Zhang Yepeng

ypzhang@nsd.pku.edu.cn
13563945570
Google Scholar Profile

Research Interests

Research Fields: Urban data science; transportation and mobility; urban and regional economics; infrastructure and environmental sustainability, with a focus on road networks, urbanization, and freight/human mobility.

Specialized Skills :

- ***Urban Data Science & Mobility***: construction of road and transit networks, travel-time and accessibility measures, and micro-level mobility indicators from GPS/OD data;
- ***Deep Learning Modeling & Optimization***: proficient in building and tuning a variety of deep learning architectures for classification and regression, with substantial experience in data wrangling, training pipelines, and prediction tasks;
- ***GIS & Spatial Analysis***: large-scale geospatial processing (Python/R, Google Earth Engine) and urban form/infrastructure metrics;
- ***Causal Inference & Econometrics***: DID, RD, IV, and panel models in Stata/R, complemented by machine learning methods and text/LLM-based measurement tools.

Work Experience

NSD, Peking University Research Assistant **Spatial Economic Lab Personal Page** 2025.7 - 2026.6

- **Research Project**: Analysis of the impact and optimization of "Belt and Road Initiative" road construction projects on local economies. ***Urban economics, spatial economics***
- **Research Responsibilities**: Conducting large-scale geographic data analysis, indicator computation, and spatial visualization using Python, R, QGIS, and Google Earth Engine; supporting causal identification and empirical analysis for road infrastructure projects.

Education Experience

CUHK-Shenzhen, Computational Social Science, *M.S. in Statistics* 2024.9 - 2025.7

Shandong University of Finance and Economics, Financial Engineering, *B.Ec.* 2019.9 - 2023.6

Publications and Research Experience

High-speed rail, technological improvement, and urban carbon emission efficiency, *Transportation Research Part D: Transport and Environment* 2025.3

With Weijian Su, Xiabing Li, **Yepeng Zhang**, Quanfei Zhang, Tao Wang, Małgorzata Magdziarczyk, Adam Smolinski.
Main Findings: Using a DID model, this study demonstrates that HSR connectivity significantly enhances urban carbon emission efficiency.
Responsibilities: Compiled and cleaned city-level panel data; conducted empirical analysis and mechanism identification; visualized spatial characteristics before and after HSR implementation.

Power of competition: Unravelling the impact of China's fair competition review system implementation on firm innovation, *Economic Analysis and Policy* 2023.11

With Zumian Xiao, Shuhan Wang, **Yepeng Zhang**, Dengkui Si.
Main Findings: Using China's 2016 Fair Competition Review System as a natural experiment, this study finds that intensified competition spurred firm innovation by tightening financing for monopolistic firms.
Responsibilities: Developed innovation metrics using LDA topic modeling and TextCNN sentiment classification on managerial disclosures (2000–2021), trained with ChatGPT-3.5 tokens; contributed to empirical analysis and robustness checks.

Proximity Advantage: Bank-Firm Distance and Corporate Digital Transformation, *Finance and Economics* 2024.3

With Weijian Su, Quanfei Zhang, **Yepeng Zhang**.
Main Findings: This study finds that greater distance from banks promotes corporate digital transformation, especially by encouraging the use of digital financial applications.
Responsibilities: Measured digitalization indices using text analytics on annual reports; implemented empirical models and mechanism tests.

Research assistant work for project **Moving to Opportunity for Polluting: Intra-City Evidence from China′s Land Market**, *now R&R at Journal of Development Economics*

Main Findings: Constructing grid level economic indicators (pixel level of Nighttime Light, GDP, Settlement Volume, Settlement Build-up) to support the empirical analysis

Working Papers

The Spatial Consequence of Hukou Reform, *SSRN abstract id: 5337216* 2025.7

With Da Fang, Qin Yu, Fan Zhang, **Yepeng Zhang**.

Main Findings: Using China's 2014 Hukou reform as a quasi-experiment, this DID study shows that relaxing migration rules led to more fragmented urban spatial forms.

Responsibilities: Co-first author; responsible for conceptualization, data processing, core metrics construction, empirical analysis, and code maintenance.

Adding Fuel to the Fire: Does Temperature Affect the Carbon Emission Intensity of Firms?, *Energy Economics* **(Under Review)** 2025.4

With Qihang Li, Quanfei Zhang, **Yepeng Zhang**, Da Fang.

Main Findings: This study finds that high temperatures increase firms' carbon emission intensity by reducing labor productivity and spurring mechanization, creating a "vicious cycle." Market-based carbon trading policies can help break this cycle.

Responsibilities: Data collection and cleaning; indicator construction; robustness and heterogeneity tests; literature review.

How Does Urban Morphology Affect Entrepreneurship?, *Applied Economics* **(Under Review)** 2025.9

With Qihang Li, Linman Zheng, **Yepeng Zhang**, Da Fang.

Main Findings: This study finds that cities with more compact spatial forms have higher levels of corporate entrepreneurship, facilitated by better resource allocation, innovation, and agglomeration.

Responsibilities: Developed core metrics and instrumental variables; processed spatiotemporal big data and implemented empirical analysis.

Carbon Emissions within Compact Cities: Evidence from China′ s Urban Forms, *Journal of Urban Planning and Development* **(Revise and Resubmit)** 2025.10

With Qihang Li, Linman Zheng, Haoxiang Zhang, **Yepeng Zhang**.

Main Findings: This study shows that compact urban form significantly reduces industrial carbon emissions through agglomeration, technology diffusion, and better resource allocation.

Responsibilities: Indicator construction; dataset preparation; core econometric estimation and mechanism analyses.

The Impact of Digital Transformation on Corporate Financing Constraints: A Measurement Approach Based on an LDA Topic Model *Outstanding Graduation Thesis*

Main Findings: Constructed a firm-level digital transformation index based on text mining of annual reports and examined its impact on financing constraints using panel regression models.

Responsibilities: Sole author; responsible for indicator design, empirical analysis, and thesis writing.

Research Projects

Road Network Assessment, Optimization, and Urbanization Analysis in Developing Countries July 2025 – Present

- **Road Network Data Construction and Accessibility Analysis:** Annotated and corrected vector road data for developing countries using satellite imagery to build a standardized road network database. Calculated travel time between major economic nodes based on a raster-based cost-weighted algorithm to quantify regional transportation accessibility.
- **Market Access and Macroeconomic Impact Research:** Integrated multi-source raster data, including land use, topography, and nighttime light data, to construct a market potential indicator. Employed spatial econometric methods to evaluate the causal effects of road construction and upgrades on regional economic growth, industrial layout, and urbanization processes.
- **Road Network Optimization Algorithm Design and Simulation:** Developed an optimal path planning model targeting logistics efficiency maximization. Proposed optimization strategies for the existing road network layout from a social welfare perspective to provide decision support for infrastructure planning.
- **Country-specific Case Studies (Mozambique, Zambia, Pakistan, China, Central Asia):**
  - *Mozambique (N6 Corridor):* Investigated the impacts of the N6 road expansion project on local urbanization, economic development, employment, and school enrollment. Used Google Earth Pro to inspect and evaluate road upgrades, implemented shortest-path algorithms, and computed commuting time costs between key locations to build core accessibility indicators.
  - *Zambia:* Constructed region-specific geospatial indicators and harmonized multi-period road network data to track network extension over time. Calculated travel-time costs between major economic, political, and cultural centers, and related them to patterns of urbanization, economic development, and population agglomeration.
  - *Pakistan:* Analyzed the staged development of Pakistan's road network and its temporal evolution. Combined built-up capacity, population distribution, and economic development measures within urban areas to study the link between road expansion and urbanization, and explored potential underlying mechanisms.
  - *China:* Compiled multi-source data and modeling results to reconstruct nearly 50 years of urban built-up area expansion and intra-urban indicators. Linked these to road network development to deepen understanding of the drivers, consequences, and distinctive features of China's urbanization process.
  - *Central Asian Countries and Belt and Road Trade:* Matched China's customs export data with rail and road corridors to analyze how Belt and Road infrastructure has strengthened trade connections between China and Central Asian countries and supported the growth and development of cities along these routes.

Research on Travel Behavior Patterns and Environmental Effects Based on Micro-level Big Data July 2025 – Present

- **Truck GPS Trajectory Mining and Feature Identification:** Processed hundreds of millions of truck GPS points; used map-matching to align trajectories with highways and identify OD, stops, and travel speeds.
- **High-Frequency Human Mobility Dynamics Analysis:** Cleaned large-scale human OD data and applied time-series analysis to characterize spatiotemporal mobility patterns and the effects of external shocks (e.g., policies, weather).
- **Finely Resolved Spatial Quantification of Traffic Emissions:** Built a 1km×1km national highway grid, linked truck traffic density with satellite NO_2 data, and used panel regressions to estimate freight contributions to local air pollution.
- **Project Cases:**
  - *Beijing Trucks and Emissions:* Linked Google Earth Engine grid-level air pollution data with truck trajectories; constructed indicators of truck behavior and emissions under different policies, time periods, and shocks to profile freight–environment interactions.
  - *People Flow Patterns from Navigation Data:* Analyzed Amap-style grid-aggregated navigation records with flexible time windows (hours to days) to track evolving travel patterns, combining trip counts with landmarks and locations to describe joint spatiotemporal flows.

Urban Spatial Structure Analysis and Indicator Construction Based on Multi-source Spatiotemporal Data January 2023 – June 2025

- **Multi-source Spatial Data Fusion and Indicator Construction:** Fused DMSP/OLS nighttime lights, LandScan population, OSM roads, and Amap POIs to construct key spatial metrics such as the Urban Boundary Compactness Index (Shape Index) and Spatial Population Concentration (SPC).
- **Spatial Analysis and Applied Research:** Used these indicators plus elevation and statistical yearbook data to study how urban spatial form (e.g., compactness) affects economic activity and to derive valid instrumental variables for causal inference.
- **Infrastructure Accessibility and Density Calculation:** Computed highway accessibility, road network density, and public facility densities for industrial enterprises and urban areas to support infrastructure-related research.

Research on Impact Mechanisms of Housing Prices on Manufacturing Investment from Factor Mobility Perspective: funded by *2024 National Social Science Foundation Project* 2024.8

- **Project Overview:** Studied how Chinese housing price fluctuations affect factor mobility in manufacturing enterprises and economic outcomes, showing cost, financing, and expectation channels and the need to rebalance real estate and the real economy.
- **Primary Contributions:** Authored Chapter 5 "Data Processing and Indicator Calculation," analyzing drivers of housing prices and building integrated indicators of urban development, population spatial distribution, and government sentiment using raster/vector (population, nighttime lights) and textual big data.

Measuring AI Exposure: Analysis of LLM Impact on China's Labor Market 2024.7 - Present

- **Model Deployment and Fine-tuning:** Deployed and fine-tuned LLMs with domain-adapted text and instructions to enhance robustness and accuracy in semantic matching and indicator construction tasks.
- **AI Exposure Measurement and Application:** Built generative AI exposure metrics by linking DeepSeek API outputs with China's Occupational Classification Standard, providing a basis for assessing generative AI's potential impact on China's labor market.

Technical Skills

English Proficiency: IELTS: 7

Programming: Python, R, Stata, Matlab

Data Analysis: Data science/web scraping (***pandas, scrapy***); GIS (***Geopandas, QGIS***); visualization (***matplotlib, ggplot***); ML/DL modeling (***sklearn, pytorch***); ML/DL in text analysis; applied econometrics and causal inference in Stata (DID, RD, IV).

Conferences & Competitions

2025 Interdisciplinary Contest in Modeling (*Evaluation of Key Transportation Infrastructure in Baltimore*), *Student Advisor*, **Finalist** 2025.5

2022 Interdisciplinary Contest in Modeling, **Honorable Mention** 2022.5

2021 Interdisciplinary Contest in Modeling, **Honorable Mention** 2021.5

Presenter, Camphor Economics Seminar (Development Economics) 2024.9

Presenter, Sixth Urban Economics Forum 2024.12

Presenter, Fourth China Industrial Economics Conference 2022.7

Presenter, 21st China Economics Annual Conference 2021.12

Participant, Fourth Future Economists Forum 2021.12