

知名度因子及流动性因子对股票收益率的影响

李 任

(南开大学金融学院, 天津 300071)

摘要: 本文以 2021 年 3 月 1 日至 2022 年 11 月 15 日期间由新三板精选层上市北交所的企业股票为例, 研究股票收益率与股票流动性、知名度 (该股票是否在投资者潜在投资组合中) 之间关系, 探究知名度因子、流动性因子能否为股票定价提供依据。在总体面板数据上, 本文发现股票知名度越高、流动性越高, 股票超额收益率越低, 且结论显著; 进一步对北交所开市 (2021 年 11 月 15) 前后两个时期数据分组研究作为稳健性检验, 发现在北交所开市前时期结论与总体数据结论相同且显著; 在北交所开市后时期结论不显著。

关键词: 知名度因子; 股票流动性; 多因子模型

1 文献综述

1.1 股票知名度差异

Merton (1987) 提出了不完全信息下的资本市场均衡模型 (Capital Equilibrium with Incomplete Information)^[1], 构建出在每个投资者仅知道一个可得到的证券资产组合的假设下资本市场均衡的两期模型, 并探讨这种特殊类型的不完全信息对均衡资产价格结构的影响。在模型中, 如果一些投资者只知道市场所有股票的一个适当子集, 并且只持有他们所知道的股票, 那么这些投资者将无法有效地配置股票以使其资产组合充分多样化, 并承担非系统性风险, 因此, 他们会要求一个更高的风险溢价——影子成本 (shadow cost)。这种由于投资者对股票认知导致的对其收益率要求差异, 可以近似度量股票的知名度差异。

Kadlec 和 McConnell (1994)^[2] 在研究不同细分市场上股票价格差异时, 给出了 shadowcost 的一种计算公式: $\lambda_k = \frac{\delta \sigma_k^2 x_k (1 - q_k)}{q_k}$ 。其中, λ_k 表示企业 k 股票的 shadowcost, δ 表示总体风险厌恶系数, σ_k^2 表示企业 k 股票收益的月度数据方差, x_k 表示企业 k 的价值与证券市场总市值的比值, q_k 表示企业 k 投资者数量与市场总投资者数量的比值。

此外, 利用文本挖掘、词频统计等手段也可以构造或者直接得到词频数等指标来度量股票的知名度、热度差异。

1.2 股票流动性

针对股票流动性与股票收益之间是否存在联系, 早有学者研究过, Amihud 和 Mendelson (1986)^[3] 提出了股票收益与股票流动性之间关系的假设: 股票的超额收益 (即传统意义上的风险溢价) 是预期企业股票非流动性的增函数。即流动性较差的资产, 其市场价格比流动性较之更好的资产更低, 投资者对该资产要求更高的风险溢价。Amihud 在研究股票流动性与收益的截面与时间序列效应时, 给出了非流动性比率 ILLIQ 以衡量企业股票的流动性。

此外, 还可以通过股票成交量、换手率衡量股票的流动性。还有更为精细的衡量股票流动性的指标如: 股票买卖价差 (bid-ask spread)、基于信息交易的频率、逐笔交易对总体市场的影响等, 但这些高频交易指标需要大量的微观市场结构数据, 难以获取, 并且在数据基础上还需要构建长期的序列数据。

2 模型与假设检验

根据过往资产定价模型，股票收益率具有明确的因子表达形式，即：

$$E(R_t) = \sum_{i=1}^N f_{it} X_{it} \quad (1)$$

其中， $E(R_t)$ 为任何股票的期望收益率； X_{it} 为对股票收益率产生影响的因子； f_{it} 为对应因子针对股票产生的收益，即对应因子对股票收益率的影响程度，也即因子暴露。

根据表达式(1)，可对股票收益率的影响因子进行探究，可构建股票知名度因子、流动性因子对股票超额收益的影响表达式：

$$Exret_{i,t} = Intercept_{i,t} + b_1 \times Shadowcost_{i,t} + b_2 \times ILLIQ_{i,t} + b_3 \times CapSize_{i,t} + b_4 \times CapAge_{i,t} \quad (2)$$

其中，下标*i*表示对应股票，*t*表示对应时期， $Exret_{i,t}$ ， $Shadowcost_{i,t}$ ， $ILLIQ_{i,t}$ ， $CapSize_{i,t}$ ， $CapAge_{i,t}$ 分别表示股票*i*在第*t*期超额收益率、知名度因子（影子成本）、流动性因子（非流动性比率）、市值大小、上市时长。 b_1 ， b_2 ， b_3 ， b_4 为对应因子对超额收益率的影响。上述指标的含义与计算方法将在标题 3 数据描述与实证方法中说明。

结合标题 3 数据描述与实证方法中对知名度因子、流动性因子的解释，分析可知企业股票的影子成本 $Shadowcost$ 、非流动性比率 $ILLIQ$ 越大，投资者对其要求的收益率也应该越大，可提出假设如下：

假设 H1：股票知名度越高，流动性越高，股票超额收益率越低。

具体即，知名度因子（影子成本 $Shadowcost$ ）、流动性因子（非流动性比率 $ILLIQ$ ）对股票超额收益率 $Exret$ 影响显著且为正，即表达式系数 b_1 ， b_2 均显著且为正。

3 数据描述与实证方法

3.1 数据描述

本文采用 2021-03-01 至 2022-11-15 期间由新三板精选层上市北交所的企业股票数据，数据来源于 choice 数据库、RESSET 数据库，且全部采用日频数据。根据这些数据构造如下变量：

Exret：股票的日度超额收益率，即股票日度收益率减去当日的无风险利率。股票日度收益率、日度无风险利率来源于 RESSET 新三板数据库持有期收益板块。

Shadowcost：股票的知名度因子，采用 Merton 在不完全信息下资本市场均衡模型（Capital Equilibrium with Incomplete Information）中构造的影子成本指标。投资者会因企业在股票市场中的是否为自己所知对股票产生不同等的收益率要求，对于投资者认知到的企业股票，投资者可以利用其充分多样化自身资产组合，因而会要求相对低的收益率；对于不为投资者所知的企业股票，投资者会要求一个相对更高的收益率。这种因知名与否而产生的收益率要求差别，即为影子成本（ $Shadowcost$ ）。其计算方法为股票当日流通市值乘以股票登记在册股东数除以股票所在市场总市值与股票所在市场投资者户数。即 $shadowcost_i =$

$\frac{FirmCap_i \times FirmSahreholder_i}{MarketCap_i \times MarketInvestor_i} \times R_{std,i}$ 。其中， $shadowcost_i$ 为企业股票 *i* 的影子成本； $FirmCap_i$ ， $FirmSahreholder_i$ 分别为对应企业的市值（单位：万元）、登记在册股东数量（单位：人）； $MarketCap_i$ ， $MarketInvestor_i$ 分别为对应企业所在股权市场的总市值（单位：万元）、投资者户数（单位：万人）； $R_{std,i}$ 为对应企业股票收益率的标准差。结合 $shadowcost$ 的定义，可知股票知名度越高，其 $Shadowcost$ 越小。

ILLIQ: 股票流动性因子, 采用 Amihud 构造的股票非流动性指标。非流动性指标的计算方法为: 股票当日收益率的绝对值除以当日成交额 (单位为万元), 为了保证在数量级上非流动性指标与其他变量可比, 调整比值, 将其乘以 10^6 。即 $ILLIQ = \frac{|R|}{VOL} \times 10^6$ 。结合 Amihud 的定义, 股票流动性越差, 其非流动性比例 *ILLIQ* 越大。

CapSize: 股票的规模因子, 采用股票当日市值指标 (单位: 万元) 的对数值。

CapAge: 股票的交易时长因子, 采用股票在新三板市场或北交所市场的挂牌交易月数值。

表 1 给出了新三板精选层上市北交所股票的全部面板数据的描述性统计。其中 A 组是原始数据下各变量的描述性统计, B 组的是对异常值进行缩尾(winsorized) 处理后各变量的描述性统计。在 B 组中将每一个变量的原始数据中高于或低于其均值 2.5 个标准差的数值设置成为均值加上或减去 2.5 个标准差, 以降低某些极值数据 (由观测误差、股票的特殊性等等原因造成) 对结果可能产生的影响。

将 B 组与 A 组相比较, 发现 A 组中 shadowcost、ILLIQ 变化范围分别为 0.002534 至 213.9663 与 0 至 419.7219, 而均值分别为 2.290747、1.072412, 其极值点大小与均值水平相比明显是不合理的; 而 B 组 shadowcost、ILLIQ 变化范围分别为 0.002534 至 21.72331、0 至 9.871811, 相对合理。

表 1: 全体变量面板数据描述性统计

变量	样本数	均值	标准差	最小值	25%分位	50%分位	75%分位	最大值
A 组: 原始数据样本								
exret	18348	0.000479	0.030412	-0.26167	-0.01078	-7.1E-05	0.008635	0.303833
ILLIQ	18348	1.072412	4.399699	0	0.147832	0.440075	1.05347	419.7219
shadowcost	18348	2.290747	9.716282	0.002534	0.143491	0.390993	1.337541	213.9663
capAge	18348	6.811805	3.870662	0	4	7	10	15
capSize	18348	11.85806	0.923639	10.36457	11.26064	11.65173	12.23132	16.02587
B 组: 缩尾 (winsorized) 数据样本								
exret	18348	-0.00032	0.023212	-0.06035	-0.01078	-7.1E-05	0.008635	0.061303
ILLIQ	18348	0.95143	1.52809	0	0.147832	0.440075	1.05347	9.871811
shadowcost	18348	1.633144	3.705438	0.002534	0.143491	0.390993	1.337541	21.72331
capAge	18348	6.803573	3.853675	0	4	7	10	14.55313
capSize	18348	11.80596	0.764433	10.36457	11.26064	11.65173	12.23132	13.70534

考虑北交所开市的时间点 2021 年 11 月 15 日, 将时间段划分为北交所开市前时期 (企业在新三板时期)、北交所开市后时期 (企业在北交所时期), 对缩尾后的面板数据进行描述性统计, 结果如表 2。

表 2：北交所开市前后个变量面板数据描述性统计

变量	样本数	均值	标准差	最小值	25%分位	50%分位	75%分位	最大值
C 组：北交所开市前时间段数据样本								
exret	7612	0.003008	0.024734	-0.05961	-0.00819	0.000831	0.012333	0.067605
ILLIQ	7612	0.736732	0.862047	0	0.160851	0.441779	0.940766	3.735583
shadowcost	7612	2.869046	5.607687	0.002534	0.495483	1.208498	2.638148	32.34851
capAge	7612	8.651155	3.612007	1.331401	6	9	11	15
capSize	7612	11.86291	0.750335	10.56618	11.31337	11.75966	12.31999	13.70731
D 组：北交所开市后时间段数据样本								
exret	10736	-0.00263	0.021909	-0.06028	-0.01246	-0.00184	0.006433	0.056244
ILLIQ	10736	1.08321	1.872687	0	0.138143	0.438737	1.164097	12.47276
shadowcost	10736	0.704166	1.662956	0.009311	0.090785	0.193664	0.427146	8.935624
capAge	10736	5.520492	3.48367	0	3	6	8.25	12
capSize	10736	11.76535	0.771208	10.36457	11.22903	11.60474	12.11904	13.6994

观察 C 组、D 组数据，发现在北交所开市后，企业超额收益率、shadowcost 均值明显降低，而 ILLIQ 数据在 25%、50%分位数均降低，但其均值、75%分位数、最大值均升高。考虑到个别股票的特殊性对其流动性有所影响，即可近似认为北交所开市降低了股票超额收益率、shadowcost，部分降低了股票的非流动性，这与现实逻辑相符，北交所相较于新三板市场具有更高规格、知名度、投资者人数也更多。

3.2 实证方法

Fama-Macbeth 回归方法^[4]是对股票因子进行实证研究的常用办法，这种方法本质为线性回归，但其可充分利用面板数据，可排除残差在截面上的相关性对标准误的影响。我们采用 Fama-Macbeth 回归方法研究缩尾后的面板数据。

Fama-Macbeth 回归步骤为，第一步：在时间序列上，对每个股票用超额收益与各因子值（shadowcost、ILLIQ）进行时间序列回归，得到回归系数（因子暴露）；第二步：在每个时间截面上，用超额收益率对因子暴露进行截面回归，得到回归系数（因子收益）。不断重复两个步骤，得到每一截面的因子收益。

2021 年 3 月 1 日至 2022 年 11 月 15 日，共 18348 个日度数据，所以对每个风险因子（Shadowcost、ILLIQ、CapSize、CapAge）的系数 b_1 ， b_2 ， b_3 ， b_4 ，Fama-Macbeth 回归后将产生 18348 维的时间序列，将每个时间序列的均值作为对应风险因子对股票超额收益率的影响系数，并用其样本 t 值检验其对超额收益率的影响是否显著。

4 实证分析

对缩尾后全体面板数据进行 Fama-Macbeth 回归，回归结果如表 3。

表 3：B 组数据 Fama-macbeth 回归结果

变量	影响系数值	标准差	T 统计量	P 统计量
B 组：缩尾（winsorized）数据样本				
Intercept	0.0012	0.0081	0.1463	0.8837
CapAge	9.088E-5	0.0001	0.7698	0.4414
CapSize	4.44E-6	0.0006	0.0069	0.9945
ILLIQ	0.0018	0.0009	1.967	0.0492
shadowcost	0.0002	0.0001	2.0195	0.0435

在全体缩尾后数据样本中 Shadowcost、ILLIQ 系数均为正，且 P 统计量均小于 0.05，即 Shadowcost、ILLIQ 对收益率影响在 5%水平下显著，满足**假设 H1**。即总体而言，股票知名度越高，流动性越高，股票超额收益率越低。股票知名度、流动性对其定价产生影响。

5 稳健性检验

最后，对缩尾后北交所开市前面板数据、缩尾后北交所开市后面板数据，进行 Fama-Macbeth 回归作为稳健性检验，即对上述分组 C、D 组分别进行 Fama-macbeth 回归，其中 D 组中，由于企业均为 2021 年 11 月 15 日由新三板上市北交所，其上市时间 CapAge 完全相等，对股票收益率影响均相同，可在回归表达式中舍去 CapAge。Fama-Macbeth 回归结果如表 4。

表 4：C、D 组数据 Fama-macbeth 回归结果

变量	影响系数值	标准差	T 统计量	P 统计量
C 组：北交所开市前时间段数据样本				
Intercept	0.0004	0.0083	0.0438	0.9651
CapAge	9.508E-5	0.0001	0.803	0.422
CapSize	9.401E-5	0.0007	0.1424	0.8867
ILLIQ	0.002	0.0009	2.0913	0.0365
shadowcost	0.0002	7.916E-5	2.1338	0.0329
D 组：北交所开市后时间段数据样本				
Intercept	-0.0009	0.0044	-0.2052	0.8374
CapSize	-7.95E-05	0.0004	-0.2082	0.8351
ILLIQ	-0.0013	0.0007	-1.7514	0.0799
shadowcost	-0.0001	0.0002	-0.6064	0.5442

C 组数据回归结果中，Shadowcost、ILLIQ 系数均为正，且 P 统计量均小于 0.05，即 Shadowcost、ILLIQ 对收益率影响在 5%水平下显著，满足**假设 H1**，即在北交所开市前，企业在新三板市场时期，也有股票知名度越高，流动性越高，股票超额收益率越低。而 D 组数据中，shadowcost、ILLIQ 的系数为负，且均不显著，未通过**假设 H1**。

结合之前的描述性统计结果，在北交所开市后，受部分股票的特殊性影响，且考虑到北交所开市初期个人投资者投机心理等因素的影响，ILLIQ 有不合常理的 75%分位数与最大值，这些离群值对回归结果有一定影响，可能导致回归结果有偏。但总体而言，在缩尾后的全部面板数据、北交所开市前的面板数据回归结果均接受**假设 H1**，即企业知名度越高，流动性

越大，超额收益率越小；同时描述性统计结果也表明，企业在上市北交所后，除极端值外，在 25%、50%分位数上，其 shadowcost、ILLIQ 总体是减小的，且企业的超额收益率是减小的。

参考文献：

- [1] Merton, Robert C., 1987, Presidential address: A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42(3), 483–51.
- [2] Kadlec, Gregory B., and John J. McConnell, 1994, The effect of market segmentation and illiquidity on asset prices: Evidence from exchange listings, *Journal of Finance* 49(2), 611–63.
- [3] Amihud Y . Illiquidity and stock returns: Cross-section and time-series effects [J]. *Journal of Financial Markets*, 2002, 5 (1): 31 – 56.
- [4] Fama E F, MacBeth J D . Risk, return, and equilibrium: Empirical tests [J]. *Journal of Political Economy*, 1973, 81 (3): 607 – 636.