

# 基于半参数可加模型的余额宝 收益率影响因素分析

孙春燕,马馨悦,刁海涛,丰茂芳

(南京财经大学应用数学学院,江苏南京210023)

**摘要:**对2013年10月—2018年3月的余额宝七日化收益率的影响因素进行分析,一方面从投资策略中选取现金占净比(NCR)作为余额宝的内部影响因素,从运行机制选取余额宝个性化指标:银行间同业拆放利率(SHI)和沪深300指数(CSI),另一方面从外部环境选取影响因素:消费者价格指数(CPI)、人民币兑美元汇率(ER)和狭义的货币供应量( $M_1$ ),建立半参数可加回归模型。结果表明:NCR、SHI和CSI均有线性影响作用,其中SHI的影响表现为正效应,NCR、CSI表现为负效应;CPI对其具有波动性的非线性作用; $M_1$ 和ER对其具有交互作用,且在ER处于较低水平、 $M_1$ 处于中间水平时,余额宝收益率较低,ER与 $M_1$ 同高时,余额宝收益率较高,二者同低时,余额宝的收益率可达到极大值。

**关键词:**互联网金融;余额宝收益率;半参数回归;可加模型;交互作用

**中图分类号:**F832.2 **文献标识码:**A **文章编号:**1672-6049(2018)04-0062-10

## 一、引言与文献综述

从20世纪末开始,互联网信息技术呈现快速发展的势头,人们的思维方式和生活方式也随之发生了巨大改变。在金融领域,互联网技术和金融业务的结合产生了一种新型产品和业务模式——互联网金融。目前,互联网金融可大致分成如下四类:传统金融业务的互联网化、基于互联网平台开展的金融业务、全新的互联网金融模式以及金融支持的互联网化。

由于近几年互联网金融发展风头正劲,大众理财思维已由传统产品转向购买互联网平台的理财产品,其中2013年6月基于支付宝平台的天弘基金余额宝的出现,被认为开创了我国互联网平台理财元年<sup>[1]</sup>。因其具有低风险、高流动性、1元起投且随买随取的特征,余额宝快速成为了当下互联网金融投资产品的代表。随着余额宝的成功,越来越多的宝宝类产品也应运而生,如腾讯理财通的理财宝,小米旗下的活期宝等。这些产品逐渐改变了大众的理财观念,也加剧了理财类产品的竞争。

国外并没有明确定义互联网金融的概念,相关理论研究集中在互联网银行、P2P等互联网金融模式上。Cambra *et al.*<sup>[2]</sup>应用部分最小二乘法,发现社会效应和感知风险是互联网银行的在线客户进行投资的影响因素,且感知风险促进了客户投资,而社会效应则减少客户投资。Huang<sup>[3]</sup>研究发现在线P2P借

收稿日期:2018-04-19;修回日期:2018-06-18

基金项目:国家社会科学基金项目(16BTJ030);江苏省研究生科研创新计划项目(KYCX17\_1206)

作者简介:孙春燕(1965—),女,湖北天门人,南京财经大学应用数学学院副教授,硕士生导师,研究方向为应用统计与经济建模;马馨悦(1995—),女,安徽亳州人,南京财经大学应用数学学院硕士研究生,研究方向为应用统计与经济建模;刁海涛(1994—),女,山西大同人,南京财经大学应用数学学院硕士研究生,研究方向为应用统计与经济建模;丰茂芳(1992—),女,山东临沂人,南京财经大学应用数学学院硕士研究生,研究方向为数理统计。

贷平台的迅速增长取决于三个关键因素: 互联网的深度渗透、大量资金供应和未满足的金融需求。

国内关于互联网金融的研究很多, 在研究方法上大致可分为定性研究和定量研究。在定性研究方面, 周宇<sup>[4]</sup>主要说明互联网金融的崛起代表了现代社会的金融变革, 其不仅颠覆了传统的金融支付方式和交易方式, 还挑战了传统金融业务、机构、模式和监管方式。刘冬<sup>[5]</sup>以互联网金融——余额宝为例, 对比“美国版余额宝”PayPal 货币市场基金, 详细阐明了余额宝今后的发展道路以及其可能面临的监管问题。廖愉平<sup>[6]</sup>强调要加强对互联网金融的代表产品诸如余额宝、P2P、第三方支付等的监管研究。在定量研究方面, 陈耀辉等<sup>[7]</sup>建立贝叶斯分位数回归模型, 得到余额宝收益变动会带来流动性风险、投资风险、市场风险。周亮等<sup>[8]</sup>发现余额宝收益率的自身滞后项, 回购利率、日内波动及长短期利差的滞后项均会影响其收益率。

综上, 互联网金融较传统金融而言, 其创新性能促进经济发展, 带动金融领域变革, 但其也会带来未知的金融风险。关于互联网金融的研究主要集中在以下几个方面:

(1) 已有研究多关注互联网金融这种全新的金融模式的特点, 大多从理论角度分析该模式产生的原因、运行机制、对传统金融行业的冲击及其带来的风险。鲜有文章对该模式进行系统全面的分析, 且有关互联网金融理财产品的定量研究也少之又少。

(2) 国内有关余额宝收益率的研究, 大都从两方面着手, 一方面关注其自身具有的滞后性, 建立相应的 VAR 模型等。另一方面从单个变量入手探索其对余额宝收益率的影响。缺少对余额宝收益率影响因素的归纳和整体把握。

(3) 关于互联网金融或者余额宝的研究得到的结论多依赖于许多不变性条件, 不具有一般性也不够全面。

针对上述已有研究存在的不足, 选取余额宝为研究对象, 其作为一种新型互联网货币基金的代表, 对它进行研究不仅能探索出这一类产品的特点, 还能认识该产品对传统金融行业的冲击。本文尝试用半参数可加模型, 从内部机理及外部环境两方面着手, 对余额宝收益率的影响因素进行分析, 以期丰富互联网金融的定量研究, 为互联网金融的风险测度及监管提供建议, 并为市场投资者在进行理财产品选择时提供参考。

## 二、模型形式设定与计量方法

### (一) 模型形式设定

#### 1. 变量选取的理论依据

影响余额宝收益率的经济指标有很多, 可从两个方面着手考虑, 一方面是余额宝个性化指标, 另一方面是外部环境的指标。

因余额宝的投资组合是其收益率的主要来源, 现结合天弘基金官网公布的 2018 年余额宝第一季度投资组合报告(表 1) 选取其个性化指标。

在传统的金融市场中, 理财产品一般遵循“二八定律”, 即 20% 的高净值客户贡献总收益的 80%, 其往往忽略微小客户的投

资效益。而对于互联网金融理财产品而言, 其主要投资者为微小客户, 投资金额低, 但流量大, 多具有“长尾效应”。余额宝作为互联网理财产品的开拓者, 其牢牢抓住了“长尾效应”的特点, 发现微小客户市场潜力, 扩大自己的客户规模, 从而获取更多收益。微小客户作为普通人进行投资时, 往往关注投资收益的稳定性, 长久性, 金融市场的巨大波动均可能引起普通客户的恐慌, 从而终止投资。在余

表 1 余额宝的资产组合情况

序号	项目	金额 (元)	占基金总资产 的比例(%)
1	固定投资收益	96 706 041 894. 52	5. 71
	其中: 债券	96 706 041 894. 52	5. 71
	资产支持证券	-	-
2	买入返售金融资产	560 798 787 353. 19	33. 08
3	银行存款和结算备付金合计	1034 458 870 659. 18	61. 03
4	其他资产	3 123 697 716. 31	0. 18
5	合计	1 659 087 397 623. 20	100. 00

余额宝的投资策略中,为考虑微小客户需求,其将投资重点放在银行存款和结算备付金上,该部分收益主要是基金公司跟银行签订的协议,银行按约定给予利息,具有高度稳定性。因此,从客户来源和投资策略两方面均可看出,银行存款和结算备付金占基金总资产的比例,即现金占净比(NCR)是余额宝收益率的重要内部影响因素,选其作为自变量具有必要性。

由表1可知,在余额宝的投资组合中,银行存款和结算备付金、返售金融资产及固定投资(债券),三部分投资所占比例高达99.82%,这说明余额宝收益基本来源于此。首先,就银行存款和结算备付金而言,其收益主要受上海银行间同业拆放利率(Shanghai Interbank Offered Rate,简称Shibor)影响。一方面,银行内部的定价定息主要参考Shibor,当Shibor上升时,说明银行缺乏资金,需调高利息来吸引更多资金注入,从而引起余额宝收益的增加。另一方面,Shibor是金融市场的基准利率,会造成银行和基金市场的利差使同业差价资金流入基金市场,基金供给增加,从而使余额宝收益率降低。其次,返售金融资产的本质是一种债券,可与固定投资(债券)合为一类,其收益均可能受到证券市场的影响。沪深300指数(CSI)能较为全面反映我国证券市场情况。CSI较高时说明证券市场走势较好,余额宝的资金规模中投资于债券的部分将会获得较高的收益。但普通投资者也可能更愿意将资金更多地投入传统证券市场,从而减少对余额宝的购买,进而影响余额宝收益率。综上,由余额宝投资组合选取Shibor和CSI作为余额宝收益率的个性化影响因素。

从外部环境来看,余额宝的本质是一种货币基金,其收益率会受到我国经济环境的影响。我国经济环境可通过通货膨胀指标、汇率和货币供应量来反映。

首先,通货膨胀的不确定性会使投资者面临更高的风险,对传统金融市场来说,当投资者意识到实际收益率会降低时,其投资的积极性也会随之降低,进而减少对传统理财产品的购买。部分投资者可能会转投风险较低的互联网金融产品,进而影响余额宝的收益率。结合已有研究,选取同比的消费者价格指数(CPI)作为通货膨胀率指标。其次,汇率的调整对整个社会的经济影响很大。当汇率下降,本国货币升值,会吸引国外资金流入,然后转换为基础货币增强我国货币市场流动性,从而对货币基金的收益率造成影响。康风华和王宏勇<sup>[9]</sup>的研究也佐证了汇率会对货币基金收益率产生影响这一观点。最后,央行通过调整货币供应量来影响国民经济运行。朱雁春等<sup>[10]</sup>指出当货币供应量充分时,会刺激生产,投资者对市场的预期会好转,从而加大自己闲置资金投资力度。但随着货币供应量增加,市场流动性会增强,各大基金公司和银行的供给充足,从而在供求关系下影响货币基金收益率。余额宝现已成为我国第一大货币基金,因此汇率和货币供应量也必将对其收益率产生影响。在众多汇率中,人民币兑美元汇率(ER)最具代表性。狭义的货币供应量( $M_1$ )是最能反映个体、团体、企业的货币政策指标,而余额宝的投资者中个体用户所占比例高达99%。综上,结合已有研究,选取CPI、ER和 $M_1$ 作为余额宝收益率外部影响因素。

现金占净比(NCR)、上海银行间同业拆放利率(Shibor)、沪深300指数(CSI)、消费者价格指数(CPI)、汇率(ER)及狭义的货币供应量( $M_1$ )这六个指标的选取,能较为全面反映余额宝内部和外部的影响因素,各因素对收益率的具体影响还需建模探讨。

根据上述理论分析及研究目的,选取2013年10月到2018年3月的各变量数据进行研究,其中余额宝七日年化收益率、一个月品种的Shibor报价、ER、CSI均为日度数据,CPI选取同比月度数据,狭义货币供应量 $M_1$ 由央行每月定时发布,NCR由余额宝季度报告公示。为统一变量单位,对日度数据计算月平均,季度数据进行插值转化为月度数据,并去掉部分数据带有的百分号,对各数据取对数建模。编程选取R软件。各变量的命名及来源见表2。

## 2. 模型形式的选择

上述分析为模型中的变量选择提供了一定的理论依

表2 变量说明

变量(对数处理)	符号	数据来源
余额宝收益率	$LYEB$	天天基金网
现金占净比	$INCR$	
沪深300指数收益率	$ICSI$	国家统计局
消费者价格指数	$ICPI$	
货币供应量	$IM_1$	中国人民银行
汇率	$IER$	
Shibor	$ISHI$	

据,但对于各变量关系的确切形式及作用机制,还需结合样本数据的特征进一步选择合适的模型来刻画。本文借助回归模型来探讨各因素对余额宝收益率的影响。在回归模型中,由于变量的具体影响形式表现为线性和非线性,其中在非线性形式中,如果可转化为线性形式仍可归为线性回归,但对于难以确定的非线性关系,可引入非参数回归模型。若自变量对因变量的影响既有线性又有不确定的非线性,则可建立半参数可加模型。现介绍如下三种模型:

(1) 模型 I——线型回归模型

$$y = \beta_0 + \beta_1 \times INCR + \beta_2 \times LSHI + \beta_3 \times ICSI + \beta_4 \times ICPI + \beta_5 \times LM_1 + \beta_6 \times IER + \varepsilon \quad (1)$$

其中  $\varepsilon$  为随机误差项,需满足零均值独立同分布的假定,  $\beta_0$  为常数项,  $\beta_1, \beta_2, \dots, \beta_6$  为模型待估参数。

(2) 模型 II——非参数回归模型

非参数回归模型可反映自变量对因变量的非线性影响。本文主要使用具有可加效应的广义可加非参数回归模型,将其定义为模型 II,表达式为:

$$LYEB = f_1(INCR) + f_2(LSHI) + f_3(ICSI) + f_4(ICPI) + f_5(LM_1) + f_6(IER) + \varepsilon \quad (2)$$

这种非参数回归模型与线型回归模型形式类似,模型的估计仅考虑样本数据,避免了线性模型可能产生的模型误设问题。但其可能将线性项作为非线性项处理,从而无法揭露某变量对因变量的真实影响。

(3) 模型 III——半参数可加回归模型

为了克服线性模型及非参数模型的缺点,本文将线性模型与广义可加模型相结合,建立半参数可加回归模型<sup>[11]</sup>。此外,半参数可加模型还允许非参数部分具有一些交互项,更好的解释自变量的影响。交互项包括参数部分和参数部分的交互、参数与非参数部分的交互及非参数部分与非参数部分的交互。

半参数可加模型较于一般的参数回归模型具有如下三个方面的优点:一是适用范围更广,在不易判断因变量的分布时仍然适用;二是可以有效避免数据维数较高时造成的“维数灾难”;三是能具体探索出自变量对因变量的线性或非线性影响。具体而言,半参数可加模型一般描述的是较长时期内时间序列在动态变化中的相对波动关系。其中非线性关系是对长期变动过程中的短期波动特征做出显性拟合,而线性关系描述了长期总体上可能存在的线性关系特征<sup>[12]</sup>。

在进行回归模型之前,需要对因变量进行正态性检验来判断其是否满足半参数可加模型的条件,若因变量不能通过正态性检验,建立的半参数可加模型才有意义。

观察数据是否为正态分布,最直观的方法就是画 QQ 图,若样本服从正态分布,则图中散点应该集中在 45° 线附近。此外,将数据分布图与正态分布图作对比也是检验是否服从正态分布的一个方法,本文将两种方法结合对因变量数据进行正态性检验(见图 1、图 2)。

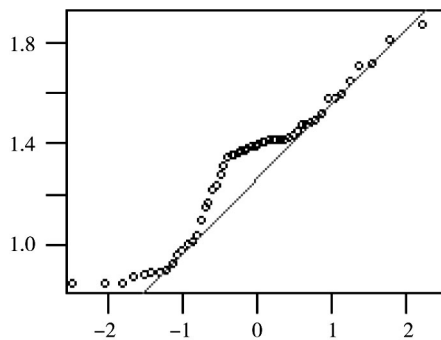


图 1 余额宝收益率 QQ 图

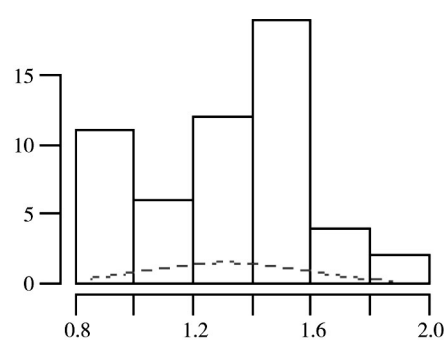


图 2 余额宝收益率直方图和分布图

由图 1 可以看出,样本的散点图明显偏离直线,图 2 可以看出收益率分布直方图也明显不符合其相应的正态分布,两图均表明余额宝收益率不服从正态分布。这也证明了半参数可加模型对余额宝

收益率影响因素的适用性。

实际回归建模时,可能会碰到多重共线性或多重共曲线性问题,半参数可加模型允许自变量建立交互项,不仅能有效解决共线性或共曲线性的影响,还能避免剔除变量引起的模型解释能力下降。根据相关系数矩阵(R软件 cor 函数可直接得出)和共曲线性相关矩阵(R软件 concavity 函数可直接得出)可判断各自变量有无可能存在多重共线性或共曲线性<sup>[13]</sup>。

通过线性相关系数,共曲线相关系数及各变量分别对余额宝收益率的影响趋势图共同来确定模型中线性项的自变量,非线性项的自变量,及可能具有交互作用的自变量。

先计算各变量间的相关系数矩阵  $\rho_{linear}$ :

$$\rho_{linear} = \begin{matrix} IYEB \\ INCR \\ LSHI \\ ICSI \\ ICPI \\ LM_1 \\ IER \end{matrix} \begin{pmatrix} 1 & 0.38 & 0.89 & 0.12 & -0.05 & -0.42 & -0.57 \\ & 1 & 0.36 & -0.00 & 0.00 & -0.66 & -0.46 \\ & & 1 & 0.19 & -0.02 & -0.27 & -0.40 \\ & & & 1 & -0.13 & 0.00 & -0.02 \\ & & & & 1 & -0.13 & -0.05 \\ & & & & & 1 & 0.85 \\ & & & & & & 1 \end{pmatrix}$$

由相关系数矩阵可知,其中  $\rho(IYEB, LSHI) = 0.89$ ,说明 IYEB 与 LSHI 存在着较强的线性关系,可建立 LSHI 为线性项。 $\rho(LM_1, IER) = 0.85$ ,说明二者也具有很强的线性关系,可以考虑建立  $LM_1 \times IER$  的线性交互项。

各变量得出的共曲线性  $\rho_{cur}$  矩阵为:

$$\rho_{cur} = \begin{matrix} INCR \\ LSHI \\ ICSI \\ ICPI \\ LM_1 \\ IER \end{matrix} \begin{pmatrix} 1 & 0.25 & 0.18 & 0.07 & 0.12 & 0.48 \\ & 1 & 0.12 & 0.26 & 0.26 & 0.23 \\ & & 1 & 0.30 & 0.32 & 0.14 \\ & & & 1 & 0.17 & 0.13 \\ & & & & 1 & 0.74 \\ & & & & & 1 \end{pmatrix}$$

由共曲线性矩阵可以看出,最可能存在共曲线性的变量为  $LM_1, IER$ 。结合  $\rho_{linear}$ ,建模时可以先尝试  $LM_1, IER$  的非线性交互项  $f_i(LM_1)f_j(IER), i \neq j$ 。

综上,Shibor 为线性项, INCR、CSI 及 CPI 可先建立非参数项,  $M_1$  及 ER 的共线性、共曲线性均显著,可先建立非参数交互项。因此模型 III 形式可暂定为:

$$IYEB = \beta_0 + \beta_1 \times LSHI + f_1(INCR) + f_2(ICSI) + f_3(ICPI) + f_4(LM_1) \times f_5(IER) + \varepsilon \quad (3)$$

## (二) 计量方法

对模型 III 半参数可加模型求解时,会综合涉及模型 I、II 求解所用计量方法。对线性项求解时使用普通最小二乘估计,对非线性项使用样条估计。因最小二乘估计已经成熟并被应用广泛,本文在此将重点对样条估计进行说明,其基本思想为:通过使用样条在节点内拟合一个方程,再在节点处将这些逐段回归所得的曲线连接起来。对于出现过度拟合的曲线,可通过光滑样条来减轻。

### 1. 光滑样条及估计方法

以一元非参数回归  $y = f(x) + \varepsilon$  为例说明光滑样条及其估计方法。用  $x_i$  表示样本观测值。在  $x_i$  点对应的观测值为  $y_i$ ,模型的拟合值为  $\hat{y}_i$ ,则样条估计与 OLS 估计类似,都是使模型的残差平方和  $Q(f)$  最小,即满足

$$\min Q(f) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n [y_i - f(x_i)]^2 \quad (4)$$

若仅关注上式,会导致样条对应的节点数增加,从而增加模型的变量。为了权衡模型的拟合结果

和变量个数,现对上式施加一个惩罚项,使光滑样条估计的判断准则变为:

$$\min Q(f) = \sum_{i=1}^n [y_i - f(x_i)]^2 + \lambda \int [f''(x)]^2 dx \quad (5)$$

其中  $\lambda \int [f''(x)]^2 dx$  称为粗糙惩罚,  $f''(x)$  是回归方程  $f(x)$  的二阶导数,其现实意义是曲线  $f(x)$  斜率的变化率,  $f''(x)$  越大,曲线  $f(x)$  波动越大,越粗糙。引入惩罚项的目的是为了兼顾模型拟合优度与光滑程度。 $\lambda$  为光滑参数,取值非负,用来控制回归曲线的光滑程度。

### 2. 光滑参数的选取

在确定了某一变量对被解释变量的影响是非线性的情况下,往往通过图像来反映,而做出的影响曲线光滑程度取决于光滑参数  $\lambda$ 。若所得曲线较为粗糙,则所研究内容也不具有可靠性,因此光滑参数  $\lambda$  的选取对光滑样条估计起着决定性的作用。在实际的建模时,可以通过自动光滑技术进行  $\lambda$  的选取。常用的自动选取  $\lambda$  的思想有交叉验证法或者广义交叉验证法。

交叉验证法(CV)是求满足交叉验证得分最小的  $\lambda$ ,即:

$$CV(\lambda) = n^{-1} \sum_{i=1}^n [y_i - \hat{f}_i(x_i; \lambda)]^2 \quad (6)$$

对样条回归模型来说,CV方法的计算量太大。因此将其转化为能够降低运算量的广义交叉验证法(GCV),其思想为最小化:

$$GCV(\lambda) = \frac{RSS}{(1 - n^{-1}df)^2} = \frac{\sum_{i=1}^n [y_i - \hat{f}_i(x_i; \lambda)]^2}{(1 - n^{-1}df)^2} \quad (7)$$

其中  $df$  表示光滑样条的自由度,近似等于光滑样条回归中的参数个数。在实际操作中,往往控制  $df$  而不是  $\lambda$ 。

### 三、模型计量结果及分析

在模型 III 中,  $LNCR$ 、 $ICSI$ 、 $ICPI$  的非参数影响未知,交互项  $M1$  及  $ER$  的形式未定。现对模型 I 进行求解,探索各变量是否为线性项,结果见表 3。

由表 3 可以看出,  $LSHI$  和  $IER$  均能通过回归系数显著性检验,说明对二者可建立线性项。对未能通过线性模型显著性检验的  $LNCR$ 、 $ICSI$  及  $LM_1$  建立表达式(1)的模型进行求解,探索出各变量的具体影响形式。

$$LYEB = \beta_0 + \beta_1 \times LSHI + f_1(LNCR) + f_2(ICSI) + f_3(ICPI) + f_4(IER) + \varepsilon \quad (8)$$

式(8)非参数估计部分结果见表 4。

由表 4 可知,  $ICPI$  和  $LM_1$  的非线性效应显著,可建立非线性项。但  $LNCR$ 、 $ICSI$ 、 $IER$  的非参数效应不明显。结合表 3,  $LNCR$ 、 $ICSI$  的线性非线性效应均不显著,为保证对余额宝影响因素探究的全面性,半参数可加模型建立的过程中可以不剔除两个变量,将二者仍作为线性项处理。由于  $LM_1$  的非线性作用较  $IER$  的线性作用更显著,二者具有交互作用,因此建立二者的非线性交互项。综上,最终半参数可加模型形式可确定为:

$$LYEB = \beta_0 + \beta_1 \times LNCR + \beta_2 \times LSHI + \beta_3 \times ICSI + f_1(ICPI) + f_2(IER) + f_3(IER) + \varepsilon \quad (9)$$

表 3 模型 I 的回归分析结果

变量	参数估计	标准误	t 值	p 值
截距	15.532	14.338	1.083	0.284
LNCR	-0.047	0.151	-0.313	0.756
LSHI	0.883	0.075	11.755	1.35E-15***
ICSI	0.037	0.043	-0.849	0.400
ICPI	-2.596	2.971	-0.874	0.387
LM <sub>1</sub>	0.001	0.200	0.006	0.995*
IER	-1.632	0.757	-2.156	0.036

表 4 模型 III 的非参数回归部分结果

变量	估计自由度	参考自由度	f 值	p 值
$f_1(LNCR)$	1	1	1.315	0.260
$f_2(ICSI)$	4.615	5.576	1.706	0.180
$f_1(ICPI)$	7.268	8.23	4.547	0.001***
$f_1(LM_1)$	7.682	8.484	12.344	4.71E-10***

对式(9)进行求解,其参数项的估计结果见表5,非参数估计结果见表6。

由表5可得  $INCR$  的回归系数  $\beta_1 = -0.082$ , 说明现金净占比每增加一个单位会造成余额宝收益率下降约 0.082 个单位。即余额宝稳定投资的增加会造成其收益率的降低。 $ISHI$  的回归系数  $\beta_2 = -0.373$ , 说明市场基准利率每增加一个单位,大约增加 0.358 单位的余额宝收益率。 $Shibor$  的增加说明银行给予协议存款和结算备付金的利息会增加,从而直接影响余额宝收益率。回归系数  $\beta_3 = -0.021$ , 说明  $ICSI$  每增加一个单位,会使余额宝收益率降低约 0.021 个单位。

沪深 300 指数反映证券市场情况,回归结果表明证券市场走势较好,会造成余额宝的收益率下降。

由表6可知,各  $f_1(ICPI)$  及交互项  $f_2(IM_1)f_3(IER)$  非参数估计结果均显著,对应的  $P$  值表示平滑函数是否显著的减少了模型误差。 $ICPI$  对  $IYEB$  的非线性影响见图3。 $IM_1, IER$  对  $IYEB$  的非线性交互影响见图4。

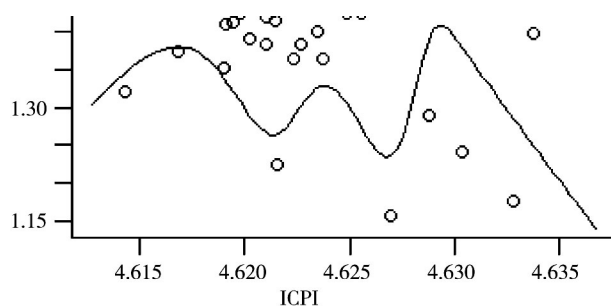


图3  $ICPI$  对  $IYEB$  的非线性影响

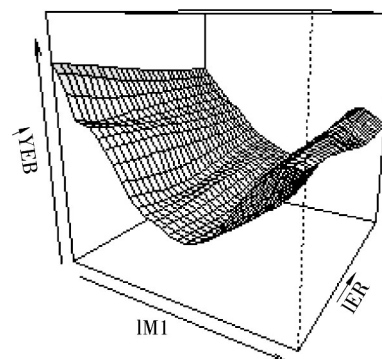


图4  $IM, IER$  对  $IYEB$  的非线性交互影响

由图3看出在其他条件不变的情况下  $ICPI$  对余额宝收益率的影响波动程度很大,明显拒绝线性关系可能,但  $ICPI$  的变动,会将相应的余额宝收益率大致限制在范围  $[1.2, 1.4]$  间,当  $ICPI > 4.633$  左右时,余额宝的收益率才大幅跳动到区间外,在  $ICPI \approx 4.630$  时,余额宝收益率可能达到极大值。

由图4可以看出  $IM_1, IER$  交互项对  $IYEB$  的交互项作用明显,且在汇率处于较低水平,货币供应量处于中间水平时,会造成余额宝收益率较低,汇率与货币供应量均较低或较高时,余额宝收益率才比较高,二者同时低时可以让余额宝的收益率达到极大值。

综上,本文所建立的最终模型式(9)的求解结果为:

$$IYEB = 1.265 - 0.082 \times INCR + 0.373 \times ISHI - 0.021 \times ICSI + f_1(ICPI) + f_2(IER) + f_3(IER) + \varepsilon \quad (10)$$

下面对模型结果进行评价,通过拟合优度(可决系数)及对应残差项是否符合正态分布来判断所建模型拟合的好坏。其中式(10)的拟合优度为  $R^2 = 0.967$ ,残差检验结果见图5。

图5反映出残差检验的结果,其中左上图反映的是残差的理论分位数与实际分位数,若其均分布在实线附近,则说明残差服从正态分布;右上图表示的残差和预测值之间的散点图,若模型建立的较好,二者应无明显关系,即右上图需呈现凌乱分布;左下图是残差分布的直方图,该部分能粗略看出分布是否为正态分布;右下图是拟合值与真实值得散点图,当模型建立较好时,显然二者需要无限接近,

即近似呈 45 度直线。由图 5 左可以看出残差近似服从均值为 0 的正态分布,图 5 右可以看出模型的拟合值与真实值呈现一条 45 度直线,说明拟合值与真实值二者高度相等。综上所述,所建立的余额宝收益率影响因素的半参数可加模型(式(10))不仅拟合优度很好,还能通过残差检验,因此该模型具有较强的适用性,可以较好的反映自变量对余额宝收益率的影响。

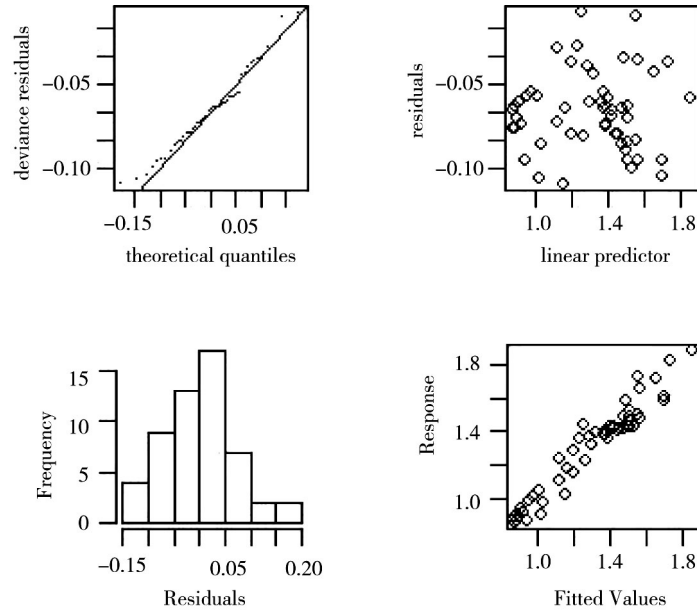


图 5 式(10)残差分析结果

#### 四、结论与启示

本文基于半参数建模的思想,对余额宝七日年化收益率的影响因素建立半参数可加模型并进行了分析研究。具体结论有:现金占净比、上海银行间同业拆借利率和沪深 300 指数收益率对余额宝收益率的影响为线性效应,银行间同业拆借利率的影响相对较大且为正效应,现金占比的线性影响较小且为负效应,沪深 300 指数的影响最小为负效应;消费者价格指数对收益率的影响呈波动性的非线性效应,货币供应量和汇率对余额宝收益率存在非参数交互作用,且二者同时低时,余额宝的收益率可达到极大值。

现金占净比表示的大小反映余额宝的稳定投资所占比例。由回归结果分析可知,该因素对余额宝收益率产生负面影响,说明该部分的投资越高,余额宝收益率越低,产生这一影响的原因可能因为对投资产品来说,收益与风险常呈现正相关,一般风险越低收益越低,协议存款这种稳定的投资,风险较小,所以产生的收益较少,从而降低余额宝收益。由余额宝的季度报告显示,现金占净比由最初的将近 90% 已经下降到 2018 年一季度的 61.03%,这说明天弘基金公司可能发现这一规律,为获得更大的收益,将余额宝资金池更多的投向债券、固定资产等。

银行基准利率 Shibor 通过银行协议存款和结算备付金直接影响余额宝收益率。余额宝的季度报告显示其主要投资收益来源于此。回归结果显示 Shibor 越高,余额宝投资收益越高,从而用户投资收益率越高。Shibor 可能通过余额宝的投资策略增加其收益率,也有可能通过对传统金融市场的影响降低其收益率,回归系数为 0.373,说明 Shibor 对余额宝收益率的正向效应大于负面冲击,且回归系数不高也暗示了负面冲击的存在。

沪深 300 指数是证券市场的代表性指数,余额宝的部分收益来源于证券投资,从这一方面来说,当指数增加时,余额宝收益率会随之增加,表现为正效应。但对于整个市场来说,余额宝只是一种新兴的货币基金,沪深 300 走高同时意味着传统的金融市场形势较好,部分投资者会转向投资收益更高的股票或者债券,带走部分余额宝投资,从而降低余额宝的收益率,表现为负效应。通过  $\beta_2 = -0.021$



$<0$  可看出,沪深 300 指数在研究期间内,会降低余额宝的收益率,说明负效应的影响更大,但  $\beta_2$  数值较小也侧面证明了正效应的存在。这同时能够说明传统金融市场和互联网金融市场,都不能忽略长尾客户产生的影响。

消费者价格指数反映通货膨胀,其对余额宝的收益率影响呈现波动趋势也与现实情况相符,一方面,国内通货膨胀率高,部分投资者会对理财产品有相对较低的预期收益,从而影响其投资,余额宝作为互联网金融理财产品,资金量也会受到相应的影响,从而影响资金收益,另一方面,一部分投资者只会回避高风险的理财投资产品,转而投向风险较小的稳值收益的余额宝,从而使余额宝资产增加,余额宝的投资收益也会增加,因此 CPI 对余额宝收益的影响表现的是一种波动趋势。

汇率与货币供应量的协同作用表现在,当二者同高时,货币市场流动性大,会带动余额宝收益会增加,二者同低时,投资者对市场的预期不会更坏,从而加大对理财产品的投资,因此这两方面无论从余额宝自身收益来源还是资金规模加大来说,都会对余额宝收益率产生正向影响。

基于上述分析,本文对余额宝未来的发展以及投资者的选择提出以下建议:

第一,减少对协议存款的过度依赖。由于余额宝对其主要收益来源银行协议存款的依赖性很强,且由本文结论可知,Shibor 对余额宝收益率具有较大的影响,而 Shibor 的变动与银行协议存款直接相关,因此为了预防银行协议存款可能带来的风险,余额宝应权衡各项投资组合的分配管理,拓宽投资渠道,降低对协议存款的投资比例,减少对其依赖性,引入新的投资理念,保证自己的可持续发展。

第二,加强流动性风险管理意识,建立长效监管机制。余额宝虽然推动了国内利率市场化进程,但其自身会受到利率市场化的影响。随着利率市场化进程的加速,余额宝的套利空间越来越小,相关部门的监管程度越来越大,余额宝已经很难维持原有的高水平收益。建议余额宝在政策允许范围内,用好备付金资源,可以将平台开放给银行,让银行来做风险控制,寻找之间的利益平衡点,促进发展制度化和规范化,实现与银行的互补,加强流动性风险防范能力。

第三,重视自身的长期投资收益率<sup>[14]</sup>。余额宝一直以来以其稳定的较高收益吸引着众多投资者,但就当前互联网理财产品市场来说,余额宝并不是收益率最高的理财产品,再加上各大银行相继推出了回报率较高的理财产品,一旦这些产品能够灵活便捷的使用,用户就会将资金从余额宝转投到其他理财产品。因此余额宝应在采取一定措施保证其收益优势的同时,要坚持“以客户为中心”,开发新的产品,不断增加客户依赖性,提高竞争力,从而扩大发展。

第四,对于投资者,要充分认识基金投资存在的风险,清楚收益来源,避免盲目将货币基金预期收益率与储蓄利率相比较,明确认识“余额宝”不同于银行存款,尤其是其安全性,要根据自己的实际情况选择合适的金融理财产品,理性投资,避免盲目跟风。

综上所述,本文建立的关于余额宝收益率研究的半参数可加模型能够很好的探索出代表宏观环境的汇率(ER)和消费者价格指数(CPI)、代表市场环境的银行间市场基准利率(Shibor)、货币供应量(M1)以及沪深 300 指数(CSI)对余额宝收益率的影响,不仅丰富了对余额宝收益率的量化研究,还能对研究其他互联网金融产品提供参考,同时为市场各参与者提供良好的投资决策参考。

#### 参考文献:

- [1]庄雷.余额宝与国债市场收益率波动的实证研究[J].经济与管理,2015(3):74-79.
- [2]CAMBRA FIERRO J, MELERO I, SESE F J. Online customer-initiated contacts and the development of profitable relationships[J]. Electronic commerce research & applications, 2017, 26(4): 13-22.
- [3]HUANG R H. Online P2P lending and regulatory responses in China: opportunities and challenges[J]. European business organization law review, 2018, 19(1): 63-92.
- [4]周宇.互联网金融:一场划时代的金融变革[J].探索与争鸣,2013(9):67-71.
- [5]刘冬.互联网金融新模式之余余额宝的发展与监管[D].上海:华东政法大学,2014.

- [6] 廖榆平. 我国互联网金融发展及其风险监管研究—以 P2P 平台、余额宝、第三方支付为例 [J]. 经济与管理, 2015 (2): 51-57.
- [7] 陈耀辉, 白兰. 基于贝叶斯分位数回归模型的余额宝收益风险分析 [J]. 长江大学学报(社科版), 2017(3): 82-90.
- [8] 周亮, 袁永智. 余额宝收益率、回购利率与资本市场 [J]. 金融理论与实践, 2017(10): 81-86.
- [9] 康凤华, 王宏勇. 复旦人民币汇率指数的分形分析与市场风险测度 [J]. 南京财经大学学报, 2016(1): 57-63.
- [10] 朱雁春, 苏文兵, 王亚星. 货币政策、企业资本结构变动及经济后果—基于三大货币政策区间的探讨 [J]. 山西财经大学学报, 2015(3): 34-47.
- [11] KEELE L. Semiparametric regression for the social sciences [M]. America: John Wiley and Sons, 2008: 151-186.
- [12] 逯进, 苏妍. 人力资本、经济增长与区域经济发展差异—基于半参数可加模型的实证研究 [J]. 人口学刊, 2017 (1): 89-101.
- [13] 向伟, 宁魏青. 广义可加模型在出生缺陷影响因素分析中的应用及 R 语言实现过程 [J]. 中国妇幼保健, 2014 (29): 4711-4715.
- [14] 陈耀辉, 张军. 企业脆弱性测度方法及应用研究—基于我国上市房地产业数据 [J]. 南京财经大学学报, 2017(2): 71-81.

(责任编辑: 黄明晴; 英文校对: 陈芙蓉)

## Analysis of Influencing Factors of Yu'E Bao Yield based on Semi-Parametric Additive Model

SUN Chunyan, MA Xinyue, DIAO Haitao, FENG Maofang

(School of Applied Mathematics, Nanjing University of Finance and Economics, Nanjing 210023, China)

**Abstract:** The monthly average data of the seven day yield of the Yu'E Bao from October 2013 to March 2018 are selected for the research. There are two perspectives to begin the study. One is to select Net cash ratio (NCR) as the Yu'E Bao inner index from the investment portfolio of the fund. At the same time, Shanghai Interbank Offered Rate (SHI) and the Shanghai & Shenzhen index (CSI) are chosen as the personalized index from the operation mechanism. And the other is the external environment, including the consumer price index (CPI), the RMB/US dollar exchange rate (ER) and the narrow monetary supply (M1). Then the semi-parametric additive model is established and the result is obtained that among the factors affecting the yield of the Yu'E Bao. It shows that NCR, SHI and CSI have the linear effect. The effect of SHI is positive to Yu'E Bao, while the effects of the other two are negative. And the nonlinear effect of CPI is fluctuating. Also, the interaction between M1 and ER is that when ER is at a low level and M1 is at the middle level, the yield of Yu'E Bao is relatively low, when ER and M1 are both at high level, the yield of the Yu'E Bao is higher, when the two are both at low level, the yield of the Yu'E Bao can reach the maximum value.

**Key words:** internet finance; the yield of Yu'E Bao; semi-parametric regression; additive model; interaction