

电子商务发展与城市数字技术创新

李林悦

浙江理工大学经济管理学院, 浙江 杭州

收稿日期: 2026年2月5日; 录用日期: 2026年2月14日; 发布日期: 2026年3月11日

摘要

本文基于2012~2022年中国279个地级及以上城市的面板数据, 以国家电子商务示范城市设立为准自然实验, 运用双重机器学习模型系统评估了电子商务发展对城市数字技术创新的政策效应。研究发现, 电子商务示范城市政策显著促进了城市数字技术创新水平的提升, 该结论在经过一系列稳健性检验后依然成立。机制分析表明, 电子商务发展主要通过促进高质量发明专利的申请与授权来推动城市数字技术创新, 而对实用新型专利的影响不显著, 说明电商政策在提升创新“量”的同时更加注重创新“质”的提升。本研究为理解数字经济时代下电子商务与城市创新之间的内在联系提供了经验证据, 也为地方政府优化电商政策、推动高质量创新提供了决策参考。

关键词

电子商务, 数字技术创新, 双重机器学习, 城市创新质量

E-Commerce Development and Urban Digital Technology Innovation

Linyue Li

School of Economics and Management, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: February 5, 2026; accepted: February 14, 2026; published: March 11, 2026

Abstract

This study employs panel data from 279 Chinese prefecture-level and above cities spanning 2012~2022. Utilizing the establishment of national e-commerce demonstration cities as a quasi-natural experiment, a double machine learning model is applied to systematically evaluate the policy effect of e-commerce development on urban digital technology innovation. The findings reveal that the e-commerce demonstration city policy significantly promotes the level of urban digital technology innovation, a conclusion that remains robust after a series of rigorous tests. Mechanism analysis indi-

cates that e-commerce development primarily drives urban digital technology innovation by fostering the application and grant of high-quality invention patents, while its impact on utility model patents is insignificant. This suggests that the e-commerce policy enhances not only the “quantity” but, more importantly, the “quality” of innovation. This research provides empirical evidence for understanding the intrinsic link between e-commerce and urban innovation in the digital economy era, and offers decision-making references for local governments to optimize e-commerce policies and promote high-quality innovation.

Keywords

E-Commerce, Digital Technology Innovation, Double Machine Learning, Urban Innovation Quality

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在数字经济快速发展的时代背景下，电子商务已成为推动经济转型升级、促进市场循环与产业生态重构的关键力量。作为数字经济的核心组成部分，电子商务不仅改变了传统交易模式，也深刻影响着城市创新系统的结构与效能[1]。近年来，中国陆续推出一系列电子商务示范城市政策，旨在通过制度创新与场景赋能，激发城市创新活力，尤其是数字技术领域的创新突破[2][3]。然而，现有研究多关注电子商务对经济增长、就业或消费的直接影响，较少从因果推断视角系统评估其对城市数字技术创新的政策效应，尤其是缺乏对创新质量的结构分析。

既有文献在政策评估方法上多依赖于传统计量模型，难以有效处理高维控制变量与非线性关系，可能导致估计偏误。近年来，双重机器学习等现代因果推断方法的发展为解决上述问题提供了新思路，引入正交化思想，能够在高维混淆变量存在的情况下实现政策效应的无偏估计[4]。本文尝试将双重机器学习模型引入电子商务政策评估中，以更科学、更精细的方式识别电子商务发展对城市数字技术创新的真实影响。

本文可能的边际贡献在于：第一，在研究方法上，首次将双重机器学习应用于电子商务与城市数字技术创新的政策评估中，提升了估计的准确性与稳健性；第二，在研究内容上，不仅关注创新的数量，更通过区分发明专利与实用新型专利，揭示电子商务对创新质量的异质性影响；第三，在政策启示上，为地方政府如何通过电子商务政策引导高质量创新提供了实证依据，有助于推动数字经济与城市创新体系的深度融合。

2. 研究设计

2.1. 模型构建

本文旨在探究国家设立电子商务试点对城市数字技术创新的政策效应评估，近年来，为了探索更为灵活和强大的方法，以提高政策评估的准确性和可靠性，学者们开始关注在因果推断领域新方法的应用与发展。Chernozhukov 等提出的双重机器学习模型通过引入正交化思想[5]，有效解决了高维混淆变量对政策的干扰问题。该模型不仅能够捕捉变量间的复杂非线性关系[4]，还能够有限样本条件下实现对政策处置效应的无偏估计[6]，为政策评估提供了更为精确和可靠的分析工具。首先，双重机器学习的部分线性模型构建如下：

$$patent_{it} = \theta_0 eccity_{it} + g(X_{it}) + U_{it}$$

$$E(U_{it} | eccity_{it}, X_{it}) = 0$$

其中, i 为城市; t 为年份; $patent_{it}$ 表示被解释变量, 为城市数字技术创新水平; $eccity_{it}$ 表示处置变量, 为国家电子商务示范城市的虚拟变量, 设置列为试点城市后 1 年为 1 (考虑存在滞后效应), 否则为 0; θ_0 表示政策对应的处置系数。 X_{it} 为控制变量集合, 本文将采用机器学习算法得到其估计 $\hat{g}(X_{it})$ 。 U_{it} 为误差项, 条件均值为 0。 利用辅助样本估计 $\hat{g}(X_{it})$, n 为样本容量, 直接对上述公式进行估计, 可以得到处置系数估计量为:

$$\hat{\theta}_0 = \left(\frac{1}{n} \sum_{i \in I, t \in T} eccity_{it}^2 \right)^{-1} \frac{1}{n} \sum_{i \in I, t \in T} eccity_{it} (Y_{it+1} - \hat{g}(X_{it}))$$

然而上式所得到的并非无偏估计量, $\hat{\theta}_0$ 难以收敛于 θ_0 , 可进一步考察其估计偏误:

$$\begin{aligned} \sqrt{n}(\hat{\theta}_0 - \theta_0) &= \left(\frac{1}{n} \sum_{i \in I, t \in T} eccity_{it}^2 \right)^{-1} \frac{1}{\sqrt{n}} \sum_{i \in I, t \in T} eccity_{it} U_{it} \\ &\quad + \left(\frac{1}{n} \sum_{i \in I, t \in T} eccity_{it}^2 \right)^{-1} \frac{1}{\sqrt{n}} \sum_{i \in I, t \in T} eccity_{it} [g(X_{it}) - \hat{g}(X_{it})] \end{aligned}$$

其中, $\alpha = \left(\frac{1}{n} \sum_{i \in I, t \in T} eccity_{it}^2 \right)^{-1} \frac{1}{\sqrt{n}} \sum_{i \in I, t \in T} eccity_{it} U_{it}$, 服从均值为 0 的正态分布,

$b = \left(\frac{1}{n} \sum_{i \in I, t \in T} eccity_{it}^2 \right)^{-1} \frac{1}{\sqrt{n}} \sum_{i \in I, t \in T} eccity_{it} [g(X_{it}) - \hat{g}(X_{it})]$ 为正则化偏差项。 双重机器学习采用机器学习及其正则化算法估计具体函数形式 $\hat{g}(X_{it})$, 不可避免地引入“正则偏误”, 虽能防止估计量方差过大, 但也导致了其不具无偏性, 具体表现为 $\hat{g}(X_{it})$ 向 $g(X_{it})$ 的收敛速度较慢, $n^{-\varphi g} > n^{-1/2}$, 因此, 随着 n 趋向于无穷大, b 也趋向于无穷大, $\hat{\theta}_0$ 难以收敛于 θ_0 。

若采用机器学习算法对上述模型中的 $g(X_{it})$ 直接求解, 则此时得到的估计系数 $\hat{\theta}_0$ 为正则化估计量, 有限样本下有偏, 因此, 进一步构建的辅助回归如下:

$$eccity_{it} = m(X_{it}) + V_{it}$$

$$E(V_{it} | X_{it}) = 0$$

其中, 函数形式 $m(X_{it})$ 未知, V_{it} 为误差项, 条件均值为 0。 因此, 本文采用机器学习算法估计 $\hat{m}(X_{it})$, 并以此构建残差估计 $\hat{V}_{it} = eccity_{it} - \hat{m}(X_{it})$, 然后采用同样算法估计主回归中的 $\hat{g}(X_{it})$, 获得 $Y_{it} - \hat{g}(X_{it}) = \theta_0 eccity_{it} + U_{it}$, 并将 \hat{V}_{it} 看作 BDB_{it} 的“工具变量”进行回归, 获得无偏的系数估计量如下:

$$\check{\theta}_0 = \left(\frac{1}{n} \sum_{i \in I, t \in T} \hat{V}_{it} eccity_{it} \right)^{-1} \frac{1}{n} \sum_{i \in I, t \in T} \hat{V}_{it} (Y_{it} - \hat{g}(X_{it}))$$

此时, $\check{\theta}_0$ 的收敛速度将取决于 $\hat{g}(X_{it})$ 、 $\hat{m}(X_{it})$ 向 $g(X_{it})$ 、 $m(X_{it})$ 的收敛速度。 两次机器学习估计一方面有利于排除处置变量 D_{it} 中由混淆变量集合 X_{it} 导致的影响, 另一方面, 则能加快 $\check{\theta}_0$ 的收敛速度, 进而获得有限样本下的准确估计。 另外, 为提高模型估计的稳定性和可靠性, 本文在回归中采用 5 折分割拟合的方法处理回归样本。

此外, 为验证控制变量与数字技术创新之间的复杂关系, 本文首先进行了系统的非线性检验。 如表 A1 所示, 经济发展水平($pgdp$)的二次项系数为 -0.196 且在 1% 水平上显著, 政府干预程度(gov)的二次项

系数为-2.222 且在 5%水平上显著, 两者均呈现明显的倒 U 型非线性关系。进一步比较发现, 非线性模型的拟合优度($R^2 = 0.820$)高于线性基准模型($R^2 = 0.815$)。这些非线性关系的存在表明, 传统线性模型难以准确刻画控制变量对创新的影响, 可能因函数形式误设导致政策效应估计偏误。因此, 本文采用双重机器学习模型, 通过机器学习算法灵活拟合控制变量的非线性函数, 以更准确地识别电子商务示范城市政策对数字技术创新的因果效应。

2.2. 数据来源

本文选取 2012~2022 年中国 279 个地级及以上城市作为实证研究样本。样本数据来源分为两类: 一是城市数字技术创新申请专利数量数据主要来自国家知识产权局; 二是城市层面数据来自《中国城市统计年鉴》、国泰安数据库(CSMAR), 关于样本的缺失值通过结合地方统计年鉴或者插值法补齐, 对数据严重缺失的城市进行了删除处理。

2.3. 变量设定

2.3.1. 被解释变量

被解释变量 *patent* 为城市数字技术创新, 用各地级市数字技术创新申请专利数量衡量。参考李雪琴的做法, 根据国家知识产权局印发的《关键数字技术专利分类体(2023)》与国际专利分类参照关系表的 IPC 号信息, 将属于云计算、大数据、物联网、工业互联网、区块链、人工智能、虚拟现实、增强现实等数字技术产业对应 IPC 的专利识别为数字专利[7]。筛选并加总得到各地级市历年的数字技术创新申请专利总量, 加 1 取自然对数, 以此衡量城市数字技术创新水平。

2.3.2. 处置变量

处置变量 $eccity_{it} = Treat_t \times Post_t$ 为国家电子商务城市政策的虚拟变量。参考唐跃桓等的做法, 根据设立国家电子商务示范城市的时间确定其数值[8]。考虑到政策从出台到落地执行、再到微观主体调整其研发决策, 存在一定的实施与响应时滞。我们在后续实证分析中, 将政策产生实际处理效应的起始年份设定为官方公布年份的次年, 即将处理组城市划分为三个队列, 其处理起始时间分别设定为 2013 年、2015 年和 2018 年。具体来说, $Treat_t$ 为当城市在 t 年设立国家电子商务示范城市时, 该指标在 $t+1$ 年及后续年份取 1, 否则取 0; $Post_t$ 变量为年份虚拟变量, 在各地级市政策实施当年及之前取 0, 实施 $t+1$ 及以后则取 1。总的来说, 如果城市 i 在第 $t+1$ 年被设立为国家电子商务示范城市, 则在第 $t+1$ 年及之后, $eccity_{it}$ 取 1, 否则为 0。

2.3.3. 控制变量

本文加入一系列城市层面的控制变量, 具体包括, 经济发展水平(*pgdp*, 人均 GDP 的自然对数)、产业结构水平(*ind*, 第二产业和第三产业生产总值与 GDP 之比)、金融发展水平(*fin*, 年末金融机构各项贷款总额与金融机构存贷款总额之比)、科教支持强度(*stc*, 科教支出与财政总支出之比)、政府干预程度(*gov*, 政府财政支出与 GDP 之比)、社会消费水平(*con*, 消费品销售总额与 GDP 之比)、对外开放程度(*ope*, 进出口总额与 GDP 之比)、城镇化水平(*urb*, 城镇人口与总人口之比)。在实证分析中, 本文还加入了城市和年份固定效应, 以进一步控制个体和时间层面不可观测因素的影响。主要变量的描述性统计特征, 见表 1。

3. 实证结果

3.1. 基准回归

本文采用双重机器学习模型估计了城市电子商务发展对数字技术创新的影响效应。同时, 为了保证

结果的稳健性, 本文分别选择套索回归、弹性回归和支持向量机这三种算法进行回归, 回归结果如表 2 第 (1)~(3)列所示。结果显示, 无论是选用哪种算法, *eccity* 的系数均正向显著, 即城市电子商务发展确实能够显著促进数字技术创新。为了方便起见, 后文分析中均选择可解释性较强的套索回归算法进行计算。

Table 1. Descriptive statistics of variables

表 1. 变量的描述性统计特征

变量	OBS	Mean	Sd	Min	Max
<i>patent</i>	2995	6.035	1.840	0.693	12.079
<i>eccity</i>	2955	0.171	0.377	0.000	1.000
<i>pgdp</i>	2995	10.818	0.563	9.007	13.056
<i>ind</i>	2995	0.882	0.077	0.501	1.000
<i>fin</i>	2995	0.409	0.065	0.171	0.849
<i>stc</i>	2995	0.194	0.041	0.048	0.372
<i>gov</i>	2995	0.202	0.100	0.044	0.916
<i>con</i>	2995	0.390	0.107	0.000	1.013
<i>ope</i>	2995	0.177	0.277	0.000	2.491
<i>urb</i>	2995	0.573	0.145	0.182	1.000

Table 2. Baseline regression results

表 2. 基准回归结果

变量	套索回归 (1)	弹性网络 (2)	支持向量机 (3)
<i>eccity</i>	0.094** (0.043)	0.097** (0.043)	1.360*** (0.072)
控制变量	是	是	是
时间固定	是	是	是
城市固定	是	是	是
样本量	2955	2955	2955

注: **、*和*分别表示在 1%、5%和 10%水平上显著; 括号内为稳健标准误, 下表同。

3.2. 稳健性检验

3.2.1. 安慰剂检验

本文还进行了随机化安慰剂检验以避免某些随机现象的影响作用, 借鉴白俊红等的做法, 随机选取城市作为实验组, 其余作为对照组, 将此过程重复 500 次, 并画出 500 次安慰剂检验的估计系数、对应的 *p* 值以及基准回归所得到的估计系数[9]。结果显示, 安慰剂检验的估计系数主要集中于 0 值附近, 通过了安慰剂检验, 基准结论是成立的, 见图 1。

3.2.2. 剔除异常值影响

由于回归样本中可能存在异常值, 这些异常值可能导致估计结果出现偏差, 为了减少异常值对回归结果的干扰, 本文对基准回归中除处置变量外的所有变量进行了缩尾处理, 分别采用了 1%和 99%分位点的缩尾方法, 将高于最高分位点和低于最低分位点的数值替换为相应分位点的值, 并在此基础上重新进

行回归分析。表 3 的第(1)是具体的回归结果。结果表明，剔除异常值后，本文的结论依然显著。

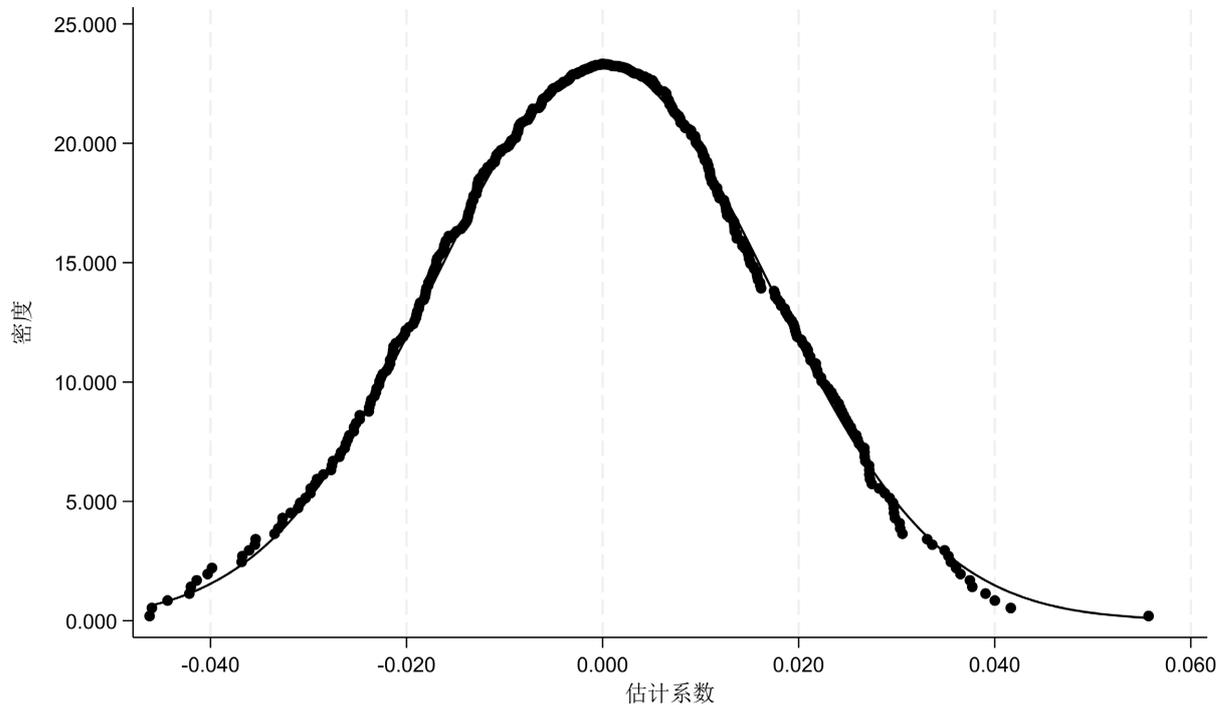


Figure 1. Placebo test

图 1. 安慰剂检验

Table 3. Robustness tests

表 3. 稳健性检验

变量	缩尾处理 (1)	排除其他政策干扰		改变样本分割	
		(2)	(3)	(4)	(5)
<i>eccity</i>	0.072* (0.040)	0.093** (0.043)	0.042** (0.028)	0.114*** (0.043)	0.082* (0.042)
大数据战略		是			
宽带中国			是		
控制变量	是	是	是	是	是
时间固定	是	是	是	是	是
城市固定	是	是	是	是	是
样本量	2955	2955	2955	2955	2955

3.2.3. 排除其他政策干扰

考虑到在本文的样本期间内还实施了其他政策，为保证政策效应估计的准确性，对其他相似政策进行了控制，如 2018 年的国家“大数据战略”以及 2014 年开始的“宽带中国”试点政策。据此，本文构建了“大数据战略”和“宽带中国”的政策虚拟变量加入回归分析。表 3 的第(2)列和第(3)列为分别控制两项政策的结果显示，在考虑其他政策的干扰后，电子商务发展对城市数字技术创新的政策效应显著性不变，说明本文结论的稳健性。

3.2.4. 重设双重机器学习模型

为避免双重机器学习模型设定偏误对结论产生影响，本文改变双重机器学习模型的样本分割比例，由先前的 1:4 改为 1:7，探究样本分割比例对本文结论的可能影响，研究结果如表 3 第(4)列和第(5)列所示，结果依然稳健。

3.3. 机制分析

在数字经济时代，电子商务发展作为畅通市场循环、重塑产业生态的关键驱动力，其价值不仅在于交易规模的提升，更应着眼于创新结构的优化。城市电子商务生态的完善能够通过构建高效的数据流通与应用场景，显著降低数字技术研发的市场验证成本，并形成对高质量创新主体的识别与集聚效应。城市本地企业在电商平台提供的实时市场需求反馈及商业化应用场景支持下，更聚焦于底层算法与关键技术攻关，其创新产出因此集中于体现实质进步的发明专利。这一过程同时促进了高端创新要素的集聚，从而系统地优化城市数字创新生态质量。

为了进一步研究电子商务发展对城市数字技术创新的影响机制，表 4 进一步给出了实证结果。第(1)列和第(2)列的结果显示，*eccity* 的系数在 1% 的水平上显著为正，表明电子商务示范城市政策的实施对城市数字经济授权和申请的发明专利产生了显著的促进作用；第(3)列和第(4)列的结果显示，*eccity* 的系数不显著，表明城市电子商务发展对授权和申请的实用新型专利作用不显著。产生这个结果原因可能是设立国家电子商务示范城市一政策工具通过优化数据要素配置和强化市场需求牵引，对城市数字经济的创新质量产生了“择优性”的促进作用，政策红利主要流向了高质量、高含金量的创新活动，推动创新成果向技术含量更高、保护期更长的发明专利集中，相对弱化了以适应性改良为特征的实用新型专利产出。总之，电子商务发展不仅促进了城市数字技术创新的“量”，更显著地提升了城市创新的“质”，体现了数字经济领域的高质量发展理念。

Table 4. Mechanism analysis

表 4. 影响机制

变量	申请发明 (1)	授权发明 (2)	申请实用新型 (3)	授权实用新型 (4)
<i>eccity</i>	0.180*** (0.047)	0.277*** (0.052)	0.049 (0.037)	0.035 (0.039)
控制变量	是	是	是	是
时间固定	是	是	是	是
城市固定	是	是	是	是
样本量	2955	2955	2955	2955

4. 研究结论

在数字经济发展的关键阶段，城市电子商务生态的完善成为驱动高质量数字技术创新的重要制度基础。作为数字经济中发展规模最大、覆盖范围最广、创业创新最活跃的重要组成部分，电子商务正在为构建城市创新发展格局发挥重要作用。因此，本文基于城市面板数据，以国家设立电子商务示范城市为准自然实验，研究电子商务发展对城市数字技术创新的影响效应。

本文研究结论：第一，通过基准回归，发现设立国家电子商务示范城市，有助于推动当地电子商务发展，提高城市数字技术创新水平。第二，无论是剔除异常值、排除其他政策干扰，还是重设机器学习模型等通过一系列的稳健性检验，验证了基准回归的结论依然显著。第三，本文进一步研究了电子商务

发展对城市数字技术创新的影响机制，发现城市电子商务发展通过提高数字经济相关发明专利的数量，进而促进城市数字技术创新水平，而对实用新型专利的促进作用不显著。这一结论表明电商发展形成的快速迭代与颠覆性创新竞争范式，使研发主体更倾向于追求实质性技术突破，这种“质量筛选”机制推动了创新资源向高质量活动集聚，相对弱化了以适应性改良为特征的实用新型专利产出。进一步说明了电子商务的发展不仅促进了城市创新的“量”，更显著地提升了城市数字技术创新的“质”。

参考文献

- [1] 蔡嘉怡, 龚利, 叶爱山, 金彩灵, 刘芳. 我国电子商务转型升级研究[J]. 合作经济与科技, 2026(4): 51-53.
- [2] 刘乃全, 邓敏, 曹希广. 城市的电商化转型推动了绿色高质量发展吗?——基于国家电子商务示范城市建设的准自然实验[J]. 财经研究, 2021, 47(4): 49-63.
- [3] 乔智, 谭淳丰, 王思雨, 杨志国. 电子商务发展与审计师风险感知——基于国家电子商务示范城市的准自然实验[J]. 中国注册会计师, 2025(12): 47-55.
- [4] Yang, X., Wu, H., Ren, S., Ran, Q. and Zhang, J. (2021) Does the Development of the Internet Contribute to Air Pollution Control in China? Mechanism Discussion and Empirical Test. *Structural Change and Economic Dynamics*, **56**, 207-224. <https://doi.org/10.1016/j.strueco.2020.12.001>
- [5] Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., et al. (2018) Double/Debiased Machine Learning for Treatment and Structural Parameters. *The Econometrics Journal*, **21**, C1-C68. <https://doi.org/10.1111/ectj.12097>
- [6] 王茹婷, 彭方平, 李维, 王春丽. 打破刚性兑付能降低企业融资成本吗? [J]. 管理世界, 2022, 38(4): 42-64.
- [7] 李雪琴, 郑酌基, 韩先锋. 乘“数”而上: 政府数据治理赋能企业数字创新[J]. 数量经济技术经济研究, 2024, 41(12): 68-88.
- [8] 唐跃桓, 黎静霖, 杨其静. 电子商务与企业跨地区交易: 交易成本经济学的视角[J]. 经济研究, 2025, 60(1): 74-90.
- [9] 白俊红, 张艺璇, 卞元超. 创新驱动政策是否提升城市创业活跃度——来自国家创新型城市试点政策的经验证据[J]. 中国工业经济, 2022(6): 61-78.

附录

Table A1. Nonlinearity test results of core control variables (*pgdp, gov*)

表 A1. 核心控制变量非线性检验结果(*pgdp, gov*)

变量	套索回归 (1)	弹性网络 (2)
一次项系数	0.399*** (0.079)	1.220** (0.580)
二次项系数	-0.196*** (0.042)	-2.222** (0.957)
时间固定	是	是
城市固定	是	是
样本量	2955	2955
R ²	0.820	0.816