



2012.06.26

R-Breaker 模型后续测试及优化

——程序化交易系列研究之六

	杨喆 (分析师)	蒋瑛琨 (分析师)	耿帅军 (研究助理)
	021-38676442	021-38676710	021-38674875
	yangzhe@gtjas.com	jiangyingkun@gtjas.com	gengshuaijun@gtjas.com
证书编号	S0880511010020	S0880511010023	S0880111110128

本报告导读:

模型数据敏感性测试、参数敏感性测试、改进 R-Breaker 模型等

摘要:

前期报告回顾

- 我们在前期报告中对 R-Breaker 模型进行了详细介绍和沪深 300 指数期货的数据测试。在 2010 年 4 月 16 日至 2011 年 12 月 16 日的测试区间内,总交易 377 笔,盈利交易比例为 48.5%,合计盈利 1644.4 个指数点。扣除交易及冲击成本后,盈利 1313.8 个指数点,折合 394140 元,假设 50 万初始资金的累计收益率为 78.8%,最大回撤率为 6.25%。

本期报告要点

- 在本报告中,我们从多方面对该模型进行更详细的测试和分析,包括数据质量的改进、采用指数数据对结果的影响、采用不同频率的数据对结果的影响、不同止损位的测试、触发条件敏感性测试、趋势与反转交易分类测试等。在进行了不同方面的敏感性测试后,我们针对测试结果对模型进行了改进,最后我们将 Improved R-Breaker Model (IRBM) 测试结果展示给大家。

本期报告结论

- 经过多方面测试,我们得到的几点结论如下:
 - 模型测试效果对数据质量高度依赖
 - 指数测试收益显著低于同期期货
 - 不同参数组合的测试结果显示,绝大多数参数组合可以让我们取得较为稳定的收益。仅有 19.2% 的参数获得的收益率为负。有 63.9% 的参数可以获得 10% 以上的收益率。
 - 采用 1 分钟数据可以获得比采用 5 分钟数据更优的收益。
 - 一定的止损,能提升整体收益并且降低风险。
- 改进模型 Improved R-Breaker Model (IRBM):** 针对 R-Breaker 模型高度数据依赖和条件依赖的特点,为了更好地反应 R-Breaker 模型的思想,我们对 R-Breaker 模型进行适度优化。经过改进后,IRBM 模型将不带杠杆收益率提升到了 75% (年化 37.7%),最大回撤降低至 3.9%,总交易笔数有所上升,胜率有所下降,但整体收益与风险特性显得更优。

相关报告

《全球顶尖程序化交易模型研究——程序化交易系列研究之五》	2012.02.01
《股指期货量价结合交易策略探索——程序化交易系列研究之四》	2010.08.06
《股指期货改进布林交易模型 (IBTM)——程序化交易系列研究之三》	2010.04.27
《股指期货即日交易模型 (DTM)——程序化交易系列研究之二》	2010.03.17
《程序化交易的模型和应用——程序化交易系列研究之一》	2010.03.15

目 录

1. 前期报告回顾和本期要点	3
2. R-Breaker 模型条件敏感性测试	4
2.1. 模型测试效果对数据质量高度依赖.....	4
2.2. 指数测试收益显著低于同期期货.....	5
2.3. 模型触发条件敏感性测试.....	8
2.4. 数据频率对于测试结果的影响.....	9
2.5. 不同止损位对交易结果的影响.....	10
2.6. 趋势与反转交易分类测试.....	10
3. 改进 R-Breaker 模型(IRBM).....	12

1. 前期报告回顾和本期要点

我们在前期报告中对 R-Breaker 模型进行了详细介绍和沪深 300 指数期货的数据测试。我们选取期货连续合约的 5 分钟高频数据，交易一手合约。假定万分之三的交易及冲击成本、1% 的止损线和 30 个点的波幅过滤。在 2010 年 4 月 16 日至 2011 年 12 月 16 日的测试区间内，总交易 377 笔，盈利交易比例为 48.5%，合计盈利 1644.4 个指数点。扣除交易及冲击成本后，盈利 1313.8 个指数点，折合 394140 元，最大回撤 81150 元。假设 50 万初始资金只交易一手合约，期指推出以来的累计收益率为 78.8%，最大回撤率为 6.25%。

图 1 模型累计收益率变化情况（2010Q2-2011Q4）



数据来源：国泰君安证券研究

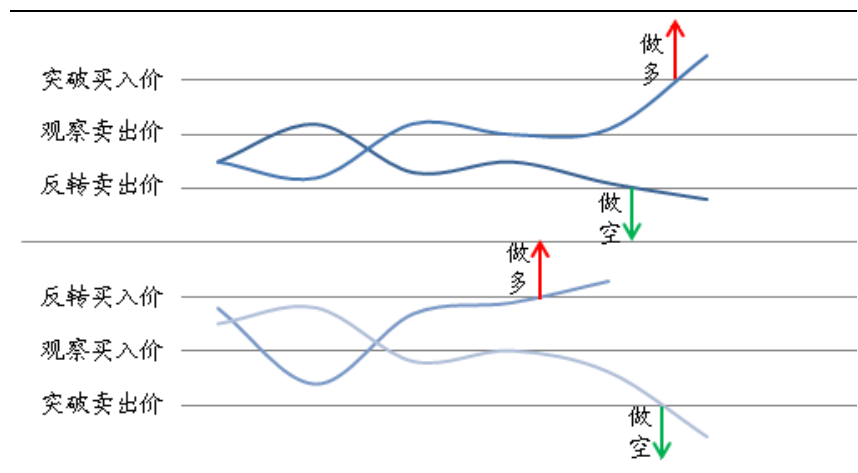
在本报告中，我们首先将考虑数据质量对测试结果的影响。而后我们以数据质量改进后的测试结果为比较基准，从多方面对该模型进行更详细的测试和分析，包括采用指数数据对结果的影响、采用不同频率的数据对结果的影响、不同止损位的测试、触发条件敏感性测试、趋势与反转交易分类测试等。在进行了不同方面的敏感性测试后，我们针对测试结果对模型进行了改进，最后我们将优化后的模型 Improved R-Breaker Model (IRBM) 测试结果展示给大家。

我们先对模型原理进行回顾，R-Breaker 是一种短线交易策略，它结合了趋势和反转两种交易方式。其基本原理如下：

1. 根据前一个交易日的收盘价、最高价和最低价数据通过一定方式计算出六个价位，从大到小依次为：突破买入价、观察卖出价、反转卖出价、反转买入价、观察买入价、突破卖出价。
由这六个价格来形成当前交易日盘中交易的触发条件。对计算方式的调整，可以调节六个价格间的距离，进一步改变触发条件。
2. 追踪盘中价格走势，实时判断触发条件。具体条件如下：
 - ◆ 当日内最高价超过观察卖出价后，盘中价格出现回落，且进一步跌破反转卖出价构成的支撑线时，采取反转策略，即在该点位（反

- 手、开仓) 做空;
 - ◆ 当日最低价低于观察买入价后, 盘中价格出现反弹, 且进一步超过反转买入价构成的阻力线时, 采取反转策略, 即在该点位(反手、开仓) 做多;
 - ◆ 在空仓的情况下, 如果盘中价格超过突破买入价, 则采取趋势策略, 即在该点位开仓做多;
 - ◆ 在空仓的情况下, 如果盘中价格跌破突破卖出价, 则采取趋势策略, 即在该点位开仓做空。
3. 设定止损条件。当亏损达到设定值后, 平仓。
 4. 设定过滤条件。当前一个交易日波幅过小, 该交易日不进行交易。
 5. 在每日收盘前, 对所持合约进行平仓。
 6. 可使用 1 分钟、5 分钟或 10 分钟等高频数据进行判断。

图 2 R-Breaker 模型



数据来源: 国泰君安证券研究

2. R-Breaker 模型条件敏感性测试

2.1. 模型测试效果对数据质量高度依赖

由于 R-Breaker 模型是采用日内高频数据进行交易的策略, 并且交易的触发条件一部分是根据盘中数据设定的, 而我们比较不同数据源的高频数据, 它们之间存在着明显的差异, 数据对于结果的最终影响有多大, 这是我们需要解答的问题。

为了对该问题进行解答, 我们分别选取了来自万德、天软和彭博的数据进行测试。对数据进行初期删选后, 我们剔除了彭博数据, 因为数据异常样本较多, 而万德和天软的数据虽有差异, 但异常数据较少。

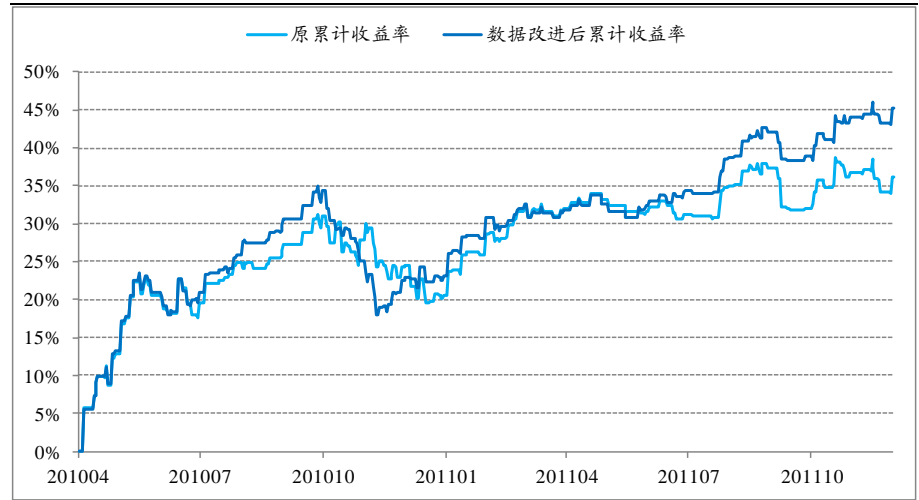
为了保证质量, 我们首先对获取的初始数据进行了预处理, 包括了去除重复和无效样本、补足缺失样本、检验高频数据与日数据之间关系等。

采用数据质量较高的数据源进行测试, 以及对源数据进行预处理, 可以明显改善模型测试效果, 测试结果也能够更真实地反应模型的适用性和局限性。这样, 我们便可以对模型的各种特性进行更深入了解, 从而有

效提高实际交易中的盈利情况。

下面是对数据源质量改善后进行测试的情况。我们仍然选取 5 分钟数据频度的沪深 300 指数股指期货连续合约进行模拟测试，系统参数完全不变，结果发现测试结果差异显著。

图 3 数据质量改善后模型累计收益率变化情况（2010Q2-2011Q4）



数据来源：国泰君安证券研究

经过对数据源的筛选和数据的预处理，数据质量改善，数据稳定性提高，这也直接导致了模型判断机制的有效性得到提高，避免了由于噪音数据而导致的判断和交易失误。交易策略在测试区间内，总交易 258 笔，盈利交易比例为 50.4%，利用点数计算的不带杠杆累计收益率为 45.3%。文章后续部分，如不特殊提及，均以此结果为比较基准。

表 1: 数据质量改善后 R-Breaker 模型测试结果（比较基准）

	累计收益率	最大回撤	夏普比率	总交易笔数	日均交易笔数	盈利交易比例	平均盈利/亏损
基准	45.3%	12.5%	1.1	258	0.6	50.4%	1.5

数据来源：国泰君安证券研究

注：表中指标均为按指数点计算之结果，下同。

2.2. 指数测试收益显著低于同期期货

为检验模型的适用性，我们首先用 08 年到 10 年股指期货推出前的沪深 300 指数 5 分钟高频数据进行测试，系统参数完全不变，R-Breaker 模型共交易 504 笔，取得了 50.1% 的正收益，盈利交易比例为 41.5%。这其中，趋势交易 289 笔，贡献了 119.3% 的盈利；反转交易 215 笔，贡献了 -19.3% 的盈利。可见，在 08 年到 10 年这段时间，由于市场的单边趋势明显，R-Breaker 模型在指数上的盈利全部来自于对于趋势机会的把握，而反转交易整体则略亏。

图 4 沪深 300 指数测试结果 (2008Q1-2010Q1)



数据来源: 国泰君安证券研究

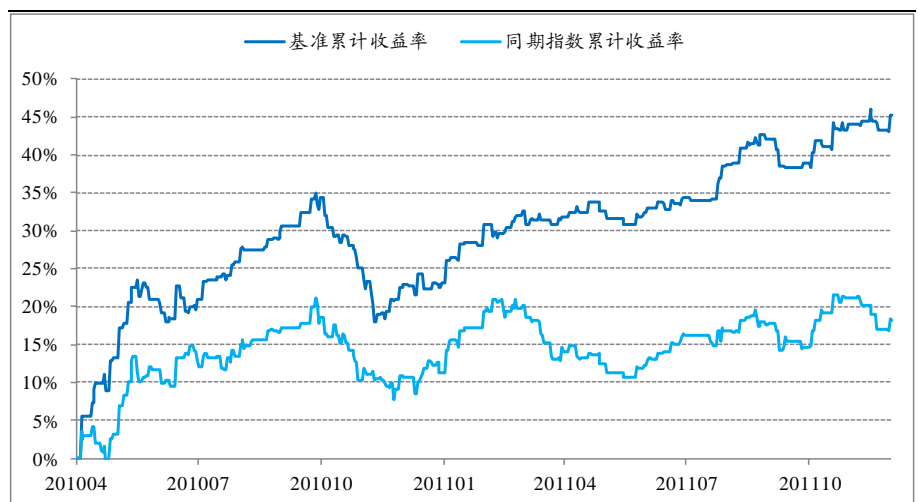
表 2: 沪深 300 指数测试结果 (2008Q1-2010Q1)

	累计 收益率	最大 回撤	夏普 比率	总交易 笔数	日均交 易笔数	盈利交 易比例	平均盈 利/亏损
指数	50.1%	11.4%	0.7	504	0.9	41.5%	1.6

数据来源: 国泰君安证券研究

我们又选取与指数期货推出以来的同时期沪深 300 指数 5 分钟高频数据进行测试, 从数据可以看到, 尽管指数上测试也有正收益, 但是收益要显著低于同期期货。股指期货推出以来, 交易策略在指数上共交易 305 笔, 盈利交易比例为 48.5%, 无杠杆累计收益率为 18.3%。

图 5 沪深 300 指数测试结果 (2010Q2-2011Q4)



数据来源: 国泰君安证券研究

表 3: 沪深 300 指数测试结果 (2010Q2-2011Q4)

	累计 收益率	最大 回撤	夏普 比率	总交易 笔数	日均交 易笔数	盈利交 易比例	平均盈 利/亏损
指数	18.3%	11.1%	0.4	305	0.7	48.5%	1.5
基准	45.3%	12.5%	1.1	258	0.6	50.4%	1.5

数据来源: 国泰君安证券研究

期货和指数间的这种差异主要是因为 R-Breaker 模型是数据高度依赖的日内交易模型。尽管沪深 300 指数和同期股指期货在长期走势上保持一致, 但高频数据的细微差异都可能导致模型给出截然不同的交易指令, 从而导致交易结果完全不同。

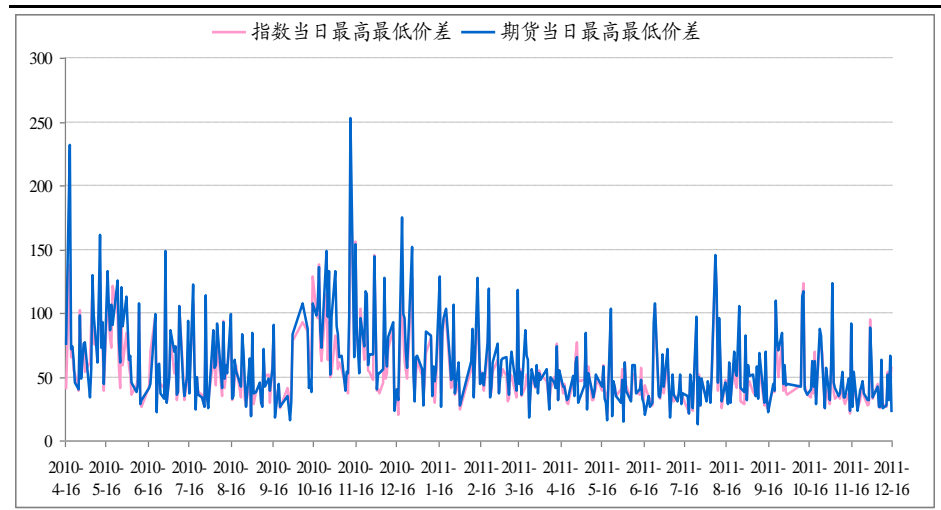
我们对比分析指数和期货数据的差异, 主要有以下几点:

1. 由于现货市场和期货市场的交易时间不同, 一个交易日对应的五分钟数据, 期货数据有 54 个, 对应的沪深 300 指数有 48 个。
2. 期货市场与现货市场波动情况存在显著差异。

观察沪深 300 指数和对应的股指期货的日价差 (最高价-最低价), 发现期货的日价差的平均值、标准差、峰值都显著大于指数。

采用 5 分钟高频数据计算指数和期货的日内实际波动率, 期货的波动率的平均值、标准差、峰值也同样高于指数。

图 6 沪深 300 指数期货日价差显著高于指数 (2010Q2-2011Q4)



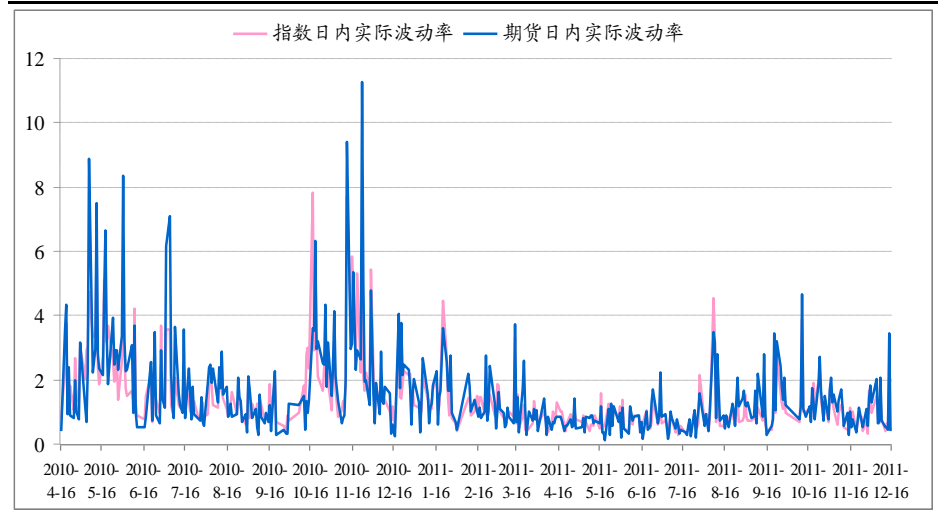
数据来源: 国泰君安证券研究

表 4: 指数和期货日价差的统计量情况

标的	均值	标准差	最大值	最小值
指数	54.7	26.9	212.7	20.2
期货	59.5	32.6	252.4	12.6

数据来源: 国泰君安证券研究

图 7 沪深 300 指数期货日内实际波动率高于指数 (2010Q2-2011Q4)



数据来源：国泰君安证券研究

表 5: 指数和期货日内实际波动率的统计量情况

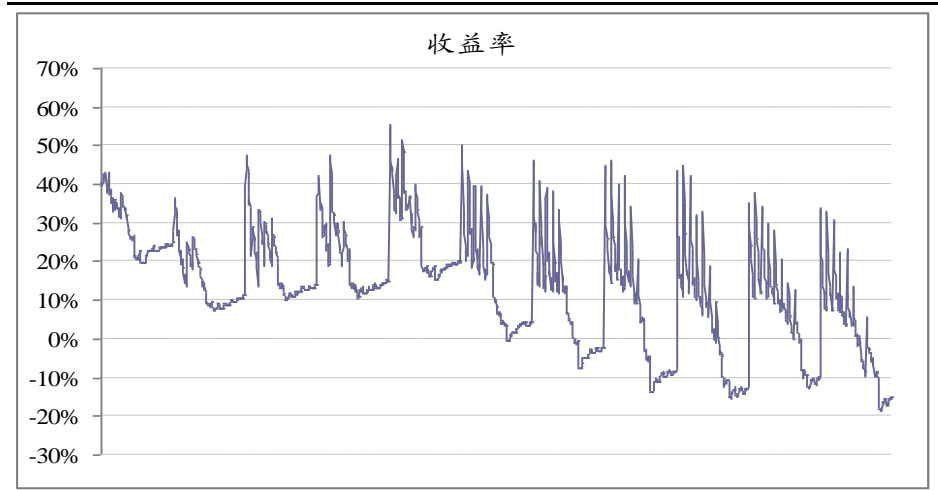
标的	均值	标准差	最大值	最小值
指数	1.4	1.1	7.8	0.3
期货	1.5	1.4	11.3	0.1

数据来源：国泰君安证券研究

2.3. 模型触发条件敏感性测试

在上篇报告中我们提到，交易系统曲线匹配的特性越强，越不可能在未来发生作用。如果过度优化，系统未来的可靠性将降低。因此在优化过程中，要选择参数值的适当区间，而不是最佳参数值，以避免过度优化问题。下面我们便在不改变其它条件的情况下，通过不断改变触发条件来观察其对收益的影响。图中横坐标为不同的触发条件，纵坐标则表示在不同条件下取得的累计收益率。

图 8 不同触发条件下的累计收益率情况



数据来源：国泰君安证券研究

通过测试，我们希望能找到结果较为稳定的一个参数区间，同时避免那

些收益较差的区间。

测试结果显示，绝大多数参数组合可以让我们取得较为稳定的收益。仅有 19.2% 的参数获得的收益率为负。有 63.9% 的参数可以获得 10% 以上的收益率。

表 6: 不同触发条件下的累计收益率分布

累计收益率	>0%	>10%	>20%	>30%	>40%	>50%
样本占比	80.8%	63.9%	32.4%	13.5%	3.7%	0.2%

数据来源：国泰君安证券研究

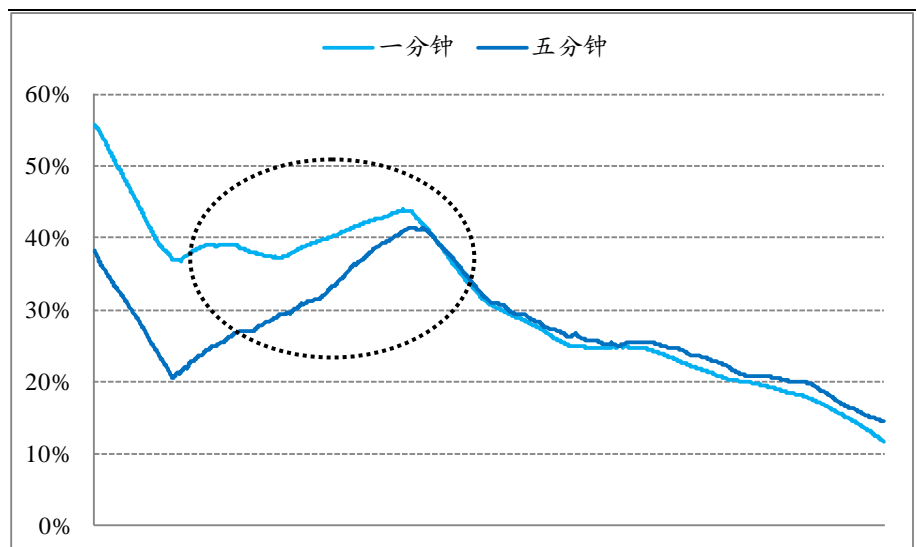
2.4. 数据频率对于测试结果的影响

我们在前文中提到过 R-Breaker 模型是数据高度依赖的交易模型，因此数据质量越高，模型越能反应其思想。其中数据频率也是有可能对结果造成影响的因素。因此我们分别使用 1 分钟数据和 5 分钟数据对模型进行测试，比较其结果。

下图为采用 1 分钟数据和 5 分钟数据得到的累计收益率情况。图中横坐标为不同的触发条件，纵坐标则表示在不同条件下取得的累计收益率。

数据显示，两种数据均有占优的时候，但优势的程度不一样。5 分钟数据优于 1 分钟数据的情形下，收益率提升并不明显，最多的时候提升了 2.7%；而 1 分钟数据优于 5 分钟数据的情形下，收益率却相差很大，1 分钟数据最多的时候产生的收益率能比 5 分钟数据高 18.3%。因此整体来看我们认为采用 1 分钟数据更优。也就是说，随着数据频率的提高，采用更加精细的数据，可以使得数据高度依赖的 R-Breaker 模型有更好的判断，从而使模型效果也更加优异。

图 9 采用不同数据频率的模型收益比较



数据来源：国泰君安证券研究

2.5. 不同止损位对交易结果的影响

在程序化交易系统中，系统给出离场交易信号通常有以下四种情况：止盈、止损、趋势终止和收市。R-Breaker 是一种结合了多种终止规则的交易策略。在这里，我们主要针对止损位的设置来进行一些思考。在程序化交易中，止损位是该设置地较为宽松还是较为严格没有一个肯定的答案。通常情况下，我们认为日内交易策略，设置的止损位应该严格一些。因为系统每天都会进行交易，严格的止损设置，会避免由于频繁交易胜率降低而导致的收益率大幅下降。而对于交易周期较长的系统来说，止损位应该设置的宽松些。因为长线交易系统捕捉的是市场的长期趋势，短期的市场波动不应成为交易终止的条件。

我们通过对 R-Breaker 的止损位进行敏感性测试，来观察其对交易结果带来的影响。我们保持其它系统参数不变，止损位分别设置为 0.1%、0.3%、0.5%、0.8%、1.5%、2% 和无止损情形。

表 7: 不同止损位下的测试结果比较

止损位	0.1%	0.3%	0.5%	0.8%	1.5%	2%	Inf
累计收益率	45.4%	50.8%	49.6%	44.7%	41.2%	38.5%	37.6%
交易次数	258	258	258	258	258	258	258
盈利交易比例	27.1%	41.1%	46.9%	49.6%	51.2%	51.2%	51.2%
止损次数	174	114	72	39	11	7	0
收益率最大回撤	5.3%	6.6%	7.8%	11.0%	12.5%	14.5%	16.6%
平均盈利/亏损	5.3	2.5	1.9	1.6	1.4	1.3	1.3

数据来源：国泰君安证券研究

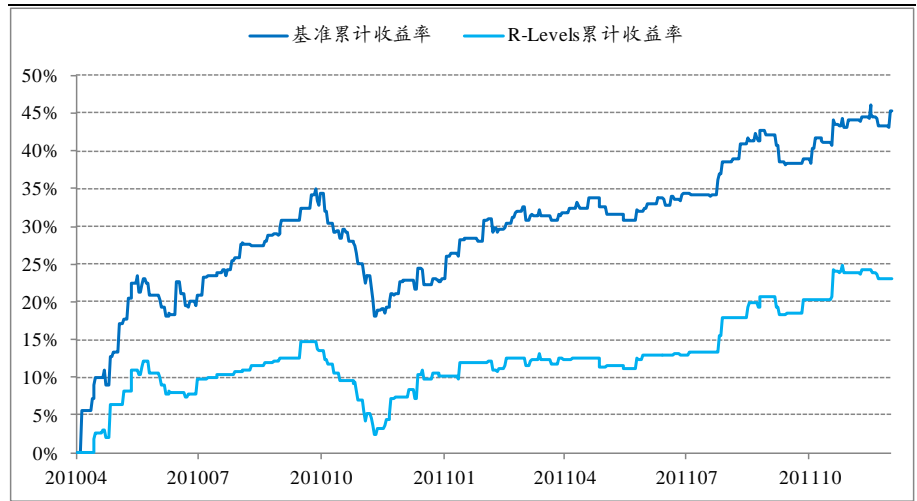
经过测试我们发现，随着止损位设定的越来越严格，胜率显著下降，但是收益率先升后降，也就是说，一定的止损，能提升整体收益。并且止损越严格，最大回撤越小。选择 0.3% 的止损，将使收益率和最大回撤配比最佳。

2.6. 趋势与反转交易分类测试

R-Breaker 模型结合了趋势和反转两种交易策略。其创始人 Richard Saidenberg 还推出了只包含反转策略的交易模型 R-Levels。下面我们通过测试 R-Levels 模型和只进行趋势交易的 R-Breaker 模型，来观察这两种策略的表现。

我们依然使用沪深 300 指数股指期货连续合约的 5 分钟高频数据来进行模拟测试。经过统计，R-Levels 模型共交易 123 笔，盈利交易比例为 55.3%，无杠杆累计收益率为 23.1%。只进行趋势交易的 R-Breaker 模型共交易 175 笔，盈利交易比例为 46.9%，无杠杆累计收益率为 18.9%。

图 10 R-Levels 模型测试结果 (2010Q2-2011Q4)



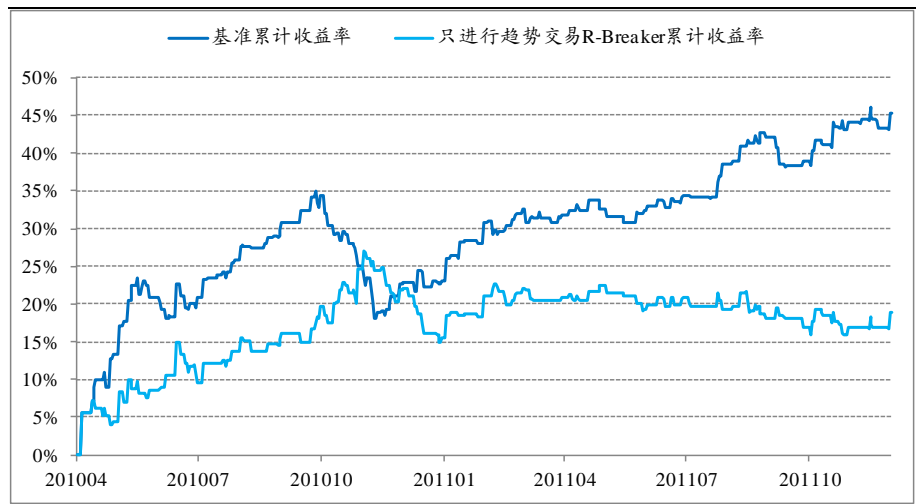
数据来源: 国泰君安证券研究

表 8: R-Levels 模型测试结果

	累计 收益率	最大 回撤	夏普 比率	总交易 笔数	日均交 易笔数	盈利交 易比例	平均盈 利/亏损
R-Levels	23.1%	10.7%	0.5	123	0.3	55.3%	1.3
基准	45.3%	12.5%	1.1	258	0.6	50.4%	1.5

数据来源: 国泰君安证券研究

图 11 只进行趋势交易的 R-Breaker 模型测试结果 (2010Q2-2011Q4)



数据来源: 国泰君安证券研究

表 9: 只进行趋势交易的 R-Breaker 模型测试结果

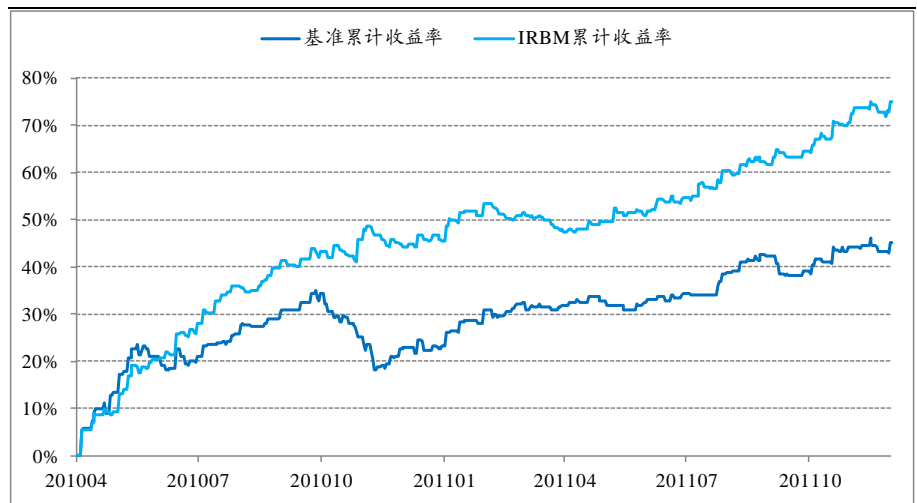
	累计 收益率	最大 回撤	夏普 比率	总交易 笔数	日均交 易笔数	盈利交 易比例	平均盈 利/亏损
趋势 R	18.9%	9.6%	0.4	175	0.4	46.9%	1.5
基准	45.3%	12.5%	1.1	258	0.6	50.4%	1.5

数据来源: 国泰君安证券研究

3. 改进 R-Breaker 模型(IRBM)

针对 R-Breaker 模型高度数据依赖和条件依赖的特点，为了更好地反应 R-Breaker 模型的思想，我们对 R-Breaker 模型进行适度优化，其中包括对数据质量的改进，过滤价格数据中的噪音数据，采用 1 分钟高频数据，对大类市场阶段进行判断等，以更好地切合模型思想。经过改进后，IRBM 模型的测试结果如下图和表格所示。**IRBM 模型将收益率提升到了 75% (年化 37.7%)，最大回撤降低至 3.9%，总交易笔数有所上升，胜率有所下降，但整体的收益与风险特性显得更优。**

图 12 IRBM 模型测试结果 (2010Q2-2011Q4)



数据来源：国泰君安证券研究

表 10: IRBM 模型测试结果

	累计收益率	最大回撤	夏普比率	总交易笔数	日均交易笔数	盈利交易比例	平均盈利/亏损
IRBM	75.0%	3.9%	2.2	308	0.8	40.3%	3.1
基准	45.3%	12.5%	1.1	258	0.6	50.4%	1.5

数据来源：国泰君安证券研究

本公司具有中国证监会核准的证券投资咨询业务资格
分析师声明

作者具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，本报告清晰准确地反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响，特此声明。

免责声明

本报告仅供国泰君安证券股份有限公司（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本公司的当然客户。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。

本报告的信息来源于已公开的资料，本公司对该等信息的准确性、完整性或可靠性不作任何保证。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。

本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。

本公司利用信息隔离墙控制内部一个或多个领域、部门或关联机构之间的信息流动。因此，投资者应注意，在法律许可的情况下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。在法律许可的情况下，本公司的员工可能担任本报告所提到的公司的董事。

市场有风险，投资需谨慎。投资者不应将本报告为作出投资决策的惟一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向专业人士咨询并谨慎决策。

本报告版权仅为本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制、发表或引用。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“国泰君安证券研究”，且不得对本报告进行任何有悖原意的引用、删节和修改。

若本公司以外的其他机构（以下简称“该机构”）发送本报告，则由该机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该机构以要求获悉更详细信息或进而交易本报告中提及的证券。本报告不构成本公司向该机构之客户提供的投资建议，本公司、本公司员工或者关联机构亦不为该机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

评级说明
1. 投资建议的比较标准

投资评级分为股票评级和行业评级。以报告发布后的 12 个月内的市场表现作为比较标准，报告发布日后的 12 个月内的公司股价（或行业指数）的涨跌幅相对同期的沪深 300 指数涨跌幅为基准。

2. 投资建议的评级标准

报告发布日后的 12 个月内的公司股价（或行业指数）的涨跌幅相对同期的沪深 300 指数的涨跌幅。

	评级	说明
股票投资评级	增持	相对沪深 300 指数涨幅 15%以上
	谨慎增持	相对沪深 300 指数涨幅介于 5%~15%之间
	中性	相对沪深 300 指数涨幅介于-5%~5%
	减持	相对沪深 300 指数下跌 5%以上
行业投资评级	增持	明显强于沪深 300 指数
	中性	基本与沪深 300 指数持平
	减持	明显弱于沪深 300 指数

国泰君安证券研究

	上海	深圳	北京
地址	上海市浦东新区银城中路 168 号上海银行大厦 29 层	深圳市福田区益田路 6009 号新世界商务中心 34 层	北京市西城区金融大街 28 号盈泰中心 2 号楼 10 层
邮编	200120	518026	100140
电话	(021) 38676666	(0755) 23976888	(010) 59312799
E-mail:	gtjaresearch@gtjas.com		