

# 基本面量化之二：国债的 Beta



报告日期：2018 年 1 月 4 日

章顺 高级分析师(基本面量化)  
从业资格号：F0301166  
Tel: 8621-63325888-3902  
Email: [shun.zhang@orientfutures.com](mailto:shun.zhang@orientfutures.com)

### ★摘要：

《基本面量化之一：国债的重要因子》中，我们选取了 58 个重要因子，通过这些因子可以对当前国债的基本面有个整体的把握。基本面量化的第二篇中，我们将这 58 个重要因子作为研究的初始数据，进行了国债定价与预测的实证研究：

- 第一，根据因子公布的时间和频率，筛选出符合建模需求的因子，共有 40 个因子；
- 第二，利用相关性分析以及主成分分析对数据进行初步的分析，国债收益率具备预测的可行性；
- 第三，通过 Complete Subset Regressions (CSR) 方法初步筛选模型因子、建立方程组、构建预测模型并形成子策略进行样本内和样本外的测试；
- 第四，根据宏观预测的特性，我们推出预测模型自动跟踪和主观赋权跟踪两种方式来模拟交易的结果。

实证研究的结果表明：

- 第一，包含经济增长、通货膨胀、流动性和银行负债结构信息的预测模型能够获得 65%左右的胜率；
- 第二，30 个月的样本内回测和 38 个月的样本外测试表明模型具备一定的适应性和预测性；
- 第三，多个子策略的模式能够在市场出现拐点时，避免出现净值大幅波动的情形。

关于宏观预测的研究，一直是市场争论的焦点，多数模型在后续跟踪阶段的预测能力下降，我们可能会尝试建立动态的预测模型并充分运用子策略进行测试跟踪，后续会进一步探讨政策变量的影响。接下来，我们将基于基本面的 Beta 策略推广至商品期货和外汇，继续秉承“基本面为主，量化为辅”的研究思路。

**重要事项：**本报告版权归上海东证期货有限公司所有。未获得东证期货书面授权，任何人不得对本报告进行任何形式的发布、复制。本报告的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，也不保证所包含的信息和建议不会发生任何变更。我们已力求报告内容的客观、公正，但文中的观点、结论和建议仅供参考，报告中的信息或意见并不构成交易建议，投资者据此做出的任何投资决策与本公司和作者无关。  
**有关分析师承诺，见本报告最后部分。并请阅读报告最后一页的免责声明。**

## 目录

1、 导读：基本面量化之国债的 Beta .....	4
2、 研究方法 & 数据 .....	4
2.1、 研究方法——Complete Subset Regressions (CSR) .....	5
2.2、 数据 .....	5
2.2.1、 数据初选 .....	5
2.2.2、 相关性分析 .....	8
2.2.3、 主成分分析 .....	10
3、 实证研究 .....	12
3.1、 利用 CSR 进行因子的筛选 .....	13
3.2、 国债 Beta 模型的样本内外测试 .....	16
3.3、 国债 Beta 策略的跟踪方案 .....	20
4、 结论及展望 .....	20
5、 参考文献 .....	21

## 图表目录

图表 1: 经济增长框架内的重要指标 .....	6
图表 2: 通货膨胀框架内的重要指标 .....	6
图表 3: 流动性框架内的重要指标 .....	7
图表 4: 商业银行投资行为框架内的重要指标 .....	8
图表 5: 自变量与国债收益率的相关性 .....	8
图表 6: 自变量相关系数的大小和显著性 .....	9
图表 7: 主成分分析结果 .....	10
图表 8: 主成分的碎石图 .....	11
图表 9: 主成分分析中贡献度 .....	11
图表 10: CSR 实证结果 (nbest=6) .....	13
图表 11: CSR 实证结果 (nbest=6, v1-v35) .....	13
图表 12: CSR 实证结果 (nbest=6, v36-v40) .....	14
图表 13: CSR 初选重要指标 .....	15
图表 14: CSR 样本内实证结果 (30 个月, nbest=8, 月末预测起点) .....	16
图表 15: CSR 样本内实证结果 (30 个月, nbest=8, 月末预测起点) .....	17
图表 16: CSR 样本内 72 个回归方程 AIC 值 (30 个月, nbest=8, 月末预测起点) .....	18
图表 17: CSR 样本内 19 个子策略组合累计净损益 (30 个月, nbest=8, 月末预测起点) .....	18
图表 18: CSR 样本内和样本外 2 个子策略累计净损益 (68 个月, 月末预测起点, 无止盈止损) .....	19
图表 19: CSR 样本内 2 个子策略累计净损益 (30 个月) .....	19
图表 20: CSR 样本外 2 个子策略累计净损益 (38 个月) .....	19
图表 21: 优选子策略样本内和样本外测试结果 .....	20

## 1、导读：基本面量化之国债的 Beta

在《基本面量化之一：国债的重要因子》中，我们从 305 个指标中筛选出 58 个指标，这 58 个指标是影响国债收益率的重要指标，我们在《基本面量化之二：国债的 Beta》中将通过这 58 个指标捕捉国债的 Beta，构建国债收益率的预测模型。

资产收益率的定价研究是金融研究的重要课题，如 CAPM、Fama-French 三因子模型以及后续发展的四因子模型、流动性调整的资产定价模型等等，都是在资产定价领域中贡献显著的研究方向。

在金融市场中，因子定价的实务研究在股票市场中应用最为广泛，典型的方向就是股票的 Beta 策略以及相关的 Alpha 策略。然而，面对日益兴起的大宗商品期货市场以及逐渐对外开放的债券市场，有关因子定价的研究并不多，我们认为关于债券和大宗商品的因子定价研究将是市场研究中的一个新方向，而且我们从基本面的层面挖掘因子将获得较强的逻辑支撑。

国内债市方面，有关债券 Beta 策略的研究较少，而国外针对债券因子的研究较多，Nektarios Aslanidis 等人在 2017 年的一篇文章中提到利用宏观的变量来预测债券的 Beta，在此思路的启发下，我们认为可以利用国债基本面的指标来预测国债的 Beta，其中国债基本面的指标就是我们在《基本面量化之一：国债的重要因子》中筛选出来的指标。

需要强调的一点是，我们研究的方法参考 Elliott, G., A. Gargano, and A. Timmermann (2013)，此方法不同于多数的 Beta 策略研究方法，优于常见的回归模型。

## 2、研究方法 & 数据

因子定价模型是利用相关的因子来解释资产收益率的变化，一般的形式为：

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots$$

其中  $y$  为资产收益率， $x_i$  ( $i=1,2,3,\dots$ ) 为因子， $\beta_j$  ( $j=1,2,3,\dots$ ) 为各因子载荷， $\alpha$  为常数项。如果根据因子的数量来划分，因子模型可以分为单因子模型和多因子模型。典型的单因子模型有 CAPM，多因子模型包括 APT 模型以及 Fama-French 多因子模型。

从前人的研究成果来看，回归分析是资产定价研究的重要方法，不过共线性和过拟合一直是困扰回归分析的问题，为了减弱这些问题的影响，我们引入 Complete Subset Regressions (CSR) 方法，Elliott, G., A. Gargano, and A. Timmermann (2013) 将 CSR 方法与一元线性回归、岭回归、弹性网回归等方法进行比较，发现 CSR 的优势明显。

## 2.1、研究方法——Complete Subset Regressions (CSR)

Complete Subset Regressions (CSR) 是由 Elliott, G., A. Gargano, and A. Timmermann (2013) 提出的一种新的预测方法，将所有潜在的自变量组成一个集合，然后在集合内选择固定数目的自变量全排列地构建回归方程，这样就会得到一个由固定数目自变量的回归方程组成的方程组。

举一个例子，假设我们要预测一个变量  $y_{T+1}$ ，潜在的自变量有  $k$  个，构建的因子模型为单因子模型，那么最终形成的方程组包含  $k$  个回归方程。如果用单个的方程式来表示  $y$  的预测结果，那么具体的式子可以写成如下的形式：

$$y_{T+1} = \frac{1}{k} \left( \sum_{i=0}^k (\beta_i x'_{iT} + \varepsilon_i) \right)$$

实际上，最终的  $y$  值预测通过方程组中所有回归方程的预测值的加权平均来计算。

## 2.2、数据

### 2.2.1、数据初选

实证研究中涉及到的数据主要是基本面相关的变量和 10 年期国债收益率。10 年期国债收益率的数据是日频的，可以转换为低频的数据，而其他基本面相关的变量的频次存在较大的差异。考虑到预测的频次，我们认为预测的频次定为月度较为合适，那么自变量的频次需要统一调整为月度。

在自变量月度数据的调整中，我们统一将月中公布的数据日期调整为当月最后一天。数据对齐后，所有自变量和因变量的数据日期统一定位为当月最后一天。

考虑到建模过程中，数据的时间区间和公布的时间都是影响模型构建的因素，因此时间区间不够或者公布时间延迟较长的因子不纳入到模型中。首先，模型构建中回归分析的时间序列最低需要六年左右的数据，包含样本内和样本外两部分；另外，数据公布的时间不能超过下个月。

**图表 1: 经济增长框架内的重要指标**

项目	指标	公布日期	时间区间
经济增长	固定资产投资完成额累计同比 (v1)	月频, 下月中旬	1992 年至今
	制造业固定资产投资完成额累计同比 (v2)	月频, 下月中旬	2003 年至今
	乡村消费品零售总额累计同比 (v3)	月频, 下月中旬	2010 年至今
	工业增加值: 电力、燃气及水的生产和供应业累计同比	月频, 下月中旬	2013 年至今
	工业增加值累计同比 (v4)	月频, 下月中旬	1998 年至今
	工业企业主营业务收入累计同比 (v5)	月频, 下月下旬	1996 年至今
	房地产开发投资完成额累计同比 (v6)	月频, 下月中旬	1999 年至今
	房屋新开工面积累计同比 (v7)	月频, 下月中旬	1992 年至今
	财新中国 PMI (v8)	月频, 下月月初	2005 年至今
	PMI 从业人员 (v9)	月频, 当月月底	2005 年至今
	PMI 原材料库存 (v10)	月频, 当月月底	2005 年至今
	PMI 供货商配送时间 (v11)	月频, 当月月底	2005 年至今
	PMI 产成品库存 (v12)	月频, 当月月底	2005 年至今

资料来源: Wind, 东证衍生品研究院

经济增长框架内的重要因子中, 电力、燃气及水的生产和供应业工业增加值累计同比的数据时间区间偏窄, 因此需要提出该因子, 而其他因子在时间区间和公布时间方面都满足建模的要求。

**图表 2: 通货膨胀框架内的重要指标**

项目	指标	公布日期	时间区间
通货膨胀	CPI 不包括食品和能源(核心 CPI)当月同比	月频, 下个月的上旬	2013 年至今
	CPI 服务当月同比	月频, 下个月的上旬	2013 年至今
	CPI 非食品当月同比	月频, 下个月的上旬	2013 年至今
	发改委口径 36 个城市平均零售价: 富强粉	日频, 延迟一周内	2015 年至今
	发改委口径 36 个城市平均零售价: 桶装花生油	日频, 延迟一周内	2015 年至今
	发改委口径 36 个城市平均零售价: 桶装菜籽油	日频, 延迟一周内	2015 年至今
	发改委口径 36 个城市平均零售价: 猪肉	日频, 延迟一周内	2015 年至今
	发改委口径 36 个城市平均零售价: 胡萝卜的价格	日频, 延迟一周内	2015 年至今
	统计局口径市场价: HRB400 螺纹钢	日频, 延迟一周内	2014 年至今
	统计局口径市场价: 5#角钢	日频, 延迟一周内	2014 年至今
	统计局口径市场价: 石油苯(工业级)	日频, 延迟一周内	2014 年至今
	统计局口径市场价: 1/3 焦煤	日频, 延迟一周内	2014 年至今
	统计局口径市场价: 黄玉米二等	日频, 延迟一周内	2014 年至今
	统计局口径市场价: 草甘膦(95%原药)	日频, 延迟一周内	2014 年至今
	统计局口径市场价: 尿素(小颗粒)	日频, 延迟一周内	2014 年至今

统计局口径 50 个城市平均价：豆腐 (v13)	周频，延迟两周左右	2009 年至今
统计局口径 50 个城市平均价：一级散装菜籽油 (v14)	周频，延迟两周左右	2009 年至今
统计局口径 50 个城市平均价：标准粉 (v15)	周频，延迟两周左右	2009 年至今
统计局口径 50 个城市平均价：5L 桶装大豆油 (v16)	周频，延迟两周左右	2009 年至今

资料来源：Wind，东证衍生品研究院

通货膨胀框架内，统计局口径 50 个城市平均价中豆腐、一级散装菜籽油、标准粉和 5L 桶装大豆油在时间区间和公布时间两方面满足建模要求。

**图表 3：流动性框架内的重要指标**

项目	指标	公布日期	时间区间
流动性	存款类机构 3 个月质押式回购加权利率	日频	2014 年 12 月至今
	3 个月 SHIBOR (v17)	日频	2006 年 10 月至今
	社会融资规模累计同比 (v18)	月频、下月中旬	2002 年 1 月至今
	金融机构短期贷款余额同比 (v19)	月频、下月中旬	1994 年 1 月至今
	金融机构各项存款余额同比 (v20)	月频、下月中旬	1978 年 12 月至今
	中法 1 年利差 (v21)	日频	2002 年至今
	中美 2 年利差 (v22)	日频	2002 年至今
	中德 2 年利差 (v23)	日频	2002 年至今
	美元指数 (v24)	日频	1971 年至今
	欧元兑美元 (v25)	日频	1978 年至今
	美元兑日元 (v26)	日频	1971 年至今
	官方储备资产中外汇储备 (v27)	月频，下月上旬	1950 年至今
	货币当局国外资产中外汇(中央银行外汇占款) (v28)	月频，下月中旬	1997 年至今
	公共财政支出累计同比 (v29)	月频，下月中旬	1990 年至今
	债券发行量中政策性银行债累计同比 (v30)	月频，下月上旬	1997 年至今
	中小型存款类金融机构人民币存款准备金率(月) (v31)	月频，当月月底	1985 年至今
	大型存款类金融机构人民币存款准备金率(月) (v32)	月频，当月月底	1985 年至今
	金融机构超额存款准备金率(超储率)	季度，延迟两个月左右	2001 年至今
	商业银行国债托管量同比 (v33)	月频，下月上旬	1997 年至今
	其他存款性公司对其他金融机构债权同比 (v34)	月频，下月中旬	1999 年至今

资料来源：Wind，东证衍生品研究院

流动性框架内，存款类机构 3 个月质押式回购加权利率和金融机构超额存款准备金率(超储率)在时间区间和数据公布时间方面不满足建模要求，保留其他指标作为模型的自变量。



图表 4：商业银行投资行为框架内的重要指标

项目	指标	公布日期	时间区间
商业 银行 投资 行为	(中债 10 年国债收益率-中债 10 年国开债到期收益率) (v35)	日频	2002 年至今
	(中债 10 年国债收益率-中债 10 年商业银行普通债到期收益率(AAA)) (v36)	日频	2009 年至今
	(中债 10 年国债收益率-中债 10 年企业债到期收益率(AAA)) (v37)	日频	2006 年至今
	(中债 10 年国债收益率-中债 10 年铁道债到期收益率) (v38)	日频	2008 年至今
	住户活期存款/非金融企业及机关团体各项贷款 (v39)	月频, 下月中旬	2004 年至今
	非金融企业活期存款/非金融企业及机关团体各项贷款 (v40)	月频, 下月中旬	2011 年至今

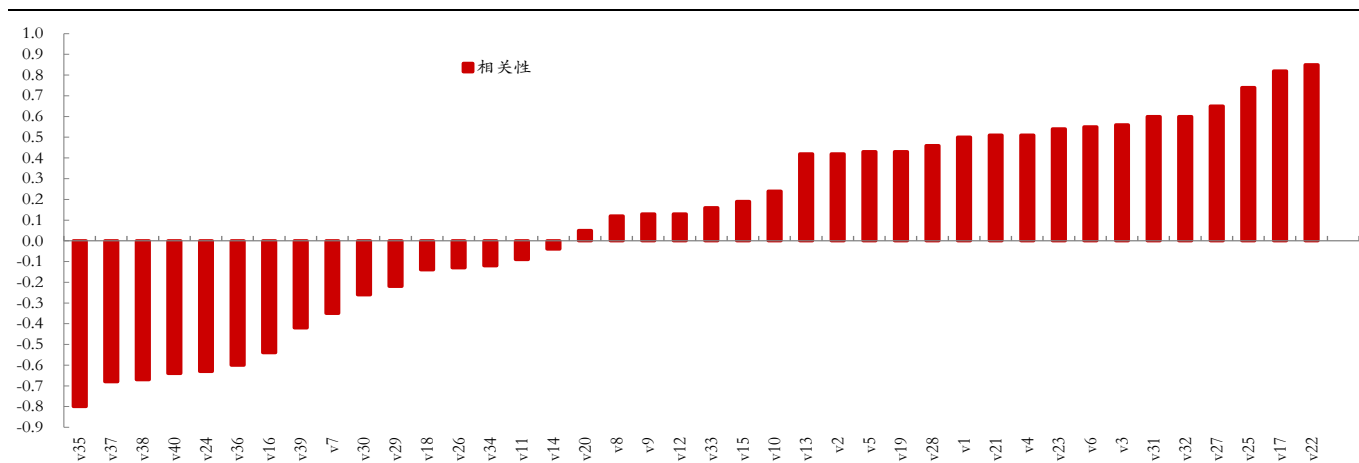
资料来源：Wind，东证衍生品研究院

商业银行投资行为框架内的指标都满足建模的要求。根据数据建模的要求，经济增长内筛选出来的指标有 12 个，通货膨胀框架内满足条件的指标有 4 个，流动性框架内 18 个，商业银行投资行为框架内包含 6 个，经过初步筛选出来的指标共有 40 个。

## 2.2.2、相关性分析

经过初选，我们得到 40 个指标，为了便于展示，我们将 40 个指标进行了标识，依次标记为 v1-v40，与上述表格中的指标一一对应。

图表 5：自变量与国债收益率的相关性



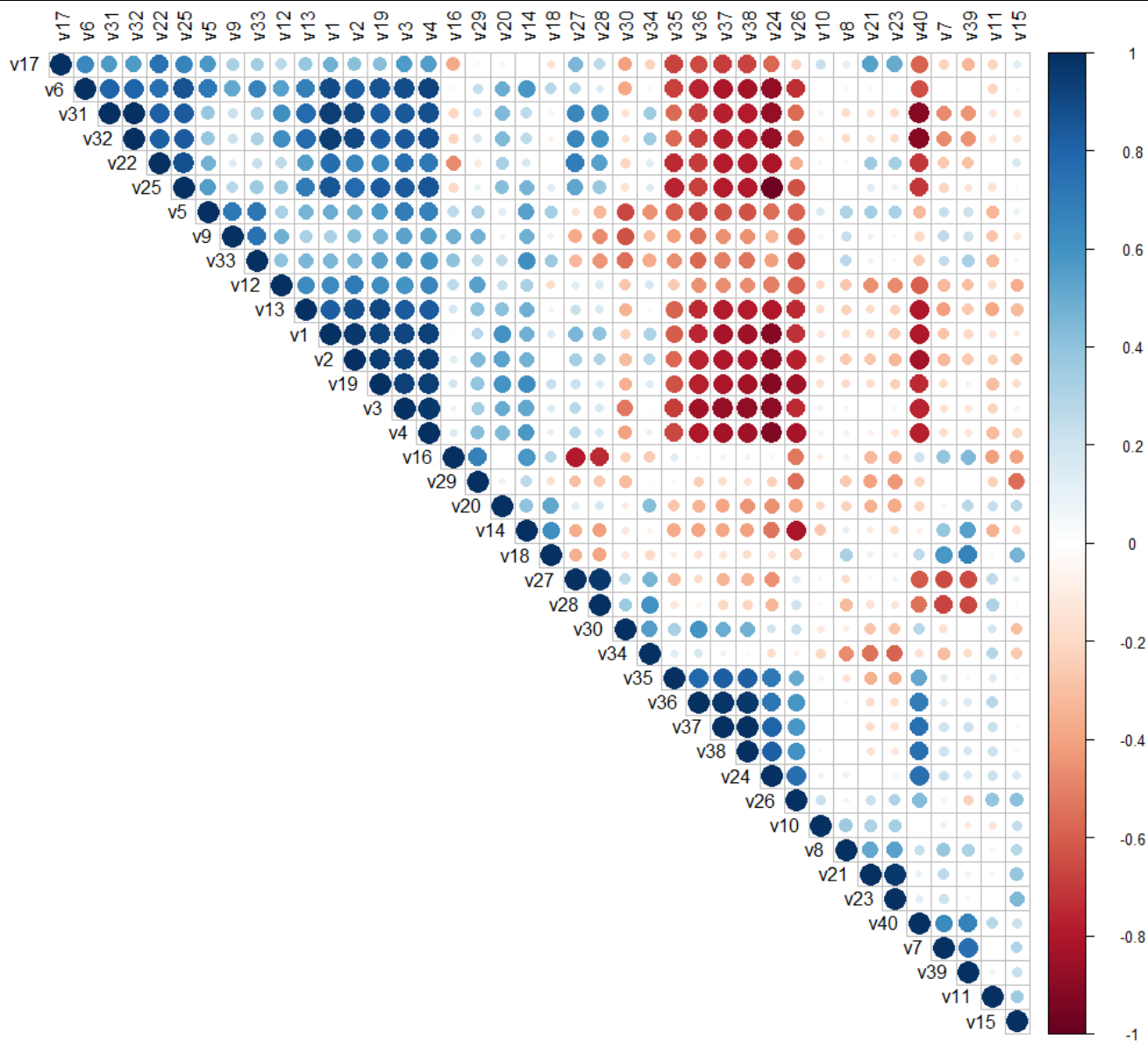
资料来源：Wind，东证衍生品研究院

相关性分析的结果表明国债收益率具有预测的可行性，部分因子可能具备预测国债收益率波动的特点。不过，通过自变量间的相关性分析的结果观察，发现部分自变量之间存在较强的相关性，因此在做 CSR 分析时极有必要进行回归模型的筛选。



观察自变量相关性分析的结果，国债与其他债券的利差和经济指标呈显著负相关，欧元兑美元与投资类等经济指标显著正相关，欧元兑美元与债券利差显著负相关，经济指标间一般呈现正相关的关系。

图表 6：自变量相关系数的大小和显著性



资料来源：Wind，东证衍生品研究院

### 2.2.3、主成分分析

通过主成分分析，我们可观察是否存在一些变量比其他变量包含的信息更多。

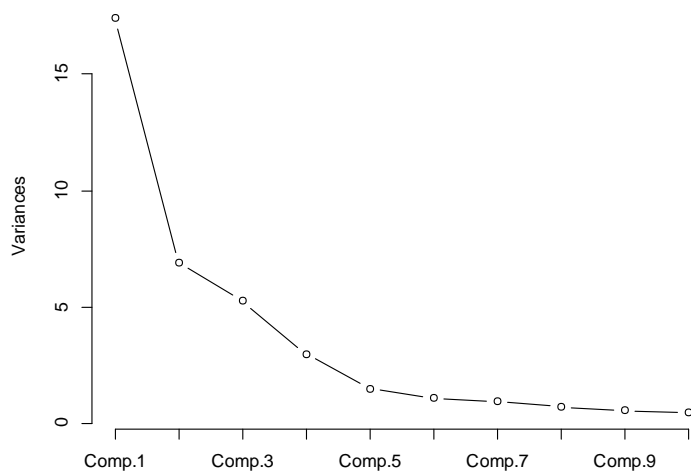
主成分分析结果表明 Comp.1、Comp.2、Comp.3、Comp.4、Comp.5 包含了自变量大部分的信息。主成分碎石图反映 Comp.5 之后的曲线趋向平缓。因此，我们可以从 Comp.1、Comp.2、Comp.3、Comp.4、Comp.5 查看自变量的贡献度。

图表 7：主成分分析结果

components	Standard deviation	Proportion of Variance	Cumulative Proportion
Comp.1	4.170	0.435	0.435
Comp.2	2.626	0.172	0.607
Comp.3	2.294	0.132	0.739
Comp.4	1.730	0.075	0.814
Comp.5	1.222	0.037	0.851
Comp.6	1.047	0.027	0.878
Comp.7	0.985	0.024	0.902
Comp.8	0.857	0.018	0.921
Comp.9	0.751	0.014	0.935
Comp.10	0.685	0.012	0.947
Comp.11	0.616	0.009	0.956
Comp.12	0.548	0.007	0.964
Comp.13	0.477	0.006	0.969
Comp.14	0.461	0.005	0.975
Comp.15	0.424	0.004	0.979
Comp.16	0.409	0.004	0.983
Comp.17	0.351	0.003	0.986
Comp.18	0.315	0.002	0.989
Comp.19	0.297	0.002	0.991
Comp.20	0.280	0.002	0.993

资料来源：Wind，东证衍生品研究院

图表 8：主成分的碎石图



资料来源：Wind，东证衍生品研究院

从 Comp.1、Comp.2、Comp.3、Comp.4、Comp.5 中各自变量的贡献度可以看出，并没有贡献度较为突出的变量，因此通过主成分分析难以找出模型的因子。

图表 9：主成分分析中贡献度

Items	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
v1	-0.229	0.107	0.108		
v2	-0.227	0.116	0.137	-0.105	
v3	-0.232	-0.165	0.106	-0.195	-0.105
v4	-0.234				
v5	-0.156	-0.215	-0.123	0.103	0.151
v6	-0.221	0.167	0.208	0.17	-0.156
v7	-0.282	0.205	0.146	-0.137	-0.153
v8	-0.165	-0.228	0.185	0.320	0.431
v9	-0.117	-0.235	-0.169	-0.140	0.152
v10	-0.185	-0.126	0.694	0.312	-0.319
v11	0.124	0.266	-0.331	-0.529	-0.302
v12	-0.153	0.196	-0.158	-0.169	0.110
v13	-0.209	-0.206	0.231	-0.165	-0.131
v14	-0.122	-0.232	0.130	0.185	0.275
v15	-0.272	0.279	-0.406	0.193	0.171
v16	-0.279	0.253	-0.116	-0.154	-0.142
v17	-0.151	-0.261	-0.154	-0.129	0.299

v18	-0.217	0.391	0.352	-0.155	0.261
v19	-0.227	-0.121	-0.151	0.198	
v20	-0.113	0.109	0.409	-0.270	-0.154
v21	-0.388	0.157	-0.112	-0.125	-0.147
v22	-0.189	0.109	-0.199	-0.105	-0.129
v23	-0.403	0.119	-0.114	-0.111	-0.100
v24	0.225	-0.180	0.119	0.109	0.143
v25	-0.213	0.106	0.203	-0.108	0.102
v26	0.183	0.147	-0.170	-0.200	0.109
v27	0.326	-0.125	0.115	-0.134	0.149
v28	0.355	0.116	-0.144	-0.126	0.106
v29	-0.107	0.252	-0.238	-0.105	0.196
v30	0.181	0.114	0.180	0.490	-0.120
v31	-0.217	0.139	0.106	0.111	0.138
v32	-0.217	0.139	0.106	0.111	0.138
v33	-0.135	-0.24	0.399	0.257	-0.205
v34	0.233	0.207	0.223	0.197	-0.197
v35	0.179	0.169	-0.119	0.138	0.154
v36	0.208	0.127	0.163	0.236	-0.173
v37	0.219	0.102	0.16	0.188	-0.149
v38	0.221	0.147	0.183	-0.203	-0.103
v39	-0.281	0.312	-0.168	0.100	-0.180
v40	0.200	-0.154	0.161	-0.187	0.113

资料来源：Wind，东证衍生品研究院

### 3、实证研究

Complete Subset Regressions 即全子集回归，可以对所有可能的线性回归模型进行检验。全子集回归可以选择展示所有可能的结果，也可以展示  $n$  个不同子集大小的最佳模型。

实证研究主要分为以下几个阶段：

- 第一、通过 CSR 得到局部最优的回归模型，进行模型因子的初选；
- 第二、利用初选的因子建模，将得到方程组进行样本内和样本外的测试；
- 第三、针对方程组内各方程形成子策略，检验子策略和策略集合的收益风险特征。

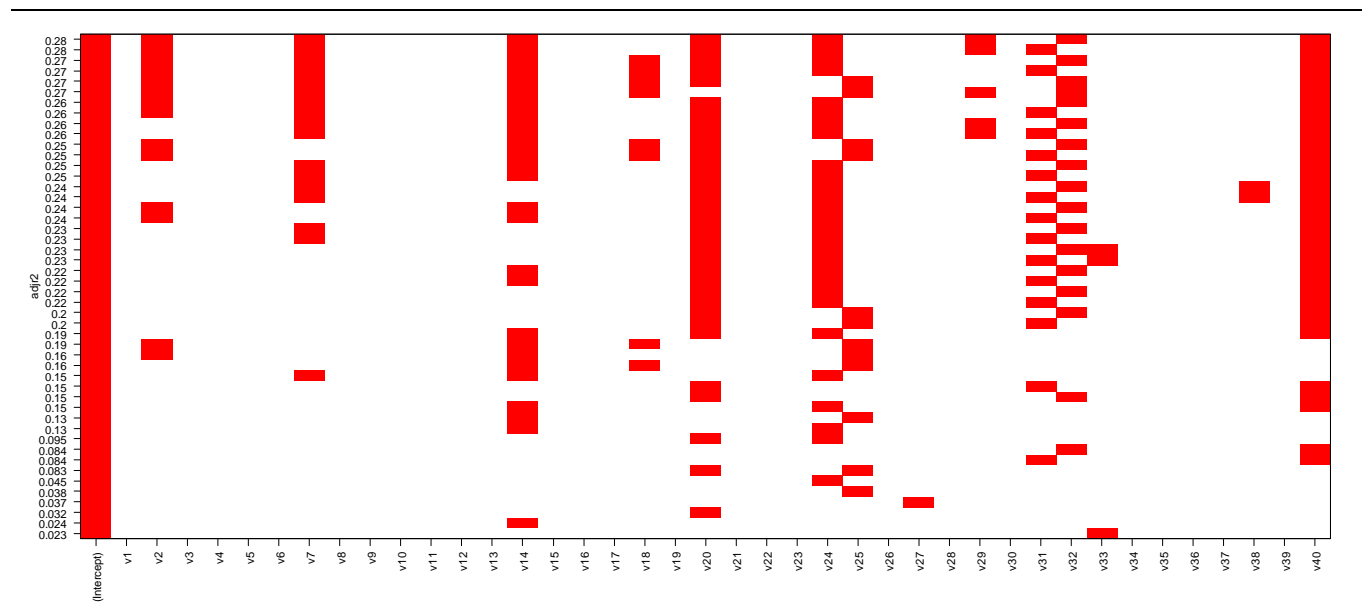
### 3.1、利用 CSR 进行因子的筛选

CSR 实证部分具体进行自变量和因变量的回归分析。

首先我们利用滞后一期的食品价格类高频数据和其他月频数据来对下一期的 10 年期国债收益率变动进行回归。利用 CSR 方法分析，自变量调整为变化率的形式，因变量调整为差分的形式。

当设定 nbest=6 时，我们发现调整的 R 方最大在 0.3 左右，48 个回归方程显示 v2、v7、v14、v18、v20、v24、v25、v27、v29、v31、v32、v33、v38、v40 是影响国债收益率变化的重要变量。下图所示，横坐标为各自变量的代号，纵坐标对应不同回归方程的调整的 R 方，红色方块展示了不同回归方程内的自变量。

图表 10：CSR 实证结果 (nbest=6)



资料来源：Wind，东证衍生品研究院

从上图可以看出，各回归方程中的因子重合度较高。直观地看，v2、v7、v14、v20、v24、v25、v31、v32、v40 是多个回归方程的因子。

图表 11：CSR 实证结果 (nbest=6, v1-v35)

6 subsets of each size up to 8																																			
Selection Algorithm: exhaustive																																			
	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10	v11	v12	v13	v14	v15	v16	v17	v18	v19	v20	v21	v22	v23	v24	v25	v26	v27	v28	v29	v30	v31	v32	v33	v34	v35
1 ( 1 )																									*										
1 ( 2 )																									*										
1 ( 3 )																										*									
1 ( 4 )																				*								*							
1 ( 5 )														*							*														
1 ( 6 )																																			*
2 ( 1 )														*												*									
2 ( 2 )														*											*										
2 ( 3 )																				*					*										

2	( 4 )							*
2	( 5 )							*
2	( 6 )				*		*	
3	( 1 )	*			*		*	
3	( 2 )				*		*	
3	( 3 )		*		*		*	
3	( 4 )				*			*
3	( 5 )				*			*
3	( 6 )			*			*	
4	( 1 )				*		*	*
4	( 2 )				*		*	*
4	( 3 )				*		*	*
4	( 4 )				*		*	*
4	( 5 )			*	*		*	
4	( 6 )	*		*	*		*	
5	( 1 )		*		*		*	*
5	( 2 )		*		*		*	*
5	( 3 )				*		*	*
5	( 4 )				*		*	*
5	( 5 )			*	*		*	*
5	( 6 )			*	*		*	*
6	( 1 )		*		*		*	*
6	( 2 )		*		*		*	*
6	( 3 )		*		*		*	*
6	( 4 )		*		*		*	*
6	( 5 )	*		*	*		*	*
6	( 6 )	*		*	*		*	*
7	( 1 )	*	*		*		*	*
7	( 2 )	*	*		*		*	*
7	( 3 )		*		*		*	*
7	( 4 )		*		*		*	*
7	( 5 )	*		*	*		*	*
7	( 6 )	*		*	*		*	*
8	( 1 )	*	*		*		*	*
8	( 2 )	*	*		*		*	*
8	( 3 )	*	*		*		*	*
8	( 4 )	*	*		*		*	*
8	( 5 )	*	*		*		*	*
8	( 6 )	*	*		*		*	*

资料来源：Wind，东证衍生品研究院

为了尽可能避免遗漏重要的变量,我们设定  $n_{best}=10$  时,发现调整的 R 方最大仍然在 0.3 左右,但是重要的自变量范围有所扩大,  $v_2$ 、 $v_7$ 、 $v_{12}$ 、 $v_{14}$ 、 $v_{18}$ 、 $v_{19}$ 、 $v_{20}$ 、 $v_{24}$ 、 $v_{25}$ 、 $v_{27}$ 、 $v_{29}$ 、 $v_{31}$ 、 $v_{32}$ 、 $v_{33}$ 、 $v_{36}$ 、 $v_{38}$ 、 $v_{40}$  是影响国债收益率波动的重要变量。从重要变量的逻辑看,投资、通货膨胀、流动性、资金策略和银行负债结构对于国债收益率的影响更显著。

通过 CSR 分析后, 自变量经过进一步地筛选, 得到的变量明显减少, 然后在此基础上建立预测模型。

图表 12: CSR 实证结果 (nbest=6, v36-v40)

	v36	v37	v38	v39	v40
2 ( 4 )					*
2 ( 5 )					*
3 ( 4 )					*
3 ( 5 )					*
3 ( 6 )					*
4 ( 1 )					*
4 ( 2 )					*
4 ( 3 )					*
4 ( 4 )					*

4 ( 5 )	*
5 ( 1 )	*
5 ( 2 )	*
5 ( 3 )	*
5 ( 4 )	*
5 ( 5 )	*
5 ( 6 )	*
6 ( 1 )	*
6 ( 2 )	*
6 ( 3 )	*
6 ( 4 )	*
6 ( 5 )	*
6 ( 6 )	*
7 ( 1 )	*
7 ( 2 )	*
7 ( 3 )	*
7 ( 4 )	*
7 ( 5 )	*
7 ( 6 )	*
8 ( 1 )	*
8 ( 2 )	*
8 ( 3 )	*
8 ( 4 )	*
8 ( 5 )	*
8 ( 6 )	*

资料来源：Wind，东证衍生品研究院

图表 13：CSR 初选重要指标

项目	指标	备注
CSR 初选重要指标	制造业固定资产投资完成额累计同比 (v2)	指标更新时间 最迟在下月中 旬，另外频次 最高为日频， 其次为周频和 月频。
	房屋新开工面积累计同比 (v7)	
	PMI 产成品库存 (v12)	
	统计局口径 50 个城市平均价：一级散装菜籽油 (v14)	
	社会融资规模累计同比 (v18)	
	金融机构短期贷款余额同比 (v19)	
	金融机构各项存款余额同比 (v20)	
	美元指数 (v24)	
	欧元兑美元 (v25)	
	官方储备资产中外汇储备 (v27)	
	公共财政支出累计同比 (v29)	
	中小型存款类金融机构人民币存款准备金率(月) (v31)	
	大型存款类金融机构人民币存款准备金率(月) (v32)	
	商业银行国债托管量同比 (v33)	



(中债 10 年国债收益率-中债 10 年商业银行普通债到期收益率(AAA))(v36)

(中债 10 年国债收益率-中债 10 年铁道债到期收益率) (v38)

非金融企业活期存款/非金融企业及机关团体各项贷款 (v40)

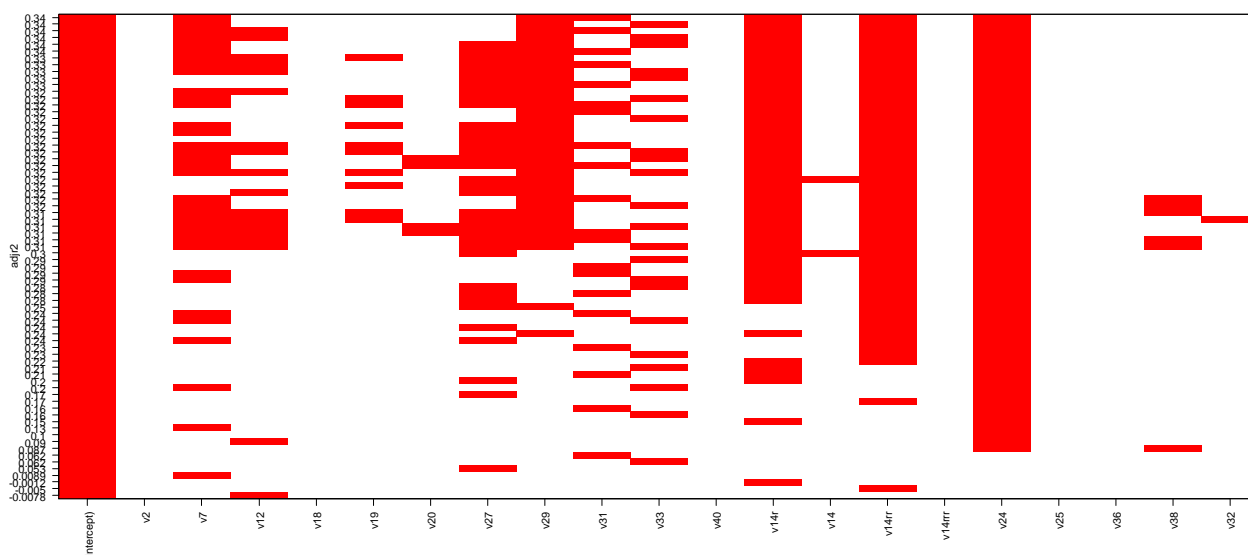
资料来源: Wind, 东证衍生品研究院

### 3.2、国债 Beta 模型的样本内外测试

从 CSR 初选指标的更新频次和时间来看,除了高频指标外,其他的月度指标都在下月 20 日前发布,因此可以将预测的起点定在每月 20 日,预测区间为每月 20 日后的 20 个自然日。

CSR 样本内测试的结果显示,调整的 R 方最高在 0.34 左右,72 个回归方程中部分方程调整的 R 方在较低的水平,因此需要进一步筛选回归方程。另外,涉及到通胀监测的高频数据,我们将和 v14 类似的 v14r、v14rr、v14rrr 高频指标加入到回归测试的范围中,具体包括 50 个城市豆腐平均价、50 个城市标准粉平均价、50 个城市 5L 桶装大豆油平均价。

图表 14: CSR 样本内实证结果 (30 个月, nbest=8, 月末预测起点)



资料来源: Wind, 东证衍生品研究院

72 个回归方程构成的方程组中部分方程的拟合效果并不好,而且回归系数的显著性也存在差异,因此需要进一步的筛选,筛选的标准参考调整的 R 方、AIC 和 P 值。

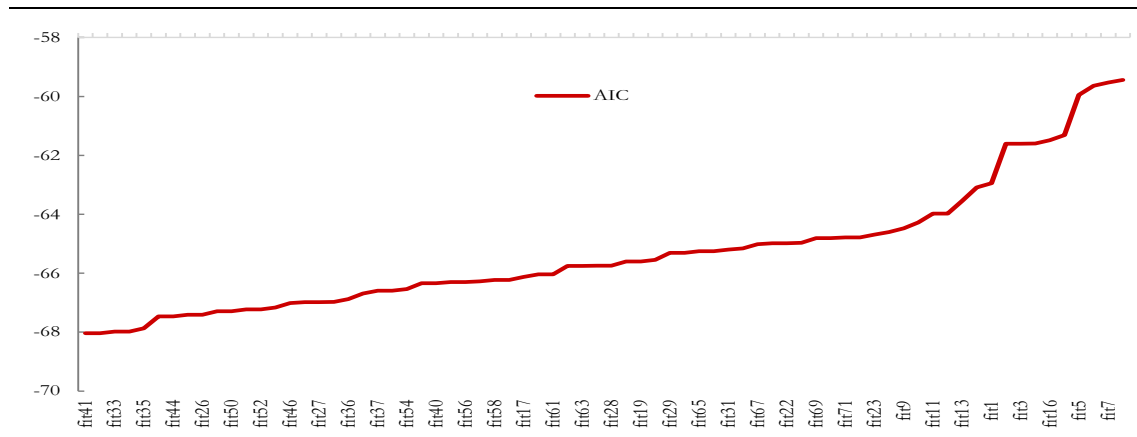
另外,针对一元回归和多元回归,根据回归方程中因子数的大小进行横向的比较。比如,三元线性回归的方程不与一元线性回归的方程进行比较。

图表 15: CSR 样本内实证结果 (30 个月, nbest=8, 月末预测起点)

8 subsets of each size up to 9																				
Selection Algorithm: exhaustive																				
	v2	v7	v12	v18	v19	v20	v27	v29	v31	v32	v33	v40	v14r	v14	v14rr	v14rrr	v24	v25	v36	v38
1 (1)																*				
1 (2)										*										
1 (3)									*											
1 (4)							*													
1 (5)	*																			
1 (6)												*								
1 (7)														*						
1 (8)		*																		
2 (1)						*										*				
2 (2)														*		*				
2 (3)										*						*				
2 (4)									*							*				
2 (5)												*				*				
2 (6)	*															*				
2 (7)		*														*				
2 (8)																*				*
3 (1)						*								*		*				
3 (2)									*					*		*				
3 (3)										*				*		*				
3 (4)											*			*		*				
3 (5)									*		*			*		*				
3 (6)									*		*			*		*				
3 (7)						*				*		*		*		*				
3 (8)	*								*		*			*		*				
4 (1)									*		*		*	*		*				
4 (2)									*		*		*	*		*				
4 (3)						*					*		*	*		*				
4 (4)						*	*				*		*	*		*				
4 (5)	*								*		*		*	*		*				
4 (6)	*								*		*		*	*		*				
4 (7)						*				*		*	*	*		*				
4 (8)	*					*			*		*		*	*		*				
5 (1)						*		*	*		*		*	*		*				
5 (2)						*	*		*	*		*	*	*		*				
5 (3)						*	*		*	*		*	*	*		*				
5 (4)						*	*		*	*		*	*	*		*				
5 (5)	*					*	*		*	*		*	*	*		*				
5 (6)	*					*	*		*	*		*	*	*		*				
5 (7)						*	*		*	*		*	*	*		*				
5 (8)						*	*		*	*		*	*	*		*				
6 (1)	*					*	*		*	*		*	*	*		*				
6 (2)	*					*	*		*	*		*	*	*		*				
6 (3)						*	*	*	*	*		*	*	*		*				
6 (4)						*	*	*	*	*		*	*	*		*				
6 (5)	*					*	*	*	*	*		*	*	*		*				
6 (6)						*	*	*	*	*		*	*	*		*				
6 (7)				*		*	*	*	*	*		*	*	*		*				
6 (8)		*				*	*	*	*	*		*	*	*		*				
7 (1)	*	*				*	*	*	*	*		*	*	*		*				
7 (2)	*	*				*	*	*	*	*		*	*	*		*				
7 (3)	*					*	*	*	*	*		*	*	*		*				
7 (4)	*					*	*	*	*	*		*	*	*		*				
7 (5)	*	*				*	*	*	*	*		*	*	*		*				
7 (6)	*			*		*	*	*	*	*		*	*	*		*				
7 (7)	*			*		*	*	*	*	*		*	*	*		*			*	
7 (8)	*			*		*	*	*	*	*		*	*	*		*			*	
8 (1)	*	*	*	*		*	*	*	*	*		*	*	*		*				*
8 (2)	*	*				*	*	*	*	*		*	*	*		*				
8 (3)	*	*				*	*	*	*	*		*	*	*		*				
8 (4)	*		*			*	*	*	*	*		*	*	*		*				
8 (5)	*		*			*	*	*	*	*		*	*	*		*				
8 (6)	*				*	*	*	*	*	*		*	*	*		*				
8 (7)	*				*	*	*	*	*	*		*	*	*		*				
8 (8)	*	*	*		*	*	*	*	*	*		*	*	*		*				
9 (1)	*	*	*	*	*	*	*	*	*	*		*	*	*		*				
9 (2)	*	*	*	*	*	*	*	*	*	*		*	*	*		*				
9 (3)	*	*	*	*	*	*	*	*	*	*		*	*	*		*			*	
9 (4)	*	*	*	*	*	*	*	*	*	*		*	*	*		*				*
9 (5)	*	*	*	*	*	*	*	*	*	*		*	*	*		*				
9 (6)	*	*	*	*	*	*	*	*	*	*		*	*	*		*				
9 (7)	*	*	*	*	*	*	*	*	*	*		*	*	*		*			*	
9 (8)	*	*	*	*	*	*	*	*	*	*		*	*	*		*			*	

资料来源: Wind, 东证衍生品研究院

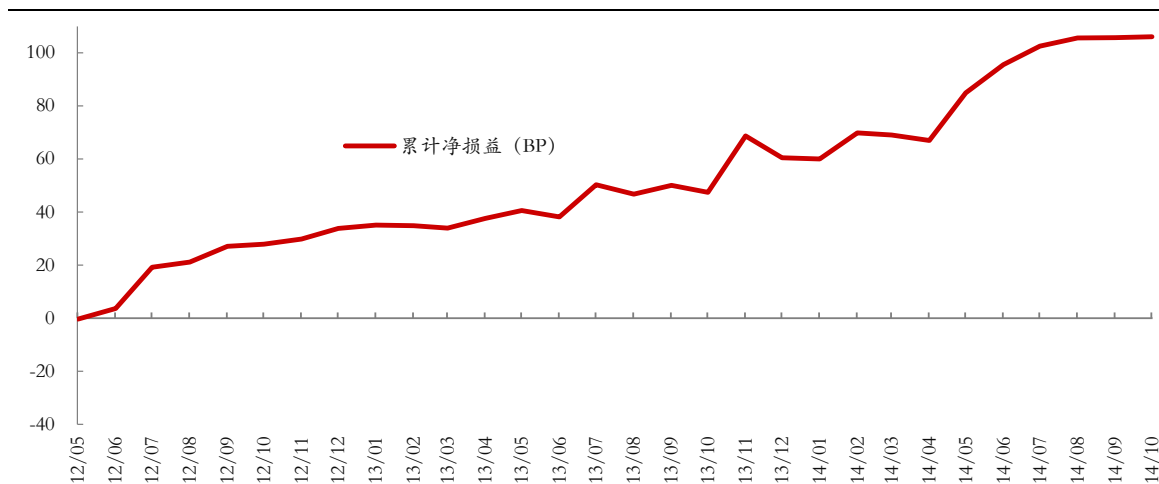
图表 16: CSR 样本内 72 个回归方程 AIC 值 (30 个月, nbest=8, 月末预测起点)



资料来源: Wind, 东证衍生品研究院

经过回归方程的筛选, 我们选取其中的 19 个回归方程建立组合策略, 样本内测试时均匀分配各个策略的资金规模, 样本内测试时间为 30 个月, 另外不考虑止损止盈, 直接计算测试期内的 10 年期国债到期收益率的变动, 模型通过当月 20 日之前公布的最新数据预测当月 20 日以后 20 个自然日的利率涨跌。

图表 17: CSR 样本内 19 个子策略组合累计净损益 (30 个月, nbest=8, 月末预测起点)

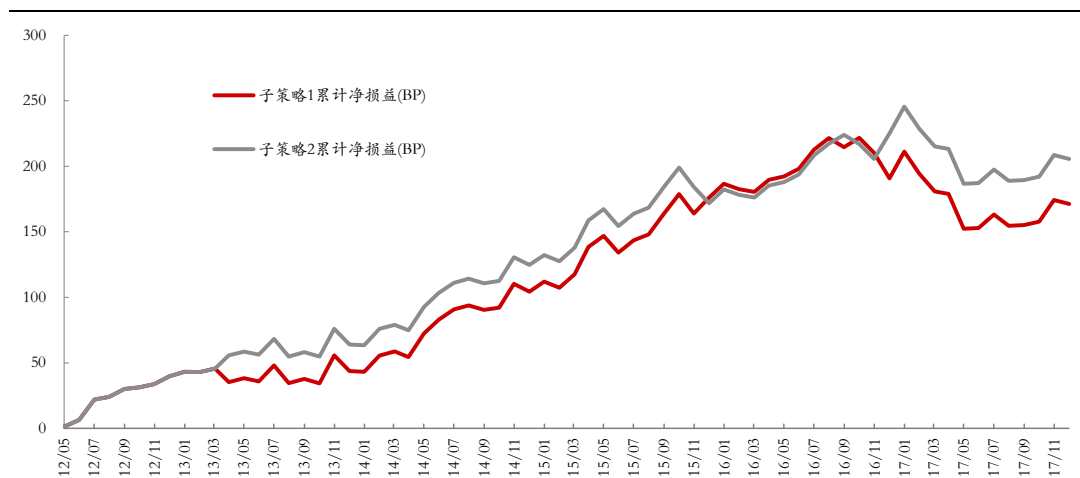


资料来源: Wind, 东证衍生品研究院

19 个回归方程的测试结果表明部分方程测试的结果重合度较高, 影响回归的因子主要集中在 v27、v29、v40、v14、v14rrr。然而, 再利用 19 个回归方程进行样本外的回测, 回测结果并不理想, 为此我们进一步对回归方程进行了筛选, 除了回归方程自身的特点之外, 还会考虑自变量包含的信息, 剔除结果明显趋同的回归方程, 选取了两个回归方程并纳入到最终的预测方程组中。

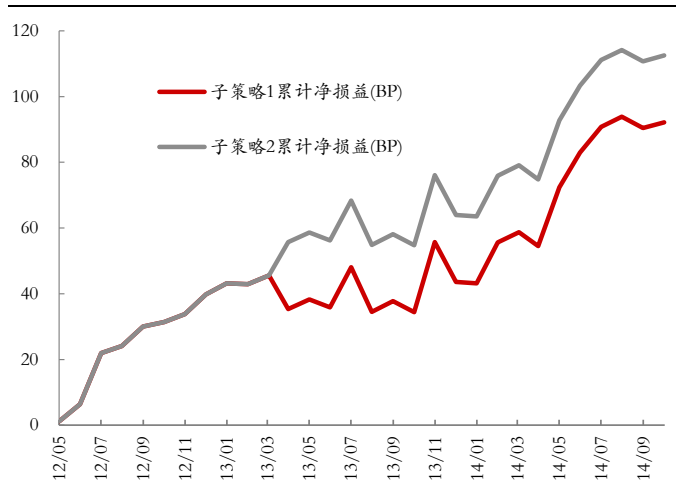
将优选后的两个回归方程作为两个子策略，利用样本内回测的模型参数作为预测模型的参数，预测样本外 38 个月的结果，样本内和样本外两个子策略测试的累计净损益如下图所示。

图表 18: CSR 样本内和样本外 2 个子策略累计净损益 (68 个月, 月末预测起点, 无止盈止损)



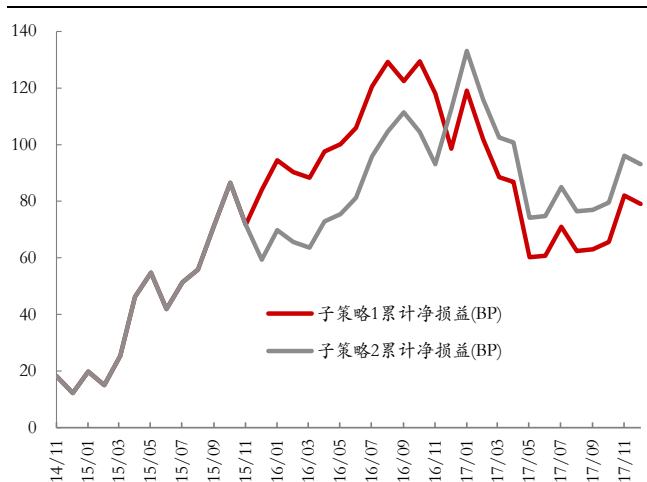
资料来源: Wind, 东证衍生品研究院

图表 19: CSR 样本内 2 个子策略累计净损益 (30 个月)



资料来源: Wind, 东证衍生品研究院

图表 20: CSR 样本外 2 个子策略累计净损益 (38 个月)



资料来源: Wind, 东证衍生品研究院

我们的预测模型每月 20 日发出信号，可以归为宏观预测模型，预测胜率是考察模型的重要方面，因此并未设置回测的止盈止损，具体测试结果的量化指标如下表所示。

图表 21：优选子策略样本内和样本外测试结果

策略	样本内	样本外	胜率	年化收益 (BP)	累计收益 (BP)
子策略 1	2012 年 5 月至 2014	2014 年 11 月至	65%	30	171
子策略 2	年 10 月	2017 年 12 月	66%	36	206

资料来源：Wind，东证衍生品研究院

官方储备资产中外汇储备、公共财政支出累计同比、非金融企业活期存款/非金融企业及机关团体各项贷款、统计局口径 50 个城市一级散装菜籽油和 5L 桶装大豆油平均价是优选后的子策略的因子。经济增长、通货膨胀、流动性以及银行负债结构的信息都包含在模型中，30 个月的样本内回测和 38 个月的样本外测试表明模型具备一定的适应性和预测性。

从样本内和样本外测试结果，可以发现一个突出的问题，两个子策略在 2017 年上半年出现了较为明显的回撤，我们认为导致模型误判的主要原因是金融监管的信息未反映在模型中，基本面已经不能充分解释此期间的利率波动。关于金融监管，我们在《基本面量化之一：国债的重要因子》中已有一些分析，难以量化且替代性指标滞后是监管信息不易建模的重要原因，但是考虑到监管对于国债的影响较大，我们会在后续的研究进一步深入探讨政策监管对于国债的影响。

### 3.3、国债 Beta 策略的跟踪方案

作为宏观预测模型，国债 Beta 策略模型的跟踪方案分为两种，一种是通过优选后的两个子策略预测国债收益率的变动，二是通过 19 个回归方程组成的方程组预测国债收益率的变动。

优选后的两个子策略可以通过既定的模式不断更新预测结果，而 19 个回归方程组成的方程组既可以监控单个方程的预测结果，也可以主观对不同的回归方程进行赋权，对影响下一期利率变动的重要因素赋予较大的权重。

## 4、结论及展望

《基本面量化之一：国债的重要因子》中，我们选取了 58 个重要因子，通过这些因子可以对当前国债的基本面有个整体的把握。《基本面量化之二：国债的 Beta》中，我们将这 58 个重要因子作为研究的初始数据，进行了如下的实证研究：

- 第一、根据因子公布的时间和频率，筛选出符合建模需求的因子，共有 40 个因子；
- 第二、利用相关性分析以及主成分分析对数据进行初步的分析，国债收益率具备预测的可行性；
- 第三、通过 Complete Subset Regressions (CSR) 方法初步筛选模型因子、建立方程组、构建预测模型并形成子策略进行样本内和样本外的测试；

第四、根据宏观预测的特性，我们推出预测模型自动跟踪和主观赋权跟踪两种方式来自模拟交易的结果。

实证研究的结果表明：

第一，包含经济增长、通货膨胀、流动性和银行负债结构信息的预测模型能够获得 65% 左右的胜率；

第二，30 个月的样本内回测和 38 个月的样本外测试表明模型具备一定的适应性和预测性；

第三，多个子策略的模式能够在市场出现拐点时，避免出现净值大幅波动的情形。

两个子策略在 2017 年上半年出现了较为明显的回撤，我们认为导致模型误判的主要原因是金融监管的信息未反映在模型中，我们会在后续的研究进一步深入探讨政策监管对于国债的影响。

关于宏观预测的研究，一直是市场争论的焦点，而且多数模型在后续跟踪阶段的预测能力下降，因此，我们可能会尝试建立动态的预测模型并充分运用子策略的模式进行测试跟踪。接下来，我们将基于基本面的 Beta 策略推广至商品期货和外汇，继续秉承“基本面为主，量化为辅”的基本面量化研究的思路。

## 5、参考文献

- [1] Nektarios Aslanidis, Charlotte Christiansen and Andrea Cipollini (2017). Predicting Bond Betas using Macro-Finance Variables Working Paper.
- [2] Barillas, F. (2015). Can we exploit predictability in bond markets? Working Paper.
- [3] Bauer, D. M. and J. D. Hamilton (2016). Robust bond risk premia. Working Paper.
- [4] Cochrane, J. H. and M. Piazzesi (2005). Bond risk premia. American Economic Review 95(1), 138-160.
- [5] Elliott, G., A. Gargano, and A. Timmermann (2013). Complete subset regressions. Journal of Econometrics 177(2), 357-373.
- [6] Ludvigson, S. C. and S. Ng (2009). Macro factors in bond risk premia. Review of Financial Studies 22(12), 5027-5067.
- [7] Sarno, L., P. Schneider, and C. Wagner (2016). The economic value of predicting bond risk premia. Journal of Empirical Finance 37, 247 -267.
- [8] Ludvigson, S., and S. Ng(2010): “A Factor Analysis of Bond Risk Premia” in Handbook of Empirical Economics and Finance, ed. by A. Uhla, and D. E. A. Giles, pp. 313-372. Chapman and Hall, Boca Raton, FL.

[9]Altavilla, C., R. Giacomini, and R. Costantini (2014). Bond returns and market expectations.  
Journal of Financial Econometrics 12(4), 708-729.



### 期货走势评级体系（以收盘价的变动幅度为判断标准）

走势评级	短期（1-3 个月）	中期（3-6 个月）	长期（6-12 个月）
强烈看涨	上涨 15%以上	上涨 15%以上	上涨 15%以上
看涨	上涨 5-15%	上涨 5-15%	上涨 5-15%
震荡	振幅-5%-+5%	振幅-5%-+5%	振幅-5%-+5%
看跌	下跌 5-15%	下跌 5-15%	下跌 5-15%
强烈看跌	下跌 15%以上	下跌 15%以上	下跌 15%以上

### 上海东证期货有限公司

上海东证期货有限公司（简称东证期货）是东方证券股份有限公司全资子公司，注册资本达 15 亿元，系国内四家期货交易所的结算会员。

东证期货专注于金融期货和商品期货的研究与服务，提供权威、及时的研发产品服务和投资策略；专注于信息技术的创新，创建安全、快捷的交易通道，开发多样化、个性化的交易系统；专注于构筑全面的风险管理和客户服务平台。

东证期货管理团队管理经验丰富，业绩出众，在业内享有盛誉。人才管理及激励机制完善，公司拥有硕士学历以上人员占比 30%，具有海外证券和期货经历的高端人才占比 10%。

2010 年，东证期货发展迅猛，成绩斐然，成为业内进步最快、最受瞩目的期货公司之一。2011 年初，东证期货荣获 2010 年度中国金融期货交易所年度会员金奖，同时获投资者教育奖、客户管理奖、技术管理奖和功能发挥奖等四项单项大奖；荣获上海期货交易所优胜会员第七名，铜、橡胶和燃料油三项企业服务奖；荣获大连商品交易所优秀会员第九名；东证衍生品研究院（原东证期货研究所）荣获大连商品交易所、和讯网第二届全国“十大期货研发团队”农产品团队全国第二名、化工团队全国第五名；荣获郑州商品交易所行业进步奖等。

东证期货全年无风险事故，充分体现了公司稳健经营，稳步发展的经营宗旨。

## 分析师承诺

### 章顺

本人具有中国期货业协会授予的期货执业资格或相当的专业胜任能力，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接收到任何形式的报酬。

## 免责声明

本报告由上海东证期货有限公司（以下简称“本公司”）制作及发布。

本研究报告仅供本公司的客户使用。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。

本研究报告是基于本公司认为可靠的且目前已公开的信息撰写，本公司力求但不保证该信息的准确性和完整性，客户也不应该认为该信息是准确和完整的。同时，本公司不保证文中观点或陈述不会发生任何变更，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。本公司会适时更新我们的研究，但可能会因某些规定而无法做到。除了一些定期出版的报告之外，绝大多数研究报告是在分析师认为适当的时候不定期地发布。

在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，也没有考虑到个别客户特殊的投资目标、财务状况或需求。客户应考虑本报告中的任何意见或建议是否符合其特定状况，若有必要应寻求专家意见。本报告所载的资料、工具、意见及推测只提供给客户作参考之用，并非作为或被视为出售或购买投资标的的邀请或向人作出邀请。

在任何情况下，本公司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，投资者需自行承担风险。

本报告主要以电子版形式分发，间或也会辅以印刷品形式分发，所有报告版权均归本公司所有。未经本公司事先书面授权，任何机构或个人不得以任何形式复制、转发或公开传播本报告的全部或部分内容，不得将报告内容作为诉讼、仲裁、传媒所引用之证明或依据，不得用于营利或用于未经允许的其它用途。

如需引用、刊发或转载本报告，需注明出处为东证衍生品研究院，且不得对本报告进行任何有悖原意的引用、删节和修改。

---

## 东证衍生品研究院

地址：上海市中山南路 318 号东方国际金融广场 2 号楼 22 楼

联系人：梁爽

电话：8621-63325888-1592

传真：8621-33315862

网址：[www.orientfutures.com](http://www.orientfutures.com)

Email：[research@orientfutures.com](mailto:research@orientfutures.com)