

潜藏因子的信息含量: 来自中国国债市场的证据

郑振龙, 陈 蓉, 黄海峰

(厦门大学 金融系, 厦门 361005)

摘 要 潜藏因子是指对当前利率曲线没有影响或影响甚微, 但对未来利率具有预测作用的信息. 本文采用 Duffee 的潜藏因子分析框架, 发现中国市场上中短期利率的五因子模型的方差比仅为 0.41, 较好地证实了潜藏因子的存在性. 接着从风险溢价因子中剔除利率曲线信息后得到潜藏因子, 并且证实了潜藏因子包含着丰富的经济信息含量, 并集中体现为通货膨胀和宏观流动性的信息, 特别是通货膨胀信息, 但与经济增长类指标联系微弱. 最后, 本文发现中国市场上的潜藏因子对未来利率变化具有良好的预测力.

关键词 期限溢价; 潜藏因子; 信息含量

Information content of the hidden factor: Evidence from the China's treasury bond market

ZHENG Zhen-long, CHEN Rong, HUANG Hai-feng

(Department of Finance, Xiamen University, Xiamen 361005, China)

Abstract The hidden factor is the information that has no or little effect on the current interest rate curve but affects the expected interest rates. Using the hidden factor idea proposed by Duffee, we find that the variance ratio of the five-factor model in China's treasury bonds market is only 0.41, which confirms the existence of the hidden factor. Then we exclude the interest rate curve information from the risk premium factor to get the hidden factor, and confirm that the hidden factor in China does contain a wealth of economic information content. It is closely linked with the inflation and the macro-liquidity, especially the inflation indexes. But the hidden factor has little to do with the economic growth information. At last we find the hidden factor in China has rather great power on predicting future interest rate.

Keywords term premium; hidden factor; information content

1 导论

正如人们关心股票价格是否可预测一样, 利率是否可预测是固定收益领域永恒的研究课题. 近十年来, “宏观金融” (macro finance) 是这一方向的主流研究视角之一, 即在利率信息之外引入各种宏观经济变量, 无论是国外还是国内, 研究者们均发现其能够在相当大的程度上预测未来的利率.

然而, 在这一研究热潮中, 有一个十分微妙却又非常重要的基础性问题被大多数人忽略了: 弱式有效和金融变量服从马尔可夫过程¹是金融的经典假设, 在利率曲线之外的经济信息对未来具有预测性, 是否意味着在固定收益领域, 这一经典假设已经不再成立? 反之, 若利率仍然服从弱式有效和马尔可夫过程, 当前的利率期限结构已经包含了未来利率变动的所有预期信息, 其他经济信息已经充分体现在当前利率曲线中, 就无需进行此类“宏观金融”的研究了.

收稿日期: 2014-08-19

资助项目: 国家自然科学基金 (71371161, 71471155); 国家自然科学基金青年项目 (71101121); 国家自然科学基金地区项目 (71261024)

作者简介: 郑振龙 (1966-), 男, 福建人, 博士, 国务院学科评议组成员, 国务院政府特殊津贴专家, 国家“百千万”人才工程入选者, 厦门大学闽江学者教授, 博士生导师, 研究方向: 资产定价、金融工程和风险管理, 电话: 13906038903, E-mail: zlzhen@xmu.edu.cn; 陈蓉 (1976-), 女, 汉, 福建人, 博士, 教授, 研究方向: 资产定价、金融工程和风险管理, 电话: 13860126618, E-mail: aronge@xmu.edu.cn; 黄海峰 (1987-), 男, 汉, 重庆人, 硕士, 研究方向: 资产定价、金融工程与风险管理, E-mail: hhfnum1@sohu.com.

1. 若某个变量遵循马尔可夫过程, 则它的期望值只跟该变量的当前值有关, 跟该变量的其它信息无关.

本文运用 Duffee^[1] 的潜藏因子分析框架在中国国债市场上对这一有趣的问题进行了研究. 结果发现, 在运用五因子模型刻画我国国债利率时, 运用利率曲线本身的五个主成分只能刻画利率变动方差的 41%, 我国国债市场上的确存在着潜藏因子, 其对未来利率具有显著的预测力, 但却不反映在当前的利率期限结构中, 中国市场上的利率并不满足马尔可夫性质, 引入其他经济变量预测未来利率的确是有意义的. 我们还发现, 我国国债市场上的潜藏因子包含着丰富的经济信息含量, 并集中体现为通货膨胀和宏观流动性的信息, 当前通货膨胀越严重, 宏观流动性越紧张, 潜藏因子越高, 但潜藏因子与经济增长类指标联系微弱.

本文是首篇对中国利率潜藏因子现象进行研究的论文, 尽管提取和研究潜藏因子的实证过程相当复杂, 其结果却较好地证实了中国市场上潜藏因子的存在性、信息含量和预测力, 也较好地回答了马尔可夫性质和弱式有效性这一基础性问题.

此外, 我们的论文也为利率动态模型的建模提供了新的建议: 常见的三因子模型可能在拟合静态利率曲线时表现良好, 但在解释期限溢酬和利率变化时, 却可能遗漏潜藏因子信息, 最好要增加因子数量, 或直接对潜藏因子建模.

最后, 我们的研究采用了与现有潜藏因子相关研究不同的思路. 本文不像 Joslin, Pribsch 和 Singleton^[2]、Barillas^[3] 那样, 先验地假定某些宏观经济变量会对未来利率有预测作用, 并且符合潜藏因子的性质, 然后再进行研究以证实潜藏因子的存在性和经济内涵. 我们认为这样的人为设定可能会导致模型误设, 因此是先基于潜藏因子性质提取出潜藏因子, 再反向去寻找该指标所隐含的经济含义, 发现效果不错, 是一个值得探索的研究角度.

下文的结构安排如下: 第 2 部分为文献综述, 第 3 部分介绍了本文研究的基本框架和方法, 第 4 部分是对我国银行间国债市场的实证研究, 第 5 部分为结论.

2 文献综述

利率可预测性的开创性研究要追溯到 Fama 和 Bliss^[4] 以及 Campbell 和 Shiller^[5], 他们分别发现远期利率与短期利率之差、长短期利差能够在一定程度上预测未来利率. Cochrane 和 Piazzesi^[6] 则进一步对远期利率进行线性组合, 发现相应构造出的 CP 因子对未来利率的预测最高能达到 44%. 很显然, 无论是远期利率与短期利率之差、长短期利差还是 CP 因子, 都是当前利率期限结构中所蕴含的信息, 上述三篇文献是运用当前利率期限结构的信息预测未来利率变动的经典文献.

值得一提的是, 在此类研究中, 常以如下定义的期限溢酬 (term premium, TP)²作为未来利率预期变动的代理变量:

$$\begin{aligned} TP_{t,t+n} &= E_t \left(\ln \frac{p_{t+1}}{p_t} \right) - R_{t,t+1} \\ &= nR_{t,t+n} - (n-1) E_t (R_{t+1,t+n}) - R_{t,t+1} \end{aligned} \quad (1)$$

其中, p_t 和 p_{t+1} 分别表示 $t+n$ 时刻到期的零息债在 t 时刻和 $t+1$ 时刻的价格, R 为即期利率, 其下标的第一项和第二项分别表示该利率对应期间的起始时刻和到期时刻, $E_t (R_{t+1,t+n})$ 表示预期的未来利率. 可以看出, 式(1)是 n 年期零息债持有一年的预期收益率与一年期即期利率之差, 因此常常也被称为“预期的持有期超额收益”, Dai 和 Singleton^[9] 证明了其就是投资者对利率风险所要求的风险溢酬. 国内研究中陈蓉、郑振龙^[10] 也指出在一个有效市场中, 均衡状态下远期价格与即期利率预期之间的差额, 都应该是标的资产的系统性风险溢酬. 由于式 (1) 中包含预期利率, 其又与买卖零息债的实际操作简单对应, 学术研究中常以此指标代理未来利率的预期变动.

一些研究者拓展了这一领域的研究. 他们的思路是: 上述文献多采用回归分析的方法, 作为解释变量的利率期限结构信息也直接源于某些期限的长短期利率或远期利率, 而利率曲线本身是由无穷多个不同到期期限的收益率组成的, 仅使用部分期限的数据进行研究可能会丢失信息. 而若使用瞬时利率动态模型, 只需对少数因子建模就能够刻画整条利率期限结构, 就能较好地涵盖利率曲线的全面信息, 其中仿射 (线性) 模型因其处理上的方便性质而被广泛使用. 例如 Duffee^[11] 发现三因子高斯仿射模型可以很好的描述美国国债市场的利率期限结构; 范龙振^[12] 用三因子广义高斯仿射模型对我国利率期限结构数据进行了研究, 发现模型基

2. 期限溢酬是利率风险对应的风险溢酬. 常见的期限溢酬计算方法有三种, 基于远期利率的期限溢酬、基于即期利率的期限溢酬和基于持有期收益率的期限溢酬. 学术研究中最常见的是第三种, 即本文所采用的定义. 期限溢酬的定义与分类在 kim^[7] 以及郑振龙和吴颖玲^[8] 中均有详细介绍.

本能够刻画上交所利率期限结构的相对变化。

尽管多因子动态利率模型的引入为利率预测的研究开拓了新思路, 但一个新的问题接踵而来: 多因子模型中的各因子 (状态变量) 缺乏实际的经济含义. Ang 和 Piazzesi^[13] 为解决这一问题提供了开创性的研究. 他们在期限结构无套利的 VAR 模型中引入通货膨胀因子和实体经济因子, 与利率曲线因子一起解释利率的变动, 结果发现相对于仅包含利率曲线因子的模型而言, 额外包括了宏观经济变量的模型具有更好的预测效果: 宏观经济变量解释了 85% 的短中期利率变动, 无套利约束和宏观经济变量的引入有效提升了样本外预测的效果. 其他不少研究者也都证实了宏观经济变量对期限溢酬和未来利率变动的良好预测效果, 如 Ludvigson 和 Ng^[14]、Piazzesi 和 Swanson^[15]、Wright^[16]、Joslin, Priebisch and Singleton^[2]、范龙振和张处^[17]、董莉莎和朱映瑜^[18] 等. 除了宏观经济因子之外, 另一些研究者还发现滞后的远期利率、人口因素、市场方差风险溢酬、异质信念等也对期限溢酬和未来利率变动具有预测能力, 如 Cochrane 和 Piazzesi^[6] 等.

在上述经济因子的探索热潮中, Duffee^[1] 独辟蹊径, 对上述具有预测力的经济因子是否已经体现在当前的利率期限结构中进行了思考. 如果是, 显然只需要运用当前的利率期限结构信息对未来进行预测即可; 如果不是, 是否意味着利率期限结构不再遵循公认的马尔科夫过程? Duffee^[1] 证实了潜藏因子的存在性, 其对期限溢酬和未来利率具有预测作用, 但却不会体现在当前的利率期限结构中. 我们认为, Duffee^[1] 的研究具有十分重要的价值, 其提出了一个不同于传统马尔科夫模型的全新研究视角, 也为近年来宏观金融等与微观利率期限结构的融合研究提供了重要的理论支持, 因为倘若不存在潜藏因子, 就意味着宏观金融等信息已经充分体现在微观市场价格中, 也就无需进行此类研究了.

在 Duffee^[1] 之后, 研究者在构建多因子模型时, 开始引入利率曲线之外的信息 (yields-plus), 如 Joslin, Priebisch 和 Singleton^[2]、Barillas^[3] 等. 但目前此类研究的主要做法仍然是先验地设定某些经济变量可能对未来利率的变化具有预测作用, 并要求这些经济变量既是状态变量的函数, 同时又符合潜藏因子的性质. 尽管研究表明额外信息的确能够改善预测效果, 但这样的人为设定可能会导致模型误差.

事实上, 我们认为, Duffee^[1] 可以为多因子动态利率模型、因子经济内涵、期限溢酬以及预测利率等领域研究提供更多的启发和思路. 本文就是一个尝试. 与前述先验设定的研究思路不同, 本文转换视角, 首先运用 Duffee^[1] 的潜藏因子思路, 提取出对期限溢酬和未来利率具有预测作用的潜藏因子, 再将其映射到拥有具体经济含义的宏观经济变量和货币政策变量上, 反向探索潜藏因子所拥有的信息含量, 就可规避模型设定等问题, 为利率的可预测性和动态利率模型的经济含义提供新的证据. 与 Duffee^[1] 的研究区别还在于, Duffee^[1] 研究的重心是验证潜藏因子的存在性, 而本文关注的重点是通过借鉴 Duffee^[1] 的潜藏因子研究框架去分析其背后隐含的宏观经济信息含量.

3 研究框架

在本节中, 我们将简要介绍 Duffee^[1] 所分析的潜藏因子的理论存在性, 并介绍本文实证工作所遵循的方法和步骤: 首先验证潜藏因子在中国市场中的存在性, 之后运用适合中国市场的方法提取隐含信息, 最后分析其信息含量和对利率的预测能力.

3.1 潜藏因子的理论存在性

Duffee^[1] 在高斯仿射模型框架下深入讨论了潜藏因子的理论合理性. 其基本逻辑如下: 在高斯仿射模型框架下, t 时刻的即期利率向量 R_t 可表示为状态变量的线性形式,

$$R_t = A + BX_t \quad (2)$$

其中, X_t 为服从一阶马尔科夫性质的状态变量向量, A 和 B 为系数矩阵. 如果矩阵 B 可逆, 我们有

$$X_t = (B)^{-1} (R_t - A) \quad (3)$$

在此种情况下, 由于状态向量 X_t 服从一阶马尔科夫过程, 即期利率向量 R_t 自然也服从一阶马尔科夫过程, 而马尔科夫过程的性质意味着对未来利率的预测仅和当前时刻的利率期限结构有关.

在 Duffee^[1] 之前, 绝大多数研究都将矩阵 B 可逆看成是理所应当的假设. 然而事实上, 并没有任何理论保证 B 一定可逆. Duffee^[1] 考察了矩阵 B 不可逆意味着什么. 假设影响即期利率的两个相互独立的因子为瞬时利率 r_t 和其他因子 h_t , 前者代表利率期限结构自身的信息, 后者代表利率期限结构之外的影响因子.

假设在现实测度下, 这两个状态变量的动态过程为

$$\begin{pmatrix} r_{t+1} \\ h_{t+1} \end{pmatrix} = \mu + \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} \begin{pmatrix} r_t \\ h_t \end{pmatrix} + \sum \varepsilon_{t+1} \quad (4)$$

其中, ε_{t+1} 是对状态变量的外生冲击. 显然, 瞬时利率在现实测度下服从两因子的一阶马尔科夫过程. 如果 $k_{12} \neq 0$, t 时刻关于未来的瞬时利率 r_{t+1} 的预期值就同 r_t 和 h_t 都有关; 也就是说, $k_{12} \neq 0$ 意味着 h_t 包含了不影响同期利率期限结构、却影响未来利率预期的信息, h_t 是一个潜藏因子.

在经济上如何理解这样的现象呢? 由于当前的利率期限结构同时受到利率预期和期限溢价 (即利率风险溢价) 的影响, 如果潜藏因子同时影响了利率预期和期限溢价, 而且两者恰好相互抵消, 就会出现这样的情况. 例如, 宏观经济向好, 可能使得利率预期上升, 但同时降低了投资者的风险厌恶和所要求的风险溢价, 这一宏观经济因素就可能会影响未来的利率预期, 但不影响当前的利率曲线.

当 h_t 是一个潜藏因子时, 式 (3) 中的两个状态变量在风险中性测度下的随机过程只能设定为

$$\begin{pmatrix} r_{t+1} \\ h_{t+1} \end{pmatrix} = \hat{\mu} + \begin{pmatrix} \hat{k}_{11} & 0 \\ \hat{k}_{21} & \hat{k}_{22} \end{pmatrix} \begin{pmatrix} r_t \\ h_t \end{pmatrix} + \sum \hat{\varepsilon}_{t+1} \quad (5)$$

当 h_t 是一个潜藏因子时, 矩阵 B 不再可逆, 并且可以得到 n 年期零息债持有 1 年的长债短持策略的持有期超额收益 (即期限溢价) 为

$$\begin{aligned} TP_{t,t+n} &= nR_{t,t+n} - (n-1)E_t(R_{t+1,t+n}) - R_{t,t+1} \\ &= nA_n - (n-1)A_{n-1} + (1 - \hat{k}_{11})^{-1} [1 - (\hat{k}_{11})^n - (1 - (\hat{k}_{11})^{n-1})k_{11} - 1]r_t \\ &\quad - (1 - \hat{k}_{11})^{-1} (1 - (\hat{k}_{11})^{n-1})k_{12}h_t \end{aligned} \quad (6)$$

可以看出, 期限溢价不仅依赖于当前的利率期限结构信息 r_t , 也受到当前潜藏因子 h_t 的影响. 式 (4) 和 (6) 共同表明, 在一定的条件下, 可能存在潜藏因子, 其会同同时影响未来利率预期和期限溢价, 但却并不影响当前的利率期限结构.

事实上, 上述分析仅是一个例子和一种可能的情形, 用以证实潜藏因子的理论存在性. 在现实中, 可能并不存在一个纯粹的潜藏因子, 而是在某些因子中包含着部分“隐含”信息, 这些隐含信息同时影响未来利率预期和期限溢价, 但因相互抵消等原因, 并未体现在当前的利率期限结构中, 这些因子可称为“半潜藏因子”.

3.2 验证潜藏因子的现实存在性

在证实潜藏因子在理论上的可能性之后, Duffee^[1] 提出可以采用如下的方差比, 以验证潜藏因子的现实存在性:

$$\frac{Var(TP_{t,t+n} | Z_t)}{Var(TP_{t,t+n} | X_t)} \quad (7)$$

其中 X_t 是 m 个驱动利率变化的因子向量, Z_t 是从利率期限结构中提取出的前 m 个主成分向量, 其个数与状态变量的维数相同. 上述方差比越小, 越表明 X_t 中含有 Z_t 之外的信息, 的确存在着能够预测未来利率的信息“隐藏”在利率曲线之外.

为计算得到式 (7) 的方差比, 需要先设定因子数量 m . Litterman 和 Scheinkman^[19] 早已指出期限结构一般由水平、斜率和凸性三因子描述, 所以要提取潜藏因子, 模型因子个数至少为 4; 而因子个数过多又可能出现过度拟合的问题, 因此本文主要考察四因子和五因子模型, 并与三因子模型进行对比.

其次, 如式 (6) 所示, 期限溢价 TP 可以表达为因子模型中各状态变量及其参数的表达式, 因此只要估计出上述状态变量的时间序列和参数, 就可以计算得到式 (7) 方差比中的分母项 $Var(TP_{t,t+n} | X_t)$. 对于分子项 $Var(TP_{t,t+n} | Z_t)$, 由于期限结构的前 m 个主成分本身就是不同到期期限的即期利率的线性组合, 因此同样可以运用各状态变量的时间序列和参数计算得到³.

那么如何估计这些状态变量及其参数呢? 本文的做法是仅对三因子、四因子和五因子模型进行一般化的设定, 然后运用卡尔曼滤波和极大似然方法估计出与实际数据最为吻合的各状态变量的参数, 得到各状态变量的时间序列、均值和方差, 求出方差比后即可判断潜藏因子的现实存在性.

3.3 提取潜藏因子

如果从方差比的计算中发现了隐含信息的存在性, 我们自然想要进一步提取潜藏因子, 以考察其性质和对未来利率的预测能力.

3. 具体公式可由 Duffee^[1] 的式 (26) 与附录 B 进一步得到.

由于本文并未直接设定多因子模型中的特定状态变量为“纯粹潜藏因子”, 因此也无法直接提取得到, 而需要采用其他较为巧妙的方法.

从式 (6) 可知, 期限溢酬同时受到非潜藏因子 (即当前的利率期限结构信息) 和潜藏因子的共同影响. 那么, 如果从期限溢酬受因子影响的部分中剔除当前利率期限结构的信息, 就可以得到潜藏因子的对应信息.

基于这一逻辑, 我们需要先得到期限溢酬中受状态变量影响的部分. 从鞅定价基本原理可知, 若假设风险价格向量 λ_t 服从如式 (8) 的实质仿射形式⁴(essentially affine form), 即

$$\begin{aligned}\lambda_t &= \lambda_0 + \lambda_1 X_t \\ \lambda_0 &= (\lambda_{0,L} \lambda_{0,S} \mathbf{0}_{1 \times (m-2)})' \\ \lambda_1 &= \begin{pmatrix} \lambda'_{1,L} \\ \mathbf{0}_{(m-1) \times m} \end{pmatrix}\end{aligned}\quad (8)$$

期限溢酬就可以表达为状态变量的线性组合, 其中, λ_t 为风险价格, $\lambda_{0,L}$ 、 $\lambda_{0,S}$ 分别是水平风险和斜率风险⁵的固定补偿部分, $\lambda_{1,L}$ 为针对水平风险的 m 维向量.

式 (8) 的含义是, 投资者对斜率风险所要求的风险价格为固定值, 而对水平风险所要求的风险价格呈时变特征, 对其他风险因子则不要求风险溢酬. 除了能够使期限溢酬表达为状态变量的线性组合, 大大便利实证研究, 从已有的研究来看, 这样设定亦是可接受的: 首先, 从现有研究来看, 上述对风险价格的设定是基本合理的, 例如郑振龙和吴颖玲^[8] 用简单的回归分析发现水平、斜率因子的风险溢酬都显著, 但斜率因子影响很小; 其次, 其他状态变量对应的风险价格设定为 0 可能过于严格, 但一般公认, 过于复杂的风险价格设定往往导致过度拟合, 反而可能弊大于利.

具体地说, 在式 (8) 的设定下, 我们将期限溢酬中与状态变量相关的部分定义为 RP (risk premium), 其表达式为

$$RP_t \equiv \lambda'_{1,L} X_t \quad (9)$$

为剔除 RP 中的当前利率期限结构信息, 得到潜藏因子的信息, 可以采用回归式

$$RP_t = b_0 + b' Z_t + H_t \quad (10)$$

其中, Z_t 依然表示的是利率期限结构的前 m 个主成分. 显然回归得到的残差部分即可视为我们所要提取的隐含信息.

Duffee^[1] 即是采用以上方式对潜藏因子进行提取, 但在本文的实证中, 由于样本期内的主成分平稳性不一致, 无法进行式 (10) 的回归, 因此我们直接采用以下公式计算得到潜藏因子:

$$\begin{aligned}H_t &= RP_t - E(RP_t | Z_t) = \lambda'_{1,L} (X_t - E(X_t | Z_t)) \\ &= \lambda'_{1,L} (X_t - E(X_t) - Var(X_t) B' P' Var(Z_t)^{-1} Z_t)\end{aligned}\quad (11)$$

其中, B 是观测方程式 (2) 中的因子载荷矩阵, P 则是将利率向量 R_t 转化为主成分 Z_t 的系数矩阵. 将卡尔曼滤波估计出的参数和滤出的状态变量时间序列代入式 (11), 即可得到潜藏因子的时间序列.

3.4 探索潜藏因子的信息含量

在估计得到潜藏因子的时间序列之后, 我们探索了潜藏因子的信息含量和对未来利率的预测能力. 我们首先用潜藏因子对主要经济变量进行回归, 以考察潜藏因子蕴涵着怎样的经济信息, 然后运用 VAR 模型进一步探索了它们之间的相互作用, 最后, 我们考察了潜藏因子对未来利率的预测能力.

4 实证分析

4.1 数据描述

本文使用我国银行间国债 1 到 6 年期⁶即期利率周数据作为研究对象, 样本期为 2005 年 1 月到 2012 年 5 月, 基本涵盖了完整的升息与降息期, 经历过货币政策的松紧变化. 因为涉及到假期, 利率每年的周数据量并非 52 个. 由于周三的数据量最为完整, 所以本文将每周三数据作为样本点, 共得到 365 周的数据, 并将利率全部转化为连续复利形式.

4. 郑振龙、柯鸿和莫天瑜 [20] 对风险价格形式及对应的仿射模型的特点进行了具体的探讨.

5. 在后文的实证中, 多因子模型中的前三个状态变量对应着利率期限结构的前三个主成分, 即水平因子、斜率因子和曲度因子.

6. 我国债券市场流动性较差, 数据质量较差, 早期尤甚, 因此主要根据数据可得性选用中短期利率, 以减少流动性等其他因素的影响.

在经济变量方面, 由于大量文献都将通货膨胀与经济增长看作影响利率变化的重要因素, 因此本文选取了样本期内的 CPI 和 PPI 作为通货膨胀的代表性指标, 并选取工业增加值同比、固定资产投资同比和社会消费品零售总额同比作为经济增长的代表性指标, 考察潜藏因子与这些经济变量的关系. 此外, 为全面分析起见, 我们额外引入了 M1 同比、M2 同比、社会融资总量同比和货币政策调整 4 个变量, 考察潜藏因子与这些货币供应量和宏观流动性指标的关系, 其中货币政策调整是虚拟变量, 采用的是 2005 年 1 月-2012 年 5 月的货币政策调整信息: 当月若出现加息或存款准备金率上调, 变量值为 1, 当月出现降息或存款准备金率下调, 则变量值为 -1, 当月没有出现货币政策调整操作则变量值为 0; 我们还引入了 PMI 和消费者信心指数, 考察潜藏因子与市场信心的关系; 最后, 我们还考察了人民币/美元汇率和沪深 300 指数等资产价格与潜藏因子的关系. 所有宏观数据的频率为月度, 同比增长形式的变量都先转化为对数收益率形式, 汇率和沪深 300 指数均取对数收益率, ADF 检验非平稳的变量通过差分将其平稳化. 本文的数据来源为 Wind 数据库.

4.2 卡尔曼滤波与参数估计

如前所述, 本文利用卡尔曼滤波估计参数和滤出状态变量时间序列.

基于式 (2), 卡尔曼滤波的观测方程为:

$$R_t = A + B X_t + \eta_t \quad \eta_t \sim N(0, I\sigma_\eta^2) \quad (12)$$

其中, η_t 为观测误差向量, 此处本文沿用常见假设, 即所有期限利率的观测误差具有同方差 σ_η^2 . 对于状态变量, 本文亦沿用常见假设, 根据式 (8) 将状态方程设定为

$$\begin{aligned} X_{t+1} &= \mu + K X_t + \Sigma \varepsilon_{t+1} \\ &= (\hat{\mu} + \lambda_0) + (\hat{K} + \lambda_{1,L}) X_t + \Sigma \varepsilon_{t+1}, \quad \varepsilon_{t+1} | X_t \sim N(0, I) \end{aligned} \quad (13)$$

其中 Σ 为对角阵, 且对角元素按降序排列, $\hat{\mu}$ 设置为 0⁷. 这样, 待估参数为: 瞬时利率对状态变量的参数向量 ζ^8 、 Σ 的对角元素、以及 K 、 σ_η 、 λ_0 和 $\lambda_{1,L}$.

表 1 报告了五因子⁹仿射模型的 44 个参数值. 状态变量所服从的向量自回归过程特征根均在单位圆内, 满足平稳性条件.

表 1 卡尔曼滤波参数估计结果

ζ_0	ζ_1	ζ_2	ζ_3	ζ_4	ζ_5
0.0242	0.0734	2.0666	-0.6796	1.2628	1.8278
K	0.7017	0.9854	-0.2624	0.5914	-0.1461
	0.1593	0.1863	0.6421	-0.1770	-0.3340
	-0.3149	1.1338	-0.5445	-0.8166	0.0591
	0.1233	-0.4254	0.2659	0.9831	0.1295
	-0.32038	1.1795	-0.6186	-0.6094	0.2737
σ_η	0.0014				
diag Σ	0.6419	0.1403	0.0875	0.0364	0.0082
λ_0	1.1216	-0.1745			
$\lambda_{1,L}$	1.0359	1.3586	0.5659	3.1022	7.1132

注: 本文所有数值都保留 4 位小数, 下同.

为检验上述卡尔曼滤波估计的可靠性, 我们对真实利率值与模型拟合值进行了对比, 观察估计得到的参数在样本内对利率拟合的程度. 图 1 给出的是 5 年期国债即期利率的对比图. 可以看出, 样本内的模型拟合效果相当不错.

4.3 验证与提取潜藏因子

在完成参数估计和状态变量滤出后, 如第 3 节所述, 我们通过计算三因子、四因子和五因子模型的方差比来验证潜藏因子的存在性, 并对模型进行选择. 表 2 报告了方差比的结果.

7. 不少文献把整个 μ 都设为 0, 对结果影响不大.

8. $r_t = \xi_0 + \xi_1' X_t$, 待估参数 ζ 即表示 ζ_0 和 ζ_1 组成的向量, 在五因子模型下, ζ_1 为 5 维向量.

9. 我们实际上做了三因子、四因子和五因子, 在 4.3 节可以看到, 五因子的方差比指标最佳, 因此这里仅报告五因子结果.

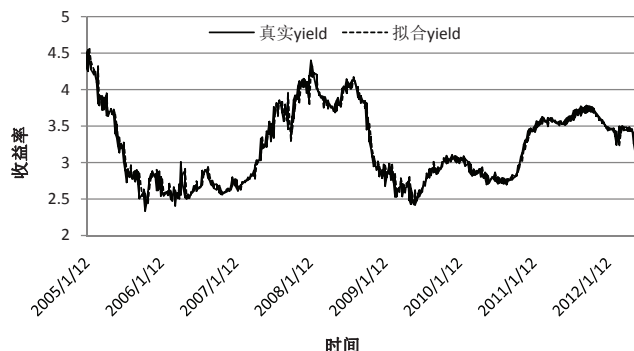


图 1 5 年期国债即期利率实际与拟合利率走势图 单位 (%)

可以看出, 使用三因子模型时, 方差比几乎接近于 1, 说明在只用三个状态变量建模时, 与直接用利率期限结构提取出的前三个主成分建模的信息含量几乎是一样的, 并没有体现出在利率曲线之外的隐含信息. 但当我们增加一个因子时, 方差比降到 0.8558, 潜藏因子的特性已有所体现; 当仿射模型设定为五因子时, 方差比大幅降低为 0.4068, 也就是说, 造成期限溢价变动的信息有一半以上隐藏在了当前利率曲线之外. 上述结果表明: 通过五因子模型可以清晰地发现不同于传统三因子模型的期限溢价性质, 本文重点关注的潜藏因子的确存在. 借助于方差比的结果, 我们将模型的维数选取为五因子.

证实了潜藏因子的存在后, 我们运用式 (11) 对潜藏因子进行提取, 以进一步分析其信息含量. 为证实提取得到的信息的确具有隐含的特征, 我们用样本内所有到期期限的即期利率分别对 RP (即期限溢价中与状态变量相关的部分) 和潜藏因子进行回归, 图 2 报告了回归得到的各期限的因子载荷.

可以明显看出, RP 同当期利率呈现显著的负相关关系, 而潜藏因子的因子载荷在图形中则几乎呈现水平, 即潜藏因子与当期的利率期限结构看不出明显关联, 这说明在提取潜藏因子时, 我们已经很好地剔除了利率期限结构的信息, 潜藏因子对当期期限结构影响甚微, 表现出我们所期待的“潜藏”性质.

4.4 潜藏因子的信息含量

在提取得到潜藏因子之后, 接下来我们首先对其本身进行简单分析, 然后引入宏观经济变量, 考察潜藏因子的信息含量. 由于宏观数据为月度数据, 而潜藏因子为周度数据, 本节中取每月最后一周的潜藏因子与宏观数据相匹配.

4.4.1 潜藏因子时间序列

图 3 描绘了我们提取得到的潜藏因子在时间序列上的表现. 可以看出, 在 2007–2008 期间, 潜藏因子明显相对较高, 2009 年 4 万亿宽松刺激政策之后, 潜藏因子显著降低, 这显然意味着潜藏因子是与经济市场状况存在合理联系的, 蕴涵着一定的经济信息. 对潜藏因子的进一步时间序列初步分析表明, 潜藏因子是平稳序列, 但自相关程度较高, 这些性质无论在经济上还是统计上显然都是合理的.

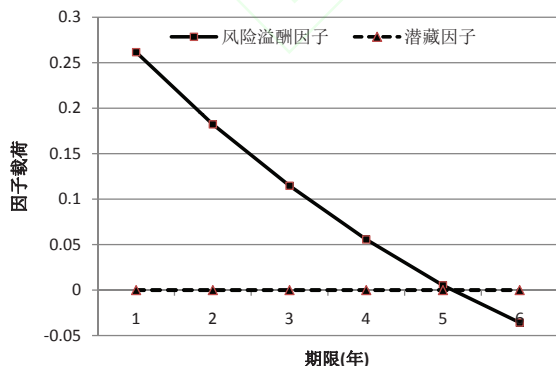


图 2 利率对潜藏因子和 RP 的载荷

4.4.2 回归分析

由于其简单有效的优点, 线性回归分析在解释期限溢价的文献中被广泛使用. 本文首先利用简单的单变量回归对潜藏因子的信息含量进行初步分析. 表 3 报告了以潜藏因子为因变量, 对各经济变量同期值¹⁰进行



图 3 潜藏因子时间序列

10. 事实上, 我们对各经济变量的滞后值也进行了回归分析, 其结论基本相同.

单变量线性回归的主要结果. 为了控制自相关与异方差的影响, 所有的统计量都已经过 Newey-West 调整.

表 3 潜藏因子的信息含量

指标名称	系数	p 值	调整 R^2
CPI	0.3673	0.0000***	43.90%
PPI	0.1173	0.0017***	12.63%
汇率	-0.9358	0.0126**	9.30%
沪深 300 指数	-0.0341	0.0574*	5.66%
M2	-0.0988	0.0056***	5.61%
M1	-0.1553	0.0310**	3.58%
货币政策调整	0.4767	0.0848*	3.44%
社会融资总量	-0.0024	0.2622	0.09%
PMI	-0.0335	0.5037	-0.49%
消费者信心指数	-0.0536	0.4274	-0.74%
固定资产投资完成额	-1.2622	0.7584	-1.00%
社会消费品零售总额	0.0232	0.6041	-1.04%
工业增加值	0.0065	0.8475	-1.10%

注: 加粗部分为系数显著异于零的结果. *, ** 和 *** 分别表示 10%、5% 和 1% 的显著性水平, 下同.

表 3 的结果是按调整 R^2 排序的, 其相当明晰地揭示了潜藏因子的信息含量:

(1) 从信息内容看, 我国国债市场 6 年以下利率的潜藏因子中主要包含的是通货膨胀与宏观流动性的信息, 而非经济增长的宏观信息. 这从 CPI、PPI、M1、M2 和货币政策调整的回归系数显著、回归 R^2 较大, 而工业增加值指标、投资、消费和信心指数的系数均不显著、回归 R^2 极低可以看出. 汇率与沪深 300 指数的显著性也可以在同一个逻辑下得到解释: 在中国现有的特定外汇环境下, 汇率的变动总与热钱的流动和货币供应密切相关; 宏观流动性的松紧亦是中国股市的涨跌的重要影响因素. 总之, 我们初步可以得出结论: 我国的利率期限结构潜藏因子的确蕴涵着经济信息, 并且集中体现为通货膨胀和宏观流动性的信息, 这显然是符合经济直觉的.

(2) 从各因子的影响方向看, 表 3 中的回归系数基本都是合理的, 并具有内在一致性, 可以帮助我们深入理解潜藏因子的信息含量: CPI/PPI 的回归系数显著为正, 说明通货膨胀越严重, 潜藏因子越高; M1/M2 的回归系数显著为负, 说明货币供应越宽松, 潜藏因子越低, 货币供应越收紧, 潜藏因子越高; 货币政策调整的回归系数显著为正, 说明货币政策越收紧, 潜藏因子越高; 汇率的回归系数显著为负, 可以理解为汇率越下降 (即人民币越升值), 热钱流入的可能性越大, 货币政策收紧的可能性越高, 潜藏因子越高; 沪深 300 指数的回归系数显著为负, 可以理解为股市的上升往往源于宏观流动性的宽松, 此时潜藏因子较低. 总之, 从回归系数的方向我们可以进一步判断, 银根趋紧时潜藏因子较高, 银根放松时潜藏因子较低, 这与图 3 的基本走势显然也是一致的.

(3) 在对潜藏因子具有解释力度的各经济变量中, 与潜藏因子关系最为密切的是通货膨胀类信息, CPI 和 PPI 的回归调整 R 方分别达到了 43.91% 和 12.63%, 且显著性水平都小于 1%; 其他变量的回归调整 R 方则都在 10% 以下, 其中 M2 的影响最明显, 其回归系数的显著性水平低于 1%, 其余几个变量的显著性水平都在 1% 以上. 事实上, 当我们将潜藏因子对表 3 中单变量回归显著的经济变量进行多元回归时, 只有 CPI 依然在万分之一的水平下显著, 其他变量都不再显著, 再次说明潜藏因子主要蕴含的就是通货膨胀的信息.

4.4.3 向量自回归分析

上一小节中的回归分析初步证实了潜藏因子所蕴含的经济信息, 但其仅仅建立在单向的回归解释上, 并且主要涉及单变量. 而潜藏因子应当是隐含的宏观信息的综合体现, 自然联想到考察潜藏因子与众多宏观变量之间的相互影响. 本节我们运用 VAR(向量自回归模型) 将潜藏因子与各变量作为一个整体进行更深入的研究.

根据上一节单变量回归的结果, 本节 VAR 系统选择的变量本应为所有回归系数显著的经济变量, 但考虑到 CPI 和 PPI、M1 和 M2 的相关性很高, 存在重复信息, 因此最终进入 VAR 系统的变量为: 潜藏因子 H、CPI、M2、货币政策调整 MA、汇率 FX 和沪深 300 指数 HS300. 在综合考虑 AIC、SC 和 HQ 等准则后, VAR 模型的阶数定为 1 阶. 由于本文重点考察的是经济变量对潜藏因子的影响, 此处仅给出以潜藏因子

作为被解释变量的方程结果¹¹, 如式 (14) 所示.

$$H_t = 4.7156^{***} + 0.3646^{***}H_{t-1} + 0.2252^{***}CPI_{t-1} + 0.0272M2_{t-1} + 0.2181MA_{t-1} + 0.0995FX_{t-1} - 0.0016HS300_{t-1} \quad (14)$$

$$\bar{R}^2 = 46.11\%$$

VAR 模型的 AR 根均小于 1, 表明上述 VAR 模型是稳定的.

可以看出, 只有 CPI 的滞后值对潜藏因子具有显著影响, 系数依然为正, 而其他变量的滞后期系数都不再显著. 相应的 Granger 因果检验表明¹², 滞后一阶 CPI 是潜藏因子的 Granger 原因, 但潜藏因子并非 CPI 的 Granger 原因.

图 4 报告了各经济变量对潜藏因子的脉冲响应以及方差分解的结果. 从脉冲响应来看, 潜藏因子对其自身的一个标准差新息的反应最为强烈, 但到第 4 期之后就渐趋于消失; 在其他经济变量中, 潜藏因子对 CPI 的反应相对最强, 之后趋于消失. 从方差分解来看, 潜藏因子自身对其预测误差的贡献度达 60%, CPI 次之, 稳定在接近 20%. 总之, VAR 模型的结果均再次证实了潜藏因子的主要信息含量是通胀信息.

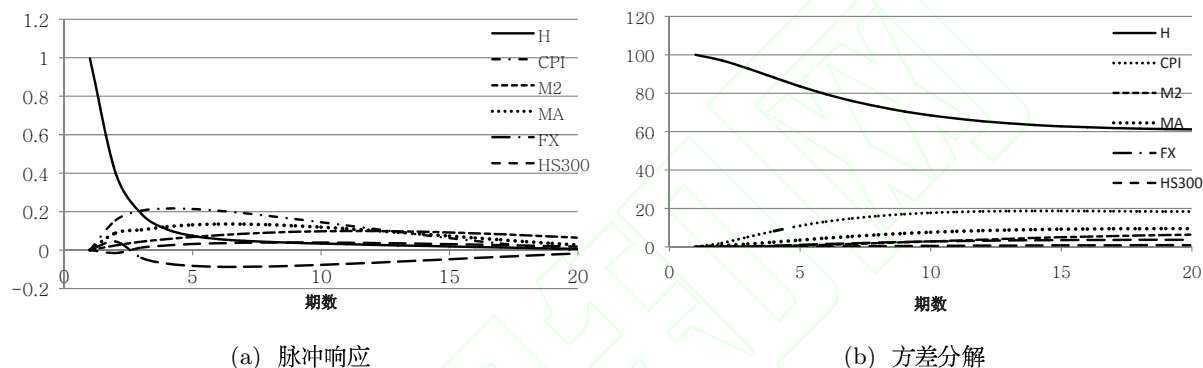


图 4 各经济变量对潜藏因子的脉冲响应与方差分解

4.4.4 潜藏因子的预测能力

在确认潜藏因子的信息含量之后, 我们还希望了解潜藏因子的预测能力, 其究竟能在多大程度上预测未来的真实利率变动? 我们计算了事后的真实一年持有期超额收益¹³, 即

$$RX_{t,t+n} = \ln \frac{P_{t+1}}{P_t} - R_{t,t+1} \quad (15)$$

然后将其对潜藏因子和其他经济变量¹⁴进行了回归, 包括持有期超额收益对潜藏因子的单变量回归 (模型 1)、对表 3 中显著的经济变量的多元回归 (模型 2)、对潜藏因子和表 3 中显著的经济变量的多元回归 (模型 3)、对表 3 中不显著的经济变量的多元回归 (模型 4)、对潜藏因子和表 3 中不显著的经济变量的多元回归 (模型 5)、对所有经济变量的多元回归 (模型 6)、对潜藏因子和所有经济变量的多元回归 (模型 7). 表 4 报告了结果.

表 4 的结果首先非常清晰地揭示了这样一个事实: 无论是单变量回归还是多元回归, 潜藏因子对未来利率的真实变化均具有显著的预测力; 而且, 从模型 3 与模型 2、模型 5 与模型 4、模型 7 与模型 6 的对比可以看出, 在增加潜藏因子之后, 拟合优度均得到了提高. 将表 4 与图 2 的结果相结合, 我们再次证实本文提取得到的确是潜藏因子, 其对未来利率具有预测效果, 但却不反映在当前的利率期限结构中, 中国市场上的利率期限结构中的确存在着隐含信息.

其次, 潜藏因子的回归系数为正, 这意味着当期的潜藏因子越高 (低), 一年后的债券价格 P_{t+1} 越高 (低), 即一年后的利率越低 (高). 由于前文我们已经发现潜藏因子与当期通货膨胀和银根松紧正向相关, 潜藏因子的回归系数为正意味着中国的利率在一年的时间上存在着“反转”现象: 当期通货膨胀越高, 银根越紧, 潜藏因子越高, 一年后的利率却将会越低, 反之亦然.

第三, 在引入其他经济变量之后, 我们发现, 在原先与潜藏因子关系显著的经济变量中, 只有原先相关性较为微弱的“沪深 300 指数”和“货币政策调整”对一年期债券超额收益具有预测力, 可以认为潜藏因子已

11. 系数右上角括号内的星号表示其系数显著性, *, ** 和 *** 分别表示 10%、5% 和 1% 的显著性水平.

12. 为节省篇幅, 此处不再详细列出 Granger 因果检验的结果.

13. 注意与式 (1) 不同, 式 (15) 是真实的持有期超额收益, 在本文实证中, 是 2~6 年期债券持有 1 年的超额收益的均值.

14. 与前文相同, 在 CPI 和 PPI 中我们仅选用了 CPI, 在 M1 和 M2 中仅选用了 M2.

经涵盖了 CPI 等关系密切变量的主要信息; 而原先与潜藏因子没有显著相关性的其他经济变量中, “固定资产投资完成额”和“社会融资总量”对一年期债券超额收益具有预测力, 也就是说, 这些指标的影响相对来说较为长期, 尽管不影响当前的潜藏因子, 但却对一年的利率变化具有影响和预测力; 同时, 包含所有变量的模型 4 的调整 R^2 最大, 达到 47.59%. 也就是说, 潜藏因子主要包含了通货膨胀和银根松紧的信息, 并且对未来利率变化具有显著的预测力, 但尚未涵盖所有具有预测力的经济信息.

表 4 潜藏因子的预测能力

预测变量	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6	模型 7
潜藏因子	0.9931***		0.8061**		1.0048***		0.6765**
CPI		0.5047**	0.1485			0.6149***	0.3041
M2		0.0284	-0.0437			-0.0500	-0.1127
货币政策调整		-1.2701***	-1.0951**			-1.6859***	-1.5835***
汇率		0.1223	0.1953			-0.2024	-0.0497
HS300		-0.0749***	-0.0704***			-0.0704***	-0.0698***
社会融资总量				-0.0027	-0.0011	0.0067	0.0063*
工业增加值				-0.0029	-0.0799	-0.1071	-0.1145
固定资产投资完成额				24.9750**	22.6713**	30.9662***	26.9584**
社会消费品零售总额				0.0348	-0.0545	-0.1160	-0.0846
PMI				-0.0566	0.01666	0.1368	0.1786
消费者信心指数				-0.1701	-0.0874	0.0545	0.0184
调整 R^2	25.85%	27.57%	36.72%	2.89%	30.5%	41.19%	47.59%

最后, 尽管与潜藏因子的预测力没有直接关系, 我们仍然可以从其他具有预测力的变量的回归系数中发现一定的规律: “货币政策调整”的回归系数显著为负、“社会融资总量”的回归系数显著为正、“固定资产投资完成额”回归系数显著为正说明当期货币政策放松、宏观流动性宽松、投资活跃时, 未来的利率越低, 这与潜藏因子呈现出的“反转”现象是一致的.

5 结论

本文以 2005 年 1 月 -2012 年 5 月的我国银行间国债市场为研究对象, 研究了我国国债利率期限结构中的潜藏因子问题, 发现了一些有意思的结论:

首先, 方差比的计算结果表明, 当我国的国债利率模型从三因子拓展至五因子时, 方差比从 0.99 大幅下降到 0.41, 这一方面说明在我国国债市场上的确存在着潜藏因子, 中国市场上的利率并不满足马尔科夫性质; 另一方面, 这也提醒我们, 在使用常见的三因子仿射模型研究期限溢酬和利率预期时需更加谨慎. 如果是为了进行利率曲线的拟合, 三因子模型有很良好的效果; 但如果是为了对期限溢酬和利率变化进行解释与预测, 三因子模型可能会遗漏掉大量的信息. 改进的思路可以像本文一样增加因子个数以体现潜藏因子的性质, 也可用如 Joslin, Priebisch 和 Singleton^[2]、Barillas^[3] 一样直接引入宏观变量与潜因子联合建模.

其次, 对潜藏因子与经济变量的回归分析和 VAR 分析表明, 潜藏因子的确包含着丰富的经济信息含量, 并且集中体现为通货膨胀和宏观流动性的信息, 尤其是 CPI 的信息, CPI 是潜藏因子的 Granger 原因, 当前通货膨胀越严重, 宏观流动性越紧张, 潜藏因子越高; 但潜藏因子与经济增长类指标联系微弱.

最后, 实证研究表明, 我国国债市场上的潜藏因子对未来利率的真实变化的确具有显著的预测力, 我们所提取得到的信息的确呈现出“隐含”的性质: 其对未来利率具有预测效果, 但却不反映在当前的利率期限结构中. 此外, 我们还发现中国的利率在一年的时间上存在着“反转”现象: 当期利率水平越高, 一年后的利率水平越低.

参考文献

- [1] Duffee G R. Information in (and not in) the term structure[J]. Review of Financial Studies, 2011, 24: 2895-2934.
- [2] Joslin S, Priebisch M, Singleton K J. Risk premiums accounting in macro-dynamic term structure models[R]. Working Paper, Stanford University, 2009.
- [3] Barillas F. Unspanned risk premia in the term structure of interest rates[R]. Working Paper, 2011.
- [4] Fama E F, Bliss R R. The information in long-maturity forward rates[J]. The American Economic Review, 1987, 77(4): 680-692.

- [5] Campbell J Y, Shiller R J. Yield spreads and interest rate movements: A bird's eye view[J]. *Review of Economic Studies*, 1991, 58(3): 495–514.
- [6] Cochrane J H, Piazzesi M. Bond risk premia[J]. *American Economic Review*, 2005, 95: 138–160.
- [7] Kim D H. The bond market term premium: What is it, and how can we measure it?[J]. *BIS Quarterly Review*, 2007: 27–40.
- [8] 郑振龙, 吴颖玲. 中国利率期限溢酬: 后验信息法与先验信息法 [J]. *金融研究*, 2009, 10: 68–82.
Zheng Zhenlong, Wu Yingling. Interest rate risk premium: Ex-post and ex-ante methods[J]. *Journal of Financial Research*, 2009, 10: 68–82.
- [9] Dai Q, Singleton K J. Expectation puzzles, time-varying risk premia, and affine models of the term structure[J]. *Journal of Financial Economics*, 2002, 63: 415–441.
- [10] 陈蓉, 郑振龙. 无偏估计、价格发现与期货市场效率 —— 期货与现货价格关系研究 [J]. *系统工程理论与实践*, 2008, 28(8): 2–11.
Chen Rong, Zheng Zhenlong. Unbiased estimation, price discovery and market efficiency: Relationship between futures prices and spot prices[J]. *Systems Engineering — Theory & Practice*, 2008, 28(8): 2–11.
- [11] Duffee G. Term premia and interest rate forecasts in affine models[J]. *Journal of Finance*, 2002, 57(1): 405–443.
- [12] 范龙振. 交所利率期限结构的三因子广义高斯仿射模型 [J]. *管理工程学报*, 2005, 1: 81–86.
Fan Longzhen. Modeling the term-structure of yields in the SSE with three-factor gaussian essential affine model[J]. *Journal of Industrial Engineering and Engineering Management*, 2005, 1: 81–86.
- [13] Ang A, Piazzesi M. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables[J]. *Journal of Monetary Economics*, 2003, 50: 745–87.
- [14] Ludvigson S C, Ng S. Macro factors in bond risk premia[J]. *Review of Financial Studies*, 2009, 22: 5027–67.
- [15] Piazzesi, Monika, Swanson E. Futures prices as risk-adjusted forecasts of monetary policy[J]. *Journal of Monetary Economics*, 2008, 55: 677–691.
- [16] Wright J H. Term premia and inflation uncertainty: Empirical evidence from an international panel dataset[J]. *American Economic Review*, 2009, 101: 1514–1534.
- [17] 范龙振, 张处. 中国债券市场债券风险溢酬的宏观因素影响分析 [J]. *管理科学学报*, 2009, 12(6): 116–124.
Fan Longzhen, Zhang Chu. Explanation of macro economic variables on bond risk premia in China[J]. *Journal of Management Sciences in China*, 2009, 12(6): 116–124.
- [18] 董莉莎, 朱映瑜. 宏观经济变量对中国国债风险溢价影响的实证研究 [J]. *南方金融*, 2011, 414: 9–12.
Dong Lisha, Zhu Yingyu. A panel-based empirical study of the influence of macroeconomic variables on risk premium of bonds[J]. *South China Finance*, 2011, 414: 9–12.
- [19] Litterman R, Scheinkman J. Common factors affecting bond returns[J]. *Journal of Fixed Income*, 1991, 1(1): 54–61.
- [20] 郑振龙, 柯鸿, 莫天瑜. 利率仿射模型下的利率风险价格形式实证研究 [J]. *管理科学学报*, 2010, 第 9 期: 4–15.
Zheng Zhenlong, Ke Hong, Mo Tianyu. Empirical research on specification of interest risk price under framework of affine DTSM[J]. *Journal of Management Sciences in China*, 2010, 9: 4–15.