

史博：大数据将颠覆传统投资

□本报记者 张昊

大数据投资潜力无限

史博表示，传统的主动投资和被动投资其实都是利用数据来发掘投资标的，而互联网时代为利用数据进行投资提供了更多的可能。互联网大数据，不仅仅能够反映企业的经营状况，更能反映市场参与者对于各种数据的反应。

中国证券报：利用互联网大数据进行的金融投资，与传统金融投资相比有何不同？

史博：传统投资分成两大类：主动投资和被动投资。主动投资又可以分成两大方向：基于基本面投资和基于技术分析投资。从某种角度而言，传统投资都是基于数据的投资。在基本面分析中，投资者注重的是企业经营数据的分析，企业历史经营效益、企业收入、企业毛利率、净资产回报等指标，都属于企业日常经营过程中产生的数据。技术分析更注重的是投资过程中不同投资者投资行为产生的数据，即由投资者共同形成的交易量的数据、股价变化的数据等等，这是技术分析进

行投资的基础。因此，主动投资本质也是基于数字的投资。

在互联网时代，利用网络财经大数据进行投资对传统投资而言不但是一种补充，甚至是一种超越。因为投研基于的数据不仅仅是传统投资中企业财务报表的数据，也不仅仅是投资者在投资过程中形成的股票交易量、股票交易价格变化的数据，它同时也会特别关注投资者对交易数据如何进行反映，例如投资者对某个事件以及对公司相关报道的观点是什么，都会产生新一轮的数据，这就是基于数据的反映。所以，大数据投资是对现有投资有效的补充，在这个补充基础上，大数据投资有可能对我们传统投资产生颠覆性影响。随着数据量的增加，数据的分析也变得越来

i指数历史表现优异

近年来，互联网飞速发展，论坛、微博、微信等互联网媒介深刻影响着人们的生活，而互联网数据对于金融投资的价值也在逐渐显现。南方新浪大数据系列指数将传统的基于价值、成长、企业质量的因子与互联网大数据所能给予的股票热度预期、成长预期、估值提升预期相结合，利用南方基金量化研究平台将这些因子进行优化。从历史数据回溯来看，南方新浪大数据系列指数与现有的传统指数相比有着明显的优势。

中国证券报：对于互联网大数据的应用，国外有没有相关的研究？

史博：大数据时代，数据已经从对事物的简单描述逐渐演变成人类行为模式的全面刻画，近年来基于社交平台和互联网搜索中用户行为被记录的数据，还被证明能够很好地映射出市场主体的情绪。

美国印第安纳大学近年的一项研究成果更表明从Twitter信息中表现出来的情绪指数与道琼斯工业指数的走势之间相关性高达87%。2014年牛津大学期刊发表的一篇文章表明，通过搜索分析投资者在网络发帖和评论中表现出来的观点，能够很好地反映多空态度，同样也能够有效地预测未来股市的收益。美国市场大量有关金融和大数据成功结合的案例鼓舞了投资管理人。作为最火的社交网络，Twitter拥有众多互联网使用群体的信息，因而其被大量的投资管理人视为大数据的最佳来源，并依此开发了相应的对冲产品，CAYMAN ATLANTIC公司就是其中的翘楚之一。CAYMAN ATLANTIC在其官网上披露了旗下对冲基金的历史收益数据。从2012年7月该基金成立至2014年6月，基金的累计收益率达到了48.36%，年化收益率为21.8%。各种指标揭示了该基金在非常低的风险下，创造了相当高的收益。

中国证券报：南方新浪大数据系列指数将怎样利用互联网提供的因子来研究金融投资？

史博：南方基金现有的量化股

票投资平台是基于价值、成长、企业质量等因子，这些因子会筛选股票的行业属性，筛选股票财务指标数据以及一些市场反映的数据。现在，南方基金将把新浪财经大数据也结合进来，将个股的热度、新闻报道以及微博相关文章反映的市场情绪反映进来，同时找出股票的热度预期、成长预期、估值提升预期，这些预期将会影响股价将来的反映，把这些因子构建在一起，形成一个不同因子的权重，对权重进行动态优化调整，构成整个大数据指数构造的基础。

在指数的构成过程中，样本选择空间非常大，全市场所有股票中扣除三类股票：ST股票、*ST股票以及上市不满一年的股票（目前市场情况下上市不满一年数据量有时候不够充足，利用大数据平台分析偏差比较大）。

选股因子包括三大类：企业财务因子、大数据因子、市场驱动因子。财务因子包括最新的市盈率指标、净资产收益率情况、企业年度营业收入、净利润增长率以及对这些指标变化的预测。市场驱动因子包括股票近期的换手率、波动率、股票价格变化率以及成交量变化率。

对于新浪大数据因子，将根据新浪财经频道下的股票页面访问热度计算单个股票的热度得分、根据财经频道下的新闻报道正负面影响计算单个股票新闻报道得分、根据股票在微博上的正负面文章影响计算单个股票微博得分，综合上述得分并根据历史回测优化结果作为大数据因子得分。

南方基金利用自身强大的量化研究平台，将上述因子进行优化后，最终创造出南方新浪大数据系列指数（i指数）。

通过对i指数的历史收益表现进行分析可以看到，从2001年至今，对大数据100、300等权重指数进行回溯，并将回溯指数和创业板指数、中小板指数、深证100、沪深300等进行了比较分析，结果显示，大数据100指数过去5年累计收益有222.4%，大数据300指数的收益为141.58%，远远跑赢其他对比指数。

作为国内互联网大数据与金融投资相结合的最早应用之一，南方新浪大数据系列指数已于近日在深交所正式挂牌，南方基金权益投资总监史博就此接受了中国证券报记者的专访。史博表示，南方基金将传统的基于价值、成长、企业质量的因子与互联网大数据所能给予的股票热度预期、成长预期、估值提升预期相结合，利用南方基金量化研究平台将这些因子进行优化，最终诞生了南方新浪大数据系列指数。大数据投资是对现有投资有效的补充，在这个补充基础上，大数据投资有可能对传统投资产生颠覆性影响。作为对传统投资模式的有效提升，南方基金有信心通过对互联网大数据的研发和应用，为持有人创造更多的价值。



史博，南方基金总裁助理兼权益投资总监，公司投资决策委员会委员；特许金融分析师(CFA)，中国人民银行研究生部国际金融学硕士。1998年至2002年任职于博时基金，期间曾在Dryden Wealth Management Co., Ltd. (London) 从事研究工作；2002年6月加入泰达荷银基金，曾担任研究部总经理、首席策略分析师；2004年7月至2005年2月，任泰达周期基金经理；2005年2月至2007年6月任职于中国人寿资产管理有限公司股票投资部，任高级投资经理；2007年6月重新加入泰达荷银基金，担任投资副总监；2007年7月至2009年5月任泰达首选基金经理；2008年8月至2009年9月任泰达市值基金经理；2009年4月至2009年9月任泰达品质基金经理；2009年10月加入南方基金；2011年2月至今，任南方绩优基金经理；2014年2月至今，任南方新优享基金经理。

共享大数据投资新时代

本次南方新浪大数据系列指数的推出，是南方基金、新浪和深圳证券信息公司强强联合的结果。史博表示，基于这样一个策略指数，南方基金希望将其产品落地化，在不远的将来针对该指数开发相应的指数产品，投资者有望从基金产品的创新中受益，共享大数据投资的新时代。

中国证券报：南方新浪大数据系列指数由三方联合推出，这样的组合有何优势？未来南方基金会不会推出跟踪这一指数的基金产品？

史博：南方基金、新浪和深圳证券信息公司都是各自领域的翘楚，新浪财经拥有海量财经大数据库；南方基金作为国内最早一批基金公司，一直以其稳健的投资理念以及跟随时代的创新精神吸引着大量投资者，稳健而良好的业绩以及符合投资者需求的产品，使得公司规模一直位居中国基金公司第一梯队；深圳证券信息公司是国内最早开展指数业务的专业化机构，自2002年以来，先后开发了深证100指数、中小板指数、创业板指数等著名指数。南方基金作为国内首只制定指数的基金公司，其2006年发布的中证南方小康指数是国内首只基金公司制定的指数，以跟踪该指数标的的ETF基金及其联接基金于2010年成功募集上市，投资运作良好，可以看出南方基金在指数编制和管理方面的深厚经验。

三方强强携手，力争将互联网金融的创新推入新纪元。基于这样一个策略指数，南方基金希望将其产品落地化，在不远的将来针对该指数开发相应的指数产品，投资者将有望从基金产品的创新中受益，共享大数据投资的新时代。

目前，国内股票市场交易较为频繁，短期交易盛行，价格波动较大。因此，国内股票市场上投资者情绪在投资者行为中的决定作用更加突出，对投资的指导意义也尤为明显。大数据系列指数的发布能够丰富和完善指数体系，为市场提供新的投资工具，并有助于满足投资者多样化的投资需求。

责编：鲁孝年
美编：尹建