

张业鹏

yepengzhang@link.cuhk.edu.cn

13563945570

个人博客主页: [Application Exploration of LLM & Fine-tuning](#)

研究兴趣

研究领域: 过往研究涵盖城市与区域经济学、文本大数据分析、以及深度学习建模与优化等方向,

兴趣特长: 熟悉地理信息系统 (GIS) 分析与空间数据处理, 具备扎实的数据科学基础; 擅长应用机器学习与大语言模型 (如 LLM) 于文本挖掘与因果推断场景

教育背景

香港中文大学 (深圳), 计算社会科学, 统计学理学硕士

2024.9 - 2025.7

山东财经大学, 金融工程, 经济学学士

2019.9 - 2023.6

论文发表

High-speed rail, technological improvement, and urban carbon emission efficiency, *Transportation*

Research Part D: Transport and Environment

2025.3

- 主要贡献: 进行数据整理、数据清洗, 并整理形成面板数据; 进行实证分析和检验, 探寻高铁开通与城市碳效率之间的潜在影响机制; 可视化城市在高铁开通前后的空间数据特征。

Power of competition: Unravelling the impact of China's fair competition review system implementation on firm innovation, *Economic Analysis and Policy*

2023.11

- 主要贡献: 基于 2000 年至 2021 年的管理层评论数据, 采用 LDA 主题建模, 并结合 TextCNN 方法进行文本情感分类 (使用 ChatGPT3.5 的 token 进行训练), 衡量上市公司创新水平。

近水楼台先得月: 银企距离与企业数字化转型, *财经科学*

2024.3

- 主要贡献: 基于上市公司年报与管理层评论文本进行文本大数据分析, 测度出企业的数字化转型指数; 探究上市公司与银行的最短距离与其数字化转型之间的潜在影响机制, 进行实证分析和检验。

Adding Fuel to the Fire: Does Temperature Affect the Carbon Emission Intensity of Firms? *Journal of Environmental Management Under Review*

2025.4

- 主要贡献: 进行数据搜集、数据的整理和清洗、相关指标的构造和测算; 进行稳健性与异质性检验, 整理相关的文献资料和理论知识, 完善潜在影响机制的检验。

城市空间形态如何影响创业水平? *Working Paper*

2024.12

- 主要贡献: 本文融合空间经济学与创业生态系统理论, 构建“城市形态—生态要素配置—创业行为”分析框架, 探究城市几何形态对创业水平的影响机制。本人主要负责指标的构建, 工具变量的测算, 以及相关时空大数据的分析和处理。

The Spatial Consequence of Hukou Reform *SSRN*

2025.7

本文研究发现, 放松人口迁移限制导致实施改革的中小城市空间形态显著恶化, 表现为城市形状不规则性、碎片化程度加剧。这一效应在财政自给能力弱、土地财政依赖度高、地理条件受限 (如可用土地少或坡度陡峭) 的城市尤为突出。研究揭示了户籍改革通过土地财政激励影响城市形态的因果路径, 为发展中国家的城镇化政策提供警示。

数字化转型对企业融资约束的影响研究——基于 LDA 主题模型的数字化转型测度方法 优秀毕业论文

本文基于 A 股上市公司数据, 运用 LDA 主题模型量化企业数字化转型, 实证发现: **数字化转型显著缓解融资约束**, 主要通过提升企业创新绩效与内控质量实现; 该效果在民营企业、小企业和高新技术企业中更突出, 且受互联网金融发展、行业竞争及市场不确定性的正向调节。研究为政策支持企业数字化提供了依据。

项目经历

GIS 在经管领域的应用研究

2024.7 至今

- 基于夜间灯光数据的城市边界紧凑度及潜在城区足迹指数测算与应用
 - 城市形状指数计算：根据 1998 至 2013 年的 DMSPP 夜间灯光数据和中国城市行政边界数据，测算衡量城市运行效率的紧凑度指数
 - 工具变量与潜在城区足迹指数计算：整合 1985-1999 年城市统计年鉴数据、1992-2022 年欧空局土地利用数据及中国高程数据，计算了城市紧凑度的工具变量及潜在城区足迹指数。
- 基于人口密度栅格数据的人口分布集中度测算与应用
 - 人口分布集中度 (SPC) 指数构建：基于 LandScan 全球人口密度数据及 2000-2013 年夜间灯光数据，构建了衡量城区内部人口分布集中度的空间人口集中度 (SPC) 指数。
- 基于路网矢量与高德 POI 数据的公共基础设施规模分析与应用
 - 交通基础设施可达性与密度分析：
 - * 基于工业企业坐标数据与 OSM 路网数据，计算工业企业的高速公路可达性。
 - * 结合夜间灯光数据界定的城区范围，计算城区内部路网密度。
 - 城区公共设施密度测算
 - * 基于高德 POI 数据 (2012-2020 年) 及夜间灯光数据界定的城区范围，计算城区内相关公共基础设施的分布密度。

2024 国家社会科学基金项目: 要素流动视角下房价对制造业投资影响机制研究

2024.8

- 课题概述：项目主要研究中国房价波动对制造业企业要素流动的影响及其经济后果，经过严密的论证，核心结论如下：房价通过成本、融资和预期渠道影响制造业要素流动，需通过政策平衡房地产与实体经济发展。
- 主要贡献：主要参与撰写了第五章“数据处理与指标计算”部分。在该章节中，我重点分析了影响房价的综合因素，并基于人口空间分布数据、夜间灯光数据等栅格与矢量数据，结合文本大数据，测算出一系列反映城市经济发展水平、人口空间分布特征及政府情绪变化的综合指标，为深入剖析城区内房价的复杂影响因素奠定了坚实的数据基础。

人工智能暴露度测算——LLM 对中国劳动力市场的冲击分析

2024.7 至今

- 模型部署与微调：部署大语言模型 (LLM)，并采用适配的文本数据结合指令进行微调，显著提升了其在语义匹配与指标测算任务中的稳健性与准确性。
- AI 暴露度测算与应用：调用 DeepSeek API，结合《中华人民共和国职业分类大典》信息，构建生成式人工智能 (AI) 暴露度测算指标，为评估生成式 AI 对中国劳动力市场的潜在影响提供了创新性分析依据。

个人能力

英语水平: IELTS: 6.5

编程与软件工具: Python, R, Stata, Matlab

数据分析与建模: 数据科学与网络爬虫 (*pandas, numpy, scrapy*); GIS(*Geopanda, sf, terra, QGIS*); 可视化 (*matplotlib, ggplot2*); 机器学习与深度学习建模 (*sklearn, pytorch*)

会议与比赛经历

2025 年美国大学生数学建模竞赛 Student Advisor, Finalist (特等奖提名)	2025.5
2022 年美国大学生数学建模竞赛, Honorable Mention	2022.5
2021 年美国大学生数学建模竞赛, Honorable Mention	2021.5
参与第 9 期香樟经济学 (济南) “云 Seminar” 暨 “发展经济学” 专题研讨会并担任主讲人	2024.9
合作论文入选第六届城市经济学者论坛	2024.12
参与第四届中国工业经济学会青年论坛并担任主讲人	2022.7
参加第二十一届中国经济学年会并担任主讲人	2021.12
合作论文入选第四届未来经济学家论坛	2021.12