

大数据应用背景下的蓝筹 T+0 初探

策略累积收益曲线



基于集合竞价试盘者行为的 A 股 T+0 策略

报告摘要:

● 国内外 T+0 交易制度现状

就海外主要股票市场而言，T+0 是一项通行制度，主要的股票市场中，仅中国大陆未实行 T+0 交易制度。

通过台湾地区及韩国推出 T+0 交易制度前后市场的变化情况可以看出，T+0 交易制度的推出会极大地促进股市的成交量（成交额）。从短期（一个月）来说，有可能会加大股市的日内震荡；但是从超过一个月以上的的时间尺度上来看，日内震荡会趋于同 T+0 制度推出前一致的水平。同时，T+0 制度的推出一一般会作为股市的利好消息，推动股市上涨。

国内在上世纪 90 年代初也曾执行过 T+0 交易，但由于当时股市投机过度，T+0 交易制度被认为助长了市场炒作而被取消。目前，监管层正在研究并将适时推出 A 股 T+0 的交易制度。

● A 股 T+0 与期货 T+0 的本质区别

通过对比传统 CTA 策略的单笔交易平均收益率，我们指出由于 A 股存在较高的交易成本，A 股市场的 T+0 交易策略思路一定和期货市场是有所差别的。A 股 T+0 的核心在于选股——预测并选择当日有可能大幅波动的个股。基于此，我们在大数据应用背景下，开发了基于集合竞价试盘者行为的 A 股 T+0 交易策略。

● 基于集合竞价试盘者行为的 A 股 T+0 策略

开盘集合竞价前 5 分钟，由于交易制度的允许，可以进行任意挂撤单，从而存在很多“虚单”。我们希望通过在这些“虚单”中进行数据挖掘，寻找当日有可能向上大幅波动的个股。通过全市场两千多只股票数年的 tick 数据，在不进行期货对冲的情况下，我们仅仅通过做多个股（组合）就实现了绝对收益的获取，为投资者提供了全新的投资思路。作为 T+0 交易方案推出前的策略储备，我们从市场分析，到模型建立，到实证分析，到技术处理，给出一套完整的解决方案。

● 核心假设风险:

本篇报告测算数据均来自 A 股 T+1 市场。T+0 交易制度推广后，有可能会改变整个市场的微观结构，对本策略造成一定的影响，请投资者注意相关风险。

策略交易数据统计

年化收益率	19.72%
平均单笔收益率	0.34%
胜率	57.83%
盈亏比	1.04
最大回撤率（2010 年至今）	-14.24%

分析师： 安宁宁 S0260512020003



0755-23948352



ann@gf.com.cn

相关研究:

联系人： 张超



020-87555888-8646



zhangchao@gf.com.cn

目录索引

一、海外股票市场 T+0 交易制度.....	4
(一) 海外股票市场 T+0 交易情况.....	4
(二) T+0 到 T+1 对市场的影响.....	4
二、国内 T+0 交易回溯与展望.....	8
三、A 股 T+0 策略焦点：交易成本.....	8
四、基于试盘者行为的个股 T+0 交易策略.....	10
(一) 集合竞价中的试盘者行为.....	10
(二) 大数据下的试盘者行为挖掘.....	14
(三) 基于试盘者行为的 T+0 交易策略.....	15
(四) 融资买入，增厚收益.....	18
(五) 止盈止损控制系统性风险.....	20
(六) 大数据下的模型调试.....	21
五、总结.....	23

图表索引

图 1: 台湾加权指数日均成交额及日内震荡幅度 (1994 年 1 月前后)	5
图 2: 韩国 KOSPI 指数日均成交额及日内震荡幅度 (1997 年 1 月前后)	6
图 3: 台湾加权指数日均成交额及日内震荡幅度 (2014 年 1 月前后)	7
图 4: SMT 与 EMDT 策略历史累积收益率 (2%交易成本)	9
图 5: SMT 与 EMDT 策略历史累积收益率 (0.26%交易成本)	10
图 6: 示意图: 竞价涨停后小幅回落, 但开盘价仍接近涨停	12
图 7: 示意图: 竞价涨停后部分回落, 但开盘仍有小幅上涨	13
图 8: 示意图: 竞价涨停后大幅回落, 平开甚至低开	14
图 9: 试盘者交易策略累积收益曲线	15
图 10: 试盘者交易策略收益率分布	16
图 11: 被选中个股的市值分布	17
图 12: 被选中个股的行业分布	17
图 13: 试盘者交易策略累积收益曲线 (2 倍杠杆)	19
图 14: 不加止盈止损的累积收益曲线	21
表 1: 试盘者交易策略回测统计数据	15
表 2: 试盘者交易策略回测统计数据 (2 倍杠杆)	20

一、海外股票市场 T+0 交易制度

(一) 海外股票市场 T+0 交易情况

“T+0”交易，又称日内回转交易，是指投资者就同一个标的（如股票）在同一个交易日内各完成一次买进和卖出的行为。日内回转交易可以概括为三种情况：1、当日买进某股票后，再于当日卖出；2、当日卖出某股票后，再于当日买进；3、投资者原来持有某股票，于当日卖出后，再于当日买进。A股投资者习惯于把以上第一种情况称为“T+0”。

就海外主要股票市场而言，T+0 是一项通行制度，主要的海外市场都允许采用 T+0 回转交易，如纽约证券交易所、纳斯达克、伦敦证券交易所、东京证券交易所、香港证券交易所等。主要的股票市场中，仅中国大陆未实行 T+0 交易制度。

另一方面，各个股票市场根据自己市场环境的不同，对 T+0 制度的政策有不少差别，总体来说，各种监管性政策都是为了控制风险，防范过度投机和维护市场的基本平稳。

先看美国股市，美国纽交所和纳斯达克市场中，T+0 通常只能在信用账户内进行，现金账户的所有者不被定义为典型回转交易者。所谓典型回转交易者的判断标准，是在连续 5 个交易日内，回转交易达到 4 次及 4 次以上。而且美国证监会要求典型回转交易者的账户最低净值为 25000 美元。对于典型回转交易者，规定每天的交易规模，也限于其上一个交易日所持股票市值与维持保证金的 4 倍。假设该客户所进行的是平进平出的交易，那么也就意味着他在一个交易日内理论上只能有两轮完整的回转交易。

台湾地区股市从 1994 年开始实施 T+0 制度，但是仅限于信用账户。直到 2014 年 1 月，才开放现金账户的 T+0 制度。配合现金账户 T+0 制度的推出，台交所出台了一系列投资者适当性准则和风控措施。如要求投资人必须开户满 3 个月，最近一年内委托买卖成交达十笔以上，并且在同证券经纪商签订授权同意书和风险告知书之后，才可从事回转交易。同时，证券经纪商根据相关规则和投资者账户资金情况设定投资者单日回转交易买卖额度，并进行严格前端控制。而且，证券经纪商应于每日收盘后，对投资人当日回转的损益进行评估，从而增减其单日买卖额度或回转交易额度。证券经纪商发现投资人前一个交易日回转交易累计亏损达到当日买卖额度的 50% 时，应暂停其回转交易。此外，台交所对标的股票采取试点先行的策略，首批恢复现金账户回转交易的上市股票共 150 只，约占台交所上市股票总数的 10%，占总市值的 75%。根据 2014 年 1 月 6 日的统计数据，允许回转交易的股票大多属于排名前 25% 大盘股，其平均市值为 1339 亿新台币，而台交所上市股票平均市值仅为 153 亿新台币。

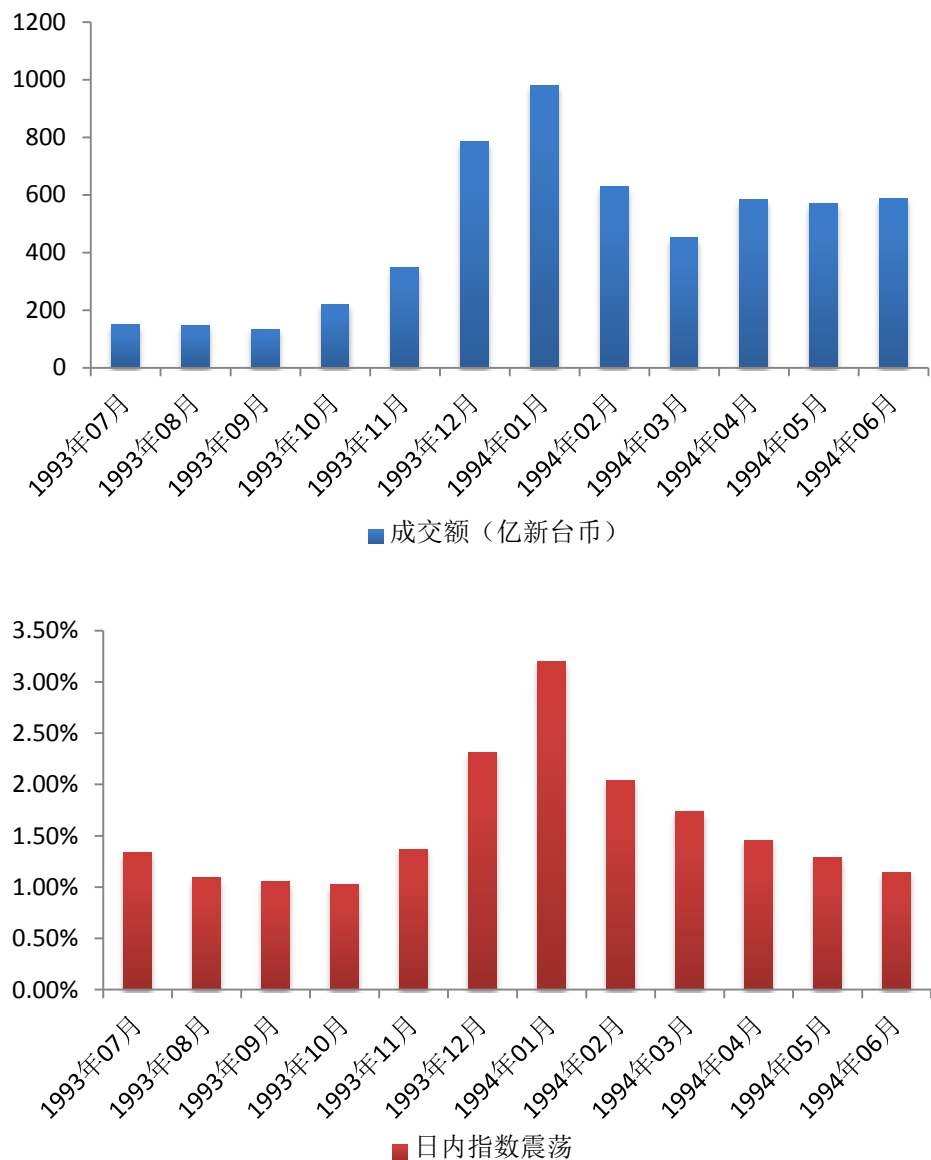
此外，日本的东京交易所对 T+0 交易的规定是一笔资金一天内只能做一次回转交易。而香港联交所是实行不计次数，不限身份，不管品种的“三不”原则，被公认为是实行 T+0 回转交易最彻底的市场。

(二) T+0 到 T+1 对市场的影响

从市场影响来说，一般认为，相比 T+1 交易制度，T+0 会更有可能增加市场交易量，并加剧市场波动。本节我们从三个海外市场 T+0 制度实行的案例来验证这方面的影响。

首先是 1994 年 1 月 1 日，台湾地区股市实施资券相抵的信用账户 T+0 制度。交易制度改革前后各六个月的市場成交额和日内震荡幅度如图 1 所示。注意到台湾股市从 1993 年 10 月份开始开启了一波大牛行情，台湾加权指数从 10 月初的 3800 点上涨到 1994 年 1 月底的超过 6000 点，然后震荡上升，最终在 1994 年 8 月份超过 7000 点。可以看到，期间股市成交量得到了急剧放大，从 1993 年年中的日均成交额 100 多亿新台币变化到 1994 年年中的日均成交额近 600 亿新台币。这一方面和大牛市有关，另一方面和 T+0 制度的推出不无关联。对应的，加权指数的日内震荡幅度在 1994 年 1 月份达到了高峰，不过此后回落，在 6 月份时候和 1 年前就基本持平。

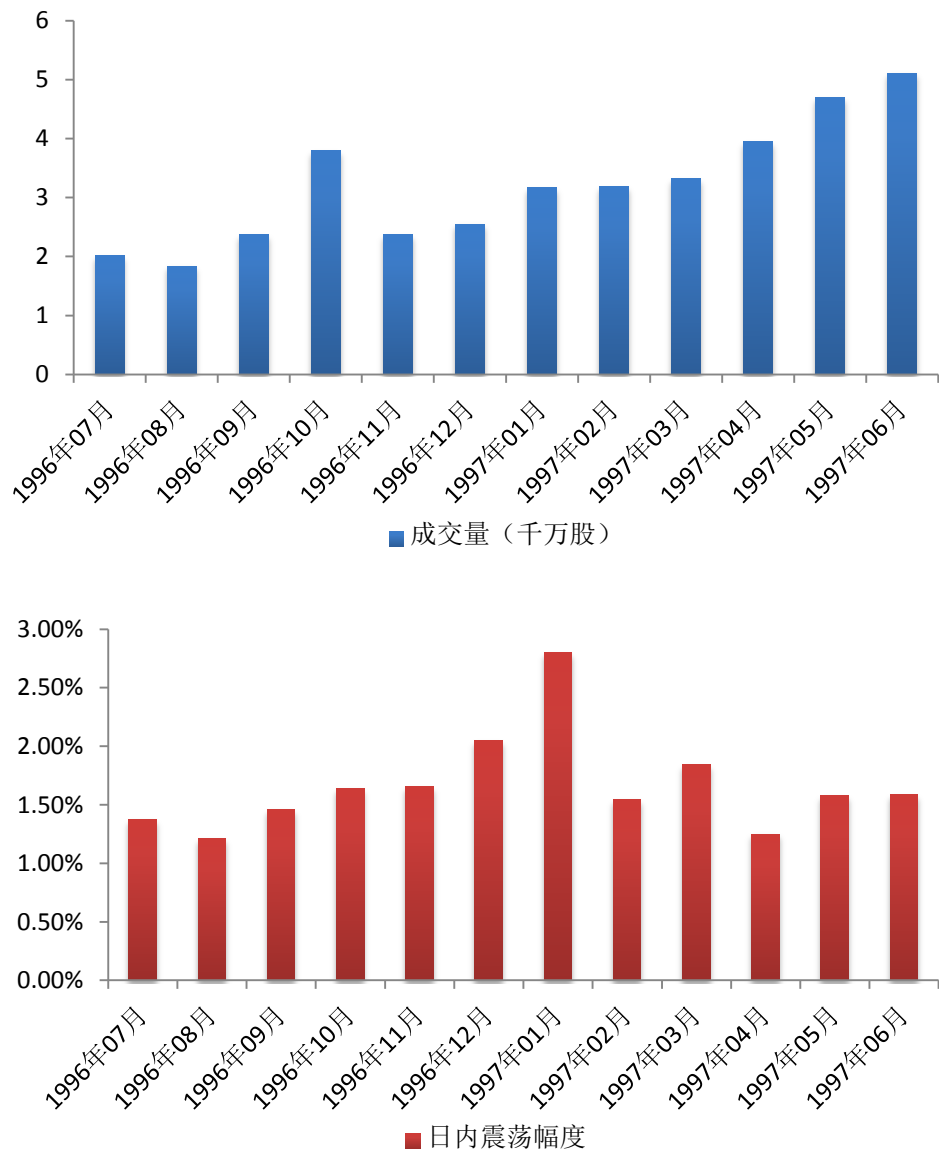
图1: 台湾加权指数日均成交额及日内震荡幅度 (1994年1月前后)



数据来源：广发证券发展研究中心，Wind 资讯

第二个案例是1997年1月初，韩国股市实施T+0制度。交易制度改革前后各六个月的市場成交量和日内震荡幅度如图2所示。同样可以看到，期间股市成交量得到了明显放大，从1996年下半年的日均成交量2000多万股变化到1997年上半年的日均成交量近4000万股。对应的，KOSPI指数的日内震荡幅度在1997年1月份达到了高峰，不过此后震荡幅度基本上和T+0实施前没有区别。另外，KOSPI指数在T+0制度实施的半年内，从1997年1月初的650点上涨为6月底的745点。

图2：韩国KOSPI指数日均成交额及日内震荡幅度（1997年1月前后）

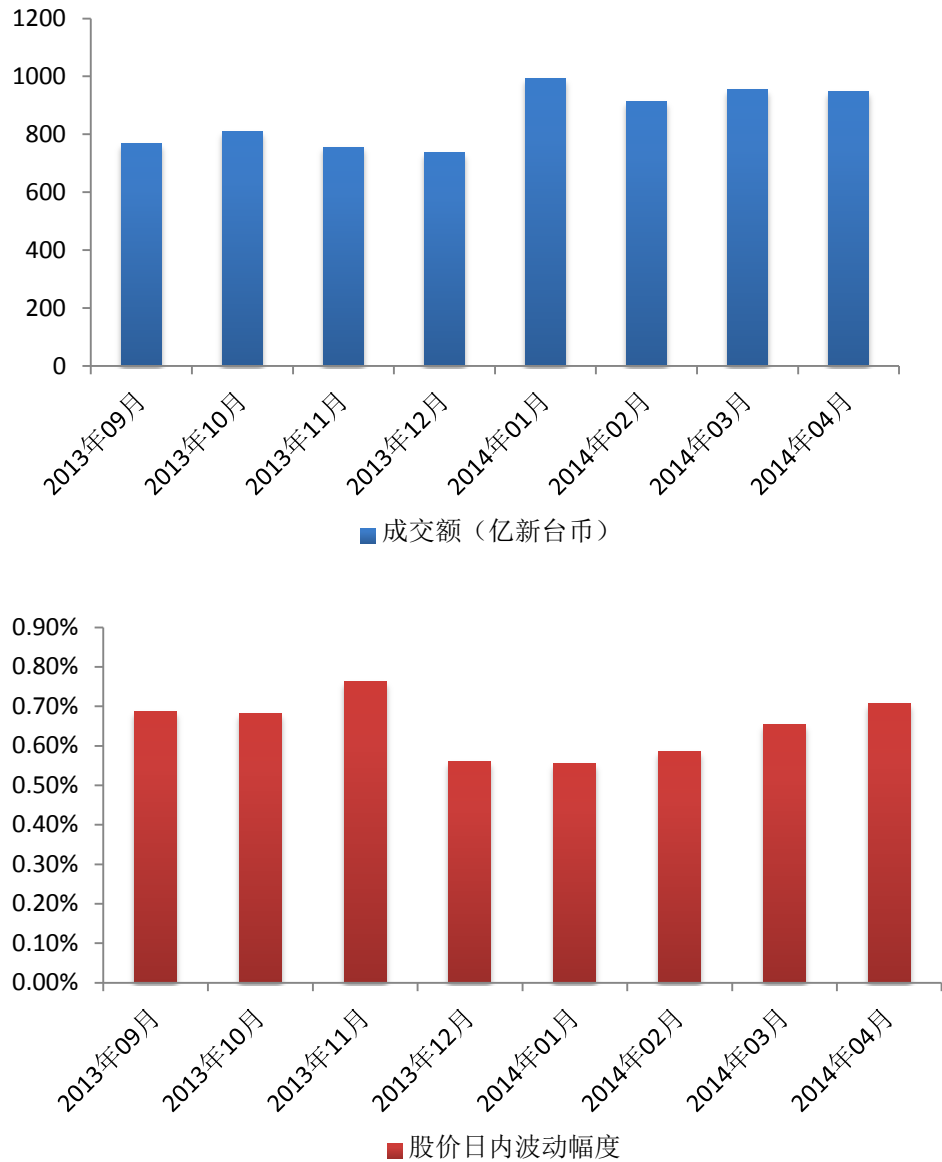


数据来源：广发证券发展研究中心，Bloomberg

第三个案例是最近发生的，还是在台湾地区股市。2014年1月6日，台湾证券交易所恢复了现金帐户下的回转交易。交易制度改革前后各四个月的市場成交量和日内震荡幅度如图3所示。同样可以看到，期间股市成交量得到了明显放大，从2013年后4个月的日均成交额约760亿新台币变化到2014年前四个月的日均成交额达到约950亿新台币。对应的，加权指数的日内震荡幅度在2014年1月前后看不到由T+0

推出带来的明显变化。另外，台湾加权指数从2014年1月6日的8500点上涨到4月底的8791点，期间曾突破9000点，总体而言略有上涨。

图3: 台湾加权指数日均成交额及日内震荡幅度 (2014年1月前后)



数据来源: 广发证券发展研究中心, Wind 资讯

由以上三个案例可以看出, T+0 交易制度的推出会极大地促进股市的成交量(成交额)。从短期(一个月)来说, 有可能会加大股市的日内震荡; 但是从超过一个月以上的时间尺度上来看, 日内震荡会趋于同 T+0 制度推出前一致的水平。换句话说, T+0 制度的推出并不会从加剧股市中长期的日内震荡。同时, T+0 制度的推出一般会作为股市的利好消息, 推动股市上涨。

二、国内 T+0 交易回溯与展望

自从 1990 年上海证券交易所和深圳证券交易所成立以来，经过多年的发展，我国股票市场已经取得了长足的发展。近年来，A 股市场对 T+0 交易制度的呼声见涨，这方面的话题已经成为股票市场各方关注的焦点。

事实上，A 股市场曾允许执行过回转交易，上海交易所在 1992 年 12 月 24 日至 1994 年 12 月 31 日，深圳交易所在 1993 年 11 月 22 日至 1994 年 12 月 31 日期间，都曾执行过 T+0 回转交易。但由于当时股市投机过度，T+0 交易制度被认为助长了市场炒作而被取消。继而，1999 年实施的《证券法》规定“当日买入的股票，不得当日卖出”，严格禁止了股票 T+0 交易。

与 A 股市场相对应的是，沪深 300 股指期货在 2010 年 4 月 16 日上市交易，可以进行 T+0 交易，并允许做空。根据上海证券交易所关于《上海证券交易所交易规则》的最新修订（2013 年 10 月），上交所场内交易的货币 ETF、债券 ETF、黄金 ETF 自 2013 年 12 月 9 日正式开始实行 T+0 交易制度。股票类 ETF 在一二级市场之间也可以跨市场进行 T+0 回转交易。

总的来说，经过过去十多年的改革与治理，中国证券市场已发生了较大变化，机构投资者规模有所壮大，逐渐成为市场的中坚力量；ETF、股指期货、融资融券业务的推出和逐渐成熟也为 A 股市场的 T+0 交易打好了基础。而且，2005 年《证券法》修订，取消了有关“当日买入的股票，不得当日卖出”的规定，A 股推行 T+0 交易制度已经不存在法律障碍。

目前，监管层正在研究 A 股的 T+0 的交易制度。根据上交所资本市场研究所 2014 年 5 月最新发布的《沪市上市公司 2013 年年报整体分析报告》，交易所将继续研究推进大盘蓝筹股“T+0”交易等差异化交易机制。从监管层提倡蓝筹股价值投资的角度来看，未来很有可能会采取类似台交所的政策，选取蓝筹股作为 T+0 制度的初步尝试标的。

三、A 股 T+0 策略焦点：交易成本

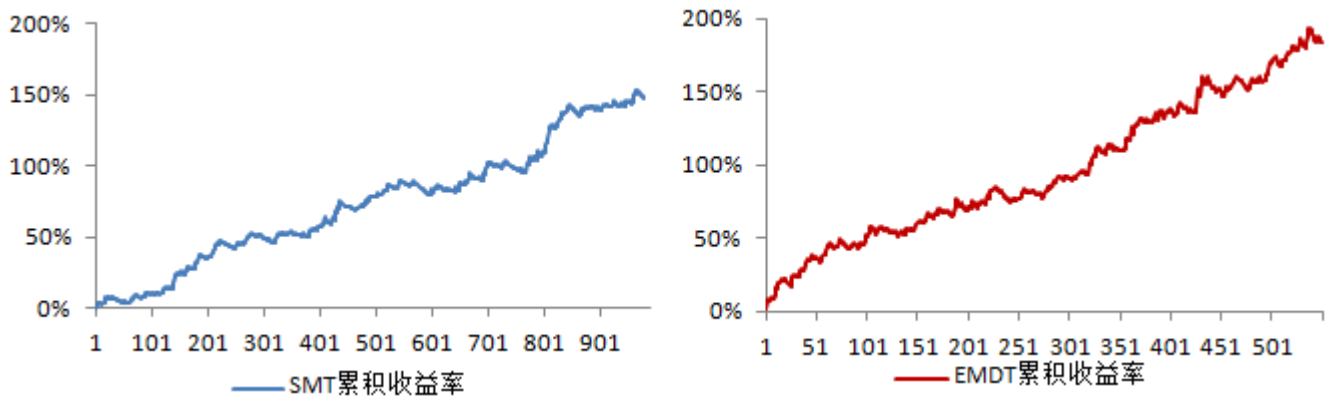
一旦 A 股 T+0 交易制度得以实施，会出现大量相关的交易策略。对于这些交易策略影响最大的因素是什么？我们认为是交易成本。

我们之前在股指期货交易策略系列研究中开发了大量的日内交易策略，我们来看一下它们对交易成本的敏感程度。

以 SMT 策略和 EMDT 策略为例，如果没有关注过这两个策略的投资者可以暂且不用关心这两个策略的算法和逻辑，我们仅仅来看一下这两个策略的历史表现。

SMT 交易策略的交易频率是每天交易一次（即多空一个来回），EMDT 策略的交易频率大约为每两天交易一次，在双边 2% 的股指期货交易成本下，两个策略从 2010 年 4 月股指期货上市到 2014 年 5 月的累积收益率（复利计算）分别约为 148% 和 183%，平均单笔交易的收益率分别为 9.4% 和 19.5%。在 2% 的双边交易成本下，两个策略的累积收益表现如图 4 所示。

图4：SMT与EMDT策略历史累积收益率（2%交易成本）



数据来源：广发证券发展研究中心，天软科技

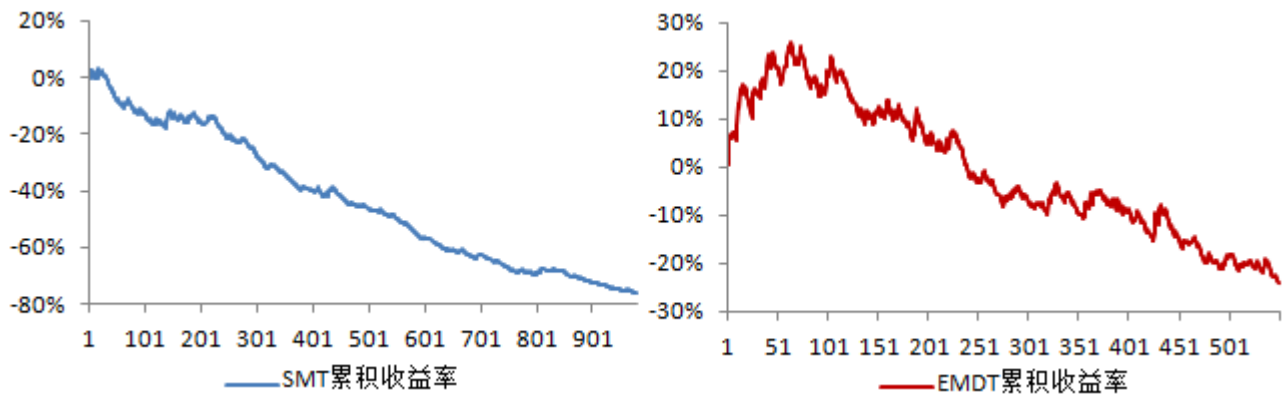
但是在 A 股市场，这样的平均单笔收益率并无法承担其相应的交易成本。A 股的交易成本主要包括佣金、印花税和冲击成本。

根据中国证监会、原国家计划和发展委员会（现为国家发展和改革委员会）、国家税务总局联合发出的《关于调整证券交易佣金收取标准的通知》，从 2002 年 5 月 1 日开始，A 股、B 股、证券投资基金的交易佣金实行最高上限向下浮动制度。证券公司向客户收取的佣金（包括代收的证券交易监管费和证券交易所手续费等）不得高于证券交易金额的 0.3%，也不得低于代收的证券交易监管费和证券交易所手续费等。

我国证券交易的印花税税率标准曾多次调整。根据相关规定，财政部、国家税务总局从 2008 年 4 月 24 日起，调整证券（股票）交易印花税税率为 0.1%。即对买卖、继承、赠与所书立的 A 股、B 股股权转让书据，由立据双方当事人分别按 0.1% 的税率缴纳证券（股票）交易印花税；2008 年 9 月 19 日，证券交易印花税调整为只对出让方按 0.1% 税率征收，对受让方不再征收。

假设不计冲击成本，那么目前 A 股市场的股票交易成本应该是 0.1% 印花税加上双边交易佣金。虽然目前市场上已出现少量互联网低佣金开户业务，但对于机构投资者而言，普遍交易佣金仍然在单边 8%% 水平。也就是说，印花税和佣金作为双边交易成本已达到 0.26%。对比这一数值，之前两个策略平均单笔 9.4% 和 19.5% 的收益率（已扣掉 2%%）水平显得无力适从。更为直观地，我们画出从 SMT 和 EMDT 策略在 0.26% 双边交易成本下，从 2010 年 4 月到 2014 年 5 月的累积收益率（复利计算），如图 5 所示。

图5: SMT与EMDT策略历史累积收益率(0.26%交易成本)



数据来源: 广发证券发展研究中心, 天软科技

在较高交易成本的压制下, 上述两个表现优异的 CTA 策略在 A 股市场中出现大幅亏损。

因此我们这里要强调, A 股市场的 T+0 交易策略思路一定和期货市场交易策略的思路是有差别的, 生搬硬套只能带来负面效果。如果是大量股票构建的组合, 其波动特征将与指数类似。所以我们认为, 个股 T+0 策略绝不适用于所有股票。

熟悉 CTA 中低频趋势交易的投资者都知道, 指数波动越大, 赚钱的机会越多, 统计上的单笔交易收益率也会越高。我们这里的出发点类似——选取波动率大的个股作为 T+0 交易的标的证券, 期望单笔交易收益率从统计上能够覆盖掉股票的交易成本。

但是如何找到日内高波动率的个股呢? 这将成为个股 T+0 交易策略的核心。我们的第一个思路是按照类似于指数波动率聚集特征去寻找, 但是这一方法无法确定股价的波动区间, 或者说较难对开仓点位做出判断。因此, 我们在这篇报告中, 利用 A 股市场集合竞价规则, 进行数据挖掘, 根据集合竞价中试盘者的行为预测和选择当天股价有可能大幅拉升的个股, 并通过构建小样本组合进行个股 T+0 交易策略的建模和回测。

四、基于试盘者行为的个股 T+0 交易策略

(一) 集合竞价中的试盘者行为

1、集合竞价规则

目前, 我国证券交易所采用两种竞价方式: 集合竞价方式和连续竞价方式。

所谓集合竞价, 是指对在规定的时间内接受的买卖申报一次性集中撮合的竞价方式。根据我国证券交易所的相关规定, 集合竞价确定成交价的原则为:

(1) 可实现最大成交量的价格。

(2) 高于该价格的买入申报与低于该价格的卖出申报全部成交的价格。

(3) 与该价格相同的买方或卖方至少有一方全部成交的价格。

如果有两个以上申报价格符合上述条件的，深圳证券交易所取距前收盘价最近的价位为成交价；上海证券交易所则规定使未成交量最小的申报价格为成交价格，若仍有两个以上使未成交量最小的申报价格符合上述条件的，其中间价为成交价格。

集合竞价的所有交易以同一价格成交。

然后进行集中撮合处理。所有买方有效委托按委托限价由高到低的顺序排列，限价相同者按照进入交易系统电脑主机的时间先后排列。所有卖方有效委托按照委托限价由低到高的顺序排列，限价相同者也按照进入交易系统电脑主机的时间先后排列。依序逐笔将排在前面的买方委托与卖方委托配对成交。也就是说，按照价格优先、同等价格下时间优先的成交顺序依次成交，直到成交条件不满足为止，即所有买入委托的限价均低于卖出委托的限价，所有成交都以同一成交价成交。集合竞价中未能成交的委托，自动进入连续竞价。

每个交易日 9:15 至 9:25 为开盘集合竞价时间；另外，深交所股票在 14:57 至 15:00 还存在一个收盘集合竞价过程。本篇报告中，我们主要通过两个交易所开盘集合竞价中的一些试盘行为进行选股。

2、集合竞价中的试盘行为

根据交易所规定，开盘集合竞价的订单收集阶段分为两个阶段，其中 9:15 至 9:20 阶段允许撤销已经提交的订单；9:20 至 9:25 阶段不允许撤销已经提交的订单。除此之外，在 9:25 至 9:30 分期间，交易系统不接受任何申报，在这段时间内提交至交易所的订单，将被当作废单处理。

也就是说，在每个交易日 9:15 至 9:20 这段时间内，可以进行任意委托及撤单，这就会导致市场存在部分“虚单”。从直观市场行为上来看，也就是我们每个交易日都可以看得到的一些集合竞价挂撤单试盘现象。

集合竞价阶段往往隐含着主力资金当日运作意图的部分信息。因此，投资者认真、细致地分析集合竞价情况，有助于敏锐发现并能抓住集合竞价中出现的某些稍纵即逝的机会。特别是对于控盘度高的股票，我们会发现它们在开盘集合竞价前五分钟，也就是 9:15 至 9:20 阶段会偶尔出现被拉涨停或者被砸跌停，而后冲板资金又撤单、价格出现回归的情况。

由于我们希望寻找的是当日有可能大幅拉升波动的股票，因此我们尝试捕捉上述集合竞价期间出现过“假涨停”的个股。

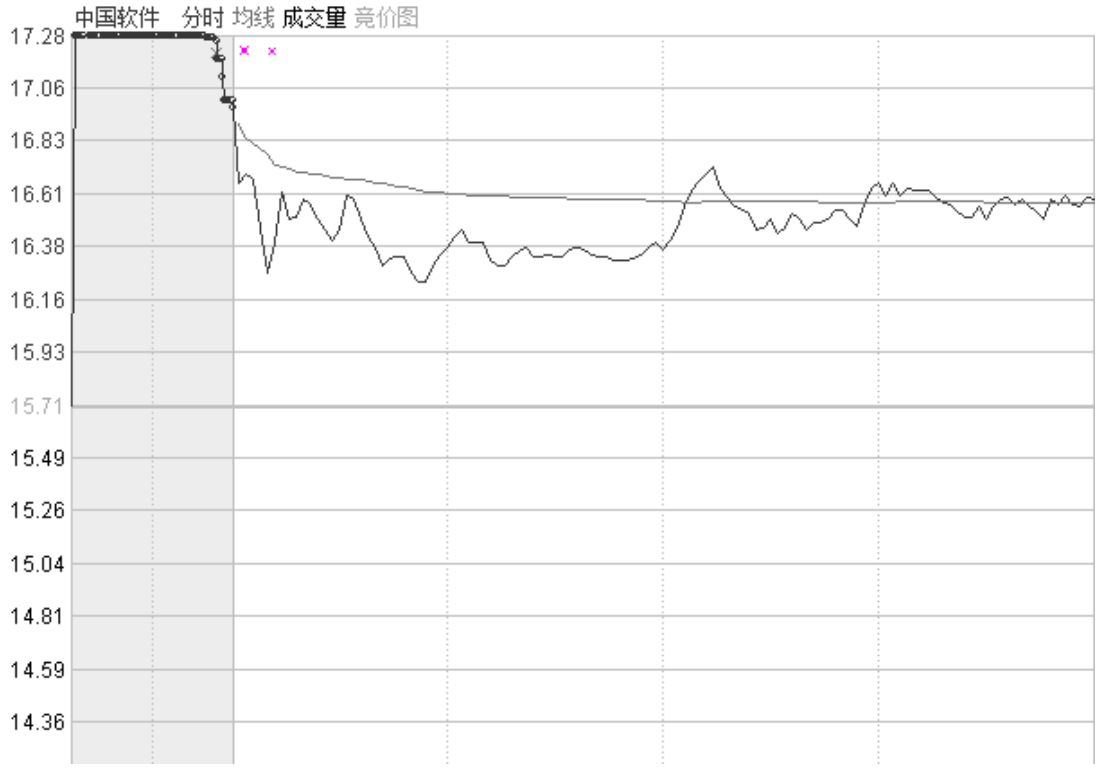
集合竞价冲击涨停又出现价格回落的个股，通过价格回落程度不同又可以分为三类。

(1) 竞价涨停后小幅回落，但开盘价仍接近涨停。

这一类股票大多是存在消息面或政策面利好，但利好强度又不足以巨量封住涨停。在开盘集合竞价阶段，出现一定数量的资金将价格打至涨停；但此时卖方又出现一定程度的出货。由此，集合竞价阶段未能封住涨停，但以大幅上涨开盘。这种情况的开盘会有两种盘面出现，一种是开盘后继续强势封住涨停，另一种是抛压较大股价回落。对于前一种情况，T+0 交易的盈利空间已非常小，而对于后一种情况，

日内可能出现亏损的幅度较大。因此出现这种情况时，我们不建议买入。

图6: 示意图: 竞价涨停后小幅回落, 但开盘价仍接近涨停

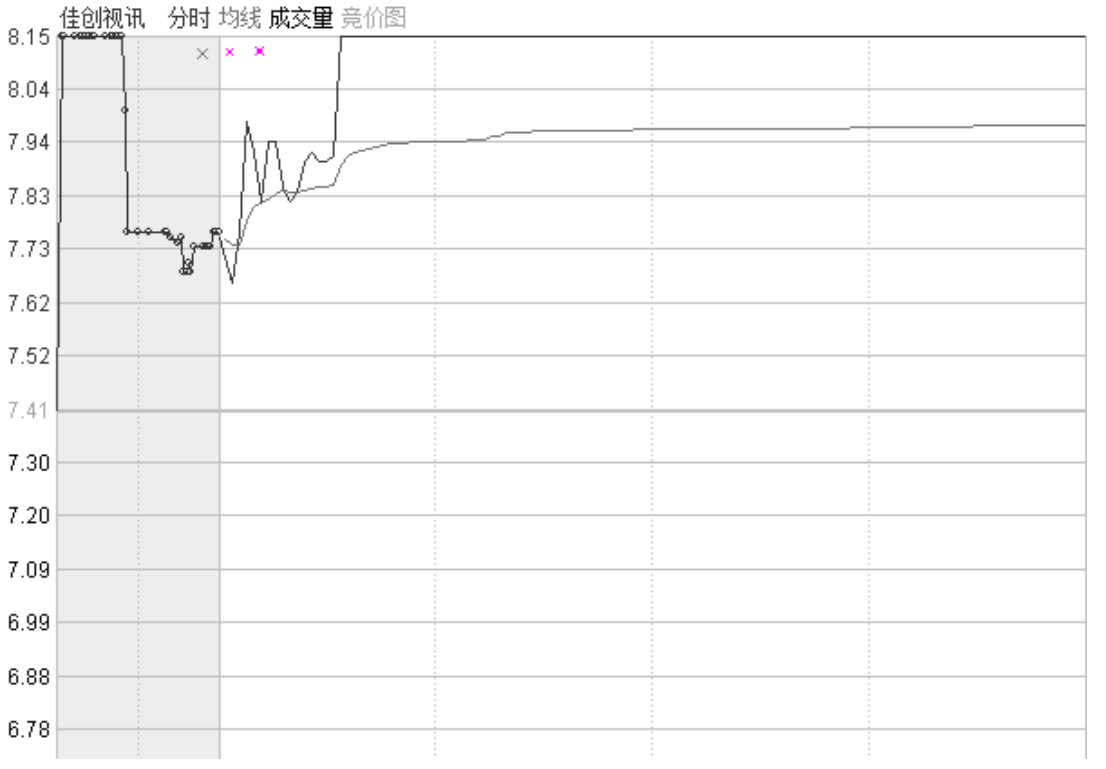


数据来源: 通达信

(2) 竞价涨停后部分回落, 但开盘仍有小幅上涨。

出现这种情况的股票不一定要有利好伴随, 但这一类股票往往是主力资金控盘度较高的个股。其 9:20 前挂出涨停价试盘, 主要是为了试探卖盘抛压的大小。如果抛压较大, 一般试盘的委托单撤单后开盘价都不会太高, 对应下面要讲的第三种情况; 而如果抛压较小, 又有一定的跟风盘出现, 在主力资金撤单后仍然会以小幅上涨开盘。此时主力资金很有可能低位强势拉升, 吸引更多跟风盘进入, 这就为 T+0 波段交易提供了良好机会。

图7: 示意图: 竞价涨停后部分回落, 但开盘仍有小幅上涨

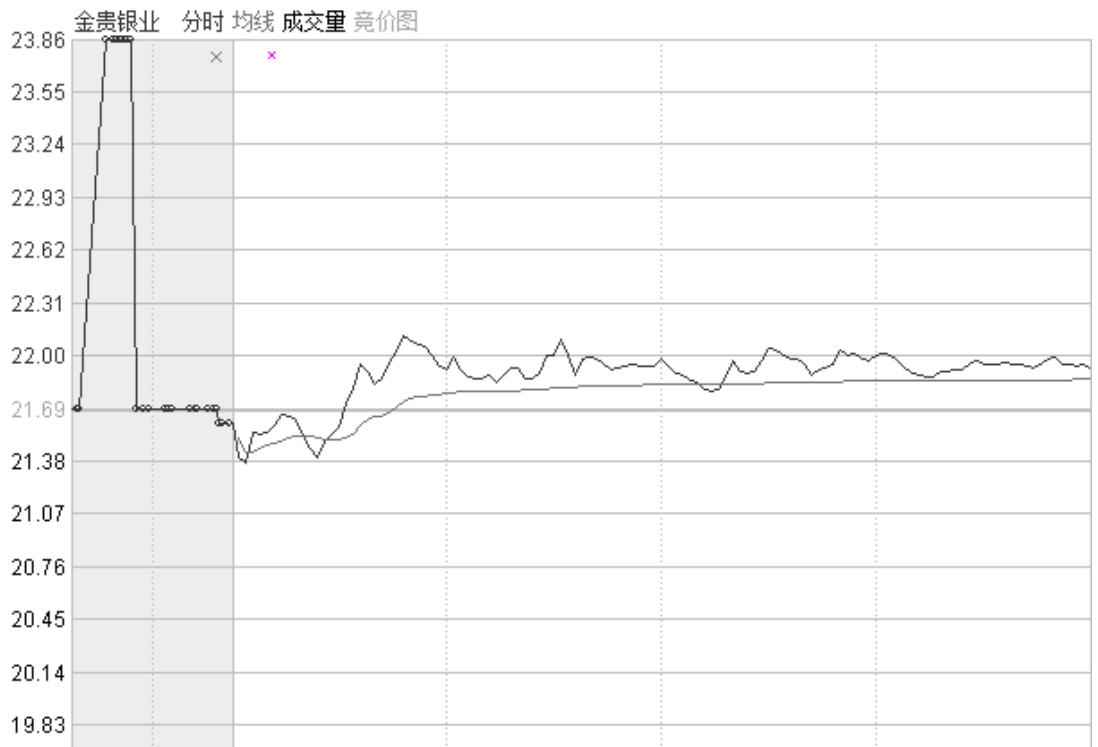


数据来源: 通达信

(3) 竞价涨停后大幅回落, 平开甚至低开。

集合竞价阶段打涨停后, 开盘价大幅回落, 这种情况的个股有几种可能: 一是上面所说的主力资金试盘, 抛压大跟风少, 决定观望; 二是主力资金定期监控市场抛压, 并无抬升股价之意; 三是主力资金只是为了吸引一下市场注意力, 也没有抬升股价的打算, 但至于为什么要吸引市场对个股的注意力并不好归因。但无论是由于何种因素, 这种竞价涨停后大幅回落, 平开甚至低开的个股往往在当日都表现平平, 并不是 T+0 交易的理想标的。

图8: 示意图: 竞价涨停后大幅回落, 平开甚至低开



数据来源: 通达信

综合上述三类开盘集合竞价出现过涨停, 而开盘未涨停情况的分析, 我们认为开盘小幅上涨的第(2)种情况在当日最有可能出现股价大幅拉升, 有机会可以为T+0交易创造良好条件。因此, 我们将通过定量的方法, 测算该类股票T+0交易的实证效果。

(二) 大数据下的试盘者行为挖掘

上述个股T+0交易策略总结起来很简单: 开盘买入集合竞价中涨停、回落, 并小幅高开的个股组合, 日内出货平仓了结。

但是这样一个简单的策略, 在回测过程中并不轻松, 原因在于回测数据量太大——要用到全市场两千多只股票数年的tick数据(特别是集合竞价阶段, 每一次价格变化将形成一笔tick数据)。我们要感谢大数据时代的来临, 为我们提供了更多的技术工具去解决这一问题。

最早提出大数据时代到来的是麦肯锡: “数据, 已经渗透到当今每一个行业和业务职能领域, 成为重要的生产因素。人们对于海量数据的挖掘和运用, 预示着新一波生产率增长和消费者盈余浪潮的到来。”

业界(最早为IBM)定义的大数据为: 数据量大(Volume)、数据种类多样(Variety)、要求实时性强(Velocity)、商业价值高(Value), 即大数据的4V特性。符合这些特性的, 叫大数据。

本篇报告介绍的个股T+0策略, 在建模过程中所用到的数据体量大、策略实时

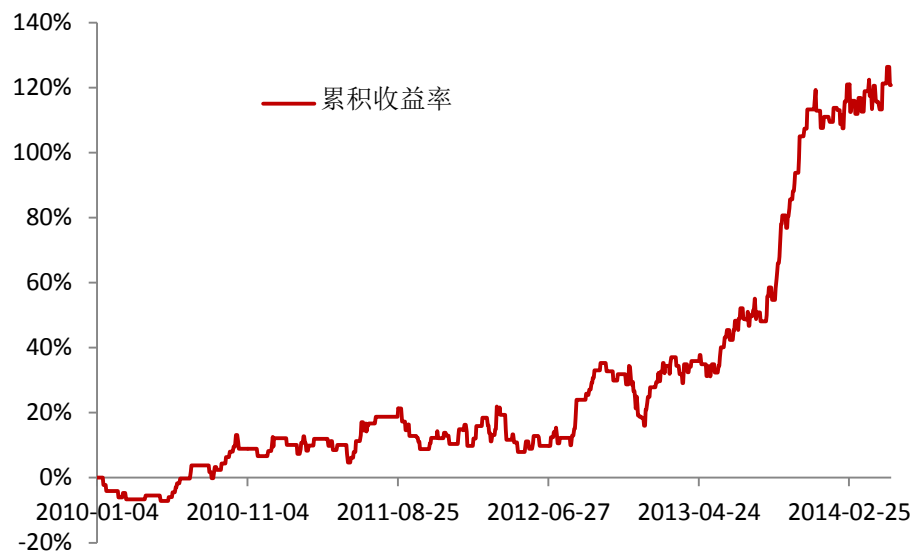
性要求高、能在二级市场带来丰厚收益，基本符合大数据背景，特别是在参数优化这一过程，更体现出了显著的大数据特征（后文会详细描述）。

（三）基于试盘者行为的 T+0 交易策略

我们定量实证的策略是：集合竞价触及过涨停板、并且开盘价涨幅在 2%至 6%之间的股票，按照资金等权配在当日要购买的股票组合中，并且在实证中假设全部以开盘价买入。盘中设置 1.5%的止盈位和-1.5%的止损位。如果触及止盈止损，卖出股票后当天不再操作；如果没有触及止盈止损，就在尾盘按收盘价把所有股票卖掉。

回测时间从 2010 年 1 月 4 日开始，至 2014 年 5 月 16 日结束。股票交易成本按双边 0.3%计算。回测累计收益曲线（按复利计算）如图 9 所示。

图9：试盘者交易策略累积收益曲线



数据来源：广发证券发展研究中心，天软科技

相关交易数据如表 1 所示。在 1056 个交易日中总共有 249 个交易日发生交易，也可以看出在较为严格的开盘涨幅限制下，发生交易的次数并不是特别多，大约每 4 个交易日才会进行 1 次交易。但是平均单笔收益率和刚才 CTA 策略的平均单笔收益率相比有了很大的提升，扣掉 0.3%交易成本后平均还可以保留 0.34%的单笔收益率。

表 1：试盘者交易策略回测统计数据

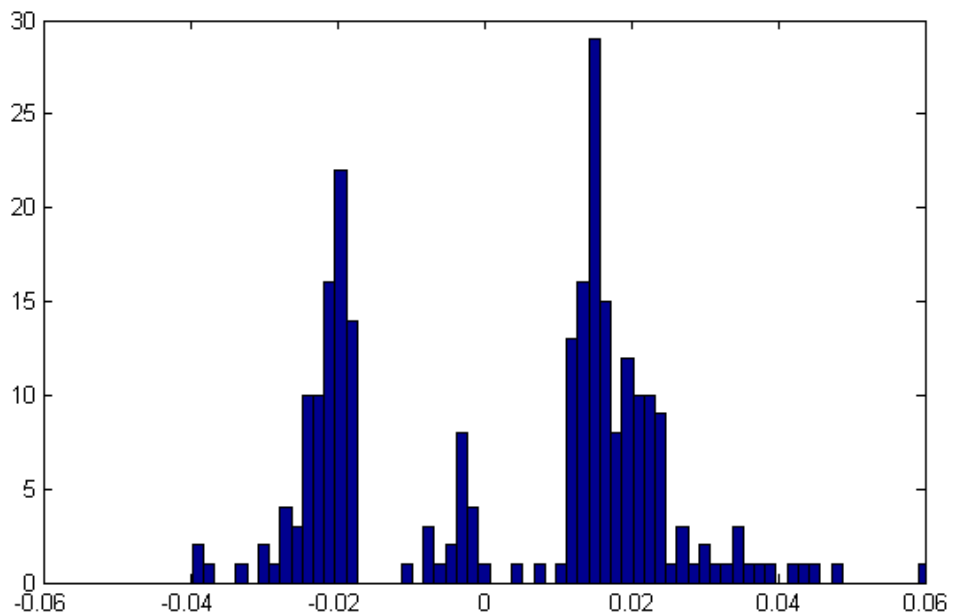
回测时间	2010 年 1 月 1 日至 2014 年 5 月 16 日
回测天数	1056
交易总次数	249
累积收益率	120.79%
年化收益率	19.72%
平均单笔收益率	0.34%
获胜次数	144

失败次数	105
胜率	57.83%
单次获胜收益率	1.95%
单次失败亏损率	-1.87%
盈亏比	1.04
最大回撤	-14.24%
最大连胜次数	12
最大连亏次数	5
单日最大盈利率	6.07%
单日最大亏损率	-3.84%

数据来源：广发证券发展研究中心，天软科技

由于受到计算速度的约束，我们通过 tick 数据监控开盘集合竞价情况，而在开盘后采用 1 分钟收盘价对策略进行测算，因此导致平均单笔盈利和平均单笔亏损的收益率都在一定程度上超出了 $\pm 1.50\%$ 的范围。所有有发生交易的收益率分布如图 10 所示。

图10：试盘者交易策略收益率分布



数据来源：广发证券发展研究中心

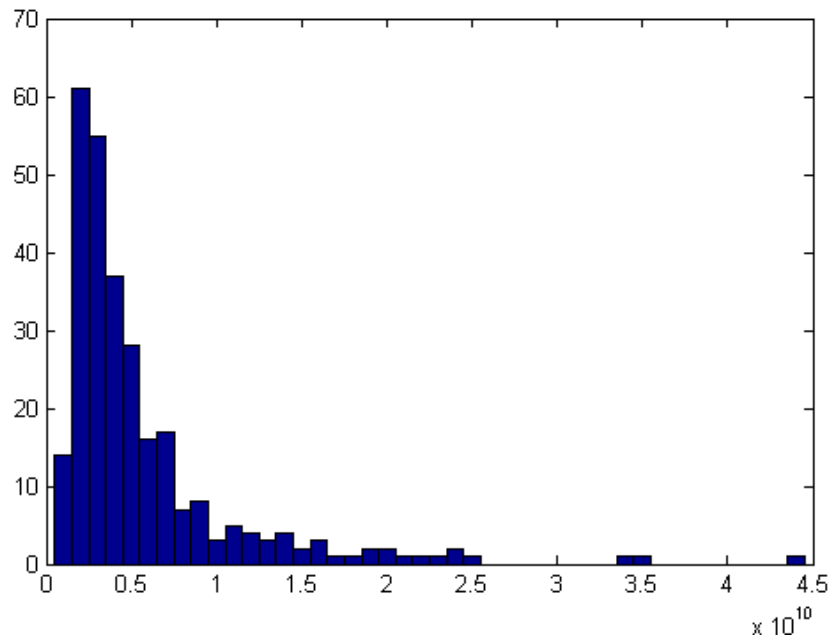
从图 10 的收益率分布可以看出，大多数股票在交易过程中都出现了止盈或者止损，这也验证了我们确实按照起初的设想，抓到了日内可能出现较大波动的个股，并获得了绝对收益。

另外，我们对四年多时间中策略选出的个股比较关心。

在四年多的时间中，一共有 249 个交易日发生交易。在这 249 个交易日中，前后总共选出过 302 只*次股票。去掉重复出现过的股票，共计 282 只股票。因此其实每次交易的组合里也就只有一两只股票而已。

这 282 只股票的市值分布如图 11 所示。

图11: 被选中个股的市值分布

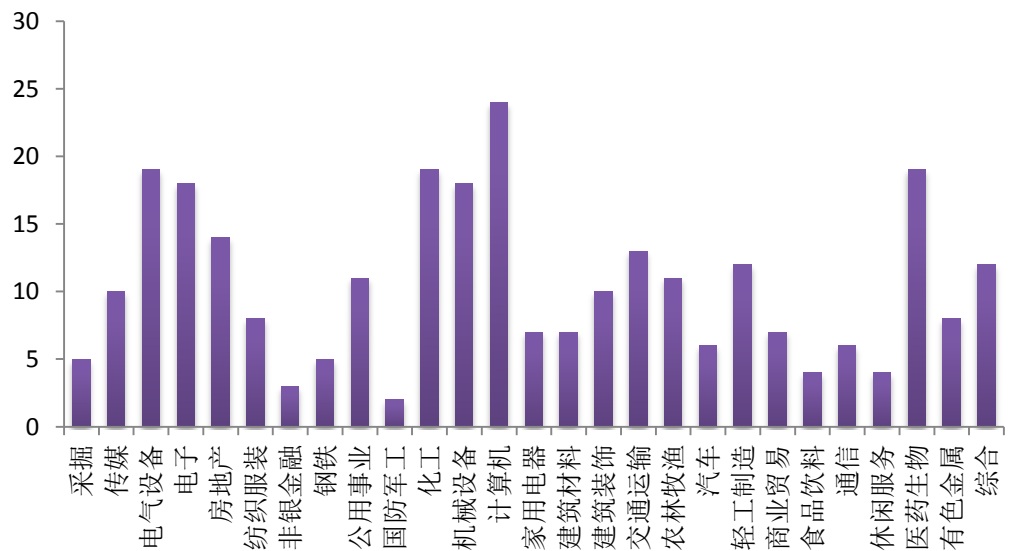


数据来源: 广发证券发展研究中心, wind 资讯

可以看出, 大多数股票为市值处于 100 亿元以下的中小市值股票, 这也是它们更易被主力资金控盘的内在原因。

这 282 只股票的申万一级行业分布情况如图 12 所示。

图12: 被选中个股的行业分布



数据来源: 广发证券发展研究中心, wind 资讯

其中被选中次数较多的行业包括计算机、医药生物、电气设备、电子、化工和机械设备。

(四) 融资买入，增厚收益

从图 9 的累积收益曲线来看，在单纯股票做多的情况下，策略的表现稳定性一直较好。因此，我们试图通过融资加杠杆的方式放大收益。

这里，股票 T+0 交易的一个优势展现出来——T+0 零融资成本。

按照广发证券融资融券业务相关规定，T 日融资买入股票，并在 T 日归还现金的；或在 T 日融券卖出股票，并在 T 日进行还券的，免收融资融券相关费用。

之前我们研究过的很多策略中，只要涉及到融资融券的，很多模型效果都不是很好，主要归因于高额的融资融券成本。而在 T+0 交易方式下，可以通过融资融券放大交易杠杆并忽略这一部分交易成本的影响。

在融资融券业务中，需要注意将维持担保比例和保证金比例这两个风险警戒指标保持在一定范围内。

1、维持担保比例

维持担保比例是指客户担保物价值与其融资融券债务之间的比例。其计算公式为

$$\text{维持担保比例} = \frac{\text{现金} + \text{信用证券账户内证券市值}}{\text{融资买入金额} + \text{融券卖出证券数量} \times \text{市价} + \text{利息及费用}} \quad (1)$$

特别注意三点：

- (1) 当维持担保比例 $\geq 300\%$ 时，可以提取超出部分的担保物。
- (2) 当维持担保比例 $< 150\%$ 时，需要追加保证金到 150% 以上。
- (3) 当维持担保比例 $< 130\%$ 时，未按约定追加保证金，将被强制平仓。

假设我们有 100 元现金作为担保品进行融资，实现融资买入股票共计 x_1 元，则初始维持担保比例等于 $(100 + x_1) / x_1$ 。由于是日内交易，只要保证日内维持担保比例大于等于 150% ，所以按照最差的情况——股票以上涨 6% 开盘，开盘买入后盘中跌停，即 $(100 + x_1 / 1.06 * 0.9) / x_1 \geq 150\%$ ，可以解得 $x_1 \leq 153\%$ ，即大约可以融资买入 150 元的股票。

若将提供担保的 100 元现金再买入股票，则总计买入的股票为 250 元。也就相当于提供了 2.5 倍的杠杆，这是在维持担保比例约束下可提供的最大杠杆。

2、保证金比例

保证金比例为投资者交付的保证金与融资、融券交易金额的比例。保证金比例用于控制投资者初始资金的放大倍数，相当于杠杆控制器。其又可分为融资保证金比例和融券保证金比例两种。我们这里仅关注融资保证金比例，其计算公式为

$$\text{融资保证金比例} = \frac{\text{保证金}}{\text{融资买入证券数量} \times \text{买入价格}} \quad (2)$$

交易所规定的融资保证金比例不得低于 50%。

在此要求下，假设我们有 100 元现金作为担保品进行融资，实现融资买入股票共计 x_2 元，再用 100 元担保现金全部买入股票。由于股票的保证金最低折算率为 65%，则初始融资保证金比例等于 $100 \times 65\% / x_2$ 。同样按照最坏的情况打算，即股票以上涨 6% 开盘，开盘买入后盘中跌停，则融资保证金比例应满足 $100 \times 65\% \times 0.9 / 1.06 / x_2 \geq 50\%$ 。因此可以得到融资保证金比例至少要满足 $x_2 \leq 110$ 。加上保证金买入的 100 元股票，相当于提供了 2.1 倍杠杆。

还有一种方式，是用现金作为担保物，现金不再进行股票购买，这样保证金的折算率较高，可以按照 100% 折算（即不折算）。这种情况下，假设融资买入股票共计 x_3 元，融资保证金比例为 $100 / x_3 \geq 0.5$ ，则 $x_3 \leq 200$ ，也能提供 2 倍杠杆。

所以在保证金比例约束下可提供的最大杠杆约为 2 倍。

所以综合维持担保比例和保证金比例两个约束条件，股票 T+0 交易在合理风险控制下可取的最大杠杆倍数为

$$L = \min \left\{ L_{\text{维持担保比例约束}}, L_{\text{保证金比例约束}} \right\} \approx 2 \quad (3)$$

即上述试盘者 T+0 策略可取 2 倍杠杆放大收益。在这样的杠杆放大下，上述策略的累积收益曲线变为

图13：试盘者交易策略累积收益曲线（2倍杠杆）



数据来源：广发证券发展研究中心，天软科技

相关交易数据如表 2 所示。

表 2: 试盘者交易策略回测统计数据 (2 倍杠杆)

回测时间	2010 年 1 月 1 日至 2014 年 5 月 16 日
回测天数	1056
累积收益率	338.13%
年化收益率	39.90%
交易总次数	249
平均单笔收益率	0.68%
获胜次数	144
失败次数	105
胜率	57.83%
单次获胜收益率	3.90%
单次失败亏损率	-3.74%
盈亏比	1.04
最大回撤	-27.06%
最大连胜次数	12
最大连亏次数	5
单日最大盈利率	12.14%
单日最大亏损率	-7.68%

数据来源: 广发证券发展研究中心, 天软科技

(五) 止盈止损控制系统性风险

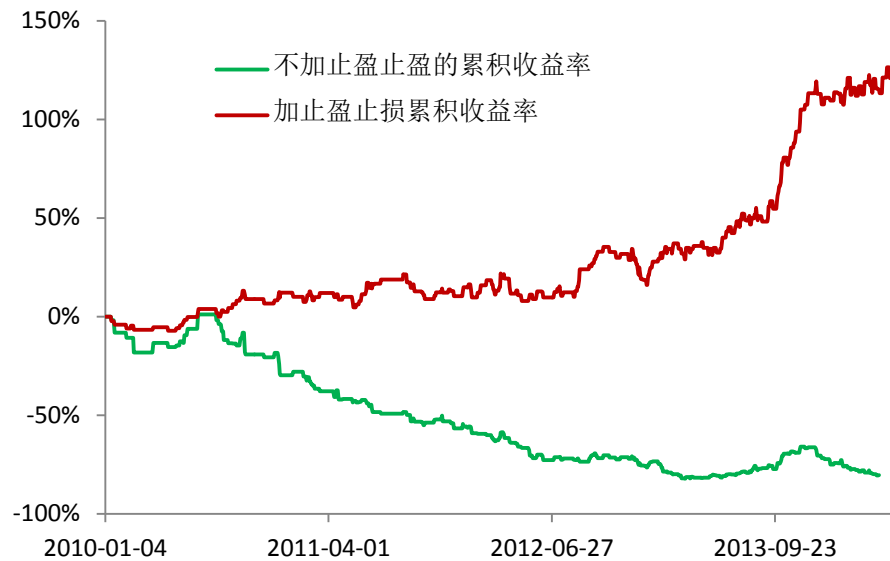
由于目前 T+0 交易制度还没有正式推出, 因此很多投资者会感兴趣上述策略能否应用于 T+1 交易, 即按照之前的规则, 在 T 日开盘买入, T+1 日开盘卖出。

这里我们首先测算了 T+0 交易不设止盈止损, 开盘买入后拿到尾盘, 按照收盘价平仓的结果。测算累积收益曲线与图 9 相比, 发生了显著的变化, 如图 14 所示。

收益情况发生如此大的变化, 究其原因, 是因为 2010 年至今, 市场总体处于下跌行情, 因此个股也难免受到系统性下跌风险的影响。很多个股集合竞价触及涨停并以上涨局面开盘, 在盘中也出现大幅拉高, 但由于大盘不济, 最后也难免回落下跌。

因此, 找到波动大的股票组合后, 通过加入止盈止损策略, 有利于在一定程度上锁定收益, 规避系统性下跌风险。而针对前述试盘者 T+1 交易是否可行的问题, 在此也没有必要继续追究下去了。

图14: 不加止盈止损的累积收益曲线



数据来源: 广发证券发展研究中心, 天软科技

(六) 大数据下的模型调试

针对上一节的止盈止损问题, 有人会提出: 策略采用了 $\pm 1.5\%$ 的止盈止损, 这个参数为什么要这样取? 是怎么优化出来的?

对于这个问题, 我们可以负责的说, 上述模型没有经过参数优化, 因此在测算的时候也没有做样本内外的区分。这里的参数不光指止盈止损点, 还包括开盘涨幅区间。

没有优化参数的原因在于这个模型的数据量实在太大了。试盘者模型要用到全市场两千多只股票 4 年多的 tick 数据, 初步估计, 数据量大概在 2GB。如果回测时间更长, 则需要更大的数据量。

单纯几个 GB 的数据测算倒是没有太大问题, 例如我们通过 matlab 调用天软数据库 (非落地) 的方式, 计算上述策略, 测算一遍的时间大约为 40 小时 (Windows Server 2003 x64, Xeon E5620 2.4G 8 核 16 线程, 16G)。

但是如果加入若干参数的调试和寻优过程, 相当于加入了若干个维度的搜索过程, 这项工作将成为一项巨大的工程。由于时间限制, 本篇报告中我们没有做这样的尝试, 但是我们这里分别从软件和硬件方面提出两种解决问题的思路——1、改进参数搜索算法; 2、改进大数据计算方式。

1、改进参数搜索算法

按照传统寻优方式, 可以对样本空间进行深度优先搜索或广度优先搜索。广度优先是从初始状态一层一层向下找, 直到找到目标为止。深度优先是按照一定的顺序前查找完一个分支, 再查找另一个分支, 以至找到目标为止。

但不论深度优先还是广度优先, 它们其实都是一种穷举算法, 在时间复杂度方

面并没有得到改善。如果有 n 个参数，每个参数有 k_n 个待测元素，那么策略调试的

时间复杂度将增加 $\prod_{i=1}^n k_i$ 倍。比如之前我们的策略跑一次需要大约 40 个小时，那么

当我加入止盈、止损、开盘价下界、开盘价上界四个参数后，假设每个参数要测试 10 个候选值，那么如果按照上述穷举算法，一共需要测算 10^4 个可能的参数组合，整体的测算时间将达到 40 万个小时 ≈ 45 年！这显然是不可取的。

在这种大数据模型测算背景下，采用启发式搜索算法作为寻优算法的改进是有必要的。

启发算法包括遗传算法、模拟退火算法、蚁群算法、粒子群算法等。

遗传算法 (Genetic Algorithm) 是一类借鉴生物界的进化规律 (适者生存, 优胜劣汰遗传机制) 演化而来的随机化搜索方法。它是由美国的 J. Holland 教授 1975 年首先提出, 其主要特点是直接对结构对象进行操作, 不存在求导和函数连续性的限定; 具有内在的隐并行性和更好的全局寻优能力; 采用概率化的寻优方法, 能自动获取和指导优化的搜索空间, 自适应地调整搜索方向, 不需要确定的规则。

模拟退火算法 (Simulated Annealing) 最早的思想由 N. Metropolis 等人于 1953 年提出。1983 年 S. Kirkpatrick 等成功地将退火思想引入到组合优化领域。它是基于 Monte-Carlo 迭代求解策略的一种随机寻优算法, 其出发点是基于物理中固体物质的退火过程与一般组合优化问题之间的相似性。模拟退火算法从某一较高初温出发, 伴随温度参数的不断下降, 结合概率突跳特性在解空间中随机寻找目标函数的全局最优解, 即在局部最优解能概率性地跳出并最终趋于全局最优。

蚁群算法 (Ant Colony Optimization), 又称蚂蚁算法, 是一种用来在图中寻找优化路径的机率型算法。它由 Marco Dorigo 于 1992 年在他的博士论文中提出, 其灵感来源于蚂蚁在寻找食物过程中发现路径的行为。其主要特点是: 通过正反馈、分布式协作来寻找最优路径。

粒子群算法, 也称粒子群优化算法 (Particle Swarm Optimization), 是近年来发展起来的一种新的进化算法。粒子群算法属于进化算法的一种, 和模拟退火算法相似, 它也是从随机解出发, 通过迭代寻找最优解, 它也是通过适应度来评价解的品质, 但它比遗传算法规则更为简单, 它没有遗传算法的“交叉”和“变异”操作, 它通过追随当前搜索到的最优值来寻找全局最优。

无论是上述算法还是其他一些启发式算法, 随着时代的演变它们也在不断演绎出新的版本, 这里我们并不打算详细介绍每一个算法。我们只是希望提出这样一种想法, 供读者在金融市场利用海量数据进行建模和测算时进行参考, 节省大数据时代宝贵的每一分钟。

2、改进大数据计算方式

在之前的测算中, 我们通过 matlab 利用 COM 对象远程调用天软数据库, 并对模型进行计算。这种算法是在没有进行参数寻优的基础上进行的。如果模型调试过程中涉及到参数寻优, 应该首先将海量数据落地到本地硬盘, 再进行计算, 这样就不会重复远程调用数据库, 进行了大量的网络传输时间。

除此之外, 在处理海量数据的过程中, 还有一些技术性的问题可以参考。接下

来我们简单了解一下常见大数据分析的解决方案。

技术方面，大数据主要包括储存和处理两个方面，考虑到量化投资的实际需求，我们这里主要介绍处理方面的解决方案。

海量数据处理往往采用分布式计算。Hadoop 是一个开放源码的分布式数据处理系统架构，主要面向存储和处理结构化、半结构化或非结构化、真正意义上的大数据（通常成百上千的 TB 甚至 PB 级别数据）应用。其分布式文件系统 HDFS 具有高容错性而且对硬件系统的要求不高。

Hadoop 提供的 MapReduce（和其他一些环境）是处理大数据分布式计算的理想解决方案。在 Map 阶段，MapReduce 能将大数据问题分解成多个子问题，将它们分配到成百上千个处理节点之上，在各个子节点上同时进行运算，然后在 Reduce 阶段，将各个节点的计算结果汇集到一个小数据集当中，从而更容易分析得出最后的结果。这一标准化的解决方案极大的提高了分布式计算程序的开发和运行的效率。

Hadoop 可以运行在低成本的硬件产品之上，通过扩展可以成为商业存储和数据分析的替代方案。它已经成为很多互联网巨头，比如 AOL、eHarmony（美国在线约会网站）、易趣、Facebook、Twitter 和 Netflix 大数据分析的主要解决方案。也有更多传统的巨头公司比如摩根大通银行，也正在考虑采用这一解决方案。

除了计算架构，服务器（阵列）、数据库管理模式、数据仓库的平台也成为制约大数据处理的关键因素，这里我们就不一一列举。

总而言之，随之国内量化市场的发展，海量数据处理已成为未来发展的一大趋势。机构投资者在软硬件方面提前布局，有利于在量化市场的竞争格局中抢先一步，制胜高点。

五、总结

本篇报告在大数据应用的背景下开发了基于集合竞价试盘者行为的股票 T+0 交易策略，在不进行期货对冲的情况下，仅仅做多个股（组合）就实现了绝对收益的获取，为投资者提供了全新的投资思路。作为 T+0 交易方案推出前的策略储备，我们从市场分析，到模型建立，到实证分析，到技术处理，给出一套完整的解决方案。我们相信，未来即便市场推出 T+0 交易制度，相关量化策略在 A 股较高交易成本的约束下，也并不容易开发和实现。但我们仍然会努力进行挖掘和开发，尝试寻找和完成更多的有效策略。

风险提示

本篇报告测算数据均来自 A 股 T+1 市场。T+0 交易制度推广后，有可能会改变整个市场的微观结构，对本策略造成一定的影响，请投资者注意相关风险。

广发证券—行业投资评级说明

- 买入： 预期未来 12 个月内，股价表现强于大盘 10%以上。
持有： 预期未来 12 个月内，股价相对大盘的变动幅度介于-10%~+10%。
卖出： 预期未来 12 个月内，股价表现弱于大盘 10%以上。

广发证券—公司投资评级说明

- 买入： 预期未来 12 个月内，股价表现强于大盘 15%以上。
谨慎增持： 预期未来 12 个月内，股价表现强于大盘 5%-15%。
持有： 预期未来 12 个月内，股价相对大盘的变动幅度介于-5%~+5%。
卖出： 预期未来 12 个月内，股价表现弱于大盘 5%以上。

联系我们

	广州市	深圳市	北京市	上海市
地址	广州市天河北路 183 号 大都会广场 5 楼	深圳市福田区金田路 4018 号安联大厦 15 楼 A 座 03-04	北京市西城区月坛北街 2 号 月坛大厦 18 层	上海市浦东新区富城路 99 号 震旦大厦 18 楼
邮政编码	510075	518026	100045	200120
客服邮箱	gfyf@gf.com.cn			
服务热线	020-87555888-8612			

免责声明

广发证券股份有限公司具备证券投资咨询业务资格。本报告只发送给广发证券重点客户，不对外公开发布。

本报告所载资料的来源及观点的出处皆被广发证券股份有限公司认为可靠，但广发证券不对其准确性或完整性做出任何保证。报告内容仅供参考，报告中的信息或所表达观点不构成所涉证券买卖的出价或询价。广发证券不对因使用本报告的内容而引致的损失承担任何责任，除非法律法规有明确规定。客户不应以本报告取代其独立判断或仅根据本报告做出决策。

广发证券可发出其它与本报告所载信息不一致及有不同结论的报告。本报告反映研究人员的不同观点、见解及分析方法，并不代表广发证券或其附属机构的立场。报告所载资料、意见及推测仅反映研究人员于发出本报告当日的判断，可随时更改且不予通告。

本报告旨在发送给广发证券的特定客户及其它专业人士。未经广发证券事先书面许可，任何机构或个人不得以任何形式翻版、复制、刊登、转载和引用，否则由此造成的一切不良后果及法律责任由私自翻版、复制、刊登、转载和引用者承担。