Environmental Emissions, Economic Linkages, and Public Health: An Input-Output Based Data Science Approach

Jiacheng Zheng

PhD Application Presentation
karcenzheng17@gmail.com

June 5, 2025

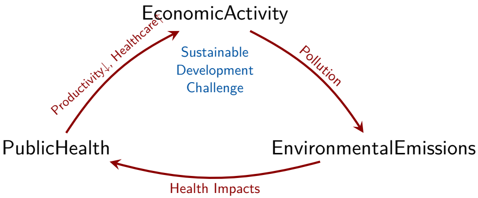
Outline

The Interconnected Challenges of Development

Economic activities are fundamental to societal well-being, but often lead to:

- **Environmental Degradation:** Increased emissions of CO_2, particulate matter ($PM_{2.5}$), SO_x, NO_x, etc.
- **Public Health Burdens:** These pollutants are major contributors to respiratory diseases, cardiovascular problems, cancers, and premature mortality.

Understanding the pathways from **economic production** $\rightarrow$ **emissions** $\rightarrow$ **human exposure** $\rightarrow$ **health impacts** $\rightarrow$ **economic costs** is crucial for sustainable development.



Research Proposition

My research proposes an **Input-Output (I-O) based framework**, augmented by data science techniques, to quantify these linkages and inform evidence-based sustainable development policy.

My Journey Towards Integrated Assessment

- **Foundational Education:**
  - Bioengineering (SDU): Provides understanding of human physiology and health impacts.
  - Chemical Engineering (DUT): Strong quantitative and process-oriented analytical skills.
- **Core Expertise in Economic Modeling & Data Science:**
  - **Input-Output Analysis:** Extensive experience as Research Assistant (SDU) with Dr. Jamal Khan and Prof. Yuan Li, working on I-O models, sectoral linkages (e.g., "Sectoral growth dynamics...in Pakistan", "Linkages...in Chinese financial sector"). Presented I-O research at the IIOA Conference.
  - **Data Science:** Proficient in programming (Python, R, Stata), statistical analysis, ML/NLP techniques, and data visualization applied to economic and environmental data (SDU, UCL climate change work).
  - *This PhD aims to synthesize these distinct skill sets to address the complex environment-economy-health nexus.*

My background is uniquely suited to pioneering research in this interdisciplinary domain.

Strong Quantitative and Analytical Training

Shandong University (SDU) - Bioengineering (2023 – 2026 expected)

GPA: 3.384 / 4.0. Rank: 6 / 50.

Relevant coursework: General Biology, Biochemistry, Microbiology, and Cell Biology.

Mathematics-related grades:

- Calculus I: **98/100**
- Linear Algebra: **90/100**

Dalian University of Technology (DUT) - Chemical Engineering (2019 – 2021)

GPA: 3.52 / 4.0. Rank: 17 / 230.

Key mathematical and physical science skills, demonstrated by:

- Calculus II: **92/100**
- Probability and Mathematical Statistics Theory: **88/100**
- College Physics (Classical Mechanics and Optics): **98/100**
- College Physics (Electromagnetism and Quantum Mechanics): **100/100**

Direct Experience in Input-Output Modeling and Data Analysis

- **Research Assistant, SDU (Institute of International Studies, 2023-Present):**
  - **Focus:** Input-Output modeling, sectoral analysis, economic dynamics, data science applications (ML/NLP).
  - *Key Projects directly relevant to proposal:*
    - "Sectoral growth dynamics, domestic linkages, and GVC movement in Pakistan: A single and multi-regional input-output modeling approach." (Under Review)
    - "Linkages and structural changes in the Chinese financial sector, 1996–2018: A network and input–output approach." (Published SCED, IIOA Conference 2025)
  - *Skills gained:* Constructing and analyzing I-O tables, structural decomposition analysis, network analysis of economic systems, programming for large-scale data.
- **Research Assistant, University College London (Climate Change Economics, 2024):**
  - *Contribution:* IPCC database development; analysis for "The effectiveness of unilateral vs. multilateral carbon adjustment mechanisms..."
  - *Relevance:* Experience with **environmental economic data, emissions accounting, and climate policy analysis**, complementing I-O skills.

HKUST Experience

- **Research Intern, The Hong Kong University of Science and Technology (HKUST), 2025.02–2025.05:**
  - *Focus:* Data science, large-scale data alignment, and AI research collaboration.
  - *Collaboration:* Worked closely with **Jiaming Ji** (Apple Scholar, Peking University) and **Boyuan Chen** (Peking University) on NeurIPS 2025 submission.
  - *Relevance:* Strengthened quantitative research skills, exposure to advanced machine learning pipelines and multi-institutional scientific collaboration; experience applicable to modeling tasks in carbon accounting and global value chain restructuring.

Demonstrated Output in Relevant Fields (Selected)

My research contributions highlight my capabilities in I-O analysis and quantitative economics:

- Khan, J., Li, Y., & **Zheng, J.** (Under Review). *Sectoral growth dynamics, domestic linkages, and GVC movement in Pakistan: A single and multi-regional input-output modeling approach*. Review of Development Economics.
- Khan, J., Li, Y., & Mahsud, Q. J. (2024). *Linkages and structural changes in the Chinese financial sector, 1996–2018: A network and input–output approach*. Structural Change and Economic Dynamics, 70, 33–44. (Acknowledged).
- Xiao, X., Khan, J., & **Zheng, J.** (Under Review). *Structural changes and sectoral linkages: Unveiling the ICT sector's role in China and India*. Telecommunication Policy.
- **Conference:** 31st International Input-Output Association Conference, Malé, Oral Presentation, *Linkages and Structural Changes in the Chinese Financial Sector, 1996–2018*, 2025.

These demonstrate my ability to conduct and disseminate complex I-O based research. Full list in CV.

Technical Skillset

Modeling and Quantitative Analysis

- **Input–Output Analysis (Advanced):** Model construction, Leontief inverse, structural decomposition, GVCs Decomposition.
- Statistical inference, hypothesis testing.
- Economic network analysis.

Programming and Software

- **Python (Advanced):** pandas, NumPy, scikit-learn, matplotlib.
- **R (Intermediate):** Data visualization.
- **MATLAB (Intermediate):** Matrix modeling and simulations.

Data Science and Applied Domains

- Data wrangling and preprocessing.
- Applied machine learning: regression, classification.
- Visualization of complex and high-dimensional datasets.

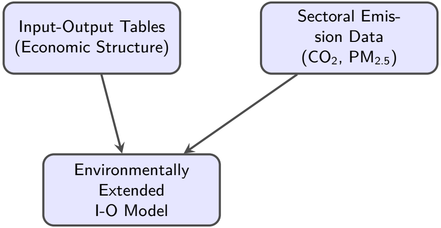
Languages and Communication

- Chinese (Native), English (IELTS 7.0 overall).

Proposal: Linking Emissions, Economy, and Health via I-O & Data Science

Central Aim: To develop and apply an integrated framework using Environmentally Extended Input-Output (EEIO) models, health impact assessment methodologies, and data science techniques to quantify the public health burdens attributable to sectoral emissions (CO_2 and other key air pollutants like $PM_{2.5}$, SO_x, NO_x).

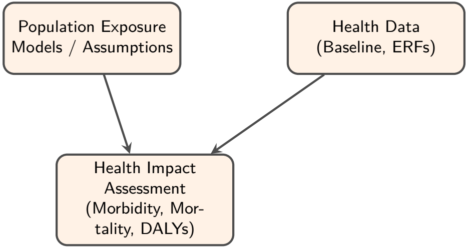
Integrated Assessment Framework: Step 1



Key Inputs:

- Economic structure via Input-Output tables
- Emissions data (CO2, PM2.5, etc.) by sector
- Integration into environmentally extended I-O model for footprint analysis

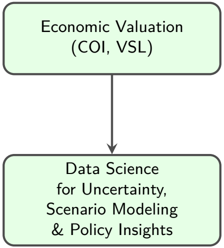
Integrated Assessment Framework: Step 2



Health Module:

- Population exposure modeled by emissions footprints
- Health baseline and exposure-response functions (ERFs)
- Resulting health impacts: morbidity, mortality, DALYs

Integrated Assessment Framework: Step 3



Economic & Policy Module:

- Monetizing health impacts (Cost of Illness, Value of Statistical Life)
- Applying data science for uncertainty quantification, scenario modeling
- Informing sustainable development policies

Guiding the Investigation

This research will address the following KRQs:

1. **KRQ1 (Methodological Integration):** How can national/regional I-O tables be effectively integrated with detailed sectoral environmental emission accounts (CO_2, $PM_{2.5}$, SO_x, NO_x) and health impact assessment models to create a robust Environment-Economy-Health analytical framework?
2. **KRQ2 (Empirical Quantification - Sectoral Contributions):** Which economic sectors are the most significant contributors (directly and indirectly via supply chains) to the public health burden attributable to these emissions in a chosen case study region (e.g., China, or a specific province/megacity cluster)?
3. **KRQ3 (Economic Valuation of Health Impacts):** What are the monetized values of the health impacts (e.g., mortality, morbidity, DALYs converted to economic costs) linked to sectoral emissions, and how do these compare to the economic output or value-added of these sectors?
4. **KRQ4 (Policy & Data Science Application):** How can data science techniques (e.g., sensitivity analysis, machine learning for predictive sub-models, advanced visualization) enhance the framework's utility for evaluating policy scenarios (e.g., carbon pricing, sectoral technology upgrades, pollution control measures) aimed at mitigating health impacts?

A Phased Research Approach

1. **Phase 1: EEIO Model Construction & Emission Accounting:**
   - Compile I-O tables (e.g., China's national/provincial tables, or global MRIO like GTAP, WIOD).
   - Develop satellite accounts for emissions (CO_2, $PM_{2.5}$, SO_x, NO_x) by I-O sector using national statistics, EDGAR, CEDS, or regional inventories.
   - Calculate direct and total (Leontief-inverse based) emission intensities and footprints.
2. **Phase 2: Health Impact Assessment Module:**
   - Define population exposure pathways (e.g., linking sectoral emissions to ambient concentrations, potentially using simplified dispersion or intake fraction methods if detailed modeling is out of scope).
   - Collect demographic data and baseline health statistics for the study region.
   - Utilize established Exposure-Response Functions (ERFs) from WHO, IHME Global Burden of Disease studies, or relevant epidemiological literature to link pollutant exposure to health outcomes (e.g., attributable deaths, YLLs, incidence of specific diseases).

A Phased Research Approach

1. **Phase 3: Economic Valuation of Health Damages:**
   - Apply economic valuation techniques: Cost of Illness (COI), Value of a Statistical Life (VSL), or DALY-to-monetary conversion, considering regional context.
2. **Phase 4: Data Science Application, Scenario Analysis & Policy Implications:**
   - Employ statistical methods for uncertainty and sensitivity analysis of the integrated model.
   - Use data science tools for managing large datasets, visualization, and potentially ML for sub-component modeling (e.g., predicting emission changes under policy).
   - Analyze policy scenarios (e.g., carbon tax, renewable energy adoption, pollution abatement technologies in key sectors) and their co-benefits for health.

Building the Data Foundation

Table: Potential Data Sources for the Integrated Model

Data Type	Description	Potential Sources
Input-Output Tables	National, Regional, or Global MRIO	NBS China, OECD, WIOD, GTAP, Eora
Emission Data	CO_2, $PM_{2.5}$, SO_x, NO_x by sector	National Env. Statistics, EDGAR, CEDS, REAS, MEIC
Health Data	Baseline mortality/morbidity rates, DALYs	WHO, IHME GBD, National Health Stats
Population Data	Demographics, spatial distribution	National Census, World Bank, SEDAC
Exposure-Response	Coefficients linking exposure to health outcome	WHO, GBD studies, Peer-reviewed literature
Economic Valuation	VSL, COI parameters	OECD, Local studies, Literature review
Geospatial Data	Administrative boundaries, population density	Natural Earth, GADM, National mapping agencies

Data collection, harmonization, and gap-filling (using statistical methods) will be crucial early steps. My data science skills will be essential here.

Research Timeline and Objectives

- **Year 1:**
  - Strengthen theoretical foundation related to environmental health, focusing on mechanisms of air pollutants (CO_2, $PM_{2.5}$, etc.) impacting public health.
  - Address gaps in undergraduate coursework in environmental epidemiology, statistical modeling, and machine learning to enhance data processing and programming skills.
- **Year 2:**
  - Develop predictive tools based on input-output tables to estimate emission-exposure-health impact pathways quantitatively.
- **Year 3:**
  - Integrate multi-source data and global value chain analysis tools, utilizing AI technologies such as ChatGPT to streamline health impact assessment and automate policy report generation.

Research Timeline and Objectives

- **Year 4:**
  - Prepare doctoral dissertation titled "Optimization of Low-Carbon Value Chains under Multimodal Data Fusion: Health Impact Assessment of Gas Emissions";
  - Propose policy-relevant carbon emission and health impact accounting standards;
  - Aim to publish breakthrough findings in top-tier journals such as *Nature Sustainability*.
- **Year 5:**
  - Complete dissertation defense and compile research outcomes;
  - Publish a monograph, *Machine Learning Empower Economic and Environmental Health Research*;
  - Apply for postdoctoral positions to further pursue interdisciplinary research on gas emissions and public health.

Advancing Knowledge and Informing Policy

- **Methodological Advancements:**
  - A refined, operational integrated EEIO-Health assessment framework applicable to various regions.
  - Novel applications of data science techniques within this integrated modeling context.
- **Empirical Insights:**
  - Quantification of health burdens and economic costs attributable to sectoral emissions in the chosen case study.
  - Identification of critical sectors and supply chain pathways for targeted policy interventions.
- **Policy Relevance:**
  - Evidence base for designing co-benefit policies (climate/pollution mitigation and health improvement).
  - Tool for ex-ante assessment of different development and environmental policy scenarios.
- **Scholarly Output:** Publications in high-impact journals (e.g., Environmental Science & Technology, Lancet Planetary Health, Ecological Economics, Journal of Industrial Ecology, Energy Policy).

My strong background in I-O analysis, coupled with data science skills, is critical for achieving these contributions.

Synergies and Long-Term Vision

Why This PhD Program?

- **Faculty Expertise:** Eager to work with Dr. Juexiao Zhou due to your expertise in healthcare and machine learning background.
- **Research Environment:** You are free to work anywhere on campus during working hours. What matters to me is your research progress, your ability to graduate successfully, and your future career development, not how many hours you spend physically in the lab.
- **Specific Strengths:** Productivity & Technical Expertise, Science Taste and Direction Control, Unlimited Computational Resources, Global Collaboration Network, Money and Career Support.

Synergies and Long-Term Vision

My Future Goals

- To become a leading researcher in the field of integrated environmental-economic-health assessment.
- To develop and apply quantitative tools that support evidence-based policymaking for sustainable and healthy societies.
- To contribute to academic institutions or international organizations working on climate change, public health, and sustainable development.

This PhD is a pivotal step towards achieving these aspirations, and I am confident that my skills and research vision align with your program's strengths.

Thank You for Your Time and Consideration

Never Go Home!

Questions?

Jiacheng Zheng
karcenzheng17@gmail.com
https://github.com/Karcen