

文章编号:1003-207(2020)05-0079-10

DOI:10.16381/j.cnki.issn1003-207x.2020.05.008

基于动态模型平均的大豆期货价格预测研究

熊 涛¹, 鲍玉昆²

(1. 华中农业大学经济管理学院, 湖北 武汉 430070; 2. 华中科技大学管理学院, 湖北 武汉 430074)

摘 要:针对大豆期货价格波动的复杂性及影响因素的多元性, 本文将动态模型平均理论引入大豆期货价格分析与预测研究中, 通过动态选择解释变量和系数时变程度, 在有效控制模型和系数不确定性的同时, 最大限度综合利用大豆期货市场内外部信息, 以提高大豆期货价格预测准确度。具体的, 本文提出一套基于动态模型平均理论的大豆期货价格影响因素与预测分析框架, 从期货市场和经济环境等两方面准确地识别出大豆期货价格影响因素的时变特征, 进而构建大豆期货价格预测模型, 并通过预测误差指标和 Diebold—Mariano 检验法评估其与基准模型的预测能力。研究结果表明, 动态模型平均理论在有效剖析大豆期货价格影响因素的时变特征的同时, 能明显提升大豆期货价格预测准确度。

关键词:农产品期货; 预测模型; 动态模型平均; 时变参数模型

中图分类号:F830 **文献标识码:**A

1 引言

我国大豆期货合约于 1993 年 12 月在大连商品交易所上市交易, 是我国最早上市的农产品期货之一。自上市以来, 大豆期货成交量逐年增加, 2016 年, 大豆期货合约的成交额和成交量分别高达 2448 亿元和 651 万手。经过近 25 年的发展, 我国大豆期货价格已成为重要的大豆市场价格表现指标, 对大豆生产经营者提供了较准确的远期价格波动信息, 也是大豆产业链从业者进行套期保值决策的重要依据。然而, 大豆期货在发挥其市场功能的同时仍难以摆脱衍生品固有的高风险特征, 保证金交易和强行平仓制度等, 使得大豆期货价格波动会造成套期保值者与投机者获利或损失成倍放大, 使用不当不仅容易诱发金融市场的极端风险, 也将危害大豆生产经营活动。因此, 深入分析大豆期货价格波动规律、明晰大豆期货价格决定机制、进而准确预测大豆期货价格走势, 对于大豆产业链套期保值者规避价

格风险, 期货投机者理性投资以及有关政府管理部门科学合理发挥大豆期货的市场功能具有重要意义^[1]。

人们在进行套期保值和追逐风险收益的同时, 越发关注农产品期货市场潜在的价格风险及可能带来的经济损失和对农业生产经营的冲击, 因此大量学者积极投身到农产品期货市场预测研究中, 相继提出多种预测理论与方法。早期主要以统计学为学科基础的自回归 (Wen 和 Wang^[2])、移动平均 (Weng 和 Shy^[3])、指数平滑模型 (Weng 和 Shy^[3]) 为主, 模型简单且具有较好的解释性, 但因其本质视数据规律服从线性特征而很难准确地预测复杂多变的农产品期货价格。鉴于此, 从非线性特征 (Onour 和 Sergi^[4])、价格跳跃 (Schmitz 等^[5])、长记忆性 (Power 和 Turvey^[6]; Wu 等^[7])、时频分析 (Power 和 Turvey^[6]; Yang 和 Zhang^[8]; He 和 Wen^[9])、混合模型 (王书平和朱艳云^[10]; 杨科和田凤平^[11]; Xiong 等^[12])、高频交易信息 (Huang 等^[13]; 杨科和田凤平^[14]) 等视角对农产品期货价格进行预测建模成为近年来的热点。例如, Power 和 Turvey^[6] 运用小波分析对农产品期货价格的长记忆性、长依赖性与可预测性进行研究。Schmitz 等^[5] 整合随机波动率、价格跳跃、季节性和持有随机成本, 构建考虑价格跳跃的随机波动率模型对农产品期货价格进行预测。He 和 Wen^[9] 从量价分析的视角, 运用小波相

收稿日期:2017-10-26; 修订日期:2018-03-13

基金项目:国家自然科学基金资助项目 (71501079, 71771101, 71571080, 71673103); 中国博士后科学基金资助项目 (2015M570648); 中央高校基本科研业务费专项资金资助项目 (2662015PY026)

通讯作者简介:鲍玉昆 (1974—), 男 (汉族), 湖北襄阳人, 华中科技大学管理学院, 教授, 博士生导师, 研究方向: 预测理论与方法, E-mail: yunkunbao@hust.edu.cn.

干分析对农产品期货市场的可预测性与市场有效性进行研究。杨科和田凤平^[14]对我国农产品期货市场已实现波动率的长记忆性与区制转换性进行分析,进而构建长记忆马尔科夫区制转换模型对已实现波动率进行预测。然而,上述研究局限于从农产品期货价格本身的波动特征构建预测模型,忽视了农产品期货市场的内外部因素对该价格的影响。

作为典型的粮油兼用农产品,大豆具有粮食作物、饲料作物和能源作物等多重属性,使得大豆期货价格易受供需因素与宏观经济因素等的影响。因而,有学者从多变量时间序列分析出发,探讨某种(类)因素对农产品期货价格波动的影响,进而构建预测模型(Garcia^[15]; Sanders 等^[16]; Karali 和 Power^[17]; 刘庆富和张金清^[18]; Zhang 和 Qu^[19])。例如, Garcia^[15]运用同步估计方法研究期货期权对农产品期货价格的预测性能。基于 Spline—GARCH 模型, Karali 和 Power^[17]分析宏观经济因素对农产品期货价格波动率的短期与长期影响机理。 Sanders 等^[16]对美国商品期货交易委员会发布的交易员持仓报告对农产品期货价格波动的预测性能进行研究。 Zhang 和 Qu^[19]构建 ARJI—AGARCH 与 AR—MA—GARCH 模型研究国际原油价格冲击对中国农产品期货价格的影响。刘庆富和张金清^[18]采用基于蒙特卡洛马尔科夫链模拟的贝叶斯技术研究大豆和小麦期货隔夜信息对日内交易的预测能力。但是,上述研究局限于探讨某种(类)因素对农产品期货价格的影响,并在预测建模中受限于固定解释变量及系数的预测模型,系统全面地考察期货市场和宏观经济环境等因素对农产品期货价格的时变性影响程度,进而构建预测模型相对空白,也更加复杂,需要统筹考虑解释变量及其回归系数时变性。

2010 年, Raftery 等^[20] 初创性地提出动态模型平均(Dynamic Model Averaging, DMA)理论,以解决包含大量解释变量的多变量时间序列预测问题。动态模型平均理论根据预测表现动态选择解释变量和系数时变程度,在有效控制模型和系数不确定性的同时,最大限度地综合利用各种内外部信息(郭永济等^[21])。因此,动态模型平均可以有效降低单变量预测模型不能兼顾不同外部环境对于目标时间序列的影响所导致的预测结果的不确定性。目前,动态模型平均已在通货膨胀(崔百胜^[22]; Koop 和 Tole^[23]; Filippo^[24]; 郭永济等^[21])、外汇市场(Gupta 等^[25]; Bruyn 等^[26])、商品市场(Buncic 和 Moretto^[27]; Bork 和 Möller^[28]; Wei 等^[29]; Risse 和

Ohl^[30])等预测场景得到成功的应用。动态模型平均能够有效地捕捉解释变量及其回归系数的时变特征,并在此基础上构建预测模型,对于大豆期货价格这类波动特征复杂化与影响因素多元化的预测建模问题,动态模型平均理论在识别大豆期货价格影响因素的时变特征、进而提升大豆期货价格预测性能和稳定性上具有重大潜力。

鉴于此,本文提出一套基于动态模型平均理论的大豆期货价格影响因素与预测分析框架。本文以大连商品交易所上市的大豆 1 号期货价格为研究对象;考虑到大豆的多重属性与我国大豆较高对外依存度,本文从期货市场和和经济环境等方面选择美国大豆期货价格、中国大豆期货注册仓单量、换手率和成交量、中国大豆现货价格、美国西德克萨斯(WTI)轻质原油现货价格和美元兑人民币汇率等 7 个变量作为解释变量。通过分析以上 7 个解释变量对大豆期货价格的时变性影响特征,并在此基础上构建大豆期货价格预测模型,探究允许解释变量及其回归系数随时间变化的大豆期货价格预测效果。

本文的主要贡献如下:(1)首次将动态模型平均理论引入到农产品期货市场价格分析与预测研究中,拓展动态模型平均理论应用领域的同时,为农产品期货市场分析提供一种新的分析框架。(2)提出一套基于动态模型平均理论的大豆期货价格影响因素与预测分析框架,通过解决解释变量及其回归系数的时变性问题,构建准确的大豆期货价格预测模型。(3)针对我国大豆期货价格影响因素的多元性问题,本文从期货市场与经济环境两方面选取 7 个经济变量用做大豆期货价格预测,并首次明晰以上 7 个经济变量对大豆期货价格的时变性影响特征,增进对我国大豆期货价格决定机制的了解。(4)基于预测误差指标和 Diebold—Mariano 检验法,本文构建的基于动态模型平均的大豆期货价格预测模型明显优于基准模型,为大豆期货价格预测提供一种新的建模框架。

2 模型设定与数据描述

2.1 模型设定

自 Raftery 等^[20] 提出动态模型平均理论以来,该理论已成为复杂经济环境下探究经济变量影响因素的时变特征及预测经济变量未来走势的重要基础性分析框架。本节对动态模型平均理论进行简要阐述,首先从时变参数模型(Time Varying Parameter, TVP)入手。TVP 允许回归系数随时间变化,

能有效的捕捉各解释变量对被解释变量的时变影响程度,进而提升建模与预测准确度,已被成功的应用于宏观经济预测等诸多领域。然而,在经济预测实践中,经济变量中的结构突变,可能导致某些时点需要大量的解释变量而另一些时点仅需要少量的解释变量,TVP不能解决此类解释变量的动态选择问题,也难以捕捉结构突变对预测建模的影响(Koop和Tol[23])。此外,待估参数的增加将导致TVP出现过拟合现象,降低预测的推广性(Koop和Tol[23])。

动态模型平均理论能高效地解决以上问题。假设存在 m 个解释变量,则将有 $K=2^m$ 个时变参数模型($k=1,2,\dots,K$):

$$\begin{aligned} y_t &= Z_{t-1}^k \theta_t^{(k)} + \varepsilon_t^{(k)}, \\ \theta_{t+1}^{(k)} &= \theta_t^{(k)} + \eta_t^{(k)}, \end{aligned} \quad (1)$$

其中 $\varepsilon_t^{(k)}$ 服从 $N(0, H_t^{(k)})$, $\eta_t^{(k)}$ 服从 $N(0, Q_t^{(k)})$,以上 K 个模型包含的解释变量及其系数均不同。在 $t-1$ 时刻, $\pi_{t-1,k}$ 表示模型 k 是进行 y_t 预测的正确模型的概率。并设定 $\pi_{t-1}^* = \max\{\pi_{t-1,k}\}_{k=1}^K$ 为各时点模型选择概率的最大值, $1(\lambda)$ 为指示函数,即当 λ 为真时 $1(\lambda)=1$,当 λ 为否时 $1(\lambda)=0$ 。那么,动态模型平均的各时点最终预测值由以上 K 个时变参数模型的预测值及模型选择概率的加权平均值所得,即 $\hat{y}_t^{DMA} = \sum_{k=1}^K \pi_{t-1,k} \hat{y}_t^{(k)}$;而动态模型选择的各时点最终预测值由各时点选择概率最大的模型的预测值所得,即 $\hat{y}_t^{DMS} = \sum_{k=1}^K \hat{y}_t^{(k)} 1(\pi_{t-1}^* = \pi_{t-1,k})$ 。要得到以上最终预测值,需解决各时变参数回归模型及其选择概率的估计问题。

针对第一个问题,可利用卡尔曼滤波更新方法,采用递归方法进行预测。设

$$\begin{aligned} \theta_{t-1}^{(k)} | Y^{t-1} &\sim N(\hat{\theta}_{t-1}^{(k)}, \sum_{i=1}^{t-1} \varepsilon_i^{(k)}) \\ \theta_t^{(k)} | Y^{t-1} &\sim N(\hat{\theta}_t^{(k)}, \sum_{i=1}^t \varepsilon_i^{(k)}) \end{aligned} \quad (2)$$

其中 $\sum_{i=1}^t \varepsilon_i^{(k)} = \sum_{i=1}^{t-1} \varepsilon_i^{(k)} + Q_t^{(k)}$, $Y^{t-1} = \{y_1, \dots, y_{t-1}\}$ 。从 $t=0$ 时刻开始,利用卡尔曼滤波更新公式(2),进而根据如下预测分布进行预测。

$$\begin{aligned} y_t | Y^{t-1} \\ \sim N(x_{t-1} \hat{\theta}_{t-1}^{(k)}, H_t^{(k)} + x_{t-1}^k \sum_{i=1}^{t-1} \varepsilon_i^{(k)} x_{t-1}^{(k)'}) \end{aligned} \quad (3)$$

由公式(3)可见,如上预测过程的实现需对 $Q_t^{(k)}$ 和 $H_t^{(k)}$ 进行有效估计。Raftery等[20]运用遗忘因子方法估计 $Q_t^{(k)}$,如公式(4)所示:

$$\sum_{i=1}^t \varepsilon_i^{(k)} = \sum_{i=1}^{t-1} \varepsilon_i^{(k)} / \lambda \quad (4)$$

或者 $Q_t^{(k)} = (1 - \lambda^{-1}) \sum_{i=1}^{t-1} \varepsilon_i^{(k)}$ 其中 λ 为遗忘因子,并满足 $0 < \lambda \leq 1$,其含义为在估计系数 $\theta_t^{(k)}$ 时,过去 j 期的观测值所占的预测信息权重为 λ^j 。 λ 值越大,表明回归系数 $\theta_t^{(k)}$ 时变速度越慢。根据Koop和Tol[23]的建议,本文设定 $\lambda = 0.99$ 。本文的研究对象大豆期货价格为日度数据, $\lambda = 0.99$ 表示10个交易日之前的数据预测下一期观测值的预测表现最多达到上一期之前数据预测下一期观测值表现的90%。通过对遗忘因子 λ 赋予不同的数值,可探究时变回归系数是否较固定系数能获得更高的预测准确度。接下来,本文运用指数加权移动平均方法对 $H_t^{(k)}$ 进行估计:

$$\hat{H}_t^{(k)} = \frac{1 - \kappa}{1 - \kappa^t} \sum_{j=1}^t \kappa^j (y_{t-j} - Z_{t-j}^k \hat{\theta}_{t-j}^{(k)})^2 \quad (5)$$

其中 κ 为衰减因子。根据文献(RiskMetrics[31]),本文设定 $\kappa = 0.94$,其表示20个交易日之前预测误差最多达到上一期预测误差权重的30%;

至此,通过引入遗忘因子 λ 和衰减因子 κ ,有效地解决了各时变参数回归模型的估计问题。接下来探讨模型选择概率的估计问题。设 $\pi_{t-1,k}$ 表示在 $t-1$ 时刻, $\pi_{t-1,k}$ 表示模型 k 是进行 y_t 预测的正确模型的概率。在模型 $k(k=1, \dots, K)$ 的预测密度和转移概率 $p_k(y_{t-1} | y_{t-2}, y_{t-3}, \dots, y_1)$ 可知的情况下,如果采用马尔科夫链描述 K 个模型的演化过程,则得到

$$\pi_{t-1,k} = \sum_{i=1}^K \pi_{t-1,i} p_{ik} \quad (6)$$

$$\text{其中 } \pi_{t-1,k} = \frac{\pi_{t-1,k} p_k(y_{t-1} | y_{t-2}, y_{t-3}, \dots, y_1)}{\sum_{l=1}^K \pi_{t-1,l} p_l(y_{t-1} | y_{t-2}, y_{t-3}, \dots, y_1)}。$$

但是该方法的计算负担非常重,降低了其实际应用价值。鉴于此,Raftery等[20]将公式(6)进行如下修正:

$$\pi_{t-1,k} = \pi_{t-1,k}^\alpha / \sum_{l=1}^K \pi_{t-1,l}^\alpha \quad (7)$$

其中 $\alpha(0 < \alpha \leq 1)$ 为Raftery等[20]引入的另一个遗忘因子,与 λ 具有相似的功能。根据Koop和Tol[23]的建议,本文设定 $\alpha = 0.99$ 。通过公式(8),可更方便的理解遗忘因子 α 对模型选择的影响。

$$\pi_{t-1,k} \propto \prod_{i=1}^{t-1} [p_k(y_{t-i} | y_{t-i-1}, \dots, y_1)]^{\alpha^i} \quad (8)$$

其中, $p_k(y_{t-i} | y_{t-i-1}, \dots, y_1)$ 为模型 k 的预测密度,该值由卡尔曼滤波所得,用于衡量模型的预测表

现。在 $t-1$ 时刻,如果模型 k 对 t 时刻的观测值预测效果较好,则对该模型 k 赋予较大的权重。

2.2 变量描述及预测误差评价

(1) 变量描述

大豆期货合约价格序列的选取和影响因素的确定是本文数据选择的关键。黄大豆 1 号期货合约以食用品质非转基因大豆为标的物,黄大豆 2 号期货合约以榨油品质转基因、非转基因大豆为标的物。由于我国实行严格的转基因大豆管理制度,使得黄大豆 2 号合约自 2004 年上市以来成交量持续低迷,故本文采用近交割日法生成黄大豆 1 号期货合约日度收盘价格序列,作为被解释变量。

借鉴已有研究农产品期货价格预测理论与实证文献,本文从期货市场与经济环境两方面选取 7 个经济变量用做中国大豆期货价格预测。美国大豆期货作为全球最重要的大豆期货合约,显著影响我国大豆期货价格波动(夏天和程细玉^[32]);而大豆期货注册仓单量、换手率和成交量等交易行为变量也将明显影响大豆期货价格波动(Etienne 等^[33]),因而期货市场方面包括美国大豆期货价格、中国大豆期货注册仓单量、换手率和成交量等 4 个变量。我国大豆现货价格对期货价格存在显著的因果关系(夏天和程细玉^[2]);而作为典型的粮油兼用农产品,大豆期货价格与国际原油价格也存在明显的动态关联

性(Zhang 和 Qu^[19]);此外,作为对外依存度高达 90% 的农产品,外汇市场也将影响我国大豆期货价格(熊正德等^[34]),因而经济环境方面包括中国大豆现货价格、WTI 原油现货价格、美元兑人民币汇率等 3 个变量。

除了以上 7 个经济变量以外,滞后 1 期至 3 期的大豆期货价格也做选作解释变量(根据施瓦兹信息准则确定),但该滞后 3 期解释变量的回归系数不随时间变化。以上所有变量均为日度数据,来自 Wind 资讯金融终端,时间跨度为 2008 年 10 月 10 日至 2017 年 6 月 15 日,共 1878 个样本,其中 2008 年 10 月 10 日至 2009 年 7 月 29 日作为初始样本区间(共 100 个样本),为模型预测提供先验信息。从 2009 年 7 月 30 日开始向前一步滚动预测,即根据初始样本区间估计 2009 年 7 月 30 日的大豆期货收盘价,然后将 2009 年 7 月 30 日的大豆期货收盘价真实值加入初始样本,估计 2009 年 7 月 31 日的大豆期货收盘价,以此类推,直到估计出 2017 年 6 月 15 日的大豆期货收盘价,因此,滚动预测区间的样本量为 1778。由于时变参数模型的估计要求解释变量与被解释变量均满足平稳性,因此本文首先运用 ADF 检验探究各变量的平稳性,进而对非平稳变量进行相应的平稳性转换。平稳性检验结果和平稳性转换方法见表 1。

表 1 变量的平稳性检验结果和平稳性转换方法

变量类型	变量名称	原始序列 ADF 检验 P 值	一阶差分序列 ADF 检验 P 值	T Code
被解释变量	中国大豆期货价格	0.611	0.000***	2
	美国大豆期货价格	0.525	0.000***	2
	中国大豆期货注册仓单量	0.405	0.001***	2
	中国大豆期货换手率	0.001***	—	1
解释变量	中国大豆期货成交量	0.001***	—	1
	中国大豆现货价格	0.354	0.000***	2
	WTI 原油现货价格	0.278	0.001***	2
	美元兑人民币汇率	0.6041	0.001***	2

注:*** 表示在 99% 置信度上显著。转换代码 T code 分别表示为:1—不做转换,2—做一阶差分。

(2) 预测误差评价

为有效评价模型预测表现,本文选取均方根误差(Root Mean Square Error, RMSE)和对称平均绝对百分比误差(Symmetric mean absolute percentage error, SMAPE)作为模型预测的评价指标, RMSE 和 SMAPE 越小,表示预测准确度越高。RMSE 和

SMAPE 表达式如下:

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2} \quad (9)$$

$$SMAPE = \frac{1}{T} \sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{(y_t + \hat{y}_t)/2} \times 100 \quad (10)$$

其中, y_t 与 \hat{y}_t 分别为大豆期货价格在 t 时刻的真实值与预测值, T 为滚动预测区间的样本量。

除了上述预测误差指标,本文运用 Diebold—Mariano(DM)检验法(Diebold 等, 1995)进一步探究基于动态模型平均的大豆期货价格预测模型与基准模型的预测表现是否存在显著差异。在 DM 检验法中,损失函数设定为均方预测误差(mean square prediction error),原假设为被检验预测模型 te 的均方预测误差优于参照预测模型 re 的均方预测误差。具体的,DM 统计量定义为

$$S_{DM} = \frac{\bar{g}}{\sqrt{(\hat{V}_g/T)}} \quad (11)$$

其中, $\bar{g} = (\sum_{t=1}^T g_t)/T$, $g_t = (y_t - \hat{y}_{te,t})^2 - (y_t - \hat{y}_{re,t})^2$, $\hat{V}_g = \gamma_0 + 2 \sum_{t=1}^{\infty} \gamma_t$, ($\gamma_t = \text{cov}(g_{t+1}, g_t)$), γ_0 为 g_t 的方差, $\hat{y}_{te,t}$ 和 $\hat{y}_{re,t}$ 分别表示被检验预测模型 te 和参照预测模型 re 对于大豆期货价格 y_t 的预测值, T 为滚动预测区间的样本量。

3 实证分析

3.1 大豆期货价格影响因素的时变特征

如上节所述,本文选择美国大豆期货价格、中国大豆现货价格、WTI 原油现货价格、中国大豆期货注册仓单量、换手率、成交量和美元兑人民币汇率等 7 个经济变量(解释变量)用于我国大豆期货价格预测。因此,基于动态模型平均理论,本文共需构建 $2^7 = 128$ 个预测模型,各预测模型的解释变量及其系数各不相同,能够准确地捕捉各解释变量对大豆期货价格影响程度的时变特征。本节着重描述大豆期货价格影响因素的时变特征,下节将分析基于动态模型平均的大豆期货价格预测模型与基准模型的预测表现。

动态模型平均理论可以根据模型的预测表现在不同的时点选择最优的解释变量,因此,各预测时点的平均解释变量个数存在明显差异。本文设定 $Size_k$ 为模型 k 的解释变量个数, $\pi_{t|t-1,k}$ 为模型 k 是进行 y_t 预测的正确模型的概率,其中 $k = 1, 2, \dots, K$, $K = 8192$ 为预测模型总数。那么,各预测时点的平均解释变量个数定义如下:

$$E(Size_t) = \sum_{k=1}^K \pi_{t|t-1,k} Size_k \quad (12)$$

动态模型平均可以根据模型的预测表现在不同的时点选择最优的解释变量和时变系数,所以,在各时点的平均解释变量个数也将发生变化。图 1 描述了各时点的平均解释变量个数。可以看出,各时点平均解释变量的个数变化较大,解释变量个数集中

在 3 和 4 左右,均值为 3.2。考虑到解释变量总数为 7 个,动态模型平均倾向于选择较少的解释变量,在保证预测准确度的前提下可简化预测模型,提升预测模型的可解释性。

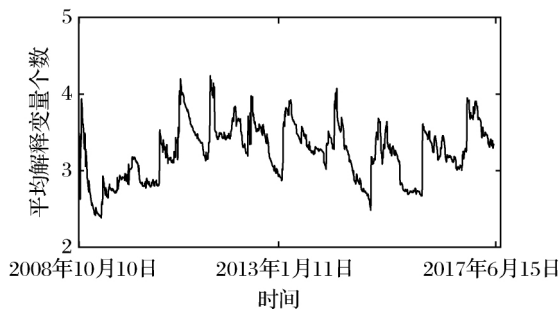


图1 各时点平均解释变量个数

更进一步,可根据各时点预测模型选择概率的大小,将选择概率最大的预测模型所选用的解释变量个数提取出来,以表示各时点最优预测模型的解释变量个数,如图 2 所示。可以发现,在各时点,最优预测模型的解释变量个数集中在 1 至 3 个,均值为 1.8,进一步佐证了动态模型平均倾向于使用较少的解释变量进行预测建模。此外,有 213 个时点(占比 11.9%),最优预测模型不选用任何经济变量,仅根据滞后 1 期至 3 期的大豆期货价格进行预测建模。此外,分别仅有 29 个和 1 个时点,最优预测模型选用了 5 个和 6 个经济变量进行预测建模,而未有选择所有 7 个经济变量的情况。

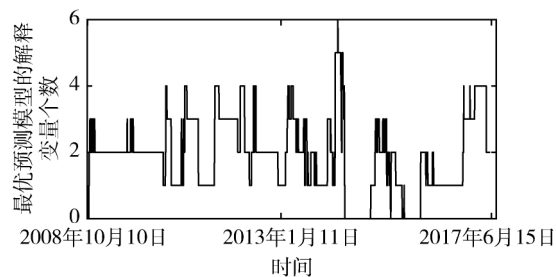


图2 各时点最优预测模型的解释变量个数

图 3 展示了各解释变量在各时点被用于预测建模的选择概率。由图 3 可知,总体而言,各解释变量均呈现出明显的时变特征,选择概率波动剧烈,但变化模式不尽相同。首先聚焦期货市场方面,美国大豆期货价格的选择概率经历了高位运行、低位运行和高位震荡的变化阶段。具体的,在 2008 年低至 2012 年低的近 845 个交易日期间(占比 47.5%),美

国大豆期货价格的选择概率近似于 1,说明在此期间美国大豆期货价格对我国大豆期货价格的影响程度很高;但随后其选择概率急剧下降,维持在 0.4 左右;其选择概率低位运行近 3 年后,于 2016 年初迅速上升至 0.8 左右,并持续高位震荡。较美国大豆期货价格,我国大豆期货注册仓单量的选择概率呈

现出完全不同的变化态势,注册仓单量的选择概率在 2008 年底于 0.7 迅速下跌至 0.02,此后呈现出震荡上扬的变化态势,并于 2013 年初达到最大值 0.92,此后平稳下跌至 2017 年中的 0.38。我国大豆期货换手率和成交量的选择概率呈现出相似的震荡上行变化态势,波动尤为剧烈。

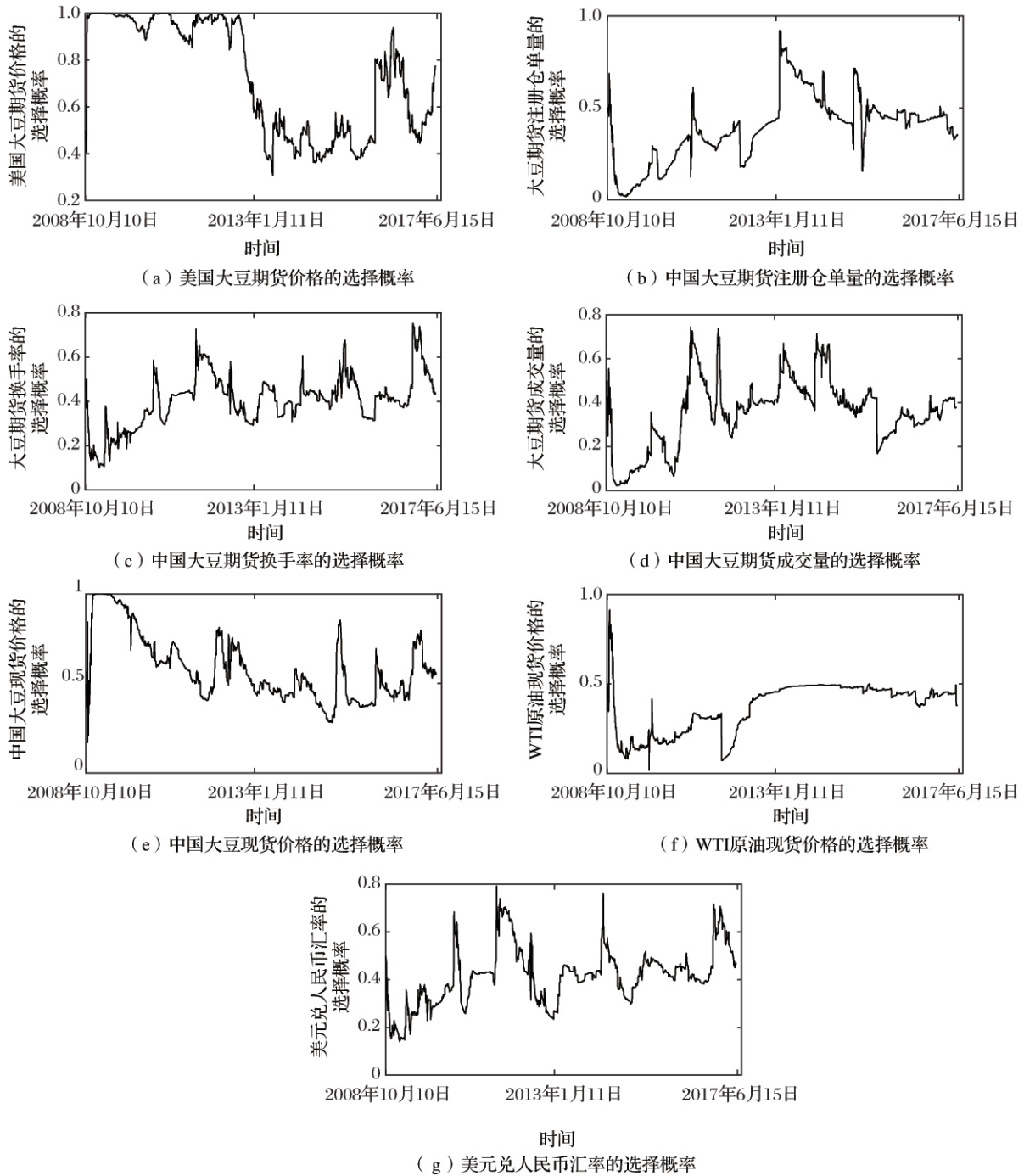


图 3 各解释变量在各时点的选择概率

聚焦经济环境方面,中国大豆现货价格的选择

概率呈现出震荡下行的变化态势,其在 2008 年底至

2009 年初的近 196 个交易日期间(占比 11.1%)近似于 1,随后震荡下行,保持在 0.4 至 0.7 的区间内波动。这说明,总体而言,近十年来我国大豆现货价格对期货价格的影响程度逐年下跌。相对于其他解释变量,WTI 原油现货价格的选择概率呈现出完全不同的变化态势,其在样本初期达到最大值 0.91,继而急剧下跌至 0.08 左右,此后平稳上升,虽然在 2011 年底经历了一个短暂的急剧下跌阶段,但迅速回到稳步上升的模式,维持在 0.45 左右。美元兑人民币汇率的选择概率波动较大,相继经历了低位震荡、高位震荡、急剧下跌和平稳上涨的变化阶段,最小值为 0.07,最大值达到 0.73。

3.2 大豆期货价格预测

如第二章所述,本文构建的基于动态模型平均的大豆期货价格预测模型的参数设定如下: $\lambda = \alpha = 0.99$ 。为评估该模型的预测表现,本文选择如下 6 个预测模型作为基准模型,具体如下:

- (1)动态模型选择: $\lambda = \alpha = 0.99$;
- (2)动态模型平均: $\lambda = \alpha = 0.95$;
- (3)动态模型选择: $\lambda = \alpha = 0.95$;
- (4)贝叶斯模型平均: $\lambda = \alpha = 1$;
- (5)时变参数模型(TVP);
- (6)零漂移随机游走模型;

模型(1)为动态模型选择,其各时点最终预测值由各时点选择概率最大的模型的预测值所得,参数取值与动态模型平均一致。模型(2)和(3)分别为动态模型平均和动态模型选择,但参数取值为 $\lambda = \alpha = 0.95$,与本文构建的基于动态模型平均的大豆期货价格预测模型的参数取值不同,旨在进行稳健性分析,以探究参数取值的不同是否显著影响预测表现。模型(4)为贝叶斯模型平均,其使用模型后验概率加权平均单一模型,是一种特殊的动态模型平均(当 $\lambda = \alpha = 1$ 时,动态模型平均转变为贝叶斯模型平均)。模型(5)为时变参数模型(TVP),其不允许解释变量时变,但允许解释变量的系数时变,在各预测时点,TVP 的解释变量包括全部 7 个经济变量和滞后 1 期至 3 期的大豆期货价格。模型(6)为不带漂移项(零漂移)的随机游走模型。

各模型的预测表现如表 2 所示。由表 2 可得到如下结论:(1)动态模型平均($\lambda = \alpha = 0.99$)与动态模型选择($\lambda = \alpha = 0.99$)的预测表现明显优于其他预测模型,预测准确度相对较高,但以上两个模型的预测表现相当。(2)无论参数设定为 $\lambda = \alpha = 0.99$ 或 $\lambda = \alpha = 0.95$,动态模型平均与动态模型选择的

预测表现均优于贝叶斯模型平均、TVP 和随机游走模型,这说明动态模型平均理论能够根据预测表现动态选择解释变量和系数时变程度,在有效控制模型和系数不确定性的同时,最大限度地综合利用各种内外部信息,进而明显提升预测准确度。(3)随机游走模型的预测表现最差,可能的原因是其忽视了重要影响因素对大豆期货价格的解释性。

表 2 各模型的预测误差

预测模型	预测误差指标	
	RMSE	SMAPE
动态模型平均 $\lambda = \alpha = 0.99$	51.157	0.758
动态模型选择 $\lambda = \alpha = 0.99$	53.241	0.749
动态模型平均 $\lambda = \alpha = 0.95$	58.847	0.841
动态模型选择 $\lambda = \alpha = 0.95$	58.682	0.832
贝叶斯模型平均 $\lambda = \alpha = 1$	61.294	0.849
时变参数模型	62.318	0.847
零漂移随机游走模型	65.114	0.862

更进一步,本文分别将动态模型平均($\lambda = \alpha = 0.99$)和动态模型选择($\lambda = \alpha = 0.99$)作为被检验预测模型,其他模型作为参照预测模型,进行 DM 检验,以探究基于动态模型平均的大豆期货价格预测模型与其他模型的预测表现是否存在显著性差异,DM 检验结果如表 3 所示。

表 3 DM 检验结果

参照预测模型	被检验预测模型	
	动态模型平均 $\lambda = \alpha = 0.99$	动态模型选择 $\lambda = \alpha = 0.99$
动态模型平均 $\lambda = \alpha = 0.99$	—	0.28
动态模型选择 $\lambda = \alpha = 0.99$	-0.35	—
动态模型平均 $\lambda = \alpha = 0.95$	-0.81	-0.68
动态模型选择 $\lambda = \alpha = 0.95$	-0.74	-0.62
贝叶斯模型平均 $\lambda = \alpha = 1$	-1.38*	-1.11
时变参数模型	-1.46*	-1.43*
零漂移随机游走模型	-1.86**	-1.78**

注: * 和 ** 分别表示在 0.1 和 0.05 的显著性水平下显著。

由表 3 可得到如下结论:(1)在 0.1 的显著性水平下,动态模型平均($\lambda = \alpha = 0.99$)或动态模型选择($\lambda = \alpha = 0.99$)的预测表现均显著优于时变参数模型和随机游走模型。(2)在 0.1 的显著性水平下,

动态模型平均($\lambda = \alpha = 0.99$)的预测表现显著优于贝叶斯模型平均;动态模型选择($\lambda = \alpha = 0.99$)的预测表现虽然优于贝叶斯模型平均,但不存在显著性差异。(3)无论参数设定为 $\lambda = \alpha = 0.99$ 或 $\lambda = \alpha = 0.95$,在0.05的显著性水平下,动态模型平均与动态模型选择的预测表现均无显著差异。

4 结语

准确的大豆期货价格预测指出了未来给定时间窗口中价格的波动幅度,不仅为度量极端价格风险、优化交易策略提供重要信息,还可以藉此设计有效的期货市场风险管理工具。针对大豆期货价格波动的复杂性及影响因素的多元性,本文将动态模型平均理论引入农产品期货价格分析与预测研究中,通过动态选择解释变量和系数时变程度,在有效控制模型和系数不确定性的同时,最大限度综合利用各种内外部信息,进而提高大豆期货价格预测准确度。

具体的,本文构建基于动态模型平均的大豆期货价格预测模型,考虑到大豆的多重属性与我国大豆较高的对外依存度,本文从期货市场和经济环境等两方面选择美国大豆期货价格、中国黄大豆1号注册仓单量、换手率和成交量、中国大豆现货价格、WTI原油现货价格和美元兑人民币汇率等7个影响因素用做中国黄大豆1号期货价格预测,通过准确地识别出大豆期货价格影响因素的时变特征,进而构建大豆期货价格预测模型,并通过预测误差指标和DM检验法评估其与基准模型的预测能力。研究表明,大豆期货价格的影响因素存在明显的时变特征,但变化模式不尽相同。通过准确地捕捉影响因素的时变特征,基于动态模型平均的大豆期货价格预测模型的预测表现明显优于基准模型。

参考文献:

- [1] 林学贵. 棉花价格波动溢出效应[J]. 中国管理科学, 2016, 24(S1): 504—508.
- [2] Wen D U, Wang H H. Price behavior in China's wheat futures market[J]. China Economic Review, 2004, 15 (2): 215—229.
- [3] Weng Y, Shy J. The comparison of the forecasting models of the raw material futures' prices[J]. Journal of Information & Optimization Sciences, 2010, 31(1): 129—145.
- [4] Onour I A, Sergi B S. Modeling and forecasting volatility in the global food commodity prices[J]. Agricultural Economics, 2011, 57(3): 132—139.
- [5] Schmitz A, Wangand J Z, Kimn H. A jump diffusion model for agricultural commodities with bayesian analysis[J]. Journal of Futures Markets, 2014, 34(3): 235—260.
- [6] Power G J, Turvey C G. Long—range dependence in the volatility of commodity futures prices: Wavelet—based evidence[J]. Physica A Statistical Mechanics & Its Applications, 2010, 389(1): 79—90.
- [7] Wu F, Myers R J, Guanand Z, et al. Risk—adjusted implied volatility and its performance in forecasting realized volatility in corn futures prices[J]. Journal of Empirical Finance, 2015, 34: 260—274.
- [8] Yang L, Zhang D. Can futures price be a powerful predictor? Frequency domain analysis on Chinese commodity market [J]. Economic Modelling, 2013, 35(5): 264—271.
- [9] He L, Wen X. Predictability and market efficiency in agricultural futures markets: A perspective from price—volume correlation based on wavelet coherency ananlysis [J]. Fractals—complex Geometry Patterns & Scaling in Nature & Society, 2015, 23(2): 1550003.
- [10] 王书平, 朱艳云. 基于多尺度分析的小麦价格预测研究[J]. 中国管理科学, 2016, 24(5): 85—91.
- [11] 杨科, 田凤平. 结构突变条件下农产品期货市场波动率的预测[J]. 中山大学学报(自然科学版), 2014, 53 (2): 59—72.
- [12] Xiong T, Li C, Bao Y, et al. A combination method for interval forecasting of agricultural commodity futures prices[J]. Knowledge—Based Systems, 2015, 77 (C): 92—102.
- [13] Huang W, Huang Z, Matei M, et al. Price volatility forecast for agricultural commodity futures: The role of high frequency data[J]. Journal for Economic Forecasting, 2012, 15(4): 83—103.
- [14] 杨科, 田凤平. 农产品期货市场波动率的动态特征及其预测模型[J]. 经济评论, 2014, (4): 50—67.
- [15] Garcia P. Options—based forecasts of futures prices in the presence of limit moves[J]. Applied Economics, 2007, 39(2): 145—152.
- [16] Sanders D R, Irwin S H, Merrin R P. Smart money: the forecasting ability of CFTC large traders in agricultural futures markets[J]. Journal of Agricultural & Resource Economics, 2009, 34(2): 276—296.
- [17] Karali B, Power G J. Short— and long—run determinants of commodity price volatility[J]. American Journal of Agricultural Economics, 2013, 95(3): 724—738.
- [18] 刘庆富, 张金清. 中国商品期货隔夜信息对日间交易的预测能力[J]. 管理科学学报, 2013, 16(11): 81—

- 94.
- [19] Zhang C, Qu X. The effect of global oil price shocks on China's agricultural commodities[J]. *Energy Economics*, 2015, 51: 354—364.
- [20] Raftery A E, Kárny M, Ettler P. Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill[J]. *Technometrics A Journal of Statistics for the Physical Chemical & Engineering Sciences*, 2010, 52(1): 52.
- [21] 郭永济, 丁慧, 范从来. 中国通货膨胀动态模型预测的实证研究[J]. *中国经济问题*, 2015, (5): 3—15.
- [22] 崔百胜. 基于动态模型平均的中国通货膨胀实时预测[J]. *数量经济技术经济研究*, 2012, (7): 76—91.
- [23] Koop G, Tole L. Forecasting the european carbon market[J]. *Journal of the Royal Statistical Society*, 2013, 176(3): 723—741.
- [24] Filippo G, D. Dynamic model averaging and CPI inflation forecasts: A comparison between the Euro Area and the United States [J]. *Journal of Forecasting*, 2015, 34(8): 619—648.
- [25] Gupta R, Hammoudeh S, Kim W J, et al. Forecasting China's foreign exchange reserves using dynamic model averaging: The roles of macroeconomic fundamentals, financial stress and economic uncertainty[J]. *North American Journal of Economics & Finance*, 2014, 28(C): 170—189.
- [26] Bruyn R D, Gupta R, Eyden R V. Can we beat the random—walk model for the South African Rand—U. S. dollar and South African Rand—UK pound exchange rates? Evidence from dynamic model averaging [J]. *Emerging Markets Finance & Trade*, 2015, 51(3): 502—524.
- [27] Buncic D, Moretto C. Forecasting copper prices with dynamic averaging and selection models [J]. *Social Science Electronic Publishing*, 2014, 33: 1—38.
- [28] Bork L, Møller S V. Forecasting house prices in the 50 states using dynamic model averaging and dynamic model selection [J]. *International Journal of Forecasting*, 2015, 31(1): 63—78.
- [29] Wei Y, Liu J, Lai X D. Which determinant is the most informative in forecasting crude oil market volatility: Fundamental, speculation, or uncertainty? [J]. *Energy Economics*, 2017, 68: 141—150.
- [30] Risse M, Ohl L. Using dynamic model averaging in state space representation with dynamic Occam's window and applications to the stock and gold market [J]. *Journal of Empirical Finance*, 2017, 44: 158—176.
- [31] Morgan J P, Reidens J P. *RiskMetrics—Technical Document* [M]. New York, 1996.
- [32] 夏天, 程细玉. 国内外期货价格与国产现货价格动态关系的研究——基于 DCE 和 CBOT 大豆期货市场与国产大豆市场的实证分析 [J]. *金融研究*, 2006, (2): 110—117.
- [33] Etienne X L, Irwin S H, Garcia P. Price explosiveness, speculation, and grain futures prices [J]. *American Journal of Agricultural Economics*, 2015, 97(1): 65—87.
- [34] 熊正德, 文慧, 凌语蓉. 基于时频分析的农产品期货市场与外汇市场联动关系研究 [J]. *中国管理科学*, 2013, (S1): 255—263.

Soybean Future Prices Forecasting based on Dynamic Model Averaging

XIONG Tao¹, BAO Yu-kun²

(1. College of Economics and Management, Huazhong Agricultural University, Wuhan 430070, China;

2. School of Management, Huazhong University of Sciences and Technology, Wuhan 430074, China)

Abstract: In view of the complexity of soybean futures price fluctuation and the diversity of influencing factors, soybean futures price forecasting is conducted by introducing the dynamic model averaging theory. It should be noted that this technique is capable of dynamically choosing the explanatory variables and coefficient of variation, which will maximize the utilization of various information to control the models effectively and coefficients uncertainty, and finally improve the forecasting performance. More specifically, an analysis framework on influencing factors of soybean futures price is proposed. The time—vary characteristics of soybean futures price's influencing factors are identified from the perspective of futures markets and economics environment, and then a forecasting model for soybean futures price is constructed. Furthermore, covering an out—of—sample period from July 30, 2009 to June 15, 2017, the forecasting performance of the proposed forecasting model is evaluated and compared with six benchmarks on the basis of accuracy measures and Diebold—Mariano test. The experimental results show that the dynamic model av—

eraging can effectively identify the influence degree of each explanatory variables on soybean futures price, and at the same time, outperform the Bayes model averaging, time—varying parameter model, and random walk in soybean futures price forecasting. Policymakers should be cognizant of the fact that there are many potential predictors that can help to forecast the Chinese soybean futures price, and the predictive powers of these predictors vary over time.

Key words: agricultural futures; forecasting model; dynamic model averaging; time — varying parameter model