基金经理能够战胜市场吗?

——基于双重自举抽样法的证据

廖长友[1],李玉敏[2],刘星意[3]

(1. 西华大学 管理学院,四川 成都 610039;2. 西华大学 国际经济与管理研究院,四川 成都 610039;
3. 北京大学 软件与微电子学院,北京 102600)

摘要:在评价基金经理的投资能力时,自举抽样法能够有效控制运气因素对基金业绩的影响,因而得到广泛运用。自举抽样有不同的实现方法,基于双重自举抽样法,计算在假设检验中不同自举抽样实现方法的两类错误,并确定更为可靠的自举抽样方法,在此基础上检验国内市场基金经理是否能够战胜市场。研究发现,按时间对所有基金的回归残差序列同时自举抽样并且同时不对因子收益联合自举抽样的实现方法是更为可靠的自举抽样方法。进一步发现,在 2011—2015 年间,基金经理没有获得超过市场基准的超额收益,没有证据表明基金经理战胜了市场;而在 2016—2020 年间,基金经理整体上获得了显著的超额收益,基金经理战胜了市场。

关键词:基金经理;双重自举抽样法;市场有效性;投资能力

中图分类号:F830.91　　文献标志码:A　　文章编号:1672-6049(2022)05-0044-11

一、引言

巴菲特曾在 1984 年举例说明运气在投资业绩中可能发挥的作用。假设当时全美国 2.7 亿人都参加猜扔硬币游戏的赌博,游戏赌注为 1 美元,猜错的人输 1 美元并出局,猜对的人赢得 2 美元并在下一轮继续押注所有赌资,如果又猜对则赌资翻倍,以此类推。显然,每次赌博猜对的概率是 50%,到第 10 轮,全美约有 26 万人全部猜对,到第 20 轮,全美约有 250 人全部猜对,每人将会获得1 048 576美元。那么,我们能够相信这 250 人具有猜硬币的神力吗?显然不能。可以证明,对于一个特定的个人,参加 20 轮的猜硬币赌博,连续猜对 20 轮的可能性微乎其微。然而,当有足够数量的人一起参加猜硬币赌博,则出现全部都猜对的人几乎是必然发生的事件。这个例子告诉我们,仅凭运气而非能力的投资者也有可能取得亮眼的业绩①。

截至 2021 年 5 月,国内基金市场共有公募基金 8 061 只,私募基金有 24 427 只②。市场上每年都会出现取得骄人业绩的基金。例如,在 2020 年华夏成长基金取得了 336% 的净值收益率,而同期上证

收稿日期:2022-02-12;修回日期:2022-07-27

基金项目:国家自然科学基金青年项目"中国企业女性领导多重身份双刃剑效应及身份管理策略实证研究"(72102186)

作者简介:廖长友(1970—　),男,重庆永川人,经济学博士,西华大学管理学院副教授,研究方向为产业组织与证券投资基金;李玉敏(1995—　),女,河南灵宝人,西华大学国际经济与管理研究院硕士研究生,研究方向为资产定价;刘星意(1998—　),女,四川成都人,北京大学软件与微电子学院硕士研究生,研究方向为金融科技。

①该例由巴菲特于 1984 年在《聪明的投资者》出版发行 30 周年庆祝大会中提出。

②数据来自中国证券投资基金业协会网站。

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

综合指数仅上涨了 14%。这个明星基金的存在意味着基金经理战胜了市场吗？由于不能确定明星基金经理是凭自己的投资能力还是凭运气取得超额业绩，该答案并不确定。因此，当我们判断基金经理是否能够战胜市场时，必须考虑运气的因素。

对基金的超额收益，即运用定价因子模型估计基金的 alpha，进行显著性假设检验，是判断基金经理是否具有投资能力的常见做法。然而，由于存在样本随机性以及 alpha 一般不服从正态分布，在假设检验中研究者有可能犯两类错误。第一类错误是，当原假设为真时，错误地拒绝原假设，即将没有投资能力的基金经理判断为有投资能力；第二类错误是，当备择假设为真时，错误地拒绝备择假设，将原本有投资能力的基金经理判断为没有投资能力。当对多达几千只基金的 alpha 做假设检验时，由于很难获得 alpha 的联合概率分布，通过参数建模判断是否存在具有投资能力的基金经理就变得异常困难。自举抽样法能够在统计量的概率分布未知的条件下对总体参数进行估计与推断，能够有效控制运气因素，因而是一种非常有吸引力的分析基金经理投资能力的方法[1]。

Kosowski *et al.*[2] 和 Fama and French[3] 是将自举抽样法用于分析基金经理投资能力的早期最重要的研究。他们试图通过假设检验回答这样一个问题：市场上存在具有投资能力的基金经理吗？其原假设和备择假设分别是：

原假设：$\alpha_1 = 0$，$\alpha_2 = 0$，…，$\alpha_N = 0$。

备择假设：至少有一只基金的 α 不等于 0。

不同的自举抽样法[2-3] 通过构造分位数统计量实现对假设的检验，但得到的结论大相径庭。Kosowski *et al.*[2] 表明市场上存在具有投资能力的基金经理；而 Fama and French[3] 却没有发现基金经理能够取得显著的超额收益，即基金经理整体上不具有投资能力。是什么原因导致二者的结论[2-3] 存在巨大差异呢？哪一个结论更为可靠呢？在上述假设检验的具体环节中，Kosowski *et al.*[2] 和 Fama and French[3] 只有抽样方法存在差异[4]。显然，自举抽样实施方法不同导致了不同的结论。因此，自举抽样方法的可靠性决定了两者结论的可靠性。

在假设检验中，研究者不可避免地会犯两种错误，即错误地拒绝原假设的第一类错误和错误地拒绝备择假设的第二类错误。假设检验中的两类错误是判断假设检验结果可靠性最重要的依据。在假设检验中，如果一种自举抽样方法能够导致在原假设为真时拒绝原假设的概率接近于 100%，或者在备择假设为真时拒绝原假设的概率接近于 0，则这种自举抽样方法没有可靠性。只有在检验的过程中实施可靠的自举抽样方法，才能够对结果做出正确的判断。

然而，如何计算假设检验中通过实施不同抽样方法所导致的两类错误的发生概率呢①？Harvey and Liu[5] 提出了双重自举抽样方法（double bootstrap）。这种方法为研究者计算多重假设检验中两类错误发生的概率提供了解决办法。本文运用 Harvey and Liu[5] 的双重自举抽样方法，基于中国基金市场的数据，在计算 KTWW 方法[2] 和 FF 方法[3] 的两类错误发生概率的基础上，筛选更为可靠的自举抽样实施方法，进而用于检验基金经理是否具有投资能力。本文的贡献主要体现在以下两个方面：

第一，运用双重自举抽样法对比分析了 KTWW 方法和 FF 方法的两类错误，为二者的可靠性提供了证据。由于二者结论截然不同，研究者从数据样本、抽样方法以及推断方法等不同的角度考察了产生差异的原因[4-5]，Huang *et al.*[6] 则从理论上检查了 KTWW 方法的基本型和 FF 方法的基本型的第一类错误和检验功效。本文运用 Harvey and Liu[5] 的双重自举抽样方法考察了 KTWW 方法和 FF 方法的两类错误。研究发现，在假设检验中，FF 方法的扩展 I 型和扩展 II 型是更为可靠的自举抽样方法。本文的研究为 Huang *et al.*[6] 的研究提供了经验证据，也为正确认识 KTWW 方法和 FF 方法结论的差异性提供了新的证据。

第二，运用 FF 方法的扩展 I 型和扩展 II 型这两种与国内现有相关研究完全不同的方法，重新检

①在不引起误解的情况下，为行文简洁，本文提到“两类错误的发生概率”有时仅表达为“两类错误”。

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

查了 2011—2020 年间不同时间段内国内基金经理的投资能力。国内现有的关于基金经理是否具有投资能力的研究采用的方法各异，结论也各不相同。在 2011—2015 年间，没有证据表明市场上的基金经理具有投资能力；而在 2016—2020 年间，基金经理表现出了显著的投资能力，基金经理战胜了市场。本文为研究国内基金经理是否具有投资能力提供了新的方法和证据。

自有效市场理论提出以来，学者们运用不同的方法对市场是否有效进行了各种检验。基金经理是否具有投资能力，是否能够获得超额收益，能够为检验市场是否有效提供直接证据，是学术界经久不衰的重要话题。本文的研究提供了关于基金经理投资能力的证据，对深入认识中国金融市场的有效性具有重要意义。

二、文献回顾

自 Jensen [7] 以来，大量的学者对基金经理投资能力做了深入的研究 [8-13]。Kosowski *et al.* [2] 基于对每只基金的回归残差以及因子收益率进行独立有放回的重复自举抽样，并通过构建基金 alpha 或其 t 值分位数的统计分布进行检验，结论拒绝原假设，发现市场上确实存在少量具有投资能力的基金。随后，运用 KTWW 方法，Kosowski *et al.* [14] 发现，排名靠前的对冲基金的业绩并不能由运气因素解释，市场上确实存在部分具有投资能力的基金经理。Cuthbertson *et al.* [15] 研究了 935 只英国的单位信托基金和开放式基金的业绩，也发现少量基金具有真实而非运气导致的投资能力。

Fama and French [3] 在相同的时间点上对所有基金的回归残差以及因子收益率同时进行有放回的重复自举抽样。FF 方法 [3] 没有拒绝原假设，表明市场上基金未能取得超额收益，基金经理没有战胜市场。运用 FF 方法 [3]，Yan and Zheng [16] 研究美国市场上 18 113 个异象投资策略，结果发现约有 10% 的策略能够获得超额回报，据此认为市场上存在范围较广的错误定价。Crane and Crotty [17] 研究了指数型基金的业绩，发现这些基金的超额业绩并非由于基金经理的股票选择能力或选时能力所致。在国内市场上，王珏和张新民 [18] 发现至少有 10% 的基金具有选股能力，而超过半数的基金没有选股能力。甘甜和冯硕 [19] 也发现了少数基金经理具有选股能力的证据。

KTWW 方法 [2] 和 FF 方法 [3] 截然不同的结论引起了研究者的兴趣。Blake *et al.* [20] 基于英国的数据，对比研究了 KTWW 方法 [2] 和 FF 方法 [3] 的差异及其对基金经理投资能力判断结果的影响。他们发现，在考虑基金费用的条件下，无论是运用 KTWW 方法 [2] 还是 FF 方法 [3]，都未能发现基金经理具有投资能力的证据。而在不考虑基金费用的条件下，运用 KTWW 方法 [2] 确实发现部分基金经理具有投资能力，而运用 FF 方法 [3] 依然没有发现任何能够表明基金经理具有投资能力的证据。这些研究表明，自举抽样方法的差异对基金经理是否具有投资能力的判定具有关键性的影响。还有其他学者也运用自举抽样法研究了基金经理的投资能力 [21-24]。

Harvey and Liu [25] 认为，应该在比较两类错误的基础上对 KTWW 方法和 FF 方法做出评价。他们发现，KTWW 方法存在第一类错误的发生概率较高，即倾向于将没有投资能力的基金判定为具有投资能力。而 FF 方法却由于存在抽样不足的问题①，导致第二类错误的发生概率较高，即倾向于将有投资能力的基金判定为没有投资能力。Harvey and Liu [25] 的发现为 KTWW 方法和 FF 方法存在差异的原因提供了解释。

Huang *et al.* [6] 第一次从理论上证明并比较了 KTWW 方法的基本型和 FF 方法的基本型的第一类错误和检验功效，发现 KTWW 方法基本型的一类错误受到基金收益观测值数量不足的影响且检验功效较低。FF 方法基本型由于对因子和残差同时按照时间自举抽样，导致抽样后的因子收益和残差之

①抽样不足是指在一个时间段内在相同的时间同时对所有基金的月收益观测值抽样时，由于部分基金在样本区间内的观测值数量小于样本区间内的月份数，抽样时这部分基金在没有观测值的月份可能被多次抽样，最终抽样形成的基金观测值数量小于基金实际的观测值数量。例如，在 2016—2020 年间的时间段内共有 60 个月，某基金自 2017 年 1 月成立开始，到 2020 年 12 月共有 48 个月收益观测值，如果按样本区间内的时间自举抽样需要抽取该基金 2016 年 6 月的观测值，但基金在该月没有观测值，最终导致该基金自举后的收益观测值总数小于 48 个。

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

间存在相关性，因此是一种不可行的自举抽样方法，最终导致很难拒绝原假设且几乎没有检验功效。因而进一步将 Hotelling T^2 统计量运用于假设检验，发现占比 1% 至 1.6% 的基金经理具有投资能力。但是 Huang *et al*.[6] 并未比较 KTWW 方法和 FF 方法的其他具体抽样方法的一类错误以及检验功效。本文运用 Harvey and Liu[5] 的双重自举抽样法，全面比较 KTWW 方法和 FF 方法的各种具体抽样方法的两类错误，并从中筛选出具有较低一类错误和较高检验功效的抽样方法。

三、自举抽样方法

(一) KTWW 方法[2] 的自举抽样方法

1. 基本型

首先，基于实际样本数据估计基金的 alpha 及各因子的 β 系数，并形成回归残差序列。其次，对每只基金的回归残差序列独立地随机自举抽样，因子收益的时间序列保持不变。最后，用因子收益乘以该基金的相应因子的 β 系数再加上自举抽样后得到的残差，即可得到去掉 alpha 的伪收益率，用于估计基金的自举抽样后 alpha 及其 t 值。

2. 扩展型

首先，同基本型一样，基于实际样本数据估计基金的 alpha 及各因子的 β 系数，并形成回归残差序列。其次，对因子收益率进行独立自举抽样，形成新的因子收益序列，作为随后回归时所有基金使用的因子收益序列。第三，对每只基金的回归残差序列独立地自举抽样。最后，用因子收益乘以该基金相应因子的 β 系数再加上自举抽样后得到的残差，即可得到基金去掉 alpha 的伪收益率，用于估计基金的自举抽样后的 alpha 及其 t 值。

由于基金投资时存在羊群效应[26]、信息和投资策略共享[27] 以及面对市场流动性冲击，基金收益存在横截面相关性。KTWW 方法[2] 对每个基金的回归残差序列都做了独立随机自举抽样，因此丢掉了各基金收益存在截面相关的特征。FF 方法[3] 认为，KTWW 方法[2] 的这个缺陷导致其做出了有偏推断。

(二) FF 方法[3] 的自举抽样方法

1. 基本型

首先，基于实际样本数据估计基金的 alpha 及各因子的 β 系数，并形成回归残差序列。其次，为了保持自举抽样后各基金收益的横截面相关性，FF 方法[3] 基本型在每次自举抽样时，先对样本区间内的时间做有放回的重复自举抽样，然后按时间自举抽样结果实现对因子收益及所有基金的残差同时自举抽样。最后，用该因子的 β 系数乘以自举抽样后的对应因子收益，再加上自举抽样后得到的残差，即可得到去掉 alpha 的伪收益率，用于估计基金自举抽样后的 alpha 及其 t 值。

2. 扩展Ⅰ型

FF 方法[3] 的扩展Ⅰ型在每次自举抽样时，首先对样本区间内的时间做有放回的重复自举抽样，然后按时间自举抽样结果对所有基金的残差抽样，但并不对因子收益进行自举抽样。用因子收益乘以该基金的对应因子 β 系数再加上自举抽样后得到的残差，即可得到去掉 alpha 的伪收益率，用于估计基金自举抽样后的 alpha 及其 t 值。

3. 扩展Ⅱ型

FF 方法[3] 的扩展Ⅱ型同样先对样本区间内的时间(月份)做有放回重复自举抽样，按时间自举抽样结果对所有基金的回归残差抽样，但对因子收益进行独立的按时间自举抽样并确保随后回归时所有基金使用的因子收益都是相同的自举抽样结果①。用因子收益乘以该基金的对应因子 β 系数再加上自举抽样后得到的残差，即可得到去掉 alpha 的伪收益率，用于估计自举抽样后 alpha 及其 t 值。

由此可见，KTWW 方法和 FF 方法的自举抽样方法具有显著区别。正是这种差别导致了通过

①对样本区间内的时间(月份)做重复有放回的自举抽样，形成两次不同的按时间抽样的结果，一次用于对残差抽样，一次用于对因子收益抽样。

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

KTWW方法和 FF 方法的自举抽样进行多重假设检验的两类错误存在差异。值得注意的是，FF 方法的扩展Ⅰ型仅仅按时间自举抽样结果实现对所有基金的回归残差同时抽样，FF 方法的扩展Ⅱ型按时间自举抽样的不同结果分别对因子收益和回归残差抽样，因此，FF 方法的扩展Ⅰ型和Ⅱ型既保持了基金截面收益存在相关性的特征，同时也避免了 FF 方法基本型会导致因子收益序列和基金残差序列相关的缺陷。

（三）Harvey and Liu[5] 的双重自举抽样方法

为了计算通过 KTWW 方法和 FF 方法的自举抽样方法进行多重假设检验的两类错误发生概率，Harvey and Liu[5] 提出了双重自举抽样法。双重自举抽样法通过对基金的实际样本数据进行两次自举抽样，在每次抽样后都确定具有投资能力的基金，将第二次抽样后确定的具有（不具有）投资能力的基金和第一次抽样后确定的具有（不具有）投资能力的基金对比，研究者就可以计算第一类错误和第二类错误发生的概率。对比 KTWW 方法和 FF 方法两类错误发生的概率，研究者可以确定更为可靠的自举抽样方法①。

四、基金业绩的计算及其整体表现

（一）数据来源及样本选取

本文的数据来自 RESSET 金融数据库。根据基金投资风格，RESSET 将基金分为股票型、债券型、货币型以及积极配置型等若干类别。本文主要分析基金经理的投资能力，因此仅将 RESSET 中投资风格为股票型、激进配置型、偏股型以及灵活配置型的基金纳入样本②。研究样本中排除了各种指数型基金、LOF 基金、分级基金以及 ETF 基金。同时，考虑到 QDII 基金的投资标的与投资国内的基金存在较大差异，基金业绩比较基准完全不同，在样本中排除了 QDII 基金。本文使用的样本涵盖的时间范围为 2011 年 1 月至 2020 年 12 月，此外将成立时间不足 3 年的基金排除。截至 2020 年 12 月，本文研究的样本包涵2 043只基金，共 143 535 个观测值。

（二）定价因子模型

近几十年来，研究者已经发展出了多种基金业绩的评价方法。其中，CAPM 模型、Fama-French 模型和 Fama-French-Carhart 模型已经成为估计基金超额收益的常用方法，所得超额收益分别以 α^{CAPM}、α^{FF} 和 α^{FFC} 表示③。以 Fama-French 模型为例，通过回归计算基金的超额收益如式（1）所示。

$$r_{i,t} - r_{f,t} = \hat{\alpha}_i^{FF,b} + \hat{\beta}_{i,RMRF}(r_{m,t} - r_{f,t}) + \hat{\beta}_{i,SMB}SMB_t + \hat{\beta}_{i,HML}HML_t + \hat{e}_{i,t} \quad (1)$$

在（1）式中，$r_{i,t}$ 是基金在 t 月的净值收益率，$t \in (T_{i0}, \cdots, T_{i1})$。$T_{i0}$ 和 T_{i1} 分别是第 i 只基金第 1 个和最后 1 个收益观测值的取得时间（月份）。根据基金成立时间的不同，T_{i0} 和 T_{i1} 之间的时间长度也不同。$r_{m,t}$ 是沪深 A 股流通市值加权收益率。$r_{f,t}$ 是无风险利率，用上海银行间同业拆借市场 3 个月期限的同业拆借利率代表。SMB 是规模因子收益，是小规模公司股票组合的收益率与大规模公司股票组合的收益率之差。HML 是账面市值比因子收益，是高账面市值公司股票组合的收益率与低账面市值公司股票组合的收益率之差。$\hat{e}_{i,t}$ 是回归残差序列。

此外，估计 Fama-French-Carhart 模型需要动量因子。动量因子收益（MOM）为过去 2 至 12 个月高收益股票组合的收益率与低收益股票组合的收益率之差④。

（三）自举抽样的实现方法

以 FF 方法扩展Ⅰ型为例，阐述自举抽样的实现方法。首先，对每只基金运用（1）式估计{ $\hat{\alpha}_i^{FF}$，

①双重自举抽样法的具体步骤请见 Harvey and Liu[5]。

②2014 年，证监会要求凡是基金名称中标明为“股票型”的基金，其资产组合中的股票持有比例不得低于 80%。一部分股票型基金为了规避股票持仓比例的限制，便将其基金类型变更为混合型基金，导致股票型基金数量大幅减少。为了将更多的基金纳入研究样本，本文将激进配置型、偏股型以及灵活配置型基金纳入样本。

③近年来，研究者提出了其他评价基金业绩的方法。如 Hou *et al.*[28] 的四因素模型，Fama and French[29] 的五因素模型，但这些模型在国内市场的适用性仍存争议。

④本文使用的无风险利率和因子收益率的数据均来自 RESSET 金融数据库。

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

$\hat{\beta}_{i,RMRF}$，$\hat{\beta}_{i,SMB}$，$\hat{\beta}_{i,HML}$}，并得到回归残差序列 $\hat{e}_{i,t}$，$t \in (T_{i0}, \cdots, T_{i1})$。在第 b 次实施 FF 方法扩展 Ⅰ 型抽样时，仅按时间（月份）对每只基金的残差序列同时抽样，得到{ $\hat{e}^{b}_{p,t_\varepsilon}$，$t_\varepsilon \in (s^{b}_{T_{i0}}, \cdots, s^{b}_{T_{i1}})$ }，其中，$(s^{b}_{T_{i0}}, \cdots, s^{b}_{T_{i1}})$ 是按照有放回重复随机抽样的方法取自 $[T_{i0}, \cdots, T_{i1}]$。同时因子收益序列保持不变。

其次，按照（2）式构建基金去掉 alpha 的伪收益率序列 $r^{b}_{i,t}$。

$$r^{b}_{i,t} = \hat{\beta}_{i,RMRF}(r_{m,t} - r_{f,t}) + \hat{\beta}_{i,SMB}SMB_t + \hat{\beta}_{i,HML}HML_t + \hat{e}^{b}_{p,t_\varepsilon} \tag{2}$$

在（2）式中，$t \in (T_{i0}, \cdots, T_{i1})$，$t_\varepsilon \in (s^{b}_{T_{i0}}, \cdots, s^{b}_{T_{i1}})$，显然，$r^{b}_{i,t}$ 的超额收益 $\alpha^{FF,b}_{i}$ 应该为 0。但是，当运用（1）式做 $r^{b}_{i,t}$ 关于各因子的回归时，由于随机抽样可能会抽到较大的残差，有可能得到正的 $\alpha^{FF,b}_{i}$，当然，也可能抽到较小的残差，因此得到负的 $\alpha^{FF,b}_{i}$。当对所有的基金伪收益序列分别做回归后，就得到第 b 次自举抽样时每只基金的 $\alpha^{FF,b}_{i}$，{ $i = 1, \cdots, N$ }，并进一步计算第 b 次抽样后 $\alpha^{FF,b}_{i}$ 的各分位数。重复上述抽样 1 000 次，即可得到 $\alpha^{FF,b}_{i}$ 各分位数的截面统计分布。

（四）基金业绩的描述性统计

对基金的 alpha 做描述性统计，结果表明，在 2011—2015 年间，以 Fama-French 三因素模型计算的 α^{FF} 的均值为 0.006 6%，中位数为 -0.01%，年均值分别为 0.079% 和 -0.16%。显著性检验表明，α^{FF} 的均值与 0 没有差异，基金整体上没有取得超过市场基准的超额收益。而在 2016—2020 年间，α^{FF} 的月平均值为 0.43%，中位数为 0.36%，年均值分别为 5.2% 和 4.32%。显著性检验表明，α^{FF} 的均值显著大于 0，基金整体上取得超过市场基准的超额收益。通过对比上述两个时段的 α^{FF} 后发现，在 2016—2020 年间基金的超额收益显著高于 2011—2015 年间的超额收益，两者相差的年均值约为 5%。

进一步发现，α^{CAPM} 和 α^{FF}、α^{FFC} 的相关系数在 0.71 至 0.89 之间，而 α^{FF} 和 α^{FFC} 的相关系数均在 0.98 以上。这说明运用 Fama-French 模型和 Fama-French-Carhart 模型计算的基金业绩高度趋同。因此，为节约篇幅，本文主要基于 α^{FF} 检验各种抽样方法的两类错误，并进一步检验基金经理的投资能力。在随后的稳健分析过程中，将基于 α^{CAPM} 进行相应检验。

（五）基金业绩的整体表现

将所有研究样本内的基金看作一个资产组合，并考察这个资产组合的超额收益，这能够为检验基金经理从整体上是否能够战胜市场提供直接证据。在 2011—2020 年间，计算所有基金在每月的平均收益率①，得到所有基金的平均收益率月度时间序列。运用定价因子模型，可以估计所有基金的超额收益。具体结果见表 1。表 1 中给出了在各个定价因子模型下全部基金的 alpha 及 β 系数，并在其下方给出了假设检验的 t 值。

由表 1 可见，在 2011—2015 年间，即使在 10% 的显著性水平下，运用三种定价因子模型计算的超额收益均不显著。这表明，从整体来看，基金没有取得超额收益，基金经理并未给投资者带来额外回报。在 2016—2020 年间，基金经理的业绩表现完全不同。以 α^{FF} 来看，超额收益率为 5%，t 值为 3.57。这表明基金经理能够为投资者带来超过市场基准收益率的额外回报，基金经理战胜了市场，为投资者创造了价值。进一步考察 2011—2020 年间基金经理的超额收益，当显著性水平为 5% 时，以 α^{CAPM} 和 α^{FF} 度量的超额收益分别是 0.41% 和 0.36%，且在统计上是显著的，当显著性水平为 10% 时，α^{FFC} 也是显著的。可见，从整体上来看，尽管受 2011—2015 年间基金总体收益较差的影响，导致基金经理在 2011—2020 年间获得的超额收益有所下降，但仍然有证据表明整体上基金经理战胜了市场，是具有投资能力的。

①本文只计算了所有基金每月的简单算术平均收益率。在假设检验过程中，每个基金都被同等对待。而在计算简单算术平均收益率时，每只基金的收益也被同等对待。因此，通过检验所有基金的简单算术平均收益率，能够为假设检验给出一致的结果。同时也计算并检验了所有基金按净资产加权的平均超额收益率，所得结果与表 1 结果一致。

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表 1　基金的超额收益

时间	变量	α	$\hat{\beta}_{pRMRF}$	$\hat{\beta}_{pSMB}$	$\hat{\beta}_{pHML}$	$\hat{\beta}_{pMOM}$	R^2
2011—2015 年	α^{CAPM}	0.002 1	0.81				0.82
		0.60	16.53				
	α^{FF}	−0.000 2	0.75	0.11	−0.42		
		−0.07	23.15	1.46	−4.66		
	α^{FFC}	−0.000 2	0.72	0.11	−0.44	0.19	0.95
		−0.11	24.59	1.72	−5.63	4.64	
2016—2020 年	α^{CAPM}	0.006 0	0.72				0.89
		3.24	22.31				
	α^{FF}	0.004 2	0.66	−0.22	−0.43		0.95
		3.57	30.62	−5.59	−9.42		
	α^{FFC}	0.003 5	0.64	−0.14	−0.42	0.12	0.97
		3.20	29.56	−3.04	−9.92	3.32	
2011—2020 年	α^{CAPM}	0.004 1	0.77				0.84
		2.10	25.32				
	α^{FF}	0.003 6	0.72	−0.01	−0.46		0.92
		2.59	32.22	−0.27	−8.91		
	α^{FFC}	0.002 2	0.67	0.06	−0.43	0.21	0.95
		1.90	33.74	1.50	−9.90	6.95	

五、假设检验中 KTWW 方法和 FF 方法的两类错误

运用 Harvey and Liu[5] 的双重自举抽样方法，计算在假设检验中运用 KTWW 方法以及 FF 方法的两类错误，所得结果见表 2。从表 2 可见，在 2011—2015 年间，在各种抽样方法中 FF 基本型的一类错误和二类错误总是处于最高水平。例如，当 p_0 为 1% 时，FF 基本型的一类错误为 11.26%，比其他抽样方法高出约 5 个百分点；FF 基本型的二类错误为 0.44%，比其他抽样方法高出 4 至 8 个基点。其次，在各种抽样方法中，FF 扩展 I 型总是具有最低的一类错误和二类错误。例如，当 p_0 为 1% 时，FF 扩展 I 型的一类错误为 6%，二类错误为 0.36%，较 FF 基本型的一类错误和二类错误分别下降了 5.26% 和 0.08%，是各种抽样方法中的最低水平。此外，我们发现，FF 扩展 II 型的两类错误仅次于 FF 扩展 I 型，是所有抽样方法中具有次低两类错误的抽样方法。由此可见，在 2011—2015 年间，FF 扩展 I 型是更为可靠的抽样方法。

从表 2 也可以得到相同的结论，在 2016—2020 年间，KTWW 基本型的两类错误总是低于 FF 基本型的两类错误。如当 p_0 设为 35% 时，KTWW 基本型的第一类错误和第二类错误分别为 6.4% 和 9.9%，均低于 FF 基本型的 6.8% 和 10.6%。当比较所有抽样方法的两类错误时，FF 基本型总是具有最高的一类错误和二类错误，而 FF 扩展 I 型和扩展 II 型是第一类错误最低的抽样方法。第二类错误最低的仍然是 FF 扩展 I 型。由此可见，在 2016—2020 年间，FF 扩展 I 型是更为可靠的抽样方法。

表 2 中展示了 2011—2020 年内各种抽样方法的两类错误，所得结果与表 2 中的 2011—2015 年间及 2016—2020 年间一致。

Huang *et al.*[6] 发现，相对于 FF 基本型，KTWW 基本型具有较低的一类错误和更高的检验功效。本文的发现为 Huang *et al.*[6] 的结论提供了实证证据。同时，在比较了各种具体自举抽样方法后发现，KTWW 基本型不是最可靠的自举抽样方法，而 FF 扩展 I 型和扩展 II 型才是具有更低一类错误和更高检验功效的抽样方法。

FF 扩展 I 型没有对因子收益率自举抽样，而总是在相同时间点对所有基金的残差同时自举抽样。因此，采用 FF 扩展 I 型自举抽样形成的数据，因子收益和自举后形成的基金收益率之间不存在相关性，但同时又保持了自举后基金收益率具有截面相关性的特点。而 KTWW 方法对每只基金的回归残差（以

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

及因子收益率）进行独立自举抽样，丢掉了基金收益率之间存在截面相关的特性，因此，Fama and French[3] 认为，基于 KTWW 的两种具体抽样方法[2] 的推断总是存在偏差①。此外，我们发现，FF 扩展 II 型的两类错误总是更接近 FF 扩展 I 型的两类错误，是除了 FF 扩展 I 型以外更可靠的抽样方法。

表 2　假设检验中 KTWW 方法和 FF 方法的两类错误

时间	p_0	Type Ⅰ					Type Ⅱ				
		KTWW	KTWW Ⅰ	FF	FF Ⅰ	FF Ⅱ	KTWW	KWW Ⅰ	FF	FF Ⅰ	FF Ⅱ
2011—2015 年	1%	0.065 4	0.068 6	0.112 6	0.060 9	0.066 3	0.003 6	0.004 0	0.004 4	0.003 6	0.004 0
	5%	0.221 6	0.228 1	0.249 4	0.214 8	0.219 9	0.015 3	0.016 8	0.019 0	0.015 2	0.016 7
	10%	0.317 0	0.323 7	0.328 6	0.312 2	0.317 0	0.029 8	0.032 2	0.036 6	0.029 7	0.032 1
2016—2020 年	35%	0.064 1	0.065 5	0.067 7	0.059 7	0.059 5	0.099 1	0.106 2	0.105 8	0.096 5	0.104 3
	40%	0.077 3	0.078 7	0.080 0	0.072 5	0.072 4	0.115 6	0.123 2	0.122 2	0.112 2	0.120 5
	45%	0.090 0	0.091 4	0.091 9	0.085 1	0.084 9	0.134 6	0.142 6	0.140 8	0.130 2	0.139 1
2011—2020 年	30%	0.043 7	0.044 3	0.049 6	0.042 0	0.042 5	0.086 6	0.090 0	0.093 4	0.085 2	0.089 4
	35%	0.056 3	0.057 0	0.061 4	0.054 4	0.054 8	0.101 8	0.105 5	0.109 1	0.099 9	0.104 6
	40%	0.069 9	0.070 6	0.074 1	0.067 9	0.068 1	0.118 1	0.122 1	0.125 7	0.115 6	0.120 6

六、基金经理投资能力的再检验

（一）实证结果

通过评估假设检验中运用各种自举抽样方法后的两类错误，我们发现 FF 扩展 I 型和 FF 扩展 II 型是更为可靠的方法。因此，本节运用这两种抽样方法，检验基金经理的投资能力，所得结果见表 3。表 3 第 1 列是基于基金实际收益率数据采用 Fama-French 模型估计的超额收益（α^{FF}）的分位数，第 2 列是 α^{FF} 的实际 p 值的分位数，第 3 列和第 4 列分别是基于 FF 扩展 I 型和 FF 扩展 II 型自举抽样方法估计的 α^{FF} 的分位数统计分布基础上计算取实际 t 值各分位数的概率，即自举抽样 p 值。

由表 3 可见，在 2011—2015 年间，当显著性水平是 5% 时，实际 alpha 的 94% ~99% 分位数的 p 值小于 5%，拒绝（1）式的原假设。因此，如果不考虑运气的因素，基金经理表现出投资能力；然而，无论采用的自举抽样方法是 FF 扩展 I 型还是 FF 扩展 II 型，自举抽样 p 值在 0.28 至 0.44 之间，不能拒绝原假设。这表明当去掉运气因素之后，基金经理并不具有投资能力。另一方面，实际 α^{FF} 的 1% 分位数的 p 值为 0.01，而自举抽样的 p 值为 0.57，这表明当去掉运气因素之后，基金的超额收益也并不显著。此外，10% ~90% 分位数的实际 alpha 的 p 值和自举抽样 p 值均大于 5%，不能拒绝原假设。由此我们发现，在 2011—2015 年间，去掉运气因素以后，基金经理没有投资能力，基金经理并未战胜市场。

然而，当考察在 2016—2020 年间基金经理的投资能力时发现了完全不同的证据。当显著性水平是 5% 时，实际 α^{FF} 的 60% ~99% 分位数都是显著的，拒绝原假设。可见，如果不考虑运气的因素，这些证据支持基金经理具有投资能力的结论。但是，无论采用的自举抽样方法是 FF 扩展 I 型还是 FF 扩展 II 型，自举抽样的 p 值都接近 0。这表明即使去掉了基金业绩中的运气因素，原假设也被拒绝，基金经理同样表现出显著的投资能力。另一方面，实际 α^{FF} 的 1% 分位数的 p 值为 2.5%，而自举抽样的 p 值为 0.87，这表明当去掉运气因素之后，没有基金取得显著的超额收益。由此发现，在 2016—2020 年间，基金经理确实具有投资能力。

当考察在 2011—2020 年间基金经理的投资能力时，研究发现，尽管受到 2011—2015 年间较差业绩的影响，仍然有证据表明基金经理具有投资能力。

①Fama and French[3] 的原文是 "The important point, however, is that whatever inclusion rules are used, failure to account for the joint distribution of fund returns, and of fund and explanatory returns, biases the inferences of Kosowski *et al*. (2006)[2] toward positive performance."

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表 3　基金 alpha 的 t 统计量的分位数横截面检验

分位数	2011—2015				2016—2020				2011—2020			
	alpha	p 值	FF I	FF II	alpha	p 值	FF I	FF II	alpha	p 值	FF I	FF II
1%	-0.012 8	0.009 6	0.570 1	0.570 5	-0.007 6	0.024 9	0.866 5	0.866 0	-0.007 3	0.031 6	0.989 2	0.990 4
5%	-0.007 6	0.054 1	0.632 4	0.632 7	-0.003 4	0.190 0	0.985 9	0.985 6	-0.003 2	0.192 0	1.000 0	1.000 0
10%	-0.006 1	0.097 1	0.500 8	0.493 3	-0.001 3	0.342 8	0.996 1	0.995 5	-0.001 1	0.378 4	1.000 0	1.000 0
20%	-0.004 3	0.186 1	0.388 7	0.382 9	0.000 7	0.415 8	0.999 3	0.999 5	0.000 9	0.392 7	1.000 0	1.000 0
30%	-0.002 4	0.282 3	0.358 2	0.356 1	0.002 0	0.264 0	0.999 6	0.999 7	0.002 0	0.244 6	1.000 0	1.000 0
40%	-0.001 1	0.396 7	0.478 6	0.479 6	0.002 7	0.168 6	0.999 8	0.999 7	0.002 8	0.152 0	1.000 0	1.000 0
50%	-0.000 1	0.488 5	0.568 6	0.566 6	0.003 6	0.101 4	0.000 1	0.000 3	0.003 7	0.088 7	0.000 0	0.000 0
60%	0.001 2	0.399 0	0.468 7	0.465 8	0.004 8	0.054 5	0.000 1	0.000 2	0.004 8	0.045 9	0.000 0	0.000 0
70%	0.002 5	0.2902	0.382 6	0.381 8	0.006 3	0.022 4	0.000 0	0.000 2	0.006 2	0.018 5	0.000 0	0.000 0
80%	0.004 1	0.200 6	0.472 2	0.473 4	0.008 2	0.007 9	0.000 0	0.000 2	0.008 0	0.006 5	0.000 0	0.000 0
90%	0.006 2	0.089 5	0.323 7	0.317 1	0.010 4	0.001 7	0.000 0	0.000 1	0.010 3	0.001 1	0.000 0	0.000 0
91%	0.006 4	0.085 3	0.409 0	0.408 3	0.010 8	0.001 3	0.000 0	0.000 0	0.010 7	0.001 0	0.000 0	0.000 0
92%	0.006 7	0.077 1	0.425 6	0.420 8	0.011 1	0.001 0	0.000 0	0.000 0	0.010 9	0.000 8	0.000 0	0.000 0
93%	0.007 5	0.064 8	0.371 2	0.362 0	0.011 4	0.000 8	0.000 0	0.000 0	0.011 3	0.000 5	0.000 0	0.000 0
94%	0.007 7	0.049 6	0.292 9	0.282 9	0.011 9	0.000 6	0.000 0	0.000 0	0.011 8	0.000 4	0.000 0	0.000 0
95%	0.008 0	0.042 6	0.314 2	0.307 7	0.012 7	0.000 3	0.000 0	0.000 0	0.012 5	0.000 2	0.000 0	0.000 0
96%	0.009 2	0.035 8	0.392 5	0.383 2	0.013 5	0.000 2	0.000 0	0.000 0	0.013 1	0.000 2	0.000 0	0.000 0
97%	0.009 7	0.024 8	0.320 6	0.315 2	0.014 2	0.000 1	0.000 0	0.000 0	0.013 9	0.000 1	0.000 0	0.000 0
98%	0.010 5	0.018 0	0.437 8	0.441 4	0.015 7	0.000 1	0.000 1	0.000 1	0.015 2	0.000 1	0.000 0	0.000 0
99%	0.010 5	0.018 0	0.437 8	0.441 4	0.015 7	0.000 1	0.000 1	0.000 1	0.015 2	0.000 1	0.000 0	0.000 0

为什么在不同时间段内国内基金经理表现出不同的投资能力呢？本文认为其中主要有两个原因。

第一，为了应对全球金融危机，2008 年开始政府陆续出台了一系列财政和货币刺激政策，国内经济快速增长。然而，从 2011 年开始，政府收紧了财政和货币政策，国内经济受到影响，股票市场也受到冲击。另一方面，在 2012 年经历了政府换届，换届前后的宏观经济政策面临诸多不确定性，这会增加投资的难度。研究表明，经济周期以及宏观经济政策的不确定性都会影响基金业绩[30]。因此，基金经理在 2011—2015 年间很难获得超额收益。

第二，国内股市在 2014 年下半年至 2015 年底期间经历了暴涨暴跌的过程，股市波动性大幅度提高。而从 2016 年开始，股市波动性明显下降。在剧烈波动的市场里，投资者及基金经理的投资难度增加，很难获得超额收益，无法表现出投资能力。

（二）稳健性检验

上文基于 Fama-French 三因素模型估计的 alpha（α^{FF}）分析了基金经理的投资能力。本文进一步基于 CAPM 模型估计的 α^{CAPM} 分析基金经理的投资能力，所得结果与前文一致①。

七、结论与启示

在评价基金经理投资能力时，自举抽样法能够有效控制运气因素对基金业绩的影响，因而得到了广泛运用。通过假设检验，Fama and French[3] 和 Kosowski *et al.*[2] 基于不同的自举抽样实现方法检验了基金经理的投资能力，得到了完全不同的结论。在假设检验中，由于抽样随机性等因素的影响，总是会出现两类错误，即原假设为真时，错误地拒绝原假设的一类错误，以及备择假设为真时，未能拒绝原假设的第二类错误。Harvey and Liu[5] 提出的双重自举抽样法为研究者计算假设检验中两类错误的

①稳健性分析结果可向作者索取。

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

发生概率提供了解决办法。运用 Harvey and Liu[5] 的双重自举抽样法，计算了在假设检验中不同自举抽样实现方法的两类错误，并确定了最优的自举抽样方法。在此基础上检验了国内市场基金经理的投资能力。本文发现，在假设检验中，采用 FF 扩展 I 型抽样方法的两类错误更低，因而是最为可靠的自举抽样方法。这是因为 FF 扩展 I 型仅在相同时间点上对所有基金的回归残差同时自举抽样，保留了自举后基金收益率的截面相关性。同时，采用这种抽样方法后，定价因子收益和自举抽样后形成的基金收益之间不存在相关性。进一步在假设检验中运用 FF 扩展 I 型的自举抽样方法，检验了基金经理的投资能力。本文发现，在去除基金业绩中的运气因素后，尽管在 2011—2015 年间基金的业绩相对较差，然而在 2011—2020 年的长达 10 年的时间内，基金经理整体上取得了超过市场基准收益的超额收益，表现出了显著的投资能力，基金经理战胜了市场。本文的结论表明：(1) 国内基金经理确实具有投资能力，能够为基金持有人创造财富，基金经理以业绩证明了自己存在的价值。(2) 根据有效市场假说，在信息充分流动、制度健全、充分竞争的金融市场里，投资者无法获得超额收益。然而，国内基金经理能够获得超额收益的事实表明，监管机构必须进一步完善相关制度，打击内幕交易、严惩市场操纵行为，不断提高金融市场的有效性。

参考文献：

[1] HOROWITZ J L. The bootstrap [J]. Handbook of econometrics, 2001, 5: 3159-3228.

[2] KOSOWSKI R, TIMMERMANN A, WERMERS R, et al. Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis [J]. Journal of finance, 2010, 61(6): 2551-2595.

[3] FAMA E F, FRENCH K R. Luck versus skill in the cross section of mutual fund returns [J]. Social science electronic publishing, 2010, 65(5): 1915-1947.

[4] BLAKE D, CAULFIELD T, IOANNIDIS C, et al. Improved inference in the evaluation of mutual fund performance using panel bootstrap methods [J]. Journal of econometrics, 2014, 183(2): 202-210.

[5] HARVEY C R, LIU Y. False (and missed) discoveries in financial economics [J]. Journal of finance, 2020, 75(5): 2503-2553.

[6] HUANG H, JIANG L, LENG X, et al. Bootstrap analysis of mutual fund performance [J/OL]. Journal of econometrics, 2020. https://ssrn.com/abstract=3513979.

[7] JENSEN M C. The performance of mutual funds in the period 1945—1964 [J]. Journal of finance, 1968, 23(2): 389-416.

[8] ELTON E J, GRUBER M J, DAS S, et al. Efficiency with costly information: a reinterpretation of evidence from managed portfolios [J]. The review of financial studies, 1993, 6(1): 1-22.

[9] CARHART M M. On persistence in mutual fund performance [J]. Journal of finance, 1997, 52(1): 57-82.

[10] DANIEL K, GRINBLATT M, TITMAN S, et al. Measuring mutual fund performance with characteristic-based benchmarks [J]. Journal of finance, 1997, 52(3): 1035-1058.

[11] JIANG G J, YAO T, YU T. Do mutual funds time the market? Evidence from portfolio holdings [J]. Journal of financial economics, 2007, 86(3): 724-758.

[12] 汪光成. 基金的市场时机把握能力研究 [J]. 经济研究, 2002(1): 48-55+95.

[13] 廖长友. 我国基金持有人选时能力的实证分析 [J]. 中南财经政法大学学报, 2011(2): 121-126.

[14] KOSOWSKI R, NAIK N Y, TEO M. Do hedge funds deliver alpha? A bayesian and bootstrap analysis [J]. Journal of financial economics, 2007, 84(1): 229-264.

[15] CUTHBERTSON K, NITZSCHE D, O'SULLIVAN N. UK mutual fund performance: skill or luck? [J]. Journal of empirical finance, 2008, 15(4): 613-634.

[16] YAN X S, ZHENG L. Fundamental analysis and the cross-section of stock returns: a data-mining approach [J]. Review of financial studies, 2017, 30(4): 1382-1423.

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

[17] CRANE A D, CROTTY K. Passive versus active fund performance: do index funds have skill? [J]. Journal of financial and quantitative analysis, 2018, 53(1): 33-64.

[18] 王珏，张新民. 基于 bootstrap 分析方法的我国基金经理选股能力研究 [J]. 中国软科学，2013(11): 139-150.

[19] 甘甜，冯硕. 基于 bootstrap 分析方法的中国股票型公募基金业绩实证研究 [J]. 中国证券期货. 2018(4): 65-80.

[20] BLAKE D, CAULFIELD T, IOANNIDIS C, et al. New evidence on mutual fund performance: a comparison of alternative bootstrap methods [J]. Journal of financial and quantitative analysis, 2017, 52(3): 1279-1299.

[21] BARRAS L, SCAILLET O, WERMERS R. False discoveries in mutual fund performance: measuring luck in estimated alphas [J]. Journal of finance, 2010, 65(1): 179-216.

[22] CAO C, CHEN Y, LIANG B, et al. Can hedge funds time market liquidity? [J]. Journal of financial economics, 2013, 109(2): 493-516.

[23] FERSON W, MO H. Performance measurement with selectivity, market and volatility timing [J]. Journal of financial economics, 2016, 121(1): 93-110.

[24] CHORDIA T, GOYAL A, SARETTO A. Anomalies and false rejections [J]. Review of financial studies, 2020, 33(5): 2134-2179.

[25] HARVEY C R, LIU Y. Luck versus skill in the cross-section of mutual fund returns reexamining the evidence [J]. Journal of finance, 2022, 77(3): 1921-1966.

[26] WERMERS R. Mutual fund herding and the impact on stock prices [J]. Journal of finance, 1999, 54(2): 581-622.

[27] CUJEAN J. Idea sharing and the performance of mutual funds [J]. Journal of financial economics, 2020, 135(1): 88-119.

[28] HOU K W, XUE C, ZHANG L. Digesting anomalies: an investment approach [J]. The review of financial studies, 2015, 28(3): 650-705.

[29] FAMA E F, FRENCH K R. A five-factor asset pricing model [J]. Journal of financial economics, 2014, 116(1): 1-22.

[30] KACPERCZYK M, VAN NIEUWERBURGH S, VELDKAMP L. Rational attention allocation over the business cycle [J]. Econometrica, 2015, 99(3): 546-559.

（责任编辑：刘淑浩；英文校对：葛秋颖）

Can Fund Managers Beat Market? Evidence from Double Bootstrap

LIAO Changyou[1], LI Yumin[2], LIU Xingyi[3]

(1. School of Management, Xihua University, Chengdu 610039, China;
2. School of International Economics and Management, Xihua University, Chengdu 610039, China;
3. School of Software and Microelectronics, Peking University, Beijing 102600, China)

Abstract: In the evaluation of fund manager's investment ability, bootstrap can effectively control the influence of luck factor on fund performance, so it has been widely used. Bootstrap has different sampling methods. Based on the double bootstrap method, this paper calculates two types of errors of different sampling methods in hypothesis testing, and determines a more reliable sampling method. On this basis, the paper test whether domestic market fund managers can beat the market. It has been found that, it is more reliable to simultaneously bootstrap sampling the series of regression residuals for all funds by time without joint bootstrap sampling of factor returns. The paper further finds out that during 2011—2015, fund managers did not obtain excess returns that exceeded the market benchmark, and there is no evidence that fund managers beat the market. While during 2016 to 2020, fund managers as a whole achieved significant excess returns, and fund managers beat the market.

Key words: fund manager; double bootstrap sampling method; market efficiency; investment ability

(C)1994-2022 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net