

证券研究报告



分析师：

任瞳

[rentong@xyzq.com.cn](mailto:rentong@xyzq.com.cn)

S0190511080001

王武蕾

[wangwulei@xyzq.com.cn](mailto:wangwulei@xyzq.com.cn)

S0190517070008

高智威

[gaozhw@xyzq.com.cn](mailto:gaozhw@xyzq.com.cn)

S0190517080001

郑兆磊

[zhengzhaolei@xyzq.com.cn](mailto:zhengzhaolei@xyzq.com.cn)

S0190517070008

报告关键点

- 本期三篇文献涵盖：多因子选股、文本挖掘、资产配置三个领域。第一篇报告通过回测验证了公司不完全信息的影子成本与股票收益的关系。第二篇报告重点研究了正面情绪、负面情绪、综合情绪、有无新闻等对不同持有期的股票收益的预测作用。第三篇报告提出了对股市、债市长期预期收益的建模预测方法，并对美国市场的股市真实风险溢价进行了实证研究。

相关报告

《瞳观天下·兴业证券资产配置建议 2018 年 3 月》2018-03-01

《市场情绪偏乐观——兴业期权水晶球预测日报 20180301》2018-03-01

《市场情绪谨慎——兴业期权水晶球预测日报 20180228》2018-02-28

## 西学东渐——海外文献推荐系列之十

2018 年 05 月 26 日

投资要点

- 西学东渐，是指从明朝末年到近代，西方学术思想向中国传播的历史过程。西学东渐不仅推动了中国在科学技术和思想文化方面的发展，也有力地促进了社会与政治的大变革。在今天，西学东渐仍有其重要的现实意义。作为 A 股市场上以量化投资为研究方向的卖方金融工程团队，我们在平日的工作中，常常深感海外相关领域的研究水平之高、内容之新。而这也正促使我们通过大量的材料阅读，去粗取精，将我们认为最有价值的海外文献呈现在您的面前！为了保证品质，我们的文献推荐系列报告将以双周形式推出，每期推荐 1~3 篇不等的优秀文章。
- 本期三篇文献涵盖：多因子选股、文本挖掘、资产配置三个领域。第一篇报告通过回测验证了公司不完全信息的影子成本与股票收益的关系。第二篇报告重点研究了正面情绪、负面情绪、综合情绪、有无新闻等对不同持有期的股票收益的预测作用。第三篇报告提出了对股市、债市长期预期收益的建模预测方法，并对美国市场的股市真实风险溢价进行了实证研究。

**风险提示：**文献中的结果均由相应作者通过历史数据统计、建模和测算完成，在政策、市场环境发生变化时模型存在失效的风险。

## 目 录

文献一：Does investor recognition predict returns? .....	- 3 -
推荐原因 .....	- 3 -
数据与研究方法 .....	- 3 -
结论 .....	- 3 -
我们的思考 .....	- 4 -
文献二：News vs. Sentiment: Predicting Stock Returns from News Stories .....	- 5 -
推荐原因 .....	- 5 -
数据与研究方法 .....	- 5 -
结论 .....	- 6 -
我们的思考 .....	- 6 -
文献三：What Risk Premium is “Normal”? .....	- 7 -
推荐原因 .....	- 7 -
数据与研究方法 .....	- 7 -
结论 .....	- 7 -
我们的思考 .....	- 8 -

## 报告正文

### 文献一：Does investor recognition predict returns?

作者：Andriy Bodnaruk, Per Ostberg

来源：Journal of Financial Economics

推荐人：高智威

#### 推荐理由

传统金融模型通常基于无摩擦市场假设和完全信息假设，往往难以解释复杂金融市场中的合理变化。在 Merton 模型(Merton, 1987)中，投资者所掌握的信息是不完整的，并根据他们所获悉的证券组合来优化他们持有的投资组合。市场对每只股票的需求取决于被告知股票的投资者的比例。具有较小股东基数的公司的资产应该因此具有更低的价格和更高的回报。Merton 表明，这种回报溢价取决于公司非完全信息的影子成本，而该影子成本进一步取决于非系统性风险、相对市场规模和股东基数。本文中，Andriy Bodnaruk 等人利用瑞典股票的数据库，证明了股票收益与影子成本呈正相关。

#### 数据与研究方法

本文使用瑞典证券登记处收集的从 1995 年 6 月到 2000 年 12 月个人投资者的股票数据。用 SIS Ägarservice AB 数据库来补充这些信息，它包含了通过信托和外国控股公司持有的股份的最终所有者的信息。本文掌握了持有超过 500 股的所有股东的信息，同时也包含了较小持股的股东。

第一步：根据 Merton 模型计算影子成本以及其他描述性统计数据。

第二步：本文给出基于影子成本 ( $\lambda_1, \lambda_2$ ) 构建的投资组合的结果。在每年六月底和十二月底利用影子成本大小以等权方式构建四个投资组合，根据之后六个月的数据计算每个投资组合的月平均超额收益。本文计算了每一个投资组合的月度超额收益的平均数和中位数等统计量。每个投资组合至少包含 38 个股票。

第三步：利用 Fama-MacBeth 回归分析了影子成本与未来收益的关系。

第四步：本文基于前文对影子成本与回报之间的分析，利用影子成本做多高影子成本公司和做空低影子成本公司构建策略，分别用等权和市值加权的方式构建投资组合，计算不同投资组合的平均月回报率以及相关统计量。本文使用四因子模型从承担更多风险的角度来对利用影子成本进行交易获得的高回报率进行解释。

第五步：探索了影响影子成本的变量和预期回报之间的关系。

第六步：研究影响股东基数的变量，以及这些变量与回报之间的关系。

#### 结论

本文通过回测检验发现，不完全信息的影子成本与收益呈正相关。此外，相

比股东基数高的公司，股东基数低的公司可以提供更大的回报。本文中基于影子成本的交易策略可以赚取高额回报，这个交易策略的超额收益可以根据四因子模型来解释。本文的主要贡献是提供了具有代表性的证据来证明影子成本对公司股票回报率的影响。

### 我们的思考

我国股票市场成熟度与一般发达国家股票市场还存在差距，投资者信息不对称的情况时常出现。我们可以将 Merton 模型推广应用到我国的股票市场之中，有机会获得一个较大超额收益的交易策略。

## 文献二: News vs. Sentiment: Predicting Stock Returns from News Stories

作者: Steven L. Heston, Nitish Ranjan Sinha

来源: Financial Analysts Journal 73(2017) 67-83

推荐人: 郑兆磊

### 推荐理由

机器学习和文本挖掘已经越来越广泛的应用到当今金融领域。特别是对于文本信息的研究及实证分析受到许多学者的重视，但是已有的绝大多数研究仅限于新闻对股价的短期预测研究，数据量也较为有限。本文则从一个更长的时间窗口，使用更大量的数据，运用神经网络对新闻情绪进行判断。在此基础上验证正面情绪、负面情绪、综合情绪、有无新闻对不同时间窗口的股价预测作用，对我们量化选股有一定的借鉴意义。

### 数据与研究方法

#### 1、数据源

本文使用的原始数据来自于 Thomson Reuters 数据库，选用了 2003 年至 2010 年间带有公司标识的 900754 篇新闻报道。所有数据由数据库中的情绪数据集 Thomson Reuters NewsScope Data 和文本数据集 Thomson Reuters News Archive 合并而成。

#### 2、研究步骤

##### 2.1、情绪判断

应用神经网络对新闻情绪进行判别。然后在考察时段内，针对于每个公司：将正面情绪加总得到正面情绪因子、负面情绪加总得到负面情绪因子、正面情绪减负面情绪得到综合情绪因子。

作者通过分位数组方法验证不同的形成期和持有期下综合情绪因子对股价的预测作用，通过线性回归的方式论证不同情绪、有无新闻对股价的预测作用。

##### 2.2、分位数组测试

在日频上，对情绪因子做五分位组合测试，构建多空组合，测试其对持有期为 1-10 天的收益预测作用。周频上，对情绪因子做十分位组合测试，构建多空组合，测试其对持有期为 1-13 周的收益预测作用。

##### 2.3、线性回归测试

$$r_{i,t} = \alpha_{k,t} + \gamma_{k,t} \text{If\_news}_{i,t-k} + \beta_{k,t} \text{Positive}_{i,t-k} + \theta_{k,t} \text{Neutral}_{i,t-k} + \delta_{k,t} \text{Negative}_{i,t-k} + \varepsilon_{i,t}$$

符号说明： $r_{i,t}$  为股票  $i$  在第  $t$  周的收益； $\text{If\_news}_{i,t-k}$  是一个虚拟变量，为针对有新闻的公司的  $k$  阶滞后； $\text{Positive}_{i,t-k}$ 、 $\text{Neutral}_{i,t-k}$  和  $\text{Negative}_{i,t-k}$  为滞后  $k$  周发布的新闻的情绪结果； $\alpha_{k,t}$  为不发布任何新闻的公司的在滞后第  $k$  期的等权收益；

$\gamma_{k,t}$  为有中性新闻的公司对于不发布任何新闻的公司的超额收益； $\beta_{k,t}$ 、 $\theta_{k,t}$  和  $\delta_{k,t}$  为多元化的投资组合的超额回报。

通过线性回归的方式测试正面情绪、负面情绪、有无新闻对股价的预测作用。

## 结论

1、日频新闻情绪对未来一至两天的股票收益有一定的预测作用，而周度新闻情绪能够对一个季度的股票收益具有显著的预测作用。作者认为差异存在的原因有两点：a)、日频的新闻的噪声更大，所以预测作用更弱；而周度的新闻可以从某种程度上抵消一部分噪音；b)、日度的新闻更稀疏，周度的新闻相对更密集，也更加有效。

2、新闻中体现更多积极情绪的公司将在一周内有更高的股票收益，新闻中更多消极情绪的的公司将在未来的一个季度内表现则相对较差。

3、发布中性新闻的公司比不发布任何新闻的公司股票收益高。

## 我们的思考

随着人工智能时代的到来，量化投资领域也必将面临着新一轮的变革和考验。从目前来看，大数据和相应的机器学习的算法应用有着不可估量的作用。而从目前国内量化从业者的体验来看，比较关心的问题有两点：1、如何对 UGC(用户自主生成)内容以及新闻或者上市公司公告做情绪判断？；如何对各种文本内容构建知识图谱，从而形成传导机制。本文为我们在情绪判断方面提供了较好的思路，是值得尝试的方向。



## 文献三：What Risk Premium is “Normal”?

作者：Robert D. Arnott and Peter L. Bernstein

来源：Financial Analysts Journal

推荐人：王武蕾

### 推荐原因

随着养老目标基金的渐行渐近，资管机构对资产配置尤其是长期战略配置的研究需求明显增加。由于国内资本市场成立历史时间相对较短、估值波动大等特征使得投资者往往忽视对其长期收益风险特征的研究，包括预期收益来源的拆分与计算，但是这些数据又是进行战略配置的必备基础。本文作者基于近 200 年美国资本市场的历史数据，提出构建股市和债市真实预期收益率的模型，最终得到预期的股票风险溢价数据，作者详细分析了预期收益率模型的构建方法，并回测了该模型在美国市场的预测能力，其分析方法对国内市场进行相关研究有很好的参考意义。

### 数据与研究方法

由于本文回顾的历史时间长达 200 年，且使用的数据包括资本市场、企业基本面和宏观经济等多维度的数据，因而数据来源多元化。早期数据主要来自于 NBER、Ibootsen Associates 以及相关研究报告、新闻媒体（Wall street journal）等，近年来的数据则来自于 Bloomberg 数据库。

本文首先详细阐述了预期风险溢价跟历史风险溢价的重要区间，且以 1926 年为起点作为案例研究，说明了站在当时给出的预期风险溢价与实际风险溢价的差异，并指出估值调整、幸存者偏差、监管改革等因素是导致预期风险溢价与实际风险溢价出现偏差的重要原因。接下来，文章重点介绍了预期风险溢价的计算方法，包括两个步骤：第一，计算股市和债市真实预期收益，并回测预期收益对实际收益的预测能力；第二，计算股市风险溢价，并回测预期风险溢价对真实风险溢价的预测能力。

### 结论

我们将重点阐述作者对股市、债市真实预期收益率和预期风险溢价的建模过程以及实证结果：

1. 文章将股市的长期实际预期收益率拆分为真实股息率、真实股息增长率和估值变动三个项目，通过实证研究发现真实股息增长率显著低于真实 GDP 的增速，而跟真实人均 GDP 的增速相接近（股息增长相对略低），并进一步将真实股息增长率拆分为预期实际人均 GDP 增速和稀释调整效应两项，通过该计算方法得到的预期收益率跟事后实际收益率的相关性超过 40%，解释力  $R^2$  为接近 20%。在过去的近 200 年中，美股的真实收益为 6.8%，其中股息率贡献 5%，人均 GDP 增速为 1.7%，股息增速相对人均 GDP 增速的稀释调整项为 0.6%，估值上升的贡献为

0.7%，作者指出考虑到当前股息率水平的下降以及估值水平，认为美股未来的实际增速应该在 2%-4% 区间；

2. 文章将债券的长期实际收益拆分为名义收益率减预期通货膨胀，考虑到当前已有通胀连结债券(TIPS)，则可使用 TIPS 利率作为预期通货膨胀率，而在 TIPS 发行之前则用历史三年的通货膨胀率作为未来的预期，依次计算出实际债券收益。在过去仅 200 年中，美债的真实收益为 3.7%，其中名义利率为 4.9%，而预期通胀率为 1.2%；

3. 基于对股市和债市预期收益的计算，可以得到理性预期下的股市真实风险溢价，作者计算出长期股市的真实风险溢价为 2.4%，远低于名义 5% 的风险溢价水平，作者指出依据历史数据计算出的 5% 风险溢价有相当部分来自于历史的偶然调整以及股市估值的调整，但在理性预期下未来并不能持续。

### 我们的思考

本文完成于 2000 年末，从当时计算的结果来看，无论是对股市的预期收益还是预期风险溢价的计算结果，从事后来看都是相对准确的。借助本文的分析方法，我们可以类似的对国内股市债市的长期收益和风险溢价进行更准确的预测分析，为长期战略配置打下坚实基础。

**风险提示：文献中的结果均由相应作者通过历史数据统计、建模和测算完成，在政策、市场环境发生变化时模型存在失效的风险。**



## 分析师声明

本人具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

## 投资评级说明

投资建议的评级标准	类别	评级	说明
报告中投资建议所涉及的评级分为股票评级和行业评级(另有说明的除外)。评级标准为报告发布日后的12个月内公司股价(或行业指数)相对同期相关证券市场代表性指数的涨跌幅。其中：A股市场以上证综指或深圳成指为基准，香港市场以恒生指数为基准；美国市场以标普500或纳斯达克综合指数为基准。	股票评级	买入	相对同期相关证券市场代表性指数涨幅大于15%
		审慎增持	相对同期相关证券市场代表性指数涨幅在5%~15%之间
		中性	相对同期相关证券市场代表性指数涨幅在-5%~5%之间
		减持	相对同期相关证券市场代表性指数涨幅小于-5%
		无评级	由于我们无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使我们无法给出明确的投资评级
	行业评级	推荐	相对表现优于同期相关证券市场代表性指数
		中性	相对表现与同期相关证券市场代表性指数持平
		回避	相对表现弱于同期相关证券市场代表性指数

## 信息披露

本公司在知晓的范围内履行信息披露义务。客户可登录 [www.xyq.com.cn](http://www.xyq.com.cn) 内幕交易防控栏内查询静默期安排和关联公司持股情况。

## 使用本研究报告的风险提示及法律声明

兴业证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本报告仅供兴业证券股份有限公司(以下简称“本公司”)的客户使用，本公司不会因接收人收到本报告而视其为客户。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。

本报告所载资料的来源被认为是可靠的，但本公司不保证其准确性或完整性，也不保证所包含的信息和建议不会发生任何变更。本公司并不对使用本报告所包含的材料产生的任何直接或间接损失或与此相关的其他任何损失承担任何责任。

本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可升可跌，过往表现不应作为日后的表现依据；在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告；本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现。过往的业绩表现亦不应作为日后回报的预示。我们不承诺也不保证，任何所预示的回报会得以实现。分析中所做的回报预测可能是基于相应的假设。任何假设的变化可能会显著地影响所预测的回报。

本公司的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司没有将此意见及建议向报告所有接收者进行更新的义务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告并非针对或意图发送予或为任何就发送、发布、可得到或使用此报告而使兴业证券股份有限公司及其关联子公司等违反当地的法律或法规或可致使兴业证券股份有限公司受制于相关法律或法规的任何地区、国家或其他管辖区域的公民或居民，包括但不限于美国及美国公民(1934年美国《证券交易所》第15a-6条例定义为本「主要美国机构投资者」除外)。

本报告的版权归本公司所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。未经授权的转载，本公司不承担任何转载责任。

## 特别声明

在法律许可的情况下，兴业证券股份有限公司可能会持有本报告中提及公司所发行的证券头寸并进行交易，也可能为这些公司提供或争取提供投资银行业务服务。因此，投资者应当考虑到兴业证券股份有限公司及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。

## 兴业证券研究

上海	北京	深圳
地址：上海浦东新区长柳路36号兴业证券大厦15层	地址：北京西城区锦什坊街35号北楼601-605	地址：深圳福田区中心四路一号嘉里建设广场第一座701
邮编：200135	邮编：100033	邮编：518035
邮箱：research@xyzq.com.cn	邮箱：research@xyzq.com.cn	邮箱：research@xyzq.com.cn