

中证指数研究报告

债券指数抽样复制方法研究

目录

一、前言.....	2
二、相关指数跟踪方法回顾.....	2
（一）完全复制法.....	3
（二）抽样复制法.....	3
（三）多因素风险模型最优化法.....	4
（四）衍生品复制法.....	5
（五）随机情景模拟法.....	5
三、分层抽样复制方法.....	5
四、分层抽样方案设计.....	7
五、结果分析.....	7
（一）指数走势与样本周转率比较.....	7
（二）不同抽样方法持有期相对收益比较.....	10
（三）不同抽样方法跟踪误差比较分析.....	11
六、跟踪误差产生的原因.....	15
七、主要结论.....	17

图表目录

图 1：分层抽样复制标准划分.....	4
图 2：跟踪误差分解示意图.....	5
图 3：流动性 100 指数与 10 种抽样复制组合走势.....	8
图 4：不同抽样方法调样过程中相同债券权重因子变化绝对值.....	10
图 5：基准指数与抽样复制债券组合样本区间年化收益率（%）.....	11
图 6：样本期间不同抽样复制方法年化跟踪误差比较.....	12
图 7：样本期间不同抽样复制方法累积跟踪误差比较.....	12
图 8：随机、流动性、主成份无权重累积跟踪误差.....	13
图 9：随机、流动性、主成份有权重累积跟踪误差.....	13
图 10：流动性抽样无权重与有权重累积跟踪误差.....	13
图 11：主成份抽样无权重与有权重累积跟踪误差.....	13
图 12：流动性规划无权重与有权重累积跟踪误差.....	14
图 13：主成份规划无权重与有权重累积跟踪误差.....	14
图 14：随机、流动性、主成份无权重跟踪误差（月度移动平均值）.....	14
图 15：随机、流动性、主成份有权重跟踪误差（月度移动平均值）.....	15
图 16：不同抽样复制组合与基准指数平均修正久期、凸性差异.....	16
图 17：不同抽样复制组合与基准指数平均到期收益率、票面利率差异.....	16
表 1：指数与抽样复制债券组合走势统计.....	9
表 2：抽样复制债券组合与基准指数持有期收益差异比较.....	11

一、前言

指数化投资也成为被动投资，是以长期收益和优先管理为目标。其核心的方式是通过各种优化复制策略，从指数成分券中选取有代表性的标的券，用以跟踪标指数的走势，从而达到降低管理成本、提高收益的目的。相比较主动投资策略，指数化投资具有以下几类优势：一是指数化投资可用较小资金获得一个非系统性风险高度分散的投资组合，这在非单边市场中、在长期投资的区间中都是对投资者较为有利的一项特点；二是指数化投资由于是被动投资，因此其运营成本和交易成本较为低廉；三是指数化投资的资金可以被全部用来跟踪指数，资金利用率较主动管理的基金要高，且因为纯指数化投资完全复制指数，其透明度也非常高。因此，指数化投资逐渐获得投资者的青睐。

目前，我国指数化投资主要应用于股票市场。相对于股票市场而言，债券市场中的指数化投资并不多见，其主要原因是由于我国债券市场体制方面的原因，我国债券市场流动性、透明度和成交连续性等方面，严重落后于发达国家。另一方面，由于我国债券市场投资者结构较为单一，银行占据着债券市场绝对数量的份额，使得我国债券市场不能满足不同投资者的风险偏好。未来，随着我国债券市场体制、规模、效率等方面的改善，债券指数和指数化投资必然逐步获得市场认可，并迎来高速发展时期。

本次研究以工行中票 100 指数为研究样本，具体探讨不同抽样复制方法在指数构建、投资组合样本选取、事后组合调整等方面的实际应用，并选取出最优的抽样复制策略。虽然，无论对于债券或者是股票而言，抽样复制方法是资金管理者自身的主观行为，但鉴于目前我国债券市场指数化投资策略市场认可度不高等原因，指数公司可以通过抽样复制方法，进一步丰富现有的指数产品库，引领市场对指数化投资的关注度。因此，本次研究的核心目标是通过检验不同抽样复制策略，为进一步丰富现有指数产品和未来指数的创新奠定基础。并通过对指数构建、跟踪复制、事后调整方案等一系列解决方案，为投资者提供更加全面的市场服务。

二、相关指数跟踪方法回顾

指数化投资核心的内容就是构造并保持一个能尽量拟合指数收益的跟踪组合，即如何跟踪指数的问题。指数化投资理论发展至今，产生了各种不同跟踪指数的方法，按其基本原理不同，可以分四大类：（1）完全复制法；（2）抽样复制法；（3）多因素风险模型最优化法（非线性最优化法）；（4）衍生品复制法；（5）随机情景模拟法。

（一）完全复制法

该方法试图采用完全复制指数样本股的方法构建跟踪组合,即将目标指数中包含的所有成份股全部作为一个跟踪组合,并以指数编制中自然权重作为跟踪组合中每个债券的权重,以此方式构造跟踪组合的方法称为全样本复制法。全样本复制指数虽然能够使得跟踪组合与指数具有最优的拟合,但是,如果指数成份股过多,在计算跟踪组合中各债券投资比例额的时候,会有过多无法实现买卖的零散股产生,导致产生一定的跟踪偏差,而且,维持或调整全样本复制跟踪指数会带来较大的交易费用,最终也必将产生较大的跟踪误差。所以除非所跟踪指数是一个具有较少样本股的成份股指数,或者指数基金本身已经具有非常庞大的规模,可以对全样本复制中可能出现的不利因素忽略不计,否则通常不采用全样本复制跟踪的方法。

（二）抽样复制法

抽样复制法指的是采用一定的随机或非随机抽样方法从构成指数的成份股中抽取若干债券构成用于复制指数跟踪组合的方法。相对全样本复制的方法,该方法在控制跟踪误差的基础上进一步减少了组合维护及再平衡的所需的费用,所以抽样法复制指数具有较强的可操作性。抽样法复制指数按照其抽样原理可以进一步分为三类:简单随机抽样复制,大权重样本复制,分层抽样复制。

1、简单随机抽样复制

根据简单随机抽样原则,从指数成份股中随机抽取指定数量的债券或对指定的债券组合赋予随机的投资比例权重,按以此方法构造跟踪组合跟踪指数称为简单随机抽样复制指数。该方法通常不用于指数跟踪的实践,可主要作为其它非全样本复制跟踪指数的方法的一个比较和参照的基准,用于验证该方法的有效性。举例说明,如果采用某种部分样本股复制指数的跟踪方法,其跟踪效果和随机抽样相同数量债券的跟踪效果相当,那么可在一定程度上说明该复制指数的方法并没有显著的优化跟踪效果。

2、大权重样本复制

大权重样本复制是众多抽样复制法中最简单的一种,该方法抽取指数样本权重最大的若干只债券,并以它们在指数中的自然权重作为投资比例向量构成跟踪组合。隐含在该方法背后的原理就是:指数的大部分收益率可以被一些数量相对少的权重较大的债券的收益所解释,所以仅需通过买卖指数成份债券中的这些权重大的债券就可以在很大程度上实现指数化投资的目标,而且相对于买卖小权重债券,这些债券的流动性风险也会更小,减少了交易成本,这也进一步减少了跟踪误差。

3、分层抽样复制（单元格匹配法）

分层抽样复制法是国际上较为流行的指数复制方法,同时也是构建其他跟踪

复制方法的基础。其核心就是根据不同债券基本要素分类基础上,按照一定规则抽取标的券的方法。该方法先按一定标志对总体中各单位进行分类,通过分类把总体中标志值比较接近的单位归为一组,尽量减少各组内的差异程度,再从各组中按简单随机抽样的原则抽取样本单位,得到的样本通常对总体具有更大的代表性,因而抽样误差也就相对缩小了,分层抽样复制指数运用了相同的原理。

在层化取样的指数化方法中,指数中所包含的债券被分成不同的单元,每一单元里的债券都具有某种相同的特征,一般用以下几种债券特征来分解指数:(1)久期;(2)息票率;(3)到期期限;(4)市场板块(国债,公司债或抵押债);(5)信用等级;(6)可提前赎回因素;(7)偿债基金”。例如,按照图 1 可以将所有债券分为 $3 \times 3 \times 3 = 27$ 个单元格,进而对各个单元格进行抽样。

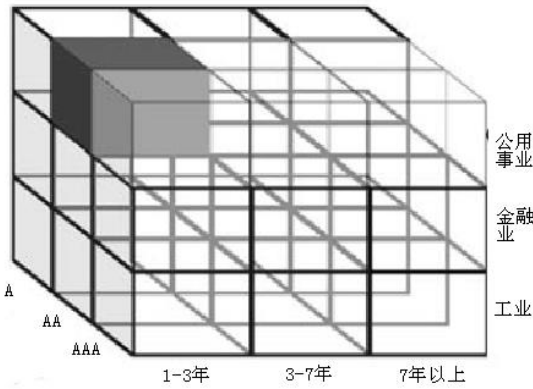


图 1: 分层抽样复制标准划分

对于分层抽样法来说,要针对每个单元格的债券进行基本要素计算,例如:平均收益率、平均久期、平均利差久期、平均凸性、债券数量、发行者数量、流动性、单元格所占比重等方面,进而通过匹配各个单元格的基本要素进行抽样。此种方法的优点是复制精度高、跟踪误差小。其缺点是:(1)复制结果随机性较高,难以控制跟踪误差;(2)事后不能有效动态分析跟踪误差的来源;(3)在可控跟踪误差基础上,不能进行有效的积极管理策略的制定。

(三) 多因素风险模型最优化法

多因素风险模型最优化法是基于计算机程序的数量化方法,其最优化方法可以采用二次规划、遗传算法等不同的方法进行。其核心思想是:(1)确定影响债券价格不同风险因素。例如,期限结构风险、部门风险、信用风险、期权风险等方面,对历史数据进行协方差矩阵分析,找出各个风险因子的方差与相关系数。

(2)对所挑选出的债券历史收益率变化与风险因子进行线性回归,得到各个债券对不同风险的敏感性度量。(3)确定目标方程与约束条件。(4)通过不同的优化算法进行目标方程的最大化或最小化,从而得出标的债券的权重。

该方法的优点在于:(1)跟踪误差较小;(2)权重的确定完全量化,程序化

程度较高（如巴克莱指数的“OPTIMIZER”系统），具有客观性特点。（3）跟踪误差可以进行定量分析（如图 2），进一步可以分析系统性和非系统性（特殊）跟踪误差的来源，从而找出导致跟踪误差的因素，如期限不匹配、信用不匹配、部门结构不匹配等方面，并可以得到定量结果，形成系统性的风险提示报告，有助于指数投资人进行系统风险的管理，也有助于未来定期样本调整方案的确定。另一方面，该方法的缺点在于：（1）计算方法复杂；（2）对历史数据质量要求较高，对不同风险因子的方差与相关性分析缺乏数据积累；（3）数据处理量巨大。

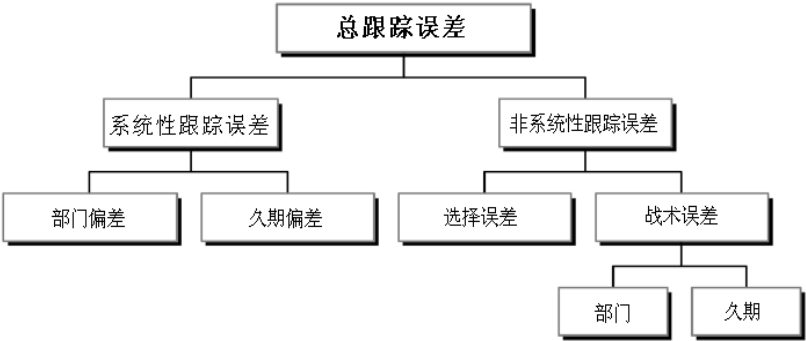


图 2：跟踪误差分解示意图

（四）衍生品复制法

债券衍生品（如国债期货、利率互换）具有流动性高、交易成本低、杠杆交易等特点，可以快速地对由期限风险引起的跟踪误差进行资产配置转换。其次，对于指数基金需要大量现金留存以保证为预期的基金赎回风险来说，其跟踪误差偏离程度一般较大，因此，可以利用衍生品保证金交易的特点，用来快速修复由于持有大量现金引起的跟踪误差。最后，对于初期运营的指数基金来说，由于暂时的标的债券在市场无法获得，可以通过灵活的衍生品复制策略进行资产配置，尤其是针对期限长、信用级别较低的债券品种来说，其市场流动性一直存在较大问题。

（五）随机情景模拟法

以上几种方法，完全是基于对历史数据的分析方法，对数据的质量要求较高。同时，这些方法缺乏对市场前瞻性的考虑与分析。基于以上的分析，国际上提出了更为复杂的随机情景模拟法，通过对债券市场不同风险因子进行情景模拟，并给定未来不同情景发生的概率分布，综合计算出未来标的债券的价格分布。最后根据最优化方法计算各个标的债券的权重。该方法的主要优点在于：通过主动调整风险暴露因子，可以制定细化的量化投资策略。

三、分层抽样复制方法

本次研究主要以分层抽样复制为主要研究方向，通过对不同分层抽样的研究，

分析各种方法的优缺点。但其核心仍是分层抽样方法，总结起来，分层抽样策略主要有以下几个步骤：

步骤 1：分层指标的确定。分层抽样复制策略的核心是如何对指数成分债券进行合理有效的分层，不同的分层方式会导致抽样复制的效果不同。一个好的分层方式是分层指标能够全面、有效、合理地代表成分债券的特征，并且各个指标没有明显的相关性。同时，要保证成分债券能够均匀地分布在每个盒子中，不至于出现有的盒子集中过多债券，而有的盒子却没有债券情况的出现。分层指标过多会导致各个指标相关性过高，同时会使盒子数量过大，进而会出现债券分布不均的情况。相反，分层指标过少会不能有效区分不同债券在不同市场环境下的走势特征，进而使抽样复制策略不能有效跟踪标的指数。因此，分层指标的确定是整个分层复制策略的核心，需要进行大量的分析与检验。针对不同指数样本，前期要合理地确定不同分层指标，使之能够合理地代表成分债券的特征。

步骤 2：抽样债券数量的确定。通过把不同债券放进不同指标盒子当中，下一步就是对债券的选取。每个盒子当中债券选取的数量要依据盒子债券数量占整个债券指数债券数量和所规定的抽样数量而定，这样能够保证不同特征债券在抽样债券总数量相对比例不变，从而使抽样效果更为接近标的指数。

步骤 3：抽样债券的选取。在确定每个盒子的抽样数量以后，就可以进行债券的选取。通常对于股票指数来说，每个盒子股票其基本特征较为接近，同时股票的交易方式和价格连续性等方面较好，并不需要在针对每个盒子股票做额外的区分。因此，可以采用随机抽样的策略从而实现股票的选取，这可以有效避免人为因素的干预，同时也可以使由于抽样复制过程中所积累的系统性误差大大降低。然而，对于债券来说，由于债券市场与生俱来的价格连续性和透明度较低，使债券的选取不能简单通过随机办法来进行。主要原因是：一是通过随机方式选取会包含流动性较差的债券，这会使指数复制跟踪较为困难；二是在指数定期样本调整时，流动性较差债券不能及时买入卖出，从而会加大指数的跟踪误差；三是当遇到市场大幅波动时，流动性较差的债券所积累的系统性风险较大，会受到市场较大影响，加剧了指数波动性和跟踪复制的难度。因此，一般来讲，对于债券指数的跟踪复制要求一般选取流动性较好的债券，这主要是因为流动性较好的债券能够及时反映市场的变化，并且能够代表不同权重的市场走势。因此，可以较好的跟踪标的指数的走势。然而，对于流动性指标的选取，是一项极为复杂和困难的事情，需要长期关注。

步骤 4：抽样债券权重因子的确定。权重因子的确定是整个抽样复制过程的最后一步，确定权重因子是为了使抽样债券所占比重等于标的指数债券所占的各自比重，从而使抽样债券可以进一步匹配标的指数的风险特征。一般来讲可以有两种方法可以实现：一是每个盒子所挑选出债券市值总和占比等于标的债券盒子债券市值总和占比，并使所挑选出来的债券比重等比例扩大；二是更为复杂和科

学的方法，在方法一基础上同时满足每个债券相对权重保持不变的限制，从而实现权重因子的确定。本文所选用的方法为第二种方式。

四、分层抽样方案设计

本次研究主要目的是测试分层抽样复制策略在实际应用过程当中的效果，重点检验不同分层抽样复制方法的实际结果，通过比较不同方法，找出最优的跟踪复制方法。本次研究计划通过以下 10 种不同分层抽样方法进行比较研究，分别是：（1）随机无权重抽样；（2）随机有权重抽样；（3）流动性无权重抽样；（4）流动性有权重抽样；（5）主成份无权重抽样；（6）主成份有权重抽样；（7）流动性整数规划无权重抽样；（8）流动性整数规划有权重抽样；（9）主成份整数规划无权重抽样；（10）主成份整数规划有权重抽样。

本次选择的样本为工行中票流动性 100 指数，其中包含 100 只流动性高的银行间中期票据，计划在样本空间里选出 50 只作为复制债券，追溯时间为 2011 年 1 月 1 日到目前。按照分层抽样步骤：首先我们确定了以剩余期限和信用评级为维度的二维抽样指标的确定。剩余期限划分为：1-2 年、2-3 年、3-4 年和 4 年以上；信用评级分为：AAA、AA+和 AA。一共将整个样本划分为 12 个格子；其次，每个格子样本抽取数量和权重因子依照步骤 3 与步骤 4 加以确定；最后，是样本的选取。选取方法分为 5 大类共 10 种方法：一是随机抽样。每个格子按照所选取的标准，随机选取出指定的数量组成抽样结果；二是流动性抽样。以指数公司对样本空间流动性打分为标准，优先选取流动性最好的指数，直到满足样本数量；三是以流动性主成份打分为标准，优先选取流动性最好的指数，直到满足样本数量；四和五是以线性规划为基础，按照流动性最大为目标条件，进行整数规划，选取出规定样本数量。

五、结果分析

指数化投资不同于主动投资，主动投资主要关注的是绝对收益，而指数化投资更加注重相对收益和跟踪误差的结果。跟踪误差是衡量一个跟踪复制债券组合的重要的好坏标准，较低的跟踪误差能够较好的表现指数的走势特征，并且能够及时地反馈市场的变化，而较高的跟踪误差表明跟踪复制的债券组合不能表现出指数的基本特征，在很多情况下会与指数走势相背离，并且在市场剧烈波动的情况下，会加剧自身的风险累积。因此，本次 10 种不同的指数复制策略分别以指数走势、相对收益与跟踪误差为主要评判标准。

（一）指数走势与样本周转率比较

图 3 展示了中票流动性 100 指数的走势与不同抽样复制策略的走势，从图中我们可以直观看出：以剩余期限和信用级别两个维度可以较好地刻画样本债券的

基本特征,并且通过分层抽样复制策略,复制债券组合可以较好的跟踪基准指数。同时,图 3 也代表了指数与不同方法的累积收益率走势,我们可以看出,最上方两条指数分别是随机无权重与随机有权重两条债券组合走势,从样本区间内累积收益率来看,随机抽样的累积收益率为最高。但需要指出的是,一方面,指数投资并不以绝对收益率来判定好坏,虽然随机抽样累积收益率最高,但是这两条在组合在样本区间内的波动率也是所有方法中最高的,且高于中票 100 指数的波动率;另一方面,由于随机抽样过程中,流动性较差的债券也会包含在该组合当中,而本次研究没有考虑到交易成本因素,因此,如果加入交易成本因素以后,随机抽样的效果会远低于以流动性选样的方法。同时,流动性较差的债券,在遇到市场行情大幅波动情况下,也会出现大幅折价的情况。例如今年 6 月以来出现的银行间市场资金紧张的情况,随机抽样的债券组合价格下降也是最快的,随机抽样无权重组指数由 2013 年 6 月 3 日的 109.313 点下降到 2013 年 6 月 25 日的 107.9115,降幅 1.3%,而同期中票 100 指数值下降了 1.2%。

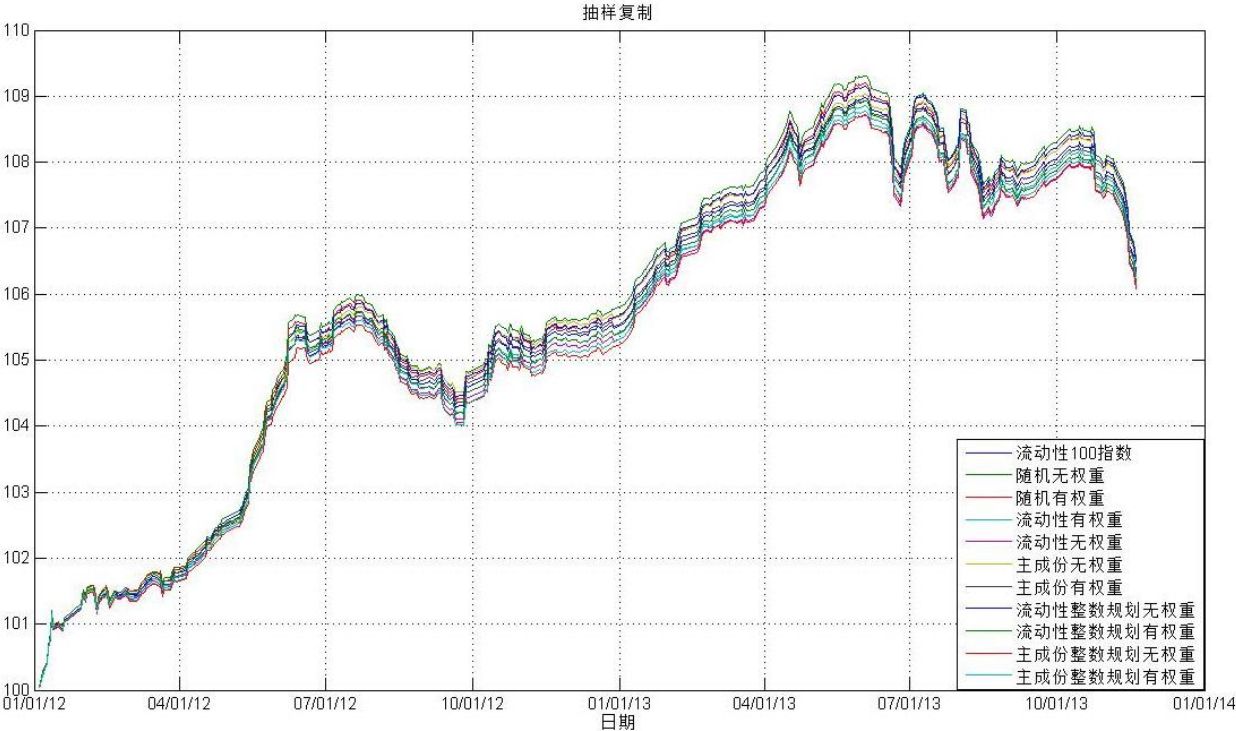


图 3：流动性 100 指数与 10 种抽样复制组合走势

表 1：指数与抽样复制债券组合走势统计

	工行中票 100 指数	随机抽样 无权重	随机抽样 有权重	流动性抽样 无权重	流动性抽样 有权重	主成份抽样 无权重	主成份抽样 有权重	流动性规划 无权重	流动性规划 有权重	主成份规划 无权重	主成份规划 有权重
最大值	108.9	109.3	109.1	108.7	108.7	109.0	108.9	109.1	108.9	108.7	108.8

最小值	100.0	100.1	100.0	100.0	100.0	100.0	100.0	100.0	100.1	100.0	100.0
极差	8.928	9.26	9.14	8.72	8.67	8.98	8.88	9.108	8.877	8.670	8.812
样本标准差	2.384	2.448	2.407	2.315	2.289	2.376	2.368	2.430	2.322	2.313	2.347
周转率	42.43%	69.72%		44%		49.7%		43.14%		48.84%	

表 1：指数与抽样复制债券组合走势统计

从表 1 描述了基准指数与抽样复制组合走势特征，我们可以看出：一方面，随机抽样的抽取方法的极差与波动率高于其他的流动性选取方法；另一方面，样本区间内，通过流动性和主成份选样的债券组合的波动性也低于中票 100 指数和随机选样方法。这两方面表明：流动性较高的债券组合会降低组合的波动率，因为流动性较高的债券能够快速反映市场的信息，其风险能够得到充分的释放，而流动性较差的债券，遇到市场波动较大的时候，价格变化的幅度也远高于流动性好的债券，因而其波动率也较高。

其次，在流动性和主成份抽样复制的四种方法中，通过计算权重因子的组合的波动率低于无权重因子的组合，因此，通过计算权重因子也会降低样本区间组合的波动率。

再次，从表 1 各种分层抽样复制方法周转率结果来看，以指数公司流动性筛选方法选样的周转率最低，而随机抽样的周转率最高，以主成份流动性选样的方法居中。该结论似乎与我们预期的结果并不完全一致，主要是主成份流动性选样周转率较原方法高。理论上，主成份流动性选样由于包含较多解释变量，同时对各解释变量能够定量的给出权重因子，因此，主成份方法的稳定性和连续性理论上应该高于我们原方法。我们分析主成份方法周转率过高的主要原因有以下两点：一是中票 100 指数自身的周转率较高，因此，导致主成份方法的稳定性不能有效发挥；二是中票 100 指数自身的选样标准就是以原方法进行的，主成份只是在此基础上重新进行了流动性排序，属于二次筛选。因此，主成份方法的有效性和连续性会受到原方法的影响，不能实现我们的最初预期。未来，应该在原始数据基础上分别进行流动性的筛选，从而能够客观地比较两种不同方法的优劣，进而优化我们流动性选样的方案。

最后，下图描述了随机、流动性和主成份抽样方法 7 次调样过程中，相同债券权重因子变化绝对值，该值表示了调样过程中相同债券在相邻期间里权重因子的变化的程度，权重因子变化绝对值越高代表了债券组合调整规模、次数以及调

样成本越高。从图中我们可以看出：平均来讲，随机抽样权重因子的变化绝对值最大，流动性抽样次之、主成份抽样绝对值越小。这表明主成份抽样方法在不同样本期间里，相同债券权重因子的调整成本小于流动性抽样和随机抽样两种方法。但是，从个券权重因子调整数目来看，对于每种方法，即使债券仍然留在下一期样本空间里，其权重因子仍然需要逐个调整，这必然增加复制组合调整成本，降低抽样复制的可操作性。未来，需要在控制调整样本基础上，优化抽样复制组合，从而达到降低实际交易成本的目的。

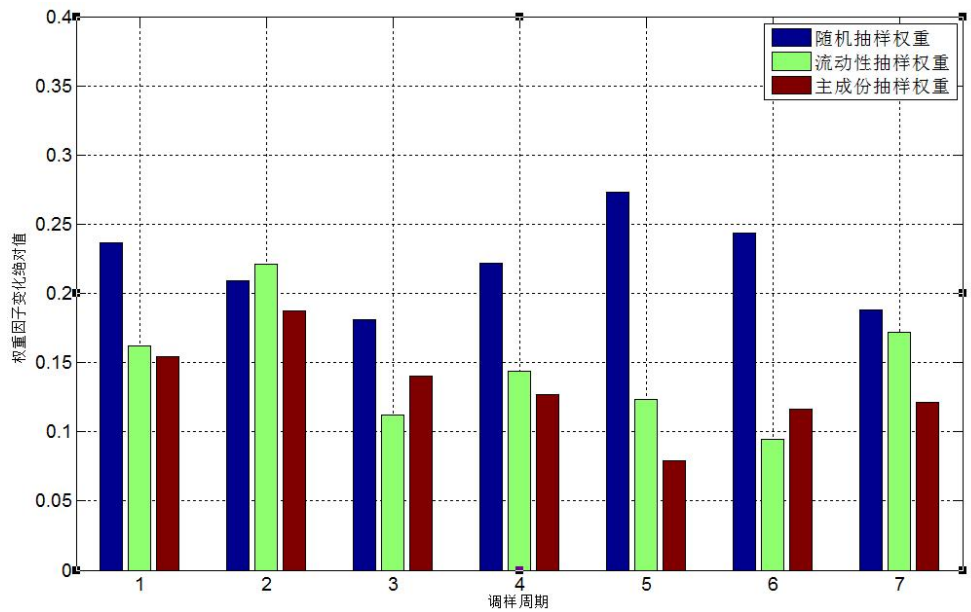


图 4: 不同抽样方法调样过程中相同债券权重因子变化绝对值

（二）不同抽样方法持有期相对收益比较

表 2 与表 3 分别给出基准指数与抽样复制组合样本区间收益率均值与差异比较。从图 2 与表 3 我们可以看出：除了随机抽样方式选出的债券组合以外，以流动性为筛选条件的各种抽样方式的持有期收益均低于基准中票 100 指数收益。其隐含的逻辑是：流动性与持有期收益互为权衡关系，即流动性高的组合其相对收益率较低，而流动性较低的组合其相对收益率较高。其结果符合关于流动性溢价理论，该理论认为：相同信用级别的债券利差可以部分用流动性溢价来解释，其

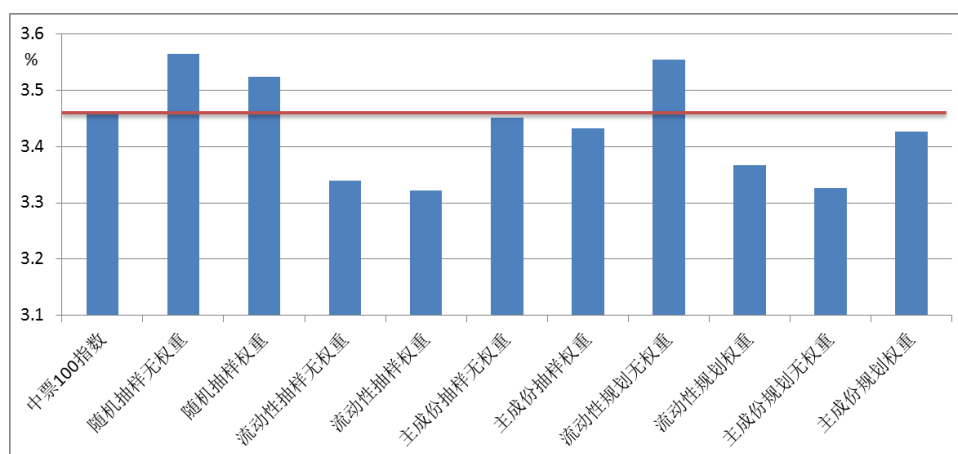


图 5：基准指数与抽样复制债券组合样本区间年化收益率（%）

流动性溢价可以平均补偿 13-20 基点，这意味着流动性高的债券，其价格越高，但其持有期收益率会下降。因此，对于没有进行流动性筛选的随机选样债券组合来说，其包含一部分流动性价差的债券品种，导致其持有期的收益率会高于以流动性筛选为条件的抽样复制组合。

	随机抽样无权重	随机抽样权重	流动性抽样无权重	流动性抽样权重	主成份抽样无权重	主成份抽样权重	流动性规划无权重	流动性规划权重	主成份规划无权重	主成份规划权重
均值 (%)	0.00040	0.00025	-0.00047	-0.00054	-0.000277	-0.000108	0.00036	-0.00036	-0.000517	-0.000133
年化均值 (%)	0.10407	0.06427	-0.12043	-0.13843	-0.00843	-0.02783	0.09397	-0.09383	-0.13353	-0.03433
偏度	0.1833	-0.013	0.0277	0.1166	-0.1938	-0.31023	0.43999	-0.9924	0.7406	0.3286
峰度	13.916	12.64	7.4423	11.987	6.7548	12.4224	9.0742	9.4514	12.1917	8.6463

表 2：抽样复制债券组合与基准指数持有期收益差异比较

另一方面，对两种不同流动性选样的抽样债券组合持有期收益来看，以主成份选样方法的债券组合其持有期收益率同样高于以原方法选样的债券组合，如果按照流动性溢价理论解释，我们可以简单推论认为：原方法流动性筛选方式在此次抽样复制选样过程中同样优于主成份流动性筛选方式，即原方法所选取的债券流动性要高于主成份分析方法。但是正如前面对周转率的解释一样，主成份流动性选样对数据要求较高，本次的结果是建立在中票 100 指数原流动性选样基础上。因此，是否原方法在判定流动性指标上一定优于主成份分析方法，有待于下一步继续跟踪分析。

（三）不同抽样方法跟踪误差比较分析

跟踪误差是指组合收益率与基准收益率之间差异的标准差，反映了被动投资

的相对风险。跟踪误差是衡量指数化投资的核心指标之一，对固定收益指数投资来讲，纯被动投资的跟踪误差一般要求小于 25BP，而各种增强型的指数化投资的跟踪误差一般在 50-200BP，高于 200BP 的投资跟踪误差展现了主动投资的特性。

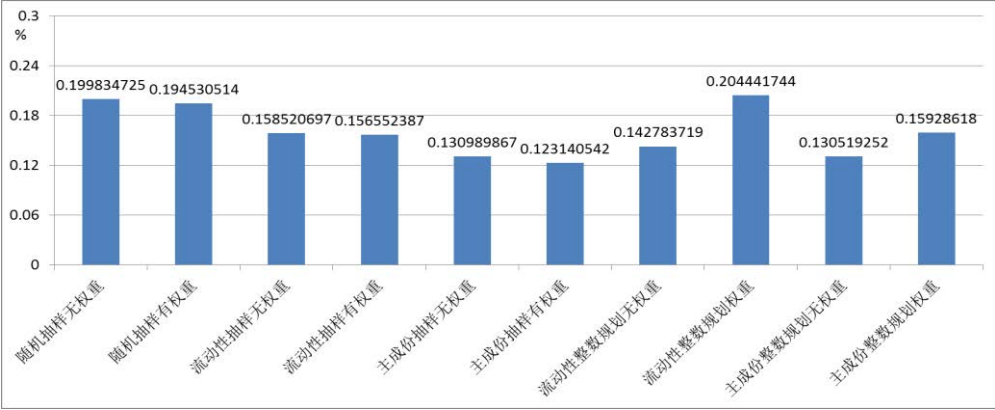


图 6：样本期间不同抽样复制方法年化跟踪误差比较

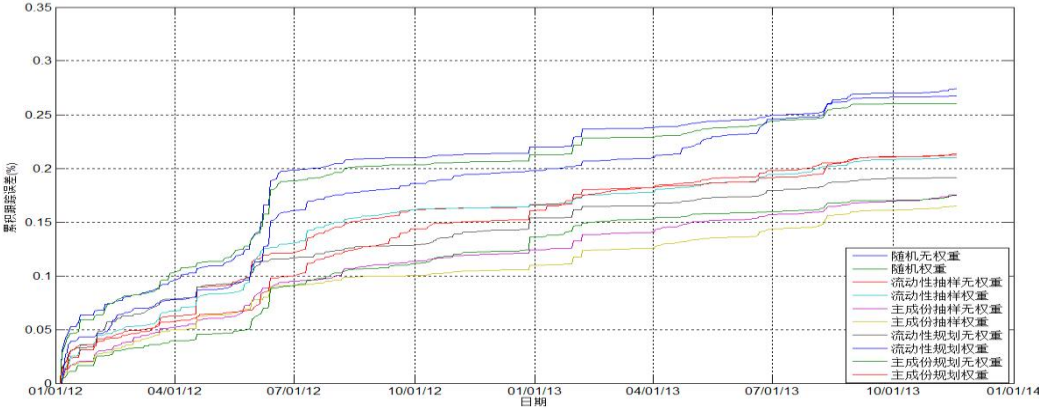


图 7：样本期间不同抽样复制方法累积跟踪误差比较

图6展示了各种不同抽样复制方法对中票 100 指数样本区间的年化跟踪误差。我们可以看到：一方面，采用流动性抽样与主成份抽样方法的年化跟踪误差均小于其他方法，而采用随机复制的跟踪误差和流动性规划方法的年化跟踪误差最高，分别达到 0.1998%和 0.2044%，另一方面，在随机抽样、抽样流动性和主成份方法中，计算权重因子的方法均小于无权重因子的方法，最小的为采用主成份抽样有权重因子的方法，年化跟踪误差为 0.123%（12.3BP），这表明，通过计算权重因子，可以较好地跟踪基准指数，有效地降低跟踪误差。最后，在没有采用分层复制的四种整数规划的方法中，计算权重因子的跟踪误差却均大于没有计算权重因子的跟踪误差，这表明，分层抽样方法能够更好的匹配指数走势，提高复制精度，而权重因子不能独立地提高复制的精度。

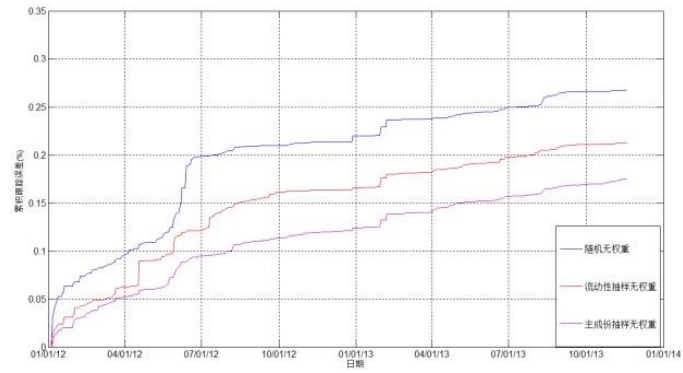


图 8：随机、流动性、主成份无权重累积跟踪误差

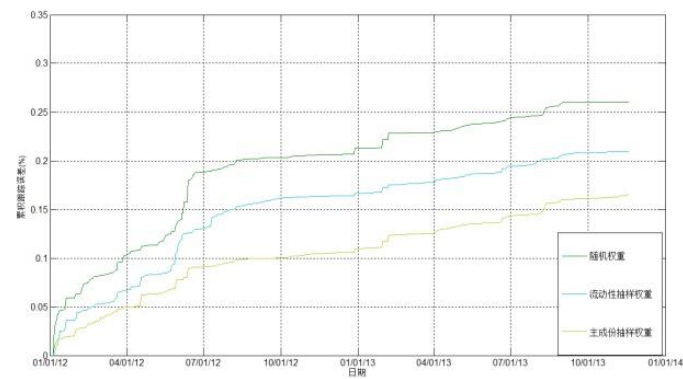


图 9：随机、流动性、主成份有权重累积跟踪误差



图 10：流动性抽样无权重与有权重累积跟踪误差

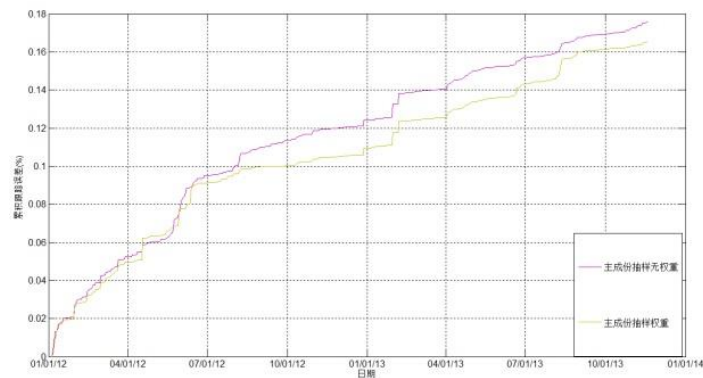


图 11：主成份抽样无权重与有权重累积跟踪误差

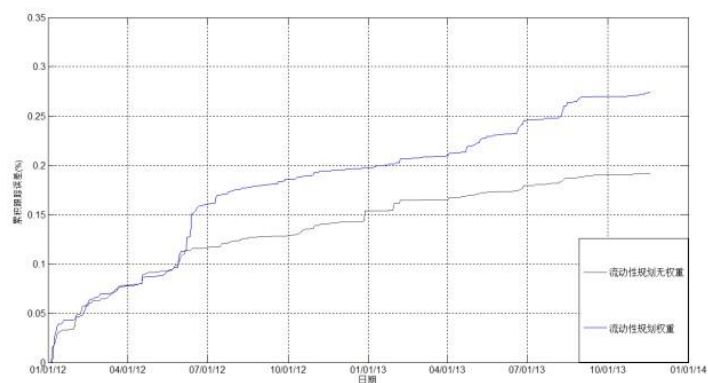


图 12：流动性规划无权重与有权重累积跟踪误差



图 13：主成份规划无权重与有权重累积跟踪误差

图 7 与图 8-13 展示了各种方法在样本区间里的累积跟踪误差走势。与前面的结论相一致，随机有（无）权重的方法累积跟踪误差最高，流动性抽样有（无）权重的方法次之，而主成份抽样有（无）权重方法累积跟踪误差最低。同时，各种方法（除整数规划方法以外）在整个区间内，有权重因子的跟踪误差均小于没有权重因子的方法。

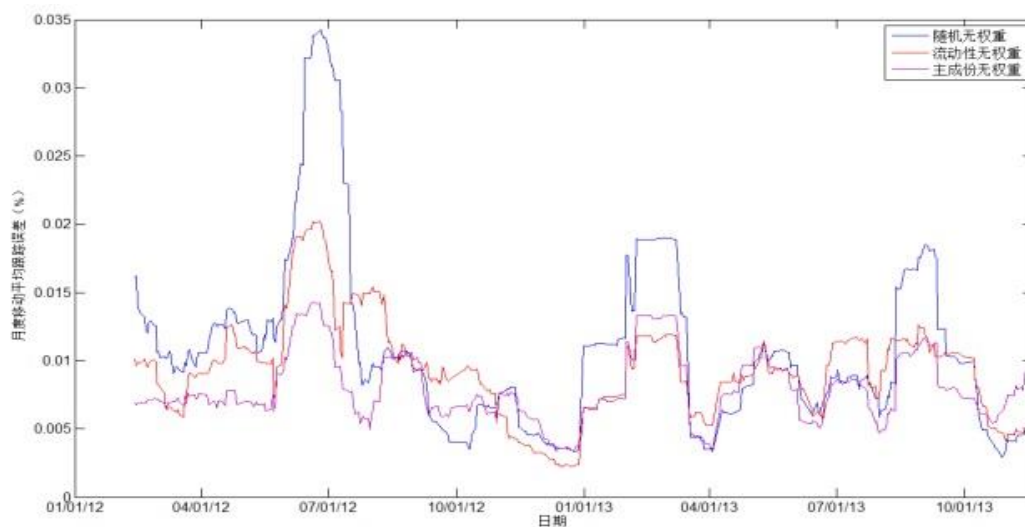


图 14：随机、流动性、主成份无权重跟踪误差（月度移动平均值）

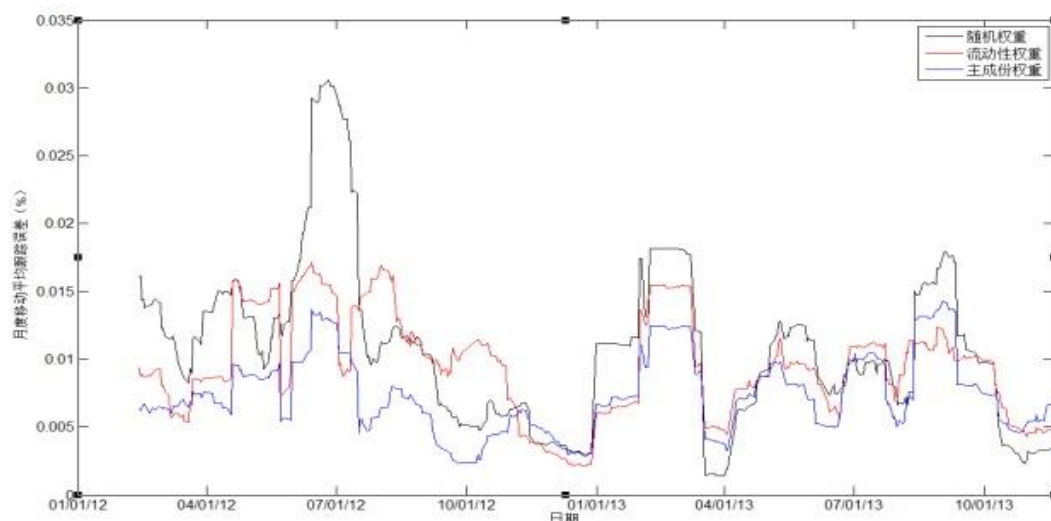


图 15：随机、流动性、主成份有权重跟踪误差（月度移动平均值）

图 14 和图 15 展示了随机、流动性、主成份三种方法跟踪误差的月度移动平均时间序列。从图上我们可以看出两点：一是采用随机方法的月度移动平均跟踪误差最高，流动性方法次之，主成份最低；二是在 2011 年 6 月、2012 年 3 月、7 月和 9 月期间，移动平均跟踪误差均出现大幅走高的情况，在这些时段里，我们发现市场收益率曲线均发生大幅波动的情况，移动平均跟踪误差大幅上升表明了不同抽样方法在这些时间段里，由于其自身特征与基准指数特征的不同而产生的走势不一致。出现跟踪误差大幅走高的原因有很多，本次研究并未涉及，未来需要进一步深入分析，这有助于抽样指数建立过程中对跟踪误差的有效控制。

六、跟踪误差产生的原因

由于分层抽样复制方法只是选取指数部分成分债券作为跟踪标的，不可避免地会产生与指数不同的走势特征，这是跟踪误差产生的原因之一，同时交易过程中产生的交易成本，也是产生跟踪误差的主要原因之一。虽然分层抽样方法不能根本消除跟踪误差，但通过对误差产生原因的分析，可以进一步控制误差的范围，市值能够有效的匹配基准指数的走势。国际上，对固定收益指数跟踪误差的分析主要采用多因子固定收益模型（Multi-Factor Fixed Income Models），该模型不仅可以用来复制指数组合，同时也能够有效地对跟踪误差进行分解，并通过最优化技术进行有效的成分权的调整（如 Barclays “Optimizer” 系统）。未来，需要进一步进行跟踪研究。

本部分主要从定性的角度来展现跟踪误差与抽样复制组合特征的关系。从久期差异、凸性差异、到期收益率差异和平均票面利率差异等方面找出误差产生的原因。图 16 和 17 展示了随机、流动性抽样与主成份抽样组合与基准中票 100 指数在平均久期、平均凸性、平均到期收益率和平均票面利率四个方面的差异。每个图分为 3 个部分，分别是随机抽样、流动性抽样与主成分抽样，同时每个部

分左面是无权重因子抽样，右边代表有权重因子抽样，从图中我们看出：正是由于久期、凸性、到期收益率与票面利率的差异，造成了我们抽样复制的结果与基准指数走势的差异，即出现跟踪误差的主要来源。同时，我们可以从图中得出以下的结论：

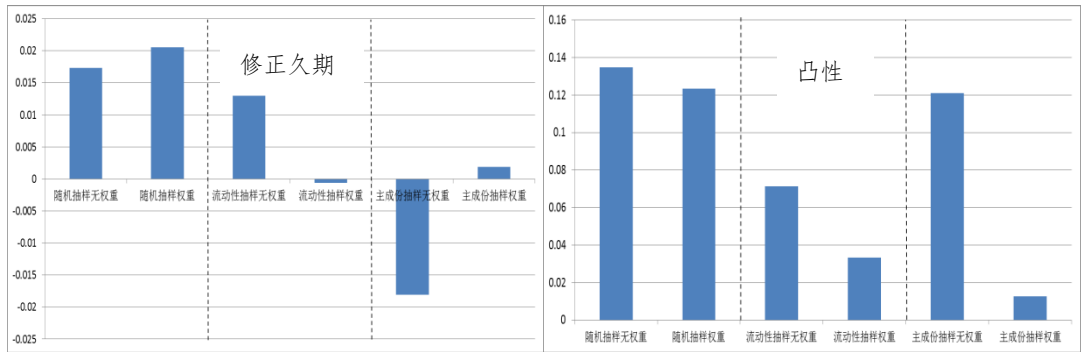


图 16：不同抽样复制组合与基准指数平均修正久期、凸性差异

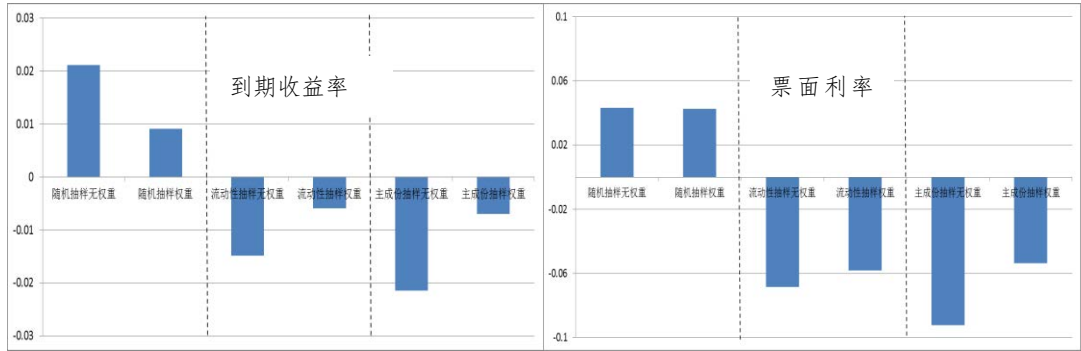


图 17：不同抽样复制组合与基准指数平均到期收益率、票面利率差异

首先，3 种方法有权重与无权重对比来看，通过计算权重因子能够有效降低复制组合与基准指数基本特征的差异，并有效降低跟踪误差。例如，久期差异图中，流动性无权重抽样平均久期高于基准指数 0.013，而通过计算权重因子的抽样方法使得久期差异大幅降低（接近于 0）。再如，平均到期收益率差异图中，主成分无权重抽样方法平均到期收益率低于基准指数 2.1BP, 而通过计算权重因子，使得平均到期收益率差异降低到 0.7BP。虽然从图中我们可以看出权重因子的计算可以降低组合和基准指数的特征差异，但我们尚不能确定其核心作用机制，我们预期是由于权重因子可以有效地确保了分层抽样与基准指数相对权重相同，同时也保证了同一层次里不同券种的相对比例保持不变的原因，使得有权重因子调整的债券组合更加接近于基准指数的特征。

其次，由于流动性溢价的存在，流动性高的债券价格一般高于流动性低的债券，所以其到期收益率与票面利率也低于流动性差的债券。因此，通过流动性和主成分抽样复制的债券组合的平均到期收益率和票面利率均低于含有流动性较差债券的基准指数，并且这也是导致两种债券组合的年化收益率均低于基准指数

年化收益率的主要原因。这也再次验证了流动性与持有收益的互为权衡的关系(Tradeoff)。

最后,修正久期和凸性是衡量利率风险的1阶和2阶因子,对于任何指数复制的债券组合,久期与凸性的不匹配是导致组合走势与基准指数相背离的主要原因。因此,在纯被动指数投资时,抽样复制过程应强化久期的匹配。同时应尽可能提高债券组合的凸性,因为凸性本身是有价值的,可以有效提高组合收益并且降低组合的下跌风险。

七、主要结论

本次研究以工行中票100指数为研究样本,具体探讨不同抽样复制方法在指数构建、投资组合样本选取、事后组合调整等方面的实际应用,并在此基础上,选取出最优的抽样复制策略。通过比较分析,我们得出以下几点主要结论:

一是从持有期收益率来看,流动性和持有期收益率存在着互为权衡的关系。即包含有流动性较差的随机抽样组合持有期收益最高,主成份抽样组合次之,流动性抽样组合最低。其结果符合关于流动性溢价理论。这意味着流动性高的债券,其价格越高,但其持有期收益率会下降。因此,对于没有进行流动性筛选的随机选择债券组合来说,其包含一部分流动性较差的债券品种,导致其持有期的收益率会高于以流动性筛选为条件的抽样复制组合。

二是从无权重和有权重的抽样组合来看,有权重因子的抽样组合在跟踪基准指数走势、降低跟踪误差和匹配基准指数特征方面均好于无权重的抽样组合。在随机抽样、抽样流动性和主成份方法中,计算权重因子的方法均小于无权重因子的方法,最小的为采用主成份抽样有权重因子的方法,年化跟踪误差为0.123%(12.3BP),这表明,通过计算权重因子,可以较好地跟踪基准指数,有效地降低跟踪误差。最后,在没有采用分层复制的四种整数规划的方法中,计算权重因子的跟踪误差却均大于没有计算权重因子的跟踪误差,这表明:分层抽样方法能够更好的匹配指数走势,而权重因子不能独立地提高跟踪指数的精度。

三是从周转率来看,流动性抽样组合周转率最低(44%),主成份次之(49.7%),随机抽样最高(69.72%)。该结论似乎与我们预期的结果并不完全一致,主要是主成份流动性选择周转率较原方法高。理论上,主成份流动性选择由于包含较多解释变量,同时对各解释变量能够定量的给出权重因子,因此,主成份方法的稳定性和连续性理论上应该高于我们原方法。我们分析主成份方法周转率过高的主要原因有以下两点:一是中票100指数自身的周转率较高,因此,导致主成份方法的稳定性不能有效发挥;二是中票100指数自身的选样标准就是以原方法进行的,主成份只是在此基础上重新进行了流动性排序,属于二次筛选。因此,主成份方法的有效性和连续性会受到原方法的影响,不能实现我们的最初预期。

四是从跟踪误差角度来看，主成份抽样组合年化跟踪误差最低、流动性抽样组合次之，随机抽样组合最高。以流动性和主成份流动性抽样组合包含了市场流动性较好的债券，流动性高的债券组合能够及时的反映市场的信息，快速地调整市场价格，可以较好地释放自身的风险，因此，在遇到市场大幅波动的同时，能够较好地匹配基准指数的走势，降低自身跟踪误差偏离程度。相反，以随机抽样为代表的非流动性筛选，流动性较差的债券由于自身风险不能得到有效释放，其组合的跟踪误差会远高于流动性筛选的方法。

五是从跟踪误差月度移动平均值来看，各种抽样复制方法在 2011 年 6 月、2012 年 3 月、7 月、9 月均出现移动平均值大幅增加，这表明了不同抽样复制方法在久期凸性等特征方面与基准指数的差异所致，其市场波动率的大幅变化，会使各种抽样复制组合的走势与基准指数走势相背离。出现跟踪误差大幅走高的原因有很多，本次研究并未涉及，未来需要进一步深入分析，这有助于抽样指数建立过程中对跟踪误差的有效控制。

六是从事后跟踪误差检验来看，不同抽样组合的久期、凸性、平均到期收益率和平均票面利率的差异导致了与指数走势的差异和跟踪误差的产生。同时，权重因子的计算可以极大地提高指数复制的效果，降低组合与基准指数在久期等方面的差异，同时大幅降低跟踪误差，提高抽样复制的精度。最后，久期匹配是提高组合跟踪效率的核心要素之一，而凸性由于其自身对投资者的价值，较高的凸性可以提高组合的收益和降低组合下跌的风险。

综合来看：首先，流动性抽样组合与主成份抽样组合在跟踪复制基准指数方面各有优势，流动抽样组合的年化收益率和跟踪误差方面较主成份抽样组合稍差，而在样本周转率指标来看，流动性抽样周转率低于主成份抽样组合。由于本次研究并没有考虑组合交易成本因素，如果加入交易成本，二者的差异将会更小，未来应值得我们进一步研究；其次，流动性与持有期收益率互为权衡的关系值得我们重视，过多强调高流动性筛选方法，会降低组合的平均到期收益率和平均票面利率水平，进而降低组合的持有其收益；再次，权重因子的计算能够提高抽样组合复制基准指数的效果，是降低组合跟踪误差的必要条件之一；最后，对债券组合的事后跟踪误差、风险度量以及定期调整的分析，是控制跟踪误差，降低交易成本的主要方法。未来通过对因子模型的应用，可以定量地实现对组合定期调整的优化处理，提高组合复制基准指数的精度，以及提高投资者对抽样复制组合方法的可操作性。