

《保险资金入市创新驱动效应：发展“耐心资本”的视角》附录

附录 1 主要变量度量方式、描述性统计和分组均值差异检验

附表 1 主要变量的度量方式及描述性统计

变量名称	变量含义	均值	标准差	观测值
企业创新	上市企业集团公司当年合计的专利申请总数加一取自然对数	0.8434	1.3610	44743
实质性创新	上市企业集团公司当年合计的发明专利申请总数加一取自然对数	0.6632	1.1647	44743
策略性创新	上市企业集团公司当年合计的实用新型专利和外观设计专利申请总数加一取自然对数	0.4989	1.0756	44743
保险资金入市	当上市公司前十大股东中有保险资金时，取值为1，否则取值为0	0.1560	0.3629	44783
企业规模	主营业务收入加一取自然对数	20.997	1.8128	44588
资产负债率	总负债/总资产	0.4430	0.2008	44783
总资产收益率	净利润/总资产	0.0334	0.0609	44783
经营现金流量回报率	经营活动产生的现金流量净额/总资产	0.0499	0.0684	44783
固定资产比	固定资产净额/总资产	0.2262	0.1653	44783
董事会规模	董事会规模加一取自然对数	2.2491	0.1801	44775
董事会独立性	独立董事占比	0.3740	0.0534	44773
股权集中度	第一大股东持股比例	0.3416	0.1500	44783
重大公共卫生事件	上市公司所在城市发生新冠疫情当年及之后取值为1，否则取值为0	0.2858	0.4518	44783

为初步考察保险资金入市与企业创新之间的关系，本文将全样本依据是否有保险资金持股划分为两组，并对主要变量进行分组均值差异检验。检验结果如附表 2 所示。数据显示，有保险资金持股组的企业创新均值为 1.15，高于无保险资金持股组的均值 0.79，且组间差异在 1%的水平上显著。这一结果提供了初步的经验证据，表明保险资金入市可能对企业创新具有积极的促进作用，不仅支持了本文的假设 H1，也为后续的多变量回归分析奠定了基础。

附表 2 分组均值差异检验

变量	有保险资金	无保险资金	均值差异	T值
企业创新	1.1532	0.7860	-0.3672***	-20.8159
实质性创新	0.9204	0.6156	-0.3048***	-20.1858
策略性创新	0.7164	0.4587	-0.2577***	-18.4640
企业规模	21.6689	20.8723	-0.7966***	-34.1236
资产负债率	0.4605	0.4397	-0.0208***	-7.9601
总资产收益率	0.0502	0.0302	-0.0200***	-25.3990
经营现金流量回报率	0.0600	0.0481	-0.0120***	-13.4572
固定资产比	0.2366	0.2243	-0.0123***	-5.7362
董事会规模	2.2797	2.2435	-0.0363***	-15.5113
董事会独立性	0.3748	0.3739	-0.0009	-1.3555
股权集中度	0.3707	0.3363	-0.0344***	-17.6770
重大公共卫生事件	0.2178	0.2983	0.0805***	13.7196

附录 2 内生性处理结果

（一）工具变量法

附表 3 列（1）报告了工具变量法第一阶段的回归结果，可以看出工具变量对保险资金入市的估计系数在 1%水平上显著且系数方向为负，这一结论也符合预期，即距离保险资金投资机构越近的上市企业，被保险资金持股的可能性越大，满足工具变量的相关性要求。列

(2)中保险资金入市对企业创新的影响显著且系数方向为正。此外,列(2)中 *Kleibergen-Paap rk LM* 统计量的 *p* 值在 1%水平上显著,拒绝“工具变量识别不足”的原假设;列(2)中 *Kleibergen-Paap rk Wald F* 统计量大于 Stock-Yogo weak ID test critical values 中 10%偏误的临界值(16.38),拒绝原假设,说明不存在弱工具变量问题。工具变量法回归结果说明,在削弱了模型内生性问题后基准回归结果依然稳健。

附表 3 基于工具变量法的内生性处理

变量	(1)	(2)
	保险资金入市 第一阶段回归	企业创新 第二阶段回归
IV (ln(Dist+1)*year)	-0.0097*** (-28.4334)	
保险资金入市		0.2936*** (3.3678)
控制变量	Yes	Yes
企业固定效应	Yes	Yes
年份固定效应	Yes	Yes
样本量	44017	44017
<i>F</i>	808.46	14.2703
<i>Kleibergen-Paap rk LM</i>		206.931***
<i>Kleibergen-Paap rk Wald F</i>		808.4571
10% maximal IV size		16.38

注:考虑到本文大类行业数量较少(18个),为保证工具变量第一阶段 *F* 统计量分布的可靠性,在工具变量回归中将稳健标准误聚类至企业层面。

(二) 倾向得分匹配 (PSM)

附表 4 呈现了四种方法匹配后的回归结果,在不同匹配方法下,保险资金入市的系数均显著为正,表明基准回归结果稳健。

附表 4 基于倾向得分匹配法的内生性处理

企业创新	(1) 1:1近邻匹配	(2) 1:4近邻匹配	(3) 核匹配	(4) 卡尺匹配
保险资金入市	0.0583*** (4.3908)	0.0503*** (3.3712)	0.0388*** (3.6150)	0.0388*** (3.6150)
控制变量	Yes	Yes	Yes	Yes
企业固定效应	Yes	Yes	Yes	Yes
年份固定效应	Yes	Yes	Yes	Yes
样本量	12152	24162	44548	44548
<i>R</i> ²	0.7735	0.7535	0.6836	0.6836

(三) Heckman 二阶段法

Heckman 二阶段法的回归结果如附表 5 所示。列(1)的第一阶段回归结果显示,同年同行业上市公司前十大股东中是否有保险资金的平均状况的影响系数显著为正,说明工具变量与保险资金入市之间存在相关性。附表 5 列(2)的第二阶段回归结果显示,逆米尔斯比率的影响系数不显著,说明基准回归可能不存在严重的反向因果和样本自选择问题。并且,第二阶段中的保险资金入市影响系数显著为正,说明在控制了可能存在的反向因果和样本自选择问题后,基准回归结果依旧稳健。

附表 5 基于 Heckman 二阶段法的内生性处理

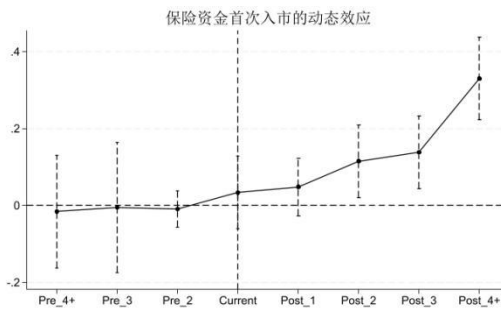
变量	(1) 保险资金入市 第一阶段	(2) 企业创新 第二阶段
保险资金入市		0.0357*** (3.1446)
<i>AvIns</i>	3.6328*** (44.3951)	
逆米尔斯比率		-0.0287 (-1.1215)
控制变量	Yes	Yes

企业固定效应	No	Yes
年份固定效应	No	Yes
样本量	44580	44548
Pseudo R ² / R ²	0.1108	0.6837

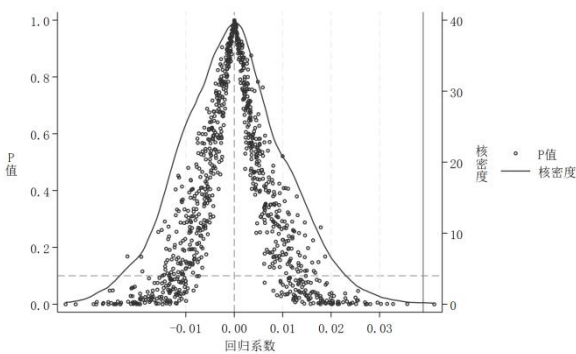
注：第（1）列汇报 Pseudo R²，第（2）列汇报 R²。

（四）构造准自然实验

附图 1 是多时点 DID 的平行趋势检验结果。从图中可以看出，在保险资金首次入市前，控制组与处理组之间不存在显著差异，回归系数均在 0 附近波动且统计上不显著，表明处理组和控制组在事前并未表现出显著的创新趋势差异，满足平行趋势假设。在保险资金入市之后的第二年，保险资金入市对企业创新的系数开始显著为正，原因在于企业创新（尤其是突破性创新）是一个长周期、高风险的过程，保险资金的资源支持转化为实际的专利产出需要一定的孵化周期。从附表 6 的回归结果可知，交互项系数在列（1）和列（2）中显著为正，而在列（3）策略性创新中不显著，与基准回归结果一致。这表明本文的主要结论在经过双重差分模型削弱内生扰动后依然成立。



附图 1 平行趋势检验与动态效应



附图 2 安慰剂检验

附表 6 多时点 DID 分析结果

变量	(1) 企业创新	(2) 实质性创新	(3) 策略性创新
<i>DID</i>	0.1601** (2.3573)	0.1446** (2.6257)	0.0826 (1.4746)
控制变量	Yes	Yes	Yes
企业固定效应	Yes	Yes	Yes
年份固定效应	Yes	Yes	Yes
样本量	26326	26326	26326
R ²	0.7051	0.6989	0.6753

附录 3 稳健性检验结果

（一）安慰剂检验

附图 2 中的黑色曲线为重复 1000 次得到的估计系数和核密度分布函数,空心圆点为估计系数对应的 P 值分布,右侧黑色竖线为正文基准回归中的真实估计系数(0.0388)。从附图 2 可以看出，虚假的处理变量回归系数基本服从均值为 0 的正态分布,而正文基准回归中 *Insurance* 的真实系数明显位于核密度函数外。这一结果表明，基准回归结果并未受到遗漏重要不可观测因素的干扰，本文结论具有较高的可信度。

（二）双重机器学习及 PPML 回归

在经济学分析中,对变量取自然对数后再进行回归分析,有助于改善数据的分布特性(如减少偏度和峰度)，但当数据存在零值时，直接取自然对数会导致样本损失。传统的“加 1 后取对数”方法虽可规避计算错误,却可能因人为调整数据结构而引入估计偏误，尤其在小样本或零值较多的情况下，可能影响因果推断的可靠性(Chen and Roth, 2024)。针对上述问

题，本文采用双重机器学习模型以及泊松伪最大似然估计（PPML）进行稳健性检验。双重机器学习具备两方面的优势：一方面，通过分阶段回归有效剥离混淆变量的影响；另一方面，可嵌入随机森林、Lasso、梯度提升等多种机器学习算法，自动捕捉变量间潜在的非线性关系，同时无需对零值作特殊处理，也避免了对函数形式施加强假设。基于双重机器学习框架，本文分别使用随机森林、Lasso 回归、梯度提升、神经网络与支持向量机五种方法，重新估计保险资金入市对企业创新的影响。同时为了确保结果的经济含义稳健，本文进一步采用了 Chen and Roth (2024) 推荐的泊松伪最大似然估计（PPML）方法。PPML 回归不需要对因变量取对数，能够直接处理零值问题，且其系数可以直接解释为半弹性。附表 7 结果显示，无论是双重机器学习模型还是 PPML 回归的估计系数均显著为正。这表明，即使采用更为灵活、能够更好地处理零值与非线性的估计策略，保险资金入市对企业创新的促进作用仍然十分显著，进一步支持了基准回归结果的稳健性。

附表 7 双重机器学习及 PPML 回归

企业创新	(1) 随机森林	(2) Lasso回归	(3) 梯度提升	(4) 神经网络	(5) 支持向量机	(6) PPML回归
保险资金入市	0.0774*** (5.6510)	0.0524*** (4.1904)	0.1298*** (7.4927)	0.0218** (2.1857)	0.1872*** (10.2742)	0.0660*** (7.9502)
控制变量	Yes	Yes	Yes	Yes	Yes	Yes
企业固定效应	Yes	Yes	Yes	Yes	Yes	Yes
年份固定效应	Yes	Yes	Yes	Yes	Yes	Yes
样本量	44548	44548	44548	44548	44548	42993

（三）其他稳健性检验

附表 8 展示了其他稳健性检验的回归结果，均支持了基准回归结果的稳健性。

附表 8 其他稳健性检验结果

变量	(1) 替换解释变量 企业创新	(2) 替换被解释变量 <i>rd pers</i>	(3) <i>rd expe</i>	(4) 增加固定效应 企业创新	(5) 剔除特殊样本 企业创新
保险资金入市		0.0010** (2.1882)	0.0002* (2.0159)	0.0346*** (3.3817)	0.0348*** (2.9894)
<i>InsRate</i>	0.0506*** (3.4871)				
控制变量	Yes	Yes	Yes	Yes	Yes
企业固定效应	Yes	Yes	Yes	Yes	Yes
年份固定效应	Yes	Yes	Yes	Yes	Yes
年份-行业固定效应	No	No	No	Yes	No
样本量	44548	25939	20497	44442	35629
R ²	0.6837	0.9000	0.9182	0.7096	0.6935

附录 4 FC 指数构造过程

参考 Hadlock and Pierce (2009)、况学文等(2010)、张悦玫等（2017）、顾雷雷等（2020）、陈峻等（2020）建立衡量企业融资约束程度的模型：

$$P(QUFC = 1 \text{ 或 } 0|Z_{i,t}) = \frac{e^{Z_{i,t}}}{1+e^{Z_{i,t}}} \quad \text{其中:}$$

$$Z_{i,t} = \alpha_0 + \alpha_1 size_{i,t} + \alpha_2 lev_{i,t} + \alpha_3 (\frac{CashDiv}{ta})_{i,t} + \alpha_4 MB_{i,t} + \alpha_5 (\frac{NWC}{ta})_{i,t} + \alpha_6 (\frac{EBIT}{ta})_{i,t} \quad (A1)$$

其中：*size*：表示企业资产规模，总资产的自然对数；

lev：表示企业财务杠杆率，资产负债率=总负债/总资产；

CashDiv：公司当年发放的现金股利；

MB: 表示企业市账比=市场价值/账面价值;
NWC: 净营运资本=营运资本-货币资金-短期投资;
EBIT: 息税前利润;
ta: 总资产

第一步, 按照年度对公司规模、公司年龄、现金股利支付率三个变量进行标准化处理

$(y_i = \frac{x_i - \bar{x}}{s}, \text{ 其中, } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}, \text{ 并根据标准化后的变量均值对上市公司进行排序 (升序), 分别以上下三分位点作为融资约束的分界点, 确定融资约束虚拟变量 QUFC, 大于 66\% 分位的上市公司定义为低融资约束组, QUFC=0, 小于 33\% 分位的上市公司定义为高融资约束组, QUFC=1.}$

第二步, 对模型 (1) 进行 Logit 回归, 拟合企业每一年度的融资约束发生概率 P, 并将其定义为融资约束指数 FC (取值在 0 到 1 之间), FC 越大, 企业的融资约束问题越严重。

附录 5 分组分析结果

附表 9 报告了基于企业创新投入和市场份额的分组分析结果。附表 10 报告了基于企业生命周期的分组分析结果。

附表 9 企业创新投入与市场份额分组分析结果

企业创新	(1) 创新投入较高组	(2) 创新投入较低组	(3) 市场份额较大组	(4) 市场份额较小组
保险资金入市	0.0627*** (4.0217)	0.0218 (1.2678)	0.0201 (1.5527)	0.0417* (1.9249)
控制变量	Yes	Yes	Yes	Yes
企业固定效应	Yes	Yes	Yes	Yes
年份固定效应	Yes	Yes	Yes	Yes
样本量	27055	17080	22245	22048
R ²	0.7167	0.7286	0.7471	0.5925

附表 10 企业生命周期分组分析结果

企业创新	(1) 成长期	(2) 成熟期	(3) 衰退期
保险资金入市	0.0452* (1.9228)	0.0538*** (3.1109)	0.0119 (0.3660)
控制变量	Yes	Yes	Yes
企业固定效应	Yes	Yes	Yes
年份固定效应	Yes	Yes	Yes
样本量	17796	17135	8257
R ²	0.7134	0.7467	0.7161

参考文献

Hadlock, C. J. and Pierce, J. R., 2010, "New Evidence On Measuring Financial Constraints: Moving Beyond The KZ Index", *The Review Of Financial Studies*, 23(5), pp. 1909~1940.
 况学文、施臻懿和何恩良, 2010, 《中国上市公司融资约束指数设计与评价》, 《山西财经大学学报》第 5 期, 第 110-117 页。
 张悦玫、张芳和李延喜, 2017, 《会计稳健性、融资约束与投资效率》, 《会计研究》第 9 期, 第 35-40+96 页。
 顾雷雷、郭建鸾和王鸿宇, 2020, 《企业社会责任、融资约束与企业金融化》, 《金融研究》第 2 期, 第 109-127 页。
 陈峻和郑惠琼, 2020, 《融资约束、客户议价能力与企业社会责任》, 《会计研究》, 第 8 期, 第 50-63 页。