

进度汇报：投资者情绪和债券风险溢价以及out-of-sample

沈煜豪 杨宇昊

所有数据和STATA处理代码都在[杨宇昊的Github](#)上。都进行了非常详尽的注释。

总回归模型

$$rx_{t+1}^{(n)} = b_{n,CP}CP_t + b_{n,LN}LN_t + b_{n,BW}BW_t + \xi_{t+1}^{(n)} \quad (1)$$

其中, $rx_{t+1}^{(n)}$ 为从一年期债券的收益中获得的超额收益

b 为系数

CP_t 为单一回报预测因子

LN_t 为宏观因素

BW_t 为情绪因素

$\xi_{t+1}^{(n)}$ 为误差项

rxba的来源

根据文献的公式, \overline{rx}_{t+1} 从以下公式计算

$$\overline{rx}_{t+1} = (1/4) \sum_{n=2}^5 rx_{t+1}^{(n)} \quad (2)$$

其中, $rx_{t+1}^{(n)}$ 的数据为小组其他成员所取得。目前, rxba的数据处理已经完成。难点在于为了以后的合并所做的日期处理。处理方法已上传Github. 我们采取的徐哥和鹏总的更长时间跨度的 \overline{rx}_{t+1} .

CPt的来源

CP_t 的数据为

$$\overline{rx}_{t+1} = \lambda_0 + \lambda_1 y_t^{(1)} + \lambda_2 f_t^{1 \rightarrow 2} + \lambda_3 f_t^{2 \rightarrow 3} + \lambda_4 f_t^{3 \rightarrow 4} + \lambda_5 f_t^{4 \rightarrow 5} + u_{t+1} \quad (3)$$

这个回归中的拟合值, 但是徐哥和鹏总小组已经计算好了CP, 我们将直接使用。

LNt来源

LN_t 来源于以下回归的拟合值

$$\overline{rx}_{t+1} = \theta_0 + \theta_1 F_{1t} + \theta_2 F_{1t}^3 + \theta_3 F_{3t} + \theta_4 F_{4t} + \theta_6 F_{8t} + v_{t+1} \quad (4)$$

原文使用的 Sydney Ludvigson 的网站已经失效，善用搜索找到了他的新网站

<https://www.sydneyludvigson.com/data-and-appendixes/>

下载了数据 `updated_LN_Macro_Factors_2018AUG`，包含了所需要的所有 Factor。

ISSUE

回归模型论文里提到的

with LN_t defined as the fitted values from (12) but with the variable F_{2t} eliminated from the regression model.

有两种可能：

1. 回归模型和拟合里直接剔除 F_{2t} ，**目前采用这种方法**
2. 回归模型中带 F_{2t} 计算系数，而计算拟合值（LN）的时候不带上 F_{2t}

BWt来源

BW_t 来源于以下回归的拟合值

$$\overline{rx}_{t+1} = \phi_0 + \phi_1 S_t^\perp + \phi_2 S_t^{\perp 2} + \phi_3 \Delta S_t^\perp + \tau_{t+1} \quad (5)$$

根据

Laborda and Olmo 2014

文章所述， S_t^\perp 在 [Investor sentiment data \(annual and monthly\)](#) 这篇数据中获得。

综合处理

就是跑了开头的回归。放在了 All-in-one 文件夹中。

out-of sample

out-of-sample 的内容都放在文件夹里了，`final-out` 是小组成员给的解释变量和被解释变量的信息，`do` 文件是代码（如果想要在本地跑，需要改一下 `do` 文件里的路径）。判断预测的准则是均方根误差 `RMSE`。原理和文字分析交给了豪哥。