SPATIAL ANALYSIS OF URBANIZATION AND ECONOMIC DEVELOPMENT IN THAILAND

BY

NUTCHAPON PRASERTSOONG

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN ECONOMICS

FACULTY OF ECONOMICS
THAMMASAT UNIVERSITY

ACADEMIC YEAR 2022

SPATIAL ANALYSIS OF URBANIZATION AND ECONOMIC DEVELOPMENT IN THAILAND

BY

NUTCHAPON PRASERTSOONG

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN ECONOMICS

FACULTY OF ECONOMICS

THAMMASAT UNIVERSITY

ACADEMIC YEAR 2022

THAMMASAT UNIVERSITY
FACULTY OF ECONOMICS

DISSERTATION

BY

NUTCHAPON PRASERTSOONG

ENTITLED

SPATIAL ANALYSIS OF URBANIZATION AND ECONOMIC DEVELOPMENT IN THAILAND

was approved as partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Economics

on July 31, 2023

Chairman	(Associate Professor Sutee Anantsuksomsri, Ph.D)
Member and Advisor	(Associate Professor Nattapong Pattanapong, Ph.D.)
Member	(Assistant Professor Monthien Satimanon, Ph.D.)
Member	(Assistant Professor Wasin Siwasarit, Ph.D.)
Member	(Associate Professor Kaewkwan Tangtipongkul, Ph.D.)
Dean	(Assistant Professor Supachai Srisuchart, Ph.D.)

Thesis Title	SPATIAL ANALYSIS OF URBANIZATION AND SOCIO-ECONOMIC DEVELOPMENT IN THAILAND
Author	Nutchapon Prasertsoong
Degree	Doctor of Philosophy (Economics)
Major Field/Faculty/University	Faculty of Economics Thammasat University
Thesis Advisor	Associate Professor Nattapong Pattanapong, Ph.D
Academic Year	2022

ABSTRACT

This study aims to guide regional and urban development policy by addressing Thailand's extreme spatial inequality and inadequate urban planning. The study seeks to reduce regional economic disparity and promote sustainable urban development by investigating the mechanisms governing the spatial distribution of economic activities. This study employed both conventional ground surveys and satellite data to overcome the limitations that previous research efforts have incurred.

Chapters 2 and 3 examined the causal relationship between economic outcomes and urbanization. Chapter 2 utilized a two-stage estimation method, incorporating clay content in the soil as an instrumental variable. The findings reveal that wage differentials were significantly influenced by workers' education, experience, and agglomeration externalities in larger cities. The study recommends creating multiple regional cities to generate agglomeration externalities.

Based on the recommendations of Chapter 2, Chapter 3 investigated the driving factors of urbanization over the past two decades. The study shows that urbanization in Thailand was driven by urban sector productivity growth, the share of workers with higher education, and favorable environmental factors such as water availability. Spatial spillover of urbanization was also observed. The findings suggest maximizing the economic potential of each region to support polycentric urbanization through productivity enhancement mechanisms.

Chapter 4 forecasted urban land expansion and economic growth in Ban Chang district, a rapidly developing city in Thailand. Geospatial data from Google Earth Engine and a web-based application developed for the study were utilized. A recursive dynamic model was applied to achieve the research's goals and demonstrated an average spatial accuracy of 92 percent. The model's ability to forecast urban land expansion and economic growth at the district level was also showcased. This framework, relying on open data and open-source software, enables cost-effective monitoring and forecasting of urbanization and economic development, facilitating a proactive approach to urban planning.

The study highlights the need to address spatial inequality and inadequate urban planning through evidence-based approaches, recommending strategies such as creating regional cities, productivity enhancement mechanisms, and proactive urban planning processes.

Keywords: urbanization, spatial analysis, agglomeration, wage disparity, productivity, satellite data, land-use change model, recursive dynamic model, Thailand

ACKNOWLEDGEMENTS

I want to express my heartfelt gratitude to all those who have contributed to the completion of this study.

I want to thank my advisor, Assoc. Prof. Dr. Nattapong Pattanapong, for his invaluable guidance, continuous support, and expertise throughout the research process. His insightful feedback and constructive criticism have been shaping the direction and quality of this study. I am grateful for the role of my advisor, whose guidance and mentorship have propelled me toward achieving academic excellence and have also played a significant role in my personal growth and maturity. His unwavering support, encouragement, and experience have shaped me not only as a researcher but also as an individual. I am truly fortunate to have had such a resolute and influential advisor by my side throughout this journey.

In addition to my advisor, I would like to express my gratitude to the other members of my thesis committee, Assoc. Prof. Dr. Kaewkwan Tangtipongkul, Asst. Prof. Dr. Monthien Satimanon, Asst. Prof. Dr. Wasin Siwasarit, and Assoc. Prof. Dr. Sutee Anantsuksomsri. Their valuable suggestions and constructive comments have played a significant role in developing and refining this thesis. Their expertise and insights have contributed immensely to the overall quality of the research.

Besides individuals, I would like to express my sincere gratitude to the scholarship of Thammasat University for financial support during my study as a Ph.D. student and thanks to the graduate program faculty members for facilitating administrative tasks throughout my study in this program. I also want to express my appreciation to the institutions providing geospatial data access, particularly Google Earth Engine service. Their commitment to open data and open-source software has made this study possible and has significantly contributed to its success.

I greatly appreciate the principles of Buddhism, as preached and elaborated by Siddhartha Gautama, as well as the esteemed Thai Buddhist monks such as Buddhadāsa Bhikkhu and Phra Paisal Visalo. Their teachings serve as a profound source of guidance, akin to a compass for a sailor voyaging the treacherous oceans, leading individuals towards the right path amidst life's trials and uncertainties.

Completing this dissertation owes much to the invaluable guidance provided by the teachings of Buddhism.

I want to extend my heartfelt appreciation to Pramuan Pengchan, Ph.D., the author of the acclaimed book "Walk to Freedom," as his thoughts and life philosophy have profoundly impacted my spiritual experience and perspective of the world. His remarkable actions and teachings have transformed my spirit and influenced my decision to pursue a Doctor of Philosophy degree.

Lastly, I am deeply grateful to my family for their love and understanding. Their constant support, both emotionally and financially, as well as their belief in my abilities, have been the driving force behind my academic pursuits.

Although it is impossible to mention everyone by name, I sincerely appreciate the collective efforts of all those who have played a part, directly or indirectly, in completing this study. Thank you for your contributions and support.

Nutchapon Prasertsoong

TABLE OF CONTENTS

	Page
ABSTRACT	(1)
ACKNOWLEDGEMENTS	(3)
LIST OF TABLES	(10)
LIST OF FIGURES	(12)
LIST OF ABBREVIATIONS	(14)
CHAPTER 1 INTRODUCTION	1
CHAPTER 2 URBANIZATION AND AGGLOMERATION IN THAILAND	6
2.1 Introduction	6
2.2 Literature Review	12
2.2.1 Theoretical Literature	12
2.2.2 Empirical Literature	16
2.3 Theoretical Framework	19
2.3.1 Ciccone and Hall Model	19
2.3.1.1 Externalities	19
2.3.1.2 Intermediate Product Variety	21
2.3.1.3 Capital and Total Factor Productivity	23
2.3.2 Foundation of Agglomeration Externalities	25
2.4 Methodology	27
2.4.1 Mincer Equation	28
2.4.2 First-Stage Regression	28

2.4.3 Second-Stage Regression 29
2.5 Data 29
2.5.1 Dependent Variable 30
2.5.2 Independent Variable in the First-Stage Regression 30
2.5.3 Independent Variable in the Second-Stage Regression 31
2.5.3.1 Urbanization Economies 31
2.5.3.2 NTL Density 32
2.5.3.3 Localization Economies 33
2.5.3.4 Instrumental Variable 35
2.6 Result 37
2.6.1 Descriptive Statistics 37
2.6.2 Mincer Regression Results 42
2.6.3 Second-Stage Regression Results 44
2.6.4 NTL Density Regression Results 48
2.6.5 Policy Recommendations 50
2.7 Conclusion 51

CHAPTER 3 URBAN LAND EXPANSION AND ECONOMIC DEVELOPMENT IN THAILAND 53

3.1 Introduction 53
3.2 Literature Review 59
3.2.1 Theoretical Literature 59
3.2.2 Empirical Literature 62
3.3 Theoretical Framework 66
3.3.1 Model of Urbanization and Structural Change 66
3.3.1.1 Household Utility Function 66
3.3.1.2 Production Function 67
3.3.1.3 Budget Constraint 67
3.3.1.4 Wage Equalization Between Sectors 68
3.3.1.5 Urbanization and Comparative Statics 69

3.4 Methodology 70
3.4.1 Fixed-Effects Model 71
3.4.2 Spatial Lag Model with Spatial Fixed Effects 72
3.4.3 Spatial Error Model with Spatial Fixed Effects 72
3.4.4 Dynamic Panel-Data Model 72
3.5 Data 73
3.5.1 Dependent Variable 73
3.5.2 Explanatory Variable 74
3.5.2.1 Tradable GPP Per Capita 74
3.5.2.2 Resource GPP Per Capita 75
3.5.2.3 Environmental Variables 75
(1) Normalized Difference Vegetation Index (NDVI) 75
(2) Normalized Difference Water Index (NDWI) 75
(3) Normalized Difference Drought Index (NDDI) 76
(4) Land Surface Temperature 76
(5) Precipitation 76
3.5.2.4 Share of Worker With Higher Education 77
3.6 Result 78
3.6.1 Descriptive Statistics 78
3.6.2 Regression Results 79
3.6.3 Policy Recommendations 86
3.7 Conclusion 87

CHAPTER 4 PREDICTING LOCAL ECONOMIC DEVELOPMENT AND URBAN LAND EXPANSION: CASE STUDY OF BAN CHANG DISTRICT, RAYONG PROVINCE, THAILAND 88

4.1 Introduction 88
4.2 Study Area: Ban Chang District 89
4.3 Literature Review 93
4.3.1 Land-Use Change Model With Socio-Economic Factor 93

4.3.2 Empirical Literature	96
4.3.3 Empirical Literature in Thailand	101
4.4 Foundation of Land-Use Change Modeling	102
4.4.1 A Cellular Model of Urban Land-Use	102
4.5 Methodology and Data	105
4.5.1 Methodology	105
4.5.1.1 Modelling Procedure	105
4.5.1.2 Forecasting Demand for Urban Land	106
4.5.1.3 Determining Supply for Urban Land	108
4.5.1.4 CLUE-S Allocation Rule	109
4.5.1.5 Accuracy Assessment	110
4.5.2 Data	111
4.5.2.1 Geospatial Data	111
4.5.2.2 Economic Data	114
4.6 Result	115
4.6.1 Estimated Results of simultaneous equations	115
4.6.2 Estimated Results of Logistic Regression	116
4.6.3 Urban Growth of Ban Chang Between 2016 and 2020	117
4.6.4 Model Validation	118
4.6.5 Projected Urban Land Expansion	119
4.6.6 Projected Economic Growth of Ban Chang	120
4.6.7 Policy Recommendations	121
4.7 Conclusion	122

CHAPTER 5 CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusion	123
5.2 Policy Implication and Recommendations	125
5.3 Limitations and Suggestion for Further Studies	127

REFERENCES 130

APPENDICES

APPENDIX A 145
APPENDIX B 147
APPENDIX C 149
APPENDIX D 151
APPENDIX E 152
APPENDIX F 155
APPENDIX G 158
APPENDIX H 160
APPENDIX I 161

BIOGRAPHY 174

LIST OF TABLES

Tables	Page
2.1 Top 10 Most Densely Populated Provinces in Thailand, 1995 and 2018	7
2.2 Top 10 Provinces That Experienced a Decline in Population Between 1995 and 2018	9
2.3 Top 10 Provinces With Highest Median Monthly Wage in Thailand, 2002 and 2018	10
2.4 Summary of Statistics in the First-Stage Regression	38
2.5 Summary of Statistics in the Second-Stage Regression	38
2.6 Results of Mincer Regression (Equation (2.36))	43
2.7 Results of Second-Stage Regression Using Population Density to Quantify Agglomeration	45
2.8 Durbin-Wu-Hausman Test of Second-Stage Regression Using Population Density to Quantify Agglomeration	47
2.9 Stock and Yogo Weak Instrument Test of Second-Stage Regression Using Population Density to Quantify Agglomeration	48
2.10 Results of Second-Stage Regression Using NTL Density to Quantify Agglomeration	48
2.11 Durbin-Wu-Hausman Test Using NTL Density to Quantify Agglomeration	50
2.12 Stock and Yogo Weak Instrument Test Using NTL Density to Quantify Agglomeration	50
3.1 Top 10 Provinces With Highest Increase in Urban Area Between 2000 and 2020	55
3.2 Top 10 Provinces With Highest Increase in Gpp (CVM) Between 2000 and 2020	56
3.3 List of Empirical Literature	63
3.4 Summary of Data	77
3.5 Descriptive Statistics of Variables in the Regression	78

3.6 Results of Non-spatial Panel Regression 79
3.7 Result of Spatial Lag Panel Regression 81
3.8 Results of Spatial Error Panel Regression 82
3.9 Results of Dynamic Panel-Data Regression 83
4.1 Previous Empirical Literature That Study Land-Use Change and Socio-Economic Factors 97
4.2 Confusion Matrix 110
4.3 Estimated Results of Equation (4.7) And Equation (4.8) Using 2SLS 115
4.4 Estimated Results of Equation (4.13) Using Logistic Regression 116
B.1 Main Specifications of Geospatial Data That Are Used in the Analysis of Chapter 2 147
B2. Main Specifications of Economic Data That Are Used in the Analysis of Chapter 2 148
D.1 Main Specifications of Geospatial Data That Are Used in the Analysis of Chapter 3 151
H.1 Main Specifications of Geospatial and Economic Data That Are Used in the Analysis of Chapter 4 160
I.1 Source Code for Simultaneous Equation 161
I.2 Source Code for Land-Use Change Simulation Between 2017–2020 163
I.3 Source Code for Land-Use Change Projection in 2021 169

LIST OF FIGURES

Figures	Page
2.1 Distribution of Population Density in Thailand, 1995 and 2018	8
2.2 Distribution of Median Monthly Wage in 2002 and 2018	11
2.3 Scatter Plot of Population Density and Median Wage in 2002 and 2018	11
2.4 Visualization of Raw and Transformed NTL Data in 2021	33
2.5 Visualization of Raw and Transformed Data Percentage of Clay Content in the Soil	36
2.6 Visualization of Hourly Wage and Population Density in 2007 and 2017	39
2.7 Scatter Plot of Population Density and Hourly Wage	41
2.8 Scatter Plot of NTL Density and Hourly Wage	42
2.9 Spatial Distribution of Average Spatial Residual	44
3.1 Urban Land Expansion in Thailand Between 2000 and 2020 in Square Kilometer	54
3.2 Distribution of an Increase in Urban Area Between 2000 and 2020 in Square Kilometer	56
3.3 Distribution of an Increase in GPP (CVM) Between 2000 and 2020	57
3.4 Relationship Between Change in Urban Land Area and Change in GPP (CVM) At Provincial Level (In Natural Log Term)	58
4.1 Geographical Location of Ban Chang District	90
4.2 Relationship Between Rayong's NTL Density and Rayong's GDP	91
4.3 NTL Density of Ban Chang District	91
4.4 Overview of Urban Land-Use Change Prediction Process	106
4.5 Gui of Constructed Application	112
4.6 Statistics of Urban Land Area and Land Use Map of Ban Chang District Generated by the Constructed Application	112
4.7 Observed Land Use Maps of Ban Chang District in 2016 and 2020	113

4.8 Slope, Elevation, and Distance to Existing Urban Area of Ban Chang District 113
4.9 Land-Use Change Pattern of Ban Chang District Between 2016 and 2020 117
4.10 Spatial Accuracy From 2017–2022 118
4.11 Urban Land Demand of Ban Chang District 119
4.12 Predicted Land-Use of Ban Chang, 2021 and 2025 120
4.13 Predicted NTL of Ban Chang District From 2016–2025 121
A.1 Scatter Plot of Labor Productivity and the Median Hourly Wage 145
A.2 Scatter Plot of Labor Productivity and the Average Hourly Wage 146
C.1 BiLISA Cluster Map 149
C.2 Moran Scatter Plot 150
E.1 Similarity Between Land-Use Map From Dynamic World and Land-Used Map From Land Development Department in Percent, 2016–2021 152
E.2 Illustration of Similarity Between Land-Use Map From Dynamic World and Land-Used Map From Land Development Department, 2016–2021 153
F.1 Projection of Ban Chang's GDP (Proxied by NTL Density) From VECM, 2021–2025 155
F.2 Projection of Ban Chang's Urban Land From VECM (The Most Likely Scenario), 2021–2025 156
F.3 Fanchart of the Distribution of Projected Ban Chang's GDP (Proxied by NTL Density) From VECM, 2021–2025 156
F.4 Fanchart of the Distribution of Projected Ban Chang Urban Area From VECM, 2021–2025 157
G.1 Correlation Between GDP and the Facebook Movement Index 158
G.2 Predicted GDP (Proxied by NTL) Of Ban Chang District and (Natural) Logarithm of the Number of Poor People (Monetary Poverty), 2018–2022 159

LIST OF ABBREVIATIONS

Symbols/Abbreviations	Terms
NESDC	National Economic and Social Development Council
GDP	Gross Domestic Product
GPP	Gross Provincial Product
LFS	Labor Force Survey
IC	Industrial Census
NSO	National Statistical Office
BMR	Bangkok Metropolitan Region
MODIS	Moderate-Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
CLUE	Conversion of Land Use and its Effects
CLUE-S	Conversion of Land Use and its Effects at Small regional extent
NTL	Nighttime Light
OLS	Ordinary Least Squares
IV	Instrument Variable
ISIC	International Standard Industrial Classification
DMSP/OLS	Defense Meteorological Program/Operational Linescan System
VIIRS/DNB	Visible Infrared Imaging Radiometer Suite/Day-Night Band
CVM	Chain Volume Measure
NDDI	Normalized Difference Drought Index
NDVI	Normalized Difference Vegetation Index

NDWI	Normalized Difference Water Index
CHIRPS	Climate Hazards Group Infrared Precipitation with Station
CA	Cellular Automata
GIS	Geographic Information System
2SLS	Two-Stage Least-Squares
GUI	Graphical User Interface
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
VECM	Vector Error Correction Model

CHAPTER 1
INTRODUCTION

Geography matter in Economics. People hardly live sparsely but find themselves settling in communities in one way or another. Businessman rarely establishes their enterprises randomly across space but struggle to set them up in cities.

The distribution of natural resources exhibits significant disparities, with soil fertility varying across landscapes and rainfall being more abundant in certain regions. The climate in tropical areas is considered more conducive to sustaining life than the Arctic, making certain places more favorable than others. The interaction between households, producers, and natural factors leads to the clustering of global economic activities. This phenomenon paves the way for the laissez-faire economy's success since the commencement of the Industrial Revolution and the unprecedented territorial division of the present time.

The concentration of economic activities is the backbone of the success of the modern capitalist economy. Clustering, even in primitive form, improves human welfare by protecting human lives from wilderness and wars. The city is a brilliant innovation that helps humanity create innovative technology that overcomes the constraints of constant return to scale in production, allowing capital accumulation and growth to occur (Glaeser, 2013). Geographic concentration and innovation could also be self-reinforcing, a process often associated with increasing return to scale or agglomeration externalities.

On the other hand, the unprecedented economic integration between regions and the recent advancement in telecommunication technology has made some places remarkably more prosperous than others and caused a decline in economic activity of some areas, worsening territorial inequality and violating the Pareto efficiency. From a national point of view, the dwindling economic opportunities in specific areas imply that people who cannot move to flourishing cities must endure persistent poverty, unemployment, and despair. As governments in both developed and developing countries fail to rejuvenate the cities and towns in the declining areas, people in the "places that don't matter" have resorted to the recent extreme populism voting as their remedy for economic hardship, an action that hurt overall national or global development but serve as the vengeance of left-behind people (Rodríguez-Pose, 2018). From a city point of view, poorly planned urbanization could lead to various problems that urban populations in developing countries are currently experiencing. These problems include congestion, growth of crowded communities, crime, pollution, and health issue (Arfanuzzaman & Dahiya, 2019).

Understanding the theoretical mechanisms of urbanization through dedicated studies is crucial due to city growth's positive and negative impacts. Developing a sound economic model or conducting rigorous research to depict real-world situations accurately is the initial stride toward formulating effective policies that address prevailing urban issues. Urban policies designed to combat or alleviate problems without such endeavors would lack proper guidance, leading to potential failures.

Such academic demand has led to the birth of urban economics. This economic doctrine explicitly incorporates location choice into the utility and profit maximization problem of households and producers, answering why households and firms concentrate in certain areas. According to O'Sullivan (2012), the study of urban economics can be divided into six key areas. The first area studies how the interaction between agglomeration and dispersion force generates big and small cities. It also explores various causes of urban and economic growth. The second area studies how the household and firm decision affects the urban land use pattern within the cities. The third area aims to make urban transportation more efficient. The fourth specifically studies the negative side of the cities–the concentration of crime and (urban) poverty in the cities. The fifth area addresses housing policy. Lastly, the sixth area of urban economics studies how local government expenditures and taxes affect the location choice of the household.

Given that urban economics covers a broad subject of study. This thesis focuses on the first and second areas, studying why cities differ in size, what causes urban growth, and how urban land use patterns respond to changing economic conditions. The first and second area of urban economics is rich in history. Economic concentration's byproducts have attracted social theorists from the Age of Enlightenment and contemporary economists to reason about a particular economic activity flourishing in a specific area, where it is taking place, how land use patterns would change, and what economists should do about it.

Adam Smith is the most well-known economist who considers the influence of geography on economic performance. In addition to his famous idea of division of labor and the invisible hand, Smith recognized that location and urban scale make some regions more advantageous to others (Ioannou & Wójcik, 2022). After Adam Smith, the German economist Johann Heinrich von Thünen, who found the classic von Thünen model, showed how economic activities would distribute across space. According to his model, an urban center of his hypothetical town would become a market where agricultural products are traded. The centrality of the urban center ensures that land rent is the highest in the town. The Central-Place Theory developed by Walter Christaller and August Lösch hypothesized that transaction cost would be minimized and benefit from economy of scale would be maximized at the core of each hexagonal lattice, where each core represents a city in a hierarchical order ranging from the metropolis to minor towns.

Based on the idea of thinkers in the late nineteenth century, Masahisa Fujita, Paul Krugman, and Anthony J. Venables are the contemporary economist who developed the "New Economic Geography" in their highly celebrated book "The Spatial Economy: Cities, Regions, and International Trade." Fujita, Krugman, and Venables attempted to answer the question of "where economic activities occur and why" by deriving a model that could mimic the actual distribution of city size by considering the role of population growth, transportation cost, and agglomeration externalities. The New Economic Geography made two significant contributions. First, it showed how an interaction between the agglomeration (centripetal) and dispersion (centrifugal) forces generates the spatial distribution of city size, where some cities become more attractive to workers than others. Second, it implies that cities and economic growth would go hand in hand as advanced economic activities would take place, allowing cities to grow bigger.

While the first implication of New Economic Geography drew many researchers to empirically examine whether agglomeration and dispersion forces affect regional inequality and productivity of cities, understanding whether economic growth affects city expansion is another concern as many developing countries have been experiencing both economic boom and rapid urbanization. Moreover, understanding the drivers of urbanization might pave the way to a new regeneration policy that could successfully urbanize regional cities.

According to statistics from the World Bank, the economies of many countries have gone through a period of remarkable change over the last fifty years. Globally, GDP (Gross Domestic Product) per capita has risen from 457 US dollars in 1960 to 10,918 US dollars in 2020. As the world economy transformed, many countries successfully shifted from the agricultural-based economy to the industrial and service-based economy. Rapid urbanization has been taking place alongside this trend. More people move to cities at an unprecedented rate. The share of the urban population rose shapely from 1.1 billion in 1960 to 4.4 billion in 2020. The World Bank projected that urban populations would reach 7 billion by 2052. In other words, 70 percent of the global population will live in cities in the next three decades (World Bank, 2021).

Thailand shares a similar story to the experience of other countries. Economically, Thailand's GDP per capita has tremendously risen from 100 US dollars in 1960 to 7,186 US dollars in 2020. Her urban population has robustly risen from around 5.4 million in 1960 to 36 million in 2020. However, urbanization in Thailand highly concentrates in Bangkok Metropolitan Region (BMR). Research suggests Thailand has the highest regional inequality globally (Short & Pinet-Peralta, 2009). The extreme concentration in Thailand has also resulted in a poorly organized growth of urban land, worsening environmental, health, economic, and social problem (Iamtrakul, Padon, & Klaylee, 2022).

The extreme territorial inequality triggers a demand for polycentric growth patterns and better urban planning policies from intellectuals, policymakers, politicians, and civil society. Political parties often promote inequality and poverty reduction by "redistributing wealth to the provinces" as their political campaign, reflecting the concern about spatial inequality in Thai society. The prolonged tension between rich and poor regions in Thailand was manifested in the 2011 general election and the recent decade-long political crisis between Bangkok's middle-classes and rural workers from North and Northeast regions (Hewison, 2014). Thai scholars believe that the urbanization of regional cities and better urban planning policies would not only defuse the tension between rich and poor regions but also promote sustainable economic development, allowing Thai people to attain a higher level of well-being and happiness. Had the government failed to provide proper policies on urbanization beforehand, such rapid urbanization could result in unsustainable economic and urban development, exacerbating regional inequality and social problem.

Based on the author's best knowledge, a comprehensive study that explores the linkages between urbanization and economic development in Thailand needs to be included. Conducting such research would be an essential initial step toward shaping regional development policies in the future.

This study aims to fill this gap by answering three major questions. First, how the interaction between agglomeration and dispersion forces make some cities more successful. Second, how economic development affect urbanization in Thailand. Lastly, how the interaction between economic growth and urbanization allows researchers to predict pattern of urbanization and economic growth at district lelvel.

The first question was extensively addressed in Chapter 2, where a large micro-data from conventional ground survey and satellite data was gathered and analyzed via the latest technique currently employed in the field. Specifically, Chapter 2 answered whether positive agglomeration externalities of big cities make workers in some regions better off than others. Based on the evidence found in Chapter 2, an analysis of the driver of cities' growth was conducted in Chapter 3.

Like Chapter 2, the combination of conventional government and satellite data was analyzed via one of the most rigorous regression techniques in spatial econometrics. Based on economic theory, Chapter 3 uncovered the underlying linkage between rising productivity growth and urbanization, creating numerous implications for regional development policies. Based on Chapters 2 and 3 results, the self-enforcing mechanism of urbanization and economic development were considered in predicting urbanization at the cell level in Chapter 4. Chapter 4 aims to provide a flexible modeling framework that could incorporate economic data into land-change modeling and predict future urban expansion at the cell level under certain economic conditions.

CHAPTER 2
URBANIZATION AND AGGLOMERATION IN THAILAND

2.1 Introduction

In theory, urbanization is considered the primary driver of economic expansion. A higher population density in urban areas facilitates a broader range of diverse and advanced economic activities. As knowledge is exchanged among educated individuals, productivity naturally increases, leading to faster economic growth. According to Fujita, Krugman, and Mori (1999), socioeconomic factors such as population growth and reduced transportation costs have significant roles in developing cities and the economy, particularly in a monocentric form.

On the other hand, urbanization encompasses various challenges, such as crime, pollution, disease, and environmental degradation, as it involves converting natural land into built environments. According to Harari (2020), rapid and unregulated urban growth without proper planning can lead to urban sprawl. Urban sprawl refers to the expansion of urban areas with low population density, which often results in long commuting distances and limited transportation alternatives. The negative aspects of urban living are often referred to as congestion costs or centrifugal forces. Evaluating the combined impact of agglomeration forces (benefits of urban concentration) and congestion forces is a crucial initial step in understanding the overall effects of agglomeration externalities on the urbanization pattern.

Over the past six decades, Thailand has witnessed a significant shift of people from rural areas to urban centers as part of its urbanization process. However, are few cities that experience rapid increases in population. This trend is depicted in Table 2.1 and Figure 2.1, which illustrate the changes in population density across Thailand between 1995 and 2018. The top three provinces with the highest population density in 1995 remained unchanged in 2018. These provinces are Bangkok, Nonthaburi, and Samut Prakan.

However, contrasting with the previous point, significant population density growth occurs in Chon Buri, Rayong, and Chachoengsao during the same period. These provinces are recognized for their industrial parks and flourishing tourism industry, making them attractive destinations for workers. The high demand for labor in these sectors entices individuals from provinces primarily engaged in agricultural activities to migrate in search of employment opportunities.

Table 2.1

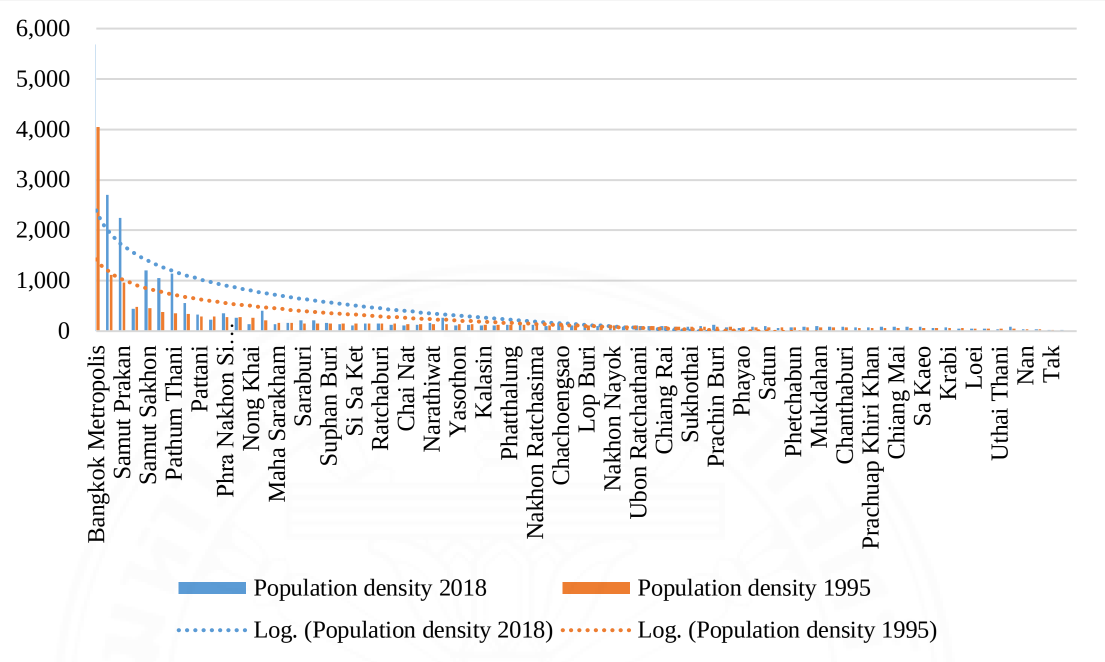
Top 10 Most Densely Populated Provinces in Thailand, 1995 and 2018

	Province	Population density 1995		Province	Population density 2018
1	Bangkok	4047.86	1	Bangkok	5683.70 (-)
2	Nonthaburi	1113.58	2	Nonthaburi	2705.79 (-)
3	Samut Prakan	963.03	3	Samut Prakan	2245.35 (-)
4	Samut Songkhram	483.16	4	Samut Sakhon	1202.32 (+1)
5	Samut Sakhon	461.55	5	Pathum Thani	1136.38 (+2)
6	Phuket	384.91	6	Phuket	1056.53 (-)
7	Pathum Thani	359.61	7	Nakhon Pathom	553.28 (+1)
8	Nakhon Pathom	345.72	8	Samut Songkhram	439.73 (-4)
9	Pattani	291.10	9	Chon Buri	403.75 (+5)
10	Sing Buri	287.18	10	Phra Nakhon Si Ayutthaya	356.06 (+1)

Note. (i) The table was created from data provided by the National Economic and Social Development Council (NESDC). (ii) The symbol (-) indicate stable rank, (+) indicate moving up in rank, (-) indicate moving down in rank.

Figure 2.1

Distribution of Population Density in Thailand, 1995 and 2018



Note. The figure was created from data provided by the NESDC.

The impact of industrialization is not uniformly distributed across the country, resulting in contrasting economic outcomes. This disparity is clearly illustrated in Table 2.2, highlighting the top 10 provinces that experienced a population decline. Most of these provinces are in the northeastern region of the country. Sing Buri, Si Sa Ket, Surin, and Yasothon have a significant population decline.

In addition to the northeastern region, population decline is also taking place in some provinces in the Central and Northern regions, such as Sing Buri, Samut Songkhram, Chi Nat, Ang Thong, Phayao, Phrae, Nakhon Sawan, and Phichit. A population decline might be attributed to the limited presence of significant agricultural, industrial, and tourism development in these areas.

Table 2.2

Top 10 Provinces That Experienced a Decline in Population Between 1995 and 2018

	Province	Change in population density between 1995 and 2018
1	Sing Buri	-54.23
2	Si Sa Ket	-43.55
3	Samut Songkhram	-43.44
4	Chai Nat	-27.45
5	Surin	-24.67
6	Maha Sarakham	-24.05
7	Yasothon	-22.47
8	Buri Ram	-21.81
9	Amnat Charoen	-20.52
10	Ang Thong	-19.71

Note. The table was created from data provided by the NESDC.

Urban economic theory suggests that provinces with higher population densities are likely to have higher productivity because of agglomeration externalities. In a perfectly competitive labor market, productivity can be estimated by looking at labor wages. The median provincial wage is a representative measure of the average worker's wage. Interestingly, Thailand's labor wage change is more dynamic than population density. As shown in Table 2.3, Bangkok and Nonthaburi, Thailand's two most densely populated provinces, consistently hold the top positions in worker income.

Workers who resided in the BMR and the Eastern Economic Corridor have been experiencing a consistent increase in their median wages over the past 17 years. Some provinces become part of the top 10 provinces with the highest average monthly wages in Thailand, while others drop out of this list. Chachoengsao, Loei, Chon Buri, Chumphon, Prachin Buri, Mukdahan, and Prachuap Khiri Khan provinces significantly climbed the ladder. This income rise could be attributed to establishing industrial parks and expanding the tourism industry in these provinces.

Table 2.3

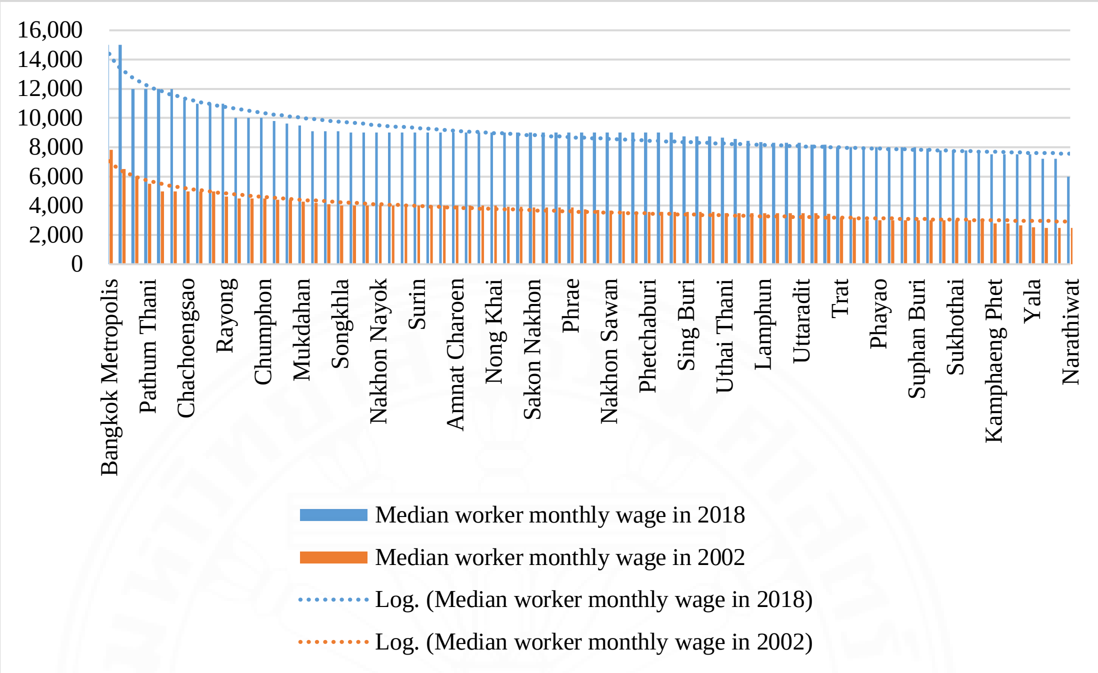
Top 10 Provinces With Highest Median Monthly Wage in Thailand, 2002 and 2018

	Province	Median worker monthly wage in 2002		Province	Median worker monthly wage in 2018
1	Nonthaburi	7,830	1	Bangkok	15,000 (+1)
2	Bangkok	6,500	2	Nonthaburi	15,000 (+1)
3	Pathum Thani	6,000	3	Samut Prakan	12,000 (+1)
4	Samut Prakan	5,500	4	Pathum Thani	12,000 (-1)
5	Phra Nakhon Si Ayutthaya	5,000	5	Chon Buri	12,000 (+1)
6	Chon Buri	5,000	6	Phuket	12,000 (+2)
7	Rayong	5,000	7	Chachoengsao	11,250 (+8)
8	Phuket	5,000	8	Phra Nakhon Si Ayutthaya	11,000 (-3)
9	Songkhla	5,000	9	Saraburi	11,000 (+2)
10	Nakhon Pathom	4,628	10	Rayong	11,000 (-3)

Note. (i) The table was created from data provided by the NESDC. (ii) The symbol (-) indicate stable rank, (+) indicate moving up in rank, (-) indicate moving down in rank.

Figure 2.2

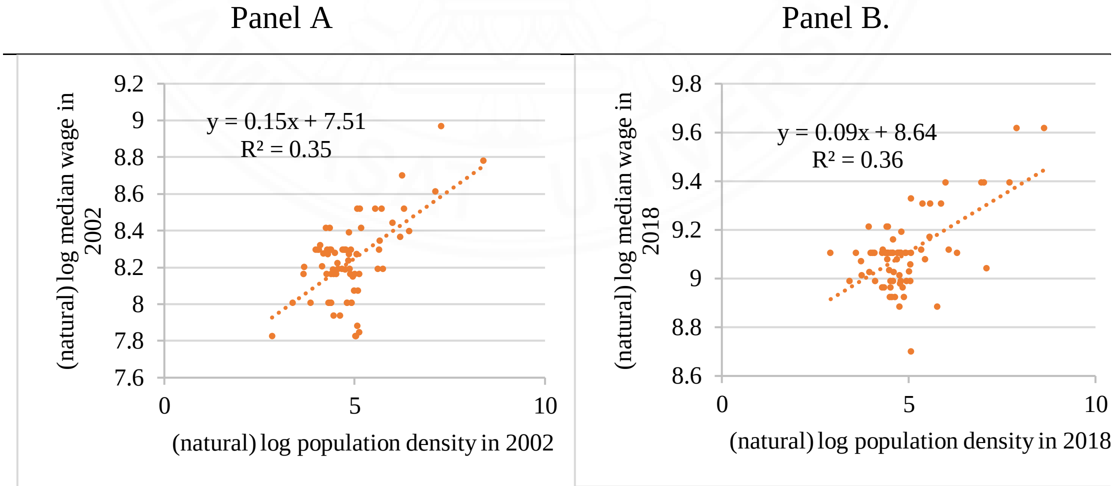
Distribution of Median Monthly Wage in 2002 and 2018



Note. The table was created from data provided by the NESDC.

Figure 2.3

Scatter Plot of Population Density and Median Wage in 2002 and 2018



Note. (i) The table was created from data provided by the NESDC and the National Statistical Office (NSO) of Thailand. (ii) Panel A shows the relationship between median wage and population density in 2002. Panel B shows the relationship between median wage and population density in 2018.

Figure 2.3 visualizes the relationship between population and wages in 2002 and 2018. The horizontal axis represents the natural logarithm of population density, and the vertical axis represents the median monthly wage for workers. The regression line in the graph has a slope of 0.14 for 2002 and 0.09 for 2018. The R-squared values are 0.35 for the year 2002 and 0.36 for the year 2018. The figure clearly illustrates a positive relationship between population density and the median wage of workers in Thailand. In both years, population density alone accounts for approximately 35 percent of the variation in workers' median wages. These findings, along with the other figures and tables, highlight the significant role of density and its impact on productivity in Thailand over the past few decades. Certain provinces continue to attract workers from other regions, leading to their growth and prosperity at the expense of other provinces. Thus, analyzing how agglomeration externalities influence regional disparity in Thailand is crucial to developing an effective regional development policy.

Therefore, this paper aims to quantify agglomeration externalities in Thailand. The study used ground survey and remote sensing data to achieve the objective. This approach allows for reliable quantification of the impact of agglomeration externalities. Remote sensing data is another way to validate the findings derived from conventional ground survey data, enabling researchers to compare and cross-check the results obtained from both sources.

2.2 Literature Review

2.2.1 Theoretical Literature

Urban economists have studied spatial dimensions' impact on the economy since the nineteenth century. Fujita, Krugman, and Mori (1999) summarized academic progress in urban economics and regional science. Though the primary goal of Fujita, Krugman, and Mori was to provide a rigorous analysis of the development of the urban system, their work concisely summarized the evolution of theoretical literature that addressed the city's agglomeration externalities.

Even though the mainstream economic community largely ignored the locational distribution aspect of the economy, the first consideration of the spatial dimension of economic activities and agglomeration dated back to the early nineteenth century. The German economist Johann Heinrich von Thünen pioneered the field of urban economics. Von Thünen imagined an isolated town surrounded by agricultural land. He assumed that agricultural product productivity and transportation costs vary according to soil quality and geography. The difference in crop productivity and transportation costs gives rise to different intensities of cultivation. In his model, von Thünen showed that, in a perfectly competitive market, competition among farmers would lead to an equilibrium condition where the level of land rents would be highest in the town and decline to the lowest level at the farthest point of cultivation.

Despite its simplicity, von Thünen's model highlighted the basic idea of the city's agglomeration effect that led to the concentric ring pattern in which landlords in the town could get the highest rent and land owners at the city's edges get the lowest rent. Following von Thünen, Hoover (1948) proposed the basic concept known as agglomeration externalities or external economies of scale. The basic idea of external economies is that firms should be located in the same industrial district to reduce the number of possible market failures. Many potential advantages of external economies are also called centripetal force. Alonso (1964) also provided a new interpretation of von Thünen's model. Alonso reinterpreted the classic von Thünen model to suit the modern context. Alonso replaced farmers with modern workers and isolated towns with business centers. By doing so, he revolutionized the spatial economic model. His model predicted the growth of the monocentric pattern of city structure, which is analogous to the concentric ring pattern of the original von Thünen model.

In contrast to agglomeration externalities, one of the essential ideas in urban economics is the force that works against the growth of a single-center city. Mills (1967) and Henderson (1974) were among the first scholars to consider cities' adverse effects. They suggested that higher density always translates to higher commuting costs, lowering the welfare analysis of workers in the big city. While agglomeration is related to centripetal force, the negative side of external economies is called centrifugal force. When the centripetal force leads to the monocentric formation of the city, its countervailing force, the centripetal force, leads to the breakdown of the monocentric pattern of the city. These two competing forces eventually form multiple cities with industry specialization.

Regional science is another branch of economics that investigates how economic activity is affected by location. Christaller (1933) and Lösch (1940) are among the first scholars that pioneered regional science and developed the Central-Place Theory to explain why most economic activity is in cities. Similar to the agglomeration externalities of urban economics, The theory suggested that transportation costs would be minimized, and scale economies would be maximized at the multiple central places in a hexagonal lattice where each central place represents a city in hierarchical order, ranging from the urban core to small towns serving villagers.

After Christaller and Lösch, Harris (1954) applied the idea of market potential to examine the concentration of industries. He found that industries in the United States concentrated in areas where market potential–the level of purchasing power of neighboring areas–is high. Although the early studies of urban economics and regional science touched on the city's agglomeration, these models are mainly stylized and assume perfect competition. Their model's treatment of agglomeration force is primarily implicit, ad-hoc, and sometimes exogenous. However, later generations of spatial economic models solved traditional models' limitations.

The works of Fujita et al. (2001) are the milestone in regional economics. Specifically, they developed explicit models of economic geography based on the Dixit-Stiglitz model of monopolistic competition. They developed the model to fix the theoretical problems of traditional urban economics and regional science. In their model, the interaction between distance, transportation costs, and the opposing forces of centrifugal and centripetal factors was demonstrated to shape the actual formation of urban structure.

Specifically, the objective of Fujita et al.'s (2001) work is to provide a better understanding of the formation of urban systems where interaction among socio-economic factors—population growth, transportation costs, and the concept of increasing return to scale—explicitly contribute to the hierarchical system of the city. Krugman's study laid the foundation for many subsequent economic literature to incorporate the increasing return of economic geography into their models. Indeed, following their work, many researchers attribute the sources of the increasing return of economic geography—the agglomeration externalities—to linkages, thick markets, knowledge spillovers, and pure external economies generated by the cities. For example, Duranton and Puga (2001) formulated a micro-foundation model of "nursery cities," the cities act as nurseries for innovative firms to develop and test their products or services.

Ciccone and Hall (1996) developed a framework to show how density and productivity are related. They showed that a higher spatial density is linked to a higher aggregate return through local externalities and the variety of local intermediate services. The work of Ciccone and Hall (1996) is consequential; the subsequent theoretical studies of agglomeration externalities are driven by micro-foundation models that use human capital and knowledge spillover as endogenous factors. After Ciccone and Hall, Duranton and Puga (2004) derived the micro-foundation theory as a basic framework for understanding the externalities of urban agglomeration. In their model, agglomeration externalities stem from the positive effects of sharing, matching, and learning, all of which are advantages for the city—from the early work of pioneers like von Thünen and Hoover to the work of contemporary urban economists like Krugman and Fujita, the study of urban economics included the core idea of increasing returns from geography or the agglomeration effect of cities on productivity and wages.

In addition to the previously mentioned scholars, O'Sullivan (2012) provided a comprehensive summary of the key concepts related to agglomeration effects, drawing from various studies. In his book, O'Sullivan identified four main categories of agglomeration sources contributing to the phenomenon. These categories are as follows: 1. Agglomeration effect resulting from sharing intermediate inputs. 2. Agglomeration effect arising from the availability of a larger labor pool. 3. Aggregation effect facilitated by improved labor matching. 4. Aggregation effect stemming from knowledge spillovers. According to O'Sullivan's simplified model, firms in agglomerated areas benefit from these mechanisms, leading to increased productivity and the ability to offer higher wages to their workers.

2.2.2 Empirical Literature

Examining the agglomeration effect is a challenging task for researchers, especially for those making early efforts to test the theory empirically. The difficulties lie in the data's limitations and the methodological challenges. The methodological issue arises from the theoretical framework and common sense, which suggests a strong positive feedback loop between urbanization and agglomeration externalities. In effect, all research aiming to estimate the agglomeration externalities on productivity must deal with the issue of endogeneity bias since researchers must naturally be interested in causality running from agglomeration to productivity. As a result, empirical literature frequently employs two-stage least squares estimation to address such a problem. Therefore, a feasible instrument variable (IV) is needed to be found.

Estimating the agglomeration effect of urbanization brings up another important issue: what variables best represent the total concentration of cities. Empirical studies used different proxies to measure the level of urban concentration, which resulted in different outcomes. O'Sullivan (2012) classified the representation of metropolitan areas into four classes. These four categories are as follows: 1. Urban Area; 2. City Population Density; 3. Metropolitan Area 4. The principal or municipal city. Indeed, previous research on agglomeration economies has emphasized the relative relevance of localization and urbanization economies or whether agglomeration economies result from industrial concentration or city size (Melo et al., 2009). As a result, in keeping with the paper's objectives, this section of the empirical literature review focuses on the agglomeration effects of urbanization economies.

The early study that explicitly estimated the agglomeration effect of urbanization was the study of Ciccone and Hall (1996). As discussed in the previous section, Ciccone and Hall formulated a model to show the relationship between employment density and productivity. Ciccone and Hall (1996) applied the Two-Stages Least Squares estimation to deal with the endogeneity problem. To reliably estimate the agglomeration effect, They used the historical population in the mid-19th century and the distance from the eastern seaboard of the U.S. as the IVs for current employment density (total urban concentration). They discovered that a 100 percent increase in employment density increased average labor productivity by around 6 percent. Furthermore, differences in the density of economic activities may explain more than half of the variation in worker productivity across the United States.

After the study of Ciccone and Hall, many researchers followed their technique of Two-Stages Least Squares and instrumental variables to estimate the agglomeration effect on the economic level. Indeed, Ciccone estimated the urban agglomeration effect in Europe. Ciccone (2002) applied a technique and set of European socio-economic data like his previous work but changed the IV from historical population to historical total land area in the nineteenth century. He obtained a similar result.

Recent studies focus on the micro-level effect of agglomeration on individual worker productivity (Au & Henderson, 2006; Fingleton, 2003, 2006; Rosenthal & Strange, 2008; Wheeler, 2001). Nevertheless, some studies are interested in investigating how agglomeration affects the aggregate level of productivity (Brülhart & Mathys, 2008; Rice, Venables, & Patacchini, 2006).

Instead of using the historical population, some studies resorted to an alternative IV to deal with endogeneity bias. The new method for measuring urban concentration used history and geology because history and geology are thought to be determinants of human settlement prior to the formation of the modern economy (Combes, Duranton, Gobillon, & Roux, 2010; Glaeser, Kerr, & Kerr, 2015; Rosenthal & Strange, 2008). Moreover, some studies used climate variables such as temperature, precipitation, longitude, and latitude to instrument current population density (Glaeser & Gottlieb, 2009).

Recently, the studies of agglomeration externalities shifted from developed countries to developing countries, such as those in Africa, Asia, and South America. For example, using Colombian Household Survey data, Duranton estimated the agglomeration effect in Columbia. Likewise, Matano, Obaco, and Royuela (2020), using Ecuadorean Labor Surveys, showed that agglomeration could be related to the spatial wage premium of Ecuadorean workers. In Asia, Chen et al. (2020) showed an agglomeration of externalities in cities through knowledge spillover, which helped firms innovate their products and services.

Most of the previous studies found a positive agglomeration effect. However, it does not necessarily entail urbanization's positive and universal impact on economic growth. Countries in Africa do not benefit from the agglomeration externalities of urbanization predicted by the theory and empirical evidence elsewhere. For example, Brückner (2012) did not find a self-reinforcing mechanism between GDP growth and the urbanization rate in African countries. In his work, he found that GDP per capita growth did not significantly affect the urbanization rate, and, worse still, an increase in the urbanization rate had a significant negative average effect on GDP per capita growth.

In the case of Thailand, Limpanonda (2012) estimated the magnitude of agglomeration externalities via a method derived by Ciccone and Hall. Limpanonda (2012) found a significant effect of agglomeration externalities on labor productivity. Specifically, Limpanonda (2012) found that doubling population density at the provincial level increased workers' wages by 16–37 percent. Likewise, Houbcharaun (2013) used firm-level data from the Industrial Census (IC) and the New Economic Geography framework to confirm the existence of the linkage between agglomeration externalities and differences in regional productivity.

In addition to the effect of agglomeration on labor productivity, some studies investigated the effect of agglomeration externalities on firm productivity. For example, Tippakoon (2011) applied two-stage least squares estimation with lagged explanatory variables as instruments to deal with endogeneity issues, whereas Puttanapong (2018) applied a spatial econometric model to estimate the effects of industrial agglomeration or the positive spillover of industry concentration. Both studies obtained positive effects of the agglomeration externalities on labor and firm productivity.

To the author's knowledge, no study empirically combines micro-level data such as LFS and IC with satellite data to quantify the effect of agglomeration externalities on labor force productivity via the latest technique currently employed in the field.

2.3 Theoretical Framework

The theoretical framework of this paper is based on Ciccone and Hall (1996). The model serves the objectives of this paper for two reasons. First, Ciccone and Hall developed the micro-foundation model of increasing returns to economic geography, predicting causality from increasing density to increasing workers' productivity. Second, the theoretical model could be tested empirically with the data limitations of developing countries where individual identifiers through time are lacking.

2.3.1 Ciccone and Hall Model

2.3.1.1 Externalities

In the first part of the model, the producers produce goods without capital but with land and labor. Land is described by acre of space (indexed by a and assumed to be equivalent) and labor is indexed by n. The total output is indexed by q. Thus, density in the model could be explicitly seen as $\frac{q}{a}$. The elasticity of output with respect to density is a constant term, $\left((\lambda - 1)/\lambda\right)$ and the elasticity of output with respect to number of workers is also a constant term, α. In mathematic term, the production function is described by equation (2.1).

$$(n, q, a) = n^{\alpha} \left(\frac{q}{a}\right)^{(\lambda - 1/\lambda)} \tag{2.1}$$

Assume that worker at county level, n_c, is distributed equally across the acre of space, a, in the county. Therefore, total output at the county level is described by equation (2.2).

$$q_c = a_c \left(\frac{n_c}{a_c}\right)^{\alpha} \tag{2.2}$$

where subscript c is the index of county level.

The county-wide output per acre is obtained by solving equation (2.2) for output; that is, the output density at county level is described by equation (2.3).

$$\frac{q_c}{a_c} = \left(\frac{n_c}{a_c}\right)^{\gamma} \tag{2.3}$$

where γ is the product of congestion cost (α) and agglomeration force (λ) or $\gamma = \alpha\lambda$. The output at county level could be aggregated to the state level which is $Q_s = \sum_{c \in \mathcal{C}_s} n_c{}^{\gamma} a_c{}^{-(\gamma-1)}$ where $\mathcal{C}_s$ is the subset of counties covering state s.

The average labor productivity in the state is described by the following equation.

$$\frac{Q_s}{N_s} = \frac{\sum_{c \in \mathcal{C}_s} n_c{}^{\gamma} a_c{}^{-(\gamma-1)}}{N_s} \tag{2.4}$$

where N_s is the number of workers in state s, and the state density index, $D_s(\gamma)$ is defied as the following equation:

$$D_s(\gamma) = \frac{\sum_{c \in \mathcal{C}_s} n_c{}^{\gamma} a_c{}^{-(\gamma-1)}}{N_s} \tag{2.5}$$

Therefore, equation (2.4) becomes.

$$\frac{Q_s}{N_s} = D_s(\gamma) \tag{2.4.1}$$

Equation (2.4.1) implies that average labor productivity in the state depends on the density index from equation (2.5).

The density index could be decomposed into three components:

$$D_s(\gamma) = D^{\gamma-1} \left(\frac{D_s}{D}\right)^{\gamma-1} \frac{\sum_{c \in \mathcal{C}_s} n_c \left(\frac{d_c}{D_s}\right)^{\gamma-1}}{N_s} \tag{2.6}$$

where d_c is employment per acre (employment density) in county c, $\frac{n_c}{a_c}$; D_s is employment per acre (employment density) in state s; and D is employment per acre (employment density) at national level.

According to equation (2.6), the state density effect is the product of a national output density, a state output density (relative to national density) and an inequality of density across counties within the state. It is commonly assumed that the agglomeration externalities outweigh congestion cost from increasing in density that is γ must be more than one. In effect, the area with higher average density and higher inequality of density will have higher labor productivity.

2.3.1.2 Intermediate Product variety

In the second part, an increasing return to scale stems from a greater variety of intermediate products in denser areas. Let us assume that the factors of production in the economy consist of intermediate goods and labor. The production function of final goods is the second model is described by equation (2.7)

$$f(m,i) = \left[m^{\beta} i^{(1-\beta)}\right]^{\alpha} \tag{2.7}$$

where m is the number of employed workers, i is the amount of composite service input which cannot be transported outside the acre, the constant term, α, is congestion effect of the two factors of production, β is the agglomeration effect. The composite service i is described by the following equation:

$$i = \left(\int_0^z x(t)^{1/\mu}\, dt\right)^{\mu} \tag{2.8}$$

From equation (2.8), the composite service i is produced from individual differentiated services, $x(t)$, where t indexes type of individual. z describes the variety of intermediate products. The parameter μ indicate level of substitutability of the intermediate products and μ is assumed to be more than 1. The high value of μ means high monopoly power of the producer (less one product substitutes for others).

Under the assumption of Bertrand competition of the standard Spence-Dixit-Stiglitz model, μ is the mark-up price that producer charges to maximize profit. The model further assume that it takes $x + v$ unit of labor to produce x. The intermediate producer pays w to labor. To maximize profit, the intermediate producer charges a price of μw and make a profit of $(\mu - 1)wx - wv$. Assuming perfectly competitive market condition, the profit would be zero; that is, the level of output would become.

$$x = \frac{v}{\mu - 1} \tag{2.9}$$

Inserting x into equation (2.8) yields equation (2.10)

$$i = z^{\mu} x \tag{2.10}$$

Equation (2.10) explicitly reveals relationship between density and productivity; that is, producer uses zx unit of intermediate inputs to produce composite service i, the productivity of the i production process is $z^{\mu-1}$. Since $\mu > 1$, the productivity rises with number of intermediate goods. The positive relationship between density and productivity persists because denser area could accommodate more intermediate service producers.

Since the production technology follows the Cobb-Douglas production function, the share of final output paid to labor employed directly in production of final product is $\alpha\beta$; therefore, $wm = \alpha\beta f(m, i)$. The share output paid to land must be $(1-\alpha)$. Recall that the total labor employed in the acre is n. In a free entry equilibrium, the share output paid to total worker must be the share output not paid to land; that is, α. Hence, $wn = \alpha f(m, i)$. In equilibrium allocation, the relationship between labor employed in making final product, m, and the total labor employed in acre, n, is governed by the following equation:

$$m = \beta n \tag{2.11}$$

The remaining $(1-\beta)$ n share of labor is used to make intermediate goods. Given the total amount of labor devoted to make intermediate goods and the amount of each one produced, we can show that number of intermediate products is proportional to the number of employed workers by solving for z:

$$z = (1-\beta)\frac{\mu - 1}{\mu}\frac{n}{\nu} \tag{2.12}$$

Inserting equation (2.12) into equation (2.10) to solve for i, and then insert equation (2.11) into equation (2.7). The consolidated production function emerges as

$$\phi n^{\gamma} \tag{2.13}$$

where ϕ is a complicated function of the other constants.

The elasticity of employed labor in the acre with respect to output is described by equation (2.14).

$$\gamma = \alpha[1 + (1-\beta)(\mu - 1)] \tag{2.14}$$

From equation (2.14), the congestion cost, α, the agglomeration externalities, β, and degree of substitutability, μ interact with each other. We can see that effect of congestion cost is outweighed by higher enough μ and low enough β. If this is the case, a higher density of employed workers would lead to higher productivity. Moreover, if we normalized ϕ to 1 and assume uniform distribution of labor across the acre of a county. The county production function is described by equation (2.15)

$$\frac{q_c}{a_c} = \left(\frac{n_c}{a_c}\right)^{\gamma} \tag{2.15}$$

where output per workers is determined by density of workers which is proportional to γ.

2.3.1.3 Capital and Total Factor Productivity

The third part, an increasing return to scale from greater density could be described by production function with capital and total factor productivity:

$$q_c = A_s\left[(e_c n_c)^{\beta} {k_c}^{1-\beta}\right]^{\alpha} \left(\frac{q_c}{a_c}\right)^{(\lambda-1)/\lambda} \tag{2.16}$$

where $A_s$is a Hicks-neutral technology multiplier for state s and e_s is labor efficiency at the county level. n_s and $k_c$is the quantities of labor and capital employed in a county. The value of β is between 0 and 1. α represents production elasticity while λ represents elasticity from externality where α is less than 1 and λ is greater than 1. The last term, $\left(\frac{q_c}{a_c}\right)^{(\lambda-1)/\lambda}$, is meant to represent externality from production density. Assuming that n_s and k_c are equally distributed across the acres in the county. Thus, the total output in county c is described by equation (2.17).

$$q_c = a_c A_s \left[\left(\frac{e_c n_c}{a_c}\right)^{\beta} \left(\frac{k_c}{a_c}\right)^{1-\beta}\right]^{\alpha} \left[\frac{q_c}{a_c}\right]^{(\lambda-1)/\lambda} \tag{2.17}$$

where the elasticity α is less than 1.

Solving equation (2.17) for output per acre yields

$$\frac{qc}{ac} = A_s^{\lambda} \left[\left(\frac{e_c n_c}{a_c}\right)^{\beta} \left(\frac{k_c}{a_c}\right)^{1-\beta}\right]^{\gamma} \tag{2.18}$$

The elasticity of factor of production to output per acre, $\gamma,$ is the product of production elasticity, α and the elasticity from externality, λ, that is $\gamma = \alpha\lambda$. Given that α is less than 1 (represents diminishing return to scale from congestion) and λ is greater than 1 (represents increasing return from agglomeration). Thus, if γ is more than 1, agglomeration externalities dominate congestion cost effect.

Since capital is rarely available for empirical estimation, we deal with it by assuming that the rental price of capital, r, equally incur to everyone. From factor demand function, substituting the factor price for the factor quality yield equation (2.19).

$$\frac{k_c}{a_c} = \frac{\alpha(1-\beta)}{r}\frac{q_c}{a_c} \tag{2.19}$$

Inserting equation (2.19) into equation (2.18) yields equation (2.20).

$$\frac{q_c}{a_c} = \phi A_s^{\omega}\left(\frac{e_c n_c}{a_c}\right)^{\theta} \tag{2.20}$$

where the elasticities for the technology multiplier for the state, ω, and for the labor input for the county,θ, are defined by equation (2.21).

$$\omega \equiv \frac{\theta}{\alpha\beta} \tag{2.21}$$

and

$$\theta \equiv \frac{\gamma\beta}{1-\gamma(1-\beta)} \tag{2.22}$$

Ciccone and Hall further assume that labor efficiency at county level depends log-linearly on workers' average year of education at county level, h_c ,that is, $e_c = h_c^{\eta}$, where η is the elasticity of level of education. From equation (2.22), aggregating the relationship to the state level yields equation (2.23).

$$\frac{Q_s}{N_s} = \phi A_s^{\omega} D_s(\theta,\eta) \tag{2.23}$$

where $D_s(\theta,\eta)$ is described by equation 2.24

$$D_s(\theta,\eta) = \frac{\sum_{c\,\in\,C_s}\left(n_c h_c^{\eta}\right)^{\theta} a_c^{1-\theta}}{N_s} \tag{2.24}$$

Assuming that state productivity A_s is log-normally distributed around an underlying nationwide level, and that measurement error of state productivity A_s also follows a log-normal distribution with zero mean. Taking log logarithms of equation (2.23) yields equation (2.25).

$$\log\frac{Q_s}{N_s} = \log\phi + \log D_s(\theta,\eta) + u_s \tag{2.25}$$

where u_s is the sum of measurement error.

Equation (2.25) implies that workers' productivity at the state level, state output per worker, depends on density at the state level. All the three sub-models: externalities model, intermediate product variety model, and total factor productivity model yield the same conclusion that higher density leads to higher workers' productivity.

2.3.2 Foundation of Agglomeration Externalities

According to the explanation provided by Combes and Gobillon (2015), the concept of profit for a firm is defined in equation (2.26) within the framework of microeconomic theory.

$$\pi_{r,t} = p_{r,t}Y_{r,t} - \omega_{r,t}L_{r,t} - c_{r,t}K_{r,t} \tag{2.26}$$

In equation (2.26), $\pi_{r,t}$ represents the profit in region r at time t. $Y_{r,t}$ denotes the output produced in that region and time. $L_{r,t}$ represents the amount of labor, while $K_{r,t}$ stands for the capital. The variable $p_{r,t}$ represents the price of the output, $\omega_{r,t}$ refers to the wage, and $c_{r,t}$ represents the return on capital.

Drawing from the Cobb-Douglas production function and the theoretical framework introduced by Ciccone and Hall in 1996, the mathematical representation of production can be expressed as follows:

$$Y_{r,t} = \frac{A_{r,t}}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}\left(s_{r,t}L_{r,t}\right)^{\alpha}{K_{r,t}}^{1-\alpha} \tag{2.27}$$

In equation (2.27), α represents the parameter of elasticity of substitution, and $A_{r,t}$ represents the total factor productivity. In a competitive equilibrium, the optimal production is determined by the first-order condition, which can be expressed as follows:

$$\omega_{r,t} = \left(p_{reg,t}\frac{A_{r,t}}{\left(c_{r,t}\right)^{1-\alpha}}\right)^{\frac{1}{\alpha}} s_{r,t} \equiv B_{r,t}s_{r,t} \tag{2.28}$$

The equation (2.28) suggests that the wage in region r is influenced by two factors: the labor skill, denoted as $s_{r,t}$, and the composite localized productivity, represented by $B_{r,t}$. This localized productivity concept is closely related to the main idea of the advantage of big cities proposed by Buchanan (1965), which emphasizes the spillover of public goods. Moreover, Lucas (1988) introduced an alternative framework that emphasizes the importance of spatial concentration of labor and industry in facilitating the spillover of knowledge. This knowledge spillover, in turn, contributes to enhanced productivity in specific geographic areas.

The examination of empirical evidence that explores the impact of labor skill and productivity on wages can be alternatively described as:

$$y_{r,t} = Z_{r,t}\gamma + \eta_{r,t} \tag{2.29}$$

In practical terms, the impact of labor skill and composite localized productivity on wages can be explored through empirical investigation using econometric techniques. This can be achieved by examining equation (2.29), where $y_{r,t}$ represents the logarithm of wage, $Z_{r,t}$ is a vector that combines the factors influencing composite localized productivity and localized labor skill, and $\eta_{r,t}$ represents the unobserved component represented as a residual. Equation (2.29) provides a mathematical expression that enables the empirical investigation of these factors and their effects on wages.

To account for the individual characteristics of labor, equation (2.28) is modified in equation (2.30) to include the skill of person i denoted as $s_{i,t}$, which partially determines the individual wage, $\omega_{i,t}$. The modified equation (2.30) is as follows:

$$\omega_{r,t} = B_{r,t}s_{i,t} \tag{2.30}$$

Like the transition from equations (2.28) to (2.29), Equation (2.30) can also be converted into equation (2.31), establishing the framework for empirical validation using econometric methods. This modification, initially proposed by Glaeser and Maré (2001), results in the following expression:

$$y_{i,t} = u_i + X_{i,t}\theta + Z_{r(i,t),t}\,\gamma + \eta_{r(i,t),t} + \epsilon_{i,t} \tag{2.31}$$

where $X_{i,t}$ and u_i denote the characteristics of workers, and $\epsilon_{i,t}$ represents the residual.

Equation (2.32) and (2.33) is applied in empirical literature to estimate the effect of agglomeration externalities.

$$y_{i,t} = u_i + X_{i,t}\theta + \beta_{r(i,t),t} + \epsilon_{i,t} \tag{2.32}$$

$$\beta_{r,t} = Z_{r,t}\gamma + \eta_{r,t} \tag{2.33}$$

In equation (2.33), the region-time fixed effect, $\beta_{r,t}$, is introduced to quantitatively capture the agglomeration externalities. It plays a crucial role in determining the individual wage along with the individual characteristics presented in equation (2.32).

To account for the influence of diverse firms, the two-step approach can be expanded to incorporate the industry-specific effect. This is demonstrated in equation (2.34) and (2.35):

$$y_{i,t} = u_i + X_{i,t}\theta + \beta_{r(i,t),s(i,t),t} + \epsilon_{i,t} \tag{2.34}$$

$$\beta_{r,s,t} = Z_{r,t}\gamma_s + \eta_{r,s,t} \tag{2.35}$$

where s is the index for industry. Equation (2.34) introduces $\beta_{r(i,t),s(i,t),t}$ as the region-industry-time fixed effect, which contributes partially to the setting of individual wages. On the other hand, equation (2.35) investigates the effects of agglomeration externalities, represented by $Z_{r,t}$, on the region-industry-time fixed effect, $\beta_{r,s,t}$.

This study successfully combined the latest surveys and geospatial data to create comprehensive region-industry-time dimensions. As a result, the following section provided a detailed explanation of the empirical approach, which was built upon the theoretical foundations outlined in equation (2.34) and (2.35).

2.4 Methodology

Building on the studies by Groot, de Groot, and Smit (2014) and Ridhwan (2021), a two-stage estimation technique was utilized to determine the factors contributing to regional wage disparities. In the first stage, a regression analysis was performed, relating workers' wages to their individual attributes using the Mincer earnings function, which incorporates variables such as education level, age, gender, industry, province, and year (Mincer, 1974). In the second stage, another regression analysis examined the relationship between local productivity and externalities arising from urbanization and localization economies.

2.4.1 Mincer Equation

$$log(\omega_{r,t}) = \alpha + \sum_{edu} B_{1,edu} D_{i,t}^{edu} + \beta_2 age_i + \beta_3 age_i^2 + \beta_4 D_i^{gender} + \sum_{ind} B_{5,ind} D_{i,t}^{ind} + \sum_{r} B_{6,r} D_{i,t}^{r} + \sum_{year} B_{7,year} D_{i,t}^{year} + \varepsilon_{i,t} \quad (2.36)$$

Equation (2.36) presents the standard mathematical representation of the Mincer equation, incorporating workers' characteristics, industry fixed effects, regional fixed effects, and year-fixed effects. The education level of workers was controlled by introducing a dummy variable representing their highest educational attainment. The workers' experience level was approximated by including their age and age-squared to account for a nonlinear relationship between experience and wages. Gender bias was addressed by including a dummy variable for gender in the model. Additionally, industry, region, and year were dummy variables, as workers chose industries and locations based on their skills and preferences. Notably, equation (2.36) separately estimates the industry, region, and year dummy variables. However, the industry, region, and year-fixed effects are combined and estimated in the subsequent step.

2.4.2 First-Stage Regression

$$log(\omega_{r,t}) = \alpha + \sum_{edu} B_{1,edu} D_{i,t}^{edu} + \beta_2 age_i + \beta_3 age_i^2 + \beta_4 D_i^{gender} + \sum_{ind} \sum_{r} \sum_{year} \gamma_{ind,r,t} D_{i,t}^{ind} D_{i,t}^{r} D_{i,t}^{year} + \varepsilon_{i,t} \quad (2.37)$$

In order to differentiate the impact of workers' characteristics from the fixed effects of industry, region, and year, equation (2.36) was slightly modified and transformed into equation (2.37) above. Instead of estimating region and year fixed effects separately, equation (2.37) combines the fixed effects of industry, region, and year. This combined set of fixed effects, referred to as spatial residuals, accounts for the effects of workers' characteristics, industrial composition, location, and time on regional productivity (Combes, Duranton, & Gobillon, 2008; Groot et al., 2014; Ridhwan, 2021). These residual terms represent the regional productivity after adjusting for the abovementioned factors.

2.4.3 Second-Stage Regression

$$\gamma_{ind,r,t} = \alpha + \beta_1 lnDensity_{r,t} + \beta_2 Specialization_{ind,r,t} + \beta_3 Diversity_{r,t} + \beta_4 Competition_{ind,r,t} + \beta_5 Area_r + \sum_{year} \beta_{6,year} D_{r,t}^{year} + \varepsilon_{ind,r,t} \tag{2.38}$$

Equation (2.38) explains for the disparities in spatial residuals by incorporating agglomeration variable externalities. The dependent variable in equation (2.38) is derived from equation (2.37). The measurements of agglomeration externalities include representatives of urbanization economies and proxies of localization economies. The specific computational details for constructing each variable will be discussed in the subsequent section. Additionally, to account for the effects of political boundaries and province size, the area of each province is included as a control variable.

As mentioned in the literature review section, applying the Ordinary Least Squares (OLS) method to estimate equation (2.38) can lead to inconsistent estimation of urbanization economies. Equation (2.38) was estimated using the OLS and IV techniques to address this issue. In addition to the potential simultaneity bias, there is also a risk of omitted variable bias arising from unobserved regional characteristics that may affect regional productivity. Following the approach of Combes et al. (2010), Groot et al. (2014), and Ridhwan (2021), the IV technique was employed to overcome the simultaneity bias and mitigate the omitted variable bias concerns.

2.5 Data

Data from the LFS and the IC were obtained from the NSO of Thailand for this study. The LFS dataset covers the period from 1997–2021. It provides comprehensive information on important worker characteristics such as wages, hours of work, employment status, gender, age, industrial classification, workplace location (province), and sampling weights. The analysis included workers aged 15 and above, while those identified as self-employed or working without payment were excluded. It is important to note that the LFS data does not include migration information; therefore, the impact of migration is not considered in this study.

In addition to the LFS, data obtained from the IC were also employed in this study. The IC dataset provided information on individual characteristics of industrial firms, which were further categorized into 24 production sectors. Specifically, the study utilized IC data collected in 2007, 2012, and 2017. Both the LFS and the IC datasets utilized in this study adhered to the same industrial classification system, namely the International Standard Industrial Classification (ISIC) Revision 4[1]. The compatibility between the two surveys allows for their integration, resulting in a comprehensive dataset encompassing both individual and industry-specific characteristics (Appendix B provides more detail about specification and technical reference of geospatial and economic data). To the best of our knowledge, this study represents the first attempt to combine official surveys, geospatial data, and satellite-based indicators to investigate the productivity impact of urban agglomeration in Thailand.

2.5.1 Dependent Variable

In the initial regression, the hourly wage of workers was employed as the dependent variable to estimate their productivity, as demonstrated in Appendix A (the empirical validation is provided therein). To ensure consistency, the monthly wage data from the LFS was converted into an hourly wage for the workers. In the subsequent regression analysis, the dependent variable was the regional productivity derived from equation (2.37), which adjusted for workers' characteristics and the composition of industries.

2.5.2 Independent Variable in the First-Stage Regression

The explanatory variables in the second-stage regression analysis encompassed the individual attributes of workers, including their education level, age, gender, industrial classification, and workplace location (province). In order to ensure compatibility and consistency across all surveys, the education level of workers was

[1] More detail about the ISIC classification systems is publicly available at https://unstats.un.org/unsd/classifications/Econ/isic

reclassified into seven distinct categories: (1) less than elementary education, (2) primary education, (3) lower secondary education, (4) secondary education, (5) post-secondary education, (6) bachelor's degree, and (7) graduate degree (Masters and PhD).

2.5.3 Independent Variable in the Second-Stage Regression

This subsection presents the mathematical definitions of the four variables representing agglomeration externalities, which were subsequently incorporated into the second-stage regression analysis. The agglomeration externalities can be categorized into two types: urbanization economies and localization economies. Urbanization economies, as described by Combes and Gobillon (2015), refer to the overall positive external effects of increasing population density in a specific area. On the other hand, localization economies represent the positive external effects of industrial specialization within a specific geographical boundary. Following the methodologies introduced by Groot et al. (2014) and Ridhwan (2021), measurements of urbanization and localization economies were constructed using microdata from the IC and geospatial data.

2.5.3.1 Urbanization Economies

To encompass the overall impact of urbanization economies on agglomeration externalities, population density was defined as follows:

$$Density_{r,t} = \frac{Population_{r,t}}{Area_r} \tag{2.39}$$

where $Population_{r,t}$ is the total population in province *r* in year *t* obtained from the WorldPop Global Project Population Data, and $Area_r$ represents the total area (in square kilometers) of province *r*.

In technical term, the total number of populations was acquired from the WorldPop Global Project Population Data, which is publicly available through the Google Earth Engine platform service[2]. The WorldPop Global Project Population maps were generated by integrating national census data with different geospatial covariate layers. The combined dataset underwent further processing using the random

[2] The dataset can be accessed through the following link: https://developers.google.com/earth-engine/datasets/catalog/WorldPop_GP_100m_pop

forest distribution mapping method. Through these computational techniques, up-to-date and high-resolution population maps were produced, with a grid size of 100 x 100 meters.

2.5.3.2 NTL Density

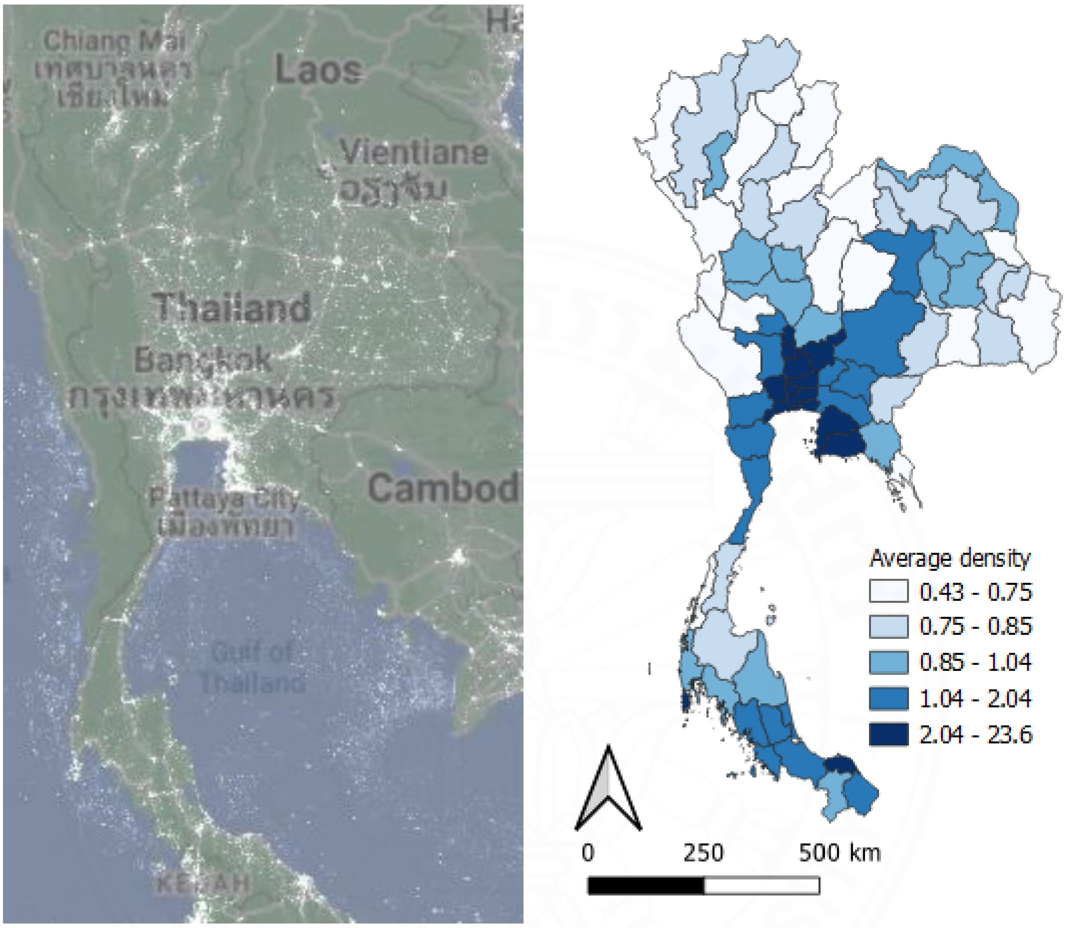
In addition to the traditional method of measuring local density discussed in Section 2.5.3.1, an alternative approach was adopted in this study by utilizing NTL density. The NTL density, spanning from 2000–2021, was sourced from the Google Earth Engine service. The comprehensive studies conducted by Huang, Yang, Gao, Yang, and Zhao (2014) and Li, Zhao, and Li (2016) demonstrate that NTL has emerged as a widely adopted satellite-based indicator across various fields. By capturing human-generated light from the DMSP/OLS and VIIRS/DNB satellites during nighttime, NTL provides a reliable measure of urban area extent (L. Chen, Hasan, & Jiang, 2022; J. Vernon Henderson & Kriticos, 2018; Zhang & Seto, 2011). Notably, research specific to Thailand by Sangkasem and Puttanapong (2022) and Puttanapong, Luenam, and Jongwattanakul (2022) established a statistically significant association between NTL density and population. Puttanapong, Martinez, et al. (2022) also corroborated this relationship using machine learning techniques.

As evidenced by numerous studies, NTL density can serve as a viable alternative measure for indicating population density on a global scale. The application of NTL has particular significance in the context of Thailand, where conventional population statistics may be subject to measurement errors due to the substantial presence of an informal sector within urban areas. Poonsab, Vanek, and Carré (2019) discovered that informal employment in urban areas constituted 34 percent of Thailand's total employment in 2017. Furthermore, the seasonal migration of informal labor between rural and urban areas, driven by crop cycles, has contributed to discrepancies in population statistics (Amare, Hohfeld, Jitsuchon, & Waibel, 2012; Puttanapong, 2008; Suttiwichienchot & Puttanapong, 2014). Consequently, the NTL index effectively captures the current population and urban density characteristics.

Figure 2.4

Visualization of Raw and Transformed NTL Data in 2021

Panel A Panel B



Note. (i) The dataset is publicly accessible at https://code.earthengine.google.com/0739f7e8922e020fbcff35fe134092e9. (ii) Panel A illustrate raw data of NTL. Panel B displays transformed NTL data.

Figure 2.4 displays maps showcasing the NTL density in Thailand, generated using the original data acquired from the Google Earth Engine service and the transformed provincial NTL index (shown in panels A and B, respectively). These maps highlight the presence of highly urbanized provinces, with Bangkok and its surrounding areas prominently featured.

2.5.3.3 Localization Economies

In addition to quantifying the impact of urbanization, this section provides a comprehensive explanation of the mathematical formulation for the three variables that represent localization economies: 1. intra-sectoral knowledge externality (also known as Marshall-Arrow-Romer spillover); 2. cross-sectoral knowledge externality (referred to as Jacobs spillover); and 3. local intra-sectoral competition externality (referred to as Porter spillover).

The impact of intra-sectoral knowledge externality (also known as Marshall-Arrow-Romer spillover) is determined based on data extracted from the IC and can be defined as follows:

$$Specialization_{ind,r,t} = \frac{E_{ind,r,t}}{E_{r,t}} \tag{2.40}$$

In equation (2.40), $E_{ind,r,t}$ denotes the total employment of industry ind in province r in year t. $E_{r,t}$ represents the total employment in province r in year t. A higher value of $Specialization_{ind,r,t}$ indicates a greater employment share of industry ind within province r.

To quantify the cross-sectoral knowledge externality, known as Jacobs spillover, Shannon's entropy was employed to estimate the level of industrial diversity. This estimation, expressed mathematically in equation (2.41), captures the degree of industrial diversity.

$$Diversity_{r,t} = -\sum_{ind}\left(\frac{E_{ind,r,t}}{E_{r,t}} ln\ \frac{E_{ind,r,t}}{E_{r,t}}\right) \tag{2.41}$$

In the equation (2.41), $E_{ind,r,t}$ represents the total employment of industry ind in province r in a specific year t, and $E_{r,t}$ denotes the total employment in province r in the year t. By calculating the summation across industries, the value of $Diversity_{r,t}$ was obtained. The $Diversity_{r,t}$ serves as a measure of the cross-sector knowledge externality within each province. A higher value of $Diversity_{r,t}$ indicates a greater diversity of industries present in province r in that year. Conversely, a lower value of Shannon's entropy suggests a higher degree of specialization within certain industries in province r.

The final type of knowledge spillover considered is the externality effect resulting from competition between industries, referred to as Porter spillover. Mathematically defined in equation (2.42), this formula quantifies Porter spillover and can be seen as a modified version of the Herfindahl-Hirschman Index. The expression is presented below:

$$Competition_{ind,r,t} = 1 - \sum_{f} \left(\frac{E_{f,ind,r,t}}{E_{ind,r,t}} \right)^2 \tag{2.42}$$

In equation (2.42), $E_{f,ind,r,t}$ represents the total employment of firm f in industry ind in a specific year t. The Herfindahl-Hirschman Index is calculated as the summation across firms within each combination of industry and province. A value close to one indicates a high level of competition among firms within specific industries. Conversely, a value close to zero suggests a high concentration of employment within a few dominant firms.

2.5.3.3 Instrumental Variable

When productivity and local density are interdependent, estimating the agglomeration effect of local density can lead to inconsistent coefficient estimations. To address this issue, the instrumental variable (IV) technique, as recommended by Wooldridge (2010), was employed. The IV technique mitigates the simultaneity bias by selecting valid instruments that satisfy relevance and exogeneity conditions. The exogeneity condition requires that the chosen instrument is not correlated with any missing variables in the model and does not directly influence the outcome variable, which is regional productivity.

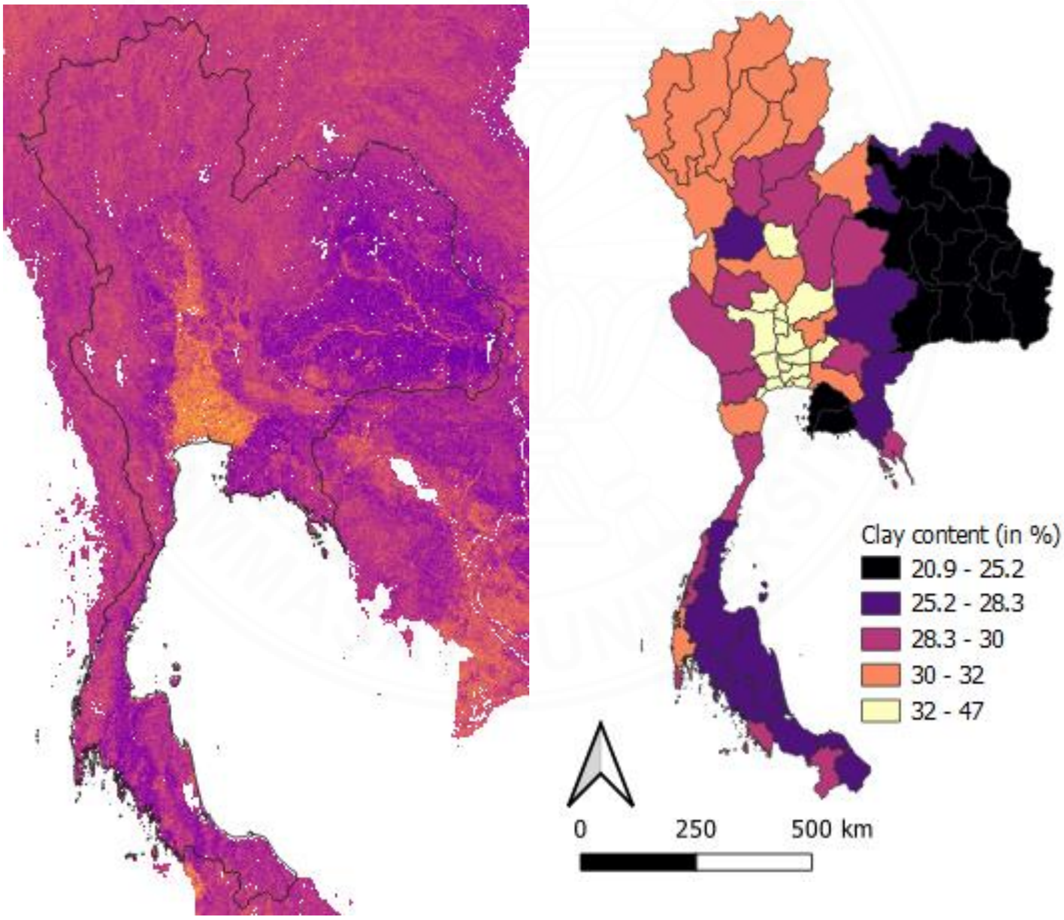
As extensively discussed in the literature by Combes and Gobillon (2015) and L. Chen et al. (2022), two conditions must be satisfied when selecting a suitable IV for empirically estimating agglomeration elasticity. Firstly, the chosen variable should exhibit the persistence of population distribution, meaning it should maintain the consistent geographic pattern of urbanization and exhibit a strong positive correlation with the current spatial density of human settlements and economic activities. Secondly, the selected variable should be mathematically orthogonal to the error term, indicating that the instrument affects productivity solely through population density and has no other influence paths.

In this study, the approach recommended to address the issue of simultaneity bias was adopted. Taking inspiration from Combes et al. (2010), the clay content in the soil was employed as the instrumental variable. In their research on productivity in the French context, Combes et al. (2010) utilized the soil content map as an instrument. They highlighted the exogeneity property of soil characteristics, explaining that natural factors determine these attributes. During the 15th century, the settlement pattern of the Tai ethnic group, the ancestors of present-day Thai people, was notably influenced by the clay content in the soil. This historical period saw rice cultivation as a prominent economic activity, and the clay content proportion in the soil primarily determined the establishment of cities. Consequently, the fertile lowlands of the Chao Phraya basin in the central region of Thailand emerged as the focal point of human settlements for several centuries (Baker & Phongpaichit, 2014).

Figure 2.5

Visualization of Raw and Transformed Data Percentage of Clay Content in the Soil

Panel A Panel B



Note. (i) The dataset is publicly accessible at https://developers.google.com/earth-engine/datasets/catalog/OpenLandMap_SOL_SOL_CLAY-WFRACTION_USDA-3A1A1A_M_v02. (ii) Panel A illustrate raw data of clay content in soil. Panel B displays transformed data of clay content in soil.

While studies showed that soil fertility can influence rural poverty in Africa, it does not play a significant role in urban economies, where poverty tends to be temporary and more dynamic (Barrett & Bevis, 2015). Similarly, in Thailand, research has indicated that soil fertility does not impact the current urban setting because economic activities in urban areas are predominantly driven by service and industrial sectors, rendering land fertility less relevant (Suphannachart, 2017; World Bank, 2021). These characteristics align with the two conditions for selecting an appropriate instrumental variable.

Figure 2.5 presents the spatial distribution of soil content percentage. Panel (A) displays the original data acquired from Google Earth Engine, while panel (B) showcases the transformed provincial average depicting the geographic distribution of clay content density. The central region stands out with its high clay content density, conducive to rice cultivation since the 15^{th} century.

2.6 Result

2.6.1 Descriptive Statistics

This section presents the descriptive statistics of the variables utilized in the first- and second-stage regression. Table 2.4 displays each variable's mean and standard deviation for the years 2007, 2012, and 2017, as well as the pooled cross-section. Over 2007–2017, there was a consistent increase in hourly wages, which served as the dependent variable. This rise in hourly wages corresponded with a noticeable growth in the proportion of workers with a tertiary education, which experienced an increase of approximately four basis points. Despite a gradual decline, workers without a bachelor's degree still constituted around 88 percent of the total workforce in 2017. This suggests that Thailand's manufacturing industry continued to employ a significant number of low-educated workers for an extended period.

Table 2.4

Summary of Statistics in the First-Stage Regression

First-Stage Regression	2007	2012	2017	Pooled
Observations	12,403	12,374	10,453	35,230
Continuous variables				
Log hourly wage	3.32	3.70	3.93	3.63
	(0.66)	(0.50)	(0.50)	(0.64)
Age	34.66	35.22	37.05	35.56
	(11.11)	(10.49)	(11.08)	(10.93)
Categorical variables				
Female	52.75%	50.83%	49.44%	51.06%
Less than primary-post secondary education	92.44%	90.36%	88.57%	90.56%
Tertiary education	7.55%	9.63%	11.42%	9.43%

Note. Standard deviations are reported in parentheses.

In 2017, the average age of the workers included in the sample was approximately 2.5 years higher compared to 2007. There was a slight change in the proportion of unemployed females in the manufacturing sector. The female labor force participation rate declined from 52.8 percent in 2007 to 49.4 percent in 2017. Table 2.5 illustrates the gradual increase in agglomeration measures, namely population density and NTL density, over time. Of particular note was the significant rise in NTL density, indicating that the density of NTL was more responsive to changes in urban areas than population density. Moreover, population density based on census data often excludes workers in the informal sector and those who migrate seasonally.

Table 2.5

Summary of Statistics in the Second-Stage Regression

Second-Stage Regression	2007	2012	2017	Pooled
Observations	714	949	966	2,629
Log population density	4.95	5.01	5.06	5.06
	(1.07)	(1.11)	(1.17)	(1.12)
Log NTL density	2.14	2.39	2.68	2.43
	(0.7)	(0.71)	(0.55)	(0.69)
Specialization	0.10	0.07	0.07	0.08
	(0.13)	(0.11)	(0.11)	(0.11)
Diversity	1.73	2.11	2.09	2.00
	(0.30)	(0.50)	(0.44)	(0.45)
Competition	0.75	0.76	0.77	0.76
	(0.24)	(0.25)	(0.24)	(0.24)

Note. Standard deviations are reported in parentheses.

In contrast to the consistent changes observed in the attributes of workers and local density, the employment composition within the Thai manufacturing sector remained relatively unchanged. The indicators of sectoral knowledge externalities, including specialization, diversity, and competition, exhibited a stable pattern between 2007 and 2017. In addition to the descriptive statistics, Figure 2.6 presents the temporal evolution of important variables (median hourly wage and population density) from 2007–2017.

Figure 2.6

Visualization of Hourly Wage and Population Density in 2007 and 2017

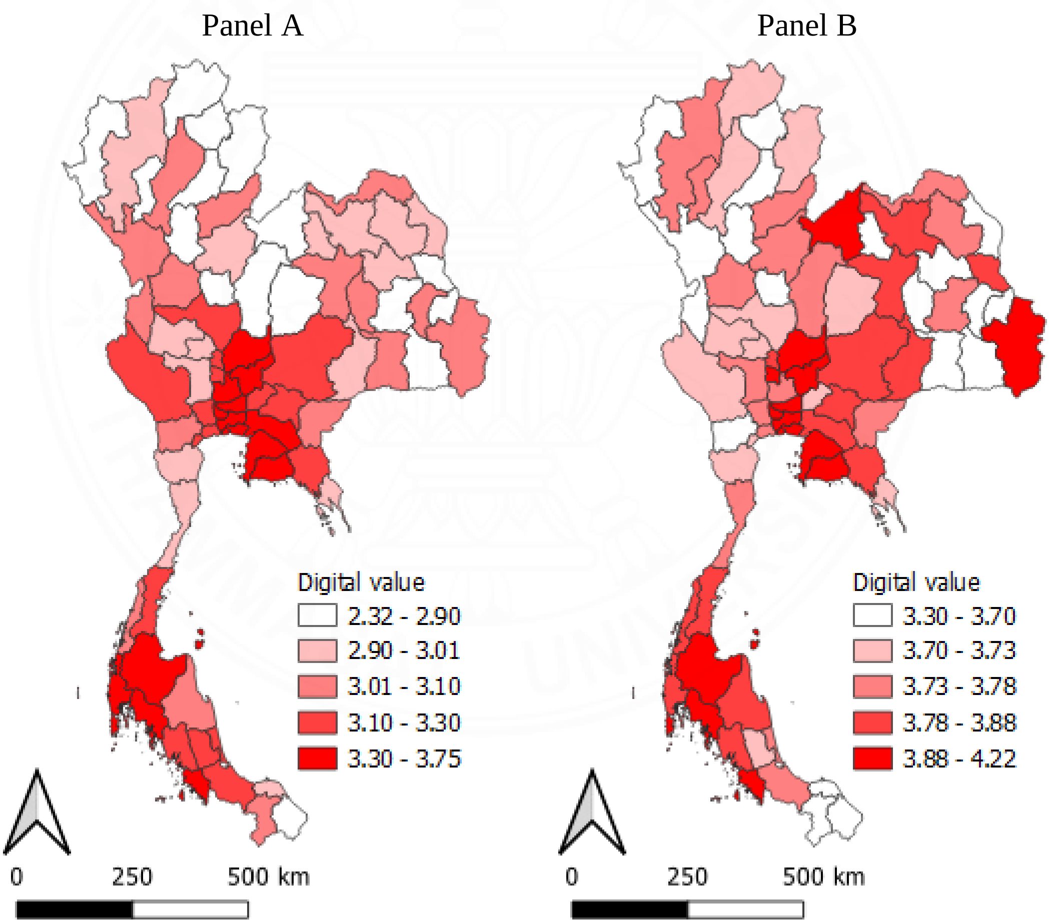
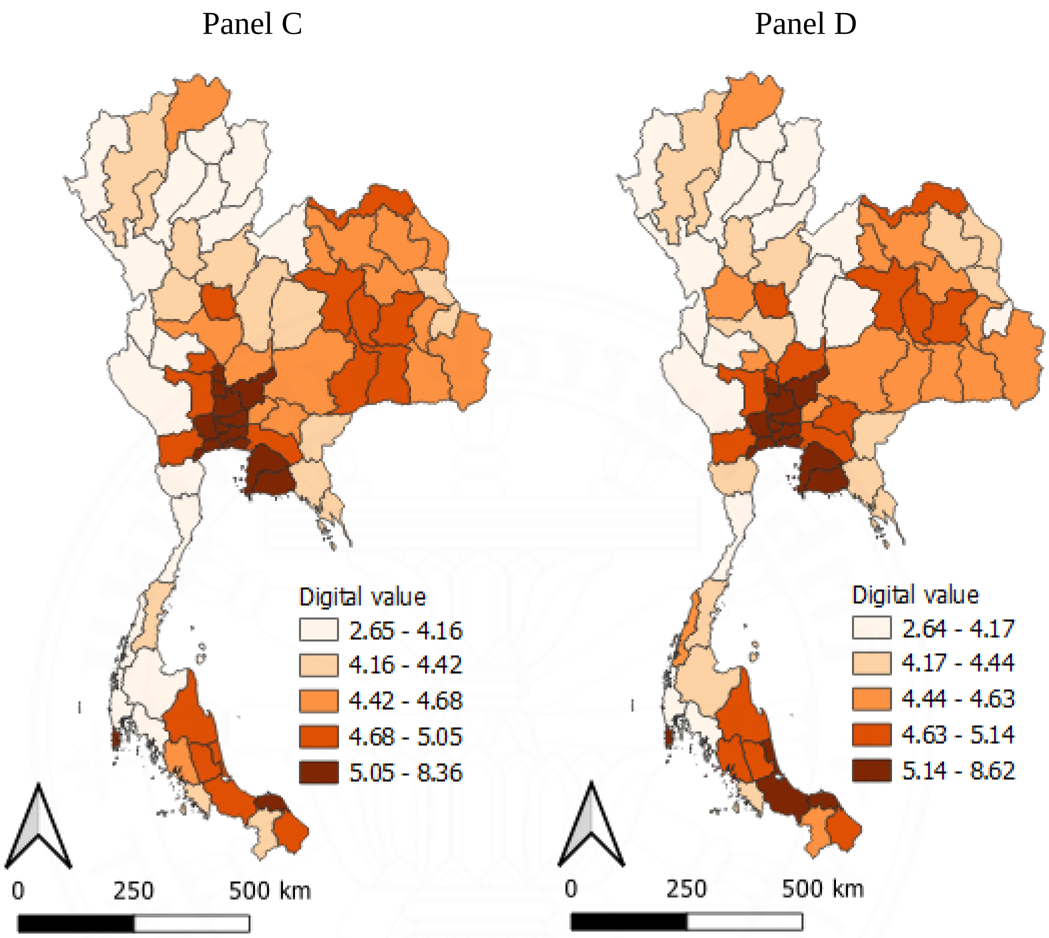


Figure 2.6

Visualization of Hourly Wage and Population Density in 2007 and 2017 (Cont.)



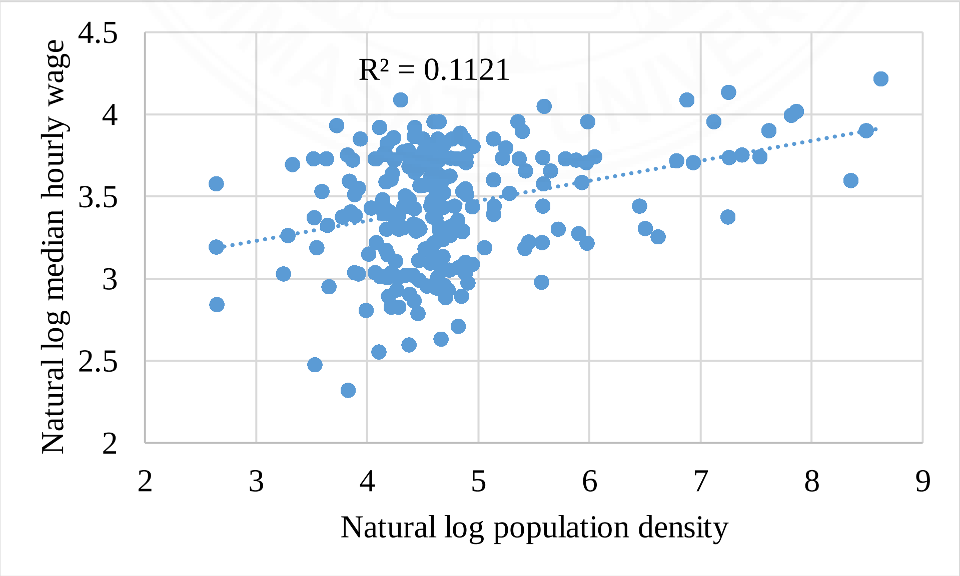
Note. (i) Panel A: Median log hourly wage in 2007; Panel B: Median log hourly wage in 2017; Panel C: Log population density in 2007; Panel D: Log population density in 2017. (ii) In panels A and B, the darkest red areas represent provinces with highest hourly wage, while the white areas represent provinces with lowest hourly wage. In panels C and D, the darkest orange areas represent provinces with highest population density, while the lighest orange areas represent provinces with lowest population density.

Panels A and B of Figure 2.6 demonstrate that provinces in Thailand's central, eastern, and southern regions exhibited the highest median hourly wages for workers. In contrast, provinces in the northern and northeastern regions had the lowest median hourly wages. Notably, there were positive developments in more recent years, as some provinces in the northeastern region advanced to a higher quintile of hourly wages. Figure 2.6 also includes panels C and D, which depict Thailand's provincial-level distribution of population density. It is evident that provinces in the central and eastern regions consistently had the highest population density in 2007 and 2017. The spatial distribution of population density aligned with the pattern observed for hourly wages, as the central and eastern regions were the most densely populated and the most productive regarding wages. In summary, Figure 2.6 demonstrates that most provinces in Thailand's central and eastern regions experienced higher hourly wages, corresponding to their increased population.

Before discussing the findings of the first- and second-stage regression, a simple correlation analysis was conducted to examine the relationship between the key agglomeration measures of local density and the logarithm of median hourly wage at the provincial level. The results are depicted in Figures 2.7 and 2.8. Figure 2.7 displays a positive correlation between population density and median hourly wage, with an R-squared value of 0.11. However, a more pronounced pattern emerged when NTL density was used to measure local density. Figure 2.8 reveals a stronger positive correlation between NTL density and median hourly wage, with an R-squared value of 0.42. In essence, NTL density could explain almost half of the variation observed in the median hourly wage.

Figure 2.7

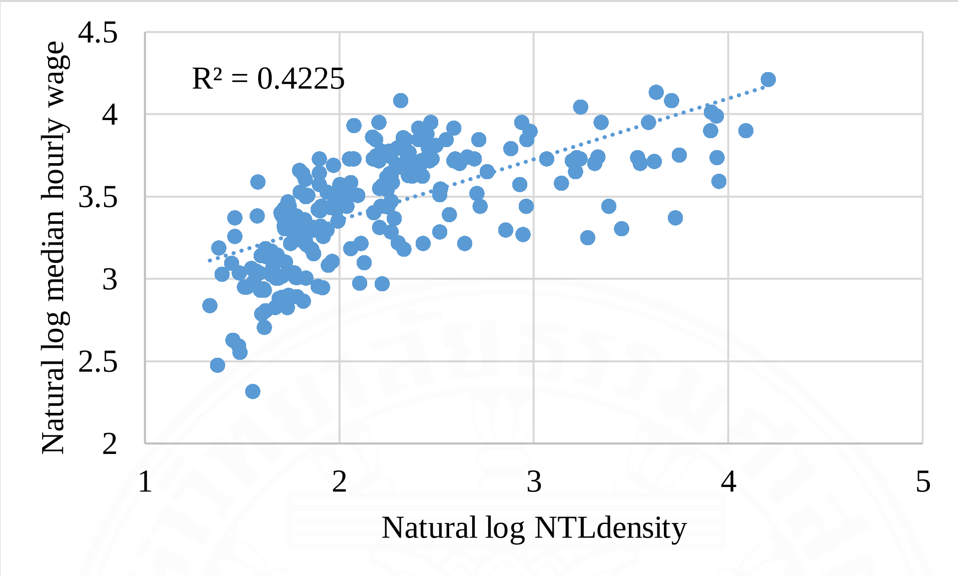
Scatter Plot of Population Density and Hourly Wage



Note. The table was created from author's calculations

Figure 2.8

Scatter Plot of NTL Density and Hourly Wage



Note. The table was created from author's calculations

2.6.2 Mincer Regression Results

The Mincer equation was utilized to investigate how the attributes of workers affect the wage gap. The findings from the estimation were provided in Table 2.6 for the years 2007, 2012, and 2017 separately, along with the results for the combined cross-section spanning from 2007–2017. The regression analysis did not consider individuals with less than a primary education or no formal schooling. The results indicated that workers' educational attainment and work experience significantly influenced their wages. Individuals with higher levels of education and more extensive work experience tended to earn higher hourly wages, suggesting that regional disparities in the composition of skills among workers contribute to spatial inequality. The study observed a gradual decline in the benefits of tertiary and primary education between 2007 and 2012 but a slight increase in 2017. The estimated coefficients for the impact of education and gender align with previous research (Tangtipongkul, 2015; Vivatsurakit & Vechbanyongratana, 2020; Warunsiri & McNown, 2010) regarding both magnitude and direction.

Table 2.6

Results of Mincer Regression (Equation (2.36))

Dependent: ln hourly wage	2007	2012	2017	Pooled
Age	0.05***	0.04***	0.03***	0.04***
	(27.31)	(10.53)	(16.62)	(38.27)
Age squared	-0.0006***	-0.0005***	-0.0003***	-0.0005***
	(-22.92)	(-17.84)	(-12.69)	(-31.77)
Female	-0.20***	-0.15***	-0.16***	-0.18***
	(-21.85)	(-19.34)	(-21.72)	(-36.63)
Education Dummies[3]				
Primary	0.15***	0.08***	0.11***	0.12***
	(9.99)	(6.57)	(8.64)	(14.93)
Lower secondary	0.27***	0.18***	0.18***	0.22***
	(16.72)	(12.97)	(14.05)	(26.03)
Upper secondary	0.40***	0.26***	0.29***	0.33***
	(24.01)	(19.36)	(22.12)	(38.68)
Post-secondary education	0.68***	0.49***	0.50***	0.56***
	(31.58)	(28.36)	(30.48)	(52.36)
Bachelor's degree	1.17***	0.94***	0.89***	1.00***
	(56.38)	(55.39)	(58.33)	(97.25)
Graduate (Master and PhD)	1.98***	1.52***	1.42***	1.63***
	(32.32)	(34.91)	(33.25)	(56.74)
Constant	2.35***	2.95***	3.29***	2.60***
	(52.22)	(73.82)	(85.74)	(107.69)
Industry dummies	Included	Included	Included	Included
Year dummies	Not Included	Not Included	Not Included	Not Included
Province dummies	Included	Included	Included	Included
R squared	0.50	0.50	0.51	0.56
Number of observations	12,403	12,374	10,453	35,230

Note. t-statistics are reported in parentheses.

* $p<0.05$, ** $p<0.01$, *** $p<0.001$.

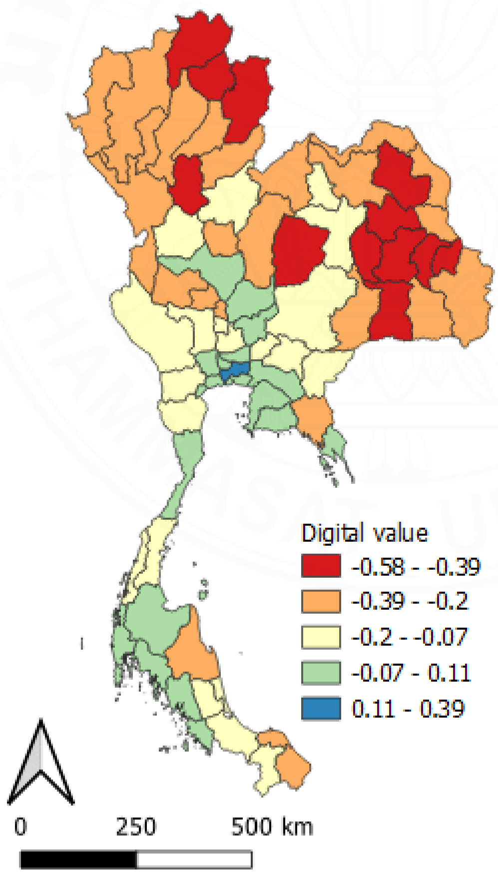
[3] Education dummies denote the highest qualification obtained. The omitted education category is workers who have less than primary education.

Tangtipongkul (2015) also documented a comparable decline in the benefits of education, providing additional evidence for the effectiveness of the Mincer equation in elucidating wage trends in Thailand. The findings emphasize that individual attributes, such as education and experience, accounted for around 50 percent of the wage variability. Notably, the coefficients derived from the analysis suggest a decreasing wage disparity between individuals with high and low levels of education over time. Nevertheless, regional wage gaps persist, indicating the necessity for a secondary regression analysis incorporating localized external factors to understand and explain regional wage variations.

2.6.3 Second-Stage Regression Results

Figure 2.9

Spatial Distribution of Average Spatial Residual



Note. The average spatial residual represents the provincial average wage of workers after considering their observed characteristics and the composition of industries. It reflects the wage level that remains unexplained by these factors and can be interpreted as the regional wage component.

Figure 2.9 exhibits the mean spatial residual among provinces, which, following the methodology employed by Groot et al. (2014), represents the average wage within each region after adjusting for observed worker attributes and sectoral distribution. Alternatively, the provincial average spatial residual could be perceived as a regional wage premium, as Ridhwan (2021) discussed. The spatial residuals were standardized using the nationwide average value.

Figure 2.9 illustrates that after considering skill and experience, BMR workers (highlighted in blue) and the eastern and southern regions (highlighted in green) received a more significant wage premium. It is worth mentioning that the spatial distribution pattern of the spatial residuals identified in this study exhibits similarities to those observed in the Netherlands and Indonesia, where Amsterdam and Jakarta, respectively, were linked to the highest wage premiums (Groot et al., 2014; Ridhwan, 2021). Table 2.7 shows the results of the second-stage regression using both OLS and IV methods. According to the OLS approach, when the population density doubles, there is an associated increase of 0.07 units in wage premium. However, the IV estimate revealed that the impact of local density on wage premiums rose to 0.14. This indicated that endogeneity did not lead to an upward bias in the OLS estimate. Similar findings were reported in other studies, where the IV estimate demonstrated an amplified effect of local density on wages (Barufi, Haddad, & Nijkamp, 2016; Combes, Duranton, & Gobillon, 2008; Groot et al., 2014; Ridhwan, 2021).

Table 2.7

Results of Second-Stage Regression Using Population Density to Quantify Agglomeration

Dependent: first-Stage Spatial Residual ($\gamma_{ind,r,t}$)	OLS	IV
Ln population density	0.11***	0.21***
	(12.14)	(6.58)
Specialization (industry share)	-0.03	0.03
	(-0.60)	(0.37)
Diversity	-0.03	-0.14**
	(-1.30)	(-3.56)
Competition (Herfindahl-Hirschman Index)	-0.05	-0.11**
	(-1.75)	(-3.16)

Table 2.7

Results of Second-Stage Regression Using Population Density to Quantify Agglomeration (Cont.)

Dependent: first-Stage Spatial Residual ($\gamma_{ind,r,t}$)	OLS	IV
Area	0.0000001	0.00001**
	(0.43)	(3.21)
Constant	-0.89***	-1.24***
	(-18.44)	(-10.78)
Year dummies	Included	Included
R squared	0.31	0.27
Number of observations	2,629	2,629

Note. t-statistics are reported in parentheses.

* p<0.05, ** p<0.01, *** p<0.001.

The impact of agglomeration in Thailand was found to have a similar magnitude to that observed in other developing nations like India, China, Colombia, and Brazil (Barufi et al., 2016; Chauvin, Glaeser, Ma, & Tobio, 2017; Duranton, 2016). Other studies indicated that the agglomeration effects in Thailand at the provincial level ranged from 16–37 percent (Limpanonda, 2012).

The impact of local density on wages in Thailand was larger than in developed countries like the Netherlands and France, where the influence of local density on wages was approximately 4–5 percent (Combes et al., 2008; Groot et al., 2014). However, various factors could influence the degree of agglomeration, including the spatial scale and model specifications (Combes and Gobillon, 2015; Grover, Lall, & Timmis, 2021). In the case of the Netherlands, the agglomeration effect doubled when the analysis shifted from municipalities to NUTS-3 regions (Groot et al., 2014). In a meta-analysis, Grover et al. (2021) demonstrated a higher overall agglomeration effect in developing countries. However, when accounting for the increasing congestion in developing countries, the size of the agglomeration effect significantly decreased to a level similar to that observed in developed economies.

Regarding the impact of localization economies, it was discovered that specialization, specifically the Marshall-Arrow-Romer spillover, did not significantly affect workers' wages. The lack of significance for specialization was also observed in Brazil's manufacturing sector (Barufi et al., 2016). Through spatial variance analysis, Combes et al. (2008) demonstrated that while overall employment density accounted for a considerable portion of regional productivity disparities, the explanatory power of specialization was relatively small. Furthermore, the results obtained through the IV estimation indicated a negative association between diversity with wage premiums, suggesting that cross-sectoral knowledge externality did not adhere to the theory of Jacobs spillover. These findings align with similar results in France, the Netherlands, Indonesia, and Brazil (Barufi et al., 2016; Combes et al., 2008; Groot et al., 2014; Ridhwan, 2021).

Concerning the final type of externality, Porter spillover, the coefficient of competition was negative and statistically significant, aligning with empirical findings from Brazil and the Netherlands (Barufi et al., 2016; Groot et al., 2014). However, the ambiguity of the localization effect was also reported by Combes and Gobillon (2015).

The outcome of the Durbin-Wu-Hausman test (shown in Table 2.8) revealed that population density is an endogenous variable. The substantial values of the Durbin-Wu-Hausman statistics reinforce our decision to employ two-stage least squares regression. To further verify the characteristics of the instrumental variable, I conducted the Stock and Yogo weak instrument test to assess the relevance of the IV, which is the percentage of clay content in the soil. As shown in Table 2.9, the instrument is deemed valid. The high F statistics indicate that the instrument effectively predicts the current population density.

Table 2.8

Durbin-Wu-Hausman Test of Second-Stage Regression Using Population Density to Quantify Agglomeration

Null hypothesis: Log population density is exogenous	**Test statistics**	**p-value**	**Verdict**
Durbin (score)	12.00	0.0005	Reject the null hypothesis
Wu-Hausman	12.02	0.0005	Reject the null hypothesis

Note. The table was created from author's calculations using Durbin-Wu-Hausman test.

Table 2.9

Stock and Yogo Weak Instrument Test of Second-Stage Regression Using Population Density to Quantify Agglomeration

Null hypothesis: Instrument is weak	**10%**	**15%**	**20%**	**Verdict**
Minimum eigenvalue statistic = 225.97				
2SLS Size of nominal 5% Wald test	16.38	8.96	6.66	Instrument is not weak
LIML Size of nominal 5% Wald test	16.38	8.96	6.66	Instrument is not weak

Note. The table was created from author's calculations using Stock and Yogo Weak Instrument test.

2.6.4 NTL Density Regression Results

The findings from the second-stage regression, presented in Table 2.10, utilized NTL density as a measure of local density. Table 2.10 represents a reevaluation of equation (2.37), where population density was replaced with NTL density. The results demonstrated that the outcomes remain consistent and robust even when substituting population density with NTL density.

Table 2.10

Results of Second-Stage Regression Using NTL Density to Quantify Agglomeration

	OLS	IV
Dependent: first-Stage Spatial Residual ($\gamma_{ind,r,t}$)		
Ln NTL density	0.25***	0.26***
	(16.27)	(6.89)
Specialization (industry share)	0.01	0.01
	(0.23)	(0.26)
Diversity	-0.07***	-0.08*
	(-3.75)	(-2.53)
Competition (Herfindahl-Hirschman Index)	-0.06*	-0.7*
	(-2.25)	(-2.17)
Area	0.000006***	0.00006*
	(3.29)	(2.07)
Constant	-0.84***	-0.84***
	(-20.07)	(-14.31)
Year dummies	Included	Included
R squared	0.34	0.34
Number of observations	2,629	2,629

Note. t-statistics are reported in parentheses.

* p<0.05, ** p<0.01, *** p<0.001.

The coefficient of NTL density in the IV model strongly resembles the estimate obtained through OLS, indicating that the OLS estimator is unbiased. Based on the IV estimation, the agglomeration effect's magnitude is larger than the previous model (0.26 versus 0.21). As NTL density captures the actual density in Thailand's urban areas, this outcome suggests that the agglomeration effect holds greater significance for informal workers, such as seasonal migrants and urban informal labor. This finding is not specific to Thailand, as similar results were observed in Colombia, where informal workers also received greater benefits from agglomeration economies (Duranton, 2016).

The impact of specialization was insignificant, indicating that it did not have a statistically significant effect. However, diversity was found to have a negative influence on spatial residuals, albeit with a smaller magnitude and significance level than the previous model. In contrast to the previous results, the effect of competition is no longer significant in this model. These regression results show the individual components contributing to agglomeration effects. Specifically, the lack of statistical significance for specialization and competition suggests that the concentration of industries and local competition in Thailand needed to be increased to generate agglomeration externalities. These findings align with comparable results observed in some developing countries, where the concentration of low-technology production and imperfect market mechanisms in certain areas were commonly observed. Therefore, it is important to consider technological development and other fundamental factors, such as enforcing competition law, as significant underlying factors driving agglomeration effects.

Table 2.11 presents the coefficient value of NTL density in the IV estimate, similar to the estimate obtained through OLS. This suggests that the issue of endogeneity bias was effectively addressed by using NTL density as a substitute for local density. The results of the Stock and Yogo weak instrument test, shown in Table 2.12, indicate that the soil's clay content percentage is not a weak instrument in the model incorporating NTL density.

Table 2.11

Durbin-Wu-Hausman Test Using NTL Density to Quantify Agglomeration

Null hypothesis: **Log population density is exogenous**	**Test statistics**	**p-value**	**Verdict**
Durbin (score)	0.02	0.89	Fail to reject the null hypothesis
Wu-Hausman	0.02	0.89	Fail to reject the null hypothesis

Note. The table was created from author's calculations using Durbin-Wu-Hausman test.

Table 2.12

Stock and Yogo Weak Instrument Test Using NTL Density to Quantify Agglomeration

Null hypothesis: **Instrument is weak**	**10%**	**15%**	**20%**	**Verdict**
Minimum eigenvalue statistic = 546.63				
2SLS Size of nominal 5% Wald test	16.38	8.96	6.66	Instrument is not weak
LIML Size of nominal 5% Wald test	16.38	8.96	6.66	Instrument is not weak

Note. The table was created from author's calculations using Stock and Yogo Weak Instrument test.

All the results indicated that employing the two-stage regression approach to analyze the dataset, which combined official surveys and geospatial indicators, could effectively quantify the factors contributing to regional wage disparities in Thailand. Despite using various proxies for population density and employing different regression techniques, the main finding remains consistent, providing robust evidence of the impact of agglomeration economies on labor productivity. Therefore, in the context of Thailand, the distinct pattern of wage concentration in Bangkok and its surrounding areas could be attributed to the powerful agglomeration effect generated by the high population density in these regions.

2.6.5 Policy Recommendations

As mentioned earlier, Thailand faces declining GDP growth and persistent spatial inequality. This study's key findings provide policy recommendations to address both issues. Specifically, the estimated coefficients of agglomeration elasticity indicate the productivity benefits derived from density-driven externalities.

Given the historical focus on growth centered around Bangkok, promoting a more balanced and multipolar growth across all regions of the country is advisable. It is highly recommended to implement a new development plan that encourages decentralized growth by establishing new cities that generate significant agglomeration externalities. The level of agglomeration elasticity identified in this study also indicates the potential for regional growth resulting from increased productivity and wages. Expanding the presence of high-agglomeration cities in multiple regions will reduce the urban dominance of Bangkok and ultimately uplift the average nationwide income.

2.7 Conclusion

This research investigates the factors contributing to the ongoing spatial inequality in Thailand. The study employed a two-stage estimation approach to analyze the impact of agglomeration externalities on regional productivity disparities at the provincial level. A novel aspect of this research is using open geospatial data and satellite-based indicators. Specifically, NTL density was employed as a substitute for local density, while soil clay content percentage was used as an instrumental variable.

The findings from the estimation indicate that differences in worker skills and the presence of agglomeration externalities contributed to regional wage variations. In addition to individual capabilities, workers in densely urbanized provinces also benefited from higher wage premiums resulting from agglomeration effects. Consequently, the concentration of urbanization increased spatial inequality nationally, as it disproportionately favored workers in highly urban areas. These results align with previous research in developing and developed economies, highlighting the need for a new policy approach that promotes more balanced economic growth by establishing regional cities that generate density-driven spillover effects.

This study has several limitations that should be acknowledged. Firstly, the industrial classification used in the analysis is not up-to-date, which may limit the accuracy and relevance of the findings. Secondly, the spatiotemporal resolutions of the data used in this study may not capture the full complexity of the phenomenon under investigation. Future research should focus on improving the spatial and temporal details of the data and incorporating new industries that may play a significant role in shaping regional dynamics. Additionally, a dynamic analytical approach considering the multi-period effects of agglomeration externalities would provide a more comprehensive understanding of the phenomenon.

CHAPTER 3
URBAN LAND EXPANSION AND ECONOMIC DEVELOPMENT IN THAILAND

3.1 Introduction

Urbanization is viewed by scholars as one of the most eminent legacies of humanity (Elmqvist et al., 2021). Cities all over the world are not only getting denser at the core but also expanding outward. Globally, the natural cities—the urbanized areas observed by satellite at nighttime—increased from 170,000 square kilometers in 1992 to 560,000 square kilometers in 2016. The expansion of natural cities mostly takes place in many developing countries in Asia. Specifically, 1,459 cities in 42 special economic zones are expanding in size. Almost half are in the People's Republic of China (Asian Development Bank, 2019). Given the current population growth and other socioeconomic conditions, the global urban land cover will expand rapidly until the 2040s (Chen et al., 2020).

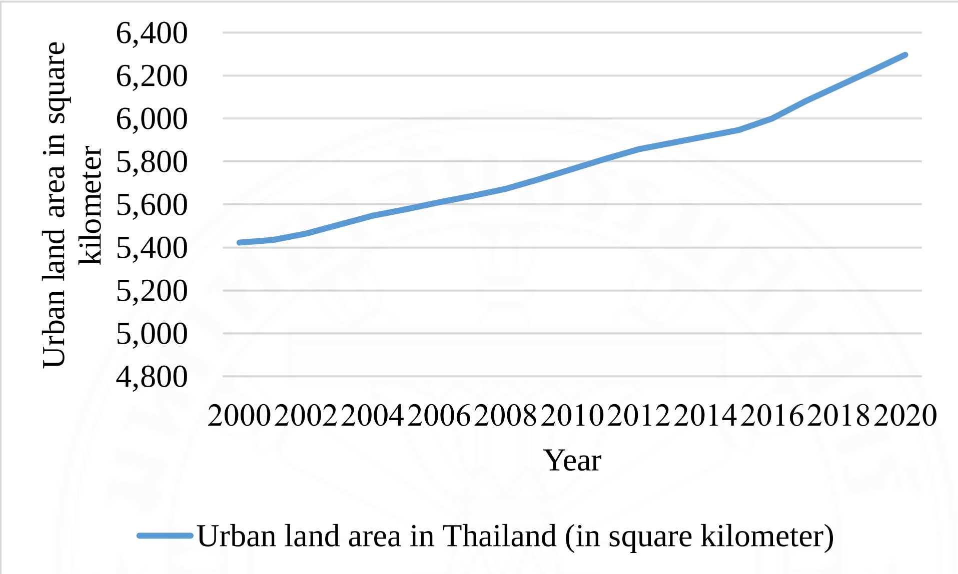
Thailand, one of Asia’s developing countries, has been experiencing rapid urbanization. Cities in Thailand are becoming denser and bigger. Implementing the first national economic and social development plan in 1960 transformed Thai society from an agriculturally dominant economy into an industrial and service-based economy. Investments in infrastructure over a broad range of the nation’s hinterland cause unprecedented rural-urban migration. Moreover, recent urban development policies such as the implementation of Smart City, Tourism City, and Special Economic Zones accelerate the expansion of the urban core into former agricultural areas (Tonguthai, 2019).

Thanks to the recent advancement of remote-sensing and cloud-computing techniques, the rapid expansion of cities in Thailand could be observed from space. Specifically, Thailand's observed urban area expansion is extracted from Terra and Aqua Terra Moderate-Resolution Imaging Spectroradiometer (MODIS) reflectance data provided by the National Aeronautics and Space Administration (NASA). Figure 3.1 shows the dynamic of urban land expansion in Thailand at the national level from

2000–2020. The data from MODIS suggests that Thailand's urban areas increased from 5,422 sq.km. in 2000 to 6,296 sq.km. in 2020.

Figure 3.1

Urban Land Expansion in Thailand Between 2000–2020 in Square Kilometer



Note. The figure was created from MODIS data.

However, despite having achieved constant urban land expansion over the last two decades, such progress is still taking place in a few provinces, reflecting the monocentric growth pattern of the urban system in Thailand. Research suggests that Thailand is one of the countries with the highest level of regional inequality globally (Rodríguez-Pose, 2018). The urban size of the biggest city in Thailand—Bangkok Metropolis—is about ten times larger than the second and third largest cities, making Thailand the country with the highest Urban Primacy Index in the world (Short & Pinet-Peralta, 2009). One research also suggests that during 1992–2019, urbanization in Thailand, measured by NTL density, was primarily concentrated in the BMR and a few provinces (Sangkasem & Puttanapong, 2022).

Table 3.1

Top 10 Provinces With Highest Increase in Urban Area Between 2000 and 2020

Rank	Province	Increase in urban area between 2000 and 2020 (in square kilometer)	Increase in urban area between 2000 and 2020 (in percent)
1	Bangkok Metropolis	144.94	16.52
2	Chon Buri	114.14	52.80
3	Samut Prakan	94.59	49.25
4	Nonthaburi	86.64	48.33
5	Pathum Thani	81.81	36.53
6	Samut Sakhon	65.79	49.25
7	Phra Nakhon Si Ayutthaya	35.67	33.84
8	Nakhon Pathom	28.03	16.29
9	Rayong	26.92	31.66
10	Chachoengsao	22.43	44.18

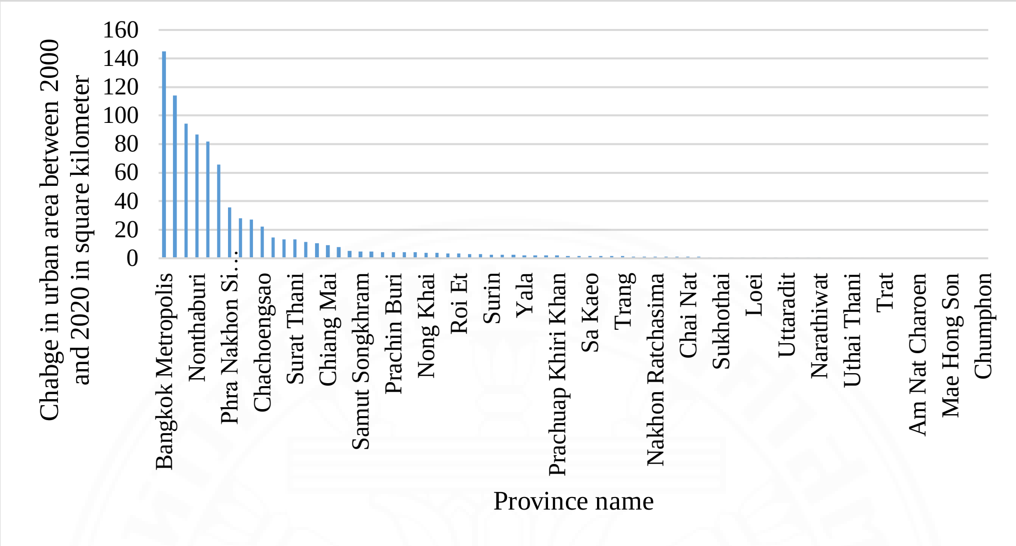
Note. The figure was created from MODIS data.

Using data obtained from MODIS, Table 3.1 shows the urban land expansion dynamic of Thailand's top primate cities. Bangkok Metropolis has experienced the highest total urban land area increase over the last two decades. The vicinity of Bangkok, namely Pathum Thani, Samut Prakan, Nonthaburi, Samut Sakhon, and Nakhon Pathom remain in the top ten provinces with the highest increase in total urban land area. The urban land areas of these six provinces increased by 615.94 sq.km. within 20 years.

Between 2000 and 2020, the provinces situated in the Eastern Economic Corridor, namely Chon Buri, Rayong, Chachoengsao, and Phra Nakhon Si Ayutthaya, which also house major industrial parks, ranked among the top ten provinces with the most tremendous growth in urban area, excluding Bangkok and its surrounding regions. The urban land areas of these six provinces increased by 199.16 sq.km. within 20 years. In other words, urban land expansion of the top ten provinces accounts for about 93 percent of total urban land expansion in Thailand.

Figure 3.2

Distribution of an Increase in Urban Area Between 2000 and 2020 in Square Kilometer



Note. The figure was created from MODIS data.

Figure 3.2 shows the distribution of urban land expansion at the provincial level. The distribution follows the long-tail distribution, which is analogous to the distribution of population density observed in Chapter 2. Few primate cities exhibit a notable change in urban areas, while the rest show no change or an insignificant increase in urban areas over the last 20 years.

The disproportionate growth of the urban land area, as shown in Table 3.1 and Figure 3.2, characterizes extreme regional inequality in Thailand, and what is worse still, wealth and job creation concentrate in the same primate cities, namely the BMR and the Eastern Economic Corridor.

Table 3.2

Top 10 Provinces With Highest Increase in GPP (CVM) Between 2000 and 2020

Rank	Province	Increase in GPP (CVM) between 2000 and 2020 (in million baht)	Increase in GPP (CVM) between 2000 and 2020 (in percent)
1	Bangkok Metropolis	1,889,235	103.10

Table 3.2

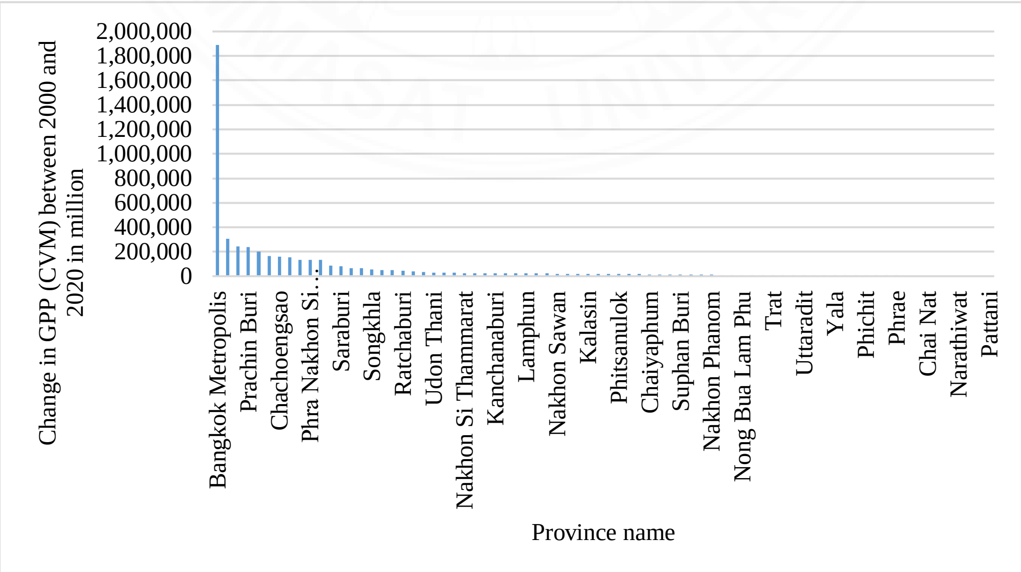
Top 10 Provinces With Highest Increase in GPP (CVM) Between 2000 and 2020 (Cont.)

Rank	Province	Increase in GPP (CVM) between 2000 and 2020 (in million baht)	Increase in GPP (CVM) between 2000 and 2020 (in percent)
2	Chon Buri	305,952	119.83
3	Rayong	242,511	97.72
4	Prachin Buri	236,163	521.41
5	Samut Prakan	199,615	64.83
6	Nonthaburi	167,009	220.93
7	Chachoengsao	160,070	182.04
8	Samut Sakhon	156,799	111.30
9	Nakhon Pathom	134,635	138.98
10	Phra Nakhon Si Ayutthaya	133,404	79.55

Note: The table was created from data provided by NESDC.

Figure 3.3

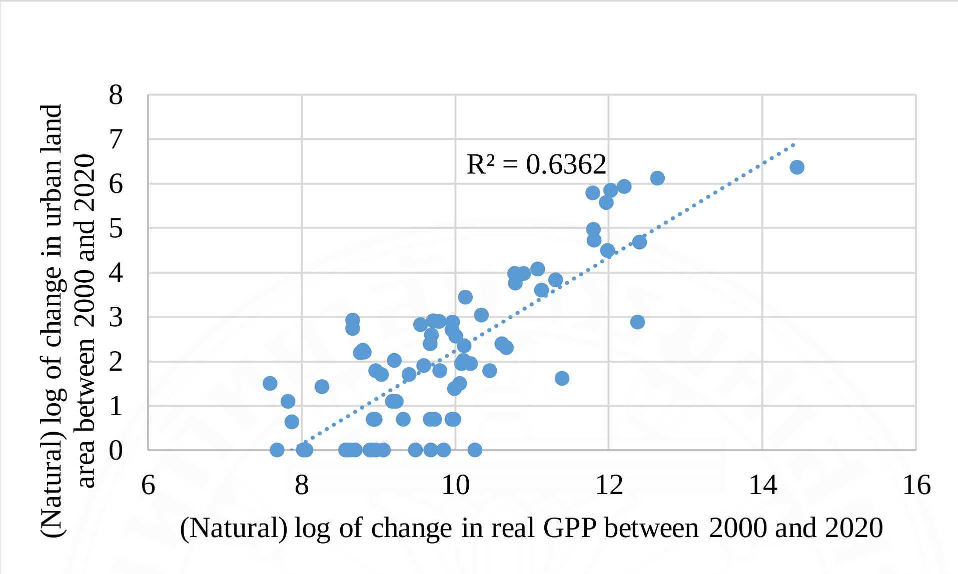
Distribution of an Increase in GPP (CVM) Between 2000 and 2020



Note: The table was created from data provided by NESDC.

Figure 3.4

Relationship Between Change in Urban Land Area and Change in GPP (CVM) At Provincial Level (In Natural Log Term)



Note: The table was created from data provided by NESDC.

Table 3.2 shows that the growth of real GPP is as disproportionate as that of the urban land area. Notably, nine out of ten provinces in Table 3.1 also appear in Table 3.2. Specifically, Bangkok Metropolis and some of its surrounding provinces (Samut Prakan, Nonthaburi, Samut Sakhon, Nakhon Pathom) received the highest benefit of real economic progress over the last two decades, which has a total value of 2.5 million baht. Like Table 3.1, most economic development has also occurred in the familiar provinces in the eastern and central regions, namely Chon Buri, Rayong, Prachin Buri, Chachoengsao, and Phra Nakhon Si Ayutthaya.

Table 3.3 shows the distribution of an increase in real GPP between 2000 and 2020. It clearly illustrates the extreme concentration of economic activity in Thailand. During 2000 and 2020, the real GPP of Thailand increased by five trillion in total. Seventy-two percent of this increase took place in BMR and Eastern Economic Corridor. Bangkok Metropolis alone accounts for 37 percent of the change.

A relationship between urban land and GPP growth is plotted in Figure 3.4. It shows that A higher rate of GPP growth correlates with a higher rate of urban land expansion. The variation in GPP growth alone could explain 63 percent of the variation in urban land growth between 2000 and 2020. This fact leaves no doubt that unequal economic growth is associated with the monocentric growth pattern of the urban system in Thailand.

As territorial inequality led to serious social conflict, retard economic growth, and political turmoil in Thailand (Hewison, 2014; Rodríguez-Pose, 2018), a new redistributive policy that promotes the growth of regional cities could defuse the tension between the few prosperous provinces and left behind regions. A better understanding of what drives urban growth is crucial to form such a policy.

Analysis from Chapter 2 suggests that workers benefited from the agglomeration externalities of big cities. Thus, identifying the driver of urbanization is essential for policymakers, particularly in the case of Thailand, where a polycentric growth pattern of the urban system is needed. Although urban economists made constant progress in reliably quantifying agglomeration externalities of megacities (Combes & Gobillon, 2015), research that investigates how economic activities and other factors drive urbanization or urban land expansion is relatively limited. Therefore, this chapter examines the role of economic growth and other possibly relevant factors, such as human capital and natural factors, on urban land expansion in Thailand.

The content in this chapter is divided into seven sections. The following section discusses the related theoretical and empirical literature. The theoretical framework of this paper is provided in the third section. The fourth section overviews the estimation method of this paper. The fifth section gives detail of the data used in the analysis. Estimation results are discussed in the sixth section. Finally, the seventh section concludes this paper.

3.2 Literature Review

3.2.1 Theoretical Literature

The extensive body of literature that analyzed the impact of urbanization on productivity developed its theoretical models based on microeconomics theory. In contrast, the literature that investigated the relationship between economic development and urbanization focused on examining the influence of economic development on urbanization at a broader scale, specifically at the national level.

Traditionally, urban economic research defines urbanization as the growth of urban areas, including expanding urban land or establishing new cities. In this context, urban economic theorists often use an increasing urban population and rural-urban migration as urbanization indicators in their models. Consequently, to maintain clarity and align with the research objective, this literature review concentrates on theoretical studies examining economic development's impact on urbanization, precisely measured as an expansion in urban size.

The urban economic study that developed a theoretical framework to explain the relationship between an increase in urban size and economic growth was envisaged by Wheaton (1974). Relying on monocentric growth theory, Wheaton (1974) derived a comparative static that explicitly showed a positive relationship between rising household income, urban expansion, and urban land demand.

After Wheaton (1974), Bertinelli and Black (2004) developed a simple model incorporating endogenous human capital where workers can choose any location and their level of human capital to maximize their welfare. Bertinelli and Black's (2004) model is based on the neo-classical endogenous growth theory, where human capital formation is the main factor of creative disruption that produces sustainable, long-term economic growth.

Bertinelli and Black (2004) also introduced an important departure from the standard economic growth theory by incorporating the concept of urbanization within their model. Unlike the conventional economic growth theory, which tends to overlook the spatial dimension, Bertinelli and Black (2004) recognized the significance of urbanization and made it an endogenous component of their model. By assuming that innovative ideas disseminate in the city, the formation of human capital leads to technological improvement, fostering long-term economic growth. According to their prediction, under certain conditions where the initial level of technology is high enough, the economy could experience unlimited growth, and the urbanization rate would converge to its maximum point in the long run. In summary, Bertinelli and Black's (2004) model showed that, through human capital formation, there is a close relationship between economic growth and urbanization.

Following the publication of Bertinelli and Black's study in 2004, Henderson (2005) provided a comprehensive overview of the theory of city formation. Henderson's work focused on examining the factors that determine the size of cities and analyzing the relationship between urbanization and economic development. In doing so, Henderson's research offered a solid understanding of how cities form and their significance in the context of economic progress.

Henderson (2005) showed that improvement in technology leads to urbanization, which is measured as an increase in city size. The rationale behind the positive effect of technological improvement lies in the hypothesized inverted U-shaped relationship between real income per worker and city size. The inverted U-shaped curve represents the trade-off between agglomeration externalities and congestion costs. According to the model, instead of offering a positive externality, the oversized city leads to rising commuting costs and diseconomies of scale. Technological change and economic growth shift the inverted-U curve upward and outward, increasing the optimal size of the city. In addition to technological change, the role of institutions also explains the pattern of the formation of the city. Democratic institutions are more likely to promote polycentric growth. At the same time, authoritarian regimes tend to establish monocentric growth poles that concentrate political and economic power in a single (and often) capital city.

Although the literature in the early 2000s laid the foundation of a causal relationship between urban expansion and rising income, most of the models are purely theoretical and lack rigid empirical testing. However, a recent study by Gollin, Jedwab, and Vollrath (2016) described the underlying mechanism that governs the causal relationship between urban expansion and GDP growth and connected theory to empirical work.

Based on the static equilibrium model, the model developed by Gollin et al. (2016) is clear and straightforward. It shows that rural-urban migration is the basis of urbanization. An increase in surplus income—an income after having purchased basic agricultural products—either from the discovery of precious natural resources or rising industrial productivity moves workers from rural areas into urban areas where they would work in industrial and service sectors of urban areas. This rural–urban migration process induced by a rising surplus income of households commences urbanization in many countries.

3.2.2 Empirical Literature

The early empirical literature that examined the relationship between the urbanization process and economic development not only studied how economic factors affect city growth or the formation of city systems (size distribution of cities) but also sought to determine the optimal size of the city. Technology change increases the optimal size of an individual city because knowledge accumulation and positive externalities of innovation help mitigate rising congestion costs associated with increasing city size (Henderson & Wang, 2007).

The type of institutions affects the size and distribution of cities. Inclusive institution leads to polycentric city size distribution, while extractive institution favors the monocentric growth of cities (Ades & Glaeser, 1995). In addition to technological progress and institutional factors, international trade and geography factors also affect city size and the distribution of city size. The New Economic Geography literature considers trade and commuting costs to be crucial in determining city size growth (Krugman & Elizondo, 1996). The optimal size of a large metropolitan area is the primary interest of early research concerning the relationship between urbanization and growth. This literature is based on the hypothesis of an inverted U-shaped curve. The oversized metropolitan region stifles economic growth by concentrating resources in a single city, resulting in diseconomies of scale (Henderson, 2003).

The recent empirical literature that studied economic development and urbanization focused on the experience of China's economic boom and rapid urban growth. According to this literature, rapid urbanization in China was characterized by urban land expansion, converting agricultural areas into built-up areas. Satellite data were often used to observe the urbanization process in China.

Outside China, Duranton (2016) showed that wage growth positively affected city growth in Columbia. Gollin et al. (2016) also empirically showed that a rising surplus income from resource export started urbanization in many developing countries in the Middle East and Africa.

However, as suggested by the theoretical literature on economic growth and urbanization, there is a self-reinforcing mechanism between urbanization and economic growth. Specifically, economic growth leads to higher income, and technological advancements, technological change, and rising wages promote urban expansion. Thus, to investigate causality from economic development to urbanization, previous empirical studies applied various techniques to deal with the endogeneity issue. Most studies found a positive effect of economic growth on urbanization.

Table 3.3 lists the empirical literature investigating the relationship between urbanization and economic development and provides detail of the data type used, year of study, methodology, and dependent and key explanatory variables in chronological order. Some studies directly investigated the relationship between urban land expansion and GDP (Bai, Chen, & Shi, 2012; Deng, Huang, Rozelle, & Uchida, 2008; Gollin et al., 2016; G. Li, Sun, & Fang, 2018; Liu, Yan, & Zhou, 2016).

Another key socio-economic variables (such as population, wage, foreign direct investment, institution, technology) were also used by researchers to explore how these factors affected urbanization process (Ades & Glaeser, 1995; Duranton, 2016; Gao, Wei, Chen, & Chen, 2014; Gao, Wei, Chen, & Yenneti, 2015; J. V. Henderson & Wang, 2007; H. Li, Wei, Liao, & Huang, 2014; Ustaoglu & Williams, 2017; Xu, Wang, Chi, & Zhang, 2020).

Table 3.3

List of Empirical Literature

Author(s)	Data	Period	Methodology	Dependent variable	Key explanatory variable
Deng et al. (2008)	PD	1987–2000	Panel regression analysis	Urban core area	GDP
Bai et al. (2012)	PD	1990–2006	Granger causality	Built-up area	GDP

Table 3.3

List of Empirical Literature (Cont.)

Author(s)	Data	Period	Methodology	Dependent variable	Key explanatory variable
Gao et al. (2014)	PD	2002–2008	Panel regression analysis	Urban land-use change	Level of technology
Li et al. (2015)	PD	1998–2008	1. Panel regression analysis	Urban land area	Fixed asset investment
Gao et al. (2015)	PD	2000–2010	1. GTWR 2. SRM	Built-up area	FDI
Liu et al. (2016)	PD	1997–2010	Granger causality	Built-up area	GDP
Duranton (2016)	PD	1993–2010	Panel regression analysis	Urban population	Wage
Gollin et al. (2016)	PD	1960–2010	Spatial panel regression	Urban population	Natural resource exports
Ustaoglu and Williams (2017)	PD	2000–2006	Seemingly unrelated regressions	Industrial built-up area	Population
G. Li et al. (2018)	PD	2005–2010	1. SAR Probit	Built-up area	GDP
Xu et al. (2020)	CS	2015	1. Quantile regression	Land-use intensity index	Population

Note. PD and CS denotes panel and cross-section data, respectively.

Most empirical literature strongly supported the causality of economic development to urban expansion. However, Bloom, Canning, and Fink (2008) did not confirm a positive relationship between economic growth and the level of urbanization at a global level between 1970 and 2000. Their regression analysis yielded an insignificant result. The insignificant relationship between urbanization and economic growth rate might stem from the fact that large parts of African countries experienced high urbanization rates with an extremely low growth rate of GDP. Bloom et al. (2008) finding is consistent with Henderson (2003)'s finding. Henderson argued that one should not regard urbanization per se as a panacea for productivity improvement since empirical evidence showed that rapid urbanization occurred during economic stagnation.

However, Bai et al. (2012) argued that the insignificant effect of urbanization and economic development reported by Bloom et al. (2008) might be due to the inappropriateness of using demographic data to proxy urbanization rate since the definition of urban population varies across countries. Thus, instead of using population to represent urbanization, Bai et al. resorted to a more apparent data choice to proxy the urbanization process—built-up area extracted from satellite data. After Bai et al. (2012), researchers utilized satellite data to examine the effect of economic development on urban land expansion in China over the past decade.

From a spatial perspective, urbanization and economic development in Thailand were found to follow the monocentric growth theory of urbanization. Although economic development in Thailand benefited rural areas, as Tonguthai (2019) argued, most of the wealth generated by economic growth was concentrated in BMR. Government policies also affected the urbanization rate in Thailand. For example, the policies on agricultural products, the concentration of investment in infrastructure, and the promotion of industrialization in Bangkok resulted in large-scale rural-urban migration. They led to the rapid growth of Thailand's capital city.

Since a quantitative examination of socio-economic factors that might affect the urban land expansion in Thailand is not adequately supplied, this paper, therefore, utilizes innovative data to study the relationship between economic development and urbanization processes in Thailand over a relatively long period of time and at higher frequency.

3.3 Theoretical Framework

The basic theoretical foundation of this paper is based on theoretical work of Gollin et al. (2016). The model was specifically developed to deal with urbanization and increasing levels of income from the discovery of natural resources and rising industrial productivity. Most importantly, the model could be assessed empirically in the context of data availability in a developing country.

3.3.1 Model of Urbanization and Structural Change

The foundation of the model stated that urbanization is driven by income shock either from natural resource export or increasing level of industrial productivity. The income shock cause workers to structurally shift away from agriculturally based economy in rural areas to industrial and service based in urban areas. The economy consists of tradable sectors and non-tradable sectors. The tradable sectors produce manufactured goods and services that can be traded internationally such as finance, insurance, and business services. Some tradable goods such as agricultural products are produced in rural area. The non-tradable sectors are composed of government and personal services as well as local retail, transportation, construction, education, and healthcare.

3.3.1.1 Household Utility Function

A household sector is assumed to have the following utility function.

$$U = \beta_f \ln\left(c_f - \overline{c_f}\right) + \beta_d \ln c_d + \beta_n \ln c_n \tag{3.1}$$

where $\overline{c_f}$ is subsistence requirement for food, c_f is food consumption (rural tradable goods), c_d is urban tradable goods (manufacturing as well as tradable services such as finance, insurance, and business services), c_n is urban non-tradable goods (government and personal services as well as local retail, transportation, construction, education, and health). The values of $\beta_f + \beta_d + \beta_n$ are between zero and one and $\beta_f + \beta_d + \beta_n = 1$.

3.3.1.2 Production Function

The production function of each sector takes the following form.

$$Y_j = A_j L_j^{1-\alpha} \tag{3.2}$$

where $j\epsilon(f, d, n)$ and the sum of number of workers in each sector equal to total number of workers in the economy. Mathematically, $L_f + L_d + L_n = 1$ and L_f, L_d, and L_n represent share of workers in each sector.

3.3.1.3 Budget Constraint

In addition to utility and production function, the household sector has the following budget constraints.

$$p_f^* c_f + p_d^* c_d + p_n c_n = m \tag{3.3}$$

$$p_f^*(c_f - \overline{c_f}) + p_d^* c_d + p_n c_n = m - p_f^* \overline{c_f} \tag{3.4}$$

where p_f^* and p_d^* denote international price of tradable goods and p_n represents domestic price of urban non-tradable goods. The LHS of equation (3.4), $m - p_f^* \overline{c_f}$, is surplus income after purchasing subsistence requirement for food.

Since urban non-tradable goods is assumed to be produced domestically, total consumption on urban non-tradable goods is equal to the total value of production which could be written in mathematic form as

$$\beta_n(m - p_f^* c_f) = p_n Y_n \tag{3.5}$$

where the LHS of equation (3.4), $\beta_n(m - p_f^* c_f)$, is total consumption on urban non-tradable goods and RHS of equation (3.5), $p_n Y_n$, is total value of production of urban non-tradable goods.

Like urban non-tradable goods, total consumption of rural and urban tradable goods must equal the total value of production as described in equation (3.6).

$$(\beta_f + \beta_d)(m - p_f^* \overline{c_f}) + p_f^* \overline{c_f} = R + p_f^* Y_f + p_d^* Y_d \tag{3.6}$$

where R is a revenue from the natural resource that can be used to pay for imports. The LHS of equation (3.6), $(\beta_f + \beta_d)(m - p_f^* \overline{c_f}) + p_f^* \overline{c_f}$, is total consumption on rural and urban tradable goods. The RHS of equation (3.6), $R + p_f^* Y_f + p_d^* Y_d$, is their total value of production plus revenue from the natural resource.

3.3.1.4 Wage Equalization Between Sectors

Assuming free labor mobility between the sectors of the economy, wage between sectors is equalized for any sector j and k.

$$(1-\alpha)p_j A_j L_j^{-\alpha} = (1-\alpha)p_k^* A_k L_k^{-\alpha} \quad (3.7)$$

From the budget constraints and production function, equation (3.5), (3.6), (3.7), (3.2) and $L_f + L_d + L_n = 1$, the number of labors in urban non-tradable sector is given by equation (3.8).

$$L_n = \beta_n \left(1 + \frac{(1-L_n)^\alpha}{\bar{A}}\left(R - p_f^* \bar{c_f}\right)\right) \quad (3.8)$$

where $\bar{A} = \left[(p_d^* A_d)^{1/\alpha} + \left(p_f^* A_f\right)^{1/\alpha}\right]^\alpha$

Equation (3.8) implies that the number of urban non-tradable sector depends on income level from either increase in resource export or productivity level. The rising income level allows workers to meet their basic needs more easily, enabling urban migration to take place.

From equation (3.8), the allocation of labor to the other sector is given by equation (3.9) and (3.10).

$$L_d = (1-L_n)\left(\frac{p_d^* A_d}{\bar{A}}\right)^{1/\alpha} \quad (3.9)$$

$$L_f = (1-L_n)\left(\frac{p_f^* A_f}{\bar{A}}\right)^{1/\alpha} \quad (3.10)$$

Like equation (3.8), equation (3.9) and (3.10) suggests that number of workers in tradable sectors depends on relative productivity level of that sector.

3.3.1.5 Urbanization and Comparative Statics

Labor allocation between sectors implies that urbanization equals the number of workers in urban tradable and non-tradable sector.

$$U = L_n + L_d \tag{3.11}$$

6.1 Proposition 1 (resource revenue and urbanization)

From equation (3.8), (3.9) and (3.10), it can be shown that.

$$(A)\ \frac{\partial U}{\partial R} > 0$$

$$(B)\ \frac{\partial L_n}{\partial R} > 0$$

$$(C)\ \frac{\partial L_d}{\partial R} < 0$$

Proposition 1 implies that increased resource revenue leads to higher demand for workers in urban non-tradable sectors, but it reduces demand for worker in rural and urban tradable sectors because the fall in tradables production can be compensated by imported good paid by the higher resource revenue. But the increase in workers in urban non-tradable sectors outweighs the fall in workers in tradable sectors, therefore the increased resource revenue leads to higher urbanization rate.

6.2 Proposition 2 (industrialization and urbanization)

When $R < p_{rd}^{*}\bar{c}_{rd}$, then given (3.8) and (3.9) it can be shown that.

$$(A)\ \frac{\partial U}{\partial p_d^{*} A_d} > 0$$

$$(B)\ \frac{\partial L_n}{\partial p_d^{*} A_d} > 0$$

$$(C)\ \frac{\partial L_d}{\partial p_d^{*} A_d} > 0$$

Proposition 2 applies to the case when resource revenue is less than income from tradables production. In this case, rising tradable sectors productivity not only shifts away workers from food production sector to tradable sectors but also drives a shift towards urban non-tradable sectors. The effect of increased tradable sectors productivity on urbanization rate in this case is straightforward; the increase in tradable sectors productivity leads to higher urbanization rate.

3.4 Methodology

In this paper, regression analyses were applied to examine whether an increase in income (either from resource revenue or productivity) significantly impacts urban land expansion at the provincial level in Thailand, as established in the prior section. However, as Thailand is an export-led growth country where income from tradable production is always higher than income from the export of natural resources, the prediction of proposition 2 is more appropriate to use as a basic theoretical foundation when constructing the regression models. However, it is still interesting to see whether the prediction of Proposition 1 is correct.

According to Proposition 2, increasing urban tradable sectors' productivity, and holding constant resource revenue, should positively and significantly affect urban growth. Since the observed urban land expansion takes time, selecting the lagged value of income, environmental factors, and human capital as explanatory variables in the regression models reveals the effects of income growth and short-term effects of variation in the last two factors on urban land expansion. Following Gollin et al. (2016), the ten years lag variables were used as explanatory variables in the regressions.

The first empirical approach is the fixed-effects panel regression model that estimated the effect of tradeable sectors' productivity on urban land expansion, holding constant resource revenue and control variables. The fixed-effects model could also control any unobserved provincial characteristics. The second empirical approach is the spatial autoregressive model with spatial fixed-effects, the basic specification of spatial panel-data models. The third empirical approach is the spatial error model with spatial fixed-effects, an extended version of the spatial autoregressive model. Both spatial autoregressive and spatial panel-data models were applied to control the spatial correlation between geographic units, which is likely to happen when sample data were collected from geographically close entities (Belotti, Hughes, & Mortari, 2017). Lastly, the dynamic panel-data model was applied to control for the dynamic expectation of workers. Mathematical representations of each model are presented below.

3.4.1 Fixed-Effects Model

$$U_{i,t} = \alpha + \beta_1 tradable_{i,t-10} + \beta_2 resource_{i,t-10} + \beta_3 climate_{i,t-10} + \beta_4 education_{i,t-10} + u_i + \varepsilon_{i,t} \quad (3.12)$$

Equation (3.12) shows the standard mathematical expression of (non-spatial) fixed-effects model. The dependent variable, $U_{i,t}$, represents urban land area of province i at time t. The urban land data was obtained from the Terra and Aqua combined MODIS. With its foundation on satellite image, the urban land cover types were measured by supervision from various sources, including the International Geosphere-Biosphere Programme, University of Maryland, Leaf Area Index, BIOME-Biogeochemical Cycles, and Plant Functional Types classification schemes. The primary classified land-use map was corrected with auxiliary information before final release.

The key explanatory variable is the term $tradable_{i,t-10}$ which is a real tradable GPP per capita of province i at time $t-10$. The real tradable GPP per capita is meant to proxy productivity of tradable sectors. The key control variable is $resource_{i,t-10}$ which is a real resource GPP per capita of province i at time $t-10$. The inclusion resource revenue as the main control variable was based on theoretical modelling discussed in the previous section.

The other control variable is $climate_{i,t-10}$ which includes a set of environmental factors such as yearly average value of NDDI, NDVI, NDWI, temperature, and precipitation of province i at time $t-10$. To avoid potential multicollinearity, each environmental factor was estimated separately. The last control variable is $education_{i,t-10}$ which is the share of worker with higher education of province i at time $t-10$. u_i represents unobserved characteristic of province i. $\varepsilon_{i,t}$ is pure residual. The detail of each variable will be discussed in the next section.

3.4.2 Spatial Lag Model with Spatial Fixed Effects

$$U_{i,t} = \rho W U_{i,t} + \beta_1 tradable_{i,t-10} + \beta_2 resource_{i,t-10} + \beta_3 climate_{i,t-10} + \beta_4 education_{i,t-10} + u_i + \varepsilon_{i,t} \quad (3.13)$$

Equation (3.13) elaborates specification of spatial autoregressive model with spatial fixed-effects. Equation (3.13) includes same set of control variables as did in equation (3.12), but the term $\rho W U_{i,t}$ was added. According to Belotti et al. (2017), W is the $n \times n$ matrix that describe the spatial arrangement of the n units. Every geographic unit i is associated with the spatial matrix W. In other words, W could be thought of as a defined neighborhood and the term ρ would capture any spatial spillovers effect.

3.4.3 Spatial Error Model with Spatial Fixed Effects

$$U_{i,t} = \beta_1 tradable_{i,t-10} + \beta_2 resource_{i,t-10} + \beta_3 climate_{i,t-10} + \beta_4 education_{i,t-10} + u_i + v_{i,t} \quad (3.14)$$

where $v_{i,t} = \lambda M v_{i,t} + \varepsilon_{i,t}$

The spatial error model is similar spatial autoregressive model, but it focuses on spatial spillovers effect that work through the error term as described in $v_{i,t} = \lambda M v_{i,t} + \varepsilon_{i,t}$. The matrix M is a spatial weight matrix that may or may not be equal to W.

The differences in specification of each model could show whether increased tradable sectors productivity has significant effect on provincial urban land expansion. They also highlight the channel through which spatial spillovers operate, namely whether through a change in dependent variable or through a change in error term. Most importantly, this practice could deliver more meaningful result as well as lively policy discussions.

3.4.5 Dynamic Panel-Data Model

$$U_{i,t} = \theta + \gamma U_{i,t-1} + \beta_1 tradable_{i,t-10} + \beta_2 resource_{i,t-10} + \beta_3 climate_{i,t-10} + \beta_4 education_{i,t-10} + u_i + \varepsilon_{i,t} \quad (3.15)$$

As migration take time to occur and cities take time to build, urban land expansion of province i at time $t-1$ could affect urban land expansion of province i at time t. Workers had observed that urban land expansion was taking place at province i at time $t-1$ and decided to move to that province in the next period, generating an inertia effect of urban land expansion. Moreover, urban land expansion is an accumulation process where the current value of urban land is accumulation of urban land in the previous period. Failing to control the dynamic mechanisms could cause biased coefficient estimators. To econometrically deal with this problem, this study also applied the dynamic panel-data model. The lagged dependent variable, $U_{i,t-1}$, was included in the model to control for a partial adjustment mechanism. The model could potentially reveal effect of urban land expansion at time $t-1$ on urban land expansion at time t. The equation (3.15) was estimated via Arellano–Bond estimation's technique.

3.5 Data

The data for this paper consists of conventional economic indicators and alternative satellite data. The conventional economic indicators obtained from the NESDC contain information on GPP of various sectors and the number of populations of each province. LFS provided by the NSO of Thailand also contains information on the characteristics of workers, which is a key explanatory variable in the regression, such as level of education.

Besides the conventional economic indicators and ground survey, various satellite data from the Google Earth Engine platform service were used as dependent and explanatory variables in the regression. These data consist of urban land areas, Normalized Difference Drought Index (NDDI), Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Land Surface Temperature (Night), and Precipitation (rainfall), which is available from 2000–2020. More details of these variables will be discussed below.

3.5.1 Dependent Variable

The urban land area at the provincial level from 2000–2020 was used as a dependent variable in the regression. The urban land area was extracted from Terra and Aqua MODIS reflectance data via supervised classification techniques. Each data point was mapped to the political boundaries of each province in Thailand. Using the total urban land area obtained from the remote sensing technique as a proxy for urbanization has two advantages.

First, when compared to using conventional data such as construction permits or populations, the satellite data better capture the true extent of urbanization in Thailand because the tempo of urbanization is directly observed from space. In contrast, the data collected by the government might include only the formally registered population and miss out on the large informal sectors in developing countries. A comparison of using conventional and satellite data as a proxy for urbanization was extensively discussed in the previous chapter.

Second, the most refined unit of analysis in Thailand is available at the provincial level. However, the political boundaries of each province in Thailand could be starkly different in size. Thus, using the total urban land as a proxy for urbanization helps attenuate statistical bias arising from the modifiable areal unit problem (MAUP).

Specifically, using urbanization rate or density at the provincial level (instead of the absolute total land area) as a proxy for urbanization might produce biased results by artificially inflating or deflating the degree of urbanization of some provinces. For example, the total land area of Nakhon Ratchasima is higher than Bangkok Metropolis. However, its urbanization rate is different because the political area of Nakhon Ratchasima is much larger than Bangkok.

Thus, regressing total urban land on a set of economic and natural factors should reveal how explanatory factors affect the tempo of urbanization at the provincial level in Thailand.

3.5.2 Explanatory Variable

3.5.2.1 Tradable GPP Per Capita

Following theoretical construction in the previous section, the tradable GPP includes GPP of the key sectors as follows: agricultural, industrial (except mining and quarrying), information and communication, financial and insurance, real estate activities, professional, scientific, and technical activities, arts, entertainment and recreation, and other service activities. The real GPP value of these sectors were divided by number of populations to reflect rising productivity.

3.5.2.2 Resource GPP Per Capita

Resource GPP per capita is GPP of mining and quarrying sector. Again, to derive per capita value, the real GPP value of mining and quarrying sector were divided by number of populations.

3.5.2.3 Environmental Variables

Glaeser, Kolko, and Saiz (2001) and Duranton (2016) argued that environmental factors could affect city growth in US and Columbia. To control for such effect, this study includes various measurements of environmental factors, including NDDI, NDVI), NDWI, Land Surface Temperature (Night), and Precipitation (rainfall). The NDDI, NDVI, NDWI, and Temperature data were provided by the Terra MODIS satellite, while Precipitation was obtained from the CHIRPS dataset. All these datasets were processed and downloaded via the Google Earth Engine platform service, as mentioned earlier. The detail of each environmental variable is discussed below.

(1) Normalized Difference Vegetation Index (NDVI)

The NDVI has a key role in drought monitoring in ecosystem. NDVI measures dryness based on the changes in chlorophyll content and spongy mesophyll of the vegetation canopy. Those changes determine the greenness of both vegetation and leaves to human perception. The chlorophyll content and spongy mesophyll are reflected via absorption of visible red radiation and NIR radiation, respectively. The red and NIR radiation could be measured via remote sensing of Terra MODIS, and Sentinel-2 satellite. The NDVI could be expressed in mathematical formula as follows:

$$NDVI = ((NIR - Red))/((NIR + Red)) \tag{3.16}$$

The value of NDVI ranges between -1 to 1 where 1 indicates highest density of vegetation and -1 indicates lowest density of vegetation.

(2) Normalized Difference Water Index (NDWI)

Based on NDVI, NDWI was used to monitor drought based on change in both water content and spongy mesophyll. The water content is reflected via absorption of SWIR radiation. The NDWI could be expressed in mathematical formula as follows:

$$NDWI = ((NIR - SWIR))/((NIR + SWIR)) \tag{3.17}$$

The value of NDWI ranges between -1 to 1 where value that lies between 0.2-1 indicates existence of water surface and value between -1 to 0.2 indicates non-aqueous surface of humidity.

(3) Normalized Difference Drought Index (NDDI)

Following Gu, Brown, Verdin, and Wardlow (2007), the NDDI is combination of NDVI and NDWI geospatial indicators. The NDWI could be expressed in mathematical formula as follows:

$$NDDI = ((NDVI - NDWI))/((NDVI + NDWI)) \quad (3.18)$$

The value of NDDI ranges between -1 to 1 where value of -1 suggests no drought condition and value of 1 indicates severe drought condition.

(4) Land surface temperature

The land surface temperature of both daytime and nighttime is measured via remote sensing of Terra MODIS satellite. The increasing temperature could lead to a decline in economic productivity, especially in developing countries (Burke, Hsiang, & Miguel, 2015). Nevertheless, land surface temperature at nighttime was selected as an explanatory variable in this study because higher nighttime temperature was found to negatively affect not only crop yield but office productivity in the Netherlands (Daanen et al., 2013; Mohammed & Tarpley, 2011).

(5) Precipitation

The precipitation is obtained from The Climate Hazards Group Infrared Precipitation with Station (CHIRPS) dataset. The CHIRPS dataset is the collection of both remote sensing via satellite and ground station data. This technique cross-checks the accuracy of precipitation data from satellite and ground stations.

Variations in precipitation could affect agricultural productivity in non-irrigated areas or trigger conflict in many developing countries that inherited extractive institutions (Amare, Jensen, Shiferaw, & Cissé, 2018; Miguel, Satyanath, & Sergenti, 2004). Although environmental factors are often thought to be exogenous or, at worst irrelevant, it could still be argued that the environment still plays some vital role in determining urban growth. For example, the availability of water was found to positively affect long-run city growth in Columbia (Duranton, 2016). The ideal temperature of January and July was also found to be a main predictor of recent city growth in the Southern US (Glaeser et al., 2001). As roughly 31 percent of workers in

Thailand were employed in the agricultural sector (World Bank 2021), environmental factors such as availability of water, amount of rainfall, and temperature (all of which are captured by satellite) could affect the tradable sector productivity and subsequent urban growth.

3.5.2.4 Share of Worker with Higher Education

Various well-known literature often considered human capital as a main driver of city growth, especially in developed countries (Glaeser, Saiz, Burtless, & Strange, 2004; Glaeser, Scheinkman, & Shleifer, 1995; Simon & Nardinelli, 2002). Higher wages and positive local amenities were often associated with higher levels of human capital in cities (Moretti, 2004; Shapiro, 2006), which subsequently attracted workers into cities. Thus, a share of workers with higher education was used to control the effect of human capital on city growth. The main data used in the regression models were briefly summarized in Table 3.4.

Table 3.4

Summary of Data

Variable type	Description	Unit	Source
Dependent variable	Urban land area	Square kilometer	MODIS
Explanatory variable of interest	Tradable GPP per capita	Million Baht per population	NESDC
Socio-economic variables	1. Resource GPP per capita	Million Baht per population	NESDC
	2. Share of worker with higher education	Percent of total workforce	LFS
Environmental variables	1. NDDI 2. NDVI 3. NDWI	-1 to 1	MODIS
	4. Land Surface Temperature (at nighttime)	Celsius	MODIS
	5. Precipitation	Millimeter of rain per year	CHIRPS

Note. This table was created from author's compilation.

3.6 Result

3.6.1 Descriptive Statistics

Table 3.5 displays each variable's mean and standard deviation for the years 2000, 2010, and 2020. The average urban land area steadily increased from 71 square kilometers in 2000 to 82.85 square kilometers in 2020. Similarly, the real tradable GPP per capita remarkably rose from 44,643 Baht a year to 66,663 Baht a year during the same period. But the change of real resource GPP per capita depicted more fluctuation.

Table 3.5

Descriptive Statistics of Variables in the Regression

	2000	2010	2020
Urban land area	71.34	75.83	82.85
	(110.32)	(120.34)	(133.36)
Real tradable GPP per capita	44,643.82	61,675.62	66,663.8
	(63624.52)	(81362.05)	(77187.77)
Real resource GPP per capita	2,719.98	3,465.806	2,792.972
	(16439.98)	(20017.47)	(14807.39)
Sum of NDDI	2.34	2.35	2.78
	(0.80)	(1.13)	(1.14)
Sum of NDVI	7.21	7.29	7.52
	(1.22)	(1.18)	(1.23)
Sum of NDWI	5.18	5.48	5.28
	(1.12)	(1.31)	(1.43)
Sum of precipitation	1,792.48	1,848.25	1,787.07
	(594.05)	(590.00)	(741.31)
Land surface temperature at night	21.49	22.93	22.73
	(1.22)	(1.35)	(0.98)
Share of worker with college degree	3.67	5.81	8.12
	(2.23)	(2.45)	(3.38)

Note. Standard deviations are in parentheses.

Compared to economic variables, the mean values of environmental factors were relatively stable but still showed some noticeable trends. The mean of NDDI and NDVI slightly increased during the past two decades, while NDWI and precipitation increased in 2010 but decreased in 2020. The temperature at night became warmer as the average land surface temperature at night increased by 1.24 degrees Celsius within 20 years. The NDDI, NDWI, and precipitation showed a significant change in standard deviation, reflecting the higher variability.

The level of human capital in Thailand significantly progressed as the share of workers with college degrees rose from 3.67 percent of the total workforce in 2002 to 8.12 percent in 2020, indicating higher accessibility to tertiary education.

3.6.2 Regression Results

The effect of increased tradable sectors' productivity on urban land expansion was evaluated via equations 3.12, 3.13, and 3.14. Since each environmental factor was estimated separately (as mentioned earlier), each regression contains four specifications dubbed M1, M2, M3, and M4, respectively. Table 3.6 shows the results of non-spatial panel regression. The results suggest that lagged value of tradable GPP per capita significantly and positively affected urban growth, which is consistent with the derived theory that tradable sectors' productivity leads to urbanization. Resource GPP per capita, though positive, remains insignificant in all four specifications. These results are not surprising because resource income was dominated by industrial income since the implementation of the first national economic and social development plan in the 1950s, which fits well with Proposition 2.

Table 3.6

Results of Non-spatial Panel Regression

Dependent: Urban land area	M1	M2	M3	M4
Real tradable GPP per capita (at t-10)	0.05***	0.06***	0.05***	0.05***
	(2.91)	(3.36)	(3.04)	(2.96)
Real resource GPP per capita (at t-10)	0.145	0.154	0.146	0.01
	(1.22)	(1.29)	(1.23)	(0.85)

Table 3.6

Results of Non-spatial Panel Regression (Cont.)

Dependent: Urban land area	M1	M2	M3	M4
NDDI (at t-10)	-2.15***			
	(-2.71)			
NDVI (at t-10)		0.27		
		(0.26)		
NDWI (at t-10)			1.57**	
			(2.01)	
Precipitation (at t-10)				0.01***
				(5.08)
Average land surface temperature at night (at t-10)	0.72*	0.40	0.59	0.39
	(1.65)	(0.95)	(1.35)	(0.93)
Share of worker with college degree (at t-10)	61.80***	66.45***	62.73***	46.49**
	(3.24)	(3.48)	(3.28)	(2.43)
Constant	62.01***	61.53***	51.45***	57.11***
	(6.60)	(4.74)	(4.57)	(6.12)
Observation	684	684	684	684
Within R-squared	0.0659	0.0547	0.0608	0.0933
Between R-squared	0.1477	0.1426	0.1066	0.0833
Overall R-squared	0.1450	0.1401	0.1046	0.0802

Note. t-statistics are reported in parentheses.

* p<0.05, ** p<0.01, *** p<0.001.

Environmental variables were found to have a significant impact on urban land expansion. Specifically, NDDI, a drought measurement, negatively affected urban land growth, while NDWI and precipitation positively affected urban land growth. Nevertheless, the impact of land surface temperature at night on urban land was less robust, showing a significant impact (at a much wider confidence interval) only in the M1 specification. Like tradeable sectors' productivity, regressions suggest that the level of human capital significantly impacted urban land expansion in all specifications.

In addition to the standard non-spatial panel regression, this chapter applied spatial panel regression techniques that explicitly control for spatial autocorrelation, which could potentially arise since the data are collected in the form of geographical units. Spatial autocorrelation was illustrated by calculating a bivariate

Local Moran's I statistics. The results of Moran's I statistics were showed in Appendix B. The application of spatial panel regression techniques controlled both spatial autocorrelation and unobserved provincial characteristics, making the results more robust.

Table 3.7

Result of Spatial Lag Panel Regression

Dependent: Urban land area	M5	M6	M7	M8
Real tradable GPP per capita (at t-10)	0.05***	0.05***	0.05***	0.05***
	(2.79)	(3.22)	(2.94)	(2.87)
Real resource GPP per capita (at t-10)	0.15	0.15	0.15	0.10
	(1.31)	(1.38)	(1.32)	(0.92)
NDDI (at t-10)	-2.01***			
	(-2.69)			
NDVI (at t-10)		0.03		
		(0.02)		
NDWI (at t-10)			1.36*	
			(1.85)	
Precipitation (at t-10)				0.01***
				(5.18)
Average land surface temperature at night (at t-10)	0.69*	0.38	0.55	0.38
	(1.68)	(0.95)	(1.35)	(0.97)
Share of worker with college degree (at t-10)	52.73***	56.07***	53.90***	39.41**
	(2.86)	(3.04)	(2.92)	(2.14)
rho	0.10**	0.12**	0.10**	0.09*
	(1.99)	(2.23)	(1.96)	(1.66)
Observation	684	684	684	684
AIC	4032.5	4039.8	4036.3	4013.3
BIC	4064.2	4071.4	4068	4045
Within R-squared	0.08	0.07	0.07	0.10
Between R-squared	0.13	0.13	0.10	0.07
Overall R-squared	0.13	0.13	0.09	0.07

Note. t-statistics are reported in parentheses.

* *p*<0.05, ** *p*<0.01, *** *p*<0.001.

The results of the spatial lag panel and spatial error regression were presented in Table 3.7 and 3.8, respectively. The signs and significance levels of each variable in Tables 3.7 and 3.8 are consistent with the result presented in Table 3.6. According to Table 3.7, tradable GPP per capita significantly affected urban growth, while resource income had no meaningful effect. The impact of natural factors on urban growth could be attributed to the same set of climate variables, and the impact of human capital on urban growth was also found to be significant.

Nevertheless, the rho and lambda coefficient depicted in Table 3.7 and 3.8 provides some meaningful discussion, as expected in the methodology section. Specifically, the rho coefficient is statistically significant in all specifications of Table 3.7, while the rho coefficient (presented in Table 3.8) is not. The result suggests that spatial spillover operated through a change in the dependent variable or through a change in the error term. In other words, urban land growth in a particular province spill over to its neighbor in a predictable way.

Table 3.8

Results of Spatial Error Panel Regression

Dependent: Urban land area	M9	M10	M11	M12
Real urban tradable GPP per capita (at t-10)	0.05***	0.05***	0.05***	0.05***
	(2.84)	(3.17)	(3.00)	(2.93)
Real resource GPP per capita (at t-10)	0.15	0.16	0.15	0.10
	(1.31)	(1.39)	(1.32)	(0.93)
NDDI (at t-10)	-2.14***			
	(-2.85)			
NDVI (at t-10)		0.17		
		(0.16)		
NDWI (at t-10)			1.55**	
			(2.08)	
Precipitation (at t-10)				0.005***
				(5.39)
Average land surface temperature at night (at t-10)	0.72*	0.40	0.58	0.39
	(1.74)	(0.98)	(1.42)	(0.98)

Table 3.8

Results of Spatial Error Panel Regression (Cont.)

Dependent: Urban land area	M9	M10	M11	M12
Share of worker with college degree (at t-10)	59.97***	62.66***	61.33***	44.69**
	(3.14)	(3.23)	(3.22)	(2.39)
lambda	0.0185	0.0377	0.0148	0.0224
	(0.29)	(0.58)	(0.23)	(0.37)
Observation	684	684	684	684
AIC	4036.3	4044.3	4040.1	4015.9
BIC	4068	4075.9	4071.7	4047.6
Within R-squared	0.07	0.05	0.06	0.09
Between R-squared	0.15	0.14	0.10	0.08
Overall R-squared	0.14	0.14	0.10	0.08

Note. t-statistics are reported in parentheses.

* $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Table 3.9

Results of Dynamic Panel-Data Regression

Dependent: Urban land area	M13	M14	M15	M16
Urban land area (at t-1)[4]	1.10***	1.11***	1.11***	1.10***
	(230.89)	(233.41)	(230.77)	(224.29)
Real urban tradable GPP per capita (at t-10)	0.01***	0.01***	0.01***	0.01***
	(3.60)	(3.70)	(3.64)	(3.56)
Real resource GPP per capita (at t-10)	-0.003	-0.004	-0.004	-0.007
	(-0.16)	(-0.16)	(-0.15)	(-0.29)
NDDI (at t-10)	0.004			
	(0.04)			
NDVI (at t-10)		-0.12		
		(-0.99)		

[4] It is worth noting that the coefficients of lagged dependent variable in M13-M16 models are slightly above 1. Empirically, it describes the characteristic of data and does not damage interpretation of regression model since the objective of the study focuses on short-term effect of economic and environmental factor on urbanization in contrast to long-term one.

Table 3.9

Results of Dynamic Panel-Data Regression (Cont.)

Dependent: Urban land area	M13	M14	M15	M16
NDWI (at t-10)			-0.03	
			(-0.32)	
Precipitation (at t-10)				0.0002
				(1.58)
Average land surface temperature at night (at t-10)	-0.19***	-0.19***	-0.19***	-0.19***
	(-4.21)	(-4.54)	(-4.38)	(-4.44)
Share of worker with college degree (at t-10)	4.32*	4.14*	4.32*	4.30*
	(1.77)	(1.69)	(1.77)	(1.76)
Constant	-4.32***	-3.37**	-4.11***	-4.45***
	(-4.44)	(-2.47)	(-3.52)	(-4.55)
Observation	532	532	532	532

Note. z-statistics are reported in parentheses.

* p<0.05, ** p<0.01, *** p<0.001.

In addition to the non-spatial and spatial panel regression, Table 3.9 displays the estimated result of dynamic panel-data regression. The regression confirms an inertia effect of urban land expansion, that is, urban land expansion in the past affected urban land expansion. The dynamic expectations of workers influenced urban land expansion in Thailand. Notably, the real urban tradable GPP per capita and share of workers with college degrees remain significant but with much lower estimated coefficient values. The environmental factors are insignificant in the dynamic panel model except for average land surface temperatures at night time, which was found to affect urban land expansion negatively.

Though signs and significance levels of each variable are remarkably similar across model specifications, the estimated coefficients of spatial panel regression are smaller than non-spatial regression. An explanation and discussion of estimated results are provided below.

These results lead to six insights. First, urban tradable sectors productivity positively affected urban land expansion in Thailand as predicted by Proposition 2, which is consistent with previous literature (Bai et al., 2012; Deng et al., 2008; Gao et al., 2015; G. Li et al., 2018; Liu et al., 2016). But the coefficient values of revenue from natural resources, though positive, are statistically insignificant in all specifications, contradicting the prediction of Proposition 1. These results suggest that rising productivity from urban tradable sectors drove urban land expansion in Thailand by allowing workers to efficiently meet their subsistence food consumption, increasing their surplus income and demand for products from the urban sector. However, the increased revenue from natural resources did not affect urban expansion in Thailand. Resource revenue represents a tiny share of total real GDP compared to revenue from urban tradable sectors and therefore did not significantly affect surplus income accumulation. This paper leaves room for future studies to investigate this issue.

Second, according to the results from standard panel regression, spatial lag, and spatial error regression, the natural factor was still relevant in predicting urban land expansion in Thailand. The estimated results show that higher water availability mattered for urban land growth prediction, suggesting that favorable natural factors affected urbanization. A higher amount of water might lead to higher yield and productivity of agricultural activities, generating higher income and demand for products and services from urban sectors. A hospitable environment might prevent workers from leaving the areas and induce population growth. However, coefficients of environmental factors turned insignificant in the dynamic regression, implying that when urban growth's autocorrelation was considered, water availability was irrelevant to urbanization. The effects of the availability of water were, therefore, at best indirect.

Third, human capital's substantial impact on urban land expansion in Thailand indicated positive externalities of higher education on urban growth. A higher share of workers with higher could attract more workers into cities and make them grow bigger. A higher number of college workers was associated with higher wages and amenities that come with higher education, such as university, park, theatre, or high-paying jobs. This migration made cities grow and encouraged agglomeration externalities, a process that would promote more migration through higher wages.

Fourth, the urban land expansion of the province in Thailand positively affected the urban land expansion of its neighboring province in a predictable way.

Fifth, it was found that controlling for spatial autocorrelation lowered estimated coefficients. The possible explanation is that the standard non-spatial regression might not be able to control for some omitted variables that drove urban growth but were captured via spatial regression. In other words, the impacts of estimated coefficients on urban growth were attenuated by spatial autocorrelation in the regression model.

Sixth, dynamic expectations of workers affect the tempo of urban land expansion in Thailand. Workers observed actual urban land expansion in a particular province and decided to move there in the next period. The previous urban land expansion was also found to have a strong impact on current urban land expansion.

3.6.3 Policy Recommendations

By synthesizing the six insights from the regression results, some policy recommendations could be made to promote urbanization in Thailand. According to the estimated coefficients of the urban tradable sector GPP per capita and human capital, higher education's positive externalities and increased urban sector productivity were Thailand's main drivers of urbanization. The dynamic expectation of workers was found to impact urbanization substantially. The impact of environmental factors, such as water availability, should be addressed.

Moreover, the results from Chapter 2 suggests that decentralized urbanization could solve both spatial inequality and sluggish economic growth in Thailand and that a polycentric urban structure in Thailand, as opposed to the current monocentric structure, should be promoted.

The urban growth promotion policy should seek to establish highly productive jobs in all regional cities. The policy should be able to redistribute highly educated workers from the BMR to regional cities. Based on the autocorrelation of urban land expansion, the policy should change the workers' expectations about the future urbanization of the prospective cities, unlocking self-reinforcing mechanisms in the urbanization process. In other words, such a policy must be long-term-oriented and be implemented with consistency and perseverance.

3.7 Conclusion

The rapid expansion of cities occurs in many developing countries, especially in Asia. Satellite data suggest that cities in Thailand are expanding. But urban land expansion in Thailand concentrates on BMR and some areas in the Eastern Economic Corridor that locate eastward of Bangkok, reflecting the monocentric growth pattern of the urban system. In addition to urban concentration, most of the economic activities in Thailand take place in the same area. Such extreme regional inequality led to serious social, economic, and political problems. Besides, analysis from Chapter 2 suggests that workers in large cities were found to be more productive. Thus, investigating the drivers of urbanization is essential for policymakers.

This chapter examined the role of economic development and other relevant factor, such as natural factors and human capital, on urban land expansion in Thailand. Like Chapter 2, the unit of analysis of this chapter is the individual province in Thailand. The non-spatial and spatial panel regression models were applied to achieve the objectives.

The estimated results suggest that increased productivity of urban tradable sectors positively affected urban land expansion by allowing workers to meet their basic needs and accumulate surplus income. It was also found that natural factors such as water availability and rainfall made cities grow bigger, implying that a hospitable environment prevented workers from leaving and attracted workers from other areas. A larger share of workers with higher education was found to significantly impact urban growth. Additionally, urban land expansion in the past was found to have a substantial impact on the present urban land expansion. Lastly, urban land expansion in one province led to land expansion in a neighboring area, indicating spatial spillover effects of urban land expansion in Thailand.

CHAPTER 4
PREDICTING LOCAL ECONOMIC DEVELOPMENT AND URBAN LAND EXPANSION: CASE STUDY OF BAN CHANG DISTRICT, RAYONG PROVINCE, THAILAND

4.1 Introduction

Urbanization is a spectacular phenomenon in modern times. Fifty-six percent of the world's population lives in cities (World Bank, 2021). This trend, however, is not declining. The United Nations projected that seven billion people would live in urban areas by 2050. Ninety percent of future urbanization is expected in Asia's low- and middle-income partners. Sustainable urbanization should therefore be emphasized to ensure that the benefits of urbanization are maximized and that its negative impact on the environment is minimal (United Nations, 2018).

Urbanization is associated with agglomeration forces and economies of scale, generating economic growth and innovation. Fujita et al. (2001) explained how the interaction between socio-economic factors, such as population growth and reduction in transportation costs from an improved road network, contribute to the growth and sustainability of the city, particularly in a monocentric manner. On the other hand, cities are also the main source of crime, pollution, disease, global warming, and environmental degradation. Urbanization could be hazardous when its rapid growth process is characterized by haphazard development and poor urban planning, resulting in urban sprawl or low-density urban development with lengthy commutes and limited urban mobility (Harari, 2020). Urbanization will become unsustainable if urban planners fail to innovate the cities, resulting in a collapse.

Like other developing countries, the number of urban inhabitants in Thailand has increased over the last six decades. The number of people living in urban areas in Thailand steadily increased from 5 million in 1960 to 37 million in 2021 (World Bank, 2021). In addition to rural-to-urban migration, rapid economic growth coincides with urbanization. Many studies suggested that urbanization fosters economic growth in Thailand through agglomeration externalities (Houbcharaun, 2013; Limpanonda, 2012; Southichack, 1998; Prasertsoong & Puttanapong, 2022). However, the urbanization process in Thailand, particularly in the BMR, is characterized by urban sprawl, the rapid, unplanned, and haphazard growth of urban land (Iamtrakul, Padon, & Klaylee, 2022). Without knowing future urban growth patterns and economic growth at the city level, such development could result in congestion, extremely long commuting hours, pollution, and a decline in the general well-being of the urban population (Arfanuzzaman & Dahiya, 2019). To prevent potential congestion, urban sprawl, and environmental problems from urbanization found in the BMR, this study, therefore, aims to predict future urban land use patterns and economic growth in one of the fastest-growing cities in Thailand, Ban Chang district, by explicitly endogenizing the local economic growth into the land-use change modeling.

The content of this chapter was divided into six sections. The following section discusses the details of the selected study area. The third section overviews the related literature. Details on methodology and data are provided in the fourth section. The fifth section discusses the result of the prediction and limitations of this study. The last section is the conclusion.

4.2 Study Area: Ban Chang District

Located in the southwestern part of Rayong, as depicted in Figure 4.1, Ban Chang is one of the eight districts of Rayong province. Rayong is one of the three provinces chosen to be the main strategic area of the Eastern Seaboard Development Program and the Eastern Economic Corridor (the other two are Chonburi and Chachoengsao). Since their implementations in 1982 and 2016, both industrialization programs have increased the income per capita of the chosen provinces and substantially expanded urban areas. Among the three provinces, Rayong province has undergone the most pronounced transformation in both aspects (Tontisirin & Anantsuksomsri, 2021).

Figure 4.1

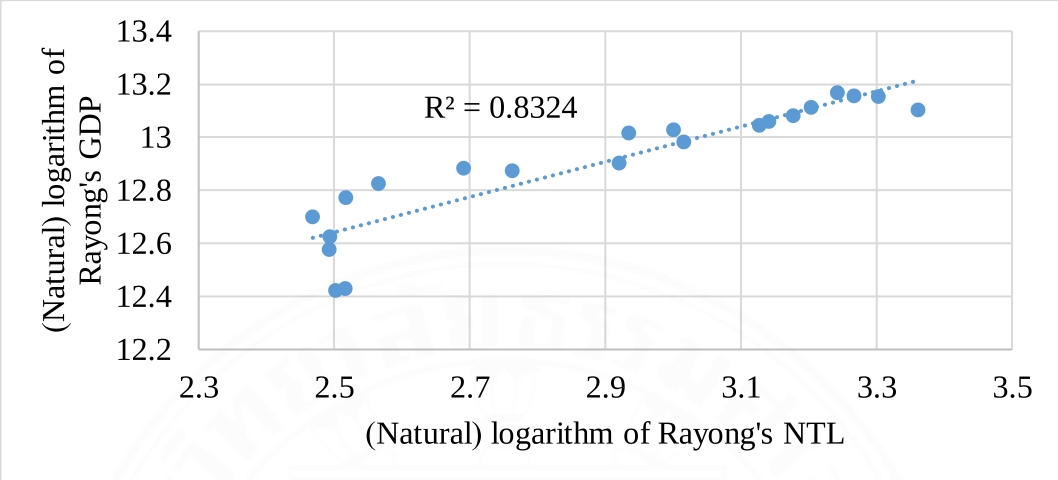
Geographical Location of Ban Chang District

Panel A Panel B

Note. A red solid line of panel A shows the political of Rayong province. A red line pattern of panel A shows the political of Ban Chang district. A satellite image of Ban Chang district is shown in panel B.

Figure 4.2

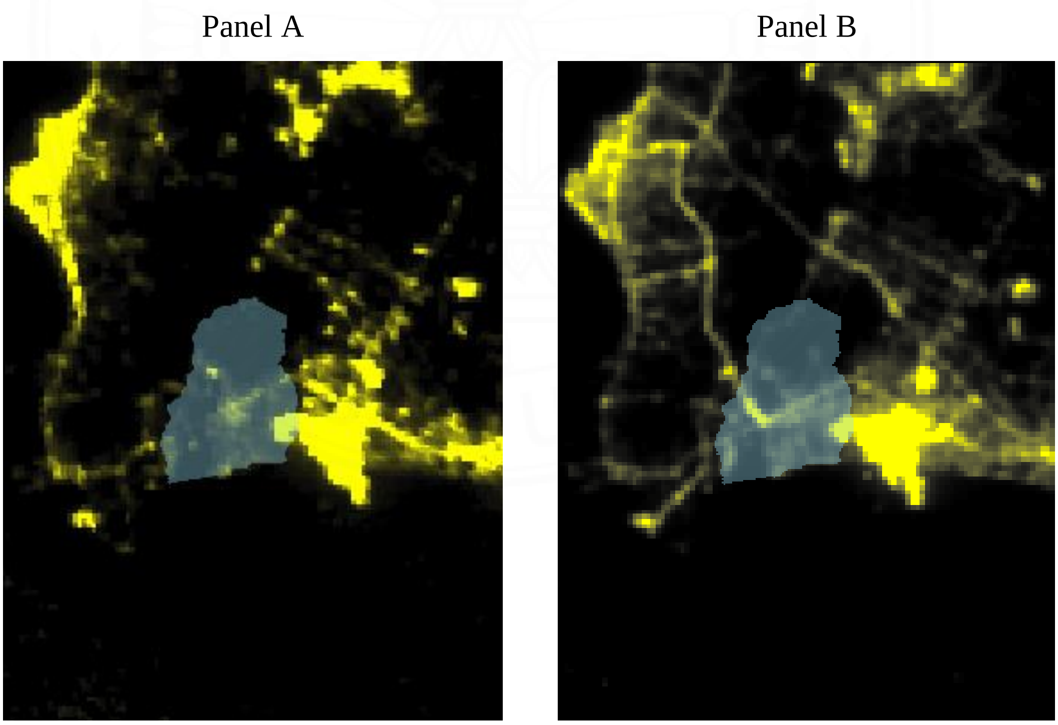
Relationship Between Rayong's NTL Density and Rayong's GDP (2000–2020)



Note. This table was created from author's calculations.

Figure 4.3

NTL Density of Ban Chang District



Note. (i) Panel A: NTL image of Ban Chang district in 2012; Panel B: NTL image of Ban Chang district in 2021. (ii) A blue mask on panel A and B represents the political of Ban Chang district. Yellow pixels represent the built-up areas. Black pixels represent the non-built-up areas such as sea, forest, and agricultural land.

As one of the eight districts in Rayong, Ban Chang is an attractive study area for five reasons. First, Ban Chang is part of the Eastern Seaboard Development Program and the Eastern Economic Corridor, making it one of Thailand's most rapidly changing cities.

Second, Ban Chang is where U-Tapao International Airport and several industrial parks are located. The airport and the industrial parks are expected to be improved to accommodate tourists and foreign direct investment in light of the Eastern Economic Corridor initiation.

Third, in 2022, the Eastern Economic Corridor and the National Telecom of Thailand chose Ban Chang district as the first model prototype of a "5G enabled" smart city in Thailand (Bangkok Post, 2022). In effect, the most up-to-date digital infrastructures would be invested in Ban Chang, accelerating the city's growth in the future.

Fourth, economically, Ban Chang is one of the fastest-growing districts in Rayong province. Observed from outer space, NTL is conventionally applied to proxy economic growth, especially in regions where official data is lacking (Henderson, Storeygard, & Weil, 2012; Martínez, 2022). Following empirical studies econometrically verifying that NTL data can proxy the local economic growth in Thailand (Puttanapong et al., 2022; Sangkasem & Puttanapong, 2022), the regression result in Figure 4.2 specifically confirms the statistically significant relationship between Rayong's NTL density and Rayong's GDP[5] during the period 2000–2020. Based on the findings, the NTL density was used as a proxy for the GDP of Ban Chang district. In particular, the NTL density obtained from the Visible Infrared Imaging Radiometer Suite (VIIRS) satellite indicates that the compound annual growth rate of the NTL density of Ban Chang during the period 2012–2020 was 10.9 percent.

Figure 4.3 illustrates an increased NTL density in Ban Chang district between 2012 and 2021, reflecting a substantial increase in economic activity and urban area.

Lastly, in addition to the airport, factories, and digital investment, Ban Chang has a large agricultural and service sector, making it a unique study area.

[5] The provincial GDP is officially recognized as GPP by the NESDC

4.3 Literature Review

4.3.1 Land-Use Change Model with Socio-Economic Factors

An analysis of urban land-use change is an interdisciplinary study involving the disciplines of economics, geography, sociology, statistics, and engineering to investigate the evolution and pattern of land-use change and predict the future direction of urban land expansion. The interdisciplinary nature of land-use change studies paves the way for diverse adoption of its theoretical basis by many fields of study to serve their objectives. Consequently, various techniques were developed or configured to fit the specific purposes of various research fields. Nevertheless, urban land change theory development can be concisely summarized below.

Since 1950, the development of the theoretical concept of land-use change modeling has evolved along two parallel paths. The first path of development is the transformation from macro-foundation to micro-foundation modeling. The recent micro-foundation urban modeling considers both heterogeneity and interaction between agents at a tinier scale. The second development path extends from the static to the dynamic model as the time dimension becomes an indispensable factor in land-use change research.

The first generation of urban growth models (the Lowry model) is static and operates at an aggregated level. Forrester (1969) proposed the dynamic model of urban growth. The Forrester model, however, is non-spatial. Based on the foundation of the Forrester model, the researcher developed a spatial urban growth model and tried to incorporate natural and socio-economic factors such as slope, elevation, population, and economic variables into the model. However, although urban growth models deal with spatial aspects, they frequently need more accuracy and tend to have a wide-ranging focus. They usually function on a city-wide scale and cannot account for intricate details at a smaller micro-level, such as individual units or cells. Consequently, these models bear a stronger resemblance to macro-economic models rather than contemporary spatial models.

Treating external socio-economic factors is also central to developing the land-use change model. Early land-use models consider the concept of land-use change implicit, i.e., those models view land-use change as an overall city growth with complete negligence of the smaller parts of the spatial perspective. The only socio-economic factors that affect land-use change are trade costs and population growth. For example, one of the most distinctive studies of the spatial economy and the formation of cities derived by Fujita et al. (2001)—New Economic Geography—theoretically showed that the population growth rate and lower transportation costs caused by better technology and infrastructure are the key factors in determining the growth of the city.

The city will become a monocentric growth pole when it has a strong centripetal force that pulls the population of the neighboring areas into that city. On the contrary, a polycentric urban structure is possible if the city possesses some centrifugal force that prevents the emergence of total agglomeration. The sustainability of either monocentric or polycentric urban structure depends on the parameters of the model as well as the dynamic of socio-economic factors such as population growth, reduction in trade cost, and agglomeration of cities. The distance to the road, the distance to a major intersection, and the distance to a transportation hub also determine the location of the new city.

Toward the mid-1980s, with the improvement of computational technology, the consideration of more micro and dynamic treatment of land-use change was taking ground. Technically speaking, the Cellular Automata (CA) and agent-based models were developed to cope with rising demand and to take both micro and dynamic aspects into account. The CA model has five main advantages. First, the CA model is spatial in nature. Second, the CA model is compatible with GIS and remote sensing techniques, which are essential features in monitoring and analyzing urban growth. Third, the CA model is dynamic. Thus, researchers could measure the spatio-temporal dynamics of urban growth. Fourth, the concept of CA model is simple. Lastly, the CA model pays attention to micro-level and bottom-up approaches.

The pioneering studies of CA-based urban growth models are the studies of White and Engelen (1993, 1997), Batty and Xie (1994), Clarke and Gaydos (1998), Li and Yeh (2002), and Barredo, Kasanko, McCormick, and Lavalle (2003). Although the CA model considers socio-economic factors when analyzing land-use change, it lacks theoretical support from a solid economic foundation as the core basis of CA stems from the concept of self-organization, which could be traced back to the idea of Von Neumann and Ulam in 1948 and John Horton Conway who invented the explicit representation of CA known as "Conway's Game of Life" in 1970.

Although the CA specifically deals with land-use change analysis, its ability to target socio-economic activities is limited. Thus, an agent-based model was developed to solve some of the limitations of the CA. Technically speaking, the agent-based model extended the capabilities of the CA model to take individual agents' decision-making into account.

The remarkable advantage of agent-based modeling is its ability to embed each agent with its attributes, goals, and behaviors. The attributes of individual agents are explicit socio-economic factors such as income. The goals and behaviors of individual agents could be different. For example, homeowners and real estate developers have different goals and behaviors. The primary distinguishing characteristic of agent-based modeling in land-use modeling is utilizing a utility function, which enhances its compatibility with microeconomic theory, where both homeowners and real estate maximize their utility and profit.

In addition to CA and agent-based modeling, Conversion of Land Use and its Effects at a Small Regional extent (CLUE-S) is an alternative method to conduct land-use change modeling. The CLUE-S is suitable for operating in small regions and requires computational capability. Like CA and agent-based models, the CLUE-S could perform spatial allocation of land-use change at fine resolution and interact with external socio-economic and biophysical driving factors. The CLUE-S is a modification of the CLUE model by Verburg et al. (2002). While the CLUE only operates at a gross level (e.g., national or continental extent) and land-use input data must be derived from a census or survey, the CLUE-S operates at a finer resolution (e.g., at pixel level), allowing researchers to utilize remote sensing data in the modeling procedure. The main advantage of the CLUE-S model is its ability to incorporate non-spatial analysis with spatial analysis explicitly. The CLUE-S leaves room for advanced econometric analysis or another structural economic modeling in land-use allocation. In other words, the CLUE-S could allocate land-use demand scenarios, calculated with economic techniques, to specific locations within the study area (Verburg et al., 2002).

4.3.2 Empirical Literature

As the model's configuration and purpose of studies vary according to their fields of study, previous empirical literature studied land-use change with various objectives and model specifications. The review of the empirical literature on land-use change in this section provides evidence of the effect of socio-economic factors on urban land expansion. This section also considers empirical studies incorporating feedback-loop mechanisms into land-use change modeling.

Previous empirical studies applied a Logistic-CA model to their studies to identify the effect of socio-economic factors on land-use change or the conversion of agricultural land into built-up land. Logistic Regression is convenient for dealing with the state of land-use change because it is specifically designed to identify factors affecting the transition probability of a binary state, i.e., whether land-use change occurred (if occurred, assign value to 1) or not occurred (if not occurred, assign value to 0). Numerous factors and variables were used in previous studies but could be grouped into four main categories.

The first category is natural factors, which usually consist of slope, elevation, soil type, and water. The second category is socio-economic factors, which often include distance to urban settlement, distance to roads, distance to transportation hubs, population density, GDP or income, and housing rent. The third category is government policy, i.e., conservation area or land development plan. The fourth and third category is the neighborhood effect, which could either be included in the model or not be included in the model. Since the purpose of this paper is to consider economic data, this section focuses on discussing socio-economic variables in land-use change modeling.

Table 4.1

Previous Empirical Literature That Study Land-Use Change and Socio-Economic Factors

Author(s)	Area of study	Methodology	Key socio-economic variable	Sign of coefficient
Verburg et al. (2002)	Klang-Langat and Sibuyan Island	CLUE-S	-	-
Braimo & Onishi (2007)	Lagos, Nigeria	Logistic Regression	Manufacturing and services value-added	Positive
Hu & Lo (2007)	Atlanta, U.S.	Logistic Regression	Income	Negative
Claessens et al. (2009)	Betic Cordilleras, Spain	1. CLUE	Geological factors	-
Le et. al. (2012)	Hong Ha, Vietnam	1.Agent-Based model	Household income	-
Arsanjani et al. (2013)	Tehran, Iran	Logistic Regression	Population Density	Positive
Munshi et al. (2014)	Ahmedabad, India	Logistic Regression	Distance from Railway Station	Negative
Li et al. (2014)	Shanghai, China	Logistic Regression	Distance to railway	Negative
Jafari et al. (2016)	Hyrcanian, Iran	Logistic Regression	Distance to major roads	Positive
Losiri et al. (2016)	Bangkok Metropolitan, Thailand	Cramer’s V	Economic growth rate	Positive

Table 4.1

Previous Empirical Literature That Study Land-Use Change and Socio-Economic Factors (Cont.)

Author(s)	Area of study	Methodology	Key socio-economic variable	Sign of coefficient
Liu and Feng (2016)	Ningbo, China	Logistic Regression	Distance to economic corridors	Negative
Gounaridis et al. (2019)	Attica, Greece	Random forest	Employment rate	-
Tripathy and Kumar (2019)	Delhi, India	CA	Distance to major roads	-
Dadashpoor et al. (2019)	Tabriz, Iran	Logistic Regression	Population growth and employment	Positive
Trisurat et al. (2019)	Nan, Thailand	Logistic Regression	Distance to major roads	Negative
Tontisirin and Anantsuksomsri (2021)	Eastern Economic Corridor, Thailand	CA–Markov model	-	-
Wang, Hu, Niu, Yan, and Zhen (2022)	Thailand	Multivariate linear stepwise regression	Urban population	Positive

Note. This table was created from author's compilation.

Table 4.1 overviews the evolution of the empirical literature on land-use change modeling with socio-economic variables in chronological order and the list of socio-economic variables used in previous literature. Table 4.1 shows a wide range of applications of Logistic Regression in land-use change modeling from the late 2000s to recent times. The prominent studies that applied Logistic Regression to their model are work of Braimoh and Onishi (2007), Hu and Lo (2007), Jokar Arsanjani, Helbich, Kainz, and Darvishi Boloorani (2013), Munshi et al. (2014), Li et al. (2014), Jafari et al. (2016), Liu and Feng (2016), and Dadashpoor et al. (2019).

However, only some studies incorporate economic variables such as GDP, wages, income, and employment into land-use change models. A few of these studies are the studies of, in chronological order, Hu and Lo (2007), Braimoh and Onishi (2007), Losiri et al. (2016), and Dadashpoor et al. (2019). The other studies focused on other socio-economic variables such as distance to a major road, railway, transportation hub, and population density. Besides Logistic Regression, some studies resorted to machine learning techniques such as random forest (Gounaridis et al., 2019) and Markov models (Tontisirin & Anantsuksomsri, 2021; Tripathy & Kumar, 2019). Some researchers relied on Cramer's V technique to identify the effects of socio-economic variables on land-use change (Losiri et al., 2016).

Table 4.1 also shows the sign coefficient of key socio-economic variables affecting land conversion into built-up areas from Logistic Regression. The sign of the estimates reveals the pattern of urbanization, which is measured as land-use change in each study area. Table 4.1 shows that the economic variables—GDP growth of value-added—positively affected the possibility of land conversion into built-up areas. However, when a level of per capita income was used, the coefficient was found to be negative but statistically insignificant (Hu & Lo, 2007).

The coefficient of other socio-economic variables ranges from negative to positive values. The effects of population density and connectivity—distance to road, distance to railway—have positive and negative values. The negative value of the coefficient indicates that the urbanization process in the specific study area is spreading outward from the urban center, i.e., the denser populated areas were less likely to be converted into a built-up area, or the built-up areas were more likely to be developed near an expanding road network (Li et al., 2014; Liu & Feng, 2016; Munshi et al., 2014). Some researchers found positive coefficients of population density and connectivity, which suggests that, instead of stretching out, urban centers in the study area of this research were becoming denser, particularly in Iran's cities (Arsanjani et al., 2013; Jafari et al., 2016).

As mentioned earlier, positive feedback between urbanization and economic growth will likely occur. Some researchers dealt with the issue of simultaneity by explicitly plugging in feedback mechanisms into land-use change modeling. For example, Claessens et al. (2009) predicted future agricultural land-use changes by incorporating a feedback mechanism between soil erosion due to agricultural activity and land-use change impact on soil erosion. Similarly, Le, Seidl, and Scholz (2012) applied agent-based models and the Land-Use Dynamics Simulator (LUDAS) to introduce a long-term feedback mechanism between land-use action and land-use decision.

Moreover, the previous studies also considered various socio-economic factors, such as community income, as a supporting environment for future land-use decisions where individual agents change their behaviors in response to the change in socio-economic and bio-physical conditions induced by land-use change. However, the positive feedback mechanism between urban land-use changes and socio-economic variables is mostly neglected by many researchers in the field.

Regarding the application of the CLUE-S model, Verburg et al. (2002) applied the CLUE-S model in an empirical study. The study area in Verburg et al.'s (2002) paper is the Klang-Langat Watershed, Malaysia, and the small island of Sibuyan, Philippines. To perform land-use conversion allocation, Verburg et al. (2002) applied Logistic Regression to create a probability surface of all land-use types and use demand for all land-use types as another input for the CLUE-S model. Additionally, many physical and socio-economic variables could be used in this study, making it more akin to CA modeling. Recently, Moulds, Buytaert, and Mijic (2015) developed an R package called lulcc, open-source and extensible software that could be applied to the CLUE-S model without proprietary restriction.

4.3.3 Empirical Literature in Thailand

Previous empirical literature that studied urban land-use change and economic development in Thailand focused on urban land expansion and socio-economic change phenomena in main economic cities such as the Bangkok Metropolitan area and the Eastern Economic Corridor (which consists of Chonburi, Rayong, and Chachoengsao provinces) (Losiri et al., 2016; Tontisirin and Anantsuksomsri, 2021). These studies applied the logistic-CA and CA-Markov model to examine the effect of socio-economic variables and government policy on land-use change in Thailand's megacities. The studies analyzed the models and found a positive effect of economic growth, connectivity, and high-tech industrial promotion policies on converting agricultural areas into built-up areas.

Besides the megacities, Trisurat et al. (2019) analyzed the impact of socio-economic change on the environment in Nan province, Thailand. They examined the impact of various scenarios on future land-use changes of crops and biodiversity. By applying Logistic Regression, land-use change model (Dyna-CLUE), and FRAGSTATS, Trisurat et al. (2019) discovered that areas with good road access are suitable for a rice paddy, and road expansion into pristine forest areas increases the risk of deforestation in Nan province. Like previous studies outside Thailand, socio-economic factors were treated as exogenous variables in the land-use change model. More recently, the study by Wang et al. (2022) applied multivariate linear stepwise regression to investigate the driving factor of urban land expansion in Thailand between 2000–2020. They found that the growth urban population significantly affected urbanization in Thailand. The growth of international tourism also led to infrastructure development which reduced scrubland.

Although the previous studies considered the impact of socio-economic factors and land-use change in Thailand, applying land-use change modeling, especially the CA in Thailand or the global research community, disregards possible simultaneous causality between urban land-use change and economic growth. To the author's knowledge, this study is the first attempt to endogenize local economic growth into the land-use change model at the city level in Thailand. This study pioneers a new modeling technique to forecast economic growth at the district level in Thailand by using geospatial data.

4.4 Foundation of Land-Use Change Modeling

The basic concept of land-use change modeling in this section is based on a CA model of computation. CA model is recognized as an efficient technique that urban researchers use to investigate cities' dynamic and self-organizational evolution. CA could replicate organisms' evolution in living things such as human cells. Urban planning researchers applied CA to the evolution of urban structures because the concept of dynamic and self-organization evolution in CA could be incorporated with other statistical techniques and socio-economic variables, analyzing urban land expansion in greater detail of spatial units possible. More importantly, the fundamental concept of CA is the foundation of later generations of land-use change modeling, such as the CLUE-S model.

4.4.1 A Cellular Model of Urban Land-Use

Initially, the cellular model of urban land-use was applied to examine fundamental concerns of urban form rather than providing accurate simulations of the growth of specific cities. The model is simple yet capable of representing urban structures and analyzing the process of urbanization through time. The model consists of cells where each cell represents a type or state of land use. The land use classification can vary in detail, ranging from precise descriptions like residential or commercial and vacant land to more generalized categories such as built-up areas, vegetation, bare land, and water bodies. Each cell state could be converted into other land use types as the city grows according to transition rules.

In general, the CA model consists of four basic elements. First, a grid cell, which is a unit in finite cellular space. Second, the state of the cell where each cell has its own state or status which characterizes the cell. Third, the neighborhood cell, which is a set of surrounding cells in the grid system. Fourth, the transition rule, which dictates the future evolution of the cell (Losiri et al., 2016). Mathematically, the CA could be written in a 3×3 matrix form as shown in equation (4.1).

$$A_{i,j}^{t} = \begin{bmatrix} a_{i-1,j-1}^{(t)} & a_{i-1,j}^{(t)} & a_{i-1,j+1}^{(t)} \\ a_{i,j-1}^{(t)} & a_{i,j}^{(t)} & a_{i1,j-1}^{(t)} \\ a_{i+1,j-1}^{(t)} & a_{i+1,j}^{(t)} & a_{i+1,j+1}^{(t)} \end{bmatrix} \tag{4.1}$$

where $a_{i,j}^{(t)}$ represents the central cell (i,j) at time t. The surrounding cells are neighboring cell. The CA depends primarily on the state of the test cell, state of neighboring cell, neighboring type, and transition rules. The future state of $a_{i,j}^{(t)}$, $a_{i,j}^{(t+1)}$, depends on specified transition rule (f) (Tripathy & Kumar, 2019).

Following Verburg et al. (2004), CA model, in general, could be expressed formally as:

$$S_{ij}^{t} = f\left(S_{ij}^{t-1}, I^{t}, A^{t}\right) \tag{4.2}$$

where S_{ij}^{t} and S_{ij}^{t-1} denote the state of the cells (i,j) at time t and $t-1$, respectively. Like equation (4.1), the S_{ij}^{t} represents center cell. The state of the cells could be, for example, built-up area, vegetation, water, and bare soil. I^{t} represents vector of model input, which are spatial attributes and boundary conditions; A^{t} represents all parameters; and f represents transition rules.

In addition, probability rule could be used to determine conversion potential of a given location. Given that the model inputs consist of socio-economic factors, physical factors, neighborhood conditions, spatial constraints and stochastic disturbance factors, the probability of cells conversion is described by the equation (4.3) (Feng, Liu, Tong, Liu, & Deng (2011); White & Engelen (1993); Wu (2002)).

$$P_{ij}^{t} = (P_{l})_{ij}(P_{\Omega})_{ij}\, con() P_{r} \tag{4.3}$$

where P_{ij}^{t} is conversion probability of cell (i,j) at time t; $(P_{l})_{ij}$ is suitability conversion of cell (i,j) at time t which is calculated from socio-economic and physical factors via statistical analysis techniques. $(P_{\Omega})_{ij}$ defines the development density of a small cellular neighborhood. $con()$ represents constraint on land conversion. Lastly, P_{r} is the stochastic disturbance factors.

From equation (4.3), the $(P_{l})_{ij}$ is defined in general term, that is, $(P_{l})_{ij}$ could follows the standard Logistic Regression in determining suitability conversion or it could be determined by recent machine learning techniques such as random forest regression. Likewise, $(P_{\Omega})_{ij}$, $con()$, and P_{r} are also generally defined.

Following X. Liu, Li, Shi, Wu, and Liu (2008) and X. Liu et al. (2014), by applying the 3×3 Moore's neighborhood, the neighborhood function $(P_\Omega)_{ij}$ could be defined by the following equation:

$$(P_\Omega)_{ij} = \frac{\sum_{3\times 3} con\left(s_{ij}^t = dev\right)}{3\times 3 - 1} \tag{4.4}$$

where $(P_\Omega)_{ij}$ represents the development density of neighborhood and $\sum_{3\times 3} con\left(s_{ij}^t = dev\right)$ is the number of developed cells in the neighborhood window. The constraint could be defined as a protection zone or other built-up development restriction area. The constraint could also be used to prevent certain conversion from taking place – for example land conversion from built-up area to water should be restricted.

To deal with complexity of urban expansion, the stochastic disturbance factors, P_r. Following White and Engelen (1993), the stochastic disturbance factors could be defined as

$$P_r = (1 + (-\ln\gamma)^\alpha) \tag{4.5}$$

where $(1 + (-\ln\gamma)^\alpha)$ is the stochastic factors; γ takes value between 0 and 1; α controls the effect of stochastic factor which is ranging from 1 to 10.

From here, the transition rule could be applied after the development probability of the center cell, P_{ij}^t, was obtained from equation (4.3). The transition rule could be defined as:

$$S_{ij}^t = \begin{cases} Developed, & P_{ij}^t \geq P_{threshold} \\ Undeveloped, & P_{ij}^t < P_{threshold} \end{cases} \tag{4.6}$$

where S_{ij}^t is the state of the cell (i,j) at time t; P_{ij}^t is conversion probability of cell (i,j) at time t which ranges between $[0,1]$; and $P_{threshold}$ is threshold value that governs conversion of the cell. The transition rule here is simple. It dictates that cells convert from a certain type to another type – for example, from bare land to built-up area if the conversion probability of the cell (i,j) is greater than the threshold value. Again, the transition rule here is simple but quite general.

Alternatively, researchers could opt for another land-use modeling technique such as CLUE-S model as mentioned earlier. According to Moulds et al. (2015)'s summary of CLUE-S model, the transition rule in CLUE-S model is simple – it allocates each cell to highest conversion probability (or suitability) of particular land-use type as determined by the statistical analysis techniques. The amount of cell conversion will depend on the predetermined land-use demand of each type.

4.5 Methodology and Data

4.5.1 Methodology

4.5.1.1 Modelling Procedure

The flowchart in Figure 4.4 illustrates an overview of the urban land-use change prediction process that was adapted from Moulds et al. (2015). The Ban Chang land-use map is the main input. At the same time, slope, elevation, and distance to the existing urban area are explanatory variables for the Logistic Regression, the selected predictive model in this study.

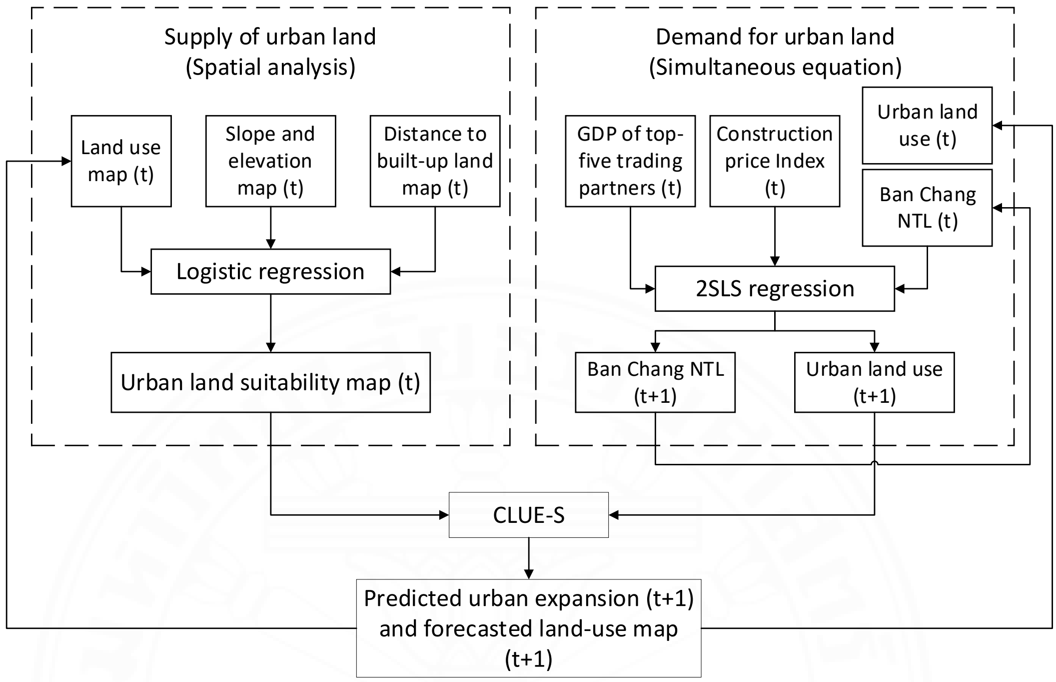
The urban probabilities (suitability) map would be created after conversion probabilities were calculated from the Logistic Regression. To generate future land-use maps, a user needs to supply the CLUE-S algorithm with the urban land demand of the next period. More details about urban land demand prediction and the CLUE-S algorithm are discussed in section 4.1.3 and section 4.1.4, respectively.

Users could use a built-in linear interpolation or a sophisticated econometrics model to project urban land demand between two periods. Nevertheless, this chapter applied a simultaneous equation model to predict urban land demand and Ban Chang GDP growth (as proxied by NTL) in the next period (t+1).

The detail of the calculation is given in section 4.5.1.2. After the urban land demand at period t+1 was obtained, the CLUE-S algorithm will allocate each cell according to the computed suitability map and urban land demand. The predicted urban land-use/land-demand and the Ban Chang GDP at time t+1 will be used to predict the same set of outcomes for the next period. The observed land-use map at period t+1 will be used to assess the accuracy of the simulated land-use map, which is calculated by generating a transition matrix between the simulated land-use map and the observed land-use map at the same time point. The modeling procedure is repeated until the simulated land-use map for 2025 is generated.

Figure 4.4

Overview of Urban Land-Use Change Prediction Process



Note. In-sample prediction started from 2016 to 2020. The out-of-sample prediction started in the year 2021 and repeated until the year 2025 was reached.

4.5.1.2 Forecasting Demand for Urban Land

As economic growth and urbanization are closely related (Gollin et al., 2016). This phenomenon is econometrically classified as a simultaneity problem. To estimate and predict future urban land demand and economic growth in the light of simultaneity, the simultaneous equation estimation was applied. The simultaneous equations are described by equation (4.7) and (4.8). Since official GDP at district level in Thailand is not available, NTL was used as a proxy of Ban Chang's GDP.

$$ntl = \theta_0 + \beta_1 ul + c_1 z_1 + u_1 \tag{4.7}$$

$$ul = \theta_1 + \beta_2 ntl + c_2 z_2 + u_2 \tag{4.8}$$

where ntl = Ban Chang's NTL, ul = Ban Chang's urban land, z_1 = GDP of top five-trading partners of Thailand, z_2 = Construction price index. The GDP of top five-trading partners of Thailand represents the purchasing power of trading countries of Thailand.

Since the study area locates in export-oriented economy, z_1 should positively and exogenously affect Ban Chang's NTL. The construction price index (z_2) represents cost of urban growth. Therefore, it should negatively and exogenously affect urbanization. It is worth noting that the environmental factors such as NDDI, NDVI NDWI, and precipitation could be included in the model as exogenous variables affecting urbanization. However, the analysis from Chapter 3 suggests that environmental factors at best indirectly affect urbanization. Hence the environmental factors are not included in the simultaneous equation.

Equation (4.7) and (4.8) show that ntl and ul are determined simultaneously and z_1 and z_2 are assumed to be exogenous variables. By solving equation (4.7) and (4.8), Ban Chang's NTL and urban land are determined as follows.

$$\begin{bmatrix} ntl \\ ul \end{bmatrix} = \begin{bmatrix} \dfrac{\theta_0 + \beta_1\theta_1 + c_1z_1 + \beta_1c_2z_2 + \beta_1u_2 + u_1}{1 - \beta_2\beta_1} \\ \dfrac{\theta_1 + \beta_2\theta_0 + c_2z_2 + \beta_2c_1z_1 + \beta_2u_1 + u_2}{1 - \beta_2\beta_1} \end{bmatrix} \tag{4.9}$$

To solve the equation (4.9), the set of parameters $\theta = \{\beta_1, \beta_2, c_1, c_1\}$ is needed to be estimated. From equation (4.7) and (4.8), z_1 and z_2 are correlated with the endogenous regressors ntl and ul and z_1 and z_2 are assumed to be exogenous variables. So z_1 and z_2 could be used as instrumental variables for gpp and ul. Once we know θ, the $\{u_1, u_2\}$ can be calculated. The Two-Stage Least-Squares (2SLS) was applied to estimate the set of parameters θ. The equations to be estimated via 2SLS are described as follows:

First-stage regression

$$\widehat{ntl} = \pi_{11}z_2 + \pi_{12}z_1 \tag{4.10}$$

$$\widehat{ul} = \pi_{21}z_1 + \pi_2z_2 \tag{4.11}$$

Second-stage regression

$$ntl = \beta_1\widehat{ul} + c_1z_1 + u_1 \tag{4.12}$$

$$ul = \beta_2\widehat{ntl} + c_2z_2 + u_2 \tag{4.13}$$

The observed urban land and NTL of Ban Chang district from 2016–2020 were used to make an in-sample forecast. However, since land use data are published annually, the number of observations would be insufficient for the regression estimation. To overcome this limitation, the annual data were transformed into monthly data through linear interpolation.

To make a projection, the exogenous variables must be populated over the entire forecast horizon, i.e., the years 2021–2025, before solving the model. According to projection from International Monetary Fund, average GDP growth rate of the top-five trading partners of Thailand (z_1) from 2022–2025 is 3.2 percent (International Monetary Fund, 2023). The construction price index (z_2) is assumed to grow at 1.1 percent annually which is the compound annual growth rate of construction price index from 2004–2020.

4.5.1.3 Determining Supply for Urban Land

To address the causality effect of the Machine Learning technique, the Logistic Regression was selected as the predictive model for this reason. For three reasons, Logistic Regression is one of the most widely used Machine Learning algorithms in land-use change modeling. First, the Logistic Regression performance is as good as other Machine Learning algorithms, such as Random Forest, when number of explanatory variables is small (Kirasich, Smith, & Sadler, 2018). Second, Logistic Regression is simple when compared to other Machine Learning techniques. Third, in addition to predicting outcomes, Logistic Regression helps researchers understand the causality running from explanatory variables to the outcome, which is a feature that satisfies the research objective of this paper. The probability that dependent variables turn into built-up or non-built-up areas is calculated by equation 4.14.

$$P = \frac{e^{(\beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3)}}{1 + e^{(\beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3)}} \tag{4.14}$$

where P is the probability of urban gain in the cell, $X_{1,i}$, $X_{2,i}$, and $X_{3,i}$ are the explanatory variables – slope, elevation, and distance to urban settlement at grid cell i, respectively.

From equation 4.14. the odd ratio can be defined as follows:

$$odd\ ratio = \frac{P}{1-P} \tag{4.15}$$

The logistic transformation of equation 4.15 can be defined as follows:

$$P' = \log(odd\ ratio) = \log\left(\frac{P}{1-P}\right) \tag{4.16}$$

The probability that dependent variables turn into built-up or non-built-up area is calculated by equation (4.17).

$$P_i' = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \beta_3 X_{3,i} \tag{4.17}$$

where dependent variable probability that non-built-up area will turn into built-up area at grid cell i and $X_{1,i}$, $X_{2,i}$, and $X_{3,i}$ are the corresponding driving factor of urbanization – slope, elevation, and distance to urban existing urban area of grid cell i, respectively.

4.5.1.4 CLUE-S Allocation Rule

According to Verburg et al. (2002), the CLUE-S allocation rule is divided into five steps. The first step is determining the number of changeable grid cells and the urban land demand. Grid cells in protected areas (such as national parks) are excluded from further analysis. All grid cells in this study can change, and the amount of urban land demand is determined from simultaneous regression.

Second, Logistic Regression calculates the conversion probability of all grid cells for each land-use type (for example, the probability that bare land converts to the built-up area). Third, the CLUE-S allocates each cell according to the highest conversion probability (or suitability) of a particular land-use type as determined by the Logistic Regression. Fourth, after the first allocation from the third step, the CLUE-S compares the primary urban land conversion amount with the predicted urban land demand. If the amount of primary urban land conversion is less or more than the predicted urban land demand, the conversion probability will increase or decrease respectively. Fifth, steps two to four are repeated to fulfill the urban land demand from the first step.

4.5.1.5 Accuracy Assessment

Table 4.2

Confusion Matrix

	Simulated			
Observed		Built-up area	Other area	Total
	Built-up area	True Positive (TP)	False Positive (FP)	TP+FP
	Other area	False Negative (FN)	True Negative (TN)	FN+TN
	Total	TP+FN	FP+TN	TP+FP+FN+TN

Note. This table shows the confusion matrix that conventionally used to access model's accuracy.

The accuracy, precision, and recall were calculated from the confusion matrix. The accuracy value ranges from zero to one. A value close to one indicates the highest accuracy, while a value close to zero indicates the lowest accuracy. Precision indicates the correctness of identifications of the predicted model, while recall indicates the correctness of actual positives. The following equations conventionally calculate accuracy, precision, and recall.

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \tag{4.18}$$

$$Precision = \frac{(TP)}{(TP + FP)} \tag{4.19}$$

$$Recall = \frac{(TP)}{(TP + FN)} \tag{4.20}$$

4.5.2 Data

4.5.2.1 Geospatial Data

In this study, a web-based application was constructed on Google Earth Engine to visualize and acquire the geospatial data[6]. A land-use map of Ban Chang district was obtained from the Dynamic World V1 dataset provided by World Resources Institute Google. The Dynamic World is a 10m near-real-time Land Use/Land Cover (LULC) dataset that includes nine land-use classes, namely water, trees, grass, flooded vegetation, crops, shrub and scrub, built-up area, bare land, and snow. For simplicity, land-use classes were reclassified into built-up and non-built-up areas. The dataset in Dynamic World V1 is processed from the Sentinel-2 L1C collection from 2016–2022. To show the validity of the land-use maps used in the land-use model. The land-used maps obtained from Google Earth Engine were compared with the land-used map officially provided by Department of Land Development from 2016-2021. The results show that similarity indices range from 78.3–84.4 percent (see Appendix C for more detail).

Figure 4.5 illustrates the constructed application's Graphical User Interface (GUI) that allows user to select the land-use map of preferred year. As shown in Figure 4.6, the percentage of built-up area is shown on the left panel. Land-use maps of Ban Chang district for the year 2016 and 2020 are depicted in Figure 4.7. In addition to the land-use map, slope, elevation and distance to existing urban areas were used as explanatory variables for the predictive model and were obtained from the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) dataset. Like land-use maps, slope, and elevation maps were downloaded from Google Earth Engine. Agriculture and Agri-Food Canada (AAFC) provide classification of slope gradient (in percent)[7].

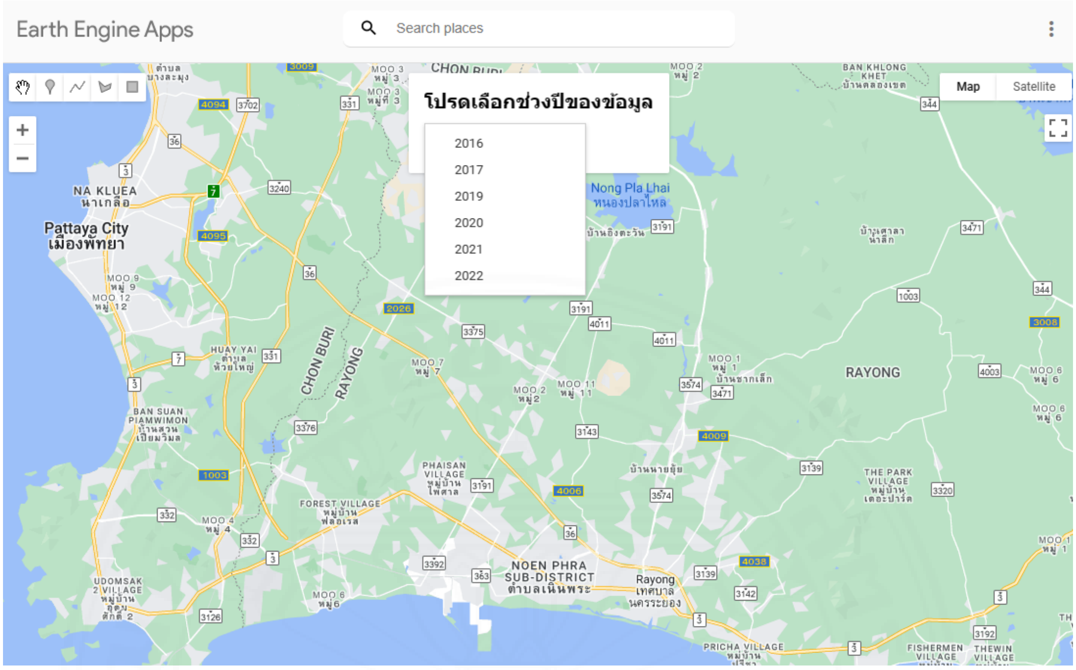
---

[6] This application is publicly accessible at https://nutchapon.users.earthengine.app/view/ban-chang-land-use-map

[7] The source of data is publicly accessible at https://sis.agr.gc.ca/cansis/nsdb/slc/v3.2/cmp/slope.html

Figure 4.5

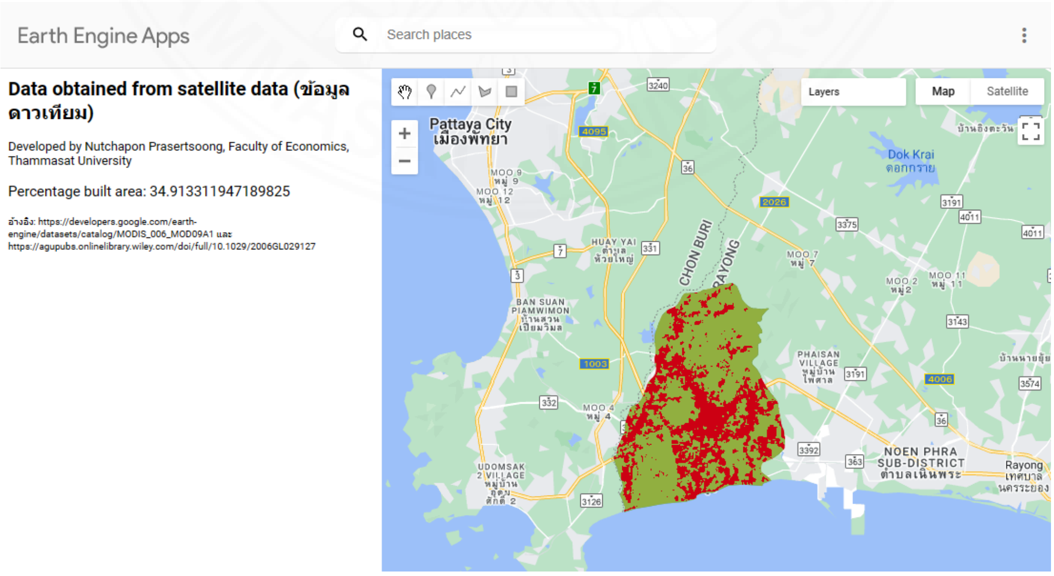
GUI of Constructed Application



Note. The application is publicly available at https://nutchapon.users.earthengine.app/view/ban-chang-land-use-map

Figure 4.6

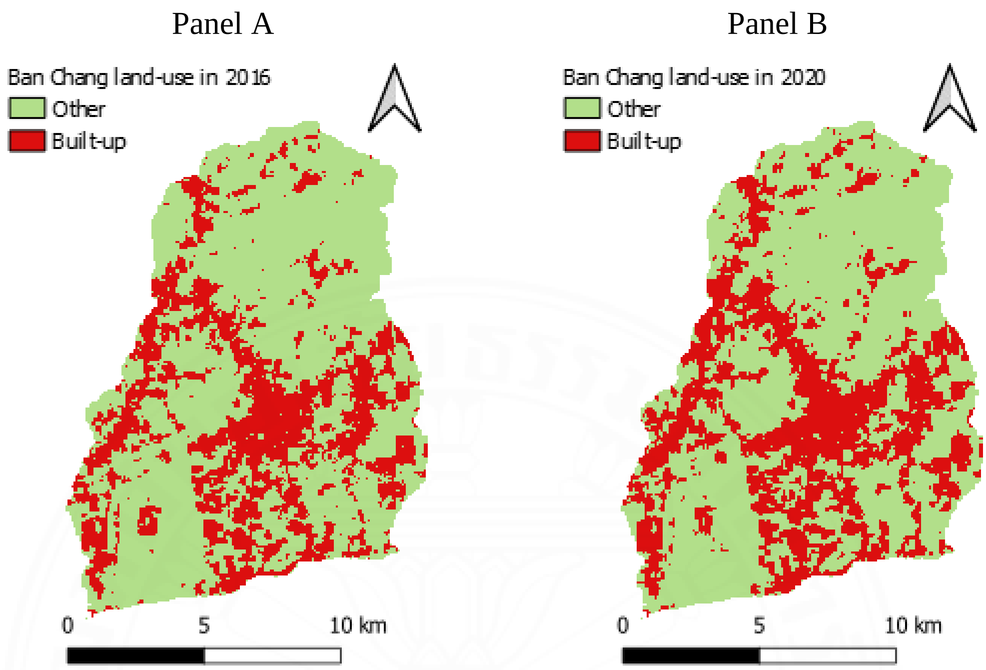
Statistics of Urban Land Area and Land Use Map of Ban Chang District Generated by the Constructed Application



Note. The application is publicly available at https://nutchapon.users.earthengine.app/view/ban-chang-land-use-map

Figure 4.7

Observed Land Use Maps of Ban Chang District in 2016 and 2020



Note. Panel A: Ban Chang's land-use in 2016; Panel B: Ban Chang's land-use in 2020.

Figure 4.8

Slope, Elevation, and Distance to Existing Urban Area of Ban Chang District

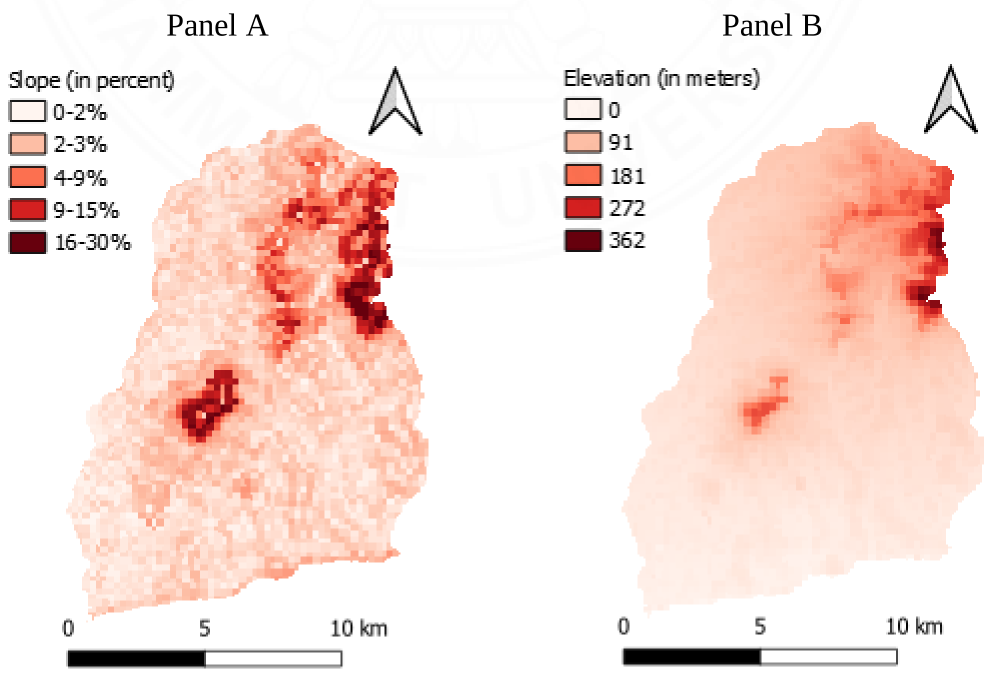
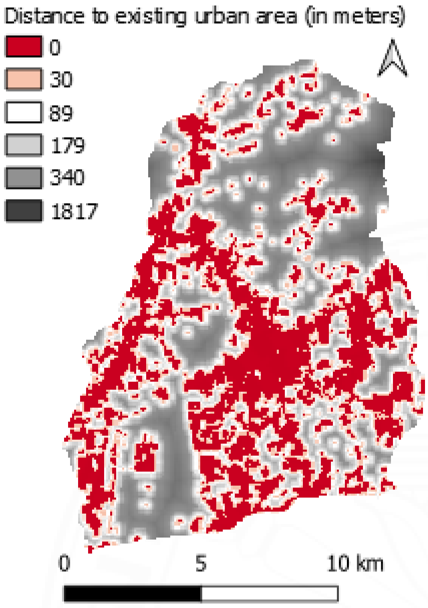


Figure 4.8

Slope, Elevation, and Distance to Existing Urban Area of Ban Chang District (Cont.)

Panel C: Distance to existing urban area in 2016



Note. Panel A: Ban Chang's slope map; Panel B: Ban Chang's elevation map; Panel C: Ban Chang's distance to existing urban area.

A higher slope increases construction costs and reduces the feasibility of the project. A slope gradient less than or equal to 3 percent is suitable for construction. A slope gradient between 4–9 percent is suitable for most building types except sports stadiums. A slope gradient greater than 9 percent significantly reduces the possibility of construction. A developer might need to apply special construction techniques in areas with a slope greater than 15 percent, significantly raising the construction cost. The example of slope, elevation, and map of and distance to existing urban area is illustrated in Figure 4.8.

4.5.2.2 Economic Data

The economic data in this study includes the following: First, the GDP of Rayong province; Second, the GDP of Thailand's top-five trading partners; and third, the construction price index. The Rayong GDP from 2000–2020 is obtained from the NESDC[8]. According to the official international trade statistics (Ministry of

[8] The data is publicly available at https://www.nesdc.go.th/main.php?filename=gross_regional

Commerce), the GDP of the top five-trading partners of Thailand includes the GDP of the United States, China, Japan, Hong Kong, and Vietnam[9]. The GDP of each country is obtained from World Bank dataset[10]. Lastly, the construction price index is publicly published by the Ministry of Commerce from the year 2000 to the present[11].

4.6 Result

4.6.1 Estimated Results of Simultaneous Equations

Using data from 2016–2020, Table 4.3 exhibits the estimated results of equation (4.7) and equation (4.8), respectively. It indicates that Ban Chang NTL affected urban land expansion and vice versa. The income of Thailand's top-five trading partners (the United States, China, Japan, Hong Kong, and Vietnam) positively affected the GDP of Rayong province and, thereby, Ban Chang district. The higher construction price (either from inflation or supply shock) negatively affected urban land expansion in Ban Chang district.

Table 4.3

Estimated Results of Equation (4.7) And Equation (4.8) Using 2SLS

Dependent: Ban Chang NTL	Estimated coefficient	Dependent variable: Urban land	Estimated coefficient
Urban land (β_1)	0.16***	Ban Chang NTL (β_2)	3.25***
	(14.54)		(47.96)
GDP of top-five trading partners of Thailand (c_1)	$5.31*10^{-13}$***	Construction price index (c_2)	-1.13***
	(18.84)		(-19.14)
Intercept (θ_0)	-21.83***	Intercept (θ_1)	161.83***
	(-30.88)		(26.91)

[9] The data is publicly available at https://tradereport.moc.go.th/TradeThai.aspx
[10] The data is publicly available at https://data.worldbank.org/indicator/NY.GDP.MKTP.CD
[11] The data is publicly available at http://www.indexpr.moc.go.th/price_present/csi/stat/other/conyear.asp

Table 4.3

Estimated Results of Equation (4.7) And Equation (4.8) Using 2SLS (Cont.)

Observation	49	Observation	49
R-squared	0.98	R-squared	0.98

Note. z statistics are reported in parentheses

*p<0.10, **p<0.05, ***p<0.01

4.6.2 Estimated Results of Logistic Regression

Table 4.4 exhibits the estimated results of equation (4.13) via Logistic Regression; the coefficient values of the slope and elevation are negative but insignificant. The coefficient of distance to urban is negative and highly significant, suggesting that a higher value of distance to existing urban amenities such as roads, electricity, water supply, and agglomeration negatively affected urban land expansion and reduced urban land suitability.

Table 4.4

Estimated Results of Equation (4.13) Using Logistic Regression

Dependent: Rayong GDP	Estimated coefficient
Slope (X_1)	-0.0005
	(-0.85)
Elevation (X_2)	-0.005
	(-0.32)
Distance to existing urban area (X_3)	-8,262.25***
	(-45.09)
Intercept	3.08***
	(31.53)
Observation	93,646
AIC	26,918.53

Note. z statistics are reported in parentheses

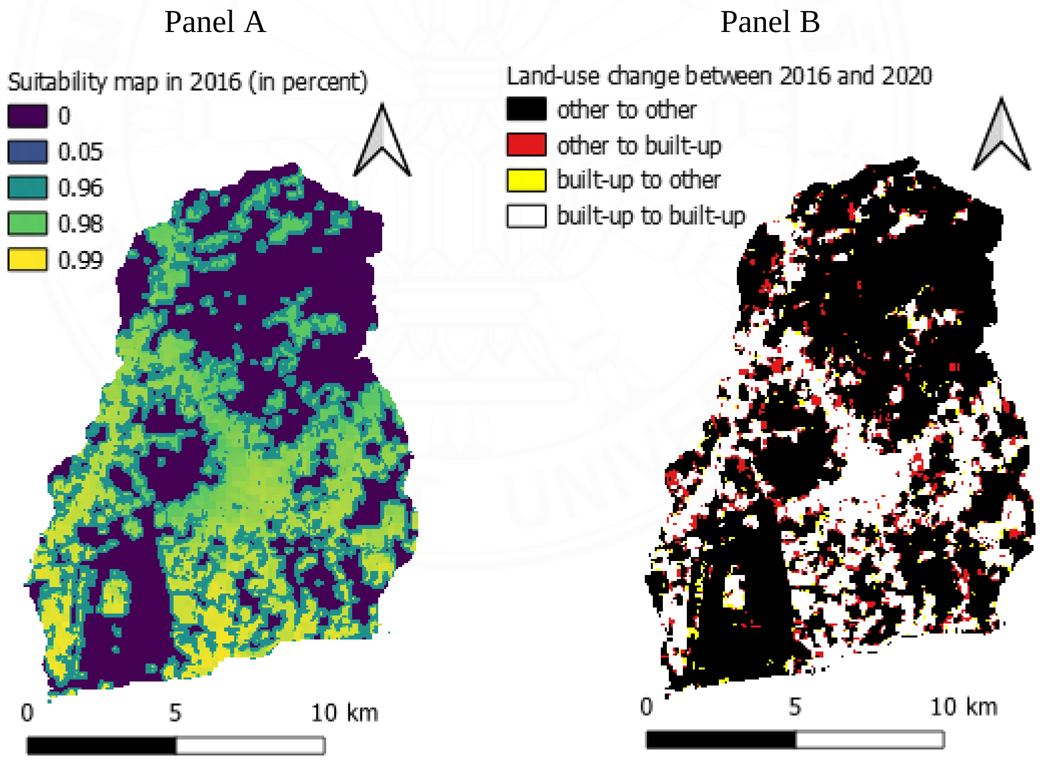
*p<0.10, **p<0.05, ***p<0.01

4.6.3 Urban Growth of Ban Chang Between 2016 and 2020

The built-up area of Ban Chang district increased from 71.8 square kilometers in 2016 to 80.8 square kilometers in 2020, or about 12.6 percent in five years. Six percent of Ban Chang's non-built area was transformed into a built-up area between 2016 and 2020. Panel A of Figure 9 shows the suitability map for 2016. It indicates that Ban Chang's built-up areas were likely to be expanded within the existing urban areas and around their edges. Panel B of Figure 4.9 spatially identifies the actual transition of land-use classes in Ban Chang district between 2020 and 2016, illustrating that the conversion of non-built land to built-up land in Ban Chang district occurred near the existing urban area.

Figure 4.9

Land-Use Change Pattern of Ban Chang District Between 2016 and 2020



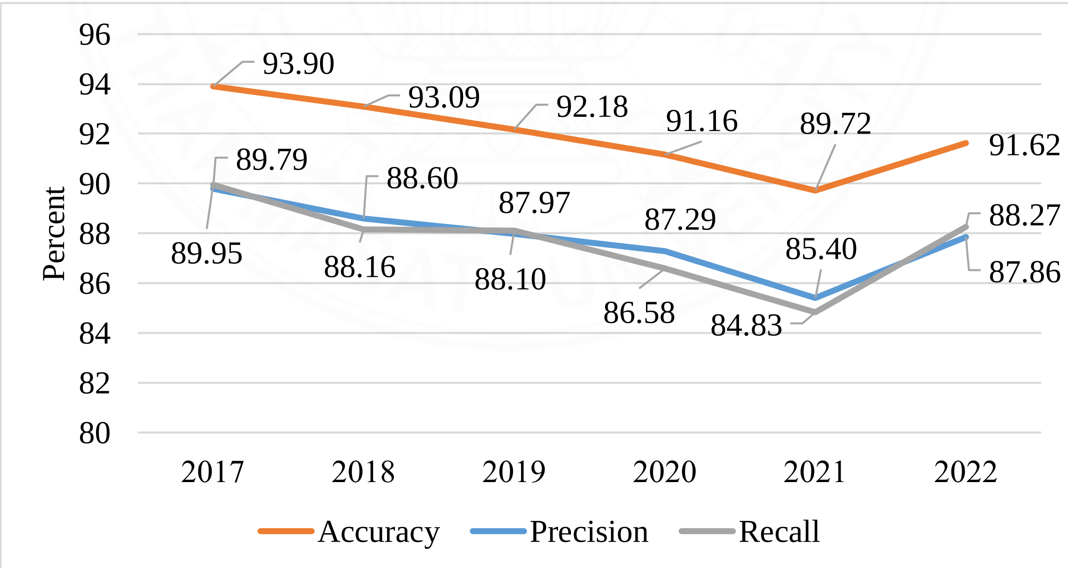
Note. Panel A: Suitability map in 2016; Panel B: Land-use change between 2016 and 2020

4.6.4 Model Validation

Figure 4.10 shows the accuracy of the simulated land-use map from 2017–2022. The land-use map for 2016 was removed from the accuracy assessment since the year 2016 is treated as the initial period in the projection. The accuracy of the prediction stood above 90 percent from 2017–2020 but slightly decreased to 89.7 percent in 2021 before increasing to 91.6 percent in 2022. The average spatial accuracy of the model between 2017–2021 was 91.9 percent. In other words, the model correctly predicted about 91.9 percent of land-use transitions, and 7.9 percent of land-use transitions were misclassified. The precision and recall rates move along with the accuracy rate. But the precision and recall rates ranged from 89 to 84 percent from 2017–2020. Based on the precision rates observed between 2017 and 2022, the model demonstrated an average accuracy of 87.8 percent in correctly forecasting urban expansion during that period. Likewise, the recall rate suggests that the model correctly identified, on average, 87.7 percent of urban land expansion from 2017–2022.

Figure 4.10

Spatial Accuracy From 2017–2022



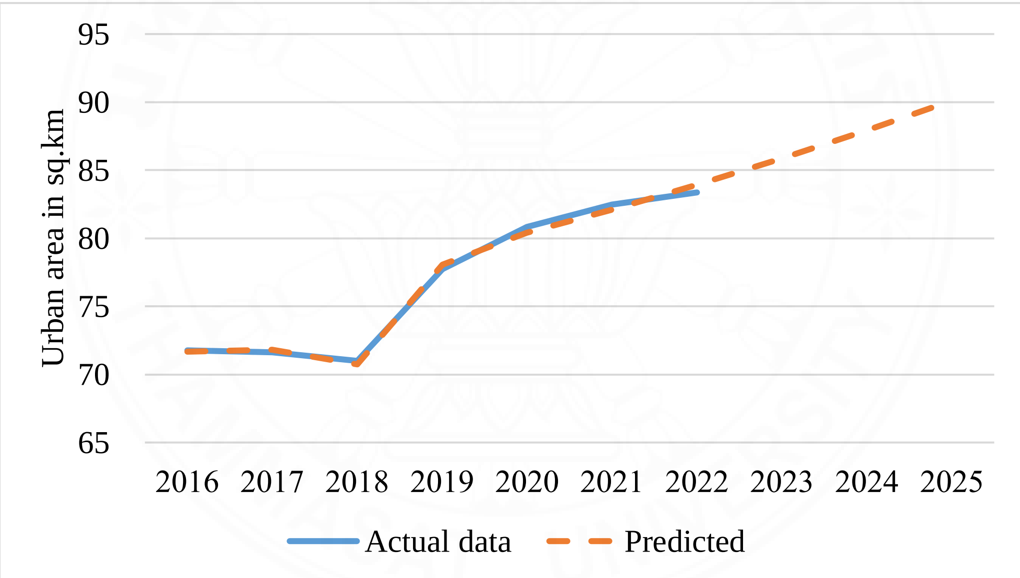
Note. This figure was created from author's calculations.

4.6.5 Projected Urban Land Expansion

Figure 4.11 depicts urban land demand estimated using the 2SLS technique. The solid blue line represents the actual urban land use from 2016–2022. The orange dashed line represents the prediction of land-use demand from 2016–2025. It clearly shows that both prediction data moved in line with the actual urban land-use from 2016–2022, reflecting the high accuracy of the predictive model. The result suggests Ban Chang's urban land would increase from 84 square kilometers in 2022 to 90 square kilometers in 2025, or by seven percent in the next three years. According to the assumption, 37.8 percent of Ban Chang is expected to be covered by the built-up area in 2025.

Figure 4.11

Urban Land Demand of Ban Chang District



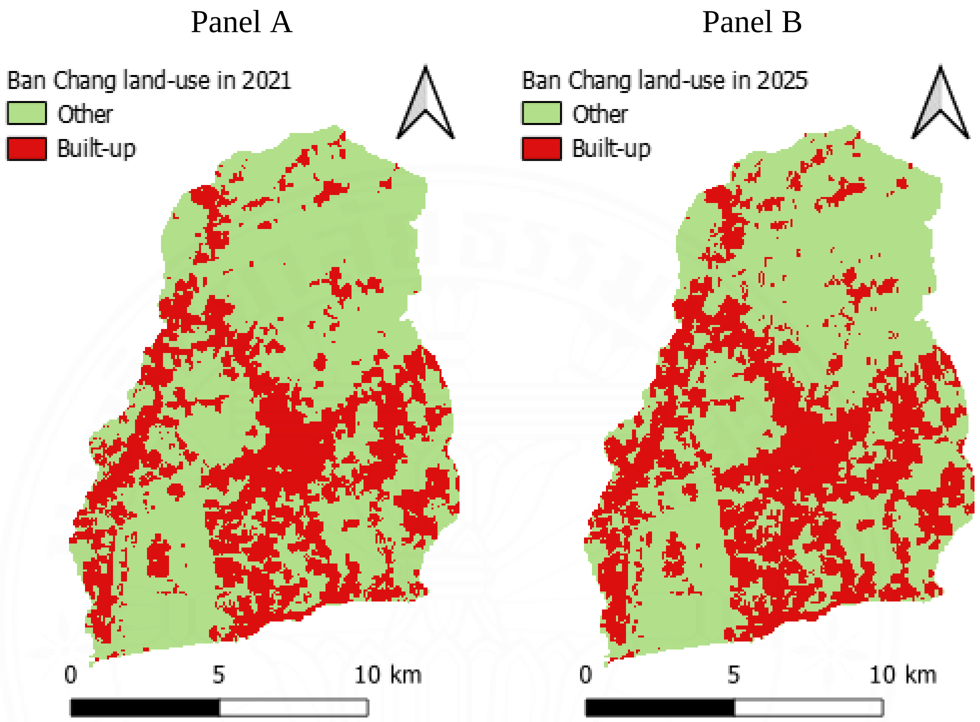
Note. This figure was created from author's calculations.

Figure 4.12 shows the projected land use map from 2021– 2025. Given the estimated future urban land demand, both pictures illustrate that urban land expansion in Ban Chang is expected to take place around the existing urban area located at the center and westward of Ban Chang district. At the same time, the north of Ban Chang would receive fewer urban settlements in the future. The urban land expansion will likely occur in the southern part of Ban Chang, where beaches are located. Notably, a slight expansion of built-up areas is expected within the U-Tapao International

Airport, reflecting recent government efforts to enhance the airport's passenger carrying capacity.

Figure 4.12

Predicted Land-Use of Ban Chang, 2021 and 2025



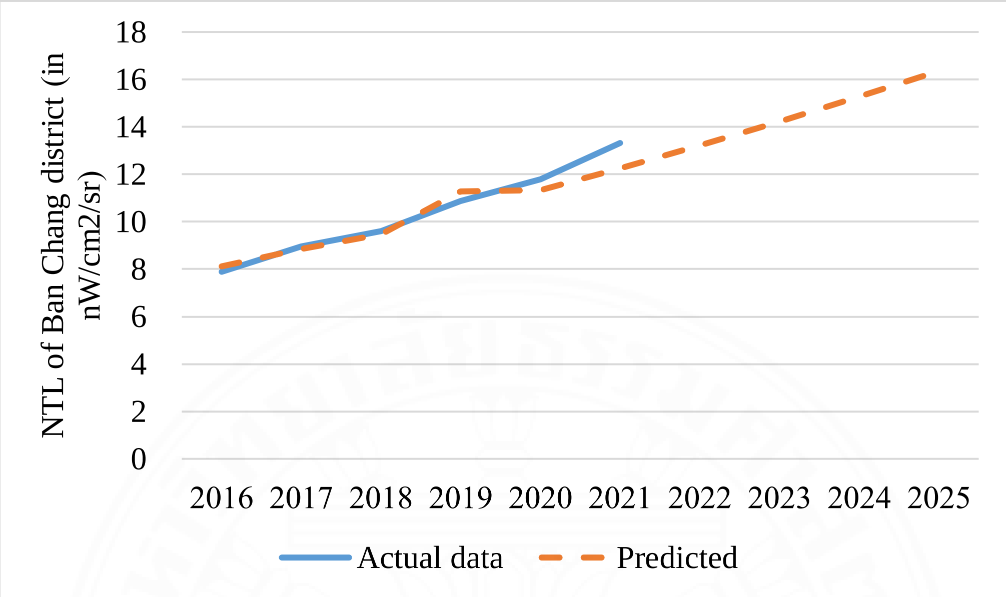
Note. Panel A: Predicted land-use of Ban Chang in 2021; Panel B: Predicted land-use of Ban Chang in 2025.

4.6.6 Projected Economic Growth of Ban Chang

Figure 4.13 displays the projected economic expansion of Ban Chang district from 2021–2025. Following the mathematical foundation elaborated in section 4.5.1.2, the projected economic growth was mutually induced by the growth of urban land in Ban Chang district and influenced by two exogenous variables: the GDP of Thailand's trading partners and the construction price index. According to the assumption in section 4.5.1.2, the model predicted that the economy of Ban Chang district would experience an average annual growth rate of 5.9 percent from 2021 to 2025. The predicted economic growth of Ban Chang district was compared with the mobility index and the number of poor people who have difficulty reaching employment and minimum level of household income. The result is shown in Figures E1 and E2 of Appendix E, respectively.

Figure 4.13

Predicted NTL of Ban Chang District From 2016–2025



Note. This figure was created from author's calculations.

4.6.7 Policy Recommendations

The results suggest that the municipality authority should focus on developing basic urban infrastructure such as roads, sewerage systems, waste management, and other related urban infrastructure in the western, central, and coastal areas of Ban Chang to accommodate the growing urban land area in those three areas. The urban planning policy should emphasize urban sustainability. For example, the new urban planning should encourage the usage of public transportation by integrating urban transportation with the city road network. Additionally, in contrast to heavily relying on major roads, the new city road network should be well-connected.

Local public transportation should also be promoted to reduce reliance on personal vehicles. Alternatively, the walkability and rideability of the city should also be endorsed, and the footpath and bicycle lane could be enlarged and expanded to encourage walking and cycling. Such a policy could reduce the possibility of traffic congestion and improve urban citizens' health and safety (Coughenour, de la Fuente-Mella, & Paz, 2019). Without proper planning to meet the growing demand that would naturally occur in certain areas with high suitability, urban land expansion could result in urban sprawl and congestion, hampering the city's sustainable development.

4.7. Conclusion

The recursive dynamic model was applied to predict local economic growth and the pattern of urban land expansion at the district level in Thailand. Open data from Google Earth Engine and open-source software were utilized. The importance of anticipation of urban land demand in urban administration and sustainable urbanization was also emphasized. According to the projection, the government and municipal authorities should focus infrastructure investments on the central, western, and southern parts of Ban Chang district, as urban land expansion is likely to occur in these three areas. Failing to provide sufficient infrastructure in an area with a high conversion probability could make urbanization unsustainable. Moreover, Ban Chang's GDP's predicted compound annual growth rate from 2022–2025 is 4.4 percent.

The major contributions of this study are twofold. First, with its foundation in open data and open-source software, other researchers could freely extend or improve the developed model. The near-real-time nature of geospatial data enables researchers or policymakers to monitor the expansion of urban areas and economic growth at any district or sub-district level in Thailand. Second, this study could answer the following three dimensions of urban and economic growth: when growth would occur, where growth would take place, and how much growth would be in the future. The methodology and the data applied in this study potentially enhance government agencies’ capacity to timely formulate both economic development and sustainable urbanization plans at the city level. Regarding urban planning policy, the government should actively anticipate the growing urban land demand in the city and formulate an urban development plan promoting sustainable urbanization. Public transportation, a well-designed road network, and people-centric urban planning should be promoted to make cities sustainable.

CHAPTER 5
CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusion

This study quantitatively examined the relationship between urbanization, spatial inequality, and economic development in Thailand. Conventional socioeconomic and geospatial data obtained from government agencies and the Google Earth Engine platform were utilized to achieve the objective. The causal relationship between urbanization and productivity differential at the provincial level was investigated in Chapter 2. Based on the finding from Chapter 2, four regression models were applied to explore the short-term effects of urban sector productivity, environmental factors, and human capital on urbanization at the provincial level in Chapter 3. Chapter 4 predicted the pattern of urban land expansion and economic growth at the district level by considering the simultaneous relationship between urbanization and economic development in Ban Chang district (Rayong province).

The unit of analysis differs across chapters in this study. Specifically, in Chapters 2 and 3, the analysis focuses on each province in Thailand, with geospatial data aggregated and examined at the provincial level. Chapter 4 of the study introduced a shift in the unit of analysis to the pixel and district levels. The analysis involved computations and econometric analysis based on individual pixels within districts. The primary objective of Chapter 4 is to predict the direction of urban growth at the pixel level and analyze the economic growth, specifically in the Ban Chang district. By employing this approach, the researchers could gain insights into the specific patterns of urban expansion and economic development at a more detailed spatial resolution.

The analysis from Chapter 2 shows that agglomeration externalities from urbanization positively affected the productivity of workers. From the national perspective, the monocentric urbanization dominated by BMR is Thailand's leading cause of spatial inequality.

The results from Chapter 3 reveals that urban sector productivity and the share of workers with higher education significantly affected urbanization in all regression specification. Higher rainfall and water availability were found to affect urbanization positively. It suggests that favorable natural endowment still influences the rural-urban migration of workers. Urbanization in one province was found to affect urbanization in the neighboring provinces. Moreover, the level of urbanization during the initial period strongly impacted the level of urbanization in the subsequent periods. Lastly, the simulation results from Chapter 4 suggest that urban land expansion is expected to occur in the central, western, and southern parts of the Ban Chang district in the next three years. In addition to urban growth, the projected economic growth rate of Ban Chang district from 2022–2025 is 4.4 percent. The findings of Chapter 4 highlight the importance of urban land growth anticipation in urban administration that could prevent problems arising from unorganized and unrestricted urban growth.

Most of the findings in all three chapters are consistent with the findings of previous studies. In Chapter 2, the magnitude of agglomeration externalities and localization economies is comparable to previous literature that conducted research in Thailand (Limpanonda, 2012) and in developed countries such as the French and Netherlands (Combes et al., 2008; Groot et al., 2014). However, the effect of specialization in this study is positive but insignificant, which contradicts the studies above but is consistent with studies conducted in Brazil (Barufi et al., 2016). The findings in Chapter 3 are consistent with existing literature. The findings in Chapter 4 are unique to the local context. However, the result from Logistic Regression is consistent with existing studies. Specifically, the higher distance value to existing urban settlements negatively affected conversion possibilities to urban land. The necessity of urban administration to serve the growing urban land demand was also highlighted by other literature that studies urban land expansion in Eastern Economic Corridor (Tontisirin & Anantsuksomsri, 2021).

5.2 Policy Implication and Recommendations

The monocentric and unorganized urbanization in Thailand is the main cause of spatial inequality and urban sprawl, worsening the general well-being of the Thai people. At the national level, Chapter 2 suggests that, in contrast to the monocentric growth, the Thai economy should alternatively be driven by polycentric urban development in all regions. Such policy potentially reduces persistent spatial inequality and promotes economic growth in Thailand.

According to the findings in Chapter 3, policymakers should create highly productive jobs that require highly educated workers in regional cities to establish polycentric development. Such a policy could lift urban sector productivity and promote urbanization in regional cities. The environmental factors should not be neglected since a higher level of drought is found to affect urbanization negatively. One possible explanation is that workers have been leaving drought regions over the past two decades and migrating into the central growth pole – Bangkok and Metropolitan Region. Based on the aforementioned logic, investing in water infrastructure could indirectly affect urbanization by preventing workers in the agricultural sector from leaving the local areas, an investment that could enlarge the local market and increase the urbanization rate. Therefore, the success of polycentric urban development depends on the government's capability to secure a constant water supply in the drought area.

Chapter 4 emphasizes the challenge of future urban administration in Thailand. Specifically, instead of passively investing in urban infrastructure to catch up with the growing demand for urban land, a new urban administration should actively anticipate the growing demand for urban land and invest in urban infrastructure in the area with the highest possibility of urban development. Active urban planning could mitigate the problems observed in the mega city in Thailand and make cities more habitable, simultaneously unlocking the full economic benefit of urbanization and reducing its negative impact on the environment and people.

In summary, the policy recommendations of this study lean toward the so-called “place-based” policy, an approach that seeks to reduce the difference in economic outcomes between regions. This study argues that polycentric development in the Thai context involves no trade-off between efficiency and equity since it potentially reduces territorial inequality and promotes economic growth at the national level. However, as Rodríguez-Pose (2018) emphasized, such policy should not result in the creation of assisted economies where the left-behind people significantly depend on government transfer. On the contrary, the new policy should seek to maximize the economic potential of all regions by considering the local context.

As the study found that polycentric development could be achieved by lifting productivity in all regions, the left-behind place should be reinvigorated by establishing policies that create new entrepreneurship, education, and innovation. Regional development policies from developing and developed countries that seek to reduce spatial inequality through new economies, education, and increasing productivity could be implemented in Thailand. For example, a carbon trading policy that generates capital flows from primate cities to the hinterland could reduce inequalities between rich and poor regions in China (Zhang, Chen, Wang, Wen, & Chen, 2023). The Chinese government's massive investment in high-speed rail was found to reduce regional economic disparity and promote local tourism. These findings are consistent with the prediction from New Economic Geography that regional convergence is possible when trade cost is sufficiently low such that firms forgo the linkage they have in the primate cities and move to areas that offer lower operating costs (Chen & Haynes, 2017; Gao, Su, & Wang, 2019). According to a study in Argentina, a policy that promotes the reduction of informality and expansion of basic education could be endorsed to reduce disparities between regions (Quiroga-Martínez & Fernández-Vázquez, 2021). In developed counties, the European Regional Development Fund positively affected local total factor productivity in a region with higher institutional quality and population density (Albanese, de Blasio, & Locatelli, 2021). In fiscal administration, a fiscal structure that incentivizes utilizing local resources for economic development should be established (Bartolini, Stossberg, & Blöchliger, 2016). The example of urban and regional development policies from both developing and developed countries could be synthesized to enhance the practicality of the urbanization policy in the Thai context.

5.3 Limitations and Suggestions for Further Studies

The study in Chapter 2 has several limitations that also serve as key areas for future research. These limitations can be categorized into four main aspects.

Firstly, there is a need to refine the spatial resolution of the study to a smaller scale, such as district or sub-district level. This would provide more detailed analytical insights into the relationships under investigation. Secondly, the frequency of data collection should be increased. Specifically, productivity measures can be derived from firms' annual or quarterly financial statements. This would provide a more accurate and comprehensive understanding of productivity dynamics.

Thirdly, the study is limited by its static analysis approach. Future research should consider employing high-frequency and continuous data sets. This would enable the development of a dynamic analytical framework, allowing for the examination of inter-temporal relationships between density and agglomeration. Lastly, the study should consider including a broader range of sectors, such as information technology. By addressing these limitations, future research can further enhance the understanding of spatial and economic dynamics, contributing to advancing knowledge in the field.

The limitations of Chapter 3 (and suggestions for future study) are two-fold. First, future studies should obtain more information on historical urban land data. For example, the Google Earth Engine platform might be utilized to extract historical urban land cover in Thailand from Landsat satellite imageries which have been available since 1987. Second, instead of analyzing at the provincial level, the spatial unit in the regression models might be reduced to a district level, providing more analytical detail. However, since official information on income at the district level in Thailand is not available, the proposed future research would be methodologically challenging and original in the Thai context.

The study's limitations in Chapter 4 and key issues for future research are three-fold. First, the historical data on land-use change needs to be improved. Future research should obtain more information on historical land-use change while maintaining the same data frequency and spatial resolution level.

Second, the current model's configuration could only be applied to an export-oriented economy. Instead of relying on the income of the international trade partners as exogenous variables, future research should generalize the model by considering the influence of regional purchasing power, such as the GDP of the neighboring provinces of the selected study area (the purchasing power of surrounding area is like the concept of "Market Access" officially asserted by Fujita et al. (2001)).

Market Access helps future research mitigate the technical issue arising from the edge around the selected area since the economic activities around the edge of the selected study area would be considered. Alternatively, the Vector Error Correction Model (VECM) could be applied to predict both value and shape of the distribution of the interested variables – NTL and urban land. An example of results from an application of VECM is provided in Appendix D.

Furthermore, it is recommended that future research maintains consistency in the unit of analysis across three chapters, specifically by analyzing at the district or pixel level throughout the study. Lastly, open-source, user-friendly software might be developed and integrated with the recursive dynamic model for a seamless computational process.

REFERENCES

Books and Book Articles

Alonso, W. (1964). *Location and land use: toward a general theory of land rent*: Harvard university press.

Baker, C., & Phongpaichit, P. (2014). *A History of Thailand* (3 ed.). Cambridge: Cambridge University Press.

Christaller, W. (1933). *Central places in Southern Germany*. Jena, Germany: Fischer (English translation by C. W. Baskin, London: Prentice Hall, 1966).

Forrester, J. W. (1969). *Urban Dynamic*: Waltham, Massachusetts: Pegasus Communications, Inc. (January 1, 1969).

Fujita, M., Krugman, P. R., & Venables, A. (2001). *The spatial economy: Cities, regions, and international trade*. Cambridge, Massachusetts: MIT press.

Glaeser, E. (2013). Triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier. *New York, NY: Penguin Books. doi, 10*, 1726-3247.

Henderson, J. V. (2005). Chapter 24 - Urbanization and Growth. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, pp. 1543-1591): Elsevier.

Hoover, E. M. (1948). *The Location of Economic Activity*. New York: McGraw-Hill.

Mincer, J. (1974). *Schooling, Earnings, and Experience*. New York: Columbia University Press.

Puttanapong, N. (2018). Monocentric Growth and Productivity Spillover in Thailand. In *Measuring Urbanization in ASEAN from Space*. Chiba, Japan: Institute of Developing Economies - Japan External Trade Organization (IDE-JETRO).

Tonguthai, P. (2019). Implicit policies affecting urbanization in Thailand. In *Urbanization and Urban Policies in Pacific Asia* (pp. 183-192): Routledge.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge, Massachusetts: MIT press.

Articles

Ades, A. F., & Glaeser, E. L. (1995). Trade and Circuses: Explaining Urban Giants*. *The Quarterly Journal of Economics, 110*(1), 195-227. doi:10.2307/2118515

Albanese, G., de Blasio, G., & Locatelli, A. (2021). Does EU regional policy promote local TFP growth? Evidence from the Italian Mezzogiorno. *Papers in Regional Science, 100*(2), 327-348. doi:https://doi.org/10.1111/pirs.12574

Amare, M., Jensen, N. D., Shiferaw, B., & Cissé, J. D. (2018). Rainfall shocks and agricultural productivity: Implication for rural household consumption. *Agricultural Systems, 166*, 79-89. doi:https://doi.org/10.1016/j.agsy.2018.07.014

Arfanuzzaman, M., & Dahiya, B. (2019). Sustainable urbanization in Southeast Asia and beyond: Challenges of population growth, land use change, and environmental health. *Growth and Change, 50*(2), 725-744. doi:https://doi.org/10.1111/grow.12297

Au, C.-C., & Henderson, J. V. (2006). Are Chinese Cities Too Small? *The Review of Economic Studies, 73*(3), 549-576. doi:10.1111/j.1467-937X.2006.00387.x

Bai, X., Chen, J., & Shi, P. (2012). Landscape Urbanization and Economic Growth in China: Positive Feedbacks and Sustainability Dilemmas. *Environmental Science & Technology, 46*(1), 132-139. doi:10.1021/es202329f

Barredo, J. I., Kasanko, M., McCormick, N., & Lavalle, C. (2003). Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. *Landscape and Urban Planning, 64*(3), 145-160. doi:https://doi.org/10.1016/S0169-2046(02)00218-9

Barrett, C. B., & Bevis, L. E. M. (2015). The self-reinforcing feedback between low soil fertility and chronic poverty. *Nature Geoscience, 8*(12), 907-912. doi:10.1038/ngeo2591

Barufi, A. M. B., Haddad, E. A., & Nijkamp, P. (2016). Industrial scope of agglomeration economies in Brazil. *The Annals of Regional Science, 56*(3), 707-755. doi:10.1007/s00168-016-0768-3

Batty, M., & Xie, Y. (1994). From Cells to Cities. *Environment and Planning B: Planning and Design, 21*(7), S31-S48. doi:10.1068/b21S031

Belotti, F., Hughes, G., & Mortari, A. P. (2017). Spatial panel-data models using Stata. *Stata Journal, 17*(1), 139-180.

Bertinelli, L., & Black, D. (2004). Urbanization and growth. *Journal of Urban Economics, 56*(1), 80-96. doi:https://doi.org/10.1016/j.jue.2004.03.003

Bloom, D. E., Canning, D., & Fink, G. (2008). Urbanization and the Wealth of Nations. *Science, 319*(5864), 772-775. doi:10.1126/science.1153057

Braimoh, A. K., & Onishi, T. (2007). Spatial determinants of urban land use change in Lagos, Nigeria. *Land Use Policy, 24*(2), 502-515. doi:https://doi.org/10.1016/j.landusepol.2006.09.001

Brückner, M. (2012). Economic growth, size of the agricultural sector, and urbanization in Africa. *Journal of Urban Economics, 71*(1), 26-36. doi:https://doi.org/10.1016/j.jue.2011.08.004

Brülhart, M., & Mathys, N. A. (2008). Sectoral agglomeration economies in a panel of European regions. *Regional Science and Urban Economics, 38*(4), 348-362. doi:https://doi.org/10.1016/j.regsciurbeco.2008.03.003

Buchanan, J. M. (1965). An Economic Theory of Clubs. *Economica, 32*(125), 1-14. doi:10.2307/2552442

Burke, M., Hsiang, S. M., & Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature, 527*(7577), 235-239. doi:10.1038/nature15725

Chauvin, J. P., Glaeser, E., Ma, Y., & Tobio, K. (2017). What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics, 98*, 17-49. doi:https://doi.org/10.1016/j.jue.2016.05.003

Chen, G., Li, X., Liu, X., Chen, Y., Liang, X., Leng, J., . . . Huang, K. (2020). Global projections of future urban land expansion under shared socioeconomic pathways. *Nature Communications, 11*(1), 537. doi:10.1038/s41467-020-14386-x

Chen, L., Hasan, R., & Jiang, Y. (2022). Urban Agglomeration and Firm Innovation: Evidence from Asia. *The World Bank Economic Review, 36*(2), 533-558. doi:10.1093/wber/lhab022

Chen, Z., & Haynes, K. E. (2017). Impact of high-speed rail on regional economic disparity in China. *Journal of Transport Geography, 65*, 80-91. doi:https://doi.org/10.1016/j.jtrangeo.2017.08.003

Ciccone, A. (2002). Agglomeration effects in Europe. *European Economic Review, 46*(2), 213-227. doi:https://doi.org/10.1016/S0014-2921(00)00099-4

Ciccone, A., & Hall, R. E. (1996). Productivity and the Density of Economic Activity. *The American Economic Review, 86*(1), 54-70. Retrieved from http://www.jstor.org/stable/2118255

Claessens, L., Schoorl, J. M., Verburg, P. H., Geraedts, L., & Veldkamp, A. (2009). Modelling interactions and feedback mechanisms between land use change and landscape processes. *Agriculture, Ecosystems & Environment, 129*(1), 157-170. doi:https://doi.org/10.1016/j.agee.2008.08.008

Clarke, K. C., & Gaydos, L. J. (1998). Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science, 12*(7), 699-714. doi:10.1080/136588198241617

Combes, P.-P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics, 63*(2), 723-742. doi:https://doi.org/10.1016/j.jue.2007.04.004

Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2010). Estimating agglomeration economies with history, geology, and worker effects. In *Agglomeration economics* (pp. 15-66): University of Chicago Press.

Combes, P.-P., & Gobillon, L. (2015). Chapter 5 - The Empirics of Agglomeration Economies. In G. Duranton, J. V. Henderson, & W. C. Strange (Eds.), *Handbook of Regional and Urban Economics* (Vol. 5, pp. 247-348): Elsevier.

Coughenour, C., de la Fuente-Mella, H., & Paz, A. (2019). Analysis of Self-Reported Walking for Transit in a Sprawling Urban Metropolitan Area in the Western U.S. *Sustainability, 11*, 852. doi:10.3390/su11030852

Dadashpoor, H., Azizi, P., & Moghadasi, M. (2019). Analyzing spatial patterns, driving forces and predicting future growth scenarios for supporting sustainable urban growth: Evidence from Tabriz metropolitan area, Iran.

Sustainable Cities and Society, 47, 101502. doi:https://doi.org/10.1016/j.scs.2019.101502

De La Roca, J., & Puga, D. (2017). Learning by Working in Big Cities. *The Review of Economic Studies, 84*(1 (298)), 106-142. Retrieved from http://www.jstor.org/stable/26160269

Deng, X., Huang, J., Rozelle, S., & Uchida, E. (2008). Growth, population and industrialization, and urban land expansion of China. *Journal of Urban Economics, 63*(1), 96-115. doi:https://doi.org/10.1016/j.jue.2006.12.006

Dewan, A. M., & Yamaguchi, Y. (2009). Using remote sensing and GIS to detect and monitor land use and land cover change in Dhaka Metropolitan of Bangladesh during 1960–2005. *Environmental Monitoring and Assessment, 150*(1), 237-249. doi:10.1007/s10661-008-0226-5

Duranton, G. (2016a). AGGLOMERATION EFFECTS IN COLOMBIA. *Journal of Regional Science, 56*(2), 210-238. doi:https://doi.org/10.1111/jors.12239

Duranton, G. (2016b). Determinants of city growth in Colombia. *Papers in Regional Science, 95*(1), 101-131. doi:https://doi.org/10.1111/pirs.12225

Duranton, G., & Puga, D. (2001). Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products. *American Economic Review, 91*(5), 1454-1477. doi:10.1257/aer.91.5.1454

Duranton, G., & Puga, D. (2004). Chapter 48 - Micro-Foundations of Urban Agglomeration Economies. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2063-2117): Elsevier.

Elmqvist, T., Andersson, E., McPhearson, T., Bai, X., Bettencourt, L., Brondizio, E., . . . Van Der Leeuw, S. (2021). Urbanization in and for the Anthropocene. *npj Urban Sustainability, 1*(1), 6. doi:10.1038/s42949-021-00018-w

Feng, Y., Liu, Y., Tong, X., Liu, M., & Deng, S. (2011). Modeling dynamic urban growth using cellular automata and particle swarm optimization rules. *Landscape and Urban Planning, 102*(3), 188-196. doi:https://doi.org/10.1016/j.landurbplan.2011.04.004

Fingleton, B. (2003). Increasing returns: evidence from local wage rates in Great Britain. *Oxford Economic Papers, 55*(4), 716-739. doi:10.1093/oep/55.4.716

Fingleton, B. (2006). The new economic geography versus urban economics: an evaluation using local wage rates in Great Britain. *Oxford Economic Papers, 58*(3), 501-530. doi:10.1093/oep/gpl006

Fujita, M., Krugman, P., & Mori, T. (1999). On the evolution of hierarchical urban systems1. *European Economic Review, 43*(2), 209-251. doi:https://doi.org/10.1016/S0014-2921(98)00066-X

Gao, J., Wei, Y. D., Chen, W., & Chen, J. (2014). Economic transition and urban land expansion in Provincial China. *Habitat International, 44*, 461-473. doi:https://doi.org/10.1016/j.habitatint.2014.09.002

Gao, J., Wei, Y. D., Chen, W., & Yenneti, K. (2015). Urban Land Expansion and Structural Change in the Yangtze River Delta, China. *Sustainability, 7*(8), 10281-10307. doi:10.3390/su70810281

Gao, Y., Su, W., & Wang, K. (2019). Does high-speed rail boost tourism growth? New evidence from China. *Tourism Management, 72*, 220-231. doi:https://doi.org/10.1016/j.tourman.2018.12.003

Glaeser, E. L., & Gottlieb, J. D. (2009). The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States. *Journal of Economic Literature, 47*(4), 983-1028. doi:10.1257/jel.47.4.983

Glaeser, E. L., Kerr, S. P., & Kerr, W. R. (2015). Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines. *The Review of Economics and Statistics, 97*(2), 498-520. doi:10.1162/REST_a_00456

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of Economic Geography, 1*(1), 27-50. doi:10.1093/jeg/1.1.27

Glaeser, Edward L., & Maré, David C. (2001). Cities and Skills. *Journal of Labor Economics, 19*(2), 316-342. doi:10.1086/319563

Glaeser, E. L., Saiz, A., Burtless, G., & Strange, W. C. (2004). The Rise of the Skilled City [with Comments]. *Brookings-Wharton Papers on Urban Affairs*, 47-105. Retrieved from http://www.jstor.org/stable/25067406

Glaeser, E. L., Scheinkman, J., & Shleifer, A. (1995). Economic growth in a cross-section of cities. *Journal of Monetary Economics, 36*(1), 117-143. doi:https://doi.org/10.1016/0304-3932(95)01206-2

Gollin, D., Jedwab, R., & Vollrath, D. (2016). Urbanization with and without industrialization. *Journal of Economic Growth, 21*(1), 35-70. doi:10.1007/s10887-015-9121-4

Gounaridis, D., Chorianopoulos, I., Symeonakis, E., & Koukoulas, S. (2019). A Random Forest-Cellular Automata modelling approach to explore future land use/cover change in Attica (Greece), under different socio-economic realities and scales. *Science of The Total Environment, 646*, 320-335. doi:https://doi.org/10.1016/j.scitotenv.2018.07.302

Groot, S. P. T., de Groot, H. L. F., & Smit, M. J. (2014). REGIONAL WAGE DIFFERENCES IN THE NETHERLANDS: MICRO EVIDENCE ON AGGLOMERATION EXTERNALITIES. *Journal of Regional Science, 54*(3), 503-523. doi:https://doi.org/10.1111/jors.12070

Gu, Y., Brown, J. F., Verdin, J. P., & Wardlow, B. (2007). A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. *Geophysical Research Letters, 34*(6). doi:https://doi.org/10.1029/2006GL029127

Harari, M. (2020). Cities in Bad Shape: Urban Geometry in India. *American Economic Review, 110*(8), 2377-2421. doi:10.1257/aer.20171673

Harris, C. D. (1954). The, Market as a Factor in the Localization of Industry in the United States. *Annals of the Association of American Geographers, 44*(4), 315-348. doi:10.1080/00045605409352140

Henderson, J. V. (1974). The Sizes and Types of Cities. *The American Economic Review, 64*(4), 640-656. Retrieved from http://www.jstor.org/stable/1813316

Henderson, J. V., & Kriticos, S. (2018). The Development of the African System of Cities. *Annual Review of Economics, 10*(1), 287-314. doi:10.1146/annurev-economics-080217-053207

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review, 102*(2), 994-1028. doi:10.1257/aer.102.2.994

Henderson, J. V., & Wang, H. G. (2007). Urbanization and city growth: The role of institutions. *Regional Science and Urban Economics, 37*(3), 283-313. doi:https://doi.org/10.1016/j.regsciurbeco.2006.11.008

Henderson, V. (2003). The Urbanization Process and Economic Growth: The So-What Question. *Journal of Economic Growth, 8*(1), 47-71. doi:10.1023/A:1022860800744

Hewison, K. (2014). Considerations on inequality and politics in Thailand. *Democratization, 21*(5), 846-866. doi:10.1080/13510347.2014.882910

Hu, Z., & Lo, C. P. (2007). Modeling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems, 31*(6), 667-688. doi:https://doi.org/10.1016/j.compenvurbsys.2006.11.001

Huang, Q., Yang, X., Gao, B., Yang, Y., & Zhao, Y. (2014). Application of DMSP/OLS Nighttime Light Images: A Meta-Analysis and a Systematic Literature Review. *Remote Sensing, 6*(8), 6844-6866. doi:10.3390/rs6086844

Iamtrakul, P., Padon, A., & Klaylee, J. (2022, 2022//). *Analysis of Urban Sprawl and Growth Pattern Using Geospatial Technologies in Megacity, Bangkok, Thailand.* Paper presented at the Geoinformatics and Data Analysis, Cham.

Ioannou, S., & Wójcik, D. (2022). Was Adam Smith an economic geographer? *GeoJournal, 87*(6), 5425-5434. doi:10.1007/s10708-021-10499-y

Jafari, M., Majedi, H., Monavari, S. M., Alesheikh, A. A., & Zarkesh, M. K. (2016). Dynamic simulation of urban expansion through a CA-Markov model Case study: Hyrcanian region, Gilan, Iran. *European Journal of Remote Sensing, 49*(1), 513-529. doi:10.5721/EuJRS20164927

Jang, S., Kim, J., & von Zedtwitz, M. (2017). The importance of spatial agglomeration in product innovation: A microgeography perspective. *Journal of Business Research, 78*, 143-154. doi:https://doi.org/10.1016/j.jbusres.2017.05.017

Jokar Arsanjani, J., Helbich, M., Kainz, W., & Darvishi Boloorani, A. (2013). Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *International Journal of Applied Earth Observation and Geoinformation, 21*, 265-275. doi:https://doi.org/10.1016/j.jag.2011.12.014

Kirasich, K., Smith, T., & Sadler, B. (2018). Random forest vs logistic regression: binary classification for heterogeneous datasets. *SMU Data Science Review, 1*(3), 9.

Krugman, P., & Elizondo, R. L. (1996). Trade policy and the Third World metropolis. *Journal of Development Economics, 49*(1), 137-150. doi:https://doi.org/10.1016/0304-3878(95)00055-0

Le, Q. B., Seidl, R., & Scholz, R. W. (2012). Feedback loops and types of adaptation in the modelling of land-use decisions in an agent-based simulation. *Environmental Modelling & Software, 27-28*, 83-96. doi:https://doi.org/10.1016/j.envsoft.2011.09.002

Li, D., Zhao, X., & Li, X. (2016). Remote sensing of human beings – a perspective from nighttime light. *Geo-spatial Information Science, 19*(1), 69-79. doi:10.1080/10095020.2016.1159389

Li, G., Sun, S., & Fang, C. (2018). The varying driving forces of urban expansion in China: Insights from a spatial-temporal analysis. *Landscape and Urban Planning, 174*, 63-77. doi:https://doi.org/10.1016/j.landurbplan.2018.03.004

Li, H., Wei, Y. D., Liao, F. H., & Huang, Z. (2015). Administrative hierarchy and urban land expansion in transitional China. *Applied Geography, 56*, 177-186. doi:https://doi.org/10.1016/j.apgeog.2014.11.029

Li, H., Wei, Y. H., & Huang, Z. (2014). Urban Land Expansion and Spatial Dynamics in Globalizing Shanghai. *Sustainability, 6*(12), 8856-8875. doi:10.3390/su6128856

Li, X., & Yeh, A. G.-O. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science, 16*(4), 323-343. doi:10.1080/13658810210137004

Liu, X., Li, X., Shi, X., Wu, S., & Liu, T. (2008). Simulating complex urban development using kernel-based non-linear cellular automata. *Ecological Modelling, 211*(1), 169-181. doi:https://doi.org/10.1016/j.ecolmodel.2007.08.024

Liu, X., Ma, L., Li, X., Ai, B., Li, S., & He, Z. (2014). Simulating urban growth by integrating landscape expansion index (LEI) and cellular automata. *International Journal of Geographical Information Science, 28*(1), 148-163. doi:10.1080/13658816.2013.831097

Liu, Y., & Feng, Y. (2016). Simulating the Impact of Economic and Environmental Strategies on Future Urban Growth Scenarios in Ningbo, China. *Sustainability, 8*(10). doi:10.3390/su8101045

Liu, Y., Yan, B., & Zhou, Y. (2016). Urbanization, economic growth, and carbon dioxide emissions in China: A panel cointegration and causality analysis. *Journal of Geographical Sciences, 26*(2), 131-152. doi:10.1007/s11442-016-1259-2

Losiri, C., Nagai, M., Ninsawat, S., & Shrestha, R. P. (2016). Modeling Urban Expansion in Bangkok Metropolitan Region Using Demographic–Economic Data through Cellular Automata-Markov Chain and Multi-Layer Perceptron-Markov Chain Models. *Sustainability, 8*(7). doi:10.3390/su8070686

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22*(1), 3-42. doi:https://doi.org/10.1016/0304-3932(88)90168-7

Martínez, L. R. (2022). How Much Should We Trust the Dictator’s GDP Growth Estimates? *Journal of Political Economy, 130*(10), 2731-2769. doi:10.1086/720458

Massimino, B., Gray, J. V., & Boyer, K. K. (2017). The Effects of Agglomeration and National Property Rights on Digital Confidentiality Performance. *Production and Operations Management, 26*(1), 162-179. doi:https://doi.org/10.1111/poms.12627

Matano, A., Obaco, M., & Royuela, V. (2020). What drives the spatial wage premium in formal and informal labor markets? The case of Ecuador. *Journal of Regional Science, 60*(4), 823-847. doi:https://doi.org/10.1111/jors.12486

Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy, 112*(4), 725-753. doi:10.1086/421174

Mills, E. S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. *The American Economic Review, 57*(2), 197-210. Retrieved from http://www.jstor.org/stable/1821621

Mohammed, A. R., & Tarpley, L. (2011). Effects of high night temperature on crop physiology and productivity: Plant growth regulators provide a management option. *Global warming impacts—case studies on the economy, human health, and on urban and natural environments,* 153-172.

Moretti, E. (2004). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics, 121*(1), 175-212. doi:https://doi.org/10.1016/j.jeconom.2003.10.015

Moulds, S., Buytaert, W., & Mijic, A. (2015). An open and extensible framework for spatially explicit land use change modelling: the lulcc R package. *Geosci. Model Dev., 8*(10), 3215-3229. doi:10.5194/gmd-8-3215-2015

Munshi, T., Zuidgeest, M., Brussel, M., & van Maarseveen, M. (2014). Logistic regression and cellular automata-based modelling of retail, commercial and residential development in the city of Ahmedabad, India. *Cities, 39*, 68-86. doi:https://doi.org/10.1016/j.cities.2014.02.007

O'Sullivan, A. (2012). *Urban Economics* (8th ed.): McGraw-Hill Education / Asia.

Poonsab, W., Vanek, J., & Carré, F. (2019). *Informal workers in urban Thailand: A statistical snapshot* (20). Retrieved from http://www.nso.go.th/sites/2014en/Survey/social/labour/LaborForce/2017/Informal_2017.pdf

Prasertsoong, N., & Puttanapong, N. (2022). Regional Wage Differences and Agglomeration Externalities: Micro Evidence from Thai Manufacturing Workers. *Economies, 10*(12). doi:10.3390/economies10120319

Puttanapong, N., Luenam, A., & Jongwattanakul, P. (2022). Spatial Analysis of Inequality in Thailand: Applications of Satellite Data and Spatial Statistics/Econometrics. *Sustainability, 14*, 3946. doi:10.3390/su14073946

Puttanapong, N., Martinez, A., Bulan, J. A., Addawe, M., Durante, R. L., & Martillan, M. (2022). Predicting Poverty Using Geospatial Data in Thailand. *ISPRS International Journal of Geo-Information, 11*(5). doi:10.3390/ijgi11050293

Quiroga-Martínez, F., & Fernández-Vázquez, E. (2021). Education as a key to reduce spatial inequalities and informality in Argentinean regional labour markets. *Regional Science Policy & Practice, 13*(1), 177-189. doi:https://doi.org/10.1111/rsp3.12369

Rice, P., Venables, A. J., & Patacchini, E. (2006). Spatial determinants of productivity: Analysis for the regions of Great Britain. *Regional Science and Urban Economics, 36*(6), 727-752. doi:https://doi.org/10.1016/j.regsciurbeco.2006.03.006

Ridhwan, M. M. (2021). Spatial wage differentials and agglomeration externalities: Evidence from Indonesian microdata. *Economic Analysis and Policy, 71*, 573-591. doi:https://doi.org/10.1016/j.eap.2021.06.013

Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge Journal of Regions, Economy and Society, 11*(1), 189-209. doi:10.1093/cjres/rsx024

Rosenthal, S. S., & Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics, 64*(2), 373-389. doi:https://doi.org/10.1016/j.jue.2008.02.006

Sangkasem, K., & Puttanapong, N. (2022). Analysis of spatial inequality using DMSP-OLS nighttime-light satellite imageries: A case study of Thailand. *Regional Science Policy & Practice, 14*(4), 828-849. doi:https://doi.org/10.1111/rsp3.12386

Shapiro, J. M. (2006). Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital. *The Review of Economics and Statistics, 88*(2), 324-335. doi:10.1162/rest.88.2.324

Short, J. R., & Pinet-Peralta, L. M. (2009). Urban Primacy: Reopening the Debate. *Geography Compass, 3*(3), 1245-1266. doi:https://doi.org/10.1111/j.1749-8198.2009.00234.x

Simon, C. J., & Nardinelli, C. (2002). Human capital and the rise of American cities, 1900–1990. *Regional Science and Urban Economics, 32*(1), 59-96. doi:https://doi.org/10.1016/S0166-0462(00)00069-7

Suphannachart, W. (2017). What drives labour productivity in the ageing agriculture of Thailand? *Advances in Management and Applied Economics, 7*(1), 89.

Suttiwichienchot, A., & Puttanapong, N. (2014). A Study on Internal Labor Movement and Policy Multiplier in Thailand. *Eurasian Journal of Economics and Finance, 2*(3), 57-68. doi:doi: 10.15604/ejef.2014.02.03.004

Tangtipongkul, K. (2015). Rates of Return to Schooling in Thailand. *Asian Development Review, 32*(2), 38-64. doi:10.1162/ADEV_a_00051

Tippakoon, P. (2011). Industrial Agglomeration and Labor Productivity: Evidence from Manufacturing Establishments in Thailand. *SOUTHEAST ASIAN JOURNAL OF ECONOMICS*, 39-76.

Tontisirin, N., & Anantsuksomsri, S. (2021). Economic Development Policies and Land Use Changes in Thailand: From the Eastern Seaboard to the Eastern Economic Corridor. *Sustainability, 13*(11). doi:10.3390/su13116153

Tripathy, P., & Kumar, A. (2019). Monitoring and modelling spatio-temporal urban growth of Delhi using Cellular Automata and geoinformatics. *Cities, 90*, 52-63. doi:https://doi.org/10.1016/j.cities.2019.01.021

Trisurat, Y., Shirakawa, H., & Johnston, J. M. (2019). Land-Use/Land-Cover Change from Socio-Economic Drivers and Their Impact on Biodiversity in Nan Province, Thailand. *Sustainability, 11*(3). doi:10.3390/su11030649

Ustaoglu, E., & Williams, B. (2017). Determinants of Urban Expansion and Agricultural Land Conversion in 25 EU Countries. *Environmental Management, 60*(4), 717-746. doi:10.1007/s00267-017-0908-2

Verburg, P. H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., & Mastura, S. S. A. (2002). Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model. *Environmental Management, 30*(3), 391-405. doi:10.1007/s00267-002-2630-x

Verburg, P. H., van Eck, J. R. R., de Nijs, T. C. M., Dijst, M. J., & Schot, P. (2004). Determinants of Land-Use Change Patterns in the Netherlands. *Environment and Planning B: Planning and Design, 31*(1), 125-150. doi:10.1068/b307

Vivatsurakit, T., & Vechbanyongratana, J. (2020). Returns to education among the informally employed in Thailand. *Asian-Pacific Economic Literature, 34*(1), 26-43. doi:https://doi.org/10.1111/apel.12284

Wang, Y., Hu, Y., Niu, X., Yan, H., & Zhen, L. (2022). Land Use/Cover Change and Its Driving Mechanism in Thailand from 2000 to 2020. *Land, 11*(12). doi:10.3390/land11122253

Warunsiri, S., & McNown, R. (2010). The Returns to Education in Thailand: A Pseudo-Panel Approach. *World Development, 38*(11), 1616-1625. doi:https://doi.org/10.1016/j.worlddev.2010.03.002

Wheaton, W. C. (1974). A comparative static analysis of urban spatial structure. *Journal of Economic Theory, 9*(2), 223-237. doi:https://doi.org/10.1016/0022-0531(74)90068-4

Wheeler, Christopher H. (2001). Search, Sorting, and Urban Agglomeration. *Journal of Labor Economics, 19*(4), 879-899. doi:10.1086/322823

Wheeler, C. H. (2006). Cities and the growth of wages among young workers: Evidence from the NLSY. *Journal of Urban Economics, 60*(2), 162-184. doi:https://doi.org/10.1016/j.jue.2006.02.004

White, R., & Engelen, G. (1993). Cellular Automata and Fractal Urban Form: A Cellular Modelling Approach to the Evolution of Urban Land-Use Patterns. *Environment and Planning A: Economy and Space, 25*(8), 1175-1199. doi:10.1068/a251175

White, R., & Engelen, G. (1997). Cellular Automata as the Basis of Integrated Dynamic Regional Modelling. *Environment and Planning B: Planning and Design, 24*(2), 235-246. doi:10.1068/b240235

Wu, F. (2002). Calibration of stochastic cellular automata: the application to rural-urban land conversions. *International Journal of Geographical Information Science, 16*(8), 795-818. doi:10.1080/13658810210157769

Xu, F., Wang, Z., Chi, G., & Zhang, Z. (2020). The impacts of population and agglomeration development on land use intensity: New evidence behind urbanization in China. *Land Use Policy, 95*, 104639. doi:https://doi.org/10.1016/j.landusepol.2020.104639

Yankow, J. J. (2006). Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. *Journal of Urban Economics, 60*(2), 139-161. doi:https://doi.org/10.1016/j.jue.2006.03.004

Zhang, Q., & Seto, K. C. (2011). Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sensing of Environment, 115*(9), 2320-2329. doi:https://doi.org/10.1016/j.rse.2011.04.032

Zhang, Y., Chen, N., Wang, S., Wen, M., & Chen, Z. (2023). Will carbon trading reduce spatial inequality? A spatial analysis of 200 cities in China. *Journal of Environmental Management, 325*, 116402. doi:https://doi.org/10.1016/j.jenvman.2022.116402

Electronic Media

Bangkok Post (Producer). (2022). 5gct, Thai-sustainability tech startup sets stage for smart cities globally. Retrieved from https://www.bangkokpost.com/thailand/pr/2348498/5gct-thai-sustainability-tech-startup-sets-stage-for-smart-cities-globally

Other Materials

Amare, M., Hohfeld, L., Jitsuchon, S., & Waibel, H. (2012). *Rural–Urban Migration and Employment Quality: A Case Study from Thailand* (309). Retrieved from Manila, Philippines: https://www.adb.org/sites/default/files/publication/29829/economics-wp-309.pdf

Anselin, L., Syabri, I., & Smirnov, O. (2002). Visualizing Multivariate Spatial Correlation with Dynamically Linked Windows. *New Tools for Spatial Data Analysis: Proceedings of the Specialist Meeting; Santa Barbara*.

Asian Development Bank. (2019). *Asian Development Outlook 2019 Update: Fostering Growth and Inclusion in Asia's Cities*. Retrieved from Manila: https://www.adb.org/sites/default/files/publication/524596/ado2019-update-theme-chapter.pdf

Bartolini, D., S. Stossberg and H. Blöchliger (2016), Fiscal Decentralisation and Regional Disparities, *OECD Economics Department Working Papers*, No. 1330, OECD Publishing, Paris, https://doi.org/10.1787/5jlpq7v3j237-en

Daanen, H., Jonkhoff, W., Bosch, P., & Broeke, H. (2013). The effect of global warming and urban heat islands on mortality, morbidity and productivity in the Netherlands. *Proceedings of the 15th international conference on environmental ergonomics, 2013*, 16-19.

Grover, A., Lall, S. V., & Timmis, J. (2021). *Agglomeration Economies in Developing Countries* (9730). Retrieved from Washington, DC., USA: https://openknowledge.worldbank.org/handle/10986/36003

Houbcharaun, A. (2013b). *Essays on economic liberalization and spatial economics.* (Doctoral). University of Hawai'i at Manoa, Honolulu, USA.

International Monetary Fund. (2023). *World Economic Outlook: A Rocky Recovery*. Washington, DC. April.

Limpanonda, S. (2012b). *Provincial disparities in Thailand: convergence, agglomeration economies and effects on poverty, 1988-2008.* (Doctoral). SOAS, University of London, London, UK.

Lösch, A. (1940). *The Economics of Location*. Jena, Germany: Fischer (English translation, New Haven, CT: Yale University Press, 1954).

Puttanapong, N. (2008). *The Computable General Equilibrium (CGE) Models with Monte-Carlo Simulation for Thailand: An Impact Analysis of Baht Appreciation.* (Doctoral). Cornell University, Ithaca, NY, USA.

Southichack, M. K. (1998). *Regional convergence and agglomeration economies: The case of Thailand*. (Doctoral). University of Hawai'i at Manoa, Honolulu, USA.

World Bank (2021). Employment in agriculture (% of total employment) (modeled ILO estimate) - Thailand. Retrieved from https://data.worldbank.org/indicator/SL.AGR.EMPL.ZS?locations=TH. Available from International Labour Organization ILOSTAT database Retrieved 19 October 2022, from The World Bank https://data.worldbank.org/indicator/SL.AGR.EMPL.ZS?locations=TH

United Nations. (2018). *2018 Revision of World Urbanization Prospects*. Retrieved from https://www.un.org/en/desa/2018-revision-world-urbanization-prospects

World Bank. (2021). Urban population. Retrieved from https://data.worldbank.org/indicator/SP.URB.TOTL?view=chart. Available from United Nations Population Division's World Urbanization Prospects: 2018 Revision Retrieved 28 March 2023, from World Bank, https://data.worldbank.org/indicator/SP.URB.TOTL?view=chart

APPENDICES

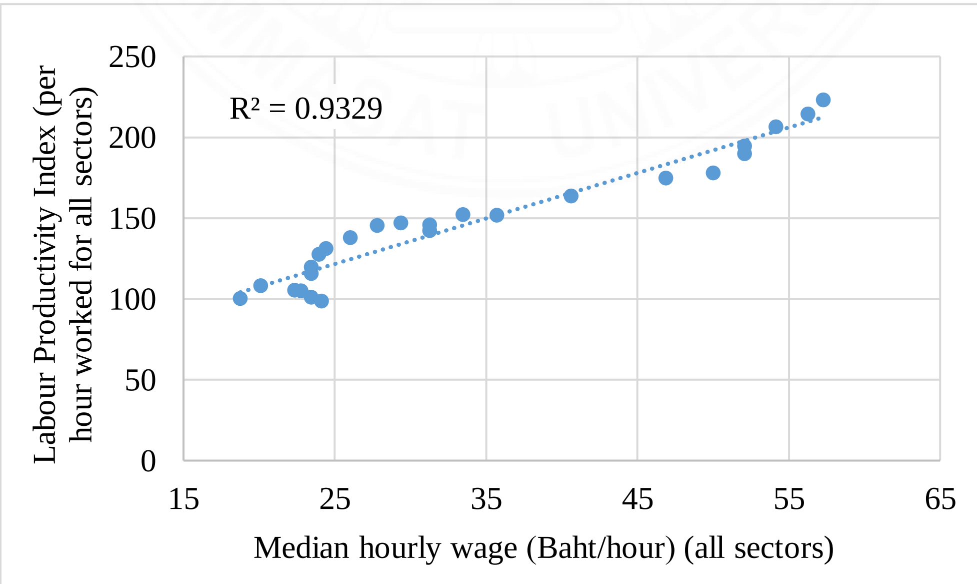
APPENDIX A
RELATIONSHIP BETWEEN WAGE AND LABOR PRODUCTIVITY IN THAILAND

Figures A.1 and A.2 present the statistically significant relationship between wage and labor productivity in Thailand over the period of 1995–2020. The labor productivity index, which measures the hourly output per worker, was calculated by the Bank of Thailand. On the other hand, the hourly wage data was obtained from the LFS These figures illustrate the observed correlation between wages and labor productivity in Thailand and provide evidence of their significant relationship.

Figure A.1 displays the regression analysis between the median hourly wage and hourly productivity, resulting in a high R-squared value. Similarly, Figure A.2 exhibits the regression analysis between the average hourly wage and hourly labor productivity, also yielding a high R-squared value. These results confirm the validity of using the hourly wage as a proxy for hourly labor productivity, indicating a strong relationship between wages and productivity in the context of the study.

Figure A.1

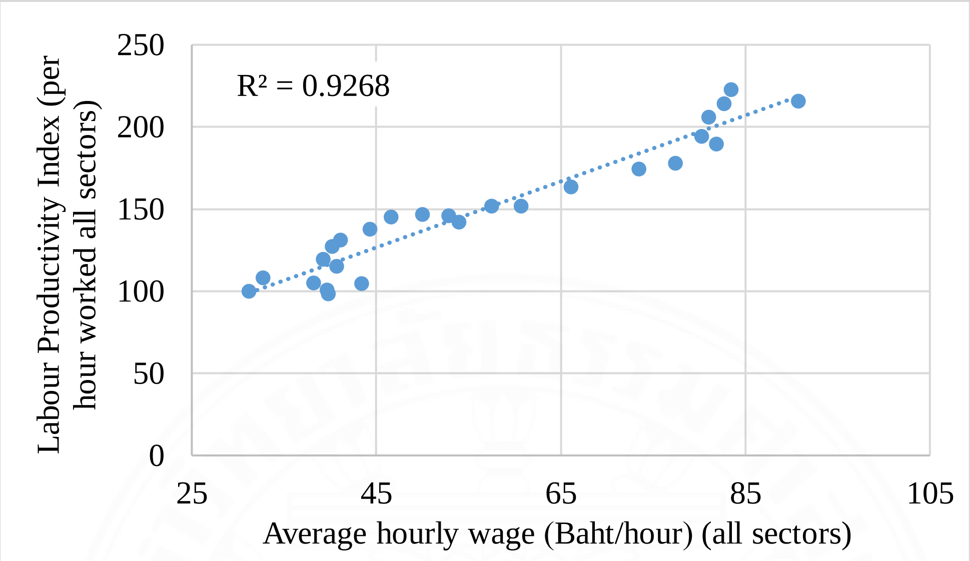
Scatter Plot of Labor Productivity and the Median Hourly Wage



Note. This figure was created from author's calculations.

Figure A.2

Scatter Plot of Labor Productivity and the Average Hourly Wage



Note. This figure was created from author's calculations.

APPENDIX B

SPECIFICATION OF DATA THAT ARE USED IN THE ANALYSIS OF CHAPTER 2

Table B.1

Main Specifications of Geospatial Data That Are Used in the Analysis of Chapter 2

Variable	Satellite/Dataset	Resolution	Unit	Frequency	Technical reference
Population	WorldPop	~92 meters	Number of people	Annual	https://developers.google.com/earth-engine/datasets/catalog/WorldPop_GP_100m_pop#bands
NTL (2007)	DMSP/OLS	~ 920 meters	nanoWatts/cm2/sr	Annual	https://developers.google.com/earth-engine/datasets/catalog/NOAA_DMSP-OLS_NIGHTTIME_LIGHTS#bands
NTL (2012 and 2017)	VIIRS/DNB	~ 460 meters	nanoWatts/cm2/sr	Monthly	https://developers.google.com/earth-engine/datasets/catalog/NOAA_VIIRS_DNB_MONTHLY_V1_VCMCFG
Percentage of clay content in the soil	EnvirometriX Ltd	250 meters	% (kg / kg)	Annual	https://developers.google.com/earth-engine/datasets/catalog/OpenLandMap_SOL_SOL_CLAY-WFRACTION_USDA-3A1A1A_M_v02

Note. This figure was created from author's calculations.

Table B.2

Main Specifications of Economic Data That Are Used in the Analysis of Chapter 2

Variable	Dataset	Range	Frequency	Technical reference
Specialization	Author's calculation from IC	[0,1]	Annual	IC from NSO (Thailand)
Diversity	Author's calculation from IC	(-∞,∞)	Annual	IC from NSO (Thailand)
Competition	Author's calculation from IC	(-∞, 1]	Annual	IC from NSO (Thailand)

Note. This figure was created from author's calculations.

APPENDIX C
BIVARIATE LOCAL MORAN'S I

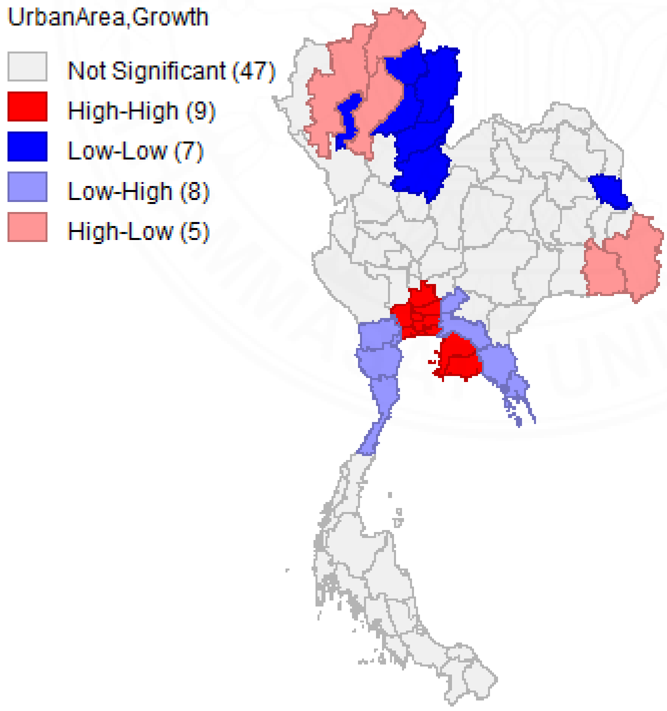
Following Anselin, Syabri, & Smirnov (2002), bivariate Local Moran's I is calculated by the following equation.

$$I_i^B = cx_i \sum_j w_{ij} y_j, \tag{C.1}$$

where w_{ij} are the spatial weights matrix which is based on K-Nearest neighbors with number of neighbor of 5, x_i is the value of one variable at the location i, y_j represents the average value of the neighboring values for another variable, and c is expressed by $\frac{1}{\sum_i (x_i - \bar{x})_i^2}$.

Figure C.1

BiLISA Cluster Map

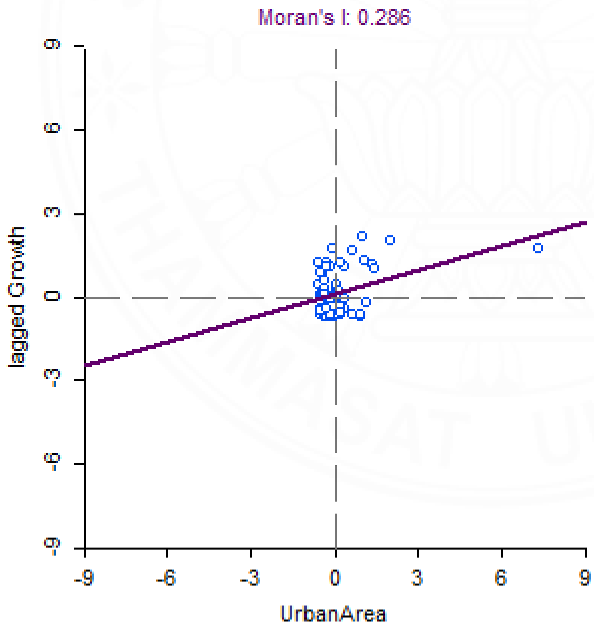


Note. This figure was created from author's calculations.

In this study, x_i represents urban area of province i in year 2000. The y_j represents growth of urban area between year 2000 and 2020 of province i. As illustrated in Figure C.1, the red areas (High-High) are the areas where urbanization has been concentrating i.e., the provinces with higher-than-average urban area in 2000 and having higher-than-average urban growth between 2000 and 2020. The blue area (Low-Low) are the areas with low urbanization i.e., the provinces with lower-than-average urban area in 2000 and having lower-than-average urban growth between 2000 and 2020. The purple areas (Low-Low) are the newly urbanized provinces i.e., the provinces with lower-than-average urban area in 2000 and but experiencing higher-than-average urban growth between 2000 and 2020. Lastly, the pink areas (High-Low) are the province with higher-than-average urban area in 2000 but having lower-than-average urban growth between 2000 and 2020.

Figure C.2

Moran Scatter Plot



Note. This figure was created from author's calculations.

Figure C.2 illustrates the spatial autocorrelation of urbanization in Thailand. It shows a positive relationship between urban areas and their subsequent growth, suggesting that highly urbanized areas are likely to be surrounded by areas with high urbanization and vice versa.

APPENDIX D

SPECIFICATION OF DATA THAT ARE USED IN THE ANALYSIS OF CHAPTER 3

Table D.1

Main Specifications of Geospatial Data That Are Used in the Analysis of Chapter 3

Variable	**Dataset**	**Resolution**	**Frequency**	**Unit**	**Technical reference**
NDDI	Terra MODIS	500 meters	8 days	-	https://developers.google.com/earth-engine/datasets/catalog/MODIS_006_MOD09A1
NDVI	Terra MODIS	500 meters	8 days	-	https://developers.google.com/earth-engine/datasets/catalog/MODIS_006_MOD09A1
NDWI	Terra MODIS	500 meters	8 days	-	https://developers.google.com/earth-engine/datasets/catalog/MODIS_006_MOD09A1
Land Surface Temperature (Night)	Terra MODIS	1 km	8 days	Celsius	https://developers.google.com/earth-engine/datasets/catalog/MODIS_006_MOD11A2
Urban area	MODIS Terra and Aqua reflectance	500 meters	Annual	Square kilometer	https://developers.google.com/earth-engine/datasets/catalog/MODIS_006_MCD12Q1
Precipitation (rainfall)	CHIRPS	~5 km	Daily	Millimeter of rain per year	https://developers.google.com/earth-engine/datasets/catalog/UCSB-CHG_CHIRPS_DAILY

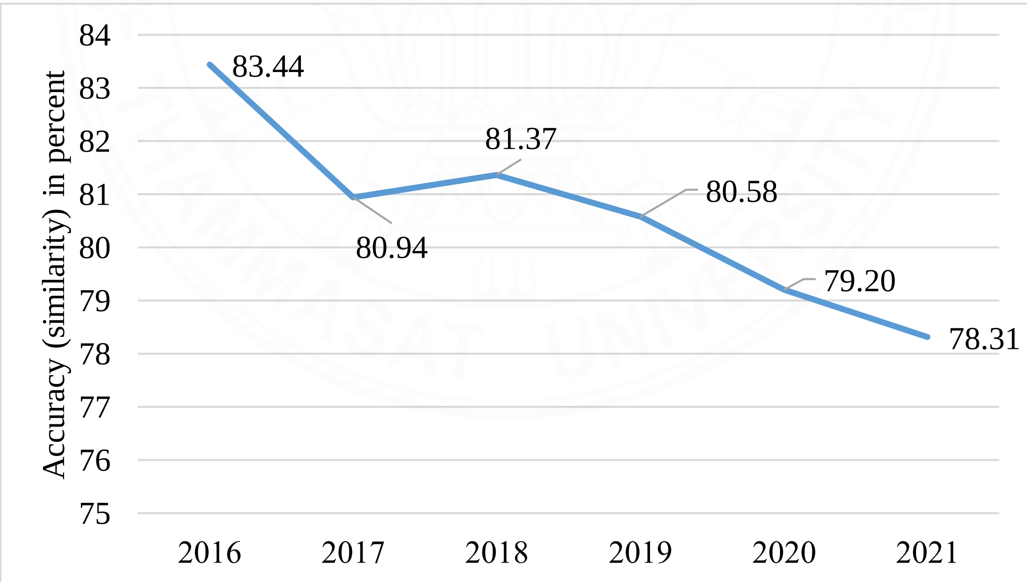
Note. This figure was created from author's calculations.

APPENDIX E
COMPARISON BETWEEN OF LAND-USE MAP FROM TWO SOURCES

Figure E.1 displays similarity between land-use map from Dynamic World and land-used map from Land Development Department in percentage. The result suggests that, on average, 80.64 percent of land cover types classified as built-up and non-built-up area from Dynamic World could be matched with the land-use maps from Land Development Department. Likewise, Figure E.2 illustrates the similarity between land-use map from Dynamic World and land-used map from Land Development Department from 2016–2021.

Figure E.1

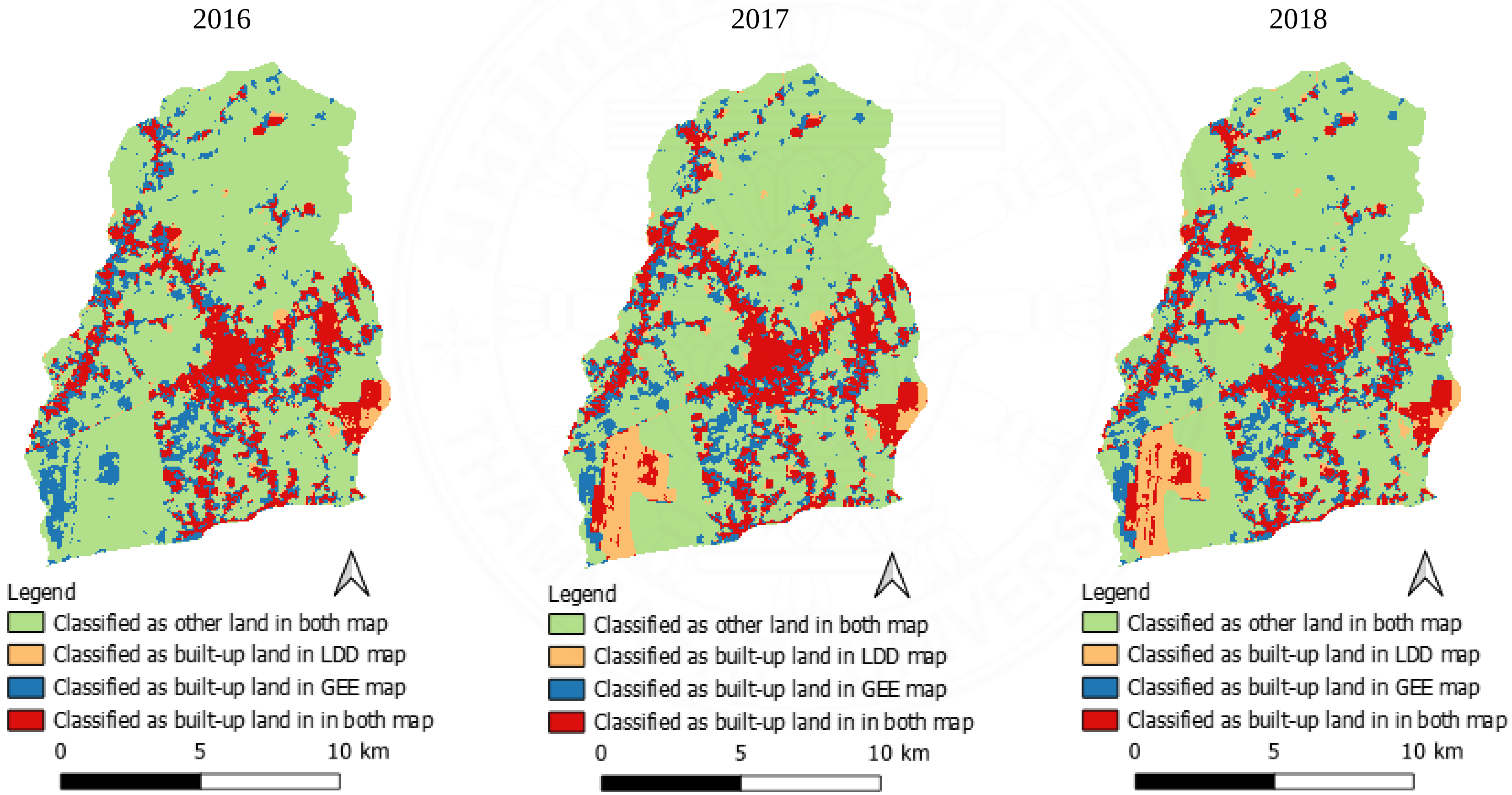
Similarity Between Land-Use Map From Dynamic World and Land-Used Map From Land Development Department in Percent, 2016–2021

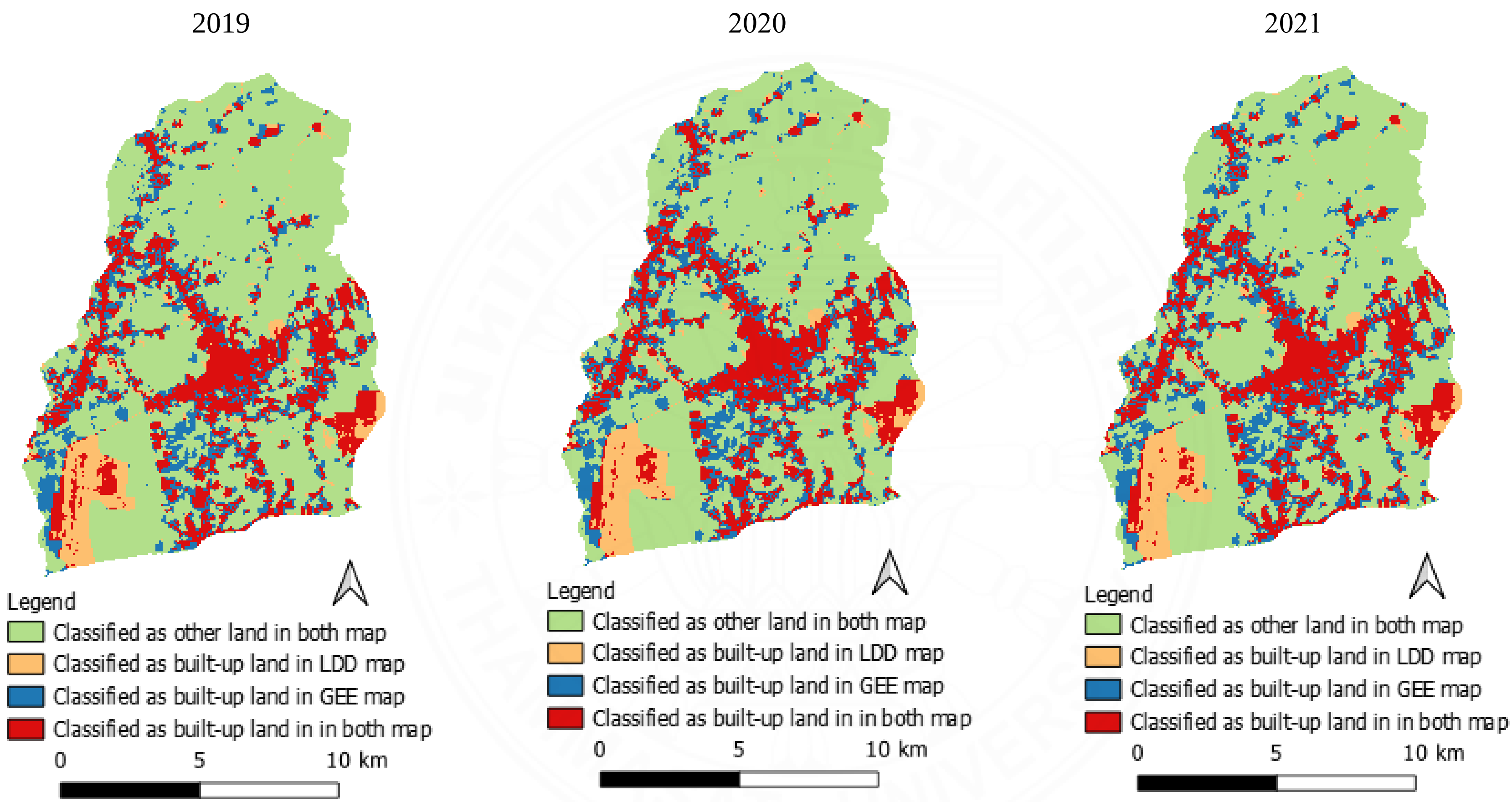


Note. This figure was created from author's calculations.

Figure E.2

Illustration of Similarity Between Land-Use Map From Dynamic World and Land-Used Map From Land Development Department, 2016–2021





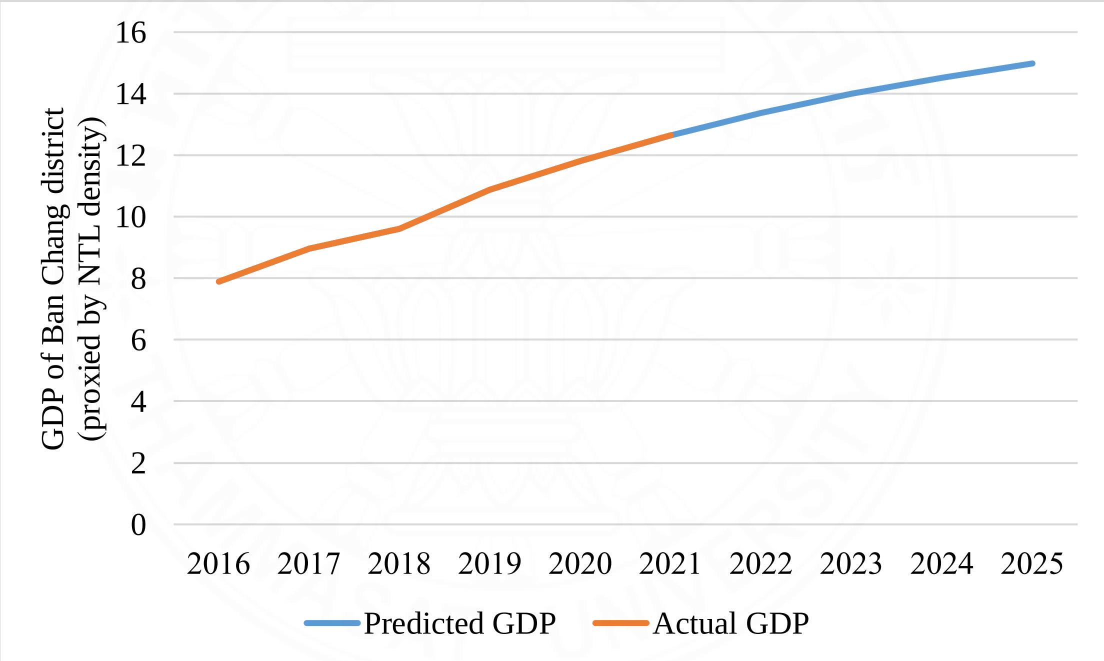
Note. This figure was created from author's calculations.

APPENDIX F
RESULTS FROM VECM

Figure F.1 and F.2 show the projected Ban Chang's GDP (proxied by NTL density) and Ban Chang's urban land from 2021–2025 from VECM. The shape of the distribution of projected variables, GDP, and urban land is shown in Figure F.3 and F.4, respectively.

Figure F.1

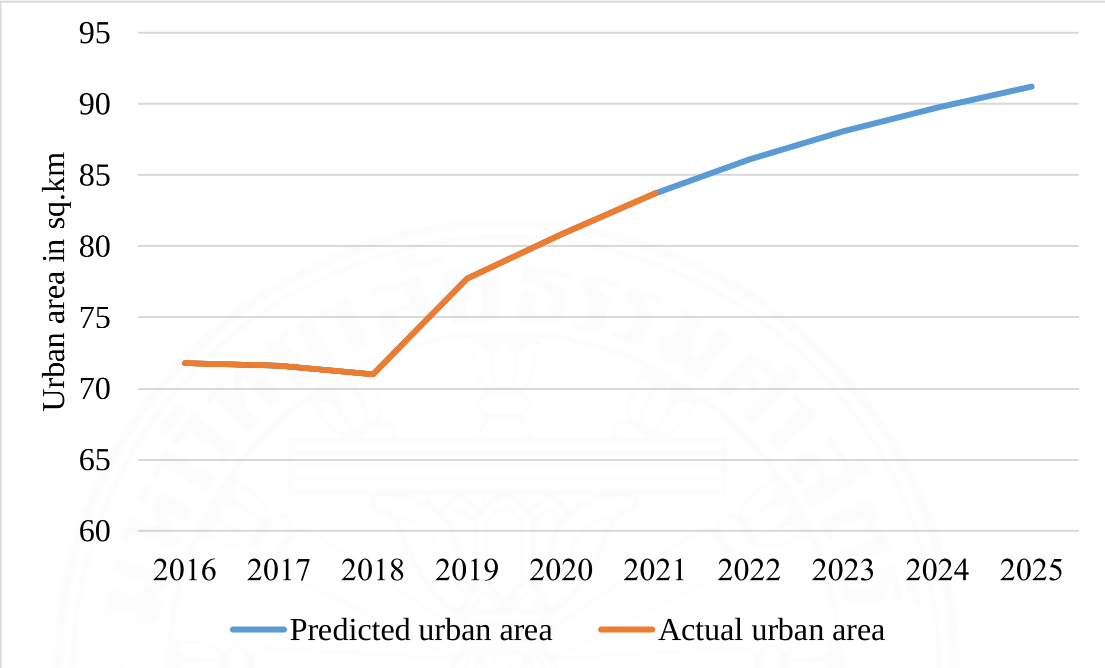
Projection of Ban Chang's GDP (Proxied by NTL Density) From VECM, 2021–2025



Note. This figure was created from author's calculations.

Figure F.2

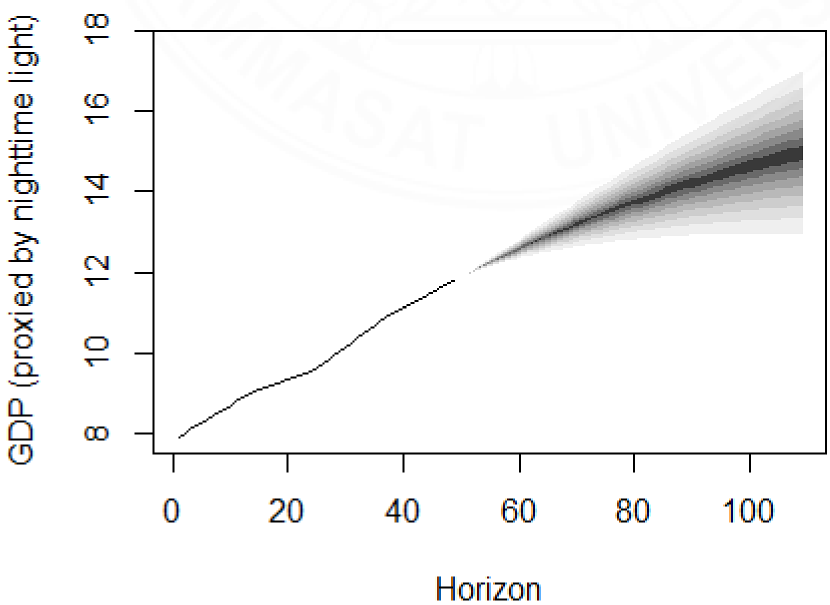
Projection of Ban Chang's Urban Land From VECM (The Most Likely Scenario), 2021–2025



Note. This figure was created from author's calculations.

Figure F.3

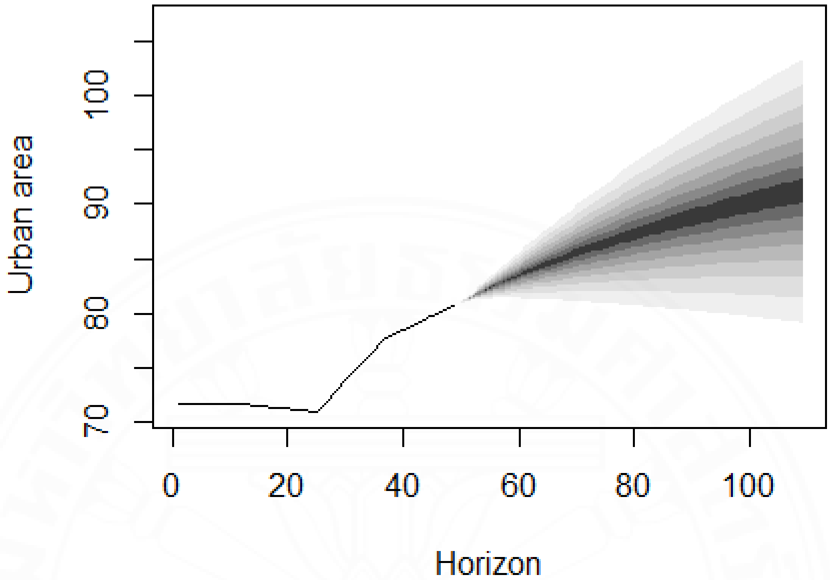
Fanchart of the Distribution of Projected Ban Chang's GDP (Proxied by NTL Density) From VECM, 2021–2025



Note. This figure was created from author's calculations.

Figure F.4

Fanchart of the Distribution of Projected Ban Chang Urban Area From VECM, 2021–2025



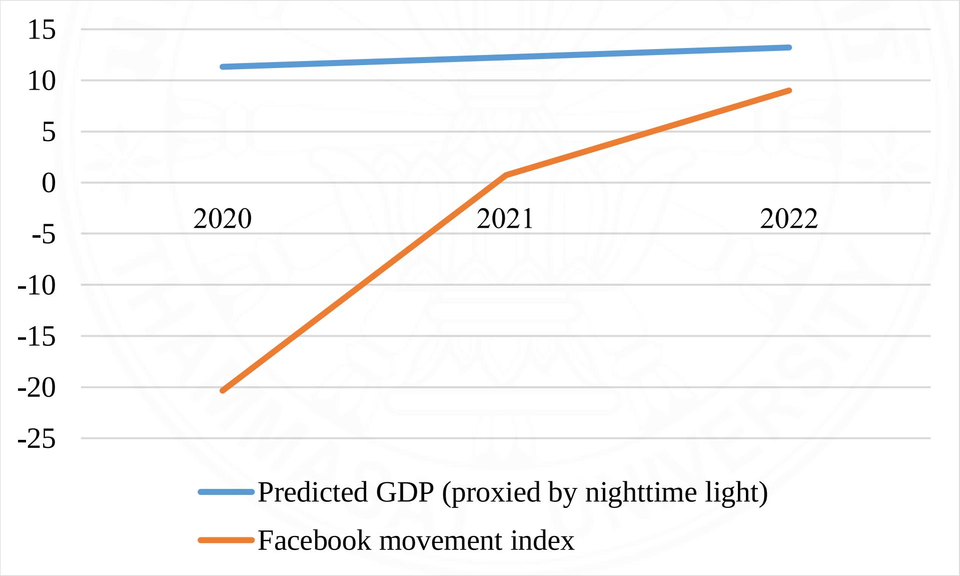
Note. This figure was created from author's calculations.

APPENDIX G
COMPARISON BETWEEN PROJECTED ECONOMIC GROWTH AND OTHER SOCIOECONOMIC VARIABLES

Figure G.1 plots the correlation between predicted GDP growth and the facebook movement index. It shows a positive relationship between predicted GDP and facebook movement index. Notably, facebook movement index substantially increased from -20 in 2020 to 9 in 2022, reflecting ease of COVID-19 lockdown between 2020 and 2022.

Figure G.1

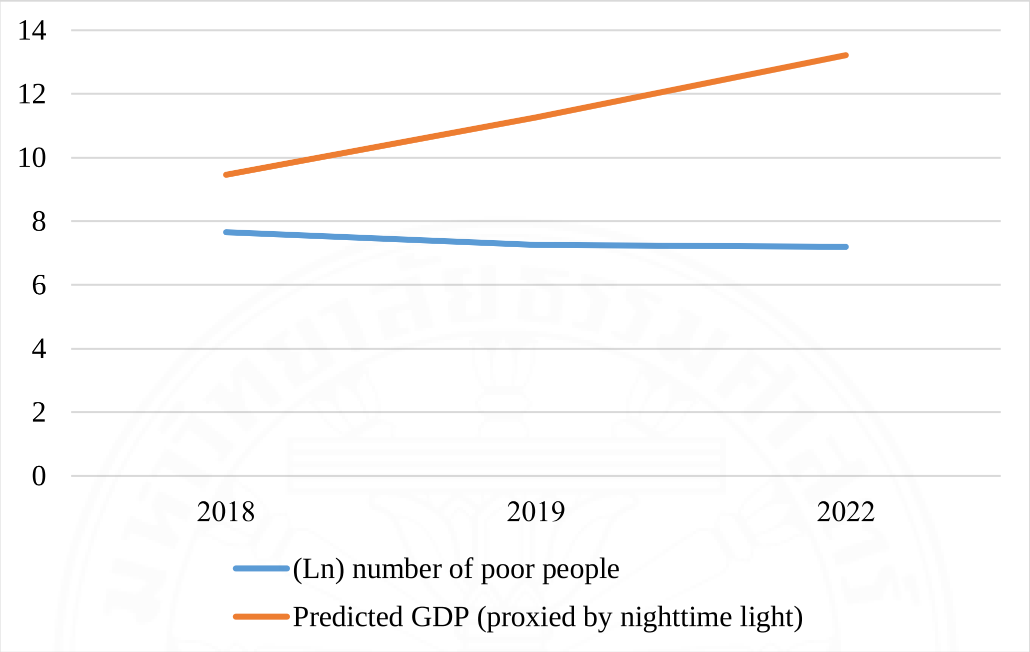
Correlation Between GDP and the Facebook Movement Index



Note. (i) Data on mobility index is publicly available at https://data.humdata.org/dataset/movement-range-maps?fbclid=IwAR2BTgi1LT6fGfR4kpxFDJBEPE0GlOp7XDVpuzNzS3bS_kVxM7t1df2qr4k. (ii) The higher value of Facebook movement index represents higher movement of population.

Figure G.2

Predicted GDP (Proxied by NTL) Of Ban Chang District and (Natural) Logarithm of the Number of Poor People, 2018–2022



Note. Data on monetary poverty between 2020-2021 is not available.

Figure G.2 plots the correlation between predicted GDP and the number of poor people in Ban Chang district. It shows the inverse relationship between local economic expansion and monetary poverty (measured by employment and household income) between 2018–2022.[12]

[12] The information on monetary poverty of Ban Chang district is publicly available at https://www.tpmap.in.th/2561/2102

APPENDIX H

SPECIFICATION OF DATA THAT ARE USED IN THE ANALYSIS OF CHAPTER 4

Table H.1

Main Specifications of Geospatial and Economic Data That Are Used in the Analysis of Chapter 4

Variable	**Dataset**	**Resolution**	**Cell size**	**Unit**	**Technical reference**
Urban area	Dynamic World	10 meters	29.78 Meters	Square kilometer	https://developers.google.com/earth-engine/datasets/catalog/GOOGLE_DYNAMICWORLD_V1#bands
Slope	USGS	~ 230 meters	29.78 Meters	Percent	https://developers.google.com/earth-engine/datasets/catalog/USGS_GMTED2010
Elevation	USGS	~ 230 meters	29.78 Meters	Meters	https://developers.google.com/earth-engine/datasets/catalog/USGS_GMTED2010
Distance to urban existing urban area	Author's calculation	10 meters	29.78 Meters	Meters	-
NTL	VIIRS/DNB	~ 460 meters	-	nanoWatts/cm2/sr	https://developers.google.com/earth-engine/datasets/catalog/NOAA_VIIRS_DNB_MONTHLY_V1_VCMCFG
GDP of top five-trading partners of Thailand	World Bank	-	-	US dollars	https://data.worldbank.org/indicator/NY.GDP.MKTP.CD
Construction price index	Ministry of Commerce	-	-	-	http://www.indexpr.moc.go.th/price_present/csi/stat/other/conyear.asp

Note. This table was created from author's calculations.

APPENDIX I
EXAMPLE OF SOURCE CODE FOR THE ANALYSIS IN CHAPTER 4

Table I.1

Source Code for Simultaneous Equation

```
clear
cd "D:\OneDrive\Desktop\Rayong district LULC"
import excel "For 2SLS.xlsx", sheet("Sheet2") firstrow

ipolate z1 time, gen(z1_ipo)
ipolate z2 time, gen(z2_ipo)
ipolate gpp time, gen(gpp_ipo)
ipolate u time, gen(u_ipo)

drop z1 z2 gpp u

rename z1_ipo z1
rename z2_ipo z2
rename gpp_ipo gpp
rename u_ipo u

export excel using "For 2SLS_v1.xlsx", sheetreplace firstrow(variables)

**# Forecast
* Import data
clear
import excel "For 2SLS_v1.xlsx", sheet("Sheet1") firstrow
tsset time

**# Train the model #**
	reg3 (gpp u z1) (u gpp z2), endog(gpp u) exog(z1 z2)
	estimates store urbansim

	forecast create urbansim_model,replace
	forecast estimates urbansim
	forecast exogenous z1
	forecast exogenous z2
	forecast exogenous time
	forecast solve, prefix(s_) begin(1) end(49)
```

Table I.1

Source Code for Simultaneous Equation (Cont.)

```
**# Projection begin here #**
* Specify how many years forward
local t=5
forvalues i=1(1)`t' {
	set obs `=_N+1'
	replace year = year[_n-1] + 1 if missing(year)
	replace month = month[_n-1] if missing(month)
	replace time = time[_n-1] + 1 if missing(time)

	local g = 0.0
	local delta = 1+`g'
	replace gpp = gpp[_n-1]*`delta' if missing(gpp)

	* Assume long-term trade partner's GDP growth of 3.2 percent
	local s = 0.032
	local alpha = 1+`s'
	replace z1 = z1[_n-1]*`alpha' if missing(z1)

	* Guess z2 in 2019 (assume 1.1 percent inflation)
	local p = 0.011
	local beta = 1+`p'
	replace z2 = z2[_n-1]*`beta' if missing(z2)

	* Assume long-term urban area growth of 0 percent
	local b = 0.0
	local beta_1 = 1+`b'
	replace u = u[_n-1]*`beta_1' if missing(u)

	* Re-forecast
	reg3 (gpp u z1) (u gpp z2), endog(gpp u) exog(z1 z2)
	estimates store urbansim_1

	forecast create urbansim_model_1,replace
	forecast estimates urbansim
	forecast exogenous z1
	forecast exogenous z2
	forecast exogenous time
	local e=49+`i'
	forecast solve, prefix(s1_) begin(1) end(`e')
```

Table I.1

Source Code for Simultaneous Equation (Cont.)

```
            * Append the projected values to original data
            local y=2020+`i'
            replace s_gpp=s1_gpp if year==`y'
            replace s_u=s1_u if year==`y'
            replace u=s1_u if year==`y'
            drop s1_gpp s1_u _est_urbansim_1 _est_urbansim
}
```

Note. This table is an example of the author's source code. Contact the author via email in the biography section for the entire script.

Table I.2

Source Code for Land-Use Change Simulation Between 2017-2020

```
## Import all necessary packages
library(lulcc)
library(gsubfn)
library(caret)
library(rgdal)
library(randomForest)
#library(rpart)
library(Hmisc)
library(raster)
library(rasterVis)
## Expand memory allowance capacity to maximum
# memory.limit(size=56000)

getwd()
setwd("/cloud/project/BanChang_mod")

## Import land-use map of study area
BanChang2016 <- raster("BanChang2016_u.tif")
BanChang2017 <- raster("BanChang2017_u.tif")
BanChang2018 <- raster("BanChang2018_u.tif")
BanChang2019 <- raster("BanChang2019_u.tif")
BanChang2020 <- raster("BanChang2020_u.tif")

## Import driving factors of study area
## bcn_2.tif = slope, bcn_3.tif = elevation, bcn_6.tif = distance to urban settlement in 2016
ef_2_b <- raster("bcn_2.tif")
ef_3_b <- raster("bcn_3.tif")
ef_4_b <- raster("bcn_6.tif")
```

Table I.2

Source Code for Land-Use Change Simulation Between 2017-2020 (Cont.)

```
## Resample to produce maps with identical extent and resolution
ef_2 <- resample(ef_2_b,BanChang2016,method='bilinear')
ef_3 <- resample(ef_3_b,BanChang2016,method='bilinear')
ef_4 <- resample(ef_4_b,BanChang2016,method='bilinear')

## Stack all maps
names(BanChang2016) <- "lu_2016"
names(BanChang2020) <- "lu_2020"
BanChang.map <- stack(BanChang2016,BanChang2020,ef_2,ef_3,ef_4)
na.omit(BanChang.map)

obs.BanChang <- ObsLulcRasterStack(x=stack(BanChang2016,BanChang2020),
                    categories=c(1,2),
                    labels=c("Other","Built"),
                    t=c(0,4))
obs.BanChang.no.na <- na.omit(obs.BanChang)
obs.BanChang.no.na

## Plot land-use map
plot(obs.BanChang.no.na)
## Produce transition matrix
crossTabulate(obs.BanChang.no.na, times=c(0,4))

## Stack explanatory variables
ev <- ExpVarRasterList(x=BanChang.map, pattern="bcn")
ev.no.na <- na.omit(ev)
ev.no.na

## Partition data to train the model
part.BanChang<- partition(x=obs.BanChang.no.na[[1]], size=0.7, spatial=TRUE)

train.data <- getPredictiveModelInputData(obs=obs.BanChang.no.na, ef=ev.no.na,
cells=part.BanChang[["train"]])
train.data.no.na <- na.omit(train.data)
all.data <- getPredictiveModelInputData(obs=obs.BanChang.no.na, ef=ev.no.na,
cells=part.BanChang[["all"]])

forms <- list(Other ~ bcn_2+bcn_3+bcn_6,
        Built ~ bcn_2+bcn_3+bcn_6)

glm.models <- glmModels(formula=forms, family=binomial, data=train.data.no.na,
obs=obs.BanChang.no.na)
```

Table I.2

Source Code for Land-Use Change Simulation Between 2017-2020 (Cont.)

```
## Create suitability maps
suitability.maps <- predict(object=glm.models, newdata=all.data, data.frame=TRUE)
points <- rasterToPoints(obs.BanChang.no.na[[1]], spatial=TRUE)
suitability.maps <- SpatialPointsDataFrame(coords=points, data=suitability.maps)
#r <- stack(rasterize(x=suitability.maps, y=obs.BanChang.no.na[[1]],
field=names(suitability.maps)))
#plot(r)

## Test ability of models to predict allocation of built and other
## Land uses in testing partition
test.data <- getPredictiveModelInputData(obs=obs.BanChang.no.na, ef=ev.no.na,
cells=part.BanChang[["test"]])
test.data.no.na <- na.omit(test.data)

glm.pred <- PredictionList(models=glm.models, newdata=test.data.no.na)
glm.perf <- PerformanceList(pred=glm.pred, measure="rch")

plot(list(glm=glm.perf))

## Test ability of models to predict location of urban gain 2016 to 2020
part.1 <- rasterToPoints(obs.BanChang.no.na[[1]], fun=function(x) x != 2,
spatial=TRUE)
test.data.1 <- getPredictiveModelInputData(obs=obs.BanChang.no.na, ef=ev.no.na,
cells=part.1, t=4)
test.data.1.no.na <- na.omit(test.data.1)
glm.pred.1 <- PredictionList(models=glm.models[[2]], newdata=test.data.1.no.na)
glm.perf.1 <- PerformanceList(pred=glm.pred.1, measure="rch")
plot(list(glm=glm.perf.1))

## Obtain demand scenario via linear interpolation
## Some routines are coupled to complex economic models that predict future or
past land use demand based on economic considerations
## However, linear extrapolation of trends remains a useful technique
dmd <- approxExtrapDemand(obs=obs.BanChang.no.na, tout=0:4)
## Note that this package is not equipped with advanced demand forecasting
technique
## However, users could create matrix that specify future land-use demand obtained
from simultaneous equation
## Use this command to modify the dmd matrix according to results from
simultaneous equation
dmd[c(1),] <- c(128150 ,55100)
dmd[c(2),] <- c(128056 ,55194)
```

Table I.2

Source Code for Land-Use Change Simulation Between 2017-2020 (Cont.)

```
dmd[c(3),] <- c(128875 ,54375)
dmd[c(4),] <- c(123230 ,60020)
dmd[c(5),] <- c(121441 ,61809)

## To see example of modified the dmd matrix run the following command
dmd

## Plot change
matplot(dmd, type="l", ylab="Demand (no. of cells)", xlab="Time point",
      lty=1, col=c("Green","Blue","Brown","Red"))
legend("topleft", legend=obs.BanChang.no.na@labels,
col=c("Green","Blue","darkgoldenrod2","Red"), lty=1)

## Create CLUE-S model object
clues.rules <- matrix(data=1, nrow=2, ncol=2, byrow=TRUE)
clues.parms <- list(jitter.f=0.0002,
                scale.f=0.000001,
                max.iter=1000,
                max.diff=50,
                ave.diff=50)

clues.model <- CluesModel(obs=obs.BanChang.no.na,
                     ef=ev.no.na,
                     models=glm.models,
                     time=0:4,
                     demand=dmd,
                     elas=c(0.2,0.2),
                     rules=clues.rules,
                     params=clues.parms)

## Perform allocation
clues.model <- allocate(clues.model)
plot(clues.model)

# Pattern validation
#clues.tabs <- ThreeMapComparison(x=clues.model,
#                            factors=2^(1:8),
#                            timestep=4)
#clues.agr <-  AgreementBudget(x=clues.tabs)
```

Table I.2

Source Code for Land-Use Change Simulation Between 2017-2020 (Cont.)

```
# Plot(clues.agr, from=1, to=2)
#clues.fom <- FigureOfMerit(x=clues.tabs)
#plot(clues.fom, from=1, to=4)

## Export result from CLUE-S model
writeRaster(clues.model@output@layers[[1]],"BanChang2016_sim.tif",
overwrite=TRUE)
writeRaster(clues.model@output@layers[[2]],"BanChang2017_sim.tif",
overwrite=TRUE)
writeRaster(clues.model@output@layers[[3]],"BanChang2018_sim.tif",
overwrite=TRUE)
writeRaster(clues.model@output@layers[[4]],"BanChang2019_sim.tif",
overwrite=TRUE)
writeRaster(clues.model@output@layers[[5]],"BanChang2020_sim.tif",
overwrite=TRUE)

## Export result from CLUE-S model (linear interpolation)
writeRaster(clues.model@output@layers[[1]],"BanChang2016_sim_li.tif",
overwrite=TRUE)
writeRaster(clues.model@output@layers[[2]],"BanChang2017_sim_li.tif",
overwrite=TRUE)
writeRaster(clues.model@output@layers[[3]],"BanChang2018_sim_li.tif",
overwrite=TRUE)
writeRaster(clues.model@output@layers[[4]],"BanChang2019_sim_li.tif",
overwrite=TRUE)
writeRaster(clues.model@output@layers[[5]],"BanChang2020_sim_li.tif",
overwrite=TRUE)

## Accuracy test ##
BanChang2017_s <- raster("BanChang2017_sim.tif")
obs.BanChangs <- ObsLulcRasterStack(x=stack(BanChang2017,BanChang2017_s),
                    categories=c(1,2),
                    labels=c("Other","Built"),
                    t=c(0,1))

## Accuracy test (linear interpolation) ##
BanChang2017_s <- raster("BanChang2017_sim_li.tif")
obs.BanChangs <- ObsLulcRasterStack(x=stack(BanChang2017,BanChang2017_s),
                    categories=c(1,2),
                    labels=c("Other","Built"),
                    t=c(0,1))
```

Table I.2

Source Code for Land-Use Change Simulation Between 2017-2020 (Cont.)

```
## Plot land-use map
#plot(obs.BanChangs)
## Produce transition matrix
crossTabulate(obs.BanChangs, times=c(0,1))

# Accuracy test was performed in excel

# Plot simulated land-use map for presentation
simulated_lu_2017 <- raster("BanChang2017_sim.tif")
simulated_lu_2018 <- raster("BanChang2018_sim.tif")
simulated_lu_2019 <- raster("BanChang2019_sim.tif")
simulated_lu_2020 <- raster("BanChang2020_sim.tif")

names(simulated_lu_2017) <- "simulated_lu_2017"
names(simulated_lu_2018) <- "simulated_lu_2018"
names(simulated_lu_2019) <- "simulated_lu_2019"
names(simulated_lu_2020) <- "simulated_lu_2020"

sim.BanChang <-
ObsLulcRasterStack(x=stack(simulated_lu_2017,simulated_lu_2018,
simulated_lu_2019, simulated_lu_2020),
                    categories=c(1,2),
                    labels=c("Other","Built"),
                    t=c(0,1,2,3))
plot(sim.BanChang)

## Export to plot suitability map
write.csv(suitability.maps@coords,"table_coords.csv")
write.csv(suitability.maps@data[["Other"]],"table_other.csv")
```

Note. This table is an example of the author's source code. Contact the author via email in the biography section for the entire script.

Table I.3

Source Code for Land-Use Change Projection in 2021

```
## Import all necessary packages
library(lulcc)
library(gsubfn)
library(caret)
library(rgdal)
library(randomForest)
#library(rpart)
library(Hmisc)
library(raster)
library(rasterVis)

## Expand memory allowance capacity to maximum
# memory.limit(size=56000)

## Get directory
getwd()
setwd("/cloud/project/BanChang_mod")

## Import land-use map of study area
BanChang2016 <- raster("BanChang2016_u.tif")
BanChang2017 <- raster("BanChang2017_sim_full.tif")
BanChang2018 <- raster("BanChang2018_sim_full.tif")
BanChang2019 <- raster("BanChang2019_sim_full.tif")
BanChang2020 <- raster("BanChang2020_sim_full.tif")

## Import driving factors of the study area
## bcn_2.tif = slope, bcn_3.tif = elevation, bcn_10.tif = distance to urban settlement
in 2020
ef_2_b <- raster("bcn_2.tif")
ef_3_b <- raster("bcn_3.tif")
ef_4_b <- raster("bcn_10.tif")

## Resample to produce maps with identical extent and resolution
ef_2 <- resample(ef_2_b,BanChang2016,method='bilinear')
ef_3 <- resample(ef_3_b,BanChang2016,method='bilinear')
ef_4 <- resample(ef_4_b,BanChang2016,method='bilinear')

## Stack all maps
names(BanChang2019) <- "lu_2019"
names(BanChang2020) <- "lu_2020"
BanChang.map <- stack(BanChang2019,BanChang2020,ef_2,ef_3,ef_4)
na.omit(BanChang.map)
```

Table I.3

Source Code for Land-Use Change Projection in 2021 (Cont.)

```
obs.BanChang <- ObsLulcRasterStack(x=stack(BanChang2019,BanChang2020),
categories=c(1,2),
labels=c("Other","Built"),
t=c(0,1))
obs.BanChang.no.na <- na.omit(obs.BanChang)
obs.BanChang.no.na

## Plot land-use map
plot(obs.BanChang.no.na)
## Produce transition matrix
crossTabulate(obs.BanChang.no.na, times=c(0,1))

## Stack explanatory variables
ev <- ExpVarRasterList(x=BanChang.map, pattern="bcn")
ev.no.na <- na.omit(ev)
ev.no.na

## Partition data to train the model
part.BanChang<- partition(x=obs.BanChang.no.na[[1]], size=0.7, spatial=TRUE)

train.data <- getPredictiveModelInputData(obs=obs.BanChang.no.na, ef=ev.no.na,
cells=part.BanChang[["train"]])
train.data.no.na <- na.omit(train.data)
all.data <- getPredictiveModelInputData(obs=obs.BanChang.no.na, ef=ev.no.na,
cells=part.BanChang[["all"]])

forms <- list(Other ~ bcn_2+bcn_3+bcn_10,
Built ~ bcn_2+bcn_3+bcn_10)

glm.models <- glmModels(formula=forms, family=binomial, data=train.data.no.na,
obs=obs.BanChang.no.na)

## Create suitability maps
suitability.maps <- predict(object=glm.models, newdata=all.data, data.frame=TRUE)
points <- rasterToPoints(obs.BanChang.no.na[[1]], spatial=TRUE)
suitability.maps <- SpatialPointsDataFrame(coords=points, data=suitability.maps)
#r <- stack(rasterize(x=suitability.maps, y=obs.BanChang.no.na[[1]],
field=names(suitability.maps)))
#plot(r)
```

Table I.3

Source Code for Land-Use Change Projection in 2021 (Cont.)

```
## Test ability of models to predict allocation of built and other
## Land uses in testing partition
test.data <- getPredictiveModelInputData(obs=obs.BanChang.no.na, ef=ev.no.na,
cells=part.BanChang[["test"]])
test.data.no.na <- na.omit(test.data)

glm.pred <- PredictionList(models=glm.models, newdata=test.data.no.na)
glm.perf <- PerformanceList(pred=glm.pred, measure="rch")

plot(list(glm=glm.perf))

## Test ability of models to predict location of urban gain 2019 to 2020
part.1 <- rasterToPoints(obs.BanChang.no.na[[1]], fun=function(x) x != 2,
spatial=TRUE)
test.data.1 <- getPredictiveModelInputData(obs=obs.BanChang.no.na, ef=ev.no.na,
cells=part.1, t=1)
test.data.1.no.na <- na.omit(test.data.1)
glm.pred.1 <- PredictionList(models=glm.models[[2]], newdata=test.data.1.no.na)
glm.perf.1 <- PerformanceList(pred=glm.pred.1, measure="rch")
plot(list(glm=glm.perf.1))

## Obtain demand scenario via linear interpolation
## Some routines are coupled to complex economic models that predict future or
past land use demand based on economic considerations
## However, linear extrapolation of trends remains a useful technique
dmd <- approxExtrapDemand(obs=obs.BanChang.no.na, tout=0:1)

## Note that this package is not equipped with advanced demand forecasting
techniques
## However, users could create matrix that specifies future land use demand
obtained from the simultaneous equation
## Use this command to modify the dmd matrix according to result from
simultaneous equation
dmd[c(1),] <- c(121441 ,61809)
dmd[c(2),] <- c(120126 ,63124)


## To see example of modified dmd matrix run the following command
dmd
```

Table I.3

Source Code for Land-Use Change Projection in 2021 (Cont.)

```
## Plot change
matplot(dmd, type="l", ylab="Demand (no. of cells)", xlab="Time point",
lty=1, col=c("Green","Blue","Brown","Red"))
legend("topleft", legend=obs.BanChang.no.na@labels,
col=c("Green","Blue","darkgoldenrod2","Red"), lty=1)

## Create CLUE-S model object
clues.rules <- matrix(data=1, nrow=2, ncol=2, byrow=TRUE)
clues.parms <- list(jitter.f=0.0002,
scale.f=0.000001,
max.iter=1000,
max.diff=50,
ave.diff=50)

clues.model <- CluesModel(obs=obs.BanChang.no.na,
ef=ev.no.na,
models=glm.models,
time=0:1,
demand=dmd,
elas=c(0.2,0.2),
rules=clues.rules,
params=clues.parms)

## Perform allocation
clues.model <- allocate(clues.model)
plot(clues.model)

## Export result from CLUE-S model
#writeRaster(clues.model@output@layers[[1]],"BanChang2020_sim_full.tif",
overwrite=TRUE)
writeRaster(clues.model@output@layers[[2]],"BanChang2021_sim_full.tif",
overwrite=TRUE)

## Accuracy test ##
BanChang2021 <- raster("BanChang2021_u.tif")
BanChang2021_s_f <- raster("BanChang2021_sim_full.tif")
obs.BanChangs <- ObsLulcRasterStack(x=stack(BanChang2021,BanChang2021_s_f),
categories=c(1,2),
labels=c("Other","Built"),
t=c(0,1))
```

Table I.3

Source Code for Land-Use Change Projection in 2021 (Cont.)

```
## Plot land-use map
#plot(obs.BanChangs)
## Produce transition matrix
crossTabulate(obs.BanChangs, times=c(0,1))

# Accuracy test was performed in excel

## Create suitability map 2020
write.csv(suitability.maps@coords,"table_coords_2020.csv")
write.csv(suitability.maps@data[["Other"]],"table_other_2020.csv")
```

Note. This table is an example of the author's source code. Contact the author via email in the biography section for the entire script.

BIOGRAPHY

Name	Mr. Nutchapon Prasertsoong
Educational Attainment	2017: Bachelor of Economics, Thammasat University (Rangsit Campus)
Scholarship (If any)	Year 2021-2023: Thammasat University (Research Fund for Doctoral Students)

Publications

Prasertsoong, N., & Puttanapong, N. (2022). Regional Wage Differences and Agglomeration Externalities: Micro Evidence from Thai Manufacturing Workers. *Economies, 10*(12). doi:10.3390/economies10120319