

公司特质信息、 R^2 与异常收益

摘要 在 CAPM 回归中, 低 R^2 的公司的价格是否更有效的反应了公司的特质信息在学术界始终存在争论。在 R^2 与公司特征变量的横截面检验基础之上, 我们通过一个零成本投资组合考察了两个竞争性的理论。结果发现, 1) R^2 与公司信息环境有显著关系, 而且对于市场信息反应不够有效的公司, R^2 会更低; 2) 低 R^2 股票组合短期的异常收益有更明显的时间序列可预测性和横截面风险, 这说明在我国股市, 低 R^2 代表了资产价格中更大的噪音成分。

关键词 R^2 异常收益 波动 方差比检验

1. 引言

Roll (1988) 首先对 R^2 进行考察, 他发现不同公司的 R^2 在横截面上存在显著差异, 并且, 这些差异不是由公开信息 (public information) 的披露造成的。Roll 认为, 这种横截面上的 R^2 差异是股票价格对私有信息 (private information) 的反应。

在此以后, 学者对于究竟更低的 R^2 意味着股票价格是否更有效的反应了投资者私有的公司特质信息进行了研究。West (1988), Leary and Porter (1981), Shiller (1981) 以及 Campbell, Lettau, Malkiel, and Xu (2001) 都认为如果股票价格更有效 (快) 的反应了私有信息, 那么该股票的 R^2 会更高, 这是因为“在折现率不变的情况下, 未来股利的变化越早进入投资者预期, 这种变化引起的变化就越小 (discount more heavily)。”(Shiller, 1981) 从而会降低该股票的波动。

但另一方面, Morck, Yeung and Yu (2000) 的研究发现, 发达国家的股市往往具有更低的 R^2 , 越成熟的股市或政治法律制度越完善的国家, 股价里包含的上市公司的特质成分越多。他们进一步认为, R^2 越低的公司, 它们的股价对于公司特质信息的反应也就越有效 (information efficiency), 而 R^2 也可以作为一个股价是否信息有效的衡量标准。在接下来的研究里, Durnev, Morck, Yeung, and Zarowin (2003, DMYZ) 发现 R^2 越低的公司, 当前收益同公司未来盈余的变化更相关; Durnev, Morck, and Yeung (2004, DMY) 则发现, R^2 越低的行业, 投资也更为有效。这些证据都直接或间接的支持了 R^2 越低的公司的股价更有效地反映了私有信息的观点。

但最近 Kelly (2005) 的研究发现, 具有更低 R^2 的公司实际上是信息环境更为恶劣的公司。机构投资者很好持有这些公司, 分析师也很少关注这类公司, 同时, 这类公司的规模要更小, 交易成本要更高。这些证据间接表明, R^2 更低的公司, 不应该是反映信息更为有效的公司。

在本文中, 我们在 Kelly (2005) 的基础上, 进一步发现, R^2 更低的公司实际上是那些对于市场信息反应不够充分有效的公司。而且低 R^2 股票组合短期的异常收益有更明显的时间序列可预测性和横截面风险。

我们的贡献在于: 1) 给出证据表明横截面上 R^2 的差异不仅仅和股价反应公司特质信息的有效性有关, 同时还和股价反应市场范围的信息的有效性有关; 2) 给出了直接的证据表明, R^2 更低的公司并不是反映信息更有效的公司, 相反, 它们对于信息的反应更缺乏效率。同时, 它们组合的横截面风险更大, 这在一定程度上也可能是阻止套利活动的原因。

全文组织安排如下: 第 2 部分提出研究假说; 第 3 部分说明数据来源、变量定义和研究研究方法; 第 4 部分进行检验并对结果分析; 第 5 部分给出简要结论。

2. 研究假说

2.1 R^2 与信息环境因素

Roll (1988) 指出, 那些无法被 CAPM 解释的特质收益可能是投资者基于私人信息的交易所致。根据 Morck, Yeung and Tu (2000) 的研究, 发达国家的股市往往具有更低的 R^2 , 越成熟的股市或政治法律制度越完善国家, 上市公司的特质成分越多, 如表 1 所示。美国有最低的 R^2 为 0.021, 而中国的 R^2 为 0.453 仅次于波兰。可见, 在 Roll 的观点之外, R^2 可能还体现了市场上私人信息融入价格的速度, 而这种速度则取决于证券市场的信息环境因素。因此具有更多私人信息或者信息环境更发达的上市公司, 具有更低的 R^2 。最近如 DMYZ (2003) 和 DMY (2004) 则进一步说明低 R^2 的公司确实具有融合了更多的公司特质信息, 他们发现这些公司的当前收益在预测其未来盈利增长方面更有效, 而且这一类公司的投资活动也倾向于更有效率。

表 1 不同国家股市 R^2 的比较

国家	平均 R^2	国家	平均 R^2	国家	平均 R^2	国家	平均 R^2	国家	平均 R^2
美国	0.021	澳大利亚	0.064	韩国	0.172	日本	0.234	马来西亚	0.429
加拿大	0.062	法国	0.075	印度	0.189	泰国	0.271	中国	0.453
英国	0.062	香港	0.150	新加坡	0.191	台湾省	0.412	波兰	0.569

说明: 根据 Morck, Yeung and Tu (2000) 的表-2 整理。数据来源为 DataStream, 以 1995 年为标准, 平均 R^2 的计算是根据 CAPM 模型对全部上市公司半月度的数据回归 R^2 的平均值。

但同时 Kelly (2005) 的研究发现 R^2 更低的公司所处的信息环境更差。我们认为, 以往基于 CAPM 的 R^2 的研究实际隐含了以下假设: 公司股票价格对于市场信息进行了有效的反应。而实际上, 公司对于市场信息的反应可能是不够有效的。如果公司对于以前的市场信息融合速度较慢, 在当期的收益中, 除了要反映当期的市场信息, 公司特质信息以外, 还要对以往的市场信息进行反应, 从而增加了股票收益的波动, 降低了当期市场信息 (收益) 对于公司当期股票收益的解释能力, 从而降低了 R^2 。由此, 我们会发现对于市场信息的反应的有效程度, 会影响到公司的 R^2 , 由此我们得到我们的第一个假说:

假说 1: 在其他条件相同的情况下, 股票的 R^2 与市场信息融合速度为正相关。

2.2 R^2 与异常收益的时间序列可预测性及横截面风险

我们进一步设计了一个直接的检验来看横截面 R^2 的差异究竟意味着什么。如果:

1. 横截面 R^2 越低（高）的公司，意味着它们包含了更多（少）的公司的特质信息，则它们的股票价格将更为有效（information efficient）（无效，information inefficient），那么它们的股票收益在时间序列上的可预测性会更低（高）。
2. 横截面 R^2 越低的公司，意味着它们的股票价格包含了更多的噪音；那么因为股票价格中的噪音更大，从而提高了套利成本，将使得这类公司的股票价格可能更难充分的反应了公司的信息（information inefficient），而导致它们的股票收益在时间序列上有更高的可预测性，而在横截面上有更高的风险。
3. 横截面 R^2 越低的公司，意味着它们的股票价格对于市场信息反应不够充分，那么这类公司的股票收益也将时间序列上有更强的可预测性。

由于对于横截面 R^2 差异的可能解释中的 1 和 2, 3 对于股票收益在时间序列上的可预测性的推论结果不同，我们可以通过检验不同 R^2 组合收益在时间序列上的可预测性来推断哪种解释更接近真实情况。于是，我们有了如下假说：

假说 2：低 R^2 的股票组合都具有更大程度的时间序列可预测性以及横截面差异。

3. 数据的来源及研究方法说明

本文研究的样本为深沪两地上市的A股。全部数据取自中国股票市场交易数据库系统(CSMAR)的交易数据库与封闭式基金数据库，研究区间为1999年1月1日到2005年12月31日。我们选择该段样本期间的原因在于后文的检验中利用到了封闭式基金投资组合的数据，而我国正式意义上的封闭式基金在1998年开始发行。

股票的价格和收益率分别采用了现金红利再投资的可比价格以及调整后收益率。研究中剔除了退市或ST/PT公司；有交易缺失的股票，在缺失期间时不纳入组合，但在其它期间则按正常计算；为了与DMYZ (2003), Kelly (2005) 等人的研究一致，我们采用了CSAMAR数据库中的分类A，共六大类，即金融业(0001)、公用事业(0002)、房地产(0003)、综合(0004)、工业(0005)和商业(0006)，并去掉了少量的代码为0001的金融类样本；为了避免IPO效应，令样本公司必须在1998.7.1之前上市。最后确定了616家上市公司进入研究。

3.1 R^2 的计算

我们以周频交易数据为基础来计算 R^2 ¹，为了避免周末效应，跟随DMYZ (2003) 的建议，本文以周三的数据为基础计算每周收益。每年度计算一次 R^2 并排序，用来构造下一年度的投资组合。具体而言， R^2 的计算是通过如下回归：

¹ Roll (1988) 的研究中利用月度数据计算了 R^2 ，但本文的研究方法是分年度计算 R^2 ，因此使用月频数据将无法保证 R^2 的精确性（每年只有 12 个数据）。DMYZ (2003) 等学者认为应该使用更高频率的数据，如周数据或者日数据。但由于日频率的交易数据会在一定程度上受市场微观结构效应的影响，如非同步交易等。因此本文选择了周频率的数据。但事实上，我们对以周和日数据的结果进行比较，基本结果是一致的。

$$r_{i,t} = \alpha_{i,t} + \beta_{i,t} r_{Mkt \neq i,t} + \gamma_{i,t} r_{Indu \neq i,t} + \varepsilon_{i,t} \quad (1)$$

这里， $r_{Mkt \neq i,t}$ 表示排除了股票*i* 以后的基于流通市值权重的市场收益率， $r_{Indu \neq i,t}$ 表示排除了股票*i* 以后的基于流通市值权重的行业收益率。这里，排除股票本身的目的在于防止收益之间的虚假相关性 (spurious correlations; DMYZ (2003)), 尤其在行业所包含的股票样本较少时。我们在研究中采用了经过自由度调整的 R^2 ，即： $R^2 = 1 - \frac{T-1}{T-k-1} \cdot \frac{s^2(e)}{s^2(r)}$ 。

由于本文报告结构均基于周数据，为了确定在中国市场上该交易频率的选择是否稳健，我们也将所有基于周数据的研究结果同样用日交易数据考察，发现基本结论是一致的。我们检验发现，不同数据计算下 R^2 排序的重合程度随着 R^2 由小到大呈U型，在 R^2 最大和最小的样本中，两种方法的排序结果相当一致，但在那些中等 R^2 的组合中，二者不太一致，这导致两种方法计算的结果只有30%~40%的重合度。但有趣的是，即便中等 R^2 程度的组合样本不一致，在基于日数据的考察中我们发现本文的假说同样成立²。

3.2 相关变量的定义

首先对假说1涉及的变量进行说明。我们以本文以市场信息融合速度说明股票融合市场信息的时间、以投资者关注度和知情者持有度来代表股票的私人信息含量，以交易成本来代表股票所面临的特质市场环境。

市场信息融合速度的代理变量为延迟度 (*Delay1* 和 *Delay2*)，这是由Hou and Moskowitz (2005) 提出的两个类似测度，他们构造该变量的目的在于考察公司收益在多长时间以内可以融合市场范围的信息。延迟越小，表示公司的股票价格能够越快的反应最新的市场信息。具体而言，该变量的计算首先对个股收益按下式回归：

$$r_{i,t} = \alpha_i + \beta_i r_{Mkt,t} + \sum_{n=1}^4 \delta^{(-n)} r_{Mkt,t-n} + \varepsilon_{i,t} \quad (2)$$

这里， $r_{Mkt,t}$ 表示基于流通市值加权的市场收益率。与 Hou and Moskowitz (2005) 的研究方法一致，我们并未剔除股票本身。通过对等式(2)的回归获得 R^2 ，然后限制上式的各 $\delta^{(-n)}$ 同时为 0，并对该约束模型回归，得到 $R^2_{\delta^{(-n)}=0, \forall n \in [1,4]}$ ，二者均经过自由度调整。此时，令：

$$Delay1 = 1 - \frac{R^2_{\delta^{(-n)}=0, \forall n \in [1,4]}}{R^2}, \quad Delay2 = 1 - \frac{\sum_{n=1}^4 n \delta^{(-n)}}{\beta + \sum_{n=1}^4 \delta^{(-n)}}。$$

需要指出，由于 *Delay* 的构造与我们的设想是反向构造的，因此在假说 1 中市场信息融合速度与 R^2 表现为正相关对应到 *Delay* 变量时应该为负相关。

² 虽然本文基于周交易数据的结果较好的支持研究假说，但令人惊讶的是，基于日交易数据的结果似乎更符合本文的研究假说。不过，考虑到难以分离出日交易数据中的微观结构效应，因此我们依然报告了周数据的结果。有兴趣的读者可以索取基于日交易数据的计算结果。

我们以公司的流通市值 (*Size*) 作为投资者关注度的测度。一般而言, 市值越大的公司, 应该在更大程度上为投资者所关注, t 年度的市值以 $t-1$ 年度 12 月份流通市值为准。

对于理性投资者持有度而言, 难以找到比较好的代理变量, 但是得益于专业的分析人士、熟练的交易技巧以及更大的资金数量, 我们认为基金交易者应该具有更大的成本优势, 而且也能在更大程度上掌握市场上有用的信息或者排除谣言的干扰。例如 Sias and Straks (1997), Chakravarty (2001) 以及 Sias (2004) 都认为机构投资者在更大程度上表现出理性交易的特征。

因此我们以某一只股票是否为机构投资者所持有来代理理性投资者持有度, 具体而言, 我们构建 $Insti$ 为一哑变量, 如果该公司的股票为基金持有, 则取值为 1, 否则为 0。检验中对各家基金在当年度投资组合报告所涉及的股票取并集。

股票的流动性测度 ($Illiq$)³, 该变量越大, 表示股票的流动性越差, 也反应了股票的交易成本越高。该变量根据 Amihun (2002) 计算, 具体而言, 在第 t 年, 股票 i 的非流动性为该月内的日交易数据计算而得:

$$Illiq_t^i = \frac{1}{Days_t^i} \sum_{d=1}^{Days_t^i} \frac{|R_{td}^i|}{V_t^i} \quad (3)$$

这里, R_{td}^i 和 V_t^i 分别指股票 i 在第 t 年的第 d 个交易日的收益率和交易量 (以十万元为单位), $Days_t^i$ 指股票 i 在第 t 年的有效交易天数。

其次, 我们根据 Lesmond, Ogden and Trzcinka (1999) 对交易成本测度的建议, 选择了股票年度的日收益为零的比例 ($ZeroRet$), 用来简化地代理交易成本。因为股票交易成本越大, 零收益比例也就越高, 这是因为交易费用较高时, 较小的期望收益将不会弥补该成本, 因此市场不发生交易, 从而收益为 0。

3.3 零成本投资组合的异常收益计算

根据 Conrad and Kaul (1998) 的研究方法, 在零成本交易策略的情况下, 将异常收益同时分解成时间性列可预测性和横截面风险两个部分以期进行更深入的考察。具体的过程如下:

- 1). 根据本年度的 R^2 , 按照由小到大分成 10 个投资组合。
- 2). 选取一定长度的时间区间, 称为投资组合形成期 (或排序期) p , 分别计算各样本证券在形成期内的收益率 $R_{i,t-p}$ 。其中, t 为任意选定的时刻。
- 3). 根据形成期收益率的高低构造赢家投资组合 W 和输家投资组合 L 。本文采用 Lo 和 Mackinlay (1990) 首创, 随后为 Conrad 和 Kaul (1998) 所延用的研究方法。

³ Hasbrouck (2002) 比较了一些流动性测度 (如 Kyle (1985) λ 测度) 与 Amihud 测度后指出, “就所考虑的代理变量而言, 似乎 Amihud 的非流动性是最好的”, 此外 Amihud 测度也非常接近微观证券研究中常用的 Amivest 流动性测度。

首先计算样本中所有证券在形成期的收益率，然后计算它们全体的等权重组合收益率 $R_{m,t-p}$ 。然后将市场所有股票分为两组：若个股收益率大于市场平均收益率则为赢家投资组合，若个股收益率小于市场平均收益率则为输家投资组合。

然后，以个股收益率相对于组合平均收益率的离散大小来调整其投资比重权数 w_{ip} 。于投资比重权数的加总为 0，即 $\sum_{i=1}^N w_{ip} = 0$ ，所以该投资策略是一种零成本的套利投资组合⁴。

4). 紧接形成期或间隔一段时间（即跳跃期）⁵，再选定另一段区间，作为投资组合的持有期（检验期） q 。由此得到交易策略 $[t-p, t+q]$ 。然后分别计算持有期内赢家投资组合的收益率 W_{t+q} 和输家投资组合的收益率 L_{t+q} ，以各持有期输家组合的收益率减去赢家组合的收益率表示该策略的收益率，即得到该策略的异常收益 $\pi(p, q)$ 。

对总异常收益而言，可分解为三部分，且与形成期 p 与持有期 q 有关：

$$E[\pi(p, q)] = -C(p, q) + O(p, q) + (q/p)\sigma^2(\mu) \quad (4)$$

关于该分解的数学推导及说明参见附录。其中， $C(p, q)$ 和 $O(p, q)$ 时间序列可预测性有关，而 $\sigma^2(\mu)$ 则仅仅取决于组合内部的横截面方差而与收益的可预测性无关。对惯性策略来说，除非 $O(p, q)$ 减去 $C(p, q)$ 的力量占优于 $(q/p)\sigma^2(\mu)$ ，否则策略将无法获利，也就是说，投资组合内个股收益率时间序列的可预测性对异常收益具有相当大的重要性。

可以看到，这种分解方法相对于以前的研究⁶来说具有如下几点好处：投资比重权数的正负号可以表现投资人对股价走势的不同信念；以个股收益率相对于市场平均收益率的离散来调整其投资比重权数，符合过去相关研究发现的“较大的变动往往会跟随一个较大的变动，较小的变动往往会跟随一个较小的变动”，即变异聚集(volatility clustering)现象；很重要的一点，就是可以将惯性策略收益加以分解，以探讨其组成内容与原因，具有重要的经济含义。

在异常收益的计算中，为了和 Conrad and Kaul 的研究方法保持一致，本文在收益的选择上使用了购买并持有策略的异常收益 (buy-and-hold abnormal return, BHAR)，而未采用累积异常收益(cumulative abnormal return, CAR)。根据 Conrad and Kaul (1993)，

⁴ 在零成本投资组合的情况下，因为投资策略是同时买卖赢家投资组合与输家投资组合，并无真实投资，故计算所得的结果为收益（回报）而非收益率 (Lo 和 MacKinlay, 1990; Lehman, 1990; Conrad 和 Kaul, 1993)，考虑到本文所关心的仅仅是收益是否显著以及各收益来源的贡献程度，因此未对结果按照持有期进行调整。

⁵ 选取一段跳跃期的原因是为了避免市场的微观结构效应，如 Lo 和 Mackingly (1990), Jegadeesh 和 Titman (1993, 1995)。目的在于减弱买卖价差 (bid-ask spread)，价格压力 (price pressure) 和滞后反应效应 (lagged reaction effects) 等因素的影响。

⁶ 理论上对于如何定义输家投资组合与赢家投资组合存在有两种方法：第一种是 De Bondt 和 Thaler (1985, 1987) 所采用的方法，这是相关文献中最常见的方法。它对各证券在形成期的收益率从高到低排序，并按照排序的结果进行分组。通常平分为 10 组（或 5 组）。收益率最高的一组就作为赢家组合，最低的一组就作为输家组合。第二种是本文延用的 Conrad 和 Kaul (1998) 方法。相对于第一种赢家输家组合的构造来说，本方法中引入了权重，这使得持有期收益成为形成期收益的线性函数，所以可以对策略收益进行分解，来进一步分析其构成。

Barber and Lyon (1997), BHAR 的优点于“更真实的反映投资者的经验”，且能够反映CAR所忽略的日期混合效应⁷。

最后，由于我国股市的有效样本期间总体长度较短，可能会导致研究构建的组合数目过少，根据Conrad and Kaul (1998) 和Jegadeesh and Titman (2002)，为避免小样本偏差 (small sample bias)，本文使用每次仅移动1个单位时间的重叠期间 (overlapping holding periods) 法⁸，对于因重叠期间法所造成的样本序列自相关与异方差现象，则按照Newey-West (1987) 的方法进行修正。在所有的形成期和持有期之间均有1周的跳跃期。

4. 实证检验结果及分析

这一部分，我们对R²组合的基本统计特征进行描述并经统计假设1和假设2。

4.1 描述性统计

基于R²组合下各变量的描述性统计如表2所示。

表 2 描述性统计

Portflio	R2	Delay1	Delay2	Ln(size)	Institute	Illiq	ZeroRet
R ² (Low)	0.154	0.124	1.135	13.5087	0.231	0.787	0.031
2	0.335	0.020	1.115	13.5299	0.201	0.472	0.032
3	0.455	0.012	1.043	13.5278	0.206	0.495	0.034
4	0.563	0.011	1.030	13.5518	0.193	0.491	0.035
R ² (High)	0.710	0.010	1.003	13.5642	0.181	0.485	0.034

说明：本表以年度R²为序构造组合，并求得各组合不同变量的均值，然后按照时间序列平均。其中，R²经过自由度调整，通过回归 $r_{i,t} = \alpha_{i,t} + \beta_{i,t}r_{Mkt \neq i,t} + \gamma_{i,t}r_{Indu \neq i,t} + \varepsilon_{i,t}$ ，然后令 $R^2 = 1 - \frac{T-1}{T-k-1} \cdot \frac{s^2(e)}{s^2(r)}$ 而得。

这里， $r_{Mkt \neq i,t}$ 表示排除了股票*i*以后的基于流通市值权重的市场收益率， $r_{Indu \neq i,t}$ 表示排除了股票*i*以后的基于流通市值权重的行业收益率，我们选用CSMAR数据库的行业代码A(0001-0005)为分类标准，但排除了少量的0001(金融)类公司。其中，Delay变量计算按Hou and Moskowitz (2005)，即首先对个股收益按下式 $r_{i,t} = \alpha_i + \beta_i r_{Mkt,t} + \sum_{n=1}^4 \delta^{(-n)} r_{Mkt,t-n} + \varepsilon_{i,t}$ 回归，获得R²，然后限制上式的各 $\delta^{(-n)}$ 同时为0，

并对该约束模型回归，得到新R²: $R^2_{\delta^{(-n)}=0, \forall n \in [1,4]}$ ，令 $Delay1 = 1 - \frac{R^2_{\delta^{(-n)}=0, \forall n \in [1,4]}}{R^2}$ ， $Delay2 = 1 - \frac{\sum_{n=1}^4 n \delta^{(-n)}}{\beta + \sum_{n=1}^4 \delta^{(-n)}}$ ；

Ln(size)为公司的流通市值，*t*年度的市值以*t*-1年度12月份取值为标准；Institute为一哑变量，如果该公司的股票为共同基金(含封闭式和开放式)持有，则取值为1，否则为0；Illiq为根据Amihun (2002)所计算，

用来代理公司股票的流动性情况，计算公式为 $Illiq_{i,t} = \frac{1}{Days_{i,t}} \sum_{d=1}^{Days_{i,t}} \frac{|r_{i,t,d}|}{V_{i,t,d}}$ ；Zero Ret则根据Lesmond,

Ogden and Trzcinka (1999)的建议所计算的股票年度的日收益为零的比例，用来简化地代理交易成本。

⁷ 有的学者认为市场异常收益的衡量应该采用CAR，如Fama (1998), Mitchell and Stafford (2000)曾指出采用CAR的好处在于：1) CAR不易受“坏模型”的影响；2) CAR使用月份的日历时间组合，能自动将横截面关系考虑进组合方差；3) CAR所估计的分布接近正态，可以使用传统的统计检验。但总体上说，CAR容易产生正的统计偏差，而BHAR则与之相反。一般来说，为了使结论更加稳妥，倾向于在反应不足的情况下使用CAR，而在反应过度的情况下使用BHAR。从有关研究来看，我国股市往往是呈现过度反应的，如王永宏和赵学军 (2001)，孔东民 (2004)以及本文的结果，所以选用BHAR得到的结果应该更稳妥。

⁸ Fuller (1976)首先证明了小样本情况下，样本均值的错误估计将会使估计出的协方差偏小。Conrad and Kaul (1998)则进一步证明了偏小的协方差会导致策略收益在分解时出现偏差，所以他们选用重叠期间的方法来尽量避免该偏差。

可以看到,随着 R^2 的不断增加, $Delay1$, $Delay2$ 都呈现除了单调递减的趋势; $Size$ 只表现出了轻微的增加; $Institute$ 则表现为递减, 这说明机构投资者的参与使得 R^2 降低; 对于交易成本变量而言, $ZeroRet$ 随着 R^2 的提高基本上表现为增加; 流动性测度 $Illiq$ 说明随着 R^2 的提高流动性呈现逐渐增加的状态, 即 R^2 增加, 流动性增大。

4.2 R^2 与信息环境

接下来对假说 1 进行检验。结果如表 3 所示。对于总样本, 我们根据 Hausman 检验值, 采取了固定效应进行回归, sub-sample1 表示 1999-2001 年度的回归结果; sub-sample2 表示 2003-2004 年度的回归结果。受限于区间长度, 两个子样本的检验为 OLS 的回归结果。

表 3 各样本期间的回归结果

Variables	total sample		sub-sample1		sub-sample2		
	Model(1)	Model(2)					
Delay1	-	-0.009		-0.244		-0.408	
		-2.97***		-6.27***		-5.28***	
Delay2	-		-0.001	0.001		-0.023	
			-0.240	1.290		-2.15**	
Ln(size)	+	0.010	0.010	0.010	0.011	0.005	0.003
		2.26**	2.26**	1.370	1.490	0.620	0.300
Inst	+	-0.021	-0.021	-0.025	-0.025	-0.018	-0.023
		-2.59**	-2.54**	-1.77*	-1.70*	-1.190	-1.480
Illiq	-	-0.001	-0.001	-0.018	-0.007	-0.003	-0.018
		-0.700	-0.740	-5.28***	-2.31**	-0.510	-2.76***
ZeroRet	+	0.682	0.671	0.371	-0.030	0.909	0.868
		3.81***	3.74***	0.900	-0.070	2.47**	2.31**
Constant		0.296	0.296	0.288	0.270	0.334	0.399
		4.77***	4.76***	2.86***	2.63***	2.97***	3.49***
诊断性统计量							
Adj-R ² (%)		0.84	0.53	3.98	1.04	5.2	2.41
Sig		6.79***	5.02***	10.17***	2.57**	7.74***	3.01**
Hausman		28.93***	28.58***	——	——	——	——

说明: 本表中 total sample 表示 1999-2005 年度的回归结果, 根据 hausman 检验值选取了固定效应的 Panel Data 回归; sub-sample1 表示 1999-2001 年度的回归结果; sub-sample2 表示 2003-2004 年度的回归结果。受限于区间长度, 两个子样本的检验为 OLS 的回归结果。

其中, 各变量定义同表 1; 检验结果的第一行表示回归系数, 第二行表示 t-value; ***, **, * 分别表示在 0.01, 0.05, 0.10 的显著性水平上拒绝系数为零的假设; Adj-R² 表示调整后的 R²; Sig 表示模型的整体显著性检验的 F-value。

从表 3 的结果可以看到, 大部分变量的估计结果都与假说 1 一致, 尤其第一个延迟测度在较高水平显著性水平上拒绝零假设; 公司规模变量在总样本期间下显著大于零, 但在两个子样本期间则不显著; 流动性测度来看, 估计结果的符号也与预期一致, 尽管总样本期间不是非常显著; 基于零收益率的交易成本测度也与得到了与假说 1 一致的结论; 最后, 基金参与持有变量的回归系数为负, 与预期不一致, 这说明机构投资者的加入导致了资产价格的反

应了更多的特质信息，从而导致 R^2 的降低。

4.3 R^2 与异常收益的分解

这一部分通过考察基于 R^2 排序组合的异常收益及其分解——时间序列和横截面。为了避免“后见之明”，我们以每年度的 R^2 排序，作为今后年度构造投资组合的基础。比如基于 2000 年的 R^2 排序，以此在 2001 年构建组合。因为总样本在 600 以上，因此我们按照分位数构造的 5 个组合，其各自的样本量大约均接近 120 多家。为了进一步控制规模效应，我们对全部样本按照规模分为三组（分别为 30%，40%，30%）然后取最大和最小的两组与 R^2 组合交叉构建新的组合，然后考察各组合在以后期间的异常收益情况。

具体而言，在规模分组的基础上，我们考察了形成期在 2, 4, 12 和 20 周的组合未来收益状况，为了能区分短期和长期，令每一个形成期所对应的持有期有更大的变化，分别取为：2, 4, 8, 16, 24, 48, 72, 96 和 120 周共九种情况。然后对这 4 种形成期下的 36 种策略又进一步区分成为 360 个基于规模和 R^2 排序的结果。我们发现不同持有期下的结论是一致的，表 4 至表 7 报告了我们的检验结果。

可以看到，对于形成期为 2, 4 的情况下，低 R^2 组合呈现了更大程度的惯性特征；对于形成期为 12, 20 的情况下，低 R^2 组合在短期呈现了更大程度的反转但在长期则同样呈现了更大程度的惯性特征。而无论是否按照规模分组，对结果影响并不大。

表 4 形成期为 2 周下 R^2 组合的异常收益及分解的均值 (单位: %)

持有期	Small Size (the bottom 30%)						Big Size (the top 30%)					
	1(Low R^2)			5(High R^2)			1(Low R^2)			5(High R^2)		
	Profit	Time	Cross	Profit	Time	Cross	Profit	Time	Cross	Profit	Time	Cross
2	-0.008	-0.382	0.373	-0.009	-0.256	0.247	0.000	-0.436	0.436	-0.010	-0.263	0.253
	0.238	0.000	0.000	0.188	0.000	0.000	0.498	0.000	0.000	0.287	0.000	0.000
4	-0.015	-0.392	0.377	-0.022	-0.272	0.250	-0.016	-0.457	0.441	-0.026	-0.282	0.255
	0.275	0.000	0.000	0.092	0.000	0.000	0.416	0.000	0.000	0.024	0.000	0.000
8	-0.022	-0.400	0.378	-0.018	-0.269	0.251	-0.022	-0.468	0.446	-0.032	-0.287	0.256
	0.312	0.000	0.000	0.262	0.000	0.000	0.372	0.000	0.000	0.017	0.000	0.000
16	-0.013	-0.392	0.379	-0.015	-0.255	0.240	-0.017	-0.436	0.419	-0.024	-0.282	0.257
	0.338	0.000	0.000	0.275	0.000	0.000	0.399	0.000	0.000	0.082	0.000	0.000
24	0.009	-0.361	0.370	0.000	-0.232	0.232	0.016	-0.393	0.409	-0.003	-0.243	0.240
	0.457	0.000	0.000	0.208	0.000	0.000	0.300	0.000	0.000	0.300	0.000	0.000
48	0.013	-0.309	0.323	0.008	-0.185	0.193	0.011	-0.304	0.316	0.007	-0.217	0.224
	0.396	0.000	0.000	0.509	0.000	0.000	0.666	0.000	0.000	0.550	0.000	0.000
72	0.002	-0.327	0.329	0.002	-0.179	0.181	-0.002	-0.298	0.296	0.000	-0.225	0.225
	0.441	0.001	0.001	0.368	0.009	0.000	0.394	0.000	0.000	0.518	0.001	0.000
96	0.012	-0.345	0.357	0.003	-0.157	0.160	0.026	-0.285	0.311	-0.027	-0.234	0.207
	0.584	0.001	0.003	0.543	0.000	0.000	0.519	0.000	0.000	0.641	0.010	0.000
120	0.042	-0.353	0.395	0.000	-0.162	0.162	0.025	-0.272	0.297	0.013	-0.179	0.192
	0.442	0.011	0.025	0.461	0.000	0.000	0.566	0.001	0.000	0.308	0.000	0.000

说明：在样本检验期间，我们以上年度的公司规模和 R^2 排序为基准来构建不同的零成本投资组合并考察其

在未来的异常收益。因为按照 R^2 (分 5 组)和规模(按照 30%, 40%, 30%分) 分组之后, 每个投资组合大约为 60 个样本, 不适合用 De Bondt and Thaler (1985) 的办法处理, 我们参照 Lo 和 Mackinlay (1990) 以及 Conrad 和 Kaul (1998) 的研究方法对各组合的全部样本赋予不同的权重来构建零成本投资策略, 该方法还能够将总收益按照赢家-输家或者时间序列预测性和横截面方差两种不同方式分解, 限于篇幅, 我们在表中报告了将总收益按照时间序列和横截面分解的结果, 即 $Profit = Time + Cross$ 。

本表中, 投资策略的形成期 (format period, 根据各样本的历史表现构建其在零成本组合中不同权重的阶段) 为 2 周, 即半个月的时间, 持有期 (hold period, 在构建了零成本组合之后, 考察异常收益的阶段) 分别考察了 2, 4, 8, 16, 24, 48, 72, 96, 120 周的情况, 为了避免市场微观结构效应的影响, 在形成期和持有期之间均有一个单位时间的跳跃期 (skip period), 为了避免小样本偏差 (small sample bias) 并增加检验效力 (power), 我们采用样本期间交叠法计算, 考虑到各年度的交易期间一般在 48 个周左右, 因此持有期在 120 以上的情况下, 会由于样本数太少不具有较强的统计意义, 故没有进一步报告更长策略持有/形成期的收益结果。所有异常收益均以购买并持有 (buy-and-hold) 方式加以报告。其中, 局限于组合构建之后的样本期间, 持有期 2, 4, 8, 16, 24 周的结果为 4 个年度而得; 持有期为 48 和 72 周的结果为 3 个年度平均而得; 持有期为 96 周的结果为 2 个年度平均而得; 持有期为 120 周的结果只有一个年度。

各 R^2 组合的第 1 列为各对应策略异常收益均值, 第 2, 3 列表示在异常收益分解之后的时间序列和横截面成分; 各持有期的第 1 行为收益大小, 第 2 行为 p -value 均值。在计算显著性的时候, 我们对 t -value 按照 Newey and West (1987) 的办法修正了自相关和异方差。

表 5 形成期为 4 周下 R^2 组合的异常收益及分解的均值

持 有 期	Small Size (the bottom 30%)						Big Size (the top 30%)					
	1(Low R^2)			5(High R^2)			1(Low R^2)			5(High R^2)		
	<i>Profit</i>	<i>Time</i>	<i>Cross</i>	<i>Profit</i>	<i>Time</i>	<i>Cross</i>	<i>Profit</i>	<i>Time</i>	<i>Cross</i>	<i>Profit</i>	<i>Time</i>	<i>Cross</i>
2	-0.015	-0.707	0.693	-0.020	-0.478	0.458	-0.013	-0.830	0.817	-0.026	-0.505	0.480
	0.194	0.000	0.000	0.135	0.000	0.000	0.399	0.000	0.000	0.023	0.000	0.000
4	-0.031	-0.730	0.700	-0.035	-0.498	0.463	-0.037	-0.860	0.823	-0.048	-0.531	0.482
	0.247	0.000	0.000	0.080	0.000	0.000	0.250	0.000	0.000	0.011	0.000	0.000
8	-0.030	-0.736	0.706	-0.023	-0.490	0.468	-0.034	-0.867	0.832	-0.049	-0.535	0.486
	0.390	0.000	0.000	0.313	0.000	0.000	0.388	0.000	0.000	0.093	0.000	0.000
16	-0.023	-0.735	0.712	-0.024	-0.477	0.454	-0.046	-0.850	0.804	-0.037	-0.520	0.484
	0.348	0.000	0.000	0.177	0.000	0.000	0.224	0.000	0.000	0.013	0.000	0.000
24	0.018	-0.660	0.678	0.014	-0.421	0.435	0.038	-0.704	0.742	0.006	-0.444	0.450
	0.290	0.000	0.000	0.254	0.000	0.000	0.301	0.000	0.000	0.350	0.000	0.000
48	0.032	-0.564	0.595	0.016	-0.340	0.356	0.026	-0.564	0.590	0.022	-0.411	0.433
	0.326	0.000	0.000	0.355	0.000	0.000	0.527	0.000	0.000	0.277	0.000	0.000
72	0.001	-0.603	0.605	0.001	-0.335	0.336	-0.004	-0.564	0.560	0.003	-0.430	0.433
	0.373	0.000	0.000	0.303	0.008	0.000	0.208	0.000	0.000	0.386	0.000	0.000
96	0.012	-0.632	0.644	0.006	-0.297	0.303	0.056	-0.541	0.597	-0.067	-0.484	0.417
	0.520	0.000	0.000	0.492	0.000	0.000	0.414	0.000	0.000	0.534	0.001	0.000
120	0.046	-0.595	0.641	0.005	-0.297	0.303	0.040	-0.485	0.525	0.021	-0.375	0.397
	0.486	0.000	0.000	0.322	0.000	0.000	0.514	0.000	0.000	0.122	0.000	0.000

说明: 本表除投资策略的形成期为 4 周之外, 其余说明请参见表 4。

表 6 形成期为 12 周下 R² 组合的异常收益及分解的均值

持 有 期	Small Size (the bottom 30%)						Big Size (the top 30%)					
	1(Low R ²)			5(High R ²)			1(Low R ²)			5(High R ²)		
	Profit	Time	Cross	Profit	Time	Cross	Profit	Time	Cross	Profit	Time	Cross
2	-0.024	-1.743	1.719	-0.019	-1.231	1.212	-0.021	-2.097	2.077	-0.028	-1.160	1.132
	0.353	0.000	0.000	0.296	0.000	0.000	0.502	0.000	0.000	0.170	0.000	0.000
4	-0.036	-1.773	1.737	-0.028	-1.245	1.217	-0.048	-2.134	2.086	-0.048	-1.181	1.133
	0.405	0.000	0.000	0.310	0.000	0.000	0.303	0.000	0.000	0.147	0.000	0.000
8	-0.019	-1.774	1.755	-0.014	-1.190	1.175	-0.068	-2.186	2.118	-0.036	-1.171	1.135
	0.457	0.000	0.000	0.152	0.000	0.000	0.239	0.000	0.000	0.077	0.000	0.000
16	0.014	-1.753	1.767	0.022	-1.090	1.111	-0.062	-2.139	2.077	0.032	-1.086	1.118
	0.129	0.000	0.000	0.108	0.000	0.000	0.186	0.000	0.000	0.346	0.000	0.000
24	0.102	-1.652	1.753	0.168	-0.829	0.997	0.162	-1.787	1.949	0.128	-0.953	1.081
	0.298	0.000	0.000	0.063	0.000	0.000	0.195	0.000	0.000	0.146	0.000	0.000
48	0.106	-1.371	1.477	0.083	-0.727	0.811	0.177	-1.279	1.456	0.152	-0.865	1.017
	0.219	0.000	0.000	0.253	0.000	0.000	0.159	0.000	0.000	0.042	0.000	0.000
72	-0.005	-1.487	1.482	-0.009	-0.826	0.817	0.056	-1.315	1.371	0.077	-0.975	1.052
	0.259	0.000	0.000	0.096	0.001	0.000	0.009	0.000	0.000	0.248	0.000	0.000
96	0.061	-1.536	1.596	0.097	-0.669	0.766	0.332	-1.204	1.537	-0.225	-1.394	1.169
	0.179	0.000	0.000	0.043	0.000	0.000	0.000	0.000	0.000	0.285	0.000	0.000
120	0.098	-1.336	1.434	0.039	-0.678	0.717	0.216	-1.101	1.317	0.086	-0.996	1.082
	0.177	0.000	0.000	0.411	0.000	0.000	0.218	0.000	0.000	0.177	0.000	0.000

说明：本表除投资策略的形成期为 12 周之外，其余说明请参见表 4。

表 7 形成期为 20 周下 R² 组合的异常收益及分解的均值

持 有 期	Small Size (the bottom 30%)						Big Size (the top 30%)					
	1(Low R ²)			5(High R ²)			1(Low R ²)			5(High R ²)		
	Profit	Time	Cross	Profit	Time	Cross	Profit	Time	Cross	Profit	Time	Cross
2	-0.012	-2.575	2.563	-0.006	-1.887	1.881	-0.025	-3.165	3.140	-0.014	-1.678	1.664
	0.496	0.000	0.000	0.218	0.000	0.000	0.470	0.000	0.000	0.188	0.000	0.000
4	-0.014	-2.587	2.573	-0.007	-1.873	1.866	-0.059	-3.230	3.171	-0.013	-1.670	1.656
	0.444	0.000	0.000	0.207	0.000	0.000	0.359	0.000	0.000	0.178	0.000	0.000
8	0.009	-2.573	2.582	0.001	-1.842	1.844	-0.103	-3.284	3.181	0.026	-1.616	1.641
	0.257	0.000	0.000	0.325	0.000	0.000	0.322	0.000	0.000	0.470	0.000	0.000
16	0.141	-2.404	2.545	0.120	-1.553	1.673	-0.077	-3.266	3.189	0.130	-1.492	1.622
	0.167	0.000	0.000	0.048	0.000	0.000	0.469	0.000	0.000	0.183	0.000	0.000
24	0.230	-2.377	2.607	0.292	-1.115	1.407	0.220	-2.653	2.873	0.247	-1.253	1.500
	0.284	0.000	0.000	0.067	0.002	0.000	0.060	0.000	0.000	0.052	0.000	0.000
48	0.112	-2.149	2.261	0.144	-1.079	1.223	0.365	-1.728	2.093	0.280	-1.210	1.490
	0.346	0.000	0.000	0.162	0.000	0.000	0.081	0.000	0.000	0.011	0.000	0.000
72	-0.039	-2.243	2.204	0.028	-1.128	1.156	0.198	-1.766	1.964	0.020	-1.491	1.511
	0.251	0.004	0.006	0.296	0.023	0.001	0.128	0.000	0.000	0.156	0.006	0.003
96	0.129	-2.294	2.423	0.169	-0.870	1.039	0.532	-2.062	2.594	-0.418	-1.925	1.507
	0.032	0.000	0.000	0.242	0.000	0.000	0.000	0.000	0.000	0.256	0.000	0.000
120	0.292	-2.022	2.314	0.088	-1.017	1.106	0.420	-1.678	2.099	0.223	-1.192	1.415
	0.259	0.000	0.000	0.216	0.000	0.000	0.005	0.000	0.000	0.137	0.000	0.000

说明：本表除投资策略的形成期为 20 周之外，其余说明请参见表 4。

接下来考察假说 2，根据表 4~7 的检验结果可以看到，尽管大部分的异常收益并不显著，但是在将其按照时间序列和横截面特征分解之后，各种情况都体现了显著的异常收益，几乎所有的 p-value 都显著的拒绝零假设。这说明，在中国市场，即使整体的收益不显著，但是其中也隐含了较大的横截面风险和时序可预测性。

通过总收益在两个角度的分解可以知道，低 R^2 的股票组合都具有更大程度的时间序列可预测性以及横截面差异。在前一部分所得到的证据表明 R^2 和公司特质信息的有效性以及股价反应市场范围的信息的有效性有关之后，我们进一步发现， R^2 较低的组合具有更大程度的内在预测性，这说明 R^2 较低并非反映信息更有效，相反，它们对于信息的反应更缺乏效率。同时，这一类组合在横截面上具有更大的风险，这将会在一定程度上进一步阻止套利，从而维持其信息反应的低效率。

5. 结论

一般而言，低 R^2 表明股价更多地反应了公司的特质信息。我们在本文认为 R^2 与公司所面临的信息环境有密切关系，并提出 2 条可假说进行检验。

通过对 R^2 与公司信息环境的检验，我们发现 R^2 确实蕴涵了市场或者公司的信息，结果表明横截面上 R^2 的差异不仅仅和股价反应公司特质信息的有效性有关，同时还和股价反应市场范围的信息的有效性有关，接下来，我们对不同 R^2 组合的异常收益情况加以检验，结果发现 R^2 更低的公司并不是反映信息更有效的公司，相反，它们对于信息的反应更缺乏效率。同时，它们组合的横截面风险更大，这在一定程度上也可能是阻止套利活动的原因。

参考文献

- Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets* 5, 31-56.
- Barber, B. and Lyon, J., 1997. Detecting Long-Horizon Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics. *Journal of Financial Economics* 43, 341-372.
- Conrad, J. and Kaul, G., 1993. Long-Term Market Overreaction or Biases in Computed Returns. *Journal of Finance* 48, 39-63.
- Conard, J. and Kaul, G., 1998. An Anatomy of Trading Strategies. *Review of Financial Studies* 11, 489-519.
- De Bondt, W. and Thaler, R., 1985. Does the Stock Market Overact? *Journal of Finance* 40, 793-808.
- De Bondt, W. and Thaler, R., 1987. Further Evidence on Investor Overreaction and Stock Market Seasonality. *Journal of Finance* 42, 557-581.
- Durnev, Artyom, Randall Morck, and Bernard Yeung. 2004. Does firm-specific information in stock prices guide capital budgeting? *Journal of Finance* 59, 65-105.
- Durnev, Art, Randall Morck, Bernard Yeung, and Paul Zarowin, 2003. Does greater firm-specific

- return variation mean more or less informed stock pricing? *Journal of Accounting Research* 41, 797-836.
- Elton Edwin J. and Gruber Martin J., 1996. Survivorship Bias and Mutual Fund Performance. *Review of Financial Studies* 9, 1097-1120.
- Fama, Eugene F., 1998. Market Efficiency, Long-Term Returns, and Behavioral Finance. *Journal of Financial Economics* 49, 283-306.
- Fuller, W., *Introduction to Statistical Time Series*, New York: John Wiley & Sons, 1976.
- Grossman, Sanford J., Joseph E. Stiglitz, 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70, 393-408.
- Hasbrouck, J., 2002. Inferring trading costs from daily data: US equities from 1962 to 2001. Unpublished working paper, New York University.
- Hou, Kewei, and Tobias J. Moskowitz, 2004. Market frictions, price delay, and the cross-section of expected returns. *The Review of Financial Studies*, forthcoming.
- Jegadeesh, N. and Titman, S., 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance* 48, 65-91.
- Jegadeesh, N., Titman, S., 1995. Overreaction, Delayed Reaction, and Contrarian Profits. *Review of Financial Studies* 8, 973-993.
- Jegadeesh, N. and Titman, S., 2002, Cross-Sectional and Time-Series Determinants of Momentum Returns. *Review of Financial Studies* 15, 143-157.
- Kelly, 2005, Information Efficiency and Firm-Specific Return Variation, working paper, Arizona State University.
- Kyle, A. S., 1985. Continuous auctions and insider trading. *Econometrica* 53, 1315-1335.
- Lehman, B., 1990. Fads, Martingales, and Market Efficiency. *Quarterly Journal of Economics* 105, 1-28.
- Lesmond, David A., Joseph P. Ogden, and Charles A. Trzcinka, 1999. A new estimate of transaction costs. *The Review of Financial Studies* 12, 1113-1141.
- Lo, A. W. and MacKinlay, A. C., 1990, When are Contrarian Profits Due to Overreaction? *Review of Financial Studies* 3, 175-205.
- Morck, Randall, Bernard Yeung, and Wayne Yu, 2000. The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics* 58, 215-260.
- Newey, W., and Kenneth D. West, 1987. A simple positive-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55, 703-708.
- Roll, Richard, 1988. R^2 . *The Journal of Finance* 43, 541-566.
- 孔东民. 2004. 中国股市反转收益的分解和“后持有期”检验. 《中国金融学》, 第 2 卷第 2 期, 109-127.
- 王永宏与赵学军. 2001. 中国股市“惯性策略”和“反转策略”的实证分析. 《经济研究》, 第6期, 56-61.

附录： 收益的分解

个股的权重向量 w_p 为：

$$w_p = \frac{1}{N} \left[R_{t-p} - \frac{1}{N} \mathbf{1}' R_{t-p} \right] = \left[\frac{1}{N} I_N - \frac{1}{N^2} \mathbf{1}' \mathbf{1}' \right] R_{t-p} \quad (\text{A.1})$$

其中， I_N 为 N 阶单位矩阵。因此，反转策略的总收益可以表示为：

$$\pi_t(p, q) = w_p' R_{t+q} = \frac{1}{N} R_{t-p}' R_{t+q} - \frac{1}{N^2} \mathbf{1}' R_{t-p} R_{t+q}' \mathbf{1} \quad (\text{A.2})$$

这里， R_{t+q} 表示投资组合在持有期的收益率向量。为了能够对总收益进行分解，分别从个股角度考虑形成期 p 和持有期 q 的收益，则式 (A1.2) 可表示为：

$$\pi_t(p, q) = \sum_{i=1}^N \left[\frac{(R_{i,t-p} - R_{m,t-p})}{N} R_{i,t+q} \right] = \frac{1}{N} \sum_{i=1}^N [R_{i,t-p} R_{i,t+q}] - R_{m,t-p} R_{m,t+q} \quad (\text{A.3})$$

对上式两端取期望：

$$\begin{aligned} E[\pi_t(p, q)] &= \frac{1}{N} \sum_{i=1}^N E[R_{i,t-p} R_{i,t+q}] + E[R_{m,t-p} R_{m,t+q}] \\ &= \frac{1}{N} \sum_{i=1}^N (\text{Cov}[R_{i,t-p}, R_{i,t+q}] + E[R_{i,t-p}] \cdot E[R_{i,t+q}]) \\ &\quad - (\text{Cov}[R_{m,t-p}, R_{m,t+q}] + E[R_{m,t-p}] \cdot E[R_{m,t+q}]) \end{aligned} \quad (\text{A.4})$$

Lo 和 Mackinlay (1990) 指出，该分解成立的前提是要求股票回报率服从联合协方差稳定的随机过程(jointly covariance-stationary stochastic process)，即股票的期望回报率在时间序列下稳定，因此有：

$$E[R_{i,t-p}] \cdot E[R_{i,t+q}] = (q/p) E^2[R_{i,t-p}], \quad E[R_{m,t-p}] \cdot E[R_{m,t+q}] = (q/p) E^2[R_{m,t-p}],$$

令 Γ_{p+q} 表示投资组合内个股收益的自协方差矩阵，且记 $E^2[R_{i,t-p}]$ 为 μ_i^2 ， $E^2[R_{m,t-p}]$ 为 μ_m^2 。

则由(A.4)式可得：

$$E[\pi_t(p, q)] = -\frac{1}{N^2} \left[\mathbf{1}' \Gamma_{p+q} \mathbf{1} - \text{tr}(\Gamma_{p+q}) \right] + \frac{N-1}{N^2} \text{tr}(\Gamma_{p+q}) + \frac{q}{p} \frac{1}{N} \sum_{i=1}^N (\mu_i - \mu_m)^2 \quad (\text{A.5})$$

其中，等式右边的推导利用了 $\sum_{i=1}^N \mu_i^2 - N \mu_m^2 = \sum_{i=1}^N (\mu_i - \mu_m)^2$ 。