

综合项目应用结果与分析

项目名称	基于投资策略预测隐性社交网络结构
小组成员	李康 程洁帆 肖扬 张旭 张昕

填表日期 2019 年 7 月 28

成绩 (50%)

1 主要结果概述

通过 EM 算法，我们估计了模型中的参数，包括：N 位基金经理的合群参数 $\{a_i\}$ , K 个省份（或直辖市）的规模参数 $\{b_k\}$ , K 个不均匀分布参数 $\{\omega_k\}$ ，股票的特征参数 $\alpha$ 和 $\beta$ ，以及基金经理私人信号的权重参数 $\lambda_0$ 和 $\lambda_1$ 。

再通过计算估计了基金经理 i 在 k 省中的关联数 $y_{ik}$ ，并对数值结果进行了阐释分析。最后，我们比较了基金经理持有的、位于不同关联度的省份的股票的收益率，发现不同的关联度确实会影响其收益率。

2 具体结果展示与分析

这里，在剔除缺损数据后，我们样本中的基金经理的数据进行建模求解。首先我们对基金经理的相关统计数据展示与分析。表 1 中给出了基金经理 i 对于在总部在 k 省的股票 j 的投资组合权重 $w_{ijk}$ ，其均值为 3.74%，标准差为 2.68%，以及基金经理 i 在 k 省中的关联数 $y_{ik}$ 。

表 1:

	均值	标准差
--	----	-----

$w_{ijk}(\%)$	3.74	2.68
$\hat{y}_{ik}$	2.61	1.27

表 2 中展示了模型估计值合群程度 $\{a_i\}$ 的统计数据。

表 2:

	均值	标准差
$a_i$	26.6	7.2

这一参数估计可以解释为我们样本中的基金经理拥有来自 8 个省份或直辖市的大约 26 个朋友，然而有一个相当大的标准差（7.2），这表明这一估计有较大的不确定性，又考虑到我们只考虑了 8 个省份，关联结果具有一定的局限性。

表 3 中我们展示了城市  $k$  的相对规模 $\{b_k\}$ 。

表 3:

	均值
北京市	0.125
广东省	0.060
浙江省	0.060
江苏省	0.060
上海市	0.048
山东省	0.055
福建省	0.048
四川省	0.052

可以看到像北京这样的大城市，相对于其他省份（直辖市）具有更大的潜在社会联系，这也是符合直觉的。另一方面，发现有一些省份的估计值相同，考虑到可能是样本数较少，且基金经理持仓的股票数也较少，导致难以得到准确的估计值。

表 4 中我们展示了分散参数 $\{\omega_k\}$ 。

表 4:

	均值
北京市	2.079
广东省	1.875
浙江省	1.952
江苏省	1.952
上海市	2.031
山东省	1.854
福建省	2.031

四川省	1.816
-----	-------

可以看到，各个省份（直辖市）的分散程度均显著大于 1，说明我们之前的假设是成立的。另一方面，分散参数与规模参数不具有显著的关系，也就是说，并不是城市越大越分散，例如，上海的相对规模并不大，但分散程度却仅次于北京。

除了社交网络的参数，我们还得到了投资组合选择参数。表 5 给出了我们的投资组合选择参数的估计，也即股票预期收益 $(\alpha + \beta X_j)$ 一式中的 $\alpha$ 和 $\beta$ ，其中 $X_j$ 是指股票的特征值。

表 5:

	均值
$\beta_{size}(size)$	0.098
$\beta_{btm}(book - to - market)$	0.120
$\beta_{mom}(momentum)$	0.036
$\alpha(constant)$	-0.214

表中的数据表明，平均来说，基金经理对于市值大、市净率高、动量大的股票具有正面的偏好。

表 6 给出了基金经理私人信号的权重参数 $\lambda_0$ 和 $\lambda_1$  ( $\lambda_{ik} = \lambda_0 + \lambda_1 y_{ik}$ )。

表 5:

	均值
$\lambda_0$	0.493
$\lambda_1$	0.052

这里 $\lambda_0 = 0.593$ ， $\lambda_1 = 0.052$ ，再联系表 1 中的 $\hat{y}_{ik} = 2.61$ ，由式 $\lambda_{ik} = \lambda_0 + \lambda_1 y_{ik}$ 可得 $\lambda_{ik}$ 的估计值为 0.629，即，平均来说，每增加一个标准差的 $\xi_{ijk}$ ，投资组合权重 $w_{ijk}$ 会增加 0.629，这也就是私人信息的影响。

最后我们进行检验，是否基金经理持有的关联度更大的省份的股票收益率会优于关联度小的省份。这里，我们对每个基金经理选择了关联较多的 2 个省份和关联较少的 2 个省份，并各自计算了这 4 个省份的股票的近 1 年的加权平均收益率，最后总汇了所有基金经理的数据，结果显示，p 值为 0.0023，即关联较多的 2 个省份的股票收益率显著高于关联较低的 2 个省份。

### 3 与现有文献结果的比较分析

对比文献 Hong and Xu (2015)，总体上有着相似的结果，也都表明了投资者在公司总部所在省份的人脉关系越好，其关于公司价值的私人信号就越准确，预期的收益也就越高。

同时对比了一些关于股票多因子模型的文献，发现在 A 股市场，市值、市净率和动量因子确实都存在一些正向  $\alpha$ ，也即说明基金经理对这些因子的偏好是正常的，我们的结果也具有一定的合理性。

## 4 结论与启示

我们采用了论文中使用的模型，在 Erdos-Renyi 模型的独立同分布关联数的基础上加入了个体影响因素，随后对基金经理的投资数据使用 EM 算法得出了表示潜在的社交信息的参数。在对结果进行观察时，我们发现基金经理在某一给定城市拥有的联系的分布相较于 Erdos-Renyi 模型来说是过度分散的。而这些处在分布尾部的经理人投资方向更具有倾向性。

此外我们检验了投资股票收益与经理人在该城市的联系的关系，证实了关联数较大的省份的股票收益是大于关联数较小的省份的。这也间接地反映了私有信息的精确性会随着投资者在该地区的联系数的增加而增加。

我们的项目结果充分反映了，**基金经理的投资与社交关系网络之间是具有匹配性的，这是在实证上对于投资关系匹配社交网络的一次有效的实验，同时也为我们进一步推测不同资本系之间的潜在社会关联提供了理论依据以及推测的方法。**本次实践最大的意义是从无到有地刻画出潜在的社交关系，机构投资的相关性也可以由此推测，所以潜在的社交网络不失为一个系统风险监控的指标。

参考文献：

Hong, Harrison G. and Xu, Jiangmin, Inferring Latent Social Networks from Stock Holdings (September 11, 2017). Journal of Financial Economics (JFE), Forthcoming. Available at SSRN: <https://ssrn.com/abstract=2634306> or <http://dx.doi.org/10.2139/ssrn.2634306>