

# 学术界、管理人、投资者视角下的因子投资

原创：石川 川总写量化 4月4日



作者：石川，北京量信投资管理有限公司创始合伙人，清华大学学士、硕士，麻省理工学院博士。知乎专栏：

<https://zhuanlan.zhihu.com/mitcshi>。

**未经授权，严禁转载。**

摘要 本文从学术界、专业管理人和普通投资者三个视角梳理因子投资。

## 1

### 引言

如今的（量化）投资界开口必谈因子投资。然而，随着所处的角色不同（学术界 vs 业界），人们对因子投资的关注点不尽相同。2018 年，Joseph Cerniglia 和 Frank Fabozzi 在 The Journal of Portfolio Management 上发表了一篇题为 Academic, Practitioner, and Investor Perspectives on Factor Investing 的文章（Cerniglia and Fabozzi 2018），从**学术界**、（专业）**管理人**和（普通）**投资者**三个视角梳理了因子投资，给人很大启发。

面对着海量关于因子投资的文献和资料，搞清楚不同视角对因子投资理解的区别和联系十分必要；它可以帮助我们更有效的甄别和获取知识，指导因子投资实践。

本文以介绍 Cerniglia and Fabozzi (2018) 为契机，从上面三个角度阐述因子投资。除了提炼该文精华之外，行文中也会加入大量相关材料供读者参考。此外，由于我之前写过不少因子投资的文章，因此也将借此机会把相关内容串联起来、做一个整理。（如果你把此文视作公众号之前所创作的主要因子投资文章的合集，那也是极好的。）

最后需要说明的是，**本文谈及的因子投资针对的是股票市场的风格因子**，而非更广义的使用因子进行大类资产配置。

## 2

### 学术界视角

一般来说，因子描述了一揽子股票所共同承担（或者暴露于的）的某种（可以是未知的）系统性风险。相应的，因子收益率是围绕该因子构建的投资组合中股票的共性收益（即系统性风险溢价）。根据这个定义，因子（收益）又被称为 **style premiums**。

#### 2.1 因子投资研究的起源

学术界对于因子的研究可以追溯到 20 世纪 30 年代。Graham and Dodd (1934) 提出了价值溢价，而这本证券分析也早已成为业界的圣经。之后的 60 年代和 70 年代，CAPM（见[《CAPM 的一小段历史》](#)）和 Arbitrage Pricing Theory (Ross 1976) 相继被提出，它们为研究因子提供了定量的分析工具。

自上世纪 70 年代以来，学者们逐渐发现按照某种风格“打包”的股票能够战胜市场。这其中最值得一提的是 Basu (1977) 发现的便宜股效应和 Banz (1981) 发现的小市值效应。由于它们和当时主流的有效市场假说相违背，因此被称为异象（anomalies）。也正是从那个时期开始，学术界开始了轰轰烈烈的挖异象大潮。

#### 2.2 Empirical Asset Pricing

对于学术界而言，研究因子的最主要目标是提出更好的 **empirical asset pricing model（定价模型）**。一个潜在的基本面或技术面风格，要想实现从成为异象、到因子再最终被加入一个多因子模型的三步走，需要经过一系列严格的计量经济学分析（见[《Anomalies, Factors, and Multi-Factor Models》](#)）。

在比较不同的多因子模型时，主流统计手段包括 GRS tests、Mean-Variance Spanning tests 以及 Bayesian approach。近年来，随着机器学习的发展，也有学者将其应用到构建因子模型上，这其中的代表是 Feng, Giglio, and Xiu (2017)。这篇文章获得了 2018 年 AQR Insight Award，它使用 two-pass cross-sectional regression（具体的，两步 LASSO）从一大堆因子中找到显著且稳定的定价因子。

谈及关于 empirical asset pricing 的研究，诺贝尔经济学奖获得者 Eugene Fama 可谓功不可没。他和诸多合作者发表的许多重磅论文奠定了不仅是因子研究，更是金融学作为一个学科的发展。即使在今天看来，这些文章依然是所有踏入这个行业的人的必修课。这其中包括大名鼎鼎的 Fama-French 三因子模型 (Fama and French 1993)、Fama-MacBeth Regression (Fama and MacBeth 1973) 以及 Fama et al. (1969) —— 事件分析 (event study) 的开山之作。

为了记录 Fama 对金融领域的贡献，John Cochrane（没错，Fama 女婿……）和 Tobias Moskowitz 挑选了 Fama 最具代表性的论文编制了一本论文集，题为 The Fama Portfolio (Fama 2017)。无疑，这里面的文章全都值得仔细品味。此外，在研究多因子模型时，回归分析（包括时序回归、截面回归）是极为重要的研究方法，[《股票多因子模型的回归检验》](#)一文对不同方法进行了详尽的说明。

## 2.3 “异象”丛生

时至今日，已有超过 400 个异象（或“因子”）被挖出 (Hou, Xue, and Zhang 2017)。在 **publication bias** 和 **multiple testing** 的不良影响下，绝大多数异象都是 **data mining** 的产物。学术界对于挖因子的狂热和浮躁的态度已经引起很多学者的警惕。John Cochrane 在 2011 年美国金融协会主席演讲时以“factor zoo”来描述当前因子研究的现状，并提出了三个至关重要的问题：

1. Which factors are independent?
2. Which factors are important?
3. Why do factors move prices?

Cochrane 的提问引发了学术界深刻的反思。最近几年，越来越多发表于顶级期刊上的文章致力于回答这些问题。Harvey, Liu, and Zhu (2016) 研究了 316 个因子、提出了一种能够利用不同因子之间相关性的全新分析框架，把代表因子显著性的 **t-statistic 阈值从 2.0 提高到了 3.0**；Harvey and Liu (2018) 则更是以 Lucky Factors 为题提出了一个基于正交化和自助法的分析框架，用来检验因子的有效性（《出色不如走运 II》、《实证研究 —— 使用正交化和自助法寻找显著因子》）。除此之外，Green, Hand, and Zhang (2017) 使用 Fama-MacBeth regression 同时分析了 94 个因子、指出真正独立的因子少之又少（《解释股票截面预期收益差异的独立因子》）。

## 2.4 样本外效果几何？

学术界关注的另一个问题是：**一个新因子被提出后，随着越来越多人使用，它在 post-publication 样本外的效果是否会打折扣**（样本内的 data mining 以及发表后被更多人知道都会降低其效果）。在这方面，McLean and Pontiff (2016) 研究了 97 个因子在被发表之后的表现，发现因子的收益率比论文中的 in-sample 降低 50% 以上。类似的，Linnainmaa and Roberts (2018) 研究了 30+ 个财务因子在 pre-sample 和 post-sample 的表现；绝大多数因子在样本外的表现令人失望（见《未知风险，错误定价，还是数据迁就？》）。

**大多数研究因子的学术论文中并没有对交易费用给予充分的考虑，造成对因子收益率的高估。**

顺带一提，由于因子投资组合一般都是多、空对冲的组合，因此如果不合理考虑做空限制，也会高估因子的收益。当然可喜的是，学者们意识到了交易费用对因子效果的冲击，并开始论文中对其加以必要的探讨。Novy-Marx and Velikov (2015) 研究了交易费用对因子效果的影响，并提出三个交易策略来降低执行这些因子投资组合的成本：（1）仅使用交易费用低的股票构建因子组合；（2）降低因子组合再平衡的频率；（3）在交易时考虑更严格的买卖价差约束。

## 2.5 多指标、多因子综合

在基金管理人进行因子投资时，无论是配置也好、给个股打分也罢，通常会把多个因子综合起来使用。**学术界对于多个因子（或指标）综合起来在解释股票收益率截面差异方面也有很多研究**。Piotroski (2000) 提出的 F-score 以及 Mohanram (2005) 提出的 G-score 是这方面的代表。对于如何使用 F-score 寻找预期差从而增强选股的效果，《寻找股票市场中的预期差》一文给出了基于 A 股的实证结果。

此外，AQR 提出的 QMJ 因子 (Quality-Minus-Junk) 也是很好的例子 (Asness, Frazzini, and Pedersen 2019; 该文最早的版本是 2014 版本)。从我自己的实证来看，QMJ 在 A 股上也是有效的。

就综合多个指标这个问题，Novy-Marx (2015) 指出，虽然同时使用多个信号无可厚非，但仍需要谨慎考虑每个因子对于提升收益的增量贡献（《出色不如走运 III》）。

## 2.6 风格因子和宏观经济的关系

学术界的最后一个视角可以归结为风格因子和宏观经济的关系。在不同的经济环境下，风格因子的表现大相径庭。研究经济环境和因子收益率的关系对于因子择时也很有帮助。

Claessens and Kose (2018) 调研了学术界的大量相关文献，是一篇很好的综述。有必要说明的是，**宏观经济和因子收益率之间的关系并没有人们预期的那么强**。Cerniglia and Fabozzi (2018) 也指出在这方面还有大量的研究要做。

### 3

## 管理人视角

对于主动型管理人来说，因子投资早已成为投资工具箱中的重要选择。一个因子的 on paper 收益率可能非常诱人，但是对于管理人来说，如何在现实中实现它是必须要解决的问题。**在实际围绕该因子构建投资组合时，必须要考虑可投资性的约束**（《因子投资 —— “被动的”主动投资》）；**成功的因子投资需要注重从理论到实践的每一个细节**（《你用因子，他也用因子；你没赚钱，他却赚钱了》）。

此外，**对于管理人来说，一个好的因子模型应该首先是一个风险模型**。相较于学术界关注因子模型在解释资产截面预期收益率差异方面的作用，管理人更关注因子模型在时序上能否和资产回归得到更高的 R-squared，从而更好的计算风险（比如获得更准确的协方差矩阵）。

在计算风险方面，Barra 模型无疑是最有名的（《正确理解 Barra 的纯因子模型》、《Barra 因子模型截面回归求解》）。这解释了为什么在学术界的模型中，我们最多只看到 Fama-French 三因子、五因子模型这些，而在 Barra 的模型中（以 CNE6 为例），我们看到了国家

因子、行业因子以及 9 大类风格因子（包括 16 个细分因子）——模型中因子总数远超学术界因子模型中的数量。

Cerniglia and Fabozzi (2018) 一文从以下四个方面阐述了管理人视角下的因子投资：

1. 越来越多资金流入对因子投资造成的冲击；
2. 因子择时；
3. 因子投资大行其道之时，如何找到仍能够带来  $\alpha$  的管理人；
4. 创新（新数据、新技术如机器学习）能带来多大优势。

### 3.1 资金流入造成冲击

当越来越多的资金涌入因子投资时就会造成因子拥挤（factor crowding）。而 rule-based（使用相似的指标排序、接近的调仓频率）的因子投资则加剧了这种负面影响。任何一个投资策略想要持续赚钱都是利用了市场在某方面的非有效性。当使用的人越来越多，市场在这方面就会变得更有效，从而降低了该因子获取收益的能力。

因子拥挤也会造成流动性冲击。一旦市场中发生冲击风格因子的事件后，持有相似头寸的管理人会竞相卖出手中的股票，由此产生的流动性危机会造成很大的亏损。2007 年 8 月，美股市场上一些非常优秀的量化对冲基金在短时间内录得了巨大的亏损。Khandani and Lo (2011) 对此进行了研究并发现，很多量化基金经理在短时间内清理了相似的头寸，巨大的抛压对流动性造成了巨大的打击、使得这些股票的价格在短时间内大幅下降。不幸的是，**严重的流动性事件是低频事件，这意味着我们很难对其造成的负面影响有效建模。**

### 3.2 因子择时

我知道因子择时是个 hot potato。所以在这小节的论述中会加点料。

之前的文章《还在对着一阶矩做因子择时？不妨试试二阶矩》从风险的角度介绍了海外顶级机构对因子择时的处理方法。今天我们从一阶矩（收益）的角度聊聊因子择时。

因子择时是个大课题，这一小节的篇幅也肯定说不完。所以先建议各位小伙伴参阅 Bender et al. (2018) 这篇文章。其领衔作者 —— State Street 的 Jennifer Bender —— 在因子投资界是响当当的人物。该文客观的综述了使用不同预测指标、在不同时间尺度下对不同风格因子择时的效果。这些预测指标包括因子估值、因子动量以及各大类宏观经济指标。

谈到因子择时，不得不提的两大门派是**因子动量 (factor momentum)** vs **因子估值 (factor valuation)**。本节重点来聊聊按估值进行因子择时是否靠谱。以下的论述来自 Lee (2017) 这篇 Factors timing factors (用因子择时因子) 的文章，颇有意思。

按因子估值择时由 Arnott et al. (2016) 提出。它指的是使用某种估值指标（比如 PB）来计算因子组合中多、空两头的估值，并以二者的比值作为因子本身的估值；认为因子未来的收益率和因子当前的估值呈现负相关 —— 即昂贵的因子在未来更有可能获得负收益。从估值的计算来看不难发现，**因子变得昂贵了意味着该因子和价值因子负相关。**

因此问题就来了：**和价值因子负相关就一定会在未来获得负收益吗？**

乍一看，这个问题的答案似乎是肯定的。由于价值因子长期有效，因此如果一个因子和它呈现负相关，那它不就是只能获得负收益了吗？然而仔细想想，答案并没有这么简单。**最简单的反例就是动量因子，该因子以和价值因子长期负相关而著称。**Asness, Moskowitz, and Pedersen (2013) 这篇著名的 Value and Momentum Everywhere 通过大量实证指出价值和动量长期存在于全球的股票、外汇、固定收益、商品期货等市场中。

为了回答这个问题，首先要分清待使用估值来择时的因子是 nonreplicable（不可复制）还是 replicable（可复制）。**如果该因子无法被现有多因子模型中的因子解释（即加入这个新因子能够提高 mean-variance 有效前沿），则它被称为一个 nonreplicable 因子；反之它是一个 replicable 因子。**

对于一个价值因子无法解释的 nonreplicable 因子，显然即便它和价值因子呈现负相关（所谓的估值高），我们也不能说这个因子要失效了。动量因子正是这样一个 nonreplicable 因子。

对于 replicable 因子，价值因子无疑能解释其获得的一部分收益。但不要忘了，该因子还在其他已知因子上有正的暴露。该因子的收益率由所有用来复制它的其他因子的收益率以及该因



子在它们上面的暴露决定；因此只因估值过高就说一个因子未来收益率要如何如何总是有些草率的。

使用估值来进行因子择时这个争议的核心问题是：**是什么让价值因子如此与众不同？为什么我们就偏偏选了它来对其他因子择时？按照这个思路，是不是也可以把所有其他因子（投资、质量、市值）都用来择时？**归根到底，每个因子都是一个投资组合，因此可以计算它的多空两头在其他因子上的取值。

颇有意思的是，在 Fama and French (2015) 这篇提出五因子模型的文章中，作者意味深长的写到：**当加入 profitability 和 investment 因子之后，价值因子变得多余了。**

#### A B S T R A C T

A five-factor model directed at capturing the size, value, profitability, and investment patterns in average stock returns performs better than the three-factor model of Fama and French (FF, 1993). The five-factor model's main problem is its failure to capture the low average returns on small stocks whose returns behave like those of firms that invest a lot despite low profitability. The model's performance is not sensitive to the way its factors are defined. With the addition of profitability and investment factors, the value factor of the FF three-factor model becomes redundant for describing average returns in the sample we examine.

© 2014 Elsevier B.V. All rights reserved.

毫无疑问，因子估值能否择时仍然会是未来争论的焦点。

### 3.3 区分 $\alpha$ 和 $\beta$ 收益

随着因子投资的盛行，一个新的挑战油然而生：**主动管理人获取的超额收益到底是  $\alpha$  还是风格  $\beta$ ？**搞清楚这个问题对于成功甄别管理人至关重要，因此它也备受 MOM / FOF 的关注。Bender, Hammond, and Mok (2014) 指出，管理人获得的超额收益中，80% 的部分可以由因子暴露解释，只有剩下 20% 才取决于获得  $\alpha$  的能力。

**由于不同风格因子的表现受到宏观经济的影响，主动配置因子  $\beta$ （涉及到因子择时和风险控制）从而获得超额收益也彰显管理人的能力。**在一项最新的研究中，Duanmu, Malakhov, and McCumber (2018) 比较了主动  $\beta$  和主动  $\alpha$  管理人，并指出**长期来看，最顶尖的主动  $\beta$  型管理人能够获得比顶尖的主动  $\alpha$  型管理人更优异的收益。**



### 3.4 创新的重要性

对于任何研究领域，创新的作用都至关重要。**对因子投资来说，创新意味着使用新的数据或者算法。**随着大数据的普及，越来越多的管理人开始搜寻新的能够成为收益源的数据（比如舆情数据、专利数据、新闻数据等）。此外，机器学习算法也被广泛的应用于因子投资实践中，包括使用机器学习算法预测基本面财务指标（见《如果我们能正确预测基本面因子》），或挖掘因子和收益率之间的非线性关系。

正是考虑到数据和技术的革新对金融领域的影响，2019 年 Institutional Portfolio Research Journals (IPR Journals) 为旗下的期刊系列增加了一位最新成员 —— Journal of Financial Data Science，旨在指导金融领域的实践者正确使用与日俱增的数据和日新月异的技术。

当然，创新的机遇总是伴随着巨大的挑战。虽然机器学习在金融领域以及因子投资领域拥有乐观的前景，但由于金融数据信噪比很低，我们也需要时刻保持冷静（见《AI 投资言过其实》、《为什么机器学习在投资领域不好使》）。Lopez de Prado (2018) 详细论述了在投资领域实践机器学习算法时可能遇到的各种坑，填补了理论和实践之间的空白。

## 4

### 投资者视角

因子投资的普及造就了越来越多的 smart beta ETF 产品。2018 年，《经济学人》(The Economist) 估计这些因子 ETFs 的总规模超过 6500 亿美元。**大量低成本的 ETFs 产品让普通投资者也能享受风格因子带来的收益 —— 当然前提是能够选择适当的 ETFs。**然而，这并不容易。

在一项针对大学本科生和研究生的实验中，Choi, Laibson, and Madrian (2009) 让参与者从众多追踪标普 500 指数的被动型 ETFs 中挑出它们认为最好的。在实验中，Choi, Laibson, and Madrian (2009) 为参与者提供了这些 ETFs 的费率以及在不同时期（故意的）的收益表现情况。作为具备金融知识的理性投资者来说，这些 ETFs 均是被动的追踪标普 500 指数，因此应该把费用的高低作为唯一的标准来选择。然而，实验者被不同阶段的 ETFs 的收益率差异所迷惑，绝大多数投资者忽略了 ETFs 的费用，而选择了收益率最高（得益于适当的数据窗口）的 ETF。

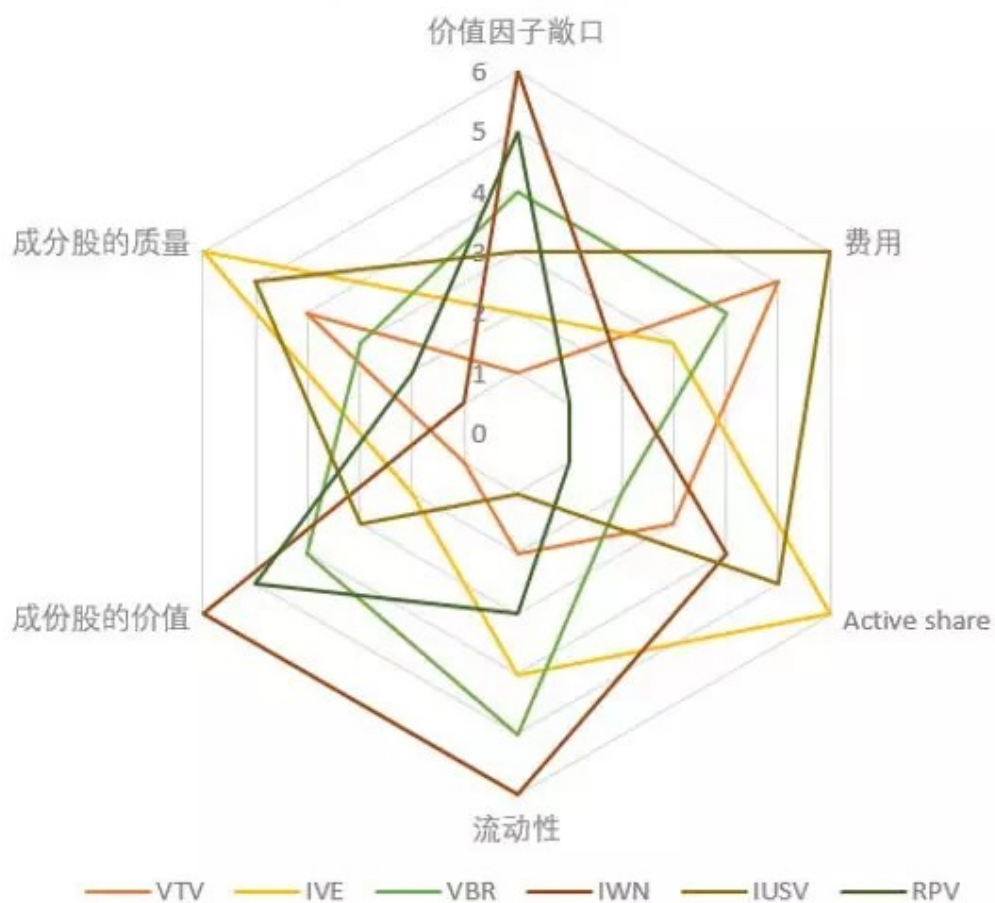
另一方面，就算是排除了投资者本身的非理性行为，琳琅满目的 ETFs 也足以让人迷失。徐杨（老虎证券合伙人）曾在 Wind 举办的万矿嘉年华介绍了美股上的价值投资实务。下表汇总了他提及的主流价值因子的 ETFs。

代码	名称	成立日期	总规模（亿美元）	持仓数量	管理费
VTV	先锋价值	2004 / 01 / 26	384.17	344	0.05 %
IWD	iShares 罗素 1000 价值	2000 / 05 / 22	377.75	733	0.20 %
IVE	iShares 标普 500 价值	2000 / 05 / 22	163.63	388	0.18 %
VBR	先锋小盘价值	2004 / 01 / 26	138.28	854	0.07 %
IWS	iShares 罗素中盘价值	2001 / 07 / 17	109.50	600	0.25 %
IWN	iShares 罗素 2000 价值	2000 / 07 / 24	100.56	1358	0.24 %
VOE	先锋中盘价值	2006 / 08 / 17	88.54	212	0.07 %
PRF	景顺 FTSE RAFI 美国 1000	2005 / 12 / 19	54.58	994	0.39 %
SCHV	嘉信美国大盘价值	2009 / 12 / 11	44.07	349	0.04 %
IUSV	iShares 核心标普美国价值	2000 / 07 / 24	37.61	683	0.04 %
VLUE	iShares Edge MSCI 美国价值因子	2013 / 04 / 16	35.31	155	0.15 %
MGV	先锋超大盘价值	2007 / 12 / 17	20.29	155	0.07 %
SPYV	SPDR 组合标普 500 价值	2000 / 09 / 25	14.31	386	0.04 %
RPV	景顺标普 500 纯价值	2006 / 03 / 01	9.19	110	0.35 %
DVP	深度价值	2014 / 09 / 22	1.50	21	0.59 %

出处：新全球资产配置（公众号）

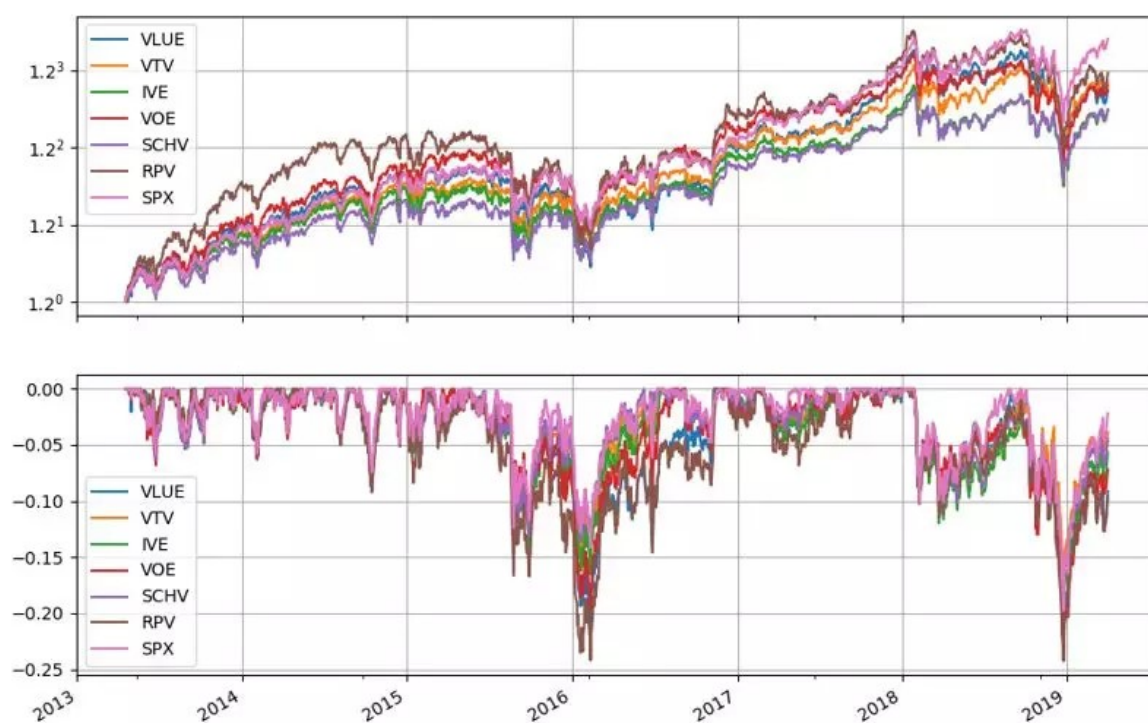
针对一些具有代表性的 ETFs，徐杨从因子敞口、费用等六个方面进行了分析。从下图中不难看出，即便都是价值因子 ETFs，它们的风格特点也截然不同——有些是更纯粹的价值投资、有些则注重于成分股的质量。

## 价值 ETFs 综合评价



出处：新全球资产配置（公众号）

我从上表中的 ETFs 中挑选了一些针对标普 500 构建价值因子的标的。下图比较了它们的表现——只能说 **not all ETFs are made equal**。



数据来源：Wind

下表总结了这些 ETFs 和标普 500 指数 (SPX) 在这段时间内的表现。由于最近几年价值投资不太好使，这些因子均跑输 SPX 本身。但即便如此，它们之间也有明显的差异。从夏普率的角度来说，最好的要数 VTV (Vanguard 价值 ETF)、排在后面的则是 IVE 和 RPV。不幸的是，对于投资者来说，事前就从这些琳琅满目的 ETFs 中挑出最好的并不容易。

Smart beta ETFs 的流行给投资者提供了越来越丰富的工具，也对投资者提出了更高的要求。在使用这些 ETFs 时，投资者需要首先明确自己的目标——是分散化风险还是获得相对于市场的超额收益。在明确目标之后，需要理解每个风格因子背后的逻辑和它代表的风险。唯有这样，才有可能享受这些标的带来的更高性价比的风险收益。

毫无疑问，因子投资已经成为了投资中的标配。

本文借 Cerniglia and Fabozzi (2018) 的花献佛，并加入了大量背景知识，从学术界、管理人和投资者三个视角介绍了因子投资。

历经几十年来的发展，因子投资从一个理论研究框架到如今如雨后春笋般诞生的大量投资工具，加深了我们对于股票市场的理解，也造福了广大参与者。而无论是身为学者、管理人还是投资者，我们需要做的是继续拥抱和适应它的发展与变革。

### 参考文献

Arnett, R. D., N. Beck, V. Kalesnik, and J. West (2016). *Timing 'Smart Beta' Strategies? Of Course! Buy Low, Sell High!* Available at SSRN: <https://ssrn.com/abstract=3040956>.

Asness, C. S., A. Frazzini, and L. H. Pedersen (2019). Quality minus junk. *Review of Accounting Studies*, Vol. 24(1), 34 – 112.

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen (2013). Value and Momentum Everywhere. *Journal of Finance*, Vol. 68(3), 929 – 985.

Banz, R.W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, Vol. 9(1), 3 – 18.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: a test of the efficient market hypothesis. *Journal of Finance*, Vol. 32(3), 663 – 682.

Bender, J., P.B. Hammond, and W. Mok (2014). Can alpha be captured by risk premia? *The Journal of Portfolio Management*, Vol. 40(2), 18 – 29.

Bender, J., X. Sun, R. Thomas, and V. Zdorovtsov (2018). The promises and pitfalls of factor timing. *The Journal of Portfolio Management*, Quantitative Special Issue, Vol. 44(4), 79 – 92.

Cerniglia, J. and F. J. Fabozzi (2018). Academic, practitioner, and investor perspectives on factor investing. *The Journal of Portfolio Management*, Vol. 44(4), 10 – 16.

Claessens, S. and M.A. Kose (2018). *Frontiers of macrofinancial linkages*. Paper no. 95, Bank for International Settlement. <https://www.bis.org/publ/bppdf/bispap95.htm>.

Choi, J.J., D. Laibson, and B.C. Madrian (2009). Why does the law of one price fail? An experiment on index mutual funds. *Review of Financial Studies*, Vol. 23(4), 1405 – 1432.

Duanmu, J., A. Malakhov, and W. R. McCumber (2018). Beta active hedge fund management. *Journal of Financial and Quantitative Analysis*, Vol. 53(6), 2525 – 2558.

*The Economist*. Maxing the factors. February 1, 2018.

Fama, E. F. (2017). *The Fama portfolio, selected papers of Eugene F. Fama*. J. H. Cochrane and T. J. Moskowitz (Eds), The University of Chicago Press.

Fama, E. F., L. Fisher, M. C. Jensen, and R. Roll (1969). The adjustment of stock prices to new information. *International Economic Review*, Vol. 10(1), 1 – 21.

Fama, E. F. and K. R. French (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, Vol. 33(1), 3 – 56.

Fama, E. F. and K. R. French (2015). A Five-Factor Asset Pricing Model. *Journal of Financial Economics*, Vol. 116(1), 1 – 22.

Fama, E. F. and J. D. MacBeth (1973). Risk, Return, and Equilibrium: Empirical Tests. *The Journal of Political Economy*, Vol. 81(3), 607 – 636.

Feng, G., S. Giglio, and D. Xiu (2017). *Taming the factor zoo*. Chicago Booth research paper no. 17-04, SSRN: <https://ssrn.com/abstract=2934020>.

Graham, B. and D. Dodd (1934). *Security analysis*. New York: McGraw-Hill.

Green, J., J. R. M. Hand, and X. F. Zhang (2017). The characteristics that provide independent information about average U.S. monthly stock returns. *Review of Financial Studies*, Vol. 30(12), 4389 – 4436.

Harvey, C. R. and Y. Liu (2018). *Lucky Factors*. Working paper, available at SSRN: <https://ssrn.com/abstract=2528780>.

Harvey, C. R., Y. Liu, and H. Zhu (2016). ... and the cross-section of expected returns. *Review of Financial Studies*, Vol. 29(1), 5 – 68.

Hou, K., C. Xue, and L. Zhang (2017). *Replicating Anomalies*. Fisher College of Business Working Paper No. 2017-03-010; Charles A. Dice Center Working Paper No. 2017-10. Available at SSRN: <https://ssrn.com/abstract=2961979>.

Khandani, A.E. and A.W. Lo (2011). What happened to the quants in August 2007? Evidence from factors and transactions data. *Journal of Financial Markets*, Vol. 14(1), 1 – 46.

Lee, W. (2017). Factors timing factors. *The Journal of Portfolio Management*, Vol. 43(5), 66 – 71.

Linnainmaa, J. T. and M. R. Roberts (2018). The history of the cross-section of stock returns. *Review of Financial Studies*, Vol. 31(7), 2606 – 2649.

Lopez de Prado, M. (2018). *Advances in financial machine learning*. Hoboken, NJ: John Wiley & Sons.

McLean, R.D. and J. Pontiff (2016). Does academic research destroy stock return predictability? *Journal of Finance*, Vol. 71(1), 5 – 32.



Mohanram, P.S. (2005). Separating winners from losers among low-book-to-market stocks using financial statement analysis. *Review of Accounting Studies*, Vol. 10(2-3), 133 – 170.

Novy-Marx, R. (2015). *Backtesting strategies based on multiple signals*. NBER Working Paper, No. 21329.

Novy-Marx, R. and M. Velikov (2015). A taxonomy of anomalies and their trading costs. *Review of Financial Studies*, Vol. 29(1), 104 – 147.

Piotroski, J. D. (2000). Value investing: the use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*, Vol. 38, 1 – 41.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, Vol. 13(3), 341 – 360.

