

证券研究报告 / 金融工程研究报告

可转债风险模型构建与应用

---可转债量化系列之二

报告摘要：

Barra 股票多因子风险模型的原理在于将股票收益拆解为多个共同因子的收益以及与股票本身相关的特异性收益。共同因子代表着股票所共享的特征，特异性收益则代表着个股未被因子解释的 alpha、特殊事件所带来的收益波动等。使用多因子模型，一方面可以实现协方差矩阵降维，减少计算所需的时间跨度，提升组合优化速度；另一方面可以对特定组合进行绩效归因，分析收益和风险的来源。

借鉴 Barra 模型的理论，使用 26 个行业因子和 10 个风格因子（估值、余额、Beta、动量、波动、流动性、剩余期限、Book-to-Price、Earning Yield、Leverage）搭建可转债风险模型 $R = \beta F + u$ ，对因子暴露数据进行去极值、标准化、正交化处理，采用加权最小二乘法（WLS）拟合计算得到因子收益与转债的特异性收益，拟合结果 R^2 均值 33.73%。

多因子风险模型将转债收益协方差矩阵的估计问题，转化为对因子收益协方差矩阵和特异性收益方差的计算， $\Sigma = \beta \Sigma_F \beta' + \Sigma_u$ ， Σ 、 Σ_F 、 Σ_u 分别代表转债收益协方差矩阵、因子收益协方差矩阵、特异性收益协方差矩阵。在任何时刻 t，采用指数加权移动平均的方式计算因子收益和特异性收益协方差矩阵，取窗口 $T=126$ ，半衰期 $h=45$ ，经过 Newy-West、Eigenfactor、波动率偏误、贝叶斯收缩等调整后，最终所得协方差矩阵偏误统计量均处于置信区间内。

为检验准确性，使用风险模型对中证转债指数组合的波动率进行计算，精确度较高，计算所得组合方差与实际方差相关系数为 0.88。同时，对于可转债基金，可以根据其重仓转债的风格因子暴露与权重占比，对收益进行风险分解，从而判断每个因子对收益的边际贡献。以规模在 50 亿左右的可转债基金 A 为例，选取 2021-2023 年 12 个报告期，过去三年该基金重仓转债的风格不断向低估值、高余额、低换手、正股低 P/B 靠近，在 Beta 因子上的暴露一直比较显著，其他因子暴露无明显变化。

结合 alpha 因子使用风险模型进行组合优化，可以在控制风险的同时实现收益。设置目标函数为 $\max_w w^T F_{\alpha} - \lambda w^T \Sigma w$ ，进行最优化求解，所得最优组合表现相对仅使用 alpha 因子所构建的策略更佳，不同风险厌恶程度(λ)下夏普比率均有所提高，最高为 1.91。未来可以加入更多 alpha 因子进行组合优化，继续提升超额收益。



组合表现	$\lambda=100$	$\lambda=500$	$\lambda=1000$
年化收益	20.27%	18.99%	17.25%
年化波动	13.30%	10.80%	9.06%
夏普比率	1.60	1.76	1.91

相关报告

《雪球产品敲入规模分布估算和市场影响点评》

--20240124

《债基回报改善，久期杠杆回升》

--20240123

《2023 年四季度公募主动权益基金持仓解析》

--20240123

《上月盈利、质量、红利因子表现相对较优》

--20240105

《寻找具有更多“优质特征”的固收+基金》

--20231226

证券分析师：王琦

执业证书编号：S0550521100001

021-61002390

wangqi_5636@nesc.cn

风险提示：以上结果基于数学模型与历史数据，存在模型失效的可能性

目 录

1.	多因子风险模型的构建	4
1.1.	因子的选择	5
1.2.	因子暴露数据预处理	6
1.2.1.	去极值与标准化	6
1.2.2.	正交化处理	6
1.3.	因子收益求解	7
1.3.1.	计算方式	7
1.3.2.	计算结果	8
2.	风险矩阵估计	9
2.1.	偏误统计量	9
2.2.	因子收益协方差矩阵	10
2.2.1.	Newey-West 调整	10
2.2.2.	Eigenfactor 调整	11
2.2.3.	波动率偏误调整 (Volatility Regime Adjustment)	15
2.2.4.	不同调整方式的优化效果	16
2.3.	特异性收益协方差矩阵	17
2.3.1.	Newey-West 调整	17
2.3.2.	结构化模型调整 (Structural Model)	18
2.3.3.	贝叶斯收缩 (Bayesian Shrinkage)	19
2.3.4.	波动率偏误调整 (Volatility Regime Adjustment)	21
2.3.5.	不同调整方式的优化效果	22
3.	风险模型的应用	23
3.1.	风险矩阵降维	23
3.2.	组合收益归因	23
3.3.	组合优化	26
4.	总结	28
5.	参考文献	28
6.	风险提示	29

图表目录

图 1: 市值分组统计 OLS 回归残差平方和累和	8
图 2: 市值分组统计 OLS 回归残差平方和均值	8
图 3: 满足前提条件且因子暴露不为空的转债数量	8
图 4: WLS 回归 R^2 分布图	8
图 5: 行业因子市值加权组合与中证转债净值对比	9
图 6: 经 Newey-West 调整后因子日收益波动偏误统计量	11
图 7: 四种组合的波动预测偏误统计量	12
图 8: Eigenfactor 调整后四种组合的波动预测偏误统计量	13
图 9: 因子特征组合偏误统计量	14
图 10: 调整后因子特征组合偏误统计量	15
图 11: 因子收益截面波动率 CSV 与调整系数 λ_F 走势	16
图 12: 每步调整后的因子日收益波动偏误统计量 (63 个交易日移动平均)	17
图 13: 协调系数 $\gamma_n=1$ 的转债占比	19
图 14: 特异性波动预测分组偏误统计量表现	20
图 15: 特异性收益截面波动率 CSV 与调整系数 λ_S 走势	21
图 16: 每步调整后的特异性收益波动偏误统计量 (63 个交易日移动平均)	22
图 17: 特异性波动预测分组未来 63 个交易日偏误统计量表现	22
图 18: 中证转债实际波动与风险模型计算结果比较	23
图 19: 2021 年转债基金 A 风格因子暴露	24
图 20: 2021 年转债基金 A 风格因子收益	24
图 21: 2022 年转债基金 A 风格因子暴露	24
图 22: 2022 年转债基金 A 风格因子收益	24
图 23: 2023 年转债基金 A 风格因子暴露	25
图 24: 2023 年转债基金 A 风格因子收益	25
图 25: 最小化方差组合净值表现	27
图 26: 不同风险厌恶程度下最优组合净值表现	27
表 1: BarraCNE5 模型风格因子	4
表 2: 转债风险模型风格因子	5
表 3: 风格因子暴露相关性	7
表 4: 各报告期转债基金 A 行业因子暴露	26
表 5: 不同风险厌恶程度下最优组合表现统计	28

1. 多因子风险模型的构建

在投资组合的构建过程中，风险管理的重要程度不亚于对超额收益的追求。马科维茨的均值方差模型奠定了风险研究的基础，随着研究的不断深入，结构化多因子风险模型逐渐发展成熟。

多因子风险模型的原理在于将股票收益拆解为多个共同因子的收益以及与股票本身相关的特异性收益。共同因子代表着股票所共享的特征，特异性收益则代表着个股未被因子解释的 alpha、特殊事件所带来的收益波动等。使用多因子模型，一方面可以实现协方差矩阵降维，减少计算所需的时间跨度，提升组合优化速度；另一方面可以对特定组合进行绩效归因，分析收益和风险的来源。

Barra 针对中国场所设计的 CNE5 多因子模型包括了国家因子、多个行业因子以及风格因子：

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} f_c + \begin{bmatrix} I_{11} & I_{12} & \cdots & I_{1M} \\ I_{21} & I_{22} & \cdots & I_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ I_{N1} & I_{N2} & \cdots & I_{NM} \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ \vdots \\ f_{1M} \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1K} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{N1} & \beta_{N2} & \cdots & \beta_{NK} \end{bmatrix} \begin{bmatrix} f_{S1} \\ f_{S2} \\ \vdots \\ f_{SK} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}$$

式中 r_N 为股票 N 的超额收益； f_c 为国家因子； I_{NM} 代表股票 N 在 M 行业上的暴露因子，即行业哑变量，每支股票截面上只在一个行业因子暴露上取值为 1，其余行业为 0， f_{1M} 代表第 M 个行业因子收益； β_{NK} 代表股票 N 在第 K 个风格因子上的暴露， f_{SK} 代表第 K 个风格因子的收益； u_N 代表股票 N 的特异性收益率。

CNE5 模型使用了 29 个行业因子，以及 10 大类风格因子，主要包括规模、beta、动量、波动率、流动性等，具体如下表：

表 1：BarraCNE5 模型风格因子

大类因子	子类因子	因子描述
Size	LNCAP	股票市值对数
Beta	BETA	超额收益与市场组合收益的回归斜率
Momentum	RSTR	指数加权历史收益均值
	DASTD	历史日超额收益波动率
Residual Volatility	CMRA	累计超额收益范围
	HSIGMA	超额收益与市场组合收益的回归残差波动率
Non-linear Size	NLSIZE	Size 的立方
Book-to-Price	BTOP	市净率倒数
	STOM	过去 1 个月平均换手率
Liquidity	STOQ	过去 3 个月平均换手率
	STOA	过去 12 个月平均换手率
	EPFWD	12 个月市盈率预测倒数
Earning Yield	CETOP	市现率(TTM)倒数
	ETOP	市盈率(TTM)倒数
	EGRLF	3-5 年盈利增长预测
Growth	EGRSF	1 年盈利增长预测
	EGRO	过去 5 年盈利增长率
	SGRO	过去 5 年销售收入增长率
	MLEV	市场杠杆
Leverage	DTOA	负债/资产
	BLEV	账面杠杆

数据来源：东北证券

本文基于 Barra 风险模型的思路，搭建可转债的多因子风险模型，对转债的收益进行拆解并估算风险矩阵，对投资组合进行优化。

1.1. 因子的选择

首先基于可转债正股的中信一级行业分类设置行业因子。由于转债市场较小，某些行业不存在或者只在短期内存在过少数几只转债，因此予以剔除，最终形成 26 个行业因子，包括：交通运输、传媒、农林牧渔、医药、商贸零售、国防军工、基础化工、家电、建材、建筑、有色金属、机械、汽车、煤炭、电力及公用事业、电力设备及能源、电子、石油石化、纺织服装、计算机、轻工制造、通信、钢铁、银行、非银行金融、食品饮料。

由于本文所涉及资产均为中国可转债，转债市场容量小且转债存续期较短，因此每个行业内转债数量有限，如果单独设置国家因子作为截距项以代表市场组合收益，可能将过多的信息剥离出了行业因子，在后续进行协方差估计时会造成较大的偏差。因此本文不再设置国家因子，后续将证明行业因子的市值加权组合即为市场组合。

风格因子层面，根据转债自身的相关性质选择了估值、余额及剩余期限因子，同时模仿 Barra 的方法论计算转债的 Beta、动量、流动性和波动率等量价指标，最终考虑到正股对转债的影响，加入少数股票 Barra 因子。基于 Barra 的处理方式，每个大类因子由小因子的加权组合形成，在保留了信息的条件下，初步解决共线性问题。最终风格因子的组成如下：

表 2：转债风险模型风格因子

大类因子	子类因子	因子描述	权重
估值	价格	转债收盘价 63 个交易日移动平均	0.33
	转股溢价率	转债转股溢价率 63 个交易日移动平均	0.33
	隐含波动率	转债隐波 63 个交易日移动平均	0.34
余额	转债余额	转债余额的自然对数	1
剩余期限	转债剩余期限	转债剩余期限，年为单位	1
Beta	转债 Beta	转债收益对中证转债收益回归的斜率	1
动量	转债动量	转债过去 63 个交易日收益率	0.5
	正股动量	正股过去 63 个交易日收益率	0.5
波动率	转债波动率	转债过去 63 个交易日日收益的波动率年化	0.9
	残差波动率	转债收益对中证转债收益回归的残差波动率年化	0.1
流动性	转债流动性	转债过去 63 个交易日平均换手率	0.5
	正股流动性	正股过去 63 个交易日平均换手率	0.5
Book-to-Price	BTOP	市净率倒数	1
Earning Yield	CETOP	市现率(TTM)倒数	0.66
	ETOP	市盈率(TTM)倒数	0.34
Leverage	MLEV	市场杠杆	0.38
	DTOA	负债/资产	0.35
	BLEV	账面杠杆	0.27

数据来源：东北证券

考虑到转债自身存续较短以及转债市场起步较晚，转债因子计算的区间长度取为一个季度。最后三大类正股因子，子类因子权重与计算方式均与 Barra 保持一致。

1.2. 因子暴露数据预处理

在计算因子收益率前，需要对原始的因子暴露数据进行清洗，以满足风险模型的约束条件，并消除共线性问题。

1.2.1. 去极值与标准化

利用绝对中位差去极值法，对截面上的因子暴露值进行调整：

$$\widetilde{\beta}_k = \begin{cases} \beta_M + 3\beta_{MAD}, & \text{if } \beta_{i,k} > \beta_M + 3\beta_{MAD} \\ \beta_M - 3\beta_{MAD}, & \text{if } \beta_{i,k} < \beta_M - 3\beta_{MAD} \\ \beta_k, & \text{else} \end{cases}$$

其中 β_k 代表某截面上 k 因子的暴露序列， $\beta_{i,k}$ 特指转债 i 的 k 因子暴露， β_M 为序列 β_k 的中位数， β_{MAD} 为序列 $|\beta_k - \beta_M|$ 的中位数，经过去极值处理后的因子暴露序列即为 $\widetilde{\beta}_k$ 。

对去极值后的因子暴露数据，进一步进行标准化处理。基于 Barra 的理论，市场组合在任何风格上都应该是中性的，即暴露为 0，因此对于每个截面上的因子暴露，将原值减去市值加权均值，并除以标准差，处理后的因子暴露将满足市场组合中性条件，同时大致符合标准正态分布：

$$\beta_k^{std} = \frac{\widetilde{\beta}_k - m_k}{\sigma_k}$$

其中 m_k 为第 k 个因子在某截面上暴露序列的市值加权均值， σ_k 为第 k 个因子在截面上暴露序列的标准差， β_k^{std} 为经过去极值和标准化处理后的因子暴露序列。

1.2.2. 正交化处理

将相关性较高的子类因子合成为大类因子，可以初步解决共线性问题，但是在大类风格因子中，仍存在少数因子之间有较高的相关性。对相关性较高的因子，进行正交化处理，可以保有因子的增量信息，同时避免共线性问题。

正交化的处理为，在每个截面上，将待处理的因子对选定因子进行 OLS 回归，得到的残差即为正交化处理后的因子暴露。如果对相关性的因子进行正交化处理，可能会使因子失去经济学意义，并丢失较多的信息，因此本文仅将波动率因子对流动性和 Beta 因子进行正交（波动率因子与两因子间的相关系数绝对值均超过 0.5）。

完成正交化处理后，因子暴露之间的相关性如下：

表 3：风格因子暴露相关性

	估值	余额	Beta	动量	波动率	流动性	剩余期限	BTOP	Earning	Leverage
估值	1.000	-0.201	-0.025	-0.084	0.040	0.012	0.035	0.064	0.034	-0.032
余额	-0.201	1.000	-0.136	0.026	-0.081	-0.216	0.102	0.072	0.001	0.068
Beta	-0.025	-0.136	1.000	0.027	-0.060	0.306	-0.030	-0.243	-0.042	-0.107
动量	-0.084	0.026	0.027	1.000	0.054	0.251	-0.113	-0.325	-0.226	-0.070
波动率	0.040	-0.081	-0.060	0.054	1.000	0.059	-0.036	-0.118	-0.050	-0.030
流动性	0.012	-0.216	0.306	0.251	0.059	1.000	-0.021	-0.326	-0.108	-0.117
剩余期限	0.035	0.102	-0.030	-0.113	-0.036	-0.021	1.000	-0.180	0.045	-0.245
BTOP	0.064	0.072	-0.243	-0.325	-0.118	-0.326	-0.180	1.000	0.284	0.306
Earning	0.034	0.001	-0.042	-0.226	-0.050	-0.108	0.045	0.284	1.000	0.042
Leverage	-0.032	0.068	-0.107	-0.070	-0.030	-0.117	-0.245	0.306	0.042	1.000

数据来源：聚源，Wind，东北证券

1.3. 因子收益求解

1.3.1. 计算方式

前文中提到，本文中不设置国家因子，因此转债风险模型表现为：

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix} = \begin{bmatrix} I_{11} & I_{12} & \cdots & I_{1M} \\ I_{21} & I_{22} & \cdots & I_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ I_{N1} & I_{N2} & \cdots & I_{NM} \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ \vdots \\ f_{1M} \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1K} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{N1} & \beta_{N2} & \cdots & \beta_{NK} \end{bmatrix} \begin{bmatrix} f_{S1} \\ f_{S2} \\ \vdots \\ f_{SK} \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}$$

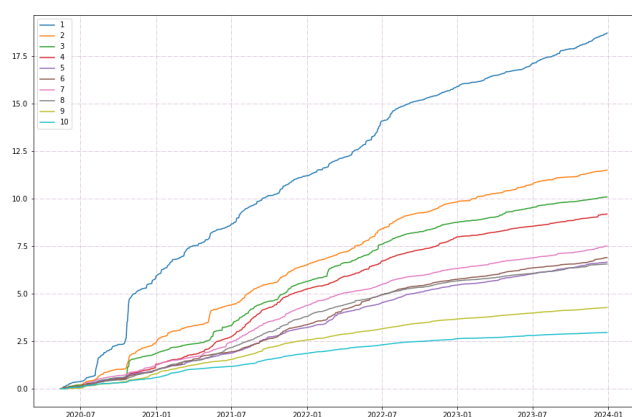
式中 r_N 为转债 N 的收益率； I_{NM} 代表转债 N 在 M 行业上的暴露因子， f_{1M} 代表第 M 个行业因子收益； β_{NK} 代表转债 N 在第 K 个风格因子上的暴露， f_{SK} 代表第 K 个风格因子的收益； u_N 代表转债 N 的特异性收益率。行业因子共有 26 个，风格大类因子共有 10 个。

在每个截面 t 时刻，选取 t 期的转债日收益序列，与 t-1 期（即 t 期期初）的行业因子以及风格因子暴露，代入上述模型，采用加权最小二乘法（WLS）拟合计算 t 期的因子收益。

由于不设置国家因子，不存在国家因子与行业因子暴露之间的共线性问题，因此在回归的过程中不需要保持市场组合在所有行业上暴露为 0。

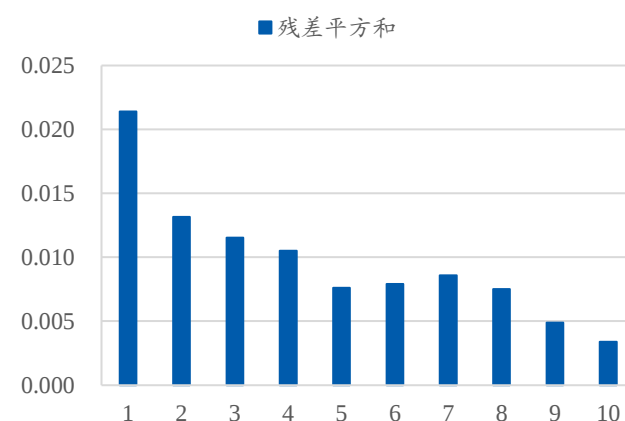
在设置 WLS 的权重时，假设转债的特异性收益波动率与转债的市值平方根成反比，即市值越小波动越大。为验证其合理性，首先使用 OLS 进行回归，按照转债的市值大小进行分组，统计各组残差平方和：

图 1：市值分组统计 OLS 回归残差平方和累和



数据来源：聚源，Wind，东北证券

图 2：市值分组统计 OLS 回归残差平方和均值



数据来源：聚源，Wind，东北证券

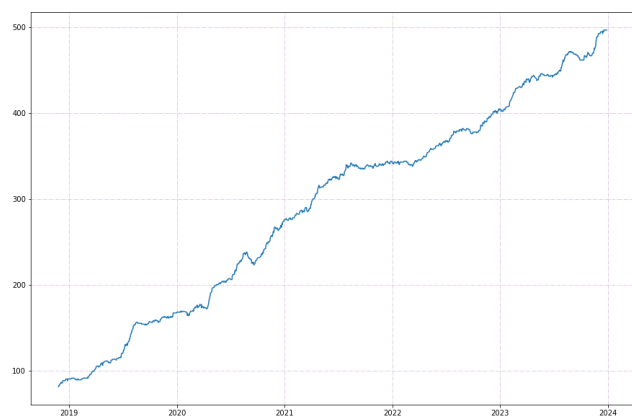
从第 10 组至第 1 组，转债的市值减小，同时 OLS 回归所得的残差平方和的时间序列累和或全样本内的均值也呈上升趋势。因此在进行 WLS 回归时，为解决异方差问题，对于大市值的转债赋予较大的权重，对小市值的转债则赋予较小的权重。

1.3.2. 计算结果

在进行因子收益计算时，为保证组合中转债具有足够的流动性，剔除了剩余期限不足 3 个月，余额不足 3 千万，信用评级在 A 以下的转债。同时，为使回归统计量具有意义，样本量应至少为变量数的 5 倍以上，因此在有效样本数据达到 200 只及以上时（即市场中至少有 200 只转债满足前提条件且因子暴露均不为空），才将回归所得的因子收益视为有效的。

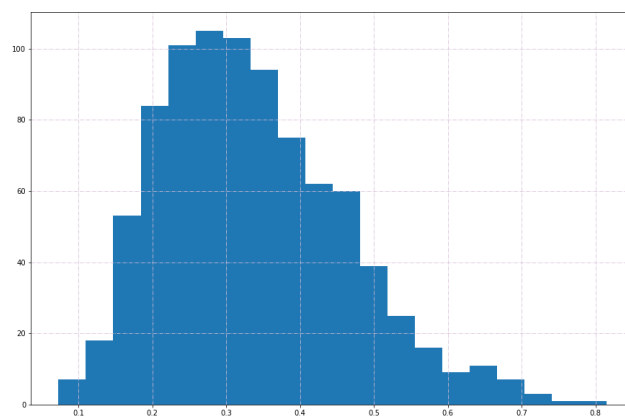
将转债日收益与因子暴露数据代入模型进行拟合，最终计算得到从 2020/05/14 开始的因子收益数据，拟合结果 R^2 均值 33.73%：

图 3：满足前提条件且因子暴露不为空的转债数量



数据来源：聚源，Wind，东北证券

图 4：WLS 回归 R^2 分布图



数据来源：聚源，Wind，东北证券

由于对风格因子进行了标准化预处理，市值加权转债组合（市场组合）对风格因子的暴露在理论上应为 0。为验证该理论，以行业市值加权合成行业因子的收益，与中证转债收益进行对比：

图 5：行业因子市值加权组合与中证转债净值对比



数据来源：聚源，Wind，东北证券

如图所示，行业因子的市值加权收益与中证转债走势基本一致，出现的差异可能来源于计算过程中对转债的筛选，证明了市场组合对于单个风格因子的暴露为 0，同时特异性收益的加权均值也为 0，论证了模型的合理性。

2. 风险矩阵估计

将模型简化表示为 $R = \beta F + u$ ，其中 β 为 $N \times K$ 因子暴露矩阵， $K=36$ 。则转债收益之间的协方差矩阵可写作：

$$\Sigma = \beta \Sigma_F \beta' + \Sigma_u$$

Σ 、 Σ_F 、 Σ_u 分别代表转债收益协方差矩阵（ N 阶）、因子收益协方差矩阵（ K 阶）、特异性收益协方差矩阵（对角矩阵， N 阶）。

多因子风险模型将转债收益协方差矩阵的估计问题，转化为对因子收益协方差矩阵和特异性收益方差的计算。在任何时刻 t ，用给定的时间窗口内的历史收益序列计算因子收益率的样本协方差矩阵和特异性收益率协方差矩阵，是对 Σ_F 、 Σ_u 的最简单估计。然而大量实证经验显示，直接使用历史数据进行简单计算难以得到准确的估计，因此需要使用不同的统计手段进行调整，本节将依次介绍。

2.1. 偏误统计量

偏误统计量是评估风险模型准确性的常用指标，用于衡量风险预测值与实际值之间的偏差。 $r_{i,t}$ 表示转债 i 在 t 期的收益， $\sigma_{i,t}$ 表示 t 期的预测波动，定义 $b_{i,t}$ ：

$$b_{i,t} = \frac{r_{i,t}}{\sigma_{i,t}}$$

如果一共有 T 期数据，定义偏误统计量 B_i ：

$$B_i = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (b_{i,t} - \bar{b}_i)^2}$$

$\bar{b}_i$ 为 T 期内 $b_{i,t}$ 的均值。

该变量的含义为，如果收益服从正态分布且预测准确，使用预测的标准差对收益进行标准化后将服从标准正态分布。当 T 足够大时， B_i 也将服从正态分布，95% 置信区间为 $[1 - \sqrt{2/T}, 1 + \sqrt{2/T}]$ ，若偏误统计量落在置信区间内，可以认为风险预测较为准确。

2.2. 因子收益协方差矩阵

由于转债自身存续时间较短，最长不过 5-6 年，同时转债市场容量较小，从 2020 年开始才具有足够进行风险分析的数据，因此本节使用日频数据进行因子收益协方差矩阵的估计，同时缩短所用的窗口期。

借鉴 Barra 模型的方式，首先采用指数加权移动平均的方式计算因子收益协方差矩阵，取窗口 $T=126$ ，半衰期 $h=45$ ，即采用过去半年的数据，且越靠近当前权重越高，对于协方差矩阵的影响越大。

指数衰减权重设置为：

$$\lambda_{t-s} = 0.5^{s/h}$$

其中 s 为该时刻距离 t 的时长，以日为单位， h 为半衰期。

使用指数加权均值对因子收益进行去均值化，计算得到的初步协方差矩阵记为 $\hat{\Sigma}_F^{raw}$ ：

$$\hat{\Sigma}_F^{raw} = F_T' F_{T,weighted} / \sum_{s=0}^T \lambda_{t-s}$$

F_T ($T \times K$) 代表从 t 开始向前倒推 T 日的去均值化因子收益时序矩阵， $F_{T,weighted}$ ($T \times K$) 代表乘上对应指数衰减权重的去均值化因子收益时序矩阵， $\sum_{s=0}^T \lambda_{t-s}$ 为整个 T 窗口期内指数权重之和。

2.2.1. Newey-West 调整

$\hat{\Sigma}_F^{raw}$ 的计算方式假设因子收益率在时序上不相关，但在实际情况中自相关性是普遍存在的，为了解决该问题，需对协方差矩阵进行 Newey-West 调整。

在进行调整前，首先需要确定因子自相关的最大滞后期，根据 Newey and West (1994) 给出的公式：

$$J = \left\lceil 4 \times \left(\frac{T}{100} \right)^{2/9} \right\rceil$$

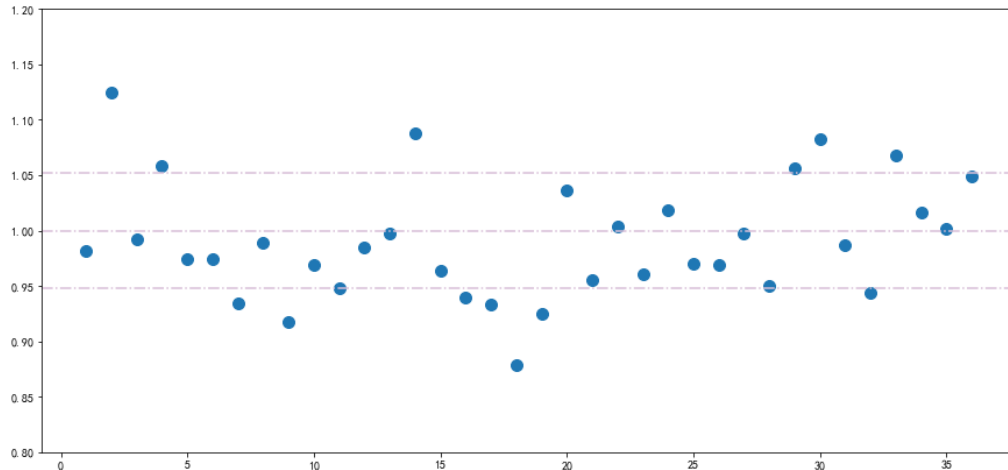
将 $T=126$ 代入，最大滞后期 J 取值为 4，即因子收益会受到过去 4 个交易日的影响。

记经过 Newey-West 调整的协方差矩阵为 $\hat{\Sigma}_F^{NW}$ ：

$$\begin{aligned} \hat{\Sigma}_F^{NW} &= \hat{\Sigma}_F^{raw} + \sum_{i=1}^J w_i \times (\Gamma_i + \Gamma_i') \\ \Gamma_i &= F_{T,i}' F_{T,i,weighted} / \sum_{s=0}^T \lambda_{t-s} \\ w_i &= 1 - \frac{i}{1+J} \end{aligned}$$

Γ_i , K 阶矩阵, 代表滞后 i 期的自协方差矩阵; $F_{T,i}$, $(T-i) \times K$, 为 F_T 所有因子收益滞后 i 期并剔除空值所得的时序矩阵; $F_{T,i,weighted}$, $(T-i) \times K$, 为 $F_{T,weighted}$ 中与 $F_{T,i}$ 时间所对应的未滞后收益时序矩阵; w_i 为 Bartlett 权重系数 (Newey and West 1987)。

图 6: 经 Newey-West 调整后因子日收益波动偏误统计量



数据来源: 聚源, Wind, 东北证券

如图, 调整后因子日收益波动率预测的偏误统计量, 大部分处于置信区间内, 由于现实资产收益并非严格服从正态分布, 因此整体的偏离处于可控范围内, 后续将会进一步调整。

值得注意的是, 由于考虑了日频收益的序列相关性, 可以将经过 Newey-West 调整后的日频协方差矩阵乘以 21 作为月度协方差矩阵, Barra 模型中便进行了这种尺度变化。但由于本文的样本区间较短, 所以在后续的调整中, 仍然以日频风险矩阵作为处理对象。

2.2.2. Eigenfactor 调整

构建风险模型的目的之一, 是为了对资产未来收益的波动进行预测, 并据此控制投资组合的风险。Menchero, Wang, and Orr (2011) 使用样本协方差矩阵作为对股票日收益波动的预测, 通过计算四种组合的偏误统计量, 来判断使用样本协方差矩阵进行预测的准确性。

四种组合分别为个股、随机组合、特征值组合和方差最小化组合。

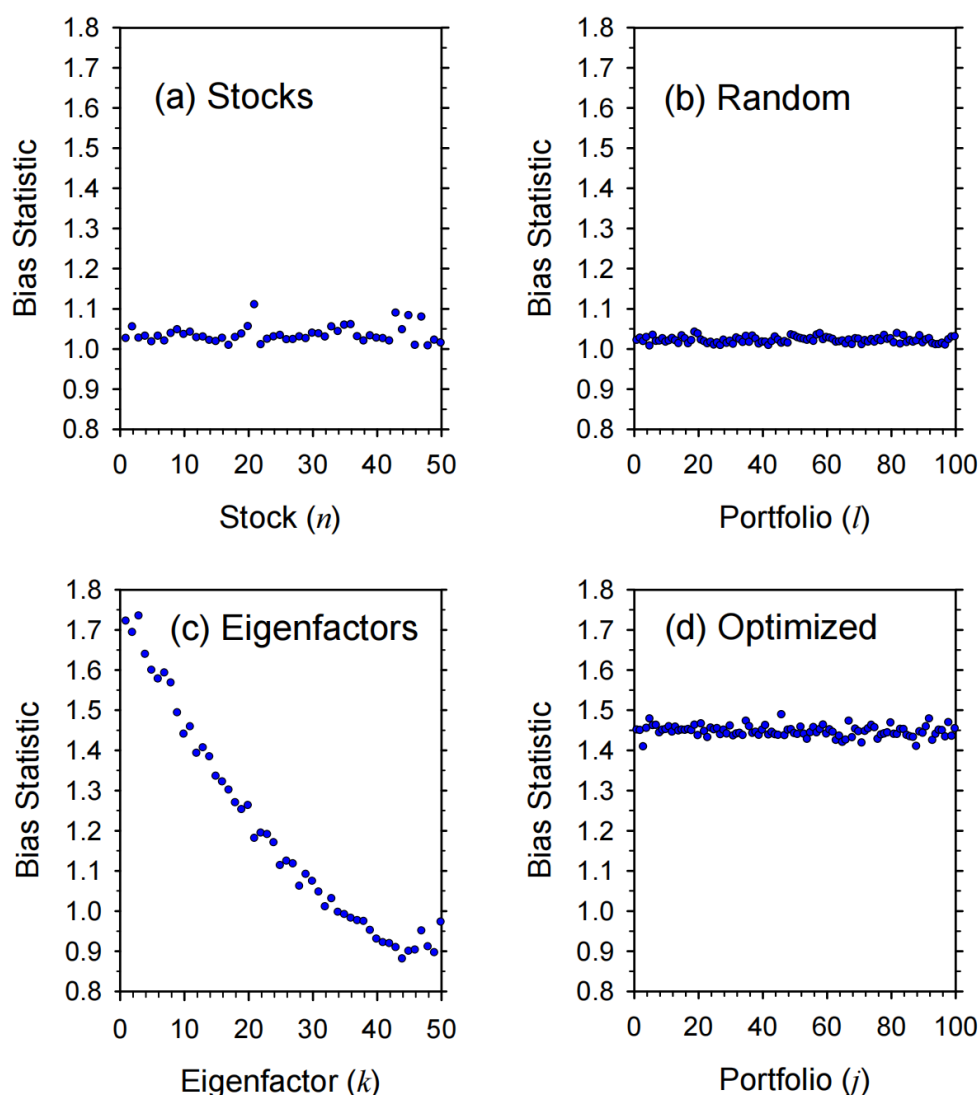
随机组合: 从正态分布中抽取随机数作为个股权重, 并使全部个股的权重之和为 0。

特征组合: 对样本协方差矩阵 $\hat{\Sigma}$ 进行特征分解, $\hat{\Sigma} = UDU'$, 其中 U 是特征向量 (eigenvector) 矩阵, 每列为一个特征向量, D 为对角阵, 对角线上元素为对应的特征值 (eigenvalue)。将上式变形为 $D = U'\hat{\Sigma}U$, 如果将因子组合视为资产, 以特征向量作为因子的权重, 所得的组合即为特征组合 (eigenfactor portfolio), 对角阵 D 上的每个元素即为对应的特征组合的方差。

方差最小化组合: 给定 α 因子暴露和样本协方差矩阵, 通过最优化求解在满足 $\alpha=1$ 条件下方差最小化的组合。

下图展示了这四种组合的偏误统计量测试, 每个点代表一个组合, 横轴从左到右, 组合方差按照从小到大排列, 纵轴为偏误统计量:

图 7：四种组合的波动预测偏误统计量



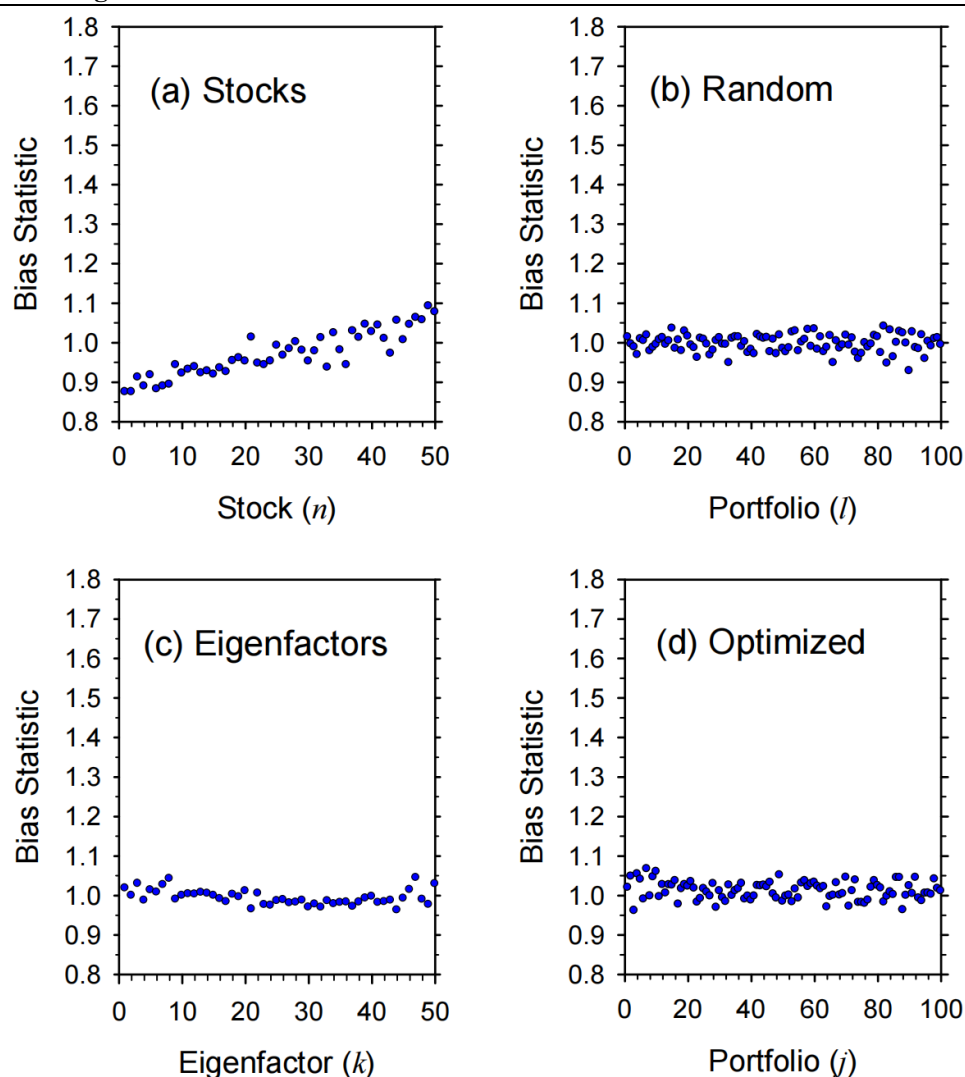
数据来源：Menchero, Wang, and Orr (2011)，东北证券

个股和随机组合的偏误统计量分布在 1 附近，说明样本协方差矩阵对这类型投资组合的风险预测较准确，但特征组合和最优化组合的预测则出现了明显的偏离。特征组合的偏误统计量分布与组合方差相关，样本方差小的组合偏误统计量大，说明其真实波动被低估，而样本方差大的组合，真实波动被高估。这也是为什么最小化方差组合同样出现方差被低估的现象，偏误统计量处于 1.4-1.5 之间。

在现实中，我们并不是随机的投资个股，而是希望通过风险模型对组合方差进行控制，因此这些特征组合是具有实际意义的。比如，方差最小的特征组合就近似于以最小化组合方差为优化目标构建的组合。所以在应用风险矩阵前，必须要进行 eigenfactor 调整，降低模型对特征组合的波动预测偏离。

下图为经过 eigenfactor 调整后，使用样本协方差矩阵对四种组合波动进行预测的偏误统计量：

图 8: Eigenfactor 调整后四种组合的波动预测偏误统计量

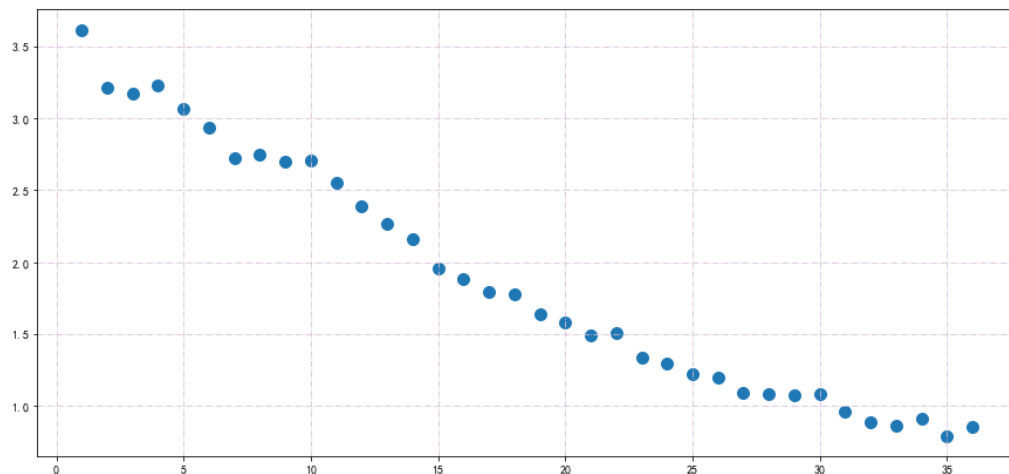


数据来源: Menchero, Wang, and Orr (2011), 东北证券

特征组合和最优化组合的偏误统计量相对于调整前有明显的改善。由于调整中改变了个股波动预测, 导致对个股的预测能力有所下降, 但鉴于现实中风险矩阵常用的场景为对组合风险进行预测, 因此 Menchero, Wang, and Orr 认为这项调整仍然是必要的。

本节借鉴以上方式, 对转债因子的风险矩阵进行 eigenfactor 调整。首先对经过 Newey-West 调整的协方差矩阵 $\hat{\Sigma}_F^{NW}$ 进行特征分解, 构建特征组合, 观察其偏误统计量是否出现相同的情况:

图 9：因子特征组合偏误统计量



数据来源：聚源，Wind，东北证券

如图，由风险因子构建的特征组合，偏误统计量仍与组合方差呈负相关，即方差小的组合波动被低估，最大的偏误统计量已经超过了 3.5，而方差大的组合波动被高估，但偏离程度相对较小。

Eigenfactor 调整的核心为 bootstrap，将 $\hat{\Sigma}_F^{NW}$ 视为“真实”的协方差矩阵，即“总体”，并以它为基础模拟计算样本协方差矩阵，则该情况下样本与总体的偏差近似于 $\hat{\Sigma}_F^{NW}$ 与真实协方差矩阵之间的偏差，从而可以对 $\hat{\Sigma}_F^{NW}$ 进行调整。

在某个截面 t ，对因子协方差矩阵进行 eigenfactor 调整的步骤如下：

- 1) 对因子协方差矩阵 $\hat{\Sigma}_F^{NW}$ 进行特征分解得到特征向量矩阵 U 和特征值对角矩阵 D 。
- 2) 令 $D(k)$ 代表 D 上的第 k 个元素，假设第 k 个特征组合的收益服从均值为 0，方差为 $D(k)$ 的正态分布。在第 m 次模拟中，基于该假设，生成所有特征组合过去 T 期的样本，作为特征组合的收益率序列，生成 $K \times T$ 收益率矩阵，记为 b_m 。
- 3) 使用 b_m 和 U 可以倒推出因子收益率矩阵 $\tilde{\lambda}_m = Ub_m$ 。
- 4) 计算得到的因子收益率矩阵计算样本协方差矩阵 $\tilde{\Sigma}_{\lambda m} = cov(\tilde{\lambda}_m, \tilde{\lambda}_m)$ 。
- 5) 再对 $\tilde{\Sigma}_{\lambda m}$ 进行特征分解得到特征向量矩阵 $\tilde{U}_m$ 和对角矩阵 $\tilde{D}_m$ 。
- 6) 由于假设已知“真实”的协方差矩阵 $\hat{\Sigma}_F^{NW}$ ，利用 $\tilde{U}_m$ 和它求出“真实”的特征组合协方差矩阵 $D_m = \tilde{U}_m' \hat{\Sigma}_F^{NW} \tilde{U}_m$ ， D_m 与 $\tilde{D}_m$ 之间的差异近似为样本与总体的差异。
- 7) 将上述步骤重复 M 次，将 M 次模拟所得的全部 D_m 与 $\tilde{D}_m$ 记录下来，比较对角线上元素的差异。 $D_m(k)$ 与 $\tilde{D}_m(k)$ 分别表示 D_m 与 $\tilde{D}_m$ 对角线上的第 k 个元素，第 k 个元素的平均偏差记为 $v(k)$ ：

$$v(k) = \frac{1}{M} \sum_{m=1}^M \sqrt{\frac{D_m(k)}{\tilde{D}_m(k)}}$$

- 8) 在现实市场中，由于收益率并非严格服从正态分布，且不稳，因此一般会在平均偏差的基础上进行额外的缩放以应对经验波动率偏误。 a 为经验系数，取值通常在 1~2 之间，缩放后的平均偏差为：

$$v_s(k) = a(v(k) - 1) + 1$$

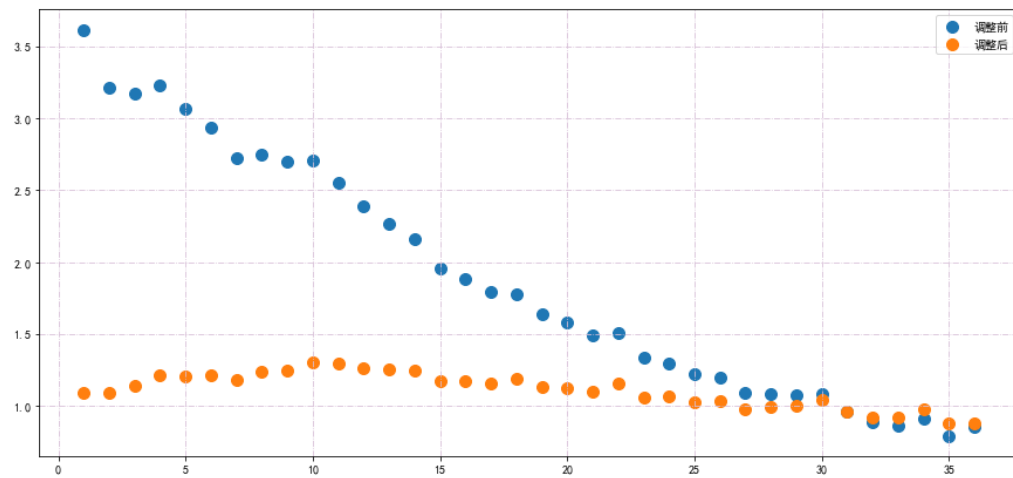
9) 利用 $v_s(k)$ 修正 $\hat{\Sigma}_F^{NW}$ 的特征值对角矩阵 D ，并利用特征向量矩阵 U 反推出修正后的因子协方差矩阵，记为 $\hat{\Sigma}_F^{Eigen}$ ：

$$D_0 = v^2 D$$

$$\hat{\Sigma}_F^{Eigen} = U D_0 U'$$

本节 eigenfactor 调整中，设置模拟 $M=1000$ 次，模拟时长 $T=63$ 个交易日，即一个季度，经验系数 $a=1.4$ ，与 Barra USE4 取值相同， $K=36$ ，为风险因子个数。下图展示调整前后因子特征组合的偏误统计量比较：

图 10：调整后因子特征组合偏误统计量



数据来源：聚源，Wind，东北证券

调整后特征组合的偏误统计量全部降至 1.5 以下，偏离程度有明显的改善，显著提升了风险估计的准确度。

2.2.3. 波动率偏误调整 (Volatility Regime Adjustment)

在前文的调整中，每个因子被视为独立的个体，Barra USE4 创新性地提出利用所有因子的信息进一步优化风险矩阵：使用截面上所有因子的风险预测来判断模型是否在某一时间段内持续高估或低估所有因子的风险，从而据此对样本协方差矩阵进行调整，降低误差，这一过程被称为波动率偏误调整 (Volatility Regime Adjustment)。

对于截面 t ，定义截面上的偏误统计量 B_t^F ：

$$B_t^F = \sqrt{\frac{1}{K} \sum_{k=1}^K \left(\frac{f_{kt}}{\sigma_{kt}} \right)^2}$$

其中 f_{kt} 为 t 期的第 k 个因子的收益率， σ_{kt} 为 t 期第 k 个因子收益的预测标准差， $K=36$ 为风险因子个数。 B_t^F 用于衡量在截面 t 上所有因子收益波动预测的偏离情况。

Barra USE4 随后以 42 天的半衰期计算 B_t^F 平方的指数加权移动均值，由于本文的样本时间区间较短，选择直接采用 B_t^F 平方的 21 天移动平均值计算调整系数 λ_F ，并对 $\hat{\Sigma}_F^{Eigen}$ 进行调整：

$$\lambda_F = \sqrt{\frac{1}{21} \sum_t B_t^{F^2}}$$

$$\hat{\Sigma}_F^{VRA} = \lambda_F^2 \hat{\Sigma}_F^{Eigen}$$

为检验调整系数的效果，定义截面波动率 $CSV_t^F = \sqrt{\frac{1}{K} \sum_k f_{kt}^2}$ ，比较 CSV21 个交易日移动平均与调整系数 λ_F 的走势：

图 11：因子收益截面波动率 CSV 与调整系数 λ_F 走势



数据来源：聚源，Wind，东北证券

在截面波动率上升的阶段，调整系数随之上升以修正可能的风险低估，同理在截面波动率下降时，调整系数同步下降以修正可能的风险高估。

2.2.4. 不同调整方式的优化效果

在经过了 Newy-West 调整、Eigenfactor 调整和波动率偏误调整后，得到最终的因子收益协方差矩阵。比较每一步所得风险矩阵的偏误统计量，可以直观地看出每种调整方式在预测因子收益波动上的优化效果。

分别计算 $\hat{\Sigma}_F^{NW}$ 、 $\hat{\Sigma}_F^{Eigen}$ 、 $\hat{\Sigma}_F^{VRA}$ 因子日收益波动偏误统计量的 63 个交易日移动平均：

图 12：每步调整后的因子日收益波动偏误统计量（63 个交易日移动平均）



数据来源：聚源，Wind，东北证券

如图，经过Newy-West调整后偏误统计量基本处于置信区间内，偏离处于可控范围。Eigenfactor 调整主要是对特征组合的波动预测进行调整，因此对因子自身波动的预测没有优化作用，相对前一步偏离没有明显改善。波动率偏误调整的优化效果最为明显，在考虑到风险模型对所有因子收益波动的系统性高估或低估后，偏误统计量一直处于置信区间内且非常接近 1。

2.3. 特异性收益协方差矩阵

特异性收益协方差矩阵 Σ_u 是风险模型的另一重要部分。假设所有转债的特异性收益率之间互不相关，则 Σ_u 为对角矩阵。与计算因子收益协方差矩阵相同，采用指数加权移动平均的方式计算特异性收益协方差矩阵，取窗口 $T=126$ ，半衰期 $h=45$ 。

指数衰减权重：

$$\lambda_{t-s} = 0.5^{s/h}$$

其中 s 为该时刻距离 t 的时长，以日为单位， h 为半衰期。

第 n 只转债在 t 时刻的特异性收益方差为：

$$\widehat{var}(n, t) = \sum_{s=0}^T \lambda_{t-s} (u_{n,t-s} - \bar{u}_n)^2 / \sum_{s=0}^T \lambda_{t-s}$$

其中 $\bar{u}_n$ 为转债 n 特异性收益在过去 T 期的指数加权均值， $u_{n,t-s}$ 为 $t-s$ 期转债 n 的特异性收益率。

由于特异性收益同样具有自相关性等问题，所以也需要进行一系列调整以优化进行波动预测的效果。

2.3.1. Newy-West 调整

考虑到特异性波动率时序的自相关性，需对其进行 Newy-West 调整。最大滞后期 J 取值仍为 4，经 Newy-West 调整后的转债 n 特异性收益方差记为 $\widehat{var}_{(n,t)}^{NW}$ ：

$$\widehat{var}_{(n,t)}^{NW} = \widehat{var}(n, t) + \sum_{i=1}^J w_i \times 2 \times (adJ_var_{n,i})$$

$$adJ_var_{n,i} = \sum_{s=0}^{T-i} \lambda_{t-s} (u_{n,t-s} - \bar{u}_n) (u_{n,t-s-i} - \bar{u}_n) / \sum_{s=0}^T \lambda_{t-s}$$

$$w_i = 1 - \frac{i}{1+J}$$

$adJ_var_{n,i}$ 代表转债 n 特异性收益滞后 i 期的自协方差, $\bar{u}_n$ 为转债 n 特异性收益在过去 T 期的指数加权均值, w_i 为 Bartlett 权重系数 (Newey and West 1987)。

同样的, 由于考虑了时序的自相关性, 可以直接将经 Newy-West 调整后的方差乘上 21 作为月度方差, 但由于本文样本的时间区间较短, 因此后续处理仍以日频的特异性收益协方差矩阵作为对象。

2.3.2. 结构化模型调整 (Structural Model)

因子收益是由市场中上百只转债的数据回归得到的, 具有连续性, 没有空值, 且不易出现异常值。但特异性收益率是针对某一只转债计算得到的, 对于新上市的或发生重大事件的转债, 其特异性收益率可能出现空值或异常波动。为保证协方差矩阵的平稳性, 使用结构化模型对特异性收益率矩阵进行调整。

结构化模型认为特异性收益波动也同样可以由风险因子解释, 因子暴露相同的转债具有相同的特异性收益波动。首先对于一段时间窗口 ($T=126$) 内的特异性收益率序列, 定义协调系数 γ_n :

$$\gamma_n = \{\min [1, \max(0, (h_n - 30)/60)]\} \{\min [1, \exp(1 - Z_n)]\}$$

$$Z_n = \left| \frac{\sigma_n - \widetilde{\sigma}_n}{\widetilde{\sigma}_n} \right|$$

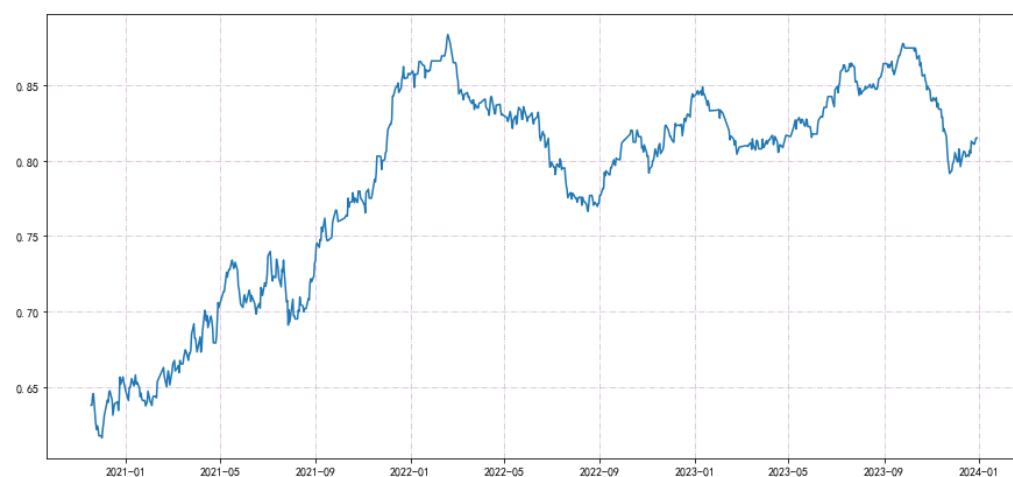
$$\widetilde{\sigma}_n = \frac{Q_{3,n} - Q_{1,n}}{1.35}$$

h_n 代表窗口内转债 n 有效数据的个数, 以 $(h_n - 30)/60$ 表示数据缺失的程度, 若有效数据过少, 则 γ_n 为 0。 Z_n 用于测定数据的肥尾程度, Z_n 超过 1 越多, 说明肥尾现象越明显, $Q_{3,n}$ 、 $Q_{1,n}$ 分别表示 3/4 和 1/4 位点的特异性收益率, σ_n 为特异性收益等权标准差。

当 $\gamma_n=1$ 时, 认为转债在这 126 个交易日窗口内无明显数据缺失和异常值, 视为数据质量优秀, 将用于计算共同因子对特异性收益率的贡献。

如图, $\gamma_n=1$ 的转债占比持续处于 60%以上, 保证了进行结构化调整的合理性。

图 13: 协调系数 $\gamma_n=1$ 的转债占比



数据来源：聚源，Wind，东北证券

在截面 t 上，使用 $\gamma_n=1$ 转债的特异性收益波动率 σ_n^{TS} （经 Newy-West 调整）自然对数对所有因子的暴露作线性回归，采用市值加权的 WLS：

$$\log(\sigma_n^{TS}) = \sum_k X_{nk} b_k + \varepsilon_n$$

X_{nk} 为转债 n 在截面上对风险因子 k 的暴露， b_k 为风险因子 k 对特异性收益率波动的贡献程度。

将因子暴露代入，则通过风险因子预测的转债 n 特异性收益波动 σ_n^{STR} 为：

$$\sigma_n^{STR} = E_0 \times \exp\left(\sum_k X_{nk} b_k\right)$$

E_0 是为了消除残差项 ε_n 所加的调整系数，本节取 1.05。

经过结构化调整的转债 n 的特异性收益波动 $\hat{\sigma}_n$ ：

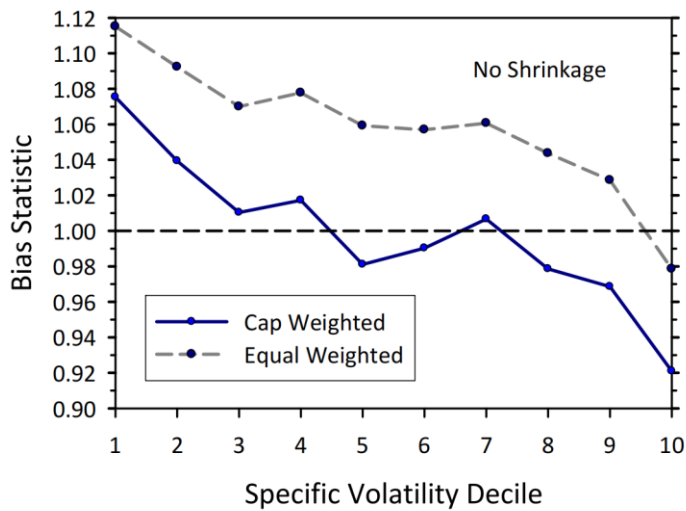
$$\hat{\sigma}_n = \gamma_n \sigma_n^{TS} + (1 - \gamma_n) \sigma_n^{STR}$$

若 $\gamma_n=1$ ，则调整后特异性波动保持不变，若 $\gamma_n<1$ ，调整后波动加入 σ_n^{STR} 项，且 γ_n 越小， σ_n^{STR} 占比越大。

2.3.3. 贝叶斯收缩 (Bayesian Shrinkage)

Barra USE4 指出，使用样本内数据计算出的特异性波动率在样本外的持续性很差，特别是波动极低或极高的股票，其样本外的波动会向均值回归。为观察该现象，每月底将股票按照特异性收益波动预测分为 10 组，计算其偏误统计量，如图：

图 14：特异性波动预测分组偏误统计量表现



数据来源：Barra USE4，东北证券

对于波动率小的分组，偏误统计量显著大于 1，说明预测方差低估了股票样本外的特异性波动率；而对于波动率大的分组，偏误统计量显著小于 1，说明预测方差高估了股票样本外的特异性波动率。

鉴于该特性，Barra USE4 使用贝叶斯收缩将先验和样本估计值结合起来，根据特异性风险回归均值的趋势进一步修正风险预测值。本节亦借鉴此方法对转债特异性风险矩阵进行调整。

在每个截面 t ，将所有转债按照市值从小到大分为 10 组，分组计算每支转债的收缩强度系数 v_n ：

$$v_n = \frac{q|\hat{\sigma}_n - \bar{\sigma}(s_n)|}{\Delta_\sigma(s_n) + q|\hat{\sigma}_n - \bar{\sigma}(s_n)|}$$

$$\Delta_\sigma(s_n) = \sqrt{\frac{1}{N(s_n)} \sum_n (\hat{\sigma}_n - \bar{\sigma}(s_n))^2}$$

$$\bar{\sigma}(s_n) = \sum_n mvw_n \hat{\sigma}_n$$

$\hat{\sigma}_n$ 为经结构化模型调整后转债 n 的预测特异性波动； $\bar{\sigma}(s_n)$ 为转债 n 所在的市值分组的特异性波动市值加权均值， mvw_n 为市值权重； $\Delta_\sigma(s_n)$ 为该市值分组特异性波动的标准差， $N(s_n)$ 代表该组转债的数量。 q 为经验系数，通过综合考虑所有特异性波动率收缩前后偏误统计量的优化效果来找到合适的值，本节中取值为 1。

$\bar{\sigma}(s_n)$ 被视为先验， $\hat{\sigma}_n$ 被视为样本， $|\hat{\sigma}_n - \bar{\sigma}(s_n)|$ 即为个券预测特异性波动与先验的偏离程度。 $|\hat{\sigma}_n - \bar{\sigma}(s_n)|$ 越大，样本的偏离程度越大，调整过程中先验的权重越大，即 v_n 越大；反之， $|\hat{\sigma}_n - \bar{\sigma}(s_n)|$ 越小，样本的偏离程度越小，可信度较高，样本自身在调整中的占比越大。

贝叶斯收缩调整后的特异性波动 $\hat{\sigma}_n^{SH}$ ：

$$\hat{\sigma}_n^{SH} = v_n \bar{\sigma}(s_n) + (1 - v_n) \hat{\sigma}_n$$

2.3.4. 波动率偏误调整 (Volatility Regime Adjustment)

与因子收益协方差矩阵相同，对于特异性收益波动方差也需要进行波动率偏误调整，利用截面上所有转债特异性波动预测的信息，判断模型预测是否在某一阶段系统性地高估或低估所有转债的特异性波动，并进行相应的调整。

对于截面 t ，定义截面上的偏误统计量 B_t^S ：

$$B_t^S = \sqrt{\sum_{n=1}^N mvw_{nt} \left(\frac{u_{nt}}{\sigma_{nt}} \right)^2}$$

u_{nt} 为 t 期转债 n 的特异性收益， σ_{nt} 为 t 期转债 n 的预测特异性收益波动， mvw_{nt} 为转债 n 的市值权重。

同样，取过去 21 个交易日的截面偏误统计量移动均值作为系数，对特异性波动进行调整：

$$\lambda_S = \sqrt{\frac{1}{21} \sum_t B_t^{S^2}}$$

$$\hat{\sigma}_n^{VRA} = \lambda_S \hat{\sigma}_n^{SH}$$

为检验调整系数的效果，定义截面波动率 $CSV_t^S = \sqrt{\sum_n mvw_{nt} u_{nt}^2}$ ，比较 CSV21 个交易日移动平均与调整系数 λ_S 的走势：

图 15：特异性收益截面波动率 CSV 与调整系数 λ_S 走势



数据来源：聚源，Wind，东北证券

与因子协方差矩阵波动率偏误调整中相同，在截面波动率上升的阶段，调整系数 λ_S 随之上升以修正可能的特异性风险低估，同理在截面波动率下降时，调整系数同步下降以修正可能的风险高估。

2.3.5. 不同调整方式的优化效果

经过了 Newy-West 调整、结构化模型调整、贝叶斯收缩和波动率偏误调整后，得到最终的特异性收益协方差矩阵。比较每一步所得风险矩阵的偏误统计量，可以直观地看出每种调整方式在预测因子收益波动上的优化效果。

分别计算每一步调整后特异性日收益波动预测的偏误统计量，并取 63 个交易日移动平均进行比较：

图 16：每步调整后的特异性收益波动偏误统计量（63 个交易日移动平均）

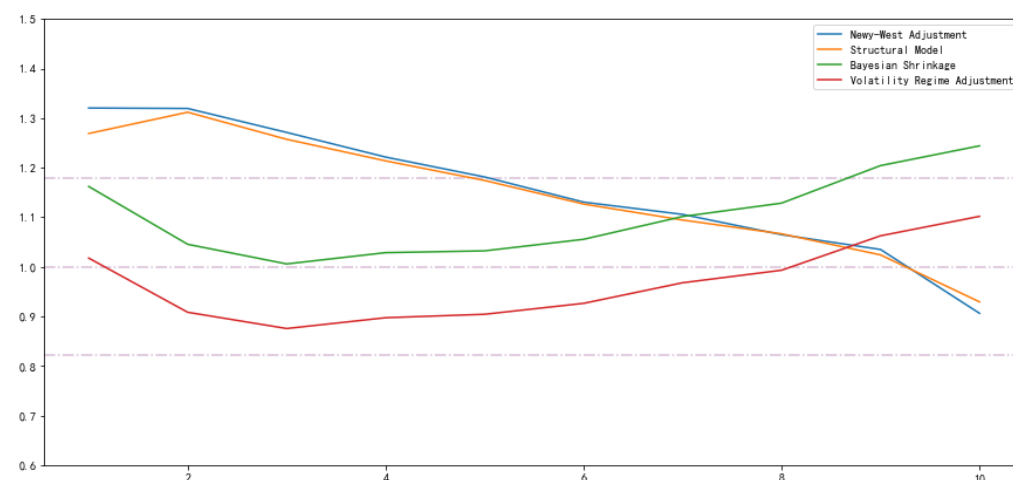


数据来源：聚源，Wind，东北证券

Newy-West 初步调整后的偏误统计量毛刺较多，即特异性收益出现异常波动导致协方差矩阵中出现异常值，在经由结构化模型调整后有明显的改善，但对预测准确性没有优化效果。贝叶斯收缩和波动率偏误调整后，偏误统计量完全落入置信区间内。

为进一步观察贝叶斯收缩和波动率偏误的优化效果，借鉴 Barra 的方法，将所有转债按照预测的特异性波动分为 10 组，计算每组未来 63 个交易日的偏误统计量均值：

图 17：特异性波动预测分组未来 63 个交易日偏误统计量表现



数据来源：聚源，Wind，东北证券

在贝叶斯调整前，和 Barra 的测试结果相同，对于方差小的组合，预测波动明显低估了样本外方差。贝叶斯收缩调整后，偏误统计量基本落入置信区间，与 1 更接近。

3. 风险模型的应用

3.1. 风险矩阵降维

在得到每个时间截面上的因子收益协方差矩阵和特异性收益协方差矩阵后，即可得到截面上所有转债收益率的风险矩阵：

$$\Sigma = \beta \Sigma_F \beta' + \Sigma_u$$

Σ 、 Σ_F 、 Σ_u 分别代表转债收益协方差矩阵（N 阶）、因子收益协方差矩阵（K 阶）、特异性收益协方差矩阵（对角矩阵，N 阶）。

如果直接使用转债全市场的收益率序列计算协方差矩阵，资产数量众多且所需的区间较长，对数据和算力的要求较高。例如，计算 500 只转债的日收益协方差矩阵至少需要 500 个交易日以上的数据，如果计算月频风险矩阵则需要 500 个月以上，对于本身就存续期较短的转债，这显然是不现实的。通过风险模型，将问题转化为对因子收益协方差矩阵和特异性收益协方差矩阵的计算，实现矩阵的降维，大大缩减了所需的数据区间长度。

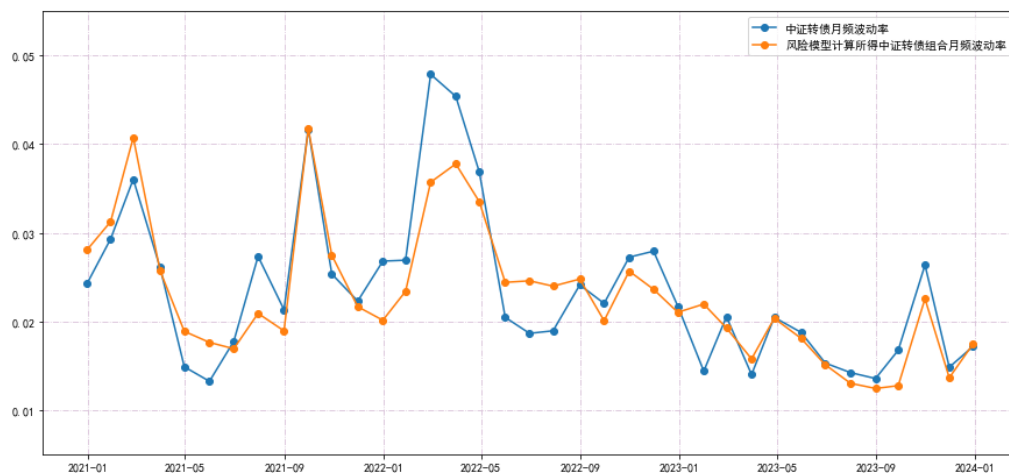
对于给定的组合，可以根据权重计算投资组合的收益波动：

$$\sigma_P^2 = w^T \Sigma w = w^T (\beta \Sigma_F \beta' + \Sigma_u) w$$

$w(N \times 1)$ 为组合权重矩阵， σ_P^2 为投资组合的收益方差。

以中证转债指数为例，根据每月底的组合权重，使用风险模型计算得到当月的收益波动，与实际波动进行对比：

图 18：中证转债实际波动与风险模型计算结果比较



数据来源：聚源，Wind，东北证券

如图，计算所得的组合波动与实际波动走势接近，相关系数为 0.88，计算较准确。在验证了风险模型可以用于计算组合波动后，即可以对组合收益进行风险归因。

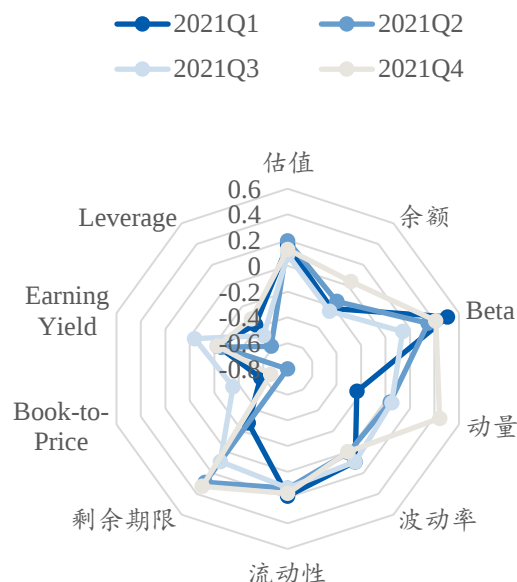
3.2. 组合收益归因

对于可转债基金，可以根据其重仓转债的风格因子暴露与权重占比，对重仓债组合的收益进行风险分解，从而判断每个因子对收益的边际贡献。

以规模在 50 亿左右的可转债基金 A 为例，选取 2021-2023 年 12 个报告期，每期披露重仓债券持仓市值占基金净值比例均在 60%以上，债券中 90%以上为可转债。使用每报告期所披露的重仓可转债持仓数据，按照持仓市值加权计算重仓转债组合的风险因子暴露与风险因子收益。

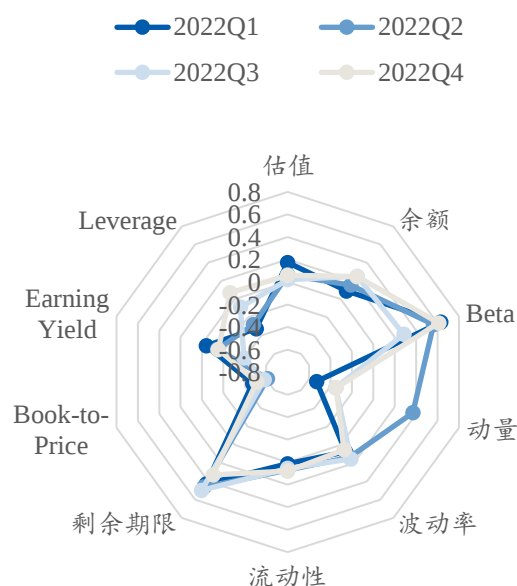
以下为各报告期风格因子的暴露与收益：

图 19：2021 年转债基金 A 风格因子暴露



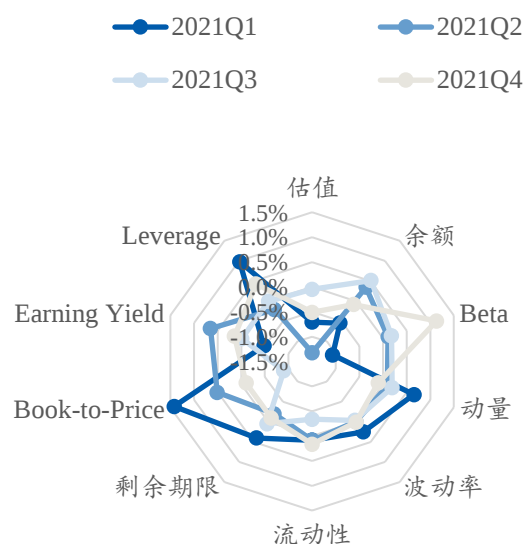
数据来源：聚源，iFind，东北证券

图 21：2022 年转债基金 A 风格因子暴露



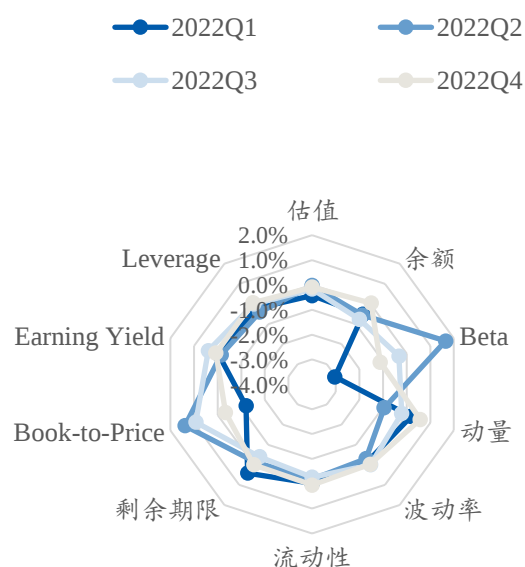
数据来源：聚源，iFind，东北证券

图 20：2021 年转债基金 A 风格因子收益



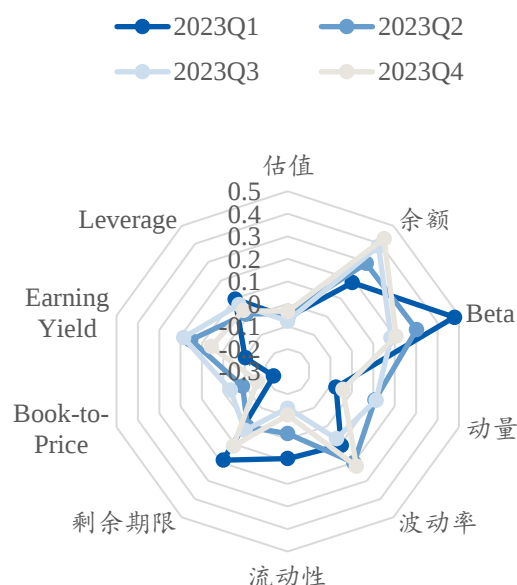
数据来源：聚源，iFind，东北证券

图 22：2022 年转债基金 A 风格因子收益



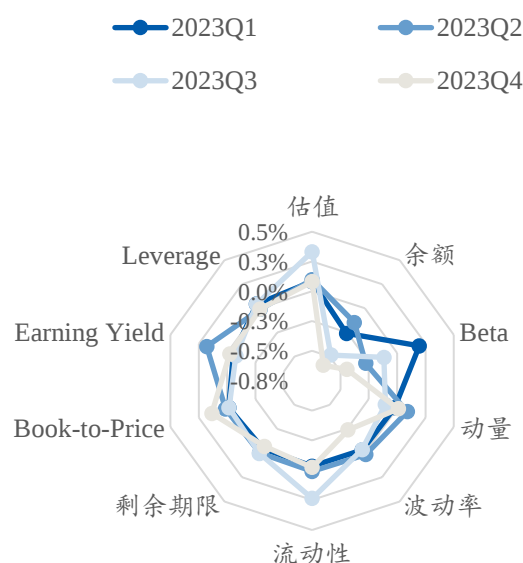
数据来源：聚源，iFind，东北证券

图 23：2023 年转债基金 A 风格因子暴露



数据来源：聚源，iFind，东北证券

图 24：2023 年转债基金 A 风格因子收益



数据来源：聚源，iFind，东北证券

过去三年，该基金重仓转债的风格不断向低估值、高余额、低换手、正股低 P/B 靠近，在 Beta 因子上的暴露一直比较显著，其他因子暴露无明显变化。

从风格因子收益贡献角度：在低估值效应明显的 2021-2022 年，由于基金重仓转债风格偏向高估值，因此估值风格因子的收益为负，在转换风格至低估值后，因子收益转正；Beta 因子收益则代表转债组合收益中受市场影响的部分，以 2023 年为例，在市场组合具有正收益的 Q1，Beta 因子暴露也为基金带来正收益，但在 Q4 市场整体下行的情况下，Beta 因子收益也转为负；Book-to-Price 因子在 2021、2022 年带来有效收益，但 2023 年失效。

统计行业因子暴露，可以直观地展示出基金重仓转债的行业分布情况。出于转债市场行业分布不均的特性，重仓转债组合在银行行业上的暴露一直位居第一；电子、电新、通信和非银行业暴露逐年减少，基础化工、有色、农林牧渔的行业暴露则有所提升。各报告期的行业因子暴露如下表：

表 4：各报告期转债基金 A 行业因子暴露

	2021Q1	2021Q2	2021Q3	2021Q4	2022Q1	2022Q2	2022Q3	2022Q4	2023Q1	2023Q2	2023Q3	2023Q4
交通运输	0.0527	0.0770	0.0619	0.0602	0.0791	0.0617	0.0708	0.1011	0.0544	0.0634	0.0539	0.0483
传媒	0.0075	0.0045	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0023	0.0000
农林牧渔	0.0313	0.0058	0.0000	0.0731	0.0919	0.0920	0.1165	0.0979	0.0886	0.0546	0.0623	0.1148
医药	0.0442	0.0665	0.0687	0.0506	0.0374	0.0000	0.0057	0.0121	0.0000	0.0223	0.0410	0.0400
商贸零售	0.0075	0.0043	0.0042	0.0032	0.0083	0.0077	0.0149	0.0055	0.0000	0.0000	0.0000	0.0000
国防军工	0.0000	0.0000	0.0000	0.0081	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0007	0.0000
基础化工	0.0447	0.0273	0.0236	0.0094	0.0050	0.0325	0.0194	0.0118	0.0730	0.0717	0.0951	0.0852
家电	0.0000	0.0107	0.0124	0.0169	0.0403	0.0103	0.0261	0.0347	0.0105	0.0124	0.0000	0.0000
建材	0.0114	0.0285	0.0244	0.0193	0.0524	0.0194	0.0072	0.0039	0.0124	0.0047	0.0013	0.0000
建筑	0.0346	0.0362	0.0457	0.0163	0.0062	0.0169	0.0252	0.0268	0.0122	0.0000	0.0018	0.0018
有色金属	0.0276	0.0106	0.0886	0.1026	0.0919	0.1308	0.0923	0.1191	0.0851	0.0546	0.0469	0.0624
机械	0.0121	0.0012	0.0152	0.0174	0.0101	0.0194	0.0118	0.0063	0.0137	0.0001	0.0000	0.0076
汽车	0.0145	0.0400	0.0230	0.0487	0.0351	0.0177	0.0231	0.0266	0.0505	0.0679	0.0407	0.0565
煤炭	0.0000	0.0132	0.0264	0.0064	0.0085	0.0171	0.0176	0.0247	0.0352	0.0132	0.0238	0.0412
电力及公用事业	0.0107	0.0604	0.0640	0.0297	0.0305	0.0097	0.0301	0.0166	0.0449	0.0638	0.0682	0.0722
电力设备及新能源	0.0618	0.0966	0.0157	0.0244	0.0000	0.1054	0.0824	0.0820	0.0218	0.0155	0.0353	0.0501
电子	0.1457	0.0983	0.1024	0.0492	0.0389	0.0306	0.0076	0.0221	0.0185	0.0385	0.0338	0.0503
石油石化	0.0310	0.0143	0.0000	0.0089	0.0444	0.0213	0.0163	0.0200	0.0000	0.0030	0.0081	0.0124
纺织服装	0.0437	0.0308	0.0450	0.0288	0.0275	0.0161	0.0178	0.0202	0.0326	0.0349	0.0250	0.0122
计算机	0.0000	0.0941	0.0279	0.1545	0.0059	0.0236	0.0100	0.0162	0.0078	0.0082	0.0102	0.0043
轻工制造	0.0267	0.0082	0.0101	0.0031	0.0044	0.0045	0.0082	0.0070	0.0090	0.0198	0.0136	0.0205
通信	0.0970	0.0527	0.0829	0.0746	0.0751	0.0689	0.0496	0.0045	0.0246	0.0210	0.0058	0.0000
钢铁	0.0000	0.0000	0.0065	0.0076	0.0156	0.0278	0.0276	0.0363	0.0478	0.0366	0.0410	0.0418
银行	0.1376	0.1436	0.1703	0.1445	0.2396	0.2453	0.2464	0.2335	0.2792	0.3232	0.2821	0.1906
非银行金融	0.1568	0.0604	0.0667	0.0265	0.0338	0.0069	0.0667	0.0666	0.0782	0.0707	0.1071	0.0876
食品饮料	0.0010	0.0151	0.0144	0.0160	0.0180	0.0143	0.0065	0.0044	0.0000	0.0000	0.0000	0.0000

数据来源：聚源，iFind，东北证券

3.3. 组合优化

在搭建了风险模型后，可以将其与 alpha 因子结合，设置不同的目标函数，进行最优化求解得到组合，从而同时考虑到收益的实现与风险控制。常见的目标函数有均值-方差、最小化方差、风险平价等。

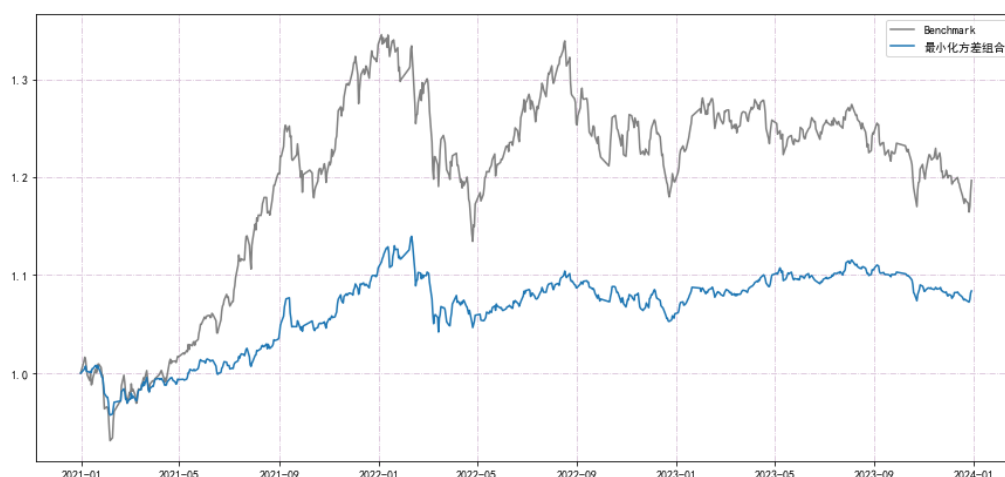
首先设置目标函数为最小化方差，加入相应约束条件：

$$\begin{aligned}
 & \min w^T \Sigma w \\
 & \sum_n w = 1 \\
 & 0 \leq w_n \leq 0.05 \\
 & |w^T F_{style}| \leq 1
 \end{aligned}$$

每月底代入预测的月度风险矩阵作为 Σ ，限制组合各风格因子暴露绝对值不大于 1，即风格偏离在 1 倍标准差内，不得做空且个券权重不超过 5%，个券权重之和为 1。

回溯区间设置为 2021/01-2023/12，以满足余额大于 3 千万、剩余期限大于 3 个月且信用评级在 A 及以上的转债为资产池，以最小化方差为目标函数，进行最优化求解，得到组合的净值表现如下：

图 25：最小化方差组合净值表现



数据来源：聚源，Wind，东北证券

基准为所有转债的等权组合，显然最小化方差组合在波动率上有明显的优化，但收益略显不足。

为在控制风险的同时保留组合收益，将 alpha 因子引入目标函数：

$$\begin{aligned} \max_w & w^T F_{\alpha} - \lambda w^T \Sigma w \\ & \sum_n w = 1 \\ & 0 \leq w_n \leq 0.05 \\ & |w^T F_{\text{style}}| \leq 1 \end{aligned}$$

本节所采用的 alpha 因子为“低估值+正股动量”策略因子（详见《可转债定价模型探索与多因子组合策略》）， λ 为风险厌恶系数，其余约束条件保持不变。

同样设置回测区间为 2021/01-2023/12，以满足余额大于 3 千万、剩余期限大于 3 个月且信用评级在 A 及以上的转债为资产池，每月底代入预测的月度风险矩阵作为 Σ ，设置不同的风险厌恶系数，对上述目标函数进行最优化求解，与原策略组合对比：

图 26：不同风险厌恶程度下最优组合净值表现



数据来源：聚源，Wind，东北证券

表 5：不同风险厌恶程度下最优组合表现统计

	年化收益	年化波动	夏普比率	最大回撤
低估值+正股动量策略	16.35%	11.30%	1.45	11.29%
lambda=100	21.27%	13.30%	1.60	11.92%
lambda=500	18.99%	10.80%	1.76	10.92%
lambda=1000	17.25%	9.06%	1.91	9.45%

数据来源：聚源，Wind，东北证券

使用组合优化器所得最优化组合，在年化收益和风险回报比表现上均优于原策略组合。伴随着风险厌恶系数的提高，组合波动降低，收益也受到了一定程度的压缩，但夏普比率呈现上升趋势，最高为 1.91，最大回撤也不断降低。原“低估值+正股动量”策略为直接选取因子值前 20% 的转债构成等权组合，本节使用目标函数进行优化，所选取的为因子值较大、自身波动较小且相关性较低的转债组合，因此达到提高收益同时降低波动的效果。

如果想得到更稳定的超额收益，可以挖掘更多 alpha 因子加入目标函数，或者直接将目标函数中的因子暴露替换为超额组合的因子暴露，将波动替换为超额收益的预测波动率，并以此进行求解。

4. 总结

本文借鉴 Barra 模型的理论，使用 26 个行业因子和 10 个风格因子（估值、余额、beta、动量、波动、流动性、剩余期限、Book-to-Price、Earning Yield、Leverage）搭建可转债风险模型，分别计算因子收益率协方差矩阵与特异性收益方差矩阵并进行相应的调整，包括 Newy-West、Eigenfactor、波动率偏误、贝叶斯收缩等，调整后滚动 63 个交易日偏误统计量均处于置信区间内。

使用风险模型对中证转债指数组合的波动率进行计算，可以实现协方差矩阵降维，缩短计算所需的时间序列长度，且精确度较高，计算所得组合方差与实际方差相关系数为 0.88。同时，对于可转债基金，可以根据其重仓转债的风格因子暴露与权重占比，对收益进行风险分解，从而判断每个因子对收益的边际贡献。

结合 alpha 因子使用风险模型，设置目标函数为 $\max_w w^T F_{\alpha} - \lambda w^T \Sigma w$ ，进行最优化求解，所得最优组合表现相对仅使用 alpha 因子所构建的策略更佳，不同风险厌恶程度下夏普比率均有所提高，最高为 1.91。未来可以加入更多 alpha 因子进行组合优化，继续提升超额收益。

5. 参考文献

Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61(4), 631-653.

Menchero, J.G., Orr, D.J., & Wang, J. (2011). The Barra US Equity Model (USE4) Methodology Notes.

Menchero, J., Wang, J., & Orr, D. J. (2011). Eigen-adjusted covariance matrices. MSCI Barra Research Paper, (2011-14).

Orr, D. J., Igor, M., & Adam, N. (2012). The Barra China Equity Model (CNE5) Empirical Notes.

6. 风险提示

以上结果基于数学模型与历史数据，存在模型失效的可能性。

研究团队简介:

王琦: 帝国理工学院数学与金融荣誉硕士, 南开大学统计学学士。2021 年加入东北证券上海证券研究咨询分公司任金融工程首席分析师, 研究方向为金融工程。曾任职于兴业财富资产管理有限公司, 任 FOF 投资经理。

贾英: 伦敦大学学院金融数学硕士, 厦门大学数学与应用数学本科。2022 年加入东北证券, 研究方向为因子选股, 现任东北证券上海证券研究咨询分公司金融工程组研究助理。

张栋梁: 复旦大学金融硕士, 南京大学金融学本科。2022 年加入东北证券, 研究方向为因子选股, 现任东北证券上海证券研究咨询分公司金融工程研究助理。

王国鑫: 伦敦国王学院金融数学荣誉硕士, 武汉大学金融学本科。2023 年加入东北证券, 研究方向为量化择时、行业轮动, 现任东北证券上海证券研究咨询分公司金融工程组研究人员。

江雨航: 加州大学洛杉矶分校金融工程硕士, 南开大学理学/经济学学士。2023 年加入东北证券, 研究方向为量化固收策略, 现任东北证券上海证券研究咨询分公司金融工程组研究人员。

田靖航: 北京大学金融硕士, 上海财经大学经济学学士。2023 年加入东北证券, 研究方向为基金研究, 现任东北证券上海证券研究咨询分公司金融工程组研究人员。

刘昱亨: 北京大学计算机硕士, 北京航空航天大学工学学士。2023 年加入东北证券, 研究方向为机器学习与衍生品量化研究, 现任东北证券上海证券研究咨询分公司金融工程组研究人员。

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, 并在中国证券业协会注册登记为证券分析师。本报告遵循合规、客观、专业、审慎的制作原则, 所采用数据、资料的来源合法合规, 文字阐述反映了作者的真实观点, 报告结论未受任何第三方的授意或影响, 特此声明。

投资评级说明

股票 投资 评级 说明	买入	未来 6 个月内, 股价涨幅超越市场基准 15% 以上。	投资评级中所涉及的市场基准: A 股市场以沪深 300 指数为市场基准, 新三板市场以三板成指 (针对协议转让标的) 或三板做市指数 (针对做市转让标的) 为市场基准; 香港市场以摩根士丹利中国指数为市场基准; 美国市场以纳斯达克综合指数或标普 500 指数为市场基准。
	增持	未来 6 个月内, 股价涨幅超越市场基准 5% 至 15% 之间。	
	中性	未来 6 个月内, 股价涨幅介于市场基准-5% 至 5% 之间。	
	减持	未来 6 个月内, 股价涨幅落后市场基准 5% 至 15% 之间。	
	卖出	未来 6 个月内, 股价涨幅落后市场基准 15% 以上。	
行业 投资 评级 说明	优于大势	未来 6 个月内, 行业指数的收益超越市场基准。	
	同步大势	未来 6 个月内, 行业指数的收益与市场基准持平。	
	落后大势	未来 6 个月内, 行业指数的收益落后于市场基准。	

重要声明

本报告由东北证券股份有限公司（以下称“本公司”）制作并仅向本公司客户发布，本公司不会因任何机构或个人接收到本报告而视其为本公司的当然客户。

本公司具有中国证监会核准的证券投资咨询业务资格。

本报告中的信息均来源于公开资料，本公司对这些信息的准确性和完整性不作任何保证。报告中的内容和意见仅反映本公司于发布本报告当日的判断，不保证所包含的内容和意见不发生变化。

本报告仅供参考，并不构成对所述证券买卖的出价或征价。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证券买卖建议。本公司及其雇员不承诺投资者一定获利，不与投资者分享投资收益，在任何情况下，我公司及其雇员对任何人使用本报告及其内容所引发的任何直接或间接损失概不负责。

本公司或其关联机构可能会持有本报告中涉及到的公司所发行的证券头寸并进行交易，并在法律许可的情况下不进行披露；可能为这些公司提供或争取提供投资银行业务、财务顾问等相关服务。

本报告版权归本公司所有。未经本公司书面许可，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，须在本公司允许的范围内使用，并注明本报告的发布人和发布日期，提示使用本报告的风险。

若本公司客户（以下称“该客户”）向第三方发送本报告，则由该客户独自为此发送行为负责。提醒通过此途径获得本报告的投资者注意，本公司不对通过此种途径获得本报告所引起的任何损失承担任何责任。

东北证券股份有限公司

网址：<http://www.nesc.cn> 电话：95360,400-600-0686 研究所公众号：dbzqyanjiusuo

地址	邮编
中国吉林省长春市生态大街 6666 号	130119
中国北京市西城区锦什坊街 28 号恒奥中心 D 座	100033
中国上海市浦东新区杨高南路 799 号	200127
中国深圳市福田区福中三路 1006 号诺德中心 34D	518038
中国广东省广州市天河区冼村街道黄埔大道西 122 号之二星辉中心 15 楼	510630

