# Thesis Title: Methodological Innovation of the Input-Output Model and Its Applications in High Resolution Time-Series Economic Analysis

Kehan He

Bartlett School of Sustainable Construction

University College London

A thesis submitted for the degree of

**Doctor of Philosophy**

## Declaration

I, Kehan He, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

# Abstract

The Input-Output Model is a powerful economic tool used for economic structure analysis and extended applications. Economists often adopts the Input-Output Model to investigate the changes in economic structures and the associated social impacts. Taking China as a case study, this thesis first describes the compilation of China's 2017 Multi-Regional Input-Output table and its application in revealing the intensifying geographical inequality of China's carbon emissions. It however exposes some of the most widely criticized drawbacks of the Input-Output Model, being the high cost in data preparation, the oversimplification of economic symbiosis, and the lack in forecast ability. Thus, conventional Input-Output Model is unfortunately unsuited to high resolution time-series analysis.

This thesis hence proposes an innovative regression algorithm on a derivative of Input-Output model, the Sequential Interindustry Model, to incorporate high-frequency time domain into the analysis of the intersectoral interactions of economic sectors by regressions on the observations of outputs of different economic sectors. To test the efficacy of the innovated algorithm, the electricity consumption data of Chongqing municipality is used as a proxy to economic activities. The result verifies the validity of the innovated algorithm in time-lagged analysis of the economic sectors' interlinkages in the investigated region, and hence facilitate short future predictions of the economy under different scenario settings. Through the inclusion of investment as a parameter and learning from extended Input-Output Model methodological literatures, further algorithm innovation is introduced to a dynamic version of SIM to account for the indirect cost of 2015 South India flood across different regions.

Based on the series of venturing in the methodological advancement of Input-Output Model, this thesis concludes by a literature review and qualitative discussion on the directions of interdisciplinary development in Input-Output Model, which is the integration of system engineering techniques in the exploration of solutions to economic cybernetics.

# Impact Statement

This thesis presents a new economic analysis tool developed using innovative algorithms for short-term economic system analysis under a general disequilibrium assumption based on the Sequential Interindustry Model (SIM), a variant of the Input Output (IO) Model. The model is compatible with high time-resolution data, such as daily electricity consumption data by sectors, to reduce data collection costs. This tool has a wide range of applications beyond academia and can benefit policymakers, businesses of all scales, and economic planning authorities.

Specifically, the use of high time-resolution electricity consumption data for economic analysis reduces data collection costs as well as provides timely analysis with enriched information on chronological interindustry linkages. The analysis result using the algorithm developed in this research can thus provide cheap and timely decision advice to various stakeholders. For example, businesses at all scales can use the analysis result to adjust their production and inventory levels accordingly. Economic planning authorities can also use this algorithm to identify market trends and adjust intervention strategies in real-time. In addition, the economic analysis tool developed in this research can be useful for policymakers and business leaders to make short-term economic scenario predictions for various economic shock events by aiding decision-making for recovery plan formulation at a macro scale.

The methodological advancement of IO model proposed in this research introduces a hybridization of big data and the economic theory of IO, thus providing a new way of thinking about economic analysis that incorporates both traditional IO theory and cutting-edge data analysis techniques. The innovative algorithms and concepts introduced in this research can inspire new research and development in economic cybernetics and data science. It can lead to the creation of new research areas and interdisciplinary collaborations between economics and other fields such as computer science, engineering, and data science, hence promote a more comprehensive understanding of economic phenomena and generate new avenues of research for scholars in the field.

In addition, a 2017 provincial Multi-Regional Input Output (MRIO) table of China has been compiled in this research based on official data released by the statistic departments of China. The accuracy of the MRIO table has also been validated against other sources. This MRIO table of China can be used in future research to serve as a database for extended analysis. Upon completion of this thesis, two other studies have knowingly adopted the 2017 China MRIO table developed in this thesis.

The skills and knowledge on IO modelling acquired during the PhD have contributed significantly to published research that bear substantial policy implications. This includes two Lancet Countdown reports that analysed the impacts of climate change on global health and provided policy recommendations for addressing the health risks associated with climate change (Romanello et al., 2021, Romanello et al., 2022), one UNDP report that summarized the pros and cons for the policy tools in fossil fuel reforming (Marcel Alers and Ben-jamin Jones, 2021), one UN policy brief that accounted for the impact of Covid-19 pandemic on global carbon emissions (He and Mi, 2022), and several other peer-reviewed journal articles.

# PhD Publications

**Peer Reviewed Journals**

Romanello, M., [...] He, K. et al. (2023). The 2023 report of the Lancet Countdown on health and climate change: the imperative for a health-centred response in a world facing irreversible harms. *The Lancet*, 402 (10419), 2346-2394

Chen, Y., He, K.*, Muhammet, D. *, Coffman D.M.* (2023) Health impacts of bike sharing system – a case study in Shanghai, *Journal of Transport & Health*. 30(2023) 101611 (*Corresponding authors)

He, K., Mi, Z., Zhang, J., Li, J., & Coffman, D. M. (2023). The Polarizing Trend of Regional CO2 Emissions in China and Its Implications. *Environmental Science & Technology*. 2023 57 (11), 4406-4414

Romanello, M., [...] He, K. et al. (2022). The 2022 report of the Lancet Countdown on health and climate change: health at the mercy of fossil fuels. *The Lancet*, 400 (10363), 1619-1654

He, K., Mi, Z., Coffman, D. M., & Guan, D. (2022). Using a linear regression approach to sequential interindustry model for time-lagged economic impact analysis. *Structural Change and Economic Dynamics*. 2022, 62: 399-406

Romanello, M., [...] He, K. et al. (2021). The 2021 report of the Lancet Countdown on health and climate change: code red for a healthy future. *The Lancet*, *398*(10311), 1619-1662.

He, K., Mi, Z., Chen, L., Coffman, D. M., & Liang, S. (2021). Critical transmission sectors in embodied atmospheric mercury emission network in China. *Journal of Industrial Ecology*, 2021, *25*(6), 1644-1656.

**Working Papers:**

Mi, Z., Yang, J., Zheng, J., Tang, L., He, K., Li, L., Cai, W., Drummond, P. (2022) Assessing the impacts of different fertility and retirement policies on China's future carbon emissions. *Nature Climate Change* (Under Reviews)

He, K., Coffman, D. M., Hou, X., Mi, Z. (2022) Analyse the Chronological Interlinkage of Economic Sectors Using Sequential Interindustry Model. *Economic System Research*. (Accepted)

**Conference Proceedings**

He, K., & Mi, Z. Carbon Implications of COVID-19. (2022). *Science-Policy Brief for the Multistakeholder Forum on Science, Technology and Innovation for the SDGs, UN DESA*

# UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- have been uploaded to a preprint server;
- are in submission to a peer-reviewed publication;
- have been published in a peer-reviewed publication, e.g. journal, textbook.

*This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.*

| **1. For a research manuscript that has already been published** (if not yet published, please skip to section 2)**:** | |
|---|---|
| **a) Where was the work published?** (e.g. journal name) | Environmental Science & Technology |
| **b) Who published the work?** (e.g. Elsevier/Oxford University Press): | ACS Publications |
| **c) When was the work published?** | 28/02/2023 |
| **d) Was the work subject to academic peer review?** | Yes |
| **e) Have you retained the copyright for the work?** | Yes |
| [If no, please seek permission from the relevant publisher and check the box next to the below statement]:<br>☐ *I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.* | |

<table>
<tr><td colspan="4">2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3):</td></tr>
<tr><td colspan="2">a) Has the manuscript been uploaded to a preprint server? (e.g. medRxiv):</td><td></td><td>If yes, which server? Click or tap here to enter text.</td></tr>
<tr><td colspan="2">b) Where is the work intended to be published? (e.g. names of journals that you are planning to submit to)</td><td colspan="2"></td></tr>
<tr><td colspan="2">c) List the manuscript's authors in the intended authorship order:</td><td colspan="2"></td></tr>
<tr><td colspan="2">d) Stage of publication</td><td colspan="2">Please select.</td></tr>
<tr><td colspan="4">3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):</td></tr>
<tr><td colspan="4">Kehan He: First author, research idea, experiment design, conducting, writing of the paper. Zhifu Mi: Research Supervision, corresponding author Jin Zhang: Collaboration in writing Jinkai Li: Collaboration in writing D'Maris Coffman: Research Supervision</td></tr>
<tr><td colspan="4">4. In which chapter(s) of your thesis can this material be found?</td></tr>
<tr><td colspan="4">Chapter II</td></tr>
<tr><td colspan="4">5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work):</td></tr>
<tr><td>Candidate:</td><td>Kehan He</td><td>Date:</td><td>Click or tap to enter a date.</td></tr>
<tr><td>Supervisor/ Senior Author (where appropriate):</td><td>Zhifu Mi</td><td>Date:</td><td>Click or tap to enter a date.</td></tr>
</table>

# UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- have been uploaded to a preprint server;
- are in submission to a peer-reviewed publication;
- have been published in a peer-reviewed publication, e.g. journal, textbook.

*This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.*

| **1. For a research manuscript that has already been published** (if not yet published, please skip to section 2)**:** | |
|---|---|
| **a) Where was the work published?** (e.g. journal name) | Structural Change and Economic Dynamics |
| **b) Who published the work?** (e.g. Elsevier/Oxford University Press): | Elsevier |
| **c) When was the work published?** | 29/03/2022 |
| **d) Was the work subject to academic peer review?** | Yes |
| **e) Have you retained the copyright for the work?** | Yes |
| [If no, please seek permission from the relevant publisher and check the box next to the below statement]:<br>☐ *I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.* | |

| | | | | |
|---|---|---|---|---|
| **2. For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)**:** | | | | |
| **e) Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv): | | Please select. | **If yes, which server?** Click or tap here to enter text. | |
| **f) Where is the work intended to be published?** (e.g. names of journals that you are planning to submit to) | | Click or tap here to enter text. | | |
| **g) List the manuscript's authors in the intended authorship order:** | | Click or tap here to enter text. | | |
| **h) Stage of publication** | | Please select. | | |
| **3. For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4**):** | | | | |
| Kehan He: First author, research idea, experiment design, conducting, writing of the paper. Zhifu Mi: Research Supervision, corresponding author D'Maris Coffman: Research Supervision Dabo Guan: Research Supervision | | | | |
| **4. In which chapter(s) of your thesis can this material be found?** | | | | |
| Chapter III | | | | |
| **5. e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work): | | | | |
| **Candidate:** | Kehan He | **Date:** | Click or tap to enter a date. | |
| **Supervisor/ Senior Author** (where appropriate)**:** | Zhifu Mi | **Date:** | Click or tap to enter a date. | |

# UCL Research Paper Declaration Form: referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- have been uploaded to a preprint server;
- are in submission to a peer-reviewed publication;
- have been published in a peer-reviewed publication, e.g. journal, textbook.

*This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.*

| | |
|---|---|
| **1. For a research manuscript that has already been published** (if not yet published, please skip to section 2)**:** | |
| **a) Where was the work published?** (e.g. journal name) | Click or tap here to enter text. |
| **b) Who published the work?** (e.g. Elsevier/Oxford University Press): | Click or tap here to enter text. |
| **c) When was the work published?** | Click or tap to enter a date. |
| **d) Was the work subject to academic peer review?** | Please select. |
| **e) Have you retained the copyright for the work?** | Please select. |
| [If no, please seek permission from the relevant publisher and check the box next to the below statement]:<br>☐ *I acknowledge permission of the publisher named under 1b to include in this thesis portions of the publication named as included in 1a.* | |

| | | | |
|---|---|---|---|
| **2. For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)**:** | | | |
| **i) Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv): | No | | **If yes, which server?** Click or tap here to enter text. |
| **j) Where is the work intended to be published?** (e.g. names of journals that you are planning to submit to) | Economic System Research | | |
| **k) List the manuscript's authors in the intended authorship order:** | Kehan He, D'Maris Coffman, Xingzhe Hou, Zhifu Mi | | |
| **l) Stage of publication** | Submitted | | |
| **3. For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4**):** | | | |
| Kehan He: First author, research idea, experiment design, conducting, writing of the paper. D'Maris Coffman: Research Supervision Xingzhe Hou: Data provision/collaboration Zhifu Mi: Research Supervision, corresponding author | | | |
| **4. In which chapter(s) of your thesis can this material be found?** | | | |
| Chapter IV | | | |
| **5. e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work): | | | |
| **Candidate:** | Kehan He | **Date:** | Click or tap to enter a date. |
| **Supervisor/ Senior Author** (where appropriate)**:** | Zhifu Mi | **Date:** | Click or tap to enter a date. |

# Acknowledgement

Earning a PhD is no simple endeavour, yet I have been privileged to receive generous support and guidance from a remarkable group of individuals. At the forefront, I am profoundly thankful to my supervisors, Prof. Zhifu Mi and Prof. D'Maris Coffman. The challenges posed by the Covid-19 pandemic made supervision tasks considerably more difficult, yet they both offered their invaluable expertise and steadfast guidance. Zhifu has been not only an enlightened mentor but also a close friend, generously sharing insights about professional and personal life. His meticulous guidance was instrumental in the completion of this thesis and his teachings extended beyond the academic realm, offering lessons on how to balance the roles of a researcher and a responsible father. D'Maris's astute advice significantly enhanced the quality of my work. Her guidance is also crucial in my career development. The constant encouragement from both supervisors has truly broadened my intellectual horizons, and their wisdom passed to me is the most rewarding part of my PhD journey.

I am deeply appreciative of the insightful critiques and valuable recommendations offered by the examiners, Prof. Dabo Guan, Prof. Edgar Hertwich, and Prof. Shen Wei. Dabo's constructive advice during my PhD upgrade was instrumental in refining my research plan. Edgar and Shen's pertinent queries and comments during my thesis examination were extremely helpful in improving my work.

My time at the Bartlett School of Sustainable Construction has been an extraordinarily enriching experience. I am grateful for the financial support, the stimulating academic environment, and the opportunity to work alongside talented and inspiring colleagues. The technical and administrative staff have also offered immense support, particularly during the challenges presented by the global Covid-19 pandemic.

My heartfelt appreciation also extends to my friends and colleagues met during this journey. Their companionship and encouraging words, particularly during trying times, have made this journey immeasurably more enjoyable. I am also grateful to the colleagues and peers I encountered at the International Input-Output Association Conferences. Their professional advice and innovative research were a great source of inspiration for the completion of this thesis.

Lastly, my deepest gratitude is reserved for my family. I thank my parents for instilling in me the virtues of hard work and perseverance, and for their unwavering love and support throughout this journey. I am grateful to my spouse, Ruizhi, for being an unwavering pillar of support amidst life's

challenges. And to my son, who is the most precious gift I received during my PhD journey.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1: Introduction

## 1.1. The Input-Output Method

Imagine a being from the cosmos gazing upon our tiny blue planet; they would be astounded by the complexity and intricacy of our ecosystem. What's particularly striking is how humans interact with one another and the environment, creating a unique, interdependent system that continually fascinates us. This system, referred to as "economics," has captivated scholars across civilizations, who dedicate their intellect to understanding its workings in the hope of satisfying humanity's ever-growing needs, while grappling with limited resources and technology.

The practice of accounting for economic production activities can be traced back to the ancient Babylonian Empire, where people documented the efficiency of agricultural activities using ratios. Over the past five thousand years, nearly every civilization has developed and adopted some form of national accounting system or philosophical concept to assist governing authorities in economic planning. Instead of being deemed as a subject to manage the king's treasury, the concept of national accounting started its formalization in 1600s as a study on the operation of the economy. As one of the pioneers in modern economic theorist and national statistics, William Petty's 'Political Arithmetik' has consolidated several key theories of classical economics, such as interest rate, theory of value, and division of labour (Hoppit, 1996). Gregory King and Charles Davenant, one generation later, have also developed the foundations of price and trade theories (Evans, 1967), which are all then included in Adam Smith's magnum opus work 'Wealth of Nations' as the beginning of classical economic studies.

Other than the contributions from thinkers of the United Kingdom, the physiocracy theory developments of French philosopher François Quesnay, such as laissez-faire and views on labour, have also contributed towards the foundation of classical economic theories. Among his various contributions as an economist, François Quesnay has also completed the ground-breaking work of consolidating the production relationships between the agricultural/mining sector and the artisanal/manufacturing sector in eighteenth-century France. He developed the *tableau économique* to represent the flow of goods and services between these sectors, providing a comprehensive understanding of the economic interdependencies at play. Quesnay's work laid the foundation for future developments in economic accounting and analysis, especially in Karl Marx's Das Kapital to explain his theory of economic circulation (Gehrke and Kurz, 2002). By mapping the connections between different sectors, Quesnay's *tableau économique* offered valuable insights into the structure and functioning of the economy, enabling more informed decision-making and policy development for the betterment of society of his time (Leontief, 1953, Coffman, 2021).

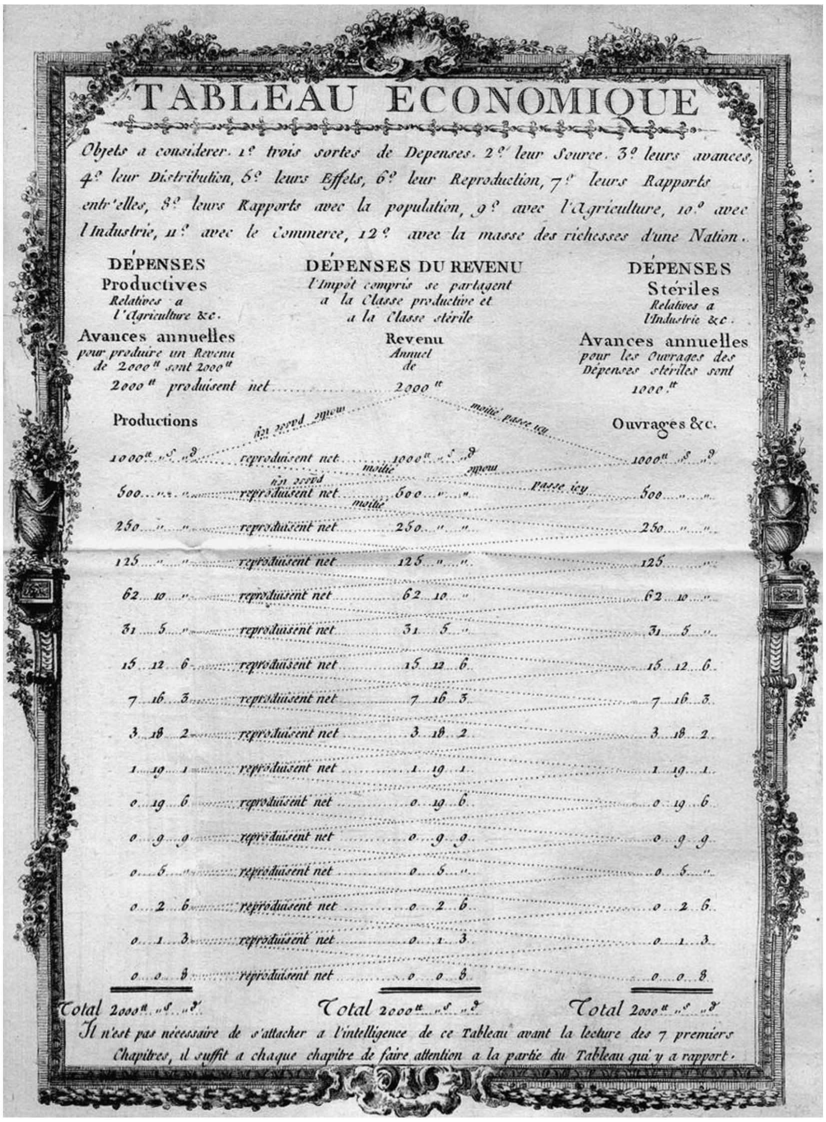



*Figure 1 The zigzag tableau économique proposed by François Quesnay (Charles, 2004).*

Building upon the concept of the *tableau économique* and its further adoption in Marxism economics, the Russian-American economist Wassily Leontief, who was later awarded the Nobel Prize of Economics, introduced the Input-Output (IO) Model in the 1930s as an economic statistical tool designed to elucidate the interdependent relationships between various sectors of the United States' economy (Leontief, 1936). Drawing from extensive surveys and governmental accounting data, the IO Model bridges the gap between micro-level production and macro-level economic flows. This innovative approach provides policymakers with a powerful instrument to analyse the functioning of past economies, enabling them to develop more targeted macroeconomic strategies.

By offering a comprehensive understanding of how different sectors interact and influence each other, the IO Model has been proven valuable in shaping economic policies that promote growth, stability, and the overall well-being of society.

Since the 1950s, statistical bureaus of governments worldwide have embraced the construction of national IO Tables as a primary method for implementing the System of National Accounts to quantitatively measure and analyse the performance of economies. To facilitate this process, the United Nations Statistical Commission has published official standards for compiling IO Tables (United Nations Statistical Division, 1999), which are the most reliable data widely used in modelling of economic structures and sectorial symbiosis. It serves as an alternative to mainstream economic researchers such as the Austrian School and Chicago School, who are fundamentally based on neoclassic economic theories and focus their research and debate on more specific issues such as setting of monetary policies and role of governments (Salin, 2023, Murphy, 2011). By combining statistical approach with a model theory outside classic economics, the IO model offers economists and policy makers a key tool to understand the functioning and well-being of nations, and even the global community.

Through IO analysis, economists can gain insights into economic analysis and planning, shedding light on the intricate inter-sectoral connections within an economy. It aids policymakers in formulating economic strategies, assessing the impact of economic shocks or policy alterations, and understanding structural relationships among different sectors. Moreover, it's instrumental in regional economic planning, optimizing resource allocation, and evaluating environmental impacts of economic activities. The IO model facilitates economic contribution and multiplier analysis, supply chain dynamics exploration, educational learning about economic interdependencies, international trade analysis, and industrial strategy development (Tan et al., 2019, Miller and Blair, 2009). Its extensive applicability makes it an invaluable tool for economists, policymakers, and analysts, allowing for a comprehensive examination of complex economic relationships and well-informed decision-making across diverse economic and policy realms.

However, due to the labour-intensive and computationally demanding nature of compiling IO Tables, the resources and capacity required may be unattainable for less developed countries (Bickel, 1987). For countries that regularly compile national IO Tables, the feasible intervals between table preparations typically range from 5 to 7 years. Even for larger countries with the capacity to produce such tables, like China, India, and the United States, the creation of multi-regional IO Tables for subnational regions can be challenging using the recommended accounting process. As a result, scholars often resort to algorithmic estimations to generate the desired data. This highlights the need for continued development of more efficient methods and tools to facilitate analysis of economic structures for all nations, ensuring that policymakers have access to the most accurate and swift information to guide their decision-making processes.

In addition to cost of data collection, another significant concern pertains to the underlying assumptions of the model (Ackerman et al., 2004). When compiling an IO table, it is assumed that an equilibrium state is reached during the statistical year, aligning with Walras' general equilibrium theory. As such, the IO model can be considered a simplified version of the Walrasian general equilibrium model (Akhabbar and Lallement, 2010). As a Walrasian general equilibrium model, the IO model simulates a balanced economic structure, meaning that the input and output, or supply and demand, are always matched (Miller and Blair, 2009). However, the validity of the general equilibrium assumption has been called into question, particularly in the context of short-term economic performance. While it may hold true in the long term as classic economic theories has suggested, it is less likely that supply and demand will be perfectly balanced within shorter time frames, such as in weeks and months. In other words, the economy is in a state of continuous instability and fluctuation (Debreu, 1974). These critiques highlight the need for further refinement and development of economic models that can more accurately capture the nuances and complexities of real-world economic systems. By addressing these limitations, economists and policymakers can gain a more comprehensive understanding of economic interactions and make more informed decisions that contribute to the overall well-being of society.

On the other hand, modern economic research has evolved from being purely theoretical to embracing evidence-based statistical research or econometrics (Frisch, 1970), as widely promoted by the Chicago School economists (Hamermesh, 2013). Conventional econometric research relies on economic data such as Gross Domestic Product (GDP), Manufacturing Index, and others. Economists, especially from the Chicago School, typically make extensive use of empirical data and mathematical models in their research and theories (Peck et al., 2011). They heavily rely on regression calculations on these economic parameters as they believe in the efficacy of quantitative analysis to understand and predict economic phenomena. This belief is criticised by other economists, specifically the Austrian School. Austrian economists tend to be more sceptical of the value of econometric models and empirical data, arguing that they can often misrepresent the complexity of human economic behaviour (Rosen, 1997). The IO model, however, is a combination of the two. As a praxeology model that accounts for industrial symbiosis, the IO model is built on statistically compiled IO tables, which serve as a reliable source of quantitative evidence to support IO analysis. If further developed in its data sources, the IO model bears the potential to reconcile the dilemma between the incapability of theoretical economics and data insufficiency of empirical economics.

## 1.2. Motivation

To ameliorate the statistical cost barrier of IO modelling and economic modelling in general, a possible approach is to integrate techniques in data science research, a subject advancing quickly owing to the rapid development of information and communications technologies. It is argued by some scholars that our increased capability in dealing with data has enabled us to collect and analyse empirical evidence at an unprecedented speed to support economic research (Varian, 2014). With the convenience of much lowered time and labour cost of data, the time resolution of the IO model can be improved to be daily or weekly instead of quarterly or annually in conventional economic research, thus empower more swift and accurate policy making.

Meanwhile, the economic system tends to fluctuate more volatilely in shorter time span, which further jeopardizes the classic general equilibrium assumption of the IO model. Since current focuses of policy formulations are mostly under long term settings, researchers generally lack the motivation to create an additional quantitative tool that examines the performance of a disequilibrium economic system in high time frequency. Hence, for the IO model to take advantage of faster data, fundamental revision and improvement are needed by the IO model to simulate rapid chronological interactions among economic sectors.

Nevertheless, the methodological development to increase the time resolution of IO model is very limited. Avelino (2017) proposed a method to disintegrate annual IO tables into intra-year tables by adjusting with intra-year GDP data and balance using the T-EURO method, but still limited its time resolution to quarterly collected GDP data. Holý and Šafr (2023) approached the same objective by using the RAS method, an alternative balancing method for IO tables. Zheng et al. (2018) and similar type of works resort to pure mathematic algorithms to extend IO tables and thus compensate for the missed data in the time series. As representations of their kinds, all approaches aim to intraplate or extrapolate annual IO tables and thus conduct analysis at higher frequencies. However, such attempts are still based on traditional economic indicator data, whose time resolution are fundamentally limited by costs and cannot be improved without the support of traditional statistic data (Su and Ang, 2022). Other researchers adopt improved versions of IO model that moderately considers time dimension in their simulations. For instance, Pagsuyoin et al. (2019) applied the Inoperability IO model as an improved IO model to simulate the daily impact induced by drought events in Massachusetts of the United States. The Inoperability IO model is also applied by Okuyama and Yu (2019) to assess the chronological economic impact of Kobe Earthquake in Japan. Hallegatte (2008) developed the Adaptive Regional IO model to simulate the monthly economic impacts of Hurricane Katrina into the future. Wang et al. (2023) used the average inventory turnover times of Chinese companies as auxiliary data to supplement IO modelling and hence assess the time lag effect of China's industry carbon emissions. Nevertheless, these modelling improvements and related works are based on strong assumptions. Compared to the empirical evidence based praxeological research in the field of data science, the IO model needs to be improved so that it can be more strongly supported by high time resolution data.

In the later stages of his research, Leontief emphasized that the IO model is not merely an economic method but can also be approached from a technical or engineering perspective (Leontief, 1991). The same idea has also been echoed among some other scholars (Tan et al., 2017). Building on this insight, it is necessary to create a bridge between the theories and models of data science and economic IO modelling.

## 1.3. Research Aim and Objectives

This research aims to examine the IO model through the lens of data science and system engineering. By incorporating innovative algorithms into the latest variant of the IO model, an economic tool based on the IO model capable of chronological interindustry analysis in a short-term disequilibrium state will be developed and implemented. The innovative model will be compatible with high-frequency data, allowing for the determination of cross-sectoral linkages over time. This will enable modellers to investigate the chronologically accumulated impact among economic sectors and offer a dynamic analysis of the economy from a perspective distinct from most current economic research. The innovative model will serve as a powerful tool for rapid and short-term economic analysis at low statistic cost. Rather than providing an account solely for the final state, this model can deliver a description of all sectors at each discrete step, better equipping policymakers to diagnose the functioning of an economy and make informed decisions.

Specifically, the following objectives will be achieved in this research:

- Demonstrate the application of the IO model by compiling the 2017 China multi-regional IO table. This will reveal the growing inequality of $CO_2$ emissions embedded in China's consumption patterns, showcasing an extended application of the IO model that is commonly utilized (Chapter 2).
- Develop an algorithm based on the Sequential Interindustry Model (SIM), a time domain variant of the IO model, capable of reverse-calculating the

IO relationships of economic sectors over a period using output observations (Chapter 3).

- Apply the developed algorithm to a case study of Chongqing, using daily sectorial electricity consumption data to demonstrate its feasibility. Conduct an analysis of time-delayed economic performance and provide short-term predictions for economic performance under various economic shock event scenarios (Chapter 4).
- Integrate the algorithm with the latest advancements in IO methodological innovation, ensuring that the impact of capital investment is considered. Conduct an analysis of the impact of the 2015 South India Flood to demonstrate the applicability of the enhanced algorithm (Chapter 5).
- Provide a qualitative discussion on the future development of economic cybernetics, drawing insights from this research exercise (Chapter 6).

Drawing inspiration from Leontief's proposal to establish an alternative measurement system (a proxy measurement, material measurement approach) alongside the widely adopted price measurement system (Leontief, 1991), organized regional electricity consumption data is organized based on a creative concept proposed by electrical engineers and IO researchers (Mu et al., 2010, Qu et al., 2017, Pasinetti, 2009). Furthermore, the concept of a dynamic IO model will be incorporated to enhance the algorithm developed, introducing non-linearity and extra parameters, such as capital investment, into the improved model. By approaching the IO model from a data science and system engineering perspective, this research aims to contribute a novel economic analysis tool that not only addresses the limitations of the general equilibrium assumption but also adapts to high time-resolution data for more efficient and accurate short-term economic predictions, hence potentially unlock new insights and applications in the field of economics.

## 1.4. Thesis outline

This thesis has been structured into six chapters. An outline of the thesis is given in Figure 2 to provide a clear view on how the chapters are organized to answer the research objectives.

Chapter 1 serves as an introduction to the research and provides a comprehensive overview of the research objectives. In this chapter, the motivations behind conducting the study and the potential implications of the research findings are clearly outlined. This sets the foundation for the rest of the thesis, helping readers understand the context and purpose of the study.

In Chapter 2, the IO model is utilized to examine the growing inequality of consumption-based $CO_2$ emissions in China. To achieve this, the conventional method is employed to compile a 2017 multi-regional IO table for China. This table is then connected to an emissions database for further environmental extension analysis. Throughout the process of compiling the IO table, the limitations and weaknesses of the conventional IO modelling are identified and summarized, providing insights into the areas where improvements can be made. This sets the stage for the development of an innovative approach in the subsequent chapters.

To address the limitations of the traditional input-output (IO) model, Chapter 3 introduces an innovative algorithm that enhances the SIM model. This enhancement allows for the theoretical computation of chronological interconnections between various economic sectors, based on the observed output and demand levels within an economy. In this chapter, a recursive calculation is carried out using the regression results obtained from simulated observations, which serves to validate the theoretical viability and robustness of the proposed algorithm.

In Chapter 4, an extensive dataset comprising daily electricity consumption data for hundreds of commercial products in Chongqing municipality, China, is collected and analysed. The data is then aggregated into eight distinct economic sectors, with each product categorized as either a final or intermediate product. This organized dataset is subsequently employed to calculate chronological interlinkages within the context of the enhanced SIM model. To assess the potential impact of sector-specific growth on Chongqing's overall economic performance, three distinct scenarios are simulated. Each scenario reflects varying growth rates for a selected sector, allowing for the evaluation of short-term performance implications for all other sectors within the municipality. This approach not only highlights the interconnected nature of modern economies but also offers valuable insights into potential policy decisions and strategic planning for Chongqing's future economic development.

Chapter 5 delves deeper into methodological advancements by incorporating the influence of capital investment into the enhanced SIM. In this chapter, estimated monthly output levels for various regions in India serve as the foundation for analysing the chronological impact of the devastating 2015 South India Flood on different regions. By comparing the results to a business-as-usual scenario, the chapter uncovers the chronological indirect costs associated with the disaster, providing valuable insights into the far-reaching consequences of such events. This innovative approach to assessing the impact of natural disasters on regional economies offers a more comprehensive understanding of the complex interplay between economic sectors, infrastructure investments, and the repercussions of unforeseen calamities.

Chapter 6 provides a summary and conclusion of the research conducted throughout this thesis. Drawing upon the insights and lessons learned, a qualitative analysis is combined with an extensive literature review to explore the potential for future advancements in economic modelling. Given that this research has adopted an alternative approach, one that aligns with general disequilibrium economic theories while also incorporating concepts from control system analysis, it could signify a novel direction for economic cybernetics research. By reflecting on the findings and methodologies presented in this thesis, Chapter 6 highlights the value of interdisciplinary approaches in advancing our understanding of complex economic systems. The fusion of traditional economic theory with modern control systems analysis has the potential to reshape the way we approach economic modelling, offering new perspectives and innovative solutions for addressing contemporary economic challenges.

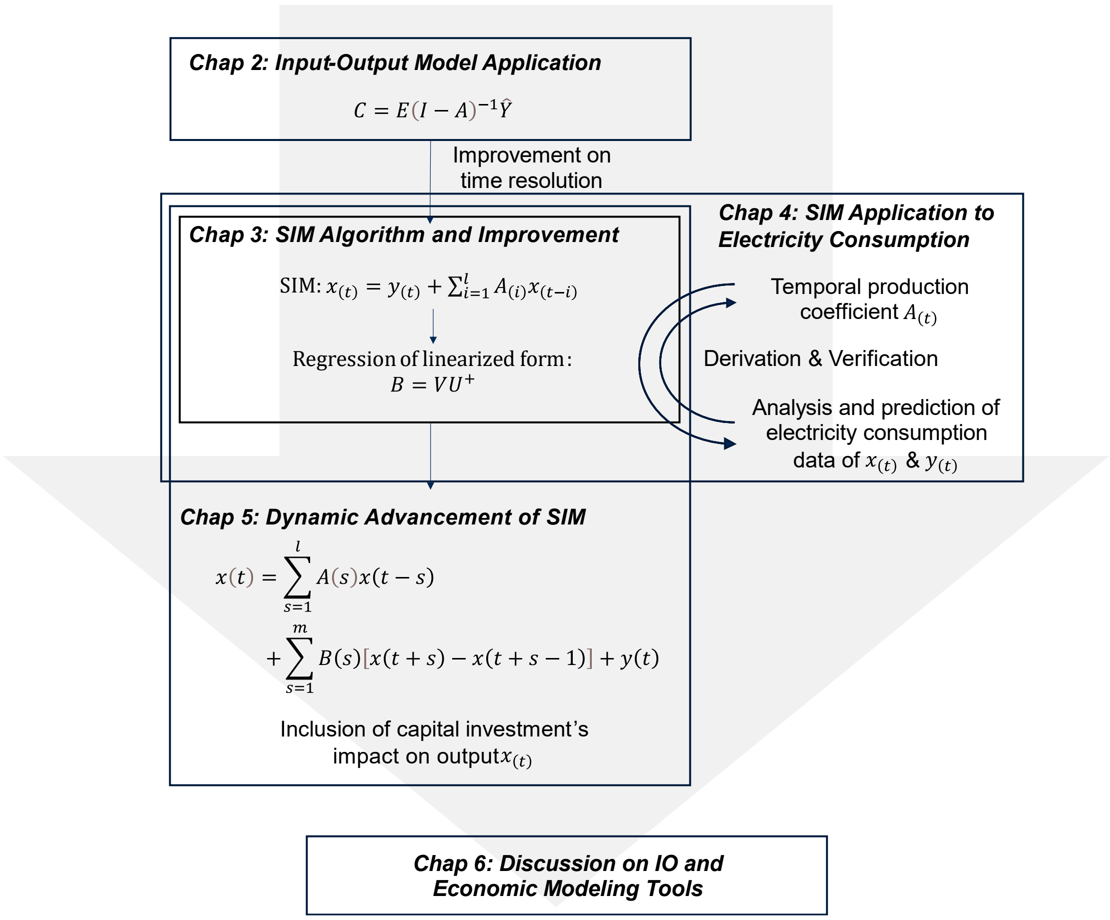



*Figure 2 Thesis structure*

# Chapter 2: Input-Output Model Application – Carbon Emission Inequality in China

In laying the groundwork for the methodological innovation central to this research, this chapter commences by revisiting the foundational principles of the IO model. To illustrate the practical applications and the analytical power of the IO model, this chapter will delve into a case study focused on the unequal distribution of China's consumption-based CO2 emissions. This case study serves a dual purpose. Firstly, it highlights the IO model's capability in dissecting and understanding complex environmental and economic issues. Secondly, it provides a real-world example of how the IO model can be employed to uncover nuanced insights into the distributional aspects of environmental impacts within a large and economically diverse country like China. This case study aims to demonstrate not only the utility of the IO model in its traditional form but also set the stage for introducing the novel methodological enhancements proposed in this research.

## 2.1. Fundamentals of Input-Output Model

### 2.1.1. The Classic Input-Output Model

The IO Model has been instrumental in analysing the interdependence between various economic sectors since its first proposal by Wassily Leontief in the 1930s (Leontief, 1936). Initially, Leontief employed the model to examine the intricate connections between diverse industrial sectors within the US economy, assessing the direct and indirect inputs needed by each sector to function efficiently. Over time, the IO Model has seen widespread adoption across numerous countries and has been implemented at various scales, ranging from global and regional levels to more localized scope. The model's enduring utility and adaptability showcase its value as a tool for understanding the complex web of relationships that underpin modern economic systems.

In Leontief's IO theory, the interdependence of economic sectors in an economy can be described by an IO table, shown below in its simple form.

| | | Industry | | | | Final Demand (y) | Total Output (x) |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | … | n | | |
| Industry | 1 | $z_{11}$ | $z_{12}$ | … | $z_{1n}$ | $y_1$ | $x_1$ |
| | 2 | $z_{21}$ | $z_{22}$ | … | $z_{2n}$ | $y_2$ | $x_2$ |
| | … | … | … | … | … | … | … |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | n | $z_{n1}$ | $z_{n2}$ | … | $z_{nn}$ | $y_n$ | $x_n$ |
| Value added (v) | | $v_1$ | $v_2$ | … | $v_n$ | | |
| Total Input (x) | | $x_1$ | $x_2$ | … | $x_n$ | | |

*Table 1 An illustration of the IO table. Inputs and outputs are recorded in monetary units.*

Table 1 represents an Input-Output (IO) table for an economy with $n$ sectors, illustrating the interactions between these sectors as they collectively strive to satisfy the final demand imposed by consumers. In the context of the IO model, the production of a product necessitates inputs from all other industries within the economy. Given that both inputs and outputs are documented in monetary units, the total output of a sector $i$ can be easily determined as the row sum, which is the sum of all intermediate outputs and final demand for that sector. This calculation is expressed as follows:

$$x_i = z_{i1} + z_{i2} + \cdots + z_{in} + y_i \tag{2.1}$$

In equation ( 2.1 ) and Table 1, $z_{in}$ is the intermediate output needed by the $n$th sector from the $i$th sector. $y_i$ is the total final demand in the $i$th sector. $x_i$ is the total output by the $i$th sector. It also equates to the column sum of the $i$th sector, suggesting that the total output in sector $i$ is equal to the input in sector $i$. Under a linear assumption, the efficiency of inputs can be expressed as a ratio of input to output in equation ( 2.2 ) below:

$$a_{ij} = \frac{z_{ij}}{x_j} \tag{2.2}$$

In equation ( 2.2 ), $a_{ij}$ is known as the technical coefficient, meaning the amount of input needed from sector $i$ to produce one unit of output in sector $j$. Hence, equation ( 2.1 ) can be rewritten as a linear system:

$$\begin{gathered} x_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + y_1 \\ \vdots \\ x_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + y_i \\ \vdots \\ x_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + y_n \end{gathered} \tag{2.3}$$

The linear system can be easily rewritten into matrix form as follow:

$$
\begin{aligned}
X &= Y + Z \\
X &= Y + AX
\end{aligned}
\qquad (2.4)
$$

Applying simple matrix algebra, equation ( 2.4 ) can be converted into ( 2.5 ) below to show a proportional relationship between final demand and total outputs:

$$X = (I - A)^{-1}Y \qquad (2.5)$$

In equation ( 2.5 ), $x$ is the total output in vector form. $y$ is the final demand in vector form. $A$ is the production coefficient matrix. $I$ is the identity matrix. $(I - A)^{-1}$, or simply $L$, is referred to as the Leontief Inverse in IO model. The Leontief Inverse describes the sum of direct and indirect input needed from all producing sectors. In other words, the IO model depicts an economic system that requires not only inputs from itself, but also a coordinated production from all sectors in the economy. On the other hand, the efficiency in sectorial production of an economy can also be revealed through the technical coefficient, thus providing a holistic account for the performance of an economy.

### 2.1.2. Multi-Regional Input-Output Model

The IO Model has undergone continuous refinement and expansion since its inception. One of the most significant developments is the formalization of the Multi-Regional Input-Output (MRIO) Model (Wiedmann, 2009). In contrast to the classic Single Regional IO Model, which concentrates on a single, closed economy, the MRIO model is employed to analyse the interdependence between various regions and countries within a specific region or across the entire world (Stadler et al., 2018b, Wang et al., 2017, Stadler et al., 2018a). The MRIO model enables economists to trace the flows of goods, services, and factors of production among distinct regions and nations. This comprehensive approach allows for the estimation of the impact of changes in production, consumption, and trade on the global economy. By capturing the intricate connections between different economies, the MRIO model offers valuable insights into the dynamics of international and interregional trade and the far-reaching consequences of economic policies and decisions.

Table 2 below depicts the MRIO table for a closed world consisting of two countries and three industries. It bears a striking resemblance to the classic Single Region IO table demonstrated in Table 1 in terms of structure, as it retains the key feature of balanced inputs and outputs, or column and row sums.

The primary distinction lies in the fact that intermediate outputs are generated not solely for domestic production, but also for production in another country. Consequently, equation ( 2.1 ) is modified as follows:

$$x_i = z_{i1}^{aa} + \cdots + z_{in}^{aa} + z_{i1}^{ab} + \cdots + z_{in}^{ab} + y_i \quad (2.6)$$

In equation ( 2.6 ), $z_{in}^{aa}$, highlighted in yellow with its similar terms in Table 2 represents the intermediate input needed in sector $i$ of country a by sector $n$ of country a, comprising the domestic intermediate inputs in a MRIO table, normally consistent with the Single Region IO table of the specific country/region. $z_{in}^{ab}$ , highlighted in green with its similar terms in Table 2, represents the intermediate input needed in sector $i$ of country a by sector $i$ of country b, comprising the foreign intermediate inputs in a MRIO table, which are collected from international and interregional trade data available.

| | | Country A Industry | | | Country B Industry | | | Final Demand (y) | Total Output (x) |
|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 1 | 2 | 3 | | |
| Country A Industry | 1 | $z_{11}^{aa}$ | $z_{12}^{aa}$ | $z_{13}^{aa}$ | $z_{11}^{ab}$ | $z_{12}^{ab}$ | $z_{13}^{ab}$ | $y_1^a$ | $x_1^a$ |
| | 2 | $z_{21}^{aa}$ | $z_{22}^{aa}$ | $z_{23}^{aa}$ | $z_{21}^{ab}$ | $z_{22}^{ab}$ | $z_{23}^{ab}$ | $y_2^a$ | $x_2^a$ |
| | 3 | $z_{31}^{aa}$ | $z_{32}^{aa}$ | $z_{33}^{aa}$ | $z_{31}^{ab}$ | $z_{32}^{ab}$ | $z_{33}^{ab}$ | $y_3^a$ | $x_3^a$ |
| Country B Industry | 1 | $z_{11}^{ba}$ | $z_{12}^{ba}$ | $z_{13}^{ba}$ | $z_{11}^{bb}$ | $z_{12}^{bb}$ | $z_{13}^{bb}$ | $y_1^b$ | $x_1^b$ |
| | 2 | $z_{21}^{ba}$ | $z_{22}^{ba}$ | $z_{23}^{ba}$ | $z_{21}^{bb}$ | $z_{22}^{bb}$ | $z_{23}^{bb}$ | $y_2^b$ | $x_2^b$ |
| | 3 | $z_{31}^{ba}$ | $z_{32}^{ba}$ | $z_{33}^{ba}$ | $z_{31}^{bb}$ | $z_{32}^{bb}$ | $z_{33}^{bb}$ | $y_3^b$ | $x_3^b$ |
| Value Added (v) | | $v_1^a$ | $v_2^a$ | $v_3^a$ | $v_1^b$ | $v_2^b$ | $v_3^b$ | | |
| Total Input (x) | | $x_1^a$ | $x_2^a$ | $x_3^a$ | $x_1^b$ | $x_2^b$ | $x_3^b$ | | |

*Table 2 An illustration of the MRIO table of two countries with three industries.*

The application of the MRIO model gained traction during the 1960s and 1970s.

In this era, economists began utilizing the MRIO framework to analyse the global economy and estimate the interdependence among countries and regions concerning trade, production, and consumption. One of the earliest and most influential studies employing MRIO analysis was conducted by American economist Walter Isard in the 1960s (Isard, 1966). Isard's ground-breaking work offered a framework for examining the interdependence between regions in terms of trade and production, ultimately laying the foundation for future research in this domain. The insights gleaned from his work have contributed significantly to our understanding of the complex relationships that exist between nations and regions within the global economic system.

### 2.1.3. Environmental Extension of Input-Output Model

In addition to its applications in economic analysis, the IO model has been adapted to incorporate emission databases, resulting in the development of the Environmentally Extended Input-Output (EEIO) model (Davis and Caldeira, 2010, Davis et al., 2011). The EEIO model considers the environmental impacts of economic activities, broadening the scope of the traditional IO model. By integrating data on environmental flows, such as greenhouse gas emissions and the use of natural resources, the EEIO model expands upon the economic flows of goods and services typically found in the conventional IO model. This environmentally conscious approach enables researchers and policymakers to gain a more comprehensive understanding of the interplay between economic activities and their ecological consequences, supporting the pursuit of sustainable development strategies.

Specifically, the EEIO requires the collection of environmental stress intensities in sectors that agrees with the specification of the IO table to be used.

$$C_{cba} = E(I - A)^{-1}\hat{Y} \qquad (2.7)$$

In equation ( 2.7 ), $E$ is the environmental intensity in vector form, with each of its elements representing the environmental intensity of the corresponding industry. $\hat{Y}$ is the diagonalized form of $Y$. The matrix product of the Leontief Inverse $(I - A)^{-1}$ and $\hat{Y}$ is the breakdown of total output $X$. Thus, the matrix multiplication in equation ( 2.7 ) gives $C_{cba}$ , the consumption-based environmental footprint as a horizontal vector.

On the contrary, if $E$ is diagonalized instead of $Y$ in equation ( 2.7 ), the equation becomes ( 2.8 ), known as the production-based environmental footprint as a vertical vector.

$$C_{pba} = \hat{E}(I - A)^{-1}Y \qquad (2.8)$$

In the context of EEIO analysis, consumption-based accounting (CBA) assigns pollution responsibilities to consumers. Unlike production-based accounting (PBA), which records emissions within territorial boundaries, CBA provides a perspective that encompasses emissions embedded in the upstream processes of products' final destinations (Peters, 2008). Initially, PBA was the widely adopted standard for pollution accounting, as it aligned with the methodology used in national scale surveys. However, this approach can lead to underestimating a country's true environmental impact, as it fails to account for emissions produced in one country but linked to the consumption of goods in another. Conversely, CBA considers the emissions associated with the production of goods and services consumed by a country, irrespective of the location of production. This method offers a more comprehensive understanding of a country's total emissions and the influence of its consumption patterns on the environment.

## 2.2. A Case Study of $CO_2$ emission inequality in China

### 2.2.1. Consumption Based $CO_2$ Emissions

In recent years, EEIO modelling and CBA for anthropogenic $CO_2$ emissions have been employed in a wide range of research (Beylot et al., 2020, Ivanova et al., 2016, He and Hertwich, 2019). These studies underscore the value of accounting for the global nature of pollution and the importance of considering consumption patterns when assessing a country's environmental impact. As the most major source of greenhouse gas, the increased $CO_2$ emission is proven to be one of the fundamental causes of earth climate change (IPCC, 2007). The catastrophic consequences of climate change include rising temperatures, sea level rise, altered weather patterns, and decreased biodiversity. Hence, knowing the sources and changing pattern of $CO_2$ emissions can help stakeholders to develop strategies for reducing them and mitigate climate change to achieve a sustainable future.

Since anthropogenic $CO_2$ emissions are driven by economic activities, countries with different levels of economic development share varied responsibility of $CO_2$ emissions, which means that consumption-based $CO_2$ emissions are distributed unequally across the globe. From the global perspective of regional economic development, $CO_2$ emissions have increasingly shifted from developed regions to developing regions, whose population generally earns lower incomes (Peters Glen et al., 2011). Recent studies confirm that $CO_2$ emissions have been relocating to developing regions with increasing speed (Lu et al., 2020, Hubacek et al., 2021, Zhang et al., 2019). Since the start of the new millennium, the $CO_2$ emissions produced by developing regions have drastically increased compared to those produced by developed regions (Fernández-Amador et al., 2016). On the other hand, some developed regions of the EU and North America have already achieved a decoupling of $CO_2$ emissions and economic growth (Hubacek et al., 2021). However, this much-lauded decoupling has often been achieved at the expense of exploiting the emissions embodied in imports from developing regions (Fan et al., 2017). Research shows that the emissions embodied in trade from developing regions to developed regions have increased drastically from 0.9 Gt $CO_2$ in 1996 to its peak of 2.1 Gt $CO_2$ in 2006, although they then quickly decreased to 1.5 Gt in 2016 (Wood et al., 2020, Mi et al., 2021). In addition, it is likely that poverty alleviation efforts will mean that those newly lifted out of poverty and near-poor individuals will increase their demands on energy consumption (Wolfram et al., 2012). Thus, poverty alleviation in developing regions can inadvertently contribute to intensified $CO_2$ emissions (Hubacek et al., 2017, Wan et al., 2022).

Being the single largest $CO_2$ emitter, China has been studied by many for its consumption-based $CO_2$ emissions. As a net exporter of $CO_2$ emissions (Wang et al., 2020b, Zhong et al., 2018), China's success at economic upgrading decreased its emissions embodied in exports from 2008 to 2015 (Mi et al., 2018). Due to its economic and geographical size, China's provinces remain varied in their levels of development. Thus, domestic trade-embodied emissions and their associated energy consumption are also considered a key research topic (Zheng et al., 2022). Recent research has identified that China's western regions are net domestic exporters of embodied $CO_2$ emissions to coastal eastern regions due to the differences in China's domestic economic structure and development (Duan et al., 2018, Zhou et al., 2018, Yang et al., 2021, Ning et al., 2019). Some lately published research account for the $CO_2$ CBA of China's provinces in 2017 (Lei et al., 2022, Dong et al., 2022). However, none of the referenced studies emphasise the intensifying domestic inequalities in consumption-based $CO_2$ emissions among China's regions as well as the possibility of their further development. Since China entered the so-called economic "new normal" in 2012 (Mi et al., 2017), the economic and $CO_2$ emission structures may have undergone alterations. Economic transformation and development in China may also imply that other countries may take up the polluting roles in the coming years. Research is thus needed to characterise and understand the rationale and scale of these changes, thus formulate further policy recommendations in accordance with the widening inequalities and overseas outsourcing trend of $CO_2$ emissions in China.

### 2.2.2. Compilation of MRIO Table

The MRIO table used in this case study was compiled using 2017 Chinese provincial IO tables published by the National Bureau of Statistics. Excluding regions and territories with no data available, 31 regions and 42 economic sectors area compiled in the 2017 Chinese MRIO table. It should be noted that the 2007 and 2012 Chinese MRIO tables have 30 regions and 30 economic sectors. The missing region in 2007 and 2012 is Tibet due to missing data. Since Tibet only composes a very small portion of consumption-based $CO_2$ emission (0.06%) and final consumptions (0.25%) in China, Tibet is excluded in the result of 2017 despite of the inconsistency with past results to maximize the information presented in this study. Besides, although the number of economic sectors for the year 2007 and 2012 is 30, inconsistent with the 42-sector specification for the year 2017, all sectors are aggregated and summed up into single regions to decrease the data resolution. Hence, in the presentation of final result, only a total number of economic output and emissions for each province is presented. Thus, the issue of inconsistent number of economic sectors is also resolved in this study.

In the construction of the 2017 Chinese MRIO table, the method utilised by Mi et al. (2017) has been taken as the reference, where the gravity model is applied to simulate interprovincial and intersectoral trade. The gravity model considers the trade between two locations to be directly proportional to the economic sizes of and inversely proportional to the distance between the two locations. Concretely, it can be expressed by equation ( 2.9 ):

$$y_i^{rs} = e^{\beta_0} \frac{\left(x_i^{rO}\right)^{\beta_1} \left(x_i^{Os}\right)^{\beta_2}}{(d^{rs})^{\beta_3}} \tag{2.9}$$

In equation ( 2.9 ), $y_i^{rs}$ represents the economic quantity of item $i$ traded from location $r$ to location $s$. $x_i^{rO}$ is the quantity of item $i$ exported by location $r$. $x_i^{Os}$ is the quantity of item $i$ imported by location $s$. $d^{rs}$ is the distance between location $r$ and $s$. In this study, the distances of provincial capitals $d^{rs}$ are used. $\beta_1$, $\beta_2$, and $\beta_3$ are the model coefficients to be obtained through regression. $e^{\beta_0}$ is the error term. To reconcile for linear regression, equation ( 2.9 ) is manipulated into equation ( 2.10 ) as shown below

$$\ln(y_i^{rs}) = \beta_0 + \beta_1 \ln\left(x_i^{rO}\right) + \beta_2 \ln\left(x_i^{Os}\right) - \beta_3 \ln(d^{rs}) + \varepsilon \tag{2.10}$$

Having the regressed coefficients, it is thus possible to model the economic flow between any two provincial sectors.

In addition to standard gravity model, impact coefficients are also introduced to model cooperation and competition relationships among provincial sectors, which is given as $c_i^{gh}$ below in equation ( 2.11 )

$$\begin{cases} c_i^{gh} = \dfrac{\mu_i^g + \mu_i^h}{\left|\mu_i^g - \mu_i^h\right| + min_{r=1,2,\ldots mn}\mu_i^r} & g \neq h \\ c_i^{gh} = 1 & g = h \end{cases} \tag{2.11}$$

In equation ( 2.11 ), $c_i^{gh}$ is the impact coefficient for item $i$ between location $g$ and $h$ for $n$ locations. It measures the strength of interaction of item $i$. $\mu_i^g$ and $\mu_i^h$ are the location quotients of item $i$ in location $g$ and $h$.

Then, trade flow obtained from gravity model is further modified into equation ( 2.12 ) to reflect cooperative and competitive relationships using impact exponents $\bar{\delta} - \delta_i$:

$$y_i^{rs\prime} = \frac{y_i^{rs}}{(c_i^{gh})^{\bar{\delta}-\delta_i}} \qquad (2.12)$$

$\delta_i$ is the proportion of total output of item $i$ that it uses as its own intermediate inputs, while $\bar{\delta}$ is its average value. Hence, the denominator of equation ( 2.12 ) will adjust the trade flow modelled from standard gravity model to reflect cooperation and competition. Final MRIO table is performed with RAS algorithm to ensure its consistency in column and row sums (Jackson and Murray, 2004).

For the $CO_2$ emission inventories, the CEADs database (Shan et al., 2020) is adopted. The CEADs database (https://www.ceads.net/) is a widely utilised database for $CO_2$ emissions in China. It follows IPCC Guidelines for National Greenhouse Gas Inventories when compiling its emission inventories (Intergovernmental Panel on Climate Change, 2019). It gives a breakdown of $CO_2$ emissions across Chinese provinces and sectors. The CEADs emission inventory is mapped in accordance with the MRIO table using a method applied by previous studies (Mi et al., 2020, Yan and Yang, 2021). The MRIO table is compiled from the regional Input-Output tables published by the Bureau of Statistics of the respective provinces. Global flow of embodied $CO_2$ emissions is calculated using data of EXIOBASE database (Stadler et al., 2018b).

### 2.2.3. Emission Gini Coefficient

Economists often use the Gini coefficient to quantitatively compare income inequalities (Dorfman, 1979). Recently, some researchers have altered the methodology for calculating Gini coefficients to investigate the $CO_2$ emission inequalities across different income groups (Sun et al., 2021, Wiedenhofer et al., 2017). Here, in this research, the variables in the Gini coefficient calculations are changed to directly show the difference in $CO_2$ emissions among Chinese provinces instead of population groups. Originally, Gini coefficient is derived from the Lorenz Curve. The larger the Gini coefficient is, the more unequally the income is distributed among the population. In a Lorenz Curve, the horizontal axis is the fraction of population, while the vertical axis is the cumulative share of income. A line of equality indicates perfectly equal distribution of income among all the population. Denoting the area between the Lorenz Curve and the line of equality as $A$ and the area between the Lorenz

Curve and the axes as $B$, the Gini coefficient is simply given by $A/(A+B)$. The emission Gini coefficient in this study changes the horizontal axis to the proportion of final consumption in China's provinces and the vertical axis to the cumulative consumption-based $CO_2$ emissions, as shown in Figure 4. Hence, the alternative version of the Gini coefficient can be calculated using equation ( 2.13 ) below:

$$G = \frac{A}{A+B} \qquad (2.13)$$

In equation ( 2.13 ), $A$ is the area between the emission Lorenz Curve and the line of equality. $B$ is the area between the emission Lorenz Curve and the axes. By changing the concept of the Gini coefficient into the format presented in equation ( 2.13 ), it is intended to reveal the inequality in emissions embodied in consumption activities across Chinese provinces.

### 2.2.4. Research Results

#### *2.2.4.1. The Intensifying Inequality of $CO_2$ Emissions*

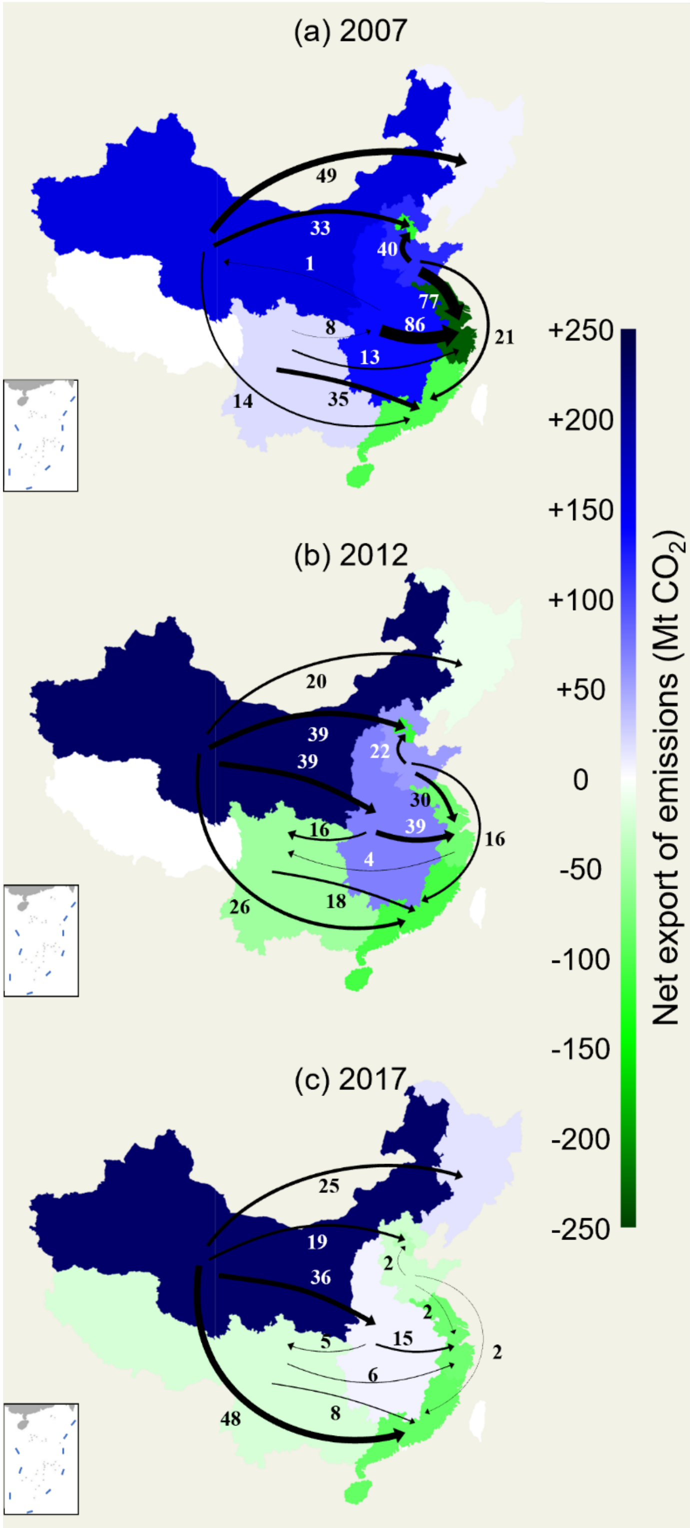



*Figure 3 The net flow of $CO_2$ emissions embodied in domestic trade among regions of China in (a) 2007, (b) 2012, and (c) 2017.*

The latest result of 2017 suggests that $CO_2$ emissions in China continues to be shifted towards the less developed northwest region, create a widening inequality of consumption-based $CO_2$ emissions. In Figure 3, China's provinces are organized into the identical 8 regions to show the flow of $CO_2$ emissions. The 8 regions are divided in accordance with the widely practiced China administrative region specification. Note that not all transregional flows are presented in this figure due to artistic constraints. In 2007, 5 regions of China were net exporters of embodied $CO_2$ emissions (coloured in blue in Figure 3). It decreased to 3 regions in 2017. Specifically, the southwest region first changed from net exporter to net importer in 2012 and remained a net importer in 2017, although its amount of $CO_2$ emissions net imported has decreased from 54 Mt to 22 Mt. Among all trade partners of southwest region, north and central coastal regions have changed from net exporters to net importers from 2012 to 2017, suggesting the strengthened industrial linkages among the newly developed regions of China. The same reverse for the north region happened later in 2017, with its net export of consumption-based $CO_2$ emissions decreased from 60 Mt to -26 Mt. The largest decrease in consumption-based $CO_2$ export from north region happened with central coastal region from 2012 to 2017 (30 Mt to 2 Mt). Although a reverse was yet to be observed in the central region, continuous and drastic decrease of $CO_2$ emissions net export can be easily seen for the decade of 2007 to 2017. From 2012 to 2017, the net export of consumption-based $CO_2$ emissions has greatly decreased from 71 Mt to 10 Mt. The central coastal region has the largest decrease (24 Mt) in net import of $CO_2$ emissions with central region from 2012 to 2017. $CO_2$ emissions export was also observed to be on the trend of polarizing towards the northwest region. Specifically, after the sharp increase in 2012, northwest region's net export of consumption-based $CO_2$ emissions remained constant at 230 Mt, significantly outnumbers the next net exporters, the northeast region (18 Mt) and the central region (10 Mt). It makes northwest region the only significant $CO_2$ exporter among all eight regions of China, a very different situation than 2007 and 2012 where north and central regions also played significant roles in producing $CO_2$ emissions for other regions of China. The northeast region showed an exceptional swaying trend, meaning it experienced two reverses in $CO_2$ emissions net exportation in 2012 and 2017 respectively. Another interesting observation is that although Beijing-Tianjin, central coastal and south coastal regions remain as the $CO_2$ emission net importers, the quantities of consumption-based $CO_2$ imported decreased by 78 Mt, 154 Mt, and 11 Mt respectively in 2017 compared to 2007, suggesting their reliance on the domestic supply chain for pollution outsourcing has decreased.

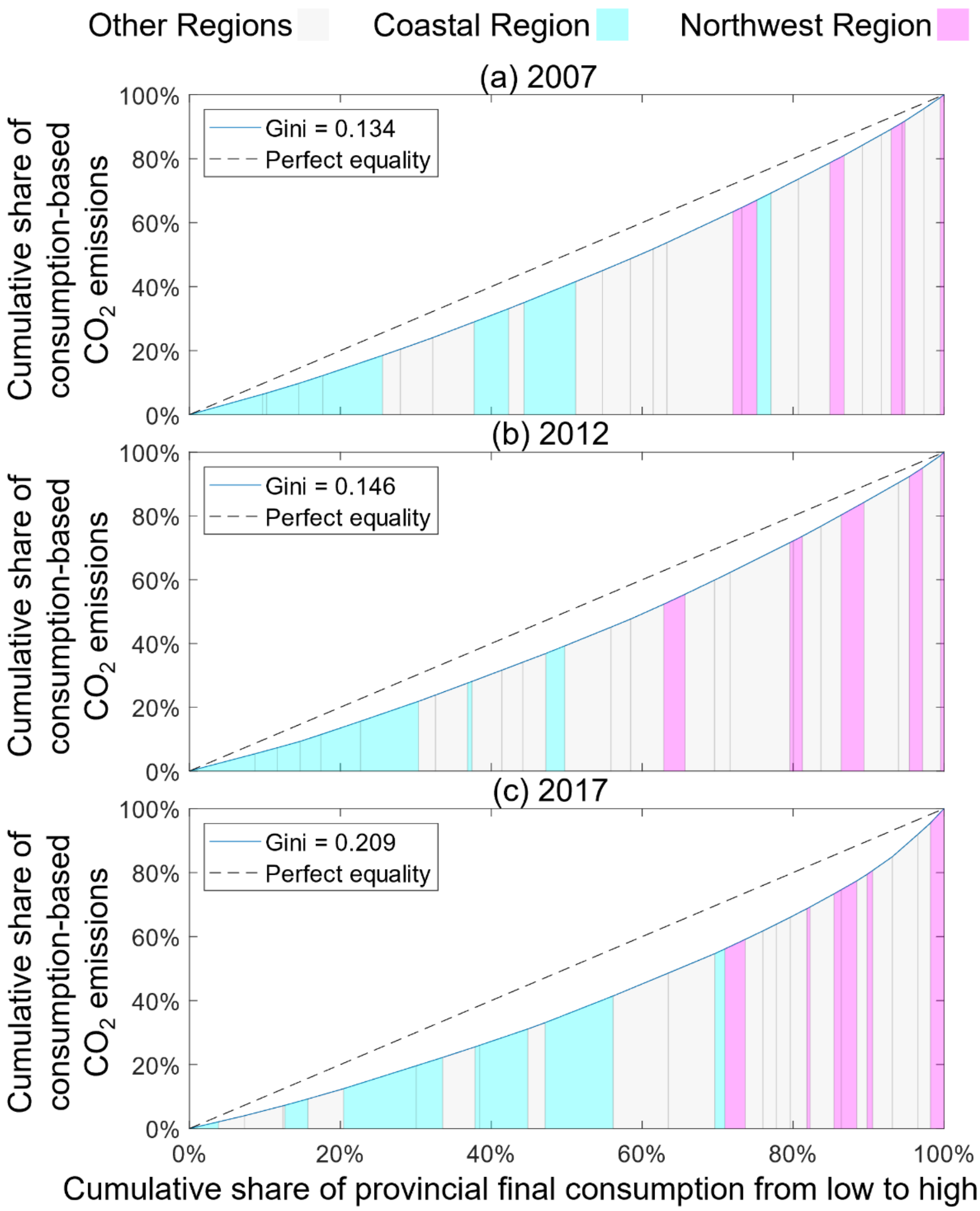



*Figure 4 Lorenz curves of Chinese provincial consumption-based $CO_2$ emissions in 2007, 2012, and 2017.*

Since the $CO_2$ emissions of China have been polarised towards the northwest region, this phenomenon suggests that the inequality in the geographical distribution of emissions is intensifying. a modified emission Gini coefficient is thus introduced to quantitatively compare the emission inequalities among China's provinces for the years 2007, 2012, and 2017. Instead of showing the distribution of income among populations, the modified emission Gini coefficient shows the distribution of consumption-based $CO_2$ emissions among final consumption across provinces. Figure 4 shows the Lorenz curves of consumption-based $CO_2$ emissions against the proportion of final consumption in China's provinces from 2007 to 2017. The Lorenz curves have been altered so that the horizontal axis is the cumulative share of provincial final consumption, and the vertical axis is the cumulative share of consumption-based CO2 emissions. i.e. each bar under the Lorenz curve has its width (horizontal axis) representing the amount of the province's final consumption, while the height (vertical axis) representing the amount of emissions produced cumulatively added with the emissions produced by provinces positioned on its left. Provinces are positioned from left to right in ascending order of emissions produced. It is clearly shown that the emission Gini coefficient among China's provinces has drastically increased from 0.134 in 2007 to 0.209 in 2017, or +56.0%. In comparison, the change in income Gini coefficient of China was only -3.5% from 2007 to 2017. The reason for this change can be attributed to the different strength in $CO_2$ emission decoupling between the developed and developing regions in China, which is discussed in more detail later in the Discussion section. In addition, although an increase in emission inequality occurred from 2007 to 2012 (+0.012), it was not tantamount to the intensification of emission inequality from 2012 to 2017 (+0.063). It shows that the emission inequality is much widened between 2012 to 2017, implying that the more developed regions have been gradually reaching emission decoupling. It coincides with the emphasis of green development by the Central Government of China in more recent years, suggesting that the developed regions are more capable in responding to the policy shift as they possess more resources to do so. The Lorenz curve for 2017 is distorted towards the right, suggesting that $CO_2$ emissions are more concentrated towards the northwest region.

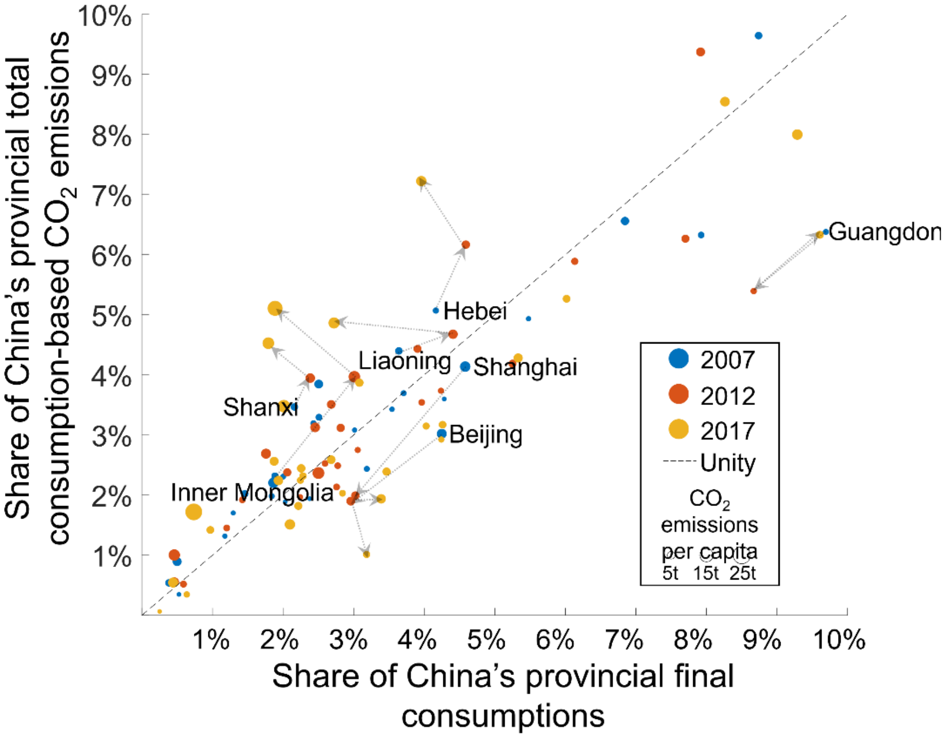



*Figure 5 The shares of consumption-based $CO_2$ emissions versus the shares of final consumptions of Chinese provinces in 2007, 2012, and 2017.*

The inequal distribution of emissions in China can also be shown by the disparities between final consumptions and consumption-based $CO_2$ emissions among provinces. While some developed provinces enjoy high level of consumptions, the $CO_2$ emissions associated are disproportionately less. Figure 5 is produced for the convenient comparison of the two quantities, where the sizes of dots are the consumption-based CO2 emission per capita. In Figure 5, provinces on the left of the unity line induced relatively more $CO_2$ emissions than they consume and vice versa for the provinces on the right of the unity line. Observation shows that Inner Mongolia, one of the northwest provinces, is a typical province with a disproportionally higher $CO_2$ emission than its consumption. The differences between the percentages of consumption-based $CO_2$ emissions and final consumptions of Inner Mongolia have increased by 2.8 percentage point from 2007 to 2017, the largest of all China's provinces, followed by Hebei (2.4), Shanxi (1.4), and Liaoning (1.4). The developed provinces of Guangdong, Beijing, and Shanghai etc., on the other hand, have higher proportion of final consumptions than consumption-based $CO_2$ emissions. The differences between the percentage of final consumptions and consumption-based $CO_2$ emissions of Guangdong have remained at 3.3 percentage point from 2007 to 2017, but gradual increases in differences can be identified in other developed provinces like Beijing (0.9) and Shanghai (1.0). Moreover, increases in disparities between the percentage of final consumptions and consumption-based $CO_2$ emissions can be seen in many more provinces. Scattered dots of year 2007 are located closer to the unity line compared to year 2017 in Figure 6, meaning that the disparities are less severe a decade ago. Another indicator suggesting a widening inequality is the number of provinces with higher proportion of consumption-based $CO_2$ emissions than final consumptions. In 2007, 17 out of 30 provinces studied have higher proportion of consumption-based $CO_2$ emissions than final consumptions (i.e. located on the left of the unity line in Figure 6). In 2017, this number has decreased to 13 out of 31 provinces studied, among which 4 of them (Hebei, Shanxi, Inner Mongolia, and Liaoning) have a disparity higher than 2 percentage point. Provinces with higher proportion of consumption-based $CO_2$ emissions than final consumptions generally have higher emission per capita. The extent of differences in emission per capita also intensified in 2017 compared to 2007. It indicates again a polarizing trend of consumption-based emissions towards the less developed provinces of China.

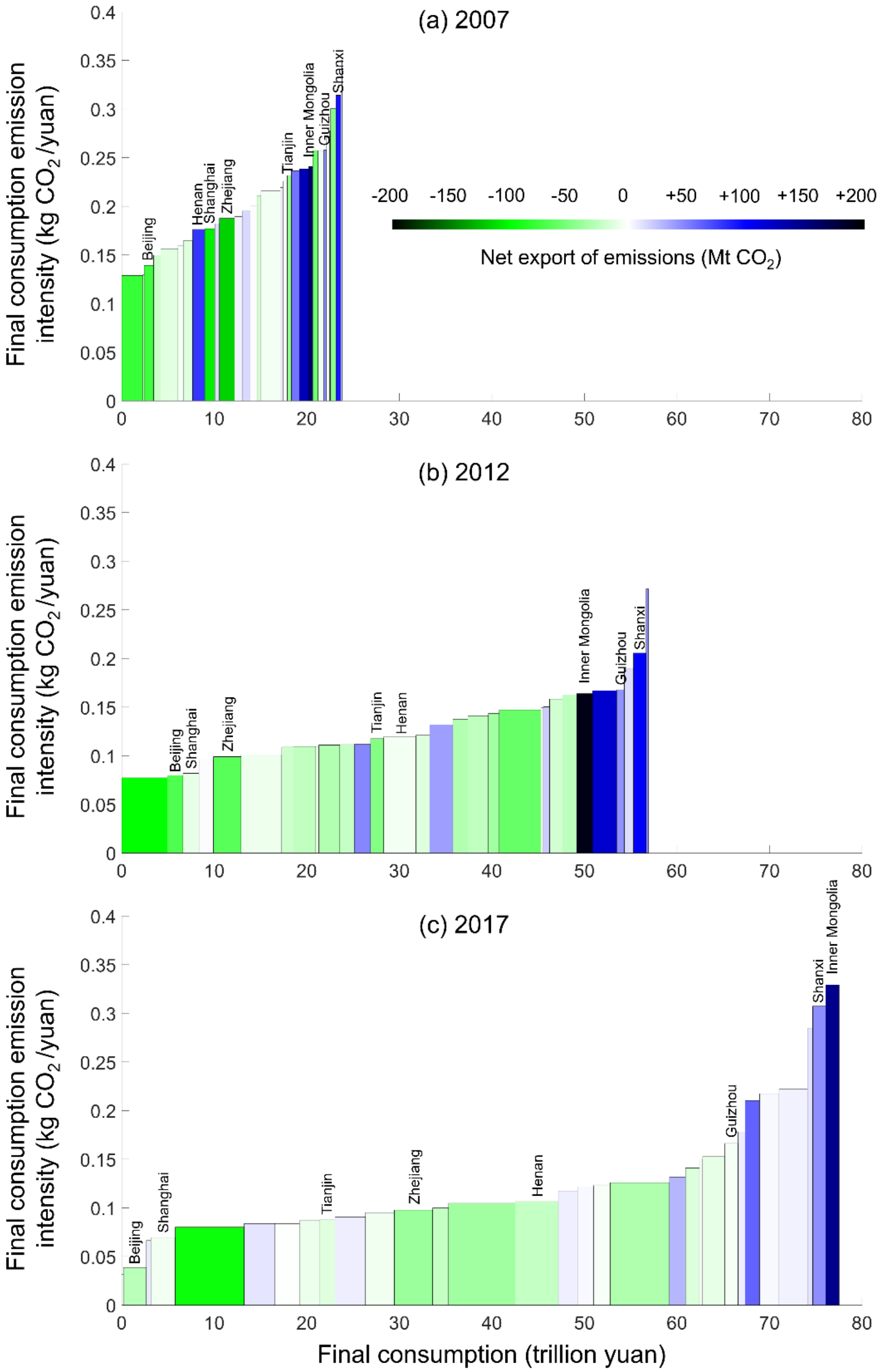



*Figure 6 The final consumptions, consumption-based $CO_2$ emission intensities, and net export of consumption-based $CO_2$ emissions of China's provinces in (a) 2007, (b) 2012, and (c) 2017.*

The discrepancy in consumptions and consumption-based $CO_2$ emissions can be analysed from the differences in emission intensity and emission embodied in domestic trades as shown in Figure 6. Each rectangular bar represents the size of consumption-based emissions of the labelled province. Heights and widths of the bars show final consumption emission intensities and final consumptions of the labelled provinces respectively. Face colour of the bars indicates net trades of consumption-based $CO_2$ emissions. Darker green means larger net imports of consumption-based $CO_2$ emissions. Darker blue means larger net exports of consumption-based $CO_2$ emissions. In general, provinces with higher emission intensities export more consumption-based $CO_2$ emissions to other provinces but consume less consumption-based $CO_2$ emissions in 2017. The opposite applies for provinces with lower emission intensities. In 2017, all 5 provinces with the highest $CO_2$ emission intensities are net exporters of consumption-based $CO_2$ emissions. However, having high emission intensities before 2017 was not equivalent to being net exporter of consumption-based $CO_2$ emissions. Among the 5 provinces with the highest $CO_2$ emission intensities in 2007, 3 of them are net importers. Inner Mongolia is the only province with an increased $CO_2$ emission intensity, while also remains as the largest net exporter of consumption-based $CO_2$ emission from 2007 to 2017. For developed provinces like Beijing, Shanghai, Tianjin, and Zhejiang, the net imports of consumption-based $CO_2$ emissions are continuously decreasing, with the decreases counting 48 Mt, 100 Mt, 30 Mt, and 78 Mt respectively from 2007 to 2017. On the contrary, net export of consumption-based $CO_2$ from developing provinces like Hebei, Henan, Shanxi, and Guizhou also drastically decreased from 2007 to 2017, amounting 123 Mt, 103 Mt, 48 Mt, and 46 Mt respectively. Nevertheless, as the largest net importer and exporter of consumption-based $CO_2$ emission in 2017, Guangdong and Inner Mongolia (96 Mt imported and 146 Mt exported consumption-based $CO_2$ respectively) have not undergone much change in their traded consumption-based $CO_2$ emissions from 2007 to 2017 (differences count 17 Mt and 13 Mt respectively). Thus, although the variance of trade embodied $CO_2$ emissions has decreased from 3688 $Mt^2$ in 2007 to 1453 $Mt^2$ in 2017, outlying data points of Inner Mongolia and Guangdong again indicate a polarizing trend of consumption-based $CO_2$ emissions' distribution among provinces in China. This polarizing trend can also be visually deduced from the increased concavities of the plots in Figure 6. The tops of the rectangles are straighter in 2007 than in 2017, suggesting that the differences between emission intensities of various provinces increases faster in 2017 than in 2007. It is another indicator to show the polarising trend of inequality in $CO_2$ emissions. Detailed results are presented in Appendix Table S1.

#### *2.2.4.2. Uncertainty Analysis*

Same to any research, this study is prone to limitations and weakness. Although economic factor is fundamentally determining to consumption level and hence consumption-based emission, there are other factors that play important roles in determining the consumption based $CO_2$ emissions across geographical locations. For instance, provinces in the north requires more heating in colder seasons, which may contribute to higher consumption-based $CO_2$ emissions as economic grows and residents' income increases. As the degree of income elasticity varies for different factors, the extent on the non-economic factors' impact on consumption-based emissions and their inequalities may be investigated in further studies.

In addition, due to differences in specifications and data source used for MRIO table compilation, uncertainties may also arise between MRIO tables used and hence the calculated consumption-based emissions. The same calculation is performed for consumption-based $CO_2$ emissions using another recently published 2017 China MRIO table compiled by CEADs (Zheng et al., 2021). For the consumption-based $CO_2$ emissions of the 31 provinces, the average difference is 22%. If the two outlying results are removed, the average in differences is further reduced to 14%. It suggests the results in this study are generally accurate and reliable, but further investigations may be needed to discuss the specific provinces that have larger discrepancies in the $CO_2$ emissions calculated.

### 2.2.5. Case Result Analysis

In this study, the changing distribution of consumption-based $CO_2$ emissions among Chinese provinces from 2007 to 2017 is revealed. The general trend of consumption-based $CO_2$ emissions flow is from inner lands to coastal regions, same as the most recent study of Dong et al. (2022) has found. Being more unique and focused, this result revealed that the unequal geographical distribution of $CO_2$ emissions intensified. Emission responsibilities shifted towards a few provinces, which were mostly net exporters of embodied emissions to other provinces. In other words, less developed provinces are becoming the so-called "pollution haven" for the more developed provinces. In past studies, the hypothesis of a "pollution haven" has been proven with evidence at the international scale (Hoekstra et al., 2016, Malik and Lan, 2016). This study quantitatively tells the interesting story that the domestic transfer of embodied pollution in China not only exists but has also intensified in line with global trends.

On the other hand, the consumption-based $CO_2$ among provinces in China also shows a polarizing trend. More and more consumption-based $CO_2$ emissions are now induced by the less developed regions in China, shown by the intensifying emission Gini coefficient. In the early 2000s, the coastal regions of China economically benefited from globalisation and China's opening up, constituting the first batch of developed Chinese regions. One possible explanation for the observation is the Flying Geese Paradigm proposed by Akamatsu (Kasahara, 2004). After acting as the outsourcing hub for other developed economies, the lower-end and more polluting industries of the coastal regions were phased out and moved to the less developed inland regions once industry upgrading was completed (Xu and Ang, 2013, Kanitkar et al., 2015, Su and Ang, 2016). This can also be explained by the “carbon leakage” phenomenon widely emphasised by the policy and science communities (Misch and Wingender, 2021). As a region becomes more developed, the cost of pollution increases due to tightened local regulations. Businesses will seek alternative locations with laxer pollution restrictions to lower the cost of production, causing the shifting of pollution sources to the less developed regions, as revealed in this study. In addition to policy drivers, the polarization trend between the developed and developing regions may be a result structural change, too. Due to technological advancement and change in energy sources, the developed regions achieve cleaner production faster than their developing counterparts, thus intensifying the polarizing trend of consumption-based $CO_2$ emissions.

##### *2.2.5.1. Lagging in Carbon Decoupling*

The transfer of pollution to less developed regions can also be linked to the U-shaped relationship between pollution and economic development. Initially, pollution continues to increase with progress in economic development. Once a region is relatively developed, pollution will start to decline after a tipping point due to increased emphasis on environmental welfare. Such a relationship has been proven in global studies (Hailemariam et al., 2020). The existence of a turning point in China’s $CO_2$ emissions is also proven with Chinese historical data (Huang and Zhao, 2018). The decoupling of carbon emissions and economic growth in the more developed coastal regions has also been verified by Zhou et al. (Zhou et al., 2017b).

For further analysis, $CO_2$ intensities against consumption per capita are plotted in Figure 7. An approximated Environmental Kuznets Curve (EKC) is added as an illustration of the hypothetical relationship between environmental degradation and economic development. Doing so illustrates an inverted U-shaped relationship, coinciding with Environmental Kuznets Curve (EKC) theory. Although most EKC research adopts production-based accounting, some literatures also verify that consumption-based emissions may also follow the inverted U-shaped relationship with economic development (Aldy, 2005, Gawande et al., 2001). It is shown in Figure 7 that developed regions such as Beijing and Shanghai already exhibit a strong decoupling between final consumption and per unit emissions, but such a decoupling trend has yet to be discovered among less developed regions such as Inner Mongolia and Ningxia.

Observation shows that the less developed regions of China still have great potential to improve their emission efficiencies. Yet, the EKC theory serves only as a potential explanation for the decoupling of $CO_2$ emissions between different regions. To determine if the EKC exists in this case, quantitative and systematic investigation is needed.

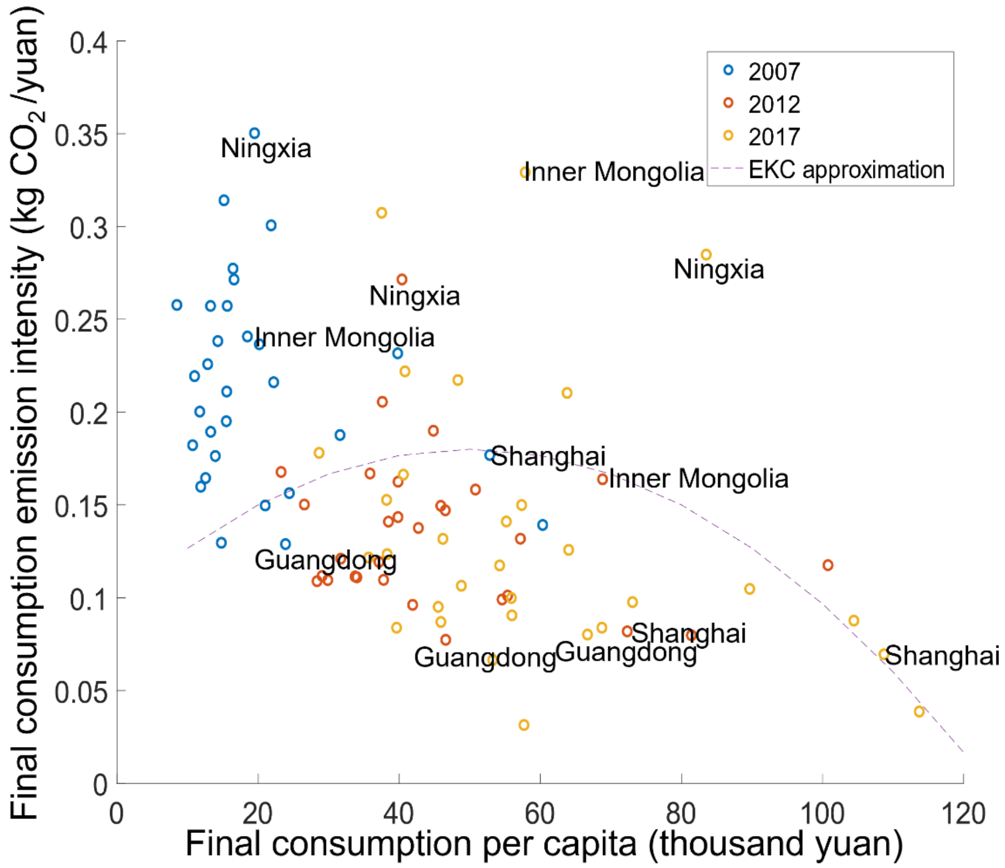



*Figure 7. Final consumption emission intensities against final consumption per capita for China's provinces.*

An effective way to achieve $CO_2$ emission mitigation is to target less developed and more emission-intensive regions, a strategy that has proven to be one of the most effective for $CO_2$ emission mitigation (Wood et al., 2020, Rao and Min, 2018). In fact, the Central Government of China has already realised the challenges in $CO_2$ emission efficiencies faced by the less developed region of the northwest (NDRC, 2021). In the upcoming 14th Five-Year Plan, policies targeted at the northwest regions of China have been devised to alleviate emission intensities to achieve China's goal of peaking $CO_2$ emissions by 2030 (Government of Inner Mongolia, 2021). In addition, this study also quantitatively shows that the increasing of $CO_2$ emissions embodied in the export from the northwest to other regions is a key contributor to the increasing polarization of $CO_2$ emissions across China. Besides focused policies on the less developed regions only, the Central Government may also consider formalized mechanism to promote coordinated cross regional policies among both developing and developed regions. For instance, Clean Development Mechanism (CDM) has always been advocated in the international setting, but less attention has been diverted for CDM with countries' borders. The Central Government may consider the implementation of similar mechanism to ensure more just allocation of emissions responsibilities among domestic players and less mitigation resource burden on the Central Government. In addition, financial tool may be an alternative for alleviating the inequality in $CO_2$ emissions. Regulatory easing and subsidies for green bond issuance from less developed regions to the developed regions may also be a viable option.

#### *2.2.5.2. Trends with the World*

However, $CO_2$ emission mitigation is not only a domestic problem but also a global challenge that requires international coordination. Pollution transfer happens across borders between China and the world as well. Evidence supports the observation that the $CO_2$ emissions embodied in China's net exports to developed countries are already decreasing (Hu and Wu, 2021, Mi et al., 2017), shifting to the developing world (Wu et al., 2021, Meng et al., 2018). Such an observation serves as empirical evidence for the argument that a further shifting of the $CO_2$ emissions embodied in exports from the less developed regions of China to the world in the near future is impending. Given the uncertainties of global geopolitical situations, further supply chain shifts from China to the rest of the world will be a very likely and imminent event (Cao, 2022).

With the world MRIO table and emission inventory of the EXIOBASE database (Stadler et al., 2018b) and the calculation by He and Hertwich(He and Hertwich, 2019), Figure 8 is produced to show how $CO_2$ emissions embodied in trade shifted from 2007 to 2017 across the world (See Appendix Table S2). To ensure the discrepancies between different data sources are minimized, the global emissions calculated from EXIOBASE are normalized with the domestic $CO_2$ emissions of China calculated in this study. In general, the $CO_2$ emissions embodied in China's net exports to the world decreased for all regions, which is in line with the findings of other studies. However, North America's (a typical developed region) $CO_2$ emissions embodied in its net imports show an increase from the world other than China from 2012 to 2017, while the net exports of embodied $CO_2$ emissions from emerging economies also show an increase from 2012 to 2017.

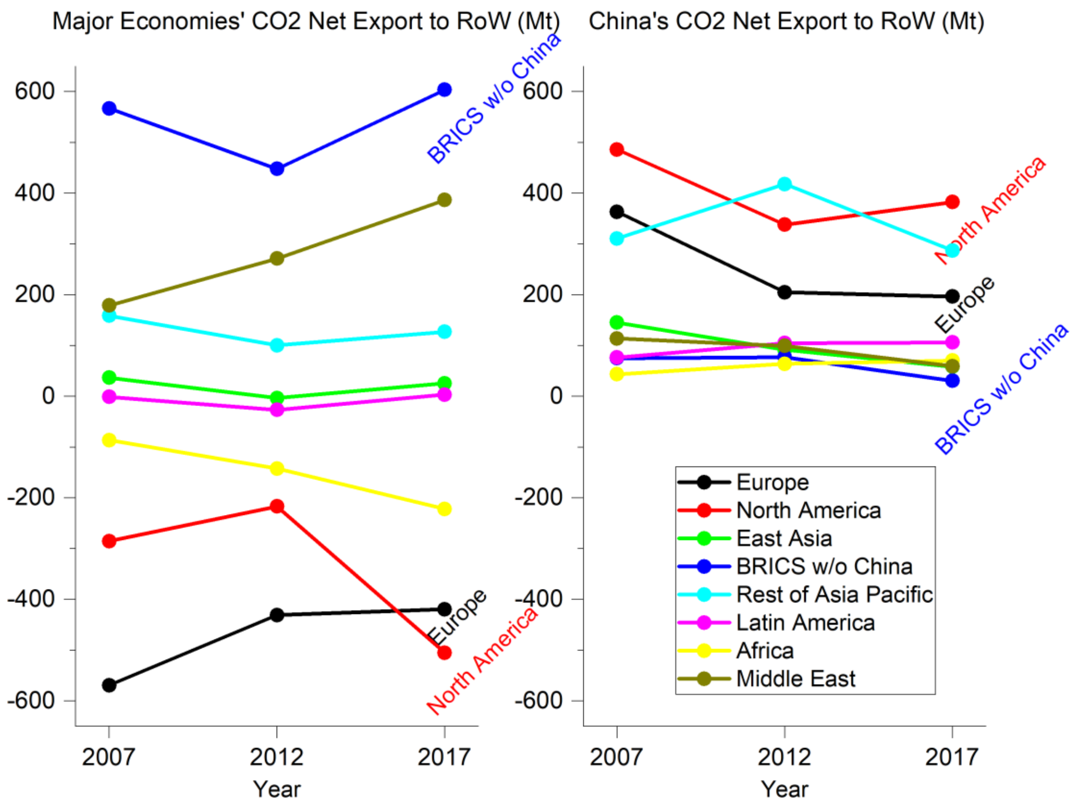



*Figure 8. Net export of $CO_2$ emissions from major world economic regions and China to the rest of the world other than China in 2007, 2012, and 2017*

Fortunately, the rapid development and deployment of clean technologies that were not available a decade ago may help developing countries achieve a clean and green reception of supply chains from China. In recent years, the cost of renewable energies has drastically decreased (IRENA, 2020). Knowledge of the best practices for sustainable investment in energy infrastructures is becoming more available and is regarded as a larger priority by global policy makers (Grubb et al., 2021). Recent evidence also shows that the exchange of clean energy technologies among countries may be able to put a stop to pollution outsourcing (Gosens, 2020). In other words, having a variety of green technology choices provides us with an alternative future to the past of perpetual emission outsourcing. It could be the key for us to achieve an equally sustainable future for all countries, regardless of the relative levels of economic development and the overall levels of economic well-being.

### 2.3. Limitations of Input-Output Method

The case study demonstrates a typical research conducted using the EEIO model. It shows that the IO model is a sufficient and essential tool to answer questions not only limited to economic structural analysis but also environmental impacts. In many other studies, researchers have also further exploited the potentials of the IO model in more research fields. For instance, the IO model can be used to facilitate the optimization of resource allocation among different sectors of a system (Liu et al., 2017, Wang et al., 2009). Some research uses the IO model to determine the optimal production levels for different products given the available resources and market demand (Lin and Polenske, 1998). Based on the IO model, some research attempts to optimize the performance of complex systems, such as the power grid (Qu et al., 2017), by balancing the trade-offs between different objectives.

Nevertheless, the limitations of the classic IO model have also been raised and long-debated by researchers. In this case study of China's carbon emission inequality, following limitations of the IO model has been demonstrated.

Firstly, the assumptions for the IO model are overly simplified (Munroe and Biles, 2005). The IO model assumes a simple linear relationship between input and output as shown in equation ( 2.4 ). This practice may be valid when accounting for past economic activities as it is only a way of attributing economic input and output contributions to different sectors. However, when it comes to more complex economic reasoning, the original IO model shows its incapability in modelling the nonlinear factors of economies, such as economic of scale, bottleneck in production, and elasticities.

Secondly, the data collection for the IO model is resource intensive (Miller and Blair, 2009). It normally needs an economic survey at the national level, which is expensive in terms of labour and time resources needed. In this case, the 2017 provincial IO tables used were only available after 2022. Conventionally, most countries publish their IO tables normally by a lag of 3 to 5 years. It thus greatly deteriorates the timeliness and accuracy of IO model in analysing economic performance.

Thirdly, the forecast ability of the IO model is limited (Israilevich et al., 1997). This study already shows that the economic structure and hence $CO_2$ emission pattern of China has undergone tremendous changes for the decade investigated. However, due to limited availability of IO tables as data sources, the resolution of the time dimension must be limited in annual terms. In other words, it is assumed in IO analysis that the technical coefficients remain unchanged within each of the investigating years. However, economic structure is constantly under change. Trade patterns from one region to another in this quarter may be very different to the last quarter as shown in abundantly available macro-economic data. Thus, the ability of IO model to provide more dynamic analysis and forecasts in a shorter time span is limited.

Fourthly, the uncertainties in the process of national survey are unavoidable (Lenzen et al., 2010). The compilation of 2017 China MRIO table in this case study is based on gravity model, which is sometimes criticized for its strong assumption on the interaction of regions and sectors by IO modellers who are more convinced with the statistic nature of IO tables. In similar research involving the disaggregation of IO tables, no matter in the time domain or in sector resolution, various strong assumptions and simplifications must be made, contributing to the uncertainties of the IO research result.

To surmount the limitations of the IO model as elucidated in this case study, methodological innovations are not just desirable but essential. Marked by the rapid evolution of computational algorithms and the burgeoning field of data science, viable alternatives are emerging to address the challenges that traditional statistical models encounter nowadays. By innovatively modifying the IO model and synergistically integrating it with big data, IO model's capacity can be significantly enhanced to map out more intricate economic interdependencies and thus offer more accurate predictions of short-term economic outcomes, thereby providing invaluable insights into the dynamic economic landscape. Moreover, the process of collecting data for IO tables stands to benefit immensely from technological advancements in data science. The adoption of automated systems for statistical compilation, leveraging high time-resolution economic data or proxy data, promises to drastically reduce inaccuracies. Consequently, the limitations highlighted in this case study can be substantially mitigated, if not entirely overcome, through these methodological advancements and technological integrations. This paves the way for a more robust, reliable, and dynamic IO model, tailored to meet the analytical demands of modern economics.

# Chapter 3: The Sequential Interindustry Model and Its Improvement

This chapter commences by revisiting several pivotal methodological innovations within the realm of the IO model. As an innovation in the time series, this chapter adopts the Sequential Interindustry Model (SIM). This chapter then details the development of an algorithm specifically designed to facilitate the effective resolution of the SIM model based on observations of economic outputs. The subsequent sections of this chapter will delve deeper into the technical intricacies of this algorithm and its practical applications in the analysis of economic data.

## 3.1. Past Development of Input-Output Model

Since the first proposal of the classic IO Model, economists have made some useful advancement in the methodology of IO modelling to tailor for the differentiated needs for economic analysis. As demonstrated in Chapter 2, the EEIO model is developed based on IO to assess the environmental stresses of the economy. Besides environmental accounting, energy (Liang et al., 2010), resource (Nakamura and Nakajima, 2005), and even health (Zhang et al., 2017b) stresses are investigated using EEIO model. Chapter 2 also shows how the MRIO model can be used to study trade activities and supply chains among regions. With global MRIO tables available, economists can easily conduct analysis related to global value chains considering both direct and indirect consumptions (Albino et al., 2002, Wang et al., 2020a). In addition to the two variants applied in Chapter 2, the IO price model factors in price formation mechanism to discuss fiscal and monetary policies concerning inflations (Przybyliński and Gorzałczyński, 2022), exchange rates (Duan et al., 2020), subsidies (Harun et al., 2018) and more. The following section describes some typical examples of most widely adopted variants of the IO model to offer a glimpse on the important methodological improvements made for the IO model.

### 3.1.1. Structural Decomposition Analysis

Structural Decomposition Analysis (SDA) is a powerful analytical tool used in economics to study the contribution of various factors to changes in economic activity. Based on the IO model, it is used to decompose the changes in an economy into underlying factors such as sectoral changes, technological improvements, and changes in demand. By differentiating the classic IO Model given in ( 2.5 ), the changes in total output can be given as follow:

$$X = (I - A)^{-1}Y$$
$$X = LY \qquad (3.1)$$

$$\Delta X = L\Delta Y + \Delta LY$$

In equation ( 3.1 ), Δ means change in the quantity, so that $\Delta X$, $\Delta Y$, and $\Delta L$ represent the changes in total outputs, final demands, and technological efficiency in relative sizes. The first term on the right-hand side of the equation represents the direct effect of changes of the demand on the economy. The second term on the right-hand side of the equation represents the indirect effect of changes in the structure of the economy. By decomposing changes in economic activity into direct and indirect effects, SDA allows researchers to identify the underlying drivers of change in an economy.

The SDA method is often further enhanced and widely applied in a variety of fields, including environmental studies (De Haan, 2001), energy policy (Alcántara and Duarte, 2004), and international trade (Xu and Dietzenbacher, 2014). SDA can provide valuable insights into the drivers of change and can inform policy decisions aimed at achieving specific economic or environmental objectives.

Taking the EEIO Model described by equation ( 2.7 ) for instance. Applying the same concept of differentiation in equation ( 3.1 ), the change in consumption-based emission can be rewritten into equation ( 3.2 ) below:

$$\Delta C_{cba} = EL\Delta Y + E\Delta LY + \Delta \mathrm{E}LY \qquad (3.2)$$

In equation ( 3.2 ), the drivers for consumption-based emission can be decomposed into three factors, adding emission intensity as another factor for the increase in emissions induced. Under the same rationale, researchers can further change the investigating variable into energy consumption, and air pollutions etc. They can also further decompose the investigated variables into more drivers such as population change, wealth accumulation and so on. However, SDA merely provides an accounting tool to tell a story of what happened in the past economic performance. Rigorous economic reasoning is lacked in SDA and thus limited applications have been extended to more complex economic analysis (Rose and Casler, 1996).

### 3.1.2. Hypothetical Extraction Method

In the attempt to measure the interindustry linkages at the comprehensive macro level, the Hypothetical Extraction Method (HEM) is proposed as another development on IO the model. Based on the classic IO model given in ( 2.5 ), HEM first assumes one or multiple industries are removed from the economy. Specifically, if the industry $k$ is to be removed for investigation, the $k$th row and column of the $A$ matrix and $k$th row of the final demand vector $Y$ are changed to zeros accordingly to form the new technical coefficient matrix $\bar{A}$ and final demand vector $\bar{Y}$. The new total output $\bar{X}$ will be given as:

$$\bar{X} = (I - \bar{A})^{-1}\bar{Y} \tag{3.3}$$

Hence, the comprehensive impact of industry $k$ on the entire economy can be simply obtained from the difference between $X$ and $\bar{X}$ as below in equation ( 3.4 ).

$$X - \bar{X} = (I - A)^{-1}Y - (I - \bar{A})^{-1}\bar{Y} \tag{3.4}$$

The same HEM concept can also be applied to study the interregional linkages be removing a region from the MRIO table (Tormo García et al.). Similarly, if emission and energy consumption extension are to be studied, the respective intensity vectors are multiplied to both sides of the equation (Zhao et al., 2015). HEM is thus widely used to investigate the hypothetical impact of a certain industry on the entire economy. However, like SDA, HEM is often considered limited in its ability to give economic reasoning (Dietzenbacher et al., 2019).

### 3.1.3. Hybrid Input-Output

The hybrid input output model (Jongdeepaisal and Nasu, 2020) is also a useful improvement from the classical IO Model, mostly used in techno economic studies to assess the interindustry impact of a specific technology on the entire economy. Based on life cycle assessment information, the input requirement of a certain technology is integrated into the technology coefficient $A$ by adding a corresponding column and row. Hence, the classic IO Model in equation ( 2.4 ) is again changed to construct a new IO relationship as follow:

$$\begin{bmatrix} X_p \\ X \end{bmatrix} = \begin{bmatrix} Y_p \\ Y \end{bmatrix} + \begin{bmatrix} Z_p & C_d \\ C_u & Z \end{bmatrix} \tag{3.5}$$

In equation ( 3.5 ), $X$, $Y$, and $Z$ bear their original meaning of total output, total demand, and intermediate input in the classical IO Model. $X_p$, $Y_p$, and $Z_p$ are the output, final demand, and internal intermediate input-output process of the investigated product $p$. $C_u$ is the upstream cut-off submatrix that describes the upstream inputs taken from the economy. $C_d$ is the downstream cut-off submatrix that describes the downstream outputs delivered for intermediate consumptions in the economy. In other words, the new IO relationship in equation ( 3.5 ) describes an economy with an additional production sector added into the economy. IO analysis can then be applied to the new economic system and hence compare the impacts that the added sector has on the entire economy.

However, the limitation for hybrid IO model is obvious - manual assignment of intermediate input is prone to high inaccuracy. In hybrid IO research, the determination of intermediate input has to rely on subjective expert judgement (Sharrard et al., 2008).

### 3.1.4. Input-Output-Based Network Analysis

A more recent methodological innovation for IO is to combine network analysis techniques with IO analysis (He et al., 2021). In network studies, nodal analysis is a commonly used technique to investigate the flow of transections among different nodes. In an IO economic system, each sector is similar to a node in a network, where nodes interact with each other to form a closed loop system. Hence, based on the betweenness analysis in network studies, the following equation is developed:

$$B_i = TJ_iTY \qquad (3.6)$$

In equation ( 3.6 ), $B_i$ means the betweenness of sector $i$, which measures the importance of the sector in terms of its role in acting as the transmitting point for upstream and downstream productions and consumptions. $T = A + A^2 + A^3 + \cdots = (I - A)^{-1}A$ represents the infinite extension of interindustry linkages towards upstream or downstream consumptions and productions. $J_i$ is a selection vector with 1 registered in the $i$th sector and 0s in other sectors. Compared to other IO analysis, the IO-based network analysis provides a different perspective in quantitatively accounting for the importance of sectors in transmitting productions, hence enable policy makers to devise policies accordingly.

IO-based network analysis thus helps us better understand the structure and symbiosis of economies. However, the IO-based network analysis only provides an account for economic symbiosis over a longer period. The intricacy of economic linkages in higher time resolution is unfortunately overlooked (Xu and Liang, 2019).

### 3.1.5. Improvements of Input-Output Model in the time domain

In addition to the efforts in extending IO modelling, some efforts are also devoted in improving IO model towards time domain analysis and thus predictions for economic performances. In fact, Leontief himself delved into the realm of the time domain of IO by introducing the dynamic IO model (Leontief, 1970, Leontief, 1953), attempting to explain the interaction between capital investment and production efficiencies on a chronological basis (Duchin and Szyld, 1985). Even though some researchers are still working towards the perfection and application of the dynamic IO model (Aulin-Ahmavaara, 2000, Rocco, 2019) to better factor in the impact of capital formation on productivity as a solution for economic prediction, such an approach has been criticized by other researchers for a number of reasons, such as omission of production factor constraints and inability to account for labour inputs (Kurz and Salvadori, 2000). Besides the classic demand driven IO model, attempts have also been made to construct a supply-driven IO model in the 1980s (Oosterhaven, 1988), but few developments have been fulfilled until its recent emphasis in disaster event analysis (Galbusera and Giannopoulos, 2018, Yagi et al., 2020), reigniting the discussion on supply-driven IO (Reyes and Mendoza, 2021). Skolka (1989) proposed Taylor Expansion on SDA to explain the induced intermediate production over time. More recent discussion has proposed the temporal Leontief inverse, which compromises between SDA and dynamic IO (Okuyama et al., 2006, Avelino et al., 2021). In addition, several other analysis tools build on the idea of chronological impact in IO model, such as dynamic inoperability (DIIM), supply bottlenecking (ARIO), and hybridization with Computational General Equilibrium (CGE) model, to analyse the impact of disasters in the long term (Mendoza-Tinoco et al., 2017, Zeng et al., 2019b, Zeng and Guan, 2020b, Guan et al., 2020). Such developments normally include non-linear characteristics, assuming a final steady state will be reached given a long-term general equilibrium.

Building on this idea, the most well developed and widely applied superset of IO model is the CGE model. As discussed by Koks et al. (2015), CGE models focuses on the long-run future equilibrium state, contrasting to the nearer future focus of IO models. As an economic model widely adopted by governments of developed countries, the CGE model uses numerical methods to simulate the behaviour and interaction of economic agents. CGE models are often used to analyse the impact of various economic policies, such as taxation, trade, fiscal and environmental policies, as well as shocks to an economy. It assumes that supply and demand in all markets reach a state of balance or equilibrium among various economic agents and their interactions across multiple markets. In the settings of most CGE model, households maximize utility by consuming goods and services, while firms maximize profits by producing goods and services using inputs like labour and capital. Governments may set policies and collect taxes to provide public goods and services. The modelled economy is also further divided into various specific markets, such as goods, services, labour, and capital markets, in which prices are determined by the interaction between supply and demand. Furthermore, production and consumption functions are introduced to quantitatively describe the relationships between inputs and outputs in the production process, as well as the preferences of consumers for different goods and services. Constraints such as budget constraints for households, production constraints for firms, and fiscal constraints for governments are introduced on economic agents to factor in more realistic interactions in an economy. Depending on the research question, CGE models can be static or dynamic, and disaggregated to represent different sectors, regions, or countries, hence capturing more detailed economic interactions

(Zhou and Chen, 2021).

However, the expensive human capital investment required for the learning, building, and using of CGE remains a significant entry barrier for many scholars. This complexity can make it difficult for non-experts to understand and evaluate the model's results (Rose, 1995). With large number of parameters based on estimations or calibrations, the accuracy of CGE is questioned by many (Burfisher, 2021), testified by the empirical evidence, such as its failure to forecast the magnitude of economic downturn in 1990s Spain (Polo and Viejo, 2015). Data limitations, measurement errors, and uncertainty in parameter values can also introduce biases and reduce the model's predictive power (Babatunde et al., 2017). Although economic theorists may produce logically sound modelling frameworks for CGE, the hardest part that directly deteriorates its accuracy in real world application is the lack of information on parameters setting (Dixon and Rimmer, 2009, Zhou and Chen, 2020). Lou (2016) discussed the possibility of using big data in existing economic models such as CGE, a complex economic model that factors in multiple economic parameters to improve its time-series economic analysis, but also admits that using CGE to analyse shorter term economic performance is not possible. A simpler version of chronological IO variant that focuses on the nearer future event would thus be a great complement to the existing toolbox of economic modelling.

## 3.2. Sequential Interindustry Model (SIM)

Jumping out of the general equilibrium assumption, Romanoff and Levine (1977) propose the SIM as a new strain of IO model innovation to pollinate the time domain analysis into the IO model in a culminative linear interaction. Similar to the dynamic material flow analysis (B. Müller, 2006), the SIM hybridises the time lag in demand propagation with the IO model by building a direct linear linkage between future and past economic activities. Some real-world SIM applications have also been made by later researchers. For instance, Okuyama et al. (2004) and Okuyama et al. (2000) use the SIM to assess the economic impact of natural disasters in Japan by using quarterly disaggregated hypothetical data. However, past attempts with the SIM proposed technical coefficient changes based only on hypothetical estimations, which greatly aggravates the uncertainties of modelling outcomes and limits further applications. This is largely because the data demands of the SIM are not easily met (Levine et al., 2007), thus hindering the SIM from fulfilling its deeper potential. It is also the reason why many IO researches merely mention and discuss the SIM in literature reviews but resolve to other IO variations as the tool for IO related modelling involving chronological analysis (Barker and Santos, 2010, Malik et al., 2014, Mendoza-Tinoco et al., 2017, Yu et al., 2013, Avelino and Hewings, 2019, Avelino, 2017).

Nevertheless, none of the existing SIM research has extended their time domain analysis to a resolution higher than monthly. The reasons that there is hardly any advancement in the SIM model towards high time-resolution innovation are as follows. 1) Researchers are too convinced that the IO model is an economic model that obeys the principle of general equilibrium, so efforts are concentrated on introducing more economic concepts, such as inoperability (Yu et al., 2013, Barker and Santos, 2010), to build a “perfect economic model”, but the fundamental interactions between economic in a physical way are sadly overlooked. If SIM is to be applied in a short run disequilibrium setting under which macroeconomic structural changes are less influential factors to be considered, SIM’s potential in regression with big data can then be further exploited. 2) The data requirement for high time-resolution analysis is costly. National IO tables are normally produced every 3-5 years. It takes great effort from IO scholars to increase the time resolution on a yearly basis (Avelino, 2017) to match the annually or quarterly updated economic indicators. Data unavailability has disincentivized model builders from working on a theoretical model with limited applications. Thus, limited methodological advancement has been made based on the concept proposed by SIM.

As Leontief pointed out at his later stage of research, the IO model is not simply an economic method but can also be understood from the technical/engineering perspective (Leontief, 1991). This research shares the same spirit of Coluzzi et al. (2011) to look at the IO model from the perspective of data science. Through an innovative algorithm, this research creatively simplified the complicated interactions among economic sectors proposed by SIM into a solvable linear system. Then, this research pollinated linear regression technique into the new algorithm, so that the best fitted technical coefficient from sectorial demands and outputs observations can be reversely calculated across the discrete time intervals, instead of hypothetically and less accurately “divide” production coefficients along the time domain like past SIM modellers did. A simulation based on dummy input has been conducted in this research to show that the algorithm proposed can accurately calculate the unknown chronological production coefficients based on simply output observations. The innovated algorithm will thus offer potential applications both in data science and in econometric interdisciplinary research. It opens the “black box” to reveal the relationships that cannot be told by conventional big data tools such as machine learning algorithms. Thus, scholars could have a new perspective to describe the structure of an economy in a chronological way. If the production coefficient of SIM is to be changed, it will be possible to investigate the impact of an external shock, such as nature disaster, on the economy in a timely dynamic way. On the other hand, if inputted with adequate proxy data, the proposed algorithm may extensively alleviate the cost associated with IO table compilation, revolutionizing the way of conducting large scale economic analysis.

To look into the SIM, recall that the classic IO model in equation ( 2.5 ) can be rewritten as follow using the Taylor expansion:

$$X = IY + AY + A^2Y + \cdots + A^nY \quad (n \to \infty) \tag{3.7}$$

The physical meaning of equation ( 3.7 ) is that the output induced by demand takes effect through layers of intermediate productions. SIM then introduces the time domain by splitting $x$ and $y$ into finite discrete time intervals as time series, meaning that $x$ and $y$ can be rewritten as $X_{(t)}$ and $Y_{(t)}$, respectively. Equations ( 3.8 ) and ( 3.9 ) illustrate the principle of this conversion.

$$\begin{aligned} X &= \sum_{t=1}^{m+n} X_{(t)} \\ X &= X_{(0)} + X_{(1)} + X_{(2)} + \cdots + X_{(m)} + \cdots + X_{(m+n)} \qquad (n \to \infty) \end{aligned} \tag{3.8}$$

$$\begin{aligned} \mathrm{Y} &= \sum_{t=1}^{m} Y_{(t)} \\ Y &= Y_{(0)} + Y_{(1)} + Y_{(2)} + \cdots + Y_{(m)} \end{aligned} \tag{3.9}$$

For each of the terms, the subscript denotes the specific demand and output at discrete time $t$. $m$ is the number of discrete time interval investigated. $n$ is the number of propagation layers that ideally approach infinity, the same as demonstrated by Taylor expansion in equation ( 3.7 ). It means that for $m$ observations of $X_{(t)}$ and $Y_{(t)}$, due to productional propagations, changes on output $X_{(t)}$ will theoretically extend into the infinitely distant future. Hence, considering the time lag feature of $X_{(t)}$ and $Y_{(t)}$, equation ( 2.5 ) will then be changed to equation ( 3.10 ) below.

$$X_{(t)} = Y_{(t)} + AX_{(t-1)} \qquad \forall\, t > 1, \quad y_{(t)} = 0 \text{ if } t > m \tag{3.10}$$

In equation ( 3.10 ), a recursive algorithm is introduced to obtain the output $X_{(t)}$ at time $t$ based on two variables: the current demand $Y_{(t)}$ at time $t$ and total output $X_{(t-1)}$ at the previous discrete time of $(t-1)$. The rationale is that what has been produced "today" will signal intermediate production "tomorrow" and propagate into the more distant future recursively. For instance, a hundred cars are manufactured and consumed "today", so that four hundred tyres are used in the inventory of car making factories. Thus, receiving the market signal, tyre manufacturers will produce four hundred tyres "tomorrow" to respond to the consumption signal sent out “today”. Since SIM discusses the state of economic structure in a very short period, it is assumed that the economy will not be able to respond by investing in capital equipment or similar means to change its structure, but to change production level in the face of differed market signal. Although the length of time duration investigated is $m$, $n$ more discrete time is extended into the future as the residue of the outputs induced by demand induced by outputs before time $m$ takes effect.

The following equations illustrate equation ( 3.10 ) in its expanded form to help understanding the mathematics and the underlying idea.

$$
\begin{aligned}
X_{(0)} &= Y_{(0)} \\
X_{(1)} &= Y_{(1)} + AX_{(0)} = Y_{(1)} + AY_{(0)} \\
X_{(2)} &= Y_{(2)} + AX_{(1)} = Y_{(2)} + AY_{(1)} + A^2Y_{(0)} \\
&\dots \dots \\
X_{(n)} &= Y_{(n)} + AX_{(n-1)} = Y_{(n)} + AY_{(n-1)} + A^2Y_{(n-2)} + \cdots + A^nY_{(0)} \\
&\dots \dots \\
X_{(m)} &= Y_{(m)} + AX_{(m-1)} = Y_{(m)} + AY_{(m-1)} + A^2Y_{(m-2)} + \cdots + A^nY_{(m-n)} \\
X_{(m+1)} &= AY_{(m)} + A^2Y_{(m-1)} + \cdots + A^nY_{(m-n+1)} \\
X_{(m+2)} &= A^2Y_{(m)} + A^3Y_{(m-1)} + \cdots + A^nY_{(m-n+2)} \\
&\dots \dots \\
X_{(m+n-1)} &= A^{n-1}Y_{(m)} + A^nY_{(m-1)} \\
X_{(m+n)} &= A^nY_{(m)}
\end{aligned}
$$

Since the row-sum of elements of production coefficients in $A$ are all smaller than 1 as input must not exceed output for any productions, $A^n$ converges to zero as $n$ becomes sufficiently large. It can thus be assumed that terms with $A$’s degree powers higher than $n$ to be neglected.

Bearing this concept in mind, the algorithm introduces next the time domain into the production coefficient $A$ in a similar manner to that of equations ( 3.8 ) and ( 3.9 ). $A$ is thus converted into equation (6) as follows:

$$A = A_{(1)} + A_{(2)} + \cdots + A_{(l)} \qquad ( 3.11 )$$

The reason for splitting $A$ is to reflect the time lag that occurs between different sectors due to technical constraints (Romanoff and Levine, 1986). Hence, the modified production coefficients not only reflect the magnitude of intermediate inputs, but also the timing for the intermediate inputs to be fulfilled. For instance, tyres in the car manufacturing industry may be made quicker than cars’ control chips due to reasons like the nature of engineering process and distances among different industry clusters. Such information cannot be captured by $A$ in traditional IO models, but its spitted version in SIM as described by equation ( 3.11 ). An earlier simple attempt by Okuyama (2004a) splits $A$ into three stages. Equation ( 3.10 ) is thus further modified into equation ( 3.12 ), the general form of SIM as follows:

$$X_{(t)} = Y_{(t)} + \sum_{i=1}^{l} A_{(i)} X_{(t-i)} \qquad \forall\, i < t \quad y_{(t)} = 0 \text{ if } t > m \qquad (3.12)$$

In an early work of Romanoff and Levine (1981), equation ( 3.12 ) is defined as the responsive production SIM where productions are responses to demand in the past. A further advancement in SIM modelling has been proposed to include anticipatory production of demand from the future. Similarly, there is also proposal to change the Leontief IO model from demand-driven to supply-driven Ghosh model, so that outputs are results of supply for productions instead of demand induced productions (de Mesnard, 2007). To avoid unnecessary complications, this research will only work with the general responsive production form of SIM.

For the purpose of illustration, the SIM process described in equation ( 3.12 ) is shown visually using the schematic diagram given in Figure 9 below. Each horizontal row shows how the output $X_{(t)}$ is comprised at the respective time $t$. Initially in the first row, it shows the state where $A_{(t)}$ and $Y_{(t)}$ have not started interacting with each other. Until it moves to the second row of $t = 0$, $Y_{(0)}$ is first multiplied by the identity matrix $I$ to produce $X_{(0)}$, as highlighted in red on the left side. It suggests that output will match the market demand for the first time interval. In the next discrete time of $t = 1$, output at the previous discrete time, $X_{(0)}$, is multiplied by $A_{(1)}$ to produce one element that is to be added to the product of $Y_{(1)}$ and $I$, according to the aforementioned assumption that output will match the instant demand. The product of $X_{(0)}$ and $A_{(1)}$ indicates that the production coefficient on the second discrete time will be applied to the output at first discrete time. It is to simulate the time-lagged response of intermediate productions to output of previous time discrete. The added sum then produces $X_{(1)}$ and is to be used in the next step of $t = 2$. Theoretically, this procedure is repeated for infinitely many processes. Essentially, readers may want to visualize the process as moving the series of $A_{(t)}$ and $X_{(t)}$ towards each other to perform a convolution-like operation, as demonstrated in Figure 9.

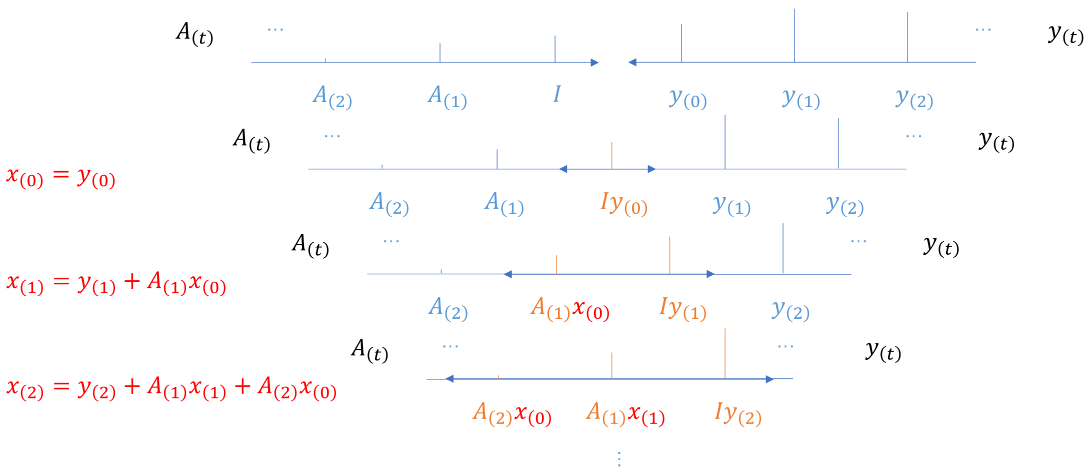



*Figure 9 The algorithm process in a visual diagram.*

For easier discussion, we constructed a SIM system with a dummy production coefficient matrix $A_{(t)}$ with two layers as shown in Table 3 below, where $A_{(1)}$ and $A_{(2)}$ shows the interaction of the two sectors in t=1 and t=2 respectively.

| | $A_{(1)}$ | | $A_{(2)}$ | |
|---|---|---|---|---|
| | Sector A | Sector B | Sector A | Sector B |
| Sector A | 0.2 | 0 | 0 | 0.7 |
| Sector B | 0 | 0.375 | 0.25 | 0 |

*Table 3 A simulated production coefficient matrix in time-series of two sectors and two layers*

Assuming two unitary demand of 1 happens for Sector A at $t = 0$ and for Sector B at $t = 1$ respectively, denoted as vectors $Y_{(0)} = \begin{bmatrix}1\\0\end{bmatrix}$ and $Y_{(1)} = \begin{bmatrix}0\\1\end{bmatrix}$, the following output of $X_{(t)}$ can be then obtained according to equation ( 3.12 ):

$$X_{(0)} = Y_{(0)} = \begin{bmatrix}1\\0\end{bmatrix}$$

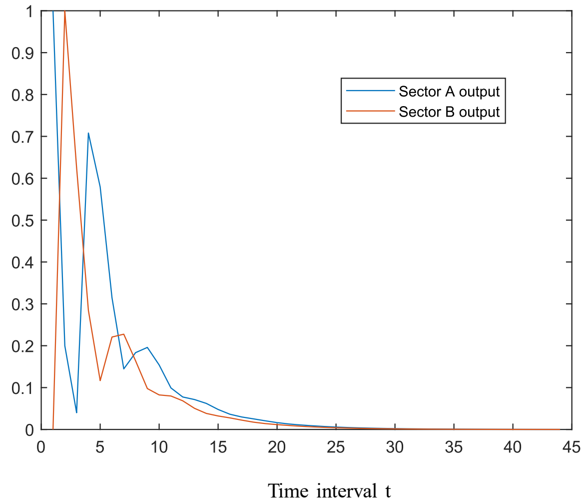
$$X_{(1)} = Y_{(1)} + A_{(1)}\, X_{(0)} = \begin{bmatrix}0\\1\end{bmatrix} + \begin{bmatrix}0.2 & 0\\0 & 0.375\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0.2\\1\end{bmatrix}$$

$$X_{(2)} = A_{(1)}X_{(1)} + A_{(2)}X_{(0)} = \begin{bmatrix}0.2 & 0\\0 & 0.375\end{bmatrix}\begin{bmatrix}0.2\\1\end{bmatrix} + \begin{bmatrix}0 & 0.7\\0.25 & 0\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0.04\\0.625\end{bmatrix}$$

$$X_{(3)} = A_{(1)}X_{(2)} + A_{(2)}X_{(1)} = \begin{bmatrix}0.2 & 0\\0 & 0.375\end{bmatrix}\begin{bmatrix}0.04\\0.625\end{bmatrix} + \begin{bmatrix}0 & 0.7\\0.25 & 0\end{bmatrix}\begin{bmatrix}0.2\\1\end{bmatrix} = \begin{bmatrix}0.708\\0.2844\end{bmatrix}$$

… …

By organizing and plotting $X_{(t)}$ in Figure 10, it is easily observed that under SIM, the final output will decay towards zeros, validating the deductions in SIM made previously.




| | Sector A output | Sector B output |
|---|---|---|
| t=1 | 1.000 | 0.000 |
| t=2 | 0.200 | 1.000 |
| t=3 | 0.040 | 0.625 |
| t=4 | 0.708 | 0.284 |
| t=5 | 0.579 | 0.117 |
| t=6 | 0.315 | 0.221 |
| t=7 | 0.145 | 0.228 |
| t=8 | 0.183 | 0.164 |
| t=9 | 0.196 | 0.098 |
| t=10 | 0.154 | 0.082 |
| t=11 | 0.099 | 0.080 |
| t=12 | 0.078 | 0.068 |
| t=13 | 0.071 | 0.050 |
| t=14 | 0.062 | 0.038 |
| t=15 | 0.048 | 0.032 |
| t=16 | 0.036 | 0.028 |
| t=17 | 0.030 | 0.022 |
| t=18 | 0.025 | 0.017 |
| t=19 | 0.021 | 0.014 |
| t=20 | 0.016 | 0.012 |
| t=21 | 0.013 | 0.010 |
| | …… | |

*Figure 10 The decaying output over time as a response to the illustrative unitary demands in sector A and B and production coefficients $A_{(t)}$.*

## 3.3. Model Innovation

Nonetheless, the chronological production coefficient $A_{(t)}$ is often unknown to modelers. The conventional approach involves making rough estimations of the time delay and the ratios of intermediate inputs at each step (Okuyama et al., 2004). This method is susceptible to inaccuracies and data insufficiency. If the coefficient $A_{(t)}$could be obtained through quantitative calculations derived from economic output observations, i.e. $X_{(t)}$ and $Y_{(t)}$, the effectiveness of the SIM model could be significantly enhanced, allowing for a more accurate representation of the complex interdependencies in economic systems and leading to better-informed decision-making and more robust policy development.

The objective of this chapter is to determine the production coefficient $A_{(t)}$ in a time series through extensive observations of total output $X_{(t)}$ and final demand $Y_{(t)}$. Although Levine et al. (2007) proposed using Z transformation to address the convolution-like problem inherent in the SIM model, there is no mathematical solution for applying Z transformation to matrix variables in the frequency domain. As a result, no methodological advancements have been made to solve the SIM model. Considering the advancements in programming capacity developed in recent years, it is natural to explore solving the SIM model in the time domain. By leveraging modern computational capabilities, the potential for obtaining a more accurate representation of the production coefficient $A_{(t)}$could lead to improved modelling of economic systems and a deeper understanding of the interdependencies between sectors. This in turn could contribute to better-informed policy decisions and more effective strategies for economic development.

Hence, the SIM model is to be solved within the time domain. In the proposed algorithm, it assumes that productivity $A_{(t)}$ is unchanged, and production activity $X_{(t)}$ is only signalled by consumption activities $Y_{(t)}$ as a responsive demand.

To reversely work out $A_{(t)}$, let us first write out equation ( 3.12 ) as follows:

$$
\begin{aligned}
X_{(0)} &= Y_{(0)} \\
X_{(1)} &= Y_{(1)} + A_{(1)}X_{(0)} \\
&= Y_{(1)} + A_{(1)}Y_{(0)} \\
X_{(2)} &= Y_{(2)} + A_{(1)}X_{(1)} + A_{(2)}X_{(0)} \\
&= Y_{(2)} + A_{(1)}Y_{(1)} + \left({A_{(1)}}^2 + A_{(2)}\right)Y_{(0)} \\
X_{(3)} &= Y_{(3)} + A_{(1)}X_{(2)} + A_{(2)}X_{(1)} + A_{(3)}X_{(0)} \\
&= Y_{(3)} + A_{(1)}Y_{(2)} + \left({A_{(1)}}^2 + A_{(2)}\right)Y_{(1)} \\
&\qquad\qquad + \left({A_{(1)}}^3 + A_{(1)}A_{(2)} + A_{(2)}A_{(1)} + A_{(3)}\right)Y_{(0)} \\
&\dots\dots
\end{aligned}
$$

It can be rewritten again into a linear system in matrix form, as shown in equation ( 3.13 ):

$$
\begin{bmatrix} I & A_{(1)} & \left({A_{(1)}}^2 + A_{(2)}\right) & \left({A_{(1)}}^3 + A_{(1)}A_{(2)} + A_{(2)}A_{(1)} + A_{(3)}\right) & \cdots \end{bmatrix}
\times \begin{bmatrix} Y_{(n)} & Y_{(n+1)} & \cdots \\ Y_{(n-1)} & Y_{(n)} & \cdots \\ \vdots & \vdots & \cdots \\ Y_{(0)} & Y_{(1)} & \cdots \end{bmatrix}
= \begin{bmatrix} X_{(n)} & X_{(n+1)} & \cdots \end{bmatrix} \qquad (3.13)
$$

Let there be $p$ sectors. If we take $(n+1) \times p$ observations of $y$ to make the second matrix on the left a symmetric matrix, then the first term in equation ( 3.13 ) is solvable by taking an inverse of the second term, as shown in equation ( 3.14 ).

$$
\begin{bmatrix} I & A_{(1)} & \left({A_{(1)}}^2 + A_{(2)}\right) & \left({A_{(1)}}^3 + A_{(1)}A_{(2)} + A_{(2)}A_{(1)} + A_{(3)}\right) & \cdots \end{bmatrix}
= \begin{bmatrix} X_{(n)} & X_{(n+1)} & \cdots \end{bmatrix} \times \begin{bmatrix} Y_{(n)} & Y_{(n+1)} & \cdots \\ Y_{(n-1)} & Y_{(n)} & \cdots \\ \vdots & \vdots & \cdots \\ Y_{(0)} & Y_{(1)} & \cdots \end{bmatrix}^{-1} \qquad (3.14)
$$

By observing the term on the left-hand side of the equation, the time subscripts and power degrees of ${A_{(t)}}^q$ in each of the elements all sum up to be the same. For instance, the element $\left({A_{(1)}}^3 + A_{(1)}A_{(2)} + A_{(2)}A_{(1)} + A_{(3)}\right)$ has its terms' time subscripts and power degrees all sum up to 3. Physically, this can be explained by the different paths taken by production operations to reach the current layer; i.e., in $\left({A_{(1)}}^3 + A_{(1)}A_{(2)} + A_{(2)}A_{(1)} + A_{(3)}\right)$, the production sectors that are "closer" to each other and linked by $A_{(1)}$ will be induced three times compared to only once by $A_{(3)}$. The algorithm calculating the combinations that satisfy such a requirement is called a partition in number theory, for which plenty of well-programmed functions in multiple programming platforms for its realization are available. Hence, knowing $A_{(1)}$ in the beginning, $A_{(2)}$ can first be obtained according to the expression of $A_{(2)}$ shown in the second row of equation set ( 3.15 ). Having $A_{(1)}$ and $A_{(2)}$, it is then feasible to obtain $A_{(3)}$ according to the third row of equation set ( 3.15 ). Thus, any $A_{(t)}$ for $t > 1$ can be obtained by recursive induction step by step.

$$\begin{aligned}
A_{(1)} &= A_{(1)} \\
A_{(2)} &= \left({A_{(1)}}^2 + A_{(2)}\right) - {A_{(1)}}^2 \\
A_{(3)} &= \left({A_{(1)}}^3 + A_{(1)}A_{(2)} + A_{(2)}A_{(1)} + A_{(3)}\right) \\
&\qquad - \left({A_{(1)}}^3 + A_{(1)}A_{(2)} + A_{(2)}A_{(1)}\right) \\
&\dots\dots
\end{aligned} \tag{3.15}$$

Equations ( 3.13 ) to ( 3.15 ) describe a mathematic algorithm to linearly solve for an ideal SIM system. To apply SIM into the real world, it needs to combine statistical and econometric research techniques with SIM to compensate for uncertainties. Linear regression is thus introduced to minimize the summation of square errors, so as to find the best fitted solution to $A_{(t)}$ based on observations of $y_{(t)}$ and $x_{(t)}$.

First, let the variables in equation ( 3.13 ) become:

$$B = \left[I \quad A_{(1)} \quad \left({A_{(1)}}^2 + A_{(2)}\right) \quad \left({A_{(1)}}^3 + A_{(1)}A_{(2)} + A_{(2)}A_{(1)} + A_{(3)}\right) \quad \cdots\right]$$

$$U = \begin{bmatrix} Y_{(n)} & Y_{(n+1)} & \cdots \\ Y_{(n-1)} & Y_{(n)} & \cdots \\ \vdots & \vdots & \cdots \\ Y_{(0)} & Y_{(1)} & \cdots \end{bmatrix}$$

$$V = [X_{(n)} \quad X_{(n+1)} \quad \cdots]$$

So that equation ( 3.13 ) becomes:

$$BU = V \tag{3.16}$$

The left-hand side of equation ( 3.16 ) is thus the estimated output based on $A_{(t)}$ and $y_{(t)}$, while the right-hand side is the real output. In equation ( 3.16 ), all variables are asymmetric matrices, unlike required by equation ( 3.14 ), to include all observations of $x_{(t)}$ and $y_{(t)}$ available. The error in estimation is thus easily obtained as $BU - V$. Since all variables are matrices, the sum of squared errors is thus given by a simple matrix operation as:

$$f(x) = (BU - V)^T(BU - V) \tag{3.17}$$

Where the superscript $T$ means transpose matrix. To find the value of $B$ that minimizes $f(x)$, the first derivative of $f(x)$ with respect to $B$ is taken, so that equation ( 3.17 ) becomes:

$$\begin{aligned} \frac{\delta f}{\delta B} &= \frac{\delta (BU - V)^T(BU - V)}{\delta B} \\ &= \frac{\delta (U^T B^T BU - 2V^T BU + V^T V)}{\delta B} \\ &= 2U^T B^T U - 2V^T U \end{aligned} \tag{3.18}$$

Then, equation ( 3.18 ) is set to zero and find the expression of $B$ at which minimizes the squared error sum given in equation ( 3.17 ). The following matrix operations are thus performed.

$$\begin{aligned} 0 &= 2U^T B^T U - 2V^T U \\ U^T B^T U &= V^T U \\ BUU^T &= VU^T \\ BUU^T(UU^T)^{-1} &= VU^T(UU^T)^{-1} \\ B &= VU^+ \end{aligned} \tag{3.19}$$

In equation ( 3.19 ), $U^+$ denotes the Moore-Penrose pseudoinverse matrix of $U$, widely used for linear regression problems (MacAusland, 2014). The best fitted $A_{(t)}$ can thus be easily obtained.

## 3.4. Simulation Results

To verify the efficacy of the innovative algorithm proposed, this section simulated a $X_{(t)}$ series based on the SIM interaction of a randomly generated 2-sector $Y_{(t)}$ series of 200 discrete time intervals and the 2-layer production system $A_{(t)}$ described in Table 3. The number of propagation layers is arbitrarily set to 40, so that the influence from demands at $Y_{(0)}$ will be negligible after output $X_{(42)}$, i.e. after 40 layers of propagation and 2 layers of production by $A_{(t)}$. Starting to take observations from $X_{(43)}$ (after 40 layers of propagation and 2 layers of production), 158 observations out of the 200 samples of $X_{(t)}$ are used to construct the $V$ matrix in equation ( 3.16 ). $U$ matrix and its pseudoinverse are also constructed using all 200 observations of $Y_{(t)}$ according to equation ( 3.19 ). Thus, the preliminary form of $A_{(t)}$ shown as $B$ in equation ( 3.16 ) is easily obtained. Performing the partition algorithm illustrated in equation ( 3.15 ) will give us the regressed production coefficient $A'_{(t)}$.

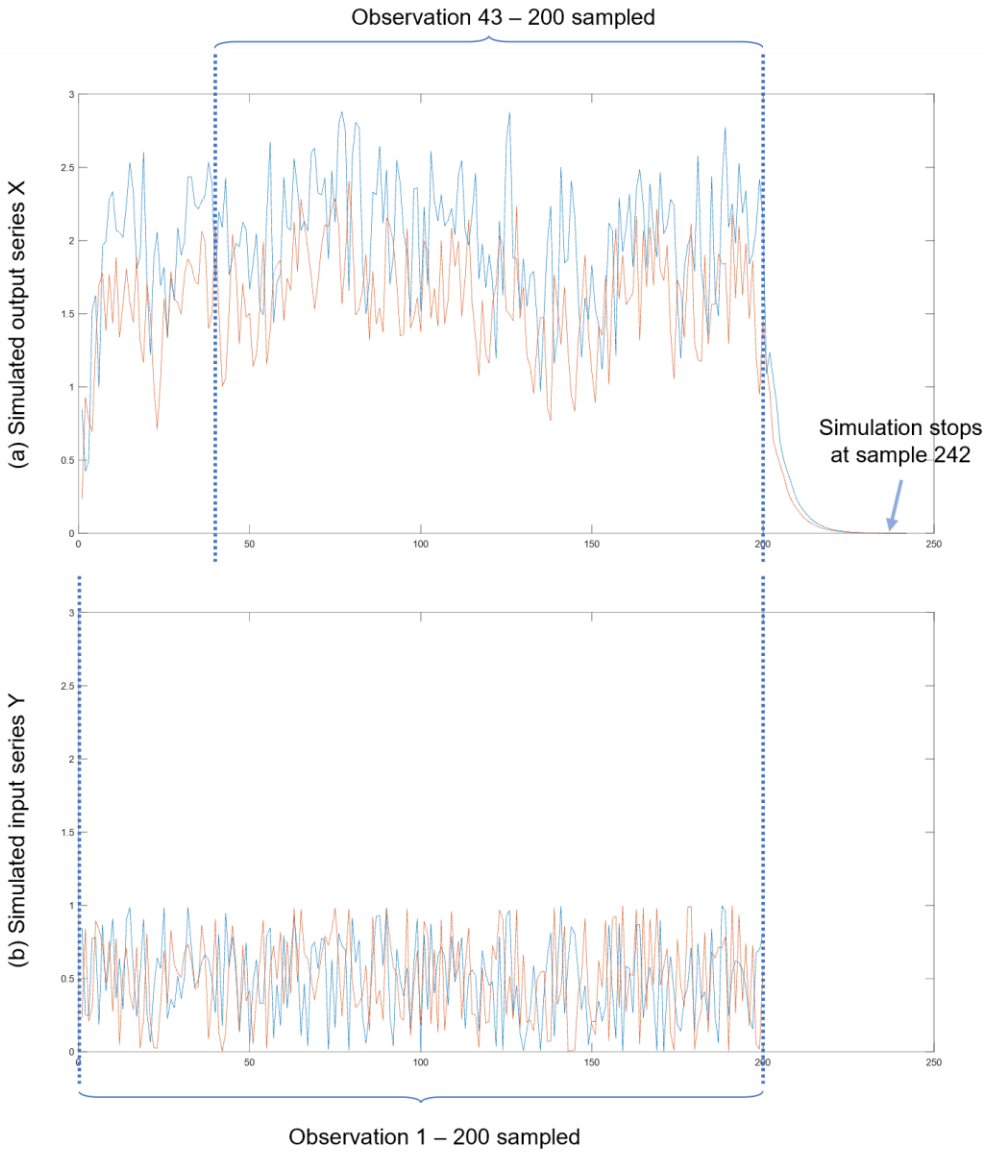



*Figure 11 (a) Simulated total output $x_{(t)}$ (b) Simulated final demand $y_{(t)}$.*

Table 4 shows the errors between the $A'_{(t)}$ obtained and the actual $A_{(t)}$. In terms of percentage error, the relative error level of the estimated production coefficient lies between 0.0029% and 0.0007%, minimal enough to be ignored and conclude that $A'_{(t)}$ and $A_{(t)}$ are basically identical, demonstrating the effectiveness of the innovated algorithm.

| | $A'_{(1)} - A_{(1)}$ | $A'_{(2)} - A_{(2)}$ |
|---|---|---|

| | Sector A | Sector B | Sector A | Sector B |
|---|---|---|---|---|
| Sector A | $-1.43\times10^{-6}$ | $-0.11\times10^{-6}$ | $-9.96\times10^{-6}$ | $-5.75\times10^{-6}$ |
| Sector B | $-1.01\times10^{-6}$ | $-7.70\times10^{-6}$ | $-7.22\times10^{-6}$ | $-2.72\times10^{-6}$ |

*Table 4 the errors between regressed production coefficient $A'_{(t)}$ and the actual production coefficient $A_{(t)}$.*

To quantitatively measure the squared mean error $e$, equation ( 3.20 ) is used:

$$e = \frac{\sum_i^p \sum_j^p \sum_l^t \left(a_{ij}'_{(l)} - a_{ij(l)}\right)^2}{q} \qquad ( 3.20 )$$

In equation ( 3.20 ), $a_{ij}'_{(l)}$ and $a_{ij(l)}$ is the $i$th column and $j$th row element at $l$th layer of $A'_{(t)}$ and $A_{(t)}$ respectively. $p$ and $t$ are the number of sectors and number of production layers respectively. In the simulation exercise, $p = 2$ and $t = 2$. $q$ is the total number of elements in $A'_{(t)}$ and $A_{(t)}$, being $2 \times 2 \times 2 = 8$ in this case. The squared mean error between $A'_{(t)}$ and $A_{(t)}$ is thus $4.66\times10^{-11}$ as calculated from Table 4, sufficiently small to conclude that $A'_{(t)}$ is a good enough approximation of $A_{(t)}$.

To further test the capability of the innovated algorithm for SIM, the squared mean errors are calculated under varied number of propagation layers. Figure 12 illustrates how the mean squared errors change with different number of propagation layers set. It is clearly seen that as the number of propagations increases, the mean square error between $A'_{(t)}$ and $A_{(t)}$ exponentially decreases, suggesting that $A'_{(t)}$ approaches the true value of $A_{(t)}$ as number of propagations increases. It makes sense as SIM models infinitely distant past demand to still have an impact on the present output, but minimal and negligible as propagation extends to the future. Increasing the number of propagations set in this algorithm factors in the increasingly minimal effect on output $X_{(t)}$ from demand $Y_{(t)}$, thus producing a more accurate $A'_{(t)}$ as solution to the system.

Another important observation is that the squared mean error drastically increases to $1\times10^{-2}$ after the number of propagation layers becomes more than 64, sufficiently large to invalidate the estimated $A'_{(t)}$. The reason is that insufficient number of sample available hinders the functioning of linear regression algorithm. When the number of propagation layers is 64, the dimension of $U$ in equation ( 3.16 ) becomes 134-by-134, a symmetric matrix just sufficient to solve for $B$. As the number of propagation layers increases beyond 64, the system described in equation ( 3.16 ) becomes an underdetermined system with no unique solution since the number of observations become less than the number of unknowns that needs to be solved. $A'_{(t)}$ thus deviates significantly from the true value $A_{(t)}$ in this specific case shown in Figure 12. If there are more observations available, more propagation layers can then be accommodated to improve the accuracy in the proposed algorithm further, so that the sudden jump shown in Figure 12 will be pushed further to the right.

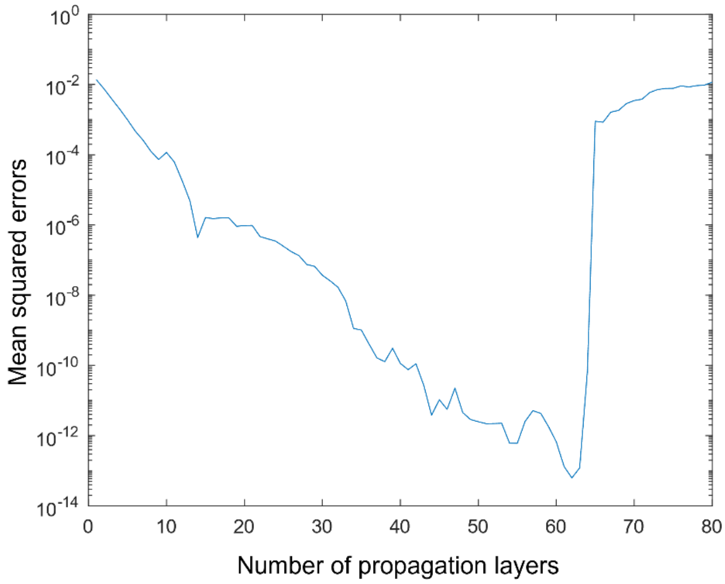



*Figure 12 Mean squared errors of the elements in the timed production coefficient $A_{(t)}$ against varied number of propagation layers taken.*

The simulation exercise in this section works on a system with known layers of production in $A_{(t)}$. In real world applications, the number of layers in production is normally unknown. It is thus sensible to vary both the number of production layers and number of propagation layers in real world applications to find the solution for best fitted $A_{(t)}$ with the minimum mean squared errors as the best approximation for chronological production coefficients.

## 3.5. Discussion

As a unique variation of IO model, SIM offers tremendous potential in economic system analysis. Unlike the other IO model variants which keep general equilibrium as an underlying assumption, the SIM is by nature a disequilibrium model that focuses on the nearer future analysis. Unfortunately, disproportionally fewer efforts and achievement in SIM methodological advancement have taken place since its first proposal in 1980s. This Chapter proposes an innovative algorithm as an important methodological advancement for the SIM. It complements with SIM by providing a way to integrate it with econometric linear regression analysis. Investigation into the chronologically extended production coefficients can provide temporal information into the interlinkages among economic sectors. For instance, although knowing the units of steel and the associated time needed to produce one unit of output in the manufacturing industry may be possible by conventional life cycle assessment methods, the cost of such an assessment will be tremendously high if comprehensively conducted at macro-level, not to mention the difficulty in constructing the sectorial interlinkages among all the industrial sectors in a similar manner as hybrid IO tables. The complementary algorithm to SIM provides a cheap and efficient way to quickly draw a comprehensive picture of chronological interlinkages based on economic activities in high time-resolution among sectors. Being more accurate than hypothetically constructed in previous studies, the linearly regressed temporal production coefficients of SIM can be reversely used for short future predictions of sectorial outputs given unscheduled events such as natural disaster. Also, the temporal production coefficients of economic sectors can be used to analyse the time lag between industries, as well as predict short future economic outcomes by sectors given consumer demands. Such a genuine approach bears great potentials in future econometric studies in terms of economic predictions. It will thus fill the gap in approaching econometric study with data science tools.

Extending the improved model beyond economic research, other complex networks similar to IO systems with delay characteristics, such as the human body metabolism system, can borrow the method presented in this study for their unique applications. Unlike macro econometric data, observation data on smaller systems are normally much easier to be obtained. Analysis like the previously proposed can be conducted to investigate how materials interact among themselves in a temporal manner to offer insight into the systems investigated.

Nevertheless, the algorithm proposed in this study does not address some of the fundamental limitations of SIM. Firstly, SIM does not deal with the “bottlenecking” issue inherent to IO model. This refers to a scenario where demand surpasses current production capacity, and industries are unable to linearly scale up their production in response. Secondly, SIM does not differentiate between changes in capital and consumer goods. Most importantly, the algorithmic development for SIM in this Chapter has primarily focused on advancing the methodological aspects of the two motivations discussed in Chapter 1. This is only part of the solution, and there remains a broader challenge to be tackled before the improved model can be used. In particular, meeting the data requirement for high-frequency IO data presents a significant hurdle that must be overcome to achieve the research objectives. Meeting this requirement is a vital next step for this research.

# Chapter 4: The Application of SIM in Electricity Consumptions

This chapter aims to test the effectiveness of the algorithm designed for the SIM using electricity consumption data from Chongqing municipality as a stand-in for economic outputs. It begins with a review of existing literature on the use of electricity consumption data to analyse economic performance. Following this, the chapter details the organization of the data and the methodology employed. The subsequent section interprets the quantitative findings, assessing Chongqing's economic structure and offering short-term economic forecasts. The chapter concludes with a discussion on the limitations of the current approach and suggestions for future research.

## 4.1. Electricity Consumption and Economics

As illustrated in the simulation in Chapter 3, the time resolutions for demand and output observations must be high enough for the algorithm developed to induce practical applications. On the other hand, the data set must also be classified into economic sectors in a similar way as the IO model. Thus, daily or hourly demand and output observations on economic sectors should be preferably used for this algorithm. Although online transaction data across different sectors in monetary term is ideally the best, it is highly unlikely that such data are available due to technical constraints at present. As a sensible compromise, proxy data with high chronological resolutions on economic activities may be a solution to the stringent data quality requirement.

Fortunately, with the rapid development of information and communication technologies, researchers can now have alternatives to traditional economic statistic data for analysis of economic performances. From online purchases to utility bills, 2.5 quintillion bytes of data are generated per day (Forbes, 2018). Given the name “big data”, many recent works have engaged in the study of the values buried in binary registers transmitted and stored in companies’ servers. A vast number of studies have utilized big data to investigate human behaviours at the microscale (Wang et al., 2018, Yuan et al., 2020). Instead of the time-intensive consumption statistics measured in monetary terms, a vast number of studies have utilized high time-frequency big data to investigate economic activities at the microscale (Wang et al., 2018, Yuan et al., 2020). At macroscale, many human activity indicators like night-time lights (Mellander et al., 2015), mobile phone usage (Šćepanović et al., 2015), and primary energy consumption (Aslan et al., 2014) are also used as proxy to analyse the functioning of economics. Compared to classic econometric studies based on economic data (i.e. gross domestic production etc.), the emerging big-data economic research overcomes the barriers in data collection. Specifically, proxy indicators can be updated much more frequently and covers much higher resolution than conventional economic indicators.

Although big data economic research bears great potentials, the underlying economic theories are unproportionally overlooked in the current so-called big data economic research, which are mostly found to be still based on conventional economic indicators such as GDP (Zhao et al., 2018, Yaacob et al., 2021). These research, strictly speaking, are not qualified to be categorized as “big data” research widely agreed upon data scientists (De Mauro et al., 2016).

Electricity data, nevertheless, are a good exception to be used as a good data source for economic studies. Electricity consumption data are a typical genre of big data that satisfy the “3V” requirements (Volume, Velocity, and Variety). Some studies adopt recently popular machine learning (ML) tools such as artificial neural networks (Zeng et al., 2019a, Rahman et al., 2016), convolutional neural networks (Dong et al., 2017), back propagation neural networks (Naimur Rahman et al., 2016), pattern sequence-based forecasting (Perez-Chacon et al., 2020, Viloria et al., 2020), and clustering analysis (Zhou et al., 2017a) for the purpose of pure prediction and pattern recognition in economic studies based on electricity consumption data. However, such an approach does not factor in the economic rationales.

A dozen of other studies have also attempted to use electricity consumption data on the analysis of macro-economic performances (Ashraf et al., 2013, Kim, 2015, Zhang et al., 2017a). In recent literatures, high time-frequency electricity consumption data has been widely used in response to study the economic impact of COVID-19-related lockdowns (Janzen and Radulescu, 2020, Fezzi and Fanghella, 2020, López Prol and O, 2020). Other electricity consumption data are used to focus prediction on the electricity market itself. For instance, Novan et al. (2020) utilized 158,112 households in Sacramento, California, to investigate the household electricity consumption relationship with temperature. Unfortunately, no literature on electricity consumption analysis can solidly integrate proper economic theories into its modelling. Qu et al. (2015) advanced one step further in analysing the patterns recognized through ML regression, but the interpretation is still far from being called “economic”.

Concretely, ML is a group of regression computing algorithms that find the statistically optimized solution to a specific question using a large data set. By applying flexible criteria to a learning model, machine learning algorithms can quickly establish a correlation between input variables and intended prediction parameters. In addition to ML’s wide application in the internet industry, more econometric research has employed ML in the hope of making economic performance predictions. Unlike traditional econometric regression analysis, machine learning research does not require pre-setting of model parameters. The number of relevant research has increased drastically in recent years in both academia and industries. As summarized by Harding and Lamarche (2021), electricity utility bills are a typical form a good source of big data for energy economic research. Harding and Lamarche (2021) also listed a few ML tools that are commonly used in this kind of research. The least absolute shrinkage and selection operator (LASSO), probably the most well-known ML tool for economists (Mullainathan and Spiess, 2017), was used to predict electricity usage based on weather forecast data (Ludwig et al., 2015). A literature review on ML tools to be used in econometrics can be found in Varian (2014). Interested readers may refer to Hastie et al. (2009) for advanced learning on these tools.

However, as Crown (2019) has pointed out, a fundamental limitation of ML is its inability to provide economic explanations for the interaction processes among input variables. While machine learning can uncover relationships and make predictions, it often lacks the interpretability that is crucial for understanding the underlying economic mechanisms at play. In a few exceptional cases, a relationship of some sort may be deduced but cannot be supported with economic reasoning (Mullainathan and Spiess, 2017). Einav and Levin (2014) argues that the integration of regression ML algorithms and economic theory will be a persistent obstacle for data science and economic interdisciplinary researchers. The “trial and error” ML applications should be thoroughly revised to accommodate economic and engineering reasoning. It highlights the need to strike a balance between leveraging advanced computational methods and preserving the explanatory power inherent in traditional econometric techniques.

As sufficiently discussed in Chapter 3, the algorithm developed for SIM (Romanoff and Levine, 1977) in this thesis is exactly looking for a data source that bears the characteristics of electricity consumption big data. In past research, the applications of the SIM have been limited to disaster recovery analysis (Okuyama et al., 2000, Okuyama et al., 2004). Chapter 3 points out that economic sectors interact with each other in response to consumers’ demands through a production network – an identical concept inherited from the IO model – that carries out production in a step-by-step manner, an interaction well modelled by SIM. Under this theory, the economic outputs at each discrete time are the result of induced productions signalled by demands at previous discrete times. Using a regression algorithm, the chronological production coefficients can then be theoretically obtained from ample observation of demands and outputs.

In this Chapter, a data set of daily electricity consumption in Chongqing municipality of China is used to investigate the chronological and inter-sectorial linkages of economic sectors. The reason Chongqing is selected as the investigated case is because it is the most easily accessible data that meets the requirements of this study. Due to the ongoing global pandemic of Covid-19 at the time of writing, connection to institutions in other regions with the same data became extremely difficult. After assessment of the data available for Chongqing, it is determined that electricity consumption data in Chongqing can be used for this study to reveal the chronological economic symbiosis in this mega city of China.

The chronological interlinkage of economic sectors in Chongqing is thus obtained and used for multiple important purposes in this study. First, it can be a quantitative indicator helping researchers to understand how much input and how much time are needed from one sector to respond to a unitary output in another. Second, the interlinkages suggested can be used for disequilibrium short-term future predictions, supplementing current economic tools that focus on long-run general equilibrium. This application encompasses both the cross-sectorial lagged induced demand and the multiplier effects from the economic perspective. The robustness testing showed that the model predictions have a certain degree of reliability. Motivated by the model's good performance, three hypothetical scenarios are further created to simulate how a change in one sector quantitatively and chronologically affects all other sectors in the following two months, showing the varied multiplier effect of economic demands of different sectors. As a good example of fusion between econometrics and data science, this study establishes a basis for further investigation between economic and engineering theories.

## 4.2. Data Collection and Cleansing

This study used the electricity consumption data of the Chongqing State Grid Research Institute. The State Grid is the monopolistic electricity supplier in Chongqing, China. Chongqing has a population of 31 million people and an area of 82,000 $km^2$. Its GDP was $362 billion in 2020. Every electricity consumer, regardless of whether commercial or household in nature, pays their electric metre fare to State Grid. When registering the metre reading, the state grid also registers commercial customers' nature of business in accordance with the Industrial Classification for National Economic Activities (GB/T 4754-2017) (UNSD: 2006, International Standard Industrial Classification of all Economic Activities, NEQ). Among the 708 sectors listed in GB/T 4754-2017, Chongqing data includes 440 sectors. It sampled the daily electricity consumption data from 1 March 2018 to 21 November 2020, with 971 data points in total across the time dimension. For this research, the data collected from all commercially registered electricity metres of Chongqing are categorized into eight sectors based on the registration information.

As explained previously, electricity consumption data can be and have been used as proxy data for economic research. It is because that there is generally a positive correlation between electricity consumption and economic activities, as both residential and commercial sectors require electricity to function. Higher electricity consumption typically signifies increased industrial production, business operations, and consumer demand in their respective sectors. Even though the unit electricity input needed by sectors differ due to the nature of production process, the scale of production is still endogenously proportional to the electricity consumption of the sector itself, thus revealing an input-output relation different from that measured in monetary term. In addition, electricity consumption data is often available on a real-time or near-real-time basis, allowing for a quicker analysis of economic trends compared to traditional indicators like GDP, which are usually released quarterly or annually.

However, there are several points to be considered when adopting electricity data as a proxy for economic activities. For instance, energy efficiency improvements, driven by technological advancements such as mass deployment of renewable energy sources like heat pumps and solar panels, can result in decreased electricity consumption even when economic activities are increasing, leading to an underestimation of economic growth. Hence, like inflation and deflation when measuring economic activities in GDP, technical factors should preferably be addressed on different sectors if the accounting is to be improved. Furthermore, non-electricity-based activities and informal economy, such as agriculture and small-scale industries, may not be accurately captured by electricity consumption data, even though they can constitute a significant part of some regions' economies. Hence, an important assumption made in this research is that the medium and large businesses' activities captured by the electricity consumption data is sufficient to cover most economic activities. Based on the rationales provided, the concept of using electricity consumption data as economic proxy can also extend to other mega cities and regions where the grid possesses the technical capacity to register businesses in different sectors.

To verify the assumption that electricity data can reflect economic clustering, this research designed a simple algorithm to look at the extent of sector synchronization and thus supply chain integration. By finding the correlations between every two sectors, a set of correlations ranging from -1 to 1 is obtained. Subtracting the correlations from 1, a set of values from 0 to 2 is obtained, where a larger value means less correlation and vice versa. Using the value as a distance and each sector as a node, a simple plot is created and shown in Figure 13 to see if certain relationships existed. Sectors that belong to the same categories are painted in the same colour. The red circles indicate clusters formed based on analysis of the sectors' labels. It is obvious to see that sectors of the same categories are located closer to each other to form industrial clusters, proving that these sectors are more correlated to each other and thus more concurrent in response to economic demand changes. Hence, it is reasonable to aggregate the sectors into larger sectors.

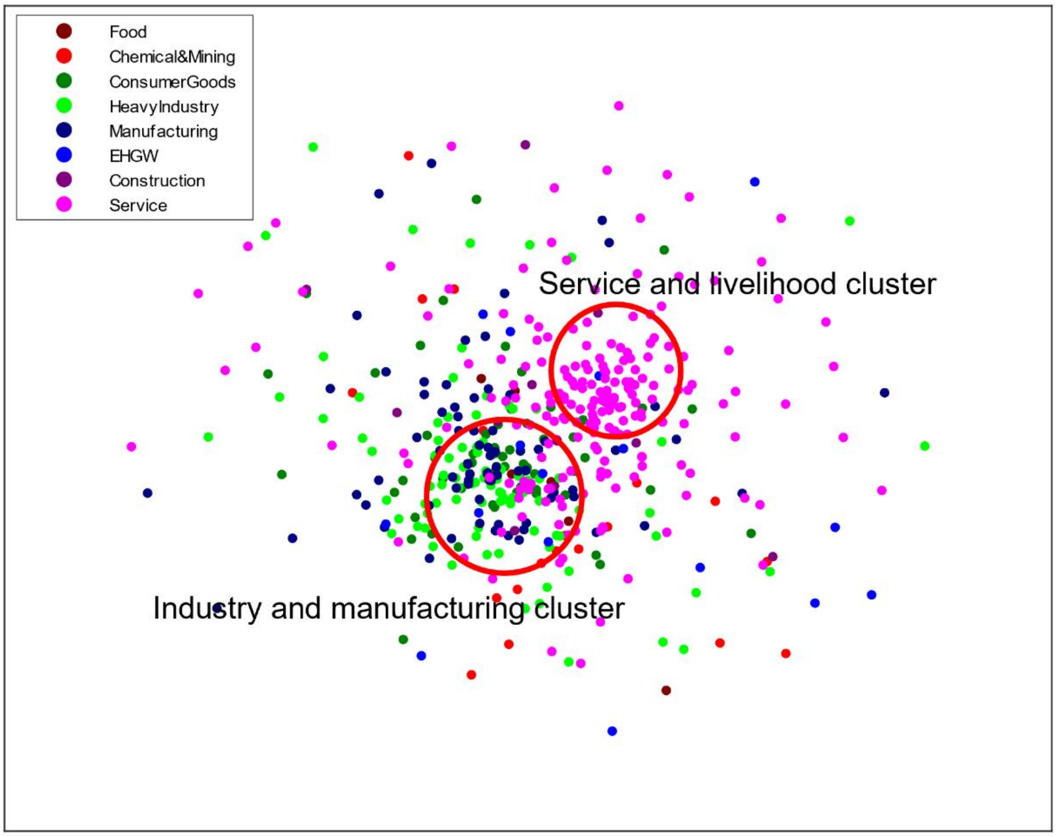



*Figure 13 An illustration of sector clustering using electricity consumption data as evidence.*

Based on the GB/T 4754-2017 specifications and interpretations, the 440 sectors are organized into 8 sectors, namely, food, chemical & mining, consumer goods, heavy industry, manufacturing, EHGW, construction, and service. The specifications are also decoded to decide if the sectors are more associated with final or intermediate consumption. In this setting, the total production was calculated as final plus intermediate production, similar as the construction of IO tables. The organized data are presented in Figure 14, where cyclical patterns of electricity consumption can be clearly observed. For instance, a drastic and persistent decrease around February 2020 in the total consumption of the heavy industry and manufacturing industry can be identified, consistent with the lockdown measures imposed in Chongqing due to the COVID-19 outbreak. For a sample of detailed specification on sector aggregation, please refer to the Appendix Table S3.

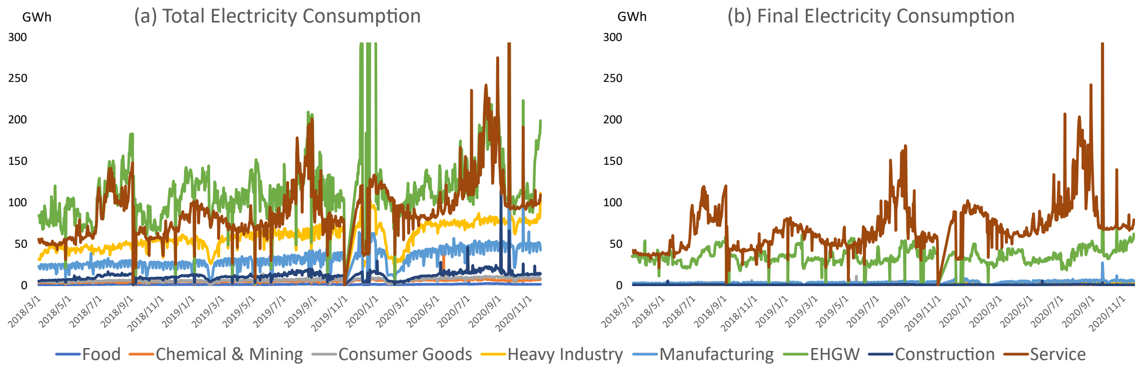



*Figure 14 Plots of chronological electricity consumption in eight sectors.*

## 4.3. Methodological Algorithm

In the case of our research, $x_{(t)}$ and $y_{(t)}$ are vector sets with 8 elements that correspond to the 8 sectors investigated. $A_{(t)}$ is a set of 8-by-8 matrices that describes how inputs from the 8 sectors induce output in the 8 sectors themselves, an identical concept to the IO model (Leontief, 1953). To better illustrate the concept of SIM in equation ( 3.12 ), we produced Table 5 to show how a single final demand at $t = 0$ produces ripple impacts in the future. Referring to both equation ( 3.12 ) and Table 5 we can hence interpret the outputs of the economic system as described by SIM. On the first day where $t = 0$, purchases are made by consumers to fulfil their demand. Producers thus supply the consumers with products from their inventory stock. The purchase thus sends out a market signal to the economy to initiate some intermediate productions at $t = 1$, given as $A_{(1)} \cdot x_{(0)}$. On the next day where $t = 2$, market signal from $t = 0$ propagates to the second layer of production coefficient to give a term $A_{(2)} \cdot x_{(0)}$. At the same time, output or intermediate purchases that happened at $t = 1$ also signal some other products to be produced in the first layer of production coefficient to give a term $A_{(1)} \cdot x_{(1)}$. Hence, the total output level at $t = 2$ is thus given as $x_{(2)} = A_{(2)} \cdot x_{(0)} + A_{(1)} \cdot x_{(1)}$. The Sankey diagram in Figure 16 is drawn to show the broken-down impacts across time and sectors if 1 unit of final demand in the heavy industry occurs at $t = 1$ based on Table 5's illustration. Each column shows the composition of output x at time t from all past time discrete. Each row shows induced outputs of output x at time t in the future.

| | | | | | |
|---|---|---|---|---|---|
| Layer no. | $x_{(0)}$<br>$= y_{(0)}$ | $x_{(1)}$<br>$= A_{(1)} \cdot x_{(0)}$ | $x_{(2)}$<br>$= A_{(2)} \cdot x_{(0)}$<br>$+A_{(1)} \cdot x_{(1)}$ | $x_{(2)}$<br>$= A_{(3)} \cdot x_{(0)}$<br>$+A_{(2)} \cdot x_{(1)}$<br>$+A_{(1)} \cdot x_{(2)}$ | …… |

| | | | | | |
|---|---|---|---|---|---|
| Induced output of $y_{(0)}$ | $I \cdot y_{(0)}$ | | | | |
| Induced output of $x_{(0)}$ | | $A_{(1)} \cdot x_{(0)}$ | $A_{(2)} \cdot x_{(0)}$ | $A_{(3)} \cdot x_{(0)}$ | |
| Induced output of $x_{(1)}$ | | | $A_{(1)} \cdot x_{(1)}$ | $A_{(2)} \cdot x_{(1)}$ | |
| Induced output of $x_{(2)}$ | | | | $A_{(1)} \cdot x_{(2)}$ | |
| …… | | | | | |

*Table 5 Illustration of SIM interactions.*

As explained by He et al. (2022), since $x_{(t)}$ can be expressed in terms of $A_{(t)}$ and $y_{(t)}$, equation ( 3.12 ) can be linearized into equation ( 4.1 )

$$\begin{bmatrix} I & A_{(1)} & \left({A_{(1)}}^2 + A_{(2)}\right) & \left({A_{(1)}}^3 + A_{(1)}A_{(2)} + A_{(2)}A_{(1)} + A_{(3)}\right) & \cdots \end{bmatrix} \times \begin{bmatrix} y_{(n)} & y_{(n+1)} & \cdots \\ y_{(n-1)} & y_{(n)} & \cdots \\ \vdots & \vdots & \cdots \\ y_{(0)} & y_{(1)} & \cdots \end{bmatrix} = \begin{bmatrix} x_{(n)} & x_{(n+1)} & \cdots \end{bmatrix} \quad (4.1)$$

We let the variables in equation ( 4.1 ) become:

$$B = \begin{bmatrix} I & A_{(1)} & \left({A_{(1)}}^2 + A_{(2)}\right) & \left({A_{(1)}}^3 + A_{(1)}A_{(2)} + A_{(2)}A_{(1)} + A_{(3)}\right) & \cdots \end{bmatrix}$$

$$U = \begin{bmatrix} y_{(n)} & y_{(n+1)} & \cdots \\ y_{(n-1)} & y_{(n)} & \cdots \\ \vdots & \vdots & \cdots \\ y_{(0)} & y_{(1)} & \cdots \end{bmatrix}$$

$$V = \begin{bmatrix} x_{(n)} & x_{(n+1)} & \cdots \end{bmatrix}$$

So equation ( 4.1 ) becomes:

$$BU = V \quad (4.2)$$

Since $U$ consists of $y_{(t)}$ and $V$ consists of $x_{(t)}$, the observations from the electricity consumption data can be reorganized. We can use a least square error regression algorithm with constraints to change $B$ to minimize the error of $(BU - V)^2$. It should be noted that the elements of the $A$ matrix should also be in the range of 0 to 1 since the electricity inputs from other sectors to produce 1 kWh output of a certain sector cannot exit 1 kWh. For obvious reason, the electricity inputs of a sector can neither be negative. Hence, elements $b_{ij}$ of the $B$ matrix, the products of elements of $A$ matrix, is set to be in the range of 0 to 1. It is because This is expressed as follows:

| | |
|---|---|
| Find the | $B$ |
| That minimizes | $(BU - V)^2$ |
| Subject to | $0 < b_{ij} < 1$ where $b_{ij}$ is any element of $B$ |

Thus, the best fitted $B$ as the sum of $A_{(t)}$ matrices' power terms can be calculated. The recursive algorithm shown below in equation ( 3.15 ) is then used to unwarp each term in $B$ and to obtain $A_{(t)}$. Some terms in $A_{(t)}$ may be smaller than 0 due to error propagated from regression. They are treated as errors and omitted in the analysis and the Sankey illustration in Figure 16.

In addition, since there is no knowledge of the optimum values of production layers and induction layers, the values of $l$ and $n$ are varied to minimize the total error $(BU - V)^2$. The total absolute error is minimal at $l = 8$ and $n = 12$. For $n > 12$, our computer ran out of calculation memory. If hardware could support, it might be possible to attempt higher power terms $n$ to further minimize regression errors.

It should be noted that the input-output relationship revealed here using electricity consumption is not necessarily proportional to monetary input from this industry for the output sector as described by the production coefficient $A$ in conventional Input-Output Model. For example, considering the case that making one bicycle requires 1 kilogram of metal and 1 kilogram of rubber, if the prices are 2 dollar/kg for metal and 1 dollar/kg for rubber, then the input output relation in the bicycle industry would be 2:1 in monetary term. At the same time, if the electricity input is 5 kWh/kg for metal and 1 kWh/kg for rubber, then the input output relation in the bicycle industry would be 5:1 in electricity term. Table 6 quantitatively compares the difference between the production coefficients of the 8 sectors in Chongqing obtained both from this research using the SIM algorithm developed and from the 2017 Input-Output Table of Chongqing. Although apparent differences can be seen in the results from the two methods, some identical key facts can be easily spotted. For instance, the intermediate output from Heavy Industry (the row of Heavy Industry) is significantly larger than other sectors in both results. Besides, the contribution from EHGW sector is significantly higher than monetary IO table, possibly as a result of the adoption of electricity consumption as the indicator for all production activities.

| Unit: kWh | Food | Chemical & Mining | Consumer Goods | Heavy Industry | Manufact uring | EHGW | Constructi on | Service |
|---|---|---|---|---|---|---|---|---|
| Food | 0.07 | 0.07 | 0.00 | 0.00 | 0.02 | 0.00 | 0.06 | 0.00 |
| Chemical&Mining | 0.08 | 0.02 | 0.00 | 0.04 | 0.02 | 0.01 | 0.08 | 0.00 |
| Consumer Goods | 0.07 | 0.04 | 0.03 | 0.02 | 0.01 | 0.01 | 0.07 | 0.00 |
| Heavy Industry | 0.17 | 0.30 | 0.48 | 0.71 | 0.56 | 0.00 | 0.01 | 0.00 |
| Manufacturing | 0.00 | 0.00 | 0.25 | 0.15 | 0.39 | 0.00 | 0.05 | 0.00 |
| EHGW | 0.00 | 0.82 | 0.00 | 0.58 | 0.00 | 0.16 | 0.11 | 0.00 |
| Construction | 0.07 | 0.04 | 0.11 | 0.07 | 0.00 | 0.01 | 0.09 | 0.01 |
| Service | 0.10 | 0.05 | 0.16 | 0.11 | 0.01 | 0.02 | 0.05 | 0.04 |

(a) Production coefficient of Chongqing obtained in this research from electricity consumption data

| Unit: Yuan | Food | Chemical &Mining | Consume r Goods | Heavy Industry | Manufact uring | EHGW | Constructi on | Service |
|---|---|---|---|---|---|---|---|---|
| Food | 0.06 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 |
| Chemical&Mining | 0.00 | 0.01 | 0.01 | 0.01 | 0.00 | 0.12 | 0.05 | 0.00 |
| Consumer Goods | 0.05 | 0.00 | 0.19 | 0.00 | 0.02 | 0.00 | 0.01 | 0.02 |
| Heavy Industry | 0.05 | 0.00 | 0.03 | 0.06 | 0.12 | 0.01 | 0.44 | 0.03 |
| Manufacturing | 0.01 | 0.00 | 0.01 | 0.00 | 0.31 | 0.01 | 0.06 | 0.02 |
| EHGW | 0.01 | 0.00 | 0.01 | 0.01 | 0.01 | 0.18 | 0.01 | 0.03 |
| Construction | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 |
| Service | 0.09 | 0.01 | 0.08 | 0.02 | 0.08 | 0.07 | 0.15 | 0.18 |

(b) Production coefficient of Chongqing obtained from its 2017 Input-Output Table

*Table 6 The input-output production coefficient obtained (a) from the electricity consumption data in this research and (b) from the 2017 Input-Output Table of Chongqing*

## 4.4. Analysis of Lagged Economic Structure

### 4.4.1. Robustness Testing

500 days of the 971 total days observed in the data set were sampled to conduct model training. By comparing the outputs simulated with the trained model and the remaining 471 observations of the total outputs of the eight sectors, Figure 15 is produced, showing the errors of both the regression trainings and predictions in percentages. To the left of the red dotted lines are the errors of the regressed model based on 500 historical observations. To the right of the red dotted lines are the errors of the predicted outcomes compared with actual remaining 471 observations. No significant change in error levels occurs after the training-prediction boundary, suggesting that the model is generalizing well to unseen data and accurately capturing the underlying relationship between the input variables and the target variable. In technical terms, overfitting occurs when the model learns the noise in the training data as patterns, leading to high accuracy on the training set but poor performance on the test set. Underfitting, on the other hand, happens when the model fails to capture the underlying patterns in the data, resulting in poor performance on both the training and test sets. The result suggests that the regressed model is neither overfitting nor underfitting, indicating the model is likely to perform well on new and unseen data. Some spikes can be seen across all eight sectors, which correspond exactly to abnormal spikes in the actual data, as shown in Figure 14, suggesting that the model is able to filter out outliers in actual observations. This reinforces the robustness of this model and analysis.

In addition, the error level is contained within -30% to 30% for most sectors, a tolerable value in comparison with other electricity output forecasting models (Ahmad et al., 2020). Among all eight sectors, the service sector has the lowest error level, while the food sector has the highest error level. A possible reason for the high error level of the food sector may be its significantly lower level of electricity consumption compared to other sectors. Throughout the 3-year observation period, the daily electricity consumption of the food sector never surpassed 0.5% of the daily total electricity consumption of eight sectors. An unproportionally higher value of electricity consumption means that a higher level of noise more likely corrupts the information from the food sector. As proposed in signal process engineering, more advanced engineering may be needed to filter out the noise and improve the pattern recognition (Tuzlukov, 2018). In contrast, the sudden shock of electricity demand plunge created by Covid-19 in early 2020 is also recorded by this simulation. As the pandemic containment measures such as lockdown do not create capacity change, the production interlinkages are assumed to remain constant and used for future predictions. Instead of a source of error, the Covid-19 demand shock serves as a good sample to boost the model's resilience.

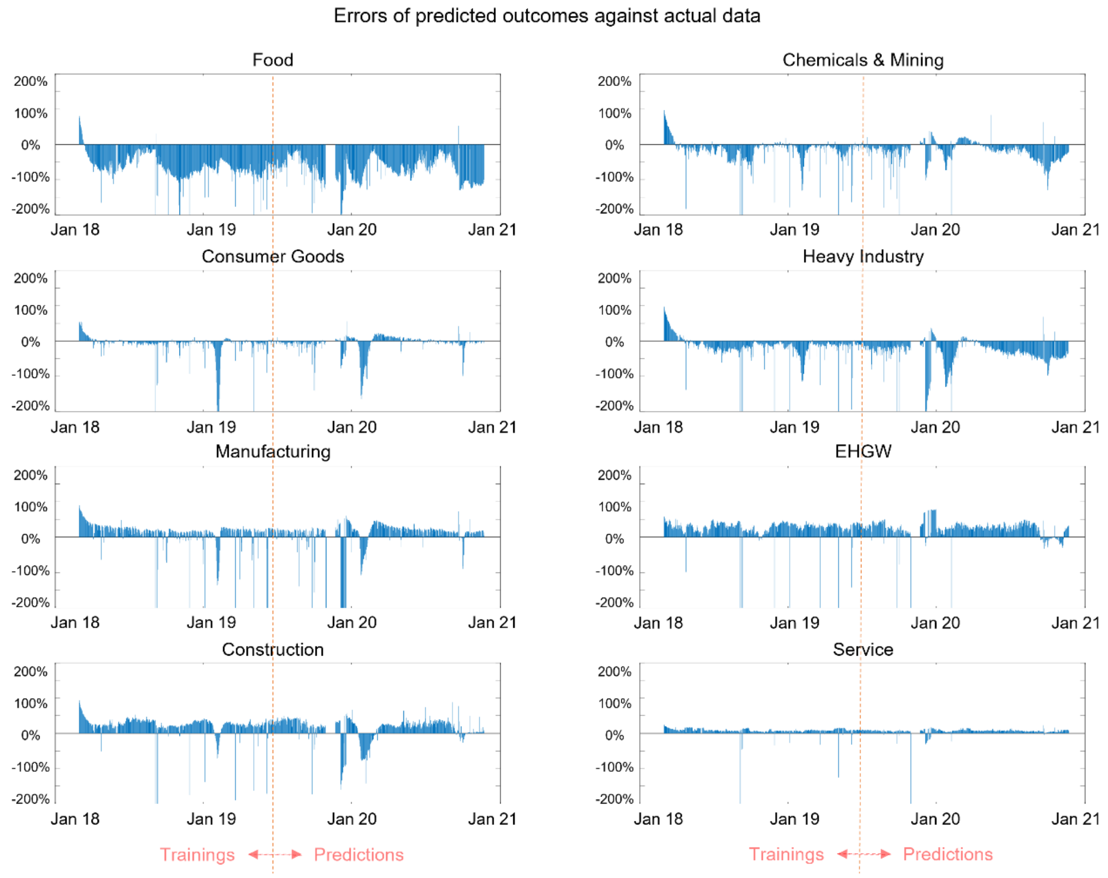



*Figure 15 The eight diagrams show the differences between the simulated outputs and actual outputs of the eight sectors in percentages.*

### 4.4.2. Chronological Interlinkages of Sectors

The SIM models how the production today will induce further production in the future. From the electricity consumption data of Chongqing, this research manages to reveal the chronological interaction of eight sectors of Chongqing. To better depict the underlying concept, a Sankey diagram is plotted to quantitatively illustrate the time lags in the induced productions from one unit demand (1 kWh in this case study) in the heavy industry sector, shown in Figure 16. A detailed explanation of the derivation and construction of the Sankey diagram can be found in Figure 16 and the corresponding paragraphs. From left to right, each of the 8 columns in varied colours represents one calendar day as a production layer. The grey transparent bands connect the intermediate outputs and later outputs induced by them, with their widths proportional to the induced quantities. In total, 7.73 units of output are generated throughout the eight sectors as a response to 1 unit of demand in the heavy industry sector. This magnifying effect of demand can be inferred as the multiplier for the heavy industry sector. The induced quantities decrease as the production layers extend into the future, in accordance with the logic that the multiplier impacts of the demand signal decay over time. In addition, it can be easily seen that heavy industry induced the largest production in the EHGW sector throughout all 8 chronological layers (2.13 units), which further induced a large proportion of outputs from the heavy industry sector (1.68 units), illustrating the close connection between EHGW and heavy industry. In comparison, the food sector is the least associated with the heavy industry sector, with 0.08 units of output induced by the 1 unit of demand from the heavy industry sector. The impact of the food sector output in layer 1 does not extend far into the future, with only 0.01 units of output induced in layer 2 of heavy industry as the largest of all other induced outputs. Furthermore, it can be observed that the consumer goods sector induced outputs into the more distant future, a feature that is not seen in other sectors. The 0.03 output units of the consumer goods sector in layer 1 induced 0.01 output units in layer 8 of the heavy industry sector, equivalent to the induced output in layer 2 of the heavy industry sector. This may be a result of the longer logistic chain and thus longer duration of demand signal propagation in the consumer goods sector, an interesting takeaway of the analysis.

The size of the vertical bars in Figure 16 also conveys valuable insights derived from the trained model. In the first layer, a total output of 1.67 kWh was produced. However, in the second layer, the output decreased to a mere 0.76 kWh. Interestingly, none of the layers following the first layer exhibit sizes larger than that of the second layer. This observation suggests that the required electricity outputs by 1kWh demand in heavy industry decrease as time progresses into the future, a phenomenon that is consistent with the physical performance of the economic system.

Since the information from the Sankey diagram is too abundant to be fully analysed here, the data used to construct Figure 16 is aggregated into days and attached in Appendix Table S4. Each column of bars in different colour codes indicates the scale of electricity consumption and hence economic outputs in the respective chorological production layer (e.g., Service_L_3 means the chorological production layer 3 of the service sector). The bands connect the intermediate outputs and later outputs induced by them. The scale of the bands and columns are proportional to the scale of electricity consumption/economic outputs induced. However, it should be noted that the chronological lag here between sectors should not be considered equivalent to the transportation time needed from one sector to another. It is because the time needed for one sector to respond to the demand change in another may be a result of interaction by multiple factors, including nature of industries, size of business, and availability of effective communication channel etc. Depending on the context, transportation may not play its role in determining the time delay as business can react to market signal with their inventories to conduct productions. On the other hand, service sector, for instance, may naturally adjust its production level much quicker in response to market signals as they rely less on inventory build-up.

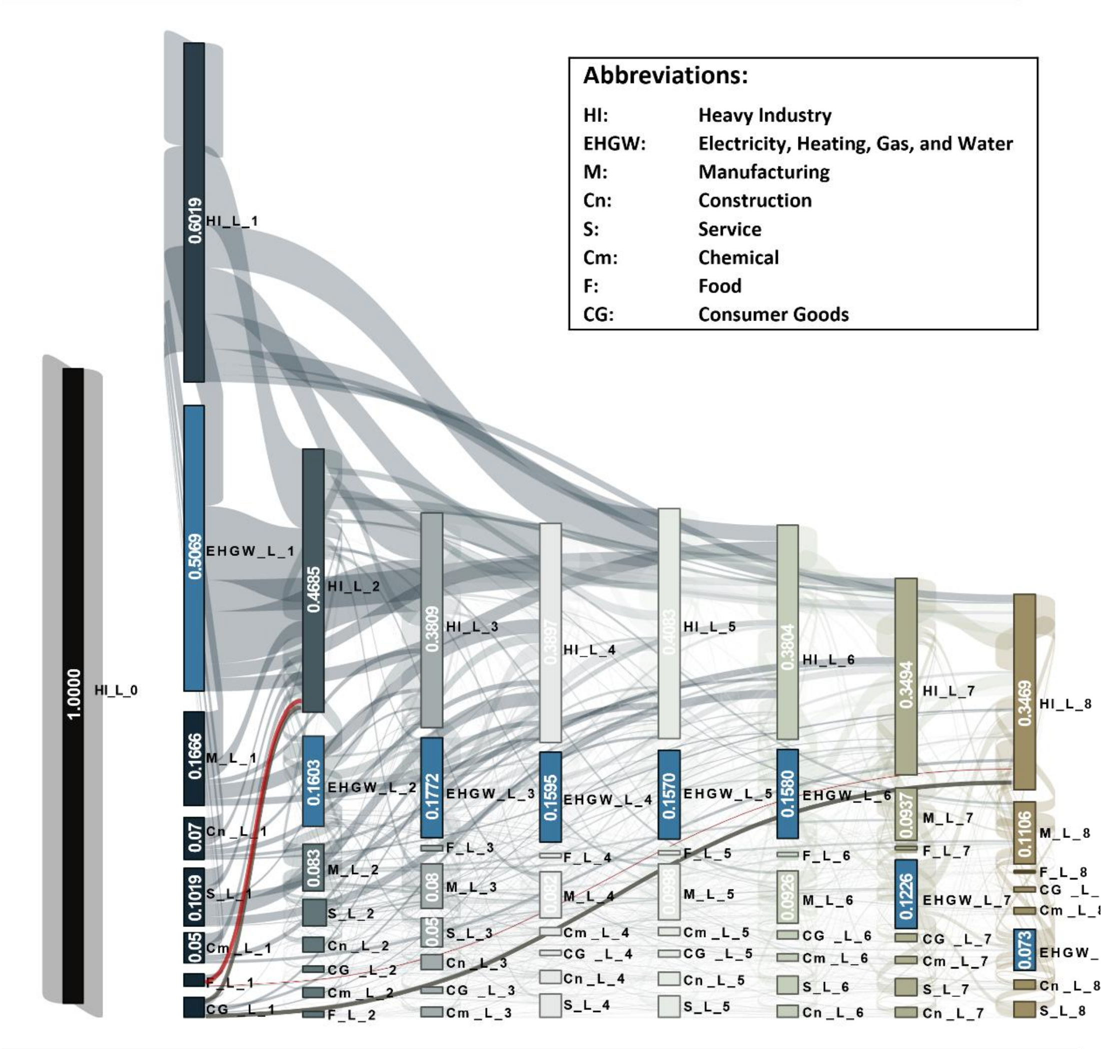



*Figure 16 Sankey diagram showing the chain of responses to one unit of demand in heavy industry in all eight sectors.*

## 4.5. Predictions under scenarios

### 4.5.1. Rationales

In the analysis with the electricity consumption data of Chongqing, only the first 500 samples of the 971 data points are used in the data set for model training to obtain $A_{(t)}$ as described in equations ( 3.15 ) and ( 3.16 ). Substituting $A_{(t)}$ into equation ( 3.12 ), the robustness of the model is further investigated and the chronological linkages of productions are thus analysed. Based on the mean and standard deviation of the growth rate of the 8 sectors in the data set, this research also simulated three scenarios of changed growth rate in the consumer goods final demand two months into the future. As described in equation ( 4.3 ), for the investigated sector:

$$\bar{\mu} = \left( \sqrt[(m-1)]{\frac{y_{(n)}}{y_{(1)}}} - 1 \right) \qquad (4.3)$$

where $\bar{\mu}$ is the daily average growth rate. $m$ is the number of samples observed, 971 in this case. Finding the value located at the 2.5 percentile over and below $\bar{\mu}$, the higher and lower daily growth rates, $\mu_h$ and $\mu_l$, of the observed samples are estimated, which were used for forecasting scenarios. Using 1000 simulations of normally distributed daily growth rates as described below in equation ( 4.4 ), a Monte Carlo simulation for future total consumption predictions across the entire 8 sectors is created.

$$\begin{gathered} y_{(t)} = y_{(t-1)}\,(1+g)\ \ where\ g \sim N(\mu, \sigma^2) \\ x_{(t)} = f\left(y_{(t)}, A_{(t)}\right) \end{gathered} \qquad (4.4)$$

in equation ( 4.4 ), $g$ represents the growth rates for all three scenarios. $\sigma^2$ is the variance of the observed samples in $y$.

### 4.5.2. Predictions Results

The chronological coefficients obtained, as discussed in the previous section, can be used for predictions of the multiplier effects on all sectors given forecasts on the final demand changes in one sector. The daily growth rate of consumer goods final demand is varied, and forecasts are made for the daily outputs of all eight sectors for the coming 80 days. The reason that 80 days is chosen as the prediction boundary is to avoid systematic uncertainties that may not be captured by this modelling. For instance, this model cannot capture the impact of sudden change to capital equipment, such as disaster events, as the regression is based only on observations in the past.

Figure 17 effectively illustrates the prediction results for various scenarios. The daily growth rates of sectors are estimated based on historical means and variances, but the mean growth rate of the consumer goods sector is varied to simulate the changes across all sectors under three scenarios. To the right of the red dotted lines, the coloured areas show the error ranges, while the black solid lines show the predicted mean outputs. In each of these scenarios, the historical observations are located on the left of the distinct red dotted lines, while the predictions extending 80 days into the future are displayed to the right of these same red dotted lines. As part of the scenario setting process, we have altered the forecast of daily growth of final consumption in the consumer goods sector, ranging from an increase of 0.41% to a decrease of -0.40% for the surging and plunging scenarios respectively. These adjustments are based on the careful calculation of past data variances. For the baseline scenarios, the mean growth rates of final demands for all sectors are derived by meticulously analysing historical calculations. Upon reaching the end of the 80-day prediction period, the calculated model predicts a difference of 3.5 GWh/day in the mean value of total electricity consumption within the consumer goods sector. Intriguingly, while the final consumption levels in the surging and plunging scenarios remain unchanged, the heavy industry (with a difference of 18.8 GWh/day), manufacturing (with a difference of 6.8 GWh/day), and EHGW (with a difference of 11.7 GWh/day) sectors exhibit more significant differences in the mean value of their forecasted total electricity consumption on the 80th day as compared to the consumer goods sector. From this observation, it can be deduced that any fluctuations in the final consumption of consumer goods tend to have a more pronounced impact on these three sectors.

On the other hand, when comparing the differences in total electricity consumption across various sectors under the plunging and surging scenarios in percentages relative to the baseline scenario, the trained model reveals that the most substantial relative difference in total electricity consumption occurs within the consumer goods sector itself at 43%. Heavy industry and manufacturing both record the second and third largest relative differences under these scenarios, at 21% each. In contrast, the EHGW sector, with a 13% relative difference, becomes the sector with the second smallest relative difference, rather than the third largest absolute difference. The disparity between absolute and relative measurements may be attributed to the varying energy intensities present in different sectors. For instance, the EHGW sector is naturally energy-intensive, while other sectors, such as service sectors, utilize less energy in their production processes. As a result, the absolute measurement may be more appropriate for predicting the electricity consumption required in each respective sector. Conversely, relative measurements can serve as a better indicator of the changes in economic activity levels.

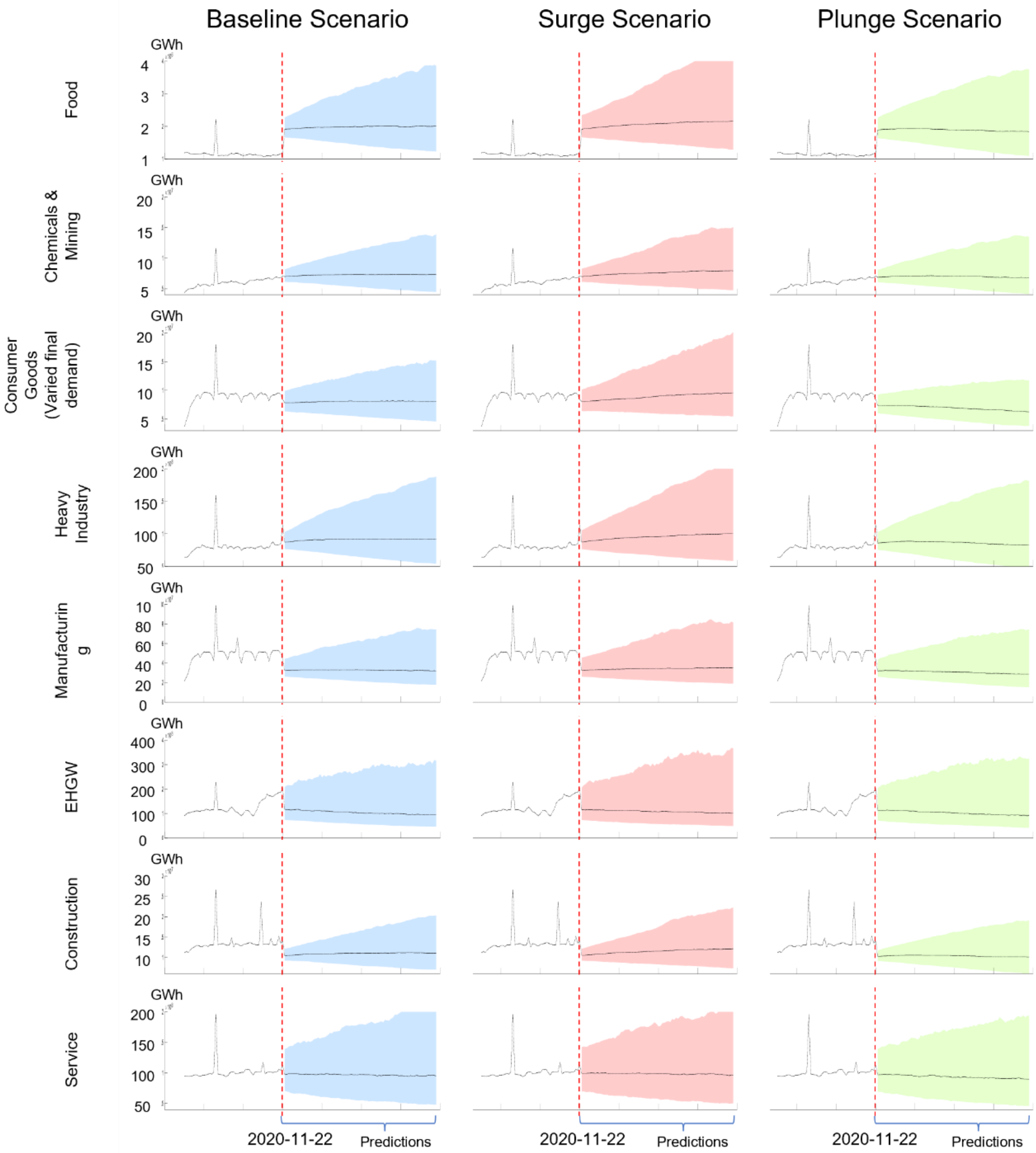



*Figure 17 The simulated outputs/electricity consumption of eight sectors under three scenarios of growth in the consumer goods sector.*

## 4.6. Discussion

In this study, the SIM is used to explore the possibility of reconciling modelling and regression, a long-debated topic in economic research. Since electricity consumption is a good approximation of economic activities, the promising result in this research implies direct application in local economic planning. For instance, the revealed chronological interconnections among sectors can help us better understand the industrial symbiosis among sectors, thus helping to predict the impacts of external shocks on the final demands. In future studies, some more technical means may be possible to reinforce the algorithm.

Specifically, the number of layers for production propagation is set to 8 due to constrain in computational power. Upgrading computer hardware may further improve the accuracy of this modelling. In addition, it may be useful to integrate real-world surveys on supply chains for reference in determining the number of layers. In addition, the classification of the final demand and intermediate output is based on empirical judgement in this research. For instance, due to possible misinterpretations and misinformation provided by consumers, systematic error may be embedded in this research and thus deteriorate model performance. Technical signal processing means, such as noise filtering, may be a way to improve the accuracy of data organization and thereby enhance model performance.

In addition, the SIM is an oversimplified economic model in which the nonlinear effects of variables such as capital constraints and price elasticity are not considered. Although the simple and linear feature of the SIM provides a more efficient estimation for production coefficients in the short run, nonlinearity should be considered for the development of the SIM to incorporate more economic theories in the next stage. Unlike other model frameworks which have factored in the relationships between changes of capital goods and productivity, SIM is not able to well model the impact on capital changes as an endogenized parameter. Later research on SIM methodological improvement can take the momentum of this study to overcome the weakness of SIM by exploring its integration with other IO variant, such as dynamic IO analysis. That being said, it may also be interesting to look at the hybridization of control system engineering and economic modelling. The nonlinear modules in the control system are good ways to model the nonlinear feedback effects of certain sectors in the whole economy. In that regard, the SIM would then evolve into a multi-input multi-output (MIMO) system identification model, but with the black box open to research and analysis on chronological processes.

# Chapter 5: Theory Advancement in SIM – A Dynamic Model

In this chapter, the SIM is enhanced by integrating production capacity into its parameters. A complementary algorithm, akin to the one developed earlier, is applied using this improved model to assess the direct and indirect economic losses resulting from the 2015 South Indian Flood. The chapter concludes with a critical analysis of the study's limitations.

## 5.1. Dynamic Input-Output Model

The original IO model explains how the economic functions under a condition of unchanged structure. However, the economic structure is constantly evolving due to changes in production capacity. While the IO model provides valuable insights into the interdependence of industries and the flow of goods and services within an economy, its limitations make it less suitable for analysing economic structural change due to changes in production capacity. Alternative models, such as CGE models, can provide a more comprehensive representation of the economy and address some of the limitations of the IO model and encompass economic theories from the IO framework.

In consideration of capital stock building and its influence on production capacity, dynamic IO model has been proposed by Leontief himself in his later stage of research (Leontief, 1970, Leontief, 1953). In a dynamic IO model, time steps are introduced into the analysis, allowing for the examination of economic interactions across multiple periods. This enables the model to capture the evolution of the economy and the effects of time-dependent policies or events (Doraszelski and Pakes, 2007). Investment and capital formation, which links current production decisions to future productive capacity, are factored into dynamic IO model to capture the accumulation of capital over time and its impact on economic growth and structural change (Johnson, 1985). Some dynamic IO models also incorporate endogenous technological change to reflect the fact that technological progress can lead to changes in production processes, input requirements, and productivity over time (Gurgul and Lach, 2018). This is important for capturing the effects of innovation and technological advancements on the economy's structure. Furthermore, final demand is often specified to change over time, accounting for shifts in consumer preferences, population growth, and other factors that influence the demand side of the economy. More importantly, dynamic IO models incorporate lagged adjustment processes, reflecting the fact that economic agents may take time to respond to changes in the economy. For example, firms may take time to adjust their production levels or investment decisions in response to changes in demand or other economic conditions. Feedback effects that occur over time, such as the impact of changes in one industry on other industries through supply chain linkages, are also captured in dynamic IO models.

In a dynamic IO model, the relationship between production capacity and investment is generalized into the following form.

$$x_{(t)} = Ax_{(t)} + B\left[x_{(t+1)} - x_{(t)}\right] + y_{(t)} \qquad (5.1)$$

At a certain time step $t$, the output level $x_{(t)}$ becomes the sum of intermediate production $Ax_{(t)}$, final demand $y_{(t)}$, and an additional capital stock compilation term $B\left[x_{(t+1)} - x_{(t)}\right]$. In the capital stock compilation term, the future output level $x_{(t+1)}$ is linked to present state. Difference in future production $x_{(t+1)}$ and present production $x_{(t)}$ is production capacity expansion, which is multiplied with the investment coefficient $B$ to denote the outputs produced at present state for the purpose of capacity expansion. $B$ in a sense also measures the efficiency in investment activities. If the discrete time form is converted into continuous time form, the dynamic IO model will become the following.

$$x = Ax + B\dot{x} + y \qquad (5.2)$$

Although it lacks explicit price mechanisms and market-clearing conditions when compared to CGE models, dynamic IO model still provides a valuable and easier tool for understanding the evolution of economic structures, the interdependence of industries, and the impact of policies and shocks over time. An innovative approach is integrating linear programming techniques to solve for economic optimization questions (Duchin and Szyld, 2006). In applied studies, economists have implemented the dynamic IO model to evaluate longer term economic growth scenarios (Cruz Jr et al., 2009). The role of capital investment and utilization in the production process can also be assessed with dynamic IO model (Blair and Miller, 2022). In more recent application of dynamic IO model, Ryaboshlyk (2006) proposed disaggregating old and new technologies in the time dimension to model technological change in a time span less than a year, better simulating the effects of technological impulses and their diffusion throughout the economy in discrete time steps.

In the assessment of economic shocks, the disaster footprints are also analysed using the variant of Dynamic Inoperability Input-Output Model (DIIM), a combination of the dynamic IO model and inoperability IO model (Barker and Santos, 2010). The DIIM model allows for the analysis of various recovery strategies, such as investments in infrastructure repair, alternative production methods, or temporary imports to replace lost production. By simulating these strategies, the DIIM can be used to assess the effectiveness of different recovery options and inform policy decisions. Based on the interindustry relationships revealed, the DIIM can also be used to analyse the resilience of the economy to disruptions, both in terms of its ability to withstand the initial shock and its capacity to recover over time. Hence, the DIIM is often used to estimate disruptions in disasters like floods, in which case industries are ranked on the basis of inoperability and economic losses to point out critical industries. Policy makers are thus enabled to allocate budget accordingly to critical sectors after disaster events (Yaseen et al., 2020).

## 5.2. Dynamic Sequential Interindustry Model

Since the first proposal of the dynamic IO Model, researchers have put continuous efforts into the methodological advancement. Aulin-Ahmavaara (1990) has made a comprehensive literature review on the methodological advancement of the dynamic IO Model so far. Among all innovations presented in Aulin-Ahmavaara's review, Johansen (1978) is one of the most significant contributions by providing a specific solution to the differential form ( 5.2 ) of the dynamic IO Model under a balanced growth rate scenario. In this case, the increase in demand is exactly matched with the increase in supply. In the further improvement of dynamic IO Model, Ten Raa (1986) has taken one step further to disaggregate production and investment activities into multiple discrete time steps, thus proposing a multiperiod model. Ten Raa's model has been applied and tested in some practical case studies (Fu and Chen, 2009). It is given below in ( 5.3 ):

$$x_{(t)} = \sum_{s=1}^{l} A_{(s)} x_{(t+1-s)} + \sum_{s=1}^{m} B_{(s)} \left[ x_{(t+s)} - x_{(t+s-1)} \right] + y_{(t)} \qquad (5.3)$$

If referring back to the SIM model proposed in Chapter 3, it can be easily noted that Ten Raa's Model approaches the multiperiod concept in a very similar way of SIM, except that a capital building term of $B\dot{x}$ is included. The capital term in Ten Raa's Model, like in the classic dynamic IO Model ( 5.1 ), links production capacity expansion and investment through the investment coefficient $B$. However, it should be noted that it is controversial to use the dynamic model in analysis demolishing or retirement of production capacity. When production decreases in the next time interval, it may be a result of either reduced or idle capacity. If capacity is only idle, investment would not be necessary to expand production in the future so that the dynamic relationship modelled by $B$ would not be valid.

In ( 5.3 ), $B$ is given as a three-dimension matrix like in SIM model, with its third dimension accounts for the discrete time steps. Hence, the impact of investment on production capacity expansion is modelled to take effect with time lags. If the capital building term is to be included in SIM, the new model will look like ( 5.4 ).

$$x_{(t)} = \sum_{s=1}^{l} A_{(s)} x_{(t-s)} + \sum_{s=1}^{m} B_{(s)} \left[ x_{(t+s)} - x_{(t+s-1)} \right] + y_{(t)} \qquad (5.4)$$

Ten Raa's Model is also different from SIM as it assumes production for outputs starts instantly with zero lags in time instead of one discrete time step later. For the consistency with past research and more convenient reference for concerned researchers, Ten Raa's Model is adopted in this chapter and given the name Dynamic SIM (DSIM) model.

## 5.3. Model Algorithm Innovation

Initially, the same recursive algorithm in Chapter 3 was attempted, in hope to obtain both production coefficient $A(t)$ and investment coefficient $B(t)$. However, the recursive algorithm can only work with the case with a single unknown of $A(t)$. If an extra unknown of $B(t)$ is to be included in the DSIM model, an alternative algorithm will be needed to solve for the unknown variables and thus conduct an analysis on the operation of the economic.

Similar to the linearization concept practiced in ( 3.13 ) of Chapter 3, this section is inspired by the "dynamic inverse" proposed by Steenge and Reyes (2020) to develop an linear algebra algorithm to solve for $A(t)$ and $B(t)$ in the DSIM model. In the work of Steenge and Reyes (2020), the capital coefficient $B$ is added to production coefficient $A$, so that the traditional Leontief inverse in IO model is modified to the "dynamic inverse". Following this practice, DSIM model in ( 5.3 ) is reorganized, so that the final demand $y(t)$ is expressed as a function of $A(t)$, $B(t)$, and $x(t)$ in ( 5.5 ).

( 5.5 )

$$y_{(t)} = -\sum_{s=1}^{l-1} A_{(s+1)} x_{(t-s)} + [I - A_{(1)} + B_{(1)}] x_{(t)} + \sum_{s=1}^{m-1} [B_{(s+1)} - B_{(s)}] x_{(t+s)} - B_{(m)} x_{(t+m)}$$

As a linear system, equation ( 5.5 ) can easily be converted into matrix form as follow in equation ( 5.6 ) to show the final demand of the economy at time $t$.

$$y_{(t)} = [-A_{(l)} \quad \cdots \quad -A_{(2)} \quad I - A_{(1)} + B_{(1)} \quad A_{(2)} - B_{(1)} \quad \cdots \quad B_{(m)} - B_{(m-} \times \begin{bmatrix} x_{(t-l+1)} \\ \vdots \\ x_{(t-1)} \\ x_{(t)} \\ x_{(t+1)} \\ \vdots \\ x_{(t+m-1)} \\ x_{(t+m)} \end{bmatrix} \qquad (5.6)$$

Hence, if the linear system is expanded to show the final demand happened in the economy at all $n$ discrete time steps observed, a matrix multiplication equation ( 5.7 ) can be obtained.

$$
\begin{bmatrix}
I-A_{(1)}+B_{(1)} & B_{(2)}-B_{(1)} & \cdots & -B_{(m)} & 0 & & \cdots & & \cdots & & 0 \\
-A_{(2)} & I-A_{(1)}+B_{(1)} & \cdots & B_{(m)}-B_{(m-1)} & -B_{(m)} & & \cdots & & \cdots & & 0 \\
 & & & & & \vdots & & & & & \\
-A_{(l)} & & \cdots & I-A_{(1)}+B_{(1)} & B_{(2)}-B_{(1)} & & \cdots & -B_{(m)} & \cdots & & 0 \\
 & & & & & \vdots & & & & & \\
0 & & \cdots & & & -A_{(l)} & \cdots & I-A_{(1)}+B_{(1)} & \cdots & B_{(m)}-B_{(m-1)} & -B_{(m)} \\
 & & & & & \vdots & & & & & \\
0 & & \cdots & & & & -A_{(l)} & -A_{(l-1)} & \cdots & I-A_{(1)}+B_{(1)} & -B_{(1)} \\
0 & & \cdots & & & & 0 & -A_{(l)} & \cdots & -A_{(2)} & I-A_{(1)}
\end{bmatrix}
\times \begin{bmatrix} x_{(1)} \\ x_{(2)} \\ \vdots \\ x_{(n-1)} \\ x_{(n)} \end{bmatrix} = \begin{bmatrix} y_{(1)} \\ y_{(2)} \\ \vdots \\ y_{(n-1)} \\ y_{(n)} \end{bmatrix} \tag{5.7}
$$

Let

$$
B = \begin{bmatrix}
I - A_{(1)} + B_{(1)} & B_{(2)} - B_{(1)} & \cdots & -B_{(m)} & 0 & & \cdots & & \cdots & & 0 \\
-A_{(2)} & I - A_{(1)} + B_{(1)} & \cdots & B_{(m)} - B_{(m-1)} & -B_{(m)} & & \cdots & & \cdots & & 0 \\
 & & & & & \vdots & & & & & \\
-A_{(l)} & & \cdots & I - A_{(1)} + B_{(1)} & B_{(2)} - B_{(1)} & & \cdots & -B_{(m)} & \cdots & & 0 \\
 & & & & & \vdots & & & & & \\
0 & & \cdots & & & -A_{(l)} & \cdots & I - A_{(1)} + B_{(1)} & \cdots & B_{(m)} - B_{(m-1)} & -B_{(m)} \\
 & & & & & \vdots & & & & & \\
0 & & \cdots & & & & -A_{(l)} & -A_{(l-1)} & \cdots & I - A_{(1)} + B_{(1)} & -B_{(1)} \\
0 & & \cdots & & & & 0 & -A_{(l)} & \cdots & -A_{(2)} & I - A_{(1)}
\end{bmatrix}
$$

$$U = \begin{bmatrix} x_{(1)} \\ x_{(2)} \\ \vdots \\ x_{(n-1)} \\ x_{(n)} \end{bmatrix}$$

$$V = \begin{bmatrix} y_{(1)} \\ y_{(2)} \\ \vdots \\ y_{(n-1)} \\ y_{(n)} \end{bmatrix}$$

So that equation above becomes:

$$BU = V$$

Hence to obtain output $U$, multiplying the inverse of $B$ on both sides we get

$$U = B^{-1}V \qquad (5.8)$$

In equation ( 5.7 ), the linear system of DSIM is transformed into a matrices multiplication process by setting an unchanging vector of $x_{(t)}$ and forming a matrix of varying coefficients $A_{(t)}$and $B_{(t)}$. Similarly, if the matrix of $A(t)$ and $B(t)$ is held as the unchanging vector and elements in the matrix of $x(t)$ are varied to form a matrix, the identical output of final demand $y(t)$ can be obtained as shown in equation ( 5.9 ).

$$\begin{aligned} &[-A_{(l)} \quad \cdots \quad -A_{(2)} \quad I - A_{(1)} + B_{(1)} \quad A_{(2)} - B_{(1)} \quad \cdots \quad B_{(m)} - B_{(m-1)}] \\ &\times \begin{bmatrix} 0 & \cdots & 0 & x_{(1)} & x_{(2)} & \cdots & x_{(l)} \\ & & & \vdots & & & \\ 0 & \cdots & x_{(n-l-3)} & x_{(n-l-2)} & x_{(n-l-1)} & \cdots & x_{(n)} \\ x_{(1)} & \cdots & x_{(n-l-2)} & x_{(n-l-1)} & x_{(n-l)} & \cdots & 0 \\ x_{(2)} & \cdots & x_{(n-l-1)} & x_{(n-l)} & x_{(n-l+1)} & \cdots & 0 \\ & & & \vdots & & & \\ x_{(m-1)} & \cdots & x_{(n-2)} & x_{(n-1)} & x_{(n)} & \cdots & 0 \\ x_{(m)} & \cdots & x_{(n-1)} & x_{(n)} & 0 & \cdots & 0 \end{bmatrix} \\ &= [y_{(1)} \quad \cdots \quad y_{(n-l-1)} \quad y_{(n-l)} \quad y_{(n-l+1)} \quad \cdots \quad y_{(n)}] \end{aligned} \qquad (5.9)$$

Again, let ( 5.9 ) be $BU = V$, so that a unique solution of $y_{(t)}$ can be obtained if $U$ is just symmetric. However, in real world situation, observations for $x_{(t)}$ and $y_{(t)}$ are normally more than required for a unique solution. Thus matrices $U$ would be asymmetric. The inverse of $U$ is hence impossible to be calculated.

Sharing the same idea of equation ( 3.19 ), the Moore Penrose Inverse is again introduced to find the best fitted solution. The Moore Penrose Inverse of $U$ is taken and multiplied on the right-hand side of ( 5.9 ) to calculate matrix $B$. Truncated at $(n-l)$ observations, a similar concept similar to the recursive induction technique as shown in ( 3.15 ) can be applied. By taking the minus of the $m$ terms of the right-side of $B$, we have a series that looks like:

$$A_{(1)} - B_{(1)} - I \quad B_{(1)} - B_{(2)} \quad \cdots \quad B_{(m-1)} - B_{(m)} \quad B_{(m)} \tag{5.10}$$

So that all investment coefficients $B(m)$ are solved by the following recursive steps:

$$\begin{gathered} B_{(m)} = B_{(m)} \\ B_{(m-1)} = \left[B_{(m-1)} - B_{(m)}\right] + B_{(m)} \\ \vdots \\ B_{(1)} = \left[B_{(1)} - B_{(2)}\right] + B_{(2)} \\ A_{(1)} - I = \left[A_{(1)} - B_{(1)} - I\right] + B_{(1)} \end{gathered} \tag{5.11}$$

Taking the minus of the $l$ terms of the left-side of $B$, the following series of elements can be obtained which directly gives the series of $A_{(t)}$.

$$A_{(l)} \quad \cdots \quad A_{(2)} \quad A_{(1)} - B_{(1)} - I \tag{5.12}$$

Finally, the term $A_{(1)}$ can be calculated as follow.

$$A_{(1)} = \left[A_{(1)} - B_{(1)} - I\right] + B_{(1)} + I \tag{5.13}$$

The proposed algorithm makes it feasible to work out the production and investment coefficients through an ample amount of observations on the outputs and final demands in an economy. Knowing these parameters permits the analysis on the operations of the economy investigated and thus make fair predictions for its performance in the short future.

## 5.4. Case of 2015 South India Flood

### 5.4.1. Background

In November 2015, a devastating rainfall has battered the Southern India States of Tamil Nadu and Andhra Pradesh. Most heavy damage was suffered by Chennai. 60% of Chennai was inundated (Johnsy and Schirinzi, 2019). The 2015 South India Flood stands as one of the most devastating natural disaster events to have struck India in recent memory, with its catastrophic consequences arising from an intricate interplay between various natural and manmade factors. On the one hand, the higher-than-normal precipitation levels can be attributed to the warming global climate, which in turn intensified the El Nino effect. This phenomenon facilitated the transportation of moisture from the Bay of Bengal, culminating in extremely heavy precipitation within an alarmingly short timeframe (Dhana Lakshmi and Satyanarayana, 2019). Geographically speaking, the coastal regions of southeast India, including cities like Chennai and its surrounding suburbs, are particularly susceptible to flooding hazards due to their relatively flat terrain. This vulnerability is further exemplified by the 2004 Indian Ocean Tsunami, which caused widespread flooding and resulted in 7,793 direct fatalities in the state of Tamil Nadu alone (Network of Emergency Physicians, 2007). On the other hand, the human element cannot be disregarded in the exacerbation of the flood's deadly impacts. Prior to the heaviest precipitation, residents of Chennai had been warned about the increased risk of flooding. However, the true severity of the impending flood remained largely unknown and unanticipated. Consequently, negligence emerged as a key human factor contributing to the disastrous outcome (Jobin et al., 2018). Moreover, the rapid expansion and development of Chennai's suburbs have also been implicated in the heightened impact of the 2015 flood (Samraj, 2017). The rampant urbanization and lack of proper urban planning led to encroachments on natural drainage systems, compromising the region's ability to effectively manage the excessive rainfall.

As evidenced by various research studies, Tamil Nadu is frequently regarded as one of the most developed regions in India, boasting extensive supply chains that stretch across numerous other regions throughout the country (Huang et al., 2021). Hence, as one of India's most significant economic centres, the extensive damages inflicted upon Chennai by the 2015 flood reverberated throughout the entire nation, causing widespread disruption (Rajan and Sridharan, 2016). In response to the catastrophe, vigorous discussions have been initiated on how to effectively recover from the disaster and mitigate future risks disaster (Mariaselvam and Gopichandran, 2016, Bremner, 2020). The Tamil Nadu State Government has estimated the direct financial loss resulting from the flood to be approximately Rs. 8,481 crore (one crore equals ten million in Indian numbering system). However, quantifying the indirect costs incurred both locally and throughout the rest of India presents a significant challenge. Accurately estimating the indirect losses from such disastrous events has proven to be a difficult task in numerous studies due to the complex nature of the interconnected economic systems and the far-reaching consequences of such calamities. This chapter will endeavour to apply real economic data to the DSIM in order to assess the indirect economic losses incurred over time. This method aims to provide a more comprehensive understanding of the overall impact of the 2015 South India Flood on the Indian economy. By utilizing this model, we can gain valuable insights into the cascading effects of the disaster and better inform policy decisions related to disaster recovery, urban planning, and risk management. Furthermore, the evaluation of indirect losses will shed light on the broader implications of natural disasters on economic activities, employment, and social well-being. By incorporating these findings into future planning and policy-making efforts, Authorities can enhance their resilience to the adverse effects of such disasters and emerge better prepared to face the challenges posed by an increasingly volatile global climate.

Interestingly, the topic of disaster footprint has emerged as a vital area of interest within the field of IO modelling. IO modelers have consistently dedicated their efforts to applying IO models for analysing the indirect costs associated with natural disasters. The most prevalent approach, such as the Adaptive Regional Input-Output (ARIO) model (Hallegatte, 2008), involves assuming a damage ratio that imposes a bottleneck effect on production capacity. This assumption is intended to simulate the production constraints resulting from a disaster event (Mendoza-Tinoco et al., 2020, Zeng and Guan, 2020a). In more sophisticated modelling studies, researchers have integrated the dynamic IO model (Leontief, 1970, Leontief, 1953) with the concept of interoperability (Santos and Haimes, 2004, Haimes and Jiang, 2001). This innovative modelling practice was subsequently formalized and dubbed the Dynamic Inoperability Input-Output Model (DIIM). The applicability of the DIIM has been extended to cover various disastrous events, including terrorism risks (Lian and Haimes, 2006, Los, 2001). By incorporating the IO economic model and its variants into disaster footprint analysis, researchers can achieve a more nuanced understanding of the intricate relationships and dependencies between various sectors of the economy in the aftermath of a catastrophe. This, in turn, enables the development of more effective strategies for mitigating the indirect costs of natural disasters and fostering a more resilient economy. The continuous advancements in IO modelling and the incorporation of dynamic elements in these models demonstrate the ongoing efforts of scholars to better comprehend and quantify the far-reaching impacts of natural disasters on economic systems.

Nonetheless, the aforementioned approaches are not without their significant drawbacks. In the research referenced earlier, there exists an underlying assumption that an equilibrium state is reached at each discrete time step. As we have argued in previous chapters of this thesis, while the equilibrium state assumption might hold true for longer periods, it is unlikely to accurately reflect the dynamics of shorter time spans. Furthermore, given that items are recorded in monetary units, outputs from different sectors are assumed to be interchangeable, with no distinctions made between them. This assumption implies that a decrease in output in one sector will result in an identical decrease in total outputs. In reality, however, a reduced amount of intermediate input in one sector could render inputs from another sector redundant in the production of the final product. This discrepancy between the model's assumptions and the actual complexities of economic systems highlights the limitations of the current modelling approaches.

To address these shortcomings, it is crucial to develop more nuanced and sophisticated models that can accurately capture the complex interdependencies between sectors and account for the varying degrees of substitutability between outputs. By incorporating these refinements into the existing models, we can achieve a more precise understanding of the dynamics at play within an economy in the aftermath of a disaster. By enhancing our modelling capabilities in these ways, we will be better equipped to assess the indirect costs of natural disasters and develop more effective strategies for mitigating their impacts on economies. Ultimately, this increased understanding will inform policy decisions and resilience-building measures, helping societies to better prepare for, adapt to, and recover from the challenges posed by natural disasters.

### 5.4.2. Method and Data

In stark contrast to existing IO modelling tools used for disaster footprint analysis, which typically model impacts based on a hypothetical inoperability ratio, the algorithm proposed in this study adopts a reverse approach. In line with the research presented in Chapters 3 and 4 of this thesis, this Chapter aims to investigate the disaster footprint of the 2015 South India Flood by employing a regression analysis of historical economic data, as explained in the previous section. By calculating the chronological production and investment coefficients, the DSIM constructs a "what-if" scenario that envisions the economic landscape had the disaster not occurred. Consequently, an estimation for local and national outputs unaffected by the shock can be obtained by adding back the direct damage to Tamil Nadu's (TN) final demand. Comparing the differences in total and final outputs between the modelled and actual scenarios will reveal the chronological impacts of the disaster event on the entire economic system of India. The DSIM also addresses the drawbacks of conventional disaster footprint analysis identified earlier. As an enhancement to IO-based disaster footprint analysis tools, DSIM incorporates the efficiency of investment into its modelling, offering an additional dimension for consideration when making policy decisions regarding recovery strategies. Furthermore, since the DSIM does not assume an equilibrium state at each time step like other IO and CGE models do, it circumvents inaccuracies that may arise when modelling disequilibrium settings over short periods. By embracing this innovative approach, the DSIM analysis carried out in this research can provide a more comprehensive and accurate understanding of the true economic consequences of the 2015 South India Flood. This deeper insight will enable policymakers to develop more effective recovery strategies and better allocate resources to mitigate the far-reaching effects of such catastrophes on economies and societies in the future.

As outlined in the previous section, in order to perform a regression analysis on the relationships among regions and sectors within a MRIO model, a series of observations on final demands and total outputs is required. However, obtaining accurate measurements of these two indicators for shorter statistical periods can be challenging due to the high costs associated with data collection and processing. In order to investigate the economic symbiosis among Indian regions and subsequently reveal the impact induced by the disastrous event, an approximation of the final demand and total output for TN and the Rest of India (RoI) is necessary. In this section, we utilize the monthly Index of Industrial Production (IIP) for manufacturing as an estimation of the intermediate production levels, while considering the total IIP as a representation of the total output levels for both TN (Evaluation & Applied Research Department, 2020) and RoI (OECD, 2022) during January 2015 and March 2016. These indices are subsequently employed to disaggregate the corresponding quantities from the 2015 India MRIO table.

Regrettably, with only 15 monthly observations available over a 15-month period, the data points provided are insufficient for an analysis for longer period. To effectively conduct a regression algorithm in a modelling system with n unknowns, it is necessary to have more than n observations on other parameters, such as final demands and total outputs in this case. Consequently, certain compromises must be made, such as aggregating the detailed economic sectors in the 2015 Indian MRIO table compiled by the CEADs database (Huang et al., 2021) into a single sector for both TN and RoI. This simplification results in a 2-by-2 dimensional MRIO matrix. By measuring the economic outputs of TN and RoI in identical monetary units, it can be reasonably assumed that they serve as substitutes for each other. Utilizing the algorithm developed in Equation ( 5.9 ), it becomes feasible to ascertain the production and investment coefficients for TN and RoI. To ensure the validity of the model, constraints and estimations are imposed on the regression process. Given that the time resolution in this study is restricted to a monthly basis, it is assumed that this duration provides economic agents with adequate time to respond to any market signals. Consequently, the layers of production and investment coefficients are set at 2, indicating that the time delay for the Indian economy to react to any market signal concludes in the second month.

Nonetheless, this study faces a comparable challenge to that encountered in Chapter 4 when utilizing the regression algorithm outlined in equation ( 5.9 ). It is evident that the elements of $A_{(t)}$ and $B_{(t)}$ must be constrained to non-negative values and less than one for practical reasons. However, imposing constraints on the objective variable during a regression algorithm using Moore-Penrose Inverse calculations is not feasible. To address this issue, linear programming techniques, combined with some simplifications, are introduced to the case study.

To begin with, our initial step involved making the most accurate estimation of $A_{(t)}$ by computing the static production coefficient $A$ using the formula derived from the classical input-output (IO) model, as illustrated in the equation ( 5.14 ) below:

$$A = Z./X \qquad (5.14)$$

Where $Z$ is the intermediate consumption matrix obtained from the 2015 India MRIO table. $X$ is the total output vector for TN and RoI. In this research, it is assumed that domestic consumptions sitting on the diagonal happens within the first time discrete (month). The traded flow of consumptions across regions occurs one time discrete later. The estimation of production delay follows the practice from many other similar studies (Okuyama et al., 2004). Hence, an initial estimation on $A_{(t)}$ is obtained.

The initial guess of $A_{(t)}$ is then substituted into the DSIM Model ( 5.3 ) to make a first guess before calculating the chronological investment coefficient $B_{(t)}$. Manipulation of equation ( 5.3 ) offers an equation with $B$ on the left to be determined, as shown below in equation ( 5.15 ).

$$B\left[x_{(t+1)} - x_{(t)}\right] = x_{(t)} - \sum_{s=1}^{l} A_{(s)} x_{(t+1-s)} - y_{(t)} \qquad (5.15)$$

The left-hand side of the equation represents the investment in capital stock accumulation, which has a direct linear relationship with the increase in total outputs up to the subsequent discrete time interval. Examining the right-hand side of the equation, we can also compute the expenditure on capital stock by subtracting the intermediate input and final consumption from the total output. To achieve the most accurate estimate of the coefficient $B$, it is essential to minimize the discrepancy between the two sides of the equation. Consequently, an optimization system is established as follows to accomplish this goal.

| | |
|---|---|
| Find the | $B$ |
| That minimizes | $B\left[x_{(t+1)} - x_{(t)}\right] - [x_{(t)} - \sum_{s=1}^{l} A_{(s)} x_{(t+1-s)} - y_{(t)}]$ |
| Subject to | $0.01 < B_{ij} < 1.00$ where $b_{ij}$ is any element of $B$ |

In this optimization framework, the boundaries for $B$ are set between 0 and 1 to represent the principle that investments are consistently positive and efficient. This implies that injecting one unit of capital into the economy will yield an output greater than one unit. Upon obtaining an initial static value for $B$, the same assumption used in $A_{(t)}$ calculation is employed: domestic consumption occurs within the same time interval, while inter-regional consumption takes place one time interval later. Consequently, an adjusted $B_{(t)}$ is derived, serving as the initial estimate of the investment coefficient $B_{(t)}$.

With the knowledge of the production coefficient $A_{(t)}$ and investment coefficient $B_{(t)}$, it becomes possible to simulate the economic symbiosis between TN and RoI in 2015. If an economic shock event impacts the final demand of any region, the model can simulate the resulting changes—both direct and indirect—across regions in a chronological manner. Specifically, the reduced demand caused by the shock event, such as the direct loss from flooding, is reintroduced to the final demand. The total output changes are then computed in reverse and compared with the actual outcomes to estimate the direct and indirect impacts on the economy. In this case study, the official estimation of Rs. 8,500 crores by the TN state government is chosen to represent the direct loss from TN's final demand. This number is selected over other estimates as it is the most authoritative source and best aligns with the trend of TN's monthly growth. Furthermore, to prevent an unrealistic growth rate in this simulation, a cap of 0.6% on the monthly growth rate is imposed by adjusting the 8.0% annual GDP growth rate of India in 2015.

### 5.4.3. Result Analysis

The regressed production coefficient $A_{(t)}$ and investment coefficient $B_{(t)}$ shed light on vital information concerning the chronological interaction between TN and RoI in 2015. Table 7 is subsequently generated to display the $A_{(t)}$ and $B_{(t)}$ values obtained in this study. Regarding $A_{(t)}$, a smaller value signifies that fewer resources and inputs are required to produce a unit output for the target region, indicating a higher efficiency in production. Similarly, a smaller value in $B_{(t)}$ implies that less resources and inputs are needed to increase output by one unit in the subsequent time interval, suggesting a higher efficiency in investment.

| | $A_{(1)}$ | | $A_{(2)}$ | |
|---|---|---|---|---|
| | Tamil Nadu | Rest of India | Tamil Nadu | Rest of India |
| Tamil Nadu | 0.2467 | 0 | 0 | 0.0152 |
| Rest of India | 0 | 0.4459 | 0.2040 | 0 |

| | $B_{(1)}$ | | $B_{(2)}$ | |
|---|---|---|---|---|
| | Tamil Nadu | Rest of India | Tamil Nadu | Rest of India |
| Tamil Nadu | 0.0100 | 0 | 0 | 0.0158 |
| Rest of India | 0 | 0.0911 | 0.5770 | 0 |

*Table 7 Simulated result of A(t) and B(t) for the economic activity interactions of Tamil Nadu Pradesh and Rest of India in the year of 2015.*

In the context of production efficiency, TN's intermediate production exported to RoI (0.0152) represents the most efficient industrial linkage, being 29.3 times more efficient than the domestically produced inputs from RoI itself (0.4459). This occurs despite the exports from TN taking a longer time to reach RoI due to various barriers, such as geographical distance. In contrast, the difference between domestic (0.2467) and imported (0.2040) production efficiencies for TN is relatively small. This reveals TN's role as an economic hub, as its products are more advantageous for RoI compared to those from other regions. A similar comparative advantage is not observed for TN's domestic production over its imports from RoI, further emphasizing the competitiveness of TN's products over imports from RoI in domestic economic activities.

The lower domestic (0.0100) and interregional (0.0158) investment coefficients for TN also demonstrate that its investment efficiency significantly surpasses that of RoI. In contrast, the considerably higher values of domestic (0.0911) and interregional (0.5770) investment coefficients for RoI indicate a clear comparative disadvantage when it comes to investing and stimulating output growth. The model developed in this study estimates that investments from TN to RoI are 36.5 times more efficient, in quantitative terms, than those from RoI to TN. This further reinforces the conclusion that TN serves as India's economic hub, playing a pivotal role in driving economic growth across all regions. The implications of this finding are far-reaching, as it highlights the importance of nurturing and supporting the development of such economic hubs to ensure sustained and balanced growth throughout the nation.

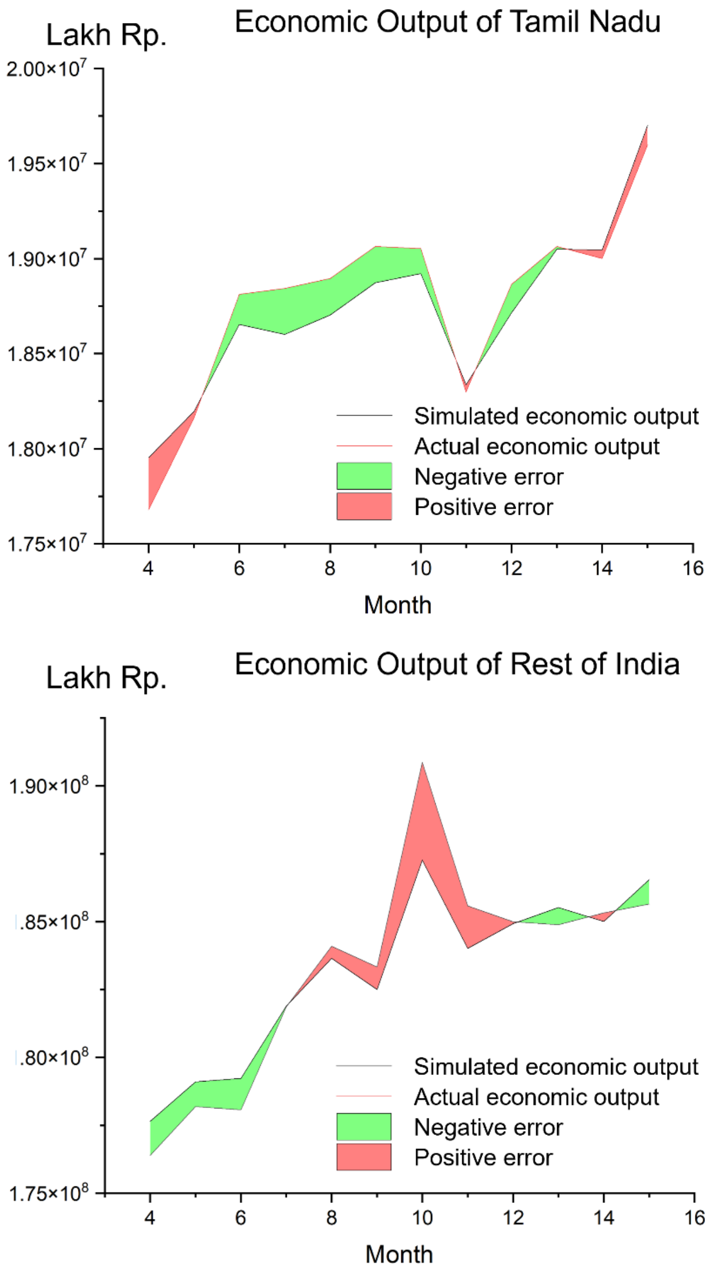



*Figure 18 Sensitivity analysis for output simulations of TN and RoI.*

In order to assess the accuracy of the developed model, a comparison is made between the simulated outputs for both TN and RoI and the actual outputs observed during the period from April 2015 to April 2016. This comparison is visually presented in Figure 18. Area shaded in red means that the simulated value for output is larger than the actual sample. Area shaded in green means that the simulated value for output is smaller than the actual sample. The number on the horizontal axis means number of months from January 2015, e.g. 6 indicates June 2015. One lakh is one hundred thousand in Indian numbering system. The model's production and investment coefficients are configured to have extended influences over a period of two discrete time steps, or months. The time step in which the final demand has a direct impact on the economy has also been taken into consideration. Within this context, the first three data points on the time dimension have not yet been completely affected by the model's mechanisms. Consequently, data points prior to April 2015 are excluded from this sensitivity analysis. The results of the analysis reveal that the percentage errors for TN range from 0.2% to 2.1%, while for RoI, the errors lie between 0.0% and 2.0%. These relatively low error levels serve as evidence that the constructed model provides an accurate representation of the economic symbiosis between TN and RoI. As a result of this accurate portrayal, scenario simulations can be employed more effectively to capture the comprehensive impacts stemming from an economic shock event.

### 5.4.4. Responding to Demand Change

To estimate the total cost of the 2015 South India Flood, an unaffected demand for TN is incorporated into the DSIM model to simulate a business-as-usual (BAU) scenario. The direct damage of Rs. 8,500 crore is added to the reduced final demand of TN in November 2015, assuming that the flood had not occurred. A monthly growth rate of 0.6% is presumed, based on the discounted annual growth rate (Evaluation & Applied Research Department, 2020). Furthermore, an additional constraint is implemented to restrict the growth of TN's final demand under the hypothetical BAU scenario from surpassing its actual level. The primary objective of this approach is to recreate a scenario in which the economy functions normally, without the disruptions caused by unexpected shock events such as the flood.

The economic outputs of both TN and RoI in time series are illustrated in Figure 19, providing a direct comparison of the flood's impact on not only TN but also other parts of the country over an extended period. Error band widths are estimated based on the percentage error obtained in the previous section. The number on the horizontal axis means number of months from January 2015, e.g. 6 indicates June 2015. A noticeable difference is observed in TN's output after November 2015 when comparing the outputs under the BAU scenario to the actual economic output of the flooded scenario (as seen in the left plot of Figure 19). For TN, the actual total economic output nearly aligns with the simulated BAU scenario by April 2016, considering a 2.1% error level based on the sensitivity analysis in the previous section. The discrepancy between the two scenarios can be interpreted as the total direct and indirect costs generated by the flood. From November 2015 until April 2016, the total cost borne by TN gradually decreased from Rs. 11,500 crore per month to Rs. 1,400 crore per month, reflecting the typical characteristics of economic recovery following disastrous events. The cumulative economic output loss for TN during the five months after the flood amounts to Rs. 30,100 crores, which is remarkably consistent with the Rs. 30,000 crore estimation from insurance firms immediately after November 2015 (Vencatesan, 2021). This demonstrates the model's effectiveness in capturing the overall economic impact of the flood.

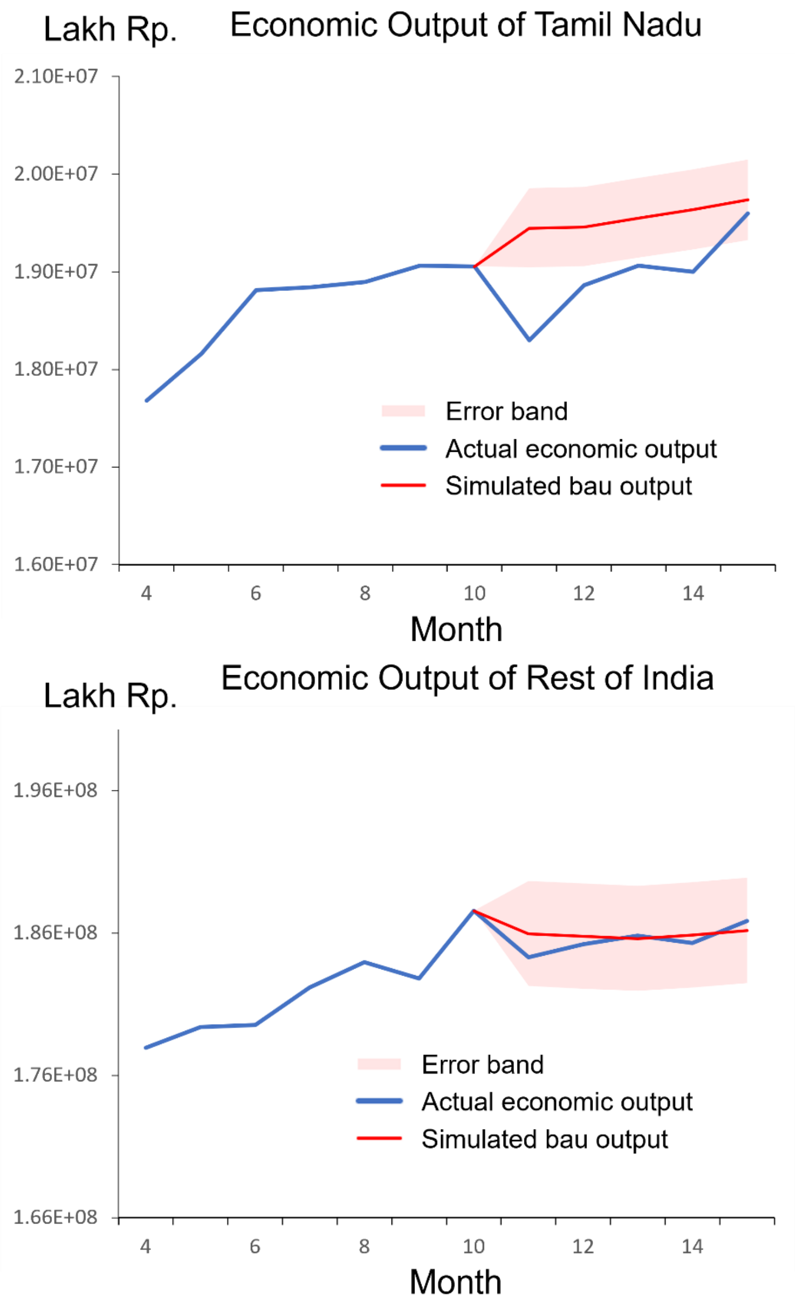



*Figure 19 The actual monthly economic outputs of TN and RoI from April 2015 to April 2016 vs simulated bau scenario using DSIM.*

Moreover, the flood's impact extends beyond local regions. The total economic output loss borne by RoI is estimated to be Rs. 16,400 crores in the month of November 2015 alone by this model, even larger in absolute terms when compared to the loss experienced by TN at the same time. If the accumulated loss from November 2015 to April 2016 is calculated, the total economic loss of

TN (Rs. 30,100 crore) surpasses the total economic loss of RoI (Rs. 18,300 crore), considering simulated outcomes that are lower than actual outputs. It is worth noting that the simulated BAU economic level of RoI converges with the actual level after December 2015, suggesting that although a spillover impact exists for RoI, it lasts for a relatively shorter period than in TN. Additionally, when comparing the sizes of the economies, RoI has undoubtedly suffered less from the 2015 South India Flood. The average loss accounts for only about 1.0% of the total economic output of RoI. In contrast, the model estimates an average loss of 3.2% to the economic output of TN, indicating a significantly higher impact compared to RoI. Most importantly, if the estimated economic loss from RoI is also included in the damage estimation, the total economic loss of the 2015 South India Flood would amount to Rs. 48,400 crores, about 1.6 times of the highest estimation (Rs. 30,000 crore) made by existing studies. This highlights the ability of the model built in this study to carry out comprehensive assessments that consider not only direct losses but also the wider ripple effects that disasters may have on interconnected economies.

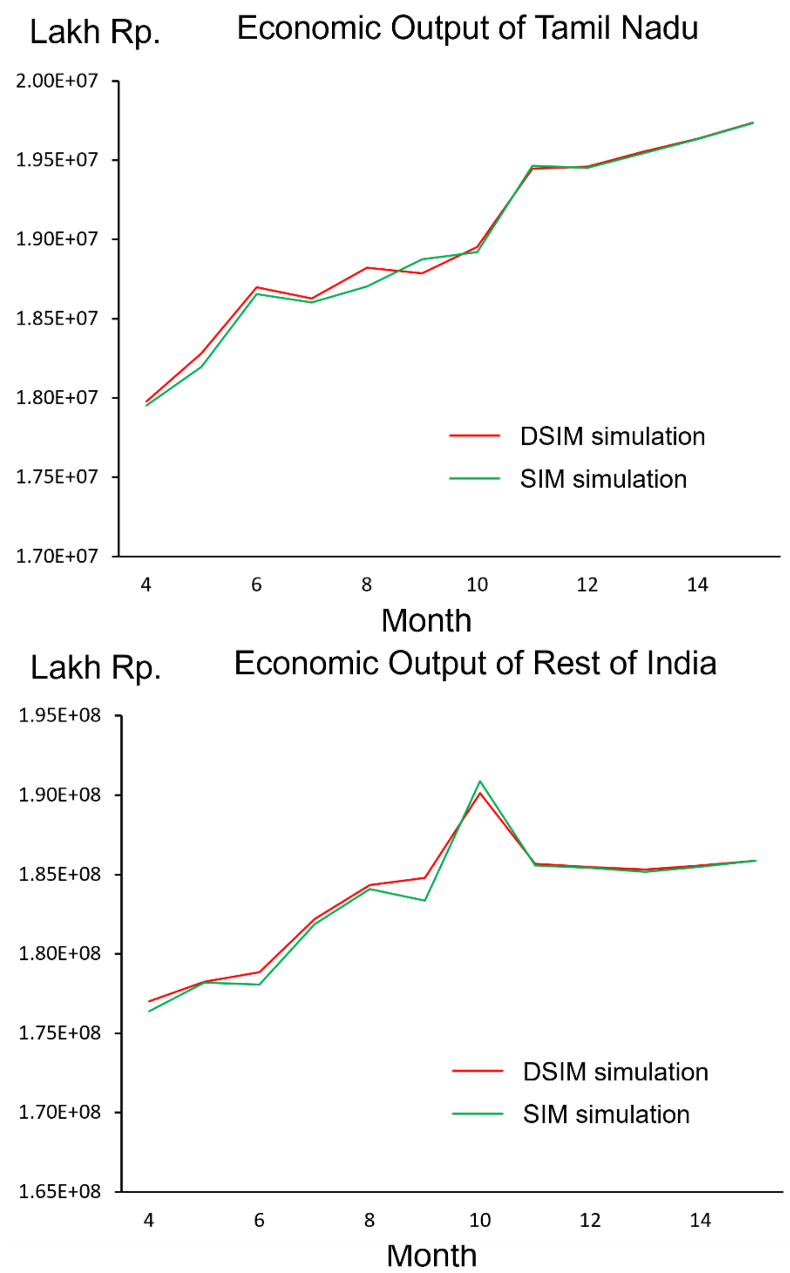



*Figure 20 A comparison of the BAU scenario economic outputs of TN and RoI simulated by DSIM and SIM model.*

In addition to simulation using the DSIM model, the BAU scenario is also simulated using the SIM model to compare the differences between the two models, as illustrated in Figure 20. The DSIM model can be easily converted to the SIM simulation by setting the investment coefficient $B_{(s)}$ in equation ( 5.4 ) to zero. This implies that no portion of the final demand fulfilled is intended to increase economic output capacity in the future. For BAU scenarios simulated by DSIM and SIM models, larger differences in both TN and RoI are observed for the months preceding the flood event in November 2015. This is because the DSIM model has factored in preloaded investment in capacity expansion, leading to a lagged increase in output that is clearly observed in TN before November 2015. For TN outputs occurring after November 2015, since both simulations are responding to an identically reconstructed increase in final demand, differences in simulated outputs are much smaller compared to earlier months. On average, the DSIM output level in TN is 0.14% higher than the SIM simulation, with a standard deviation of 0.25%. Similarly, outputs in RoI respond to the reconstructed demand in TN by preloading investment in capacity expansion into earlier months. A clear flattening trend can be observed when comparing the DSIM and SIM simulations for RoI outputs. On average, the DSIM output level in RoI is 0.18% higher than the SIM simulation, with a standard deviation of 0.34%. These comparisons highlight the ability of DSIM to consider the role of investment in capacity expansion when modelling economic interdependencies s.

All numbers produced for Figure 18, Figure 19, and Figure 20 are included in Table S5 of the Appendix.

## 5.5. Discussion

In this chapter, the algorithm developed for DSIM is applied to real data from India in an effort to assess the total indirect economic cost of the 2015 South India Flood. It is discovered that the economic cost, inclusive of potential economic growth forgone, is significantly larger than any available estimations made to date. Additionally, rather than a short-term impact, the DSIM model simulates a prolonged and spilled-over economic effect on the nation's economy. The results demonstrate that the economic output of TN only catches up with the simulated output level in April 2016, five months after the flood occurred in November 2015. Moreover, the simulation conducted in this chapter uncovers the collateral impacts induced in RoI, which are as large as those in TN when the flood first hit in November 2015. The simulated results in DSIM tend to flatten over the time investigated, illustrating the difference in adaptation from capital stock investment and its consequences on output levels if the shock event was known. Compared to other flood footprint studies, this research expands the accounting scope from direct and indirect economic losses to lost economic output growth. This approach provides a more comprehensive modelling tool for quantifying the economic losses and delay impacts associated with disaster events, shedding light on the full extent of their consequences and informing more effective disaster management strategies.

Despite the insights provided by the simulation algorithm in this study, it is important to acknowledge its limitations. As discussed in Okuyama (2004b), there is a distinction between anticipatory and responsive demands. All simulations conducted in this study assume responsive demands, implying that intermediate production only reacts to demands that have already occurred. In real-world situations, this assumption may not always hold true, as production activity changes can occur prior to actual demand shifts due to market forecasts and speculation. For instance, instead of a shock event like natural disaster, producers may respond to a policy announcement of carbon tax and reduce the production of carbon emission intensive products before the policy is implemented. In this case, anticipatory demand would be a more appropriate simulation for the interactions of economic sectors. Furthermore, the outputs in this simulation are not differentiated by economic sectors, which assumes that all outputs are substitutable for one another. In reality, economic outputs are recorded according to sectors, and outputs from each sector are not easily substitutable. Integrating sectorial information is a necessary compromise due to data source and quality limitations, which may result in oversimplification within the modelling exercise. Despite these limitations, the study still provides valuable insights into the indirect economic costs of disasters and their spill-over effects. Future research could address these limitations by incorporating anticipatory demands and sectorial differentiation into the model, leading to even more accurate and nuanced assessments of the economic impacts of disasters.

# Chapter 6: Conclusions

Ever since the first proposal of the IO model, an ample number of methodological innovations have been attempted to improve and tailor the IO model for various research objectives. Noting the insufficient effort in advancing the IO model towards time series analysis, this PhD thesis presents an innovation on the IO model that enable it for high-frequency time series analysis. In the innovated model, consumption and production activities are well formulated to incorporate delays and their accumulated impacts across the time domain. Complementary algorithms are also developed for the innovated model, so that instead of massive-scale economic surveys, alternative economic data with much higher time resolution can be applied as the observations to reconstruct an economic system that has its sectors dynamically interacting with each other in response to social demands. This research hence offers a new economic modelling tool to investigate the behaviours of an economic system over a shorter span and under a disequilibrium assumption.

Across all civilizations, human beings have explored science fundamentally in two ways – either theoretical science or experimental science. In theoretical science, researchers seek to explain natural phenomena through mathematical and conceptual models rather than direct empirical observation. It involves the use of theoretical frameworks, mathematical models, and computational simulations to develop hypotheses and predict the behaviour of complex systems. Theoretical science is an essential part of scientific research and is crucial to understanding the fundamental laws of the universe. On the contrary, experimental science is a field of scientific research that involves conducting experiments and making direct observations to gain a better understanding of natural phenomena. It involves the design, execution, and analysis of experiments to test hypotheses and investigate the behaviour of systems under controlled conditions. Experimental science is critical to advancing our knowledge of the world and plays a significant role in the development of new technologies and applications. Theoretical science and experimental science often work in tandem to create a complete picture of the natural world. Nevertheless, there are barriers in both end for the study of economics. Theoretical economists often make simplification, such as the general equilibrium assumption that demands always match supply, since the economic system is a complex system that our knowledge is limited about. As discussed in the chapters before, such simplification assumptions may not hold true if the economic models established are to be expanded over the time domain. On the other hand, experimental economics involves conducting experiments to test economic theories and hypotheses by controlling settings to observe how people behave in economic scenarios and collect data to test economic theories. For obvious reasons, conducting experiments on our economy may not be a good idea under most circumstance. It thus raises a dilemma for future economic research – economic system is too complex to be accurately described by current models, and there are limited ways to robustly test our models.

This research potentially points to a new direction in the study of economic science. Rapid advancements in technology for information collection and processing have paved the way for a third approach to scientific exploration: computational science. Characterized by the development and deployment of machine learning applications, computational science is quickly becoming prevalent in various critical fields, impacting our day-to-day lives. From weather forecasting to mechanical designs, computational science has proven to be incredibly useful in providing effective solutions at a significantly reduced cost. Some scientists advocate for the inclusion of computational science as a third pillar in our scientific knowledge structure, as illustrated in Figure 21. The availability of massive datasets and computational power now supports scientific research by providing validations for theories and models, as well as enabling analyses that were previously impossible without emerging computational tools. This thesis research exemplifies the successful application of computational science in the field of economics, showcasing the potential of harnessing advanced technology to gain deeper insights into complex economic systems and inform decision-making processes.

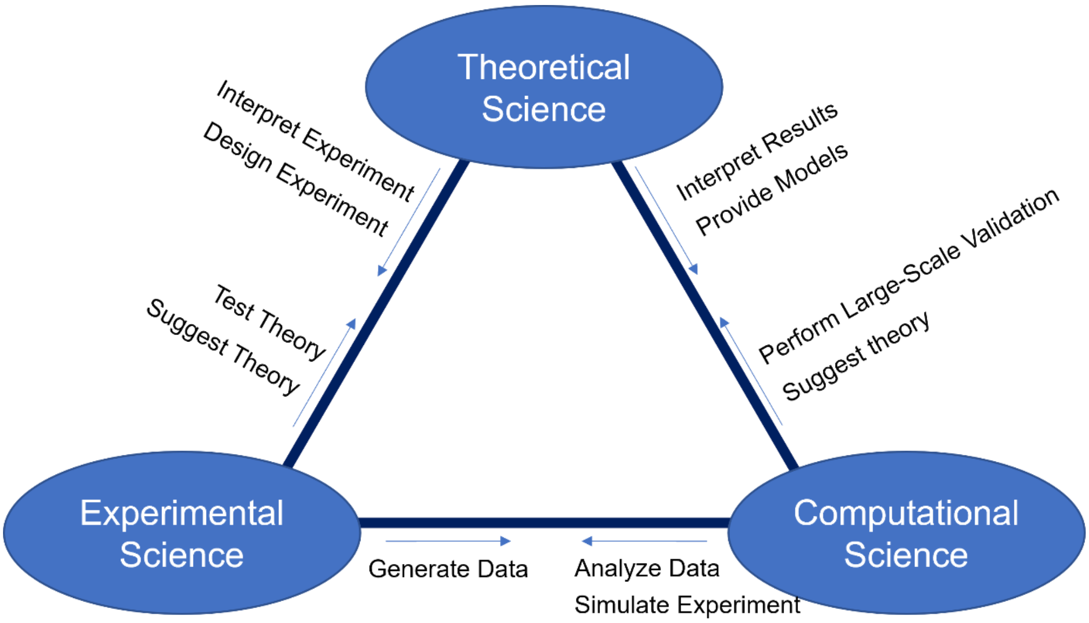



*Figure 21 The structure of science knowledge.*

## 6.1. Summary of Works

This thesis addresses directly to the gap between economic modelling and computational techniques with big data. Methodological innovation is conducted on the IO model in this thesis to tackle with economic system analysis and simulation under a non-Walrasian equilibrium setting, that supplies and demands are not set to equate each other at every time step in the investigated economic system. Specifically, this thesis proposes an innovative algorithm to combine regression algorithms with the SIM model, a variant version of IO model that considers the chronological interlinkages between economic sectors. Based on high time-frequency economic data, the proposed algorithm can reveal the chronological symbiosis of economic system, thus make nowcasting for the impact of shock events on the entire economy. As described in the section below, the research objectives listed at the beginning of this thesis are met accordingly.

### 6.1.1. Input-Output Model Application and Its Limitations

In Chapter 2, the concept and history of IO model have been comprehensively revised. Based on the classic IO model, Chapter 2 further introduced the most successfully developed and widely applied MRIO model and EEIO model. To demonstrate the capability of the two variant of IO model, a case study on the inequal distribution of consumption-based $CO_2$ emissions in China and changing $CO_2$ emissions embodied in domestic trade has been conducted in Chapter 2. The result shows that from 2007 to 2017, the consumption-based $CO_2$ emissions in China has been shifting from the more developed coastal region towards the less developed northwest inland region. Inner Mongolia is the only province with an increased $CO_2$ emission intensity, while also remains as the largest net exporter of consumption-based $CO_2$ emission from 2007 to 2017. For developed provinces like Beijing, Shanghai, Tianjin, and Zhejiang, the net imports of consumption-based $CO_2$ emissions are continuously decreasing, with the decreases counting 48 Mt, 100 Mt, 30 Mt, and 78 Mt respectively from 2007 to 2017. On the contrary, net export of consumption-based $CO_2$ from developing provinces like Hebei, Henan, Shanxi, and Guizhou also drastically decreased from 2007 to 2017, amounting 123 Mt, 103 Mt, 48 Mt, and 46 Mt respectively. Nevertheless, as the largest net importer and exporter of consumption-based $CO_2$ emission in 2017, Guangdong and Inner Mongolia (96 Mt imported and 146 Mt exported consumption-based $CO_2$ respectively) have not undergone much change in their traded consumption-based $CO_2$ emissions from 2007 to 2017 (differences count 17 Mt and 13 Mt respectively). This observation can be explained by the latency of provinces $CO_2$ decoupling with economic development.

The study further extends its analysis to $CO_2$ emissions embodied in trade between China and the world. It reveals that $CO_2$ emissions are not only migrating to the less developed parts of China, but also the less developed countries in other regions of the world. In general, the $CO_2$ emissions embodied in China's net exports to the world decreased for all regions. However, developed region's $CO_2$ emissions embodied in its net imports show an increase from the world other than China from 2012 to 2017, while the net exports of embodied $CO_2$ emissions from emerging economies also show an increase from 2012 to 2017. Hence, Chapter 2 brings up an alarming message to policy maker that other emerging economies may take over China's role of pollution haven and produce the emission intensive products for the rest of the world. On the other hand, this study has also exposed the weakness of conventional IO model: the resources, be it labour, time, or finance, needed for IO table compilation is tremendous. IO model is thus better suited to analyse economic performance in the past.

### 6.1.2. Theoretical Improvement for the Input-Output Model

Given the weakness of IO model, Chapter 3 starts with an account for some of the most successful methodological innovation for the IO model. In the literature review section of Chapter 3, Structural Decomposition Analysis, Hypothetical Extraction Method, Hybrid IO Model, and IO-based Network Analysis have been introduced in detail in terms of their rationales, derivations, and applications. In response to the weakness exposed in the previous Chapter, Chapter 3 then particularly introduced the SIM model in more details as it specifically tackles incompetency of IO model in chronological analysis. Chapter 3 then identify the new direction for SIM model's development to be the integration with computational tools and the availability of adequate data. The SIM model is reorganized into a linear matrix system and then solved with a regression algorithm to find the best fitted production coefficients from an ample number of observations for outputs and final consumptions.

To verify the efficacy of the proposed algorithm, over 200 discrete time steps of total outputs are simulated based on a set of given levels of final demand and a chronological interindustry coefficient matrix. The regressive algorithm is then applied to obtain the coefficient matrix. Comparison between the coefficient matrix obtained and the actual matrix shows that the difference between the two quantities is minimal ($10^{-6}$), so it can be safely concluded that the regressive algorithm is mathematically capable in solving the system simulated by SIM mode. Discussion is also provided in the end on the scenarios that the algorithm based on SIM can be used.

### 6.1.3. Electricity Data in Economic Analysis

To put the methodological advancement into real application, Chapter 4 investigated the economic symbiosis of Chongqing municipality of China using the algorithm developed in Chapter 3. In order to fulfil the requirement on high time-resolution and sector specific data, the daily electricity consumption data of Chongqing for more than two consecutive years is obtained and reorganized to form a series of observations on the daily outputs and demands in different sectors. Since electricity consumption is a fair reflection of economic activities, this data set is used as a proxy for economic analysis and aggregated into eight sectors. The first 500 observations are used to obtain the chronological production coefficients using the innovated regression algorithm on SIM model in Chapter 3. Validation using the remaining data set shows that the interindustry relationships obtained can predict the economic performance pretty accurately if given a certain demand level, containing the error levels within 30%. Compared to other economic prediction studies, this exercise demonstrates a relatively low error level, hence concluding the real-life efficacy of the methodological development of this research.

As a demonstration of application, three growth scenarios of demand in consumer goods are simulated to show the holistic and chronological impacts on all economic sectors in Chongqing 80 days into the future. In the scenario setting, the forecast of daily growth of final consumption of the consumer goods sector is varied from 0.41% to -0.40% for the surging and plunging scenarios, respectively. The trained model predicts a 3.5 GWh/day difference in the mean value of the total consumption of electricity in the consumer goods sector. In contrast, although the final consumption is kept unchanged in the surging and plunging scenarios, the heavy industry (18.8 GWh/day), manufacturing (6.8 GWh/day), and EHGW (11.7 GWh/day) sectors interestingly show larger differences in the mean value of forecasted total electricity consumption than the consumer goods sector, probably due to the multiplier effect of demands and outputs across sectors. This Chapter thus effectively shows the applicability of the proposed algorithm in Chapter 3 in the analysis of real economic research question.

### 6.1.4. Introducing Capital Change into Economic Modelling

In Chapter 5, the effects of capital formation and deterioration on production changes across sectors are considered in the SIM model in a differential form. Essentially, an additional capital coefficient is introduced and linked to the changes of production level, so that the new model becomes a nonlinear system. By creating an improved algorithm, the new model can be solved in a similar manner as shown in Chapter 3. The new model is then applied to assess the indirect economic loss of the 2015 South India Flood. Using the 2015 monthly industrial production index and manufacturing index of Indian regions and the 2015 Indian MRIO table, a number of observations on the output levels and final demand are simulated and used to solve for the chronological production coefficients. Due to constraints on data availability, the MRIO table of India is aggregated into two region (Tamil Nadu Pradesh and Rest of India) and one single sector with time resolution set to one month, so that the solved model suggests the chronological economic interactions between Tamil Nadu and Rest of India. Comparing with the actual outputs from April 2015 to April 2016 shows that the percentage errors ranged as 0.0%-2.1% respectively, suggesting the robustness of the model trained.

Under a business-as-usual scenario that assumes the direct loss of the disastrous event is added back to the final demand of Tamil Nadu Pradesh, Chapter 5 analyses the indirect losses induced in not only Tamil Nadu but also the rest of India. In comparison to the available estimation of lost due to 2015 South India Flood, this study estimates the direct and indirect costs induced across the whole India to be more than double. This result hence demonstrates a case that the innovated SIM model can be applied without more stringent data requirement.

## 6.2. Methodological Innovations

The most critical contribution of this study is its innovation in methodologies. Specifically, it proposes an effective way to incorporate data science with the classic IO model through a creative algorithm application. Along with other development on the IO model illustrated in Figure 22, the methodological innovation in this research advanced the SIM model and Dynamic IO model, thus enabled it to be used for high time-frequency big data analysis on the chronological interindustry linkages among regions and economic sectors. The implications are explained below.

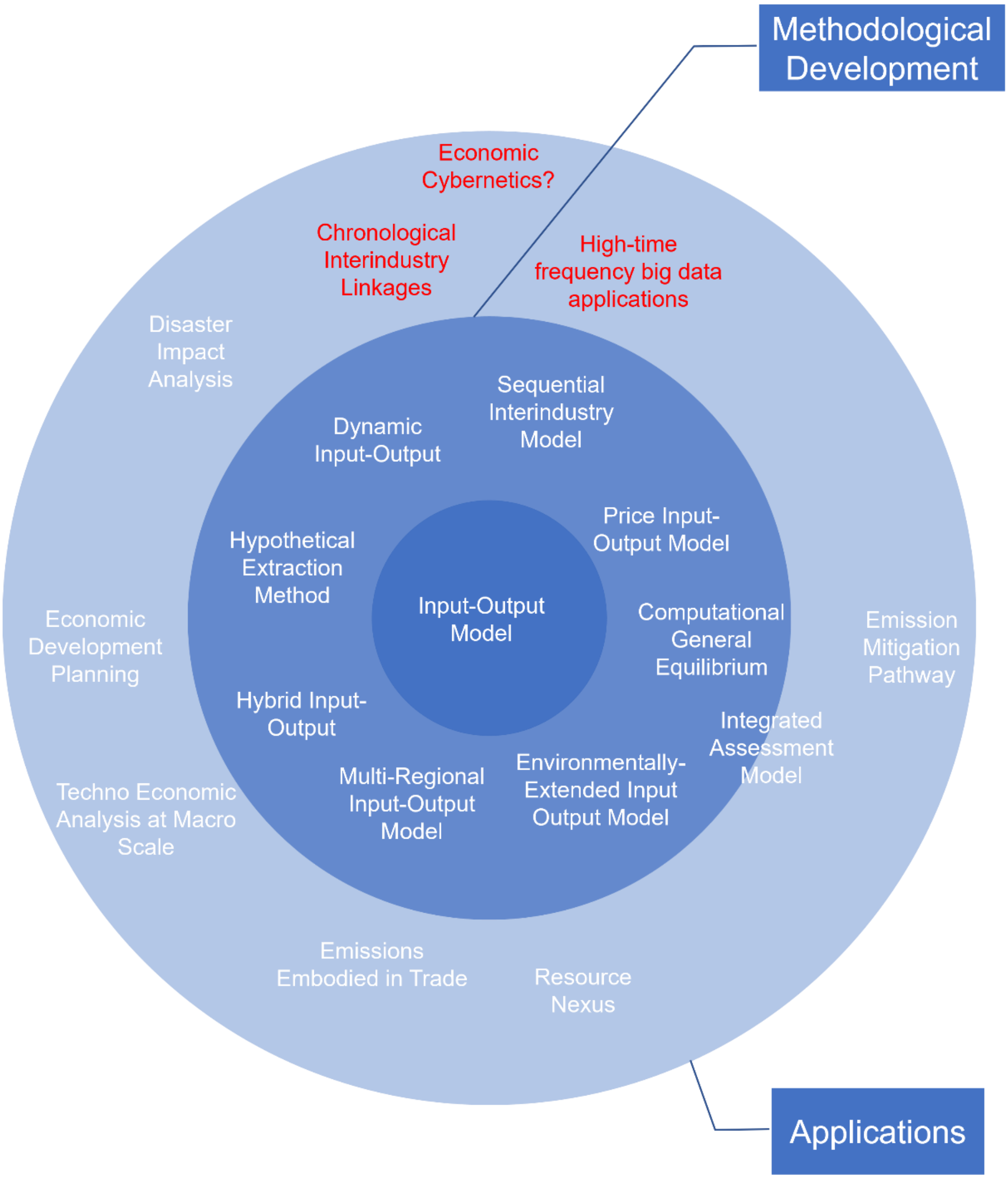



*Figure 22 A review on the development of the IO methods and their applications. Highlighted in red are the potential contributions of this thesis.*

- Providing a way of integration for data research tools and economic studies. In past research practices, data research tools that utilizes big data sets, such as machine learning approaches under heated attention recently, treats the system studies as a “black box” in return for higher prediction performance. Economic study, on the other hand, attempts to incorporate more and more economic theories and build a Grand Unified Theory, such as CGE modelling, for our economic system. The methodological contribution of this research blurs the distinct boundary of the two research directions by proposing a regressive algorithm for the SIM model, so that economic reasoning is provided for big data research to analyse how an investigated economic system performs under high time frequency.

- Enabling more dynamic scenario prediction. The current toolbox for economists in economy performance prediction is limited. Most of available economic modelling tools assumes a long-term general equilibrium condition, which is not always met in short term scenarios. On the other hand, short term future prediction involves the use of statistical models and machine learning algorithms to analyse large amounts of data and identify patterns and trends. The algorithm proposed in this research demonstrates a new economic modelling tool that can be utilized to, at an acceptable degree of accuracy, predict the short-term future performance of an economic system under demand shock events of a certain sector.
- Alerting production to demand signal. The proposed method can reveal the delayed economic impacts across different sectors under an economic shock event in a selected sector. Hence, the understanding of the chronological impacts across the economic system can be used to build an alerting system for economic impact. For instance, if a known demand shock signal is detected in the economic system, a quick analysis can be formulated to quantitatively understand which and when another sector will be affected by the shock event. Policy maker can thus use the result to react with more precise counter actions.
- Combining with available economic modelling tools. As shown in Figure 22, this research is one of the many directions in the development of IO modelling. Same as in other research, the methodological development of this research can be pollinated with other modelling tools. For instance, the economic activities can be directly linked to environmental stresses to dynamically account for the delay in economic consequences. Other economic theories such as elasticities and utility functions can be modelled into the model in this research to further expand the research scope of this method, hence answering a wider range of economic research questions.
- Applying to system modelling. The analogy of economic system to other systems is not uncommon. Thus, there are instances where modelling tools are shared between economic and other fields of research. The chronological model developed in this research may hence be applied in other research that features the functioning of a balanced system, the ecological system for instance.

## 6.3. Future Potentials

With the research conducted in this thesis laying the foundation for further investigations, it is important to consider future developments in the field of economic system modelling. As proposed and demonstrated by this research, it is crucial to borrow methodological development in other sciences. As such, it is essential to identify areas where the research on economic system modelling can be expanded upon, as well as potential challenges and limitations that may arise.

### 6.3.1. An Alternative Research Direction to General Equilibrium

As briefly touched in the previous chapters of this thesis, general equilibrium and general disequilibrium are two different concepts in economic research that describe different states of an economy. The basis of the SIM model and algorithm developed in this research is the general disequilibrium assumption.

General equilibrium refers to a state of the economy where all markets are simultaneously in balance, meaning that the supply of goods and services equals the demand for them, and all prices are at their equilibrium levels. In this state, there is no excess demand or supply in any market, and there is no incentive for market participants to change their behaviour. A classic example of the general equilibrium theory is the classic IO model.

On the other hand, general disequilibrium refers to a state of the economy where one or more markets are out of balance, leading to excess supply or demand and price distortions. In this state, market participants have incentives to adjust their behaviour to restore equilibrium, leading to changes in prices and quantities in various markets. General disequilibrium is often used to analyse the dynamics of market adjustment and the implications of market frictions. The SIM model applied in this research can be deemed as a demonstration of general disequilibrium theory.

In 20th century research of economic theories, neoclassical economists deeply believe in the general equilibrium, but consider Keynes' disequilibrium adjustment a fundamental stray from the domain of economics (Foster, 2006). On the contrary, the general disequilibrium theory disagrees with many of the assumptions made by Walrasian equilibrium. It argues that the process of the economy adjusting itself towards an equilibrium should be the focus of economic study since that is most of the cases for the states of our economies (Grossman, 1971). Disappointedly, economists gradually lost their interest in the disequilibrium theory ever since its first proposal in the 1970s. In the most recent discussions of economic theorists, the study of "non-equilibrium economics" has been re-proposed by a few economists (Berger, 2009). It relates to the alternative field of economic research, such as complexity economics.

The study of economic phenomena through the lens of complex systems theory is a new and burgeoning field known as complexity economics (Arthur, 2021).

It aims to comprehend the workings of economic systems, viewing them as intricate adaptive systems characterized by agents that engage in interaction and learning from one another. Macro-level economic phenomena are seen to arise from the intricate interactions of numerous agents at the micro-level. Complexity economics explores various intriguing topics such as network formation, the emergence of market structures, the dynamics of financial markets, and the evolution of economic institutions. To investigate these phenomena, researchers in this field employ a diverse array of mathematical and computational tools, enabling them to model, simulate, and scrutinize the properties of economic systems. Complexity economics research has a wide range of potential applications. It could contribute to the improvement of economic policies and the creation of innovative techniques for financial regulation. Additionally, it could provide valuable analysis on the influence of technological advancements on the economy. This field of study offers unique perspectives on the behaviour of economic systems and the difficulties faced in economic policy-making. A fundamental insight of complexity economics is the presence of non-linear feedback loops, path-dependence, and emergent properties that defy explanation within traditional economic theory.

### 6.3.2. Similarity with Control Theory

Interestingly, the study on the oscillation of a system with feedback and time delay is a well-studied topic in the control theory of engineering, including the causes of delay, the magnitude of delay, and the methods of compensation for delay. Control system researchers explore how delay affects different types of control systems, such as linear and nonlinear systems, and investigate various techniques to mitigate the negative effects of delay. Some of the applications of delayed feedback research include improving the stability and performance of communication networks, enhancing the control of mechanical systems, and optimizing the performance of industrial processes.

In its applications, control theory uses mathematical models to describe the process of a system over time. By writing a state equation, it describes the system's internal behaviour as a function of its current state, inputs, and time. Furthermore, the state space representation diagram in the control theory includes vectors representing the system's state and input. It also utilizes a state transition matrix describing the system's dynamics, very similar to the production coefficient matrix in IO model and SIM model. In fact, the similarities and potential synergy between control theory and IO model have been identified in earlier discussion by Livesey (1971), which also emphasized the importance of time lag and fluctuations when integrating IO models with control theory concepts. However, limited development has been consolidated along this path.

In control theory studies, state space representation diagrams are widely used to provide a convenient and systematic way to analyse the behaviour of a system. They can be used to design controllers that regulate the system's behaviour and optimize its performance, and to simulate the system's behaviour under different conditions. To illustrate the overlap between control theory and IO modelling, a state space block representation diagram below (Figure 23) is constructed for equation *( 5.2 )* in Chapter 5.

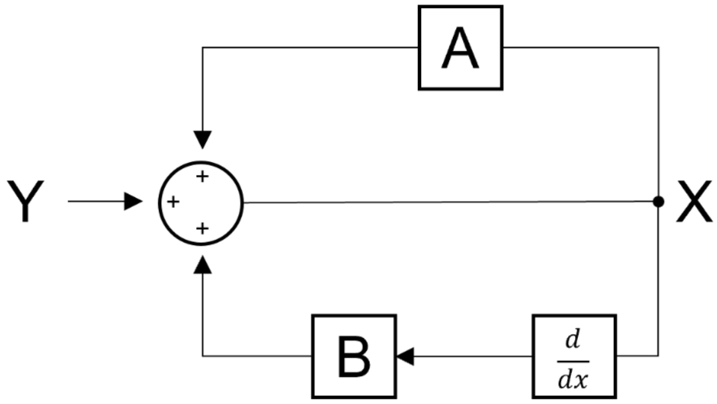



*Figure 23 A state space diagram representation of the dynamic IO system described in equation ( 5.2 )*

It is thus appropriate to conclude that there is great potential in interdisciplinary methodological innovations between IO modelling and control theory. In fact, the limitation of productivity cap mentioned in the previous section can be tackled with some techniques straight away. In control theory modelling, caps on the transfer process are typically modelled by adding saturation or limit functions to the transfer function of the system. Saturation functions limit the output of the system to a certain range or "cap," which is often specified by the modeller. For example, in a motor control system, the output torque may be limited to a certain maximum value to prevent the motor from overheating or causing damage.

In recent economic research, the borrowing of control theory concepts and technics into economic modelling is no longer an unusual practice. Recent Nobel Economic laureate Lars Peter Hansen has demonstrated in his work that how robust control theory can be introduced to facilitate macroeconomic decision (Hansen and Sargent, 2016). Many more research has applied control theory in the study of financial market performance (Lo, 1988, Geanakoplos, 1992). Hence, introducing control theory into IO modelling may bring about a potential breakthrough in macroeconomic research.

### 6.3.3. Economics Research amid Machine Learning Fanaticism

In the writing of this thesis, breakthroughs on artificial intelligence are being made. Other than the trending release of Generative Pretrained Transformer (GPT) that recently exhibits extraordinary ability in natural language processing, another key principle of artificial intelligence study is machine learning. Machine learning involves the training of models to learn from data and improve their predictive performance over time through recognition of data patterns. In that aspect, machine learning shares a handful similarity with economic modelling such as econometric studies as they 1) are both concerned with analysing data to understand the relationships between variables; 2) both make use of statistical methods and techniques to analyse data, including regression analysis, hypothesis testing, and model selection; 3) Both can be used to make predictions and forecast future trends based on historical data.

However, machine learning and economic modelling are yet to be understood as the same thing. The focus of machine learning is making predictions for decisions based on so-called applied statistics. In contrast, economic modelling investigates the relationships and understand the impact of different variables based on economic theories. While machine learning is more focused on identifying patterns and relationships in the data, economic modelling emphasis more on the reasoning underlying the data. Besides, economic models often apply economic variables such as GDP, inflation, and interest rates, while statistical learning models can be applied to a wide range of fields, including electricity consumption data utilized in Chapter 4 of this thesis.

Although the machine learning approach has gained much more attention due to its terrific performance in making predictions, some have reiterated the importance of reasoning in our understanding of the world. In the recent release of ChatGPT 4, scientists have raised the concern on the application of machine learning tools. Although violently upgrading training hardware settings can improve models' performance, it simply overlooks the reasoning behind models and treats an investigated system as a "black box". The relationships revealed by machine learning algorithms can be highly complex and impossible to interpret economically. Thus, the opacity of machine learning algorithms makes it difficult to evaluate the accuracy or validity of their outputs. Policymakers may also miss out the important information to understand how to adjust or refine economic policies based on the results.

This research sets a new direction for the interdisciplinary research between the state-of-art machine learning tools and IO economic modelling. On the one hand, this research proposed an innovative path to apply commonly used machine learning big data (daily electricity consumption data by sectors) to the economic IO model for high-time frequency analysis and short future performance prediction through machine learning style regression algorithms. Such feature greatly improved the analysis capability and predictive performance of the model. On the other hand, economic reasoning is offered to supplement the regression algorithm, so that compromise on the explanatory capability of the model is mitigated. Policy makers can thus infer more meaningful results from the regressed relationship and facilitate decision making.

### 6.3.4. Compatibility of Economic Cybernetics

Building on the discussion of this research's algorithmic contributions and its connections to control theory and machine learning, the term "economic cybernetics" has emerged as an apt descriptor for the expected outcomes of this integration.

Economic cybernetics is the field of study that applies cybernetic principles and methods to analyse and manage economic systems. The term "cybernetics" refers to the study of communication and control in complex systems, based on the idea that systems can be modelled and analysed as feedback-controlled systems. Economic cybernetics aims to optimize economic systems by creating models and algorithms to analyse the behaviour of economic agents such as producers and consumers, as well as the dynamics of markets and financial systems.

One of the key principles of economic cybernetics is feedback control (Lange, 2014), which involves using information from the output of a system to adjust its inputs and achieve desired outcomes. This can include designing policies and interventions to stabilize markets, regulate financial systems, and improve economic performance. In this view, the research method of economic cybernetics overlaps with the methodological contribution of this thesis.

Additionally, recent developments in economic cybernetics have focused on the application of advanced computational and machine learning techniques to economic modelling analysis and future economic trends predictions (Cope and Kalantzis, 2022). These developments have been driven by the availability of large amounts of high frequency economic activity data and the increasing computational power of modern computers, which have made it possible to develop more complex and sophisticated models of economic systems. To that end, the principal and objective of economic cybernetic study coincide with this thesis.

#### *6.3.4.1. The History of Economic Cybernetics*

In 1948, Norbert Wiener published "Cybernetics: Or Control and Communication in the Animal and the Machine" (Wiener, 1961) simultaneously in the United States and France, marking the establishment of the science of cybernetics. In the mid-1950s, Herbert A. Simon (Simon, 1997, Newell and Simon, 1956) was one of the prominent social scientists who studied the optimal control problem of macroeconomics. Along with his colleagues in the United States, Simon contributed to the development of the study in computer science, economics analysis, and cognitive psychology in the attempt to improve the stability of economic policies. Simon's contributions helped to shape the understanding of economic cybernetics.

Meanwhile, the Soviet Union and some Eastern European countries regarded cybernetics as pseudoscience and initially criticized it. However, by the late 1950s, they shifted their stance and began to allocate resources to pioneer the study of economic cybernetics. Since the 1950s, Soviet scientist Nikolai Veduta had worked closely with scientists from the Central Economic Mathematical Institute (CEMI) to develop mathematical models for economic planning. His book “Economic Cybernetics” published in 1971 advocated for the widespread use of cybernetic methods in planning the production, management, and overall organization of the Soviet economy (West, 2020). In 1975, Soviet economist Leonid Kantorovich won the Nobel Prize in Economics for his proposal of linear programming, which contributed to the theory of optimal resource allocation in economic planning. With much support from Kantorovich, the economist-mathematician Vasily Nemchinov played a key role in the development of the Soviet school of economic cybernetics. In 1958, he founded the first laboratory of economic-mathematical modelling in Moscow, which laid the groundwork for the establishment of the CEMI within the Academy of Sciences in 1963 (West, 2020).

Poland scientists have also contributed significantly to the early development of economic cybernetics. In 1955, the Polish Cybernetics Society was founded to act as one of the first cybernetics societies in Europe. The Polish Academy of Sciences established Poland's Central National Economic Plan System Model using applied control theory. In the early 1960s, they introduced the basic principles of economic cybernetics, publishing monographs and textbooks, and offering courses on the topic in higher financial and economic institutions. Oskar Lange, a Polish economist, was one of the pioneers of economic cybernetics. He believed that cybernetic methods could be used to design self-regulating control systems for a real-time planned economy and published his famous work Introduction to Economic Cybernetics (Lange, 2014).

In addition, the study on economic cybernetics has received attention from the Romanian regime since it was first introduced into Romania in the 1960s. The Romanian government showed interest in embracing cybernetics to improve the country's centralized planning system. Some Romanian economists, academics, and engineers had devoted into the study of economic cybernetics. In 1965, Nicolae Ceausescu, then general secretary of the Romanian Communist Party, emphasized the need to seriously study control theory and other fields that Romania was relatively backward at the time. Ceausescu wanted to modernize Romania's economy and reduce its dependence on the Soviet Union. In the 1970s, governments began to implement cybernetic methods in various sectors of the economy, including energy, transportation, and agriculture. The Central Commission for Computing and Informatics was created to oversee the development and application of computer technology and cybernetics. In 1979, Manea Mănescu, a renowned expert in economic control theory and former Prime Minister of Romania, published "Economic Cybernetics" (Mănescu, 1980). In the book, he defined economic control theory as "a management tool that enables the economy to achieve optimal, balanced, and proportional growth" and as "an extremely important branch of control theory." While there were some successes, the centralized and autocratic nature of the regime, coupled with economic challenges, ultimately limited the impact of cybernetics on Romania's economy. In the early 1970s, the socialist government of President Salvador Allende turned to economic cybernetics in order to nationalize industries, implement agrarian reforms, and redistribute wealth to address social inequalities. The most notable example of economic cybernetics in Chile during this period was Project Cybersyn (Samothrakis, 2021), a ground-breaking and ambitious project developed under the guidance of British cybernetician Stafford Beer. The aim of the project was to create a real-time computer-based control system for the Chilean economy, allowing the government to make rapid decisions and adjustments to optimize production and distribution. Despite its innovative nature, Project Cybersyn faced several challenges, including limited resources, outdated technology, and a lack of complete data from all industries. The project was never fully implemented or operational, as it was interrupted by the military coup in September 1973. Following the coup d'état, Project Cybersyn was dismantled as Chile's economic policy was shifted towards a more market-oriented and neoliberal approach. Although the project was a groundbreaking and ambitious initiative, it was never fully realized due to the political instability and the eventual military coup that led to a dramatic shift in Chile's economic policy.

The study of economic cybernetic has also raised interest in China too. For instance, Qian Xuesen was a prominent Chinese scientist who made significant contributions to the fields of rocket science, aerospace engineering, and systems engineering. He is widely regarded as the founding father of China's space program and is credited with helping to launch China's first satellite, among other achievements. At later stage of his research career, Qian developed his interest in systems engineering and system science, which has some overlap with economic cybernetics. In the 1970s and 1980s, Qian advocated for the application of systems engineering and system science principles to guide China's economic and technological development. Qian believed that systems engineering could provide a holistic and interdisciplinary approach to solving complex economic and technological problems. He emphasized the importance of considering the interrelationships between different components of an economic system and promoted the idea of optimization and efficient resource allocation. One of Qian's notable contributions to the field of systems engineering is his development of the "systems methodology," which he introduced in a series of lectures in 1980. This methodology emphasizes the importance of understanding complex systems, such as economies, from a holistic perspective and provides a framework for analysing the relationships between the various components of a system. Although Qian Xuesen's work was not specifically focused on economic cybernetics, his advocacy for systems engineering and system science principles has influenced the way China has approached economic planning and development. His work has inspired generations of Chinese scientists and engineers to apply systems thinking to various fields, including economics (Qian et al., 1993).

#### *6.3.4.2. Future Development*

In recent years, economic cybernetics has undergone significant evolution due to advances in technology, data availability, and computational methods. The accessibility of vast amounts of data has empowered researchers to explore economic systems more thoroughly. Data analytics techniques are employed to identify patterns, trends, and relationships within large datasets. Agent-based modelling has also progressed due to the theoretical development of economic research. This computational method simulates the interactions of autonomous agents, such as producers and suppliers, within an economic system. These models aid researchers in understanding the behaviour of complex economic systems by examining individual agent interactions and the emergence of macro-level patterns. Artificial intelligence and machine learning techniques have become increasingly vital in the field of economic cybernetics. They enable researchers to analyse complex economic systems, make predictions, and optimize decision-making processes. Applications of AI and machine learning in economics encompass forecasting, risk assessment, and algorithmic trading. Moreover, network theory has been applied to economic cybernetics to study the relationships between various economic agents and the structure of economic systems, investigating the connections between different agents. Complexity economics has also emerged as an interdisciplinary approach to studying economic cybernetics. It highlights the importance of understanding the complex, adaptive, and nonlinear nature of these systems by incorporating ideas and methods from other scientific fields. Complexity economics has contributed to the development of new tools and techniques for analysing economic systems, including agent-based models and network theory.

The research conducted in this thesis can be considered a contribution to the methodological toolbox of economic cybernetics, expanding the field's capacity to analyse and model complex economic systems. As thoroughly discussed in the previous section, the algorithm developed in this thesis, designed to complement and enhance the SIM model, shares numerous similarities with the most recent advancements in the realm of economic cybernetics. This cutting-edge methodological approach brings forth a new dimension to the study of economic systems. The innovative method is capable of handling high time-frequency economic activity big data, utilizing a machine learning-inspired regression algorithm to uncover the intricate chronological interactions among different industries within an economic system. By doing so, it allows for a deeper understanding of the relationships, dependencies, and influences between various industrial sectors, ultimately resulting in a more comprehensive and accurate representation of the entire system. Moreover, the implementation of this novel method has the potential to greatly improve the accuracy and predictive capabilities of economic cybernetics models. With the ability to process and analyse electricity consumption big data, the newly developed algorithm can effectively identify patterns within the complex web of economic interactions, thus leading to more informed policy decisions to develop more targeted and effective strategies for economic growth and sustainability. In addition to the direct benefits for the field of economic cybernetics, the innovative algorithm presented in this thesis may also have broader implications for other disciplines within the social sciences. By providing a robust, flexible, and adaptable methodological framework for analysing large-scale, high-frequency data, this research could potentially serve as a springboard for further advancements and interdisciplinary collaborations in areas such as econometrics, computational sociology, and data-driven public policy analysis.

Specifically, the future potential of this research may include the following. Firstly, the algorithm's ability to process and analyse large volumes of high-frequency economic data enables policymakers and central planners to make better-informed decisions, thus central planners can develop more targeted and effective strategies for resource allocation, production, and distribution. Secondly, the lack of flexibility and adaptability to changing circumstances has been a key challenge in macro-level economy planning. The algorithm developed in this research can be utilized to analyse high-frequency data and allow for real-time monitoring of economic indicators, thereby increasing the overall efficiency of the planned economy. Thirdly, the machine learning-inspired regression algorithm can help forecast economic trends and potential disruptions with greater accuracy, allowing central planners to be better prepared for potential challenges and ensuring more resilient economy planning. Fourthly, this algorithm enables a chronological economic model that has a deepened understanding of the relationships, dependencies, and influences between various industrial sectors. This information can be utilized to optimize resource allocation and facilitate better coordination between different industries. Lastly, the innovation in this study method can also help central planners to develop and evaluate different policy scenarios and their potential impacts on the economy. By simulating various short-term policy alternatives, central planners can identify the most effective strategies for achieving desired economic goals with less cost in policy development time.

## 6.4. Limitations

As an inaugural attempt on a hybridization of data science and IO modelling, this research is of course limited in various aspects.

Firstly, the quality of data used in this study may bear limitations and adversely affect the performance of the model. In the exercise conducted in Chapter 4, electricity data is used for the proposed algorithm, assuming that they are good representation of economic activities by sectors. However, this assumption has not been extensively verified and supported by peer research. In addition, due to limited capability on signal processing skills, the electricity consumption data applied to the model has not been professional cleaned to filter out abnormity data points. Thus, the result obtained from this data set may be prone to systematic errors. In Chapter 5, economic indices are used to estimate economic activities, which is an oversimplification and a compromise have to be made due to unavailability of higher quality data. In future applications, a necessary step would be to validate the model with multiple data sources, such as monetary transection at high frequencies.

The model itself is neither a perfect reflection of economic activities, so a series of new endeavour on disequilibrium economic modelling would be needed to further improve the model. In many other IO-based time series analysis of economic performances and disastrous shocks, many researchers have mentioned the need to consider bottleneck effect (Mendoza-Tinoco et al., 2020, Zeng and Guan, 2020a). By bottleneck effect, researchers point out that economic production may not always be linearly scaled up by intermediate inputs, since any economic structure has a limit on its capacity. As production expands towards its limits, a cap, or diminishing economic to scale, will be applied and distorts the linear relationship in production. In future developments, researchers may need to factor in the capacity limit into the SIM model. Techniques from signal processing, such as sigmoid damage function, may be a possible solution to address this issue.

Lastly, the application demonstrated in this research is limited by the hardware installation used. As mentioned in Chapter 4, the number of propagation layers solvable with available computing resources is limited to 12, meaning that the model built in Chapter 4 assumes demand signals that occurred in day one only extends its impact on all the other sectors 12 days into the future. This assumption may not hold true, but there are insufficient computational resources for validation or disproval. Beyond 12 layers, the personal computer used for this thesis research ran out of memory. In the practice of more recent commercial machine learning applications, much more advanced super computational hardware with tens and hundreds petaflops/sec. Since such computational power is not reachable on a personal computer’s setting, exploration on the performance of this model in a more advanced hardware setting may further improve the capability of this model.

In the popular book "Homo Deus: A Brief History of Tomorrow," Yuval Noah Harari explores the role of data and data science in shaping our future (Harari, 2016). One key idea presented by Harari is the concept of Dataism, which posits that the world is increasingly being driven by the flow of data and the ability to process and analyse that data. The importance of data science and its ability to transform various aspects of our lives, including economic systems, cannot be understated. The contribution of this thesis can undoubtedly be linked to the points made by Harari in "Homo Deus" regarding the significance of data science. The development of algorithm in this thesis for handling high-frequency economic activity data is a prime example of how data-driven approaches can potentially revolutionize our understanding and management of complex systems like economies, although it may be just a starting point of this field of research. It aligns with the principles of Dataism by enabling economic planners to make more informed, data-driven decisions in for an economy. As Harari argues in his book that the increasing reliance on data and algorithms could lead to a shift in power dynamics. Algorithms can potentially take on a more prominent role in decision-making processes. The advancements in economic cybernetics, as demonstrated by the algorithm developed in this thesis, showcase how data science can augment economic decision-making capabilities in the realm of production planning and resource allocation. Along the course set by this research, future endeavour could explore the scalability of this method to other chronological economic system models. As the IO model can be closely linked to the physical process of manufacturing, modellers may also consider the feasibility of using units other than monetary terms to denote interactions among economic sectors (Dietzenbacher et al., 2009), hence carry on perfecting the knowledge on our economy.

# References

ACKERMAN, F., NADAL, A. & GALLAGHER, K. 2004. The flawed foundations of general equilibrium. *Critical Essays on Economic Theory*.

AHMAD, T., ZHANG, H. & YAN, B. 2020. A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. *Sustainable Cities and Society,* 55**,** 102052.

AKHABBAR, A. & LALLEMENT, J. 2010. Wassily Leontief and Léon Walras: the Production as a Circular Flow.

ALBINO, V., IZZO, C. & KüHTZ, S. 2002. Input–output models for the analysis of a local/global supply chain. *International journal of production economics,* 78**,** 119-131.

ALCáNTARA, V. & DUARTE, R. 2004. Comparison of energy intensities in European Union countries. Results of a structural decomposition analysis. *Energy Policy,* 32**,** 177-189.

ALDY, J. E. 2005. An Environmental Kuznets Curve Analysis of U.S. State-Level Carbon Dioxide Emissions. *The Journal of Environment & Development,* 14**,** 48-72.

ARTHUR, W. B. 2021. Foundations of complexity economics. *Nature Reviews Physics,* 3**,** 136-145.

ASHRAF, Z., JAVID, A. Y. & JAVID, M. 2013. Electricity consumption and economic growth: evidence from Pakistan. *Economics and Business Letters,* 2**,** 21-32.

ASLAN, A., APERGIS, N. & YILDIRIM, S. 2014. Causality between energy consumption and GDP in the U.S.: evidence from wavelet analysis. *Frontiers in Energy,* 8**,** 1-8.

AULIN-AHMAVAARA, P. 1990. Dynamic Input–Output and Time. *Economic Systems Research,* 2**,** 329-344.

AULIN-AHMAVAARA, P. Dynamic input-output and capital. Presentado en: 13th International Conference on Input-Output Techniques. University of Macerata, Italia.(Agosto 2000), 2000. Citeseer.

AVELINO, A. F. & HEWINGS, G. J. 2019. The challenge of estimating the impact of disasters: Many approaches, many limitations and a compromise. *Advances in spatial and economic modeling of disaster impacts.* Springer.

AVELINO, A. F. T. 2017. Disaggregating input–output tables in time: the temporal input–output framework. *Economic Systems Research,* 29**,** 313-334.

AVELINO, A. F. T., FRANCO-SOLíS, A. & CARRASCAL-INCERA, A. 2021. Revisiting the Temporal Leontief Inverse: New Insights on the Analysis of Regional Technological Economic Change. *Structural Change and Economic Dynamics,* 59**,** 79-89.

B. MüLLER, D. 2006. Stock dynamics for forecasting material flows—Case study for housing in The Netherlands. *Ecological Economics,* 59**,** 142-156.

BABATUNDE, K. A., BEGUM, R. A. & SAID, F. F. 2017. Application of computable general equilibrium (CGE) to climate change mitigation policy: A systematic review. *Renewable and Sustainable Energy Reviews,* 78**,** 61-71.

BARKER, K. & SANTOS, J. R. 2010. Measuring the efficacy of inventory with a dynamic input–output model. *International Journal of Production Economics,* 126**,** 130-143.

BERGER, S. 2009. *The Foundations of Non-Equilibrium Economics: The Principle of Circular and Cumulative Causation*, Routledge.

BEYLOT, A., CORRADO, S. & SALA, S. 2020. Environmental impacts of European trade: interpreting results of process-based LCA and environmentally extended input–output analysis towards hotspot identification. *The International Journal of Life Cycle Assessment,* 25**,** 2432-2450.

BICKEL, R. 1987. Practical limitations of the input-output model for research and policymaking. *Educational Policy,* 1**,** 271-287.

BLAIR, P. D. & MILLER, R. E. 2022. Mixed and Dynamic Models. *Input-Output Analysis: Foundations and Extensions.* 3 ed. Cambridge: Cambridge University Press.

BREMNER, L. 2020. Planning the 2015 Chennai floods. *Environment and Planning E: Nature and Space,* 3**,** 732-760.

BURFISHER, M. E. 2021. Regulations in a CGE Model. *In:* BURFISHER, M. E. (ed.) *Introduction to Computable General Equilibrium Models.* 3 ed. Cambridge: Cambridge University Press.

CAO, A. 2022. *Lockdowns and supply chain disruption to accelerate Apple's move away from China, with India a likely beneficiary, analysts say* [Online]. South China Morning Post. Available: https://www.scmp.com/tech/big-tech/article/3176036/lockdowns-and-supply-chain-disruption-accelerate-apples-move-away?utm_source=copy_link&utm_medium=share_widget&utm_campaign=3176036 [Accessed 4/5 2022].

CHARLES, L. C. 2004. The Tableau Économique as Rational Recreation. *History of Political Economy,* 36**,** 445-474.

COFFMAN, D. M. 2021. François Quesnay, Luigi Pasinetti and the historical contexts of economic theory. *Structural Change and Economic Dynamics,* 56**,** 64-73.

COLUZZI, B., GHIL, M., HALLEGATTE, S. & WEISBUCH, G. 2011. Boolean delay equations on networks in economics and the geosciences. *International Journal of Bifurcation and Chaos,* 21**,** 3511-3548.

COPE, B. & KALANTZIS, M. 2022. The cybernetics of learning. *Educational Philosophy and Theory,* 54**,** 2352-2388.

CROWN, W. H. 2019. Real-World Evidence, Causal Inference, and Machine Learning. *Value in Health,* 22**,** 587-592.

CRUZ JR, J. B., TAN, R. R., CULABA, A. B. & BALLACILLO, J.-A. 2009. A dynamic input–output model for nascent bioenergy supply chains. *Applied Energy,* 86**,** S86-S94.

DAVIS, S. J. & CALDEIRA, K. 2010. Consumption-based accounting of CO<sub>2</sub> emissions. *Proceedings of the National Academy of Sciences,* 107**,** 5687-5692.

DAVIS, S. J., PETERS, G. P. & CALDEIRA, K. 2011. The supply chain of CO<sub>2</sub> emissions. *Proceedings of the National Academy of Sciences,* 108**,** 18554-18559.

DE HAAN, M. 2001. A Structural Decomposition Analysis of Pollution in the Netherlands. *Economic Systems Research,* 13**,** 181-196.

DE MAURO, A., GRECO, M. & GRIMALDI, M. 2016. A formal definition of Big Data based on its essential features. *Library Review,* 65**,** 122-135.

DE MESNARD, L. 2007. About the Ghosh model: clarifications. *LEG, Economy Series, Working Paper*.

DEBREU, G. 1974. Excess demand functions. *Journal of Mathematical Economics,* 1**,** 15-21.

DHANA LAKSHMI, D. & SATYANARAYANA, A. N. V. 2019. Influence of atmospheric rivers in the occurrence of devastating flood associated with extreme precipitation events over Chennai using different reanalysis data sets. *Atmospheric Research,* 215**,** 12-36.

DIETZENBACHER, E., GILJUM, S., HUBACEK, K. & SUH, S. 2009. Physical Input-Output Analysis and Disposals to Nature. *In:* SUH, S. (ed.) *Handbook of Input-Output Economics in Industrial Ecology.* Dordrecht: Springer Netherlands.

DIETZENBACHER, E., VAN BURKEN, B. & KONDO, Y. 2019. Hypothetical extractions from a global perspective. *Economic Systems Research,* 31**,** 505-519.

DIXON, P. & RIMMER, M. T. 2009. *Forecasting with a CGE model: Does it Work?*, Centre of Policy Studies (CoPS).

DONG, B., XU, Y. & LI, Q. 2022. Carbon transfer under China's inter-provincial trade: Evaluation and driving factors. *Sustainable Production and Consumption,* 32**,** 378-392.

DONG, X. S., QIAN, L. J., HUANG, L. & IEEE 2017. A CNN Based Bagging Learning Approach to Short-Term Load Forecasting in Smart Grid. *2017 Ieee Smartworld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (Smartworld/Scalcom/Uic/Atc/Cbdcom/Iop/Sci)*.

DORASZELSKI, U. & PAKES, A. 2007. Chapter 30 A Framework for Applied Dynamic Analysis in IO. *In:* ARMSTRONG, M. & PORTER, R. (eds.)

*Handbook of Industrial Organization.* Elsevier.
DORFMAN, R. 1979. A Formula for the Gini Coefficient. *The Review of Economics and Statistics,* 61**,** 146-149.
DUAN, C., CHEN, B., FENG, K., LIU, Z., HAYAT, T., ALSAEDI, A. & AHMAD, B. 2018. Interregional carbon flows of China. *Applied Energy,* 227**,** 342-352.
DUAN, Y., ZHAO, Y. & DE HAAN, J. 2020. Exchange Rate Pass-through in China: A Cost-Push Input-Output Price Model. *Open Economies Review,* 31**,** 513-528.
DUCHIN, F. & SZYLD, D. B. 1985. A DYNAMIC INPUT-OUTPUT MODEL WITH ASSURED POSITIVE OUTPUT(*). *Metroeconomica,* 37**,** 269-282.
DUCHIN, F. & SZYLD, D. B. 2006. A DYNAMIC INPUT-OUTPUT MODEL WITH ASSURED POSITIVE OUTPUT(*). *Metroeconomica,* 37**,** 269-282.
EINAV, L. & LEVIN, J. 2014. Economics in the age of big data. *Science,* 346.
EVALUATION & APPLIED RESEARCH DEPARTMENT 2020. Tamil Nadu An Economic Appraisal 2014 -15 to 2017-18. *In:* NADU, G. O. T. (ed.).
EVANS, G. H., JR. 1967. The Law of Demand — The Roles of Gregory King and Charles Davenant. *The Quarterly Journal of Economics,* 81**,** 483-492.
FAN, J.-L., WANG, Q., YU, S., HOU, Y.-B. & WEI, Y.-M. 2017. The evolution of CO2 emissions in international trade for major economies: a perspective from the global supply chain. *Mitigation and Adaptation Strategies for Global Change,* 22**,** 1229-1248.
FERNáNDEZ-AMADOR, O., FRANCOIS, J. F. & TOMBERGER, P. 2016. Carbon dioxide emissions and international trade at the turn of the millennium. *Ecological Economics,* 125**,** 14-26.
FEZZI, C. & FANGHELLA, V. 2020. Real-time estimation of the short-run impact of COVID-19 on economic activity using electricity market data. *Environmental and Resource Economics,* 76**,** 885-900.
FORBES. 2018. *How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read* [Online]. Available: https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/?sh=7a011660ba99 [Accessed].
FOSTER, J. 2006. Why is economics not a complex systems science? *Journal of Economic Issues,* 40**,** 1069-1091.
FRISCH, R. 1970. From utopian theory to practical applications: The case of econometrics. *Economic Planning Studies.* Springer.
FU, X. & CHEN, X. 2009. A multiyear lags input-holding-output model on education with excluding idle capital. *Journal of Systems Science and Complexity,* 22**,** 159-170.
GALBUSERA, L. & GIANNOPOULOS, G. 2018. On input-output economic models in disaster impact assessment. *International journal of disaster*

*risk reduction,* 30, 186-198.

GAWANDE, K., BERRENS, R. P. & BOHARA, A. K. 2001. A consumption-based theory of the environmental Kuznets curve. *Ecological Economics,* 37, 101-112.

GEANAKOPLOS, J. 1992. Common Knowledge. *Journal of Economic Perspectives,* 6, 53-82.

GEHRKE, C. & KURZ, H. D. 2002. Karl Marx on physiocracy. *Understanding'Classical'Economics.* Routledge.

GOSENS, J. 2020. The greening of South-South trade: Levels, growth, and specialization of trade in clean energy technologies between countries in the global South. *Renewable Energy,* 160, 931-943.

GOVERNMENT OF INNER MONGOLIA 2021. The Plan of Inner Mongolia Autonomous Region to Tackle Climate Change in the 14th Five-Year-Plan.

GROSSMAN, H. I. 1971. Money, Interest, and Prices in Market Disequilibrium. *Journal of Political Economy,* 79, 943-961.

GRUBB, M., DRUMMOND, P., PONCIA, A., MCDOWALL, W., POPP, D., SAMADI, S., PENASCO, C., GILLINGHAM, K. T., SMULDERS, S., GLACHANT, M., HASSALL, G., MIZUNO, E., RUBIN, E. S., DECHEZLEPRêTRE, A. & PAVAN, G. 2021. Induced innovation in energy technologies and systems: a review of evidence and potential implications for CO<sub>2</sub> mitigation. *Environmental Research Letters,* 16, 043007.

GUAN, D., WANG, D., HALLEGATTE, S., DAVIS, S. J., HUO, J., LI, S., BAI, Y., LEI, T., XUE, Q., COFFMAN, D. M., CHENG, D., CHEN, P., LIANG, X., XU, B., LU, X., WANG, S., HUBACEK, K. & GONG, P. 2020. Global supply-chain effects of COVID-19 control measures. *Nature Human Behaviour,* 4, 577-587.

GURGUL, H. & LACH, Ł. 2018. On using dynamic IO models with layers of techniques to measure value added in global value chains. *Structural Change and Economic Dynamics,* 47, 155-170.

HAILEMARIAM, A., DZHUMASHEV, R. & SHAHBAZ, M. 2020. Carbon emissions, income inequality and economic development. *Empirical Economics,* 59, 1139-1159.

HAIMES, Y. Y. & JIANG, P. 2001. Leontief-Based Model of Risk in Complex Interconnected Infrastructures. *Journal of Infrastructure Systems,* 7, 1-12.

HALLEGATTE, S. 2008. An Adaptive Regional Input-Output Model and its Application to the Assessment of the Economic Cost of Katrina. *Risk Analysis,* 28, 779-799.

HAMERMESH, D. S. 2013. Six Decades of Top Economics Publishing: Who and How? *Journal of Economic Literature,* 51, 162-172.

HANSEN, L. P. & SARGENT, T. J. 2016. *Robustness*, Princeton University

Press.
HARARI, Y. N. 2016. *Homo Deus: A brief history of tomorrow*, random house.
HARDING, M. C. & LAMARCHE, C. 2021. Small Steps with Big Data: Using Machine Learning in Energy and Environmental Economics. *Annual Review of Resource Economics,* 13**,** 469-488.
HARUN, M., MAT, S. H. C., FADZIM, W. R., KHAN, S. J. M. & NOOR, M. S. Z. 2018. The effects of fuel subsidy removal on input costs of productions: Leontief input-output price model. *International Journal of Supply Chain Management,* 7**,** 529-534.
HASTIE, T., TIBSHIRANI, R. & FRIEDMAN, J. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*.
HE, K. & HERTWICH, E. G. 2019. The flow of embodied carbon through the economies of China, the European Union, and the United States. *Resources, Conservation and Recycling,* 145**,** 190-198.
HE, K. & MI, Z. 2022. Carbon Implications of COVID-19. *Science-Policy Brief for the Multistakeholder Forum on Science, Technology and Innovation for the SDG.* UN.
HE, K., MI, Z., CHEN, L., COFFMAN, D. M. & LIANG, S. 2021. Critical transmission sectors in embodied atmospheric mercury emission network in China. *Journal of Industrial Ecology,* 25**,** 1644-1656.
HE, K., MI, Z., COFFMAN, D. M. & GUAN, D. 2022. Using a linear regression approach to sequential interindustry model for time-lagged economic impact analysis. *Structural Change and Economic Dynamics*.
HOEKSTRA, R., MICHEL, B. & SUH, S. 2016. The emission cost of international sourcing: using structural decomposition analysis to calculate the contribution of international sourcing to CO2-emission growth. *Economic Systems Research,* 28**,** 151-167.
HOLý, V. & ŠAFR, K. 2023. Disaggregating input–output tables by the multidimensional RAS method: a case study of the Czech Republic. *Economic Systems Research,* 35**,** 95-117.
HOPPIT, J. 1996. Political Arithmetic in Eighteenth-Century England. *The Economic History Review,* 49**,** 516-540.
HU, Y. & WU, W. 2021. Spatial-Temporal Variation and Driving Factors of Embodied Carbon in China-G7 Trade. *Available at SSRN 3982324*.
HUANG, L. & ZHAO, X. 2018. Impact of financial development on trade-embodied carbon dioxide emissions: Evidence from 30 provinces in China. *Journal of Cleaner Production,* 198**,** 721-736.
HUANG, Q., ZHENG, H., LI, J., MENG, J., LIU, Y., WANG, Z., ZHANG, N., LI, Y. & GUAN, D. 2021. Heterogeneity of consumption-based carbon emissions and driving forces in Indian states. *Advances in Applied Energy,* 4**,** 100039.
HUBACEK, K., BAIOCCHI, G., FENG, K., MUñOZ CASTILLO, R., SUN, L. & XUE, J. 2017. Global carbon inequality. *Energy, Ecology and*

*Environment,* 2**,** 361-369.
HUBACEK, K., CHEN, X., FENG, K., WIEDMANN, T. & SHAN, Y. 2021. Evidence of decoupling consumption-based CO2 emissions from economic growth. *Advances in Applied Energy,* 4**,** 100074.
INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. 2019. *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories* [Online]. Available: https://www.ipcc.ch/report/2019-refinement-to-the-2006-ipcc-guidelines-for-national-greenhouse-gas-inventories/ [Accessed].
IPCC, C. C. 2007. The physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,* 996**,** 113-119.
IRENA 2020. Renewable capacity statistics 2020. *International renewable energy agency*.
ISARD, W. 1966. *Methods of regional analysis*, Рипол Классик.
ISRAILEVICH, P. R., HEWINGS, G. J. D., SONIS, M. & SCHINDLER, G. R. 1997. Forecasting Structural Change With a Regional Econometric Input-Output Model. *Journal of Regional Science,* 37**,** 565-590.
IVANOVA, D., STADLER, K., STEEN-OLSEN, K., WOOD, R., VITA, G., TUKKER, A. & HERTWICH, E. G. 2016. Environmental Impact Assessment of Household Consumption. *Journal of Industrial Ecology,* 20**,** 526-536.
JACKSON, R. & MURRAY, A. 2004. Alternative Input-Output Matrix Updating Formulations. *Economic Systems Research,* 16**,** 135-148.
JANZEN, B. & RADULESCU, D. 2020. Electricity Use as a Real-Time Indicator of the Economic Burden of the COVID-19-Related Lockdown: Evidence from Switzerland. *CESifo Economic Studies,* 66**,** 303-321.
JOBIN, P., , S. & , S. 2018. Lived experience of the disaster victims of South Indian floods 2015: A concatenated disaster crisis model using phenomenographical framework analysis. *Indian Journal of Continuing Nursing Education,* 19**,** 44-53.
JOHANSEN, L. 1978. On the theory of dynamic input-output models with different time profiles of capital construction and finite life-time of capital equipment. *Journal of Economic Theory,* 19**,** 513-533.
JOHNSON, T. A continuous Leontief dynamic input-output model. Papers of the Regional Science Association, 1985. Springer, 177-188.
JOHNSY, A. C. & SCHIRINZI, G. Revisiting the South Indian floods of 2015 with Sentinel-1 data. 2019 International Conference on Data Science and Communication (IconDSC), 1-2 March 2019 2019. 1-4.
JONGDEEPAISAL, C. & NASU, S. 2020. Hybrid input-output analysis to evaluate economic impacts of biomass energy. *International Journal of Green Energy,* 17**,** 937-945.

KANITKAR, T., BANERJEE, R. & JAYARAMAN, T. 2015. Impact of economic structure on mitigation targets for developing countries. *Energy for Sustainable Development,* 26**,** 56-61.

KASAHARA, S. 2004. The Flying Geese Paradigm: A critical study of its application to East Asian regional development.

KIM, Y. S. 2015. Electricity consumption and economic development: are countries converging to a common trend? *Energy Economics,* 49**,** 192-202.

KOKS, E. E., CARRERA, L., JONKEREN, O., AERTS, J. C. J. H., HUSBY, T. G., THISSEN, M., STANDARDI, G. & MYSIAK, J. 2015. Regional disaster impact analysis: comparing Input-Output and Computable General Equilibrium models. Copernicus GmbH.

KURZ, H. D. & SALVADORI, N. 2000. The Dynamic Leontief Model and the Theory of Endogenous Growth. *Economic Systems Research,* 12**,** 255-265.

LANGE, O. 2014. *Introduction to economic cybernetics*, Elsevier.

LEI, M., DING, Q., CAI, W. & WANG, C. 2022. The exploration of joint carbon mitigation actions between demand- and supply-side for specific household consumption behaviors — A case study in China. *Applied Energy,* 324**,** 119740.

LENZEN, M., WOOD, R. & WIEDMANN, T. 2010. Uncertainty analysis for multi-region input–output models–a case study of the UK's carbon footprint. *Economic Systems Research,* 22**,** 43-63.

LEONTIEF, W. 1970. Environmental Repercussions and the Economic Structure: An Input-Output Approach. *The Review of Economics and Statistics,* 52**,** 262-271.

LEONTIEF, W. 1991. The economy as a circular flow. *Structural Change and Economic Dynamics,* 2**,** 181-212.

LEONTIEF, W. W. 1936. Quantitative input and output relations in the economic systems of the United States. *The review of economic statistics***,** 105-125.

LEONTIEF, W. W. 1953. Studies in the Structure of the American Economy.

LEVINE, S. H., GLORIA, T. P. & ROMANOFF, E. 2007. A dynamic model for determining the temporal distribution of environmental burden. *Journal of Industrial Ecology,* 11**,** 39-49.

LIAN, C. & HAIMES, Y. Y. 2006. Managing the risk of terrorism to interdependent infrastructure systems through the dynamic inoperability input–output model. *Systems Engineering,* 9**,** 241-258.

LIANG, S., WANG, C. & ZHANG, T. 2010. An improved input–output model for energy analysis: A case study of Suzhou. *Ecological Economics,* 69**,** 1805-1813.

LIN, X. & POLENSKE, K. R. 1998. Input—output modeling of production processes for business management. *Structural Change and Economic*

*Dynamics,* 9**,** 205-226.
LIU, Y., YANG, D. & Lü, G. 2017. Water Scarcity and Generalized Water Resources Management in Urumqi: Contributions from Noncompetitive Water Input-Output Model. *Journal of Water Resources Planning and Management,* 143**,** 04017038.
LIVESEY, D. A. 1971. Control theory and input-output analysis. *International Journal of Systems Science,* 2**,** 307-318.
LO, A. W. 1988. Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data. *Econometric Theory,* 4**,** 231-247.
LóPEZ PROL, J. & O, S. 2020. Impact of COVID-19 Measures on Short-Term Electricity Consumption in the Most Affected EU Countries and USA States. *iScience,* 23**,** 101639.
LOS, B. 2001. Endogenous Growth and Structural Change in a Dynamic Input-Output Model. *Economic Systems Research,* 13**,** 3-34.
LOU, F. 2016. *Big Data Economics and China Social-Economic Complex System Dynamic CGE Construction and Applications. 大数据经济学与中国经济社会复杂系统动态 CGE 模型构建及应用*, China Social Science Press 中国社会科学出版社.
LU, Q., FANG, K., HEIJUNGS, R., FENG, K., LI, J., WEN, Q., LI, Y. & HUANG, X. 2020. Imbalance and drivers of carbon emissions embodied in trade along the Belt and Road Initiative. *Applied Energy,* 280**,** 115934.
LUDWIG, N., FEUERRIEGEL, S. & NEUMANN, D. 2015. Putting Big Data analytics to work: Feature selection for forecasting electricity prices using the LASSO and random forests. *Journal of Decision Systems,* 24**,** 19-36.
MACAUSLAND, R. 2014. The Moore-Penrose inverse and least squares. *Math 420: Advanced Topics in Linear Algebra***,** 1-10.
MALIK, A. & LAN, J. 2016. The role of outsourcing in driving global carbon emissions. *Economic Systems Research,* 28**,** 168-182.
MALIK, A., LENZEN, M., ELY, R. N. & DIETZENBACHER, E. 2014. Simulating the impact of new industries on the economy: the case of biorefining in Australia. *Ecological Economics,* 107**,** 84-93.
MĂNESCU, M. 1980. *Economic Cybernetics*, Abacus Press.
MARCEL ALERS & BEN-JAMIN JONES 2021. A Guide to Carbon Pricing and Fossil Fuel Subsidy Reform. New York: UNDP.
MARIASELVAM, S. & GOPICHANDRAN, V. 2016. The Chennai floods of 2015: urgent need for ethical disaster management guidelines. *Indian Journal of Medical Ethics,* 1.
MELLANDER, C., LOBO, J., STOLARICK, K. & MATHESON, Z. 2015. Night-Time Light Data: A Good Proxy Measure for Economic Activity? *PLOS ONE,* 10**,** e0139779.
MENDOZA-TINOCO, D., GUAN, D., ZENG, Z., XIA, Y. & SERRANO, A. 2017. Flood footprint of the 2007 floods in the UK: The case of the Yorkshire

and The Humber region. *Journal of Cleaner Production,* 168**,** 655-667.

MENDOZA-TINOCO, D., HU, Y., ZENG, Z., CHALVATZIS, K. J., ZHANG, N., STEENGE, A. E. & GUAN, D. 2020. Flood Footprint Assessment: A Multiregional Case of 2009 Central European Floods. *Risk Analysis,* 40**,** 1612-1631.

MENG, J., MI, Z., GUAN, D., LI, J., TAO, S., LI, Y., FENG, K., LIU, J., LIU, Z., WANG, X., ZHANG, Q. & DAVIS, S. J. 2018. The rise of South–South trade and its effect on global CO2 emissions. *Nature Communications,* 9**,** 1871.

MI, Z., MENG, J., GREEN, F., COFFMAN, D. M. & GUAN, D. 2018. China's "Exported Carbon" Peak: Patterns, Drivers, and Implications. *Geophysical Research Letters,* 45**,** 4309-4318.

MI, Z., MENG, J., GUAN, D., SHAN, Y., SONG, M., WEI, Y.-M., LIU, Z. & HUBACEK, K. 2017. Chinese CO2 emission flows have reversed since the global financial crisis. *Nature Communications,* 8.

MI, Z., ZHENG, J., GREEN, F., GUAN, D., MENG, J., FENG, K., LIANG, X. & WANG, S. 2021. Decoupling without outsourcing? How China's consumption-based CO2 emissions have plateaued. *iScience,* 24**,** 103130.

MI, Z., ZHENG, J., MENG, J., OU, J., HUBACEK, K., LIU, Z., COFFMAN, D. M., STERN, N., LIANG, S. & WEI, Y.-M. 2020. Economic development and converging household carbon footprints in China. *Nature Sustainability,* 3**,** 529-537.

MILLER, R. E. & BLAIR, P. D. 2009. *Input-output analysis: foundations and extensions*, Cambridge university press.

MISCH, F. & WINGENDER, M. P. 2021. *Revisiting carbon leakage*, International Monetary Fund.

MU, T., XIA, Q. & KANG, C. 2010. Input-output table of electricity demand and its application. *Energy,* 35**,** 326-331.

MULLAINATHAN, S. & SPIESS, J. 2017. Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives,* 31**,** 87-106.

MUNROE, D. K. & BILES, J. J. 2005. Regional Science. *In:* KEMPF-LEONARD, K. (ed.) *Encyclopedia of Social Measurement.* New York: Elsevier.

MURPHY, R. P. 2011. The Chicago School versus the Austrian School. *Mises Daily, June,* 20.

NAIMUR RAHMAN, M., ESMAILPOUR, A. & ZHAO, J. 2016. Machine Learning with Big Data An Efficient Electricity Generation Forecasting System. *Big Data Research,* 5**,** 9-15.

NAKAMURA, S. & NAKAJIMA, K. 2005. Waste input–output material flow analysis of metals in the Japanese economy. *Materials transactions,* 46**,** 2550-2553.

NDRC. 2021. *Announcement by the National Development and Reform Commission of The People's Republic of China, No.1, 2021* [Online].

Available: https://www.ndrc.gov.cn/xxgk/zcfb/gg/202102/t20210207_1267081.html?code=&state=123 [Accessed].

NETWORK OF EMERGENCY PHYSICIANS. 2007. *Tsunami & Its Aftermath* [Online]. Available: http://www.emergencymedicine.in/nepi/articles.php?article_id=16 [Accessed].

NEWELL, A. & SIMON, H. 1956. The logic theory machine--A complex information processing system. *IRE Transactions on Information Theory,* 2**,** 61-79.

NING, Y., MIAO, L., DING, T. & ZHANG, B. 2019. Carbon emission spillover and feedback effects in China based on a multiregional input-output model. *Resources, Conservation and Recycling,* 141**,** 211-218.

NOVAN, K., SMITH, A. & ZHOU, T. 2020. Residential Building Codes Do Save Energy: Evidence from Hourly Smart-Meter Data. *The Review of Economics and Statistics***,** 1-45.

OECD. 2022. *Industrial production (indicator)* [Online]. [Accessed 01/11 2022].

OKUYAMA, Y. 2004a. Modeling spatial economic impacts of an earthquake: input-output approaches. *Disaster Prevention and Management; Bradford,* 13**,** 297-306.

OKUYAMA, Y. 2004b. Modeling spatial economic impacts of an earthquake: input-output approaches. *Disaster Prevention and Management: An International Journal,* 13**,** 297-306.

OKUYAMA, Y., HEWINGS, G. J. & SONIS, M. Sequential interindustry model (SIM) and impact analysis: application for measuring economic impact of unscheduled events. 47th north American meetings of the regional science association international, Chicago, IL, 2000.

OKUYAMA, Y., HEWINGS, G. J. D. & SONIS, M. 2004. Measuring Economic Impacts of Disasters: Interregional Input-Output Analysis Using Sequential Interindustry Model. *In:* OKUYAMA, Y. & CHANG, S. E. (eds.) *Modeling Spatial and Economic Impacts of Disasters.* Berlin, Heidelberg: Springer Berlin Heidelberg.

OKUYAMA, Y., SONIS, M. & HEWINGS, G. J. 2006. Typology of structural change in a regional economy: a temporal inverse analysis. *Economic Systems Research,* 18**,** 133-153.

OKUYAMA, Y. & YU, K. D. 2019. Return of the inoperability. *Economic Systems Research,* 31**,** 467-480.

OOSTERHAVEN, J. 1988. On the plausibility of the supply-driven input-output model. *Journal of Regional Science,* 28**,** 203-217.

PAGSUYOIN, S., SANTOS, J., SALCEDO, G. & YIP, C. 2019. Spatio-Temporal Drought Risk Analysis Using GIS-Based Input Output Modeling. *In:* OKUYAMA, Y. & ROSE, A. (eds.) *Advances in Spatial and Economic Modeling of Disaster Impacts.* Cham: Springer International Publishing.

PASINETTI, L. L. 2009. Il Tableau Économique e le economie moderne. *In:* DE VIVO, G. (ed.) *Il Tableau Économique di François Quesnay.* Milano: Fondazione Raffaele Mattioli.

PECK, J., VAN HORN, R., MIROWSKI, P. & STAPLEFORD, T. 2011. Orientation: in search of the Chicago School. *Building Chicago Economics: New Perspectives on the History of America's Most Powerful Economics Program*.

PEREZ-CHACON, R., ASENCIO-CORTES, G., MARTINEZ-ALVAREZ, F. & TRONCOSO, A. 2020. Big data time series forecasting based on pattern sequence similarity and its application to the electricity demand. *Information Sciences,* 540**,** 160-174.

PETERS GLEN, P., MINX JAN, C., WEBER CHRISTOPHER, L. & EDENHOFER, O. 2011. Growth in emission transfers via international trade from 1990 to 2008. *Proceedings of the National Academy of Sciences,* 108**,** 8903-8908.

PETERS, G. P. 2008. From production-based to consumption-based national emission inventories. *Ecological Economics,* 65**,** 13-23.

POLO, C. & VIEJO, R. 2015. On the accuracy of cge forecasts in expansion and recession: spain 1990–1997. *Economic Systems Research,* 27**,** 525-542.

PRZYBYLIŃSKI, M. & GORZAŁCZYŃSKI, A. 2022. Applying the input–output price model to identify inflation processes. *Journal of Economic Structures,* 11**,** 5.

QIAN, X., YU, J. & DAI, H. 1993. A new discipline of science—The study of open complex giant system and its methodology. *Chinese Journal of Systems Engineering and Electronics,* 4**,** 2-12.

QU, H. N., LING, P. & WU, L. B. 2015. Electricity Consumption Analysis and Applications based on Smart Grid Big Data. *Ieee 12th Int Conf Ubiquitous Intelligence & Comp/Ieee 12th Int Conf Adv & Trusted Comp/Ieee 15th Int Conf Scalable Comp & Commun/Ieee Int Conf Cloud & Big Data Comp/Ieee Int Conf Internet People and Associated Symposia/Workshops***,** 923-928.

QU, S., WANG, H., LIANG, S., SHAPIRO, A. M., SUH, S., SHELDON, S., ZIK, O., FANG, H. & XU, M. 2017. A Quasi-Input-Output model to improve the estimation of emission factors for purchased electricity from interconnected grids. *Applied Energy,* 200**,** 249-259.

RAHMAN, M. N., ESMAILPOUR, A. & ZHAO, J. H. 2016. Machine Learning with Big Data An Efficient Electricity Generation Forecasting System. *Big Data Research,* 5**,** 9-15.

RAJAN, V. & SRIDHARAN, D. 2016. A Case Study on Impact of Chennai Floods: Supply Chain Perspective. *Industrial Engineering Journal,* IX**,** 12-16.

RAO, N. D. & MIN, J. 2018. Less global inequality can improve climate

outcomes. *WIREs Climate Change,* 9**,** e513.
REYES, F. A. & MENDOZA, M. A. M. 2021. Demand-Driven and Supply-Sided Input–Output Models. *Journal of Quantitative Economics***,** 1-17.
ROCCO, M. V. 2019. Integrating Energy and Economy models based on the Dynamic Input-Output framework. *IIOA.* Glasgow, United Kingdom.
ROMANELLO, M., DI NAPOLI, C., DRUMMOND, P., GREEN, C., KENNARD, H., LAMPARD, P., SCAMMAN, D., ARNELL, N., AYEB-KARLSSON, S. & FORD, L. B. 2022. The 2022 report of the Lancet Countdown on health and climate change: health at the mercy of fossil fuels. *The Lancet,* 400**,** 1619-1654.
ROMANELLO, M., MCGUSHIN, A., DI NAPOLI, C., DRUMMOND, P., HUGHES, N., JAMART, L., KENNARD, H., LAMPARD, P., RODRIGUEZ, B. S. & ARNELL, N. 2021. The 2021 report of the Lancet Countdown on health and climate change: code red for a healthy future. *The Lancet,* 398**,** 1619-1662.
ROMANOFF, E. & LEVINE, S. H. 1977. Interregional sequential interindustry modeling: a preliminary analysis of regional growth and decline in a two region case. *Northeast Regional Science Review,* 7**,** 87-101.
ROMANOFF, E. & LEVINE, S. H. 1981. Anticipatory and Responsive Sequential Interindustry Models. *IEEE Transactions on Systems, Man, and Cybernetics,* 11**,** 181-186.
ROMANOFF, E. & LEVINE, S. H. 1986. CAPACITY LIMITATIONS, INVENTORY, AND TIME-PHASED PRODUCTION IN THE SEQUENTIAL INTERINDUSTRY MODEL. *Papers in Regional Science,* 59**,** 73-91.
ROSE, A. 1995. Input-output economics and computable general equilibrium models. *Structural Change and Economic Dynamics,* 6**,** 295-304.
ROSE, A. & CASLER, S. 1996. Input–Output Structural Decomposition Analysis: A Critical Appraisal. *Economic Systems Research,* 8**,** 33-62.
ROSEN, S. 1997. Austrian and neoclassical economics: any gains from trade? *Journal of Economic Perspectives,* 11**,** 139-152.
RYABOSHLYK, V. 2006. A dynamic input–output model with explicit new and old technologies: an application to the UK. *Economic Systems Research,* 18**,** 183-203.
SALIN, P. 2023. The Monetary Economics of the Austrian School and the Chicago School. *Quarterly Journal of Austrian Economics,* 25**,** 281-296.
SAMOTHRAKIS, S. 2021. Artificial Intelligence inspired methods for the allocation of common goods and services. *PLOS ONE,* 16**,** e0257399.
SAMRAJ, C. J. 2017. Unequal Society and Equalising Calamity: Examining Popular Discourse on Impact of Floods in Tamil Nadu. *SSRN Electronic Journal*.
SANTOS, J. R. & HAIMES, Y. Y. 2004. Modeling the Demand Reduction Input-Output (I-O) Inoperability Due to Terrorism of Interconnected

Infrastructures*. *Risk Analysis,* 24**,** 1437-1451.
ŠĆEPANOVIĆ, S., MISHKOVSKI, I., HUI, P., NURMINEN, J. K. & YLä-JääSKI, A. 2015. Mobile Phone Call Data as a Regional Socio-Economic Proxy Indicator. *PLOS ONE,* 10**,** e0124160.
SHAN, Y., HUANG, Q., GUAN, D. & HUBACEK, K. 2020. China CO2 emission accounts 2016–2017. *Scientific Data,* 7**,** 54.
SHARRARD, A. L., MATTHEWS, H. S. & RIES, R. J. 2008. Estimating Construction Project Environmental Effects Using an Input-Output-Based Hybrid Life-Cycle Assessment Model. *Journal of Infrastructure Systems,* 14**,** 327-336.
SIMON, H. A. 1997. *Administrative Behavior, 4th Edition*, Free Press.
SKOLKA, J. 1989. Input-output structural decomposition analysis for Austria. *Journal of Policy Modeling,* 11**,** 45-66.
STADLER, K., WOOD, R., BULAVSKAYA, T., SöDERSTEN, C.-J., SIMAS, M., SCHMIDT, S., USUBIAGA, A., ACOSTA-FERNáNDEZ, J., KUENEN, J., BRUCKNER, M., GILJUM, S., LUTTER, S., MERCIAI, S., SCHMIDT, J. H., THEURL, M. C., PLUTZAR, C., KASTNER, T., EISENMENGER, N., ERB, K.-H., KONING, A. D. & TUKKER, A. 2018a. EXIOBASE 3: Developing a Time Series of Detailed Environmentally Extended Multi-Regional Input-Output Tables. *Journal of Industrial Ecology,* 22**,** 502-515.
STADLER, K., WOOD, R., BULAVSKAYA, T., SöDERSTEN, C. J., SIMAS, M., SCHMIDT, S., USUBIAGA, A., ACOSTA-FERNáNDEZ, J., KUENEN, J. & BRUCKNER, M. 2018b. EXIOBASE 3: Developing a time series of detailed environmentally extended multi-regional input-output tables. *Journal of Industrial Ecology,* 22**,** 502-515.
STEENGE, A. E. & REYES, R. C. 2020. Return of the capital coefficients matrix. *Economic Systems Research***,** 1-12.
SU, B. & ANG, B. W. 2016. Multi-region comparisons of emission performance: The structural decomposition analysis approach. *Ecological Indicators,* 67**,** 78-87.
SU, B. & ANG, B. W. 2022. Improved granularity in input-output analysis of embodied energy and emissions: The use of monthly data. *Energy Economics,* 113**,** 106245.
SUN, M., CHEN, G., XU, X., ZHANG, L., HUBACEK, K. & WANG, Y. 2021. Reducing Carbon Footprint Inequality of Household Consumption in Rural Areas: Analysis from Five Representative Provinces in China. *Environmental Science & Technology,* 55**,** 11511-11520.
TAN, R. G. R., YU, K. D. S., AVISO, K. B. & PROMENTILLA, M. A. B. 2017. Input-output modeling approach to sustainable systems engineering. *Encyclopedia of Sustainable Technologies***,** 519.
TAN, R. R., AVISO, K. B., PROMENTILLA, M. A. B., YU, K. D. S., SANTOS, J. R., TAN, R. R., AVISO, K. B., PROMENTILLA, M. A. B., YU, K. D. S. & SANTOS, J. R. 2019. Introduction to Input–Output Models. *Input-Output*

*Models for Sustainable Industrial Systems: Implementation Using LINGO*, 1-8.

TEN RAA, T. 1986. Dynamic Input-Output Analysis with Distributed Activities. *The Review of Economics and Statistics,* 68, 300.

TORMO GARCíA, L., BELDA, P. R., MORILLAS JURADO, F. & CABRER-BORRáS, B. A new approach to the hypothetical extraction method: regional full extraction. *Economic Systems Research*, 1-27.

TUZLUKOV, V. 2018. *Signal processing noise*, CRC Press.

UNITED NATIONS STATISTICAL DIVISION 1999. *Handbook of input-output table compilation and analysis*, UN.

VARIAN, H. R. 2014. Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives,* 28, 3-28.

VENCATESAN, A. 2021. From Rains to Floods: A Case of Chennai in 2015. *Environment & Society*

VILORIA, A., SIERRA, D. M., CAMARGO, J. F., ZEA, K. B., FUENTES, J. P., HERNANDEZ-PALMA, H. & KAMATKAR, S. J. 2020. Demand in the Electricity Market: Analysis Using Big Data. *Intelligent Computing, Information and Control Systems, Iciccs 2019,* 1039, 316-325.

WAN, G., WANG, C., WANG, J. & ZHANG, X. 2022. The income inequality-CO2 emissions nexus: Transmission mechanisms. *Ecological Economics,* 195, 107360.

WANG, H., PAN, C., WANG, Q. & ZHOU, P. 2020a. Assessing sustainability performance of global supply chains: An input-output modeling approach. *European journal of operational research,* 285, 393-404.

WANG, P., LI, H. & XU, J. 2023. Forecasting carbon emissions of China's industrial sectors via time lag effect. *Environment, Development and Sustainability*.

WANG, S., WANG, X. & TANG, Y. 2020b. Drivers of carbon emission transfer in China—An analysis of international trade from 2004 to 2011. *Science of The Total Environment,* 709, 135924.

WANG, Y., GESCHKE, A. & LENZEN, M. 2017. Constructing a Time Series of Nested Multiregion Input–Output Tables. *International Regional Science Review,* 40, 476-499.

WANG, Y., XIAO, H. L. & LU, M. F. 2009. Analysis of water consumption using a regional input–output model: Model development and application to Zhangye City, Northwestern China. *Journal of Arid Environments,* 73, 894-900.

WANG, Z. Y., YE, X. Y., LEE, J., CHANG, X. M., LIU, H. M. & LI, Q. Q. 2018. A spatial econometric modeling of online social interactions using microblogs. *Computers Environment and Urban Systems,* 70, 53-58.

WEST, D. K. 2020. Cybernetics for the command economy: Foregrounding entropy in late Soviet planning. *History of the Human Sciences,* 33, 36-51.

WIEDENHOFER, D., GUAN, D., LIU, Z., MENG, J., ZHANG, N. & WEI, Y.-M. 2017. Unequal household carbon footprints in China. *Nature Climate Change,* 7**,** 75-80.

WIEDMANN, T. 2009. A review of recent multi-region input–output models used for consumption-based emission and resource accounting. *Ecological economics,* 69**,** 211-222.

WIENER, N. 1961. *Cybernetics Or Control and Communication in the Animal and the Machine*, M.I.T. Press.

WOLFRAM, C., SHELEF, O. & GERTLER, P. 2012. How Will Energy Demand Develop in the Developing World? *Journal of Economic Perspectives,* 26**,** 119-138.

WOOD, R., GRUBB, M., ANGER-KRAAVI, A., POLLITT, H., RIZZO, B., ALEXANDRI, E., STADLER, K., MORAN, D., HERTWICH, E. & TUKKER, A. 2020. Beyond peak emission transfers: historical impacts of globalization and future impacts of climate policies on international emission transfers. *Climate Policy,* 20**,** S14-S27.

WU, J., CUI, C., MEI, X., XU, Q. & ZHANG, P. 2021. Migration of manufacturing industries and transfer of carbon emissions embodied in trade: empirical evidence from China and Thailand. *Environmental Science and Pollution Research***,** 1-13.

XU, M. & LIANG, S. 2019. Input–output networks offer new insights of economic structure. *Physica A: Statistical Mechanics and its Applications,* 527**,** 121178.

XU, X. & ANG, B. W. 2013. Index decomposition analysis applied to CO2 emission studies. *Ecological Economics,* 93**,** 313-329.

XU, Y. & DIETZENBACHER, E. 2014. A structural decomposition analysis of the emissions embodied in trade. *Ecological Economics,* 101**,** 10-20.

YAACOB, H., ALI, Q., SARBINI, N.-A., RANI, A.-N., ZAINI, Z., ALI, N.-N. & MAHALLE, N. 2021. Economic Shocks of Covid-19: Can Big Data Analytics Help Connect the Dots. *Intelligent Automation \& Soft Computing,* 27**,** 653--668.

YAGI, M., KAGAWA, S., MANAGI, S., FUJII, H. & GUAN, D. 2020. Supply Constraint from Earthquakes in Japan in Input–Output Analysis. *Risk analysis,* 40**,** 1811-1830.

YAN, J. & YANG, J. 2021. Carbon pricing and income inequality:A case study of Guangdong Province, China. *Journal of Cleaner Production,* 296**,** 126491.

YANG, Y., DONG, S., LI, F., CHENG, H., LI, Z., LI, Y. & LI, S. 2021. An analysis on the adoption of an interregional carbon emission reduction allocation approach in the context of China's interprovincial carbon emission transfer. *Environment, Development and Sustainability,* 23**,** 4385-4411.

YASEEN, Q. M., AKHTAR, R., KHALIL, M. K. U. & USMAN JAN, Q. M. 2020. Dynamic inoperability input-output modeling for economic losses estimation in industries during flooding. *Socio-Economic Planning Sciences,* 72**,** 100876.

YU, K. D. S., TAN, R. R. & SANTOS, J. R. Impact estimation of flooding in Manila: An inoperability input-output approach. 2013 IEEE Systems and Information Engineering Design Symposium, 2013. IEEE, 47-51.

YUAN, Y. H., TSAO, S. H., CHYOU, J. T. & TSAI, S. B. 2020. An empirical study on effects of electronic word-of-mouth and Internet risk avoidance on purchase intention: from the perspective of big data. *Soft Computing,* 24**,** 5713-5728.

ZENG, A. R., LIU, S. & YU, Y. 2019a. Comparative study of data driven methods in building electricity use prediction. *Energy and Buildings,* 194**,** 289-300.

ZENG, Z. & GUAN, D. 2020a. Methodology and application of flood footprint accounting in a hypothetical multiple two-flood event. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 378**,** 20190209.

ZENG, Z. & GUAN, D. 2020b. Methodology and application of flood footprint accounting in a hypothetical multiple two-flood event. *Philosophical Transactions of the Royal Society A,* 378**,** 20190209.

ZENG, Z., GUAN, D., STEENGE, A. E., XIA, Y. & MENDOZA-TINOCO, D. 2019b. Flood footprint assessment: a new approach for flood-induced indirect economic impact measurement and post-flood recovery. *Journal of Hydrology,* 579**,** 124204.

ZHANG, C., ZHOU, K., YANG, S. & SHAO, Z. 2017a. On electricity consumption and economic growth in China. *Renewable and Sustainable Energy Reviews,* 76**,** 353-368.

ZHANG, Q., JIANG, X., TONG, D., DAVIS, S. J., ZHAO, H., GENG, G., FENG, T., ZHENG, B., LU, Z., STREETS, D. G., NI, R., BRAUER, M., VAN DONKELAAR, A., MARTIN, R. V., HUO, H., LIU, Z., PAN, D., KAN, H., YAN, Y., LIN, J., HE, K. & GUAN, D. 2017b. Transboundary health impacts of transported global air pollution and international trade. *Nature,* 543**,** 705-709.

ZHANG, Y., LI, Y., HUBACEK, K., TIAN, X. & LU, Z. 2019. Analysis of CO2 transfer processes involved in global trade based on ecological network analysis. *Applied Energy,* 233-234**,** 576-583.

ZHAO, L., CHEN, Z., HU, Y., MIN, G. & JIANG, Z. 2018. Distributed Feature Selection for Efficient Economic Big Data Analysis. *IEEE Transactions on Big Data,* 4**,** 164-176.

ZHAO, Y., ZHANG, Z., WANG, S., ZHANG, Y. & LIU, Y. 2015. Linkage analysis of sectoral CO2 emissions based on the hypothetical extraction method in South Africa. *Journal of Cleaner Production,* 103**,** 916-924.

ZHENG, H., BAI, Y., WEI, W., MENG, J., ZHANG, Z., SONG, M. & GUAN, D. 2021. Chinese provincial multi-regional input-output database for 2012,

2015, and 2017. *Scientific Data,* 8**,** 244.
ZHENG, H., FANG, Q., WANG, C., JIANG, Y. & REN, R. 2018. Updating China's input-output tables series using MTT method and its comparison. *Economic Modelling,* 74**,** 186-193.
ZHENG, J., FENG, G., REN, Z., QI, N., COFFMAN, D. M., ZHOU, Y. & WANG, S. 2022. China's energy consumption and economic activity at the regional level. *Energy,* 259**,** 124948.
ZHONG, Z., JIANG, L. & ZHOU, P. 2018. Transnational transfer of carbon emissions embodied in trade: Characteristics and determinants from a spatial perspective. *Energy,* 147**,** 858-875.
ZHOU, D., ZHOU, X., XU, Q., WU, F., WANG, Q. & ZHA, D. 2018. Regional embodied carbon emissions and their transfer characteristics in China. *Structural Change and Economic Dynamics,* 46**,** 180-193.
ZHOU, K., YANG, C. & SHEN, J. 2017a. Discovering residential electricity consumption patterns through smart-meter data mining: A case study from China. *Utilities Policy,* 44**,** 73-84.
ZHOU, L. & CHEN, Z. 2020. Are CGE models reliable for disaster impact analyses? *Economic Systems Research***,** 1-27.
ZHOU, L. & CHEN, Z. 2021. Are CGE models reliable for disaster impact analyses? *Economic Systems Research,* 33**,** 20-46.
ZHOU, X., ZHANG, M., ZHOU, M. & ZHOU, M. 2017b. A comparative study on decoupling relationship and influence factors between China's regional economic development and industrial energy–related carbon emissions. *Journal of Cleaner Production,* 142**,** 783-800.

# Appendices

Table S1. The flow of carbon emissions embodied in the trade among China's provinces.

| Province | Region | 2007 | | | | 2012 | | | | 2017 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | CBA CO2 emission intensity (kg/yuan) | Final Consumption (trillion Yuan) | CBA CO2 (Mt) | CO2 net export (Mt) | CBA CO2 emission intensity (kg/yuan) | Final Consumption (trillion Yuan) | CBA CO2 (Mt) | CO2 net export (Mt) | CBA CO2 emission intensity (kg/yuan) | Final Consumption (trillion Yuan) | CBA CO2 (Mt) | CO2 net export (Mt) |
| Beijing | Beijing–Tianjin | 0.14 | 1011.10 | 140.72 | -75.51 | 0.08 | 1684.52 | 134.34 | -68.98 | 0.04 | 2468.53 | 95.41 | -27.04 |
| Tianjin | Beijing–Tianjin | 0.23 | 443.49 | 102.74 | -42.20 | 0.12 | 1423.99 | 167.38 | -48.11 | 0.09 | 1626.10 | 142.55 | -12.70 |
| Hebei | North | 0.24 | 991.39 | 236.22 | 128.18 | 0.17 | 2612.91 | 436.10 | 119.52 | 0.22 | 3068.88 | 681.08 | 4.90 |
| Shanxi | Central | 0.31 | 514.27 | 161.53 | 92.10 | 0.21 | 1357.85 | 279.07 | 101.85 | 0.31 | 1388.66 | 426.84 | 44.29 |
| Inner Mongolia | Northwest | 0.24 | 448.71 | 108.01 | 159.18 | 0.16 | 1713.50 | 280.64 | 191.08 | 0.33 | 1462.93 | 481.50 | 145.58 |
| Liaoning | Northeast | 0.24 | 866.60 | 204.97 | 54.98 | 0.13 | 2509.36 | 330.67 | 37.88 | 0.22 | 2110.59 | 458.58 | 2.94 |
| Jilin | Northeast | 0.30 | 596.29 | 179.26 | -45.69 | 0.16 | 1397.17 | 221.11 | -17.68 | 0.14 | 1499.25 | 211.47 | -13.58 |
| Heilongjiang | Northeast | 0.26 | 596.98 | 153.55 | -5.95 | 0.16 | 1527.04 | 248.12 | -25.86 | 0.13 | 1749.95 | 230.49 | 28.63 |
| Shanghai | Central Coast | 0.18 | 1090.13 | 192.79 | -105.72 | 0.08 | 1721.43 | 140.91 | -9.37 | 0.07 | 2628.55 | 182.71 | -5.28 |
| Jiangsu | Central Coast | 0.16 | 1885.69 | 294.81 | -11.45 | 0.10 | 4383.65 | 443.10 | -6.74 | 0.10 | 7199.65 | 754.06 | -36.68 |
| Zhejiang | Central Coast | 0.19 | 1628.96 | 305.64 | -117.87 | 0.10 | 2989.16 | 296.03 | -64.96 | 0.10 | 4133.27 | 403.58 | -39.17 |
| Anhui | Central | 0.20 | 717.07 | 143.59 | 1.08 | 0.11 | 1740.69 | 194.48 | 47.53 | 0.15 | 2389.38 | 364.86 | -9.33 |
| Fujian | South Coast | 0.15 | 758.38 | 113.51 | -19.96 | 0.10 | 1570.33 | 150.98 | 1.01 | 0.08 | 2687.65 | 225.37 | -0.85 |
| Jiangxi | Central | 0.26 | 578.59 | 148.80 | -55.38 | 0.11 | 1276.76 | 139.02 | -19.55 | 0.12 | 1769.35 | 218.44 | -2.83 |
| Shandong | North | 0.22 | 2079.36 | 449.32 | -4.32 | 0.15 | 4506.05 | 662.76 | -59.63 | 0.13 | 6405.16 | 805.59 | -31.26 |
| Henan | Central | 0.18 | 1303.70 | 229.89 | 79.65 | 0.12 | 3491.38 | 416.41 | -4.41 | 0.11 | 4664.91 | 496.45 | -23.11 |
| Hubei | Central | 0.20 | 882.58 | 172.17 | 14.74 | 0.14 | 2223.17 | 313.39 | -24.44 | 0.09 | 3303.09 | 298.91 | 6.53 |
| Hunan | Central | 0.19 | 843.42 | 159.72 | 4.57 | 0.11 | 2256.86 | 250.48 | -30.35 | 0.10 | 3122.67 | 296.83 | -5.79 |
| Guangdong | South Coast | 0.13 | 2306.41 | 297.22 | -78.55 | 0.08 | 4934.63 | 381.46 | -99.79 | 0.08 | 7447.30 | 596.81 | -95.61 |
| Guangxi | Southwest | 0.16 | 565.59 | 90.39 | -3.57 | 0.11 | 1579.41 | 176.00 | -22.47 | 0.12 | 1741.86 | 211.86 | 2.87 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Hainan | South Coast | 0.13 | 124.90 | 16.18 | -0.43 | 0.11 | 334.88 | 36.70 | -8.32 | 0.07 | 493.05 | 32.76 | 7.93 |
| Chongqing | Southwest | 0.21 | 437.59 | 92.36 | -16.21 | 0.14 | 1172.53 | 168.18 | -36.59 | 0.10 | 1718.51 | 171.57 | -23.16 |
| Sichuan | Southwest | 0.16 | 1019.53 | 167.69 | -12.57 | 0.11 | 2412.50 | 264.17 | -24.16 | 0.08 | 3288.96 | 275.74 | 10.26 |
| Guizhou | Southwest | 0.26 | 307.98 | 79.37 | 42.58 | 0.17 | 809.99 | 135.87 | 42.71 | 0.17 | 1452.80 | 241.59 | -3.45 |
| Yunan | Southwest | 0.18 | 483.49 | 88.07 | 12.47 | 0.12 | 1477.14 | 178.75 | -13.17 | 0.09 | 2203.69 | 191.49 | -7.47 |
| Tibet | Southwest | - | - | - | | - | - | - | | 0.03 | 194.38 | 6.10 | -1.28 |
| Shaanxi | Northwest | 0.23 | 476.41 | 107.58 | -1.90 | 0.14 | 1602.89 | 220.51 | -30.05 | 0.12 | 2079.60 | 244.06 | 6.35 |
| Gansu | Northwest | 0.22 | 279.89 | 61.40 | 7.94 | 0.15 | 684.14 | 102.75 | 20.09 | 0.18 | 751.97 | 133.88 | 7.51 |
| Qinghai | Northwest | 0.28 | 90.70 | 25.15 | -6.75 | 0.15 | 263.02 | 39.32 | -4.43 | 0.15 | 342.95 | 51.42 | -1.67 |
| Ningxia | Northwest | 0.35 | 118.85 | 41.63 | 9.98 | 0.27 | 261.38 | 70.95 | 44.42 | 0.28 | 569.53 | 162.22 | 11.03 |
| Xinjiang | Northwest | 0.27 | 347.30 | 94.27 | -3.43 | 0.19 | 1001.14 | 190.17 | 12.97 | 0.21 | 1559.74 | 328.03 | 61.43 |

Table S2. The flow of carbon emissions embodied in the trade among major world regions.

| | 2007 | | 2012 | | 2017 | |
|---|---|---|---|---|---|---|
| $CO_2$ Net Export (Mt) | to RoW | to China | to RoW | to China | to RoW | to China |
| Europe | -601 | -384 | -439 | -209 | -356 | -167 |
| North America | -301 | -514 | -221 | -345 | -429 | -325 |
| East Asia | 39 | -154 | -3 | -93 | 22 | -50 |
| China | 1,707 | - | 1,424 | - | 1,013 | - |
| BRICS w/o China | 599 | -79 | 457 | -79 | 513 | -26 |
| Rest of Asia Pacific | 168 | -328 | 102 | -426 | 108 | -244 |
| Latin America | -1 | -81 | -27 | -107 | 3 | -90 |
| Africa | -91 | -46 | -145 | -65 | -189 | -60 |
| Middle East | 189 | -121 | 277 | -101 | 329 | -50 |

Table S3. An illustration of the aggregation for electricity sectors of Food, Chemical & Mining, and Consumer Goods.

| Intermediate(1)/Final(2) | | | Food | Chemical & Mining | Consumer Goods | Heavy Industry | Manufacturing | EHGW | Construction | Service |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Agriculture, forestry, animal husbandry and fisheries | A000 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Agriculture | 0100 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Forestry | 0200 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Animal husbandry | 0300 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Fishery | 0400 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | Agriculture, forestry, animal husbandry, fisheries and ancillary activities | 0500 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | Among them: drainage irrigation | 05A0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Mining industry | B000 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Coal mining and washing industry | 0600 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Soot and anthracite mining | 0610 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Lignite mining wash | 0620 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Other coal mining | 0690 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Oil and gas extraction | 0700 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Oil extraction | 0710 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Natural gas extraction | 0720 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | Service activities related to oil and gas extraction | 0790 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Black metal mining industry | 0800 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Iron ore mining | 0810 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Manganese ore, chromium ore mining | 0820 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Other ferrous metal ore mining | 0890 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Non-ferrous metal mining industry | 0900 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Common non-ferrous metal mining | 0910 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Precious metal mining | 0920 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Rare earth metal ore mining | 0930 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Non-metallic mining | 1000 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Sandstone mining | 1010 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Chemical mining | 1020 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Salt mining | 1030 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Asbestos and other non-metallic mining | 1090 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Other mining activities | 1100 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | Service activities related to oil and gas extraction | 1110 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | Coal mining and washing professional and auxiliary activities | 1120 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | Other mining professional and ancillary activities | 1130 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Other mining industries | 1190 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Manufacturing industry | C000 | 0 | 0 | 0.263 | 0.263 | 0.474 | 0 | 0 | 0 |
| 2 | Agricultural and side food processing industry | 1300 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Grain milling | 1310 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Feed processing | 1320 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Vegetable oil processing | 1330 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Sugar industry | 1340 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Slaughter and meat processing | 1350 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Processing of aquatic products | 1360 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Vegetables, fungi, fruits and nuts are processed | 1370 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Other agricultural and side food processing | 1390 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Food, beverage and tobacco manufacturing | 13AA | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Food manufacturing | 1400 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Baked goods are manufacturing | 1410 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Candy, chocolate and honey manufacturing | 1420 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Convenient food manufacturing | 1430 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Dairy products are manufacturing | 1440 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Canned food manufacturing | 1450 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Condiments, fermented products manufacturing | 1460 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Other food manufacturing | 1490 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Wine, beverage and refined tea manufacturing | 1500 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Alcohol manufacturing | 1510 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | The manufacture of wine | 1520 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Beverage manufacturing | 1530 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Refined tea processing | 1540 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Tobacco products industry | 1600 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | The leaves are roasted again | 1610 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Cigarettes are made | 1620 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Other tobacco products manufacturing | 1690 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Among them: agricultural and side food processing industry | 16A0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Textiles | 1700 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Cotton textile and printing and dyeing finishing | 1710 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Wool textile and dyeing finishing | 1720 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Hemp textile and dyeing finishing | 1730 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Silk textile and printing and dyeing finishing | 1740 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Household textiles manufacturing | 1750 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Knitting or crochet weaving and the manufacture of its products | 1760 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Chemical fiber weaving and printing and dyeing finishing | 1770 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | The industry is made of textiles | 1780 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Textile and clothing, apparel industry | 1800 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Weaving clothing manufacturing | 1810 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Textile fabric shoes manufacturing | 1820 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Hat manufacturing | 1830 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Knitted or crocheted garments are manufacturing | 1840 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Clothing manufacturing | 1850 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Clothing shoes and hats, leather down and its products industry | 1A00 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Leather, fur, feathers and their products and footwear | 1900 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Leather tanning processing | 1910 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Leather products manufacturing | 1920 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Fur tanning and product processing | 1930 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Feather velvet) processing and product manufacturing | 1940 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Shoe industry | 1950 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Wood processing and wood, bamboo, rattan, brown, grass products industry | 2000 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Wood processing | 2010 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Artificial plate manufacturing | 2020 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Wood products manufacturing | 2030 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Wood and wood components are processed for construction | 2031 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Wood containers manufacturing | 2032 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Cork products and other wood products manufacturing | 2039 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Bamboo, rattan, brown, grass and other products manufacturing | 2040 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Furniture manufacturing | 2100 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Wood furniture manufacturing | 2110 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Bamboo, rattan furniture manufacturing | 2120 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Metal furniture manufacturing | 2130 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Plastic furniture manufacturing | 2140 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Other furniture manufacturing | 2190 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Among them: light industry | 21A0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | Wood processing and products and furniture products industry | 2A00 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Paper and paper products industry | 2200 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Pulp manufacturing | 2210 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Paper | 2220 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Paper products manufacturing | 2230 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Printing and recording media reproduction industry | 2300 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Printing | 2310 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Binding and other printing service activities | 2320 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Copy of the recording medium | 2330 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Culture, education, work and beauty, sports and entertainment supplies manufacturing | 2400 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | The manufacture of cultural and educational office supplies | 2410 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Sporting goods manufacturing | 2420 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Musical instruments are made | 2430 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | Toy manufacturing | 2440 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Entertainment equipment and entertainment supplies manufacturing | 2450 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Arts and crafts and etiquette supplies manufacturing | 2460 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | Oil, coal and other fuel processing industries | 2500 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| … | | | | | | | | | | |

Table S4. The delayed consumption of electricity from 1 kWh of electricity consumed on the first day. Simulated for Chongqing municipality using SIM model.

| kWh | To Day 0 | To Day 1 | To Day 2 | To Day 3 | To Day 4 | To Day 5 | To Day 6 | To Day 7 | To Day 8 |
|---|---|---|---|---|---|---|---|---|---|
| From Day 0 | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| From Day 1 | 0.000 | 0.833 | 0.415 | 0.229 | 0.093 | 0.132 | 0.083 | 0.054 | 0.022 |
| From Day 2 | 0.000 | 0.000 | 0.469 | 0.182 | 0.095 | 0.049 | 0.051 | 0.044 | 0.048 |
| From Day 3 | 0.000 | 0.000 | 0.000 | 0.451 | 0.192 | 0.103 | 0.052 | 0.054 | 0.043 |
| From Day 4 | 0.000 | 0.000 | 0.000 | 0.000 | 0.440 | 0.182 | 0.094 | 0.044 | 0.051 |
| From Day 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.388 | 0.168 | 0.091 | 0.045 |
| From Day 6 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.378 | 0.172 | 0.093 |
| From Day 7 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.362 | 0.159 |
| From Day 8 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.371 |

Table S5. The monthly economic outputs of Tamil Nadu (TN) and Rest of India (RoI) in crore ($10^7$) Rupees, both simulated for business as usual (bau) scenarios using DSIM and SIM models and the actual value.

| Monthly Economic Output/Rs. Crore | DSIM bau TN | DSIM bau RoI | SIM bau TN | SIM bau RoI | SIM flooded TN | SIM flooded RoI | Actual TN | Actual RoI |
|---|---|---|---|---|---|---|---|---|
| Jan-2015 | 14,209,221 | 171,999,244 | 14,154,123 | 170,316,160 | 14,154,123 | 170,316,160 | 17,563,742 | 177,616,774 |
| Feb-2015 | 17,529,393 | 176,028,959 | 17,472,691 | 176,474,967 | 17,472,691 | 176,474,967 | 17,385,377 | 179,014,777 |
| Mar-2015 | 18,509,554 | 176,556,104 | 18,500,300 | 176,140,577 | 18,500,300 | 176,140,577 | 18,361,140 | 177,779,158 |
| Apr-2015 | 17,978,437 | 177,021,910 | 17,952,947 | 176,386,674 | 17,952,947 | 176,386,674 | 17,679,155 | 177,650,544 |
| Jun-2015 | 18,283,926 | 178,232,051 | 18,195,782 | 178,193,551 | 18,195,782 | 178,193,551 | 18,161,791 | 179,103,121 |
| Jul-2015 | 18,699,667 | 178,856,359 | 18,654,846 | 178,073,834 | 18,654,846 | 178,073,834 | 18,812,300 | 179,232,061 |
| Aug-2015 | 18,629,035 | 182,195,277 | 18,601,676 | 181,860,540 | 18,601,676 | 181,860,540 | 18,843,776 | 181,887,231 |
| Sept-2015 | 18,822,154 | 184,346,717 | 18,704,153 | 184,094,181 | 18,704,153 | 184,094,181 | 18,896,237 | 183,657,549 |
| Oct-2015 | 18,788,039 | 184,777,682 | 18,874,075 | 183,341,409 | 18,874,075 | 183,341,409 | 19,064,110 | 182,506,308 |
| Nov-2015 | 18,952,857 | 190,120,238 | 18,921,171 | 190,871,798 | 18,921,171 | 190,871,798 | 19,053,618 | 187,265,544 |
| Dec-2015 | 19,446,897 | 185,656,996 | 19,465,343 | 185,579,982 | 18,336,986 | 185,579,982 | 18,298,188 | 184,016,429 |
| Jan-2016 | 19,460,336 | 185,473,139 | 19,452,288 | 185,418,626 | 18,715,623 | 185,003,160 | 18,864,760 | 184,934,453 |
| Feb-2016 | 19,552,136 | 185,307,487 | 19,543,286 | 185,159,792 | 19,050,271 | 184,888,548 | 19,064,110 | 185,520,394 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Mar-2016 | 19,637,178 | 185,559,821 | 19,632,883 | 185,504,487 | 19,046,043 | 185,322,957 | 19,001,157 | 185,013,218 |
| Apr-2016 | 19,736,339 | 185,876,529 | 19,735,223 | 185,874,948 | 19,701,729 | 185,658,871 | 19,599,206 | 186,557,018 |