

深度报告

金融工程

数量化投资

数量化投资技术系列报告之二十五

2010年3月30日

专题报告

本报告的独到之处

■完整阐述了国信金融工程团队在主动投资领域所进行的系列研究及这些研究成果落地和应用的现状，以及未来的扩展方向，包含了我们对量化投资的思考，是集体智慧的一次集中展示。策略在指数和产品领域的应用，还将有其他报告进行综合评述。

相关研究报告：

数量化投资技术系列报告（1-24）、

海外量化技术本土化系列报告（1-3）、

算法交易与程序化交易系列报告（1-2）

交易指标与策略系列报告（1-2）

董艺婷 执笔

TEL 021-60933155

Email dongyt@guosen.com.cn

国信金融工程团队

集体智慧

独立性声明：

作者保证报告所采用的数据均来自合规渠道，分析逻辑基于本人的职业理解，通过合理判断并得出结论，力求客观、公正，结论不受任何第三方的授意、影响，特此声明。

实现量化投资盈利目标的征程 2009-2010

有效地实现收益，是我们自始至终的唯一目标，也是我们认为的，能够区别量化投资研究本土和山寨的唯一标准。

在研究中无限制地考虑细节问题对分析师来说是一件极为吃力不讨好的事情——每一个细节的加入都导致研究的边际成本和工作量以级数级增长，但是无限接近真相的唯一途径就是去了解这些细节、分解、模拟、严苛地处理投资中可能遇到的困难场景和隐形成本。

方法无优劣，决定结果的是使用的方法和场合。量化投资技术最擅长的事，是大比例、大概率地知道决策过程的依据、长短和使用场合。也许这一决策过程过于冰冷和枯燥，但实现预期的过程同样很美妙，因为我们摒弃后视镜、双色球的研究思路，因此我们有理由期待自己的智慧结晶得到未来的验证。

择时研究是股票投资永远没有最优解的难题，宏观、中观和微观视角缺一不可，宏观视角择时——易行的不靠谱，靠谱的不易行；在预测顶部拐点的时候，只分析疯狂程度不行，还要知道疯狂和疯狂的差异；金融地产是A股投资一切纠结情绪产生的源头；微观视角——高频数据乃是眼前最大的阿里巴巴宝库，投资者行为择时，是目前A股门槛最高的量化研究领域之一。

我们对每一个研究领域中所采纳的新方法都十分珍惜，因为在研究过程中，我们所共享的开发平台使得这些新方法的标准化模块将来可以快速移植到其他领域的研究中。我们的正Alpha行业配置研究和大小盘风格轮动所采用的方法，将来可以扩展应用于一切类似的对立风格的轮动研究。简单的输入和输出处理，使得这些研究成果的对接和产品封装成本大大降低，未来我们还将继续在资产配置和轮动领域批量地进行测试和策略开发。

我们所计划的市场中性策略体系中，采用传统协整关系进行个股选择只是最为基础的一部分。我们的配置策略+配对交易体系，完全可以实现又一次的“1+1>2”。将非参数方法引用到行业内选股并达到有效区分多空组合的效果，则使我们之前各个层面的策略研究具备了落地的可能。

伴随我们在算法交易、程序化交易领域的研究覆盖达到一定程度，目前我们的量化投资体系从策略开发到交易实施已经形成了一个完整的体系，在这个体系的各个层面，我们都拥有了自主开发的、经过实践检验的主打研究成果。未来一段时间内，在各项研究持续推进的同时，整合并发挥最大的优势，将成为我们新的课题。

一个没有策略储备的量化决策体系是脆弱的，随时都可能遭遇钝化和逆转，因此，我们对已有的策略不断地进行更新换代，并试图通过研究方法的多样性，达到策略和方法之间的复合，尽量分散单一策略对决策造成的影响力。我们认为，未来量化投资的核心竞争力将体现在策略的持续开发和应用能力上，一个量化团队管理的要素中，也应该包含多角度分析能力的维护，和刺激团队追寻新思路的机制，对市场永远有好奇心的天才，也是我们愿意用任何代价追寻的宝贝。在实现收益的征途上，我们有无尽的动力，2009-2010，只是一个开始。

内容目录

前言——实现收益是我们的唯一目标	5
本土还是山寨？——什么是检验量化投资的标准	5
理论还是细节？——什么是量化投资研究的瓶颈	5
择时、配置还是选股？——什么是量化投资技术擅长的事	6
择时——宏观、中观和微观视角缺一不可	8
择时是股票投资永远没有最优解的难题	8
宏观视角择时——易行的不靠谱，靠谱的不易行	8
预测顶部拐点——只分析疯狂程度不行，还要知道疯狂和疯狂的差异	9
中观视角之先行行业——金融地产是一切纠结情绪产生的源头	10
微观视角——高频数据乃是眼前最大的阿里巴巴宝库	10
投资者行为——门槛最高的量化研究领域之一	11
配置——单一策略向复合策略跨越的领域	12
正 Alpha 行业配置策略——可扩展性慢慢发挥优势	12
基于景气指数的行业选择与配置方法——结合基本面，持续优化中	15
GSRS 为代表的风格配置——未来指数基金发展的掘金工具	16
动量和反转——经典方法，创新主题	16
战术——多空、中性两用策略适应未来市场	17
GSMS——高频数据多空选股的成功范例	18
配对交易——市场中性策略实践尝试	19
CART 决策树——非参数方法+行业内选股	21
算法交易与程序化交易——实现交易模式的现代化	22
算法交易——A 股市场先行者胜	22
程序化交易——股指期货带来发展契机	24
结语：实现盈利的征途——无限未来	25

图表目录

图 1: 国信金融工程量化投资研究体系	5
图 2: 考虑各种细节问题的研究流程.....	6
图 3: 国信量化投资策略体系及应用现状.....	7
图 4: 投资绩效评估体系	8
表 1: 常见风险表征指标历史表现及适用性判断.....	9
图 5: 系统风险与市场走势对比图 (Wind 三级行业)	9
图 6: EMS 择时策略样本外检验效果图 ($R=0.525$ $T=200$)	10
图 7: EMS 择时策略样本内+样本外效果 ($R=0.525$ $T=200$)	10
图 8: 基金仓位估算的逻辑框架图	11
表 2: 一级行业的 α 值相关系数矩阵 (2001 年 - 2009 年 Q1)	12
表 3: 基于相关系数的行业聚类分析(06 年至 09 年 Q4 的 α 相关系数)	13
表 4: 行业分类轮动效果的穷举验证结果 (2002 年至 2010 年 Q1 超额收益 前 10 的分类)	13
图 9: 上证周期、上证非周期及上证 180 指数走势对比.....	14
图 10: 正 α 行业配置策略对 HS300 指数的增强效果	14
图 11: 正 α 行业配置增强资产与基准之间的相对强弱.....	14
图 12: 基于景气指数的行业选择方法.....	15
图 13: 基于景气指数的行业选择示意图	15
图 14: 超配行业组合与基准指数走势比较 (日收益率序列)	15
图 15: 中证 100/500 组合轮动示意图	16
图 16: 中证 100/200 组合轮动示意图	16
图 17: 基于基金收益缺口的反转策略历史检验效果	17
图 18: 基于基金收益缺口的动量反转策略的 Excel 应用模板	17
图 19: 招行 25 日 GSMS 排名与后 20 日 HS300 超额收益.....	18
图 20: 2005 年 4 月 8 日起始的 GSMS 组合比较.....	18
图 21: GSMS 多头组合样本外检验效果	19
图 22: GSMS 多空组合绝对收益样本外检验效果	19
表 5: 利用中证 100 和中证 500 实现量化投资策略的效果(2007-5-10 至 2009-11-20)*	19

图 23: A 股银行股及煤炭股配对交易有效性检验	20
图 24: 不同交易条件设置下的配对交易效果	20
图 25: 配对自动化交易流程	20
图 26: 配对交易组合净值与沪深 300 指数对比	20
图 27: 利用股指期货对冲配对交易组合所持现货头寸	21
图 28: 修正后的国内科技股 CART 动态树	22
图 29: 修正后的 CART 动态树决策效果	22
图 29: 民生银行上涨阶段传统 VWAP 日内走势（无反馈）	23
图 30: 民生银行上涨阶段修正 VWAP 日内走势（无反馈）	23
图 31: 民生银行下跌阶段传统 VWAP 日内走势（无反馈）	23
图 32: 民生银行下跌阶段修正 VWAP 日内走势（无反馈）	23
图 33: 震荡阶段传统 VWAP 日内走势（无反馈）	24
图 34: 震荡阶段修正 VWAP 日内走势（无反馈）	24
图 35: 应用国信金工量化体系独立运行的投资过程	25

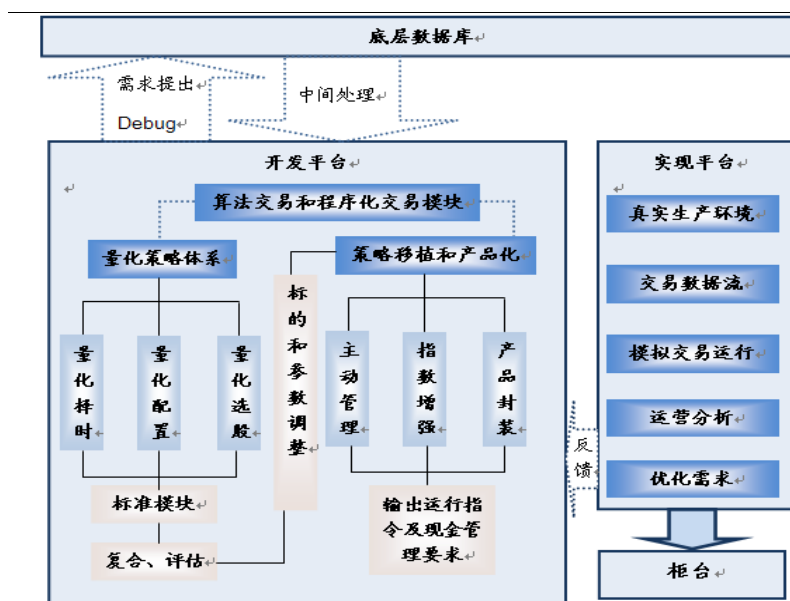
前言——实现收益是我们的唯一目标

本土还是山寨？——什么是检验量化投资的标准

作为 2009 年以来上升速度最快的金融工程团队，国信金融工程团队可以称作是 A 股量化投资在近两年中发展的缩影。从我们团队的构架变化上，可以看到 A 股市场量化投资需求在深度和广度上的前进速度——

2009 年一季度，我们只有 4 名分析师专注于寻找超额收益的主动量化投资策略，“配置”是我们的关键词；2009 年三季度，专职进行量化投资研究的分析师增长到 8 名，团队分工中出现了专门的**产品设计和指数化投资研究**，针对公募基金大量涌现的指数化投资需求，我们的服务体系开始涵盖指数化产品的各个领域——产品设计、回溯检测、数据支持、复制模块、增强策略、活跃交易商；2009 年底，团队分工中再次出现一个新的小组——**算法交易与程序化交易**；2010 年一季度，我们在统一的开发平台上对前期的量化研究进行了痛苦的整合之后，市场给了我们一次锤炼和检验的重大机遇——“**股指期货**”成为一个新的关键词。目前我们专职于量化投资领域的分析师增长到 13 名，而在策略领域，各种层面的研究成果分批出现，择时、配置、选股、交易实施、产品化——标准化流程、整合、优化，使团队的研究成果的实现、再以更好的效果实现，成为新的研究推动力。

图 1：国信金融工程量化投资研究体系



资料来源：国信证券经济研究所

这个发展历程浓缩了 A 股市场发展的一切特征——超常规的速度、低运营成本、带有浓厚本土特征的工作模式、本土培养的核心分析师、市场需求导向、实用主义——如果一定要给我们的团队一个定语，我们非常希望能够成为本土智慧的代表，但是在成熟而强大的海外量化投资体系面前，我们又时时担心和海外优秀的同行相比，我们刚刚趟出来的这条路显得多么山寨！

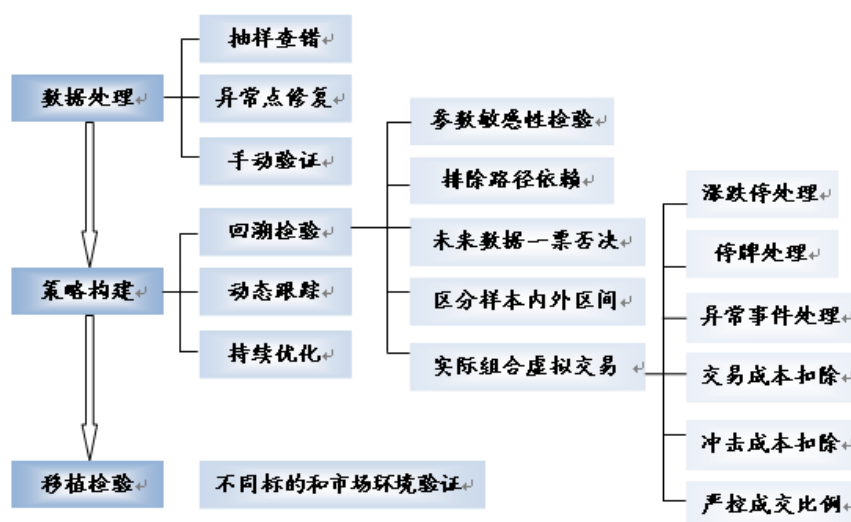
这一担忧推动我们时时回顾、紧盯我们已有的成果在实际投资中的检验效果——**有效地实现收益，是我们自始自终的唯一目标，也是我们认为的，能够区别本土和山寨的唯一标准。**

理论还是细节？——什么是量化投资研究的瓶颈

什么是量化投资研究的瓶颈？这个大而无解的问题看上去毫无意义，螺旋上升的发展过程中，任何问题都可能成为阶段性的瓶颈，理论研究深度、人才储备、软件、硬件、支撑研究投入不断扩大的盈利模式——等等；但是，我们坚持认为量化投资研究获得长久生命力的唯一秘诀在于对研究细节的把握。量化投资领域对理论研究和硬件的需求可以无限制地提高门槛，只要能够实现的预期收益可以涵盖成本的投入，最顶级的数学大师、大型计算机，都不是什么新鲜的东西，但是，所有这些投入是否能真实地转换成生产力，细节是最后的一道关。当然，我们的团队面临的很现实的问题是，我们没有数学大师，也没有小型机，细节是我们这群本土分析师唯一能够触摸的东西，那就让我们把它用到极致吧。

在研究中无限制地考虑细节问题对分析师来说是一件极为吃力不讨好的事情——每一个细节的加入都导致研究的边际成本和工作量以级数级增长，但是无限接近真相的唯一途径就是去了解这些细节、分解、模拟、严苛地处理投资中可能遇到的困难场景和隐形成本。在过去的研究中，我们有些优秀的研究成果已经为团队树立了研究细节的范本，我们所展示出来的研究成果，不允许应用修饰、选择和遮掩，至少在研究过程中必须涵盖以下流程：

图 2：考虑各种细节问题的研究流程



资料来源：国信证券经济研究所

择时、配置还是选股？——什么是量化投资技术擅长的事

什么是量化投资技术擅长的事情？在我们数量化投资系列报告之2——《数量化投资技术综述》中，对各个层面的研究内容和常见联系方式做过系统的阐述，报告中所述的内容，我们也在一一实现。但是在我们推广研究成果的过程中，遇到过各种有趣的调侃和说法——寻找拐点是上帝做的事情！基金经理怎么能容忍我们用计算机超越了他们选择的个股？如果你们的模型全部 work，那么市场将变得多么单调无趣！

图 3: 国信量化投资策略体系及应用现状



资料来源: 国信证券经济研究所

*: 得到持续应用或跟踪的研究成果

我们在这一年摸索的过程中,对量化投资技术最为直观的感受是,这是一种提高工作效率的方法,为什么要执着于投资结论来自市场数据的统计分析还是来自昨天的政府工作报告?市场数据中也包含了无数投资者对政府工作报告的理解和由此导致的行为——我们认为二者的区别无非是,思考的过程是否采纳了海量数据,是否借助了量化的方法,和程序、电脑?注意,是借助,而不是依赖。

有些量化的策略可以得到漂亮的经济学解释,而有些不行,然而解释也许只是一种巧合的对应关系,信任解释和信任方法本身,又有什么区别?有观点认为,来自价格变化的分析结论比不上来自基本面分析结合估值的结论可靠,后者是利用市场未反应的预期或反应过度来实现价值发现和纠错,但前者又何尝不是呢?在我们看来,区别无非是,批量地发现类似的机会还是独立地发现各有其趣的机会?

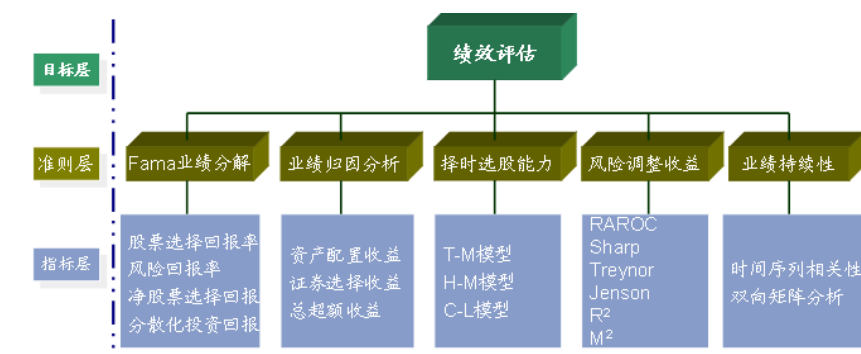
方法无优劣,决定结果的是使用的方法和场合。量化投资技术最擅长的事,是大比例、大概率地知道决策过程的依据、长短和使用场合。也许这一决策过程过于冰冷和枯燥,但实现预期的过程同样很美妙,因为我们摒弃后视镜、双色球的研究思路,因此我们有理由期待自己的智慧结晶得到未来的验证。

择时——宏观、中观和微观视角缺一不可

择时是股票投资永远没有最优解的难题

在我们的数量化投资系列报告之二——《数量化投资技术综述》中，通过对投资业绩的评估体系，分解了投资决策过程各个层面的工作。这其中，择时是一个突出的重点和难点，因为对于任何投资决策过程来说，波动就是机会，而在波动中判断方向，是一个永远没有最优解的难题。

图 4：投资绩效评估体系



资料来源：国信证券经济研究所

从我们目前已有的、经过一定时间样本外检验的择时策略中，我们发现，每种视角都有它适用的场合，在确定的场合下，可以达到特定的决策目标。各种层次的视角未必要相互影响和验证，它们可以给出不同周期和频度的建议。

宏观视角择时——易行的不靠谱，靠谱的不易行

在试图从宏观视角进行择时的摸索中，显然我们没有用上太多的量化模型和方法——二十年不到的历史，仅有寥寥几个经济周期，其中还伴随着资本市场的快速发展和变革，每一个周期都与众不同，无论把货币供应和宏观经济的各种变量如何翻来覆去地处理，都不足以得到严密的推理和结论。

在我们的数量化投资系列之十一——《A 股市场常用风险表征指标分析》中，我们列举了常用的 A 股市场择时的宏观视角，包含宏观经济指标、相对估值指标、股票市场供求关系、金融产品估值参照和非系统风险指标等五类。

这篇报告得到的结论是，宏观经济指标的异常现象出现需要特别关注；常用相对估值指标在逼近历史下限可以关注，上限几乎没有指导意义（而我们的下限也仅仅经受过一轮牛熊周期的检验）；股票市场供求关系则是一个具有分析价值的指标，但是单纯观察供给增量乃至增量的同比环比都没有太大意义，需要对新增量进行估值水平的标准化处理；需求增量确实是较为“靠谱”的指标，但至今没有一个具有强烈说服力的代理变量可供跟踪；金融产品估值水平这一类指标也十分有趣，最有效的指标出现在争议最大的权证市场；至于最后一类指标——我们至今为止仍然给予非常高的警示意义的指标，是目前 A 股行业指数非系统性风险水平及差异化，历史上出现过 5 次警示，每一次都预示了一轮上涨行情的终结。

从这个基础研究中我们得到了宏观视角择时的“潜规则”——易于观察的、人人都在看的指标，往往没有实质的应用价值，而有效的指标则潜伏与那些难以观察和分析的领域。我们可以根据历史的胜率和成熟市场的淘汰结果，在繁多的指标中选取值得长期关注和跟踪的类别，并给予较高的优先级。

表 1: 常见风险表征指标历史表现及适用性判断

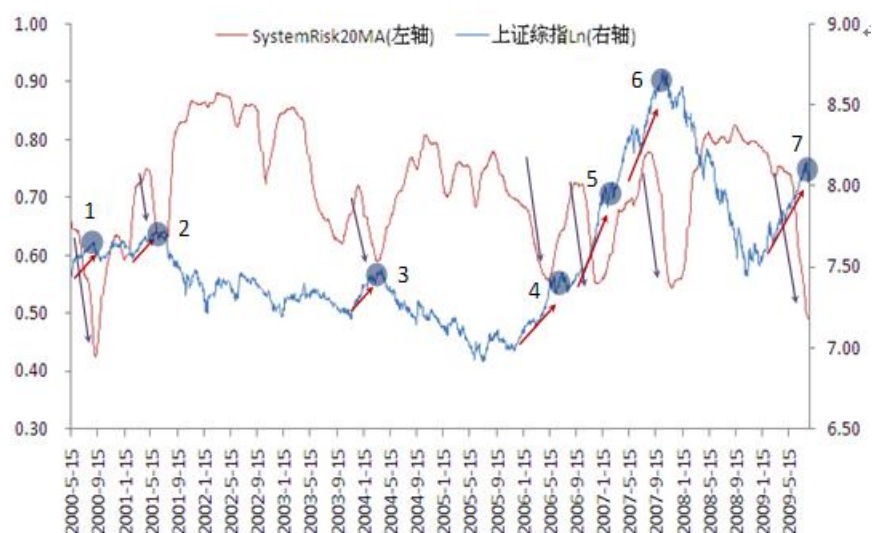
指标	历史有效性分析	是否先验	未来变数
M1、m2、m2-m1 增速	有微弱正相关，时效性较差	m1 增速超过 m2 需警惕	资本市场对外资完全放开
M1-GDP	相关性尚可	没有显著的先验性	资本市场对外资完全放开
市场平均 PE、PB	下限可关注	20 倍 PE 和 2 倍 PB 有一定先验性	多层次市场形成后下限变化
IPOPE/MarketPE	下限需警惕	低于 0.5 需高度警惕	融资制度再次改革
股票增量市值占比	没有显著相关性和量化价值	持续居高需警惕	股改结束，ipo 实行存量发行
股票市场新增开户数	持续迅猛增加时警惕乐观情绪	采样期内对市场顶部有较好先验性	机构投资者占据市场主导位置
封闭式基金到期收益率	某些市场阶段有效	否	金融产品丰富，估值体系健全
认购权证溢价率	极端市场有效	认购权证集体折价需高度警惕	认购权证消亡
行业非系统性风险	阶段性顶部有效	提前实际调整或顶部一个月左右	向美国市场靠拢

数据来源: Wind 资讯, 国信证券经济研究所

预测顶部拐点——只分析疯狂程度不行，还要知道疯狂和疯狂的差异

在我们的数量化投资系列之十四——《从系统风险的历史看当前下跌的未来》中，我们对这一得到验证的指标进行了深入和详细的分析，得到的结论是：在市场处于熊市的时候，系统风险占比一般处于高位；上涨过程中系统风险的急剧释放会对应后续的大幅调整或者顶部，无一例外；上涨过程中系统风险的急剧释放都提前于后续震荡或者见顶；系统风险的急剧释放对市场大幅下跌后的见底指示作用不明显。这一研究其实已经算是有一个中观视角的代表——在对行业的投资判断中，集中了投资者对于宏观形势和个股风险水平的判断。

图 5: 系统风险与市场走势对比图 (Wind 三级行业)



资料来源: 国信证券经济研究所

简言之，通过行业非系统性风险的跟踪，我们可以预见到上涨行情的终结点，但是不能解决下跌行情何时结束的问题。有趣的是，大量的技术指标都倾向于在市场回归理性的底部较为有效，而在上涨过程中持续钝化，可见上涨的终结不仅来源于疯狂的行为，更多的来源于疯狂行为的分化，只分析疯狂程度是不够的，还要知道疯狂和疯狂之间的差异。

中观视角之先行行业——金融地产是一切纠结情绪产生的源头

依据普林斯顿大学 Harrison Hong 等人 2004 年的研究，资本市场存在先行行业，对市场组合收益率及实体经济检测指标都有预测作用，且二者的预测能力正相关。这一结论在全球八大资本市场中的七个得到了验证。在这八个市场中具有先行效应的 14 个行业为 RLEST: 商业地产、TRANS: 交通、MINES: 矿产（金属）、UTILS: 公共设施、APPRL: 服装、RTAIL: 零售、PRINT: 出版、MONEY: 金融、PTRLM: 石化（成品油、炼油）、SRVC: 服务、LETHR: 皮革、STONE: 非金属矿产、METAL: 金属（矿产）、TV: 电视。

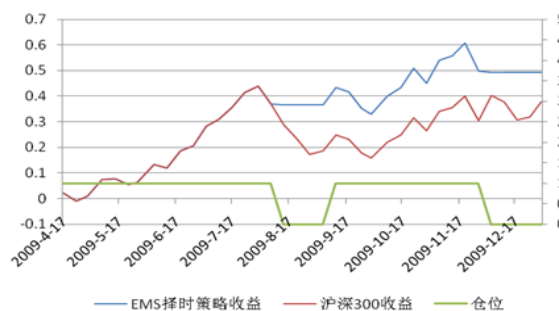
先行效应是由于某些行业富含基本面信息，且信息在市场需要扩散时间所导致的。这一提前效应不会长于两个月。数量化投资系列之十七——《A 股市场的先行行业及先行周期》中，我们对 A 股市场进行检验的结论是：A 股市场确实存在周期小于等于二个月的先行行业，目前为止策略效果最佳的先行指标为 t-1 月的金融行业收益率及 t-2 月的地产行业收益率。金融和地产仍是目前对指数影响最为显著的行业，是一切投资者纠结情绪产生的源头。

这一研究的研究方法和结论都十分简单，它的价值不在于找到过去最准确的周期，到底是八周、九周（两个月）还是十周，而是排除这一结论来自时间序列的自相关，或者来自巧合，找到一个看上去漂亮的规律并不难，难的是验证它确实存在客观的决策价值；Harrison 作为一个成功的学者给了我们很好的范例——用最笨的办法随机打乱时间序列，让计算机通过海量计算证明了他的结论是可靠的。在研究中我们也尝试了类似的验证方法，乱序抽样、穷举、海量计算，这没什么，花费这些时间完全是值得的——报酬就是我们对研究结果更强的信心，以及了解是不是有一个更优解隐藏在不为我们所知的领域。

微观视角——高频数据乃是眼前最大的阿里巴巴宝库

在我们的交易性指标与策略系列之二——《基于有效资金强弱指标（EMS）的择时策略研究》中，我们对之前一个报告中所构建的资金强弱（MS）指标进行了优化并应用于方向判断。按照通常理解，无论是传统资金流量，还是 MS、EMS 指标，当指标值为正时，我们认为资金为净流入或者主动性买入资金多于主动性卖出资金，因此很多人认为该指标越大就越应该看好后市。然而如果我们从另外一个角度来看，指标值越大说明股票的上涨阻力越来越大，因为上涨所需要付出的成交金额越大，特别是 EMS 指标对正向 MS 和负向 MS 按照有效性进行了加权，EMS 越大时，上涨单位价格所需的资金越多，上涨阻力越大；反之，EMS 越小时，下跌单位价格所需的资金越多，下跌阻力越大。

图 6: EMS 择时策略样本外检验效果图 (R=0.525 T=200)



资料来源：国信证券经济研究所

图 7: EMS 择时策略样本内+样本外效果 (R=0.525 T=200)



资料来源：国信证券经济研究所

这一策略在内部的讨论和演练中，超常规的效果激动了我们所有的人，和之前的各种方法相比，包括和一些简单的参数方法相比，它的解释力、稳定性和操作周期都更贴近实际投资的需求。高频数据的分析在我们面前展示了一个前所未有的空间，仅仅是做了一个初步的尝试，就从分析的过程和逻辑中看到了巨大的扩展和应用可能。事实上目前我们所用的 5-6 秒的，所谓 Level-1 最高频的数据，只是中频的概念，随着频度的上升，能够获取的信息量和策略意义将会更加惊人。对于量化投资在 A 股的应用而言，高频数据领域乃是眼前最大的阿里巴巴宝库，寻找咒语没有秘诀，我们需要做的只是相信这个宝库的存在，并不断地尝试各种可能。

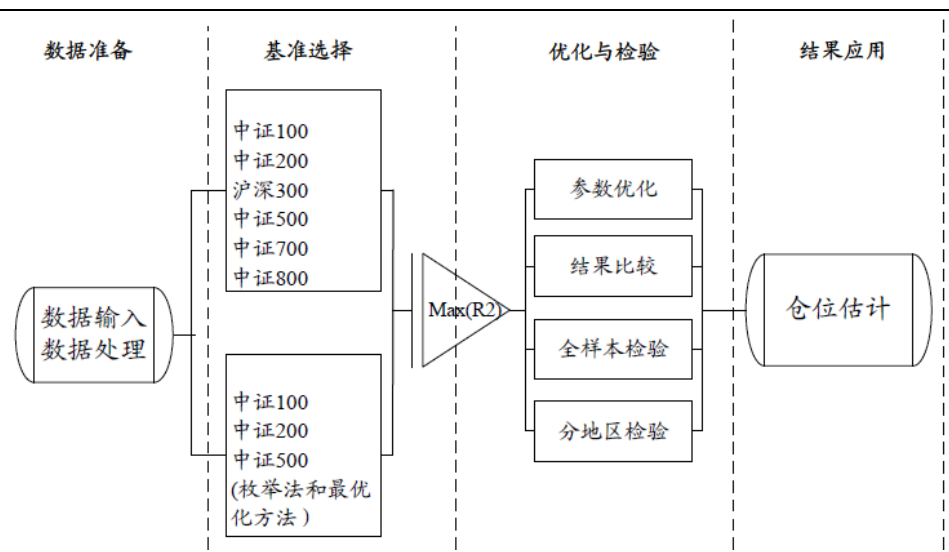
投资者行为——门槛最高的量化研究领域之一

在我们的量化投资系列之十八——《投资者行为与市场运行状况》、量化投资系列之十五——《ST 股票运行规律初探》、量化投资系列之十九——《封闭式基金价值投资分析》、《认购权证运行规律与现货市场》中，我们对投资者行为的代理变量——非系统性风险、ST 股票估值、封闭式基金估值、权证估值及公募基金仓位与分歧和市场运行之间的关系，做了系统的分析和论证。

这些报告中的前几个内容，最终没有得到特别深入的持续应用和跟踪，一个重要的原因是，低频数据的分析已经不足以得到具有高度使用价值的策略，我们希望未来能在较高的频度上实现更有价值的研究。我们认为高频数据和投资者行为分析的结合，已经成为目前 A 股量化投资研究门槛最高的领域之一，也是尤其值得花费大量时间和成本投入研究的一个领域。

在基金仓位及分歧测算这一方向，我们于 2010 年 1 月发布的《开放式股票型基金仓位估算方法研究及应用》中，创新地采用了复合基准指数作为回归基准，采用最优化方法，结合复合基准预测基金仓位比。最优化方法能够得出当前基金仓位和相对应的复合指数里中证 100、200、500 指数（大、中、小盘指数）的权重，提供当前市场在风格配置上的建议。后续的研究扩展中，包括跟踪基金仓位与大盘的相关关系，从调整两组样本的滞后期和样本区间长度两个参数出发，探讨基金仓位与大盘相关性变化所带来的择时策略。

图 8：基金仓位估算的逻辑框架图



资料来源：国信证券经济研究所

配置——单一策略向复合策略跨越的领域

正 Alpha 行业配置策略——可扩展性慢慢发挥优势

我们的量化投资系列报告之七-《基于 Alpha 的行业配置方法和投资策略》、量化投资系列之九-《正 Alpha 策略的 beta 优化》、量化投资系列之十六-《相关系数聚类对正 Alpha 策略的再优化》，对基于超额收益差异最大的行业分类及配置方法，做了系列的研究和验证，接近一年的时间过去，这一策略的持续优化和应用仍在进行中，应用和优化形成了良性循环，出现互相推动的效果。

表 2：一级行业的 alpha 值相关系数矩阵（2001 年 - 2009 年 Q1）

	采掘	交运	金融	地产	钢铁	有色	农林	公用	建筑	信息	商贸	社会	文化	综合	食品	纺织	木材	造纸	石化	电子	机械	医药	其他
采掘	1.00																						
交运	0.14	1.00																					
金融	0.38	0.08	1.00																				
地产	0.12	-0.34	0.21	1.00																			
钢铁	0.41	0.38	0.19	-0.13	1.00																		
有色	0.55	-0.18	-0.03	0.08	0.03	1.00																	
农林	-0.57	-0.17	-0.58	-0.36	-0.11	-0.31	1.00																
公用	-0.15	0.18	-0.10	-0.30	0.16	-0.28	0.09	1.00															
建筑	-0.53	-0.16	-0.40	0.10	-0.42	-0.19	0.36	0.24	1.00														
信息	-0.36	-0.04	-0.48	-0.26	-0.00	-0.34	0.51	0.19	0.03	1.00													
商贸	-0.43	-0.32	-0.45	-0.26	-0.45	0.04	0.53	-0.26	0.29	0.05	1.00												
社会	-0.47	-0.23	-0.46	0.23	-0.37	-0.18	0.34	-0.13	0.40	0.22	0.40	1.00											
文化	-0.31	-0.14	-0.24	0.04	-0.12	-0.26	0.28	-0.08	0.13	0.48	0.12	0.24	1.00										
综合	-0.77	-0.17	-0.62	-0.13	-0.29	-0.21	0.51	0.15	0.52	0.41	0.38	0.58	0.34	1.00									
食品	-0.10	-0.02	-0.06	-0.31	-0.08	0.03	0.30	-0.42	-0.12	0.06	0.57	-0.04	0.13	-0.07	1.00								
纺织	-0.66	-0.22	-0.63	-0.17	-0.41	-0.21	0.53	0.15	0.56	0.15	0.56	0.59	0.17	0.75	0.34	1.00							
木材	-0.29	-0.06	-0.16	0.22	-0.17	-0.14	0.19	-0.29	0.17	-0.01	0.32	0.47	0.10	0.19	0.13	-0.07	1.00						
造纸	-0.49	-0.08	-0.40	-0.26	-0.10	-0.23	0.58	0.01	0.01	0.47	0.30	0.27	0.22	0.46	0.17	0.75	0.01	1.00					
石化	-0.35	-0.04	-0.60	-0.48	-0.02	-0.02	0.73	0.09	0.13	0.43	0.48	0.27	0.08	0.41	0.10	0.19	0.12	0.18	1.00				
电子	-0.50	-0.07	-0.45	-0.28	-0.16	-0.17	0.42	0.31	0.20	0.64	0.13	0.32	0.29	0.63	0.22	0.46	0.15	0.37	0.02	1.00			
机械	-0.40	-0.39	-0.47	-0.04	-0.38	0.23	0.28	0.06	0.36	0.16	0.37	0.26	0.05	0.53	0.08	0.41	0.26	0.54	-0.05	0.58	1.00		
医药	-0.64	-0.31	-0.57	-0.29	-0.40	-0.29	0.64	0.24	0.45	0.42	0.50	0.47	0.26	0.60	0.29	0.63	-0.15	0.44	-0.08	0.48	0.55	1.00	
其他	-0.55	-0.35	-0.57	-0.06	-0.32	-0.11	0.59	0.11	0.39	0.48	0.44	0.55	0.18	0.63	0.05	0.53	0.02	0.44	0.15	0.18	0.23	0.33	1.00
均值	-0.22	-0.07	-0.23	-0.07	-0.08	-0.05	0.25	0.04	0.15	0.18	0.19	0.21	0.13	0.25	0.26	0.60	0.15	0.61	0.17	0.42	0.49	0.42	0.49

数据来源：Wind资讯,国信证券经济研究所

这个相关系数的矩阵，至今仍是我们的正 Alpha 行业配置的核心逻辑依据，它把市场分成了泾渭分明的两类：一类行业包括了金融、地产、钢铁、有色、采掘和交运，其余行业归为另一类。两类行业在持续轮番的 Alpha 轮动中，表现出了非常显著的风险收益特征。后来我们应用了各种各样的方法去反复验证这一分类的有效性，包括采用了聚类分析和穷举验证。

聚类方法让我们看到了行业 Alpha 收益的逐层分离，这个分离的过程中，原有的界限仍然存在显著的意义。用于全市场的这一分类方法，完全可以用于另一个独立而稳定的组合。如果组合的构成不足以形成两个稳定的群体，那么我们就用聚类的方法把它分成三类，乃至四类，只要具备足够稳定的轮动关系，对投资决策就可以形成长期可靠的指导意义。这是正 Alpha 行业配置扩展应用的第一步——用于特定投资组合，用于行业轮动的逐层分离。

表 3: 基于相关系数的行业聚类分析 (06 年至 09 年 Q4 的 Alpha 相关系数)

分类	正 Alpha 策略分类	聚类 (两类)	聚类 (三类)	聚类 (四类)	聚类 (五类)	聚类 (六类)	聚类 (七类)	聚类 (八类)
分类 结果	黑色金属	黑色金属	黑色金属	黑色金属	黑色金属	黑色金属	黑色金属	黑色金属
	交运仓储	交运仓储	交运仓储	交运仓储	交运仓储	交运仓储	交运仓储	交运仓储
	金融服务	金融服务	金融服务	金融服务	金融服务	金融服务	金融服务	金融服务
	房地产	房地产	房地产	房地产	房地产	房地产	房地产	房地产
	采掘行业	采掘行业	采掘行业	采掘行业	采掘行业	采掘行业	采掘行业	采掘行业
	有色金属	有色金属	有色金属	有色金属	有色金属	有色金属	有色金属	有色金属
	食品饮料	食品饮料	食品饮料	食品饮料	食品饮料	食品饮料	食品饮料	食品饮料
	木材家具	木材家具	木材家具	木材家具	木材家具	木材家具	木材家具	木材家具
	建筑行业	建筑行业	建筑行业	建筑行业	建筑行业	建筑行业	建筑行业	建筑行业
	信息技术	信息技术	信息技术	信息技术	信息技术	信息技术	信息技术	信息技术
	商业贸易	商业贸易	商业贸易	商业贸易	商业贸易	商业贸易	商业贸易	商业贸易
	社会服务	社会服务	社会服务	社会服务	社会服务	社会服务	社会服务	社会服务
	文化传播	文化传播	文化传播	文化传播	文化传播	文化传播	文化传播	文化传播
	综合行业	综合行业	综合行业	综合行业	综合行业	综合行业	综合行业	综合行业
	农林牧渔	农林牧渔	农林牧渔	农林牧渔	农林牧渔	农林牧渔	农林牧渔	农林牧渔
	纺织服装	纺织服装	纺织服装	纺织服装	纺织服装	纺织服装	纺织服装	纺织服装
	公用事业	公用事业	公用事业	公用事业	公用事业	公用事业	公用事业	公用事业
	造纸印刷	造纸印刷	造纸印刷	造纸印刷	造纸印刷	造纸印刷	造纸印刷	造纸印刷
	石油化工	石油化工	石油化工	石油化工	石油化工	石油化工	石油化工	石油化工
	电子行业	电子行业	电子行业	电子行业	电子行业	电子行业	电子行业	电子行业
	机械设备	机械设备	机械设备	机械设备	机械设备	机械设备	机械设备	机械设备
	医药生物	医药生物	医药生物	医药生物	医药生物	医药生物	医药生物	医药生物
	其他制造业	其他制造业	其他制造业	其他制造业	其他制造业	其他制造业	其他制造业	其他制造业

数据来源: 国信证券经济研究所

正 Alpha 行业配置策略催生了两个有趣的指数, 上证周期 (000063.SH) 和上证非周期 (000064.SH), 分别对应上证资产的两类行业。在开发指数的过程中, 为了验证我们的行业分类能够得到超额收益差最大的效果, 我们对一级行业的其他划分方法做了接近穷举的验证。

表 4: 行业分类轮动效果的穷举验证结果 (2002 年至 2010 年 Q1 超额收益前 10 的分类)

排名	一类行业名称	平均超额收益差	切换次数
1	采掘, 金融, 交运, 金属非金属, 房地产	0.034553	78.00
2	采掘, 金融, 木材家具, 交运, 金属非金属, 房地产	0.034477	85.79
3	采掘, 金融, 交运, 电气水, 金属非金属, 房地产	0.034428	88.40
4	采掘, 金融, 交运, 社会服务, 金属非金属, 房地产	0.034313	90.49
5	采掘, 金融, 木材家具, 交运, 社会服务, 金属非金属, 房地产	0.034241	92.58
6	采掘, 金融, 交运, 传播文化产, 金属非金属, 房地产	0.034239	93.55
7	采掘, 金融, 木材家具, 交运, 传播文化, 金属非金属, 房地产	0.034161	94.58
8	采掘, 金融, 交运, 金属非金属, 造纸印刷, 房地产	0.034094	95.39
9	采掘, 金融, 交运, 金属非金属, 制造业, 房地产	0.034036	97.09
10	采掘, 金融, 木材家具, 交运, 金属非金属, 造纸印刷, 房地产	0.034017	96.52

数据来源: 国信证券经济研究所

上证周期（000063.SH）和上证非周期（000064.SH）的发布，对正 Alpha 行业配置策略是一次严格的检验，检验的效果每天都通过指数的变化不断更新，我们在对这两个指数进行细化研究的过程中，也看到了更多量化策略 1+1>2 的可能。指数发布之后，表现出显著的差异，为市场行业热点的跟踪以及相应的投资收益快速实现提供了一对利器。指数的发布和未来指数产品的产生，使得两类行业轮动可以通过指数产品的交易，实现各种投资周期的轮动效果，不再局限于股票配置。我们目前所致力于深入挖掘的市场中性策略、高频数据分析和高频交易、择时体系，未来都可以通过与这两个指数及其相关的产品的结合大展拳脚。

图 9：上证周期、上证非周期及上证 180 指数走势对比

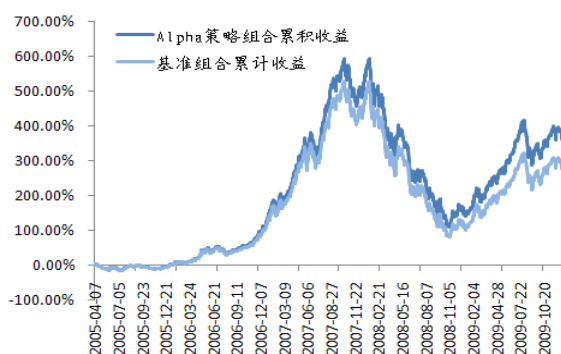


资料来源：Wind 资讯，国信证券经济研究所

除此之外，正 Alpha 行业配置策略在指数增强领域，也表现出了很好的效果和持续稳定增强的能力，我们利用正 Alpha 行业配置策略对沪深 300 指数构建了标准化的指数增强模块，细化到每一天的现金进出和每一笔交易的冲击成本控制，得到了一个几乎完美接近现实操作的增强结果。这个模块的构建，也成为我们未来利用量化策略进行指数增强的一个研究范本。

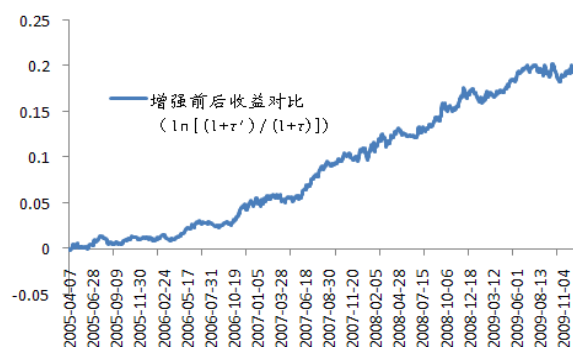
在数量化投资系列之二十四——《利用股指期货对量化策略投资组合进行 Beta 管理流程范例》中，我们展示了正 Alpha 行业配置对于沪深 300 指数的增强效果，但事实上，沪深 300 由于所覆盖的上市公司只涵盖中证规模指数中的大、中盘，这一增强效果还不是最佳，对于目前中证规模指数中同时涵盖大、中、小市值的中证 800，正 Alpha 行业策略还有更漂亮的应用空间。

图 10：正 Alpha 行业配置策略对 HS300 指数的增强效果



资料来源：国信证券经济研究所

图 11：正 Alpha 行业配置增强资产与基准之间的相对强弱

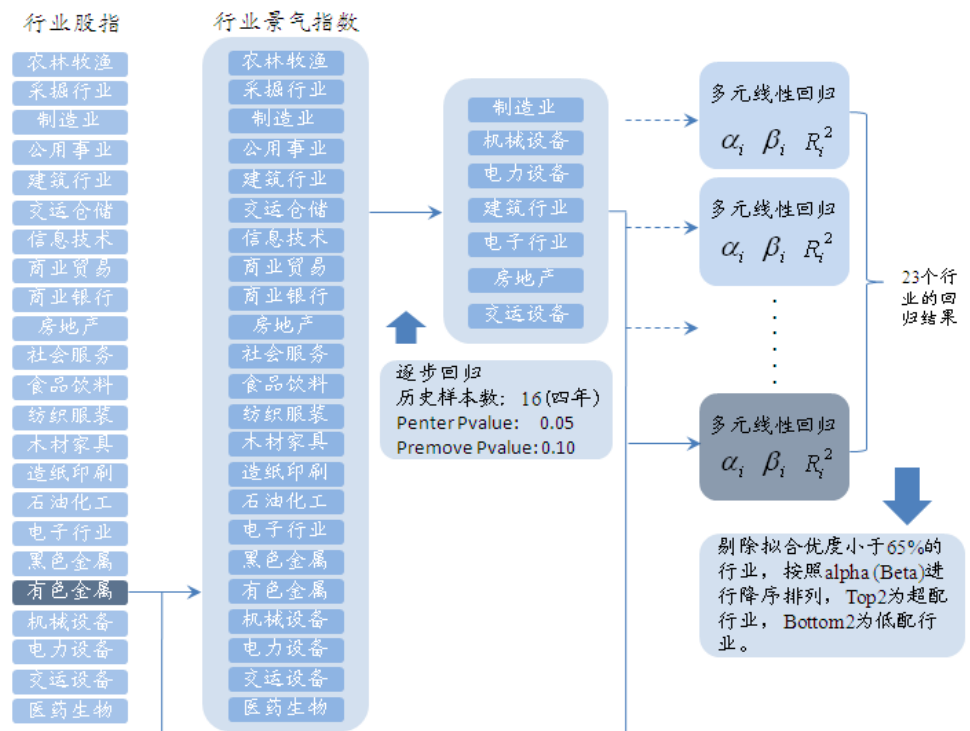


资料来源：国信证券经济研究所

基于景气指数的行业选择与配置方法——结合基本面，持续优化中

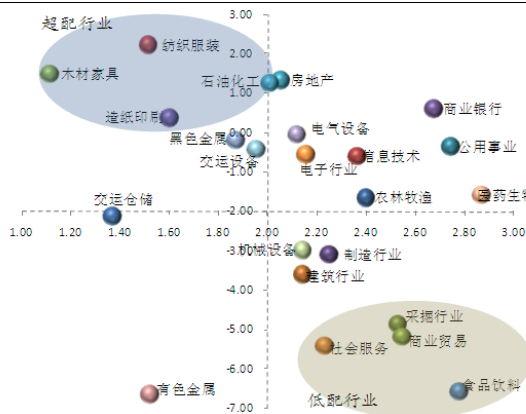
数量化投资系列报告之三——《基于景气指数的行业选择与配置方法》，是我们的系列报告中与基本面结合最为紧密的一个研究方向，核心逻辑是：先证明行业股指和行业景气指数之间存在较强的相关性，在此基础上，考虑产业链的作用，即相关行业景气指数对此行业股指的影响。最后考虑估值，剔除 40%相对估值偏高的行业，然后对剩下的行业按照 Alpha 进行降序排列，选择 Top2 行业作为超配行业；剔除 40%相对估值偏低的行业，然后对剩下的行业按照 Alpha 进行降序排列，选择 Bottom2 行业作为低配行业。

图 12：基于景气指数的行业选择方法



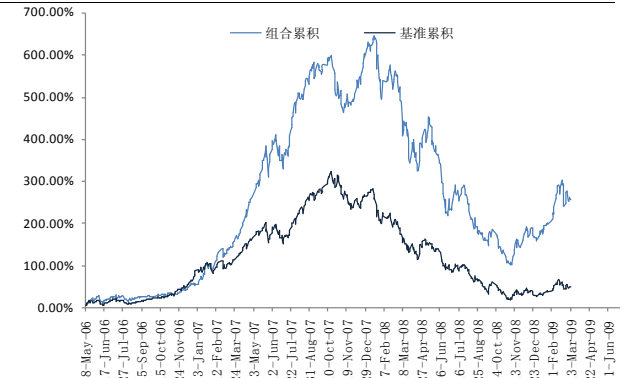
资料来源：Wind 资讯，国信证券经济研究所

图 13：基于景气指数的行业选择示意图



资料来源：国信证券经济研究所

图 14：超配行业组合与基准指数走势比较（日收益率序列）



资料来源：国信证券经济研究所

这一课题的优化，重点在于对行业进行投资评级时，在原有的景气度与估值模型的基础上，结合了宏观景气周期的判断，即投资时钟的应用，得到的结论更加贴近实际投资的需求，也得到了更为稳定的效果。在基本面相关数据的挖掘上，我们显然还有更多的领域可以涉猎，包括隐含大量信息的财务数据，可以把我们在配置上的优秀成果通过精选个股再次提升。

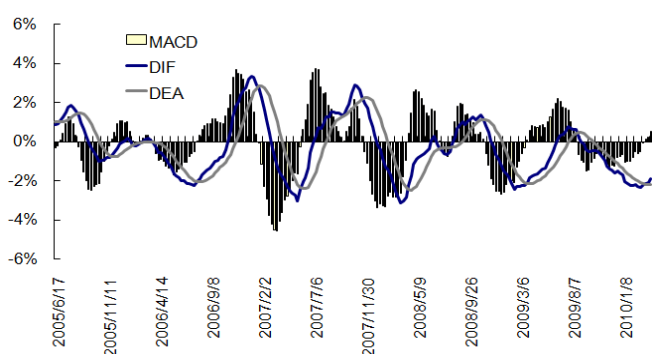
GSRS 为代表的风格配置——未来指数基金发展的掘金工具

在指数基金飞速发展的 2009 年，指数基金产品的发展进入了指数争夺的阶段，指数产品市场份额的抢占其实对应了未来指数所蕴含的投资因子在市场中受投资者认可和应用的空間，我们预计，在规模指数、行业指数多角度发掘之后，还会有众多突出风格因子的指数受到认可，因此风格配置相关的量化研究是未来指数基金，尤其是 ETF 产品发展的掘金工具。

《国信相对强弱方法在大小盘风格轮动中的应用》，是我们的数量化投资系列报告中直接针对风格轮动所进行的一次尝试，这篇报告的一个鲜明特点是引入了对传统技术分析指标的改造，应用过程中，输入和输出参数都处理成简单规范的模式。具体来说，借鉴 MACD 的思想，使用新的相对收益率序列，计算 Rn_Series 的 DIF 指标和 DEA 指标，最终构造了国信相对强弱指标 GSRS。当 GSRS 首次由负转正的时候，说明大盘股指数强于小盘股指数的趋势得到确认，这时应该超配大盘股；反之则应该超配小盘股。同时考虑到 MACD 固有的滞后后，我们还使用 PSY 来对大小盘风格转换进行辅助判断，和 GSRS 定义一样，只是输入参数不再是股价序列，而是 $Rn_Series(t)$ 。

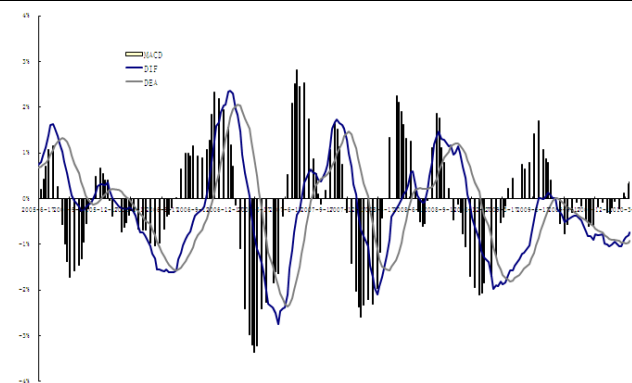
我们对每一个研究领域中所采纳的新方法都十分珍惜，因为在研究过程中，我们所共享的开发平台使得这些新方法的标准化模块将来可以快速移植到其他领域的研究中。大小盘风格轮动所采用的方法，将来可以应用于一切类似的对立风格的轮动研究。简单的输入和输出处理，使得这些研究成果的对接和产品封装成本大大降低，未来我们还将风格轮动领域批量地测试各种对立风格。

图 15：中证 100/500 组合轮动示意图



数据来源：国信证券经济研究所

图 16：中证 100/200 组合轮动示意图



数据来源：国信证券经济研究所

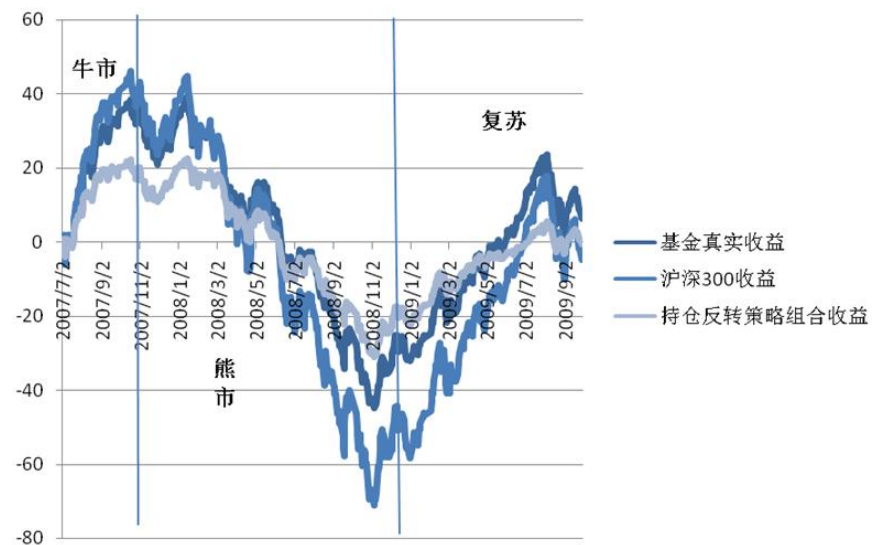
动量和反转——经典方法，创新主题

动量和反转策略在量化投资策略中属于十分经典的常用策略，我们对传统收益变量的动量和反转策略做过很多简单的尝试，未见特别有说服力的结果，常常是样本内可通过参数调整得到可靠结论（少样本的穷举论证尤其可以得到看似完美的验证效果），样本外效果飘忽不定，属于典型的后视镜或双色球策略。

数量化投资系列报告之二十一——《基于收益缺口的基金持仓动量/反转策略

研究》中，我们用动量和反转策略分析了基金收益缺口，也就是通过观测基金实际表现和以基金上一公告日公告的投资组合的虚拟资产组表现之间的差值，来观测基金的潜在行为对基金的影响。当然，我们从公开数据中只能得到基金在定期报表中所公布的组合，但是所得到的策略结果已经显示出非常值得注意的特征——基于基金收益缺口的反转策略不仅优于动量策略，其收益还显著优于基金原有持仓组合的表现，可见对于基金公布的持仓组合而言，反转策略还是有一定的用武之地的，是对已有持仓组合的一种有效改善，总体表现呈现牛市不牛，熊市不熊的态势，对特定市场环境下的投资具有很强的指导意义。

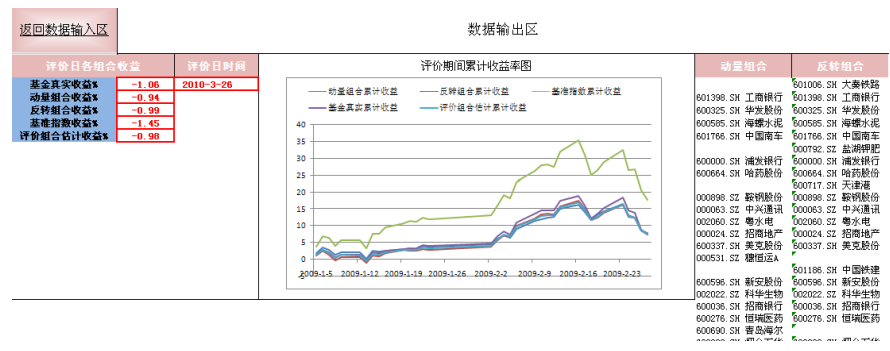
图 17：基于基金收益缺口的反转策略历史检验效果



资料来源：国信证券经济研究所

如果这个策略的研究就到此为止了，那么显然不是我们以收益为目标的研究准则所允许的，在基金日常持仓数据的辅助下，动量和反转策略至少可以为基金经理们提供两个确定的参考组合，这两个组合的日常表现与实际业绩的跟踪比较，对基金经理调整自己的投资思路和风格具有很强的实用价值。

图 18：基于基金收益缺口的动量反转策略的 Excel 应用模板



资料来源：国信证券经济研究所

战术——多空、中性两用策略适应未来市场

GSMS——高频数据多空选股的成功范例

延续我们数量化投资系列的开篇之作《资金流量指标应用分析》中分钟级数据的研究，GSMS 策略以其严密的推论过程，严谨的样本内检验和样本外检验的超高胜率，为我们量化投资策略的推广和延续提供了一个成功的范例。

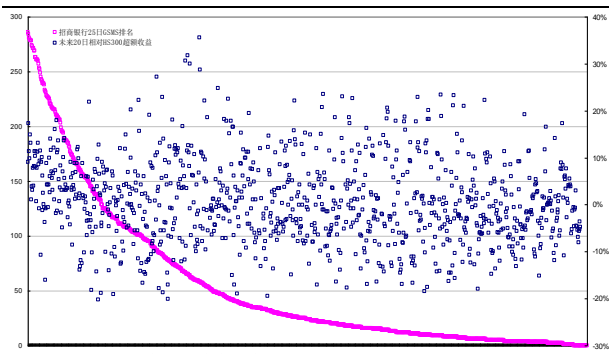
为了解决传统资金流量算法在秒级超高频数据下应用的困难，我们设计了一种新的资金强弱指标算法（Money Strength），这种算法参考实时盘口将超高频数据下区段报价前后相等时的资金强弱进行区分，同时对盘中涨停和跌停时的交易数据进行特殊处理。个股 MS 指标在对超级大盘股未来一段时间相对基准指数超额收益方面也有不错的预测效果。

盘中异常大笔交易的扰动是影响 MS、MF 等资金算法应用效果最重要因素之一。我们设计了国信资金强弱指标（GuoSen Money Strength），来解决这一问题。实证数据表明，GSMS 可以有效的削弱个股异常数据扰动的影响，该指标在对沪深 300 成份股的应用效果检验中，在 20-30 个交易日的观察和持有窗口期内，体现出极其明显的价格反转效应。遍取沪深 300 指数挂牌后的所有交易路径，观察期 GSMS 排名靠后的股票在持有期相对基准都有长期稳定的超额收益。

GSMS 所选用的 MS 算法在秒级数据下的结论与我们之前用 MF 算法在分钟级数据下的结论并无冲突和矛盾，相反两者在不同的数据频率下都找到了类似且非常稳定的统计规律，相互印证并都在实际交易检验中得到不错的效果。这为我们以后分析不同频率下的交易模式提供了非常宝贵的经验。

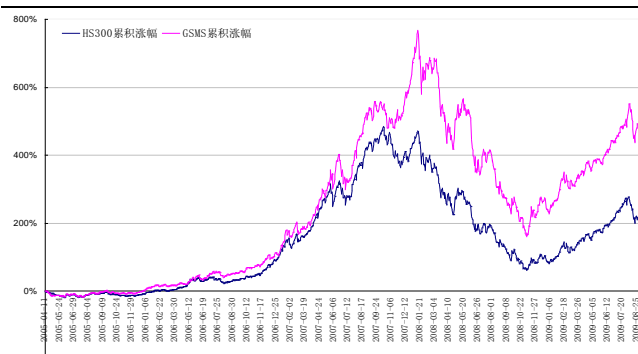
在纯样本外检验的过程中，我们从 09 年 12 月至 10 年 3 月下旬的 20 周中，取得了多头组合样本外 19 周连胜沪深 300 指数，1 周略输的优异战绩，面对这样的样本外检验结果，我们获取了把类似套路的研究分析应用到其他领域的信心。对于不同样本的宽基指数，这一策略至少可以用于寻找特定周期内相对指数的多空组合，达到增强指数的效果。

图 19: 招行 25 日 GSMS 排名与后 20 日 HS300 超额收益



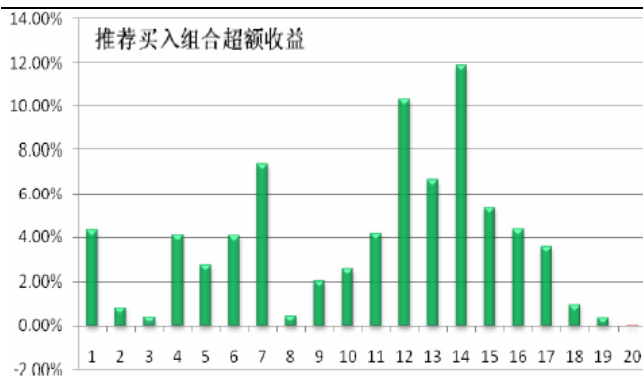
数据来源: 国信证券经济研究所

图 20: 2005 年 4 月 8 日起始的 GSMS 组合比较



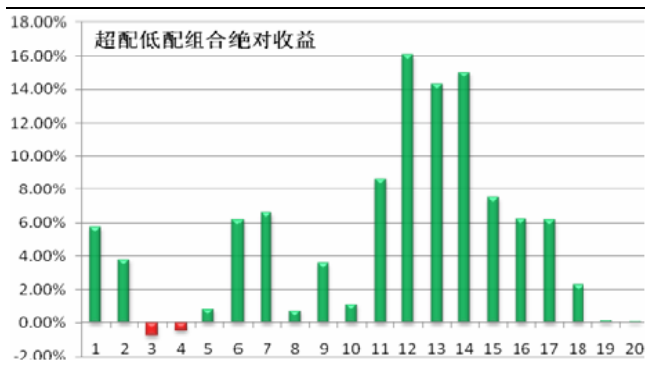
数据来源: 国信证券经济研究所

图 21: GSMS 多头组合样本外检验效果



数据来源: 国信证券经济研究所

图 22: GSMS 多空组合绝对收益样本外检验效果



数据来源: 国信证券经济研究所

配对交易——市场中性策略实践尝试

我们在海外量化技术本土化系列的开篇之作《对冲基金综述及 A 股市场应用探讨》中对海外对冲基金常用盈利模式是否在 A 股有实践的可能，做了力所能及的分析，在报告撰写之时，股指期货尚未有确定消息，因此我们判断，利用包含各种投资因子的指数基金产品为工具进行多空策略的实践，是成本最低、可行性最高的、我们可以在 A 股部分实现的对冲基金的盈利模式。报告中我们简单采用中证 800 的一、二类指数和中证规模指数中的大盘指数基金（中证 100）及小盘指数基金（中证 500），在模拟融券环境下，计算了 Alpha 动量策略可实现的盈利空间，震荡市中效果尤为显著。

表 5: 利用中证 100 和中证 500 实现量化投资策略的效果(2007-5-10 至 2009-11-20)*

切换依据	策略	基准	交易 次数	理论 收益率	双边 胜率	双边 胜率
60 日 Alpha	轮动（轮换买入正向 Alpha 指数）	中证 800	29	-7.88%	-17.91%	-8.41%
60 日 Alpha	准卖空（轮换卖出负向 Alpha 指数买入正向 Alpha 指数）	中证 800	29	5.68%	-5.68%	5.10%
20 日 Alpha	轮动（轮换买入正向 Alpha 指数）	中证 800	42	-7.27%	-21.40%	-8.03%
20 日 Alpha	准卖空（轮换卖出负向 Alpha 指数买入正向 Alpha 指数）	中证 800	42	0.30%	-15.02%	-0.50%
60 日 Alpha	轮动（轮换买入正向 Alpha 指数）	Wind 全 A	29	59.94%	42.83%	59.04%
60 日 Alpha	准卖空（轮换卖出负向 Alpha 指数买入正向 Alpha 指数）	Wind 全 A	29	46.97%	31.17%	46.17%
20 日 Alpha	轮动（轮换买入正向 Alpha 指数）	Wind 全 A	48	143.28%	101.30%	140.99%
20 日 Alpha	准卖空（轮换卖出负向 Alpha 指数买入正向 Alpha 指数）	Wind 全 A	48	58.70%	31.80%	57.27%

*: 同期 wind 全 A 指数累计收益率为 3.107%，中证 800 指数累计收益率为-0.08%。

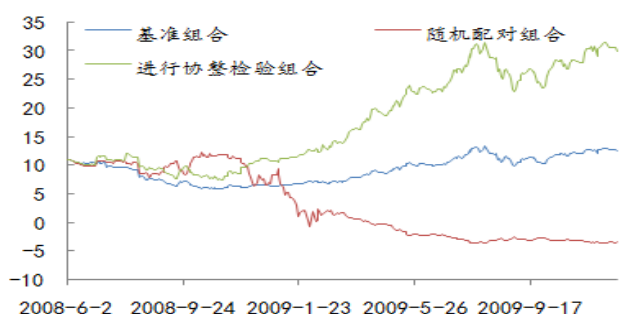
数据来源: Wind 资讯,国信证券经济研究所

融资融券和股指期货业务确立开设之后，市场中性策略在 A 股的实践即将进入较为实质的阶段，之所以说“较为”，是由于我们判断在市场发展初期，由于参与者有限及部分业务资格准入门槛较高，可能导致众多理论盈利空间的存在，但是在流动性匹配的情况下这些盈利空间可能迅速消失或转变形式。

海外量化技术本土化系列的二——《配对交易综述及其 A 股市场应用》中，我们采用全现货模拟形式，综合考虑基本面因素和技术面因素之后，我们从同质性较强的煤炭和银行业中分别精选个股 15 只和 14 只作为基准股票池，分别以波动率和相对价格倍数作为交易触发条件构建配对交易系统。

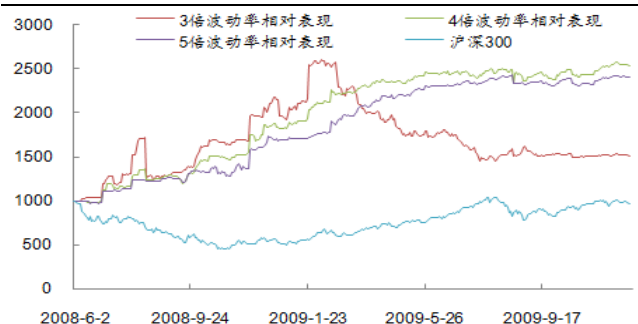
从样本内检验的效果来看，煤炭行业的效果要明显好于银行业，因为煤炭行业股票的波动高于银行业，配对交易操作空间比银行业大。在单边市中效果好，而在震荡市中效果差。究其原因，单边市中，同行业股票的上涨或者下跌具有持续性，个股之间则表现出来一定的时间先后性，对于配对交易而言是极好的交易环境；而在震荡市中，市场热点散乱，持续性都非常短，轮动速度异常迅速，在这种市态下，很难有足够进行配对套利的时间和空间。

图 23: A 股银行股及煤炭股配对交易有效性检验



数据来源：国信证券经济研究所

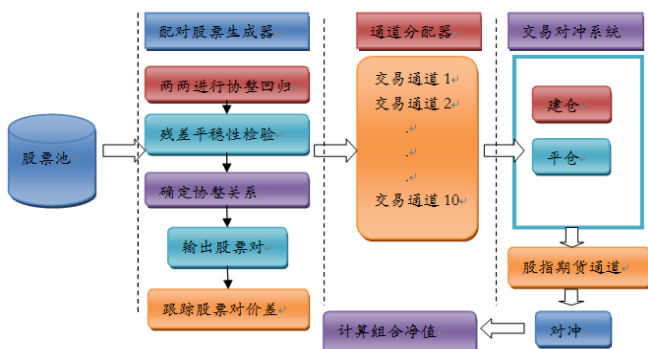
图 24: 不同交易条件设置下的配对交易效果



数据来源：国信证券经济研究所

融资融券及股指期货业务确定开设后，在海外量化技术本土化系列报告之三——《市场中性策略综述及自动化交易系统》中，我们对配对交易策略进行了较大的改造，完成了一个自动化交易模块的设计。样本内测试的结果非常接近一个自动化交易系统的特征：实际测试期是从 2008-06-02 至 2010-02-12，总共 420 个交易日。期间组合净值增长 11.32%，年化收益率为 6.38%。净值最高 1.208 元，最低 0.968 元。在所有 420 个交易日中，有 12 个交易日的净值低于面值。投资者可能的最大亏损比例为 16.87%，最大获利比例为 24.84%。400 次交易中，胜率 52.75%，组合收益与市场状况基本无关，累计收益对交易成本敏感。

图 25: 配对自动化交易流程



数据来源：国信证券经济研究所

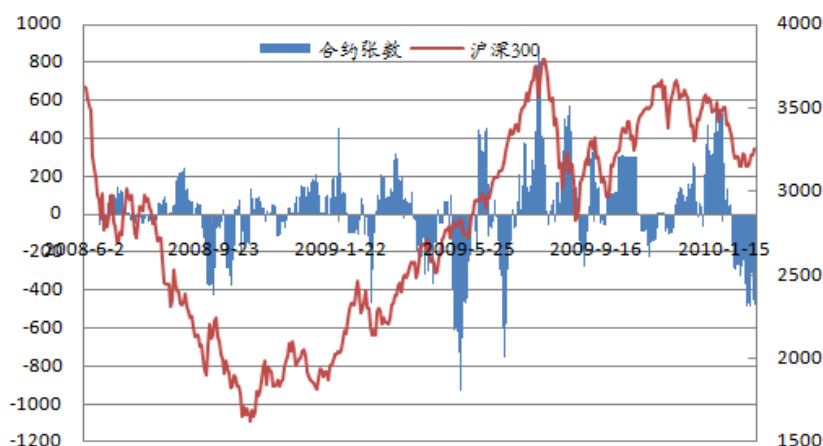
图 26: 配对交易组合净值与沪深 300 指数对比



数据来源：国信证券经济研究所

因为配对交易配对股票之间的比例是根据协整关系确定的，导致多空股票的资金规模可能不一样大，所以在建立头寸时，并不能完全对冲系统性风险。为能够有效的对冲系统性风险，我们所设计的自动交易系统在每天收盘后，计算组合的系统性风险（将所有的多空股票看作一个组合，然后计算整个组合的 β 值，计算采用前 100 个交易日的日收益率数据），然后根据系统的 β ，利用股指期货对冲组合的系统性风险。

图 27：利用股指期货对冲配对交易组合所持现货头寸



资料来源：国信证券经济研究所

我们所计划的市场中性策略体系中，采用传统协整关系进行个股选择只是最为基础的一部分。我们的配置策略+配对交易体系，完全可以实现又一次的“1+1>2”。在配置策略中我们有一套寻找超额收益差最大的两类资产的方法，通过这一方法我们可以找到股票市场中拒不见面的“参商”组合。通过它们之间的差异程度来构建配对交易品种。结合我们的高频分析，这一思路可以向最高频度的交易无限进发，通过交易机会的增加分散风险，把配置策略所能找到的配对关系在市场所允许的范围内应用到极致。除此之外，各种风格因子还可以不断加入这一策略，为组合构建的精度加分。

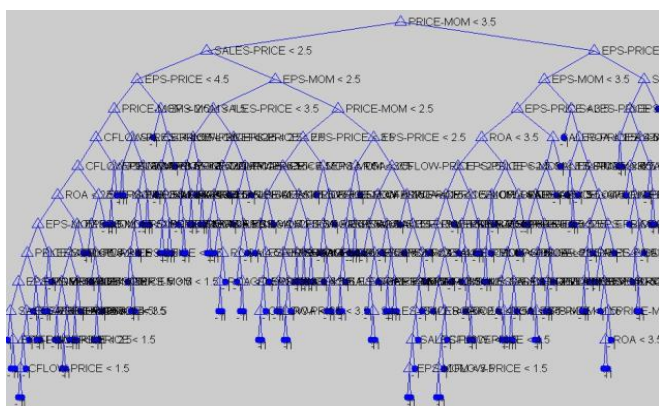
CART 决策树——非参数方法+行业内选股

数据挖掘技术具有自动发觉未知的数据模式和对未来进行预测分析的力，而根据国外学者对数据挖掘在资本市场的应用特性总结来看，决策树具有非常高的解释能力和较低的拓展能力。《基于 CART 决策树的行业选股方法》中，我们尝试对动态决策树模型面对市场环境变化的拓展性进行提升，在这一经典方法的应用中提出了创新的方法。

借鉴 Eric H, Keith L 和 Chee K 所提出的决策树方法并参考其指标，我们对国内科技股板块从 2003 年到 2009 年的数据进行静态树和动态树模型的检验，并构造相应的多空组合。检验结果表明，若不进行树的修正优化，静态树模型所选多空组合收益差为 0.64%，动态树模型所选多空组合收益差为 0.89%，考虑到交易成本、冲击成本，我们认为模型效果并不显著。

通过事前修剪、事后修剪以及节点过滤，我们对静态和动态 CART 决策树进行了修正以增加其拓展性。检验结果表明，修正后的模型多空组合收益差明显提高，而动态 CART 决策树分类后的股票多空差异显著，达到了 2.98% 的水平，且多头组合财富曲线显著超越了空头组合和科技股全样本平均表现。从选股数量上来看，多空组合所含股票不足全样本的一半，有效降低了选股难度，体现出分类的准确性。利用决策树进行行业内选股的另外一个显著优点是它包含了“一致预期”中的未来信息。

图 28: 修正后的国内科技股 CART 动态树



数据来源：国信证券经济研究所

图 29: 修正后的 CART 动态树决策效果



数据来源：国信证券经济研究所

将非参数方法引用到行业内选股并达到有效区分多空组合的效果,使我们之前各个层面的策略研究具备了落地的可能。通过对这一研究方法的掌握,我们得以在指数化配置的基础上提高品种选择带来的投资效率。从择时→配置→选股,我们已经从各个层面上都得到了具备实用价值的方法,并通过标准化模块进行了大量的研究储备。从实际投资的角度来说,现在剩下最后一个层面的问题——交易。

算法交易与程序化交易——实现交易模式的现代化

算法交易——A股市场先行者胜

目前我们关于算法交易的专题报告有两篇:《算法交易及其在 A 股的实证分析》、《VWAP 策略模拟效果及未来扩展》;但在算法交易的研究领域里,绝对不是通过报告就可以证明我们做了多少事情的。首先我们在自身的量化研究中坚持采用算法交易进行冲击成本的测算,尽管是虚拟的,但至少代表了一种态度——我们需要对所有的交易进行成本测算和控制。

我们的目标是，将实用的算法交易模块真正嵌入到投资者的交易系统中，让这一交易模式在 A 股真正得到实践。无论这其中还有多少问题需要解决，对我们而言，目前已经没有不可逾越的障碍。

算法交易，也被称为自动交易（Automated Trading）、黑盒交易（Black-box Trading）、无人值守交易（Robot Trading），是使用计算机来确定订单最佳的执行路径、执行时间、执行价格以及执行的数量一种交易方法。在海外，算法交易广泛应用于对冲基金、企业年金、共同基金以及其他一些大型的机构投资者，他们使用算法交易对大额订单进行分拆，寻找最佳的路由和最有利的执行价格，以降低市场的冲击成本、提高执行效率和订单执行的隐蔽性。任何投资策略都可以使用算法交易进行订单的执行，包括做市、场内价差交易、套利、或者纯粹的投机（包括趋势跟随交易 Trend Following）。

算法交易在交易中的作用体现在以下几个方面：1) 智能路由；2) 降低冲击成本；3) 提高执行效率；4) 减少人力成本；5) 增加投资组合收益。

我们采用两种方法实行 VWAP 下单策略:

(1)、按照历史高频计算各个周期的交易量分布函数——>将需要执行的订单按照交易量分布函数中各个周期分别执行; (没有反馈)

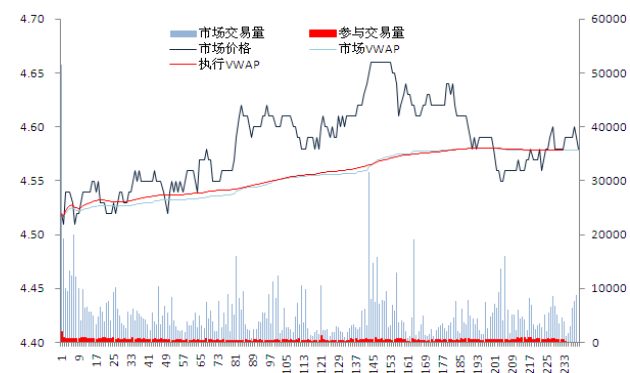
(2)、按照历史高频计算各个周期的交易量分布函数——>按照当日已经发生的交易量预测全天交易量——>分配各个剩余周期的订单数量 (循环最后两步直至订单执行完毕); (使用反馈)

这两种方法的区别在于: 传统 VWAP 要求尽量跟住市场交易量加权平均价格, 而修正 VWAP 要求尽量战胜市场交易量加权平均价格。同时, 在对交易量估计的时候, 还分为有反馈估计和无反馈估计, 无反馈估计只是用历史上日内每个周期的交易量均值作为当天交易量的估计值, 有反馈估计则会根据当日已经发生的交易量对剩余交易量进行修正。

我们目前的模拟效果具有以下特点:

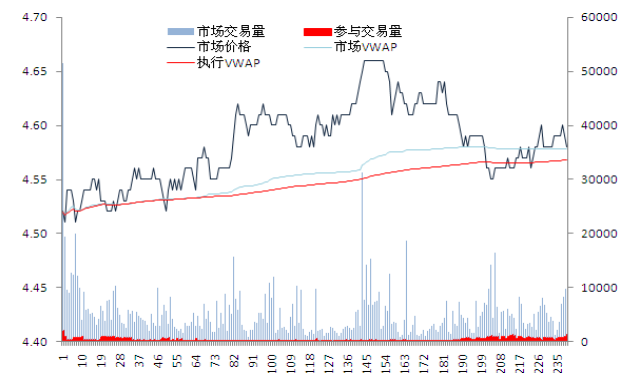
- 1、使用VWAP策略时, 大盘股的效果要好于小盘股, 这是由于流动性的差异造成的;
- 2、以1分钟为执行周期, 其VWAP效果要好于以5分钟为周期的执行效果。后续在超高频数据方面进行扩展是可行的;
- 3、策略2的效果显著优于策略1, 这说明使用反馈机制的VWAP优于不使用反馈机制的VWAP。同样的, 后续在反馈机制方面进一步深入是值得的。

图 29: 民生银行上涨阶段传统 VWAP 日内走势 (无反馈)



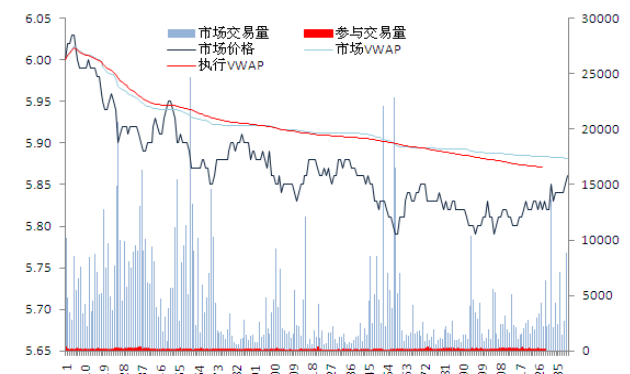
数据来源: 国信证券经济研究所

图 30: 民生银行上涨阶段修正 VWAP 日内走势 (无反馈)



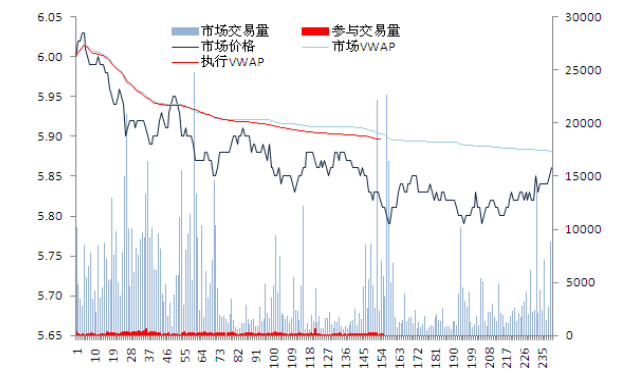
数据来源: 国信证券经济研究所

图 31: 民生银行下跌阶段传统 VWAP 日内走势 (无反馈)



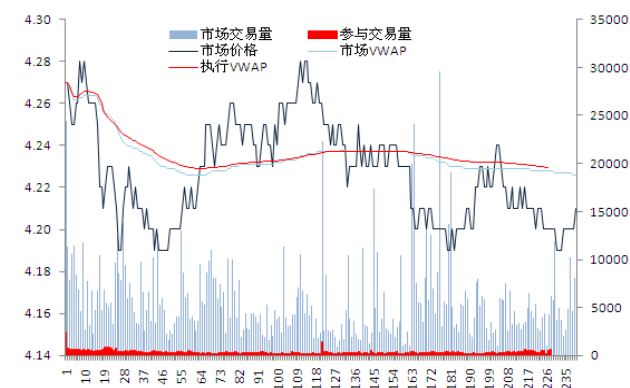
数据来源: 国信证券经济研究所

图 32: 民生银行下跌阶段修正 VWAP 日内走势 (无反馈)



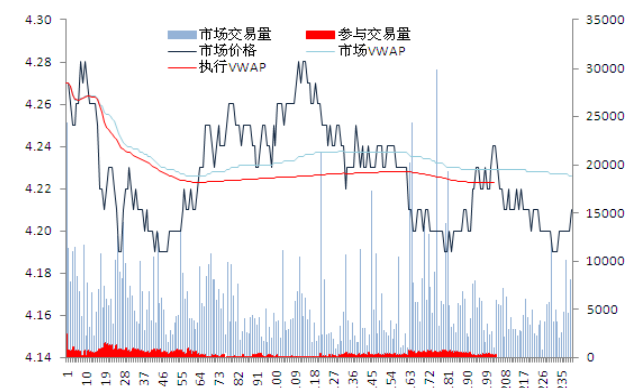
数据来源: 国信证券经济研究所

图 33: 震荡阶段传统 VWAP 日内走势 (无反馈)



数据来源: 国信证券经济研究所

图 34: 震荡阶段修正 VWAP 日内走势 (无反馈)



数据来源: 国信证券经济研究所

目前我们的算法交易系统已经非常完善, 包括 TWAP、VWAP、IS、冰山订单等模式都可以在超高频模式下得到支持, 与主流资管系统供应商的接口也没有任何技术障碍, 未来, 在我们交易优化策略的支持下, 还会持续进行组合算法交易的研究。在 A 股市场应用算法交易, 我们认为, 当前已经具备了一定的市场土壤, 既然技术已经没有障碍, 接下去需要解决的无非是商务合作模式了。

程序化交易——股指期货带来发展契机

统计显示, 程序化交易在纽交所交易量中占比一直稳定在 30% 左右, 程序化交易是国际市场常用的交易方式。对于中国市场而言, 股指期货的推出为程序化交易的发展提供了契机。在商品期货领域, 基于技术分析的程序化交易策略远高于基于基本面分析的交易策略, 而以趋势交易和模式识别为代表的程序化交易策略是其典型的代表。套利策略并不是程序化交易在期货投资中的主要策略; 程序化交易的应用种类繁多, 对于股指期货业务的应用主要集中在趋势性投资策略和套利投资策略。

对于趋势性投资策略, 不论是否是日内交易, 其交易策略的建立主要依赖价格和成交量为基础的统计指标的组合, 当价格或者成交量达到组合指标的条件时, 形成交易指令, 并自动下单。目前用作投机交易的程序化交易策略主要有: 趋势交易、模式识别、反向交易、动量交易和均值回复等。对于套利投资策略, 程序交易的价值主要体现在: 高速交易获得基差和价差的短期时间窗, 实现组合下单策略, 快速判断交易品种的流动性, 调整多个交易品种的内部匹配, 实现交易目标。

我们对程序化交易在股指期货市场中的应用建议、海外市场规则差异及海外市场应用情况作了全面的研究, 并且对分钟级的模式识别也作了深入研究, 这些研究成果将在很短的时间内向市场发布。我们相信, 伴随股指期货的诞生, 程序化交易在 A 股的推行也是指日可待。

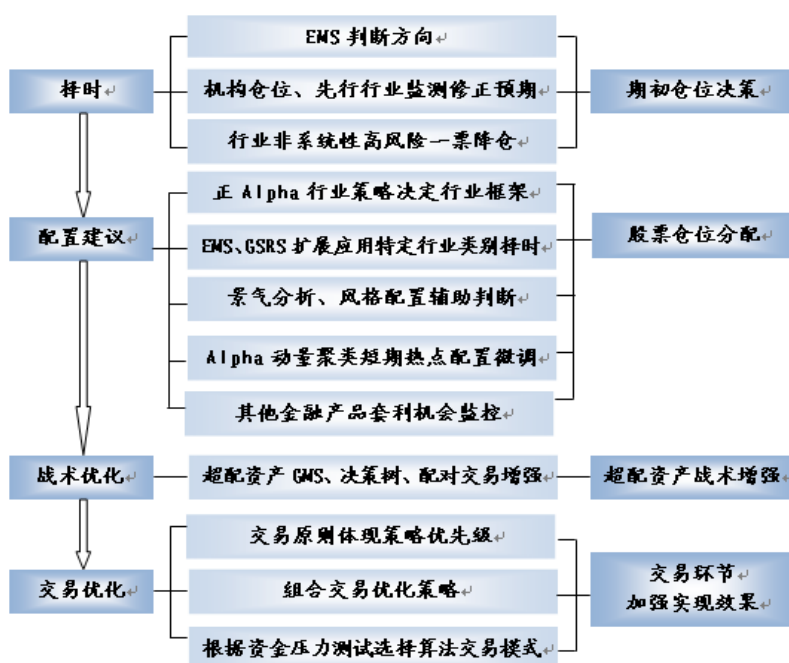
伴随对我们算法交易、程序化交易领域的研究覆盖达到一定程度, 目前我们的量化投资体系从策略开发到交易实施已经形成了一个完整的体系, 在这个体系的各个层面, 我们都拥有了自主开发的、经过实践检验的主打研究成果。未来一段时间内, 在各项研究持续推进的同时, 整合并发挥最大的优势, 将成为我们新的课题。

结语：实现盈利的征途——无限未来

作为在 A 股市场最先尝试量化投资实现体系的团队之一，我们的工作中包含了许多市场教育的成分，但是我们同时也感受到，A 股量化投资发展的速度是非常规的，因此市场的需求也在以一个非常规的速度变化着，因此，我们不能满足于 A 股量化投资实践的从无到有，而是希望能够推动这一事业从有到精的前行。

因此，我们的工作首先是完成一个独立而完整的量化投资决策体系，经过一年有余的努力，目前我们的量化投资策略至少已经覆盖了投资决策的各个层面，形成了一个能够独立运行的完整体系，也就是说，我们目前的决策体系从结构上已经可以支撑一个独立的产品在不借助外力的情况下运行，策略之间的复合应用也在不断地碰撞出新的结合点。我们将用虚拟量化基金的形式，公布我们这个体系的运行效果。

图 35：应用国信金工量化体系独立运行的投资过程



资料来源：国信证券经济研究所

一个没有策略储备的量化决策体系是脆弱的，随时都可能遭遇钝化和逆转，因此，我们对已有的策略不断地进行更新换代，并试图通过研究方法的多样性，达到策略和方法之间的复合，尽量分散单一策略对决策造成的影响力。目前策略的开发速度仍在加快，也就是在当前市场环境下，在多头策略领域内我们的开发能力还能够应对市场变化。伴随着 A 股量化投资应用的迅速崛起，量化策略的钝化速度也将加快；现有的量化策略结构，结合我们在中性策略自动化交易体系上所做的尝试，还将打开另一个更大的空间。

我们认为，未来量化投资的核心竞争力将体现在策略的持续开发和应用能力上，一个量化团队管理的要素中，也应该包含多角度分析能力的维护，和刺激团队追寻新思路的机制，对市场永远有好奇心的天才，也是我们愿意用任何代价追寻的宝贝。在实现收益的征途上，我们有无尽的动力，2009-2010，只是一个开始。

国信证券投资评级

类别	级别	定义
股票 投资评级	推荐	预计 6 个月内，股价表现优于市场指数 20%以上
	谨慎推荐	预计 6 个月内，股价表现优于市场指数 10%-20%之间
	中性	预计 6 个月内，股价表现介于市场指数±10%之间
	回避	预计 6 个月内，股价表现弱于市场指数 10%以上
行业 投资评级	推荐	预计 6 个月内，行业指数表现优于市场指数 10%以上
	谨慎推荐	预计 6 个月内，行业指数表现优于市场指数 5%-10%之间
	中性	预计 6 个月内，行业指数表现介于市场指数±5%之间
	回避	预计 6 个月内，行业指数表现弱于市场指数 5%以上

免责声明

本报告信息均来源于公开资料，我公司对这些信息的准确性和完整性不作任何保证。报告中的内容和意见仅供参考，并不构成对所述证券买卖的出价或询价。我公司及其雇员对使用本报告及其内容所引发的任何直接或间接损失概不负责。我公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行业务服务。本报告版权归国信证券所有，未经书面许可任何机构和个人不得以任何形式翻版、复制、刊登。

国信证券经济研究所研究团队

宏观		策略		交通运输	
周炳林	0755-82133339	赵 谦	021-60933153	郑 武	0755- 82130422
林松立	010-82254212	崔 嵘	021-60933159	陈建生	0755- 82130422
		廖 喆	021-60933162	岳 鑫	0755- 82130422
		黄学军	021-60933142	高 健	0755-82130678
银行		房地产		机械	
邱志承	021-68864597	方 焱	0755-82130648	余爱斌	0755-82133400
黄 飙	0755-82133476	区瑞明	0755-82130678	李筱筠	010-82254205
谈 煊	010- 82254212	黄道立	0755-82130833	黄海培	021-60933150
				陈 玲	0755-82133400
				杨 森	0755-82133343
汽车及零配件		钢铁		商业贸易	
李 君	021-60933156	郑 东	010-82254160	孙菲菲	0755-82133400
左 涛	021-60933164	秦 波	010-82252922	吴美玉	010-82252911
基础化工		医药		石油与石化	
邱 伟	0755-82133263	贺平鸽	0755-82133396	李 晨	021-60875160
张栋梁	0755-82130532	丁 丹	0755-82130678	严蓓娜	021-60933165
		陈 栋	021-60933147		
电力设备与新能源		传媒		有色金属	
皮家银	021-60933160	陈财茂	021-60933163	彭 波	0755-82133909
电力与公用事业		非银行金融		通信	
徐颖真	021-60875162	邵子钦	0755- 82130468	严 平	021-60875165
谢达成	021-60933161	田 良	0755-82130513	程 峰	021-60933167
		童成敦	0755-82130513		
造纸		家电		计算机	
李世新	0755-82130565	王念春	0755-82130407	凌 晨	021-60933157
电子元器件		纺织服装		农业	
		方军平	021-60933158	张 如	021-60933151
旅游		食品饮料		建材	
廖绪发	021-60875168	黄 茂	0755-82133476	杨 昕	021-60933168
刘智景	021-60933148	谢鸿鹤	0755-82130646		
煤炭		建筑		中小股票	
李 然	0755-82130681	邱 波	0755-82133390	陈 健	0755-82133476
苏绍许	021-60933144	李遵庆	0755-82133343	陈爱华	0755-82133397
				祝 彬	0755-82132518
				王一峰	010-82250828
				邵 达	0755- 82132098
固定收益		投资基金		量化投资	
李怀定	021-60933152	杨 涛	0755-82133339	葛新元	0755-82133332
张 旭	010-82254210	黄志文	0755-82133928	董艺婷	021-60933155
高 宇	0755-82133528	刘舒宇	0755-82131822	戴 军	021-60933166
蔺晓熠	021-60933146	彭怡萍		林晓明	021-60933154
侯慧娣	021-60875161			秦国文	0755-82133528
				程景佳	021- 60933166
				赵斯尘	021- 60875174
指数与产品设计					
焦 健	0755-82131822				
赵学昂	0755-82131822				
王军清	0755-82133297				
阳 瑾	0755-82131822				
周 琦	0755-82131822				
彭甘霖					

国信证券机构销售团队					
华南区		华东区		华北区	
万成水	0755-82133147 13923401205 wancs@guosen.com.cn	盛建平	021-60875169 15821778133 shengjp@guosen.com.cn	王立法	010-82252236 13910524551 wanglf@guosen.com.cn
邵燕芳	0755-82133148 13480668226 shaoyf@guosen.com.cn	马小丹	021-60875172 13801832154 maxd@guosen.com.cn	王晓建	010-82252615 13701099132 wangxj@guosen.com.cn
林 莉	0755- 82133197 13824397011 Linli2@guosen.com.cn	郑 毅	021-60875171 13795229060 zhengyi@guosen.com.cn	谭春元	010-82254209 13810118116 tancy@guosen.com.cn
王昊文	0755-82130818 18925287888 wanghaow@guosen.com.cn	黄胜蓝	021-60875173 13761873797 huangsl@guosen.com.cn	焦 骥	010-82254202 13601094018 jiaojian@guosen.com.cn
甘 墨	0755-82133456 15013851021 ganmo@guosen.com.cn	刘 塑	021-60875177 13817906789 liusu@guosen.com.cn	李锐	010-82254212 13691229417 lirui2@guosen.com.cn
		叶琳菲	021-60875178 13817758288 yelf@guosen.com.cn	徐文琪	010-82254210 13811271758 xuwq@guosen.com.cn
		孔华强	021-60875170 13681669123		