重加总成为低频数据从而使数据处于同一频率的做法, 但这种简单的做法会损失高频数据包含的信息, 通过使用 MIDAS 方法可以避免这一问题.

令  $y = (y_1, y_2, \dots, y_T)^T, Z = (z_1, z_2, \dots, z_T)^T, B(\theta) = (B(1; \theta), B(2; \theta), \dots, B(D; \theta))^T, u = (u_1, u_2, \dots, u_T)^T, \alpha = \beta_0 e_T, e_T$  为  $T$  维的元素都是 1 的列向量, 那么式(5)可以写成

$$y = \alpha + \beta_1 ZB(\theta) + u \quad (8)$$

和之前类似的基于 ALS 对式(8)中的参数进行估计, 则可以得到被解释变量的条件 Expectile 的估计. MIDAS-Expectile 回归的损失函数如下:

$$L(\beta_0, \beta_1, \theta; \tau) = (y - \alpha - \beta_1 ZB(\theta))^T W(\tau) (y - \alpha - \beta_1 ZB(\theta)) \quad (9)$$

式中,  $W(\tau) = \text{diag}(w_1(\tau), w_2(\tau), \dots, w_T(\tau)), w_i(\tau) = |\tau - I(y_i - \beta_0 - \beta_1 z_i^T B(\theta) < 0)|$ . 由于权重函数一般不是关于参数  $\theta$  的线性结构, 对损失函

数式(9)直接进行最小化存在一定难度, 可以使用数值方法来解决这一非线性最优化问题. 本文将使用高斯-牛顿迭代法来解决最小化式(9)的问题, 也可以称作非线性非对称最小二乘 (nonlinear asymmetric least squares, NALS) 估计. NALS 具体的估计过程如下:

令  $S(\beta_1, \theta) = \beta_1 ZB(\theta)$ , 选取适当的初始点  $(\beta_{1(0)}, \theta_{(0)})$ , 将  $S(\beta_1, \theta)$  在该点附近进行泰勒展开, 只取一阶部分:

$$\begin{aligned} S(\beta_1, \theta) &\approx S(\beta_{1(0)}, \theta_{(0)}) + \frac{\partial S(\beta_{1(0)}, \theta_{(0)})}{\partial \beta_1} (\beta_1 - \beta_{1(0)}) + \left( \frac{\partial S(\beta_{1(0)}, \theta_{(0)})}{\partial \theta} \right)^T (\theta - \theta_{(0)}) - \left( \frac{\partial S(\beta_{1(0)}, \theta_{(0)})}{\partial \theta} \right)^T \theta_{(0)} + \frac{\partial S(\beta_{1(0)}, \theta_{(0)})}{\partial \beta_1} \beta_1 + \left( \frac{\partial S(\beta_{1(0)}, \theta_{(0)})}{\partial \theta} \right)^T \theta \quad (10) \end{aligned}$$

式中,

$$\begin{aligned} \frac{\partial S(\beta_{1(0)}, \theta_{(0)})}{\partial \beta_1} &= ZB(\theta_{(0)}), \\ \left( \frac{\partial S(\beta_{1(0)}, \theta_{(0)})}{\partial \theta} \right)^T &= \beta_{1(0)} Z \left( \frac{\partial B(\theta_{(0)})}{\partial \theta_1}, \dots, \frac{\partial B(\theta_{(0)})}{\partial \theta_q} \right), \\ \frac{\partial B(\theta_{(0)})}{\partial \theta_i} &= \begin{bmatrix} \frac{\partial B(1, \theta_{(0)})}{\partial \theta_i} \\ \vdots \\ \frac{\partial B(D, \theta_{(0)})}{\partial \theta_i} \end{bmatrix}, i = 1, \dots, q. \end{aligned}$$

因此, 给定初值  $\beta_{1(0)}$  和  $\theta_{(0)}$  的条件下, 式(10)是关于未知参数  $\beta_1$  和  $\theta$  的线性形式. 那么接着将式(10)代入式(9)中从而得到参数的 ALS 估计, 将该估计值作为第一轮参数估计值, 记为  $\beta_{0(1)}, \beta_{1(1)}$  和  $\theta_{(1)}$ . 类似地接着在第一轮的 ALS 估计值的附件进行新的泰勒展开, 重复迭代直到估计值收敛, 最后便可以得到 MIDAS-Expectile 回归模型的 NALS 估计.

此外, 根据文献[5], 本文得到的 NALS 参数估计也符合相合性和渐进正态性. 由文献[5]的结论可以得到参数估计的渐近正态分布中的方差, 并可以由此构建参数估计的假设检验与置信区间.

### 1.3 似然方法

在解决某些问题时使用极大似然方法会非常方便, 为此需要首先给出 Expectile 回归模型相应的似然函数. 面对正、负随机波动项, 分别依据不同权重从正态分布中产生的这类问题, 文献[19]使用极大似然估计来进行研究, 其得到的极大似然估计结果和 ALS 是相同的. 这也表明 ALS 在处理正、负随机扰动有不同权重的正态分布数据时是合理有效的. 考虑到金融市场上人们对损失和收益有不同的态度, 通常来说市场组成以风险厌恶者为主, 因而市场对损失更为敏感. 金融数据往往也呈现出分布不对称的情况, 因此 Expectile 回归模型能够很好地描述实际的市场情况.

考虑如下模型的似然问题

$$y_t = f(x_t, \gamma) + u_t, t = 1, \dots, T \quad (11)$$

式中,  $f(\cdot)$  为任意函数, 误差项来自混合正态分布, 取值为

$$u_t = \begin{cases} \frac{u_t^*}{\sqrt{1-\tau}}, u_t \leq 0; \\ \frac{u_t^*}{\sqrt{\tau}}, u_t > 0. \end{cases}$$

其中,  $u_t^* \sim N(0, \sigma^2)$ . 那么误差项的密度函数可以表示成

$$f(u) = I(u \leq 0)\varphi(0, \sigma^2/1-\tau) + I(u > 0)\varphi(0, \sigma^2/\tau),$$

其中  $\varphi(\cdot)$  为正态分布密度函数. 容易得到对数似然函数为

$$\begin{aligned} \ln L = & -\frac{T}{2}\ln(2\pi) + \frac{T_1}{2}\ln\tau + \\ & \frac{T_2}{2}\ln(1-\tau) - \frac{T}{2}\ln(\sigma^2) - \\ & \frac{\tau}{2\sigma^2} \sum_{y_t - f(x_t, \gamma) > 0} (y_t - f(x_t, \gamma))^2 - \\ & \frac{1-\tau}{2\sigma^2} \sum_{y_t - f(x_t, \gamma) \leq 0} (y_t - f(x_t, \gamma))^2 \quad (12) \end{aligned}$$

式中,  $T_1$  代表满足  $y_t - f(x_t, \gamma) > 0$  的数据数目,  $T_2$  代表满足  $y_t - f(x_t, \gamma) \leq 0$  的数据数目. 那么接下来对于  $\sigma^2$  最大化式(12), 可得

$$\hat{\sigma}^2 = \frac{1}{T} \left[ \tau \sum_{y_t - f(x_t, \gamma) > 0} (y_t - f(x_t, \gamma))^2 + (1-\tau) \sum_{y_t - f(x_t, \gamma) \leq 0} (y_t - f(x_t, \gamma))^2 \right].$$

将其带入式(12), 得到只含参数  $\gamma$  的对数似然函数:

$$\ln L_1 = -\frac{T}{2} \ln \left[ \frac{1}{T} \sum_{t=1}^T \eta_\tau(y_t - f(x_t, \gamma)) \right] \quad (13)$$

显然, 由极大似然估计得到参数估计时最大化式(13)等价于最小化 Expectile 回归的损失函数.

基于极大似然估计给出 Expectile 回归模型的信息准则在比较不同模型效果时极为有用. 经常被研究者使用的信息准则有 AIC 与 BIC, 文献[31]给出的 AIC 形式为  $AIC = 2k - 2\ln L$  而文献[32]给出的 BIC 形式为  $BIC = \ln(n)k - 2\ln L$ , 其中  $k$  代表模型内独立参数数目,  $n$  是数据量大小. 那么自然地 Expectile 回归模型的 AIC、BIC 能够表示成

$$AIC = 2k + T \ln \left[ \frac{1}{T} \sum_{t=1}^T \eta_\tau(y_t - f(x_t, \gamma)) \right],$$

$$BIC = \ln(T)k + T \ln \left[ \frac{1}{T} \sum_{t=1}^T \eta_\tau(y_t - f(x_t, \gamma)) \right].$$

上述信息准则将被用于度量 Expectile 回归模型的相对信息损失, 损失信息越少, 模型质量越高. 这为比较不同设定下的模型提供了依据, 如在 MIDAS-Expectile 回归中可以使用信息准则确定最优的滞后阶数  $D$ .

#### 1.4 coverage 检验

当使用 NALS 得到 MIDAS-Expectile 回归模

型的参数估计后, 即可得到被解释变量的条件 Expectile, 即 EVaR 的估计. 此时一个关键的问题是如何检验条件 Expectile 估计的效果, 即检验得到的条件 Expectile 的估计是否接近真实值, 为此本文基于似然比方法构建一个 coverage 检验, 该检验与分位数的 coverage 检验有类似的思路.

文献[33]指出随机变量  $Y$  的分位数  $q_y(\alpha)$  和 Expectile  $\mu_y(\tau)$  之间存在联系, 即定义  $\tau(\alpha)$  使  $\mu_y(\tau(\alpha)) = q_y(\alpha)$ , 那么

$$\tau(\alpha) = \frac{\alpha q_y(\alpha) - \int_{-\infty}^{q_y(\alpha)} y dF(y)}{EY - 2 \int_{-\infty}^{q_y(\alpha)} y dF(y) - (1-2\alpha)q_y(\alpha)} \quad (14)$$

由分位数的定义知  $P(Y \leq q_y(\alpha)) = \alpha$ , 即有  $P(Y \leq \mu_y(\tau(\alpha))) = \alpha$  成立. 因此在样本容量为  $T$  的  $y_t$  中, 若有  $V$  个数据是大于  $\mu_y(\tau(\alpha))$  的,

$$V = \sum_{t=1}^T I(y_t > \mu_y(\tau(\alpha))),$$

则  $V$  服从二项分布  $V \sim B(T, 1-\alpha)$ , 对数似然函数为

$$\ln L(\alpha) = V \ln(1-\alpha) + (T-V) \ln \alpha + \ln \left( \frac{T}{V} \right).$$

那么通过上式能够知道  $\alpha$  的极大似然估计形式是  $\hat{\alpha} = (T-V)/T$ , 因此构建似然比检验的原假设和对立假设如下:

$$H_0: \alpha = \alpha_0 \leftrightarrow H_1: \alpha \neq \alpha_0,$$

从而似然比统计量  $\lambda$  可以定义为

$$\lambda = -2 \ln \frac{L(\alpha_0)}{L(\hat{\alpha})} = -2 \ln \frac{(1-\alpha_0)^V \alpha_0^{T-V}}{(1-\hat{\alpha})^V \hat{\alpha}^{T-V}}.$$

由似然比检验的理论知拒绝域为  $\lambda > \chi_1^2(\theta)$ ,  $\theta$  为置信水平.

当要检验  $y_t$  的条件 Expectile 的估计是否接近真实值时先得到样本  $y_t$  中大于条件 Expectile 估计值的数目  $V$  后, 再基于上述似然比方法进行假设检验. 一般情况下估计条件 Expectile 时,  $\tau$  是事先给定的, 记为  $\tau_0$ , 那么需要基于式(14)得到  $\tau(\alpha) = \tau_0$  中相对应的  $\alpha_0$ . 当分布已知时,  $\alpha_0$  很容易得到, 在表 1 中总结了一些常见分布的结果. 但在实际中, 分布往往是未知的, 计算  $\alpha_0$  时需要式(14)进行近似处理:

$$\begin{aligned} EY &= \frac{1}{T} \sum_{t=1}^T y_t, \\ q_y(\alpha_0) &= \mu_y(\tau_0) = \frac{1}{T} \sum_{t=1}^T \hat{\mu}_y(\tau_0, x_t), \\ \int_{-\infty}^{q_y(\alpha_0)} y dF(y) &= E[y_t I(y_t < q_y(\alpha_0, x_t))] = \\ &= \frac{1}{T} \sum_{y_t < \hat{\mu}_y(\tau_0, x_t)} y_t. \end{aligned}$$

因此, 得到基于上述三个近似计算后, 当分布未知时, 将它们带入到式(14)便可以近似得到  $\tau_0$  水平下对应的  $\alpha_0$ .

表 1 常见分布下的  $\alpha_0$

Tab.1  $\alpha_0$  under common distribution

$\tau$ (%)	$U(-a, a)$	$N(0, 1)$	$t(3)$	$t(5)$	$t(10)$	$t(30)$
1	9.2	4.3	2.4	3.0	3.5	4.0
3	15.0	9.1	5.6	6.8	8.0	8.8
5	18.6	12.6	8.5	10.0	11.5	12.3
10	25.0	19.5	14.5	16.6	18.3	19.0
25	36.6	33.2	29.4	31.9	32.2	32.8

## 2 数据模拟

### 2.1 数据生成过程(DGP)

在给出 MIDAS-Expectile 回归模型的参数估计的 NALS 方法和渐近性质的基础上,为了更充分地了解 NALS 估计方法的效果,本章将使用蒙特卡罗模拟进行数据模拟研究. 本文将在模拟中设置多种模型设定,从而反映实际中复杂多变的金融数据,来充分了解本文提出的方法的实际效果.

以 MIDAS-Expectile 回归模型(5)为基础,首先来介绍样本数据的产生过程. 对于解释变量  $x_t$  的数据生成过程,与之前所提相同的本文决定设定权重函数  $B(d; \theta)$  的形式为指数 Almon 多项式函数. 考虑到金融实践中时序数据滞后期越长时对当前的影响就越小,本文构建两种衰减型权重函数: 快速衰减权重函数  $\theta_1 = 6 \times 10^{-3}, \theta_2 = -5 \times 10^{-2}$ ; 缓慢衰减权重函数  $\theta_1 = 6 \times 10^{-3}, \theta_2 = -5 \times 10^{-3}$ , 两种衰减权重函数趋势如图 1 所示. 频率倍数  $n = 10, 90$ , 样本容量  $T = 200, 500$ . 参考文献[29], 高频数据  $z_t = (z_{t-1/n}^H, \dots, z_{t-n/n}^H)^T$  分别产生于 AR(1), ARCH(1)两种不同的时序模型:

$$(i) z_{t-j/n}^H = 0.5 + 0.9z_{t-j-1/n}^H + \varepsilon_{t-j/n}, \varepsilon_{t-j/n} \sim i. i. d N(0, 1);$$

$$(ii) z_{t-j/n}^H = \sigma_{t-j/n} \cdot \varepsilon_{t-j/n}, (\sigma_{t-j/n})^2 = 0.25 + 0.85(z_{t-j-1/n}^H)^2, \varepsilon_{t-j/n} \sim i. i. d N(0, 1).$$

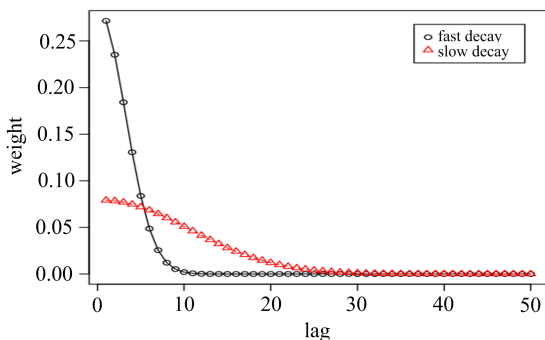


图 1 权重函数

Fig.1 The weight function

误差项  $u_t$  分别产生于两种不同的分布: 标准正态分布,  $t(5)$  分布. 设定截距项  $\beta_0 = 0.5$ , 斜率取两种:  $\beta_1 = 3, 5$ , 分别对应低信噪比、高信噪比. 本文使用 R 语言来进行数据模拟, 每个数据生成过程下模拟次数均为 1000 次. 得到模拟数据后, 基于 NALS 方法分别得到了  $\tau = 0.05, 0.1, 0.5, 0.95$  时的 MIDAS-Expectile 回归模型的估计, 并进行进一步分析.

### 2.2 模拟结果

处理高低频混合数据时,常用的做法是把高频数据等权重加总成为低频数据,然后利用 ALS 估计可以很容易地得到 Expectile 回归模型的参数估计. 但这和 MIDAS 方法,即本文提出的 MIDAS-Expectile 回归模型相比是有缺陷的,丧失了高频数据的额外信息. 本文在模拟中同时得到了在上述 DGP 下等权重加总方法和 MIDAS 方法在估计条件 Expectile 时的斜率  $\beta_1$  的估计. 比较这两种模型估计结果的有效性可以说明 MIDAS-Expectile 回归模型的优点,为此两种模型参数估计结果的均方误差:  $MSE_{ALS}$  和  $MSE_{NALS}$  先后在本文中被给出,并且取相对均方误差形式为  $re-MSE = MSE_{NALS}/MSE_{ALS}$ . 显然,如果  $re-MSE$  小于 1,说明 MIDAS 方法要比等权重加总方法在估计参数时更加有效.  $re-MSE$  越小, MIDAS 方法相对于等权重加总方法的有效性越高. 表 2 和表 3 分别给出了快速衰减和缓慢衰减权重函数下蒙特卡罗模拟得到的 MIDAS 方法估计条件 Expectile 时斜率  $\beta_1$  的估计的均方误差  $MSE_{NALS}$  和与等权重加总方法相比的相对均方误差  $re-MSE$ .

$MSE_{NALS}$  反映了使用 MIDAS 方法估计条件 Expectile 时斜率  $\beta_1$  的估计偏离真实值的误差,  $MSE_{NALS}$  越小说明误差越小,估计越准确有效. 从表 2 和表 3 可以看到在绝大多数 DGP 设定下,  $MSE_{NALS}$  都很小,这说明本文的方法,即在 Expectile 回归中引入 MIDAS 方法后得到 MIDAS-Expectile 回归模型的参数估计是十分有效的. 同时我们发现会有少数  $MSE_{NALS}$  较大(大于 1),它们都来自高频数据是由 ARCH(1)模型产生,误差项服从  $t(5)$  分布,频率倍数  $n$  较大的 DGP. 这样的 DGP 是所有设定中最复杂的,  $t$  分布是厚尾分布,更可能产生极端值;高频数据是高阶滞后的 ARCH(1)数据,具有异方差性. 经过检查模拟得到的  $\beta_1$  的估计数据,我们发现参数估计经过多次迭代,产生了个别离真实值较远的估计值,从而使  $MSE_{NALS}$  较大. 但这些极端的估计值只占 1000 次模拟结果的极少部分,排除它们的影响,估计依然很有效. 更重要的是,我们发现这些  $MSE_{NALS}$  较大的情况中,  $re-MSE$  都小于 1 (除了缓慢衰减权重函数低信噪比设定下有一个  $re-MSE$  是 1.0140,也很接近 1),说明即使在这些复杂设定下, MIDAS 方法也优于等权重加总方法.

另外,从表 2 可以看出在快速衰减权重函数下,  $re-MSE$  都小于 1,并且几乎都是偏离 1 很远,说明 MIDAS 方法和等权重加总方法相比,有效性特别高. 表 3 中  $re-MSE$  虽然没有表 2 中的那么小,但绝大多数都小于 1,并且最大值也不超过 1.25. 这表明在缓慢衰减权重函数下, MIDAS 方法至少和等权重加总方法一样有效,大多数情况下更加有效.

同时发现影响  $re-MSE$  的因素有权重函数、斜率、高频数据、误差项、样本容量  $T$ 、频率倍数  $n$ . 快速衰减权重函数比缓慢衰减权重函数  $re-MSE$  更小,高信噪比相对于低信噪比  $re-MSE$  更小,产生于 AR(1)的高频数据比产生于 ARCH(1)的高频数据

re-MSE 更小, 误差项来自标准正态分布比来自厚尾的  $t(5)$  分布 re-MSE 更小, 更大的样本容量和更大的频率倍数 re-MSE 更小. 这些类似于文献[29]所得到的结论, 不过文献[29]研究了 MIDAS 方法在最小二乘估计的条件均值回归中的效果, 本文研究的则是 MIDAS 方法在非对称最小二乘的条件 Expectile 回归中的效果. 非对称最小二乘比最小二乘更加复杂, 且在  $\tau = 0.5$  时能包含最小二乘. 同时在文献[29]中误差项只考虑了正态分布的情形, 本文则同时研究了误差项来自厚尾的  $t$  分布的情况.

总的来说, 在上述的 DGP 设定下, MIDAS 方法在 Expectile 回归中得到的参数估计非常有效, 并且比常用的等权重加总方法更加有效.

### 3 实证研究

#### 3.1 数据描述

本文将使用 MIDAS-Expectile 模型去探究加密货币收益率的 EVaR 受到各个风险因子的影响. 加密货币种类繁多, 本文选用几种主要的加密货币 (Bitcoin, Ripple, Litecoin, Stellar) 进行实证研究, 加密货币的交易数据可以在 coinmarketcap.com 获得, 图 2 展示了价格的走势, 可以明显观察到 2014 年初及 2018 初加密货币价格攀升到顶峰后的大跌. 选取这些加密货币的原因在于它们是市场价值排名前列的加密货币, 具有代表性和活跃的交易量, 同时交易历史数据足够多. 此外, 为了研究整个加密货币市场, 本文同时选用了加密货币指数 CRIX 进行分析, CRIX 的数据可以在 thecrix.de 获得.

表 2 快速衰减权重下 MIDAS 方法斜率估计的  $MSE_{NALS}$  和 re-MSE  
 Tab. 2  $MSE_{NALS}$  and re-MSE of slope estimation of MIDAS method under fast decay weight

斜率	高频数据	误差项	$T=200$				$T=500$			
			$n=10$		$n=90$		$n=10$		$n=90$	
			$MSE_{NALS}$	re-MSE	$MSE_{NALS}$	re-MSE	$MSE_{NALS}$	re-MSE	$MSE_{NALS}$	re-MSE
$\beta_1 = 3$	AR(1)	$N(0,1)$	0.0105	0.0145	0.0316	0.0081	0.0042	0.0057	0.0082	0.0019
			0.0075	0.0103	0.0147	0.0037	0.0031	0.0043	0.0065	0.0015
			0.0053	0.0072	0.0116	0.0028	0.0022	0.0030	0.0039	0.0009
		$t(5)$	0.0100	0.0133	0.0198	0.0051	0.0045	0.0061	0.0081	0.0019
			0.0363	0.0484	0.0563	0.0149	0.0150	0.0205	0.0209	0.0049
			0.0193	0.0266	0.0282	0.0071	0.0083	0.0115	0.0136	0.0032
	ARCH(1)	$N(0,1)$	0.0089	0.0125	0.0161	0.0039	0.0034	0.0048	0.0068	0.0015
			0.0329	0.0437	0.0393	0.0103	0.0163	0.0224	0.0218	0.0052
			0.1462	0.1914	0.5170	0.2799	0.0570	0.0614	0.2086	0.1121
		$t(5)$	0.1087	0.1368	0.3540	0.1933	0.0420	0.0441	0.0988	0.0491
			0.0780	0.0906	0.1922	0.0963	0.0270	0.0273	0.0699	0.0298
			0.1317	0.1671	0.4762	0.2223	0.0587	0.0655	0.1533	0.0940
$\beta_1 = 5$	AR(1)	$N(0,1)$	0.4745	0.6082	1.5230	0.6427	0.1841	0.1997	0.4273	0.2677
			0.2902	0.3620	0.9788	0.5074	0.1039	0.1115	0.2593	0.1464
			0.1235	0.1496	0.3674	0.1877	0.0437	0.0438	0.1310	0.0579
		$t(5)$	0.4713	0.5922	2.0311	0.7702	0.1884	0.2072	0.4370	0.2608
			0.0105	0.0052	0.0262	0.0024	0.0042	0.0021	0.0098	0.0008
			0.0076	0.0037	0.0214	0.0019	0.0031	0.0016	0.0072	0.0006
	ARCH(1)	$N(0,1)$	0.0053	0.0026	0.0162	0.0014	0.0022	0.0011	0.0037	0.0003
			0.0100	0.0048	0.0275	0.0026	0.0045	0.0022	0.0109	0.0009
			0.0381	0.0188	0.0585	0.0055	0.0149	0.0074	0.0243	0.0021
		$t(5)$	0.0224	0.0112	0.0387	0.0035	0.0083	0.0041	0.0164	0.0014
			0.0085	0.0042	0.0194	0.0017	0.0034	0.0017	0.0076	0.0006
			0.0359	0.0176	0.0585	0.0053	0.0164	0.0082	0.0282	0.0024
ARCH(1)	$N(0,1)$	0.1352	0.0622	0.5608	0.1260	0.0566	0.0221	0.2047	0.0420	
		0.1027	0.0464	0.4011	0.0873	0.0407	0.0150	0.1428	0.0248	
		0.0705	0.0290	0.2917	0.0492	0.0265	0.0097	0.0706	0.0097	
	$t(5)$	0.1464	0.0665	0.6403	0.1354	0.0535	0.0200	0.1418	0.0295	
		0.4841	0.2201	1.1601	0.2341	0.1869	0.0735	0.3808	0.0793	
		0.2854	0.1267	0.7656	0.1640	0.1014	0.0389	0.2584	0.0477	
ARCH(1)	$t(5)$	0.1236	0.0515	0.3137	0.0568	0.0425	0.0155	0.0935	0.0135	
		0.4741	0.2278	1.4801	0.2897	0.2001	0.0764	0.3656	0.0783	

[注] 表中每个区域中数据的第一列是  $MSE_{NALS}$ , 第二列是 re-MSE; 第一到第四行分别表示  $\tau = 0.05, 0.1, 0.5, 0.95$  下的情况, 表 3 同.

表 3 缓慢衰减权重下 MIDAS 方法斜率估计的  $MSE_{NALS}$  和 re-MSE  
 Tab. 3  $MSE_{NALS}$  and re-MSE of slope estimation of MIDAS method under slow decay weight

斜率	高频数据	误差项	$T=200$				$T=500$			
			$n=10$		$n=90$		$n=10$		$n=90$	
			$MSE_{NALS}$	re-MSE	$MSE_{NALS}$	re-MSE	$MSE_{NALS}$	re-MSE	$MSE_{NALS}$	re-MSE
$\beta_1 = 3$	AR(1)	N(0,1)	0.0094	0.6024	0.0116	0.0140	0.0038	0.3522	0.0036	0.0035
			0.0068	0.5023	0.0078	0.0087	0.0028	0.2521	0.0029	0.0028
		0.0046	0.4016	0.0047	0.0050	0.0019	0.1906	0.0018	0.0017	
		0.0092	0.6175	0.0099	0.0120	0.0038	0.3533	0.0035	0.0035	
		0.0337	1.2317	0.0416	0.0506	0.0134	0.8133	0.0146	0.0148	
		0.0187	0.8989	0.0211	0.0238	0.0076	0.5483	0.0072	0.0069	
	ARCH(1)	N(0,1)	0.0077	0.5941	0.0084	0.0087	0.0034	0.3017	0.0032	0.0029
			0.0316	1.1037	0.0473	0.0553	0.0129	0.8196	0.0144	0.0142
		0.1575	1.0256	0.9094	0.5463	0.0570	0.7893	0.2964	0.2753	
		0.1033	0.9304	0.6315	0.5415	0.4623	0.7503	0.2165	0.3355	
		0.0748	0.9385	0.4136	0.6084	0.0290	0.6486	0.1238	0.2866	
		0.1517	1.0691	0.9266	0.5236	0.0544	0.8324	0.2870	0.3200	
$t(5)$	N(0,1)	0.4595	1.2145	2.8755	0.9362	0.2017	1.1013	1.1038	0.8631	
		0.2785	1.1892	1.6276	1.0140	0.1134	1.0515	0.5667	0.7051	
	0.1276	1.0650	0.8173	1.0312	0.0489	0.8119	0.2452	0.5436		
	0.4988	1.1897	2.4250	0.8567	0.1805	1.0842	1.1272	0.8744		
	0.0094	0.2845	0.0103	0.0044	0.0038	0.1417	0.0036	0.0013		
	0.0068	0.2237	0.0071	0.0029	0.0028	0.1021	0.0029	0.0010		
$\beta_1 = 5$	AR(1)	N(0,1)	0.0046	0.1648	0.0048	0.0018	0.0019	0.0730	0.0019	0.0006
			0.0092	0.2876	0.0094	0.0041	0.0038	0.1429	0.0037	0.0013
		0.0334	0.7781	0.0325	0.0142	0.0134	0.4126	0.0131	0.0048	
	ARCH(1)	N(0,1)	0.0188	0.4950	0.0187	0.0076	0.0076	0.2488	0.0079	0.0027
			0.0077	0.2626	0.0077	0.0029	0.0034	0.1228	0.0031	0.0010
		0.0316	0.7058	0.0380	0.0159	0.1296	0.4119	0.0131	0.0047	
$t(5)$	N(0,1)	0.1546	0.7790	0.7719	0.1777	0.0570	0.4740	0.2242	0.1044	
		0.1035	0.6665	0.4927	0.1670	0.0462	0.4304	0.1435	0.0895	
	0.0747	0.6419	0.3040	0.2069	0.0290	0.3483	0.0714	0.0606		
	0.1516	0.8216	0.7058	0.1452	0.0544	0.4875	0.1901	0.0864		
	0.5153	1.1323	2.3011	0.3930	0.2028	0.9200	0.9605	0.3404		
	0.3099	1.0624	1.5523	0.4605	0.1222	0.7276	0.4749	0.2404		
ARCH(1)	N(0,1)	0.1297	0.7759	0.6203	0.3261	0.0526	0.5108	0.1552	0.1415	
		0.5075	1.1600	2.5547	0.4754	0.1894	0.8875	0.9391	0.3589	

表 4 中的 Panel A 给出了各种加密货币月度对数收益率数据的描述性统计. 对月度风险的度量有实际的需求和意义, 如目前证券投资中很大比例的投资期限还不是特别高频, 通常为月度, MSCI 公司提供的 BARRA 多因子风险模型数据对风险的估计频率就为月度. 文献[13]使用 MIDAS 模型估计股票收益率的分位数的频率是月度和季度, 本文借鉴了上述做法. 因此就传统金融资产市场和加密货币市场风险的关系这一较为宏观整体的问题, 本文将从月度频率的角度进行风险度量.

参考文献[24], 纳入本文考虑范围内的影响加密货币收益率的 EVaR 的外部变量包括上证综指 (SSEC) 和标普 500 (S&P500) 的对数收益率, 可以

反映中国和美国的股票市场加密货币收益的风险带来的影响. 本文同时还考虑基于美元指数 (USD index) 以及美元兑离岸人民币汇率来反映国际外汇市场, 特别是中美两国外汇市场的影响. 此外本文还引入了标普高盛贵金属指数、能源指数以及商品全收益指数, 代表商品市场带来的影响. 所有这些外部变量的数据都获取自 Wind, 表 4 中的 Panel B 给出了这些外部变量的日度收益数据的描述性统计.

从表 4 的 Panel A 中可以发现加密货币月度收益率均值都为正; Ripple, Litecoin, Stellar 中位数为负, 而 Bitcoin 和 CRIX 中位数为正; Ripple, Litecoin, Stellar 的标准差明显大于 Bitcoin 和



CRIX. 这些表明 Ripple, Litecoin, Stellar 收益率的波动较大, 特别是会有极端的正收益率, 从图 2 的价格走势也可以发现它们会有突然的价格暴涨. 与其不同, Bitcoin 作为市值最大的加密货币, 收益率的波动较小, 价格变动更为稳定, 其在 CRIX 指数中的较大权重使得 CRIX 更稳定. 除了 CRIX, 加密货币收益率都有显著的右偏与尖峰厚尾的分布特征, ADF 检验也表明它们的收益率序列是显著平稳的. 从 Panel B 中可以发现外部变量月度收益率均值和中位数都为正, 除了三种商品指数. 偏度表明除了美元兑人民币汇率右偏分布外, 其他外部变量都是左偏分布. 所有外部变量都有显著的尖峰厚尾的分布特征, ADF 检验也表明收益率序列是显著平稳的.

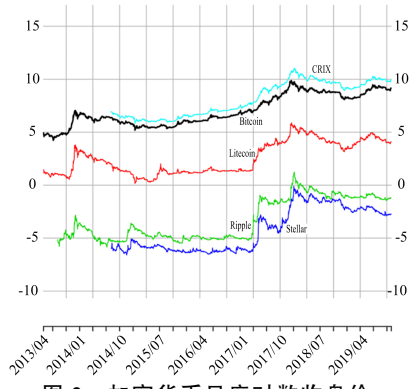


图 2 加密货币日度对数收盘价  
Fig. 2 Daily log closing price of cryptocurrency

表 4 数据概述  
Tab. 4 Data description

		begin	end	size	mean	median	std	skew	skew, 2SE	kurt	kurt, 2SE	ADF
Panel A: 加密货币收益率 (月度)	Bitcoin	2013/4	2019/5	74	5.5704	2.5632	29.6886	2.2229	3.9809	10.9601	9.9333	-3.8891**
	Ripple	2013/8	2019/5	70	6.2215	-6.3877	53.7432	1.8948	3.3039	4.1909	3.7005	-4.2535***
	Litecoin	2013/4	2019/5	74	4.4356	-3.7283	45.2669	3.3599	6.0171	17.5705	15.9245	-3.3628*
	Stellar	2014/8	2019/5	58	7.1933	-1.7758	48.7676	1.5843	2.5251	3.3140	2.6806	-4.1474***
	CRIX	2014/7	2019/5	59	5.1120	3.6896	25.6903	0.4976	0.7996	0.6834	0.5572	-2.7094
Panel B: 外部变量收益率 (日度)	S&P500	2013/4/1	2019/5/31	1554	0.0364	0.0531	0.8188	-0.4768	-3.8407	3.5810	14.4310	-12.17***
	上证综指	2013/4/1	2019/5/31	1501	0.0109	0.0548	1.4772	-1.1325	-8.9652	6.7261	26.6405	-10.585***
	贵金属指数	2013/4/1	2019/5/31	1555	-0.0187	0.0000	0.9945	-0.7962	-6.4148	9.8563	39.7321	-12.599***
	商品指数	2013/4/1	2019/5/31	1555	-0.0463	0.0000	1.1657	-0.1801	-1.4514	2.1980	8.8605	-11.296***
	能源指数	2013/4/1	2019/5/31	1555	-0.0590	0.0183	1.7947	-0.1087	-0.8755	2.6741	10.7796	-11.3***
	美元指数	2013/4/1	2019/5/31	1610	0.0104	0.0124	0.4201	-0.0933	-0.7648	2.2119	9.0723	-11.927***
	美元兑人民币	2013/4/1	2019/5/31	1610	0.0069	0.0016	0.2511	0.5297	4.3422	15.9744	65.5199	-10.198***

[注] 所有变量的收益率计算为  $r_t = 100 \ln(p_t/p_{t-1})$ . 表中统计量 skew, 2SE 是偏度的显著性指标, 计算为偏度除以其 2 倍的标准差. 若  $|skew, 2SE| > 1$ , 那么偏度显著不同于 0. kurt, 2SE 类似. 符号 \* 表示在显著性水平 10% 下显著, \*\* 在 5% 下显著, \*\*\* 在 1% 下显著.

3.2 实证结果

在这部分的实证中, 本文将每个外部变量的日度收益  $factor_{t-n}^{day}$  分别作为自变量对每种加密货币的月度收益  $crypt_t^{month}$  的 EVaR 进行回归, 即 MIDAS-Expectile 回归形式为

$$crypt_t^{month} = \beta_0 + \beta_1 \sum_{d=1}^D B(d, \theta_1, \theta_2) factor_{t-n}^{day} + u_t \quad (15)$$

表 5~表 9 给出了式 (15) 回归的参数估计结果, 揭示了其他传统金融市场与新兴的加密货币市场风险的联系. 研究金融市场间的风险传染机制极为重要, 然而关于加密货币市场的相关研究极为缺乏, 因此本文希望通过 MIDAS-Expectile 回归模型来对这一方面的研究进行补充.

表 5 Bitcoin 的 Expectile 回归结果  
Tab. 5 Expectile regression result for Bitcoin

	S&P500	上证综指	贵金属指数	商品指数	能源指数	美元指数	美元兑人民币汇率	
$\beta_0$	-14.3999 (0)	-13.7877 (0)	-13.3454 (0)	-13.7212 (0.0011)	-14.2733 (0)	-15.0167 (0)	-15.9193 (0)	
$\beta_1$	20.9282 (0.0317)	7.0045 (0.0111)	-15.3241 (0)	13.136 (0.0022)	6.3251 (0.014)	27.7691 (0)	82.5498 (0)	
$\theta_1$	-1.0351 (0.0456)	15.8611 (0.0033)	17.6158 (0)	-6.7307 (0)	-0.3865 (0.2143)	-0.8614 (0.3186)	-1.6835 (0.0748)	
Panel A: 指数 Almon 权重函数	$\theta_2$	0.0614 (0.0286)	-0.7098 (0.0037)	-2.5549 (0)	0.2817 (0)	0.0323 (0.0134)	0.3203 (0.0401)	
	p_both					0.0107		
	AIC	381.7611	382.708	382.1663	389.1497	387.5175	386.653	380.5475
	BIC	390.9774	391.9243	391.3826	398.366	396.7338	395.8692	389.7638
	coverage	0.2162 (0.9383)	0.2432 (0.6599)	0.2027 (0.7916)	0.1757 (0.7553)	0.2027 (0.983)	0.2162 (0.9541)	0.2297 (0.875)

续表 5

	S&-P500	上证综指	贵金属指数	商品指数	能源指数	美元指数	美元兑人民币汇率	
Panel B: Beta 权重函数	$\beta_1$	-3.859 (0.048)	-5.3461 (0.0032)	-7.9955 (0.0441)	-4.1282 (0.0414)	1.0436 (0.6621)	7.7134 (0.6827)	14.9581 (0.3103)
	AIC	393.903	397.1765	394.0916	393.3344	388.9758	388.7736	390.653
	BIC	403.1193	406.3927	403.3078	402.5507	398.192	397.9898	399.8693
	coverage	0.2027 (0.9272)	0.2432 (0.9341)	0.2027 (0.2477)	0.2162 (0.9489)	0.1892 (0.9521)	0.2027 (0.9881)	0.2027 (0.9652)

[注] 本表报告了在  $\tau = 0.1$  下对 Bitcoin 月度收益 EVaR 的 MIDAS-expectile 回归模型的结果, 每个外部变量分别作为自变量. Panel A 报告了权重函数为指数 Almon 多项式函数的结果, Panel B 报告了权重函数为 Beta 函数的结果. 参数估计下方的括号内是显著性检验的 p 值. p\_both 是  $(\theta_1, \theta_2)$  同时为 0 这一假设检验的 p 值. 参数估计的方差详见文献[5]. coverage 报告了收益率落在估计的 EVaR 下方的比例, 它下方的括号报告了原假设为 coverage 等于  $\alpha_0$  这一检验的 p 值. 表 6~9 同.

表 6 Ripple 的 Expectile 回归结果

Tab. 6 Expectile regression result for Ripple

	S&-P500	上证综指	贵金属指数	商品指数	能源指数	美元指数	美元兑人民币汇率		
Panel A: 指数 Almon 权重函数	$\beta_0$	-24.1613 (0)	-25.8261 (0)	-25.5437 (0)	-25.6596 (0)	-24.5576 (0)	-25.2512 (0)	-26.2119 (0)	
	$\beta_1$	-23.6827 (0)	-2.511 (0.002)	-30.9469 (0.0335)	-14.1215 (0.0212)	-11.5547 (0.0462)	-17.4889 (0.0086)	70.0225 (0.0066)	
	$\theta_1$	0.2168 (0)	0 (1)	-0.7086 (0.0086)	-0.1043 (0.4783)	0.8115 (0.1042)	-8.6962 (0)	48.4513 (0)	
	$\theta_2$	0.0006 (0)	0 (1)	0.0378 (0.0234)	0.0079 (0.4833)	-0.0344 (0.0287)	1.2494 (0)	-5.3982 (0)	
	p_both				0.0333				
	AIC	435.5164	438.0952	441.3357	437.839	438.1615	438.9038	433.6862	
	BIC	444.5104	447.0892	450.3296	446.833	447.1555	447.8978	442.6802	
	coverage	0.3 (0.7741)	0.2571 (0.8586)	0.2571 (0.3758)	0.2571 (0.8958)	0.2571 (0.7771)	0.2286 (0.8766)	0.2571 (0.8491)	
	Panel B: Beta 权重函数	$\beta_1$	-4.1814 (0.0139)	-2.5019 (0.6454)	-9.8663 (0.0114)	-7.877 (0.0393)	-5.0412 (0.0099)	-18.5903 (0.0021)	19.277 (0.0232)
		AIC	441.172	437.8583	445.7739	447.5525	447.6019	442.2484	440.1022
BIC		450.166	446.8522	454.7679	456.5464	456.5959	451.2424	449.0961	
coverage		0.2571 (0.8941)	0.2714 (0.8911)	0.2286 (0.4389)	0.2857 (0.7664)	0.3 (0.8663)	0.2429 (0.4018)	0.2429 (0.9772)	

表 7 Litecoin 的 Expectile 回归结果

Tab. 7 Expectile regression result for Litecoin

	S&-P500	上证综指	贵金属指数	商品指数	能源指数	美元指数	美元兑人民币汇率		
Panel A: 指数 Almon 权重函数	$\beta_0$	-23.5654 (0)	-18.9563 (0)	-20.6798 (0)	-21.842 (0)	-21.5859 (0)	-24.2198 (0)	-22.0247 (0)	
	$\beta_1$	14.5711 (0.0052)	5.7548 (0.0463)	-34.4053 (0.0002)	-21.8768 (0)	-2.3895 (0.0013)	28.4158 (0.0492)	83.6807 (0.0047)	
	$\theta_1$	-4.2189 (0)	10.0618 (0.0669)	1.4 (0.0015)	9.7321 (0.4626)	5.7526 (0.2476)	-4.9293 (0.0297)	0.6764 (0.0031)	
	$\theta_2$	0.4263 (0)	-0.3839 (0.0434)	-0.1713 (0.0059)	-0.4194 (0.4816)	-0.2743 (0.046)	0.8612 (0.0278)	-0.0938 (0.1599)	
	p_both				0				
	AIC	438.2896	435.8072	428.3786	444.1366	437.8821	431.8412	433.5358	
	BIC	447.5059	445.0235	437.5948	453.3529	447.0983	441.0574	442.7521	
	coverage	0.2297 (0.6943)	0.2162 (0.4155)	0.2568 (0.6032)	0.2568 (0.35)	0.2027 (0.9576)	0.2027 (0.5619)	0.2027 (0.9238)	
	Panel B: Beta 权重函数	$\beta_1$	-3.5247 (0.0107)	-7.2047 (0.011)	-9.4108 (0.1064)	-5.4619 (0.0077)	1.6304 (0.6665)	11.6311 (0.4994)	27.2914 (0.0253)
		AIC	439.2953	449.7481	443.8329	441.5946	436.8697	436.2809	440.8234
BIC		448.5115	458.9643	453.0492	450.8108	446.086	445.4972	450.0397	
coverage		0.2027 (0.8022)	0.2432 (0.9904)	0.2027 (0.2354)	0.2162 (0.8173)	0.2027 (0.935)	0.2027 (0.9757)	0.2297 (0.7959)	

表 8 Stellar 的 Expectile 回归结果  
Tab. 8 Expectile regression result for Stellar

		S&P500	上证综指	贵金属指数	商品指数	能源指数	美元指数	美元兑人民币汇率
	$\beta_0$	-22.8563 (0)	-23.353 (0)	-23.6244 (0)	-23.5805 (0)	-23.7307 (0)	-23.0567 (0)	-24.584 (0)
Panel A: 指数 Almon 权重函数	$\beta_1$	8.6245 (0.0015)	6.6691 (0.0261)	27.2723 (0.0238)	3.5155 (0)	-1.7986 (0)	-36.4781 (0.0002)	62.3649 (0.0452)
	$\theta_1$	4.0259 (0)	-1.4058 (0.0623)	-0.4378 (0.0016)	2.8 (0.2986)	-3.9759 (0.1707)	1.0331 (0.0227)	1.1943 (0.0375)
	$\theta_2$	-1.098 (0)	0.1335 (0.0132)	0.0222 (0.0214)	-0.1228 (0.3882)	0.2026 (0.0019)	-0.0375 (0.0001)	-0.1348 (0.0373)
Panel A: 指数 Almon 权重函数	p_both	0						
	AIC	354.1084	353.0661	354.3013	355.5928	355.6613	355.02	352.9317
	BIC	362.3502	361.3079	362.5431	363.8346	363.9036	363.2618	361.1735
	coverage	0.2759 (0.9159)	0.2759 (0.8353)	0.2586 (0.7806)	0.2586 (0.9927)	0.2931 (0.946)	0.2759 (0.9829)	0.2931 (0.9517)
Panel B: Beta 权重 函数	$\beta_1$	8.1921 (0.0428)	0.0623 (0.9724)	3.2773 (0.7131)	-0.5568 (0.8593)	-0.2401 (0.923)	1.8078 (0.8857)	14.0759 (0.5208)
	AIC	358.7112	355.6817	355.2189	355.6366	355.6673	355.6497	356.0014
	BIC	366.9529	363.9235	363.4607	363.8783	363.9091	363.8915	364.2432
	coverage	0.2931 (0.5717)	0.2931 (0.9996)	0.2759 (0.9705)	0.3103 (0.9972)	0.2931 (0.9861)	0.2759 (0.9863)	0.2759 (0.9259)

表 9 CRIX 的 Expectile 回归结果  
Tab. 9 Expectile regression result for CRIX

		S&P500	上证综指	贵金属指数	商品指数	能源指数	美元指数	美元兑人民币汇率
	$\beta_0$	-14.7704 (0)	-13.0162 (0.0138)	-14.8094 (0.006)	-12.3086 (0)	-12.8953 (0)	-14.2903 (0.0002)	-14.6765 (0)
	$\beta_1$	25.9409 (0)	9.9309 (0.015)	-10.0167 (0.0432)	10.128 (0)	3.4396 (0.0322)	-20.5194 (0.0134)	32.8759 (0.0244)
Panel A: 指数 Almon 权重函数	$\theta_1$	-0.4961 (0.4469)	-0.0284 (0.9576)	12.6267 (0.0111)	-1.5978 (0)	-4.5551 (0.0002)	-3.0378 (0)	-1.3071 (0.1092)
	$\theta_2$	0.0361 (0.2162)	0.0382 (0)	-0.5996 (0.0055)	0.104 (0)	0.6489 (0.0003)	0.1968 (0)	0.1023 (0.1002)
	p_both	0						
	AIC	294.1238	304.3921	306.656	300.7523	300.0051	318.1314	303.2302
	BIC	302.4339	312.7023	314.9661	309.0624	308.3153	326.4416	311.5403
	coverage	0.2373 (0.833)	0.2373 (0.5677)	0.2034 (0.9191)	0.2034 (0.9736)	0.2373 (0.8715)	0.1864 (0.7326)	0.1864 (0.8116)
Panel B: Beta 权重 函数	$\beta_1$	1.7462 (0.0252)	-2.3954 (0.517)	-10.8795 (0.0065)	5.7364 (0.1427)	2.8247 (0.3147)	-8.9821 (0.0277)	12.6989 (0.3755)
	AIC	302.9799	304.3145	318.4413	302.2068	302.5538	306.4083	308.3742
	BIC	311.29	312.6246	326.7514	310.5169	310.8639	314.7184	316.6843
	coverage	0.1864 (0.9408)	0.2203 (0.9394)	0.2203 (0.1081)	0.2203 (0.6679)	0.2034 (0.792)	0.2034 (0.9567)	0.2034 (0.8616)

为了比较不同的权重函数,本文在表 5~表 9 中的 Panel A 和 Panel B 分别给出了指数 Almon 多项式权重函数和 Beta 权重函数的回归结果. Beta 权重函数形式如下:

$$B(d, \theta_1, \theta_2) = \frac{f(d/D, \theta_1, \theta_2)}{\sum_{d=1}^D f(d/D, \theta_1, \theta_2)},$$

其中,

$$f(x, \theta_1, \theta_2) = \frac{x^{\theta_1-1}(1-x)^{\theta_2-1}\Gamma(\theta_1+\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)},$$

$$\Gamma(a) = \int_0^{\infty} e^{-t} t^{a-1} dt.$$

对比两种权重函数的回归结果,不难发现指数 Almon 权重函数的 AIC 和 BIC 比 Beta 权重函数的更小,这表明使用指数 Almon 权重函数有较小的信息损失,从而提高了模型的质量。此外,指数 Almon 权重函数的斜率估计值的绝对值比 Beta 权重函数的更大,这也使得指数 Almon 权重函数的斜率显著性检验的  $p$  值比 Beta 权重函数的更小,从而更具有统计显著性。这与文献[30]比较这两种权重函数时发现的结果一致。

为了证明在 Expectile 回归模型中引入 MIDAS 方法的必要性,本文设计了一种假设检验。面对混合频率数据的回归问题,常见的做法是将高频数据以相等的权重加总成低频数据,数据同频率后再直接回归,本文称这一方法为等权重方法。然而,数据权重不一定是相等的,特别是金融时序数据。为此研究者发展了一系列时序数据的处理方法,MIDAS 就是其中之一。正如前面所提到的指数 Almon 多项式成为 MIDAS 方法的权重函数时如果参数全为 0 时,MIDAS 方法退化为等权重。因此可以直接构建假设检验,原假设为权重函数参数全为 0,当原假设被拒绝,则说明高频数据的权重不是相等的,因此引入 MIDAS 方法能够得到更好的拟合结果。通过表 5~表 9 的 Panel A 可以看出,大部分回归中权重函数参数  $(\theta_1, \theta_2)$  显著性检验的  $p$  值都较为显著,至少有一个参数显著不为 0。至于个别回归中两个参数单独都不显著的情况,本文进一步做了原假设  $(\theta_1, \theta_2)$  同时为 0 的检验,结果也显著地拒绝两者都同时为 0。因此,MIDAS 方法可以刻画出高频数据间不同的权重结构,同时也包含了通常所用的等权重方法,在 Expectile 回归中引入 MIDAS 方法能够更加充分地提取了高频数据蕴含的信息。

在本文 MIDAS-Expectile 回归的框架下,通过对权重函数施加约束条件可以识别斜率  $\beta_1$ 。该斜率在某种程度上可以反映外部变量(本文中选取的是传统金融市场的日度收益)对加密货币市场 EVaR 的影响。在表 5~表 9 的 Panel A 中,不难发现所有回归结果的斜率都显著地不为 0。这表明加密货币市场不是孤立于传统金融市场,其他金融市场的信号将对加密货币市场风险有或正向或负向的影响,这和文献[34]的结果有所不同,文献[34]发现其他宏观变量对加密货币市场的影响并不显著。本文借助于 MIDAS-Expectile 模型,能够综合考虑高频信息,且基于 EVaR 风险测度,更加准确地反映了外生变量对加密货币风险的影响。此外,从表中可以发现一些有趣的结果,除 Ripple 外,其他加密货币的月度 EVaR 与中美股票市场日度收益都有正向关系,表明中美股票市场与加密货币市场风险有相同的变化趋势。而 Ripple 的月度 EVaR 与其他金融市场日度收益呈现负向关系,当然,除了美元兑离岸人民币汇率。而所有加密货币的月度 EVaR 都与美元兑离岸人民币汇率日度收益呈现正向关系,表明中美外汇市场对加密货币市场有正的

风险溢出。

本文选取了诸多外生变量,需要确定哪个外生变量对加密货币 EVaR 的回归起到了更好的预测作用。注意到本文的 MIDAS-Expectile 回归,参数个数与模型复杂度相同,比较同一组加密货币回归模型优劣的信息准则 AIC、BIC 可以解决这一问题。在表 5~表 9 的 Panel A 中不难发现,Bitcoin, Ripple, Stellar 三种加密货币的 EVaR 回归中美元兑离岸人民币汇率有最小的 AIC 与 BIC 而 Litecoin 的是贵金属指数,CRIX 的则是 S&P500 指数。这说明中美外汇市场对单一的加密货币有更好的预测效果,可能的原因在于加密货币的全球流通与特殊的交易机制,以及加密货币被利用作为绕过外汇管制的通道。而 CRIX 是综合性的加密货币指数,因此 S&P500 指数作为市场综合性指数对其影响最为明显。

当得到 EVaR 的估计后,需要确保该估计没有偏离真值太远,这样 EVaR 才能获得更好的应用。为此,本文提出了 Expectile 的 coverage 检验,当原假设不被拒绝时,EVaR 的估计值偏离真值不远。表 5~表 9 中给出了该检验的 coverage 值,检验的  $p$  值大部分都大于 0.5,因此原假设没有被拒绝,说明 EVaR 的估计是准确的。

## 4 结论

本文提出了 MIDAS-Expectile 回归模型,目的在于使用混合频率数据时充分提取数据的信息来得到更为精准的 EVaR 估计,为此本文也给出了一系列统计方法和技术来保证模型效果的准确与可靠性。

不等权重的思想在 ALS 和 MIDAS 方法中得到了体现,这对应于金融数据中不同数据点重要性不同的特点,这与传统假设是不同的。本文将 MIDAS 方法引入到 ALS 中形成了 MIDAS-Expectile 回归模型,该模型充分结合了 ALS 与 MIDAS 的优势,同时本文基于数值计算的技术给出了 MIDAS-Expectile 回归模型的 NALS 估计,得到了参数及条件 Expectile 的估计。另外,提供了 NALS 估计的渐近正态性,在此基础上能够完成参数估计的某些统计检验。本文还基于似然比检验构建 coverage 检验来评估条件 Expectile 的估计效果,以此确保得到的估计值不会偏离过大。本文提供了 Expectile 回归模型的似然函数,并给出了信息准则 AIC、BIC。重要的是,实证部分表明 MIDAS-Expectile 回归模型的结果有良好的性质,特别是比传统使用的等权重 Expectile 回归更有优势。

本文使用 MIDAS-Expectile 回归模型探究传统金融市场与加密货币市场风险的联系,这对于了解金融市场间的风险传染机制以及系统性风险监控有所裨益。本文的实证结果表明其他金融市场的信号将对加密货币市场风险有或正向或负向的影响,加密货币市场不是孤立于传统金融市场。此外,本文发现中美外汇市场对单一的加密货币风险有更好的预测效果,S&P500 指数作为市场综合性指数对加密货币指数影响最为明显。很多研究者将其其他统计方法和 Expectile 回归模型相结合,类似的,

本文提出的 MIDAS-Expectile 回归模型是对该模型的拓展. 在当今数字化的浪潮下, 从海量数据获取更优质的信息已成为未来发展的方向. 而相应的在金融领域中, 出现越来越多需要处理混频数据的情况. 本文提出的 MIDAS-Expectile 回归模型顺应这一趋势, 对 Expectile 回归的研究范围进行了很大的拓展, 同时提供了完备的假设检验和统计方法来保证模型的效果. 这些都为海量数据下的风险度量和管理研究提供了有利工具. 但是由于本文模型的复杂程度, 导致对数据量和数据质量的要求比较高, 同时计算也比较复杂耗时. 沿着本文的方向, 下一步有价值的研究工作可能包括但不限于: 优化模型估计的计算方法, 将本文提出的模型用于更高频率数据(分钟级、秒级)的实证研究, Expectile 回归模型的进一步拓展.

#### 参考文献(References)

- [1] NEWEY W K, POWELL J L. Asymmetric least squares estimation and testing [J]. *Econometrica: Journal of the Econometric Society*, 1987, 55 (4): 819-847.
- [2] KUAN C M, YEH J H, HSU Y C. Assessing value at risk with CARE, the conditional autoregressive expectile models[J]. *Journal of Econometrics*, 2009, 150(2): 261-270.
- [3] 谢尚宇, 姚宏伟, 周勇. 基于 ARCH-Expectile 方法的 VaR 和 ES 尾部风险测量[J]. *中国管理科学*, 2014, 22(9): 1-9.
- [4] XIE S, ZHOU Y, WAN A T K. A varying-coefficient expectile model for estimating value at risk[J]. *Journal of Business & Economic Statistics*, 2014, 32 (4): 576-592.
- [5] KIM M, LEE S. Nonlinear expectile regression with application to value-at-risk and expected shortfall estimation [J]. *Computational Statistics & Data Analysis*, 2016, 94: 1-19.
- [6] 谭常春, 操毅文, 叶五一. 基于 Expectile-based VaR 变点检测的金融传染分析[J]. *数理统计与管理*, 2018, 37(2): 371-380.
- [7] DAOUIA A, GIRARD S, STUPFLER G. Extreme M-quantiles as risk measures: From  $L^1$  to  $L^p$  optimization[J]. *Bernoulli*, 2019, 25(1): 264-309.
- [8] 许启发, 丁晓涵, 蒋翠侠. 基于 Expectile 回归的均值-ES 组合投资决策[J]. *中国管理科学*, 2018, 26(10): 20-29.
- [9] GHYSELS E, SANTA-CLARA P, VALKANOV R. The MIDAS touch: Mixed data sampling regression models[R]. CIRANO, 2004; 2004s-20.
- [10] GHYSELS E, SANTA-CLARA P, VALKANOV R. There is a risk-return trade-off after all[J]. *Journal of Financial Economics*, 2005, 76(3): 509-548.
- [11] MERTON R C. An intertemporal capital asset pricing model[J]. *Econometrica*, 1973, 41(5): 867-887.
- [12] GHYSELS E, SANTA-CLARA P, VALKANOV R. Predicting volatility: Getting the most out of return data sampled at different frequencies[J]. *Journal of Econometrics*, 2006, 131(1-2): 59-95.
- [13] GHYSELS E, PLAZZI A, VALKANOV R. Why invest in emerging markets? The role of conditional return asymmetry[J]. *The Journal of Finance*, 2016, 71(5): 2145-2192.
- [14] PETTENUZZO D, TIMMERMANN A, VALKANOV R. A MIDAS approach to modeling first and second moment dynamics [J]. *Journal of Econometrics*, 2016, 193(2): 315-334.
- [15] ANDREOU E. On the use of high frequency measures of volatility in MIDAS regressions [J]. *Journal of Econometrics*, 2016, 193(2): 367-389.
- [16] 夏婷, 闻岳春. 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*, 2018, 26(12): 1-11.
- [17] 尚玉皇, 郑挺国. 短期利率波动测度与预测: 基于混频宏观-短期利率模型[J]. *金融研究*, 2016(11): 47-62.
- [18] XU Q, WANG L, JIANG C, et al. A novel UMIDAS-SVQR model with mixed frequency investor sentiment for predicting stock market volatility [J]. *Expert Systems with Applications*, 2019, 132: 12-27.
- [19] AIGNER D J, AMEMIYA T, POIRIER D J. On the estimation of production frontiers: Maximum likelihood estimation of the parameters of a discontinuous density function [J]. *International Economic Review*, 1976, 17(2): 377-396.
- [20] 谢平, 石午光. 数字加密货币研究: 一个文献综述[J]. *金融研究*, 2015(1): 1-15.
- [21] BUCHHOLZ M, DELANEY J, WARREN J, et al. Bits and bets, information, price volatility, and demand for Bitcoin[J]. *Economics*, 2012, 312: 2-48.
- [22] VAN WIJK D. What can be expected from the BitCoin [R]. Rotterdam, Netherlands: Erasmus Universiteit Rotterdam, 2013.
- [23] KRISTOUFEK L. What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis[J]. *PLoS ONE*, 2015, 10(4): e0123923.
- [24] WALTHER T, KLEIN T. Exogenous drivers of cryptocurrency volatility: A mixed data sampling approach to forecasting[R]. St. Gallen, Switzerland: University of St. Gallen, 2018.
- [25] FRY J, CHEAH E T. Negative bubbles and shocks in cryptocurrency markets [J]. *International Review of Financial Analysis*, 2016, 47: 343-352.
- [26] URQUHART A, ZHANG H. Is Bitcoin a hedge or safe haven for currencies? An intraday analysis [J]. *International Review of Financial Analysis*, 2019, 63: 49-57.
- [27] LI X, WANG C A. The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin[J]. *Decision Support Systems*, 2017, 95: 49-60.
- [28] KOENKER R, BASSETT Jr G. Regression quantiles [J]. *Econometrica: Journal of the Econometric Society*, 1978, 46: 33-50.
- [29] ANDREOU E, GHYSELS E, KOURTELLOS A. Regression models with mixed sampling frequencies [J]. *Journal of Econometrics*, 2010, 158 (2): 246-261.
- [30] GHYSELS E, SINKO A, VALKANOV R. MIDAS regressions: Further results and new directions [J]. *Econometric Reviews*, 2007, 26(1): 53-90.
- [31] AKAIKE H. A new look at the statistical model identification [M]// *Selected Papers of Hirotugu Akaike*. New York: Springer, 1974: 215-222.
- [32] SCHWARZ G. Estimating the dimension of a model [J]. *The Annals of Statistics*, 1978, 6(2): 461-464.
- [33] YAO Q, TONG H. Asymmetric least squares regression estimation: A nonparametric approach[J]. *Journal of Nonparametric Statistics*, 1996, 6 (2-3): 273-292.
- [34] BORRI N. Conditional tail-risk in cryptocurrency markets[J]. *Journal of Empirical Finance*, 2019, 50: 1-19.