

# 行业现金流的领先滞后结构是股票市场的定价因子吗？ ——来自中国A股市场的证据

邓皓天<sup>\*</sup>，母从明<sup>†</sup>，杨金强<sup>‡</sup>  
(上海财经大学金融学院，上海200092)

**摘要：**行业现金流的领先-滞后结构是资产定价领域十分重要的问题，本文基于2003年-2022年我国A股市场的数据，计算了各行业现金流相对于GDP的交叉相关系数和领先-滞后程度，并使用Fama和Macbeth回归分析验证了行业领先-滞后程度与股票横截面收益之间的显著正向关系。单变量组合分析的结果表明领先行业的年均超额收益相较滞后行业高8.73%，且收益分析证明该领先溢价因子不由其他因子解释。此外，本文还使用GMM分别估计了包含（不包含）行业领先溢价因子的线性模型，发现领先溢价因子能够显著降低五因子模型的错误定价。这一领先溢价可由本文提出的投资现金流模型解释：滞后行业的企业能够根据领先行业的企业对外生冲击反应的“提前信息”来调整投资现金流，即行业领先溢价来源于对提前信息的补偿。

**关键词：**行业领先-滞后结构；股票横截面收益；定价因子

---

<sup>\*</sup> 邓皓天，上海财经大学金融学院博士生，haotiangeng@stu.sufe.edu.cn

<sup>†</sup> 母从明，金融学博士，副教授，湖南大学金融与统计学院，mucongming@hnu.edu.cn

<sup>‡</sup> 杨金强，数量经济学博士，教授，上海财经大学金融学院，yang.jinqiang@mail.shufe.edu.cn

## 一、引言

1960年前后，CAPM的提出拉开了实证资产定价研究的序幕，在此之后的1970年代，套利定价模型（APT）为多因子模型奠定了基础，1993年Fama和French为了解决CAPM无法解释的异常收益问题，整合了众多学者提出的市场异象，从而提出了著名的三因子模型，正式敲开了多因子研究的大门：1993年，Jegadeesh和Titman提出动量效应，Carhart等于1997年将动量因子（MOM）加入Fama和French的三因子模型；2015年，Fama和 French从股利折现模型出发，在三因子模型的基础上加入了盈利因子和投资因子，提出了五因子模型；同年，Hou等从q-理论出发，提出了包含市场、规模、投资 and 盈利的四因子模型；2017年，Stambaugh等从投资者的非理性和认知偏差出发，得到了包含管理和表现因子的四因子模型。

上述模型已经在国际市场上得到了实证检验，然而，由于中国市场与国际市场存在诸多政治制度和市场环境的差异，针对中国股票市场的因子模型有效性检验具有重要意义。目前，已有大量关于各种因子模型在我国适用性的研究：例如，2016年，赵胜民等发现Fama和French三因子模型相较五因子模型更适合我国证券市场；2017年，李志冰等则发现五因子模型优于三因子和Carhart四因子模型；2019年，Liu等对Fama和French的三因子模型进行改进，开发了一个更适合中国市场的新三因子模型。上述研究结果在一定程度上揭示了适用于我国市场的资产定价规律，也侧面表明，适合国际市场的已有因子并不一定能很好地为中国市场的资产进行定价。Croce等（2023）基于不同行业的“领先-滞后”结构，提出了全新的领先溢价因子（Lead-Lag factor，简称LL因子），他们发现美国股市中“领先”行业的股票相较“滞后”行业的股票的平均年化收益率高出3.6%。目前，我国暂无关于领先溢价对于中国市场的适用性研究，A股市场是否存在类似的行业“领先-滞后”结构？如果存在，能否被三因子或是五因子模型解释？加入新的领先溢价因子能否提升原有因子模型的解释能力？领先溢价因子能否降低因子模型的定价误差？而最重要的一点是，领先溢价因子是否只是向“因子动物园”中加入了一个新的因子（Feng等，2020）？。因此，本文将对我国A股市场的行业“领先-滞后”结构进行详细讨论，并使用各种不同的实证资产定价分析方法，对A股市场是否存在行业领先溢价进行充分检验。

本文使用2003年至2022年我国A股市场的上市公司数据，定量度量了各行业现金流对GDP的领先（滞后）程度，并对基于领先（滞后）程度构造的领先溢价因子进行了实证检验。具体而言，本文核心工作主要为以下七个方面：（1）构建了一个能够解释行业领先溢价来源的投资项目现金流模型，通过对外生冲击的不同反应速度来体现行业现金流的领先-滞后结构，证明了行业领先溢价来源于对领先行业提供的提前信息的补偿；（2）计算各行业的现金流总和与GDP的交叉相关系数，并根据不同滞后期的相关系数计算行业领先（滞后）程度；（3）对行业领先（滞后）程度进行相关性分析和持续性分析；（4）根据每一个季度的行业领先（滞后）程度对样本进行分组，通过领先和滞后组合构造零成本投资组合，并计算其平均收益率序列（即领先溢价因子）；（5）对领先溢价因子进行单样本T检验及收益分析；（6）使用Fama和MacBeth回归分析检验在控制了五因子后LL因子与股票未来收益之间的横截面关系；（7）使用GMM估计线性定价模型，验证LL因子是否能降低五因子模型的定价误差。

本文后续内容安排如下：第二部分为文献评述及理论模型，第三部分为研究设计，第四部分为实证分析，第五部分为进一步分析，第六部分为结语。

## 二、文献评述及理论模型

### （一）文献评述

许多宏观经济变量在面对外生冲击时会存在反应速度的区别，例如债券收益率对经济繁荣和衰退的反应十分迅速，而失业率则往往存在延迟，显然，不同的行业也存在类似的“领先”和“滞后”现象。行业现金流的“领先-滞后”结构（Lead-lag structure）指证券市场上不同行业的现金流在宏观经济周期中对外生冲击的差异化反应速度。这种跨行业的信息结构可能与股票的横截面收益率存在关系，从直觉来看，“领先”行业的股票平均收益率应高于“滞后”行业的股票平均收益率，这是由于“领先”行业“滞后”行业提供了确定性：持有“滞后”行业公司股票的投资者可以将“领先”行业的股票收益率视为宏观经济的风向标（正如谚语“矿井里的金丝雀”，如果金丝雀死了，矿工就知道空气有毒）。

目前，关于行业的领先-滞后结构的研究主要集中在两个方面：第一类研究集中在特定的网络结构上，例如Cohen和Frazzini（2008）认为客户-供应商结构与行业关联度之间存在关系，Gofman等（2020）认为生产网络中企业风险与其与消费者的距离之间存在关系，Menzly等（2010）认为分析师无法同时注意到上下游行业的信息从而使得部分行业存在领先溢价。第二类研究则认为这种领先-滞后结构是由规模因素导致的，主要表现在不完全信息下投资者对不同股票的关注（Merton, 1987）、分析师关注度（Brennan等, 1993）和机构投资者持股比例（Badrinath等, 1995）等与公司规模相关性较高的方面。基于这两类理论，以往的实证研究则主要通过整理供应商和客户名单和上下游关系来检验行业的领先滞后效应（Cohen和Frazzini, 2008），或是通过股票市值和交易量来判断行业的领先-滞后程度（刘煜辉和熊鹏, 2004）。而本文的研究则并非从上述两个角度切入：其一，本文定义行业领先-滞后程度来源于不同行业的企业对于外生冲击的不同内生敏感性，而并非内生横截面的领先-滞后结构；其二，本文的研究并非按照公司规模等因素进行分组，而是从行业角度入手，但这与Hou（2007）的研究并不相同，其研究并不能测度行业间的领先滞后结构，而本文的研究侧重于更长时间范围内的现金流基本面，且本文后续实证部分中的收益分析也验证了这类行业领先溢价并非是由规模因素导致的。综上所述，本文在构造行业领先-滞后程度时，不使用股票规模等作为分组依据，也并非根据投入产生或供求关系区分上下游行业，而这是由于本文认为行业领先溢价来源于领先行业企业对于滞后行业企业的提前信息溢价。

### （二）理论模型

为了说明行业领先-滞后结构是如何导致行业领先溢价的，本文提出如下模型。假设两个企业分别属于现金流“领先”和“滞后”行业，分别从时刻 $t=1$ 和时刻 $t=2$ 开始进行一个同样的项目。在每个时期，属于领先行业的企业投资的现金流为 $I_1^{Lead}, I_2^{Lead}, \dots, I_{T-1}^{Lead}$ ，而属于滞后行业的企业投资的现金流为 $I_2^{Lag}, I_3^{Lag}, \dots, I_{T-1}^{Lag}$ ，最终，两个企业同时在 $t=T$ 时刻完成投资。假设无风险利率为 $r$ ， $t=1$ 时刻领先企业的投资现金流 $I_1^{Lead}=I$ ，此后每一期投资的现金流按照无风险利率增长，即 $I_t^{Lead}=I_1^{Lead} \cdot e^{(t-1)r}=I \cdot e^{(t-1)r}$ ，其中 $t=1, 2, \dots, T-1$ 。假设在 $t=1$ 时，该项目预期在 $T$ 时刻的终值为 $K$ ：

$$\begin{aligned}
K &= \sum_{t=1}^{T-1} I_t^{Lead} \cdot e^{(T-t)r} \\
&= \sum_{t=1}^{T-1} I \cdot e^{(t-1)r} \cdot e^{(T-t)r} \\
&= (T-1)I \cdot e^{(T-1)r}
\end{aligned} \tag{1}$$

由于两个企业投资的是同一个项目，因此其每期的投资现金流同样需要满足在  $t=T$  时刻的终值为  $K$ ，即：

$$\begin{aligned}
\sum_{t=2}^{T-1} I_t^{Lag} \cdot e^{(T-t)r} &= K \\
\sum_{t=2}^{T-1} I_2^{Lag} \cdot e^{(t-2)r} \cdot e^{(T-t)r} &= (T-1)I \cdot e^{(T-1)r} \\
I_2^{Lag} &= \frac{T-1}{T-2} I \cdot e^r
\end{aligned} \tag{2}$$

因此， $I_t^{Lag} = [(T-1)/(T-2)]I \cdot e^{(t-1)r}$ ，其中  $t=2, 3, \dots, T-1$ 。假设在  $t=2$  时，外生冲击导致该项目预期在  $T$  时刻的终值发生了改变，企业实际需要达到的终值为  $K(1+\Delta_2)$ 。为了保持后续证明的一致性，可以认为在  $t=1$  时也存在外生冲击  $\Delta_1=0$  使得终值变为  $K(1+\Delta_1)=K$ 。由于  $t=1$  时投资的现金流已经发生，因此领先行业的企业需要改变此后的投资策略。改变后  $t=2$  时投资的现金流  $I_2'^{Lead}$  为：

$$\begin{aligned}
\sum_{t=2}^{T-1} I_2'^{Lead} \cdot e^{(t-2)r} \cdot e^{(T-t)r} &= K(1+\Delta_2) - I_1^{Lead} \cdot e^{(T-1)r} \\
(T-2)I_2'^{Lead} \cdot e^{(T-2)r} &= [(T-1)(1+\Delta_2) - 1]I \cdot e^{(T-1)r} \\
I_2'^{Lead} &= \frac{(T-1)(1+\Delta_2) - 1}{T-2} I \cdot e^r \\
I_2'^{Lead} &= \left( \frac{T-1}{T-2} \Delta_2 + 1 \right) I \cdot e^r
\end{aligned} \tag{3}$$

也即：

$$\frac{I_2'^{Lead}}{I \cdot e^r} = \frac{(T-1)\Delta_2 + (T-2)}{T-2} \tag{4}$$

相较原先的投资额  $I_2^{Lead}$ ，企业需要调整的现金流  $\Delta I_2^{Lead}$  为：

$$\begin{aligned}
\Delta I_2^{Lead} &= I_2'^{Lead} - I_2^{Lead} \\
\frac{\Delta I_2^{Lead}}{I \cdot e^r} &= \frac{T-1}{T-2} \Delta_2
\end{aligned} \tag{5}$$

同理，当  $t=3$  时，新的外生冲击导致终值变为  $K(1+\Delta_3)$ ，则改变后  $t=3$  时投资的现金流  $I_3'^{Lead}$  为：

$$(T-3)I_3^{Lead} \cdot e^{(T-3)r} = K(1+A_3) - I_1^{Lead} \cdot e^{(T-1)r} - I_2^{Lead} \cdot e^{(T-2)r} \quad (6)$$

$$(T-3)I_3^{Lead} \cdot e^{(T-3)r} = \left[ (T-1)(1+A_3) - 1 - \left( \frac{T-1}{T-2}A_2 + 1 \right) \right] I \cdot e^{(T-1)r}$$

$$\frac{I_3^{Lead}}{I \cdot e^{2r}} = \left[ \frac{(T-1)A_3 + (T-2)}{T-3} - \frac{(T-1)A_2 + (T-2)}{(T-2)(T-3)} \right]$$

$$\frac{I_3^{Lead}}{I \cdot e^{2r}} = \frac{[(T-1)A_3 + (T-2)](T-2) - [(T-1)A_2 + (T-2)]}{(T-2)(T-3)}$$

相较原先的投资额  $I_3^{Lead}$ ，企业需要调整的现金流  $\Delta I_3^{Lead}$  为：

$$\Delta I_3^{Lead} = I_3^{Lead} - I_3^{Lead} \quad (7)$$

$$\frac{\Delta I_3^{Lead}}{I \cdot e^{2r}} = \frac{T-1}{T-3} (A_3 - A_2)$$

同理，对于任意  $t$  时刻 ( $t=2, 3, \dots, T-1$ ) 的外生冲击，企业需要调整的现金流  $\Delta I_t^{Lead}$  为：

$$\frac{\Delta I_t^{Lead}}{I \cdot e^{(t-1)r}} = \frac{T-1}{T-t} (A_t - A_{t-1}) \quad (8)$$

而对于滞后行业的企业，延迟一期实施投资可以使其“提前”对投资现金流进行调整。在  $t=2$  时，属于滞后行业的企业可以直接调整  $I_2^{Lag}$  为  $I_2^{Lag}$ ：

$$\sum_{t=2}^{T-1} I_2^{Lag} \cdot e^{(t-2)r} \cdot e^{(T-t)r} = (1+A_2)K \cdot e^r \quad (9)$$

$$\frac{I_2^{Lag}}{I \cdot e^r} = \frac{T-1}{T-2} (1+A_2)$$

需要调整的现金流  $\Delta I_2^{Lag} = I_2^{Lag} - I_2^{Lag} = [(T-1)/(T-2)]A_2 I \cdot e^r$ ，因此，与领先行业的企业类似，对于任意  $t$  时刻 ( $t=2, 3, \dots, T-1$ ) 的外生冲击，企业需要调整的现金流  $\Delta I_t^{Lag}$  为：

$$\frac{\Delta I_t^{Lag}}{I \cdot e^{(t-1)r}} = \frac{T-1}{T-2} \cdot \frac{T-1}{T-t+1} (A_t - A_{t-1}) \quad (10)$$

假设领先行业的企业在  $t$  时刻调整投资的损耗为项目开始时间的增函数  $\eta(\cdot)$ 。当  $t=0$  时， $\eta(0)=0$ ，表示在项目开始时，调整投资现金流的损耗为0，而  $t>0$  时， $\eta(t)>0$ ，代表随着项目时间的增加，调整投资现金流存在损耗。所有时期需要调整的现金流在  $t=1$  时刻的现值  $L^{Lead}$  和  $L^{Lag}$  分别为：

$$L^{Lead} = \sum_{t=2}^{T-1} \Delta I_t^{Lead} \cdot e^{-(t-1)r} \cdot \eta(t-1) \quad (11)$$

$$= I \cdot \sum_{t=2}^{T-1} \frac{T-1}{T-t} (A_t - A_{t-1}) \cdot \eta(t-1)$$

$$\begin{aligned}
L^{Lag} &= \sum_{t=2}^{T-1} \Delta I_t^{Lag} \cdot e^{-(t-1)r} \cdot \eta(t-2) \\
&= I \cdot \sum_{t=2}^T \frac{T-1}{T-2} \cdot \frac{T-1}{T-t+1} (\Delta_t - \Delta_{t-1}) \cdot \eta(t-2)
\end{aligned} \tag{12}$$

领先和滞后行业的企业需要调整的现金流在  $t=1$  时刻的现值之差  $\Delta L$  为：

$$\begin{aligned}
\Delta L &= L^{Lead} - L^{Lag} \\
\frac{\Delta L}{(T-1)I} &= \sum_{t=2}^{T-1} \left[ \frac{\Delta_t - \Delta_{t-1}}{T-t} \cdot \eta(t-1) - \frac{T-1}{T-2} \cdot \frac{\Delta_t - \Delta_{t-1}}{T-t+1} \cdot \eta(t-2) \right]
\end{aligned} \tag{13}$$

假设对任意  $t=2, 3, \dots, T-1$ ，有  $\Delta_t - \Delta_{t-1} = \Delta$ ，则：

$$\begin{aligned}
\frac{\Delta L}{(T-1)I \cdot \Delta} &= \sum_{t=2}^{T-1} \left[ \frac{\eta(t-1)}{T-t} - \frac{T-1}{T-2} \cdot \frac{\eta(t-2)}{T-t+1} \right] \\
&\geq \sum_{t=2}^{T-1} \left[ \frac{\eta(t-1)}{T-t} - \frac{\eta(t-2)}{T-t+1} \right] = \eta(T-2) \\
\Delta L &\geq (T-1)I \cdot \Delta \cdot \eta(T-2)
\end{aligned} \tag{14}$$

又由于  $T > 2$ ，所以  $\Delta L = L^{Lead} - L^{Lag} > 0$ 。因此，相较于属于领先行业的企业，属于滞后行业的企业在外生冲击到来时调整的投资现金流的成本更小。综上所述，行业领先溢价来源于对“提前信息”的补偿，主要取决于投资项目的持续周期、不同时期外生冲击的变化幅度以及投资项目周期内不同时刻调整现金流的成本。

### 三、研究设计

#### （一）数据说明

本文在后续的实证部分所使用的数据均来源CSMAR数据库。其中，由于本文用于计算行业现金流总和的息税折旧摊销前收入（*EBITDA*）来源于企业财务报表，而我国大部分企业在2003年之前仅公布了半年报及年报，缺少第一季度和第三季度的数据，因此本文选取的样本期间为2003年1月1日至2022年12月31日。值得注意的是，由于后文中计算行业领先滞后程度时，需要以20个季度为窗口期，前后移动4个季度计算行业现金流与*GDP*的交叉相关系数，实际得到的第一个领先滞后程度的样本值对应于2010年一季度，因此，本文在后续检验因子时所使用的股票收益率样本期间为2010年1月1日至2022年12月31日。由于我国投资者主要进行交易的市场为A股市场，本文选取的上市公司样本均来源于上证A股市场（不包含科创板）和深证A股市场（不包含创业板）。我国主流的行业分类方式分为三种：中国证监会行业分类、中证行业分类以及申万行业分类。其中，中国证监会行业分类是一种管理型分类标准，其行业分布不够均匀，难以准确反映各个行业的发展状况。而对于中证行业，其一级行业仅有10个分类，公司较集中，相较于具有28个1级行业的申万行业分类更难凸显领先和滞后行业之间的差别，因此在计算行业现金流总和时，本文选用的行业划分标准为申万行业分类。而由于较新的2021版存在部分企业较早时期的数据缺失问题，因此本文选择使用2014版进行后

续分析。

与大多数文献保持一致，本文选取使用人民币三个月定期整存整取利率作为无风险利率，为了与股票月度收益率匹配，本文根据复利计算方法，将年度无风险利率转化为月度数据，即  $rfr_m = (1 + rfr_y)^{1/12} - 1$ 。在进行单变量组合分析时，由于市值加权组合赋予了具有高市值的股票更大的权重，而此类高市值股票往往具有更高的流动性，其交易成本更小，相较等权重，按照市值加权的收益率更易被投资者实现，而由于本文的研究对象不包含限售股（总股数=限售股+流通股数），因此本文在计算组合平均超额收益率时使用的市值加权方案为流通市值加权；在对领先溢价因子进行收益分析时，所使用的风险调整模型（资本资产定价模型、三因子模型和五因子模型）中的市场因子、规模因子、价值因子、盈利因子和投资因子对应的收益率数据同样按照流通市值加权计算。如无特殊说明，后文中使用五因子数据对应构造投资组合的方式均为  $2 \times 3$  类型，且相应投资组合于每年7月1日更新。本文在实施Fama和MacBeth回归分析时所使用的公司特征数据均来自CSMAR数据库提供的财务报表：其中，股票规模（*size*）为股票流通市值取对数，账面市值比（*BE/ME*）中 *BE* = 所有者权益 + 递延所得税负债 - 递延所得税资产 - 优先股账面价值，*ME* 则为 *t*-1年12月31日的流通市值；权益收益率（*OP*）为（营业收入 - 营业成本 - 销售费用 - 管理费用 - 利息支出）/ *BE*；净资产收益率（*ROE*）为净利润除以股东权益。

为了保证检验结果的有效性，本文对数据进行了如下处理：（1）剔除上市未满一年的企业；（2）剔除ST、PT股（在构造行业领先-滞后程度指标时，为了防止行业现金流在两期发生突变，无需剔除ST、PT股）；（3）在进行因子的相关检验时，剔除金融行业（包含银行和非银金融行业）的样本；（4）在计算LL因子时，剔除财务报表披露不及时（披露时间晚于组合调整时间）的企业样本，即剔除在每年8月30日后公布一季报、12月31日后公布半年报、第二年3月31日后公布三季报以及第二年6月31日后公布年报的企业。（5）为了防止极端值的影响，本文在后续分析前对数据进行了1%和99%水平上的缩尾处理。

## （二）行业领先-滞后程度指标构建

参考Croce等（2023）的构建方法，每一期的 *LL* 指标选择使用 *T* 个季度作为样本窗口期滚动计算。为了防止每个窗口期内行业现金流突变，本文首先对每个窗口期内的样本进行滚动筛选，以保证每个样本在窗口期内每个季度的数据皆可获取。根据Acharya等（2014）的选择，本文首先根据企业的息税折旧摊销前收入（= 净利润 + 所得税费用 + 财务费用 + 固定资产折旧、油气资产折耗、生产性生物资产折旧 + 无形资产摊销 + 长期待摊费用摊销）计算每个季度每个行业的现金流：

$$\begin{cases} CF_{j,t} = \sum_{i=1}^{N_{j,t}} EBITDA_{i,j,t} \\ CF_{j,t-1} = \sum_{i=1}^{N_{j,t}} EBITDA_{i,j,t-1} \end{cases} \quad (15)$$

其中， $EBITDA_{i,t}$  代表 *t* 时期行业 *j* 中企业 *i* 的息税折旧摊销前收入， $CF_{j,t}$  则代表 *t* 时期行业 *j* 的现金流总和。值得注意的是，为了防止行业现金流突变，在整个

滚动窗口期内，每个时期每个行业对应的企业数量  $N_{j,t}$  是固定的。在计算行业现金流增长前，首先对其进行CPI调整和季节性调整，本文使用虚拟变量法进行季节性调整，具体而言，对滚动窗口内的行业现金流除以对应期的CPI，而后通过纳入季节虚拟变量的回归分析得到季节性成份的估计值，最后将季节性成份从原始现金流数据中剔除即可。调整后，计算行业现金流增长和GDP增长：

$$\begin{cases} \Delta CF_{j,t} = CF_{j,t} - CF_{j,t-1} \\ \Delta GDP_t = GDP_t - GDP_{t-1} \end{cases} \quad (16)$$

而后，计算行业现金流增长和GDP增长的超前（滞后） $K$  期的交叉相关系数：

$$\rho_{j,t,k} = \text{corr}_{\{t-T \rightarrow t\}}(\Delta CF_{j,t-k}, \Delta GDP_t) \quad k = -K, \dots, 0, \dots, K \quad (17)$$

其中，超前（滞后）的时期  $K$  需小于滚动窗口期  $T$ 。例如，本文在后续计算中选择  $K = 4$ ， $T = 20$ ，则此时需要的样本时期共  $T + 2K = 20 + 2 \times 4 = 28$  个季度，如图1所示：

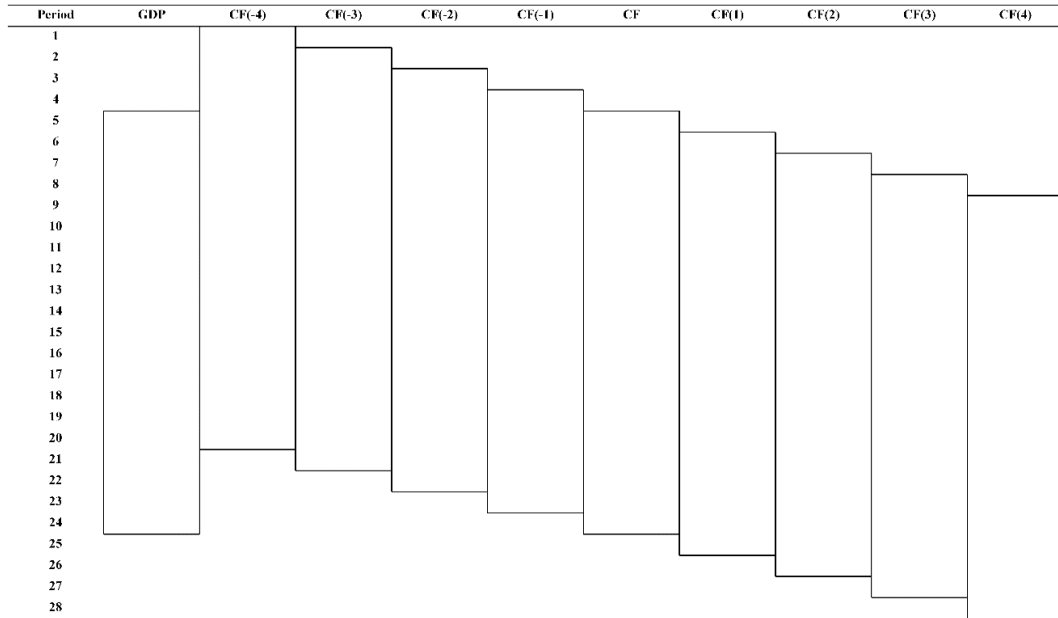


图1 交叉相关系数计算方法

本文选取的样本起始时期为2003年第一季度，因此根据上述交叉系数计算方式，所得最早的LL指标对应日期应为2003年第一季度后的28个季度，也即2010年第一季度。最后，根据交叉相关系数  $\rho_{j,t,k}$  计算行业领先-滞后程度。假设  $\Delta GDP_t$  符合  $AR(1)$  过程，即  $\Delta GDP_t = \rho \Delta GDP_{t-1} + \varepsilon_t$ （其中  $0 < \rho < 1$ ），且  $\Delta CF_{j,t-m} = -\Delta GDP_t$ ，则行业  $j$  的现金流领先了  $GDP$   $m$  个季度：

$$\begin{aligned} \Delta CF_{j,t-k} &= \Delta CF_{j,t-m+(m-k)} \\ &= -\Delta GDP_{t+(m-k)} \\ &= -\rho^{|m-k|} \Delta GDP_t + \mu_t \\ &\Downarrow \\ \rho_{j,t,k} &= -\rho^{|m-k|} \end{aligned} \quad (18)$$

又因为  $0 < \rho < 1$ ，所以行业领先（滞后）期数，也即  $LL$  指标为：



$$LL_{j,t,k} = \operatorname{argmax} \{\rho^{|m-k|}\} = \operatorname{argmax} \{|\rho_{j,t,k}|\} \quad (19)$$

然而，当  $\Delta GDP_t$  不符合  $AR(1)$  过程时，上述方式则会使得  $LL$  指标失去了其他超前（滞后）期相关系数的信息，因此，第二种  $LL$  指标的度量方式即为根据超前（滞后）期数加权的相关系数计算，相关系数绝对值越大的超前（滞后）期数对应更大的权重：

$$LL_{j,t,k} = \frac{\sum_{k=-K}^K k \cdot |\rho_{j,t,k}|}{\sum_{k=-K}^K |\rho_{j,t,k}|} \quad (20)$$

最后一种计算方式则是根据行业调整。考虑到某些行业可能领先（滞后）GDP 相同期数，但领先（滞后）程度不同，在计算  $LL$  指标时也可根据行业加权计算：

$$LL_{j,t,k} = \sum_{k=-K}^K \frac{k \cdot |\rho_{j,t,k}|}{\sum_{j=1}^N |\rho_{j,t,k}|} \quad (21)$$

相较根据不同超前（滞后）期数对应不同相关系数加权计算，这种计算方式考虑到相同领先（滞后）期对应的不同领先（滞后）程度，其分母为该期所有行业对应相关系数绝对值的和。考虑到现实情况中， $\Delta GDP_t$  一般不符合  $AR(1)$  过程，且根据第一种方法计算的  $LL$  指标分布并不均匀，大部分行业对应的领先（滞后）期数为4个季度，因此本文在后续实证分析中使用后两种计算方式所得  $LL$  指标，即  $LL(average)$  和  $LL(industry)$ ，其描述性统计结果如表1。

表1 描述性统计

	(1) <i>Mean</i>	(2) <i>SD</i>	(3) <i>Skew</i>	(4) <i>Kurt</i>	(5) <i>Min</i>	(6) <i>5%</i>
$LL(average)$	0.322	0.926	-0.293	2.213	-2.170	-1.272
$LL(industry)$	-0.000	0.236	-0.530	4.378	-1.214	-0.400
	(7) <i>25%</i>	(8) <i>Median</i>	(9) <i>75%</i>	(10) <i>95%</i>	(11) <i>Max</i>	(12) <i>N</i>
$LL(average)$	-0.365	0.374	1.115	1.642	2.247	1404
$LL(industry)$	-0.140	0.023	0.156	0.346	0.785	1404

注：上述结果均为平均横截面描述性统计，本文首先对每个指标在每个截面计算其分布特征，而后再计算样本期内横截面值的时间序列均值。

值得注意的是，上述描述性统计中样本容量为1404，这是由于该面板数据中的每一个样本代表某一季度某一个行业对应的  $LL$  指标值，而非股票。下面以中国资本市场中两个十分典型的行业：房地产行业和非银金融行业为例，展示并分析  $LL$  指标的实际变化情况。图2中展示了两个行业的  $LL(average)$  指标近10年的变化以及对应的重大事件。其中，房地产行业在近10年内的大部分时间都领先于GDP，房地产行业的繁荣可以带动消费投资，从而拉动经济增长，这也十分符合我国实际情况；而非银金融（证券、保险及多元金融等）的  $LL$  指标变化也能很好得对应我国资本市场的主要变化，例如2015年股市危机后证券市场受到的冲击、此后中美贸易战的持续升级带给中国资本市场的全面挑战以及近几年出现的新冠疫情和

俄乌战争带来的负面影响，都很好得被  $LL$  指标捕捉并体现在非银金融行业的领先程度减弱上。

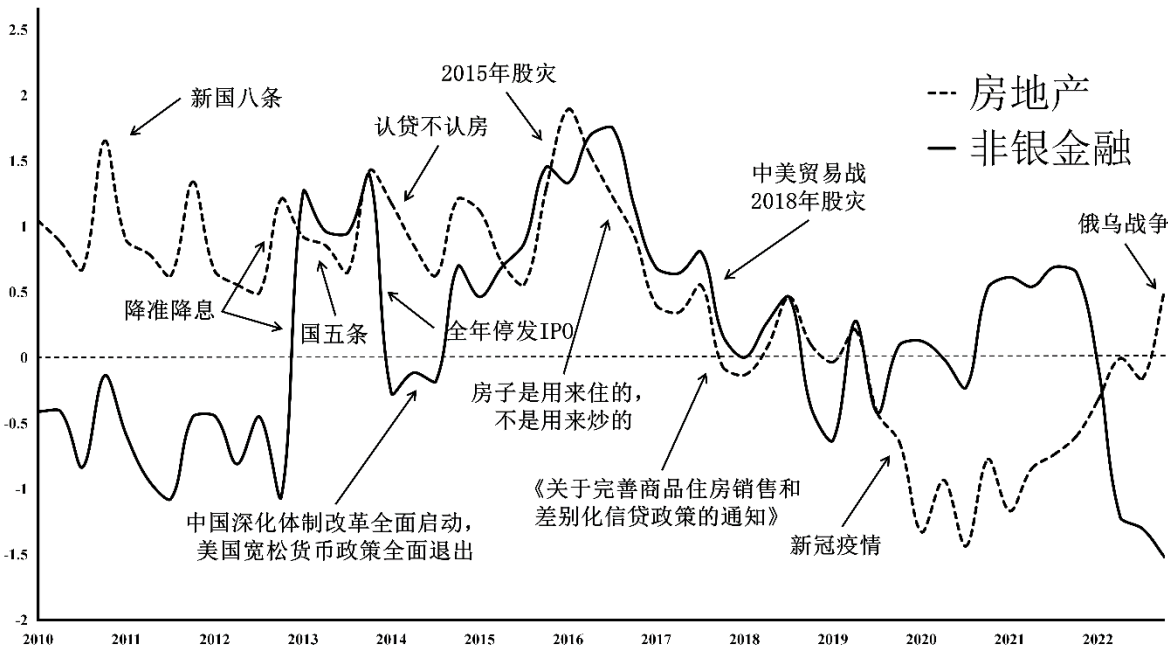


图2 房地产与非银金融行业的  $LL(average)$  指标折线图

#### 四、实证分析

##### （一）相关性分析

表2给出了  $LL(average)$ 、 $LL(industry)$ 、 $\beta$ 、 $size$ 、 $BM$ 、 $OP$  和  $ROE$  之间的相关性分析，与描述性统计类似，本文先对给定变量两两之间计算相关系数，而后再计算相关系数时间序列的横截面均值。正如预期一样， $LL(average)$ 、 $LL(industry)$  的Pearson相关系数为0.796，Spearman相关系数为0.811，前者小于后者，表明二者之间的关系是非线性的，这是由于二者对超前（滞后）期数的加权方式不同。而无论是  $LL(average)$  还是  $LL(industry)$ ，与其他五因子指标的相关系数皆很小，足够减轻在包含这些变量的多元分析中可能存在的潜在问题，更进一步看，Pearson相关系数和Spearman相关系数的结果在数值上十分接近，足以解决潜在的数据问题或者变量之间线性关系不足的问题。值得注意的是， $LL(average)$  和  $BM$  的Pearson相关系数为  $-0.127$ ，这表明领先行业的公司具有相对较小的账面市值比，但这一相关系数尚未高到能够表明这两个变量捕捉到了相同的信息。

表2 相关性分析

	$LL(average)$	$LL(industry)$	$\beta$	$size$	$BM$	$OP$	$ROE$
$LL(average)$	—	0.811	0.080	-0.019	-0.191	0.090	0.087
$LL(industry)$	0.796	—	0.022	0.010	-0.073	0.028	0.069
$\beta$	0.065	0.012	—	-0.059	-0.109	0.030	-0.053
$size$	-0.040	0.001	-0.049	—	-0.003	0.102	0.225

<i>BM</i>	-0.127	-0.048	-0.077	0.053	—	-0.051	-0.148
<i>OP</i>	0.016	0.000	0.002	-0.010	0.104	—	0.338
<i>ROE</i>	0.017	0.012	-0.007	0.034	0.001	0.005	—

注：该表展示了主要变量两两之间的年横截面相关系数，其中对角线左下方为Pearson相关系数，对角线右上方为Spearman相关系数。由于Spearman相关系数为非参数指标，因此无需对相应变量进行缩尾处理。

## （二）持续性分析

表3 持续性分析

$\tau$	<i>LL(average)</i>	<i>LL(industry)</i>
3	0.793	0.750
6	0.674	0.631
9	0.564	0.523
12	0.488	0.451

虽然某个行业的领先-滞后程度会由于某些外生事件发生改变，然而短期内一个行业的领先-滞后程度并不会发生较大突变。因此，本文在分析 *LL* 指标持续性时，近计算1年（4个季度）期间内的平均横截面相关系数。短期内的 *LL* 指标应当具有良好的持续性，以保证其确实能良好地捕获行业的领先（滞后）特征，但长期来看，*LL* 指标理应由于外生冲击而失去持续性。表3展示了针对 *LL(average)* 还是 *LL(industry)* 的持续性分析结果，对于每个时期，计算给定变量在时期  $t$  和时期  $t + \tau$  的值之间的Pearson相关系数，其中  $\tau = 3, 6, 9, 12$  个月。分析结果表明在短期内，*LL(average)* 和 *LL(industry)* 都具有较为显著的持续性，同预期一致，持续性随着滞后期数的增加而减弱，当滞后期数为1年时，即  $\tau = 12$  个月时，Pearson相关系数已经下降到不足0.5，但由于1年时间足以使一个行业从领先变为滞后，这样的持续性分析结果也完全在接受范围内。

## （三）单变量组合分析

本文以最基本的单变量组合分析来展开行业领先-滞后程度与股票未来收益率之间关系的探究。单变量组合分析是一种非参数方法，无需对变量之间关系的性质做出任何假设。首先，对每一时期的27个行业按照 *LL* 指标排序，并根据排序结果将所有样本股票分为三个投资组合，分别表示领先组合（Lead组）、滞后组合（Lag组）和中间组合（Mid组）。其中Lead组和Lag组分别包含6个领先（滞后）行业的股票，值得注意的是，这一分组既使得Lead组和Lag组各包含了20%分位数的行业，又保证了每个投资组合在每一时期至少包含了15%以上的市值，使得后续分析的结果不会因为一些缺乏流动性的小市值公司影响。接下来，本文通过做多Lead组，做空Lag组从而构造了一个Lead-Lag零成本投资组合，为了确保结果的稳健性，本文还通过做多 *LL* 指标高于中位数的行业，做空 *LL* 指标低于中位数的行业从而构造了一个Lead-Lag Strong零成本投资组合。最后，本文分别计算了这5个投资组合对应下一期股票的平均超额收益率（按流通市值加权）序列，并对每个

组合的平均超额收益率序列进行了独立样本T检验，以检验对应组合是否存在超额收益。除此之外，本文还计算了不同LL指标构造的Lead-Lag零成本投资组合的月度夏普比率（Sharpe）和波动率（Volatility）。本文在检验组合的平均超额收益等于0的原假设时，根据Newey和West（1987）五期滞后进行调整，这是由于每个组合的超额收益率数据都是一个时间跨度从2010年10月到2023年3月的时间序列数据（即 $T=12\times12+3+3=150$ 个月），其自相关和异方差性可能放大或缩小估计均值的标准误差。当采用Bartlett核方法时（ $\alpha=2/9$ ），代入 $T=150$ 计算Newey和West调整的滞后期数会产生一个4和5之间的数值（也即 $4(T/100)^\alpha\approx4.38$ ），故本文在后续分析中均根据Newey和West五期滞后进行调整。需要说明的是，组合调整频率为季度是指每年10月1日、第二年1月1日、4月1日、7月1日，分别根据当年的一季报、半年报、三季报和年报及其之前的27个季度计算所得的LL指标对行业排序和构造新的组合；组合调整频率为年度则代表每年7月1日根据去年年报及其之前的27个季度计算所得的LL指标对行业排序和构造新的组合。在调整组合时，本文删除了计算样本窗口期内无法获得的数据以防止Baquero等（2005）提到的前视偏差（Look-Ahead Bias）问题。

表4 单变量组合分析

	每季度调整一次组合		每年调整一次组合	
	LL (average)	LL (industry)	LL (average)	LL (industry)
Lead	0.017*** (2.883)	0.018*** (3.030)	0.014** (2.435)	0.012** (2.323)
Mid	0.011* (1.803)	0.013** (2.096)	0.011* (1.750)	0.011* (1.819)
Lag	0.007* (1.728)	0.009 (1.629)	0.008 (1.547)	0.008* (1.662)
Lead-Lag	0.007*** (3.946)	0.008*** (4.133)	0.006*** (3.822)	0.004** (2.404)
Lead-Lag Strong	0.006*** (3.252)	0.006*** (3.045)	0.005*** (3.696)	0.004** (2.380)
Volatility	0.028	0.034	0.026	0.026
Sharpe	0.257	0.263	0.220	0.143

注：括号中报告了根据Newey和West五期滞后调整的T统计量，\*\*\*、\*\*和\*分别代表在1%、5%和10%水平下显著。

单变量组合分析的结果见表4，其中Lead组合的平均收益率均高度显著，这并不奇怪，因为众所周知股票产生的平均收益一般高于无风险利率，而相较Lead组合，Mid组合的平均收益率便已经大幅下降了，而Lag组合甚至不能提供了高于无风险利率的超额收益。而Lead-Lag组合的结果表明，行业的领先程度与未来股票超额收益之间存在显著的正相关关系，以LL (average)为例，按季度频率进行组合调整的Lead-Lag组合的年平均超额收益高达 $(1+0.007)^{12}-1=8.73\%$ ，对应T统计量为3.946，即使按Harvey等（2016）的严格标准来看，这一结果也是十分显著的。根据LL指标每年调整投资组合相较每季度调整的Lead-Lag组合平均超额收益率有所下降并不奇怪，这是由于季频的LL指标能够更及时地捕获当期的行业领先-滞后程度，但由于存在交易费用，季频的交易策略更难被投资者实现。然而，每

年调整一次投资组合也足以说明行业领先溢价的存在。

此外，本文还描绘了根据  $LL(average)$  调整的Lead、Lag和Lead-Lag组合的累计对数超额收益，如图3所示。对数超额收益定义为1加上超额收益的资产对数，在2010年9月到2023年3月这一时期内，Lead-Lag组合的平均月度超额收益（对数超额收益）为每月0.83%，月度超额收益（对数超额收益）的标准差是3.88%，因此Lead-Lag组合超额收益（对数超额收益）的年化夏普比率约为0.74（即月度对数超额收益除以相应的标准差再乘以 $\sqrt{12}$ ）。在这一时期内，Lead-Lag组合的对数超额收益总累计值约为127%。相较Lead组合和Lag组合，Lead-Lag组合的收益一直十分稳定，至少在2015年和2018年的股灾期间，其组合收益率也并未发生十分明显的下跌。

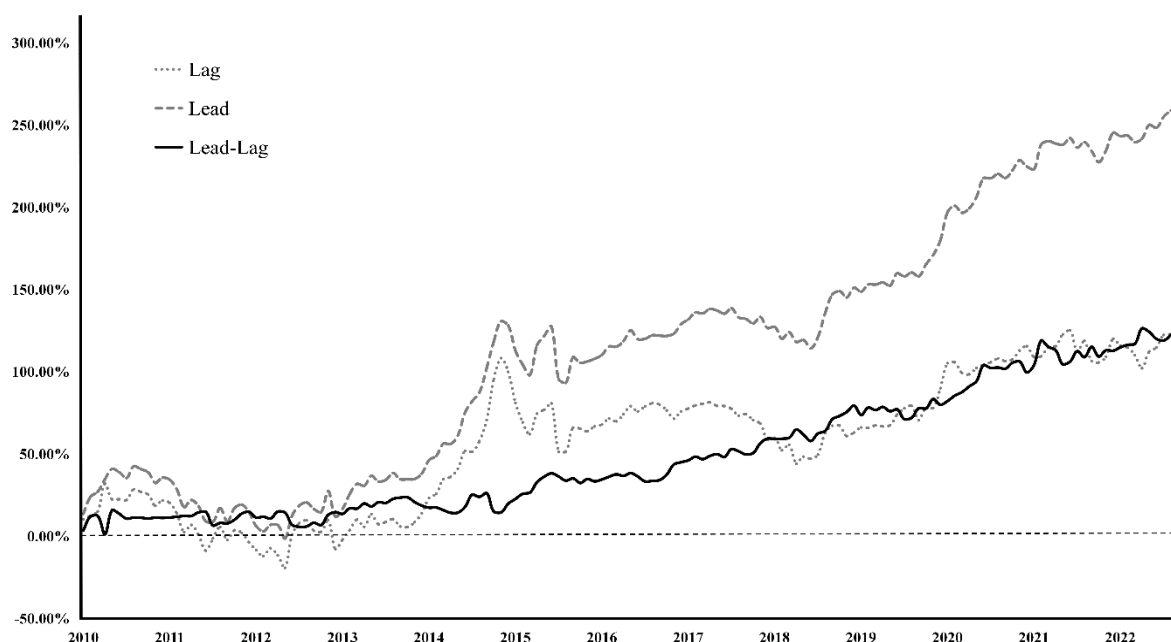


图3 Lead、Lag和Lead-Lag投资组合的累计对数超额收益

#### （四）Fama和MacBeth回归分析

接下来，本文使用Fama和MacBeth（1973）回归分析（后文中简称为FM回归）继续检验行业领先溢价因子。不同于前文中使用的单变量组合分析，FM回归可以包含较多的控制变量（多数情况下，假设被研究的控制变量和结果变量之间的关系是线性的）。具体而言，本文采用未来一个月的股票超额收益率（ $r_{t+1}$ ）作为因变量，股票 $\beta$ 、股票规模（ $size$ ）、账面市值比（ $BM$ ）、权益收益率（ $OP$ ）、净资产收益率（ $ROE$ ）以及 $LL(average)$ 或 $LL(industry)$ 指标作为自变量，在每个时期 $t$ 进行周期性横截面回归，并计算每个自变量在每个时期的斜率和截距系数的均值，检验其是否异于0。如果平均斜率系数显著，则说明在平均时期中给定的自变量和股票未来超额收益率之间存在一个横截面关系。所有自变量都在月度基础上进行了1%和99%水平下的缩尾处理，且 $T$ 检验的标准误差经过Newey和West五期滞后调整。FM回归结果见表5，列（1）为五因子指标对股票未来一期收益率的回归结果，列（2）和列（4）分别为 $LL(average)$ 和 $LL(industry)$ 指标的回归结果，列（3）

和列（5）则是在五因子的基础上分别加入  $LL(average)$  和  $LL(industry)$  指标的结果。

表5 FM回归分析

	(1)	(2)	(3)	(4)	(5)
截距	-0.001 (-0.038)	0.005 (0.767)	-0.001 (-0.037)	0.005 (0.757)	-0.002 (-0.071)
$LL(average)$	—	0.020*** (2.761)	0.018*** (2.739)	—	—
$LL(industry)$	—	—	—	0.018*** (3.070)	0.014*** (2.681)
$\beta$	0.008*** (2.487)	—	0.008*** (2.449)	—	0.008*** (2.486)
$size$	-0.000 (-0.233)	—	-0.000 (-0.229)	—	-0.000 (-0.194)
$BM$	0.002 (1.184)	—	0.002 (1.234)	—	0.002 (1.197)
$OP$	-0.000 (-0.973)	—	-0.000 (-0.982)	—	-0.000 (-1.002)
$ROE$	0.017*** (3.635)	—	0.016*** (3.609)	—	0.015*** (3.593)
$R^2$	0.051	0.008	0.058	0.005	0.057
样本容量	144897	144897	144897	144897	144897

注：括号中报告了根据Newey和West五期滞后调整的  $T$  统计量，\*\*\*、\*\*和\*分别代表在1%、5%和10%水平下显著。

相较列（1），在引入  $LL$  指标后，模型的  $R^2$  提高了。 $LL(average)$  和  $LL(industry)$  指标对应系数均为正数，且在1%水平下显著，这表明行业领先-滞后程度和股票期望收益之间存在很强的正相关关系，行业领先溢价在我国A股市场显著存在。在仅使用  $LL(average)$  作为唯一自变量的单变量回归中，其对应系数为0.020，对应的Newey和West调整的  $T$  统计量为2.761，而当加入其他因子作为控制变量时， $LL(average)$  对应系数下降到了0.018，但统计上依然高度显著，而  $LL(industry)$  的结果类似。回顾表1中对于两个  $LL$  指标的描述性统计结果， $LL(average)$  平均每个时期的横截面标准差为0.926，用其乘以FM回归中其对应的估计系数等于0.0167（ $=0.926 \times 0.018$ ），这表明  $LL(average)$  一个标准差的变化带来股票期望收益每年增加约1.67%。综上所述，FM回归分析为我国A股市场存在的行业领先溢价提供了有力证据。

## （五）定价误差分析

上述FM回归很好地验证了行业领先溢价的存在，下面本文将采用时间序列回归法以检验模型的定价误差。首先对资产用其超额收益率序列对因子收益率序列进行回归：

$$R_{i,t}^{ex} = \alpha_i + \beta_i \cdot F_t + \varepsilon_{i,t} \quad t=1, \dots, T \quad (22)$$

其中， $R_{i,t}^{ex}$  指资产  $i$  在  $t$  期的超额收益率， $F_t$  为  $t$  期的  $LL$  因子收益率，也即前文计算所得Lead-Lag组合按照流通市值加权的平均超额收益率序列。根据式（8）即可估计出  $LL$  因子在资产  $i$  上的错误定价  $\alpha_i$  和风险暴露  $\beta_i$  的估计值  $\hat{\alpha}_i$  和  $\hat{\beta}_i$ ，而后，将

估计值重新代入式（8），并在时序上对等式两边求期望可得：

$$E_T[R_i^{ex}] = \hat{\alpha}_i + \hat{\beta}_i \cdot \hat{\lambda} \quad (23)$$

考虑到式（8）中的随机干扰项  $\varepsilon_{i,t}$  可能存在的序列自相关或者异方差问题，本文使用GMM（Generalized Method of Moments）对上述模型进行检验。具体而言，本文首先估计了每个股票样本在固定滚动窗口期内对  $LL$  因子的风险暴露，并将所有样本股票根据估计的风险暴露分成30个测试组合。而后，将这些测试组合的收益率作为因变量， $LL$  因子及其他控制因子作为自变量，采用GMM进行估计，并检验联合定价误差是否等于0，GMM的估计结果见表6。

表6 定价误差分析

	(1)	(2)	(3)	(4)	(5)
$\lambda_{LL(average)}$	—	0.0042*** (4.005)	0.0042*** (2.946)	—	—
$\lambda_{LL(industry)}$	—	—	—	0.0038** (2.531)	0.0038** (2.305)
$\lambda_{MKT}$	0.0007 (0.136)	—	0.0007 (0.147)	—	0.0007 (0.129)
$\lambda_{SMB}$	0.0057 (1.235)	—	0.0055 (1.110)	—	0.0057 (1.300)
$\lambda_{HML}$	-0.0010 (-0.466)	—	-0.0010 (-0.463)	—	-0.0010 (-0.426)
$\lambda_{RMW}$	0.0009 (0.463)	—	0.0010 (0.590)	—	0.0009 (0.474)
$\lambda_{CMA}$	-0.0015 (-0.999)	—	-0.0014 (-0.976)	—	-0.0015 (-1.061)
$R^2$	0.1497	0.1805	0.4262	0.1247	0.3115
$J$ 统计量	127.21	368.78	74.75	434.01	87.35

注：括号中报告了根据Newey和West五期滞后调整的  $T$  统计量，\*\*\*、\*\*和\*分别代表在1%、5%和10%水平下显著。

表7的第（1）列报告了仅加入五因子模型的GMM估计结果，测试资产组合是根据  $LL(industry)$  分组，其结果与根据  $LL(industry)$  分组并无较大差异，其  $R^2$  为0.1497， $J$  统计量为127.21，而第（3）列和第（5）列分别加入了  $LL(average)$  和  $LL(industry)$  因子，模型的  $R^2$  分别提高到0.4262和0.3115，反应定价误差的  $J$  统计量则分别下降到74.75和87.35。综上所述， $LL$  因子确实在一定程度上降低了五因子模型的定价误差。

## 五、进一步分析

### （一）公司特征分析

为了确保Lead-Lag组合的平均超额收益不由其他因子驱动，本文检验了Lead组合与Lag组合在时间序列上其他公司特征股票  $\beta$ 、股票规模（ $size$ ）、账面市值比（ $BM$ ）、权益收益率（ $OP$ ）、净资产收益率（ $ROE$ ）上的差异。与前文中的方法类似，本文首先根据  $LL(average)$  对样本中的股票分组，分别计算其中Lead组和Lag

组对应公司特征的横截面均值序列，而后对Lead-Lag组合的每个特征横截面均值序列进行经过Newey和West五期滞后调整的独立样本  $T$  检验，检验结果见表7，Lead组合与Lag组合的公司特征并不存在显著的差异。值得注意的是，Lead-Lag组合中  $\beta$  和  $ROE$  的差异为负，这与前文中的相关性分析的结果相背，但由于其对应  $P$  值皆大于10%，且对应的相关系数均不高，这一相反的结果也不值得过分关注，而  $size$ 、 $BM$  和  $OP$  的结果均与相关性分析的结果一致，虽然其在Lead组合和Lag组合之间并无显著差异。

表7 Lead组合与Lag组合的公司特征差异分析

	Lead	Lag	Lead-Lag	$T$ 值	$P$ 值
$\beta$	1.172	1.189	-0.017	-1.045	0.296
$size$	15.404	15.410	-0.007	-0.122	0.903
$BM$	0.684	0.726	-0.042	-1.417	0.157
$OP$	0.619	0.435	0.184	0.904	0.366
$ROE$	0.043	0.057	-0.014	0.366	0.332

### （二）收益分析

为了进一步检验Lead-Lag组合的平均超额收益是否由组合对系统性风险的敏感性在横截面上的变化所驱动，本文借助三种常见的风险调整模型：单因子市场模型（CAPM）、三因子模型（FF3）和五因子模型（FF5）对组合收益进行风险调整的时间序列回归。同样，标准误差经过Newey和West五期滞后调整，收益分析的结果见表8。截距项  $\alpha$  代表组合的平均超额收益，该收益不由选定风险调整模型中的其他因子决定，而其他斜率系数则代表组合对相应因子的敏感性估计。无论是哪种风险调整模型，其因子对应系数皆不显著，而截距项  $\alpha$  的估计值与表4中的Lead-Lag组合的平均超额收益十分接近，这说明基于行业领先-滞后程度排序所得的领先溢价因子并非是由其他五因子决定。值得注意的是，市场风险因子对应系数在季频调整的组合检验中系数为负，而在年频调整时系数为正，但这并不足以影响前述分析的结果，因为无论是季度还是年度调整组合，市场风险因子系数都不能认为显著异于0。综上，可以认为领先溢价因子在经过风险调整后，仍然具有显著的风险溢价。

表8 Lead-Lag组合经风险调整后收益

调整模型	估计系数	每季度调整一次组合		每年调整一次组合	
		$LL(average)$	$LL(industry)$	$LL(average)$	$LL(industry)$
超额收益	超额收益	0.007*** (3.946)	0.008*** (4.133)	0.006*** (3.822)	0.004** (2.404)
CAPM	$\alpha$	0.007*** (3.950)	0.009*** (4.158)	0.006*** (3.829)	0.004** (2.344)
	$\beta_{MKT}$	-0.018 (-0.421)	-0.020 (-0.433)	0.035 (0.861)	0.050 (1.066)
		0.006*** (3.203)	0.009*** (3.701)	0.005*** (3.091)	0.003* (1.948)
FF3	$\alpha$	-0.042 (-0.746)	-0.011 (-0.160)	0.029 (0.574)	0.049 (0.943)
	$\beta_{MKT}$	0.183 (2.320)	0.102 (0.955)	0.138 (1.686)	0.109 (1.197)
	$\beta_{SMB}$				



FF5	$\beta_{HML}$	0.041 (0.325)	0.119 (0.714)	0.089 (0.725)	0.085 (0.755)
	$\alpha$	0.006*** (3.132)	0.009*** (3.426)	0.005*** (3.088)	0.003** (2.022)
	$\beta_{MKT}$	-0.043 (-0.943)	-0.020 (-0.339)	0.029 (0.513)	0.049 (0.937)
	$\beta_{SMB}$	0.138 (1.219)	-0.011 (-0.069)	0.146 (1.200)	0.103 (0.802)
	$\beta_{HML}$	-0.038 (-0.213)	-0.037 (-0.198)	0.110 (0.904)	0.060 (0.485)
	$\beta_{RMW}$	0.065 (0.247)	0.056 (0.151)	-0.024 (-0.085)	0.037 (0.162)
	$\beta_{CMA}$	0.210 (0.873)	0.395 (1.333)	-0.055 (-0.317)	0.067 (0.389)

注：括号中报告了根据Newey和West五期滞后调整的 $T$ 统计量，\*\*\*、\*\*和\*分别代表在1%、5%和10%水平下显著。

### （三）稳健性检验

为了保证本文的结论并非由于选用了Acharya等(2014)对于现金流衡量方式，本文还尝试使用其他指标替换息税折旧摊销前收入（ $EBITDA$ ）进行检验。本文使用的其他现金流指标有为营业利润（=营业总收入－营业总成本＋公允价值变动收益＋投资收益＋汇兑收益＋其他业务利润＋其他收益＋净敞口套期收益＋资产处置收益＋资产减值损失＋信用减值损失）、利润总额（=营业利润＋营业外收入－营业外支出）和净利润，虽然相较通过 $EBITDA$ 构造的 $LL$ 指标，这些衡量方式构造的 $LL$ 因子对五因子模型的优化作用有所下降，但总体来看结果并无较大差异。除账面市值比（ $BM$ ）外，其余指标以及涉及到按照市值加权计算时，本文选用的都是按照流通股数进行加权计算，为了结果的稳健性，本文也按照总市值加权重重新计算了各项指标，但并不影响前文中的检验结果。另外，本文还分别对上证A股主板和深证主板进行了检验，且并未发现分市场的结果和全市场有所差异。由于缩尾处理对经过加权计算的 $LL$ 指标并无太大影响，但对于其他公司特征数据则影响较大，因此本文还尝试使用未缩尾数据重新进行检验，同样，行业领先溢价因子仍然显著（限于篇幅，未报告稳健性检验的结果）。

## 六、结论

本文研究了我国A股市场存在的行业领先-滞后结构以及其对于股票横截面收益的影响。首先，本文提出了一个能够证明行业领先溢价存在的理论模型，其核心在于滞后行业的企业可以晚于领先行业公司对外生冲击的反应来调整其投资现金流，从而使得滞后行业的企业相对存在更小的风险，这便意味这领先溢价的存在。此外，本文基于行业现金流和 $GDP$ 的交叉相关系数计算了不同行业的领先滞后程度（ $LL$ 指标），并检验了 $LL$ 指标与我国A股市场的横截面收益之间的关系。具体而言，领先行业相较滞后行业存在高达年均8.73%的溢价作为提前信息的补偿，行业的领先-滞后程度是影响股票横截面收益的重要因素，行业领先程度越高，股票期望收益越高，且该结论十分稳健，不受衡量指标或样本数据的影响，此外 $LL$ 因子并非由股票规模等其他因素导致，能够有效降低五因子模型的错误定价。

## 参考文献

- [1] Fama E F, French K R. Common risk factors in the returns on stocks and bonds[J]. Journal of Financial Economics, 1993, 33(1): 3-56.
- [2] Jegadeesh N, Titman S. Returns to buying winners and selling losers: Implications for stock market efficiency[J]. Journal of Finance, 1993, 48(1): 65-91.
- [3] Carhart M M. On persistence in mutual fund performance[J]. Journal of Finance, 1997, 52(1): 57-82.
- [4] Fama E F, French K R. A five-factor asset pricing model[J]. Journal of Financial Economics, 2015, 116(1): 1-22.
- [5] Hou K, Xue C, Zhang L. Digesting anomalies: An investment approach[J]. The Review of Financial Studies, 2015, 28(3): 650-705.
- [6] Stambaugh R F, Yuan Y. Mispricing factors[J]. Review of Financial Studies, 2017, 30(4): 1270-1315.
- [7] 赵胜民, 闫红蕾, 张凯. Fama-French五因子模型比三因子模型更胜一筹吗——来自中国A股市场的经验证据[J]. 南开经济研究, 2016(02): 41-59.
- [8] 李志冰, 杨光艺, 冯永昌, 景亮. Fama-French五因子模型在中国股票市场的实证检验[J]. 金融研究, 2017(06): 191-206.
- [9] Liu J, Stambaugh R F, Yuan Y. Size and value in China[J]. Journal of Financial Economics, 2019, 134(1): 48-69.
- [10] Croce M M, Marchuk T, Schlag C. The leading premium[J]. Review of Financial Studies, 2023, hhad009.
- [11] Feng G, Giglio S, Xiu D. Taming the factor zoo: A test of new factors[J]. Journal of Finance, 2020, 75(3): 1327-1370.
- [12] Cohen L, Frazzini A. Economic links and predictable returns[J]. Journal of Finance, 2008, 63(4): 1977-2011.
- [13] Gofman M, Segal G, Wu Y. Production networks and stock returns: The role of vertical creative destruction[J]. Review of Financial Studies, 2020, 33(12): 5856-5905.
- [14] Menzly L, Ozbas O. Market segmentation and cross-predictability of returns[J]. Journal of Finance, 2010, 65(4): 1555-1580.
- [15] Merton R C. A Simple Model of Capital Market Equilibrium with Incomplete Information[J]. Journal of Finance, 1987, 42(3): 483-510.
- [16] Brennan M J, Jegadeesh N, Swaminathan B. Investment analysis and the adjustment of stock prices to common information[J]. Review of Financial Studies, 1993, 6(4): 799-824.
- [17] Badrinath S G, Kale J R, Noe T H. Of shepherds, sheep, and the cross-autocorrelations in equity returns[J]. Review of Financial Studies, 1995, 8(2): 401-430.
- [18] 刘煜辉, 熊鹏. 中国市场中股票间领先-滞后关系的规模与交易量效应[J]. 世界经济, 2004(08): 50-59.
- [19] Hou K. Industry information diffusion and the lead-lag effect in stock returns[J]. Review of Financial Studies, 2007, 20(4): 1113-1138.
- [20] Acharya V, Almeida H, Ippolito F, et al. Credit lines as monitored liquidity insurance: Theory and evidence[J]. Journal of Financial Economics, 2014, 112(3): 287-319.
- [21] Newey W K, West K D. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix[J]. Econometrica, 1987, 55(3): 703-708.

- [22] Baquero G, Ter Horst J, Verbeek M. Survival, look-ahead bias, and persistence in hedge fund performance[J]. *Journal of Financial and Quantitative Analysis*, 2005, 40(3): 493-517.
- [23] Harvey C R, Liu Y, Zhu H. ... and the cross-section of expected returns[J]. *Review of Financial Studies*, 2016, 29(1): 5-68.
- [24] Fama E F, MacBeth J D. Risk, return, and equilibrium: Empirical tests[J]. *Journal of Political Economy*, 1973, 81(3): 607-636.

## **The impact of the lead-lag structure of industry cash flows on the cross section of stock returns**

**Abstract:** The lead-lag structure of industry cash flow is an important issue of asset pricing. In this paper, we use data from China's A-share market from 2003 to 2022 to calculate the cross-correlation coefficients and lead-lag degrees of industry cash flow relative to GDP. We use Fama and Macbeth regression analysis to verify the significant positive relationship between the lead-lag degree of industry and the cross section of stock returns. The results of the univariate portfolio analysis show that the annual excess return of the leading industry is 8.73% higher than that of the lagging industry, and the return analysis proves that this leading premium factor is not explained by other factors. In addition, we use GMM to estimate linear models with and without the leading premium factor and find that it can significantly reduce the mispricing of the five-factor model. This leading premium can be explained by the investment cash flow model proposed in this paper: companies in the lagging industry can adjust their investment cash flow based on the "advance information" reflected by the external impact of companies in the leading industry, thereby obtaining the leading premium.

**Key words:** lead-lag structure of industry cash flows; the cross section of stock returns; empirical asset pricing