

DFQ 市场化国债&国开利率曲线

——宏观固收量化研究系列之（四）

东方证券
ORIENT SECURITIES

报告发布日期

2021 年 06 月 05 日

证券分析师

朱剑涛

021-63325888*6077

zhujiantao@orientsec.com.cn

执业证书编号：S0860515060001

研究结论

- 到期收益率在投资交易中最常用，但无法给债券定价，定价需要用即期收益率曲线。本文推荐用 kernel 加权的方法从付息债券价格中拟合反推即期收益率曲线。
- 国开债跟国债信用等级都接近无风险，但前者存在流动性和税收的溢价。根据国开利率曲线 = 国债利率曲线 + 溢价曲线的拆分等式，对国债利率曲线和溢价曲线用 kernel 方法建模，并把国债、国开成交数据糅合在一起同时估计两条利率曲线，可以大幅减少模型参数数量，提升参数估计准确性。
- 中债国债、国开利率曲线在估算过程中用到很多非市场交易的价格数据，给金融产品提供稳定估值的同时，牺牲了市场定价精度。本文构建的利率曲线是纯市场化曲线，只用银行间和交易所的国债、国开成交价格数据。
- 利率曲线模型参数估计是非凸优化问题，我们针对性的设计了一套技术方案，包括：Pytorch + Adam 优化算法、主成分控制模型自由度、回购利率辅助设定利率曲线初始值，引入机器学习里的交叉验证思想避免过拟合。得到的两条无风险利率曲线分别称作：**DFQ 国债利率曲线**和**DFQ 国开利率曲线**
- 2012 年之前市场有成交的券数量较少，完全市场化的 DFQ 利率曲线波动很大，2012 年之后开始稳定，十年期以下的即期利率和中债基本一致；十年期以上的超长端，由于价格数据少，市场化方法很难得到稳定利率曲线。
- 对于 10 年期以下的国开债，DFQ 利率曲线给的定价比中债更接近市场成交价，定价误差平均比中债小 15%；分年统计，DFQ 利率曲线相对中债的定价误差降幅在 2019 年时达到最高值 25.4%。
- 用 DFQ 国开利率曲线识别的收益率高估或低估的券，相对中债而言，溢价幅度更大，套利收益更稳定。按照 10bps 阈值，做多收益率高估券，做空收益率低估券的多空套利组合，在 DFQ 利率曲线下可获得 70bps 的收益，而在中债利率曲线下只有 30bps 不到。
- 近些年随着债券市场的发展成熟，流动性改善，到期收益率高估的国开债数量明显减少，平均溢价幅度降低，且溢价衰减速度加快，投资者需要工具辅助尽早发现高估券，尽快交易锁定溢价，而 DFQ 利率曲线是一个可实盘实时计算的模型，适合作为优先考虑的工具方案。

风险提示

- 量化模型失效风险
- 市场极端环境的冲击

东方证券股份有限公司经相关主管机关核准具备证券投资咨询业务资格，据此开展发布证券研究报告业务。

东方证券股份有限公司及其关联机构在法律许可的范围内正在或将要与本研究报告所分析的企业发展业务关系。因此，投资者应当考虑到本公司可能存在对报告的客观性产生影响的利益冲突，不应视本证券研究报告为作出投资决策的唯一因素。

有关分析师的申明，见本报告最后部分。其他重要信息披露见分析师申明之后部分，或请与您的投资代表联系。并请阅读本证券研究报告最后一页的免责申明。

点击进入 <http://www.hibor.com.cn>

目 录

一、利率曲线构建.....	4
1.1 即期收益率曲线常用构建方法	4
1.2 国债 & 国开债联合估计	7
1.3 模型参数估计方法	8
二、DFQ 市场化利率曲线	10
2.1 利率曲线形态比较	10
2.2 定价偏差比较	13
2.3 寻找收益率低估与高估券.....	14
2.4 利率曲线的无套利性.....	17
2.5 利率曲线隐含的未来信息.....	18
三、总结	18
风险提示.....	19
参考文献.....	19

图表目录

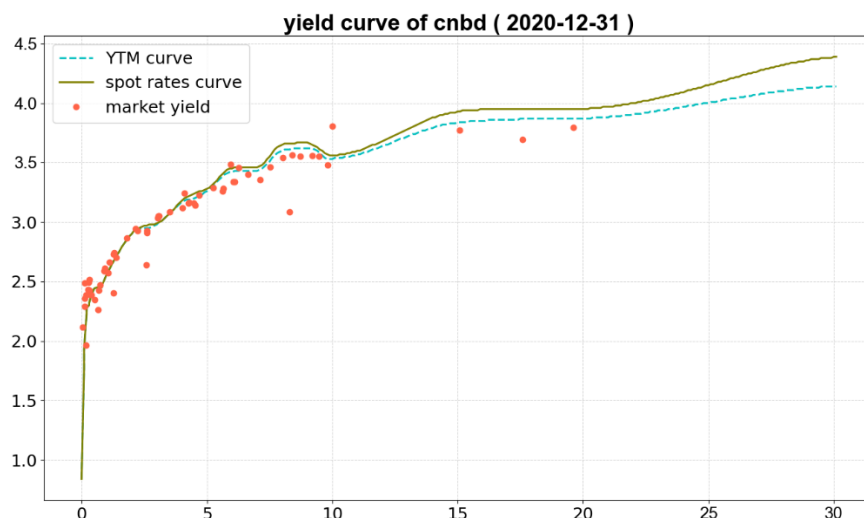
图 1: 到期收益率曲线、即期收益率曲线与债券到期收益率（国开债，2020.12.31）	4
图 2: NS 模型对中债国债即期利率曲线的拟合(20130620)	5
图 3: NS 模型对中债国债即期利率曲线的拟合(20201126)	5
图 4: 三次样条法与 Kernel 方法示例（2020.12.31）	5
图 5: bandwidth 分段线性函数，x 轴为期限	6
图 6: 中债国债利率曲线成分券各个剩余期限区间的券数量	7
图 7: 中债国债利率曲线成分券不同价格数据来源的数量占比	7
图 8: 引入回购利率数据辅助确认利率曲线初始值（国债，2013-06-20）	9
图 9: 利率曲线的欠拟合与过拟合（国开，2015-10-27）	9
图 10: 国债+国开债有市场成交的券数量	10
图 11: 国债+国开债有市场成交的券的剩余期限占比	10
图 12: 国债&国开债历史平均利率曲线（2012.01.01 – 2021.04.21）	11
图 13: 国开债利率走势（2007.01.04 – 2021.04.21）	11
图 14: 超长端国开债利率走势（2007.01.04 – 2021.04.21）	12
图 15: 不同期限利率的年化波动率（2012.01.04 – 2021.04.21）	13
图 16: DFQ 利率曲线与中债利率曲线的定价偏差（2012.01.04 – 2021.04.21）	13
图 17: DFQ 国开曲线与中债国开曲线的定价偏差（bps，2012.01.04 – 2021.04.21）	14
图 18: 多空组合的持有收益（bps，2012.01.04 – 2021.04.21）	15
图 19: 每个交易日收益率高估券的数量（2012.01.04 – 2021.04.21）	15
图 20: 到期收益率高估的国开债未来一个月的溢价衰减样本数据(bps)	16
图 21: 不同高估程度的国开债收益率溢价衰减曲线（bps，2012.01.04 – 2021.04.21）	16
图 22: 分年统计的国开债收益率溢价衰减曲线	17

一、利率曲线构建

1.1 即期收益率曲线常用构建方法

债券的到期收益率（YTM, yield to maturity）在债券分析中的作用与隐含波动率在期权市场的作用有些类似，让不同票息、不同支付日的债券可以通过这个指标比较价值高低。市场上有不少数据商提供到期收益率曲线给投资者作为定价参考，但严格意义上讲到期收益率曲线并不存在，因为同一个到期日的债券，由于票息率、票息支付日的不同，到期收益率可能完全不一样，也就是说一个 x （期限）可能对应多个 y （到期收益率），没法把 y 表示成 x 的函数。能严格用于债券定价的是即期收益率曲线（spot rate curve），也就是不同期限零息债的到期收益率（图 1），不过市场上并没有那么多不同期限的零息债券，因此只能从付息债券的价格中去拟合反推即期收益率。

图 1：到期收益率曲线、即期收益率曲线与债券到期收益率（国开债，2020.12.31）



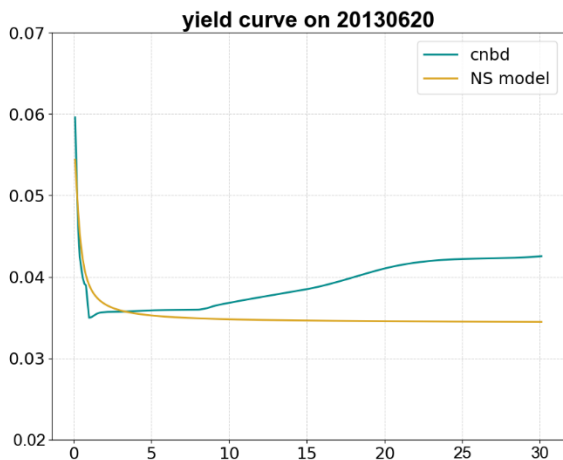
资料来源：东方证券研究所 & Wind 资讯 & 中债

利率曲线上的点有无穷多，因此要估计利率曲线必须先人为赋予曲线某种函数结构。各国央行最常用的函数结构有两种：三次样条和 Nelson-Siegel 指数函数形式，具体可参考 BIS(2005) 的技术文档。**NS 模型的结构形式非常简单**，它假设第 t 日，期限为 τ 的即期利率 $y_t^{(\tau)}$ 可表示为

$$y_t^{(\tau)} = a_t + b_t \cdot \frac{1 - e^{-\lambda \cdot \tau}}{\lambda \cdot \tau} + c_t \cdot e^{-\lambda \cdot \tau}$$

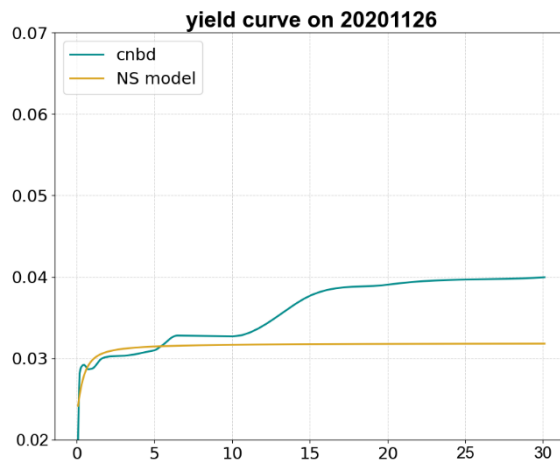
模型只有四个参数 $\{a_t, b_t, c_t, \lambda\}$ ，不同的参数组合可以大致生成常见的各种利率曲线形态。我们之前报告都采用的是此类模型来估算利率曲线，它的好处是结构简单，对债券价格数据噪音敏感性低，但缺点是局部形态不够灵活，给债券定价的误差较大。如图 2、图 3 所示，NS 模型拟合得到的曲线和中债使用三次样条方法得到的国债利率曲线（图例：cnbd）对比来看，10 年期以下，两者的形态大体一致，但中债曲线的局部细节更丰富；10 年期以上，两者走势差异明显，主要原因是 NS 模型的函数结构限制了其超长端的形态，其增强版 NSS 模型（Nelson-Siegel-Svensson）尾部形态会更丰富些，但自由度还是远不如三次样条方法。

图 2：NS 模型对中债国债即期利率曲线的拟合(20130620)



资料来源：东方证券研究所 & Wind 资讯 & 中债

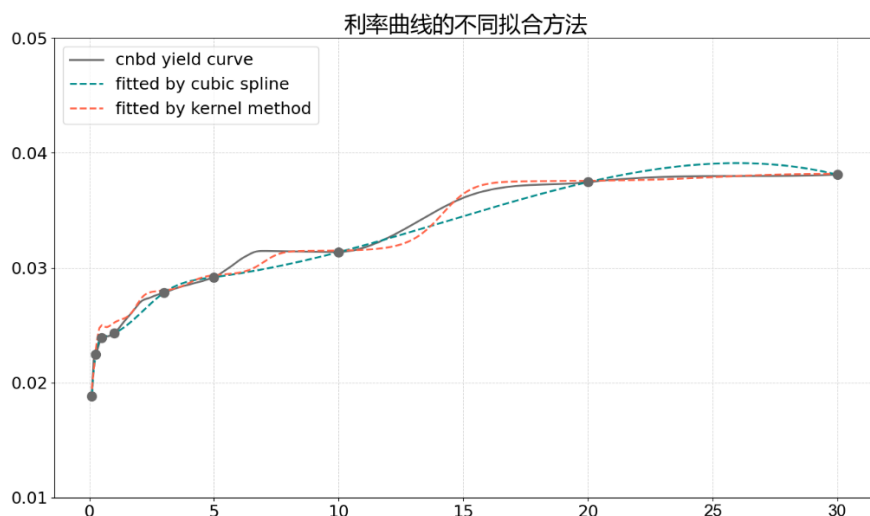
图 3：NS 模型对中债国债即期利率曲线的拟合(20201126)



资料来源：东方证券研究所 & Wind 资讯 & 中债

三次样条（Cubic Spline）方法的参数是利率曲线的一些节点（如图 4 的圆点所示），节点的数量需要事先设置，节点越密，模型的自由度越高，能描述的曲线形态越丰富，但模型对数据噪音越敏感，需要更多数据才能准确估计节点处的利率数值；节点间的利率曲线由三次多项式函数链接，保证曲线整体的光滑性；节点处的利率数值可以通过拟合债券的市场成交价格，优化求解得到。

图 4：三次样条法与 Kernel 方法示例（2020.12.31）



资料来源：东方证券研究所 & Wind 资讯 & 中债

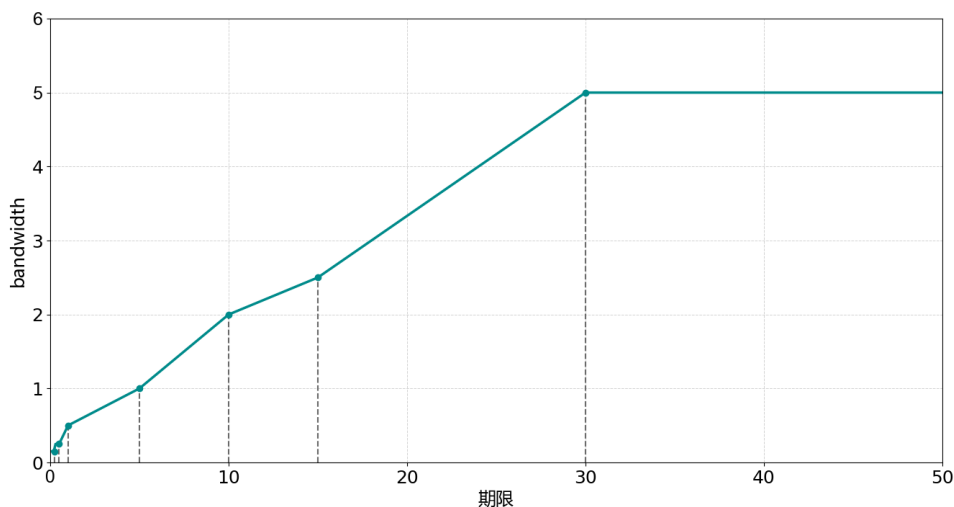
本文利率曲线估计采用的是 **Kernel 加权方法** (参考 Linton(2001))，它的参数和三次样条函数一样，是利率曲线上的一些节点。差别在于节点之间的利率曲线是通过把各个节点处的利率加权平均得到，各个节点的权重由 **Kernel 函数** 决定（报告选用是正态 kernel），计算任意期限 τ 处的即期利率时，离期限 τ 越近的节点权重越大；另外正态 kernel 函数里有 **bandwidth** 参数也需要预先设置，它可近似认为用来控制加权过程中用到的 τ 附近节点数量，**bandwidth** 越大，用到的附近节点数量越多，曲线越平滑，局部波动小；反之，曲线的局部波动会加大，局部看起来更加“歪歪扭扭”，有利于描述利率曲线局部形态，但提高定价精度的同时也会增加模型对数据噪音的敏感，降低参数估计准确性。

用数学式表达，把预先选定的 k 个期限节点，记作 $\{\tau_1, \tau_2, \dots, \tau_k\}$ ，节点处的即期利率 $\{y_{\tau_1}, y_{\tau_2}, \dots, y_{\tau_k}\}$ 是求解的目标，在 kernel 模型里，任意期限 τ 处的利率可用下式计算：

$$y_{\tau} = \frac{ker(\tau, \tau_1, h_1) \cdot y_{\tau_1} + ker(\tau, \tau_2, h_2) \cdot y_{\tau_2} + \dots + ker(\tau, \tau_k, h_k) \cdot y_{\tau_k}}{ker(\tau, \tau_1, h_1) + ker(\tau, \tau_2, h_2) + \dots + ker(\tau, \tau_k, h_k)}$$

其中 $ker(\cdot)$ 为正态 kernel 函数， $h_i, i = 1, 2, \dots, k$ 是其 **bandwidth** 参数。**Bandwidth** 参数设置的基本原则是：短端数值小，刻画短端利率的高波动；长端数值大，以获取更稳定的长端利率数值。实务中，我们用的是下列分段线性函数设置不同期限的 **bandwidth**。

图 5：bandwidth 分段线性函数



资料来源：东方证券研究所

本报告在撰写过程中，三次样条方法和 **Kernel 方法** 都有尝试，最终决定选择 **kernel 方法** 的原因一方面是 kernel 模型算法更加直观，计算高效，另一方面是读到 Liu(2019)的研究成果，作者用 kernel 方法在美国市场重构无风险利率曲线，与之前 Gurkaynak(2007) 的数据相比，kernel 方法得到的利率曲线定价更加准确，对利率市场未来变动的预测能力更强。

1.2 国债 & 国开债联合估计

国债和国开债在信用等级上都接近无风险，差异在于流动性和享受的税收政策。我们可以把国开债利率曲线形式上表示为：

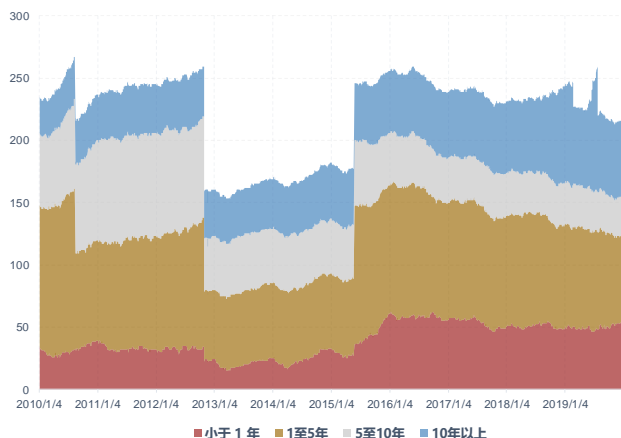
$$\text{国开债利率曲线} = \text{国债利率曲线} + \text{溢价曲线}$$

然后对国债利率曲线和溢价曲线分别用 kernel 方法建模。这种把国债和国开债数据揉合在一起来计算无风险利率曲线的方法，与把国债、国开债数据切开，分别用 kernel 方法对国债、国开利率曲线建模的方法比，数据利用效率更高。

举例来讲，假设单独用 kernel 方法给国债利率曲线建模，需设置 40 个节点，也就是 40 个参数才能较好拟合国债价格数据；类似的，国开债利率曲线也需 40 个参数；两者一共有 80 个参数。由于溢价曲线一般在 0 至 100 bps 范围内波动，变动幅度远小于利率曲线本身，若采用“国债利率曲线 + 溢价曲线”方式建模，可能只需要 10 个节点就能很好拟合溢价曲线。这样在数据量不变的情况下，需要估计的参数数量从 $40+40=80$ 降到 $40+10=50$ ，参数估计准确性能有效提升。

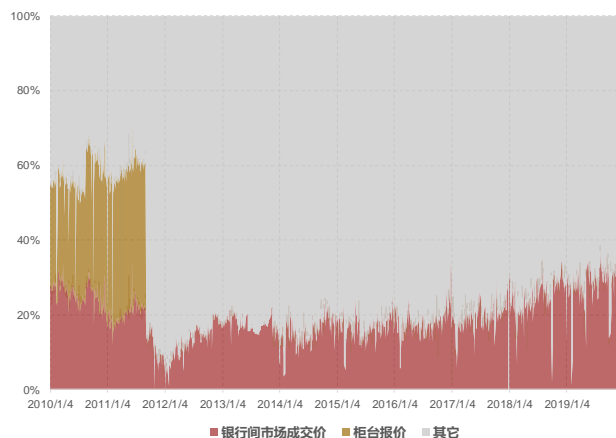
另外，如上篇专题报告《跨品种无风险利率曲线构建与应用》所述，中债国债利率曲线每天在估计过程中会用到两百多个券的价格数据，数量很多，但其中只有 20% 左右来自于银行间与交易所的成交数据，剩余 80% 来自于银行间报价和市场成员估值，主观因素和噪音因素较多；数据经过平滑性处理后，能够获得稳定的利率曲线提供给各家金融机构做债券估值；但是对于投资交易来说，过多非市场化因素的加入，可能会增加债券估值与市场成交价之间的偏差。因此，我们估算利率曲线只用银行间和交易所的成交数据，用定量算法剔除少许异常值（占比 2.5% 左右），希望这样得到的利率曲线能更准确反映出市场的定价和投资交易人员的一致预期。

图 6：中债国债利率曲线成分券各个剩余期限区间的券数量



资料来源：东方证券研究所 & Wind 资讯

图 7：中债国债利率曲线成分券不同价格数据来源的数量占比



资料来源：东方证券研究所 & Wind 资讯

只用市场成交数据估算利率曲线也有缺点，主要是早期市场上的券数据较少，价格数据质量不高，很难得到稳定连续的利率曲线；不过报告下文实证发现，随着债券市场的整体发展，从 2012 年开始，仅依靠市场成交数据，模型也可获得稳定的 10 年期利率曲线，具体见下文实证部分。

1.3 模型参数估计方法

模型参数估计也就是要找到节点 $\{y_{\tau_1}, y_{\tau_2}, \dots, y_{\tau_k}\}$ 的最优取值，使得利率曲线给的债券定价与真实市场成交价格之间的偏差尽可能小，这是一个非凸优化问题，不一定存在唯一的全局最优解，数值解法的结果与模型参数初始值设定高度相关。Liu(2019) 采用的是求解优化问题一阶条件的方式来做参数估计，但参数初始值怎么选、采用什么数值算法都没明确说明。按照作者的方案我们做了多次尝试，但优化结果不稳定，而且求解多元非线性方程组的数值算法很耗时，因此我们设计了一套自己的技术解决方案，要点如下：

1) 优化采用 Adam 算法，程序通过 Pytorch 实现。

人工神经网络的参数估计是一个典型的非凸优化问题，机器学习领域针对其发展了多种优化算法，同样可以用于利率曲线模型参数估计。本文采用的是 **Adam 算法**，它基于梯度下降过程中的“动量思想”，结合 Adagrad 和 RMSProp 两者优点，能让梯度下降更加平滑，获得更高的收敛效率。编程实现采用 Pytorch 框架。

2) 流动性好的券对利率曲线影响更大。

优化求解过程我们采用的损失函数 (Loss Function) 如下：

$$\sum_{i=1}^N w_i \left| \frac{p_i^{model} - p_i^{mkt}}{d_i} \right|$$

其中 d_i 是第 i 个券的久期， p_i^{mkt} 是其当天市场成交均价， p_i^{model} 是模型给的债券定价， w_i 是根据流动性给这个券赋予的权重，它由成交额决定：1 个亿以下的权重为 1，10 个亿以上的权重为 3，1 个亿到 10 个亿之间的，权重由 1 线性递增到 3。

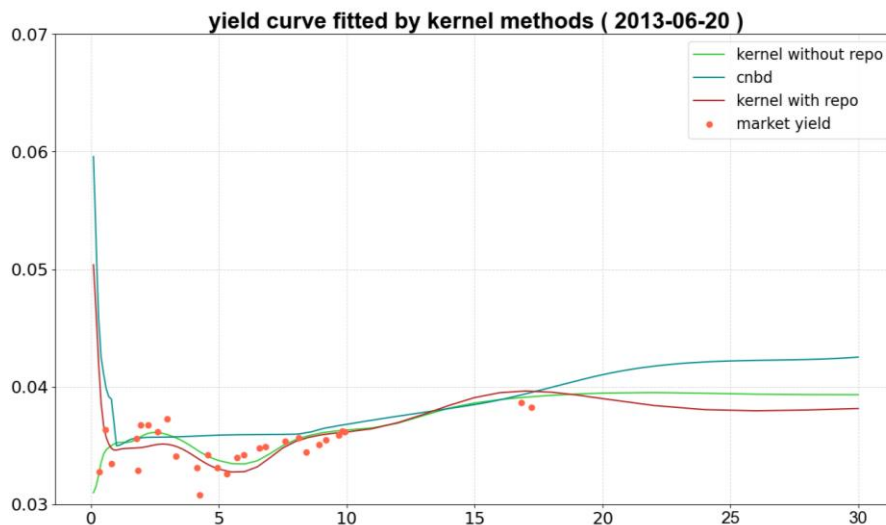
3) 通过主成分数量控制模型的自由度。

Liu(2019) 选取的节点数量多达 360 个，每月一个节点，自由度很高；如果像美国市场那样有足够多的成交数据，这个影响可能不大，但在国内市场我们尝试过这样不作任何自由度约束的方法去优化，结果很容易过拟合，得到形状“歪扭”的曲线形态。最后我们通过提取利率曲线主成分，限定主成分数量（10 个），反向拟合利率曲线的方式来控制模型的自由度，避免过拟合。

4) 用回购利率辅助确定利率曲线初始值。

在上篇专题报告《跨品种无风险利率曲线构建与应用》中，我们直接把回购作为短端无风险利率加入数据集估计利率曲线，但回购和国债是两个市场，从历史平均看，回购利率的数值大于同期限的短端国债利率，存在一定的溢价，因此加入回购利率估算得到的无风险利率曲线短端会明显高于中债国债利率曲线，增加债券的定价误差。在本篇报告里，回购利率数据只是用来辅助确定优化算法的初始值，并不参与后面拟合债券价格的优化过程。回购利率的辅助作用非常必要，因为国债和国开债市场上有些交易日会缺乏短期限的券交易，这在 2013.06.20 “钱荒” 时表现尤为明显，当日有交易的国债、国开债最短期限为 3.8 个月，无法准确反映出短端利率的陡然提升（图 8）

图 8：引入回购利率数据辅助确认利率曲线初始值（国债，2013-06-20）

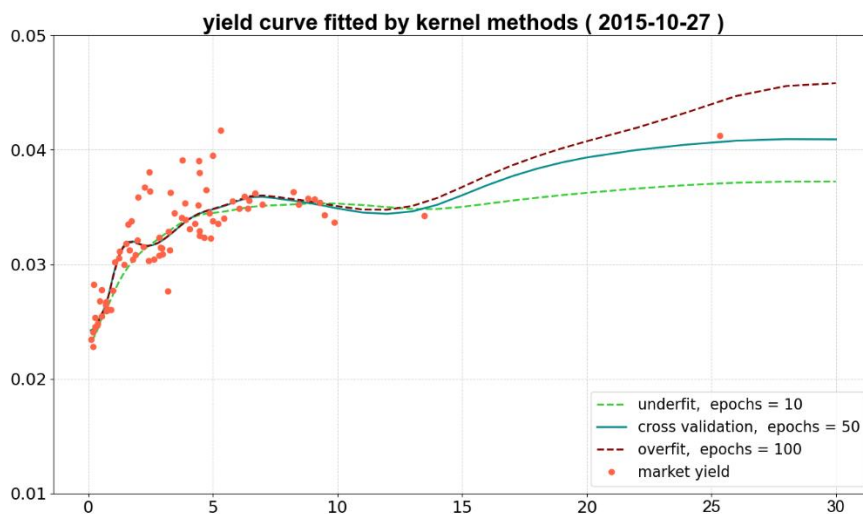


资料来源：东方证券研究所 & Wind 资讯

5) 引入机器学习里的交叉验证思想，避免模型过拟合。

优化算法求解过程中，模型通常需要对样本数据进行多次学习才能找到最优解，学习次数（epochs）如何确定，可以借鉴机器学习里的交叉验证（Cross Validation）思想。假设某个交易日市场成交 100 只券，拿出其中一只券，用剩余 99 只构建利率曲线，并对拿出的这只券做定价，看看定价误差有多少，此误差称作交叉验证误差；学习次数越多，对 99 只券价格的样本内拟合效果越好，但交叉验证误差可能变大；之后重复上面的操作，拿出另外一只券，再计算交叉验证误差；重复多次后，把得到的系列交叉验证误差取平均；学习次数的选取应该使得样本外的交叉验证误差尽可能小，避免过拟合。（图 9）

图 9：利率曲线的欠拟合与过拟合（国开，2015-10-27）



资料来源：东方证券研究所 & Wind 资讯

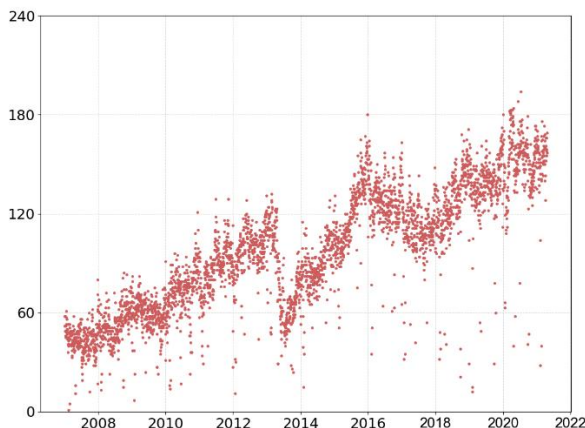
按照机器学习里的 leave-one-out 标准做法，应把所有券遍历拿出来一次，但运算太过耗时；我们的做法是把利率期限分成四个期限区间，每个区间拿出交易最活跃的 1-2 只券，做上述的交叉验证过程。交叉验证误差的概念由 Liu (2019) 提出，但作者只是事后用来度量利率曲线的定价精度，在利率曲线估计过程中并没有考虑这个问题，我们这里的做法更进一步，更符合机器学习思想，可有效避免利率曲线过拟合。

二、DFQ 市场化利率曲线

2.1 利率曲线形态比较

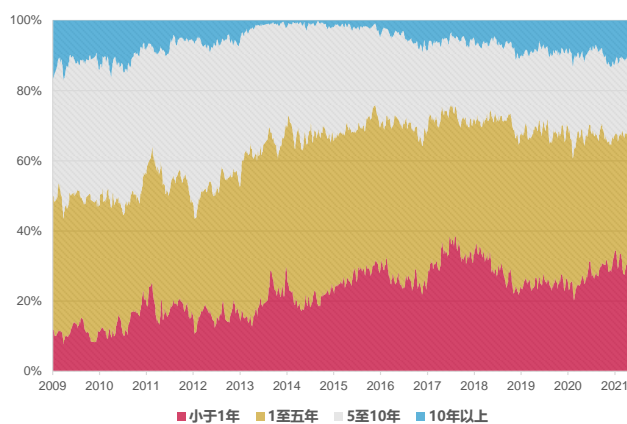
本文实证部分用到的历史数据时间段为 2007.01.04 至 2021.04.21，数据取自 Wind Filesync 数据库；每个交易日我们把银行间、交易所成交的国债和国开债糅合在一起同时估算国债和国开债利率曲线，结果称作 **DFQ 国债利率曲线**和 **DFQ 国开利率曲线**。债券价格采用当天的市场成交均价（VWAP）；如果当天成交的券数量少于 60，则会从前个交易日的成交数据中挑出流动性好且和当天无重复的券，合在一起估计利率曲线，相当于移动平均效果，不过从图 10 可以看到，2012 年之后需要“借券”的交易日数量占比很少。另外从券的期限分布看，1-5 期的券数量占比最高，10 年期以上的券数量占比很少，10%左右，因此十年期以上的利率曲线很难完全依靠市场数据做估计，这点下文有更详细讨论。

图 10：国债+国开债有市场成交的券数量



资料来源：东方证券研究所 & Wind 资讯

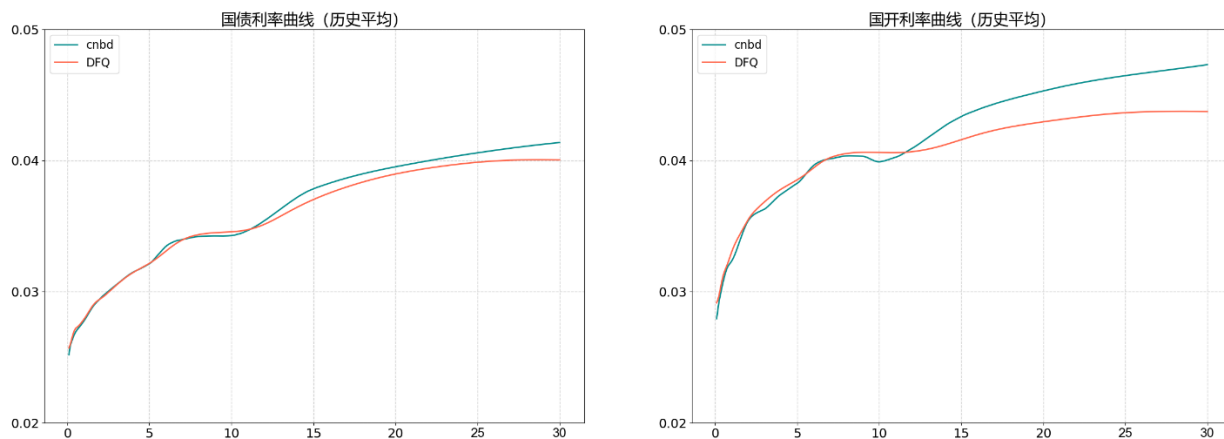
图 11：国债+国开债有市场成交的券的剩余期限占比



资料来源：东方证券研究所 & Wind 资讯

图 12 把 2012.01.04 起的中债国债、国开债利率曲线（cnbd）和 DFQ 国债、国开利率曲线在按时间序列方向进行了平均，得到平均利率曲线。历史平均来看，10 年期以内，中债利率曲线和 DFQ 利率曲线基本一致，DFQ 国开利率曲线光滑度更好；但 10 年以上的超长端，中债和 DFQ 差别较大，特别是国开债利率曲线，30 年期的即期利率差距达到 35bps。

图 12：国债&国开债历史平均利率曲线（2012.01.01 – 2021.04.21）



资料来源：东方证券研究所 & Wind 资讯

时间序列方向上，图 13 比较了三个月期、一年期、五年期、十年期国开债利率的走势。

图 13：国开债利率走势（2007.01.04 – 2021.04.21）

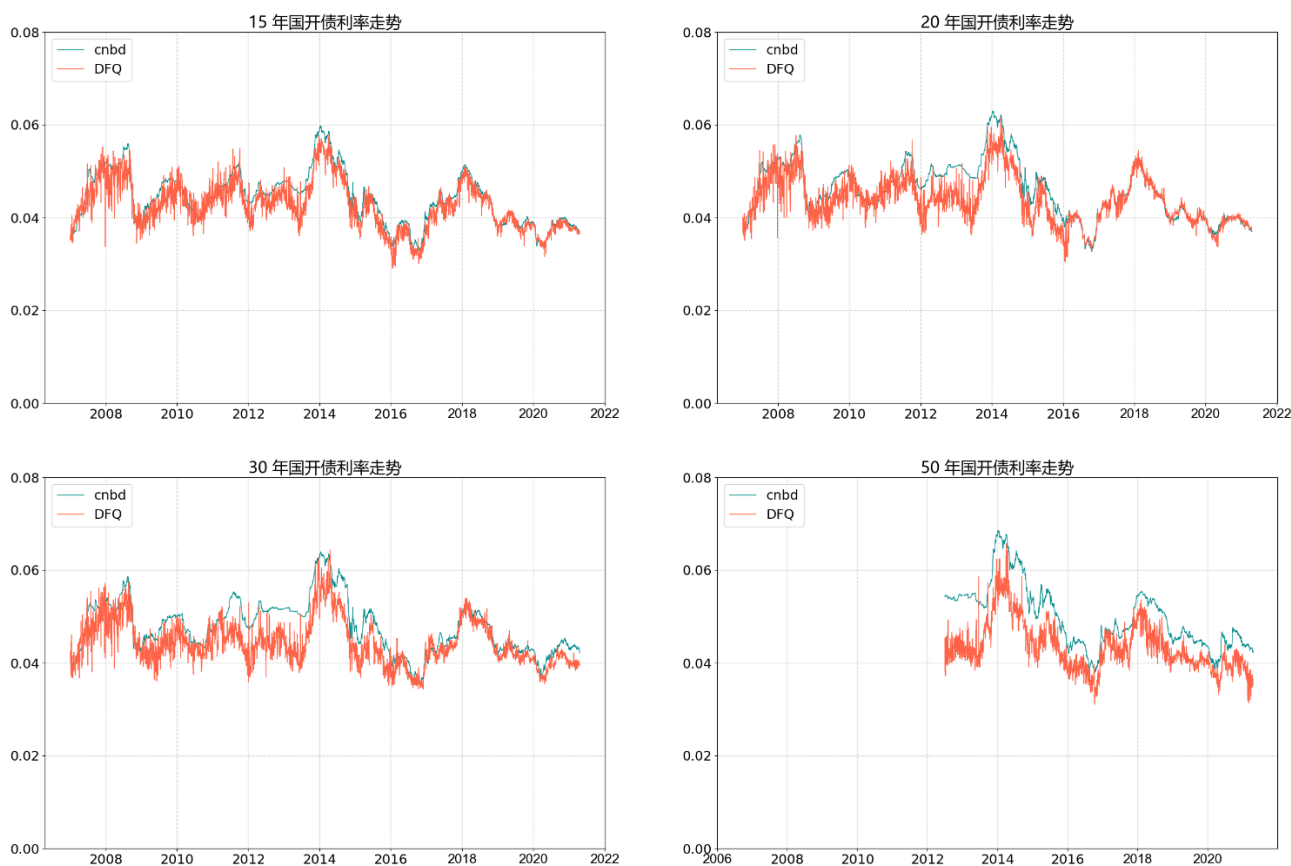


资料来源：东方证券研究所 & Wind 资讯

可以看到，2012 年之前由于市场上有成交的券数量较少，市场整体流动性和价格数据质量不高，完全基于市场成交数据估算出来的 DFQ 利率曲线波动较大，中债利率曲线要相对稳定得多，但这应该是中债在估算利率曲线时主动加了一些平滑处理，好处是能让金融机构的产品获取稳定估值，代价是牺牲利率曲线的市场定价精度。2012 年之后，十年期以下的 DFQ 利率走势基本和中债一致，这可能和机构投资者愈加壮大成熟，市场流动性提升有关，完全基于市场成交数据也能获得稳定的十年期以内利率曲线。

图 14 展示的是 10 年期以上超长端国开债利率的走势，DFQ 利率波动非常之大，这是因为 kernel 方法的自由度很高，而超长端的券数量不多，一个超长端券的价格可能要决定七八个节点处的利率，有非常多的解能够让利率曲线在这个超长端券处的拟合误差最小，但优化算法本身不能保证每个交易日都收敛到同一个最优解附近，因此看到利率走势在不停的上下抖动。中债采用的三次样条方法也是自由度很高的模型，如果完全基于市场成交数据估算利率曲线也会面临同样问题，图 14 展示的中债利率平缓走势应该也是做了主动平滑处理。不过，随着市场上超长端券数量的增加，2017 年后，kernel 方法得到的市场化 15 年、20 年、30 年国开债利率走势已经日趋稳定，和中债走势基本一致，50 年期的波动仍然很大。

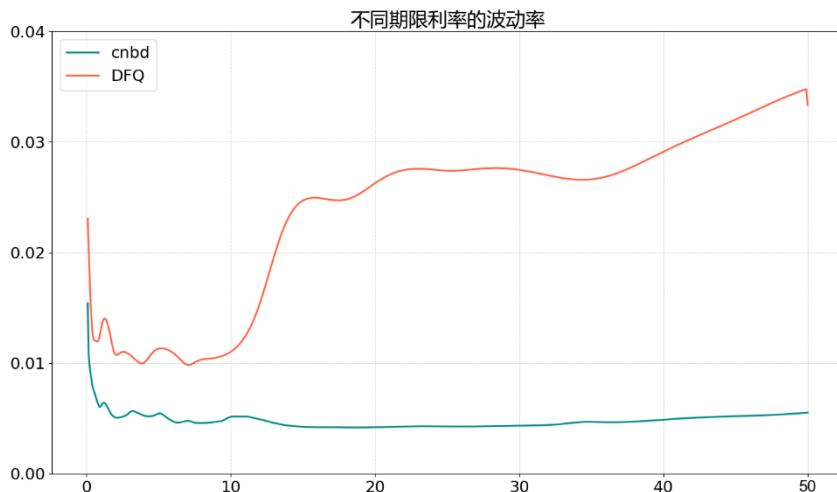
图 14：超长端国开债利率走势（2007.01.04 – 2021.04.21）



资料来源：东方证券研究所 & Wind 资讯

最后中债和 DFQ 不同期限利率的年化波动率。和前面的分析结果一致，10 年以下，DFQ 利率的波动基本上是中债的两倍，且满足“短端波动大，长端波动小”的市场经验；10 年期以上，由于有市场成交的券数量少，纯市场化方法得到的 DFQ 利率波动会陡然增加。低波动利率曲线可满足金融机构获取稳定估值的需求，但不一定能真实反映市场变化，下一节实证结果可佐证说明。

图 15：不同期限利率的年化波动率（2012.01.04 – 2021.04.21）



资料来源：东方证券研究所 & Wind 资讯

2.2 定价偏差比较

有了即期利率曲线后，投资者可以把利率债的未来现金流进行折现给债券做定价，并把它和债券的市场成交价做比较，看看定价精度如何。定价精度高的利率曲线有助于帮助投资者发掘交易中价值被“高估”或“低估”的券。

图 16：DFQ 利率曲线与中债利率曲线的定价偏差（2012.01.04 – 2021.04.21）

剩余期限	DFQ 国债利率曲线定价误差(bps)				中债国债利率曲线定价误差(bps)				误差降幅
	均值	25%分位数	50%分位数	75%分位数	均值	25%分位数	50%分位数	75%分位数	
小于1年	7.58	1.62	4.27	9.46	8.09	1.83	4.34	9.85	-6.3%
1至5年	6.36	1.48	3.87	8.34	6.74	1.83	3.91	8.56	-5.6%
5至10年	5.18	1.34	3.40	6.94	5.44	1.48	3.46	7.07	-4.9%
10年以上	3.96	0.82	2.37	5.36	4.73	1.08	2.92	6.68	-16.2%

剩余期限	DFQ 国债利率曲线定价误差(bps)				中债国债利率曲线定价误差(bps)				误差降幅
	均值	25%分位数	50%分位数	75%分位数	均值	25%分位数	50%分位数	75%分位数	
小于1年	10.55	1.97	5.42	12.62	12.04	2.48	6.26	14.18	-12.4%
1至5年	15.53	2.72	7.24	20.10	18.58	3.14	9.48	24.26	-16.4%
5至10年	14.29	1.88	5.18	13.03	16.02	2.30	6.18	14.35	-10.8%
10年以上	8.28	0.81	3.03	7.71	9.06	1.23	3.75	8.90	-8.7%

资料来源：东方证券研究所 & Wind 资讯

我们把利率曲线给的债券定价隐含到期收益率与市场价格隐含到期收益率之间的偏差绝对值称为**定价偏差**。图 16 是对历史上不同期限的国债、国开债分别在中债和 DFQ 利率曲线下的定价误差统计。由于 10 年以上券的数量少，纯用市场成交数据估算得到的 DFQ 利率曲线，10 年期以上利率不稳定，因此我们重点考察 10 年期以下券的定价偏差。从上图可以看到，不论是国债还是国开债，DFQ 利率曲线的定价误差都明显更小，国开债尤为明显，DFQ 利率曲线定价误差相对中债利率曲线定价误差的下降幅度都在 10% 以上，在数量占比最高的 1-5 年期券里，DFQ 利率曲线定价误差相对中债的降幅可达到 16.4%。

如果把 10 年期以下的国开债拿出来分年统计定价偏差（图 17），DFQ 利率曲线每年的定价误差都低于中债利率曲线，误差降幅在 2019 年最高可达 25.4%，历年误差降幅平均在 15.5%。

图 17: DFQ 国开曲线与中债国开曲线的定价偏差 (bps, 2012.01.04 – 2021.04.21)

	2012	2013	2014	2015	2016	2017	2018	2019	2020
中债国开	30.05	23.62	23.14	20.36	19.80	11.01	9.42	7.49	6.89
DFQ 国开	27.60	21.59	20.94	16.11	15.15	9.79	7.71	5.58	5.92
误差降幅	-8.2%	-8.6%	-9.5%	-20.9%	-23.5%	-11.0%	-18.2%	-25.4%	-14.0%

资料来源：东方证券研究所 & Wind 资讯

需要补充说明的是，中债在构建利率曲线时为帮助债券产品获得稳定估值，价格中引入了非市场化数据源，因此增加了利率曲线的市场定价误差；DFQ 利率曲线只用了债券的市场成交数据，但它采用了交叉验证的方式避免过拟合；如果只是单纯的追求样本内拟合误差最小，对于 kernel 加权这种自由度极高的模型，样本内拟合误差可以做到更低。

2.3 寻找收益率低估与高估券

债券按市价计算的到期收益率大幅高于或低于利率曲线给定的到期收益率，可能是受短期事件影响（例如：流动性冲击），收益率溢价会随着时间的推移减小，做多收益率高估的券，做空收益率低估的券可以帮助投资者获取溢价收益。

假设投资者每天用利率曲线给国开债做定价，并和市场成交价做比较，计算到期收益率溢价幅度；然后设定一个阈值，例如 10bps，做多到期收益率高估 10bps 以上的券，同时做空到期收益率低估 10bps 以上的券，多头组合与空头组合的久期都调整为 1；如果投资者能以当天的市价构建这个多空组合，则可以锁定溢价，而且因为多空组合的整体久期为零，对短期内利率曲线的平行移动免疫，因此若多空建仓后，每日按照利率曲线的给债券定价，多空组合的持有收益变化应该很小，在建仓时锁定的溢价收益附近上下小幅波动；但若利率曲线构建过程融入了较多非市场因素，利率曲线定价偏离市场价格太多，多空组合的持有收益可能会随着时间有较大波动。

图 18 是我们用 2012.01.04 – 2020.12.31 的国开债做的实证测试，分别用中债利率曲线和 DFQ 利率曲线做定价，溢价阈值分别设为 10bps, 20bps, 30bps，每个交易日看是否有券的到期收益率溢价水平超过阈值，有则按照上述方式构建多空组合，计算多空组合持有 5、10、20、40、60、120 个交易日的收益，最后汇总统计分析。考虑到短端利率波动较大，同时 DFQ 利率曲线 10

年期以上的即期利率估计不准，因此在选择收益率高估券与低估券时，剔除了剩余期限在半年以下或 10 年以上的券。另外，也剔除了日成交额不足一个亿、或上市不足一年的券（避免新券相对旧券流动性溢价的影响）。可以很明显的看到，不论溢价阈值如何设置，基于 DFQ 利率曲线构建的多空组合能锁定的溢价收益显著更高，而且随持有期的拉长波动很小，但基于中债利率曲线算的多空组合收益，在持有市场超过 40 个交易日后就有较大波动，可能原因是利率曲线里非市场因素较多。因此从以上结果看，市场化的 DFQ 利率曲线更适合挖掘收益率大幅高估或低估的券。

图 18：多空组合的持有收益（bps，2012.01.04 – 2021.04.21）

持有期	阈值 = 10bps		阈值 = 20bps		阈值 = 30bps	
	DFQ	cnbd	DFQ	cnbd	DFQ	cnbd
5	70.3	23.2	99.5	31.4	122.4	36.0
10	70.6	24.3	99.8	32.7	122.9	37.6
20	69.8	26.6	99.4	35.5	123.4	41.0
40	69.4	31.2	100.1	41.4	125.2	47.7
60	68.3	35.4	99.5	47.0	126.3	53.9
120	69.1	51.5	101.8	63.0	129.4	68.1

资料来源：东方证券研究所 & Wind 资讯

下面重点关注收益率高估的国开债，它们可以买入持有的方式获取溢价，国开债券池的剔除规则和图 18 一致。图 19 展示的是在不同阈值下，每日到期收益率高估程度大于阈值的券数量（为了方便观察趋势，做了 20 日移动平均处理）。可以看到，随着债券市场规模的壮大，市场的定价效率在显著提高，每日能发现的高估券数量从 2017 年开始显著减少，到 2020 年底的时候，单个交易日里，到期收益率高估 10bps 以上的券平均有 2-3 个，20bps 以上的券平均只有一个。

图 19：每个交易日收益率高估券的数量（2012.01.04 – 2021.04.21）



资料来源：东方证券研究所 & Wind 资讯

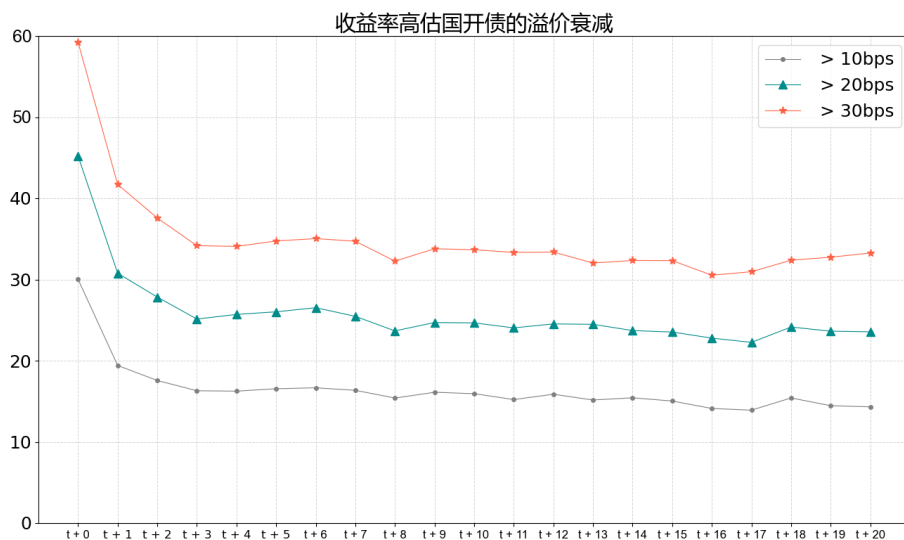
在以上回溯测试过程中,利率曲线在每天收盘后计算,之后才知每个券的到期收益率高估程度,买入交易只能等到下一个交易日进行。因此,实盘操作中一个很重要的考量是,溢价是否持续,或者说衰减速度如何。我们统计了收益率高估的券在未来 20 个交易日的收益率溢价情况,图 20 是随机截取的一段数据样本以作示例,空白之处表示当日无成交,t+0 代表的是最左侧的交易日期,然后把所有的历史数据平均,可以得到**溢价衰减曲线(图 21)**。平均来看,溢价的国开债在之后三个交易日溢价衰减速度最快,尤其是第一个交易日,溢价>10bps 的券平均溢价幅度下降 10.6bps,溢价>20bps 的券平均溢价幅度下降 14.4bps,溢价>30bps 的券平均溢价幅度下降 17.5bps;三个交易日之后溢价趋于平稳。

图 20: 到期收益率高估的国开债未来一个月的溢价衰减样本数据(bps)

交易日	代码	t+0	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10	t+11	t+12	t+13	t+14	t+15	t+16	t+17	t+18	t+19	t+20
2018/11/30	150216.IB	12.6	6.3		1.9	8.7					8.3			12.3					16.6	5.5	16.7	
2019/1/3	150216.IB	13.1	4.3	-2.3	10.9		3.4	6.4	7.1				4.2	8.8	9.4							8.3
2019/3/8	150216.IB	11.7	4.0	4.8	5.6	10.7				7.2			4.4		12.7	4.4	12.4		6.7	9.0		
2019/5/13	150216.IB	11.8	4.8								4.5	6.3		1.4					2.5			
2019/7/3	150216.IB	14.6					3.3		1.8	6.5	7.5	6.5	6.3				3.7	5.4	-0.6	-2.1		
2020/3/26	150216.IB	14.6	12.3	0.5		4.1				-2.9	6.1		5.8					2.5				3.1
2020/6/10	150216.IB	71.3		63.4						0.2				2.6	4.7	4.1	3.6	1.5				
2020/12/15	150216.IB	13.9		3.8								1.1	2.3									
2016/1/25	150217.IB	11.5	-0.5	-1.4	-2.4	-1.3	-9.2	-2.3	-9.3	-6.4	5.6	18.7	0.9	-2.8	-9.5	0.8	-9.1	-7.8	-7.4	-3.6		-9.1
2016/3/22	150217.IB	14.3	1.3	0.2	-3.5	1.7	0.5	-6.4	3.9	4.1	-1.6	-0.8	3.7	0.4	1.7	-0.6	-1.9	-2.0	-8.8	-0.3	1.0	3.3
2016/12/6	150217.IB	51.1	3.1	14.6	0.0	-1.1	3.5	1.2	-15.5	2.0	-19.5	0.1	-6.8	7.0	6.7	12.1	-11.3	0.0	-11.3		0.6	15.8
2017/1/5	150217.IB	14.7	9.4	18.1	20.2	18.6	25.7	17.1	2.7			2.7			-24.1	18.8	-1.5	0.8	-2.6	11.7	10.9	22.1
2017/2/10	150217.IB	17.6	12.3	13.9	15.0	0.2	-0.5	1.1	2.5	6.9	10.1	3.4	4.2	2.1	1.4	-0.6	2.8	5.1	0.9	0.9	-2.2	-7.2
2017/4/11	150217.IB	12.5	1.9	-0.6	0.8	6.9	0.3	0.7	2.8	-7.3	-2.1	-2.4	-1.9	-6.7	1.9	-0.9	-9.2	1.4	-1.4	1.0	3.0	1.8
2016/12/6	150218.IB	28.2	4.0	-2.5	-1.1	-0.4	0.5	5.3	-0.1	-3.7	4.2	4.2	0.7	0.4	-11.4	6.9	-10.2	-15.4	0.5	32.3	0.0	-2.3

资料来源: 东方证券研究所 & Wind 资讯

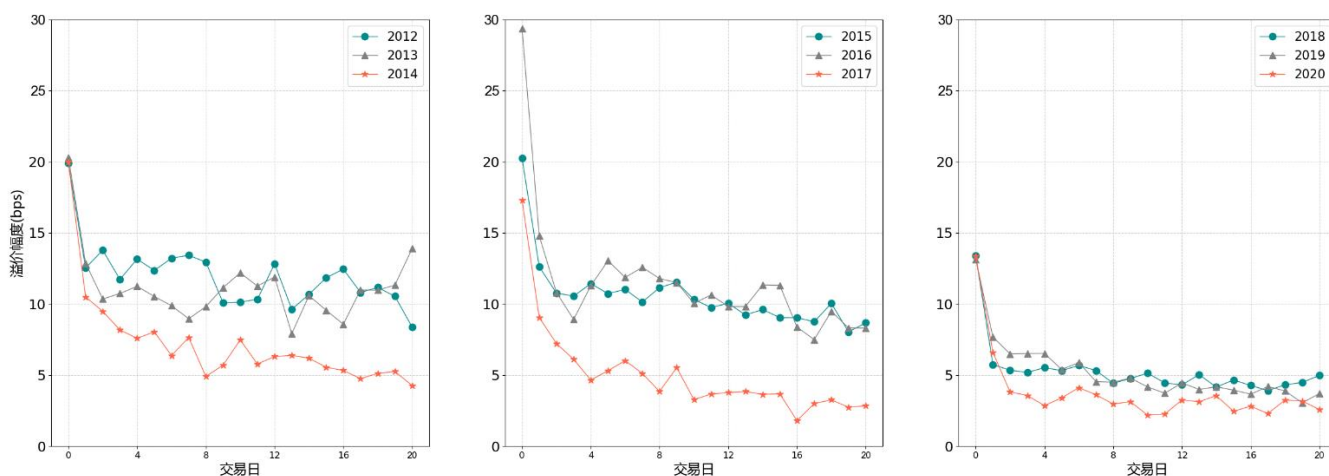
图 21: 不同高估程度的国开债收益率溢价衰减曲线 (bps, 2012.01.04 – 2021.04.21)



资料来源: 东方证券研究所 & Wind 资讯

图 21 是整个历史汇总统计的结果, 无法反应市场随时间的变化。图 22 把收益率溢价在 10bps 以上的国开债分年统计的溢价衰减曲线。可以明显看到, 随着市场定价效率的提升, 最近三年国开债溢价空间明显收窄, 且衰减速度加快, 例如 2020 年, 到期收益率在 10bps 以上的国开债平均溢价 14bps 左右, 但第二个交易日溢价陡然下降至 4bps, 留给投资者隔夜交易的空间很小。若想更有效的捕获溢价, 投资人可以考虑盘中实时估算利率曲线, 尽早发现机会, 尽快交易, 此时需要类似 DFQ 利率曲线这样完全基于市场成交价格的利率曲线估计方法。

图 22: 分年统计的国开债收益率溢价衰减曲线



资料来源: 东方证券研究所 & Wind 资讯

2.4 利率曲线的无套利性

给利率衍生品定价时, 需要在风险中性测度下用随机微分方程对利率曲线进行建模, 模拟其未来的变化路径, 而风险中性测度存在的理论前提是市场无套利, 也就是说即期利率曲线上各个期限利率对应的零息债券之间不存在套利机会。这就要求市场上有数量丰富的各种期限零息债作为工具, 配合便利的做空交易机制帮助投资者实现套利, 国内市场目前不能满足。所以基于无风险中性测度的一些 Q-Model, 例如 Vasicek, CIR, G-ATSM (Gaussian Affine term structure model) 在国内市场的使用效果都不佳。我们在之前专题报告《利率曲线: 可预测与不可预测成分》中用 G-ATSM 模型和 Andrian(2013)的方法对美国国债收益率曲线和中债国债收益率曲线进行了拟合, 前者各个期限利率的拟合误差平均都在 1bps 以内, 而后者基本都在几十个 bps, 短端甚至超过 100bps, 说明美国市场近似接近“无套利”的状态, 国内市场距离还比较大。我们当时认为中债利率曲线拟合误差大的一个重要原因是利率曲线里面掺和了非市场因素, 但这次又用纯市场化的 DFQ 利率曲线做了同样测试, 拟合误差和中债差别很小, 说明拟合误差最主要的原因还是市场机制层面的, 缺乏足够数量不同期限的零息债和做空工具。

2.5 利率曲线隐含的未来信息

在上一篇专题报告《跨品种无风险利率曲线构建与应用》中，我们用纯市场数据和 NS 模型构建了无风险利率曲线（简称，跨品种利率曲线），发现它能显著预测未来一个季度十年期零息债的风险溢价（BRP, bond risk premium），因此做这篇报告时，我们本希望能够获得和 Liu(2019)类似的结果：DFQ 利率曲线对未来 BRP 的预测能力更强，但事与愿违，实证结果显示：DFQ 利率曲线可以预测未来一年的十年期 BRP，样本外 rsquared（OSS, out of sample squared）略低于中债；但对于未来一个季度的短线十年期 BRP，DFQ 和中债一样都无法预测。

我们对模型架构做了多次调整来寻找原因，实证过程不赘述，这里简单说明一下结论。跨品种利率曲线对短线未来一个季度十年期 BRP 有预测能力的来源主要是两个：

- 1) **短端加入了回购利率数据。**如果把回购数据拿掉，任何模型结构下，利率曲线对未来一个季度的十年期 BRP 都没显著预测作用。
- 2) **NS 模型的简单架构。**如果把 NS 模型按照灵活度从低到高依次扩展，例如：放开 NS 模型形状参数约束、采用 kernel 方法替代常数项的溢价，模型获得的利率曲线对未来一个季度的十年期 BRP 预测能力都会降低，甚至消失。但用简单 NS 模型的代价是定价误差增大，高于中债利率曲线，也就是说 NS 模型得到的利率曲线和真实利率曲线可能有较大偏差，那么基于 NS 利率曲线算得的 BRP 和真实 BRP 也可能有较大偏差。

基于以上结果，我们建议在用 Cochrane（2005）方法预测短线十年期 BRP 时，只有预测值幅度非常大，100bps 以上时，模型给出的预测才是一个可靠信号。

三、总结

本报告研发的 DFQ 利率曲线，完全基于国债、国开的市场成交数据估算而得，定价误差显著低于中债，更适合用来给债券定价，发掘收益率高估或低估的券，套取溢价；并且 DFQ 利率曲线可以实盘实时估算，能让投资者更早发现机会，更快交易，锁定溢价。

一条完美的利率曲线应该定价误差很小，同时利率期限结构能充分反映投资交易人员的一致预期，对利率曲线未来变动做出预测。但从本文结果看，目前这两者是鱼与熊掌，不可兼得，定价精度高的利率曲线对未来的预测能力弱，对未来预测能力强的曲线定价精度低，两者何时能统一，还有赖于债券市场的进一步发展成熟。

风险提示

1. 量化模型基于历史数据分析得到，未来存在失效风险，建议投资者紧跟模型表现。
2. 极端市场环境可能对模型效果造成剧烈冲击，导致收益亏损。

参考文献

- [1]. Andrian, T., Crump, R.K., Moench, E., (2013), "Pricing the term structure with linear regressions", Journal of Financial Economics, 110(1): 110-138.
- [2]. BIS, (2005), "Zero-coupon yield curves: technical documentation", Monetary and Economic Department, Bank of International Settlements.
- [3]. Cochrane, J.H., Piazzesi, M., (2005), "Bond risk premia", American Economic Review, 95:138-160.
- [4]. Gürkaynak, R. S., Sack, B., Wright, J. H., (2007), "The U.S. treasury yield curve: 1961 to the present". Journal of Monetary Economics, 54: 2291-2304.
- [5]. Linton, O., Mammen, E., Nielsen, J. P., Tanggaard, C., (2001). "Yield curve estimation by kernel smoothing methods", Journal of Econometrics, 105:185-223.
- [6]. Liu, Y., Wu, J.C., (2019), "Reconstructing the Yield Curve", Journal of Financial Economics (JFE), Forthcoming, Available at SSRN: <https://ssrn.com/abstract=3286785>

分析师申明

每位负责撰写本研究报告全部或部分内容的研究分析师在此作以下声明：

分析师在本报告中对所提及的证券或发行人发表的任何建议和观点均准确地反映了其个人对该证券或发行人的看法和判断；分析师薪酬的任何组成部分无论是在过去、现在及将来，均与其在本研究报告中所表述的具体建议或观点无任何直接或间接的关系。

投资评级和相关定义

报告发布日后的 12 个月内的公司的涨跌幅相对同期的上证指数/深证成指的涨跌幅为基准；

公司投资评级的量化标准

买入：相对强于市场基准指数收益率 15%以上；

增持：相对强于市场基准指数收益率 5% ~ 15%；

中性：相对于市场基准指数收益率在-5% ~ +5%之间波动；

减持：相对弱于市场基准指数收益率在-5%以下。

未评级 —— 由于在报告发出之时该股票不在本公司研究覆盖范围内，分析师基于当时对该股票的研究状况，未给予投资评级相关信息。

暂停评级 —— 根据监管制度及本公司相关规定，研究报告发布之时该投资对象可能与本公司存在潜在的利益冲突情形；亦或是研究报告发布当时该股票的价值和价格分析存在重大不确定性，缺乏足够的研究依据支持分析师给出明确投资评级；分析师在上述情况下暂停对该股票给予投资评级等信息，投资者需要注意在此报告发布之前曾给予该股票的投资评级、盈利预测及目标价格等信息不再有效。

行业投资评级的量化标准：

看好：相对强于市场基准指数收益率 5%以上；

中性：相对于市场基准指数收益率在-5% ~ +5%之间波动；

看淡：相对于市场基准指数收益率在-5%以下。

未评级：由于在报告发出之时该行业不在本公司研究覆盖范围内，分析师基于当时对该行业的研究状况，未给予投资评级等相关信息。

暂停评级：由于研究报告发布当时该行业的投资价值分析存在重大不确定性，缺乏足够的研究依据支持分析师给出明确行业投资评级；分析师在上述情况下暂停对该行业给予投资评级信息，投资者需要注意在此报告发布之前曾给予该行业的投资评级信息不再有效。

免责声明

本证券研究报告（以下简称“本报告”）由东方证券股份有限公司（以下简称“本公司”）制作及发布。

。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。本报告的全体接收人应当采取必要措施防止本报告被转发给他人。

本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证券研究报告。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证券研究报告之外，绝大多数证券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证券或其他投资标的的邀请或向人作出邀请。

本报告中提及的投资价格和价值以及这些投资带来的收入可能会波动。过去的表现并不代表未来的表现，未来的回报也无法保证，投资者可能会损失本金。外汇汇率波动有可能对某些投资的价值或价格或来自这一投资的收入产生不良影响。那些涉及期货、期权及其它衍生工具的交易，因其包括重大的市场风险，因此并不适合所有投资者。

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，投资者自主作出投资决策并自行承担投资风险，任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

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协议授权，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。

经本公司事先书面协议授权刊载或转发的，被授权机构承担相关刊载或者转发责任。不得对本报告进行任何有悖原意的引用、删节和修改。

提示客户及公众投资者慎重使用未经授权刊载或者转发的本公司证券研究报告，慎重使用公众媒体刊载的证券研究报告。

东方证券研究所

地址：上海市中山南路 318 号东方国际金融广场 26 楼

电话：021-63325888

传真：021-63326786

网址：www.dfzq.com.cn