

知乎

首发于
川流不息

Book-to-Market 整容记：哪家 BM 比较强？



石川

量化交易 话题的优秀回答者

已关注

28 人赞同了该文章

本文由搞事情因子小组创作。小组成员：llanglli（因子动物园）；刀疤连（Chihiro Quantitative Research）；石川

摘要

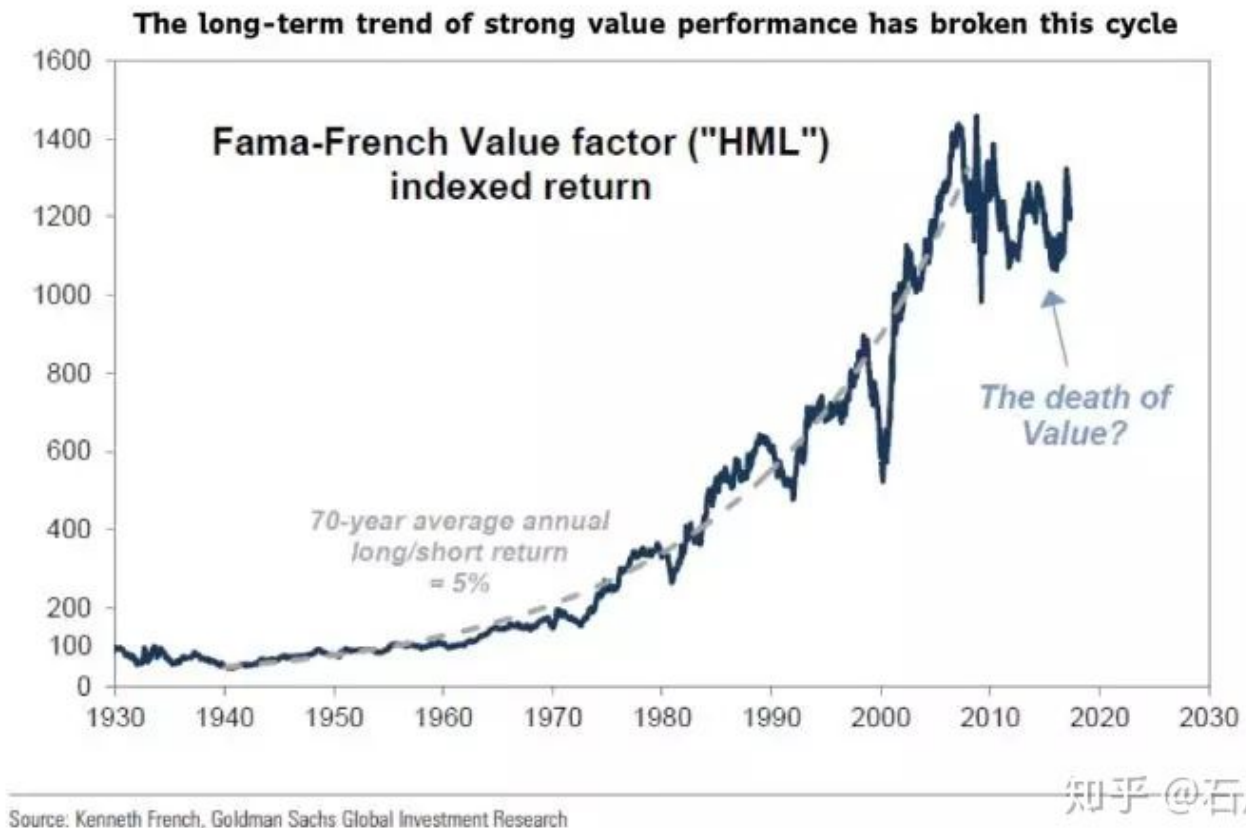
近年来学术界对 BM 进行了很多改造。本文针对 A 股进行相关实证。结果表明，利用 R&D 和 SG&A 改造后的 BM 可以在股票定价时提供增量信息。

1 引言

Fama and French (1993) 三因子在学术界和业界影响深远，使得用 Book-to-Market Ratio (BM) 来构建的 HML 因子被投资者熟知。然而，仅仅依据简单如 BM 这样的指标真的可以构建一个优秀的价值因子吗？

早在 80 年前，Graham 和 Dodd 就提醒投资者**不应该使用 Book Value 来衡量公司的内在价值**。但随着因子投资的深入人心（见《基本面分析≠基本面量化投资?》），以 BM 来选股的主策略和使用 BM 来构建的风格指数还是层出不穷，BM 这个指标也早已变得异常拥挤。下图显示了

知乎

首发于
川流不息

面对这种情况，学术界和业界开始重新审视 BM 背后的逻辑——BM 中到底哪部分代表了公司面临的风险，因而和未来的收益率挂钩；哪部分又仅仅是噪音而已。带着这个问题，学者们展开了对 BM 轰轰烈烈的改造大潮，很多最新的发现相继被发表于会计学和金融学的顶级期刊。

本文介绍这些对 BM 的改造并在 A 股上进行深度实证。由于知识和经历有限，我们无法覆盖所有相关研究，而会把关注点放在那些剖析 BM 内在逻辑的改造上，检验哪个改造后的 BM 指标具有最强的解释力。

本文的行文组织如下：第二节介绍几种对 BM 的改造逻辑；第三节说明实证数据和因子构造方法；第四节在 A 股中对改造前后的 BM 进行 Portfolio Sort Test；第五节比较不同候选 BM，考察哪种改造的效果更好；第六节检验挑选出的候选 BM 因子在 A 股上是否被定价；第七节总结全文。

这是 [搞事情因子小组] 的第二次合作。希望文中深度实证分析能给各位带来一些启发。

2 学术界对 BM 的改造

本小节按论文发表时间顺序梳理下列学术界对 BM 的改造方法。



知乎

首发于
川流不息

Frankel and Lee (1998)	使用历史收益法计算公司的内在价值	• Intrinsic Value-to-Market
Chan, Lakonishok, and Sougiannis (2001)	使用 R&D 以及 SG&A 等投入修正 Book Value	• R&D-to-Market • R&D+SG&A-to-Market
Penman, Richardson, and Tuna (2007)	剔除杠杆影响，计算经营活动对应的 BM	• NOA Book-to-Market
Ball et al. (2019)	将 Book Value of Equity 拆分成 Retained Earnings 和 Contributed Capital 分别计算	• Retained Earnings-to-Market • Contributed Capital-to-Market

出处：搞事情因子小组

知乎 @石州

2.1 Frankel and Lee (1998)

最初，Fama and French (1992) 系统地研究了 BM，并利用它构建了三因子模型中的 HML (Fama and French 1993)。Fama 和 French 对 BM 的解释是它是公司 distress risk 的一个代理变量，不过关于它是否以及从多大程度上代表了公司面临的风险一直有争论。

Frankel and Lee (1998) 是一篇较早的对 BM 进行改造的文章（其中的一位作者 Charles Lee 小伙伴们一定都不陌生）。该文另辟蹊径，并没有从风险代理变量的角度审视 BM，而是尝试利用会计学的知识对公司进行更准确的估值。

从剩余收益模型出发，该文建立了公司内在价值和 BM 的关系：

$$V_t = B_t + \sum_{i=1}^{\infty} \frac{E_t[(ROE_{t+i} - r_e)B_{t+i-1}]}{(1 + r_e)^i}$$

其中，下标 t 表示时间，B 为 Book Value， r_e 为公司的 cost of equity capital。上式说明，公司的内在价值由两部分组成：Book Value 和未来剩余收益的现值。如果能够对上式第二部分进行较准确的预测，基于 V 计算的 VM 会比 BM 更准确的反映公司的估值水平（这个 VM 称为 Intrinsic Value-to-Market）。

为了计算 V，除了公式中的 r_e 、ROE 这些变量之外，还有一个隐含的变量是股息率，因为它会影响未来的 Book Value。此外，上式是一个无穷级数，如果要在实际中使用它来选股，必须要把它转化成有限项。经过数学推导，Frankel and Lee (1998) 将上式简化成两个版本——使用历史 ROE 和使用分析师预测 ROE 的版本。

在使用历史 ROE 的版本中，该文使用最新 ROE 外推两期作为未来 ROE 的 naïve estimate，得到的 V 如下：

$$V_t = B_t + \frac{ROE_t - r_e}{1 + r_e} B_t + \frac{ROE_t - r_e}{(1 + r_e)r_e} B_{t+1}$$



知乎

首发于
川流不息

$$V_t = B_t + \frac{F'ROE_t - r_e}{1 + r_e} B_t + \frac{F'ROE_{t+1} - r_e}{(1 + r_e)^2} B_{t+1} + \frac{F'ROE_{t+2} - r_e}{(1 + r_e)^2 r_e} B_{t+2}$$

在 Frankel and Lee (1998) 基于美股的实证中，两位作者主要考察了使用分析师预测 ROE 的版本，指出它的效果要远远好于 BM。

2.2 Chan, Lakonishok, and Sougiannis (2001)

Chan, Lakonishok, and Sougiannis (2001) 在 Journal of Finance 上撰文讨论了费用化的研发 (R&D) 投入可能会干扰 PE 和 PB 等估值指标。在美国，科技类公司拥有举足轻重的地位。例如 1999 年末科技和制药类公司占标普 500 指数总市值的 50%。科技类公司的一大特点就是研发投入特别大，甚至会超过净利润。研发投入长期来看有利于公司发展，潜在的收益巨大，但市场反应了研发投入吗？

如果一个市场是有效的，当前的股票价格已经包含了研发投入的信息，研发投入和未来股价之间并没有什么关系。但事实可能并非如此。一方面研发周期往往时间较长，研究成果落地也充满不确定性，因此市场很难对研究投入进行合理估值；另一方，投资者习惯使用 PE 和 PB 等简单相对估值指标。由于会计学中的审慎原则 (conservative accounting)，会计上会把研发投入做费用化处理，导致上述估值指标偏大、造成股票“被高估”的错觉。

除此之外，Mohanram (2005) 在构造 G-SCORE 时在 R&D 之外也讨论了 SG&A。对于广告支出和研发投入的费用化处理，会造成市场往往会低估这部分投入的价值。由于这些费用在将来可能会带来利润和现金流，因此存在超额收益。

最后，纽约大学的 Baruch Lev 教授在 Lev (2019) 中表达了类似的观点。他指出应在计算 BM 时考虑 R&D 和 SG&A。Lev (2019) 认为，在当今的企业经营中无形资产投资大行其道，而美国公认会计准则对无形资产的计量方法存在重大缺陷，因此账面价值严重误导了对公司净资产科目的核算，导致市净率的计算非常不准确。

研发和广告等开支不应按照财会人员以为的那样被计入当期费用科目，而应被视作是为了未来成长所做的投资，其本质上也是一种资产，应被算做账面价值的一部分。Lev (2019) 对所有美国上市公司的全部 R&D 和部分 SG&A 作了资本化处理，对账面价值进行调整从而改造并优化了 BM。其研究表明，考虑了费用化 R&D 和 SG&A 的 BM 指标显示出强劲的生命力。

基于上述文献，本文加入两个与 R&D 和 SG&A 相关的改造版 BM，它们分别为 R&D 调整后账面价值市值比和 R&D + SG&A 调整后账面价值市值比，在后文记为 R&D 和 R&D+SG&A。

2.3 Penman, Richardson, and Tuna (2007)

Penman, Richardson, and Tuna (2007) 从资产负债表出发对 BM 进行了改造。考虑下面这个^{简化版}的资产负债表：



知乎

首发于
川流不息

+ 经营性资产 (OA)	+ 经营性负债 (OL)
+ 金融性资产 (FA)	+ 金融性负债 (FL)
= 总资产 (TA)	+ 股东权益 (B)
	= 总资产 (TA)

知乎 @石川

资产负债表的左侧可以分为经营性资产和金融性资产，二者之和为总资产；而资产负债表的右侧负债部分可以分为经营性负债和金融性负债；二者之和再加之股东权益就等于总资产——会计恒等式。

下面把上表中左右各一项调换一下位置：把经营性负债从右侧挪到左侧，将金融性资产从左侧挪到右侧，则变成下面这个样子：

经营性项目	金融性项目
+ 经营性资产 (OA)	+ 金融性负债 (FL)
- 经营性负债 (OL)	- 金融性资产 (FA)
= 净经营资产 (NOA)	+ 股东权益 (B)
	= 净经营资产 (NOA)

知乎 @石川

经过上述变换，表中左侧变成了“经营性项目”、右侧变成“金融性项目”；其中经营性项目中，经营性资产和经营性负债的差值为净经营资产 (NOA)；而右侧金融性负债减去金融性资产为净债务 (Net Debt, 简称 ND)。从上述关系出发有，得到 Book Value of Equity 的表达式：

$$B = NOA - ND$$

与之相对应的，我们可以写出以上三部分对应的市值的关系：

$$P = P^{NOA} - P^{ND}$$

其中，P 表示 Market Value of Equity； P^{NOA} 和 P^{ND} 分别表示净经营资产的市值（代表了 enterprise value）和净债务的市值。根据会计学准则，债务的 book value 一般就是其 market value，因此有 $P^{ND} = ND$ 。接下来，将 $B = NOA - ND$ 左右两侧同时除以 P 可得：



知乎

首发于
川流不息

$$= \frac{P^{NOA}}{P} \times \frac{NOA}{P^{NOA}} - \frac{ND}{P}$$

利用 $P = P^{NOA} - P^{ND}$ 以及 $P^{ND} = ND$ ，经过简单的代数运算，上述 BM 最终可以写成：

$$\frac{B}{P} = \frac{NOA}{P^{NOA}} + \frac{ND}{P} \left(\frac{NOA}{P^{NOA}} - 1 \right)$$

Penman, Richardson, and Tuna (2007) 对上式的解读为，BM 是上式右侧两部分的加权平均：其中第一项是 Enterprise Book-to-Market，而第二项则受融资活动影响。当投资者按 BM 来判断公司价值时，该指标不仅反映了来自经营活动的贡献，也包括了来自融资活动的杠杆水平。因此，BM 又称作 levered BM，它同时反映了上市公司的经营风险和财务风险。

由于上面公式中存在非线性，因此财务杠杆对于 BM 的影响是 mixed。Penman, Richardson, and Tuna (2007) 表明，BM 中真正和未来收益率相关的是第一项，即和经营活动对应的 Enterprise Book-to-Market。他们以此作为对 BM 的，即 NOA/P^{NOA} 。

2.4 Ball et al. (2019)

2019 年，会计学领域大牛 Ray Ball 领衔在 Journal of Financial Economics 上发表了一篇对 BM 改造的文章 (Ball et al. 2019)。该文认为 Book Value of Equity 主要由两部分组成，即 retained earnings (未分配利润) 和 contributed capital (资本公积)：

$$\text{Book Value of Equity} \approx \text{Retained Earnings} + \text{Contributed Capital}$$

先来看 contributed capital。Ball et al. (2019) 指出，其反应了上市公司和股东之间的净资本交易量 (net capital transaction；计算可用 net issuances)，并不能很好地揭示公司所面临的风险；它的高低仅仅表明投资者的风险偏好，而并不直接和股票的未来收益率挂钩。另一方面，学术界的很多研究表明，net issuances 和未来的收益率呈现出负相关的关系。因此，结合上面的论述，未来收益率和 contributed capital 相关性往往很低，甚至是负相关的。

再来看 BM 里的另一项 retained earnings。它代表了过去一段时间上市公司累积净利润和累积分红之差。Ball et al. (2019) 认为正是 BM 中的这部分才是和未来的收益率呈现正相关。这其中有两个原因。

首先，retained earnings 反映的是过去一段时间累积的净利润，因此如果除以时间，它反映的实际上是一个平均的概念。由于会计数据比较严重的滞后性（比如到了 2017 年中才能知道 2016 年的情况）以及利润数据的回复的特性，均值能够很好的平抑掉不同财年数据的波动以及滞后性影响，因此更好的反映长期来看公司稳定、真实的盈利水平。

驱动是相似的，因此 Ball et al. (2019) 认为 BM 中的 retained earnings 这部分才和收益率正相关。

依照上述论述，Ball et al. (2019) 认为 Book Value of Equity 中的 retained earnings 才是导致 BM 有效的直接原因。他们使用 Retained Earnings-to-Market 对 BM 进行了改造，并通过大量针对美股的实证分析证明了上述猜想。

类似的，Meredith (2016) 也指出上市公司的回购和增发会干扰到 BM 作为估值因子的价值。和其他估值指标相比，回购可能会过分夸大 BM、增发可能会低估 BM。近年来，回购和增发变得越来越常见和频繁，可能是 BM 表现不好的罪魁祸首之一，这和 Ball et al. (2019) 的结论相似。

3 数据和因子构造

本文的目的是在 A 股上对改造后的 BM 以及原始 BM 进行实证分析。对于原始 BM，这里考虑两个版本，分别为 Fama and French (1993) 的学术版，即使用 Most Recent Annual (MRA) 数据每年更新一次（记为 MRA 数据计算的 BM）；以及仿照 Asness and Frazzini (2013) 的灵活版本，即使用 Most Recent Quarter (MRQ) 数据更新的 BM（记为 MRQ 数据计算的 BM），它可以更快捕捉到最新的信息。

在计算 R&D 时，由于研发有两种处理方法，资本化放在资产负债表中的“开发支出”和“无形资产”中，费用化放到利润表中的管理费用下，前者已经进入账面资产，因此只考虑费用化后的研发投入。最后，在计算 R&D+SG&A 时，SG&A 相当于国内的销售费用和管理费用，且研发费用已经包含在了管理费用中。

下表给出了两种原始 BM 以及六种改造 BM 的具体计算公式。

因子	中文名	计算公式
MRA	最近年度账面价值市值比	账面价值市值比 = 所有者权益 / 总市值
MRQ	最近报告期账面价值市值比	账面价值市值比 = 所有者权益 / 总市值
Retained Earnings	最近报告期留存收益市值比	留存收益市值比 = 留存收益 / 总市值 留存收益 = 未分配利润 + 盈余公积金
Contributed Capital	最近报告期实缴资本市值比	实缴资本市值比 = 实缴资本 / 总市值 实缴资本 = 股本 + 资本公积金 - 库存股
NOA	最新报告期净经营资产除以净经营资产市场价值	净经营资产除以净经营资产市场价值 = 净经营资产 / 净经营资产市场价值 净经营资产 = 所有者权益 + 金融负债 - 金融资产 净经营资产市场价值 = 总市值 + 金融负债 - 金融资产
R&D	最近 12 个月 R&D 调整后账面价值市值比	R&D 调整后账面价值市值比 = (费用化研发投入 + 所有者权益) / 总市值
R&D+SG&A	最近 12 个月 R&D 和 SG&A 调整后账面价值市值比	R&D 和 SG&A 调整后账面价值市值比 = (销售费用 + 管理费用 + 所有者权益) / 总市值
Intrinsic Value	最近年度剩余收益模型内在价值市值比	剩余收益模型内在价值市值比 = 剩余收益模型内在价值 / 总市值 为了简化计算，这里采用历史外推的方法。假设未来两年 ROE 和本年相同；另外，假设股利支付率为 30% (参考莫尼塔的日报报告)

出处：tushare，搞事情因子小组

相关系数。总体来看，由于均是从原始 BM 改造而来，改造后的 BM 和原始 BM 相关性较高，尤其是 NOA、R&D 以及 R&D+SG&A，而 Intrinsic Value 和原始 BM 相关性最低。值得一提的是，根据定义，利用 Intrinsic Value 改造的 BM 指标仅使用年报数据；而其他改造版本使用的是季报数据，这部分解释了上述相关性的差异。

	MRA	MRQ	Retained Earnings	Contributed Capital	NOA	R&D	R&D+SG&A	Intrinsic Value
MRA	1.00	0.81	0.56	0.59	0.72	0.82	0.80	0.29
MRQ	0.81	1.00	0.70	0.76	0.85	0.98	0.96	0.42
Retained Earnings	0.56	0.70	1.00	0.21	0.56	0.68	0.67	0.63
Contributed Capital	0.59	0.76	0.21	1.00	0.70	0.75	0.75	0.13
NOA	0.72	0.85	0.56	0.70	1.00	0.86	0.85	0.31
R&D	0.82	0.98	0.68	0.75	0.86	1.00	0.98	0.40
R&D+SG&A	0.80	0.96	0.67	0.75	0.85	0.98	1.00	0.40
Intrinsic Value	0.29	0.42	0.63	0.13	0.31	0.40	0.40	1.00

出处：tushare，搞事情因子小组

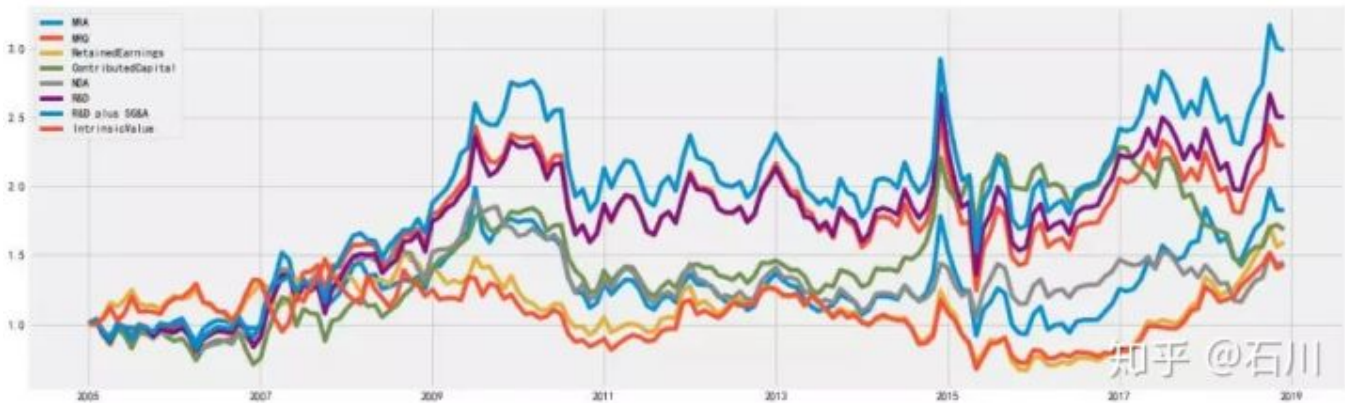
知乎 @石川

接下来进行 Portfolio Sort Test。

4 Portfolio Sort Test

在实证期内的每个月末，按照每个候选 BM 指标取值将个股从小到大排列、等分为 10 组，做多指标最大的第 10 组，同时做空指标最小的第 1 组，以此构造该 BM 指标的多空对冲因子组合。在这个过程中，剔除掉次新股、停牌股、风险警示股、一字停牌股，并同时考虑等权和市值加权两种情况。

按市值加权时，上述 8 个 BM 的因子组合在实证期内的累积净值走势如下图所示。

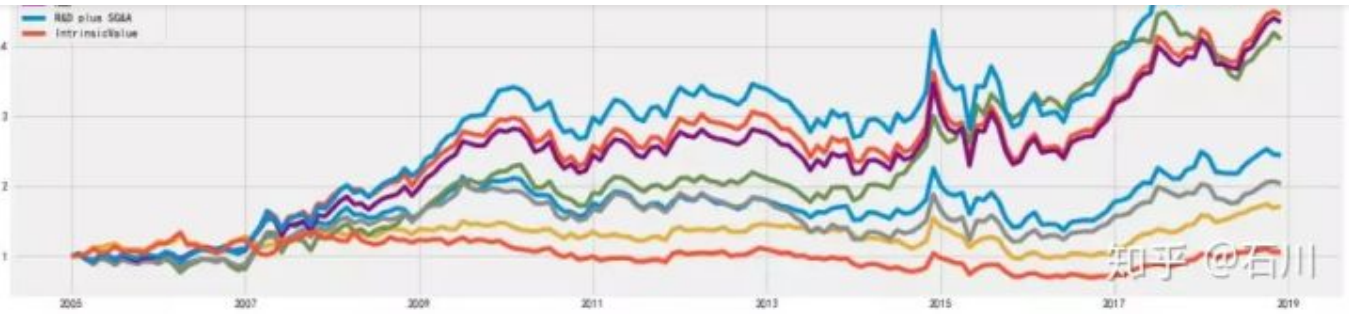


知乎 @石川

等权时，这些 BM 的因子组合的累积净值走势如下。



知乎

首发于
川流不息

从结果中可以看到，无论是等权重还是市值加权，R&D+SG&A 的表现最为优秀，而垫底的均是 Intrinsic Value 改造的版本。此外，R&D 版本能略微跑赢使用 MRQ 数据计算的原始 BM，但效果并不十分显著。最后，所有因子在市值加权下均不如等权，说明小市值会对影响 BM 因子组合的表现。

以等权构建的因子为例，下表给出了这 8 个 BM 因子组合的 ICIR。除了 Intrinsic Value 外，其他所有因子的 IC 和 IR 均高于 Fama and French (1993) 版的 BM（即使用 MRA 数据计算的 BM）。其原因是因为 MRA 只考虑年报和年末数据，使得决策判断的信息来源过于滞后。

	IC	ICSTD	IR	t-statistic
MRA	0.0290	0.1342	0.2159	2.7906
MRQ	0.0492	0.1462	0.3368	4.3519
Retained Earnings	0.0310	0.1225	0.2531	3.2708
Contributed Capital	0.0471	0.1492	0.3159	4.0828
NOA	0.0363	0.1596	0.2278	2.9439
R&D	0.0492	0.1526	0.3226	4.1693
R&D+SG&A	0.0529	0.1522	0.3478	4.4940
Intrinsic Value	0.0222	0.1149	0.1933	2.4985

出处：tushare，搞事情因子小组

知乎 @ 石川

与使用 MRQ 数据计算的原始 BM 相比，R&D 的 IC 和它十分接近，但波动略高；全部改造 BM 中，只有 R&D+SG&A 的 IC 和 IR 均优于 MRQ。

5 哪家 BM 比较强？

上一节通过 Portfolio Sort Test 检验了不同候选 BM 多空对冲因子组合的风险收益特征。本节就来让它们直接 PK 一下，看看哪个版本的 BM 比较强——即能够解释更多的 α 。

为了比较不同的 BM 指标，一个直接的方法是将它们的因子组合作为解释变量，通过时序回归或者截面回归来考察它们解释回归方程 LHS 资产收益率的能力。该方法的问题是，不同版本的 E 因子组合相关性很高。虽然高相关性不会破坏 OLS 的无偏性，且也不太可能出现完美共线性直接

本文采用 Gibbons, Ross, and Shanken (1989) Test (GRS Test) 以及不同版本 BM 因子组合两两回归相互解释这两种方法来比较它们。首先来看 GRS Test。

在 GRS Test 中，我们采用因子动物园维护的中国版 Fama and French (2015) 五因子 (FF5) 为基础，每次挑选一个候选 BM 指标构建 HML 组合并和 FF5 中的其他四个因子 (MKT、SMB、RMW、CMA) 组成五因子模型，并以其他候选 BM 的因子组合作为资产，考察该五因子模型对这些资产的解释能力。为了行文紧凑性，下面直接给出 GRS Test 的结果；GRS Test 的具体步骤详见附录。

由于在实证中我们采用了等权和市值加权两种构建因子投资组合的方法，因此在 GRS Test 中也针对这两种方法进行检验。当使用等权或市值加权的 BM 因子时，相应的 FF5 模型也是等权或市值加权。

先来看市值加权的结果（下表）。当把使用 MRA 数据计算的原始 BM 加入 FF5 模型后，该模型对其他候选 BM 的解释力度最高 (J-statistic 最低)。排在其之后的是两个和 R&D 相关的，即 R&D + SG&A 及 R&D。此外，使用 MRQ 数据计算的原始 BM 也有不错的效果。而其他的几种改造方法在本文针对 A 股的实证中效果并不好。

用在 FF5 中的 BM 因子	J-statistic	p-value
MRA 版本原始 BM	3.9495	0.7856
MRQ 版本原始 BM	6.8805	0.4414
Retained Earnings-to-Market	8.0820	0.3254
Contributed Capital-to-Market	9.0586	0.2485
NOA (Enterprise Book-to-Market)	14.0326	0.0506
R&D-to-Market	5.9582	0.5446
R&D+SG&A-to-Market	5.1528	0.6413
Intrinsic Value-to-Market	13.1773	0.0679

出处：搞事情因子小组、因子动物园。GRS Test 结果；**市值加权** BM 因子以及 FF5 因子。

下表显示了使用等权候选 BM 因子和 FF5 因子模型的 GRS Test 结果，从中可以得出类似的结论。两个 R&D 相关的候选 BM 均取得很好的结果。不过在等权时，解释力度最高的却要数 Retained Earnings-to-Market。



知乎

首发于
川流不息

MRA 版本原始 BM	12.1980	0.0942
MRQ 版本原始 BM	11.8811	0.1045
Retained Earnings-to-Market	8.4727	0.2928
Contributed Capital-to-Market	15.4033	0.0312
NOA (Enterprise Book-to-Market)	19.6749	0.0063
R&D-to-Market	11.5587	0.1160
R&D+SG&A-to-Market	9.9086	0.1938
Intrinsic Value-to-Market	23.5310	0.0014

出处：搞事情因子小组、因子动物园。GRS Test 结果；**等权** BM 因子以及 FF5 因子。 @石川

接下来看看候选 BM 因子组合两两相互解释的情况。每次从候选 BM 中挑出一个作为解释变量，用其余 BM **逐一**作为被解释变量进行时序回归，考察这些被解释变量的 α 收益率和其 t-statistic。这意味着，对于每个候选 BM 指标，我们用它的因子组合去解释另外 7 个指标的因子组合、得到 7 个 α 收益率和对应的 t-statistics；使用这些 α 绝对值的均值以及 t-statistics 绝对值的均值大小来评价被选为解释变量的 BM 的效果。

先来看通过市值加权构建 BM 因子组合的检验结果。下表中每一行的 BM 为解释变量，每一列的 BM 为被解释变量；表中每个单元格里有两个数值，第一行的数值为 α ，第二行括号内数值为它的 t-statistic；表中的最后一列为 $|\alpha|$ 均值和 $|t\text{-statistic}|$ 均值。

	被解释变量								绝对值 均值
	MRA	MRQ	RE	CC	NOA	R&D	R&D + SG&A	IV	
MRA	N/A	0.0019 (0.7745)	0.0020 (0.4545)	0.0008 (0.2401)	-0.0004 (-0.1366)	0.0023 (0.9982)	0.0035 (1.4448)	0.0024 (0.4713)	0.0019 (0.6457)
MRQ	-0.0007 (-0.2984)	N/A	0.0016 (0.3486)	-0.0007 (-0.2758)	-0.0016 (-0.6559)	0.0006 (0.6984)	0.0018 (1.3661)	0.0019 (0.3745)	0.0013 (0.5740)
RE	0.0035 (0.6941)	0.0052 (0.9530)	N/A	0.0047 (0.8796)	0.0024 (0.5079)	0.0059 (1.0580)	0.0066 (1.2672)	-0.0001 (-0.0555)	0.0041 (0.7736)
CC	0.0017 (0.4641)	0.0026 (0.8632)	0.0041 (0.8101)	N/A	0.0001 (0.0412)	0.0030 (1.0693)	0.0041 (1.4153)	0.0045 (0.8189)	0.0029 (0.7832)
NOA	0.0022 (0.7005)	0.0035 (1.1903)	0.0032 (0.6746)	0.0018 (0.5618)	N/A	0.0040 (1.3593)	0.0051 (1.7007)	0.0032 (0.6209)	0.0033 (0.9726)
R&D	-0.0012 (-0.5215)	-0.0004 (-0.5248)	0.0016 (0.3454)	-0.0012 (-0.5103)	-0.0020 (-0.8035)	N/A	0.0013 (1.0616)	0.0021 (0.4045)	0.0014 (0.5959)
R&D + SG&A	-0.0021 (-0.8901)	-0.0014 (-1.0416)	0.0009 (0.1985)	-0.0020 (-0.7418)	-0.0027 (-1.0563)	-0.0010 (-0.7495)	N/A	0.0015 (0.2966)	0.0017 (0.7106)
IV	0.0047 (0.8503)	0.0062 (1.0740)	0.0011 (0.4776)	0.0054 (1.0071)	0.0032 (0.6432)	0.0070 (1.1792)	0.0078 (1.3719)	N/A	0.0051 (0.9433)

出处：搞事情因子小组。**市值加权** BM 因子两两回归检验结果。每一行的候选 BM 为解释变量，每一列的候选 BM 为被解释变量；表中每个单元格里有两个数值，第一行的数值为 α ，第二行括号内数值为 t-statistic；表中的最后一列为 α 绝对值均值和 t-statistic 绝对值的均值。为了节约表格空间，BM 符号采用了缩写，它们的含义如下：MRA：MRA 版本的原始 BM；MRQ：MRQ 版本的原始 BM；RE：Retained Earnings-to-Market；CC：Contributed Capital-to-Market；NOA：NOA Enterprise Book-to-Market；R&D：R&D-to-Market；R&D + SG&A：R&D + SG&A-to-Market；IV：Intrinsic Value-to-Market。

果和市值加权的 GRS Test 结果十分接近。

当采用等权构建 BM 的因子组合时，结果如出一辙，以 $|\alpha|$ 均值从低到高排序的前三名仍然是使用 MRQ 数据计算的原始 BM、R&D 以及 R&D+SG&A。

		被解释变量							绝对值 均值	
		MRA	MRQ	RE	CC	NOA	R&D	R&D + SG&A		IV
解释 变量	MRA	N/A	0.0039 (2.0232)	0.0012 (0.3887)	0.0045 (1.4460)	-0.0007 (-0.3184)	0.0038 (1.8729)	0.0055 (2.4672)	-0.0010 (-0.2839)	0.0029 (1.2572)
	MRQ	-0.0022 (-1.1909)	N/A	0.0002 (0.0759)	0.0007 (0.3100)	-0.0045 (-3.8328)	-0.0002 (-0.2946)	0.0016 (1.7359)	-0.0019 (-0.5320)	0.0016 (1.1389)
	RE	0.0041 (1.1187)	0.0080 (1.9431)	N/A	0.0101 (2.0847)	0.0037 (0.8477)	0.0082 (1.0580)	0.0098 (2.2739)	-0.0027 (-1.7537)	0.0067 (1.7024)
	CC	-0.0003 (-0.0979)	0.0019 (0.8129)	0.0039 (0.8101)	N/A	-0.0030 (-1.2727)	0.0014 (0.6626)	0.0030 (1.4279)	0.0018 (0.4590)	0.0022 (0.8280)
	NOA	0.0019 (0.9933)	0.0049 (4.3079)	0.0023 (0.7101)	0.0050 (2.1508)	N/A	0.0048 (4.0683)	0.0064 (4.6292)	-0.0001 (-0.0280)	0.0036 (2.4125)
	R&D	-0.0019 (-1.0356)	0.0003 (0.5626)	0.0008 (0.2330)	0.0007 (0.3235)	-0.0043 (-3.6018)	N/A	0.0017 (2.4547)	-0.0013 (-0.3553)	0.0016 (1.2238)
	R&D + SG&A	-0.0030 (-1.4160)	-0.0012 (-1.2812)	0.0004 (0.1068)	-0.0007 (-0.3322)	-0.0057 (-4.0014)	-0.0015 (-2.0960)	N/A	-0.0017 (-0.4428)	0.0020 (1.3824)
	IV	0.0062 (1.5601)	0.0100 (2.3096)	0.0029 (2.0841)	0.0104 (2.1632)	0.0055 (1.2085)	0.0100 (2.2218)	0.0115 (2.5780)	N/A	0.0081 (2.0179)

出处：搞事情因子小组。等权 BM 因子两两回归检验结果。每一行的候选 BM 为解释变量，每一列的候选 BM 为被解释变量；表中每个单元格里有两个数值，第一行的数值为 α ，第二行括号内数值为 t-statistic；表中的最后一列为 α 绝对值的均值和 t-statistic 绝对值的均值。为了节约表格空间，BM 符号采用了缩写，它们的含义如下：MRA: MRA 版本的原始 BM；MRQ: MRQ 版本的原始 BM；RE: Retained Earnings-to-Market；CC: Contributed Capital-to-Market；NOA: Net Operating Assets-to-Market；R&D: R&D-to-Market；R&D + SG&A: R&D + SG&A-to-Market；IV: Intrinsic Value-to-Market。

综合本小节的两种比较方法可以得出如下结论：在本文考虑的实证范围内，未经改造的原始 BM（以 MRQ 数据计算）似乎已经足够优秀、战胜了绝大多数从海外论文中引入的改造方法。除此之外，R&D 相关的两个改造似乎能在原始 BM 之上提供一定的增量贡献。

6 改造版 BM 是否被定价？

接下来使用 Fama and MacBeth (1973) Regression 检验不同的 BM 因子是否被定价。

综合之前的分析结果，两个和 R&D 相关的因子效果最好。本节以 R&D+SG&A 改造版为例，计算其风险溢价。为了控制原始 BM 以及其他风格因子，本文考虑以下四种模型设定：

- **模型一**：仅使用 R&D+SG&A 改造版指标作为唯一的解释变量；
- **模型二**：使用 R&D+SG&A 改造版指标以及 FF5 中除市场外的其余三个风格因子对应的公司特征指标（firm characteristics，即对数市值、ROE、change of total assets）作为解释变量；
- **模型三**：使用 MRQ 版原始 BM 指标，以及用该 BM 指标对 R&D+SG&A 进行正交化后的指标作为解释变量；

从上面的说明不难看出，在 FM Regression 中，我们参照了 Barra 的做法，使用公司特征指标而非时序回归系数作为解释变量（见《Which Beta ?》）。此外，对于市场因子也采用 Barra 的做法，即在截面回归中加入截距项代表市场因子。

FM Regression 的实证区间和本文前面的检验相同、股票池为剔除了金融和非银的全部 A 股。在每月末进行截面回归时，对所有指标首尾各 0.5% 的部分进行 winsorization 去极值处理。下表给出了结果。每个单元格内第一行的数值为该指标月收益率均值、第二行括号内的数字为经 Newey and West (1987) 调整后的 t-statistics。

解释变量	模型一	模型二	模型三	模型四
R&D+SG&A	0.7138 (1.3045)	1.0359 (1.8935)		
经 BM 正交化后的 R&D+SG&A			2.1016 (2.0293)	1.5767 (1.6541)
BM			0.7277 (1.2701)	1.1457 (1.9586)
ln(size)		-0.7760 (-3.8100)		-0.7857 (-3.8439)
ROE		0.6260 (2.5879)		0.6522 (2.7920)
ΔTotal Assets		0.3079 (0.9288)		0.2526 (0.7304)
Adj R-squared	0.0142	0.0421	0.0179	0.0445

出处：搞事情因子小组。Fama and MacBeth (1973) Regression 结果。每个单元格内第一行的数字为解释变量的月收益率均值、第二行括号内的数字为经 Newey and West (1987) 调整后的 t-statistics。

上述四个模型中，我们最感兴趣的是模型二和模型四。模型二说明，当使用 R&D+SG&A 改造后的 BM 代替原始 BM 时，它可以获得较为显著的风险溢价（t-statistic = 1.8935）。模型四说明，当考虑到了 BM 之后，正交化后的 R&D+SG&A 版本仍然对解释个股的截面预期收益差异有一定的增量贡献，不过其显著性有所下降（t-statistic = 1.6541）。

除上述现象外，另一个有意思的发现是在模型三和模型四中，取决于其他风格因子是否被控制，BM 和正交后的 R&D+SG&A 的效果发生了反转。当没有其他控制变量时，正交后的 R&D+SG&A 比 BM 获得了更高且更显著的风险溢价；而当其他风格因子作为控制变量加入后，R&D+SG&A 的风险溢价降低而 BM 的风险溢价更高。这与这两个指标和控制变量之间的相关性以及潜在 EIV 问题有关，在此就不展开讨论了。

7 结语

知乎

首发于
川流不息

本文介绍了学术界主流的 BM 改造方法，并针对 A 股进行了深度实证。**结果表明，利用 R&D 和 SG&A 改造后的 BM 可以在股票定价时提供增量信息。**但也不可否认的是，本文的结论并不 sexy，即我们没有在 A 股上找到一个明显碾压原始 BM 的改造方法。以学术界的 publication bias 来看，这篇文章难以让人“眼前一亮”。

OK，写了这么多当然不是为了“自我否定”！

就本文结果而言，在当下的 A 股，原始的 BM 较其他改进方法依然有效。虽然文中介绍的改造方法均发表于金融学或会计学顶刊，但它们似乎在 A 股“水土不服”。然而在因子投资已经十分拥挤的今天，从基本面分析出发加工、改造因子是一条值得尝试的道路。

本文以严谨的科研态度为出发点，报告了考察的全部改造方法以及它们的效果。比起刻意掩盖 multiple testing 的尝试、仅介绍一个样本内惊艳的改造方法，我们认为这种系统、全面的因子分析、比较方法对于因子投资的实践更有价值。在未来，我们也愿和各位小伙伴一同努力，探寻适合 A 股的 BM 改造方法。

感谢阅读，全文完。

A 附录：GRS Test

GRS Test 的步骤为：

1. 挑出第 n ($n = 1, 2, \dots, 8$) 个候选 BM 指标构建 HML，和 FF5 中的其他四个因子组成 FF5 模型，记为 FF5 _{n} ，其中下标 n 代表以第 n 个 BM 因子构建的 FF5。
2. 使用其他 BM 指标的因子组合作为资产。
3. 使用 FF5 _{n} 作为因子模型，对其他 BM 因子组合资产进行 GRS Test，得到 test statistic J_n ，下标 n 对应第 n 个 BM 因子。
4. 对所有候选 BM，重复上述 1 - 3 步，得到 8 个 J_n ， $n = 1, 2, \dots, 8$ 。
5. 所有 J_n 中取值最小的 J_{n^*} 的下标取值 n^* 就是解释力度最高的 BM 因子。

在上述第 3 步中，按以下公式构建满足卡方分布的 test statistic J （自由度是 7 即资产的个数）：

$$J = T \frac{\alpha' \Sigma^{-1} \alpha}{1 + \mu_f' \Sigma_f^{-1} \mu_f} \sim \chi_7^2$$

上式中：

- T 是收益率时序期数。
- Σ_f 为因子样本协方差矩阵；在本例中，因子模型 FF5 _{n} 中共有 5 个因子，因此 Σ_f 是：

知乎

首发于
川流不息

- α 为资产的回归截距向量 (7×1 阶)，其求法为，使用 FF5_n 因子作为解释变量，逐一对每个资产 j 进行时序回归，得到的截距项就是该资产不能被 FF5_n 解释的 α_j ；所有资产的 α_j 放在一起就是 α 向量。
- Σ 为残差协方差矩阵：在上述每个时序回归中，得到了资产 j 的残差 ϵ_j 向量；使用资产的残差向量计算出它们的样本协方差矩阵为 Σ (7×7 阶)。

从上述 test statistic 的含义不难看出，我们实际上是在检验当给定五因子模型后，剩余候选 BM 的因子组合作为资产的定价错误是否联合为零。 J_n 越小说明该因子模型（包含了选定的某个候选 BM）越能够解释其他 BM；反之， J_n 越大则说明模型效果越差。

参考文献

- Asness, C. and A. Frazzini (2013). The devil in HML' s details. *The Journal of Portfolio Management*, Vol. 39(4), 49 – 68.
- Ball, R., J. Gerakos, J. T. Linnainmaa, and V. Nikolaev (2019). Earnings, retained earnings, and book-to-market in the cross section of expected returns. *Journal of Financial Economics*, forthcoming.
- Chan, L. K. C., J. Lakonishok, and T. Sougiannis (2001). The stock market valuation of research and development expenditures. *The Journal of Finance*, Vol. 56(6), 2431 – 2456.
- Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *The Journal of Finance*, Vol. 47(2), 427 – 465.
- 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.
- Frankel, R. and C. M. C. Lee (1998). Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics*, Vol. 25(3), 283 – 319.
- Gibbons, M. R., S. A. Ross, and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica*, Vol. 57(5), 1121 – 1152.
- Lev, B. (2019). Is the Market-to-Book ratio still relevant? Yes, but... Available at: levtheendofaccountingblog.wordpress.com...
- Meredith, C. (2016). Price-to-Book' s growing blind spot. OShaughnessy Asset Management Research Note.
- 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.



知乎

首发于
川流不息

- Penman, S. H., S. A. Richardson, and I. Tuna (2007). The book-to-price effect in stock returns: accounting for leverage. *Journal of Accounting Research*, Vol. 45(2), 427 – 467.

免责声明：文章内容不可视为投资意见。市场有风险，入市需谨慎。

原创不易，请保护版权。如需转载，请联系获得授权，并注明出处。已委托“维权骑士”（[维权骑士_免费版权监测/版权保护/内容多平台分发](#)）为进行维权行动。

发布于 2019-08-01

价值投资 尤金·法马 (Eugene Fama) 资产定价

▲ 赞同 28 ▼ 1 条评论 分享 收藏 ...

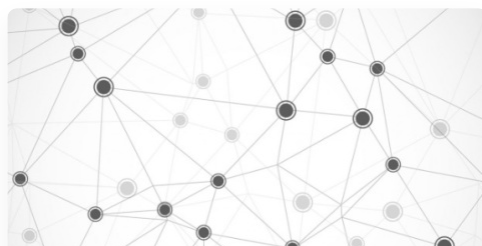
文章被以下专栏收录

**川流不息**

北京量信投资管理有限公司是一家在中国基金业协会备案登记的专业私募基金管理人...

[关注专栏](#)

推荐阅读

**获取 α 的新思路：科技关联度**

石川

**多种机器学习算法在多因子策略预测中的有趣比较和进一步思...**

zipli...

发表于Zipli...

**『机器学习之股指期货』**

陈霖



知乎



首发于
川流不息

写下你的评论...



llanglli

22 天前

yeah

👍 赞

