

## 编者按

本期的两篇文章分别从非金融企业和散户投资者的角度对市场参与者的金融投资行为进行相关研究，张成思教授基于近年来中国非金融企业日益热衷于金融资产投资的背景下，对中国非金融企业的金融投资行为影响机制进行了探究，具体的，作者基于微观企业面对实业投资和金融投资两大类资产的投资组合选择背景，在已有的关于非金融企业的金融投资行为影响机制的研究的基础上，拓展金融投资无风险的假设，同时考虑金融投资风险和固定投资风险，并利用中国A股非金融上市公司面板数据进行实证检验，结果表明，中国非金融企业的金融投资行为的显著驱动因素是固定资产投资的风险占比，而不是金融资产与固定资产的投资回报率缺口，这在政策角度的启示是，降低实业投资的不确定性是避免过度金融投资，指导企业“脱虚务实”的对症良方。第二篇文章中，孟勇和代志新教授的研究对象是散户投资者，事实上，与世界上许多股市不同，我国股市的大部分交易都是由散户投资者(而非机构投资者)进行的，因此散户投资者在我国金融市场是具有重要地位的，作者通过在我国17个不同省份随机访问的方式，收集散户投资者的代表性样本，在此基础上研究我国散户投资者的社会人口特征，并通过计量经济学的分析方法探究自我报告的投资者情绪方面是否存在性别差异，以此填补相关研究在中国散户投资的性别差异和人口特征方面的空白，结果显示：1) 96.33%的散户投资者年龄在60岁以下，只有11.55%的是退休人员或学生(2) 投资者通常受过良好的教育(3) 投资者的投资规模大多低于10万元(88.09%)，投资/收入比率低于0.5(95.4%)，这意味着散户投资者对金融市场的影响可能被夸大了。该研究还发现投资者在情绪上的一个强有力的性别差异，即男性散户投资者更理性、更容易承受损失，并且是乐观的。这一研究结果为我国金融市场的散户投资者提供了更清晰的图景。

# 中国非金融企业的金融投资行为影响机制研究

张成思，郑宁

(中国人民大学财政金融学院，中国财政金融政策研究中心，100872)

**摘要：**本文尝试构建中国非金融企业的金融投资行为影响机制模型。我们基于微观企业面对实业投资和金融投资两大类资产的投资组合选择背景，拓展已有研究关于假设微观企业投资组合模型中金融投资无风险的假设，同时考虑金融投资风险和固定投资风险，推演出与现实情况更加贴近的理论模型。本文证明，拓展模型与传统模型的机制存在本质区别。基于中国 A 股非金融上市公司面板数据的实证结果表明：中国非金融企业的金融投资行为的显著驱动因素是固定资产投资的风险占比，而不是金融资产与固定资产的投资回报率缺口。

**关键词：**金融投资；上市企业；固定资产投资；投资组合

## 一、引言

企业的投资行为受到诸多因素影响，一般而言，不同企业的个体投资决策行为并不一定受到决策层或者学术界的普遍关注。然而，非金融企业的金融投资上升而实业投资持续下降的情况（即非金融企业出现金融化趋势）则值得注意：毕竟非金融企业是实业部门的代表，如果一国实体部门都普遍追求金融投资，都热衷于金融运营，甚至金融投资主导了企业的运转，那么这一现象就值得严重关切。本文正是基于近年来中国非金融企业日益热衷于金融资产投资的背景下引出所要研究的主题，即中国非金融企业的金融投资行为影响机制问题。特别是作为市场主力的上市公司，其金融投资行为究竟受到哪些因素影响？回答这一问题成为理解中国代表性实业部门微观行为特征的重要内容。

从相关文献发展的脉络层次来看，关于非金融企业金融投资行为的一支文献集中讨论全球范围内新自由主义思潮发展带来的公司治理理念和企业投资行为的变化，关注微观企业日渐金融化的趋势。相关研究发现，非金融企业越来越多地参与金融市场和金融交易（Stockhammer 和 Grafli, 2010），而且非金融企业的利润来源渠道越来越多地倚靠金融渠道获得（Krippner, 2005），从而导致经济发展的核心从生产制造部门（还有一些外延的服务部门）变为金融部门（Foster, 2007）。在这一领域的研究中，非金融企业的金融投资行

为受到股东价值论的影响（如 Froud 等人，2000）。特别是 20 世纪 70 年代以后西方资本市场的大发展带来了，从而改变了非金融企业的公司治理理念，由追求长期增长转变为短期股东价值最大化，表现为非金融企业增加金融交易而减少生产性投资，即形成金融化特征。国内有一支文献实际上也研究了非金融企业的金融化现象，不过是从非金融企业的影子银行化角度进行分析，包括非金融企业影子银行活动的识别（王永钦等人，2015）、融资结构对非金融企业从事影子银行业务的影响（韩珣等人，2017）、非金融企业金融化程度与杠杆率变动的关系（吴军等人，2018）等。

无论是从新自由主义思潮视角考察非金融企业的投资结构变化，还是从影子银行角度分析非金融企业的金融化现象，本质上都是非金融企业的金融化或者说金融投资行为问题。虽然以上研究都与本文的研究主题具有较强相关性，不过这些研究的关注点与本文主题有所不同：这些研究分析的是企业金融投资行为形成的结果和状态（即金融化以及影子银行化），当然也包括这种结果的度量和识别等，对理解企业金融化问题的后端具有重要意义；而本文的研究目标则是非金融企业进行金融投资的影响因素和驱动机制，是企业金融化问题的前端。

事实上，针对中国非金融企业的金融投资问题，多数文献集中于对中国非金融企业进行金融投资会带来什么样的影响和冲击等问题，而且研究结论出奇的一致，都认为非金融企业的金融投资行为对企业的各方面影响都是坏的、负面的甚至是破坏性的，包括对企业经营收益率的冲击（宋军和陆旸，2015），对企业经营性业务的全要素生产率的显著抑制（刘笃池等人，2016），对实体企业技术创新能力的严重破坏（谢家智等人，2014），对企业进行实业投资具有排挤效应（张成思和张步县，2016），长期来看会显著抑制企业进行技术创新的动力（王红建等人，2016，2017），而且不仅没有缓解反而增加了企业的融资难度（徐军辉，2013），最终还损害了实体企业未来主业的发展（杜勇等人，2017），导致市场扭曲和经济结构失衡（罗来军等人，2016）。

仔细阅读以上文献，能够感受到实体企业金融投资行为导致的负面效应，但是却不容易找到企业选择金融资产投资的影响因素和机制。为厘清这一问题，我们需要调整一下待研究问题的逻辑：既然非金融企业热衷于金融投资会带来如此多的破坏性影响，那么比分析其破坏性影响的表现、范围和程度可能更重要的一个问题是什么？应该是中国非金融企业为什么进行金融资产投资的问题。也就是说，我们需要深入研究非金融企业的金融投资行为机制。只有弄清楚这一问题，才有可能对症下药，对企业金融投资行为进行针对性的引导、干预或者管理，进而才能立足国家实业部门长久发展战略层面制定相应宏观政策。从这个层面看，研究非金融企业的金融投资行为影响机制可能更具有现实意义和战略意义。

目前，有部分文献开始注意到中国非金融企业从事金融资产投资的影响机制问题。例如，谢家智等人（2014）从纯实证角度分析什么因素驱动了中国制造业企业的金融化投资行为。他们并没有提供一个基础的理论框架来阐释金融投资行为的可能影响要素，而是有所选择的

(ad hoc) 指出制造业创新发展能力弱化、金融投资收益虚高以及机构投资者的公司治理对企业投资行为的控制因素共同驱动和强化了制造业的金融投资行为。由于该文献缺乏理论框架约束,同时影响因素的指标本身存在争议,特别是“金融投资收益虚高”这一变量并不能反映金融投资与实业投资的收益率缺口信息,而且也得不到数据验证,从而弱化了这一研究结论的可信性。

与此相比,宋军和陆旸(2016)的研究则重点更加集中,分析更加细致,结论更加具有针对性。他们基于 2007-2012 年期间中国 A 股上市的非金融企业会计报表获得企业金融资产和金融收益,重点研究了非金融企业的主业经营收益率如何影响金融资产投资(配置),其研究思路与 Orhangazi(2008)比较相似。宋军和陆旸(2016)的一个亮点和创新是他们从金融资产配置与主业经营收益的非线性关系构建模型。注意,这篇文献把企业金融投资的影响因素重点放在经营业绩上,所以本质上实证模型的设立也是基于这一定向因素的纯实证模型,而不是从企业投资组合的基础理论出发。

另外,江春和李巍(2013)也曾经接近非金融企业金融投资影响因素这一问题,他们通过中国 A 股上市公司非金融企业 2003 至 2012 年的微观数据探寻非金融企业持有金融资产的决策因素和可能原因,发现非金融企业随着公司规模增大持有金融资产的比例就会上升,作者认为非金融公司进行金融投资主要是为了投机目的。这一研究虽然提出了一个好问题,但是却没有真正提供非金融企业金融投资行为的影响因素,只是从变量的时序特征做出主观判断,缺乏相对严谨的理论逻辑(模型框架)论证,也缺乏科学的计量分析。

从相对科学和规范的层面看, Demir(2009)是基于底层理论来研究非金融企业金融投资行为的代表性文献。作者构建了一个企业投资组合选择模型,分析了南美三国(阿根廷、墨西哥和土耳其)微观企业面板数据,得出结论认为非金融企业进行金融投资的影响因素主要是金融资产和固定资产投资收益率缺口(即收益率差)和固定资产投资的绝对风险水平:收益率缺口越大企业的金融投资占比越高,固定资产投资的风险水平越高企业的金融投资占比也会越高。

这一结论初看上去既符合直觉又很有解释力,似乎可以直接用来分析中国对应的问题。然而,本文并非是简单地套用 Demir(2009)的模型然后将南美国家数据换成中国数据进行复制研究,而是要强调指出既有文献理论模型存在的重要问题,即假定金融资产投资的回报率恒定不变且金融投资无风险,阐释这一问题对理论模型终解表达式的实质影响,并对这一问题进行修正和完善,从而构建更具竞争力的投资组合模型,进而实证分析中国非金融企业的金融投资行为机制问题。有鉴于此,本文的研究将首先着重构建微观企业的投资组合选择模型,强调金融资产投资与固定资产投资都存在风险,然后进行理论模型推导,获得理论模型的解析表达式后,以理论模型为基础进行实证模型设立,最后基于中国 A 股非金融上市企业 2006 至 2016 年期间的面板数据,实证检验企业金融资产投资行为的影响因素。

根据本文的理论模型和实证研究,我们的结论与已有文献不同。我们发现,中国非金融

上市公司的固定资产投资相对风险而非绝对风险主导了企业的金融投资行为,而金融资产与固定资产投资收益率缺口并非企业金融投资行为的显著影响因素。这些不同发现反映了理论模型的底层设定与现实情况相符与否至关重要,决定了我们能否获得相关问题的科学结论。

本文以下部分的结构安排如下:第二部分是微观企业投资组合选择模型的理论推演,我们重点对比已有文献中金融投资有无风险的假设对理论模型解析表达式和经济涵义的影响,获得本文实证分析的理论模型基础;第三部分是基于理论模型的实证模型设立、假说分析和变量的描述性统计;第四部分汇报实证检验结果,并进行对比分析和稳健性检验;第五部分总结全文,并基于研究结果提出正确引导中国非金融企业合理进行金融投资的对策方略。

## 二、理论模型的对比与竞争

随着上世纪八九十年代开始兴起的金融自由化思想,现代企业特别是上市公司经营理念正在发生微妙的变化,企业投资决策过程中日益关注股东价值最大化问题。与此对应,企业的投资决策本质上就是面对可选资产的投资组合选择问题。对绝大多数上市公司来讲,这种投资组合选择所面临的资产按大类可以划分为实业资产投资(即固定资产投资)和金融资产投资。Demir(2009)发表于 *Journal of Development Economics* 的文章认为,微观企业日益受到金融市场投资吸引,特别是相对金融投资而言,固定资产投资面临的不确定性(风险)可能更高,所以实体企业进行金融投资既有现实动力又有可行性。

为此,要考察实体企业的金融投资行为机制,投资组合选择模型提供了很好的理论框架基础。接下来,我们首先简要阐释了已有文献(Demir, 2009)关于实体企业进行投资选择的理论模型,用以指出此模型假设存在的现实问题,进而对已有理论进行拓展,提出更具竞争力的理论框架。我们要证明,底层模型的假设不同,会导致终解表达式截然不同,从而造成对实体企业金融投资行为机制的错误理解,并且会造成实证结论出现误导性。

Demir(2009)的理论模型设计是在 Tobin(1965)、Huang 和 Litzenberger(1988)、Tornell(1990)以及 Le 和 Zak(2006)的研究基础上,对实体企业两期投资组合选择进行优化设计。首先,假定代表性企业可以选择的投资标的可以划分为固定资产和金融资产,企业的投资效用来自于他们将可以用财富或者说资本投资于对应资产所获的收益。当然,要实现模型推导的可行性,还需要假设金融资产和固定资产的品种各自具有同质性,可以各自被视为一大类(即可以加总为一大类),且人口数量在两期内没有显著变化。

注意, Demir(2009)进一步的重要假设是  $t$  期金融资产投资  $I_t^f$  没有风险(上标  $f$  为 finance 的缩写,下标  $t$  表示时间),对应收益率在分析期内就是恒定的,因此金融投资收益率下标省略  $t$ , 记作  $r^f$ 。按照类似的符号标记规律,企业的固定资产投资可以写成  $I_t^k$ , 固定资产

投资有风险，收益率随时间波动，所以固定资产投资收益率有下标  $t$ ，记作  $r_t^k$ ，假设其服从正态分布，即  $r_t^k \sim N(\mu, \sigma^2)$ 。

进一步假定企业拥有的初始资本（Demir 称之为初始财富） $W_0$  可以在期初同时用于固定资产和金融资产两类投资。另外还要假设企业效用函数为严格增函数，且为连续凹函数，并使用  $E(\cdot)$  和  $Var(\cdot)$  表示期望和方差。

基于以上假设，代表性企业的预期效用的标准优化问题就可以写成：

$$\text{Max } E \sum_{t=0}^{\infty} \beta^t U(W_t) \quad (1)$$

约束条件为

$$W_t = (1 + r_t^k) I_t^k + (1 + r^f) I_t^f \quad (2)$$

因为初始资本  $W_0 = I_0^f + I_0^k$ ，所以等式（2）可以写成

$$W_t = W_0(1 + r^f) + I_t^k(r_t^k - r^f) \quad (3)$$

由于假设了金融投资风险为 0，所以根据协方差的斯坦因引理（Stein's Lemma），即  $\text{cov}(f(x), y) = E(f'(x))\text{cov}(x, y)$ ，企业的最优资产配置条件可以写成如下非常简单的形式，即：

$$I_t^{k*} = \frac{E(r_t^k - r^f)}{\gamma \text{Var}(r_t^k)} \quad (4)$$

其中  $\text{Var}(r_t^k)$  表示固定资产投资回报率对应的方差， $\gamma$  是效用函数期望式的二阶导数与一阶导数比值的负数，代表了企业不变风险厌恶水平。

另外，因为企业投资的总资本  $K_t^\alpha$  包括固定资产和金融资产，所以有：

$$K_t^\alpha = I_t^{k*} + I_t^f \quad (5)$$

将（5）代入（4）中可以得到均衡状态时金融投资的表达式，即：

$$I_t^f = K_t^\alpha - \frac{E(r_t^k - r^f)}{\gamma \text{Var}(r_t^k)} \quad (6)$$

我们可以将等式（6）两侧同时除以总投资资本  $K_t^\alpha$ ，得到金融投资占比的解析表达式，即：

$$\frac{I_t^f}{K_t^\alpha} = 1 - \frac{1}{K_t^\alpha} \times \frac{E(r_t^k - r^f)}{\gamma \text{Var}(r_t^k)} \quad (7)$$

需要再次强调的是，我们之所以能够推导出等式（7）这样的简洁表达式，关键是金融投资

无风险的假设。我们稍后会证明，这一假设一旦放松，金融投资占比的影响因素具体表达形式就不再像等式（7）这样简单。

Demir（2009）声称，基于等式（7）两侧取自然对数可以进一步推导得到企业投资占比更加显性的影响机制表达式，即：

$$\ln\left(\frac{I_t^f}{K_t^\alpha}\right) = -\ln(E(r_t^k - r^f)) + \ln(\gamma) + \ln(\text{Var}(r_t^k)) + \ln(K_t^\alpha) \quad (8)$$

从等式（8）可以看出，实体企业的金融投资占比主要由于两大因素决定，一是受固定资产与金融投资投资收益率差负向影响，二是受固定资产投资的风险水平影响。据此，Demir（2009）关于实体企业金融投资占比的实证模型设立形式可以概括为：

$$\left(\frac{I^f}{K^\alpha}\right)_{it} = \alpha_1 \text{Rgap}_{it}^{k-f} + \alpha_2 \text{Risk}_{it}^k + \text{控制变量} + \text{随机扰动项} \quad (9)$$

简单地说，就是企业金融投资行为受到投资收益率差  $\text{Rgap}_{it}^{k-f}$ （上标  $k-f$  表示固定资产投资回报率减去金融投资回报率）和固定资产投资面对的风险水平  $\text{Risk}_{it}^k$ 。

然而，即使不考虑 Demir 理论模型推导过程存在的问题，金融投资回报率恒定且金融投资无风险这一假设也显然与事实情况并不相符。企业的金融投资不仅仅限于国债类无风险产品，还包括诸如股票、期货等各种风险资产。因此，在企业投资组合选择模型的设立过程中，必须对金融资产投资的风险给予考虑，同时需要放松金融投资回报率恒定的假设，注意此时金融资产回报率符号下标此时就需要有时间  $t$ 。

需要注意的是，在放松了金融资产无风险假设之后，金融资产和固定资产投资回报率间是否存在相关性将影响投资组合的收益和风险，故此时可以根据相关性是否为 0 进行分类讨论。在本文的实证分析样本中，两类资产回报率的相关系数经过计算仅为 0.001，所以在理论推导层面我们直接以二者不相关的假设作为起点推导模型，以简化模型最终表达式的形式。至于更为细致的推导过程和分类讨论的情形，我们在附录中给出，下文中仅保留最为核心的推导部分，以节省篇幅。

我们假设固定和金融资产投资的回报率均服从正态分布，即：

$$\begin{cases} r_t^f \sim N(E(r_t^f), \text{Var}(r_t^f)) \\ r_t^k \sim N(E(r_t^k), \text{Var}(r_t^k)) \end{cases} \quad (10)$$

在此条件下，我们重新推导实体企业的投资优化问题。通过推导一阶条件并整理可以得到金融资产与固定资产投资回报率差（期望值）的关系，即：

$$E[r_t^f - r_t^k] = \alpha \left[ \text{Var}(r_t^f) I_t^f - \text{Var}(r_t^k) (K_t^\alpha - I_t^f) \right] \quad (11)$$

进而可以得到金融资产投资的最优配置为行为等式：

$$\frac{I_t^f}{K_t^a} = \frac{\text{Var}(r_t^k)}{\text{Var}(r_t^f) + \text{Var}(r_t^k)} + \frac{E[r_t^f - r_t^k]}{\alpha K_t^a [\text{Var}(r_t^f) + \text{Var}(r_t^k)]} \quad (12)$$

相应的实证模型则可以概括为:

$$\left(\frac{I^f}{K^a}\right) = \beta_1 \left(\frac{Rgap^{f-k}}{\text{total risk}}\right) + \beta_2 \left(\frac{Risk^k}{\text{total risk}}\right) + \text{控制变量} + \text{随机扰动项} \quad (13)$$

其中  $\beta$  表示系数。显然, 尽管在优化过程中仅增加了金融资产投资的风险项, 但推导出来的结果却与之前的结果在涵义上大相径庭 (对比等式 (12) 与 (7) 或等式 (13) 与 (9)): 首先, 影响实体企业的金融投资行为并不是单纯的固定资产投资风险, 而是固定资产投资的风险占比; 其次, 金融投资行为也不是单纯由两种投资回报率之差驱动, 而是经过总风险调整后的投资收益率之差。

### 三、基于理论模型的实证假说及变量说明

根据理论模型 (12) 可以看到, 企业的金融投资行为主要受到两个核心因素影响, 一是经过风险调整后的收益率缺口, 另外一个为固定资产投资的风险占比。在保持核心变量和理论模型一致的基础上, 我们还分别从融资约束水平、杆杆率、资产规模、资产有形性、成长性、股权结构特征等方面综合考虑了企业财务指标对企业投资选择的影响, 并引入相应的控制变量, 以避免遗漏变量问题。这样, 用来刻画企业面板数据的实证模型可以写成如下的简约形式:

$$fk_{it} = \beta_1 gap_{it}^{f-k} + \beta_2 risk_{it}^k + \chi \mathbf{V}_{it} + \beta_3 d_t + \beta_4 c_i + \eta_{it} \quad (14)$$

其中下标  $i$  表示企业,  $t$  对应于时间,  $d_t$  用来刻画时间固定效应,  $c_i$  刻画个体固定效应。 $\eta_{it}$  是模型设立随机扰动项,  $\beta$ 、 $\chi$  为各变量对应的系数,  $\mathbf{V}_{it}$  为控制变量矩阵。下面对变量进行具体说明:

$fk$  表示金融资产投资占比, 使用企业金融资产除以企业总资产进行计算, 根据我国企业会计准则的相关内容以及张成思和张步昙 (2016) 的设计, 企业的金融资产包括货币资金、持有至到期投资、交易性金融资产、投资性房地产、可出售金融资产、应收股利和应收股息七项。

$gap$  项是金融资产投资回报率与固定资产投资回报率之间的差再除以总投资风险, 即经过风险调整后的收益率缺口。这一指标的具体计算首先需要获得金融资产投资的回报率和固定资产投资回报率, 其次需要分别获得金融投资和固定资产投资风险再进行加总获得总投资风险。



参考张成思和张步昙（2016）的设计，金融投资和固定资产投资回报率具体计算公式如下：

$$r^f = (\text{投资净收益} + \text{公允价值变动损益} + \text{汇兑净收益} - \text{对联营和合营企业的投资净收益} + \text{利息收入} - \text{利息支出}) / (\text{货币资金} + \text{持有至到期投资} + \text{交易性金融资产} + \text{投资性房地产} + \text{可供出售的金融资产} + \text{应收股利} + \text{应收利息})$$

其中分子部分去除“对联营和合营企业的投资净收益”，原因在于这部分投资收益本质上是实体企业间的资金往来，是经营活动而非金融投资，不涉及金融市场，因此应该除去，出于同样的原因，分母中对应地未加入“长期股权投资”科目。

$$r^k = (\text{营业收入} - \text{营业成本} - \text{营业税金及附加} - \text{期间费用} - \text{资产减值损失}) / (\text{营运资本} + \text{固定资产} + \text{无形资产等长期资产的净值})$$

需要说明的是，我国的企业会计准则中并没有“固定资产收益率”的严格定义，因此本文用主营业务收益率来代理固定资产收益率。

对于投资风险，参考 Demir（2009）的设计，我们对  $r^f$  和  $r^k$  分别设立 GARCH(1, 1) 模型：

$$\begin{aligned} r_t &= c + u_t \\ \sigma_t^2 &= \delta_0 + \delta_1 \sigma_{t-1}^2 + \delta_2 u_{t-1}^2 \end{aligned}$$

其中  $r_t$  指代金融（或固定）资产投资回报率， $c$  和  $\delta_0$  分别是均值方程和方差方程中的常数项， $\delta_1$  和  $\delta_2$  为系数， $u_t$  是随机扰动项。方差方程中， $\sigma_t^2$  是  $u_t$  的条件方差，即本文用于度量  $\text{Var}(r_t^f)$  和  $\text{Var}(r_t^k)$  的值。通过 GARCH(1, 1) 模型得到金融资产和固定资产回报率在每个时点上的投资风险后，进行加总即获得总投资风险，然后通过收益率和风险变量计算出  $gap$  变量。

控制变量的设计，我们参考 Hovakimian（2009）、Shi 和 Zhang（2017）、张成思和张步昙（2016）的设计，共引入 8 个可能影响企业投资选择的控制变量，分别是：

$fc$ =经营性现金净流量/企业总资产，代理企业的融资约束水平；

$lev$ =企业总负债/所有者权益，即企业的财务杠杆率；

$asset$ =ln（企业总资产），代理企业的资产规模。需要说明的是：尽管在  $gap$  的计算中已经包含了总资产科目，但纳入回归后并不能直接看出企业规模对投资的影响，因此有必要重新引入企业资产规模进行控制；

$tangi$ =有形资产/企业总资产，即企业的有形资产比率；

$sales$ =营业收入同比增长率，代理企业的成长性；

$fslack$ =（现金及现金等价物余额+0.7 × 应收账款+0.5 × 存货-应付账款）/固定资产净额，该指数代理企业的流动性水平；

$bm$ =所有者权益/企业总市值，代理企业的资产价值；

*share*=公司前十大股东持股比例之和，代理企业的股权特征。

另外，为了从总体上了解各变量间的相关性特征，我们计算出各主要变量的相关系数，汇报在表 1 中。从表 1 中可见，金融资产回报率和固定资产回报率间的相关系数仅为 0.001，故从样本总体的角度可以视为无相关性。在控制变量中，*fc*、*fslack* 和 *share* 和 *fk* 呈正相关关系，其余控制变量则为负相关关系。当然，这一系数矩阵是基于样本的总体平均情形得到，不一定能够准确地展示解释变量和被解释变量基于企业个体的特征，故解释变量的系数方向为正为负需要通过回归分析进一步确定。

表1 各主要变量相关系数矩阵

	$fk$	$r^f$	$r^k$	$risk$	$gap$	$fc$	$lev$	$asset$	$tangi$	$sales$	$fslack$	$bm$	$share$
$fk$	1.000												
$r^f$	-0.018	1.000											
$r^k$	-0.002	0.001	1.000										
$risk$	-0.027	0.011	0.007	1.000									
$gap$	0.006	0.001	0.000	-0.002	1.000								
$fc$	0.006	-0.007	-0.005	-0.015	-0.001	1.000							
$lev$	-0.012	-0.001	0.000	0.008	0.000	-0.013	1.000						
$asset$	0.145	-0.023	-0.017	-0.044	0.005	0.078	0.019	1.000					
$tangi$	0.026	-0.006	-0.006	-0.018	0.000	0.010	0.000	0.062	1.000				
$sales$	-0.006	0.000	0.000	0.008	0.000	0.000	0.001	-0.002	0.000	1.000			
$fslack$	0.032	0.000	0.000	0.004	0.000	-0.029	0.018	0.009	0.013	0.000	1.000		
$bm$	-0.046	-0.050	-0.013	-0.104	0.002	0.016	0.000	0.409	0.061	-0.002	0.019	1.000	
$share$	0.103	-0.001	-0.017	0.010	-0.012	0.036	-0.002	0.220	0.017	0.008	-0.003	0.046	1.000

注：所有数值保留至小数点后三位。

为进一步规避样本极端值的影响，我们对模型（14）中除了  $fk$  之外的所有变量均进行了 winsorize 处理，将样本数值限定至 2.5% 至 97.5% 的分位区间内。同时给出各主要变量的描述性统计，如表 2 所示。从表 2 中可见，被解释变量  $fk$  均值为 0.2114，也即中国实体企业的金融投资占比水平约为 21%，标准差较大说明这一比值在不同企业间差异较大，最小值为 0 和最大值为 1 分别对应企业不持有金融资产和持有资产均为金融资产的情形。

表 2 各主要变量描述性统计

变量名称	均值	标准差	最小值	最大值
$fk$	0.2114	0.1583	0.0000	1.0000
$r^f$	0.0270	0.0681	-0.0215	0.3366
$r^k$	0.0753	0.1310	-0.2379	0.4451
$risk$	0.6000	0.3847	0.0000	1.0000
$gap$	-0.0001	0.0028	-0.2099	0.3316
$fc$	0.0304	0.0694	-0.1242	0.2061
$lev$	1.2824	1.2195	0.0477	5.5561
$asset$	21.5644	1.2948	19.0194	24.6988
$tangi$	0.4162	0.2278	-0.0807	0.8690
$sales$	0.1569	0.4125	-0.5107	1.7741
$fslack$	4.6284	12.1394	-0.0868	66.8531
$bm$	0.3518	0.2213	0.0170	0.9597
$share$	0.5589	0.1557	0.2494	0.8606

注：所有数值保留到小数点后四位。

两类资产的回报率则表现出了不同的总体特征，具体来说，金融资产投资回报率平均要远低于固定资产投资回报率，且其标准差却小于固定资产投资回报率的标准差，这表明，平均来看，尽管投资于金融资产回报不如投资于固定资产，但投资风险却较小，这一特征暗示了实体企业投资于金融资产也许并非出于对高收益的追求，而是与投资风险有密不可分的关联。进一步观察核心变量的总体指标， $risk$  的均值为 0.6，表明企业投资于金融资产和固定资产的总风险主要由固定资产投资风险构成，最小值为 0 和最大值为 1 分别对应企业仅持有金融资产和完全不持有金融资产的情形； $gap$  均值为 -0.0001，表明对多数企业个体而言，投资于固定资产得到的回报要高于投资于金融资产的回报。控制变量中， $fc$  均值为 3% 左右，但从极值数值的差异来看，不同企业面临的融资约束不尽相同，且水平悬殊。 $lev$  的均值约为 1.28，最大值超过了 5，过高的财务杠杆率逐渐成为政策制定部门的关注重点。

实体企业的资产规模同样存在巨大差异，极差达到了 5 左右，由于企业总资产经过自然对数处理，故在数值上每相差  $n$  个单位就意味着实际资产水平相差  $e^n$  倍。 $tangi$  的均值约为 41.6%，这一指标主要取决于企业的主营业务类型，若企业的生产经营依赖于厂房机器设备等有形资产，则该指标相对较高，而软件开发等无形资产占比较小的行业则相对较低。 $sales$  的均值约为 16%，标准差较大，表明不同企业的成长性差异巨大。 $fslack$  指数均值约为 4.6，这一指标主要用于衡

量企业的财务宽松程度，当企业持有较多流动性资产时，这一指数将较高，至于企业持有高流动性资产的目的则有多种可能，或出于规避未来可能出现的流动性危机，或用于短期投资等。*bm* 均值为 0.35 左右，标准差为 0.22 左右，表明不同企业间的资产质量差异较大。*share* 均值高达约 56%，表明中国实体企业的股权结构仍然高度集中。

在解释完各变量的计算方式和具体含义后，我们简短地说明实证模型对应的一些假说。根据本文的理论推导和 Demir 的实证假说，我们重点关注核心变量的系数：*gap* 项系数估计值预计为正向，因为金融资产投资回报率比固定资产投资回报率越高，则企业越有动力进行金融资产投资。其次，*risk* 项是固定资产投资风险占比，因此对应系数估计值预计也为正向，因为如果固定资产投资风险相对越高，那么企业就会越不愿意投资于固定资产扩张生产或者制造，而会在投资组合选择上更青睐金融资产。

当然，以上系数估计的方向只是对应于理论框架提出的实证假说，至于是否能够得到实证结果的验证并不一定清楚。举一个简单的例子来说，如果企业在其金融投资行为机制中极端风险厌恶，更加关注风险因素而对投资收益率关注的权重很小，那么实证结果就可能会出现风险项系数显著而收益率缺口项不显著。相反，如果企业属于激进型，更关心投资收益率，则可能出现收益率缺口项系数估计值显著而风险项系数估计值不显著。因此，我们需要通过实证分析来检验中国实体企业的金融投资行为机制特征。

## 四、实证分析

### 1. 基础模型的估计结果

为了估计实证模型 (14)，我们通过 Wind 资讯数据库获得中国 A 股上市非金融企业（即上市公司中除金融业之外的所有企业）自 2006 年至 2016 年的财务数据，剔除 ST 类企业和晚于 2010 年之后才上市的公司，共有样本企业 1902 家。Wind 资讯数据库提供了上市企业的各类财务指标，包括经营性数据和投资类数据等，为我们按照之前说明的方式计算各个变量指标提供了方便。

在计算好实证模型的关键变量指标之后，我们需要对模型进行估计方法的判断。面板数据中可能包含了无法准确观测到的个体效应，一般而言，个体效应是不随时间变化的，反映了不同个体的固有特质。个体效应存在与否、个体效应与解释变量间是否相关以及模型是否存在内生性问题决定了应选择何种估计方法来得到模型的系数估计值。具体可以通过以下几个步骤确定最优的估计方法：首先，若个体效应不存在，则可以直接使用混合最小二乘法进行估计；若个体效应存在且与解释变量相关，则应使用固定效应模型，若二者不相关则使用随机效应模型。其次，在确定了个体效应是否存在以及存在的方式之后，还应进一步检验模型是否存在内生性问题，从而引入工具变量用广义矩估计方法（以下略称为 GMM）得到更有效的系数估计值。再次，若确定了模型存在内生性问题，还应对引入的工具变量进行过度识别约束检验，考察工具变量是否有效。以上步骤是确定模型估计方法的必要过程，但其中我们重点关注模型的内生性问题。基于以上分析，本文在实证过程中引入了一系列检验对模型的设定进行了细致的考察，

具体而言：

(1) 对于模型的随机扰动项中可能存在的不随时间变化的个体效应，可以通过 Breusch 和 Pagan 拉格朗日乘数检验来判断，其原假设为不存在个体效应，在下文表格中用“LM-p”一行来表示检验统计的 p 值；

(2) 然后通过 Hausman 检验判断是否应使用固定效应模型，原假设为使用随机效应模型更有效。下文用“Hausman-p”表示结果；

(3) 判断模型是否存在内生性问题可以通过 Durbin-Wu-Hausman 检验来考察 (Durbin, 1954; Wu, 1973; Hausman, 1978)，但该检验是基于模型扰动项服从独立同分布的假设进行的，而这一假设往往无法满足，所以此时 Durbin-Wu-Hausman 检验的结果不再可信，为此本文引入了异方差稳健的 C 统计量来考察模型的内生性问题 (Hayashi, 2000)，C 统计量实际上是两个 Sargan-Hansen 统计量之差：其中被减数考虑了模型的内生性问题并引入对应工具变量，而减数则假设变量均为外生来进行估计；这里所说的内生性检验实质是通过判断两种不同估计方式得到的系数估计值的协方差矩阵之差是否为正定矩阵，从而确定哪种估计方法更有效，若结果显示为非正定矩阵，则被减数所对应的估计方法可以得到更有效的结果，即应该引入工具变量进行系数估计。若模型不存在异方差问题，C 统计量得到的结果实际上与 Durbin-Wu-Hausman 检验得出的结果是一致的，但相比 Durbin-Wu-Hausman 检验更具一般性和科学性，该内生性检验的原假设为模型不存在内生性问题。下文中用“endo-p”来表示；

(4) 若模型存在内生性问题，就还需要对引入的工具变量进行过度识别约束检验考察工具变量的有效性。过度识别约束检验使用了异方差稳健的 Hansen J 统计量，其原假设为工具变量集是有效的，下文用“HansenJ-p”表示这一结果；此外表格中还报告了模型识别检验的 p 值，下文中用“underid-p”来表示。该检验实质是检验原变量矩阵和工具变量矩阵的乘积是否满秩，若该乘积非满秩，则说明使用的工具变量使得模型无法被正确识别，其原假设为模型无法被正确识别，使用的统计量为 Kleibergen-Paap rk LM 统计量 (Kleibergen 和 Paap, 2006)。

尽管我们可以通过一系列统计检验来直观地判断模型是否存在内生性问题，但从理论上讲，由于本文采用的变量均由企业的财务数据计算得到，复杂的报表勾稽关系使得被解释变量和解释变量间不可避免地相互影响，同时，模型随机扰动项中可能包含的遗漏变量也可能与其他变量存在相关性，这些原因导致了模型必然存在内生性问题。结合理论分析和统计检验的结果，我们将本文所有的解释变量均视作内生变量（同方差情形下的 *risk* 视作外生变量），相应引入各变量的滞后项作为工具变量，同时结合统计检验的结果灵活调整工具变量阶数的设置，以得到科学稳健的系数估计值。

此外，由于本文的风险项 *risk* 是根据 GARCH (1, 1) 模型计算得到，当我们使用不同的模型设计形式时将得到不同的风险时间序列，因此仅仅使用单一计算形式欠缺说服力，对此本文还采用了时间窗口内的样本方差计算出对应的 *risk* 作为 GARCH 模型结果的对比，以增强结果的稳健性。需要指出的是，采用时间窗口内的样本方差等同于假设资产回报率在时间窗口服从同方差分布，此时可以视作投资在时间窗口内的平均风险，下文中同“同方差”来指代这一计算方式得到的结果。同时，由于通过计算方式得到的 *risk* 不随时间变化，故若采用固定效应模

型将无法得到系数估计值，因此下文涉及同方差计算方式的模型均采用随机效应模型，同时引入工具变量进行 GMM 估计，对应 Hausman 检验的  $p$  值则不再报出。最后，综合考虑模型扰动项可能存在的异方差和序列相关问题，本文在所有回归中使用了聚类（cluster）异方差稳健的标准误，以确保模型的系数估计值是异方差和序列自相关稳健的。

回归结果汇报于表 3 中。观察核心解释变量 *risk* 和 *gap* 的系数估计值不难发现，无论是使用 GARCH 模型计算得到的风险项，还是使用同方差假设下进行计算得到的投资风险，结果一致显示：*risk* 的系数显著为正，而 *gap* 的系数则不具有统计显著性。这一结果表明，中国非金融企业的金融投资行为主要是出于规避生产经营业务风险的目的，而并非由不同资产投资回报率之差驱动，这与已有文献中针对其他发展中国家的研究结果形成鲜明对比：特别是 Demir（2009）的文献结果显示不仅风险项对实体企业金融投资行为有显著影响，而且金融资产与固定资产的收益率之差也显著影响了企业的金融投资选择结果。

当然，我们在理论模型的对比部分已经说明，由于假设金融投资收益恒定且没有风险，已有文献的理论模型并不能真实刻画实体企业的金融投资行为，对应的实证结果自然不再具有可信性和可比性。工具变量的回归结果也完全一致，具体来说，表征融资约束的 *fc* 变量以及财务宽松指数 *fslack* 的系数显著为正，说明当企业现金流充足，且流动性资金相对于固定资产较多时，将正向刺激金融投资活动，加大对金融资产的投资比例；资产有形性指标 *tangi* 和杠杆率 *lev* 的系数项则显著为负，表明当企业有形资产占比较高时，企业将有更多的实物抵押品作为贷款支持来获取经营资金，因此通过金融资产投资获取额外资金的需求较小，但是较高的负债率将限制企业进一步投资的能力；*asset* 以及主营业务收入增长率 *sales* 的系数项显著为负表明，当企业主营业务收入经营成长较快时，将倾向于进一步投资于固定资产扩大主营业务，相应金融资产投资比例较低，反过来，当企业在主营业务经营上难以继续成长时，将选择投资于金融资产来获得额外收益，长期来看，随着资产规模的不断扩大，企业已经达到稳定经营状态，此时对金融资产投资的需求也有所下降；账面市值比 *bm* 和股权集中度指标 *share* 的系数显著为正，说明当企业的资产质量较低，企业难以吸引投资者投资，此时将倾向于投资于金融资产来获取额外资金，同时较高的股权集中度将促使企业对金融资产的投资比例。

表 3 面板数据模型（14）的 GMM 估计结果

	GARCH风险	同方差
<i>risk</i>	0.0433*** (0.0140)	0.0223*** (0.0082)
<i>gap</i>	-30.3646 (22.5904)	8.5844 (19.0162)
<i>fc</i>	1.3659*** (0.1844)	1.0386*** (0.0956)
<i>lev</i>	-0.0397** (0.0176)	-0.0257*** (0.0076)

	GARCH风险	同方差
<i>asset</i>	-0.1019 <sup>***</sup> (0.0178)	-0.0412 <sup>***</sup> (0.0049)
<i>sales</i>	-0.0685 <sup>**</sup> (0.0324)	-0.0062 <sup>*</sup> (0.0037)
<i>tangi</i>	-0.3765 <sup>**</sup> (0.1854)	-0.1064 <sup>*</sup> (0.0595)
<i>fslack</i>	0.0050 <sup>***</sup> (0.0008)	0.0037 <sup>***</sup> (0.0004)
<i>bm</i>	0.1794 <sup>***</sup> (0.0420)	0.0777 <sup>***</sup> (0.0166)
<i>share</i>	0.4079 <sup>***</sup> (0.0671)	0.1523 <sup>***</sup> (0.0217)
<i>constant</i>	2.3278 <sup>***</sup> (0.4164)	1.0133 <sup>***</sup> (0.1188)
样本数量	29,589	29,397
年份虚拟变量	是	是
LM-p	0.0000	0.0000
Hausman-p	0.0000	
endo-p	0.0000	0.0000
underid-p	0.0237	0.0000
HansenJ-p	0.2050	0.2731

注：对于 GARCH 风险模型，内生变量包括 *risk*、*gap*、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*，其中 *risk*、*fc*、*lev* 的工具变量选择其滞后二期项，*gap* 的工具变量选择其滞后二到三期项，*asset*、*sale* 的工具变量选择其滞后三期项，*tangi* 的工具变量选择为其滞后五期项，*fslack*、*bm*、*share* 的工具变量选择为其滞后一期项；对于同方差模型，内生变量包括 *gap*、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*，其中 *fc*、*lev* 的工具变量选择其滞后二期项，*gap*、*sales* 的工具变量选择其滞后一到二期项，*asset* 的工具变量选择其滞后三期项，*tangi* 的工具变量选择为其滞后五期项，*fslack*、*bm*、*share* 的工具变量选择为其滞后一期项。小括号中报告的是聚类异方差稳健标准误；\*\*\*、\*\*和\*表示统计量在 1%、5%和 10%水平下显著；所有数值保留 4 位小数（下同）。

## 2. 稳健性分析

为了确保实证结果的稳健性和基本结论的可信性，我们在模型（14）的基础上增加额外两种模型，并据此做了稳健性估计。

### （1）滞后效应模型和动态模型

首先我们考虑基准模型（14）中右手侧各个自变量对因变量（即金融投资占比）的影响具有时滞效应，即估计以下模型：

$$fk_{it} = \beta_1 gap_{it-1}^{f-k} + \beta_2 risk_{it-1}^k + \beta_3 fc_{it-1} + \beta_4 d_t + \beta_5 c_i + \eta'_{it} \quad (15)$$



其中  $\eta_{it}'$  表示新实证模型的设立误差项，其余各变量的含义与模型（14）完全一致。模型二是进一步将基准模型的滞后效应设定为与已有文献（Demir, 2009）完全相同的形式，即考虑如下含有部分自变量滞后项的动态面板模型（ $\eta_{it}''$  表示模型设立误差项），即：

$$fk_{it} = \alpha_1 fk_{it-1} + \alpha_2 fk_{it-2} + \beta_1 gap_{it-1}^{f-k} + \beta_2 risk_{it}^k + \beta_3 fc_{it} + \beta_4 d_t + \beta_5 c_i + \eta_{it}'' \quad (16)$$

需要说明的是，模型（16）中自变量只有  $gap$  项为滞后形式，这是 Demir 的模型设定形式，为了进行对比我们特此与其设定形式保持完全一致。

从本质上说，虽然模型（15）和（16）都考虑了滞后效应，但是各自所刻画的影响机制是不同的，这在已有各类文献中都未被提及和分析过。为此，我们在估计模型（15）和（16）之前，简略地阐释一下模型（15）和（16）在滞后项影响机制上的本质差异。要洞悉这一差异，我们可以利用滞后算子多项式的标准属性把模型（16）重新写成：

$$fk_{it} = \alpha(L)fk_{it} + \beta_2 gap_{it-1}^{f-k} + \beta_2 risk_{it}^k + \beta_3 fc_{it} + \beta_4 d_t + \beta_5 c_i + \eta_{it}'' \quad (17)$$

或者等同写成

$$(1 - \alpha L)fk_{it} = \beta_2 gap_{it-1}^{f-k} + \beta_2 risk_{it}^k + \beta_3 fc_{it} + \beta_4 d_t + \beta_5 c_i + \eta_{it}'' \quad (18)$$

其中  $\alpha(L) = \alpha_1 L + \alpha_2 L^2$ 。接下来将等式（18）左右同乘以  $(1 - \alpha(L))^{-1}$ ，即

$$fk_{it} = (1 - \alpha L)^{-1} (\beta_2 gap_{it-1}^{f-k} + \beta_2 risk_{it}^k + \beta_3 fc_{it} + \beta_4 d_t + \beta_5 c_i + \eta_{it}'') \quad (19)$$

因为滞后算子多项式  $(1 - \alpha(L))^{-1}$  总可以写成无穷个滞后算子连加的形式，即

$$(1 - \alpha(L))^{-1} = 1 + \gamma_1 L + \gamma_2 L^2 + L \quad (20)$$

其中  $\gamma_i (i=1, 2, L)$  是原始算子中系数  $\alpha$  的函数。因此，等式（19）本质上可以写成

$$fk_{it} = \sum_{n=1}^{\infty} \varphi_{1n} gap_{it-n}^{f-k} + \sum_{n=1}^{\infty} \varphi_{2n} risk_{it-n}^k + \sum_{n=1}^{\infty} \varphi_{3n} fc_{it-n} + \sum_{n=1}^{\infty} \varphi_{4n} d_{t-n} + \sum_{n=1}^{\infty} \varphi_{5n} c_{i-n} + \eta_{it}'' \quad (21)$$

在模型（21）中， $\varphi$  是滞后算子系数  $\alpha$  和原始模型中变量的系数  $\beta$  的组合。模型（21）清晰地表明，其对应的原始形式即模型（16）本质上捕捉的是所有自变量对因变量的多期滞后影响效应，同时还包含除  $gap$  变量之外的各变量当期效应，这些都与模型（15）的涵义迥异。正因为如此，对模型（15）和（16）进行估计，可以更全面地确保基础模型实证结果的稳健性。

与基础模型估计的过程类似，滞后效应模型（15）估计结果汇报在表 4 中，为了增强结果的稳健性，我们同样报告了采用同方差假设计算得到的风险项对应的估计结果。表中各变量后小括号中的数字代表相应变量的滞后项，如“ $risk(-1)$ ”代表  $risk$  的滞后 1 期项，下文采用同样的格式表示。

表4 滞后效应模型(15)的GMM估计结果

	GARCH风险	同方差
<i>risk</i> (-1)	0.0468 <sup>***</sup> (0.0166)	0.0227 <sup>***</sup> (0.0082)
<i>gap</i> (-1)	-32.7027 (31.3421)	-20.6580 (19.0006)
<i>fc</i> (-1)	0.6154 <sup>***</sup> (0.2011)	0.3408 <sup>***</sup> (0.0913)
<i>lev</i> (-1)	-0.0504 <sup>**</sup> (0.0217)	-0.0278 <sup>***</sup> (0.0080)
<i>asset</i> (-1)	-0.1228 <sup>***</sup> (0.0217)	-0.0369 <sup>***</sup> (0.0048)
<i>sales</i> (-1)	-0.0608 <sup>**</sup> (0.0308)	-0.0013 (0.0034)
<i>tangi</i> (-1)	-0.6313 <sup>***</sup> (0.2258)	-0.1650 <sup>***</sup> (0.0594)
<i>fslack</i> (-1)	0.0042 <sup>***</sup> (0.0009)	0.0027 <sup>***</sup> (0.0004)
<i>bm</i> (-1)	0.1897 <sup>***</sup> (0.0490)	0.0459 <sup>***</sup> (0.0160)
<i>share</i> (-1)	0.4591 <sup>***</sup> (0.0832)	0.1294 <sup>***</sup> (0.0207)
<i>constant</i>	2.9725 <sup>***</sup> (0.5123)	1.0143 <sup>***</sup> (0.1129)
样本数量	27,688	27,497
年份虚拟变量	是	是
LM-p	0.0000	0.0000
Hausman-p	0.0000	/
endo-p	0.0000	0.0000
underid-p	0.0574	0.0000
HansenJ-p	0.2059	0.4708

注：对于 GARCH 风险模型，内生变量包括 *risk*、*gap*、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share* 的滞后一期项，其中 *risk*、*fc*、*lev* 滞后一期项的工具变量选择其滞后二期项，*gap* 滞后一期项的工具变量选择其滞后二到三期项，*asset*、*sale* 滞后一期项的工具变量选择其滞后三期项，*tangi* 滞后一期项的工具变量选择为其滞后五期项，*fslack*、*bm*、*share* 滞后一期项的工具变量选择为其滞后一期项；对于同方差风险模型，内生变量包括 *gap*、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share* 的滞后一期项，其中 *fc*、*le* 滞后一期项的工具变量选择其滞后二期项，*gap*、*sales* 滞后一期项的工具变量选择其滞后一到二期项，*asset* 滞后一期项的工具变量选择其滞后三期项，*tangi* 滞后一期项的工具变量选择为其滞后五期项，*fslack*、*bm*、*share* 滞后一期项的工具变量

选择为其滞后一期项。

对于模型动态(16),由于解释变量引入了左手侧变量的滞后1期和2期项,因此可能会导致模型同时存在内生性和自相关问题,故应选择系统广义矩估计方法(下文略写为系统GMM)来得到更稳健有效的系数估计结果,同时通过调整工具变量的形式,消除模型自相关问题导致的估计偏差(Arellano和Bover,1995)。相比于静态面板,影响动态面板估计结果可靠性的问题主要来源于工具变量有效性和模型扰动项的序列自相关问题,因此在判断模型设定正确性的层面我们主要关注两类检验的结果:一是模型扰动项的序列自相关检验,该检验原假设是不存在自相关。由于误差项的一阶差分之间显然是序列相关的,因此一阶自相关检验往往是显著拒绝原假设的,故对于该检验我们主要考察二阶自相关检验的结果,若检验结果显示不存在序列自相关,则可以认为差分方程中相应工具变量符合外生性条件,从而认为模型估计结果是可靠的,下表中用“AR(1)-p”和“AR(2)-p”表示结果;二是工具变量的过度识别约束检验,该检验原假设为工具变量集有效,在下表中我们报告了异方差稳健的Hansen-J统计量的结果(Hansen,1982)。基于以上说明,表5汇报了动态模型(16)的估计结果。

表5 动态模型(16)的系统GMM估计结果

	GARCH风险	同方差
<i>fk(-1)</i>	0.4330*** (0.0923)	0.6740*** (0.1913)
<i>fk(-2)</i>	0.2649*** (0.0655)	0.1656 (0.1450)
<i>risk</i>	0.0195** (0.0086)	0.0395** (0.0164)
<i>gap(-1)</i>	8.3932 (7.0871)	165.3085 (104.6797)
<i>fc</i>	0.8129*** (0.0983)	0.7650*** (0.2200)
<i>lev</i>	0.0011 (0.0071)	-0.0410** (0.0170)
<i>asset</i>	-0.0901*** (0.0057)	-0.1078*** (0.0116)
<i>sales</i>	-0.0151*** (0.0040)	0.1756*** (0.0391)
<i>tangi</i>	-0.1881*** (0.0360)	-0.4356*** (0.0966)
<i>fslack</i>	0.0014*** (0.0003)	0.0011*** (0.0004)

	GARCH风险	同方差
<i>bm</i>	0.2067 <sup>***</sup> (0.0233)	0.2845 <sup>***</sup> (0.0292)
<i>share</i>	0.2935 <sup>***</sup> (0.0308)	-0.2288 <sup>***</sup> (0.0781)
<i>constant</i>	0.0000 (0.0000)	2.6083 <sup>***</sup> (0.2929)
样本数量	34,823	34,823
年份虚拟变量	是	是
AR(1)-p	0.0002	0.0135
AR(2)-p	0.1020	0.0933
HansenJ-p	0.4641	0.1057

注：对于 GARCH 风险模型，内生变量包括 *fk* 的滞后项、*risk*、*gap* 的滞后一期项、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*，其中 *fk* 的滞后项的工具变量选择对应的滞后五到六期项，*risk*、*fc*、*asset*、*tangi* 的工具变量选择其滞后三期项，*gap* 的滞后一期项的工具变量选择其滞后二到三期项，*lev* 的工具变量选择其滞后二期项，*sales*、*fslack*、*bm*、*share* 的工具变量选择为其滞后一期项；对于同方差风险模型，内生变量包括 *fk* 的滞后项、*gap* 的滞后一期项、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*，其中 *fk* 的滞后项的工具变量选择对应的滞后五到六期项，*fc*、*asset*、*lev* 的工具变量选择其滞后三期项，*gap* 的滞后一期项的工具变量选择其滞后六期项，*tangi* 的工具变量选择其滞后五期项，*sales*、*fslack*、*share* 的工具变量选择为其滞后二期项，*bm* 的工具变量选择其滞后一期项。

考虑解释变量时滞性的模型（15）的回归结果表明，无论使用哪一种风险的计算方式，均与基准模型保持一致，具有较强稳健性。动态模型（16）的结果进一步表明，中国非金融企业的金融投资行为具有时间上的延续性，同时核心解释变量的系数估计结果也与基准模型一致，说明长期来看，非金融企业进行金融投资主要是为了规避经营风险，而不是由两类资产的收益率缺口驱动。但值得注意的是，使用不同风险计算方式的情形中，控制变量的系数估计结果出现了不一致的情形：*sales* 和 *share* 的变量的系数相反，且均具有统计显著性。考虑到动态模型系数的实际含义是右手侧变量对左手侧变量的多期滞后效应，当我们考虑时间窗口内的平均风险时，有较高成长性的企业将更多投资于金融资产以对冲长期经营风险，同时，股权集中度系数估计值符号的差异表明，当企业考量风险角度不同时，出于不同的治理目标，对金融资产的投资需求也有所不同：当企业考虑长期平均风险时，股权结构较为集中的企业对金融资产的投资需求反而较低。其余控制变量的系数值均有较强的稳健性。

## （2）进一步讨论

在综合考虑解释变量的滞后效应模型的基础上，我们进一步从以下六个角度进行稳健性分析，以增强基准模型结果的可靠性和科学性：

第一是参考 Demir（2009）的设计，使用未经过总风险和总资产调整的 *gap* 替换模型中的 *gap*，重新进行模型（14）、（15）、（16）的估计，并和本文的设计进行对比。注意到 Demir（2009）

一文中的回归模型(9)是基于金融资产无风险的假设进行推导得到的,与本文的推导过程和结果存在明显的差异,而差异核心正是资产回报率缺口的形式,故我们将本文的 *gap* 替换为 Demir (2009) 中 *gap* 的计算形式,同时保持其他变量不变,从而验证理论模型基于不合理的假设将导致模型实证结果出现差异甚至不再可信。同时,由于在前文回归结果中已经验证了使用通过 GARCH(1,1) 计算的风险项进行的回归结果具有很强的稳健性,故在接下来的稳健性检验中不再报告使用同方差假设的相关结果,其余回归过程均与基准模型部分保持一致,结果汇报于表6和表7中。

结合表6和表7的结果,当使用 Demir (2009) 定义的 *gap* 进行回归时,静态模型中 *gap* 的系数表现出了高度显著的负向效应,这与理论框架和直观判断均相悖,结果存在一定疑问。而无论是静态模型还是动态模型, *risk* 项的系数均表现出了统计显著性,表明风险规避因素是驱动实体企业金融化的主要因素。其余工具变量的系数估计结果均与基准模型保持一致。可见,从不合理的假设出发推导得到的实证模型,将导致出现结果不具说服力的情形,同时也进一步表明,在考虑资产回报率缺口的时候,不能忽略资产投资风险的复合影响。

表6 不同 *gap* 定义下模型 GMM 回归结果

	基准模型		使用 Demir 定义的 <i>gap</i>	
	模型(14)	模型(15)	模型(14)	模型(15)
<i>risk</i>	0.0433*** (0.0140)	0.0468*** (0.0166)	0.0182** (0.0074)	0.0463*** (0.0119)
<i>gap</i>	-30.3646 (22.5904)	-32.7027 (31.3421)	-0.0777*** (0.0280)	-0.0199 (0.0396)
<i>fc</i>	1.3659*** (0.1844)	0.6154*** (0.2011)	0.6407*** (0.1046)	1.4036*** (0.1842)
<i>lev</i>	-0.0397** (0.0176)	-0.0504** (0.0217)	-0.0159*** (0.0060)	-0.0218*** (0.0070)
<i>asset</i>	-0.1019*** (0.0178)	-0.1228*** (0.0217)	-0.0569*** (0.0071)	-0.0678*** (0.0085)
<i>sales</i>	-0.0685** (0.0324)	-0.0608** (0.0308)	-0.0109*** (0.0034)	-0.0121*** (0.0042)
<i>tangi</i>	-0.3765** (0.1854)	-0.6313*** (0.2258)	-0.1297** (0.0620)	-0.2370*** (0.0746)
<i>fslack</i>	0.0050*** (0.0008)	0.0042*** (0.0009)	0.0041*** (0.0005)	0.0035*** (0.0005)
<i>bm</i>	0.1794*** (0.0420)	0.1897*** (0.0490)	0.1110*** (0.0181)	0.1234*** (0.0219)
<i>share</i>	0.4079*** (0.0671)	0.4591*** (0.0832)	0.2500*** (0.0301)	0.2698*** (0.0350)

	基准模型		使用 Demir 定义的 <i>gap</i>	
	模型 (14)	模型 (15)	模型 (14)	模型 (15)
<i>constant</i>	2.3278 <sup>***</sup> (0.4164)	2.9725 <sup>***</sup> (0.5123)	1.3097 <sup>***</sup> (0.1659)	1.5982 <sup>***</sup> (0.1971)
样本数量	29,589	27,688	31,207	29,307
年份虚拟变量	是	是	是	是
LM-p	0.0000	0.0000	0.0000	0.0000
Hausman-p	0.0000	0.0000	0.0000	0.0000
endo-p	0.0000	0.0000	0.0000	0.0000
underid-p	0.0237	0.0574	0.0000	0.0000
HansenJ-p	0.2050	0.2059	0.6198	0.6127

注：上述分类中，模型（14）中内生变量包括 *risk*、*gap*、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*，模型（15）的内生变量对应为模型（14）中各内生变量的滞后一期项。基准模型中，模型（14）中 *risk*、*fc*、*lev* 的工具变量选择其滞后二期项，*gap* 的工具变量选择其滞后二到三期项，*asset*、*sale* 的工具变量选择其滞后三期项，*tangi* 的工具变量选择为其滞后五期项，*fslack*、*bm*、*share* 的工具变量选择为其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致；使用 Demir 定义的 *gap* 中，*asset* 的工具变量选择其滞后二期项，*tangi* 的工具变量选择其滞后四期项，*gap* 的工具变量选择其滞后一到二期项，其余内生变量的工具变量选择其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致。

表7 不同 *gap* 定义下模型（16）系统 GMM 回归结果

	基准模型	使用 Demir 定义的 <i>gap</i>
<i>fk(-1)</i>	0.4330 <sup>***</sup> (0.0923)	0.4065 <sup>***</sup> (0.1095)
<i>fk(-2)</i>	0.2649 <sup>***</sup> (0.0655)	0.2686 <sup>***</sup> (0.0785)
<i>risk</i>	0.0195 <sup>**</sup> (0.0086)	0.0158 <sup>**</sup> (0.0078)
<i>gap(-1)</i>	8.3932 (7.0871)	-0.0093 (0.0252)
<i>fc</i>	0.8129 <sup>***</sup> (0.0983)	0.7635 <sup>***</sup> (0.0723)
<i>lev</i>	0.0011 (0.0071)	-0.0023 (0.0071)
<i>asset</i>	-0.0901 <sup>***</sup> (0.0057)	-0.0884 <sup>***</sup> (0.0056)
<i>sales</i>	-0.0151 <sup>***</sup> (0.0040)	-0.0155 <sup>***</sup> (0.0039)
<i>tangi</i>	-0.1881 <sup>***</sup>	-0.1959 <sup>***</sup>

	基准模型	使用 Demir 定义的 <i>gap</i>
	(0.0360)	(0.0361)
<i>fslack</i>	0.0014 <sup>***</sup>	0.0015 <sup>***</sup>
	(0.0003)	(0.0003)
<i>bm</i>	0.2067 <sup>***</sup>	0.1974 <sup>***</sup>
	(0.0233)	(0.0216)
<i>share</i>	0.2935 <sup>***</sup>	0.2887 <sup>***</sup>
	(0.0308)	(0.0326)
<i>constant</i>	0.0000	1.8476 <sup>***</sup>
	(0.0000)	(0.1175)
样本数量	34,823	34,823
年份虚拟变量	是	是
AR(1)-p	0.0002	0.0017
AR(2)-p	0.1020	0.1236
HansenJ-p	0.4641	0.1648

注：上述模型中内生变量包括 *fk* 的滞后项、*risk*、*gap* 的滞后一期项、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*。基准模型中，*fk* 的滞后项的工具变量选择对应的滞后五到六期项，*risk*、*fc*、*asset*、*tangi* 的工具变量选择其滞后三期项，*gap* 的滞后一期项的工具变量选择其滞后二到三期项，*lev* 的工具变量选择其滞后二期项，*sales*、*fslack*、*bm*、*share* 的工具变量选择为其滞后一期项；使用 Demir 定义的 *gap* 中，*fk* 的滞后项的工具变量选择对应的滞后五到六期项，*risk*、*asset*、*tangi* 的工具变量选择其滞后三期项，*fc*、*lev* 的工具变量选择其滞后二期项，*gap* 的滞后一期项的工具变量选择其滞后一到二期项，其余内生变量的工具变量选择其滞后一期项。

第二是根据企业的资产规模大小进行划分。为此我们首先计算各个企业在时间窗口内的平均总资产，然后取该均值截面数据的中位数作为划分大小规模企业的标准，不小于中位数的企业样本视作大企业，小于中位数的企业视作中小企业。之后分别利用两个企业样本对模型（14）、（15）、（16）重新进行估计。静态模型（14）、（15）的回归结果汇报于表 8 中，动态模型（16）的结果则汇报于表 9 中。结合表 8 和表 9 的结果不难发现，规模较大的企业的金融化并非由固定资产投资风险占比和总风险调整后的资产回报率缺口推动，而是更多地受到企业规模 *asset*、成长性 *sales*、财务宽松水平 *fslack* 以及股权集中度 *share* 的影响。而中小企业的金融投资占比则显著地受到 *risk* 的正向影响，表明对中小企业而言，持有金融资产主要是出于规避固定资产投资风险的目的，同时没有证据显示资产回报率缺口推动企业的金融化。控制变量中值得关注的是，静态模型中，中小企业样本中 *fc* 变量的系数值均大于大企业样本，表明在短期内，中小企业投资于金融资产的意愿受同期盈余现金流量影响更大，但从动态模型的结果来看却恰好相反，表明若考虑多期滞后的总效应，大企业的投资意愿对现金流水平更为敏感。其余工具变量的系数估计结果均非常稳健。

表8 按企业规模大小划分的GMM估计结果

	大企业		中小企业	
	模型(14)	模型(15)	模型(14)	模型(15)
<i>risk</i>	0.0009 (0.0137)	0.0103 (0.0146)	0.0902 <sup>***</sup> (0.0301)	0.0870 <sup>**</sup> (0.0380)
<i>gap</i>	-10.6447 (15.6205)	-0.7068 (18.1579)	-43.4166 (31.4961)	-51.7584 (38.3311)
<i>fc</i>	1.2008 <sup>***</sup> (0.1741)	0.6268 <sup>***</sup> (0.1850)	1.7341 <sup>***</sup> (0.4227)	0.9563 <sup>*</sup> (0.5175)
<i>lev</i>	-0.0202 (0.0194)	-0.0371 (0.0251)	-0.0663 <sup>*</sup> (0.0343)	-0.0873 <sup>*</sup> (0.0449)
<i>asset</i>	-0.0659 <sup>***</sup> (0.0142)	-0.0772 <sup>***</sup> (0.0163)	-0.1700 <sup>***</sup> (0.0481)	-0.2209 <sup>***</sup> (0.0601)
<i>sales</i>	-0.1240 <sup>***</sup> (0.0368)	-0.0969 <sup>***</sup> (0.0351)	-0.0074 (0.0588)	-0.0080 (0.0619)
<i>tangi</i>	-0.1418 (0.2026)	0.0026 <sup>***</sup> (0.0007)	-0.6930 <sup>*</sup> (0.3762)	-1.1869 <sup>**</sup> (0.5202)
<i>fslack</i>	0.0030 <sup>***</sup> (0.0007)	-0.3859 (0.2510)	0.0064 <sup>***</sup> (0.0016)	0.0062 <sup>***</sup> (0.0019)
<i>bm</i>	0.0497 (0.0335)	0.0405 (0.0390)	0.5158 <sup>***</sup> (0.1535)	0.6697 <sup>***</sup> (0.2029)
<i>share</i>	0.3223 <sup>***</sup> (0.0621)	0.3360 <sup>***</sup> (0.0706)	0.5251 <sup>***</sup> (0.1509)	0.6841 <sup>***</sup> (0.2004)
<i>constant</i>	1.5295 <sup>***</sup> (0.3476)	1.8965 <sup>***</sup> (0.4179)	3.7066 <sup>***</sup> (1.0911)	5.0856 <sup>***</sup> (1.4064)
样本数量	15,352	14,400	14,237	13,288
年份虚拟变量	是	是	是	是
LM-p	0.0000	0.0000	0.0000	0.0000
Hausman-p	0.0000	0.0000	0.0000	0.0000
endo-p	0.0000	0.0000	0.0000	0.0000
underid-p	0.0000	0.0003	0.0417	0.0571
HansenJ-p	0.9190	0.6754	0.8763	0.7504

注：上述分类中，模型（14）中内生变量包括 *risk*、*gap*、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*，模型（15）的内生变量对应为模型（14）中各内生变量的滞后一期项。大企业样本中，模型（14）中 *risk*、*fc*、*lev* 的工具变量选择其滞后二期项，*gap* 的工具变量选择其滞后二到三期项，*asset*、*sale* 的工具变量选择其滞后三期项，*tangi* 的工具变量选择其滞后五期项，*fslack*、*bm*、*share* 的工具变量选择其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致；中小企业样本中，模型（14）中，*risk*、*fc*、*lev* 的工具变量选择其滞后二期项，*gap* 的工具变量选择其滞后二到三期项，*asset*、*sale* 的工具变量选择其滞后三期项，*tangi*



的工具变量选择其滞后五期项, *fslack*、*bm*、*share* 的工具变量选择其滞后一期项; 模型 (15) 中各内生变量的工具变量设计形式与模型 (14) 一致。

表 9 按企业规模大小划分的模型 (16) 系统 GMM 估计结果

	大企业	中小企业
<i>fk</i> (-1)	0.3450** (0.1378)	0.4046*** (0.1543)
<i>fk</i> (-2)	0.4059*** (0.0538)	0.2397** (0.1075)
<i>risk</i>	-0.0076 (0.0089)	0.0579*** (0.0121)
<i>gap</i> (-1)	11.2924 (24.8259)	8.5595 (7.6985)
<i>fc</i>	1.6224*** (0.3466)	0.7013*** (0.0726)
<i>lev</i>	0.0158 (0.0289)	0.0064 (0.0102)
<i>asset</i>	-0.0468*** (0.0158)	-0.0532*** (0.0100)
<i>sales</i>	-0.1144** (0.0537)	-0.0496* (0.0300)
<i>tangi</i>	-0.0147 (0.3088)	-0.0574 (0.0410)
<i>fslack</i>	0.0026*** (0.0008)	0.0013** (0.0005)
<i>bm</i>	0.1167*** (0.0393)	0.1540*** (0.0494)
<i>share</i>	0.3441** (0.1665)	0.3493*** (0.0549)
<i>constant</i>	0.7920** (0.3746)	0.0000 (0.0000)
样本数量	18,085	16,738
年份虚拟变量	是	是
AR(1)-p	0.0001	0.0362
AR(2)-p	0.0812	0.2672
HansenJ-p	0.2724	0.3260

注: 上述模型中内生变量包括 *fk* 的滞后项、*risk*、*gap* 的滞后一期项、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*。大企业样本中, *fk* 的滞后项的工具变量选择对应的滞后一到二期项, *risk*、*fc*、*asset*、*tangi*、*fslack*、

*bm*、*share* 的工具变量选择为其滞后一期项, *sales*、*lev* 的工具变量选择其滞后二期项, *gap* 的滞后一期项的工具变量选择其滞后二到三期项,; 中小企业样本中, *fk* 的滞后项的工具变量选择对应的滞后五到六期项, *risk*、*lev* 的工具变量选择其滞后三期项, *fc*、*asset*、*sales*、*tangi*、*fslack* 的工具变量选择其滞后二期项, *gap* 的滞后一期项的工具变量选择其滞后二到三期项, *share*、*bm* 的工具变量选择为其滞后一期项。

第三是根据金融危机前后进行时间窗口的划分。由于本文使用的样本区间覆盖了 2008 至 2009 年金融危机时段, 在 2010 年之前, 金融市场波动较大, 政策制定部门也出台了相应的措施, 实体部门不可避免地遭到波及, 这些政策、市场的冲击最终将反映到具体的财务数据上, 因此, 我们将样本划分为 2006 至 2009 年和 2010 年至 2016 年两段, 分别对静态模型和动态模型重新进行估计, 以探究金融危机前后实体部门金融化的驱动特征。静态模型的结果汇报于表 10, 动态模型的结果汇报于表 11。

从表 10 和表 11 的结果来看, 在金融危机之前, 由于本文剔除了晚于 2010 年上市的企业, 故在 2006 年至 2009 年间上市的企业数据有部分出现不连续或波动较大的情形, 加之金融危机造成的市场动荡等因素, 最后导致模型系数估计并未显示出十分显著的结果。静态模型中仅有 *fc*、*asset* 等变量在 10% 的显著性水平显示出统计显著性, 值得注意的是, *fc* 的系数值为负, 表明在这一阶段, 企业即便当期有较充裕的现金流, 也不会用于投资金融资产, 而是减持金融资产, 这一事实反映了当时金融市场的动荡, 同时, 在动态模型中 *fc* 的系数显著为正, 且数值相比金融危机后较大, 表明这一时段内, 企业的金融投资水平与企业的现金流水平密切相关。若从整个时间段内来看, 现金流水平的波动会极大地影响企业持有金融资产的意愿, 这进一步表现出了金融危机时期市场的高度不稳定性。2010 年至 2016 年时段的样本回归结果则与基准模型部分高度一致, 表明实体企业的金融投资行为是由风险规避因素而非资产的回报率缺口所驱动。

表 10 按金融危机前后划分的 GMM 回归结果表

	2009 年之前		2009 年之后	
	模型 (14)	模型 (15)	模型 (14)	模型 (15)
<i>risk</i>	0.0922 (0.0669)	-0.0541 (0.0762)	0.0500** (0.0223)	0.0457** (0.0223)
<i>gap</i>	1.5495 (1.0423)	-0.0036 (0.3647)	-16.6503 (15.1672)	-1.3893 (14.3382)
<i>fc</i>	-1.2021* (0.6162)	0.5208 (0.4487)	1.6048*** (0.3073)	0.7974*** (0.2894)
<i>lev</i>	-0.0133 (0.0139)	0.0014 (0.0097)	-0.0888* (0.0494)	-0.1052** (0.0447)
<i>asset</i>	-0.0563* (0.0324)	-0.0271 (0.0322)	-0.1660*** (0.0455)	-0.1874*** (0.0423)
<i>sales</i>	0.0359	-0.0330	-0.0795* (0.0455)	-0.0674* (0.0423)

	2009年之前		2009年之后	
	模型(14)	模型(15)	模型(14)	模型(15)
	(0.0334)	(0.0274)	(0.0449)	(0.0400)
<i>tangi</i>	-0.0367	-0.0003	-1.0039*	-1.3387***
	(0.1441)	(0.0010)	(0.5486)	(0.4851)
<i>fslack</i>	0.0023*	-0.0161	0.0071***	0.0067***
	(0.0012)	(0.0525)	(0.0016)	(0.0015)
<i>bm</i>	0.1404	-0.2549*	0.2950***	0.3032***
	(0.1454)	(0.1325)	(0.1051)	(0.0890)
<i>share</i>	0.0751	0.2064**	0.6835***	0.7591***
	(0.0991)	(0.0906)	(0.1734)	(0.1664)
<i>constant</i>	1.5726**	0.7569	4.4120***	4.4967***
	(0.7171)	(0.6456)	(1.1824)	(1.0323)
样本数量	8,298	6,821	25,452	24,945
年份虚拟变量	是	是	是	是
LM-p	0.0000	0.0000	0.0000	0.0000
Hausman-p	0.0000	0.0000	0.0000	0.0000
endo-p	0.0000	0.0002	0.0000	0.0000
underid-p	0.0194	0.0252	0.0018	0.0140
HansenJ-p	0.2550	0.4537	0.1150	0.3165

注：上述分类中，模型（14）中内生变量包括 *risk*、*gap*、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*，模型（15）的内生变量对应为模型（14）中各内生变量的滞后一期项。2009年之前，模型（14）中 *tangi*、*fc* 的工具变量选择其滞后二期项，*gap* 的工具变量选择其滞后一到二期项，其余内生变量的工具变量选择其滞后一期项，模型（15）中各内生变量的工具变量设计形式与模型（14）一致；2009年之后，模型（14）中，*risk*、*fc*、*lev* 的工具变量选择其滞后二期项，*gap* 的工具变量选择其滞后二到三期项，*asset*、*sales* 的工具变量选择其滞后三期项，*tangi* 的工具变量选择为其滞后五期项，*fslack*、*bm*、*share* 的工具变量选择为其滞后一期项，模型（15）中各内生变量的工具变量设计形式与模型（14）一致。

表 11 按金融危机前后划分的模型（16）系统 GMM 估计结果

	2009年之前	2009年之后
<i>fk(-1)</i>	0.7091*** (0.0624)	0.4925*** (0.0963)
<i>fk(-2)</i>	0.1455 (0.0947)	0.2426*** (0.0699)
<i>risk</i>	0.1798 (0.1594)	0.0328*** (0.0095)
<i>gap(-1)</i>	-64.5986 (53.8834)	1.7448 (6.5881)

	2009年之前	2009年之后
<i>fc</i>	2.0339** (0.8296)	0.7147*** (0.1012)
<i>lev</i>	-0.0228* (0.0136)	0.0192*** (0.0074)
<i>asset</i>	-0.0110 (0.0583)	-0.0969*** (0.0088)
<i>sales</i>	0.0013 (0.0282)	-0.0158*** (0.0045)
<i>tangi</i>	0.0786 (0.1367)	-0.1397*** (0.0383)
<i>fslack</i>	0.0046*** (0.0017)	0.0007*** (0.0002)
<i>bm</i>	-0.0809** (0.0373)	0.3432*** (0.0525)
<i>share</i>	-0.0258 (0.3013)	0.3140*** (0.0390)
<i>constant</i>	0.1122 (1.1441)	1.9424*** (0.1785)
样本数量	8,540	26,283
年份虚拟变量	是	是
AR(1)-p	0.3556	0.0001
AR(2)-p	0.7119	0.1695
HansenJ-p	0.5876	0.1003

注：上述模型内生变量包括 *fk* 的滞后项、*risk*、*gap* 的滞后一期项、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*。2009 年之前，*fk* 的滞后项的工具变量选择对应的滞后一到二期项，*risk*、*fc*、*lev*、*asset*、*tangi* 的工具变量选择其滞后二期项，*gap* 的滞后一期项的工具变量选择其滞后二到三期项，*sales*、*fslack*、*bm*、*share* 的工具变量选择为其滞后一期项；2009 年之后，*fk* 的滞后项的工具变量选择对应的滞后五到六期项，*risk*、*fc*、*asset*、*tangi*、*fslack*、*bm* 的工具变量选择其滞后三期项，*lev* 的工具变量选择其滞后二到三期项，*gap* 的滞后一期项的工具变量选择其滞后二到三期项，*share* 的工具变量选择为其滞后一期项，*sales* 的工具变量选择其滞后二到三期项。

第四是根据企业所在行业的类型进行划分。本文的实体企业样本是根据中国证券业监督管理委员会颁布的上市公司行业分类指引(2012年10月26日，第31号公告)中除了第J类“金融业”之外的所有企业，其中也包含了与金融市场关联较多、投资收益高且风险相对较低的房地产业。由于房地产行业和其他实体行业的主营业务存在本质的区别，其金融化的驱动特征可能也有所差异，因此我们将样本划分为房地产行业和非房地产实体企业，分别对静态、动态模型重新进行估计。结果汇报于表12和表13中。

结合静态模型和动态模型的结果可见,对于房地产企业,规避固定资产投资风险显然不是推动其金融化的驱动力,注意到静态模型中,模型(14)中收益率缺口 *gap* 项系数显示出了较弱的显著性,表明投资金融资产的超额回报可能是其进行金融投资的推动因素,但动态模型中并未表现出这一特征。控制变量里,在静态模型中 *fc* 的系数均显著为正,表明房地产企业金融资产投资水平主要受到企业现金流水平的影响,然而观察动态模型的结果,我们发现 *tangi*、*bm*、*share* 三个变量的系数均与基准模型符号相反,表明有形资产比率越高,房地产企业将加大金融资产投资,原因可能在于房地产企业通过抵押实物资产获取的贷款资金主要用于投资性房地产,从而使金融资产占比升高;同时,较高的账面市值比意味着企业的资产(房地产)质量较差,故企业减持金融资产;而股权结构较集中的企业相较于股权分散的企业而言面临更大的经营风险,此时企业也有意愿降低对金融资产的持有。总体来说,出于房地产企业的特殊性,其持有金融资产的动机与其他实体企业有较大的差异,而非房地产实体企业的回归结果则与基准模型结果保持一致,此处不再赘述。

表 12 按企业行业类型划分的 GMM 估计结果

	房地产业		非房地产实业	
	模型 (14)	模型 (15)	模型 (14)	模型 (15)
<i>risk</i>	0.0394 (0.0281)	0.0533 (0.0333)	0.0403*** (0.0139)	0.0504*** (0.0194)
<i>gap</i>	19.4651* (11.7058)	-0.5695 (14.4869)	-12.7791 (19.9901)	-17.2201 (46.3423)
<i>fc</i>	0.3607** (0.1448)	0.4034** (0.1886)	0.8350** (0.1843)	1.8737*** (0.3190)
<i>lev</i>	-0.0270* (0.0153)	-0.0277 (0.0217)	-0.0641*** (0.0177)	-0.0814*** (0.0251)
<i>asset</i>	0.0117 (0.0191)	-0.0049 (0.0226)	-0.0693*** (0.0146)	-0.0821*** (0.0194)
<i>sales</i>	0.0377* (0.0218)	0.0382 (0.0287)	-0.0261*** (0.0089)	-0.0233* (0.0121)
<i>tangi</i>	0.1695 (0.1410)	0.0743 (0.1526)	-0.6121*** (0.1761)	-0.8317*** (0.2525)
<i>fslack</i>	0.0005 (0.0005)	0.0002 (0.0004)	0.0075*** (0.0011)	0.0074*** (0.0018)
<i>bm</i>	-0.0555 (0.0593)	-0.0517 (0.0559)	0.1828** (0.0381)	0.2258*** (0.0533)
<i>share</i>	-0.0533 (0.0821)	-0.0509 (0.0893)	0.3869*** (0.0633)	0.4195*** (0.0930)
<i>constant</i>	-0.0340	-0.1664	1.7156***	2.1566***

	房地产业		非房地产实业	
	模型 (14)	模型 (15)	模型 (14)	模型 (15)
	(0.4090)	(0.4288)	(0.3624)	(0.4849)
样本数量	2,172	2,055	27,668	25,885
年份虚拟变量	是	是	是	是
LM-p	0.0000	0.0000	0.0000	0.0000
Hausman-p	0.0381	0.0052	0.0000	0.0000
endo-p	0.0044	0.0016	0.0000	0.0000
underid-p	0.5857	0.5562	0.0404	0.1209
HansenJ-p	0.3470	0.2978	0.3870	0.2155

注：上述分类中，模型（14）中内生变量包括 *risk*、*gap*、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*，模型（15）的内生变量对应为模型（14）中各内生变量的滞后一期项。房地产业样本中，模型（14）中 *gap* 的工具变量选择其滞后一到二期项，*tangi* 的工具变量选择其滞后三期项，其余内生变量的工具变量选择其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致；非房地产实业企业样本中，*risk*、*lev* 的工具变量选择其滞后二期项，*tangi* 的工具变量选择其滞后五期项，*gap* 的工具变量选择其滞后二到三期项，其余内生变量的工具变量选择其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致。

表 13 按企业行业类型划分的模型（16）系统 GMM 估计结果

	房地产业	非房地产实业
<i>fk</i> (-1)	0.6642 <sup>***</sup> (0.2322)	0.4877 <sup>***</sup> (0.0917)
<i>fk</i> (-2)	0.0578 (0.1999)	0.1992 <sup>***</sup> (0.0640)
<i>risk</i>	-0.0075 (0.0288)	0.0208 <sup>**</sup> (0.0089)
<i>gap</i> (-1)	-50.9684 (36.2292)	7.4866 (6.2073)
<i>fc</i>	0.0304 (0.1609)	0.9185 <sup>***</sup> (0.0972)
<i>lev</i>	0.0175 (0.0111)	0.0046 (0.0070)
<i>asset</i>	0.0368 <sup>*</sup> (0.0188)	-0.0940 <sup>***</sup> (0.0059)
<i>sales</i>	-0.0208 <sup>*</sup> (0.0124)	-0.0168 <sup>***</sup> (0.0043)
<i>tangi</i>	0.3441 <sup>***</sup> (0.1257)	-0.1783 <sup>***</sup> (0.0363)

	房地产业	非房地产实业
<i>fslack</i>	-0.0003 (0.0002)	0.0017*** (0.0005)
<i>bm</i>	-0.2723*** (0.1041)	0.2144*** (0.0234)
<i>share</i>	-0.2322** (0.0978)	0.3122*** (0.0311)
<i>constant</i>	-0.5829 (0.3544)	1.9289*** (0.1232)
样本数量	2,294	32,529
年份虚拟变量	是	是
AR(1)-p	0.0003	0.0000
AR(2)-p	0.7039	0.5499
HansenJ-p	0.7837	0.3889

注：上述模型中内生变量包括 *fk* 的滞后项、*risk*、*gap* 的滞后一期项、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*。房地产企业样本中，*fk* 的滞后项的工具变量选择对应的滞后五到六期项，*risk*、*asset*、*tangi* 的工具变量选择为其滞后三期项，*lev* 的工具变量选择其滞后二期项，*gap* 的滞后一期项的工具变量选择其滞后四到五期项，其余内生变量的工具变量选择其滞后一期项；非房地产实业企业样本中，*fk* 的滞后项的工具变量选择对应的滞后五到六期项，*risk*、*fc*、*asset*、*tangi* 的工具变量选择其滞后三期项，*lev* 的工具变量选择其滞后二期项，*gap* 的滞后一期项的工具变量选择其滞后二到三期项，其余内生变量的工具变量选择其滞后一期项。

第五是根据上市年月将企业划分为“年轻”企业和“年长”企业。具体来说，若某企业在某一年 6 月份及之前上市，则忽略月份在时间上的影响，仅用年份来表示上市时间，若企业是在 6 月份之后上市的，则在该年份的基础上加上 0.5 表示上市时间，例如，A 企业上市时间为 1994 年 4 月，则上市时间为 1994 年，B 企业上市时间为 2001 年 9 月，则上市时间为 2001.5 年。然后便可以计算出每个企业在样本时间窗口内每个时点对应的上市年龄，对于晚于某时间点才上市的企业，该时间点及之前的数据视为缺失，直到该企业上市的时间点再从 0 开始计算年龄。得到所有企业的上市年龄后，我们计算各企业在样本时间窗口内的平均年龄，并以该平均年龄的中位数为分界点，小于该分界点的企业归类为“年轻”企业，大于该分界点的归类为“年长”企业。重新回归得到的结果汇报于表 14 和表 15 中。

观察核心变量的系数估计值，我们不难发现结果与根据企业资产规模进行划分的回归结果十分相似：年轻企业和中小企业配置金融资产是由风险规避因素所驱动，而年长企业和大企业则未体现出这一特征，原因在于：企业的年龄和资产规模一般呈正相关关系，随着经营时间增加，企业的规模也趋于扩张。同时，对于上市年龄较短的企业，其融资的渠道多样性和稳定性往往不如上市时间较长的企业，因此持有金融资产主要是为了规避固定资产投资风险，而非出于对收益的追逐。控制变量的系数估计结果也和前文类似，具有很强的稳健性，此处不再赘述。

表 14 根据企业上市年龄划分的 GMM 估计结果

	年轻企业		年长企业	
	模型 (14)	模型 (15)	模型 (14)	模型 (15)
<i>risk</i>	0.0771 <sup>***</sup> (0.0152)	0.0983 <sup>***</sup> (0.0226)	0.0100 (0.0139)	0.0071 (0.0161)
<i>gap</i>	-9.3784 (8.5297)	-12.5909 (9.4533)	-9.3683 (10.5806)	-8.8922 (10.0893)
<i>fc</i>	0.9760 <sup>***</sup> (0.2103)	2.5023 <sup>***</sup> (0.5363)	0.6387 <sup>***</sup> (0.1294)	1.1991 <sup>***</sup> (0.1903)
<i>lev</i>	-0.0440 <sup>***</sup> (0.0108)	-0.0638 <sup>***</sup> (0.0170)	0.0022 (0.0104)	0.0067 (0.0124)
<i>asset</i>	-0.0636 <sup>***</sup> (0.0123)	-0.0812 <sup>***</sup> (0.0179)	-0.0331 <sup>***</sup> (0.0127)	-0.0617 <sup>***</sup> (0.0212)
<i>sales</i>	-0.0371 <sup>***</sup> (0.0105)	-0.0460 <sup>***</sup> (0.0149)	0.0115 (0.0087)	-0.0341 (0.0253)
<i>tangi</i>	-0.3218 <sup>***</sup> (0.0970)	-0.6066 <sup>***</sup> (0.1480)	0.1352 (0.1488)	0.1108 (0.1840)
<i>fslack</i>	0.0078 <sup>***</sup> (0.0009)	0.0078 <sup>***</sup> (0.0011)	0.0018 <sup>***</sup> (0.0007)	0.0012 <sup>*</sup> (0.0007)
<i>bm</i>	0.1624 <sup>***</sup> (0.0260)	0.2305 <sup>***</sup> (0.0416)	0.0087 (0.0417)	0.0249 (0.0538)
<i>share</i>	0.3153 <sup>***</sup> (0.0411)	0.3604 <sup>***</sup> (0.0566)	0.0021 (0.0501)	0.0848 (0.0862)
<i>constant</i>	1.4378 <sup>***</sup> (0.2919)	1.9322 <sup>***</sup> (0.4314)	0.8706 <sup>***</sup> (0.3154)	1.5397 <sup>***</sup> (0.4776)
样本数量	14,886	13,940	15,138	14,169
年份虚拟变量	是	是	是	是
LM-p	0.0000	0.0000	0.0000	0.0000
Hausman-p	0.0000	0.0000	0.0039	0.0000
endo-p	0.0000	0.0000	0.0000	0.0000
underid-p	0.0729	0.0769	0.0815	0.0911
HansenJ-p	0.9733	0.9717	0.9488	0.5111

注：上述分类中，模型（14）中内生变量包括 *risk*、*gap*、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*，模型（15）的内生变量对应为模型（14）中各内生变量的滞后一期项。对于年轻企业样本，模型（14）中 *risk*、*asset* 的工具变量选择其滞后二期项，*gap* 的工具变量选择其滞后二到三期项，*tangi* 的工具变量选择为其滞后三期项，*fc*、*lev*、*sales*、*fslack*、*bm*、*share* 的工具变量选择为其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致；对于年长企业样本，*risk* 的工具变量选择其滞后二期项，*asset* 的工具变量选择其滞后三期项，*tangi* 的工具变量选择其滞后六期项，*gap* 的工具变量选择其滞后二到三期项，其余内生变量



的工具变量选择其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致。

表 15 根据企业上市年龄划分的模型（16）系统 GMM 估计结果

	年轻企业	年长企业
<i>fk</i> (-1)	0.5002*** (0.1046)	0.7141*** (0.0908)
<i>fk</i> (-2)	0.0849 (0.0708)	0.2333*** (0.0753)
<i>risk</i>	0.0294** (0.0126)	-0.0068 (0.0174)
<i>gap</i> (-1)	-12.7022 (11.7378)	13.6606 (11.6483)
<i>fc</i>	0.9223*** (0.1305)	0.9828*** (0.1281)
<i>lev</i>	-0.0545** (0.0247)	-0.0011 (0.0095)
<i>asset</i>	-0.1532*** (0.0167)	-0.0335** (0.0166)
<i>sales</i>	-0.0298*** (0.0074)	-0.0093 (0.0241)
<i>tangi</i>	-0.4675*** (0.1539)	-0.1236** (0.0494)
<i>fslack</i>	0.0024* (0.0014)	0.0007 (0.0005)
<i>bm</i>	0.3809*** (0.0487)	0.1354*** (0.0328)
<i>share</i>	0.4807*** (0.0659)	0.1068*** (0.0395)
<i>constant</i>	0.0000 (0.0000)	0.6479* (0.3610)
样本数量	15,840	18,983
年份虚拟变量	是	是
AR(1)-p	0.0001	0.0020
AR(2)-p	0.5580	0.6239
HansenJ-p	0.1411	0.1110

注：上述模型中内生变量包括 *fk* 的滞后项、*risk*、*gap* 的滞后一期项、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*。年轻企业样本中，*fk* 的滞后项的工具变量选择对应的滞后四到五期项，*risk* 的工具变量选择其滞后二期项，*asset*、*tangi* 的工具变量选择其滞后三期项，*gap* 的滞后一期项的工具变量选择其滞后一到三期项，其余工具变量选择为其滞后一期项；年长企业样本中，*fk* 的滞后项的工具变量选择对应的滞后五到六期项，*risk*、*lev*、*fslack* 的工具变量选择其滞后二期项，*fc*、*sales*、*tangi* 的工具变量选择其滞后三期项，*gap* 的滞后一期项的工具变量选择其滞后二到三期项，其余内生变量的工具变量选择其滞后一期项。

第六是根据企业所在经济区域划分样本重新进行回归分析。根据中国国家统计局于2011年6月13日发布的《东西中部和东北地区划分方法》，可将中国的经济区域划分为东部、中部、西部和东北四大地区，其中，东部包括：北京、天津、河北、上海、江苏、浙江、福建、山东、广东和海南；中部包括：山西、安徽、江西、河南、湖北和湖南；西部包括：内蒙古、广西、重庆、四川、贵州、云南、西藏、陕西、甘肃、青海、宁夏和新疆；东北包括：辽宁、吉林和黑龙江。依据这一标准，我们根据各企业的注册所在省份分别进行地区归类，最后得到东部地区共1183家企业，中部地区共293家企业，西部地区共304家企业，东北地区共122家企业。分别对各个地区的样本重新进行回归分析，得到的结果汇报于表16和表17中。注意，东北地区样本的Hausman检验未能拒绝原假设，故应使用随机效应模型，同时引入工具变量解决内生性问题。

表16中核心变量的结果显示，东部地区和西部地区的企业金融化主要由固定资产投资风险占比所驱动，中部地区则是由经风险调整后的金融资产和固定资产的投资收益率之差所驱动，东北地区的企业则未体现出这些特点。动态模型中则仅有东部地区企业样本的金融投资占比受到风险规避因素的驱动，其他地区的核心变量系数估计结果均未显示出统计显著性。控制变量的系数估计方向与基准模型的结果保持一致，尤其是东部地区样本，除了sales变量之外，其他变量的系数估计方向和显著性基本与基准模型相同，而其他三个地区的系数估计在统计显著性上有不同的表现。值得注意的是，无论是哪个地区的样本，fc和fslack系数均为高度显著正向，表明无论位于哪一个经济区域，企业的金融投资占比均不同程度地受到内部现金流及现金等价物的约束。

表 16 根据企业所处经济区域划分的 GMM 估计结果

	东部地区		中部地区		西部地区		东北地区	
	模型 (14)	模型 (15)	模型 (14)	模型 (15)	模型 (14)	模型 (15)	模型 (14)	模型 (15)
<i>risk</i>	0.1426*** (0.0429)	0.1870*** (0.0574)	0.0286 (0.0258)	0.0306 (0.0289)	0.0635** (0.0264)	0.0811** (0.0396)	-0.0170 (0.0251)	-0.0179 (0.0277)
<i>gap</i>	-58.4624 (55.4244)	-14.4013 (19.8004)	17.1863** (7.4795)	26.3236* (15.7254)	2.6391 (4.0664)	14.3171* (7.7507)	12.2829 (7.6718)	7.3164 (10.1316)
<i>fc</i>	1.1203*** (0.3041)	2.1515*** (0.5119)	0.5981** (0.2465)	1.0126*** (0.3385)	0.2963 (0.2452)	1.4141*** (0.4589)	0.7517*** (0.2787)	1.1619** (0.4823)
<i>lev</i>	-0.1102** (0.0516)	-0.1537** (0.0725)	-0.0182 (0.0122)	-0.0239* (0.0132)	-0.0019 (0.0102)	-0.0271 (0.0173)	-0.0047 (0.0076)	-0.0139 (0.0091)
<i>asset</i>	-0.1270** (0.0543)	-0.1695** (0.0696)	-0.0311 (0.0194)	-0.0316 (0.0212)	-0.0350** (0.0144)	-0.0509** (0.0219)	-0.0226** (0.0107)	-0.0240** (0.0107)
<i>sales</i>	0.0053 (0.0485)	0.0130 (0.0495)	0.0054 (0.0192)	0.0212 (0.0197)	-0.0223 (0.0345)	0.0371 (0.0372)	-0.0121 (0.0190)	-0.0279 (0.0186)
<i>tangi</i>	-1.1939** (0.6066)	-1.7512** (0.8275)	-0.2228* (0.1277)	-0.3253** (0.1468)	0.1109 (0.1090)	-0.2119 (0.2487)	0.0783 (0.0842)	-0.0129 (0.0836)
<i>fslack</i>	0.0054*** (0.0016)	0.0067*** (0.0022)	0.0053*** (0.0020)	0.0045** (0.0020)	0.0035*** (0.0012)	0.0027** (0.0013)	0.0032*** (0.0008)	0.0033*** (0.0009)
<i>bm</i>	0.2848** (0.1249)	0.3850** (0.1602)	0.0889** (0.0422)	0.1127** (0.0495)	0.0499 (0.0411)	0.1344* (0.0792)	-0.0188 (0.0398)	-0.0251 (0.0395)

	东部地区		中部地区		西部地区		东北地区	
	模型 (14)	模型 (15)	模型 (14)	模型 (15)	模型 (14)	模型 (15)	模型 (14)	模型 (15)
<i>share</i>	0.5765 <sup>***</sup>	0.7130 <sup>***</sup>	0.2081 <sup>***</sup>	0.2182 <sup>***</sup>	0.0544	0.1310	0.1005 <sup>*</sup>	0.0585
	(0.1675)	(0.2278)	(0.0713)	(0.0811)	(0.0621)	(0.1151)	(0.0597)	(0.0597)
<i>constant</i>	3.0169 <sup>**</sup>	3.6677 <sup>**</sup>	0.7935 <sup>*</sup>	0.8455 <sup>*</sup>	0.7993 <sup>***</sup>	1.2358 <sup>**</sup>	0.5737 <sup>**</sup>	0.7114 <sup>***</sup>
	(1.3570)	(1.8042)	(0.4474)	(.4807)	(0.3093)	(0.5060)	(0.2335)	(0.2388)
样本数量	18,119	16,937	4,919	4,626	5,124	4,541	2,166	2,043
年份虚拟变量	是	是	是	是	是	是	是	是
LM-p	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Hausman-p	0.0000	0.0000	0.0065	0.0021	0.0001	0.0038	0.5696	0.5818
endo-p	0.0000	0.0000	0.0000	0.0000	0.0029	0.0000	0.0000	0.0000
underid-p	0.0277	0.0227	0.0823	0.1492	0.0441	0.0009	0.0000	0.0000
HansenJ-p	0.2556	0.3701	0.9546	0.9685	0.4649	0.8650	0.1742	0.9468

注：上述分类中，模型（14）中内生变量包括 *risk*、*gap*、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*，模型（15）的内生变量对应为模型（14）中各内生变量的滞后一期项。对于东部地区企业样本，模型（14）中 *risk*、*sales* 的工具变量选择其滞后三期项，*gap* 的工具变量选择其滞后二到三期项，*tangi* 的工具变量选择为其滞后五期项，其余工具变量选择为其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致；对于中部地区企业样本，*risk* 的工具变量选择其滞后二期项，*asset* 的工具变量选择其滞后三期项，*tangi* 的工具变量选择其滞后四期项，*gap* 的工具变量选择其滞后一到二期项，其余内生变量的工具变量选择其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致；对于西部地区企业样本，*risk* 的工具变量选择其滞后三期项，*asset*、*sales* 的工具变量选择其滞后三期项，*tangi* 的工具变量选择其滞后四期项，*gap* 的工具变量选择其滞后二到三期项，其余内生变量的工具变量选择其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致；对于东北地区企业样本，*risk* 的工具变量选择其滞后二期项，*tangi* 的工具变量选择其滞后三期项，*gap* 的工具变量选择其滞后二到三期项，其余内生变量的工具变量选择其滞后一期项；模型（15）中各内生变量的工具变量设计形式与模型（14）一致。

表 17 根据企业所处经济区域划分的模型 (16) 系统 GMM 估计结果

	东部地区	中部地区	西部地区	东北地区
<i>fk</i> (-1)	0.4881*** (0.0742)	0.2069** (0.0974)	0.5547** (0.2371)	0.3437** (0.1630)
<i>fk</i> (-2)	0.2195*** (0.0573)	0.2436*** (0.0828)	0.2060 (0.1749)	0.4434*** (0.1329)
<i>risk</i>	0.0212* (0.0110)	0.0196 (0.0201)	0.0163 (0.0316)	0.0245 (0.0505)
<i>gap</i> (-1)	11.0763 (7.9279)	-5.9345 (4.8117)	1.0684 (3.0427)	6.3771 (11.0592)
<i>fc</i>	0.7738*** (0.1052)	0.7084*** (0.1725)	0.9541*** (0.2527)	0.8754* (0.5230)
<i>lev</i>	0.0042 (0.0163)	0.0122 (0.0176)	0.0111 (0.0247)	-0.0040 (0.0152)
<i>asset</i>	-0.1043*** (0.0106)	-0.0551*** (0.0111)	-0.0813*** (0.0239)	-0.0586*** (0.0193)
<i>sales</i>	-0.0220*** (0.0061)	-0.0138* (0.0081)	0.0102 (0.0193)	0.0138 (0.0392)
<i>tangi</i>	-0.2627*** (0.0756)	0.0333 (0.0944)	-0.0951 (0.1114)	-0.1273 (0.1473)
<i>fslack</i>	0.0018*** (0.0007)	0.0036*** (0.0013)	0.0013 (0.0023)	0.0000 (0.0009)
<i>bm</i>	0.2185*** (0.0390)	0.0810 (0.0504)	0.2586 (0.1849)	0.1809** (0.0883)
<i>share</i>	0.3521*** (0.0463)	0.1625** (0.0665)	0.1589*** (0.0545)	0.1762 (0.1072)
<i>constant</i>	2.1815*** (0.2112)	1.0989*** (0.2057)	0.0000 (0.0000)	0.0000 (0.0000)
样本数量	21,301	5,490	5,736	2,296
年份虚拟变量	是	是	是	是
AR(1)-p	0.0007	0.0012	0.0463	0.0483
AR(2)-p	0.1965	0.1062	0.8149	0.2016
HansenJ-p	0.1351	0.1221	0.1110	0.1339

注：上述模型中内生变量包括 *fk* 的滞后项、*risk*、*gap* 的滞后一期项、*fc*、*lev*、*asset*、*sale*、*tangi*、*fslack*、*bm* 以及 *share*。东部地区企业样本中，*fk* 的滞后项的工具变量选择对应的滞后四到五期项，*risk*、*lev*、*asset*、*tangi* 的工具变量选择其滞后三期项，*gap* 的滞后一期项的工具变量选择其滞后二到四期项，其余工具变量选择为其滞后一期项；中部地区企业样本中，*fk* 的滞后项的工具变量选择对应的滞后四到五期项，*risk* 的工具变量选择其滞后二期项，*asset*、*lev*、*tangi* 的工具变量选择其滞后三期项，*gap* 的滞后一期项的工具变量选择其滞后一到二期项，其余内生变量的工具变量选择其滞后一期项；西部地区企业样本中，*fk* 的滞后项的工具变量选择对应的滞后四到五期项，*risk* 的工具变量选择其滞后二期项，*fc*、*asset*、*lev*、*bm* 的工具选择其滞后三期项，*tangi* 的工具变量选择其滞后五期项，*gap* 的滞后一期项的工具变量选择其滞后二到三期项，其余内生变量的工具变

量选择其滞后一期项；东北地区企业样本中， $fk$  的滞后项的工具变量选择对应的滞后四到五期项， $risk$  的工具变量选择其滞后二期项， $fc$ 、 $sales$ 、 $fslack$  的工具选择其滞后三期项， $tangi$  的工具变量选择其滞后五期项， $gap$  的滞后一期项的工具变量选择其滞后二到三期项，其余内生变量的工具变量选择其滞后一期项。

## 五、结论

随着现代企业经营理念的不断发展，特别是股东价值论的兴起，非金融企业不仅面临如何进行实业投资的问题，而且越来越多地面临金融资产投资问题。然而，毕竟非金融企业的主营业务并非金融投资，因此非金融企业从事金融投资的行为也日益受到各界普遍关注。同时，“脱实就虚”、“不务正业”、“资金空转”等媒体词汇生动地表达了包括国家政策制定部门对非金融企业从事金融资产投资的密切关注，从一定程度上也折射出各界对非金融企业进行金融投资会带来负面影响的印象，而这一印象似乎又被近年来的一系列学术研究所印证。然而截至目前，尚无相关研究深入分析过企业进行金融资产投资的原因，传统印象和大多数人的主观判断认为这是由于金融投资更容易“赚钱”。

事实果真如此吗？要回答这一问题，就亟需厘清中国非金融企业从事金融投资的驱动机制，通过严谨的理论逻辑和实证分析，掌握究竟什么因素影响了微观企业的金融投资行为，这样各项政策才能有的放矢、宏观调控才有针对性。为此，本文构建了中国非金融企业的金融投资行为机制模型。我们基于微观企业面对实业投资和金融投资两大类资产的投资组合选择背景，拓展已有研究关于假设微观企业投资组合模型中金融投资无风险的假设，同时考虑金融投资风险和固定投资风险，推演出与现实情况更加贴近的理论模型。

本文证明，拓展模型与传统研究的模型机制存在本质区别，而这种区别可以带来实证分析结果的迥然不同。我们进一步基于中国 A 股非金融上市公司的面板数据进行实证检验，结果与以往的主观判断并不相同。实证结果表明：中国非金融企业的金融投资行为，主要受到固定资产投资的风险占比驱动，而金融资产与固定资产投资回报率缺口对其没有显著影响，稳健性检验的结果也支持这一结论。因此，金融投资更容易“赚钱”而导致实体企业“脱实向虚”这种传统印象和主观判断并非真实情况。

从政策层面看，本文的经济涵义也非常明确：产业政策和宏观政策都需要更多关注微观企业的实业投资环境，降低实业投资的不确定性（风险）是完善企业投资决策行为、避免过度金融投资的对症良方。如果仅从对企业金融投资的种类、模式和范围进行限制，可能不仅会抑制我国现代化金融市场的正常发展，而且也无法有效引导非金融企业在实业投资和金融投资上实现合理均衡。

### 参考文献

- 杜勇、张欢、陈建英(2017):《金融化对实体企业未来主业发展的影响:促进还是抑制》,《中国工业经济》第12期。
- 韩珣、田光宁、李建军(2017):《非金融企业影子银行化与融资结构——中国上市公司的经验证据》,《国际金融研究》第10期。
- 江春、李巍(2013):《中国非金融企业持有金融资产的决策因素和含义:一个实证调查》,《经济管理》第11期。
- 刘笃池、贺玉平、王曦(2016):《企业金融化对实体企业生产效率的影响研究》,《上海经济研究》第8期。
- 罗来军、蒋承、王亚章(2016):《融资歧视、市场扭曲与利润迷失——兼议虚拟经济对实体经济的影响》,《经济研究》第4期。
- 宋军、陆旸(2016):《非货币金融资产和经营收益率的U形关系——来自我国上市非金融公司的金融化证据》,《金融研究》第6期。
- 王红建、曹瑜强、杨庆、杨笋(2017):《实体企业金融化促进还是抑制了企业创新——基于中国制造业上市公司的经验研究》,《南开管理评论》第1期。
- 王红建、李茫茫、汤泰劼(2016):《实体企业跨行业套利的驱动因素及其对创新的影响》,《中国工业经济》第11期。
- 王永钦、刘紫寒、李嫦、杜巨澜(2015):《识别中国非金融企业的影子银行活动——来自合并资产负债表的证据》,《管理世界》第12期。
- 吴军、陈丽萍(2018):《非金融企业金融化程度与杠杆率变动的关系——来自A股上市公司和发债非上市公司的证据》,《金融论坛》第1期。
- 谢家智、王文涛、江源(2014):《制造业金融化、政府控制与技术创新》,《经济学动态》第11期。
- 徐光伟、孙峥(2015):《货币政策信号、实际干预与企业投资行为》,《财经研究》第7期。
- 徐军辉(2013):《中国式影子银行的发展及其对中小企业融资的影响》,《财经科学》第2期。
- 张成思、张步昙(2016):《中国实业投资率下降之谜:经济金融化视角》,《经济研究》第12期。
- Arellano, M. and Bover O. "Another Look at the Instrumental Variable Estimation of Error-Components Models." *Journal of Econometrics*, 1995, 68(1), pp. 29-51.
- Breusch, T. S. and Pagan, A. R. "The Lagrange multiplier test and its applications to model specification in econometrics." *Review of Economic Studies*, 1980, 47, pp. 239-253.
- Demir, F. "Financial Liberalization, Private Investment and Portfolio Choice: Financialization of Real Sectors in Emerging Markets." *Journal of Development Economics*, 2009, 88(2), pp. 314-324.
- Durbin, J. "Errors in Variables." *Review of the International Statistical Institute*, 1954, 22(1), pp. 23-32.
- Foster, J. "The financialization of capitalism." *Monthly Review*, 2007, 58(11), pp. 1-14.
- Froud, J. ; Haslam, C. ; Johal, S. and Williams, K. "Shareholder value and financialization:

- consultancy promises, management moves.” *Economy and Society*, 2000, 29(1), pp. 80-120.
- Hansen, L. P. “Large Sample Properties of Generalized Method of Moments Estimators.” *Econometrica*, 1982, 50, pp. 1029-1054.
- Hausman, J. “Specification Tests in Econometrics.” *Econometrica*, 1978, 46(4), pp. 1251-1272.
- Hayashi, F. *Econometrics*. Princeton University Press: Princeton, 2000.
- Hovakimian, G. “Determinants of Investment Cash Flow Sensitivity.” *Financial Management*, 2009, 38 (1), pp. 161-183.
- Huang, C. and Litzenberger, R. H. *Foundations for Financial Economics*. Netherlands: North-Holland, 1973.
- Kleibergen, F. and Paap R. “Generalized Reduced Rank Rests using the Singular Value Decomposition.” *Journal of Econometrics*, 2006, 133(1), pp. 97-126.
- Krippner, G. “The financialization of the American economy.” *Socio-Economic Review*, 2005, 3(2), pp. 173-208.
- Lashitew, A. A. “The Uneven Effect of Financial Constraints: Size, Public Ownership, and Firm Investment in Ethiopia.” *World Development*, 2017, 97, pp. 178-198.
- Le, Q. V. and Zak P. J. “Political Risk and Capital Flight.” *Journal of International Money and Finance*, 2006, 25(2), pp. 308-329.
- Orhangazi, Ö. “Financialisation and capital accumulation in the non-financial corporate sector: A theoretical and empirical investigation on the US economy: 1973–2003.” *Cambridge Journal of Economics*, 2008, 32(6), pp. 863–886.
- Sargan, J. D. “The Estimation of Economic Relationships using Instrumental Variables.” *Econometrica*, 1958, 26, pp. 393-415.
- Shi, J. and Zhang, X. “How to explain corporate investment heterogeneity in China's new normal: Structural models with state-owned property rights.” *China Economic Review*, 2018, 50, pp. 1-16.
- Stockhammer, E. and Grafl, L. “Financial uncertainty and business investment.” *Review of Political Economy*, 2010, 22(4), pp. 551-568.
- Tobin, J. “Money and Economic Growth.” *Econometrica*, 1965, 33(4), pp. 671-684.
- Tornell, A. “Real vs. Financial Investment: Can Tobin Taxes Eliminate the Irreversibility Distortion?” *Journal of Development Economics*, 1990, 32(3), pp. 419-444.
- Wu, D. “Alternative Tests of Independence between Stochastic Regressions and Disturbances.” *Econometrica*, 1973, 41(2), pp. 733-750.

#### 附录：两资产跨期投资组合选择模型详细推导过程

为了探究中国实体企业的金融资产投资问题，我们利用投资组合选择模型来构建基础框架，同时为了简化表述过程，我们首先设定几个基本假设：

假设 1. 市场参与主体为代表性企业，其  $t$  期的效用函数  $U(W_t)$  是关于  $t$  期总财富  $W_t$  的增函数，且为二阶可导的连续凹函数；

假设 2. 假定企业个体的资产可以划分为两类资产：金融资产  $I_t^f$  和固定资产  $I_t^k$ ，二者之和



为  $t$  期的总投资资本  $K_t^\alpha$ ;

假设 3. 设  $t$  期金融资产的投资回报率为  $r_t^f$ , 固定资产的投资回报率为  $r_t^k$ 。

基于以上三个基本假设, 我们首先对 Demir (2009) 的理论推导进行简要概括说明, 之后在其基础上放松不合理的假设, 从而推导出更一般的理论表达式。为了得到更为简化的解析表达式, Demir (2009) 在上述三个假设的基础上增加以下几条基本设定: 一是假设金融资产无风险, 此时假设 3 中的金融资产投资回报率可简化记作  $r^f$ , 表明该回报率不随时间波动; 二是假设固定资产投资回报率  $r_t^k$  服从正态分布:  $r_t^k \sim N(\mu, \sigma^2)$ ; 三是进一步假定企业在  $t$  期的初始财富  $W_0$  提供了投资于固定资产和金融资产的资本; 四是假设市场上企业的总数量恒定不变, 且标准化为 1。

为简化表述, 下文中我们使用  $E(\cdot)$  和  $Var(\cdot)$  分别表示变量的期望和方差。

根据以上设定, Demir (2009) 将两资产跨期投资选择问题概括为以下优化问题:

$$\text{Max } E \sum_{t=0}^{\infty} \beta^t U(W_t) \quad (1)$$

其中  $\beta$  为恒定折现因子。约束条件为

$$W_t = (1 + r_t^k)I_t^k + (1 + r^f)I_t^f \quad (2)$$

将初始资本  $W_0 = I_t^f + I_t^k$  代入等式 (2), 进一步整理得到

$$W_t = W_0(1 + r^f) + I_t^k(r_t^k - r^f) \quad (3)$$

利用等式(1)和(3), 根据斯坦因引理(Stein's Lemma), 也即  $\text{cov}(f(x), y) = E(f'(x))\text{cov}(x, y)$ ,

Demir (2009) 解出企业的最优固定资产配置表达式如下:

$$I_t^k = \frac{E(r_t^k - r^f)}{\gamma \text{Var}(r_t^k)} \quad (4)$$

其中  $\gamma$  是效用函数的二阶导数与一阶导数比值的负值, 代表企业恒定不变的风险厌恶水平。

注意到  $t$  期投资于市场上的总资本  $K_t^\alpha$  为企业个体的固定资产和金融资产之和再乘以企业的总数量, 由于企业总数量标准化为 1, 故可省略, 则:

$$K_t^\alpha = I_t^k + I_t^f \quad (5)$$

将 (5) 代入 (4) 中即得到最优金融资产投资的表达式:

$$I_t^f = K_t^\alpha - \frac{E(r_t^k - r^f)}{\gamma \text{Var}(r_t^k)} \quad (6)$$

在等式 (6) 两侧同时除以总资本  $K_t^\alpha$ , 得到的就是市场整体金融化的代理变量——金融投资占比的决定因素表达式:

$$\frac{I_t^f}{K_t^\alpha} = 1 - \frac{1}{K_t^\alpha} \times \frac{E(r_t^k - r^f)}{\gamma \text{Var}(r_t^k)} \quad (7)$$

Demir (2009) 声称: 进一步对等式 (7) 两侧取自然对数, 可以近似地得到企业投资影响机制更为显性的表达式:

$$\ln\left(\frac{I_t^f}{K_t^\alpha}\right) = -\ln(E(r_t^k - r^f)) + \ln(\gamma) + \ln(\text{Var}(r_t^k)) + \ln(K_t^\alpha) \quad (8)$$

据此, Demir (2009) 中设立的实证模型可以概括为以下形式:

$$\left(\frac{I^f}{K^\alpha}\right)_{it} = \alpha_1 Rgap_{it}^{k-f} + \alpha_2 Risk_{it}^k + \text{控制变量} + \text{随机扰动项} \quad (9)$$

等式 (9) 表明企业的金融投资行为受投资收益率差  $Rgap_{it}^{k-f}$  (上标  $k-f$  表示固定资产投资回报率减去金融投资回报率) 和固定资产投资面对的风险水平  $Risk_{it}^k$  影响。

然而, 即使不考虑 Demir 理论模型推导过程存在的问题, 金融投资回报率恒定且金融投资无风险这一假设也显然与事实情况并不相符。企业的金融投资不仅仅限于国债类无风险产品, 还包括诸如股票、期货等各种有风险资产。因此, 在企业投资组合选择模型的设立过程中, 必须对金融资产投资的风险给予考虑, 同时需要放松金融投资回报率恒定的假设, 注意此时金融资产回报率符号下标此时就需要有时间  $t$ 。由此可见, Demir (2009) 的理论推导不可避免地存在一定的问题, 其最终的模型表述和实际理论表达式也不一致, 因此其结果的可靠性存疑。为此, 本文基于前文的假设 1 至 3, 增加以下基本设定:

假设 4. 企业个体的效用函数为标准常系数绝对风险厌恶效用函数 (Constant Absolute Risk Aversion Utility Function, 下文中缩写为 CARA):

$$U(W_t) = -e^{-\gamma W_t} \quad (10)$$

其中  $e$  是自然底, 其他变量的符号及意义均沿用 Demir (2009) 的设计。

则结合式 (1)、(2)、(10) 我们可以重新书写企业的效用函数:

$$U(W_t) = -e^{-\gamma[(1+r_t^f)I_t^f + (1+r_t^k)I_t^k]} \quad (11)$$

由于我们允许金融资产投资存在风险, 故此时两类资产投资回报率的相关性将影响投资组合的收益和风险, 故接下来我们根据金融资产投资回报率和固定资产投资回报率是否相关两种情形进行分类讨论, 力求推导出企业投资行为更一般的理论表达式:

### (一) 金融资产投资回报率和固定资产投资回报率不相关的情形

假设固定和金融资产投资的回报率均服从正态分布:

$$\begin{cases} r_t^f \sim N(E(r_t^f), \text{Var}(r_t^f)) \\ r_t^k \sim N(E(r_t^k), \text{Var}(r_t^k)) \end{cases} \quad (12)$$

则式 (11) 中的指数部分实际上是两个独立的随机变量的线性组合, 这样的组合服从正态分布, 同时由于两个资产投资回报率不相关, 此时我们可以写出式 (11) 指数部分的正态分布形式:

$$-\gamma[(1+r_t^f)I_t^f + (1+r_t^k)I_t^k] \sim N\left(-\gamma\left[\left(1+E(r_t^f)\right)I_t^f + \left(1+E(r_t^k)\right)I_t^k\right], \gamma^2\text{Var}(r_t^f)(I_t^f)^2 + \gamma^2\text{Var}(r_t^k)(I_t^k)^2\right) \quad (13)$$

进一步对式(13)求数学期望可得到:

$$E[U(W_t)] = -e^{-\gamma[(1+E(r_t^f))I_t^f + (1+E(r_t^k))I_t^k] + \frac{1}{2}[\gamma^2 \text{Var}(r_t^f)(I_t^f)^2 + \gamma^2 \text{Var}(r_t^k)(I_t^k)^2]} \quad (14)$$

因为CARA是关于 $t$ 期总财富 $W_t$ 的单调递增函数:

$$U'(W_t) = \gamma e^{-\gamma W_t} > 0, \forall t$$

且CARA符号恒为负值,取期望之后符号不变,因此求解多期的总期望效用的优化问题可转化单独求解每一期的期望效用优化问题,之后进行多期加总。在这样的条件下,式(1)可简化为以下优化问题:

$$\text{Max} \beta^t E[U(W_t)]$$

再结合(5)、(14),可整理为求解下述优化问题:

$$\text{Max} [(1+E(r_t^f))I_t^f + (1+E(r_t^k))(K_t^a - I_t^f)] - \frac{1}{2} \gamma [\text{Var}(r_t^f)(I_t^f)^2 + \text{Var}(r_t^k)(K_t^a - I_t^f)^2] \quad (15)$$

对 $I_t^f$ 求导得到一阶条件,并令其为0,得到:

$$E[r_t^f - r_t^k] = \alpha [\text{Var}(r_t^f)I_t^f - \text{Var}(r_t^k)(K_t^a - I_t^f)] \quad (16)$$

整理得到企业金融资产投资的最优配置行为等式:

$$\frac{I_t^f}{K_t^a} = \frac{\text{Var}(r_t^k)}{\text{Var}(r_t^f) + \text{Var}(r_t^k)} + \frac{E[r_t^f - r_t^k]}{\alpha K_t^a [\text{Var}(r_t^f) + \text{Var}(r_t^k)]} \quad (17)$$

相应的实证模型则可以概括为:

$$\left(\frac{I_t^f}{K_t^a}\right) = \beta_1 \left(\frac{\text{Rgap}^{f-k}}{\text{total risk}}\right) + \beta_2 \left(\frac{\text{Risk}^k}{\text{total risk}}\right) + \text{控制变量} + \text{随机扰动项} \quad (18)$$

其中 $\beta$ 表示系数。显然,虽然在优化过程中仅仅增加了金融资产投资的风险项,但是推导出来的结果却与之前的结果在涵义上大相径庭(对比等式(17)与(7)或等式(18)与(9)):首先,实体企业的金融投资行为并不是单纯地由固定资产投资的风险影响,而是固定资产投资的风险占比;其次,金融投资行为也不是单纯由两种投资回报率之差影响,而是经过总风险调整后的投资收益率之差。

## (二) 金融资产投资回报率和固定资产投资回报率相关的情形

假设金融资产和固定资产均为风险资产,二者的收益率服从二维联合正态分布:

$$\begin{pmatrix} r_t^f \\ r_t^k \end{pmatrix} \sim N \left( \begin{pmatrix} E(r_t^f) \\ E(r_t^k) \end{pmatrix}, \begin{pmatrix} \text{Var}(r_t^f) & \rho \sqrt{\text{Var}(r_t^f) \cdot \text{Var}(r_t^k)} \\ \rho \sqrt{\text{Var}(r_t^f) \cdot \text{Var}(r_t^k)} & \text{Var}(r_t^k) \end{pmatrix} \right) \quad (19)$$

其中 $\rho$ 为两类资产回报率的相关系数。

此时根据二维联合正态分布的性质可知,式(11)中指数部分中两个资产投资的回报率的线性组合同样也服从正态分布,即:

$$\begin{aligned}
& -\gamma \left[ (1+r_t^f)I_t^f + (1+r_t^k)I_t^k \right] \sim \\
& N \left( -\gamma \left[ (1+E(r_t^f))I_t^f + (1+E(r_t^k))I_t^k \right], \gamma^2 \left[ \text{Var}(r_t^f)(I_t^f)^2 + 2\rho I_t^f I_t^k \sqrt{\text{Var}(r_t^f) \cdot \text{Var}(r_t^k)} + \text{Var}(r_t^k)(I_t^k)^2 \right] \right) \\
(20)
\end{aligned}$$

据此对式(11)求数学期望可得到:

$$E[U(W_t)] = -e^{-\gamma \left[ (1+E(r_t^f))I_t^f + (1+E(r_t^k))I_t^k \right] + \frac{1}{2}\gamma^2 \left[ \text{Var}(r_t^f)(I_t^f)^2 + 2\rho I_t^f I_t^k \sqrt{\text{Var}(r_t^f) \cdot \text{Var}(r_t^k)} + \text{Var}(r_t^k)(I_t^k)^2 \right]} \quad (21)$$

结合(5)、(21),与情形(一)同理,优化问题式(1)可简化为求解下述优化问题:

$$\begin{aligned}
& \text{Max} \left[ (1+E(r_t^f))I_t^f + (1+E(r_t^k))I_t^k (K_t^\alpha - I_t^f) \right] - \\
& \frac{1}{2}\gamma \left[ \text{Var}(r_t^f)(I_t^f)^2 + 2\rho I_t^f (K_t^\alpha - I_t^f) \sqrt{\text{Var}(r_t^f) \cdot \text{Var}(r_t^k)} + \text{Var}(r_t^k) (K_t^\alpha - I_t^f)^2 \right] \\
(22)
\end{aligned}$$

对 $I_t^f$ 求导得到一阶条件并令其为0,可以得到:

$$\begin{aligned}
E[r_t^f - r_t^k] &= \gamma \left[ \text{Var}(r_t^f)I_t^f + \rho K_t^\alpha \sqrt{\text{Var}(r_t^f) \cdot \text{Var}(r_t^k)} \right. \\
& \left. - 2\rho I_t^f \sqrt{\text{Var}(r_t^f) \cdot \text{Var}(r_t^k)} - \text{Var}(r_t^k)(K_t^\alpha - I_t^f) \right] \\
(23)
\end{aligned}$$

整理即得到更一般的企业最优金融资产配置的表达式:

$$\begin{aligned}
\frac{I_t^f}{K_t^\alpha} &= \frac{\sqrt{\text{Var}(r_t^k)} \left( \sqrt{\text{Var}(r_t^k)} - \rho \sqrt{\text{Var}(r_t^f)} \right)}{\text{Var}(r_t^f) - 2\rho \sqrt{\text{Var}(r_t^f) \cdot \text{Var}(r_t^k)} + \text{Var}(r_t^k)} \\
&+ \frac{E[r_t^f - r_t^k]}{\gamma K_t^\alpha \left[ \text{Var}(r_t^f) - 2\rho \sqrt{\text{Var}(r_t^f) \cdot \text{Var}(r_t^k)} + \text{Var}(r_t^k) \right]} \\
(24)
\end{aligned}$$

需要特别注意的是,式(17)是考虑了金融资产和固定资产收益率间的相关性而得出的,实际上当 $\rho=0$ 时,式(24)即可简化式(17)。

# 散户投资行为性别差异的计量分析——基于最新调查数据

孟勇<sup>1</sup>

(山西财经大学统计学院)

代志新<sup>2,3</sup>

(中国人民大学财政金融学院, 中国财政金融政策研究中心, 100872)

**摘要:** 虽然投资者的性别差异的重要性已经引起学界的很多讨论, 但是关于中国散户投资的性别差异和人口特征的研究还存在空白。本文在我国 17 个不同的省份随机访问了 2129 位散户投资者来尝试对此作出回答。研究发现: (1) 96.33% 的散户投资者年龄在 60 岁以下, 只有 11.55% 的是退休人员或学生, 这与我国的老年女性投资者 (60 岁以上) 和大学毕业生的投资案例形成了鲜明的对比; (2) 投资者通常受过良好的教育 (72.03% 的人拥有高于中等教育的学位); (3) 投资者的投资规模大多低于 10 万元 (88.09%), 投资/收入比率低于 0.5 (95.4%), 这意味着散户投资者对金融市场的影响可能被夸大了。进一步, 研究发现投资者在情绪上的一个强有力的性别差异: 与女性散户投资者相比, 男性散户投资者更理性、更容易承受损失, 并且是乐观的。

**关键词:** 我国; 性别差异; 散户投资者; 情绪; 调查

**JEL 分类号:** G11, J16

## 一、引言

我国的金融市场不仅是我国经济最重要的组成部分, 而且在世界经济中影响力也越来越大。我国金融市场最重要的特征之一就是散户投资者的高参与率。与世界上许多股市不同的是, 我国股市的大部分交易都是由散户投资者(而非机构投资者)进行的。

---

本文受到国家社会科学基金“关于主观资产组合模型的行为金融理论建构与方法研究”(11BJY013)支持,

作者简介: 孟勇, 男, 山西偏关人, 经济学博士, 教授, 研究方向: 数量经济学。山西省太原市坞城路 696 号山西财经大学统计学院, [m7025y@163.com](mailto:m7025y@163.com), [13603516019](http://13603516019)

代志新, 男, 经济学博士, 助理教授, 研究方向: 行为经济学, 实验经济学。北京市中关村大街 59 号中国人民大学财政金融学院, [dzxruc@ruc.edu.cn](mailto:dzxruc@ruc.edu.cn)。

尽管散户投资者在我国金融市场具有一定的重要性,但在文献中仍有一些显著的分歧。例如,文献通常侧重于散户投资者的投资偏差,比如周末效应<sup>[1]</sup>,过于自信<sup>[2]</sup>,本地/家庭的偏见<sup>[3]</sup>,也倾向于散户投资者关注驱动<sup>[4]</sup>和处置效应<sup>[5]</sup>,然而几乎没有证据显示散户投资者的人口统计特征是什么。

本文认为,对散户投资者的形象进行描述是很重要的。即,他们是谁及他们说什么,这些都可以提供重要的信息来预测其交易行为(De Bondt 在 1998 年也研究过美国散户的人口统计特征,但样本量只有 45 个散户)<sup>[6]</sup>。因此,应该在真实和异构的主题中收集更多的现场数据。这些数据的收集应该是对实验室收集的证据和真实交易数据的补充(而不是替代)。

本文感兴趣的第二个问题是,在自我报告的投资者情绪方面是否存在性别差异,比如对未来股票市场的理性、损失容忍度和乐观程度(见 Bandopadhyaya 和 Jones 在 2006 年对相关研究的调查)<sup>[7][8][9][10]</sup>。一方面,投资者情绪可能在金融市场中发挥重要作用<sup>[11]</sup>。另一方面,尽管性别差异是有据可查的(对性别差异的偏好进行回顾见 Croson 和 Gneezy 在 2009 的论文)<sup>[12]</sup>,但新兴市场投资者情绪是否存在性别差异的问题却鲜为人知(Dwyer et al.(2002)对美国近 2000 名共同基金投资者的调查发现,女性在共同基金投资决策中所表现出的风险要比男性低)<sup>[13]</sup>。

为了解决上述两个问题,在 2015 年 2 月至 2015 年 3 月随机访问了我国 17 个不同省份的 2129 名散户投资者。研究的数据显示,(1)96.33%的散户投资者年龄在 60 岁以下,只有 11.55%的人是退休人员或学生,这与我国的老年女性投资者(60 岁以上)和大学毕业生的投资案例形成了鲜明的对比;(2)他们通常受过良好的教育(72.03%的人拥有高于中等教育的学位);(3)他们的投资规模大多低于 10 万元(88.09%),投资/收入比率低于 0.5(95.4%),这意味着他们对金融市场的影响可能被夸大了。研究也发现了一个稳健的性别效应:男性散户投资者(1)更理性,(2)更乐观,(3)更容易承受损失。

本文的其余部分内容如下:在第 2 节中,介绍方法和数据。在第 3 节,给出研究结果。最后,在第 4 节,进行讨论和总结。

## 二、方法和数据

为了收集散户投资者的代表性样本,在 2015 年 2 月和 2015 年 3 月,本文在我国 17 个不同省份随机访问了 2129 个研究对象,调查由四部分组成:第一部分的目的是区分被调查对象是否为散户投资者;第二部分关注的是散户投资者参与股票市场的原因以及他们是如何做出决定的;第三部分研究散户投资者如何根据假设情况做出主观判断;最后一部分提出了一些传统的社会人口问题,如年龄、性别、收入、教育等。此外,调查还提出了一些问题,以衡量他们的理性、容忍度和乐观程度。本文的方法的优势在于,它提供了一种直接衡量投资者情绪的方法,可以产生一个非常精确的指标(Burghardt, 2011)<sup>[14]</sup>。

本文总共发放了 3000 份问卷,最终收到了 2460 份问卷。然后剔除无效问卷,最后有效问卷为 2129 份。为了鼓励受访者给出高质量的答案并提高参与率,我们为每个被受访的对象准备了小礼物。本文认为调查的随机样本是来自整个我国人口的真正随机抽样,因此表明了一般投

投资者的特征（附录 A 提供了每个省份的样本分布）。当然，并不是所有的受访者都回答了全部问题。然而，高回收率（70.96%）是对降低选择性偏差的最佳保障。例如，Shiller 和 Pound 认为，高回收率在某种意义上比大样本量更重要<sup>[15]</sup>，这通常被许多测量员所忽视。

### 三、结果

#### （一）描述性分析

表 1 描述了散户投资者的概况。在可观察的特征中，报告了有关性别、年龄阶层、教育阶层、职业、收入、投资规模、投资/收入比率以及投资经验时间长短的信息。从这张表可以看出，96.33%的散户投资者不到 60 岁，更多集中在 31 - 45 岁。他们的教育水平通常高于平均水平，例如，在样本中，只有 5.19%的散户投资者的学历低于高中，而在第六次全国人口普查中，这一数字是 38.79%。对于职业的分布，学生和退休人员的比例分别为 4.36%和 7.19%。只有 8.14%的散户投资者的收入高于 30 万元。而在股市中，只有 4.6%的散户投资者其投资额超过了收入的一半。超过 10 年的投资经验占比为 8.5%。

表 1 样本的描述性统计

样本量	2129
<b>个人特征</b>	
男性 (%)	59.31%
<b>年龄分布</b>	
18-30 岁 (%)	23.61%
31-45 岁 (%)	43.54%
46-59 岁 (%)	29.18%
60 岁以上 (%)	3.67%
<b>教育分布</b>	
初中及以下 (%)	5.19%
高中或同等 (%)	22.78%
大专学历 (%)	29.51%
本科学历 (%)	36.52%
硕士学历及以上 (%)	6.01%
<b>职业分布</b>	
公务员 (%)	9.49%
企、事业单位 (%)	36.91%
企业家 (%)	23.73%
三种类型的外资企业 (%)	10.88%
失业 (%)	2.78%

样本量	2129
专业投资者 (%)	3.74%
退休 (%)	4.36%
学生 (%)	7.19%
个体经营 (%)	0.91%
<b>收入分布</b>	
≤50,000 元 (%)	31.18%
50,000 元—100,000 元	41.49%
100,000 元—300,000 元	19.18%
300,000 元—500,000 元	4.24%
≥500,000 元	3.90%
<b>投资规模分布</b>	
≤20,000 元 (%)	29.44%
20,000 元—50,000 元	37.27%
50,000 元—100,000 元	20.37%
100,000 元—200,000 元	6.92%
200,000 元—500,000 元	3.07%
500,000 元—1,000,000 元	1.92%
≥1,000,000 元	1.01%
<b>投资/收入比例分布 (%)</b>	
低于 10%	38.02%
10%—30%	40.47%
31%—50%	16.73%
51%—100%	3.87%
高于 100% (债务方面)	0.73%
<b>投资经验时间长短分布</b>	
少于 1 年 (%)	22.73%
1—5 年 (%)	46.66%
6—10 年 (%)	22.10%
11—15 年 (%)	6.15%
多于 15 年 (%)	2.35%

表 2 报告了决策相关问题的结果，结果表明存在较高水平的异质性。例如，投资股市的原因是多种多样的。有些人只是有一些闲钱，有些人对交易感兴趣，有些人甚至只是看到他们的朋友赚了钱，所以他们也想尝试一下。第二个问题是关于信息资源，表明散户投资者并不知情：他们的主要信息资源来自公共媒体、朋友或其他散户投资者。



表 2 散户投资者决策的相关问题

样本量	2129
<b>你为什么投资股票市场?</b>	
有些多余的钱来赚取利润	999
对股市交易感兴趣	733
看到别人赚钱, 想试试	796
存款利率太低	662
<b>你从哪里得到决策信息?</b>	
公共媒体	1052
朋友	928
其他散户投资者	788
专业咨询服务	554
内部信息	219
<b>股票决策中各指标重要性的平均值 (S. D.)</b>	
(注意, 10 最重要, 1 最不重要)	
1 交易量	7.17 (2.31)
2 每股收益 (EPS)	6.92 (2.26)
3 红利政策	6.46 (2.36)
4 资产负债率	6.37 (2.31)
5 行业繁荣	6.85 (2.25)
6 专家或分析公司的判断	5.92 (2.40)
7 来自有经验朋友的判断	5.82 (2.45)
8 政府政策与经济形势	6.77 (2.62)
<b>未来 2-3 年经济前景的判断</b>	
增长率持续下降 (%)	29.37%
增长率保持不变 (%)	48.33%
增长率加速 (%)	22.31%

图 1 和图 2 表明了在某些假设的情况下, 随着价格上升或下降概率的增加的选择演化。

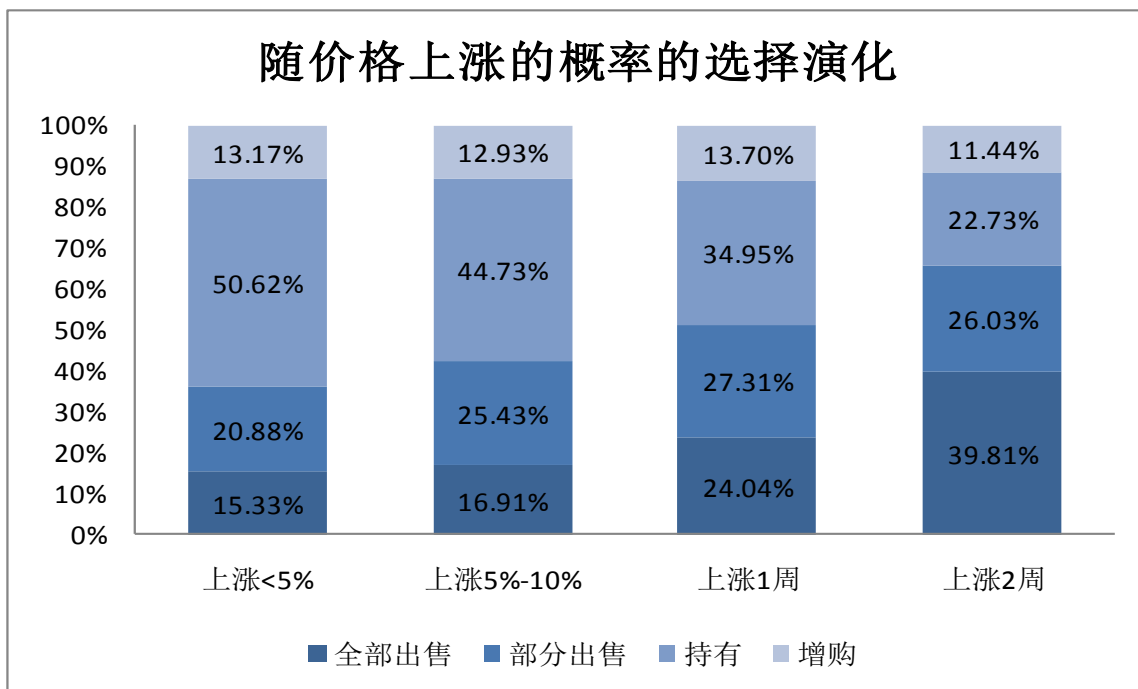


图 1 随价格上涨百分比的选择演化

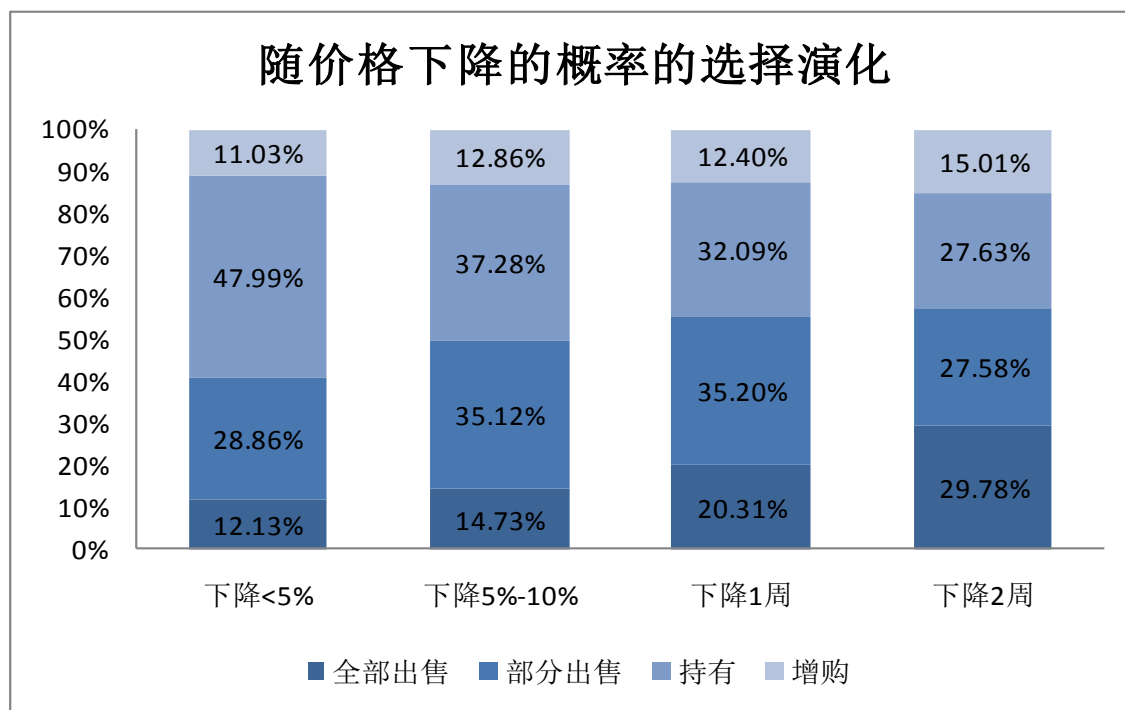


图 2 随价格下降百分比的选择演化

在图 1 中，以两种不同的方式来描述问题：首先，询问散户投资者如果价格上涨百分比低于 5% 或 5%-10%，他们会怎么做。从图中可以看到，部分出售的选择从 20.88% 小幅增加到

25.48%。当以时间为基点时，如果价格持续从一周增加到两周，有更多的散户投资者选择全部出售。相反也是类似的。当价格以百分比下跌时，更多的散户投资者选择部分出售，而如果价格下跌超过两周，他们倾向于全部出售。这可能意味着散户投资者有一些时间偏好，即，他们的风险厌恶程度随时间而增加，而非百分比。

本文也对专家意见如何影响投资者的决策选择感兴趣。图3以专家意见为基础，绘制出随着价格上涨的概率增加的选择演化。这表明，随着公共媒体专家预测价格上涨的概率的增加，散户投资者呈现购买更多的显著趋势。

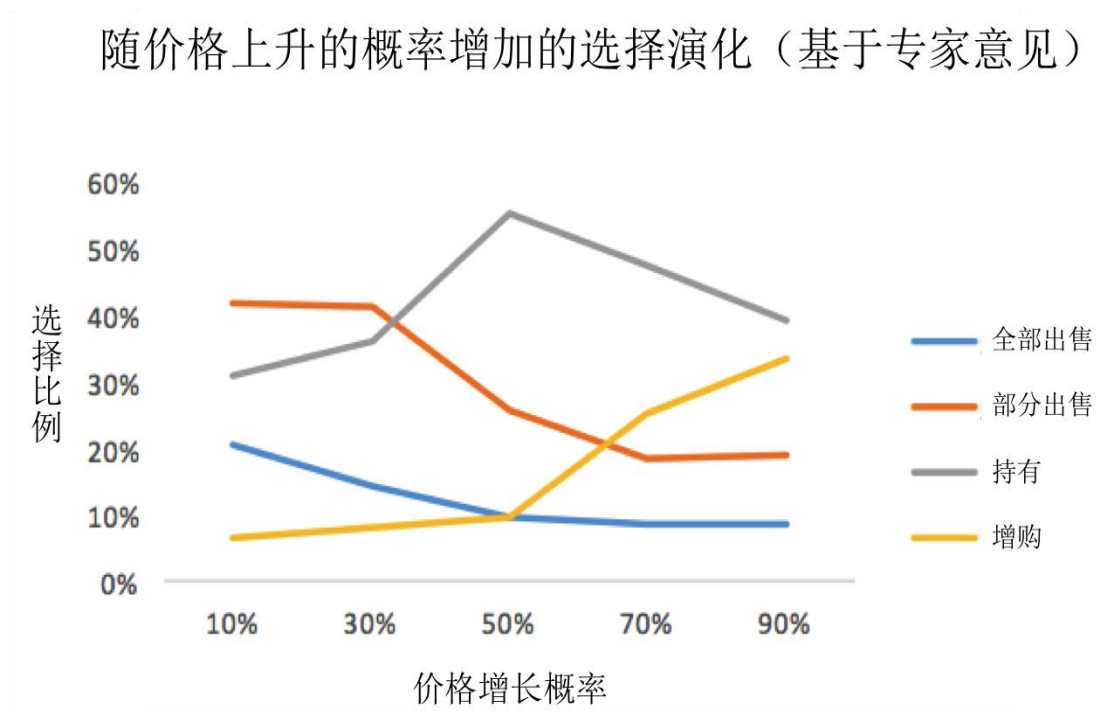


图3 随价格上升的概率增加的选择演化（基于专家意见）

为了检验调查的散户投资者是否与其回答一致，问卷中提问了类似的价格上涨和下降选择的问题，唯一的区别是这些问题是以价格为基点，而不是百分比和时间。从图4和图5可以看出，随着价格的上涨/下降，散户投资者倾向于选择全部出售，这与之前的研究结果是一致的。

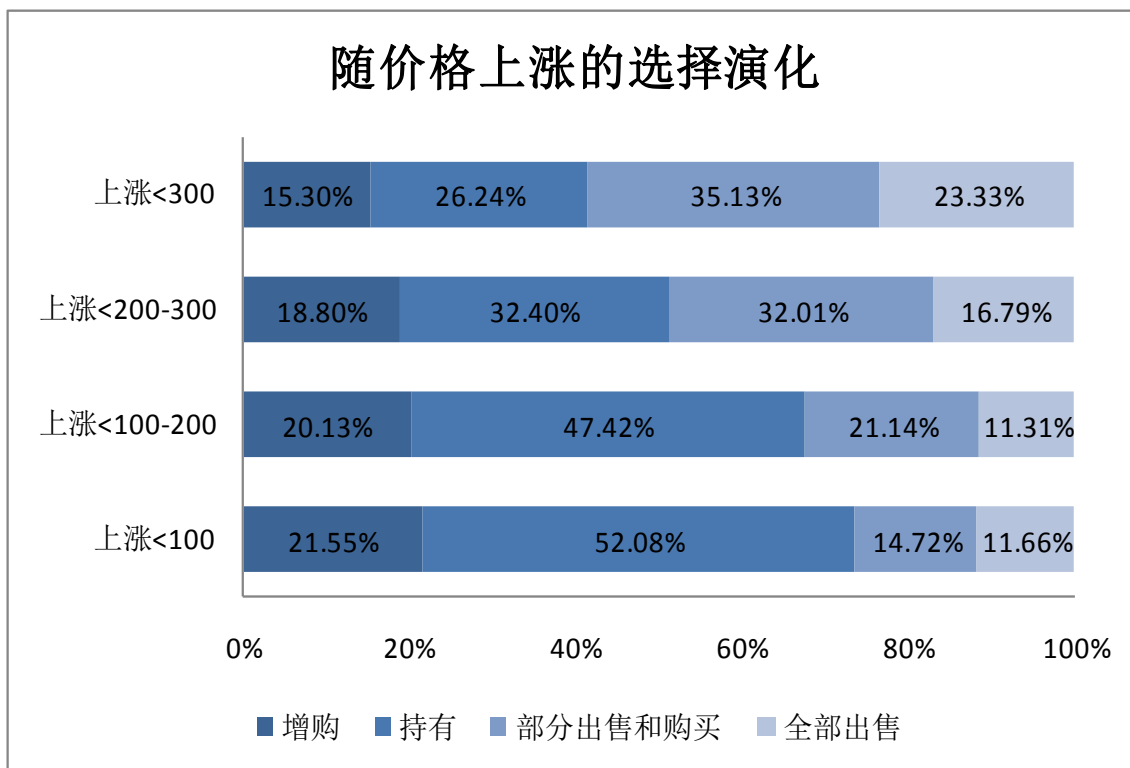


图4 随价格上涨的选择演化（稳健性检验）

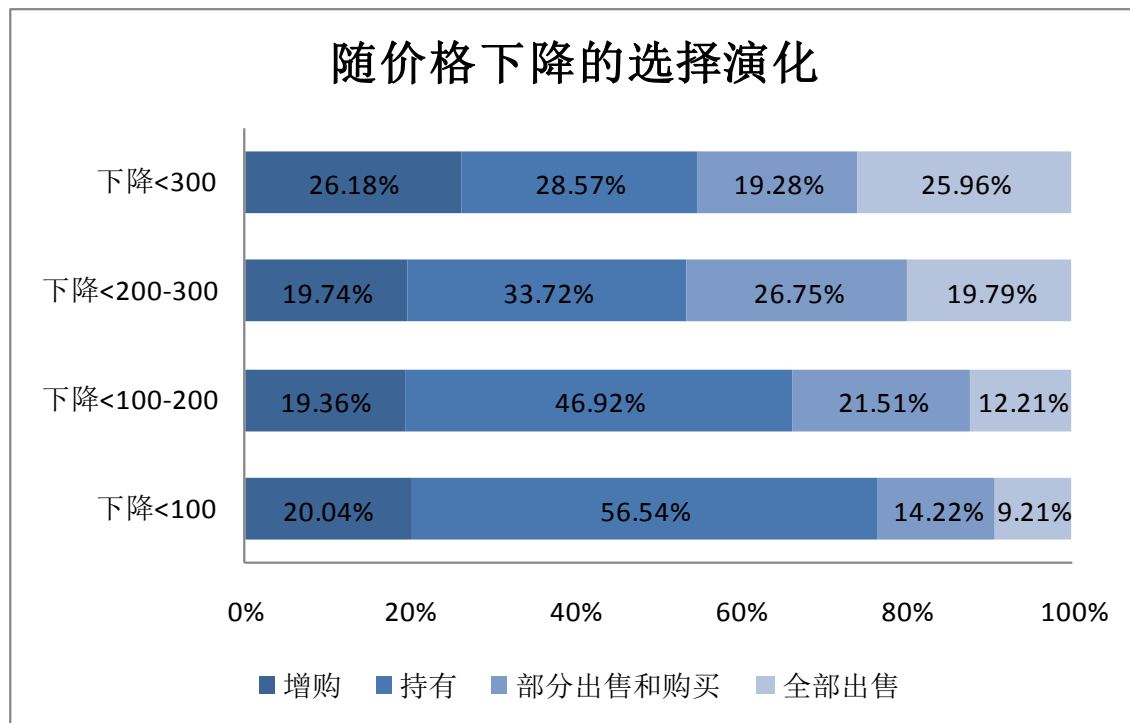


图5 随价格下降的选择演化（稳健性检验）

## (二) 计量经济学分析

这一小节分析个人情绪。为了同时确定所有外生特征的共同作用，本文选取因变量分别为理性、损失容忍度和乐观度进行回归估计。所有的估计结果均提供了稳健的标准误，修正了个人之间可能存在误差项的相关性。唯一的样本限制是在给定的回归中遗漏了变量的值。利用这些变量作为因变量，可以通过计量经济学分析来检验变量的决定因素。

表3 散户投资者理性的决定因素

因变量： 若理性则为1， 否则为0	(1)	(2)	(3)
性别	0.041** (0.020)	0.046** (0.020)	0.053** (0.021)
年龄	0.032*** (0.012)	0.024* (0.014)	0.015 (0.016)
教育程度	-	0.009 (0.011)	0.008 (0.011)
收入	-	0.016 (0.011)	0.011 (0.013)
专业投资者	-	<i>Ref.</i>	<i>Ref.</i>
公务员	-	-0.009 (0.058)	0.008 (0.063)
员工	-	-0.049 (0.052)	-0.015 (0.056)
干部	-	-0.056 (0.053)	-0.027 (0.058)
三类型外资企业	-	-0.026 (0.057)	0.002 (0.062)
失业	-	-0.050 (0.079)	-0.042 (0.085)
退休	-	0.057 (0.069)	0.096 (0.073)
学生	-	-0.009 (0.064)	0.031 (0.070)
创业	-	-0.112 (0.121)	-0.063 (0.123)
风险偏好	-	-	-0.026 (0.023)
资历	-	-	0.037***

因变量： 若理性则为1， 否则为0	(1)	(2)	(3)
			(0.013)
投资规模	-	-	-0.008 (0.011)
短期投资者	-	-	<i>Ref.</i>
长期投资者	-	-	-0.003 (0.033)
中期投资者	-	-	0.013 (0.028)
混合型投资者	-	-	0.003 (0.029)
增长率不变	-	-	<i>Ref.</i>
增长率下降	-	-	-0.029 (0.024)
增长率上升	-	-	0.001 (0.026)
Pseudo R <sup>2</sup>	0.0058	0.0112	0.0180
Log likelihood	-1142.1723	-1106.8301	-967.55956
样本量	2046	1993	1764

注：Probit 模型。报告的是平均边际效应。括号中为标准误差，\*\*\*p<0.01，\*\*p<0.05，\*p<0.1

### 1. 理性

用来论证散户投资者理性倾向的一个简单问题是：有两种可能的结果，“1：买两支股票，一个赚100元，另一个损失25元；2：只买一只股票，赚75元，你认为哪种选择会让你感觉更好？A.选项1；B.选项2；C.无差异”。本文将答案C分类为更有可能是理性的，编码为1，选项A和B为0。并使用Probit模型来估计散户投资者有哪些特征更有可能使他们更理性（Logit模型提供了类似的估计）。自变量包括了一些人口统计学变量，如性别、年龄、教育水平、收入和职业。还包括其他一些变量，比如投资者投资经验时间长短(资历)、风险态度、自我报告的投资者类型以及他们对未来经济增长率的判断。

表3报告了平均边际效应。回归结果显示，男性散户投资者更有可能是理性的。当控制其他变量时，年龄和收入不再显著。高等教育水平的散户投资者更理性。与专业投资者相比，其他类型的员工同专业投资者理性程度并为表现出显著差异。股票市场的经验越多，就越有可能是理性的。

表4 散户投资者损失容忍度的决定因素

因变量:	(1)	(2)	(3)
损失容忍度			
性别	0.112** (0.044)	0.117*** (0.044)	0.095** (0.046)
年龄	0.064** (0.027)	0.021 (0.031)	-0.048 (0.035)
教育程度	-	0.066*** (0.024)	0.046 (0.025)
收入	-	0.091*** (0.025)	-0.048 (0.029)
专业投资者	-	<i>Ref.</i>	<i>Ref.</i>
公务员	-	0.232* (0.135)	0.285** (0.139)
员工	-	-0.064 (0.116)	-0.004 (0.123)
干部	-	-0.038 (0.119)	0.044 (0.126)
三类型外资企业	-	0.026 (0.127)	0.145 (0.136)
失业	-	-0.220 (0.172)	-0.176 (0.183)
退休	-	0.236 (0.165)	0.306* (0.172)
学生	-	-0.093 (0.146)	0.042 (0.153)
创业	-	0.341 (0.260)	0.286 (0.252)
风险偏好	-	-	0.139*** (0.051)
资历	-	-	0.119*** (0.033)
投资规模	-	-	0.121*** (0.025)
短期投资者	-	-	<i>Ref.</i>
长期投资者	-	-	0.229*** (0.073)
中期投资者	-	-	0.157***

因变量:	(1)	(2)	(3)
损失容忍度			(0.060)
混合型投资者	-	-	0.199 <sup>***</sup> (0.063)
增长率不变	-	-	Ref.
增长率下降	-	-	-0.113 <sup>**</sup> (0.052)
增长率上升	-	-	0.015 (0.058)
截距项	2.125 <sup>***</sup> (0.063)	1.818 <sup>***</sup> (0.149)	1.523 <sup>***</sup> (0.164)
样本量	2040	1986	1759
R <sup>2</sup>	0.007	0.034	0.089

注：具有稳健控制的 OLS 模型。括号中为标准误差，\*\*\*p<0.01，\*\*p<0.05，\*p<0.1

## 2. 损失容忍度

问卷提问投资者能提供多少投资损失来作为衡量损失容忍度的直接指标。数值越高，损失容忍度越大。由于这个值可以被认为是连续的，因此本文使用了具有稳健控制的 OLS 模型。自变量与表 3 中的变量相同。

表 4 表明，男性散户投资者比女性投资者更容易承受损失。随着年龄、教育水平和收入的增加，散户投资者变得更加能承受损失（当控制其他变量时，年龄和收入不显著）。同专业投资者相比，公务员的损失容忍度有更高水平，而其他类型的工人与专业的散户投资者之间没有显著差异。有趣的是，风险偏好者对损失的容忍度有显著的影响。与其他类型的散户投资者相比，短期投资者的损失承受能力显著较低。与那些认为未来经济增长率将保持不变的人相比，那些认为经济增长率将会下降的人的损失承受能力较低。

表 5 散户投资者乐观度的决定因素

因变量:	(1)	(2)	(3)
乐观度			
性别	0.097 <sup>**</sup> (0.044)	0.086 <sup>*</sup> (0.044)	0.105 <sup>**</sup> (0.047)
年龄	0.031 (0.027)	0.010 (0.033)	-0.008 (0.037)
教育程度	-	-0.015 (0.024)	-0.018 (0.026)
收入	-	0.029	0.045



因变量:	(1)	(2)	(3)
乐观度		(0.025)	(0.031)
专业投资者	-	<i>Ref.</i>	<i>Ref.</i>
公务员	-	0.083	0.092
		(0.138)	(0.152)
员工	-	0.122	0.137
		(0.120)	(0.134)
干部	-	0.073	0.052
		(0.123)	(0.136)
三类型外资企 业	-	0.191	0.217
		(0.133)	(0.147)
失业	-	0.010	0.034
		(0.191)	(0.208)
退休	-	0.029	0.061
		(0.163)	(0.179)
学生	-	-0.057	-0.048
		(0.142)	(0.157)
创业	-	0.405	0.323
		(0.311)	(0.316)
风险偏好	-	-	0.030
			(0.053)
资历	-	-	-0.003
			(0.031)
投资规模	-	-	-0.018
			(0.025)
短期投资者	-	-	<i>Ref.</i>
长期投资者	-	-	0.154 <sup>**</sup>
			(0.074)
中期投资者	-	-	0.090
			(0.063)
混合型投资者	-	-	0.095
			(0.067)
增长率不变	-	-	<i>Ref.</i>
增长率下降	-	-	0.027
			(0.055)
增长率上升	-	-	0.248 <sup>***</sup>
			(0.060)

因变量:	(1)	(2)	(3)
乐观度			
截距项	2.007 <sup>***</sup> (0.063)	1.957 <sup>***</sup> (0.155)	1.838 <sup>***</sup> (0.176)
样本量	2042	1988	1760
R <sup>2</sup>	0.003	0.009	0.024

注：具有稳健控制的 OLS 模型。括号中为标准误差，\*\*\*p<0.01，\*\*p<0.05，\*p<0.1

### 3. 乐观度

乐观的散户投资者可能会有不同的表现，因此本文探究他们的乐观态度。为了测量这一点，根据过去的事实，即 2007 年和 2008 年的市场指数上升到 3000 点，然后是 6000 点，向研究对象提问他们认为未来的市场指数是多少。较高的数值表示较高的乐观程度。本文使用稳健控制的 OLS 回归，并包含了与前 2 个回归中相同的自变量。

表 5 报告了估计结果。从表 5 中可以观察到性别效应，即男性个体投资者比女性投资者更乐观。年龄、教育程度和收入没有影响。与短期投资者相比，长期投资者明显更为乐观。那些认为经济增长率将继续增长的散户投资者，比那些认为经济增长率在未来保持不变的人要乐观得多。

## 四、讨论与结论

本文从散户投资者对他们行为的偏见以及他们损失的程度进行了探究。讨论了一个基本问题：他们是谁？为此，本文使用了来自我国 17 个不同省份的代表性样本。研究结果揭示了我国散户投资者的社会人口特征，这与传统公共媒体的典型案例不同，后者可能误导人们对我国金融市场的理解。

此外，本文还测量了可能对预测散户投资者行为有重要影响的变量，如理性、损失容忍度和乐观程度，运用计量经济学的分析方法研究了这些变量的决定因素。考虑到我国市场客户的独特构成，将这些发现作为未来研究的基本步骤。事实上，本文只是研究的第一步，把它看作是真实世界中散户投资者的统计描述。

一个可能的关键问题是，调查问题是否可以在实际行为方面具有有意义的解释。Dohmen 等人（2011）在德国结合了 22000 名受试者的调查数据和一个现场实验<sup>[16]</sup>，发现在调查数据中对一个普遍的风险问题的回答是一个可靠的预测实际风险行为的指标，甚至控制了大量的可观测数据。他们的结论是，一个简单的定性调查方法可以形成风险态度的有意义的测量。事实上，结合了德国零售经纪人 Dorn 和 Huberman 的调查反应和交易记录，发现自我报告的调查数据有助于解释交易行为<sup>[17]</sup>。

本文研究结果可能对决策者、银行、经纪公司有重要意义，特别是为我国金融市场的散户投资者提供了更清晰的图景。本文的研究尚属第一步，下一步研究将通过结合散户投资者所说的和所做的，来检验本文中的变量是否与其实际行动相关联。

**参考文献:**

- Abraham, A., & Ikenberry, D.L..The Individual Investor and the Weekend Effect[J]. Journal of Financial and Quantitative Analysis, 1994, 29(02), 263-277.
- Barber, B. M., & Odean, T.. Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment[J]. The Quarterly Journal of Economics, 2001, 116(1), 261–292.
- Karlsson, A., & Nordén, L.. Home Sweet Home: Home Bias and International Diversification among Individual Investors[J]. Journal of Banking & Finance, 2007, 31(2), 317- 333.
- Barber, B. M., & Odean, T.. All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors[J]. Review of Financial Studies, 2008, 21(2), 785-818.
- Barberis, N., & Xiong, W.. What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation[J]. The Journal of Finance, 2009, 64(2), 751-784.
- De Bondt, W. F.. A Portrait of the Individual Investor[J]. European economic review, 1998, 42(3), 831-844.
- Bandopadhyaya, A., & Jones, A. L.. Measuring Investor Sentiment in Equity Markets[J]. Journal of Asset Management, 2006, 7(3-4), 208–215.
- Kumar, A., & Lee, C.. Retail Investor Sentiment and Return Comovements[J]. [The Journal of Finance, 2006, 61(5), 2451-2486.
- Schmeling, M.. Investor Sentiment and Stock Returns: Some International Evidence[J]. Journal of Empirical Finance, 2009, 16(3), 394-408.
- Baker, Malcolm, & Wurgler, Jeffrey. Investor Sentiment in the StockMarket[J]. Journal of Economic Perspectives, 2007, 21, 129–152.
- Qiu, L., & Welch, I..Investor Sentiment Measures (No. w10794)[R]. National Bureau of Economic Research. 2004.
- Croson, R., & Gneezy, U.. Gender Differences in Preferences[J]. Journal of Economic Literature, 2009, 47(2), 448–474.
- Dwyer, P. D., Gilkeson, J. H., & List, J. a.. Gender Differences in Revealed Risk Taking: Evidence from Mutual Fund Investors[J]. Economics Letters, 2002, 76(2), 151–158.
- Burghardt, M.. Retail Investor Sentiment and Behavior: An Empirical Analysis[M]. Springer Science & Business Media. 2011.
- Robert J. Shiller, John Pound Seasholes, Survey evidence on diffusion of interest and information among investors[J].Journal of Economic Behavior & Organization,1989,12(1),47-66.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences[J]. Journal of the European Economic Association, 2011, 9(3), 522-550.
- Dorn, D., & Huberman, G. Talk and Action: What individual Investors Say and What They Do[J]. Review of Finance, 2005, 9(4), 437–481.

## 附录 A 每一省份样本分布

总观测	观测值	百分比
样本的地理分布:		
1 山西	135	6.34%
2 河北	130	6.11%
3 山东	131	6.15%
4 北京	129	6.06%
5 上海	127	5.97%
6 江苏	128	6.01%
7 广东	125	5.87%
8 河南	122	5.73%
9 江西	121	5.68%
10 湖北	126	5.92%
11 湖南	126	5.92%
12 陕西	128	6.01%
13 吉林	130	6.11%
14 黑龙江	124	5.82%
15 辽宁	129	6.06%
16 安徽	134	6.29%
17 青海	84	3.95%
合计	2129	