Three Essays on Spatial Economic Analysis in Mexico

Kondo, Keisuke

(Degree)
博士（経済学）
(Date of Degree)
2014-03-25
(Date of Publication)
2015-03-01
(Resource Type)
doctoral thesis
(Report Number)
甲第6141号
(URL)
https://hdl.handle.net/20.500.14094/D1006141

博士論文

平成 25 年 12 月

神戸大学大学院経済学研究科
経済学専攻
指導教員　浜口伸明
近藤恵介

博士論文

Three Essays on Spatial Economic Analysis in Mexico

「メキシコにおける空間経済分析に関する 3 編の論文」

平成 25 年 12 月

神戸大学大学院経済学研究科
経済学専攻
指導教員　浜口伸明
近藤恵介

Doctoral Dissertation

Three Essays on Spatial Economic Analysis in Mexico

Keisuke Kondo
Graduate School of Economics, Kobe University

December 2013

Contents

List of Tables

List of Figures

Acknowledgements

It would not have been possible to write this doctoral dissertation without the help and support of the many kind people around me. Here, I would like to express my deepest gratitude to some of those who, in particular, have supported my doctoral research and academic career.

First and foremost, I would like to thank my advisor, Professor Nobuaki Hamaguchi at the Research Institute for Economics and Business Administration of Kobe University, for patiently giving me helpful advice, appropriate direction, and encouragement. My research interests span a wide range of topics, from spatial economics and spatial econometrics, to development economics, labor economics, and Latin American studies. Despite this broad focus, he has continuously supported my doctoral research with proper guidance based on his wide and deep knowledge. I could not have written this dissertation without his continuous support.

My special thanks are extended to my advisor at the Instituto Tecnológico Autónomo de México (ITAM), Professor Kensuke Teshima. He offered me an excellent opportunity to engage in research at the ITAM, and his invaluable help enabled me to complete the research project. I was greatly inspired by weekly seminars and workshops. Furthermore, I was able to deepen my research ideas by engaging in vigorous discussions with him.

I would also like to thank Professor Alfredo Erquizio Espinal at the University of Sonora, Mexico for his helpful comments and suggestions on Chapter 2 of this dissertation. He was willing to accept me when I visited him, kindly explained important issues on Mexico's business cycles, and provided me with his original data and information. I am deeply indebted to him for the information presented in Chapter 2 of this dissertation.

I am grateful to the co-directors of the Ph.D. Program in Economics at the ITAM, Professor Andrei Gomberg and Professor Joyce Sadka, for enabling me to visit the ITAM for my doctoral research. Indeed, my dissertation is based on the research carried out in Mexico. They kindly helped me whenever I had questions about student affairs at the

ITAM. I would also like to thank Professor Jorge Alonso Ortiz at the ITAM for listening to my questions about the research and giving me proper suggestions. Furthermore, I would like to thank the librarians of the ITAM, Mr. Luis Felipe Mújica Martínez and Mr. Josbel Rockdrigo González Lora, for their behind-the-scenes support. I used to study all the day at the library, and they always encouraged me to achieve my aims. They greeted me with a smile on every occasion we met. I also thank Ms. Luz María Martínez Sánchez, Ms. Irma Cervantes Aguilar, and Ms. Guadalupe Romero Pineda, who work for the department of student affairs at the ITAM. During my frequent visits to request the necessary documents for the Mexican government scholarship, they took care of me with kindness and patience. I express my gratitude to all those who helped and supported me during my stay at the ITAM.

My doctoral research was enabled by a scholarship granted by the Government of Mexico through the Ministry of Foreign Affairs. I extend my sincere thanks to them for this financial support, without which this dissertation would not have been possible. I also thank the staff of the Embassy of Mexico in Japan, Ms. Kayo Shimodaira and Ms. Emy Kameta, for offering me continuous support in tackling the procedures of the scholarship and student visa until my departure for Mexico. Furthermore, I thank a staff of the Ministry of Foreign Affairs of Mexico, Ms. Verónica de Jesús Méndez, for supporting me during my stay in Mexico.

I would like to thank Professor Hisashi Tanizaki at Osaka University and Professor Akio Namba, a member of my dissertation committee at Kobe University for allowing me to participate in their respective seminars. I was able to expand my knowledge of econometric methods through their helpful advice. Thanks to them, I was able to employ advanced econometric methods such as spatial econometrics in Chapters 1 and 2, and Bayesian econometrics in Chapter 2. Furthermore, I would like to thank Professor Yoichi Matsubayashi, also a member of my dissertation committee, for listening very carefully to my presentations of the doctoral research and giving valuable comments and suggestions. I am also grateful to Professor Toshihiro Okubo at Keio University, a co-author of my first published paper in an academic journal (Keisuke Kondo and Toshihiro Okubo, "Interregional labour migration and real wage disparities: Evidence from Japan," *Papers in Regional Science*, forthcoming). Through working with him, I learned how to develop a research project and write an academic paper, and these experiences were invaluable during the process of researching and writing this doctoral dissertation. Moreover, I would like to thank the staff of the Graduate School of Economics at Kobe University for their wholehearted support throughout my academic career.

I am grateful to my friends with whom I have studied together in the same seminars, Dr. Yang Xi and Dr. Hiroshi Goto, for the many discussions about our research after each seminar. For my research, I took a cue from these experiences, which helped me to develop my research ideas.

Last but not least, I would like to thank my parents for supporting me with great patience throughout my life and particularly as a doctoral student.

Preface

Overview of Doctoral Dissertation

In recent years, the economies of agglomeration have come back into the spotlight not only in developed countries but also in developing countries (e.g., World Bank, 2009). This is because agglomeration is considered to play an important role in fostering economic development at the regional and local levels. Benefits from agglomeration have been known implicitly for a long time, and researchers have attempted to search for the source and nature of agglomeration. For example, Marshall (1890) observes that agglomeration enables better matching between job seekers and firms, better linkages between suppliers of intermediate and final goods, and more active knowledge creation and spillover in agglomerated regions, making regions more attractive and, thus, denser.

This doctoral dissertation sheds light on the role of geographical space in economic activities. It pays special attention to three points of view that are important when considering the role of space in economic activities. First, this dissertation emphasizes the spatial interaction arising from migration and trade. Unlike traditional assumptions of international trade, the mobility of labor plays an important role in endogenously forming a core region of economic activities. This is the same point that Krugman (1991a) focused on particularly in the field known as the New Economic Geography (NEG). Second, this dissertation elucidates spatial dependence across regional economies. It is clear that regional economies are not independent but interdependent. In such a spatial structure, even a region-specific shock affects neighboring economies through spatial spillover effects, which include feedback loops or interplay effects. Third, this dissertation focuses on the benefits from agglomeration and examines the extent to which our economic activities benefit from agglomeration. It is said that agglomeration gives rise to positive and/or negative externalities in myriad ways. These three points correspond to the three chapters of the dissertation, respectively; discussions about space will be further deepened in each context.

The importance of space has mostly been ignored in both mainstream economic and econometric theories. Although the importance of spatial issues itself might have been recognized by a majority of economists, they have had no way to model the spatial aspect of economies (Krugman, 1995). As mentioned in Krugman (1991a), it was necessary to move away from the standard approach that assumed constant returns to scale and perfect competition. Krugman (1991b) indeed overcame these problems and endogenously described how economic activities concentrate in one region by offering a unifying general equilibrium framework with particular attention to increasing returns to scale, monopolistic competition, transport costs, and mobile workers across regions. As a result, his seminal paper opened a new era of the NEG in the early 1990s. Currently, the seminal textbooks of the NEG are easily found (e.g., Fujita et al., 1999; Baldwin et al., 2003; Combes et al., 2008b; Fujita and Thisse, 2013). I am, therefore, greatly indebted to the researchers who have enabled the historical development of spatial economics over the past two centuries, as documented in Fujita (2010).

It should be noticed that little attention has been paid to space in the mainstream econometric theory as well. Even if regional data are used, we have imposed a strict assumption that each regional unit is independent by using random sampling. Unlike individual or household data, independence across observations does not hold, even though we carefully consider random sampling in the regional data, which finally leads to inconsistent and/or inefficient estimators. To tackle these issues, many empirical economists and econometricians are currently paying more attention to spatial econometrics. Anselin (2010) gives an overview of the evolution of spatial econometrics over the past three decades, concluding that there is room for the further development of spatial econometrics. In this doctoral dissertation, I make extensive use of spatial econometric techniques.

This dissertation analyzes the Mexican economy. As mentioned in Krugman and Livas-Elizondo (1996), Mexico has been experiencing a dynamic change in economic activities since trade liberalization in the 1980s and 1990s. This has given rise to drastic changes in the country's domestic distribution patterns of firm location, employment, and industrial structure. Regional and local economies have displayed grater interdependence with increasing migration, improving infrastructure, and ongoing globalization. Space is, therefore, a key and essential factor for a better understanding of the current Mexican economy. Throughout this doctoral dissertation, I attempt to uncover how space affects economic activities in the Mexican economy.

An additional advantage of focusing on the Mexico is that the Mexican government has improved statistical data since the early 2000s. Even micro-data are open to researchers from around the world. Indeed, I use a micro-data set of Mexican workers in Chapter 3. The Mexican government intends to make good use of the results of research conducted by experts, and I hope that my doctoral research makes a significant contribution to their policy making processes.

Abstracts of Chapters

This doctoral dissertation consists of three chapters. Chapter 1 deals with spatial interaction arising from migration and trade. Chapter 2 focuses on spatial dependence in regional business cycles and spatial spillover effects originating from region-specific shocks. Chapter 3 explores how workers employed in denser spaces benefit from agglomeration when firms have branch networks. Space is a keyword throughout this doctoral dissertation. Detailed summaries of the chapters follow.

Chapter 1 theoretically and empirically analyzes the relationship between regional unemployment rates and agglomeration by introducing the standard search and matching framework into a new economic geography model. Furthermore, we incorporate agglomeration externalities into a search and matching framework. After our theoretical analysis, we empirically examine relationships between regional unemployment rates and agglomeration and between matching efficiency and agglomeration by using Mexican data. An important prediction of our theory is that regional unemployment rates can be positively or negatively correlated with agglomeration under negative agglomeration externalities on matching efficiency. We empirically find that denser areas have comparatively low unemployment rates even under negative agglomeration externalities on matching efficiency. Considering our theoretical predictions, we conclude that in Mexico, the agglomeration effect lowering the unemployment rates is much stronger than that increasing the rates.

Chapter 2 investigates how a region-specific shock propagates outward toward neighboring regions when regional business cycles are spatially dependent. For this purpose, we model business cycles by introducing a spatial autoregressive process into a Markov switching model. The advantage of this model is that it enabled us to numerically simulate spatial spillover effects. Therefore, we were able to demonstrate how the economic crisis in 2008–2009 spread across Mexican states. We find that business cycles across these states were spatially dependent and that a regime switch from expansion to recession caused conditions in the neighboring economies to deteriorate.

Chapter 3 studies how workers earn higher wages in denser areas using micro-data for Mexican labor. Previous studies show that higher population density raises wages. How-

ever, it has been unclear how workers receive benefits from agglomeration when firms have branch networks. To clarify this issue, we focus on branches of the Mexican commercial banks. If workers directly benefit from local agglomeration, higher wages are paid in branches located in denser areas. If workers indirectly benefit from agglomeration through branch networks, the banks' fixed effects are positively correlated with population density. Furthermore, we analyze heterogeneous effects of agglomeration on wages between high- and low-skilled workers. We find that banks locating branches in denser areas tend to pay workers better, suggesting that workers indirectly benefit from agglomeration through banks' branch networks. Thus, branch networks would provide additional agglomeration effects, especially for workers employed in less dense areas. We find that high-skilled workers are likely to enjoy direct benefits from local agglomeration, whereas low-skilled workers are not within the banks' branch networks. To make matters worse, low-skilled workers can be affected negatively by agglomeration owing to congestion effects.

Chapter 1

Regional Unemployment Rates in an Agglomeration Economy: A Theoretical and Empirical Analysis*

1.1 Introduction

Since the publication of Krugman (1991b), new economic geography (NEG) studies have examined the agglomeration mechanism of economic activities, with particular attention to the increasing returns to scale, monopolistic competition, transport costs, and mobile labor across regions (e.g., Fujita et al., 1999). With regarding to regional labor markets and agglomeration, Marshall (1890) observes that the concentration of economic activities facilitates the job search and matching between employers and job seekers in terms of industry-specific skills. Similarly, Rosenthal and Strange (2001), investigating the determinants of agglomeration, find that labor market pooling fosters agglomeration. Despite such observations and studies, only limited attention has been paid to job search and regional unemployment issues in the NEG literature.[1] Thus, we do not fully understand the underlying mechanism acting between regional unemployment rates and the agglomeration of

*I would like to specially thank Kensuke Teshima and Yasuhiro Sato for their very helpful comments and suggestions. I also thank Jorge Alonso, Hiroshi Goto, Nobuaki Hamaguchi, Naoto Jinji, Hisaki Kono, Toshihiro Okubo, Mun Se-il, Koji Shintaku, and all the participants in the CIE Brown Bag Seminar at the Instituto Tecnológico Autónomo de México (ITAM), the third spring meeting of the Japan Society of International Economics at Fukuoka University, and the Brown Bag Lunch Seminar at Kyoto University for their useful comments and suggestions. Any remaining errors are naturally my own. I am grateful to those at the ITAM for all the support that I received during my stay there. This research was carried out under a scholarship granted by the Government of Mexico through the Ministry of Foreign Affairs of Mexico.

[1]Note that the NEG literature has also contributed to uncovering wage inequality from the perspective of geographical networks. For example, many empirical papers have shown that market potential leads to higher regional nominal wages (e.g., Redding and Venables, 2004; Hanson, 2005; Hering and Poncet, 2010). However, these studies are based on theoretical models under perfectly competitive labor markets.

economic activities.

In the recent NEG literature, attempts have been made to tackle job search and unemployment issues.[2] For example, Epifani and Gancia (2005) and Francis (2009) developed a dynamic NEG model by introducing a search and matching mechanism.[3] Their models predict a lower unemployment rate in agglomerated regions in the long-run. On the other hand, motivated by that unemployment rates in high-density regions seem to be higher than in low-density regions from developed countries data, vom Berge (forthcoming) extended Krugman's (1991) model by introducing a search and matching framework.[4] His model shows that the unemployment rates in agglomerated regions are comparatively high.[5] However, as mentioned by Zierahn (forthcoming), when NEG models show full agglomeration under the spatial equilibrium, it indicates that unemployed workers do not live in the periphery region.[6] That is, under full agglomeration, the unemployment rate in the periphery region virtually becomes zero (or cannot be defined), whereas it is always positive in the agglomerated region. As such, the results obtained from *full* agglomeration models cannot exactly capture situations in the periphery regions. Therefore, we investigate the relationship between regional unemployment rates and agglomeration by using an NEG model with *partial* agglomeration.

Following the framework proposed by vom Berge (forthcoming), we develop a multi-region model of Helpman (1998) by incorporating a search and matching mechanism.[7] Unlike Krugman (1991b), Helpman (1998) lays more emphasis on the dispersion force arising from non-tradable local services. For example, the concentration of economic activities raises the prices of land and housing owing to the increased demand for them. Consequently, this type of dispersion force leads to partial agglomeration. Thus, focusing on Helpman's (1998) model, we offer fresh insight into the regional distribution of unemployment rates in

[2]Some theoretical mechanisms that generate unemployment need to be introduced (e.g., efficiency wage or search and matching frameworks). This paper employs the search and matching model proposed by Pissarides (2000). Rogerson et al. (2005) offer a review of this literature. In the literature of international trade, Helpman and Itskhoki (2010) developed an international trade model to analyze the effect of labor market rigidity on trade flow. Unlike those studies, we focus on the trade model dealing with migration between regions.

[3]Unlike the search and matching framework, Zierahn (forthcoming) introduces the efficiency wage and congestion costs due to agglomeration into Krugman's (1991) model.

[4]Unlike vom Berge (forthcoming), we find both positive and negative relationships between unemployment and agglomeration, expressed as population size or population density in empirical studies. See for example Simon (1988), Izraeli and Murphy (2003), and Chiang (2009).

[5]vom Berge (forthcoming) introduces regions into the model developed by Ziesemer (2005), who extended Pissarides (2000, Chap. 3) model by introducing monopolistic competition.

[6]Agricultural workers still live there in the case of Krugman-type models.

[7]An extension of Helpman (1998) can be found in Pflüger and Tabuchi (2010). They assume that a firm uses land as a production input.

an agglomeration economy. Furthermore, to analyze how transport costs affect the relationship between regional unemployment rates and agglomeration, we carry out a numerical analysis of the theoretical model.[8]

A contribution of this paper is to incorporate agglomeration externalities into a search and matching framework. As observed in Marshall (1890), denser areas seem to promote job matching between job seekers and firms. However, this is not necessarily true in the current economy. Recent empirical studies provide two contradictory evidences. Hynninen and Lahtonen (2007) show a positive relationship between matching efficiency and population density, whereas Kano and Ohta (2005) show a negative one. Therefore, our theoretical model assumes positive or negative agglomeration externalities on matching efficiency. Consequently, our model is able to describe a wide variety of relationships between unemployment rates and agglomeration.

Our study also contributes to the literatures of development economics and wage curve. Beginning with Harris and Todaro (1970), the literature of development economics has studied urban unemployment and migration. Given the exogenously high wage in urban area, Harris and Todaro (1970) showed that urban unemployment rate increases on account of excessive workers immigrating into a city in response to higher expected wage. Therefore, a positive relationship between wages and unemployment rates can be expected.[9] In contrast, the literature of the wage curve, beginning with Blanchflower and Oswald (1994), has studied the negative relationship between regional wages and unemployment rates.[10] Our theoretical model therefore attempts to uncover this contradictory observation.

We specifically describe three relationships between nominal wages, unemployment rates, and agglomeration across regions, as clearly illustrated in Figure 1.1. Note that a consensus already exists on the positive relationship between wages and agglomeration (e.g., Combes et al., 2008a; Mion and Naticchioni, 2009; Combes et al., 2010; de la Roca and Puga, 2012), and this always holds in our model. In addition, previous studies show that agglomeration has a decreasing effect on unemployment rates in the production side.[11] Further, if agglomeration is assumed to have positive/negative externalities on matching

[8] Although NEG models provide insightful policy implications, their theoretical and numerical analyses are usually limited to two-region cases to avoid mathematical difficulties, which are also known as *three-ness* (Combes et al., 2008b, Chap. 4). Although we build a multi-region model for the theoretical part of our study, our numerical analysis is restricted to a case of two symmetric regions.

[9] Contrary to the prediction of Harris and Todaro (1970), Suedekum (2005) showed a lower unemployment rate and higher wage in agglomerated region by endogenously expressing higher urban wage within the NEG framework.

[10] See Card (1995) for a literature review of the wage curve.

[11] See also Suedekum (2005) and Zierahn (forthcoming).

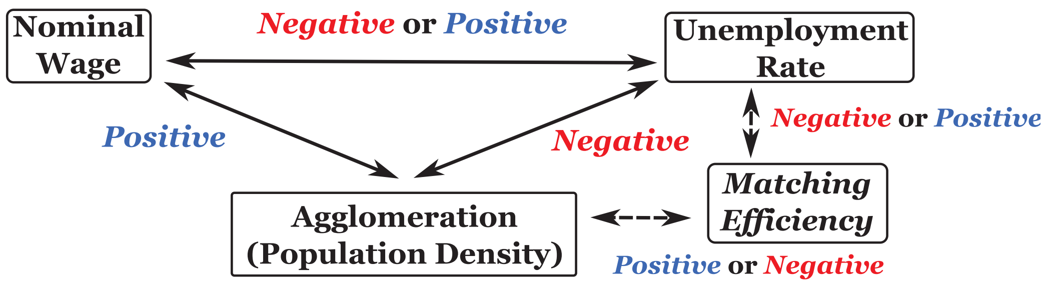


Figure 1.1: Relationship between Wage, Unemployment Rate, and Agglomeration

efficiency, it would also lead to negative/positive effects on the unemployment rate. Consequently, the advantage of our model is that we explain both the positive and negative relationships between nominal wages and unemployment rates, while also endogenously explaining the higher wages in agglomerated regions. Therefore, we believe that this paper makes a valuable contribution to the Harris–Todaro model and the wage curve literature.

This paper also includes an empirical analysis of the relationship between regional unemployment rates and agglomeration. We use Mexican municipal data and control for spatial dependence within the municipal data by using spatial econometric methods. We also estimate the matching function to examine the relationship between matching efficiency and agglomeration. Finally, we draw a conclusion about the relationship between unemployment rates and agglomeration by taking into account both the estimation results.

As mentioned in Krugman and Livas-Elizondo (1996), Mexico has experienced a dynamic change in economic activities since the trade liberalization in the 1980s and 1990s. In the meantime, this movement brought about drastic changes in the country's domestic distributional employment pattern. According to Hanson (1998), the Mexico–US border states attracted more manufacturing workers. For example, Hanson (1998) shows that the share of regional employment in the Mexico-US border states was 21.0% in 1980 but 29.8% in 1993. On the other hand, the manufacturing workers tend to leave the Mexico City metropolitan area (their share came down from 46.4% in 1980 to 28.7% in 1993). However, little attention has been paid to the relationship between regional unemployment rates and agglomeration in the Mexican literature; therefore, we try to examine whether the agglomerated regions have higher or lower unemployment rates.

The remainder of this paper is organized as follows. Section 2 builds a multi-region model of Helpman (1998) consisting of a standard search and matching framework. Section 3 numerically analyzes a case of two symmetric regions. Section 4 details the empirical strategy used for this study. Section 5 explains the data used. Section 6 discusses the estimation results. Finally, Section 7 concludes.

1.2 The Model

Following vom Berge (forthcoming), we extend the multi-region model of Helpman (1998) by introducing a search and matching framework. We consider an economy with R regions having both manufacturing and land/housing sectors. The manufacturing sector is monopolistically competitive, and each firm produces one variety of a differentiated good under increasing returns to scale. Labor is a unique production input. On the other hand, the land/housing sector is perfectly competitive; land endowment in each region is fixed, so that the supply of land/housing services is also given; consumers have their own land equally. There are two types of workers, the employed and the unemployed. We assume that both types of the worker are mobile across regions in the long-run, and that there are no migration costs. We introduce job search and matching frictions into the regional labor markets. Unemployed workers search for jobs in their own living regions, and spatial job search is not allowed. For the present purpose, we focus on steady state analysis.

1.2.1 Matching Function

We first assume that there are search and matching frictions in the regional labor markets. The number of matches existing between the job seekers and vacancies is determined by the following matching function:

$$m_i L_i = A_i m(u_i L_i, v_i L_i), \quad i = 1, 2, \ldots, R \tag{1.1}$$

where m_i is the matching rate, u_i is the unemployment rate, v_i is the vacancy rate in terms of labor, A_i is the matching efficiency, and L_i is the labor force, with the subscript i indicating region i. Note that job matches are made only within region i. We further assume that the matching function is increasing in both variables, homogeneous of degree one, concave, and twice continuously differentiable, and that $m(u_i L_i, 0) = m(0, v_i L_i) = 0$.[12] As mentioned earlier, we assume that agglomeration of economic activity has externalities on the matching efficiency A_i; our specification is as follows:

$$A_i = A(L_i/\bar{S}_i)^{\xi}, \tag{1.2}$$

[12] See Petrongolo and Pissarides (2001) for details of the matching function, including empirical findings.

where A is constant, $\bar{S}_i$ represents land endowment (or fixed supply of land/housing services), and ξ is the elasticity of agglomeration to matching efficiency. Thus, $L_i/\bar{S}_i$ can be interpreted as a kind of population density in region i.

Given the matching function (1.1), the rates at which vacancies are filled and unemployed workers leave unemployment can be expressed respectively as

$$q_i(\theta_i) \equiv \frac{A_i m(u_i L_i, v_i L_i)}{v_i L_i} \quad \text{and} \quad \theta_i q_i(\theta_i) \equiv \frac{A_i m(u_i L_i, v_i L_i)}{u_i L_i},$$

where $\theta_i \equiv v_i/u_i$ denotes the labor market tightness. From the above assumptions, we can easily verify that both $q_i(\theta_i) > 0$ and $q_i'(\theta_i) < 0$ hold for a given value of A_i.

1.2.2 Consumer and Worker

For simplicity, we assume a static consumer problem; consumers do not save any part of their income but spend all of it in each period.[13]

Further, each consumer has identical Cobb–Douglas preferences for two goods; that is,

$$\mathbb{U}_i = \frac{1}{\mu^\mu (1-\mu)^{1-\mu}} M_i^\mu H_i^{1-\mu}, \tag{1.3}$$

where $0 < \mu < 1$ is the expenditure share for manufactured goods, M_i is the composite consumption of manufactured goods in region i, and H_i is the consumption of land/housing services in region i.[14] The composite consumption of manufactured goods is given by the constant elasticity of substitution (CES) function

$$M_i = \left(\sum_{j=1}^{R} \int_0^{n_j} m_{ji}(\nu)^{(\sigma-1)/\sigma} \mathrm{d}\nu \right)^{\sigma/(\sigma-1)},$$

where $m_{ji}(\nu)$ is region i's consumption of variety ν produced in region j, n_j the number of varieties produced in region j, and $\sigma > 1$ the elasticity of substitution between any two varieties. The budget constraint of region i is given by $G_i M_i + p_i^H H_i = Y_i$, where G_i is the price index for manufactured goods, p_i^H is the price of land/housing services, and Y_i is the regional income.

From utility maximization, we obtain the following demand functions:

$$H_i = \frac{(1-\mu)Y_i}{p_i^H}, \quad M_i = \frac{\mu Y_i}{G_i}, \quad \text{and} \quad m_{ji}(\nu) = \mu p_{ji}(\nu)^{-\sigma} G_i^{\sigma-1} Y_i, \tag{1.4}$$

[13]This simplification, however, does not change the essential results of our model.
[14]We modify the methodology of Pflüger and Tabuchi (2010) to describe a land/housing market.

where $p_{ji}(\nu)$ is region i's consumer price for variety ν imported from region j; the price index in region i takes the following form:

$$G_i = \left(\sum_{j=1}^{R} \int_0^{n_j} p_{ji}(\nu)^{1-\sigma} \mathrm{d}\nu \right)^{1/(1-\sigma)}. \tag{1.5}$$

By substituting demand functions (1.4) into utility function (1.3), we obtain the indirect utility $\mathbb{V}_i$ of an individual living in region i:

$$\mathbb{V}_i = \frac{I_i}{G_i^{\mu}(p_i^H)^{1-\mu}}, \tag{1.6}$$

where I_i is the income of the individual living in region i. Indirect utility can be interpreted as the real income, that is, the individual's income I_i deflated by the cost-of-living index $G_i^{\mu}(p_i^H)^{1-\mu}$.

As mentioned earlier, there are two types of workers in the economy, the employed and the unemployed. Let $\mathbb{V}_i^e$ and $\mathbb{V}_i^u$ denote the indirect utilities of the employed and the unemployed, respectively. We assume that while the employed earns w_i, the unemployed receives unemployment benefit z from the government. The unemployment benefit is exogenously given. The government imposes a tax τ on all the workers in order to finance the unemployment benefits. Further, we assume that the rate of interest r is common across all regions. Thus, the steady state Bellman equations for the employed and the unemployed are, respectively, given as follows:

$$\begin{aligned} rE_i &= \mathbb{V}_i^e + \delta(U_i - E_i), \\ rU_i &= \mathbb{V}_i^u + \theta_i q_i(\theta_i)(E_i - U_i), \end{aligned} \tag{1.7}$$

where E_i and U_i are the present discounted values (PDV) of the expected real income stream for the employed and the unemployed, respectively, and δ is the job destruction rate. In the long-run, individuals decide to migrate depending on the expected PDV from continuing to live in the region.

1.2.3 Producer Behavior

We assume that the prices of all the varieties produced within a region are identical in view of the same production technology used and therefore denote the price of all the varieties produced in region i as p_i. We assume that a manufactured good is traded between regions i and j with iceberg transport cost T_{ij}. Thus, if one unit of any variety of manufactured goods is shipped from region i to region j, only $1/T_{ij}$ of the unit arrives. A variety of manufactured goods produced in region i is sold at price p_i in that region. If this variety is shipped from region i to region j, the delivered price is

$$p_{ij} = p_i T_{ij}, \quad T_{ij} = T_{ji} \geq 1, \quad T_{ii} = 1, \quad i, j = 1, 2, \ldots, R.$$

The total amount of goods that a firm produces to satisfy the consumption demand of all the regions therefore becomes

$$x_i = \sum_{j=1}^{R} m_{ij} T_{ij}. \tag{1.8}$$

Next, all the firms require not only fixed and marginal labor input for producing the varieties but also recruiters for hiring their workers.[15] Thus, the total labor input in region i is

$$\ell_i = F + c x_i + \gamma N_i \tag{1.9}$$

where F and c are respectively the fixed and marginal labor requirements for production, γ is the marginal labor requirement for recruiting per vacancy, and N_i is the number of vacancies that a firm needs to post. The first two terms correspond to the standard Dixit–Stiglitz assumption of increasing returns to scale. The third term indicates that a firm needs to hire recruiters to keep their workers from decreasing because the workers quit their jobs at a job destruction rate of δ. The same wage w_i is paid to both workers and recruiters. The total cost is therefore $w_i \ell_i$.

A vacant job is filled with a probability of $q_i(\theta_i)$, and an occupied job is destructed with a probability of δ. Thus, the dynamics of total labor input is given by

$$\dot{\ell}_i = q_i(\theta_i) N_i - \delta \ell_i. \tag{1.10}$$

A firm maximizes the PDV of its expected profit with respect to the produced quantity x_i and number of vacancies N_i as follows:[16]

$$\begin{aligned} &\max_{x_i, N_i} \int_0^\infty e^{-rt} \left[p_i(x_i) x_i - w_i(F + c x_i + \gamma N_i) \right] \mathrm{d}t \\ &\text{s.t.} \quad \dot{x}_i = \frac{1}{c} \left[\left(q_i(\theta_i) - \gamma\delta \right) N_i - \delta(F + c x_i) \right] \\ &\lim_{t \to \infty} \left[\lambda(t) e^{-rt} x_i(t) \right] = 0 \end{aligned} \tag{1.11}$$

[15] This formulation is developed by vom Berge (forthcoming), following Pissarides (2000, Chap. 3) and Ziesemer (2005).

[16] Using (1.9), (1.10), and the envelop theorem, we obtain the dynamic equation on production $\dot{x}$ in (1.11).

where $p_i(x_i)$ is the mill price in region i, and $\lambda(t)$ the Lagrange multiplier. Solving the current value Hamiltonian, we obtain the optimal mill price with a constant markup on marginal costs as follows:

$$p_i = \frac{\sigma}{\sigma - 1} c w_i \left(1 + \frac{r\gamma}{q_i(\theta_i)}\right)\left(1 - \frac{\gamma\delta}{q_i(\theta_i)}\right)^{-1}. \tag{1.12}$$

Note that this price is higher than that of the standard Dixit–Stiglitz monopolistic competition model because the multiplication of the second and third terms is greater than one. Intuitively, the marginal cost consists of three parts. The first two terms give the workers' wage for producing the additional quantity x_i and the expected cost of hiring a worker, and the third term captures the cost of hiring the workers engaged in production and recruitment.[17] If the job search cost is zero ($\gamma = 0$), this price takes the same form obtained for the standard Dixit–Stiglitz model.

Let V_i and J_i be the PDVs of the expected profit from the vacant and occupied jobs respectively. Then, the steady state Bellman equation for a vacancy is given by

$$rV_i = -\gamma \tilde{w}_i + q_i(\theta_i)(J_i - V_i), \tag{1.13}$$

where $\tilde{w}_i \equiv w_i/p_i$ is the real wage defined in terms of firm.

All the profit opportunities from creating new jobs are exploited in equilibrium, and the value of the vacant jobs becomes zero ($V_i = 0$). Hence, the equilibrium condition yields

$$J_i = \frac{\gamma \tilde{w}_i}{q_i(\theta_i)}. \tag{1.14}$$

From this equation, since $1/q_i(\theta_i)$ is the expected duration of a vacant job, the expected profit from a new job is equal to the expected cost of hiring a worker in equilibrium.

1.2.4 Wage Bargaining

In a wage bargaining process, we endogenize the labor market tightness θ_i. Each firm in a standard search and matching model is assumed to have only one job. Although a firm in our model employs multiple workers, we consider the bargaining process in a similar

[17] To understand the third term, we manipulate (1.17) to obtain

$$\frac{\delta(F + cx_i)}{q_i(\theta)N_i} = 1 - \frac{\gamma\delta}{q_i(\theta_i)} < 1.$$

The left-hand side shows how the quitting workers engaged in production are filled up from among the newly hired workers, implying that a part of the newly hired workers are engaged in recruitment.

manner.[18] Following Pissarides (2000, Chap. 3), we assume that the wages of workers are fixed in Nash bargains, in which the firm gets involved with each worker separately, considering the wages of all the other workers as given. This assumption results in a one-to-one relationship between a worker and a job. The total surplus arising from a job match (i.e., the net benefit of the worker and the firm from the unemployed worker starting to work and the firm producing additional goods) is shared through Nash bargaining between the worker and the firm:

$$\tilde{w}_i = \arg\max (E_i - U_i)^{\beta}(J_i - V_i)^{1-\beta},$$

where $0 \leq \beta \leq 1$ is the bargaining power of the workers. From the first-order condition, the result of the bargaining is given by

$$(1-\beta)(E_i - U_i)J_i' = \beta(J_i - V_i)E_i'.$$

By substituting (1.7) and (1.14) and imposing the equilibrium condition $V_i = 0$, we obtain the following equation

$$\tilde{w}_i = rU_i + \beta\left(\frac{\sigma-1}{c\sigma} - rU_i\right).$$

With some manipulations, we obtain the following relationship between the nominal wage and labor market tightness:

$$g_i(w_i, \theta_i) \equiv (1-\beta)\left(1 - \frac{z_i}{w_i}\right) - \beta\frac{\gamma\left[r + \delta + \theta_i q_i(\theta_i)\right]}{q_i(\theta_i) - \gamma\delta} = 0. \tag{1.15}$$

This corresponds to the wage-setting curve in Pissarides (2000), but shows a nonlinear function with regard to labor market tightness and wages in our case. From the implicit function theorem, we obtain

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}w_i} = -\frac{\partial g_i/\partial w_i}{\partial g_i/\partial \theta_i} > 0,$$

where a homogeneous degree one is assumed in the matching function.[19] Since the unemployment rate u_i and labor market tightness θ_i are negatively correlated, this result indicates a negative relationship between wage and unemployment rate.[20]

[18] Stole and Zwiebel (1996a,b) consider an extended version of Nash bargaining for multiple workers, in which the firm and a worker divide the marginal surplus obtained from the firm producing goods by hiring additional worker and the worker leaving the unemployed status. This assumption reflects the case in which additional employment additionally affects the wages of the remaining workers. In this paper, we use a simpler methodology employed by Pissarides (2000, Chap. 3).

[19] Under the assumption of a homogeneous degree one in matching function, we confirm that $q_i(\theta_i) + \theta_i q_i'(\theta_i) > 0$ holds.

[20] This result implies the existence of wage curve (Blanchflower and Oswald, 1994). In case the regional

1.2.5 Short-Run Equilibrium

We now consider a short-run equilibrium, characterized by a general equilibrium in each region without migration.[21] By substituting the price in (1.12) into the current profit in (1.11) and imposing a zero-profit condition, the equilibrium output is given by

$$x_i = \frac{F(\sigma-1)}{c}\left(1+\frac{\sigma r\gamma}{q_i(\theta_i)}\right)^{-1}. \tag{1.16}$$

Note that the equilibrium output is lower than the output of a standard Dixit–Stiglitz monopolistic competition model.

Since $\dot{\ell}_i = 0$ in the steady state, by substituting (1.9), the number of vacancies in the steady state becomes

$$N_i = \frac{\delta(F+cx_i)}{q_i(\theta_i)-\gamma\delta}, \tag{1.17}$$

where we assume $q_i(\theta_i) > \gamma\delta$ so that the number of vacancies takes a positive value. Substituting the equilibrium output (1.16) and the number of vacancies (1.17) into the total labor input (1.9), we obtain the equilibrium total labor input in region i as follows:

$$\ell_i = F\sigma\left(1+\frac{r\gamma}{q_i(\theta_i)}\right)\left(1+\frac{\sigma r\gamma}{q_i(\theta_i)}\right)^{-1}\left(1-\frac{\delta\gamma}{q_i(\theta_i)}\right)^{-1}. \tag{1.18}$$

Further, from the labor market clearing condition $n_i\ell_i = (1-u_i)L_i$, the number of firms is given by

$$n_i = \frac{(1-u_i)L_i}{F\sigma}\left(1+\frac{r\gamma}{q_i(\theta_i)}\right)^{-1}\left(1+\frac{\sigma r\gamma}{q_i(\theta_i)}\right)\left(1-\frac{\delta\gamma}{q_i(\theta_i)}\right). \tag{1.19}$$

From (1.8), the total sales of the variety produced in region i amount to

$$x_i = \mu\sum_{j=1}^{R} p_i^{-\sigma}G_j^{\sigma-1}Y_jT_{ij}^{1-\sigma}. \tag{1.20}$$

Choosing the convenient units of measurement for marginal labor requirement $c = (\sigma-1)/\sigma$ and fixed labor requirement $F = \mu/\sigma$, we simplify the model outcomes. Thus, from (1.12),

---

labor markets are homogeneous with regard to job destruction rates and job matches, a negative correlation could arise between the regional unemployment rates and nominal wages. This result is quite similar to Sato (2000), who shows that even when workers are mobile, the wage curve can be observed by using a theoretical search framework assuming different productivities across the regions and a monocentric city structure.

[21] For ease of expression and interpretation, a numéraire good is not particularly set up. This is not to lose generality of our model analysis and draw model implications for numerical analysis.

(1.16), and (1.20), we obtain the NEG wage equation:

$$w_i = \Gamma(\theta_i) \left[\mu \sum_{j=1}^{R} Y_j G_j^{\sigma-1} T_{ij}^{1-\sigma} \right]^{1/\sigma}, \tag{1.21}$$

where

$$\Gamma(\theta_i) = \left(1 + \frac{\sigma r \gamma}{q_i(\theta_i)}\right)^{1/\sigma} \left(1 + \frac{r\gamma}{q_i(\theta_i)}\right)^{-1} \left(1 - \frac{\delta\gamma}{q_i(\theta_i)}\right). \tag{1.22}$$

The sum in brackets gives the RMP $\equiv \mu \sum_{j=1}^{R} Y_j G_j^{\sigma-1} T_{ij}^{1-\sigma}$, expressing the sum of the regional income discounted by the price index, and weighted by the transport cost. Even if we assume the frictions in the regional labor markets, the standard implication for NEG holds; that is, the goodness of accessibility to other markets increases the nominal wages.

Following the assumption of an identical price for all the varieties produced within a region, the price index takes the following form:

$$G_i = \left[\sum_{j=1}^{R} n_j (p_j T_{ji})^{1-\sigma} \right]^{1/(1-\sigma)}. \tag{1.23}$$

By substituting (1.12) and (1.19) into (1.23) and with normalization, we obtain

$$G_i = \left[\sum_{j=1}^{R} (1-u_j) L_j \Gamma(\theta_j)^{\sigma} w_j^{1-\sigma} T_{ji}^{1-\sigma} \right]^{1/(1-\sigma)}. \tag{1.24}$$

As mentioned earlier, wage equation, RMP, and price index are essentially identical with vom Berge (forthcoming).

The price of land/housing services p_i^H is determined at equilibrium, where land endowment (or fixed supply of land/housing services) $\bar{S}_i$ and the regional demand for land/housing services H_i are equal. Thus, the price of land/housing services in region i is as follows:

$$p_i^H = \frac{(1-\mu)Y_i}{\bar{S}_i}. \tag{1.25}$$

The regional income, Y_i, includes the income of every employed and unemployed worker living in region i. The respective disposable income of the employed and unemployed workers are given by $I_i^e = w_i + h - \tau$ and $I_i^u = z + h - \tau$, where h is the land rent and τ is the tax rate. Since all the individuals have their own land equally, the land rent is equally redistributed.[22] Thus, the land rent is given by

$$h = \frac{1-\mu}{\mu} \frac{\sum_{j=1}^{R} [w_j(1-u_j) + zu_j - \tau]L_j}{\sum_{j=1}^{R} L_j}. \tag{1.26}$$

Therefore, the regional income Y_i becomes

$$Y_i = [w_i(1-u_i) + zu_i - \tau]L_i + \frac{1-\mu}{\mu} \frac{L_i}{\sum_{j=1}^{R} L_j} \left[\sum_{j=1}^{R} \big(w_j(1-u_j) + zu_j - \tau \big) L_j \right]. \tag{1.27}$$

Further, the individual real income takes the following forms:

$$\mathbb{V}_i^e = \frac{w_i + h - \tau}{G_i^{\mu}(p_i^H)^{1-\mu}}, \quad \text{and} \quad \mathbb{V}_i^u = \frac{z + h - \tau}{G_i^{\mu}(p_i^H)^{1-\mu}}. \tag{1.28}$$

Next, we consider labor market tightness and unemployment rates. Given w_i, labor market tightness is determined in (1.15). Since the inflow and outflow of unemployment are equalized in steady state equilibrium, we obtain $\delta(1-u_i)L_i = \theta_i q_i(\theta_i) u_i L_i$. Solving this with respect to u_i, we obtain the so-called Beverage curve:

$$u_i = \frac{\delta}{\delta + \theta_i q_i(\theta_i)}. \tag{1.29}$$

The tax rate τ is determined to balance the budget for tax revenue and expenditure for unemployment benefits as follows:

$$\tau \sum_{j=1}^{R} L_j = z \sum_{j=1}^{R} u_j L_j. \tag{1.30}$$

Finally, the matching function is assumed to take the Cobb–Douglass form with constant returns to scale

$$A_i m(u_i L_i, v_i L_i) = A_i (u_i L_i)^{\alpha} (v_i L_i)^{1-\alpha}, \tag{1.31}$$

where α is the matching elasticity. This specification of the matching function is also used in the empirical analysis.

1.3 Long-Run Equilibrium: A Two-Region Case

In this section, we numerically analyze the properties of our model.[23] We limit our numerical analysis to a two-region case ($R = 2$) owing to mathematical difficulties.

[22] See Appendix 1.A for details of the derivation.

[23] Numerical analysis is conducted using the Ox Console 7.01 (Doornik and Ooms, 2006).

1.3.1 Spatial Equilibrium

We assume that workers are mobile across regions in response to the expected PDV differentials in the long-run. For convenience of notation, we denote the shares of labor force in regions 1 and 2 as $s_1 = L_1/(L_1+L_2)$ and $s_2 = 1-s_1$, respectively. The regional differentials in the expected PDVs are then expressed as follows:

$$\varDelta\omega(s_1) \equiv \omega_1(s_1) - \omega_2(s_1), \tag{1.32}$$

where the expected PDV from living in region i is expressed as $\omega_i(s_1) = (1-u_i(s_i))E(s_1) + u_i(s_i)U(s_1)$, with the PDVs of the employed and the unemployed worker living in region i given respectively as

$$E_i(s_1) = \frac{(r+\theta_i q_i(\theta_i))\mathbb{V}_i^e + \delta\mathbb{V}_i^u}{r(r+\delta+\theta_i q_i(\theta_i))} \quad \text{and} \quad U_i(s_1) = \frac{\theta_i q_i(\theta_i)\mathbb{V}_i^e + (r+\delta)\mathbb{V}_i^u}{r(r+\delta+\theta_i q_i(\theta_i))}. \tag{1.33}$$

Note that the wage w_i, price index G_i, price of land/housing services p_i^H, land rent h, labor market tightness θ_i, unemployment rate u_i, and tax τ are functions of s_i. A spatial equilibrium arises at $s_1^* \in (0,1)$ when $\varDelta\omega(s_1) = 0$, at $s_1 = 0$ when $\varDelta\omega(0) \leq 0$, or at $s_1 = 1$ when $\varDelta\omega(1) \geq 0$. Any adjustment process over time t is governed by the following differential equation:

$$\frac{\mathrm{d}s_1}{\mathrm{d}t} \equiv \dot{s}_1 = \varDelta\omega(s_1)s_1(1-s_1), \tag{1.34}$$

where the equilibrium is stable when the slope of $\dot{s}_1$ is negative. The parameter setting for the numerical analysis is shown in Table 1.1.

1.3.2 Regional Labor Markets When Agglomeration Has No Externalities on Matching Efficiency

We first consider the benchmark case in which agglomeration has no externalities on the matching efficiency. Panel (a) of Figure 1.2 illustrates the regional differentials in PDVs of the employed and the unemployed for three cases of transport costs ($T = 1.5, 1.6, 1.7$). When $T = 1.7$, we have three equilibria, two stable at $s_1 = 0.04, 0.96$ and one unstable at $s_1 = 0.50$. When $T = 1.6$, we have two stable equilibria at $s_1 = 0.11, 0.89$ and one unstable equilibrium at $s_1 = 0.50$. However, the stable equilibria shift inward. When $T = 1.5$, we have a unique and stable equilibrium at $s_1 = 0.5$.

Panel (b) of Figure 1.2 describes the unemployment differentials between regions 1 and 2 under the short-run equilibrium. When $s_1 > 0.5$, the unemployment rate in region 1 is always lower than that in region 2, where the relationship is robust under different values

Table 1.1: Parameter Setting for Numerical Analysis

Parameter	Explanation
$1 \leq T \leq 2$	Transport Cost
$\sigma = 6$	Elasticity of Substitution between Varieties
$\mu = 0.86$	Expenditure Share for Manufactured Goods
$\delta = 0.03$	Job Destruction Rate ($i = 1, 2$)
$\gamma = 0.5$	Marginal Labor Input for Recruiter per Vacancy
$\beta = 0.5$	Bargaining Power of Worker
$\bar{S}_i = 1$	Land Endowment ($i = 1, 2$)
$r = 0.01$	Interest Rate
$z = 0.4$	Unemployment Benefit
$L_1 + L_2 = 1$	Total Labor Force
$A = 0.6$	Constant of Matching Efficiency
$\alpha = 0.5$	Matching Elasticity on Job Seekers
$\xi = 0$	Elasticity of Agglomeration to Matching Efficiency (Benchmark)
$\xi = 0.02$	Elasticity of Agglomeration to Matching Efficiency (Positive)
$\xi = -0.02$	Elasticity of Agglomeration to Matching Efficiency (Negative and Weak)
$\xi = -0.06$	Elasticity of Agglomeration to Matching Efficiency (Negative and Strong)

Notes: The matching function is $A_i m(u_i L_i, v_i L_i) = A_i (u_i L_i)^{\alpha} (v_i L_i)^{1-\alpha}$, where $A_i = A(L_i/\bar{S})^{\xi}$.

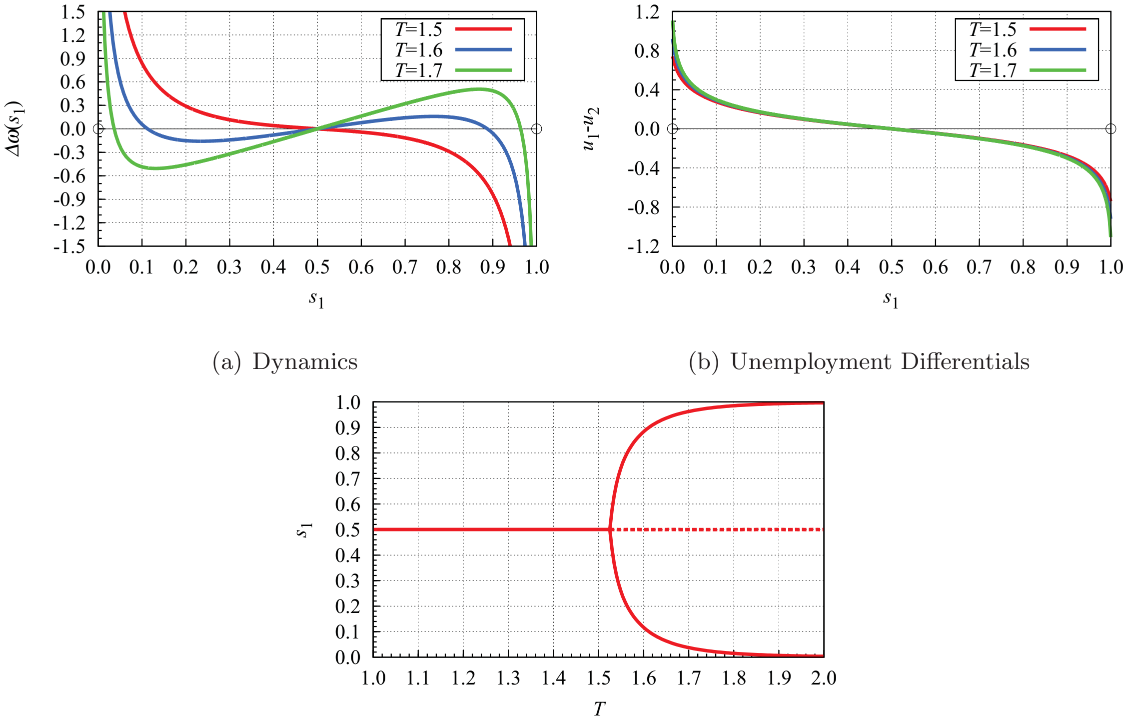
of transport costs. This result derives from the fact that the nominal wage in a denser region is always higher, resulting in a lower unemployment rate. In contrast, vom Berge (forthcoming) shows opposite results. This is because the nominal wage in a denser region is lower in the Krugman (1991b) model.

Panel (c) of Figure 1.2 summarizes the spatial equilibria with respect to transport costs. The solid and dashed lines indicate stable and unstable equilibria respectively. A partial agglomeration arises when the transport costs are high.[24] In our model, the break and sustain points coincide with each other. These points are at $T = 1.53$ in Panel (c) of Figure 1.2. Contrary to our results, vom Berge (forthcoming) shows a full agglomeration when the transport costs are low.

Following our numerical results, we discuss mainly the regional labor market outcomes in spatial equilibrium.[25] We assume that region 1 has at least half of the labor force

[24] As shown in Pflüger and Tabuchi (2010), a full agglomeration is never a stable spatial equilibrium in a typical Helpman (1998) model. Intuitively, this is because if all the workers gather in one region, the price of land/housing services in the other region becomes zero. Consequently, workers have an incentive to move to the vacant region to enjoy higher utility; thus, a full agglomeration never arises.

[25] The figures for coefficient of variation of unemployment rates, labor market tightness θ_i ($i = 1, 2$), relative nominal wage w_1/w_2, relative cost-of-living index $(G_1)^{\mu}(p_1^H)^{1-\mu}/(G_2)^{\mu}(p_2^H)^{1-\mu}$, relative price index for manufactured goods G_1/G_2, and relative price of land/housing services p_1^H/p_2^H are available in Appendix 1.B.



(a) Dynamics

(b) Unemployment Differentials

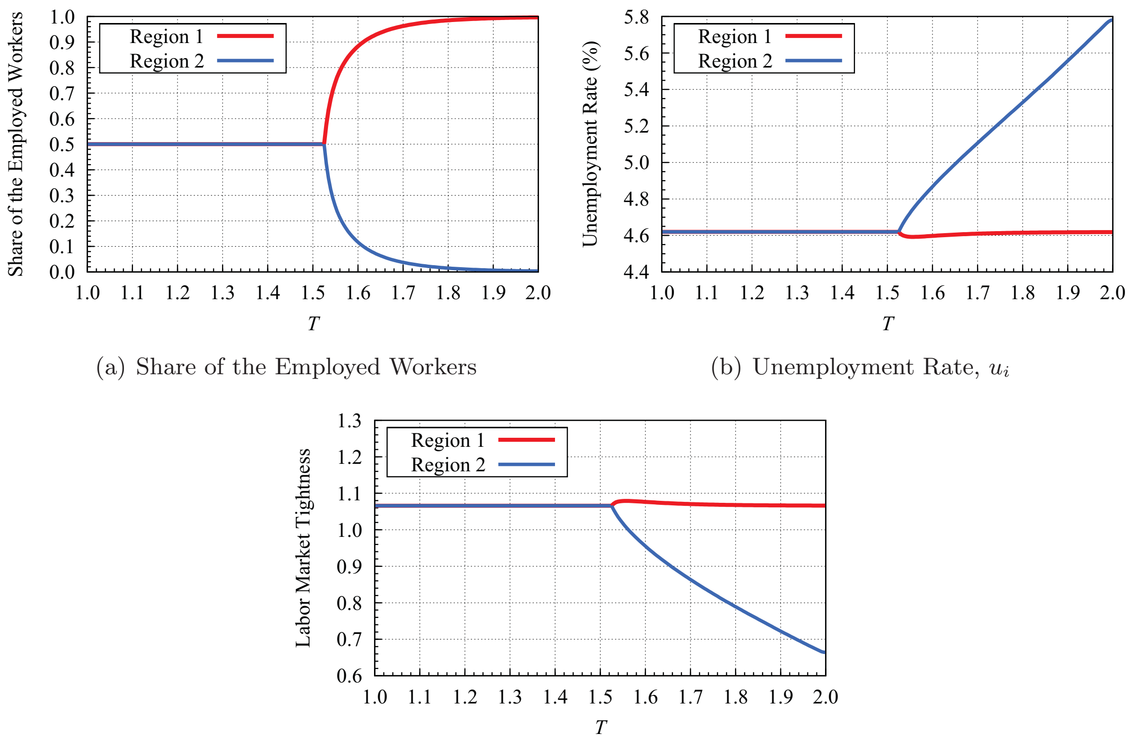
(c) Spatial Equilibrium

Figure 1.2: Results from Numerical Analysis When Agglomeration Has No Externalities on Matching Efficiency

Notes: The solid and dashed lines in Panel (b) denote stable and unstable equilibrium, respectively. The parameters used in this numerical analysis are shown in Table 1.1.

$(0.5 \leq s_1 < 1)$. Figure 1.3 illustrates how the regional shares of the employed workers, unemployment rates, and labor market tightness vary depending on transport costs.

Panel (a) of Figure 1.3 shows that when transport costs are high, region 1 has a larger share of the employed than region 2. In such a case, we call region 1 an employment cluster, a core region, or an agglomerated region. Panel (b) of Figure 1.3 presents a lower unemployment rate in the employment cluster. From the negative relationship between unemployment rate and labor market tightness, as shown in Panel (c) of Figure 1.3, labor market tightness in the employment cluster takes a higher value than that in a less dense region, suggesting that the unemployed can easily find jobs, thus lowering the unemployment rate in an agglomerated region.



(a) Share of the Employed Workers

(b) Unemployment Rate, u_i

(c) Labor Market Tightness, θ_i

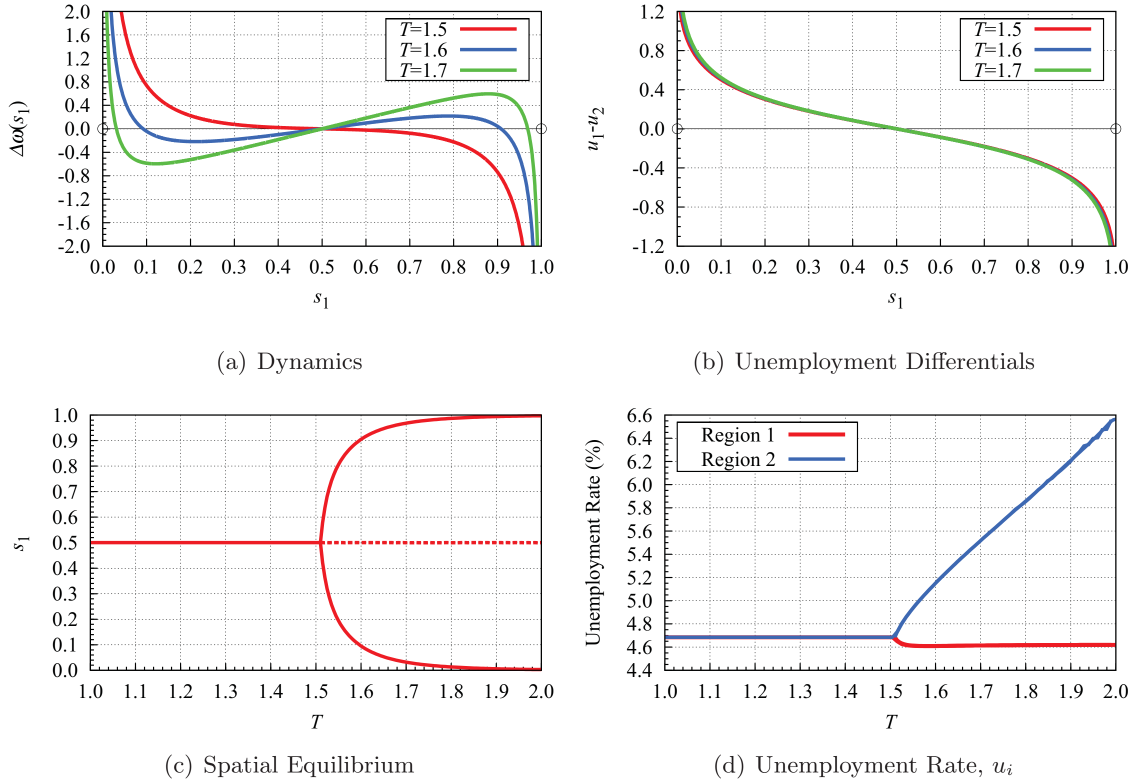
Figure 1.3: Numerical Simulation in Spatial Equilibrium When Agglomeration Has No Externalities on Matching Efficiency

Notes: The parameters used in this numerical analysis are in shown Table 1.1.

1.3.3 Regional Labor Markets When Agglomeration Has Externalities on Matching Efficiency

We further explore three cases in which the agglomeration has externalities on the matching efficiency. Figures 1.4, 1.5, and 1.6 present the results of numerical analysis for the three cases, respectively. For ease of comparison with the benchmark case, each panel of the figures corresponds to respective panels of Figure 1.2 and Panel (b) of Figure 1.3.

First, Figure 1.4 presents the results of numerical analysis for the case in which the agglomeration has positive externalities on matching efficiency. We see that the spatial distribution of workers does not change qualitatively compared to the benchmark case. However, the positive agglomeration externalities on matching efficiency lower the dispersion force from congestion costs and widen the gap in unemployment rates. In both the



(a) Dynamics

(b) Unemployment Differentials

(c) Spatial Equilibrium

(d) Unemployment Rate, u_i

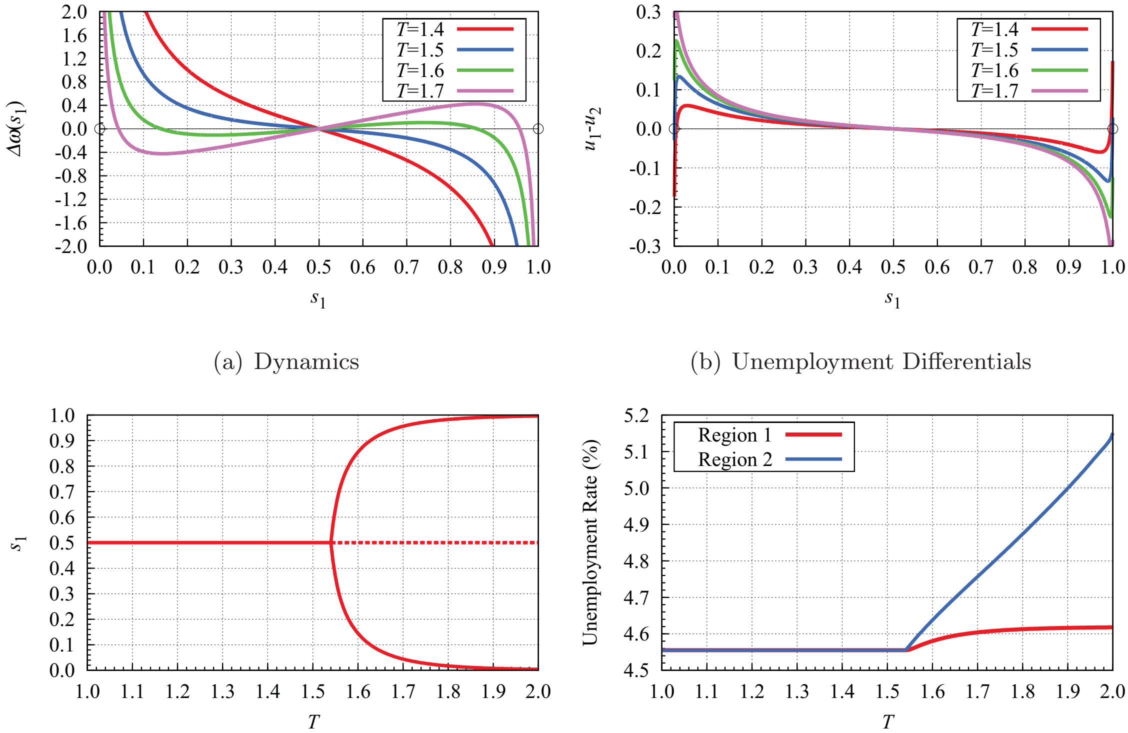
Figure 1.4: Numerical Simulation Results When Agglomeration Has Positive Externalities on Matching Efficiency

Notes: The solid and dashed lines in Panel (c) denote stable and unstable equilibria, respectively. The parameters used in this numerical analysis are shown in Table 1.1.

short- and long-run, the unemployment rate in the employment cluster is relatively low.

Second, Figure 1.5 presents the results of numerical analysis for the case in which the agglomeration has negative externalities on matching efficiency, but the relationship is comparatively weak. In this case as well, the spatial distribution of workers does not change qualitatively compared to the benchmark case. The negative and weak agglomeration externalities on matching efficiency increase the dispersion force from congestion costs and narrow down the regional gap of unemployment rates partly. Note that the unemployment rate in the employment cluster becomes either lower or higher in the short-run depending on the degree of agglomeration (s_1).

Third, Figure 1.6 presents the results of numerical analysis for the case in which the agglomeration has negative externalities on matching efficiency, but the relationship is



(a) Dynamics

(b) Unemployment Differentials

(c) Spatial Equilibrium

(d) Unemployment Rate, u_i

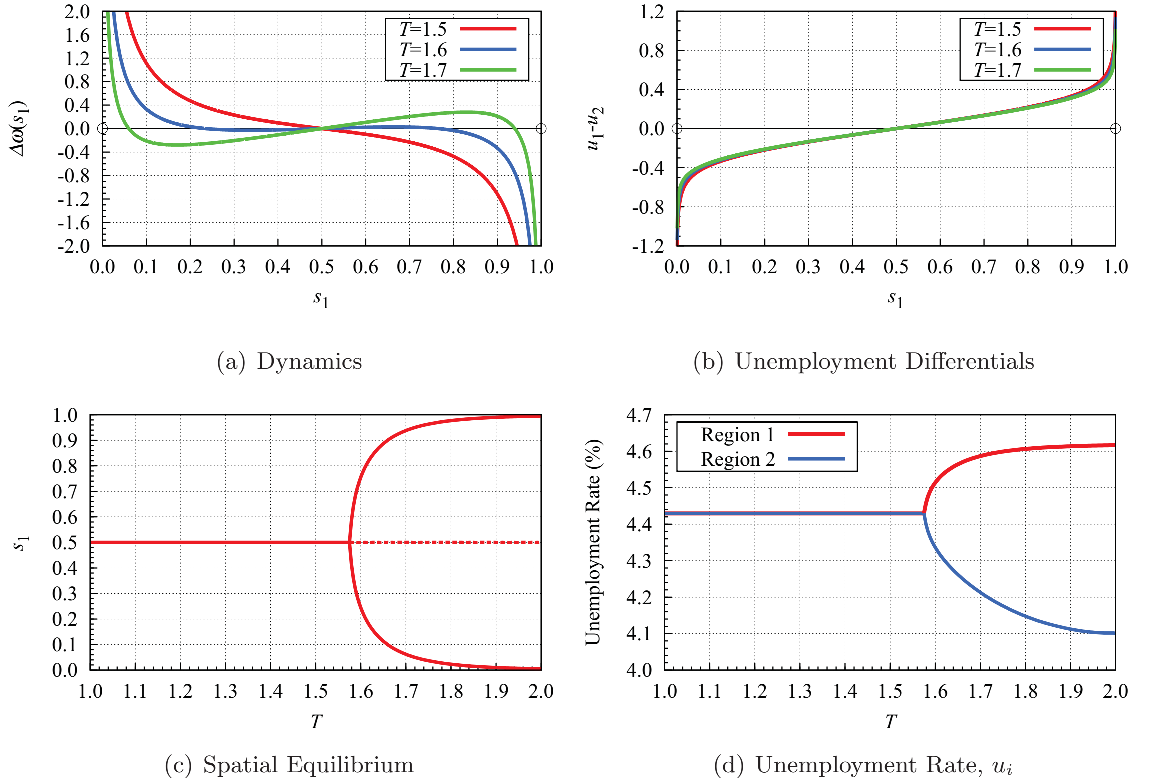
Figure 1.5: Numerical Simulation Results When Agglomeration Has Weak Negative Externalities on Matching Efficiency

Notes: The solid and dashed lines in Panel (c) denote stable and unstable equilibria, respectively. The parameters used in this numerical analysis are shown in Table 1.1.

comparatively strong. The negative and strong agglomeration externalities on matching efficiency increase the dispersion force from congestion costs and gradually widen the regional gap of unemployment rates above a certain degree of the negative relationship. The unemployment rate in the employment cluster is relatively high in the short- and long-run. Another important result is that the nominal wage in the employment cluster is always relatively high in all cases, which is consistent with the stylized facts of this literature.[26]

The theoretical predictions of this study are as follows. In the benchmark case in which agglomeration has no externalities on the matching efficiency, the unemployment rate in the employment cluster is relatively low so that agglomeration has a decreasing effect on

[26] See Appendix 1.B for numerical simulation results of the relative nominal wages in each case.



(a) Dynamics

(b) Unemployment Differentials

(c) Spatial Equilibrium

(d) Unemployment Rate, u_i

Figure 1.6: Numerical Simulation Results When Agglomeration Has Strong Negative Externalities on Matching Efficiency

Notes: The solid and dashed lines in Panel (c) denote stable and unstable equilibria, respectively. The parameters used in this numerical analysis are shown in Table 1.1.

unemployment rates in the production side.[27] However, agglomeration has a positive effect on regional unemployment rates in a search and matching process when agglomeration gives rise to negative externalities on the matching efficiency. When the negative agglomeration externalities on matching efficiency are comparatively weak, the unemployment rate in the employment cluster still remains partly low. When these externalities are comparatively strong, the unemployment rate in the employment cluster becomes higher.

Some predictions of our model are different from vom Berge (forthcoming), who incorporated a search and matching framework into Krugman's (1991) model. vom Berge (forthcoming) showed a positive relationship between regional unemployment rates and ag-

[27] This result is essentially the same as Suedekum (2005) and Zierahn (forthcoming).

glomeration through a negative relationship between nominal wages and agglomeration.[28] However, the latter relationship is clearly inconsistent with empirical evidence. The advantage of our model is that we describe a wide variety of relationships between regional unemployment rates and agglomeration, with the relationship between nominal wage and agglomeration positive. Consequently, our unifying framework contains aspects of both Harris and Todaro (1970) and Blanchflower and Oswald (1994). From our theoretical predictions, we empirically examine the relationship between unemployment rates and agglomeration, and between matching efficiency and agglomeration.

1.4 Empirical Analysis

1.4.1 Unemployment Rates and Agglomeration

First, we attempt to examine the relationship between regional unemployment rates and agglomerations. As a proxy for agglomeration, we use employment density. We use municipal data from Mexico for this analysis. More attention should be paid to spatial autocorrelation when spatially small regional units are used. In this case, the observations are closely related to each other. If the spatial dependence across observations is ignored, the estimators will be inconsistent or not efficient.[29] To solve this problem, we use spatial econometric methods. Thus, our regression models for unemployment rates are given by

$$\log(u_{i,t}) = \rho \sum_{j=1}^{R} b_{ij} \log(u_{j,t}) + \psi \log(\mathrm{Dens}^s_{i,t}) + \boldsymbol{Z}^s_{i,t}\boldsymbol{\phi} + \varepsilon_{i,t}, \tag{1.35}$$

and

$$\log(u_{i,t}) = \psi \log(\mathrm{Dens}^s_{i,t}) + \boldsymbol{Z}^s_{i,t}\boldsymbol{\phi} + e_{i,t}, \quad e_{i,t} = \lambda \sum_{j=1}^{R} b_{ij} e_{j,t} + \varepsilon_{i,t}, \tag{1.36}$$

[28] This difference arises from the sector generating a dispersion force. Krugman's (1991) model deals with freely tradable agricultural goods, but the agricultural workers are not mobile. Helpman's (1998) model deals with the land/housing sector, whose services are consumed locally. Intuitively, in a Krugman-type model, a full agglomeration emerges and no manufacturing worker lives in the periphery region. Therefore, the unemployment rate in a periphery region is virtually zero. In other words, the nominal wage given by equation (1.21) can be defined even in regions with no manufacturer and is lower in agglomerated region; so the implicit unemployment rate also can be calculated. In contrast, in a Helpman-type model, there is a partial agglomeration, and so manufacturing workers always live in the periphery region. Therefore, higher nominal wage in the core region generates higher labor market tightness, leading to a further lower unemployment rate.

[29] Regardless of endogeneity problem from employment density, OLS estimators are biased due to the omitted variable when $\rho \neq 0$. In addition, the covariance matrix of OLS estimators are no more efficient when $\lambda \neq 0$. See LeSage and Pace (2009) for detailed discussions

where $u_{i,t}$ is municipality i' unemployment rate in year t, b_{ij} is the ijth element of the spatial weight matrix (SWM), ψ is the key parameter of our interest, $\text{Dens}^s_{i,t}$ is the log of spatially smoothed employment density, $\boldsymbol{Z}^s_{i,t}$ is a row vector of spatially smoothed control variables, $\boldsymbol{\phi}$ is a column vector of parameters for control variables, and $e_{i,t}$ and $\varepsilon_{i,t}$ are error terms. The control variables include the average years of schooling, rates of male and female labor force participation, and shares of the population aged 15–24, 25–59, and 60 and above.

Note that raw municipal data are not appropriate because the commuting that flows across municipal borders are not negligible at the municipality level and the local labor markets do not necessarily coincide with the administrative areas. Therefore, we use spatially smoothed municipal data in terms of the neighboring municipalities. See Section 1.5 for calculation of the spatially smoothed variables. To control for the endogeneity problem of employment density, we estimate equations (1.35) and (1.36) by using the method of instrumental variable (IV) and generalized method of moments (GMM). Our estimation methodology is based on Kelejian and Prucha (1998).

1.4.2 Matching Efficiency and Agglomeration

We furthermore examine agglomeration externalities on matching efficiency. The estimation procedure takes a two-step approach. In the first step, we estimate the regional matching efficiencies by estimating the matching function. From the logarithm of (1.31), the regression model to be estimated is given by

$$\log(\text{Match}_{s,t}) = \alpha_1 \log(\text{Seeker}_{s,t}) + \alpha_2 \log(\text{Vacancy}_{s,t}) + a_s + \text{year}_t + \varepsilon_{s,t}, \tag{1.37}$$

where $\text{Match}_{s,t}$ is the number of matched jobs in state s at time t, $\text{Seeker}_{s,t}$ is the number of job seekers, $\text{Vacancy}_{s,t}$ is the number of vacancies, α_1 and α_2 are the elasticities of matching, $a_s = \log(A_s)$ is the state fixed effect, year_t is the year dummy, and $\varepsilon_{s,t}$ is the error term. Note that our data set of job seeker, vacancy, and matched job is at the state level owing to the data limitations, and that subscript s is used instead of i. The state fixed effect a_s represents the regional differences in matching efficiency.[30] If we assume constant returns to scale in the matching function, then $\alpha_2 = 1 - \alpha_1$. In the estimation, we test the null hypothesis of constant returns to scale.

In the second step, the estimated matching efficiency is regressed on employment density

[30]The state fixed effects are estimated by `areg` command in Stata.

as follows:

$$\hat{a}_s = \varphi + \xi \log(\text{Dens}_s) + \varepsilon_s, \tag{1.38}$$

where $\hat{a}_s$ is the estimated matching efficiency, φ is the parameter for a constant term, ξ is a parameter of our interest, the elasticity of employment density to matching efficiency in equation (1.2), Dens_s is the employment density of state s, and ε_s is an error term. Therefore, we examine the relationship between matching efficiency and agglomeration by inspecting the coefficient estimate of Dens_s.

1.5 Data

1.5.1 Unemployment Rates and Agglomeration

We use the 2000 and 2010 Mexican population censuses.[31] From the censuses, the National System of Municipal Information (*Sistema Nacional de Información Municipal*, SNIM) provides its summarized municipal data on area, labor force (the employed and unemployed), average years of schooling, labor force participation rate by gender, and the population aged 15–24, 25–59, and 60 and above.[32]

We construct our data set as follows. The unemployment rate of municipality i is calculated by the ratio of the employed to the labor force living in the municipality. Let $z^s_{i,t}$ denote the spatially local sum data of municipality i in year t, calculated as $z^s_{i,t} = \sum_{j=1}^{R} \mathbf{1}_{ij}(d) z_{j,t}$, where R stands for the number of municipalities, $z_{j,t}$ the raw data of municipality j, and $\mathbf{1}_{ij}(d)$ the ijth element of the indicator matrix, in which the ijth element takes the value of 1 if the distance between municipalities i and j is less than d km and 0 otherwise.[33] We set $d = 40$ km. Thus, the spatially smoothed employment density is $\text{Dens}^s_{i,t} = \text{Emp}^s_{i,t}/\text{Area}^s_{i,t}$, where $\text{Emp}^s_{i,t}$ and $\text{Area}^s_{i,t}$ are spatially local sums of employed worker and area, respectively, of municipality i in year t. Further, the other variables are also calculated using the same method.[34] We drop the lowermost 1% and the uppermost 99% of the distribution of unemployment rates.[35]. We use the spatially

[31] In population censuses, labor data are available for every ten years. The 1990 population census data are also used for instrumental variables. We exclude Nicolás Ruíz in the state of Chiapas from the 2000 data owing to lack of labor data. Furthermore, we found some municipalities were originally lacking in the 2000 population census data.

[32] The data are available at the following Web site (URL: `http://www.snim.rami.gob.mx/`).

[33] SNIM also offers the latitude and longitude data of municipalities, from which the bilateral distances between any two municipalities can be calculated by using the formula of Vincenty (1975).

[34] The average years of schooling is calculated as the spatially local sum of years of schooling divided by the number of municipalities within a radius of $d$km from municipality i.

[35] Observations of zero are excluded because they are included in the lowermost 1 percent. The municipality

Table 1.2: Descriptive Statistics for Unemployment Analysis

Variable	Mean	Std. Dev.	Min	Max
Year 2000				
Unemployment Rate (%)	0.953	0.608	0.031	3.815
Employment Density (person/km^2)	67.439	180.959	0.052	1386.694
Employment Density (person/km^2) in 1990	46.685	130.523	0.068	1020.300
Years of Schooling	5.473	1.178	2.910	9.140
Male Labor Force Participation Rate (%)	68.691	6.173	30.764	84.692
Female Labor Force Participation Rate (%)	25.163	6.537	6.456	40.714
Share of Population Aged 15–24 (%)	19.137	1.337	13.225	23.186
Share of Population Aged 25–59 (%)	35.143	3.740	22.531	43.705
Share of Population Aged 60 and above (%)	8.278	2.102	2.589	16.767
Year 2010				
Unemployment Rate (%)	4.037	2.626	0.067	16.266
Employment Density (person/km^2)	79.295	203.179	0.056	1572.537
Employment Density (person/km^2) in 1990	44.951	127.496	0.068	1020.300
Years of Schooling	6.689	1.149	4.081	10.360
Male Labor Force Participation Rate (%)	72.794	3.605	45.476	84.537
Female Labor Force Participation Rate (%)	27.180	8.085	4.212	48.474
Share of Population Aged 15–24 (%)	18.820	1.076	13.674	22.312
Share of Population Aged 25–59 (%)	39.422	3.661	27.211	46.881
Share of Population Aged 60 and above (%)	10.414	2.598	3.278	22.720

Notes: The numbers of observations in 2000 and 2010 are 2255 and 2387, respectively. The lowermost 1% and uppermost 99% of the distribution of unemployment rates are dropped. These municipal data are spatially smoothed except for unemployment rates.

smoothed employment density of 1990 for IV, and so use the 1990 population census as well.[36] Table 1.2 gives the descriptive statistics of municipal data by year.

For our estimation, we use distance-based SWMs, which take the following form:

$$b_{ij} = \frac{d_{ij}^{-\eta}}{\sum_{j=1}^{R} d_{ij}^{-\eta}}$$

where b_{ij} is the ijth element of an SWM, d_{ij} is the bilateral distance between municipalities i and j, R is the number of municipalities, and η is a distance decay parameter. The bilateral distance is calculated as the great-circle distance between two municipalities measured by latitude and longitude (Vincenty, 1975). The SWMs are row-standardized. In this paper, our estimation results are obtained from using distance-based SWMs ($\eta = 5$).[37]

of *Nicoláas Ruíz* located in the state of Chiapas is also excluded owing to lack of data.

[36]There is no information of municipal area in 1990 population census. Therefore, we complement municipal areas in 1990 with the 2000 population census. In that case, separated municipalities between 1990 and 2000 are added to original municipalities.

[37]Our main results do not change even if different values of η are used.

1.5.2 Matching Efficiency and Agglomeration

The yearly job seeker, vacancy, and matched job data are available from the Secretariat of Labor and Social Welfare (*Secretaría del Trabajo y Previsión Social*, STPS). The time span is from 2001 to 2011. The STPS offers services for the promotion of job matching in job placement offices (*Bolsa de Trabajo*). The data include the number of applications registered both for the first time and on subsequent occasions, the number of job vacancies, and the number of matched jobs out of the vacant jobs registered.[38] Table 1.3 presents the descriptive statistics of job seeker, vacancy, and matched job by year.

We then calculate the employment density at the state level. For this, we use the 2000 population census. In the regression analysis at the second step, the dependent variable is the matching efficiency by state estimated between 2001 and 2011. To avoid endogeneity issues, we use the employment density of 2000. For robustness, we also use the employment density of 1990 as an instrumental variable. A problem with employment density at the state level is that some states have vast uninhabitable regions, leading to underestimated employment densities. To mitigate this issue, we calculate the employment density as follows. The municipal employment density is first simply calculated and sorted by size. Then, the number of the employed in municipalities and municipal areas are summed up respectively until the share of the employed by state reaches 80%. Finally, the state employment density is calculated as the employed–area ratio.

1.6 Empirical Results

1.6.1 Unemployment Rates and Agglomeration

Table 1.4 shows the estimation results for equations (1.35) and (1.36). Columns (1) and (4) of Table 1.4 presents the ordinary least squares (OLS) estimates for 2000 and 2010, respectively. In Column (1), employment density has a significantly negative impact on unemployment rates at the 5% level, but is insignificant even at the 10% level in 2010. According to the robust LM tests for spatial dependence in the dependent variable and error terms, the null hypotheses $\rho = 0$ and $\lambda = 0$ are rejected at least at the 5% level in 2000 and 2010, respectively, and we need to control for spatial dependence.[39] The estimation results for 2000 and 2010 are given in Columns (2) and (3) and Columns (5) and

[38] A person can be hired once more depending on the type of employment (casual, temporary, or permanent).

[39] We follow the hypothesis-testing methodology for spatial dependence proposed by Anselin et al. (1996). See also Anselin (2006) for a brief summary.

Table 1.3: Descriptive Statistics of Matched Jobs, Job Seekers, and Vacancies

	Year										
Variables	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Matching:											
Mean	5419.7	6579.3	6434.8	7155.2	7212.6	7441.5	8560.1	9479.8	8549.4	8270.0	7968.5
Std. Dev.	10087.8	11791.7	10904.6	11959.3	10897.4	10893.0	11703.5	16663.9	14330.2	11337.9	11724.0
Min	607.0	748.0	790.0	1416.0	1769.0	957.0	1503.0	852.0	973.0	1232.0	920.0
Max	59102.0	68554.0	63704.0	69903.0	62221.0	61065.0	63541.0	91826.0	74760.0	49139.0	55518.0
Job Seeker:											
Mean	18199.6	20638.7	21856.6	23665.6	24071.5	24606.9	27304.4	33273.2	37362.8	32245.5	29782.7
Std. Dev.	24722.6	27026.4	28682.3	34722.2	30762.6	33742.2	35747.7	43188.4	47648.0	31384.3	28800.9
Min	4017.0	4200.0	4479.0	4970.0	4473.0	4981.0	5130.0	7170.0	6396.0	5715.0	3575.0
Max	142235.0	154099.0	162969.0	199075.0	173291.0	192490.0	196705.0	240831.0	254189.0	161809.0	146593.0
Vacancy:											
Mean	14147.8	14331.6	14174.0	15758.7	16986.2	18302.4	19405.4	24822.8	25389.1	22295.9	20064.7
Std. Dev.	31392.9	35818.6	32442.8	39032.4	37674.0	41262.8	39293.2	46842.2	41436.4	28689.2	31163.5
Min	1177.0	1415.0	1359.0	1315.0	1753.0	1462.0	1695.0	2987.0	4004.0	3727.0	3021.0
Max	178649.0	207282.0	187916.0	225711.0	217825.0	236441.0	218023.0	256896.0	219377.0	134007.0	136153.0

Notes: The number of states in each year is 32.

(6), respectively. As expected, the parameter estimates measuring spatial dependence in the dependent variable and error terms are significantly positive in both years. The coefficient estimates of employment density remain significantly negative even after controlling for spatial dependence in 2000. However, employment density is no longer significant in 2010.

For robustness, we control for the endogeneity of employment density. Table 1.5 presents the IV/GMM estimation results. In Columns (1) and (4), we control for the endogeneity of employment density but do not control for spatial dependence in the dependent variable and error terms. The Dubin–Wu–Hausman test shows that there exists an endogeneity problem in regression model. From IV/GMM estimation, we find the coefficient estimates of employment density significantly negative in 2000 and 2010. However, Robust LM Tests suggest that spatial dependence should be controlled for. As earlier, Columns (2) and (3) and Columns (4) and (5) show, respectively, the estimation results when spatial dependence in the dependent variable and error terms are controlled for. In 2000, employment density shows a significantly negative impact on unemployment rate. However, this is not the case in 2010.

Our evidence on the negative relationship between unemployment rates and agglomeration is robust for 2000, but not for 2010. When based exactly on our model, the unemployment differentials decreased as the transport costs fell, and the magnitude of the estimated coefficient became smaller and the statistical significance might not be confirmed. Another important implication is that the negative relationship between unemployment and agglomeration can be observed only when agglomeration has positive or weakly negative externalities on the matching efficiency. In the next subsection, we examine the relationship between matching efficiency and agglomeration.

1.6.2 Matching Efficiency and Agglomeration

Table 1.6 presents the estimation results of regression models (1.37) and (1.38). Column (1) shows the estimation results of the matching function. The elasticity of job match to job seeker is significant at the 5% level and takes the value of 0.33. The elasticity of job match to vacancy is also significant at the 1% level and takes the value of 0.71. The null hypothesis of constant returns to scale for the matching function is not rejected. Our estimates are consistent with the results of most of the empirical studies on the matching function. According to a survey of Petrongolo and Pissarides (2001), the estimate of plausible elasticity on job seeker lies in the range between 0.5 and 0.7.

Column (2) shows the estimation results of the agglomeration effect on matching effi-

Table 1.4: Estimation Results for Regional Unemployment Rates and Agglomeration

	Dependent Variable: $\log(u_{i,t})$					
	Year: 2000			Year: 2010		
Explanatory Variable	(1)	(2)	(3)	(4)	(5)	(6)
Employment Density	-0.058^{**}	-0.032^{***}	-0.051^{*}	-0.035	0.006	-0.035
	(0.023)	(0.011)	(0.028)	(0.022)	(0.015)	(0.028)
Years of Schooling	0.201^{***}	0.154^{***}	0.197^{***}	-0.038	-0.029	-0.023
	(0.032)	(0.021)	(0.034)	(0.034)	(0.025)	(0.039)
Labor Force Participation Rate (Male)	-2.149^{***}	-1.288^{***}	-2.005^{***}	-2.757^{***}	-1.201^{***}	-2.303^{***}
	(0.247)	(0.205)	(0.248)	(0.461)	(0.384)	(0.484)
Labor Force Participation Rate (Female)	-0.185^{*}	-0.081	-0.195^{*}	-0.180^{*}	-0.128^{*}	-0.132
	(0.100)	(0.076)	(0.102)	(0.101)	(0.076)	(0.104)
Share of Population Aged 15–24	0.747^{**}	0.383	0.753^{**}	1.033^{**}	0.732^{*}	0.920^{*}
	(0.362)	(0.246)	(0.370)	(0.484)	(0.382)	(0.528)
Share of Population Aged 25–59	1.347^{***}	0.557^{**}	1.287^{***}	2.570^{***}	1.398^{***}	2.257^{***}
	(0.341)	(0.273)	(0.357)	(0.495)	(0.392)	(0.523)
Share of Population Aged 60 and above	-0.200^{*}	-0.090	-0.156	-0.316^{***}	-0.079	-0.278^{**}
	(0.106)	(0.080)	(0.113)	(0.112)	(0.086)	(0.130)
Constant	2.027	1.916	1.568	3.080	-0.259	2.271
	(1.519)	(1.251)	(1.630)	(2.648)	(2.190)	(2.765)
Spatially Lagged Dependent Variable (ρ)		0.365^{***}			0.509^{***}	
		(0.047)			(0.052)	
Spatially Lagged Error Terms (λ)			0.153^{***}			0.245^{***}
			(0.020)			(0.019)
State Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	2255	2255	2255	2387	2387	2387
Adjusted R^2	0.243			0.253		
Robust LM Test (ρ), p-value	0.000			0.000		
Robust LM Test (λ), p-value	0.008			0.031		

Notes: Standard errors are in the parentheses. Columns (1), (2), (4), and (5) consider heteroskedastic errors. The explanatory variables are expressed in logarithm except years of schooling. Spatially smoothed municipal data are used. The instrumental variable for spatially lagged dependent variable is $\boldsymbol{B}\tilde{\boldsymbol{Z}}, \dots \boldsymbol{B}^5\tilde{\boldsymbol{Z}}$, where $\boldsymbol{B}$ is the spatial weight matrix and $\tilde{\boldsymbol{Z}}$ a matrix consisting of employment density and control variables. Robust LM Test (ρ) indicates the testing of the null hypothesis $\rho = 0$ against alternative hypothesis $\rho \neq 0$. Robust LM Test (λ) indicates the testing of the null hypothesis $\lambda = 0$ against alternative hypothesis $\lambda \neq 0$. *, **, and *** denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 1.5: Robustness Checks for Endogeneity of Employment Density

	Dependent Variable: $\log(u_{i,t})$					
	Year: 2000			Year: 2010		
Explanatory Variable	(1)	(2)	(3)	(4)	(5)	(6)
Employment Density	-0.075^{***}	-0.048^{***}	-0.072^{**}	-0.047^{**}	0.006	-0.042
	(0.024)	(0.013)	(0.029)	(0.023)	(0.016)	(0.029)
Years of Schooling	0.204^{***}	0.161^{***}	0.199^{***}	-0.035	-0.027	-0.026
	(0.032)	(0.020)	(0.034)	(0.034)	(0.025)	(0.039)
Labor Force Participation Rate (Male)	-2.160^{***}	-1.371^{***}	-2.093^{***}	-2.770^{***}	-1.166^{***}	-2.505^{***}
	(0.245)	(0.201)	(0.247)	(0.458)	(0.384)	(0.485)
Labor Force Participation Rate (Female)	-0.159	-0.071	-0.163	-0.165^{*}	-0.125^{*}	-0.151
	(0.099)	(0.075)	(0.103)	(0.100)	(0.076)	(0.104)
Share of Population Aged 15–24	0.771^{**}	0.456^{**}	0.751^{**}	1.000^{**}	0.751^{**}	0.940^{*}
	(0.359)	(0.232)	(0.368)	(0.479)	(0.383)	(0.529)
Share of Population Aged 25–59	1.408^{***}	0.670^{**}	1.414^{***}	2.600^{***}	1.359^{***}	2.409^{***}
	(0.340)	(0.270)	(0.356)	(0.493)	(0.398)	(0.524)
Share of Population Aged 60 and above	-0.222^{**}	-0.119	-0.203^{*}	-0.344^{***}	-0.073	-0.324^{**}
	(0.105)	(0.079)	(0.113)	(0.113)	(0.087)	(0.131)
Constant	1.783	1.671	1.541	3.160	-0.376	2.733
	(1.516)	(1.237)	(1.624)	(2.628)	(2.196)	(2.765)
Spatially Lagged Dependent Variable (ρ)		0.332^{***}			0.516^{***}	
		(0.043)			(0.051)	
Spatially Lagged Error Terms (λ)			0.153^{***}			0.245^{***}
			(0.020)			(0.019)
State Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	2255	2255	2255	2387	2387	2387
Robust LM Test (ρ), p-value	0.001			0.000		
Robust LM Test (λ), p-value	0.018			0.049		
Dubin–Wu–Hausman Test, p-value	0.013			0.103		

Notes: Standard errors are in the parentheses. Columns (1), (2), (4), and (5) consider heteroskedastic errors. The explanatory variables are expressed in logarithm except years of schooling. Spatially smoothed municipal data are used. The instrumental variable for employment density is the spatially smoothed 10-year lagged employment density. The instrumental variable for spatially lagged dependent variable is $\boldsymbol{B}\tilde{\boldsymbol{Z}}, \ldots \boldsymbol{B}^5\tilde{\boldsymbol{Z}}$, where $\boldsymbol{B}$ is the spatial weight matrix and $\tilde{\boldsymbol{Z}}$ a matrix consisting of the spatially smoothed employment density in 1990 and control variables. Robust LM Test (ρ) indicates the testing of null hypothesis $\rho = 0$ against alternative hypothesis $\rho \neq 0$. Robust LM Test (λ) indicates the testing of the null hypothesis $\lambda = 0$ against alternative hypothesis $\lambda \neq 0$. Dubin–Wu–Hausman Test indicates the hypothesis testing of endogeneity. *, **, and *** denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 1.6: Estimation Results for Matching Function

	(1)	(2)	(3)
	Dependent Variable		
Explanatory Variable	$\log(\text{Match}_{s,t})$	State Fixed Effects in Column (1) (OLS)	State Fixed Effects in Column (1) (GMM)
Log of Employment Density		-0.086^{***}	-0.081^{***}
		(0.026)	(0.027)
State Fixed Effects	Yes		
Log of Seeker	0.332^{**}		
	(0.150)		
Log of Vacancy	0.714^{***}		
	(0.161)		
Year Dummy	Yes		
Number of Observations	352	32	32
Adjusted R^2	0.886	0.200	
CRS Test, p-value	0.700		
Dubin–Wu–Hausman Test, p-value			0.295
F-Test (Weak IV)			1536.740

Notes: Heteroskedasticity-consistent standard errors are in the parentheses. Column (1) gives heteroskedasticity-consistent standard errors clustered by state. All regressions contain a constant term. The instrumental variable for employment density shown in Column (3) is the employment density in 1990. CRS Test indicates the hypothesis testing of constant returns to scale for the matching function. Dubin–Wu–Hausman Test indicates hypothesis testing of endogeneity. F Statistic (Weak IV) is Cragg–Donald Wald F statistic for test of weak instruments. *, **, and *** denote statistical significance at the 1%, 5%, and 10% level, respectively.

ciency. The elasticity of employment density estimated by OLS is significantly negative at the 1% level. The elasticity of employment density estimated by GMM is also significantly negative. Figure 1.7 clearly illustrates the negative relationship between matching efficiency and employment density.

To sum up, the Mexican data examined show comparatively low unemployment rates as well as matching efficiency in agglomerated regions. Taking into account the theoretical prediction that when agglomeration has negative externalities on matching efficiency it will have both positive and negative effects on regional unemployment rates, the agglomeration effect lowering the unemployment rates in Mexico is much stronger than that increasing the unemployment rates.

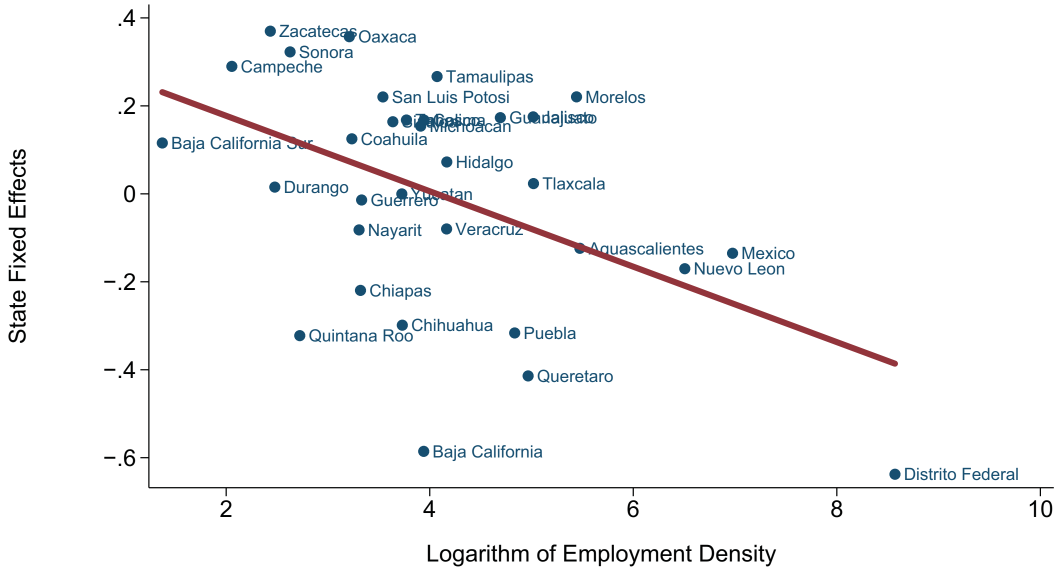


Figure 1.7: Matching Efficiency and Employment Density

1.7 Concluding Remarks

In this paper, we theoretically and empirically analyzed the relationship between regional unemployment rates and agglomeration. In the theoretical part of our analysis, we extended a multi-region model of Helpman (1998) by incorporating search and matching frictions in regional labor markets. In addition, we incorporate agglomeration externalities into a search and matching framework. In the empirical part of our analysis, we examined the relationship between regional unemployment rates and agglomeration (expressed as employment density) by using Mexican municipal data. We also estimated the matching function by using the data of job seekers, vacancies, and matched jobs.

An important prediction of our theory is that agglomeration can be positively or negatively related with regional unemployment rates under negative agglomeration externalities on matching efficiency. Thus, our theoretical framework with agglomeration externalities on matching efficiency can describe a wide variety of relationships between regional unemployment rates and agglomeration, with the relationship between nominal wages and agglomeration positive, as supported by most empirical studies. Therefore, our model can lead to predictions on unemployment rates and wages of both Harris and Todaro (1970) and Blanchflower and Oswald (1994) within a unified framework.

From our empirical results obtained with Mexican data, we found that denser areas have comparatively low unemployment rates under negative agglomeration externalities on matching efficiency. Considering our theoretical predictions, we conclude that in Mexico, the agglomeration effect lowering the unemployment rates is much stronger than that increasing the rates.

Appendix 1.A Derivation of Land Rent

The aggregate income of all regions is equal to the sum of their disposable labor income and income obtained from land/housing services:

$$\sum_{j=1}^{R} Y_j = \sum_{j=1}^{R} \left[(w_i - \tau)(1-u_i)L_i + (z-\tau)u_i L_i\right] + (1-\mu)\sum_{j=1}^{R} Y_j.$$

Thus, the aggregate income from land/housing services in the economy becomes

$$(1-\mu)\sum_{j=1}^{R} Y_j = \frac{1-\mu}{\mu}\sum_{j=1}^{R} \left[(w_i - \tau)(1-u_i)L_i + (z-\tau)u_i L_i\right]$$

Dividing this by the share of regional labor force, the aggregate land rent in region i becomes

$$\frac{L_i}{\sum_{j=1}^{R} L_j}(1-\mu)\sum_{j=1}^{R} Y_j = \frac{L_i}{\sum_{j=1}^{R} L_j}\frac{1-\mu}{\mu}\sum_{j=1}^{R} \left[(w_i - \tau)(1-u_i)L_i + (z-\tau)u_i L_i\right].$$

Furthermore, dividing this by the workers living in region i, the land rent that individuals receive becomes

$$h = \frac{1}{\sum_{j=1}^{R} L_j}(1-\mu)\sum_{j=1}^{R} Y_j = \frac{1}{\sum_{j=1}^{R} L_j}\frac{1-\mu}{\mu}\sum_{j=1}^{R} \left[(w_i - \tau)(1-u_i)L_i + (z-\tau)u_i L_i\right].$$

See Helpman (1998) for more details.

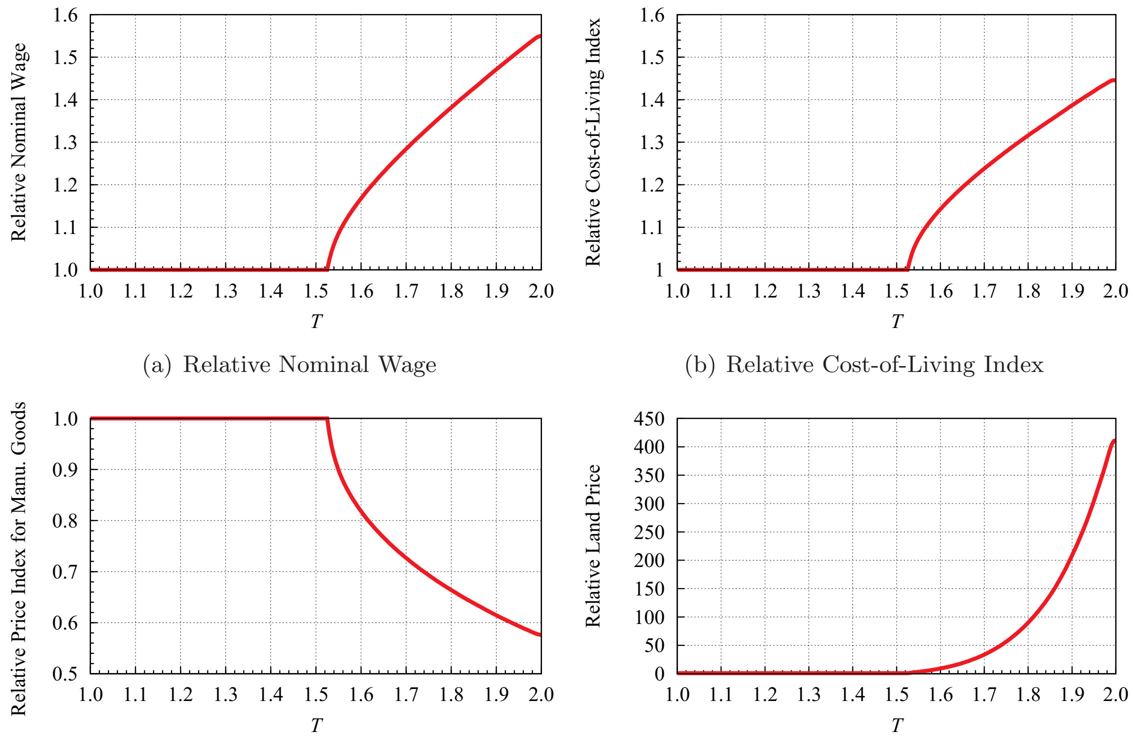
Appendix 1.B Numerical Simulation Results on Other Variables

Figure 1.8 shows the numerical simulation results of the other variables under the assumption of two symmetric regions. We assume that the agglomeration has no externalities on the matching efficiency.

Figure 1.9 shows the numerical simulation results of the other variables under the assumption of two symmetric regions. We assume that agglomeration has positive externalities on the matching efficiency.

Figure 1.10 shows the numerical simulation results of the other variables under the assumption of two symmetric regions. We assume that agglomeration has comparatively weak negative externalities on the matching efficiency.

Figure 1.11 shows the numerical simulation results of the other variables under the assumption of two symmetric regions. We assume that agglomeration has comparatively strong negative externalities on the matching efficiency.



(a) Relative Nominal Wage

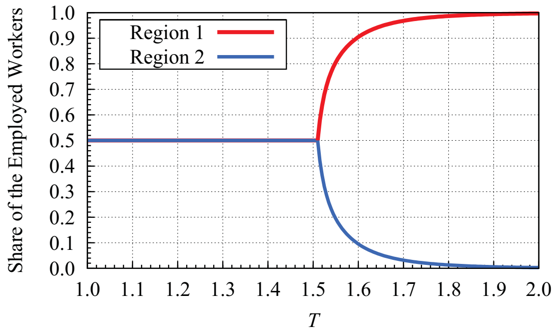
(b) Relative Cost-of-Living Index

(c) Relative Price Index for Manufactured Goods

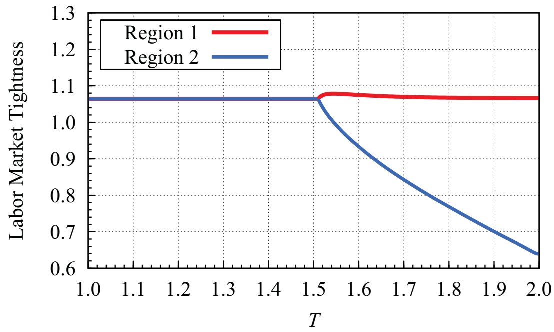
(d) Relative Price of Land/Housing Services

Figure 1.8: Numerical Simulation in Spatial Equilibrium When Agglomeration Has No Externalities on Matching Efficiency

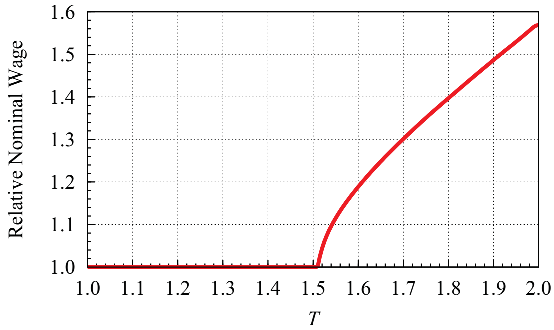
Note: Relative value of region 1 to region 2.



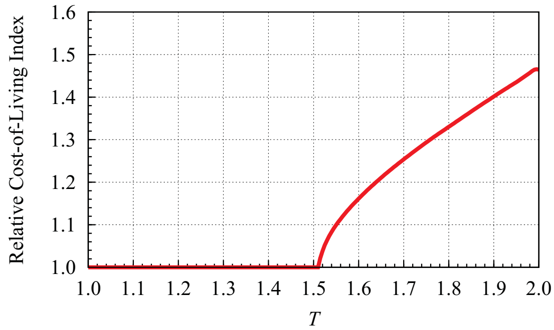
(a) Share of the Employed Workers



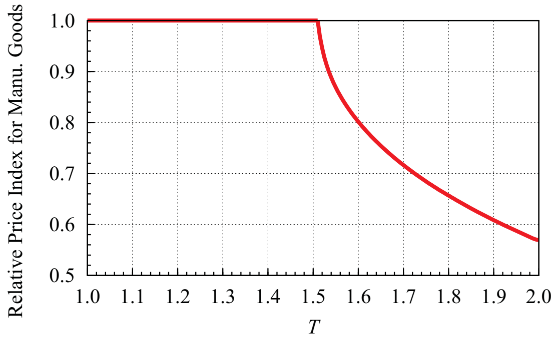
(b) Labor Market Tightness, θ_i



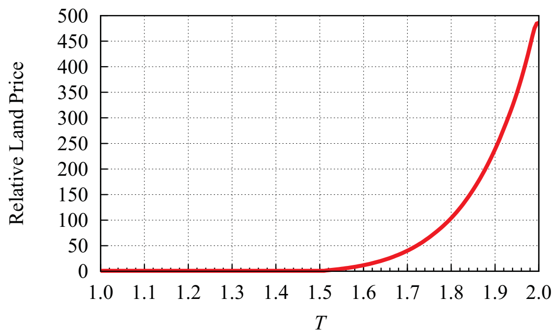
(c) Relative Nominal Wage



(d) Relative Cost-of-Living Index



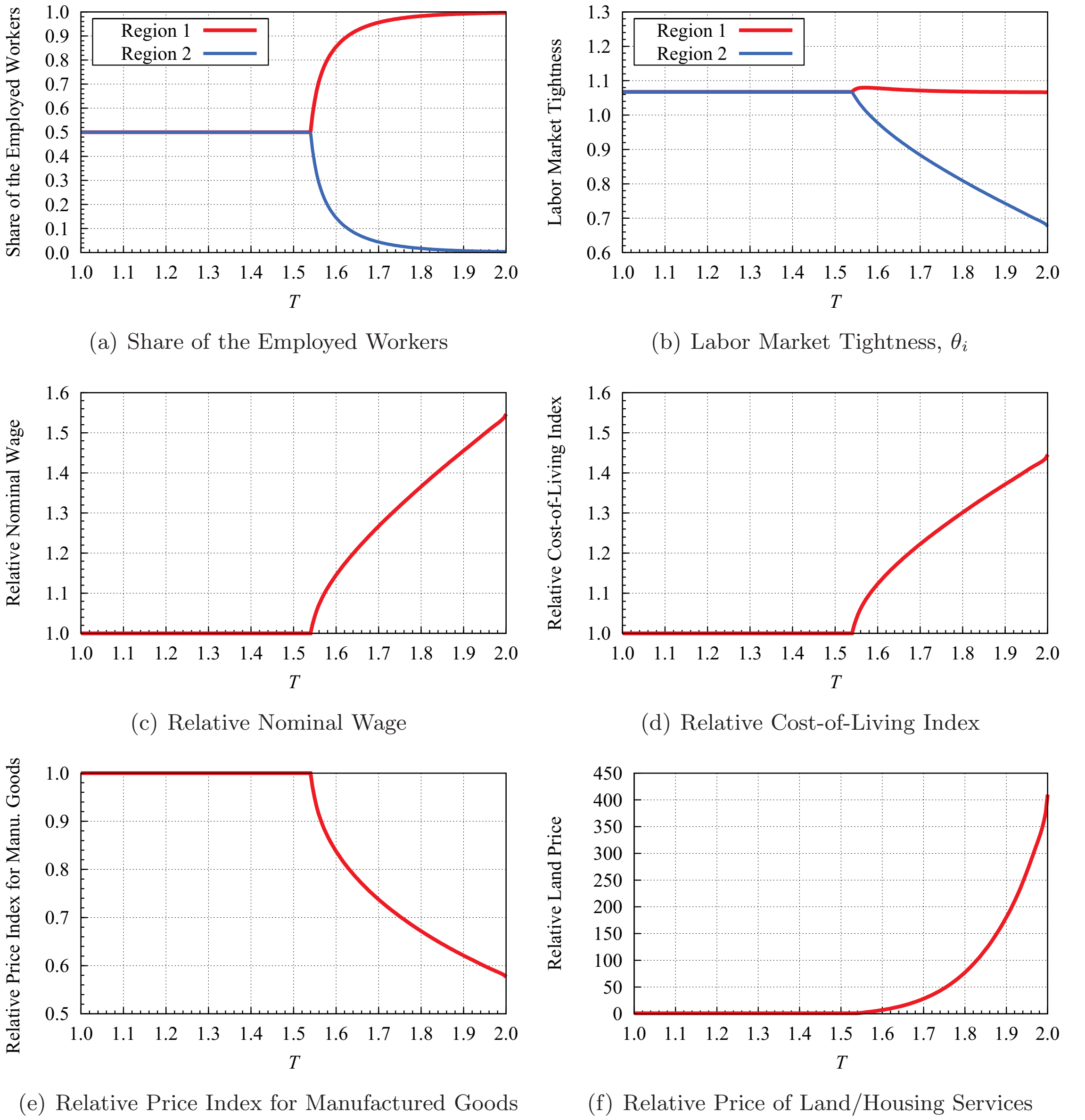
(e) Relative Price Index for Manufactured Goods



(f) Relative Price of Land/Housing Services

Figure 1.9: Numerical Simulation Results under Positive Agglomeration Externalities on Matching Efficiency

Note: Relative value of region 1 to region 2.



(a) Share of the Employed Workers

(b) Labor Market Tightness, θ_i

(c) Relative Nominal Wage

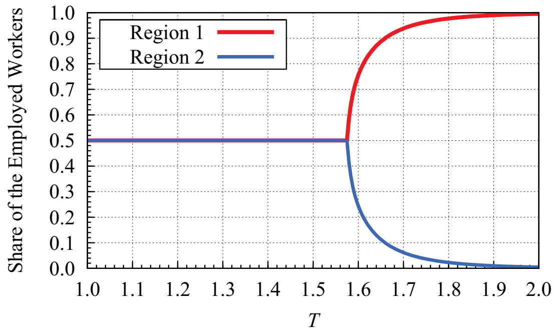
(d) Relative Cost-of-Living Index

(e) Relative Price Index for Manufactured Goods

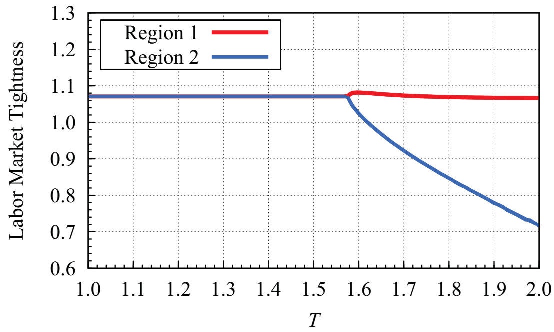
(f) Relative Price of Land/Housing Services

Figure 1.10: Numerical Simulation Results under Comparatively Weak Negative Agglomeration Externalities on Matching Efficiency

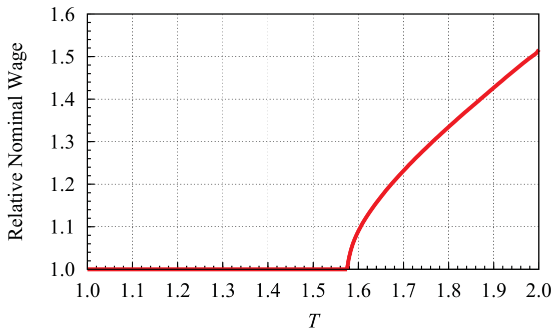
Note: Relative value of region 1 to region 2.



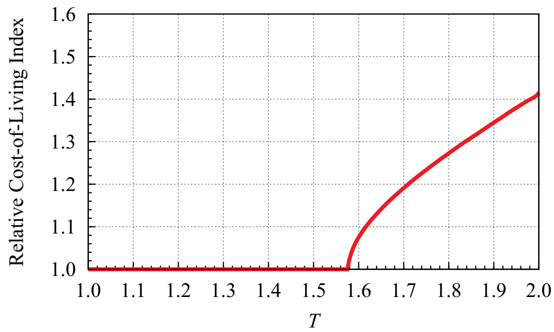
(a) Share of the Employed Workers



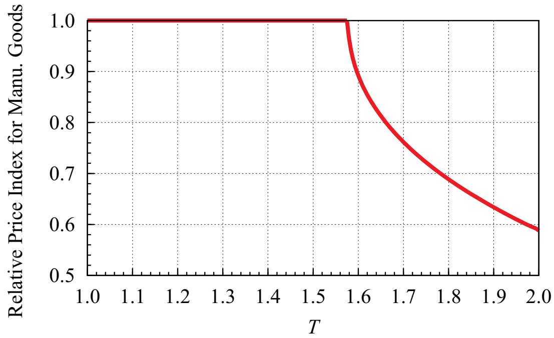
(b) Labor Market Tightness, θ_i



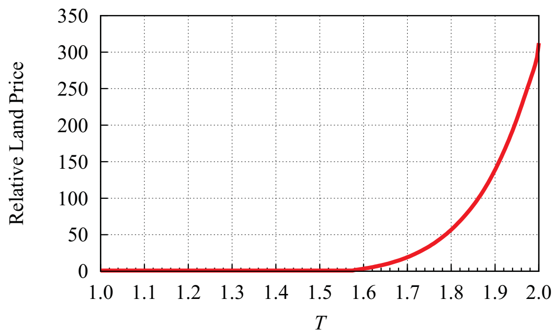
(c) Relative Nominal Wage



(d) Relative Cost-of-Living Index



(e) Relative Price Index for Manufactured Goods



(f) Relative Price of Land/Housing Services

Figure 1.11: Numerical Simulation Results under Comparatively Strong Negative Agglomeration Externalities on Matching Efficiency

Note: Relative value of region 1 to region 2.

Chapter 2

Spatial Dependence in Regional Business Cycles: Evidence from Mexican States*

2.1 Introduction

As conditions in regional economies do not necessarily coincide with national economic situations, regional business cycles tend to be highly heterogeneous. However, spatial proximity seems to characterize similarity among regional business cycles. Thus, in light of the interdependence of regional economies, we focus on the spatial spillover (or neighborhood) effects that exist across regional business cycles. In such an economic situation, a region-specific shock, for example, might propagate outward, toward neighboring economies. In recent years, the importance of conducting spatial analyses of economic activities has been emphasized from the economic stability and growth perspectives (e.g., World Bank, 2009). We, therefore, investigate spatial dependence within regional business cycles.

Although spatial similarity—in terms of business cycles and spillover effects—has attracted attention recently, the propagation process has not yet been investigated. To analyze spatial dependence in regional business cycles and spatial spillover effects, we integrate a spatially lagged dependent variable into a Markov switching model. In other words, our attempt is to provide an integrated estimation method by using the Markov switching model

*I would like to specially thank Alfredo Erquizio Espinal and Kensuke Teshima for their insightful comments and helpful suggestions. I also thank Nobuaki Hamaguchi, Yoichi Matsubayashi, Akio Namba, Tatsuyoshi Okimoto, Calros Urrutia, and participants in the 2013 Spring Meeting of the Japanese Economic Association and in the Rokko Forum at Kobe University for their useful comments and suggestions. Naturally, any errors remaining are my own. Furthermore, I am grateful for the benefits received during my stay at the Instituto Tecnológico Autónomo de México. This research was carried out under a scholarship granted by the Government of Mexico, through the Ministry of Foreign Affairs of Mexico.

and spatial econometrics. An advantage of this model is that it enables us to numerically simulate spatial spillover effects, thus enabling us to investigate the extent to which a regime switch to recession in a regional economy may cause deterioration in the economic conditions of neighboring economies.

Regional business cycles are not perfectly uniform, and thus, discussions of the national business cycle are not directly applicable. For example, in applying the Markov switching model proposed by Hamilton (1989), Owyang et al. (2005) found that business cycles across U.S. states differed considerably in terms of expansionary and recessionary phases. Furthermore, Owyang et al. (2008) investigated business cycles at the U.S. city level, and drew similar conclusions. To explain the similarities and differences within regional business cycles, Hamilton and Owyang (2012) developed a Markov switching model based on the rationale that administrative units do not necessarily coincide with economic zones. In their model, U.S. states are endogenously grouped into clusters that share similar economic characteristics by identifying common factors across states. Therefore, the authors' focus is on regional recessions within each group.[1] They found that states that have a relatively high share of oil production or agriculture are likely to be in recession separately, as compared to other U.S. states. In contrast, we focus on the spatial correlation of regional business cycles, inspired by Owyang et al. (2005). From their empirical results, it seems that we should pay more attention to the fact that state recessions appear to occur among states located close to each other. Therefore, this paper emphasizes spatial spillover effects of regional recessions.

We use quarterly indicators of state economic activity based on Mexican data. Previous studies, such as those by Owyang et al. (2008) and Hamilton and Owyang (2012), use employment data for business cycle analysis. This is because in the U.S., monthly or quarterly economic activity indicators are not available at the city or state level. However, employment data may not accurately reflect the regional economic situations owing to labor market rigidities, especially when large regional differences exist within the labor markets themselves. In contrast, in Mexico, the data on economic activity of each state have been collected every quarter and published as an official indicator since the early 2000s.[2] Capitalizing on these data, we investigate how a state recession, caused by the

[1]The approach of Hamilton and Owyang (2012) can identify regional common factors of business cycles within Markov switching model. Regarding this, another major approach to regional business cycles is to apply dynamic factor model. For example, see Kose et al. (2003), Owyang et al. (2009), and Hirata et al. (2013).

[2]Note that a business cycle should reflect co-movements of a wide range of economic activities such as output, employment, sales (Stock and Watson, 1989). Based on the spirit of Burns and Mitchell (1946), as

global economic crisis of 2008–2009, spread to the neighboring states.

In the literature on Mexico, recent studies attempt to describe business cycles across states. Note that the synchronization between the Mexican and U.S. economies has long attracted economists' attention. For example, Chiquiar and Ramos-Francia (2005) investigated business cycle synchronization between the Mexican and U.S. manufacturing sectors. Although this type of analysis does not account for regional economic activities, it is possible that such synchronization explains the differences within the Mexican states' business cycles. In that sense, Mejía-Reyes and Campos-Chávez (2011) investigated the synchronization between the Mexican states and the output of U.S manufacturing operations, finding that Mexican states that occupy a relatively high share of the manufacturing sector are more strongly affected by U.S. manufacturing production. Besides, to determine the phases of regional business cycles and investigate regional differences during the economic crisis of 2008–2009, Erquizio-Espinal (2010) calculated the coincidence index across states, in the spirit of Burns and Mitchell (1946). He found that the border states, which are closely related to the U.S. economy, are more strongly affected by the U.S. recession. As emphasized in Mejía-Reyes and Erquizio-Espinal (2012), one of the causes of the Great Recession of 2008-2009 in Mexico was different in that instead of being caused by national business cycles, it was caused by the recession that came from beyond national borders (exogenous cause). This is consistent with Torres and Vela (2003) and Sosa (2008), who investigated business cycle synchronization between U.S. and Mexico before the Great Recession of 2008-2009 in Mexico. However, a regional recession caused by U.S. economy spread across Mexican states. Although some states did not experience recession in this period, recession in their neighboring states would have caused a slowdown in their economies (endogenous cause). Looking at the endogenous cause in regional business cycles, we try to estimate the extent to which Mexican regional economies slowed down through this transmission mechanism.[3]

This paper contributes to the literature by demonstrating how a region-specific shock propagates to neighboring economies. In the literature on spatial econometrics, Anselin (1988) established spatial econometric models that were able to account for spatial dependence and heterogeneity. In this paper, we connect regional interdependence with

emphasized by Stock and Watson (1989), it would not be precise to define business cycles just in terms of fluctuations in either GDP or employment. Nevertheless, a Markov switching model that uses the growth rates of GDP has provided an insightful view on business cycles by estimating unobservable expansionary and recessionary regimes.

[3]Delajara (2011) investigated the co-movement across states during the recession period of 2008–2009. He pointed out the possibility of geographical propagation although he did not provide direct evidence. This paper provides the lacking evidence to support his discussion.

the discussion on business cycles in a macroeconomic time-series analysis by developing a Markov switching model proposed by Hamilton (1989).[4] In estimating the model, we employ a Bayesian Markov chain Monte Carlo (MCMC) method.[5] LeSage and Pace (2009) described a Bayesian estimation methodology for spatial econometric models and applied the Metropolis-Hasting (MH) algorithm to estimate a parameter measuring spatial dependence. Hence, their study is notable in that the estimation method used is characterized by the MH algorithm, rather than by Gibbs sampling.[6] In this paper, we develop the Bayesian MCMC estimation method proposed by Kim and Nelson (1998, 1999b,a) by introducing a spatial autoregressive process.

The remainder of this paper is organized as follows. In Section 2, we describe the Markov switching model, including a discussion of the spatially lagged dependent variable. In Section 3, we present the Bayesian estimation procedure using MCMC. In Section 4, we present the data. In Section 5, we discuss the estimation results. In Section 6, we provide a numerical simulation of the spatial spillover effects of a regional recession. Finally, In Section 7, we present our conclusions.

2.2 Markov Switching Model with Spatial Lag

In this paper, we construct a Markov switching model with a spatially lagged dependent variable. Let $y_{t,n}$ denote the growth rate of an indicator of economic activity for region n at date t. Using vector notation, we denote $\boldsymbol{y}_t = (y_{t,1}, \ldots, y_{t,N})^\top$ as an $N \times 1$ vector, where N represents the number of regions. Let $\boldsymbol{W}$ denote an $N \times N$ spatial weight matrix (SWM). Thus, $\boldsymbol{W}\boldsymbol{y}_t$ indicates an $N \times 1$ spatially lagged dependent variable. Let $\boldsymbol{s}_t = (s_{t,1}, \ldots, s_{t,N})^\top$ denote an $N \times 1$ vector of the indicator variable that follows a Markov chain. If $s_{t,n} = 1$, then region n is in the expansion phase at date t, whereas $s_{t,n} = 0$ means that region n is in the recession phase at date t. The Markov switching model involving spatial lag is denoted

[4] To the best of my knowledge, Ohtsuka (2010) is the first to introduce a spatial autoregressive process into a standard Markov switching model, discovering that business cycles across Japanese regions are spatially dependent.

[5] An MCMC estimation methodology for a Markov switching model was first suggested by Albert and Chib (1993). However, there were questions related to how discrete, hidden variables ought to be sampled. By developing the single-move Gibbs sampling initially proposed by Albert and Chib (1993), Kim and Nelson (1998, 1999b,a) improved sampling efficiency by using multi-move Gibbs sampling for hidden variables. In the literature on regional business cycles, Owyang et al. (2005) and Owyang et al. (2008) also adopted the method proposed by Kim and Nelson for model estimation.

[6] This difference implies the need for modifications in the discussion on model selection. See Chib (1995) and Chib and Jeliazkov (2001) for more detailed discussions.

as follows:[7]

$$\boldsymbol{y}_t = \rho \boldsymbol{W} \boldsymbol{y}_t + \boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) + \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \ldots, T, \tag{2.1}$$

where ρ is a parameter measuring spatial dependence in the dependent variable; $\boldsymbol{\mu}_0 = (\mu_{1,0}, \ldots, \mu_{N,0})^\top$ is an $N \times 1$ vector of the average growth rate controlled by the spatial lag in the recession phase; $\boldsymbol{\mu}_1 = (\mu_{1,1}, \ldots, \mu_{N,1})^\top$ is an $N \times 1$ vector of the average growth rate controlled by the spatial lag in the expansion phase; $\boldsymbol{\iota}_N$ is an $N \times 1$ vector whose all elements are ones; $\boldsymbol{\varepsilon}_t = (\varepsilon_{t,1}, \ldots, \varepsilon_{t,N})^\top$ is an $N \times 1$ vector of the error terms that follows i.i.d. $\mathrm{N}(\mathbf{0}, \boldsymbol{\Omega})$; and $\odot$ denotes element-by-element multiplication.[8]

In model (2.1), we impose a restriction that $\mu_{n,1} > \mu_{n,0}$, $n = 1, \ldots, N$. This means that the average growth rate in the expansion phase is higher than that in the recession phase. We assume that error terms are independent across regions, that is, $\boldsymbol{\Omega} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)$ is an $N \times N$ diagonal matrix. We also assume that $s_{t,n}$ follows a first-order two-state Markov chain, indicating that the transition probabilities are $\Pr(s_{t,n} = i | s_{t-1,n} = j) = p_{n,ji}$, $i, j = 0, 1$.

The SWM is a row-standardized matrix constructed from the data. In general, data on contiguity or distance are used because the SWM is assumed to be exogenous. The SWM describes a spatial spillover network across regions. The spatial autoregressive parameter ρ lies in the interval of the inverses of minimum and maximum eigenvalues of the SWM. As shown in Ord (1975), the row-standardized SWM yields $+1$ as the largest eigenvalue. Therefore, we impose the restriction $1/\omega_{\min} < \rho < 1$, where $\omega_{\min}$ is the smallest eigenvalue of the SWM.[9] See Section 2.4 for a more detailed discussion of this.

Our Markov switching model includes a spatially lagged dependent variable $\boldsymbol{W}\boldsymbol{y}_t$. Therefore, the coefficient parameter ρ measures spatial dependence in regional business

[7] Based on Hamilton (2008), a multi-regional Markov switching model involving an autoregressive process of the dependent variable can be described as follows:

$$\boldsymbol{y}_t = \boldsymbol{\Phi} \boldsymbol{y}_{t-1} + \boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) + \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t + \boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\Phi} = \mathrm{diag}(\phi_1, \ldots, \phi_N)$ is an $N \times N$ diagonal matrix of parameters measuring the *temporal* dependence in the dependent variable. Thus, our model is considered to be a Markov switching model focusing on the *spatial* dependence in the dependent variable. To see which model fits data better, we will compare these models in Appendix 2.D.

[8] One may consider another possibility of a spatial autoregressive process, that is, $\boldsymbol{W}\boldsymbol{y}_{t-1}$ instead of $\boldsymbol{W}\boldsymbol{y}_t$. Consider a simpler version of model (2.1), $\boldsymbol{y}_t = \rho \boldsymbol{W} \boldsymbol{y}_{t-1} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t$, where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_N)$. By successive iteration, we can show that $\boldsymbol{y}_t = (\mathbf{I} - \rho \boldsymbol{W})^{-1} \boldsymbol{\mu} + (\mathbf{I} - \rho \boldsymbol{W})^{-1} \boldsymbol{\varepsilon}_t$, which is equivalent to $\boldsymbol{y}_t = \rho \boldsymbol{W} \boldsymbol{y}_t + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t$. Therefore, note that simultaneous spatial autoregressive process results from dynamic spatial dependence. See LeSage and Pace (2009, Chap. 2) for a discussion on time dependence in spatial econometric models.

[9] For the SWM used in this paper, $\omega_{\min}$ always takes a negative value, as it does in most cases. See also Anselin and Bera (1998) for a more detailed discussion.

cycles. An advantage of this model is that it allows us to simulate the spatial spillover effects for a particular region-specific shock. As described in Anselin (2003), if $\rho \neq 0$, spatial spillover effects exist. Thus, based upon (2.1), we numerically investigate how a recession occurring in a Mexican state during the period 2008–2009 affected the neighboring states' economies.

2.3 Bayesian Inference

Bayesian inference, such as point and interval estimates, is based on a posterior distribution. For convenience of explanation, we define here a parameter set $\theta = (\boldsymbol{\Omega}, \boldsymbol{\mu}, \boldsymbol{p}_{11}, \boldsymbol{p}_{00}, \rho)$, where $\boldsymbol{\mu} = (\boldsymbol{\mu}_0, \boldsymbol{\mu}_1)^\top$, $\boldsymbol{p}_{11} = (p_{1,11}, \dots, p_{N,11})^\top$, and $\boldsymbol{p}_{00} = (p_{1,00}, \dots, p_{N,00})^\top$. Let $\boldsymbol{Y} = \{y_{t,n}\}$ and $\boldsymbol{S} = \{s_{t,n}\}$ denote a $T \times N$ matrix each. Thus, by using Bayes' theorem, we can derive the following relationship regarding the posterior distribution:

$$\pi(\theta|\boldsymbol{Y}, \boldsymbol{S}) \propto L(\boldsymbol{Y}, \boldsymbol{S}|\theta)\pi(\theta), \tag{2.2}$$

where $\pi(\theta|\boldsymbol{Y}, \boldsymbol{S})$ is the posterior distribution, $L(\boldsymbol{Y}, \boldsymbol{S}|\theta)$ is the likelihood function, $\pi(\theta)$ is the prior distribution, and $\propto$ represents "is proportional to". An important result here is that the posterior distribution is proportional to the likelihood function multiplied by the prior distribution. We, therefore, need to specify prior distributions and to derive the likelihood function to conduct the Bayesian inference.

2.3.1 Prior Distributions and Likelihood Function

First, we specify the prior distributions of parameters. Thus, we use an inverse gamma distribution $\mathrm{IG}(\underline{\nu}/2, \underline{\delta}/2)$ as a prior for σ_n^2 as follows:

$$\pi(\sigma_n^2) \propto \left(\frac{1}{\sigma_n^2}\right)^{\underline{\nu}/2+1} \exp\left(\frac{\underline{\delta}}{2\sigma_n^2}\right).$$

As for a prior for $\boldsymbol{\mu}_n = (\mu_{n,0}, \mu_{n,1})^\top$, we adopt a bivariate normal distribution $\mathrm{N}_2(\underline{\boldsymbol{m}}, \underline{\boldsymbol{M}})$ as follows:

$$\pi(\boldsymbol{\mu}_n) \propto \exp\left(-\frac{1}{2}\left(\boldsymbol{\mu}_n - \underline{\boldsymbol{m}}\right)^\top \underline{\boldsymbol{M}}^{-1}\left(\boldsymbol{\mu}_n - \underline{\boldsymbol{m}}\right)\right),$$

where we impose a restriction $\mu_{n,1} > \mu_{n,0}$. Prior distributions for the transition probabilities $p_{n,11}$ and $p_{n,00}$ are set to have beta distributions $\mathrm{Beta}(\underline{\alpha}_{11}, \underline{\alpha}_{10})$ and $\mathrm{Beta}(\underline{\alpha}_{00}, \underline{\alpha}_{01})$, respectively:

$$\pi(p_{n,11}) \propto p_{n,11}^{\alpha_{11}-1}(1-p_{n,11})^{\alpha_{10}-1}, \quad \pi(p_{n,00}) \propto p_{n,00}^{\alpha_{00}-1}(1-p_{n,00})^{\alpha_{01}-1}.$$

Finally, we use a uniform distribution $\mathrm{U}(1/\omega_{\min}, 1)$ as a prior for ρ.

Next, we derive the likelihood function. To evaluate it, we consider two decomposed terms by utilizing $L(\boldsymbol{Y}, \boldsymbol{S}|\theta) = L(\boldsymbol{Y}|\theta, \boldsymbol{S})p(\boldsymbol{S}|\theta)$. From the assumption that $\boldsymbol{\varepsilon}_t \sim$ i.i.d. $\mathrm{N}(\mathbf{0}, \boldsymbol{\Omega})$ and by using variable transformation, the likelihood function conditional on $\boldsymbol{S}$ is given by

$$L(\boldsymbol{Y}|\theta, \boldsymbol{S}) = \prod_{t=1}^{T}\left[(2\pi)^{-N/2}|\boldsymbol{\Omega}|^{-1/2}|\mathbf{I}_N - \rho\boldsymbol{W}|\exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}_t^{\top}\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon}_t\right)\right], \tag{2.3}$$

where $\mathbf{I}_N$ is an $N \times N$ identity matrix and $\boldsymbol{\varepsilon}_t = (\mathbf{I}_N - \rho\boldsymbol{W})\boldsymbol{y}_t - \boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) - \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t$. As for the second term, the difficulty is that the hidden variables $\boldsymbol{S}$ are directly unobservable. To evaluate the likelihood function, we follow the methodology of Kim and Nelson (1998, 1999b,a). See Appendix 2.B for more details.

2.3.2 Posterior Distributions

Having specified the likelihood function and prior distributions, we are able to conduct the Bayesian inference. From (2.2), the joint posterior distribution is given by

$$\pi(\theta|\boldsymbol{Y}, \boldsymbol{S}) \propto L(\boldsymbol{Y}|\theta, \boldsymbol{S})p(\boldsymbol{S}|\theta)\pi(\rho)\pi(\boldsymbol{\Omega})\pi(\boldsymbol{\mu})\pi(\boldsymbol{p}_{11})\pi(\boldsymbol{p}_{00}),$$

where $p(\cdot)$ is a probability mass function; an independent joint prior distribution across parameters and regions is assumed. To facilitate explanation, let us define the following vectors with respect to region n: $\boldsymbol{\mu}_n = (\mu_{n,0}, \mu_{n,1})^{\top}$, $\boldsymbol{s}_n = (s_{1,n}, \ldots, s_{T,n})^{\top}$, and $\bar{\boldsymbol{y}}_n = (\sum_{m=1}^{N} w_{nm}y_{1,m}, \ldots, \sum_{m=1}^{N} w_{nm}y_{T,m})^{\top}$. However, a difficulty concerning sampling from the posterior distribution arises because a Markov switching model includes hidden variables $\boldsymbol{S}$. Therefore, we present a more detailed discussion for the sampling methodology.

The sampling methodology is based on the conditional posterior distributions. Although we assumed an independent joint prior, the conditional posterior distributions show the same family as the priors, except for parameter ρ.[10] Since the conditional posterior distribution for ρ takes an unknown distributional form, we rely on the MH algorithm. On the other hand, those for $\boldsymbol{\Omega}$, $\boldsymbol{\mu}$, $\boldsymbol{p}_{11}$, and $\boldsymbol{p}_{00}$ take known distributional forms, and thus, we

[10] The priors for $\boldsymbol{\Omega}$, $\boldsymbol{\mu}$, $\boldsymbol{p}_{11}$, $\boldsymbol{p}_{00}$ are conditionally conjugate.

apply the Gibbs sampler for posterior sampling of these parameters.[11]

We derive conditional posterior distributions distribution for each parameter below. The conditional posterior distribution for $\rho|\boldsymbol{Y}, \boldsymbol{S}, \boldsymbol{\mu}, \boldsymbol{\Omega}$ is given by

$$\pi(\rho|\boldsymbol{Y}, \boldsymbol{S}, \boldsymbol{\mu}, \boldsymbol{\Omega}) \propto \prod_{t=1}^{T} \left[|\mathbf{I}_N - \rho \boldsymbol{W}| \exp\left(-\frac{1}{2} \boldsymbol{\varepsilon}_t^\top(\rho) \boldsymbol{\Omega}^{-1} \boldsymbol{\varepsilon}_t(\rho) \right) \right], \tag{2.4}$$

where $\boldsymbol{\varepsilon}_n(\rho) = \boldsymbol{y}_n - \rho \bar{\boldsymbol{y}}_n - \mu_{n,0}(\boldsymbol{\iota}_T - \boldsymbol{s}_n) - \mu_{n,1}\boldsymbol{s}_n$. Note that ρ is independent of $p_{n,11}$ and $p_{n,00}$. As mentioned previously, this is an unknown distributional form. We, therefore, rely on the MH algorithm. In the literature on spatial econometrics, LeSage and Pace (2009) describe a sampling method for a parameter on spatial dependence using the MH algorithm. Our sampling follows their method. See Appendix 2.A for details.

As for the sampling for the other conditional posterior distribution, although there are slight changes, the sampling strategy is the same as the one proposed in Kim and Nelson (1998, 1999b,a). Samples from these conditional posteriors are drawn via the Gibbs sampler. First, the conditional posterior distribution for $\sigma_n^2|\boldsymbol{Y}, \boldsymbol{S}, \boldsymbol{\mu}_n, \rho$ takes the following form:

$$\pi(\sigma_n^2|\boldsymbol{Y}, \boldsymbol{S}, \boldsymbol{\mu}_n, \rho) \propto \left(\frac{1}{\sigma_n^2} \right)^{\overline{\nu}/2+1} \exp\left(\frac{\overline{\delta}_n}{2\sigma_n^2} \right),$$

where $\overline{\nu} = \underline{\nu} + T$ and $\overline{\delta}_n = \underline{\delta} + \boldsymbol{\varepsilon}_n^\top \boldsymbol{\varepsilon}_n$. Note that σ_n^2 is independent of $p_{n,11}$ and $p_{n,00}$. We can see that this conditional posterior for $\sigma_n^2|\boldsymbol{Y}, \boldsymbol{S}, \boldsymbol{\mu}_n, \rho$ is distributed as an inverse gamma distribution $\mathrm{IG}(\overline{\nu}/2, \overline{\delta}_n/2)$. Next, the conditional posterior distribution for $\boldsymbol{\mu}_n|\boldsymbol{Y}, \boldsymbol{S}, \sigma_n^2, \rho$ is obtained as follows:

$$\pi(\boldsymbol{\mu}_n|\boldsymbol{Y}, \boldsymbol{S}, \sigma_n^2, \rho) \propto \exp\left(-\frac{1}{2} (\boldsymbol{\mu}_n - \overline{\boldsymbol{m}}_n)^\top \overline{\boldsymbol{M}}_n^{-1} (\boldsymbol{\mu}_n - \overline{\boldsymbol{m}}_n) \right),$$

where

$$\begin{aligned}
\overline{\boldsymbol{M}}_n &= \left(\underline{\boldsymbol{M}}^{-1} + \sigma_n^{-2} \boldsymbol{X}_n^\top \boldsymbol{X}_n \right)^{-1}, \\
\overline{\boldsymbol{m}}_n &= \overline{\boldsymbol{M}}_n \left(\underline{\boldsymbol{M}}^{-1} \underline{\boldsymbol{m}} + \sigma_n^{-2} \boldsymbol{X}^\top (\boldsymbol{y}_n - \rho \bar{\boldsymbol{y}}_n) \right), \\
\boldsymbol{X}_n &= (\boldsymbol{\iota}_T - \boldsymbol{s}_n \;\; \boldsymbol{s}_n).
\end{aligned}$$

Note that $\boldsymbol{\mu}_n$ is independent of $p_{n,11}$ and $p_{n,00}$. The conditional posterior for $\boldsymbol{\mu}_n|\boldsymbol{Y}, \boldsymbol{S}, \sigma_n^2, \rho$ is distributed as a bivariate normal distribution with mean $\overline{\boldsymbol{m}}_n$ and variance $\overline{\boldsymbol{M}}_n$, that is, $\mathrm{N}_2(\overline{\boldsymbol{m}}_n, \overline{\boldsymbol{M}}_n)$. Finally, we derive the conditional posterior distributions for $p_{n,11}|\boldsymbol{Y}, \boldsymbol{S}$ and

[11] Exactly speaking, the Gibbs sampling is a special case of the MH algorithm.

$p_{n,00}|\boldsymbol{Y},\boldsymbol{S}$ that are given by

$$\pi(p_{n,11}|\boldsymbol{Y},\boldsymbol{S}) \propto p_{n,11}^{\overline{\alpha}_{11}-1}(1-p_{n,11})^{\overline{\alpha}_{10}-1}, \quad \pi(p_{n,00}|\boldsymbol{Y},\boldsymbol{S}) \propto p_{n,00}^{\overline{\alpha}_{00}-1}(1-p_{n,00})^{\overline{\alpha}_{01}-1},$$

where $\overline{\alpha}_{11} = \underline{\alpha}_{11} + n_{11}$, $\overline{\alpha}_{10} = \underline{\alpha}_{10} + n_{10}$, $\overline{\alpha}_{00} = \underline{\alpha}_{00} + n_{00}$, $\overline{\alpha}_{01} = \underline{\alpha}_{01} + n_{01}$, and n_{ji} is the number of transitions from state j to state i. Note that $p_{n,11}$ and $p_{n,00}$ are independent of σ_n^2, $\boldsymbol{\mu}_n$, and ρ. We see that $p_{n,11}|\boldsymbol{Y},\boldsymbol{S}$ and $p_{n,00}|\boldsymbol{Y},\boldsymbol{S}$ follow the beta distributions $\mathrm{Beta}(\overline{\alpha}_{11},\overline{\alpha}_{10})$ and $\mathrm{Beta}(\overline{\alpha}_{00},\overline{\alpha}_{01})$, respectively.

2.3.3 Drawing Parameters from Posterior Distributions

We conduct the Bayesian inference via multiple-block MH sampling.[12] Besides parameters, a Markov switching model includes hidden variables $\boldsymbol{S}$. Following Kim and Nelson (1998, 1999b,a), we use multi-move Gibbs sampling for drawing $\boldsymbol{s}_n$. In summary, our sampling algorithm is as follows:

1. Set hyperparameters and the initial values for parameters.
2. Draw $\boldsymbol{s}_n^{(g)}|\boldsymbol{Y},\theta^{(g-1)}$ $(n=1,\ldots,N)$ from the multi-move Gibbs sampling.[13]
3. Draw $p_{n,11}^{(g)}|\boldsymbol{Y},\boldsymbol{s}_n^{(g)}$ $(n=1,\ldots,N)$ from $\mathrm{Beta}(\overline{\alpha}_{11},\overline{\alpha}_{10})$.
4. Draw $p_{n,00}^{(g)}|\boldsymbol{Y},\boldsymbol{s}_n^{(g)}$ $(n=1,\ldots,N)$ from $\mathrm{Beta}(\overline{\alpha}_{00},\overline{\alpha}_{01})$.
5. Draw $\sigma_n^{2,(g)}|\boldsymbol{Y},\boldsymbol{s}_n^{(g)},\boldsymbol{\mu}_n^{(g-1)},\rho^{(g-1)}$ $(n=1,\ldots,N)$ from $\mathrm{IG}(\overline{\nu}/2,\overline{\delta}/2)$.
6. Draw $\boldsymbol{\mu}_n^{(g)}|\boldsymbol{Y},\boldsymbol{s}_n^{(g)},\sigma_n^{2,(g)},\rho^{(g-1)}$ $(n=1,\ldots,N)$ from $\mathrm{N}(\overline{\boldsymbol{m}}_n,\overline{\boldsymbol{M}}_n)$.
7. Draw $\rho^{(g)}|\boldsymbol{Y},\boldsymbol{S}^{(g)},\sigma_n^{2,(g)},\boldsymbol{\mu}_n^{(g)}$ based on the MH algorithm.[14]
    (a) Draw ρ' from a truncated normal distribution $\mathrm{TN}_{(1/\omega_{\min},1)}(\rho^{(g-1)},1)$ as a proposal distribution $q(\cdot)$.

[12] Our sampling method is also called Metropolis within Gibbs sampling, which indicates a hybrid sampler of the MH algorithm and the Gibbs sampling. We may interpret that parameters are drawn from the Gibbs sampling in Steps 2–6 and from the MH algorithm in Step 7. However, in line with Chib (2001, p. 3591), we use the notation of a multiple-block MH sampling because the Gibbs sampling is a special case of the multiple-block MH sampling.

[13] See Appendix 2.B for more details. In the process of the multi-move Gibbs sampling, it is also necessary to apply the Hamilton filter. See Appendix 2.C for details of the Hamilton filter.

[14] See Appendix 2.A for more details.

Table 2.1: Prior Distributions and Initial Values

Parameter	Prior Distribution	Hyperparameters	Initial Values
σ_n^2	$\mathrm{IG}(\underline{\nu}/2, \underline{\delta}/2)$	$\underline{\nu} = 6$; $\underline{\delta} = 0.4$	$\sigma_n^{2,(0)} = 1$
$(\mu_{0n}, \mu_{0n})^\top$	$\mathrm{N}_2(\underline{\boldsymbol{m}}, \underline{\boldsymbol{M}})$	$\underline{\boldsymbol{m}} = (-0.5, 0.5)^\top$; $\underline{\boldsymbol{M}} = \mathbf{I}_2$	$(\mu_{n0}^{(0)}, \mu_{n1}^{(0)})^\top = (-0.5, 0.5)^\top$
$p_{n,00}$	$\mathrm{Beta}(\underline{\alpha}_{00}, \underline{\alpha}_{01})$	$\underline{\alpha}_{00} = 8$; $\underline{\alpha}_{01} = 2$	$p_{n,00}^{(0)} = 0.8$
$p_{n,11}$	$\mathrm{Beta}(\underline{\alpha}_{11}, \underline{\alpha}_{10})$	$\underline{\alpha}_{11} = 9$; $\underline{\alpha}_{10} = 1$	$p_{n,11}^{(0)} = 0.8$
ρ	$\mathrm{U}(1/\omega_{\min}, 1)$		$\rho^{(0)} = 0$

Notes: $n = 1, 2, \ldots, N$. IG indicates an inverse gamma distribution. N_2 indicates a bivariate normal distribution. Beta indicates a beta distribution. U indicates a uniform distribution.

(b) Calculate the acceptance probability

$$\alpha(\rho^{(g-1)}, \rho') = \min\left[\frac{\pi(\rho'|\boldsymbol{Y}, \boldsymbol{S}^{(g)}, \sigma_n^{2,(g)}, \boldsymbol{\mu}_n^{(g)})q(\rho', \rho^{(g-1)})}{\pi(\rho^{(g-1)}|\boldsymbol{Y}, \boldsymbol{S}^{(g)}, \sigma_n^{2,(g)}, \boldsymbol{\mu}_n^{(g)})q(\rho^{(g-1)}, \rho')}, 1\right].$$

(c) Generate $u \sim \mathrm{U}(0, 1)$ and determine $\rho^{(g)}$ by using the following rule:

$$\rho^{(g)} = \begin{cases} \rho', & \text{if} \quad u \leq \alpha(\rho^{(g-1)}, \rho'), \\ \rho^{(g-1)}, & \text{otherwise.} \end{cases}$$

8. Repeat Steps 2–7.

In the above algorithm, superscript (g) refers to the sample from the posterior distributions obtained in the gth iteration. Hyperparameters and initial values for parameters are shown in Table 2.1. For the MH algorithm of parameter ρ, we use a truncated normal distribution $\mathrm{TN}_{(1/\omega_{\min}, 1)}(\rho^{(g-1)}, 1)$ as a proposal distribution. In the simulation of the posterior distributions, we discard the first 2,000 draws as a burn-in period. Descriptive statistics concerning the sampled posterior distributions are based on an additional 10,000 draws.

2.4 Data

In this paper, we use seasonally adjusted quarterly indicators of economic activity by state. The National Institute of Statistics and Geography (*Instituto Nacional de Estadística y Geografía*, INEGI) provides the Quarterly Indicator of State Economic Activity (*Indicador*

Trimestral de la Actividad Económica Estatal, ITAEE) on their Web site.[15] The period for which data is available is 2003:Q1–2012:Q1 for all 31 states and the Federal District (Distrito Federal). Although the ITAEE provides only data on aggregated sectors (primary, secondary, and tertiary; financial intermediation services indirectly measured; and total economic activity), it offers more disaggregated information on state economic activity with respect to the time-series dimension. As regards the estimation of business cycles, the ITAEE is the best proxy for the Gross State Product (GSP) and hence, can be used to capture regional business cycles. We, therefore, use growth rates of the total economic activity from the ITAEE.

To estimate a Markov switching model with the spatial lag, we need to specify the SWM in advance. For robustness, we used several types of SWMs, such as those based on contiguity and distance. The contiguity-based SWM is created from the contiguity matrix, in which the nmth element takes a value of 1 if states n and m share the same border, and 0 otherwise. The distance-based SWM takes the following form:

$$w_{nm} = \frac{d_{nm}^{-\eta}}{\sum_{m=1}^{N} d_{nm}^{-\eta}},$$

where w_{nm} is the nmth element of the SWM, d_{nm} is a bilateral distance between states n and m, N is the number of states, and η is a distance decay parameter. The bilateral distance is based on the great-circle distance between two states measured by latitude and longitude (Vincenty, 1975).[16] All types of SWMs are row-standardized. Note that the SWM describes how a spatial spillover spreads across regions. In this paper, we report the estimation results using a distance-based SWM ($\eta = 4$).[17]

2.5 Estimation Results

In Table 2.2, we present the point and interval estimates of parameters drawn from the posterior distributions.[18] Our particular interest is in whether spatial dependence exists

[15] Owyang et al. (2008) and Hamilton and Owyang (2012) use employment data owing to data limitations. In the Mexican context, the quarterly indicator of state economic activity is readily available and is a more appropriate measure than employment data in the formal sector.

[16] Data on latitude and longitude of each state capital are available from *Annual Statistics of United Mexican States: Edition 2005* (*Anuario Estadístico de los Estados Unidos Mexicanos, Edicíon 2005*), published by INEGI.

[17] The estimate of ρ obtained from this value was close to the average estimate of ρ obtained among $\eta = \{1, \ldots, 10\}$. In addition, we prefer the distance-based SWM to contiguity-based one because the former can account for continuous space across regions.

[18] All numerical analyses in this paper were conducted on Ox Console 7.01 (Doornik and Ooms, 2006).

Table 2.2: Estimated Parameters

		ρ		
		Mean	Median	95% CI
Spatial Dependence		0.28	0.28	$[0.23, 0.32]$

		μ_0			μ_1		
Code	State	Mean	Median	95% CI	Mean	Median	95% CI
1	Aguascalientes	-1.25	-1.32	$[-3.12, 0.60]$	1.15	1.15	$[0.41, 1.91]$
2	Baja California	-1.34	-1.34	$[-3.08, 0.28]$	0.64	0.65	$[-0.06, 1.35]$
3	Baja California Sur	-0.27	-0.14	$[-2.19, 1.01]$	1.02	1.00	$[0.28, 1.85]$
4	Campeche	-1.35	-1.30	$[-2.47, -0.53]$	-0.16	-0.31	$[-1.26, 1.69]$
5	Coahuila	-0.70	-0.61	$[-2.53, 0.69]$	0.77	0.74	$[-0.31, 2.05]$
6	Colima	-0.49	-0.41	$[-2.20, 0.80]$	1.12	1.06	$[0.14, 2.42]$
7	Chiapas	-0.64	-0.56	$[-2.26, 0.55]$	0.57	0.52	$[-0.46, 1.88]$
8	Chihuahua	-1.03	-0.84	$[-3.34, 0.43]$	0.64	0.62	$[-0.22, 1.64]$
9	Distrito Federal	-1.70	-1.74	$[-3.77, 0.27]$	0.53	0.52	$[0.11, 0.97]$
10	Durango	-0.61	-0.49	$[-2.15, 0.30]$	0.44	0.37	$[-0.15, 1.57]$
11	Guanajuato	-0.48	-0.37	$[-2.07, 0.64]$	0.76	0.69	$[-0.04, 2.05]$
12	Guerrero	-0.63	-0.57	$[-1.94, 0.32]$	0.55	0.52	$[-0.02, 1.32]$
13	Hidalgo	-0.87	-0.76	$[-3.03, 0.68]$	0.90	0.89	$[0.08, 1.78]$
14	Jalisco	-1.40	-1.24	$[-3.95, 0.43]$	0.62	0.62	$[-0.01, 1.30]$
15	México	-1.65	-1.78	$[-3.01, 0.34]$	1.03	1.04	$[0.55, 1.47]$
16	Michoacán	-0.68	-0.53	$[-2.51, 0.48]$	0.52	0.50	$[-0.13, 1.35]$
17	Morelos	-0.41	-0.28	$[-1.97, 0.52]$	0.59	0.55	$[0.00, 1.43]$
18	Nayarit	-0.50	-0.42	$[-2.18, 0.79]$	0.85	0.81	$[-0.19, 2.15]$
19	Nuevo León	-1.11	-1.14	$[-2.92, 0.61]$	1.01	1.01	$[0.31, 1.70]$
20	Oaxaca	-0.33	-0.17	$[-2.02, 0.54]$	0.64	0.53	$[-0.04, 2.10]$
21	Puebla	-0.70	-0.60	$[-2.63, 0.77]$	1.00	0.98	$[0.08, 2.01]$
22	Querétaro	-1.79	-1.82	$[-3.30, 0.05]$	1.30	1.31	$[0.73, 1.85]$
23	Quintana Roo	-0.79	-0.66	$[-3.06, 0.95]$	1.32	1.30	$[0.17, 2.57]$
24	San Luis Potosí	-0.93	-0.88	$[-2.79, 0.61]$	0.90	0.90	$[0.16, 1.71]$
25	Sinaloa	-0.50	-0.39	$[-2.26, 0.69]$	0.75	0.72	$[-0.12, 1.88]$
26	Sonora	-0.49	-0.38	$[-2.37, 0.84]$	0.94	0.91	$[0.08, 1.92]$
27	Tabasco	-0.25	-0.11	$[-2.19, 1.09]$	1.08	1.07	$[0.33, 1.94]$
28	Tamaulipas	-0.69	-0.59	$[-2.30, 0.46]$	0.81	0.75	$[-0.05, 2.07]$
29	Tlaxcala	-0.90	-0.78	$[-2.72, 0.35]$	0.55	0.52	$[-0.17, 1.47]$
30	Veracruz	-0.36	-0.22	$[-2.08, 0.67]$	0.71	0.66	$[0.02, 1.73]$
31	Yucatán	-0.26	-0.17	$[-2.10, 1.05]$	1.23	1.20	$[0.61, 1.99]$
32	Zacatecas	-0.34	-0.22	$[-2.15, 0.90]$	0.90	0.87	$[0.08, 1.91]$

Notes: 95% CI indicates 95% credible interval.

across regional business cycles, which is tested by the parameter ρ. The point estimates, namely mean and median, take the value of 0.28 and the interval estimate, namely 95% credible interval, shows a range of $[0.23, 0.32]$. Our estimation results, therefore, show that

Table 2.3: Average Growth Rates and Sectoral Shares during Recession and Expansion Phases

	Dependent Variable: Average Growth Rate μ			
	Recession		Expansion	
Explanatory Variables	Coeff.	Std.Err.	Coeff.	Std.Err.
Agriculture, Forestry, and Fishing	0.019	(0.015)	0.000	(0.010)
Mining	0.035	(0.024)	0.048^{***}	(0.014)
Manufacture	-0.018^{*}	(0.009)	0.010^{*}	(0.006)
Construction	0.073	(0.051)	0.005	(0.034)
Electricity, Gas, and Water	-0.017	(0.031)	-0.004	(0.015)
Commerce, Restaurants, and Hotels	-0.026^{**}	(0.011)	-0.005	(0.007)
Transport, Warehouse, and Communication	0.016	(0.026)	0.045^{**}	(0.017)
Financial Service, Security, and Real Estate	0.000	(0.020)	0.016	(0.013)
Common, Social, and Personal Service	-0.035	(0.025)	-0.006	(0.015)
Number of Observations	30		30	
R^2	0.888		0.946	

Notes: Heteroskedasticity-consistent standard errors are in the parenthesis. * denotes statistical significance at the 10% level, **, at the 5% level, and ***, at the 1% level. The states of Campeche and Quintana Roo are excluded from the sample. The explanatory variables are state-wise sectoral shares to GDP, and the constant term is suppressed because the sum of these shares is equal to one. The uncentered R^2 is used because no constant term exists in the regression models.

business cycles across Mexican states are spatially dependent.[19]

From Table 2.2, we can see that average growth rates controlled by the spatial lag in the recession and expansion phases differ considerably between states.[20] To see more, we regress the average growth rates in the recession and expansion phases on the GSP shares by industry (% of GSP).[21] Table 2.3 shows the estimation results.[22] We find that the manufacturing sector tends to have a negative impact on average growth rates in the recession phase, whereas it has a positive impact in the expansion phase. As discussed in Erquizio-Espinal (2010), we can see that although the border states have been growing

[19] The short time-period in our data might cause a bias in estimates. For robustness, we also used the monthly state data of the coincidence index calculated by Erquizio-Espinal (2010) himself and used in his paper. The data span is 2003:Jan–2011:Sep. From the estimation results of model (2.1), we confirmed that spatial dependence in regional business cycles existed in his data as well because the point and interval estimates of ρ were 0.30 and $[0.27, 0.33]$, respectively.

[20] The complete estimation results are available in Appendix 2.F.

[21] In this regression, the dependent variables are average growth rates between 2003:Q1–2012:Q1. To avoid endogeneity, we follow the methodology adopted by Owyang et al. (2008), and thus, we use the values of 2003 as explanatory variables.

[22] The constant term is suppressed because the sum of shares becomes 1. Campeche and Quintana Roo are excluded because in these states, the sectors of mining and the sector of commerce, restaurants, and hotels are comparatively large, respectively.

rapidly by exporting manufactured goods, in the recession phase, their economies have been more adversely affected.[23] Besides, in the recession phase, a higher share of the construction sector mitigates a decline in the growth rate. In the expansion phase, higher shares of mining, and commerce, warehouse, and communication show positive impacts on the growth rate.

Following Owyang et al. (2005), we provide Table 2.4 that shows when each state was in recession during the period 2003:Q1–2012:Q1. Table 2.4 is constructed from the probabilities of recession by state in Figure 2.1.[24] We define that a state is in recession at date t if the probability of recession during the quarterly periods $t-1$ to t takes a value of 0.5 and over.[25] The national recession period recorded by the INEGI was 2008:Apr–2009:Jun, and thus the recessions that occurred in most of the states fell within that range as well.[26] In addition, we should note that some states were in recession during the national recession period, whereas others were not.[27] According to our estimation results, ten states (Baja California Sur, Chiapas, Nayarit, Oaxaca, Sinaloa, Sonora, Tabasco, Veracruz, Yucatán, and Zacatecas) did not experience any recession during the investigation period.[28] A relatively high proportion of agricultural activity or relatively low proportion of manufacturing in these states might explain why they did not experience recessionary conditions, even during the economic crisis of 2008–2009. The above-mentioned results are similar to those of Erquizio-Espinal (2010) and Mejía-Reyes and Erquizio-Espinal (2012), who calculated the recession resistance index by state. They found that the states of Chiapas, Oaxaca, Sinaloa, Tabasco, and Zacatecas demonstrated a relatively high level of endurance against recession.

A noteworthy result in Table 2.4 is that the time at which a state enters recession differs

[23] Erquizio-Espinal (2010) emphasizes this point as "heads and tails" of globalization. Our results also confirm his statement that economies facing globalization are likely to fluctuate.

[24] Figure 2.1 only shows probabilities of recession for selective states. The results for all the states are available in Appendix 2.F.

[25] The probability of recession is calculated by $1 - G^{-1}\sum_{g=1}^{G} s_{t,n}^{(g)}$, where G is the number of iterations, and the superscript (g) is the gth iteration.

[26] Another national recession period is 2000:Aug–2003:Sep. However, we were not able to identify state recession for that period owing to data limitations.

[27] Campeche shows a different movement from the other states. Annual growth rates of real GSP also show highly negative values, such as −5.28% in 2006–2007, −2.96% in 2007–2008, −9.44% in 2008–2009, −4.41% in 2009–2010, and −3.61% in 2010–2011. This tendency coincides with our estimation results in Table 2.4.

[28] Note that our results might not identify state recessions in their entirety. For example, Nayarit, Sinaloa, and Sonora may be considered to have been in a recession during 2008–2009. Their annual growth rates of real GSP for this period were −3.64%, −5.12%, and −5.01%, respectively. Determining whether states are in recession might depend upon the threshold of the probability of recession.

Table 2.4: State Recessions

Year	2003				2004				2005				2006				2007				2008				2009				2010				2011				2012
Quarterly	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	I
1 Aguascalientes																							■	■	■												
2 Baja California																							■	■	■	■											
3 Baja California Sur																																					
4 Campeche									■		■		■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
5 Coahuila																									■												
6 Colima																								■													
7 Chiapas																																					
8 Chihuahua																							■		■	■											
9 Distrito Federal																									■												
10 Durango																							■	■													
11 Guanajuato																								■	■												
12 Guerrero																					■	■	■	■	■	■											
13 Hidalgo																								■	■	■											
14 Jalisco																									■												
15 México															■									■	■	■											
16 Michoacán																								■													
17 Morelos																					■	■															
18 Nayarit																																					
19 Nuevo León																							■	■	■												
20 Oaxaca																																					
21 Puebla																								■	■												
22 Querétaro																							■	■	■	■											
23 Quintana Roo																								■	■	■											
24 San Luis Potosí																							■	■	■												
25 Sinaloa																																					
26 Sonora																																					
27 Tabasco																																					
28 Tamaulipas																						■	■	■	■	■	■										
29 Tlaxcala									■	■																											
30 Veracruz																																					
31 Yucatán																																					
32 Zacatecas																																					

Notes: The black bar represents state recession. We define that a state is in recession at date t when its recession probability during the period $t-1$ to t exceeds 0.5. Shaded areas correspond to the dates of recessions by INEGI.

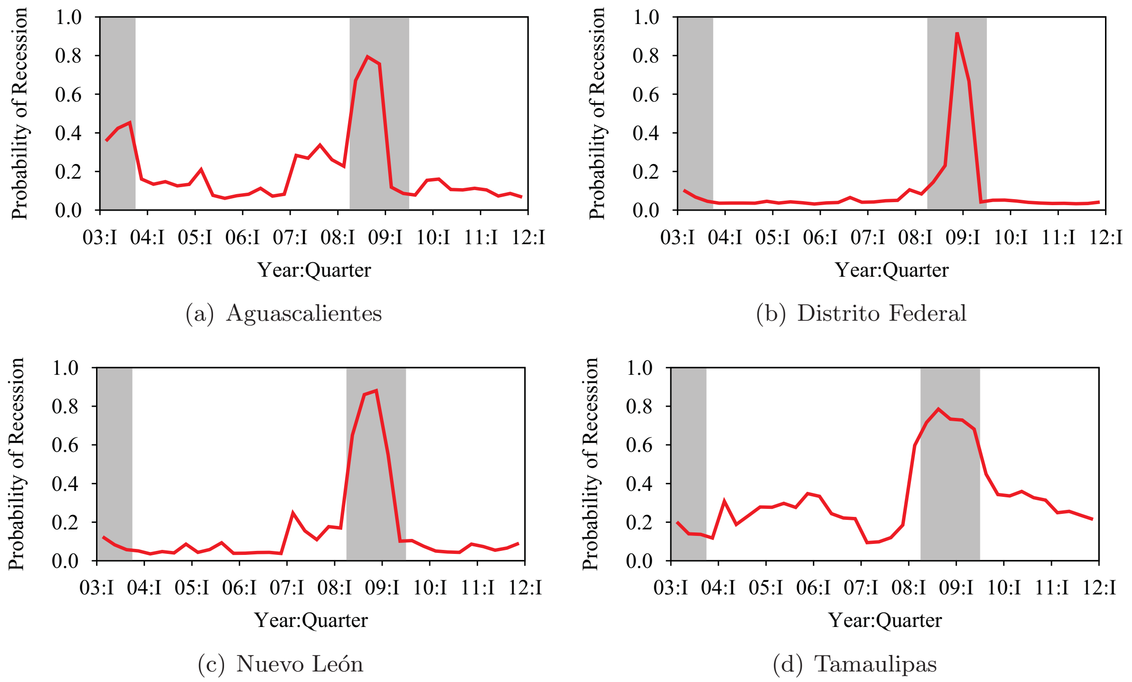


Figure 2.1: Probabilities of Recession for Selected States

Notes: Shaded areas correspond to the dates of recessions by INEGI.

between states. Figure 2.2 shows how the recession of 2008–2009 spread across states. As we have shown that inter-state business cycles are spatially dependent, Figure 2.2 also suggests that spatial proximity affects the propagation process of recession. Thus, we observe that in 2008:Q1–Q2, Guerrero, Moreros, and Tamaulipas had all entered a recessionary period. The recessions seem to propagate toward their neighboring states from 2008:Q3 through 2009:Q1. To evaluate the extent to which regional recession causes conditions in neighboring economies to deteriorate, we conduct a numerical simulation of spatial spillover effects, as discussed in Section 2.6.

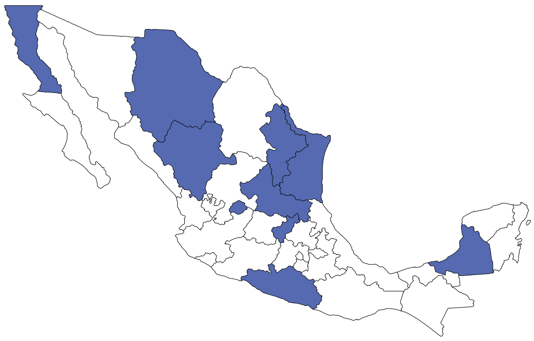
2.6 Numerical Simulations of Spatial Spillover Effects

An advantage of the spatial autoregressive model employed here is that it enables us to simulate spatial spillover effects.[29] Based on the above-mentioned estimation results, we

[29] Anselin (2003) notes that the model with a spatially lagged dependent variable has a global spillover effect that is characterized by an infinite series.

(a) 2008:Q1

(b) 2008:Q2



(c) 2008:Q3

(d) 2008:Q4

(e) 2009:Q1

(f) 2009:Q2

Figure 2.2: Regional Spread of Recession

Notes: A colored state is in recession.

conduct a numerical simulation of spatial spillover effects, thereby analyzing both the process by which and the extent to which a state's switch to a recessionary regime causes deterioration in the neighboring states' economies. First, we derive the form of spatial spillover effects that are decomposed into each wave. From the Markov switching model

(2.1), we have the following equation:

$$\boldsymbol{y}_t = (\mathbf{I}_N - \rho\boldsymbol{W})^{-1}\big(\boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) + \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t\big) + (\mathbf{I}_N - \rho\boldsymbol{W})^{-1}\boldsymbol{\varepsilon}_t,$$

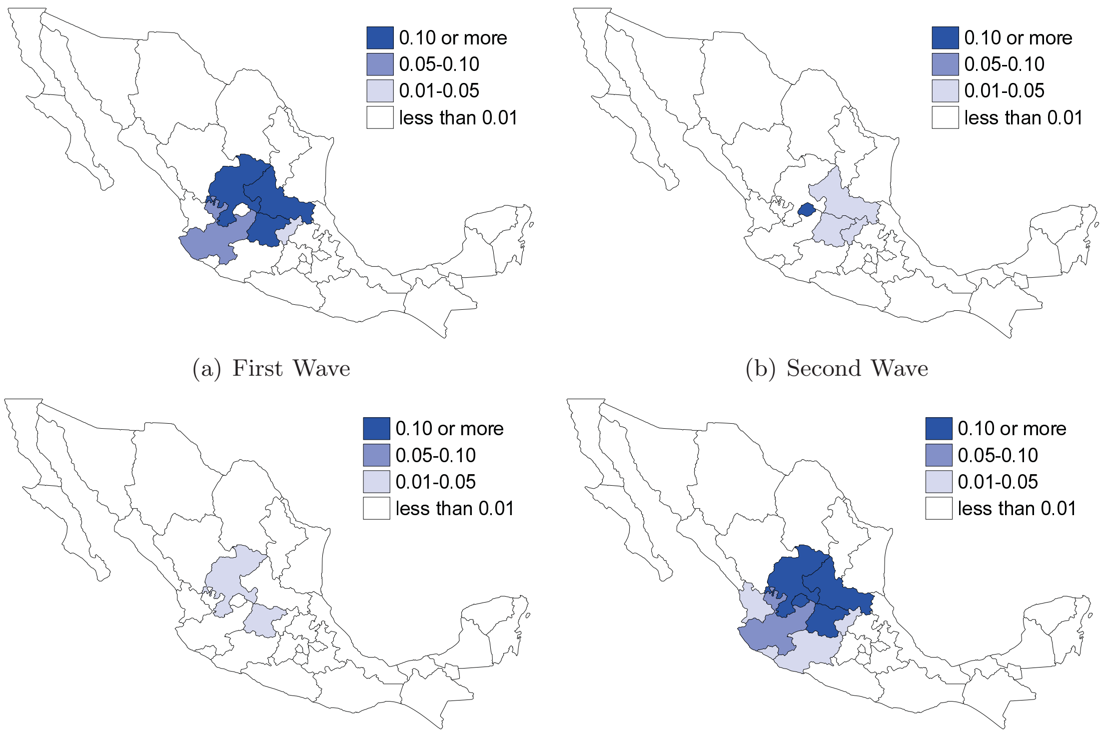
where $(\mathbf{I}_N - \rho\boldsymbol{W})^{-1}$ is a global spatial multiplier that generates spatial spillover effects (Anselin, 2003). We should note that there are no spillover effects when $\rho = 0$. Let $\Delta s_{t,n}$ denote a regime switch from $s_{t-1,n} = 1$ to $s_{t,n} = 0$ and $\Delta\boldsymbol{y}_t(\equiv \boldsymbol{y}_t - \boldsymbol{y}_{t-1})$ denote an $N \times 1$ vector of difference in growth rates. The impact of a switch to a recessionary regime can be calculated by

$$\begin{aligned}\frac{\Delta\boldsymbol{y}_t}{\Delta s_{t,n}} &= \underbrace{-(\mathbf{I}_N - \rho^*\boldsymbol{W})_n^{-1\top} \times (\mu^*_{n,1} - \mu^*_{n,0})}_{\text{Cumulative Effect}} \\ &= -\big(\mathbf{I}_N + \underbrace{\rho^*\boldsymbol{W}}_{\text{1st Wave}} + \underbrace{(\rho^*\boldsymbol{W})^2}_{\text{2nd Wave}} + \underbrace{(\rho^*\boldsymbol{W})^3}_{\text{3rd Wave}} + \dots\big)_n^\top \times (\mu^*_{n,1} - \mu^*_{n,0}),\end{aligned} \tag{2.5}$$

where $(\mathbf{I}_N - \rho^*\boldsymbol{W})_n^{-1\top}$ is the nth column vector of the matrix, and ρ^*, $\mu^*_{n,0}$, and $\mu^*_{n,1}$ are the posterior means. The first line indicates the cumulative spillover effects, and the second line shows that the cumulative effect can be decomposed by the infinite geometric series. Our numerical simulations are conducted based on equation (2.5).[30]

As a simulation exercise, we show decomposed and cumulative spillover effects of regime switches from expansion to recession. First, let us explore the case of Aguascalientes, where the manufacturing has a higher share in GSP. Second, we analyze the case of Tamaulipas, which was in recession in an earlier phase during the economic crisis of 2008–2009. Figure 2.3 shows the case of Aguascalientes. A regime switch to recession affected the economies of Guanajuato, Jalisco, Querétaro, San Luis Potosí, and Zacatecas in the first wave. Subsequently, although the impact was small, the economic deterioration caused by the shock in Aguascalientes remained in Guanajuato, Querétaro, and San Luis Potosí during the second wave. The interesting finding was that Aguascalientes itself was also affected by feedback-loop effects. During the third wave, the magnitude of economic shock diminished significantly, but was observed in Guanajuato and Zacatecas. Cumulative spillover effects are shown in Panel (d) of Figure 2.3. According to Table 2.4, Zacatecas did not experience recession in 2008–2009. However, we can see that the recession in Aguascalientes caused a slowdown in the economy of Zacatecas. Figure 2.4 shows the case of Tamaulipas. As before, a recession in Tamaulipas diffused toward the relatively distant neighboring economies of Aguascalientes, Coahuila, Hidalgo, Guanajuato, Querétaro, San Luis Potosí, Nuevo León,

[30] In this framework, the spillover effects are symmetric between the two regimes of economic recession and expansion.



(a) First Wave

(b) Second Wave

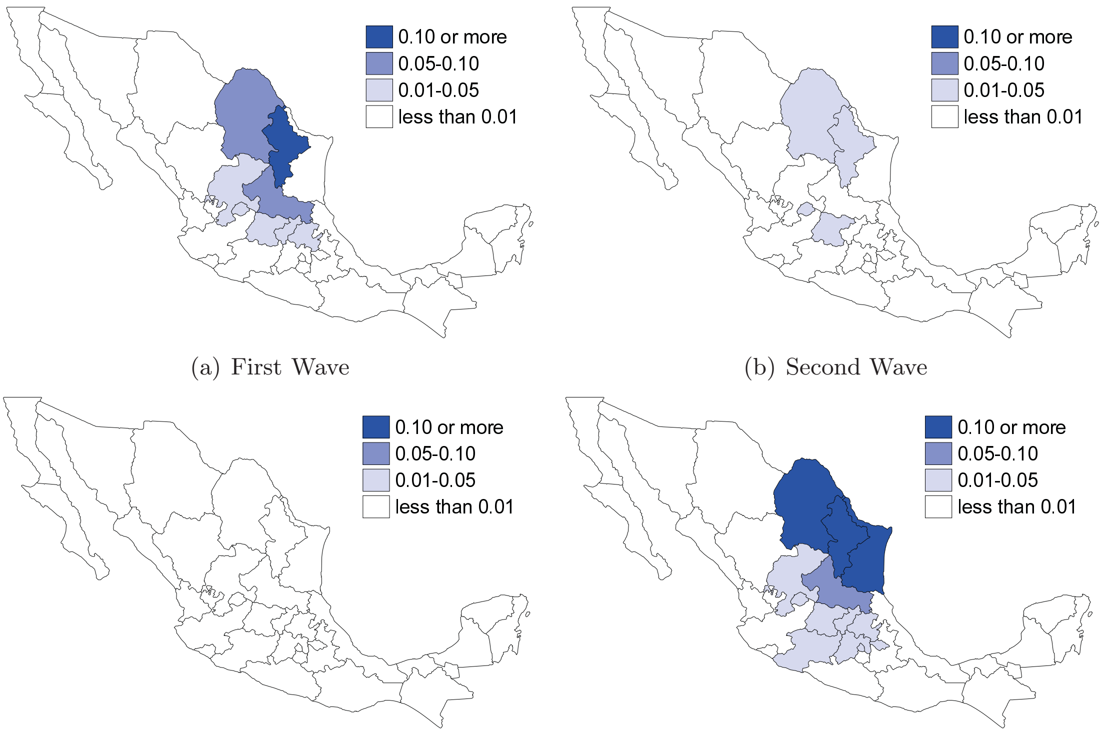
(c) Third Wave

(d) Cumulative Effect

Figure 2.3: Numerical Simulation of Spillover Effects (Case of Aguascalientes)

and Zacatecas during the first wave. Furthermore, the economic shock occurring in San Luis Potosí spread toward Aguascalientes, Coahuila, Guanajuato, and Nuevo León during the second wave. The spatial spillover effects become negligibly small during the third wave. Panel (d) of Figure 2.4 shows the cumulative spillover effects.

Furthermore, we calculated cumulative spillover effects for all states; these are summarized in Table 2.5. As discussed in Section 2.5, the states with a higher share in manufacturing to GSP, such as Aguascalientes, Baja California, Mexico, Nuevo León, Puebla, and Querétaro had greater negative impacts on the neighboring states' economies when a regime switch from expansion to recession occurred. Since the second-largest impacts were comparably small, we can see that the spillover effects diminished toward the more distant states. Note that Distrito Federal, Mexico, and Quintana Roo still had greater effects on their second-nearest neighboring states. In summary, our numerical simulation described above demonstrates that a state's regime switch from expansion to recession reduces the quarterly growth rate of the nearest neighboring state by an average of 0.28% and that of



(a) First Wave

(b) Second Wave

(c) Third Wave

(d) Cumulative Effect

Figure 2.4: Numerical Simulation of Spillover Effects (Case of Tamaulipas)

the second-nearest neighboring state by an average of 0.10%.

2.7 Concluding Remarks

The motivation for this paper derives from the idea that spatial proximity—which facilitates business with neighboring economies through commuting, migration, and trade—might result in spatial similarities in regional business cycles. In such situations, region-specific recessions would affect the neighboring economies. Thus, to investigate the regional propagation process, we introduced a spatial autoregressive process into a Markov switching model. This framework enabled a numerical simulation of the spatial spillover effects. Thus, using data from Mexican states, we conducted numerical simulations to investigate how the economic crisis occurring in a Mexican state during the period 2008–2009 affected the neighboring states' economies.

We showed that a parameter measuring spatial dependence takes a positive value; this

Table 2.5: Ranking of Cumulative Spillover Effects by State

		First			Second		
Code	Origin State	Code	Destination State	%	Code	Destination State	%
1	Aguascalientes	32	Zacatecas	−0.39	24	San Luis Potosí	−0.13
2	Baja California	26	Sonora	−0.38	8	Chihuahua	−0.08
3	Baja California Sur	25	Sinaloa	−0.26	10	Durango	−0.05
4	Campeche	31	Yucatán	−0.30	23	Quintana Roo	−0.04
5	Coahuila	19	Nuevo León	−0.44	28	Tamaulipas	−0.00
6	Colima	14	Jalisco	−0.30	18	Nayarit	−0.05
7	Chiapas	27	Tabasco	−0.34	4	Campeche	−0.01
8	Chihuahua	25	Sinaloa	−0.14	26	Sonora	−0.10
9	Distrito Federal	15	México	−0.31	17	Morelos	−0.30
10	Durango	32	Zacatecas	−0.09	18	Nayarit	−0.06
11	Guanajuato	22	Querétaro	−0.18	24	San Luis Potosí	−0.07
12	Guerrero	17	Morelos	−0.15	15	México	−0.08
13	Hidalgo	9	Distrito Federal	−0.25	29	Tlaxcala	−0.15
14	Jalisco	6	Colima	−0.18	1	Aguascalientes	−0.15
15	México	9	Distrito Federal	−0.46	17	Morelos	−0.37
16	Michoacán	22	Querétaro	−0.14	11	Guanajuato	−0.10
17	Morelos	9	Distrito Federal	−0.16	15	México	−0.14
18	Nayarit	14	Jalisco	−0.20	1	Aguascalientes	−0.07
19	Nuevo León	5	Coahuila	−0.63	28	Tamaulipas	−0.01
20	Oaxaca	21	Puebla	−0.08	29	Tlaxcala	−0.07
21	Puebla	29	Tlaxcala	−0.50	9	Distrito Federal	−0.01
22	Querétaro	11	Guanajuato	−0.47	16	Michoacán	−0.18
23	Quintana Roo	4	Campeche	−0.39	31	Yucatán	−0.30
24	San Luis Potosí	11	Guanajuato	−0.22	1	Aguascalientes	−0.18
25	Sinaloa	10	Durango	−0.14	3	Baja California Sur	−0.11
26	Sonora	8	Chihuahua	−0.14	3	Baja California Sur	−0.09
27	Tabasco	7	Chiapas	−0.37	4	Campeche	−0.02
28	Tamaulipas	19	Nuevo León	−0.14	5	Coahuila	−0.11
29	Tlaxcala	21	Puebla	−0.42	9	Distrito Federal	−0.01
30	Veracruz	29	Tlaxcala	−0.16	21	Puebla	−0.15
31	Yucatán	4	Campeche	−0.41	23	Quintana Roo	−0.04
32	Zacatecas	1	Aguascalientes	−0.28	24	San Luis Potosí	−0.05
			Average	−0.28		Average	−0.10

Notes: Based on calculation from equation (2.5).

provides evidence that spillover effects exist across Mexican states, and a region-specific shock thus causes deterioration in the neighboring states' economies through these effects. Our findings from the numerical simulations show that a regime switch from expansion to recession decreases the quarterly growth rate of economic activity for the nearest state by an average of 0.28%. However, the spatial spillover effects have only limited impact on the economies of distant states. As such, this paper emphasizes that geographical proximity does matter in regional business cycles. Therefore, our results have important implications for policymakers. For example, if a regional economy begins experiencing an economic downturn, the neighboring economies are also likely to slow down through the propagation process. Therefore, economic cooperation with neighboring state governments may be a solution for quicker recovery from a recession. In addition, spatial spillover effects are also seen in the expansion period. Knowing this will shed light on which regions fully enjoy the benefits of growth.

Finally, one limitation of this paper is that we have not looked at the sources that strengthen spatial dependence in regional business cycles, such as trade, migration, and capital flows. It would be possible to replace the geography-based SWM by the economic-distance-based SWM, although the latter SWM is no longer exogenous variable and the endogeneity must be controlled for. Thus, further research clarifying these should be undertaken.

Appendix 2.A Drawing ρ by Metropolis-Hastings Algorithm

We use a truncated normal distribution as a proposal distribution. When the random variable x has the truncated normal distribution $\mathrm{TN}_{(a,b)}(\mu, \sigma^2)$, the p.d.f. is

$$q(x) = \begin{cases} \dfrac{(1/\sigma)\phi\big((x-\mu)/\sigma\big)}{\Phi\big((b-\mu)/\sigma\big) - \Phi\big((a-\mu)/\sigma\big)}, & \text{if} \quad a < x < b, \\ 0, & \text{otherwise,} \end{cases}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density function (p.d.f.) and the cumulative distribution function (c.d.f.) of the standard normal distribution, respectively.

We use the probability integral transformation method for sampling from the truncated normal distribution. We set $\mu = \rho^{(g-1)}$, $\sigma^2 = 1$, $a = 1/\omega_{\min}$, and $b = 1$. Following Holloway et al. (2002), we introduce a tuning parameter c into the variance term, so that the acceptance rate might fall within the interval $[0.3, 0.7]$.[31] When u is distributed as a uniform distribution $\mathrm{U}(0,1)$, we can draw ρ' from $\mathrm{TN}_{(1/\omega_{\min},1)}(\rho^{(g-1)}, 1)$ as follows:

$$\rho' = \rho^{(g-1)} + c\Phi^{-1}\left(\Phi\big(1/\omega_{\min} - \rho^{(g-1)}\big) + u\left[\Phi\big(1-\rho^{(g-1)}\big) - \Phi\big(1/\omega_{\min} - \rho^{(g-1)}\big)\right]\right).$$

[31] Holloway et al. (2002) originally set the interval to $[0.4, 0.6]$, and LeSage and Pace (2009, Chap. 5) adopted the same strategy. We chose a slightly wider interval of acceptance rate. The aim of tuning the proposals is to ensure that the MH sampling moves over the entire conditional distribution. Thus, we adjust the tuning parameter c by using the following way. First, we set $c = 1$ as an initial value. Next, the tuning parameter c is adjusted by scale factor 1.01 depending on the acceptance rate ($c \times 1.01$ if the acceptance rate exceeds 0.7, while $c/1.01$ if the acceptance rate falls below 0.3).

The acceptance probability $\alpha(\rho^{(g-1)}, \rho')$ is calculated by

$$\alpha(\rho^{(g-1)}, \rho') = \min \left[\frac{\pi(\rho'|\boldsymbol{Y}, \boldsymbol{S}^{(g)}, \boldsymbol{\Omega}^{(g)}, \boldsymbol{\mu}^{(g)}) \left(\Phi\big(1-\rho^{(g-1)}\big) - \Phi\big(1/\omega_{\min} - \rho^{(g-1)}\big)\right)}{\pi(\rho^{(g-1)}|\boldsymbol{Y}, \boldsymbol{S}^{(g)}, \boldsymbol{\Omega}^{(g)}, \boldsymbol{\mu}^{(g)}) \left(\Phi\big(1-\rho'\big) - \Phi\big(1/\omega_{\min} - \rho'\big)\right)},\ 1 \right],$$

where $\pi(\rho|\boldsymbol{Y}, \boldsymbol{S}^{(g)}, \boldsymbol{\Omega}^{(g)}, \boldsymbol{\mu}^{(g)})$ is calculated from (2.4). Because a standard normal distribution is symmetric, $\phi(\rho', \rho^{(g-1)})$ and $\phi(\rho^{(g-1)}, \rho')$ are offset. Following step 7(c) in the algorithm, we judge whether ρ' is accepted or not.

Appendix 2.B Multi-Move Gibbs Sampling for $\boldsymbol{s}_n$

Kim and Nelson (1998, 1999b,a) are the first to apply a multi-move Gibbs sampling to a Markov switching model. Our explanation here is based on Kim and Nelson (1999b). For convenience of explanation, we define vectors $\tilde{\boldsymbol{s}}_n^t$ and $\boldsymbol{s}_n^t$, and a matrix $\tilde{\boldsymbol{Y}}^t$ by the following notation:

$$\tilde{\boldsymbol{s}}_n^t = \begin{pmatrix} s_{1,n} \\ s_{2,n} \\ \vdots \\ s_{t,n} \end{pmatrix}, \quad \boldsymbol{s}_n^t = \begin{pmatrix} s_{t,n} \\ s_{t+1,n} \\ \vdots \\ s_{T,n} \end{pmatrix}, \quad \tilde{\boldsymbol{Y}}^t = \begin{pmatrix} y_{1,n} & y_{1,2} & \cdots & y_{1,N} \\ y_{2,n} & y_{2,2} & \cdots & y_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ y_{t,n} & y_{t,2} & \cdots & y_{t,N} \end{pmatrix}.$$

The aim here is to obtain $p(\tilde{\boldsymbol{s}}_n^T|\tilde{\boldsymbol{Y}}^T, \theta)$. This can be rewritten as follows:

$$\begin{aligned} p(\tilde{\boldsymbol{s}}_n^T|\tilde{\boldsymbol{Y}}^T, \theta) &= p(s_{T,n}|\tilde{\boldsymbol{Y}}^T, \theta) p(\tilde{\boldsymbol{s}}_n^{T-1}|s_{T,n}, \tilde{\boldsymbol{Y}}^T, \theta) \\ &= p(s_{T,n}|\tilde{\boldsymbol{Y}}^T, \theta) \prod_{t=1}^{T-1} p(s_{t,n}|\boldsymbol{s}_n^{t+1}, \tilde{\boldsymbol{Y}}^T, \theta). \end{aligned} \tag{2.6}$$

Furthermore, the second term can be expressed as follows:

$$p(s_{t,n}|\boldsymbol{s}_n^{t+1}, \tilde{\boldsymbol{Y}}^T, \theta) \propto p(s_{t+1,n}|s_{t,n}, \theta) p(s_{t,n}|\tilde{\boldsymbol{Y}}^t, \theta),$$

where the first term on the RHS represents the transition probability. Incorporating the normalizing constant, we have the following probability mass function:

$$p(s_{t,n} = i|\boldsymbol{s}_n^{t+1}, \tilde{\boldsymbol{Y}}^T, \theta) = \frac{p(s_{t+1,n}|s_{t,n} = i, \theta) p(s_{t,n} = i|\tilde{\boldsymbol{Y}}^t, \theta)}{\sum_{j=0}^{1} p(s_{t+1,n}|s_{t,n} = j, \theta) p(s_{t,n} = j|\tilde{\boldsymbol{Y}}^t, \theta)}, \tag{2.7}$$

where $p(s_{t,n} = i|\tilde{\boldsymbol{Y}}^t, \theta)$ is calculated by using the Hamilton filter.[32] The calculation step for

[32] See Appendix 2.C for details.

(2.6) can be summarized as follows: first, we draw $s_{T,n}$ conditional on $\tilde{\boldsymbol{Y}}^T$ and θ. Second, given $s_{T,n}$, the sampling $s_{t,n}$ for $t = T-1, \ldots, 1$ is implemented by the backward recursion based on equation (2.7)

Appendix 2.C Hamilton Filter with Spatial Lag

Hamilton's (1989) filter is applied to calculate the conditional probabilities $p(s_{t,n} = i|\tilde{\boldsymbol{Y}}^t, \theta)$ for region n at date t. Based on Chib (1996, 2001), we explain how the Hamilton filter is applied in the case where a spatially lagged dependent variable is included in a Markov switching model.

Using scalar notation, the model (2.1) can be rewritten as follows:

$$y_{t,n} = \rho \sum_{m=1}^{N} w_{nm} y_{t,m} + \mu_{n,0}(1 - s_{t,n}) + \mu_{n,1} s_{t,n} + \varepsilon_{t,n}, \quad \varepsilon_{t,n} \sim \text{i.i.d. N}(0, \sigma_n^2).$$

For the conditional p.d.f. $f(y_{t,n}|s_{t,n}, \boldsymbol{y}_{t,-n}, \theta)$, the expected value and variance become $\mathrm{E}(y_{t,n}|s_{t,n}, \boldsymbol{y}_{t,-n}, \theta) = \rho \sum_{m=1}^{N} w_{nm} y_{t,m} + \mu_{n,0}(1-s_{t,n}) + \mu_{n,1} s_{t,n}$ and $\mathrm{Var}(y_{t,n}|s_{t,n}, \boldsymbol{y}_{t,-n}, \theta) = \sigma_n^2$, where the subscript $-n$ indicates that the nth element is excluded from a vector.[33] Therefore, the conditional p.d.f., which is used in the iteration process of the Hamilton filter, is given by

$$f(y_{t,n}|s_{t,n}, \boldsymbol{y}_{t,-n}, \theta) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{\left(y_{t,n} - \rho \sum_{m=1}^{N} w_{nm} y_{t,m} - \mu_{n,0}(1 - s_{t,n}) - \mu_{n,1} s_{t,n}\right)^2}{2\sigma_n^2} \right).$$

The algorithm of the Hamilton filter consists of two steps: prediction and update. The conditional p.d.f. $p(s_{t,n} = i|\tilde{\boldsymbol{Y}}^t, \theta)$ is available by the forward recursion $t = 1, 2, \ldots, T$.

1. Prediction Step: Calculate the probability

$$p(s_{t,n} = i|\tilde{\boldsymbol{Y}}^{t-1}, \theta) = \sum_{j=0}^{1} p(s_{t,n} = i|s_{t-1,n} = j, \theta) p(s_{t-1,n} = j|\tilde{\boldsymbol{Y}}^{t-1}, \theta),$$

where, when $t = 1$, $p(s_{0,n} = i|\tilde{\boldsymbol{Y}}^0, \theta)$ is replaced by the steady-state probabilities as follows:

$$\pi_{n,0} = \frac{1 - p_{n,11}}{2 - p_{n,00} - p_{n,11}}, \quad \pi_{n,1} = \frac{1 - p_{n,00}}{2 - p_{n,00} - p_{n,11}}.$$

[33] For simplicity, we impose a restriction on the calculation of expected value and variance. It is assumed that for each region n, spatial lag $\sum_{m=1}^{N} w_{nm} y_{t,m}$ is exogenously given.

2. Update Step: Calculate the probability

$$p(s_{t,n} = i|\tilde{\boldsymbol{Y}}^t, \theta) = \frac{f(y_{t,n}|s_{t,n} = i, \boldsymbol{y}_{t,-n}, \theta)p(s_{t,n} = i|\tilde{\boldsymbol{Y}}^{t-1}, \theta)}{\sum_{j=0}^{1} f(y_{t,n}|s_{t,n} = j, \boldsymbol{y}_{t,-n}, \theta)p(s_{t,n} = j|\tilde{\boldsymbol{Y}}^{t-1}, \theta)}.$$

The probabilities $p(s_{t,n} = i|\tilde{\boldsymbol{Y}}^t, \theta)$ are used in the multi-move Gibbs sampling. The probabilities $p(s_{t,n} = i|\tilde{\boldsymbol{Y}}^{t-1}, \theta)$ are also used for calculating the likelihood function in the model selection.

Appendix 2.D Model Selection

To compare different econometric models, we calculate estimates of log marginal likelihood. Chib (1995) proposes a procedure for calculating marginal likelihood in the Gibbs sampling. However, in this paper, a parameter measuring spatial dependence ρ is drawn by the MH algorithm, and thus, we employ a method proposed by Chib and Jeliazkov (2001).

The calculation of the marginal likelihood is based on the following equation:

$$m(\boldsymbol{Y}) = \frac{L(\boldsymbol{Y}|\theta)\pi(\theta)}{\pi(\theta|\boldsymbol{Y})},$$

which is termed basic marginal likelihood identity (BMI). The BMI consists of the likelihood function, prior distribution, and posterior distribution. This identity holds at any θ. In this paper, the mean of the posterior distribution θ^* is used. Thus, by taking the logarithms of the BMI and evaluating them at θ^*, we can calculate log marginal likelihood estimate as follows:

$$\log \hat{m}(\boldsymbol{Y}) = \log L(\boldsymbol{Y}|\theta^*) + \log \pi(\theta^*) - \log \hat{\pi}(\theta^*|\boldsymbol{Y}). \tag{2.8}$$

Based on (2.8), we calculate the following three terms: the likelihood function, the prior distribution, and the posterior distribution, all evaluated at θ^*.

The first term on the RHS of (2.8) is the log likelihood function. Note that the Markov switching model includes hidden variables $\{\boldsymbol{s}_t\}_{t=1}^T$. The likelihood function, therefore, takes the following form:

$$L(\boldsymbol{Y}|\theta^*) = \prod_{t=1}^{T} \left[\sum_{j=0}^{1} f(\boldsymbol{y}_t|\boldsymbol{s}_t = j, \theta^*)p(\boldsymbol{s}_t = j|\tilde{\boldsymbol{Y}}^{t-1}, \theta^*) \right].$$

The second term in the brackets must be calculated in advance. This term can be obtained from the prediction step in the Hamilton filter.

The second term on the RHS of (2.8) is the logarithm of the joint prior distribution. Since we assumed independent prior distribution across parameters and regions, the prior distribution can be obtained as follows:

$$\pi(\theta^*) = \pi(\rho^*)\left[\prod_{n=1}^{N}\pi(\sigma_n^{2*})\pi(\boldsymbol{\mu}_n^*)\pi(p_{n,11}^*)\pi(p_{n,00}^*)\right].$$

The third term on the RHS of (2.8) is the logarithm of the joint posterior distribution, which can be rewritten as follows:

$$\begin{aligned}\hat{\pi}(\theta^*|\boldsymbol{Y}) = \hat{\pi}(\rho^*|\boldsymbol{Y})\Bigg[\prod_{n=1}^{N}&\hat{\pi}(\sigma_n^{2*}|\rho^*,\boldsymbol{Y})\hat{\pi}(\boldsymbol{\mu}_n^*|\rho^*,\sigma_n^{2*},\boldsymbol{Y})\\ &\times\hat{\pi}(p_{n,11}^*|\rho^*,\sigma_n^{2*},\boldsymbol{\mu}_n^*,\boldsymbol{Y})\hat{\pi}(p_{n,00}^*|\rho^*,\sigma_n^{2*},\boldsymbol{\mu}_n^*,p_{n,11}^*,\boldsymbol{Y})\Bigg],\end{aligned}$$

where

$$\begin{aligned}\hat{\pi}(\rho^*|\boldsymbol{Y}) &= \frac{G^{-1}\sum_{g=1}^{G}\alpha(\rho^{(g)},\rho^*|\boldsymbol{Y},\boldsymbol{S}^{(g)},\boldsymbol{\Omega}^{(g)},\boldsymbol{\mu}^{(g)},p_{n,11}^{(g)},p_{n,00}^{(g)})q(\rho^{(g)},\rho^*)}{J^{-1}\sum_{k=1}^{J}\alpha(\rho^*,\rho^{(k)}|\boldsymbol{Y},\boldsymbol{S}^{(k)},\boldsymbol{\Omega}^{(k)},\boldsymbol{\mu}^{(k)},p_{n,11}^{(k)},p_{n,00}^{(k)})},\\ \hat{\pi}(\sigma_n^{2*}|\rho^*,\boldsymbol{Y}) &= \frac{1}{J}\sum_{k=1}^{J}\pi(\sigma_n^{2*}|\rho^*,\boldsymbol{\mu}_n^{(k)},p_{n,11}^{(k)},p_{n,00}^{(k)},\boldsymbol{s}_n^{(k)},\boldsymbol{Y}),\\ \hat{\pi}(\boldsymbol{\mu}_n^*|\rho^*,\sigma_n^{2*},\boldsymbol{Y}) &= \frac{1}{J}\sum_{k=1}^{J}\pi(\boldsymbol{\mu}_n^*|\rho^*,\sigma_n^{2*},p_{n,11}^{(k)},p_{n,00}^{(k)},\boldsymbol{s}_n^{(k)},\boldsymbol{Y}),\\ \hat{\pi}(p_{n,11}^*|\rho^*,\sigma_n^{2*},\boldsymbol{\mu}_n^*,\boldsymbol{Y}) &= \frac{1}{J}\sum_{k=1}^{J}\pi(p_{n,11}^*|\rho^*,\sigma_n^{2*},\boldsymbol{\mu}_n^*,p_{n,00}^{(k)},\boldsymbol{s}_n^{(k)},\boldsymbol{Y}),\\ \hat{\pi}(p_{n,00}^*|\rho^*,\sigma_n^{2*},\boldsymbol{\mu}_n^*,p_{n,11}^*,\boldsymbol{Y}) &= \frac{1}{J}\sum_{k=1}^{J}\pi(p_{n,00}^*|\rho^*,\sigma_n^{2*},\boldsymbol{\mu}_n^*,p_{n,11}^*,\boldsymbol{s}_n^{(k)},\boldsymbol{Y}).\end{aligned}$$

The superscript (g) refers to the sample from the posterior distribution in the gth iteration and (k) refers to the sample from the reduced Gibbs runs obtained in the kth iteration. Note that some of the parameters are given as a mean in the reduced Gibbs runs, and that $\rho^{(k)}$ is drawn from a proposal distribution $q(\rho^*,\rho^{(k)})$. Besides the G iterations, we need to implement additional $4\times J$ iterations for the reduced Gibbs runs. The first reduced run is for the denominator of $\hat{\pi}(\rho^*|\boldsymbol{Y})$ and $\hat{\pi}(\sigma_n^{2*}|\rho^*,\boldsymbol{Y})$; the second is for $\hat{\pi}(\boldsymbol{\mu}_n^*|\rho^*,\sigma_n^{2*},\boldsymbol{Y})$; the third is for $\hat{\pi}(p_{n,11}^*|\rho^*,\sigma_n^{2*},\boldsymbol{\mu}_n^*,\boldsymbol{Y})$; and the fourth is for $\hat{\pi}(p_{n,00}^*|\rho^*,\sigma_n^{2*},\boldsymbol{\mu}_n^*,p_{n,11}^*,\boldsymbol{Y})$. We set J to have the same number of iterations as G. Moreover, the numerical standard

Table 2.6: Log Marginal Likelihood Estimate

	Spatial Dependence ρ		Log Marginal Likelihood	
Model Fitted	Mean	95%CI	Estimate	NSE
MS-SAR				
SWM: Contiguity	0.308	$[0.262, 0.355]$	-2583.384	(0.514)
SWM: Distance ($\eta = 4$)	0.278	$[0.231, 0.323]$	-2591.102	(0.337)
MS			-2669.300	(0.207)
MS-AR(1)			-2603.581	(0.239)
AR(1)			-2646.631	(0.003)
MS-SAR-AR(1)				
SWM: Contiguity	0.333	$[0.282, 0.383]$	-2487.344	(0.293)
SWM: Distance ($\eta = 4$)	0.286	$[0.235, 0.335]$	-2496.824	(0.548)

Notes: $G = J = 10,000$. SWM indicates a spatial weight matrix. η is a distance decay parameter. 95% CI indicates 95% credible interval. NSE indicates the numerical standard errors of the marginal likelihood estimates. Models indicated in the table are as follows: MS-SAR: $\boldsymbol{y}_t = \rho \boldsymbol{W} \boldsymbol{y}_t + \boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) + \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t + \boldsymbol{\varepsilon}_t$, MS: $\boldsymbol{y}_t = \boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) + \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t + \boldsymbol{\varepsilon}_t$, MS-AR(1): $\boldsymbol{y}_t = \boldsymbol{\Phi} \boldsymbol{y}_{t-1} + \boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) + \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t + \boldsymbol{\varepsilon}_t$, AR(1): $\boldsymbol{y}_t = \boldsymbol{\Phi} \boldsymbol{y}_{t-1} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t$, MS-SAR-AR(1): $\boldsymbol{y}_t = \rho \boldsymbol{W} \boldsymbol{y}_t + \boldsymbol{\Phi} \boldsymbol{y}_{t-1} + \boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) + \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t + \boldsymbol{\varepsilon}_t$

errors of the marginal likelihood estimates are also calculated.[34] See Chib and Jeliazkov (2001) for more details.

Table 2.6 presents the log marginal likelihood estimates with numerical standard errors by using the different econometric models. First of all, it is useful to compare estimates of log marginal likelihood between Markov switching model with a spatial autoregressive process (MS-SAR) and Markov switching (MS) model because MS is a spatial case of MS-SAR when $\rho = 0$. Consequently, it is supported to take into account spatial dependence in regional business cycles. As explained in Hamilton (2008), Markov switching model with a first-order autoregressive process (MS-AR(1)) is an extended model of a first-order autoregressive process (AR(1)), which becomes MS-AR(1) if the constant term shows regime switch. Our estimation results say that MS-AR(1) fits data better than AR(1). Then, compared with MS-AR(1), we can see that MS-SAR is slightly supported. In the Mexican economy, therefore, spatial dependence in regional business cycles matters, rather than temporal dependence. Additionally, we estimated Markov switching model with a spatial autoregressive process and a first-order autoregressive process (MS-SAR-AR(1)), which is naturally supported against MS-SAR or MS-AR(1).[35]

[34]For calculation of the numerical standard errors, we need to select a lag at which the autocorrelation is small enough to be neglected. Thus, we used the lag length 40.

[35]The estimation results of MS-SAR (distance-based SWM), MS, MS-AR(1), and MS-SAR-AR(1) are available in Appendix 2.F.

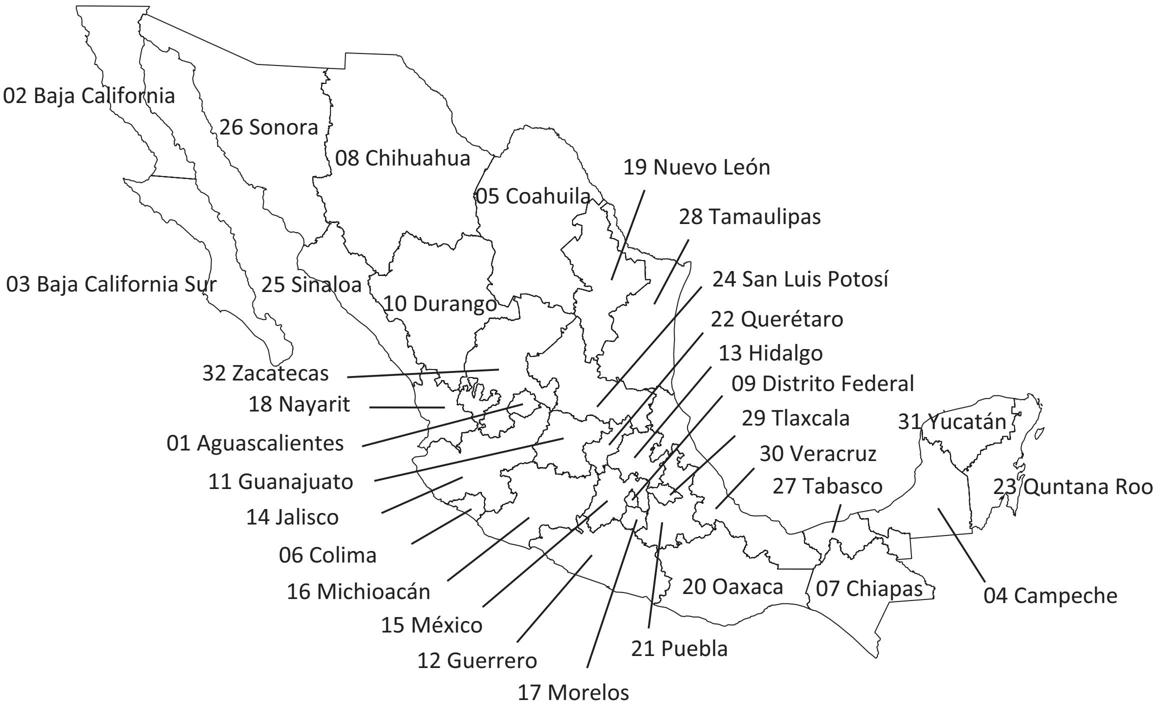


Figure 2.5: Map of Mexico

Appendix 2.E Map of Mexico

State codes and names appear in Figure 2.5.

Appendix 2.F Details of Data and Estimation Results

This appendix provides details on data and estimation results and contains five subsections.

2.F.1 Data

Figure 2.6 shows the Quarterly Indicator of State Economic Activity (*Indicador Trimestral de la Actividad Económica Estatal*, ITAEE) from 2003:Q1 to 2012Q1. Figure 2.7 shows the percentage changes of ITAEE, which are calculated by $[\log(y_{t,n}) - \log(y_{t-1,n})] \times 100$.

2.F.2 Estimation Results of Markov Switching Model with Spatial Autoregressive Process (MS-SAR); Distance-Based SWM ($\eta = 4$)

Table 2.7 shows the point estimates and interval estimates of parameters. Figure 2.8 shows the probabilities of recession, which are calculated by $1 - G^{-1} \sum_{g=1}^{G} s_{t,n}^{(g)}$, where G is the number of iterations and the superscript (g) is the gth iteration.

2.F.3 Estimation Results of Markov Switching Model (MS)

The estimation results here are obtained by estimating the standard Markov switching model:

$$\boldsymbol{y}_t = \boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) + \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t + \boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\varepsilon}_t \sim$ i.i.d. $\mathrm{N}(\mathbf{0}, \boldsymbol{\Omega})$ and $\boldsymbol{\Omega} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)$. Table 2.8 shows the point estimates and interval estimates of parameters. Figure 2.9 shows the probabilities of recession, which are calculated by $1-G^{-1} \sum_{g=1}^{G} s_{t,n}^{(g)}$, where G is the number of iterations and the superscript (g) is the gth iteration.

2.F.4 Estimation Results of Markov Switching Model with First Order Autoregressive Process (MS-AR(1))

The estimation results here are obtained by estimating the standard Markov switching model:

$$\boldsymbol{y}_t = \boldsymbol{\Phi} \boldsymbol{y}_t + \boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) + \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t + \boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\Phi} = \mathrm{diag}(\phi_1, \ldots, \phi_N)$, $\boldsymbol{\varepsilon}_t \sim$ i.i.d. $\mathrm{N}(\mathbf{0}, \boldsymbol{\Omega})$, and $\boldsymbol{\Omega} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)$. Table 2.9 shows the point estimates and interval estimates of parameters. Figure 2.10 shows the probabilities of recession, which are calculated by $1-G^{-1} \sum_{g=1}^{G} s_{t,n}^{(g)}$, where G is the number of iterations and the superscript (g) is the gth iteration.

2.F.5 Estimation Results of Markov Switching Model with Spatial Autoregressive and First Order Autoregressive Processes (MS-SAR-AR(1)); Distance-Based SWM ($\eta = 4$)

The estimation results here are obtained by estimating the standard Markov switching model:

$$\boldsymbol{y}_t = \rho \boldsymbol{W} \boldsymbol{y}_t + \boldsymbol{\Phi} \boldsymbol{y}_{t-1} + \boldsymbol{\mu}_0 \odot (\boldsymbol{\iota}_N - \boldsymbol{s}_t) + \boldsymbol{\mu}_1 \odot \boldsymbol{s}_t + \boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\Phi} = \mathrm{diag}(\phi_1, \ldots, \phi_N)$, $\boldsymbol{\varepsilon}_t \sim$ i.i.d. $\mathrm{N}(\mathbf{0}, \boldsymbol{\Omega})$, and $\boldsymbol{\Omega} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)$. Table 2.10 shows the point estimates and interval estimates of parameters. Figure 2.11 shows the probabilities of recession, which are calculated by $1 - G^{-1} \sum_{g=1}^{G} s_{t,n}^{(g)}$, where G is the number of iterations and the superscript (g) is the gth iteration.

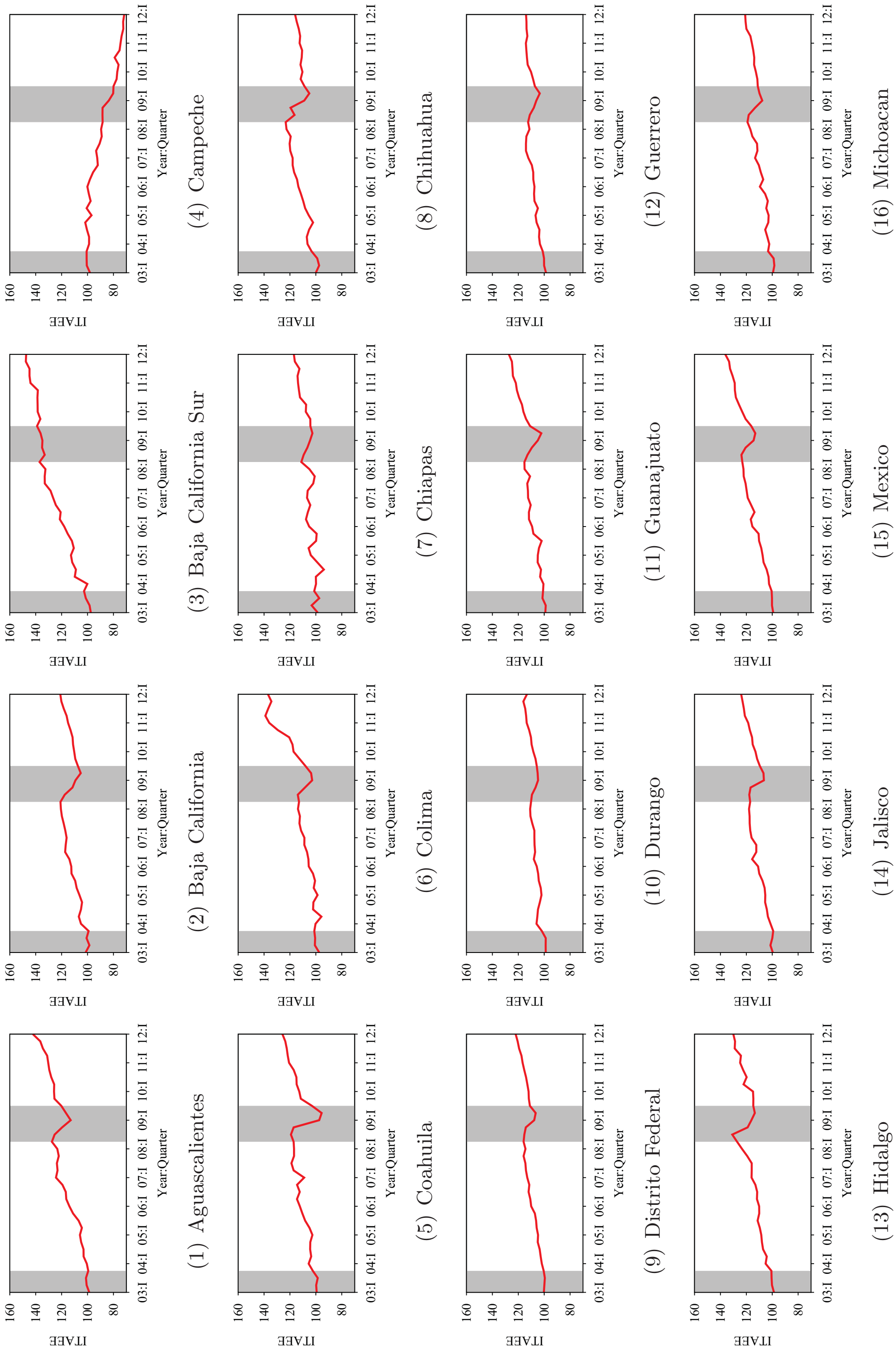


Figure 2.6: ITAEE

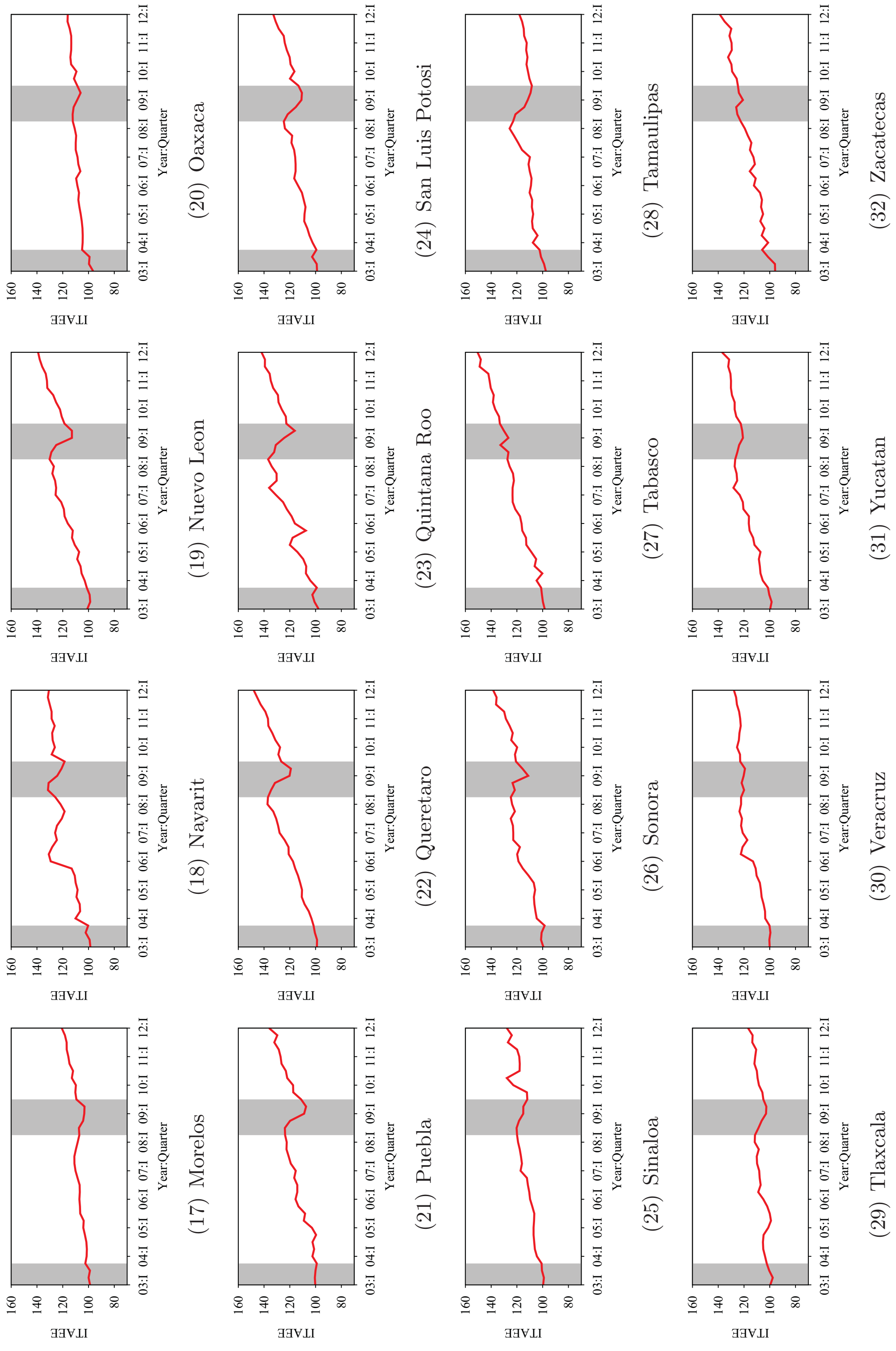


Figure 2.6: ITAEE (Continued)

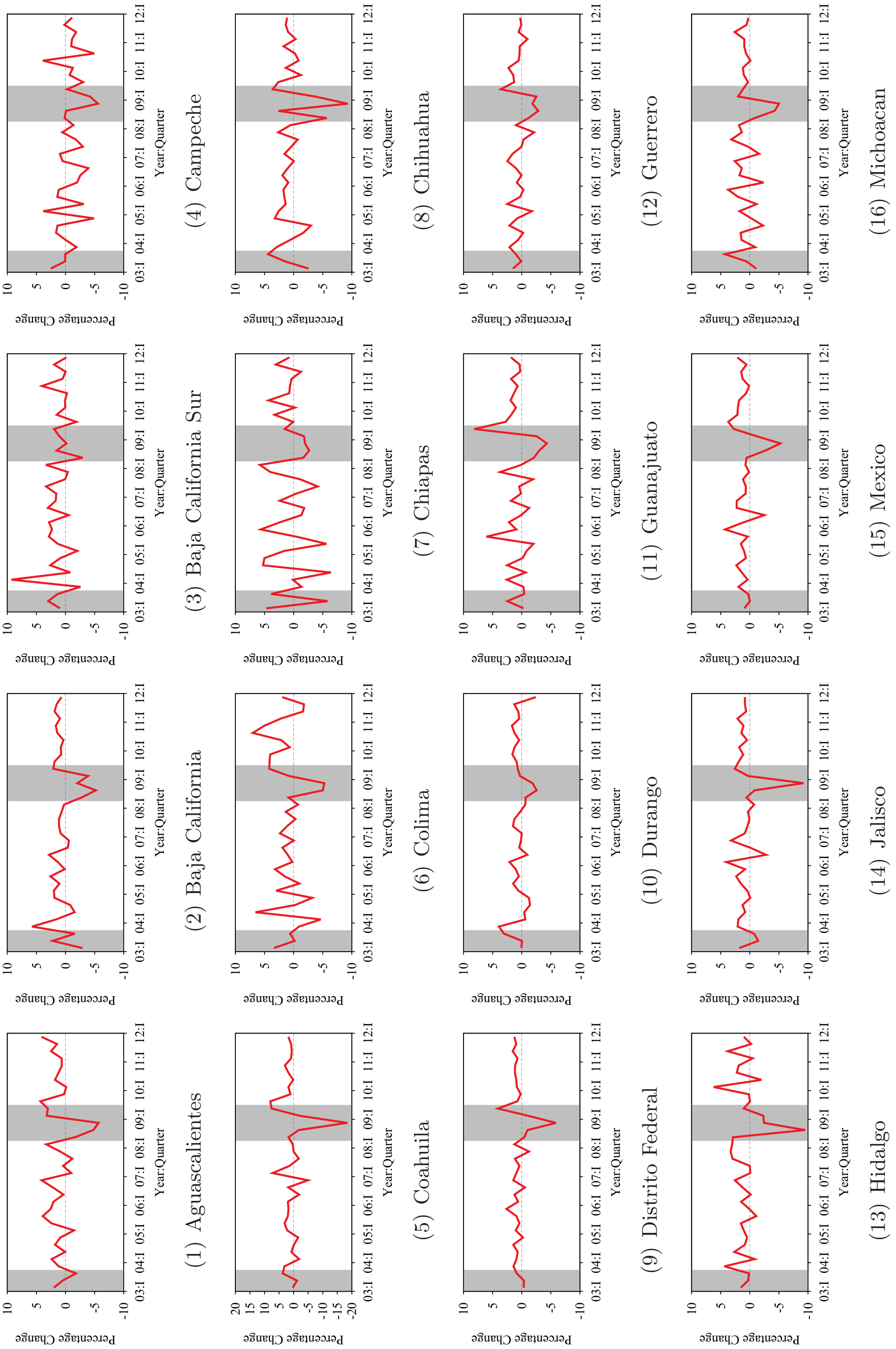
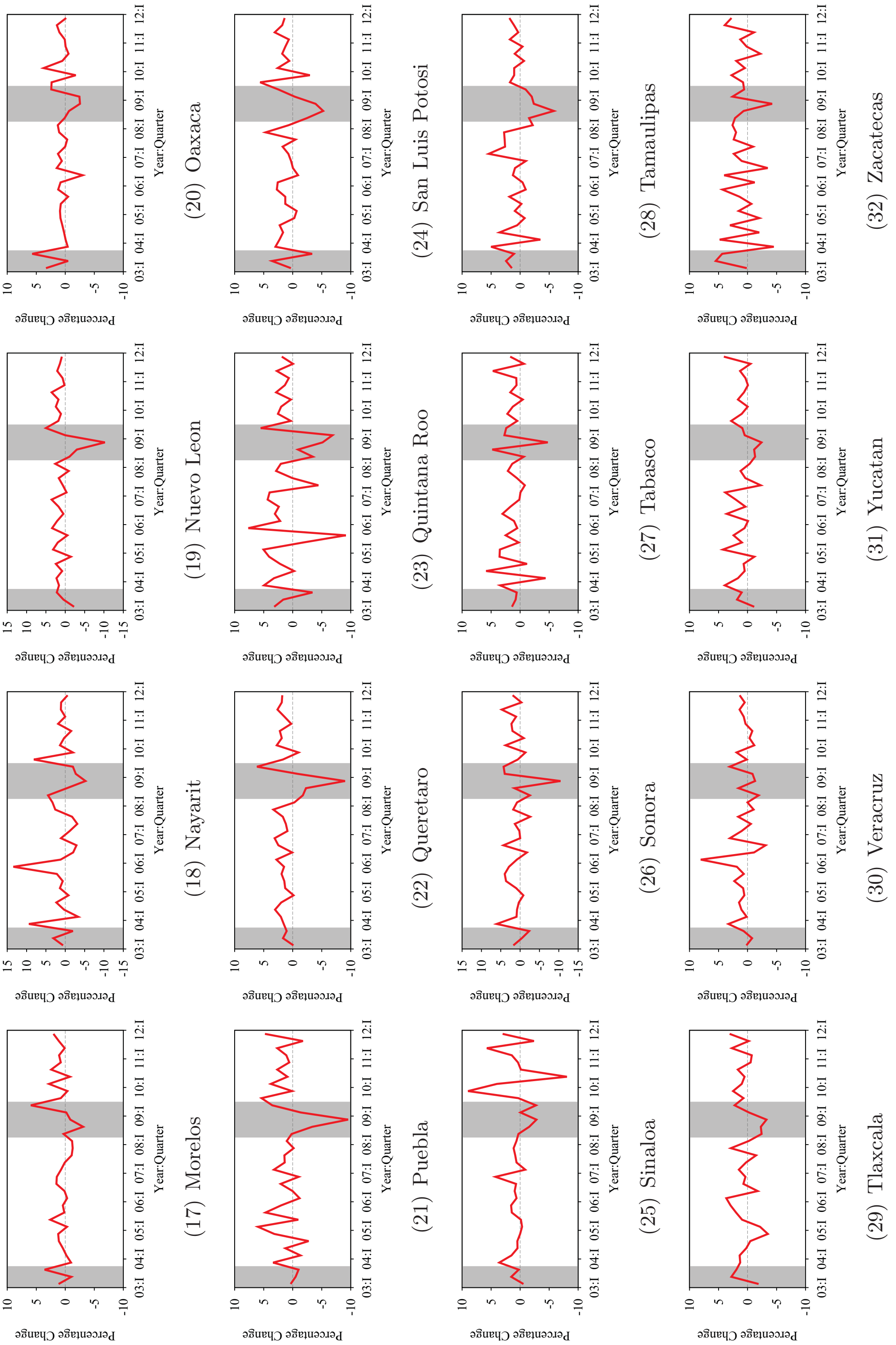


Figure 2.7: Percentage Change of ITAEE



(17) Morelos (18) Nayarit (19) Nuevo Leon (20) Oaxaca

(21) Puebla (22) Queretaro (23) Quintana Roo (24) San Luis Potosi

(25) Sinaloa (26) Sonora (27) Tabasco (28) Tamaulipas

(29) Tlaxcala (30) Veracruz (31) Yucatan (32) Zacatecas

Figure 2.7: Percentage Change of ITAEE (Continued)

Table 2.7: Estimated Parameters for MS-SAR

		ρ Mean	ρ Median	ρ 95% CI
	Spatial Dependence	0.28	0.28	[0.23, 0.32]

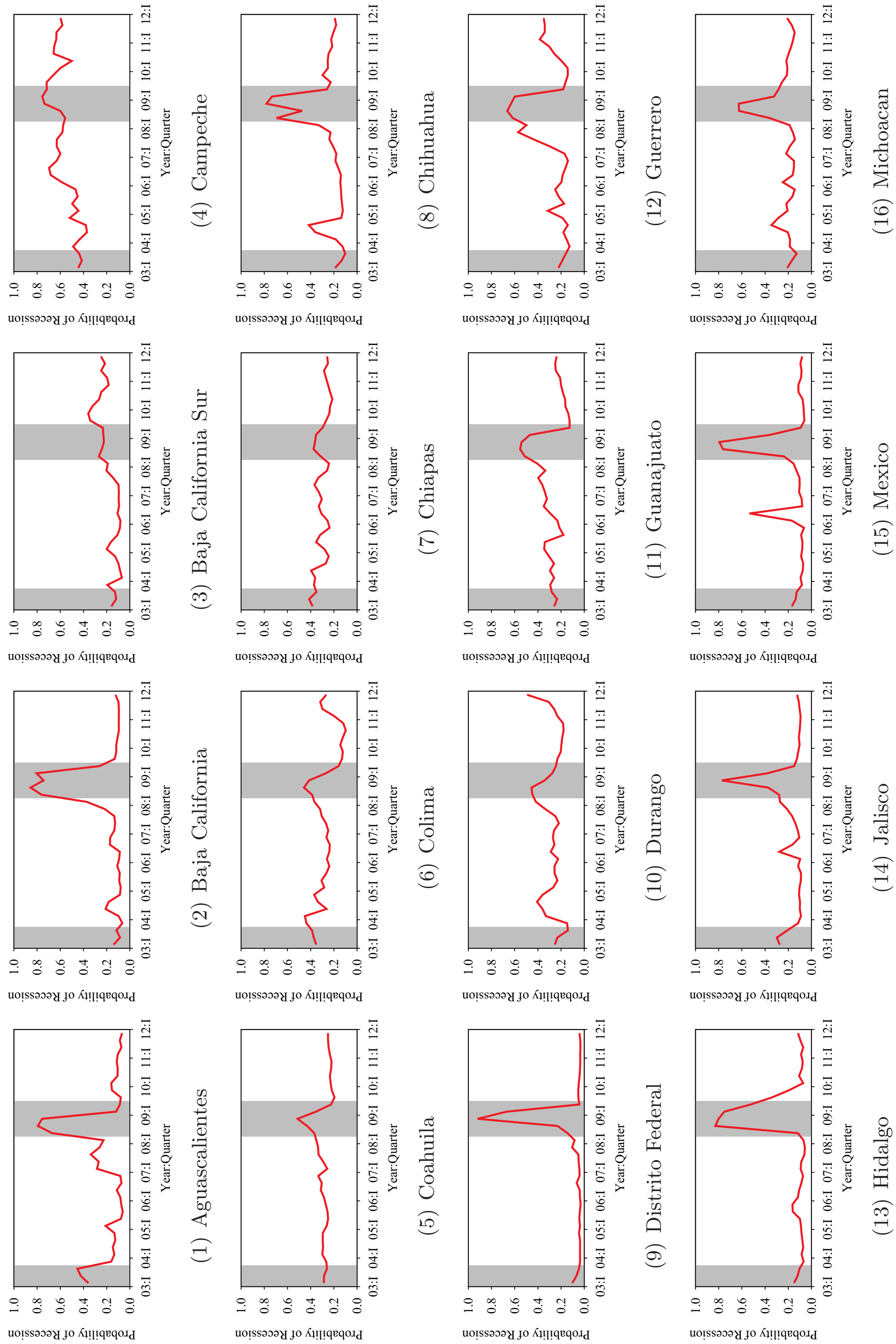
Code	State	μ_0 Mean	μ_0 Median	μ_0 95% CI	μ_1 Mean	μ_1 Median	μ_1 95% CI
1	Aguascalientes	−1.25	−1.32	[−3.12, 0.60]	1.15	1.15	[0.41, 1.91]
2	Baja California	−1.34	−1.34	[−3.08, 0.28]	0.64	0.65	[−0.06, 1.35]
3	Baja California Sur	−0.27	−0.14	[−2.19, 1.01]	1.02	1.00	[0.28, 1.85]
4	Campeche	−1.35	−1.30	[−2.47, −0.53]	−0.16	−0.31	[−1.26, 1.69]
5	Coahuila	−0.70	−0.61	[−2.53, 0.69]	0.77	0.74	[−0.31, 2.05]
6	Colima	−0.49	−0.41	[−2.20, 0.80]	1.12	1.06	[0.14, 2.42]
7	Chiapas	−0.64	−0.56	[−2.26, 0.55]	0.57	0.52	[−0.46, 1.88]
8	Chihuahua	−1.03	−0.84	[−3.34, 0.43]	0.64	0.62	[−0.22, 1.64]
9	Distrito Federal	−1.70	−1.74	[−3.77, 0.27]	0.53	0.52	[0.11, 0.97]
10	Durango	−0.61	−0.49	[−2.15, 0.30]	0.44	0.37	[−0.15, 1.57]
11	Guanajuato	−0.48	−0.37	[−2.07, 0.64]	0.76	0.69	[−0.04, 2.05]
12	Guerrero	−0.63	−0.57	[−1.94, 0.32]	0.55	0.52	[−0.02, 1.32]
13	Hidalgo	−0.87	−0.76	[−3.03, 0.68]	0.90	0.89	[0.08, 1.78]
14	Jalisco	−1.40	−1.24	[−3.95, 0.43]	0.62	0.62	[−0.01, 1.30]
15	México	−1.65	−1.78	[−3.01, 0.34]	1.03	1.04	[0.55, 1.47]
16	Michoacán	−0.68	−0.53	[−2.51, 0.48]	0.52	0.50	[−0.13, 1.35]
17	Morelos	−0.41	−0.28	[−1.97, 0.52]	0.59	0.55	[0.00, 1.43]
18	Nayarit	−0.50	−0.42	[−2.18, 0.79]	0.85	0.81	[−0.19, 2.15]
19	Nuevo León	−1.11	−1.14	[−2.92, 0.61]	1.01	1.01	[0.31, 1.70]
20	Oaxaca	−0.33	−0.17	[−2.02, 0.54]	0.64	0.53	[−0.04, 2.10]
21	Puebla	−0.70	−0.60	[−2.63, 0.77]	1.00	0.98	[0.08, 2.01]
22	Querétaro	−1.79	−1.82	[−3.30, 0.05]	1.30	1.31	[0.73, 1.85]
23	Quintana Roo	−0.79	−0.66	[−3.06, 0.95]	1.32	1.30	[0.17, 2.57]
24	San Luis Potosí	−0.93	−0.88	[−2.79, 0.61]	0.90	0.90	[0.16, 1.71]
25	Sinaloa	−0.50	−0.39	[−2.26, 0.69]	0.75	0.72	[−0.12, 1.88]
26	Sonora	−0.49	−0.38	[−2.37, 0.84]	0.94	0.91	[0.08, 1.92]
27	Tabasco	−0.25	−0.11	[−2.19, 1.09]	1.08	1.07	[0.33, 1.94]
28	Tamaulipas	−0.69	−0.59	[−2.30, 0.46]	0.81	0.75	[−0.05, 2.07]
29	Tlaxcala	−0.90	−0.78	[−2.72, 0.35]	0.55	0.52	[−0.17, 1.47]
30	Veracruz	−0.36	−0.22	[−2.08, 0.67]	0.71	0.66	[0.02, 1.73]
31	Yucatán	−0.26	−0.17	[−2.10, 1.05]	1.23	1.20	[0.61, 1.99]
32	Zacatecas	−0.34	−0.22	[−2.15, 0.90]	0.90	0.87	[0.08, 1.91]

Notes: 95% CI indicates 95% credible interval.

Table 2.7: Estimated Parameters for MS-SAR (Continued)

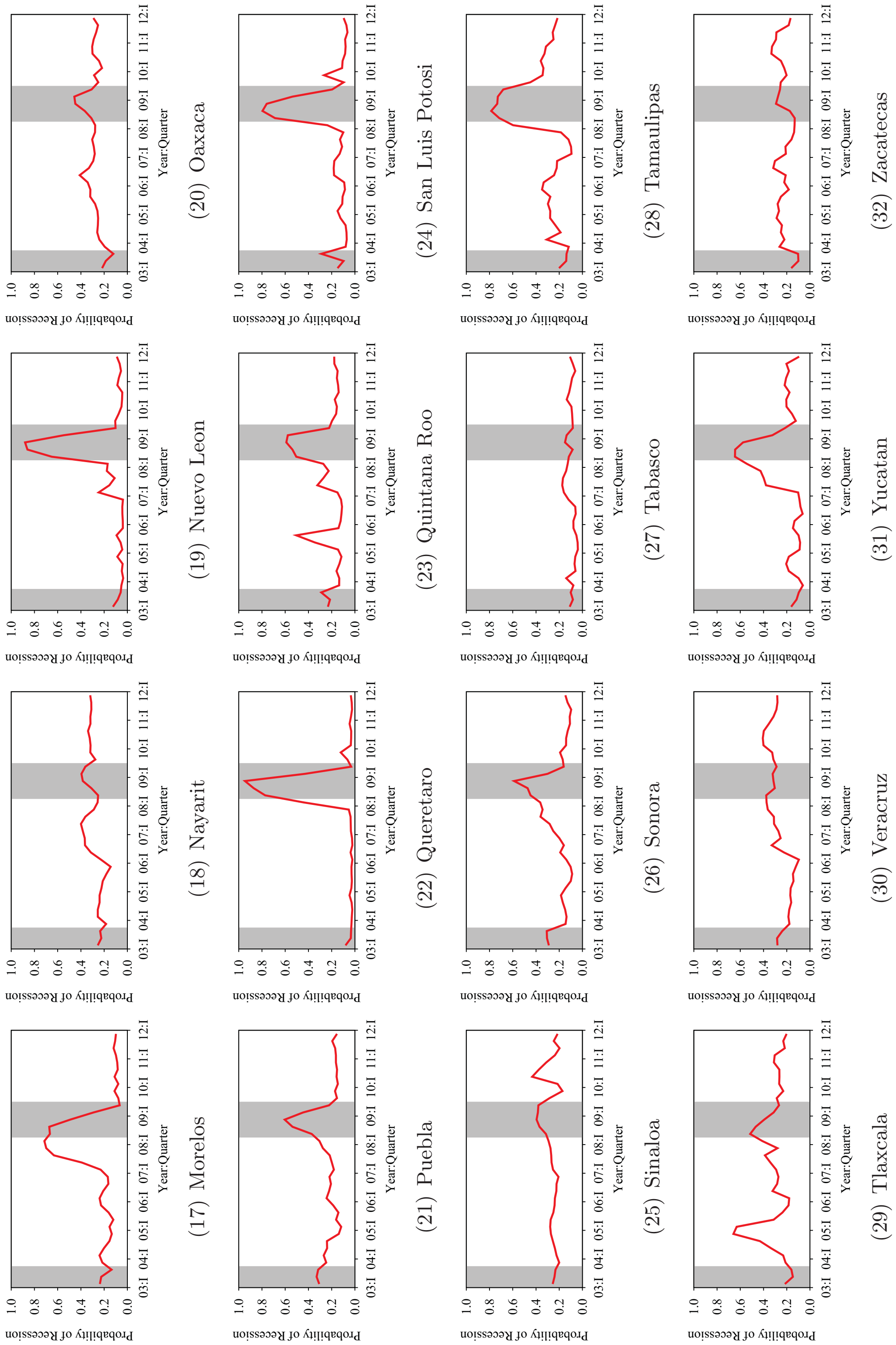
		σ^2			p_{11}			p_{00}		
Code	State	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI
1	Aguascalientes	3.15	3.00	[1.70, 5.44]	0.91	0.93	[0.73, 0.99]	0.75	0.76	[0.48, 0.95]
2	Baja California	2.94	2.83	[1.56, 5.01]	0.92	0.94	[0.76, 0.99]	0.76	0.77	[0.48, 0.95]
3	Baja California Sur	4.89	4.73	[3.14, 7.55]	0.92	0.95	[0.71, 1.00]	0.78	0.80	[0.49, 0.97]
4	Campeche	4.50	4.36	[2.74, 7.16]	0.86	0.88	[0.59, 0.99]	0.84	0.87	[0.55, 0.99]
5	Coahuila	12.62	12.19	[7.99, 19.77]	0.91	0.93	[0.67, 1.00]	0.79	0.81	[0.50, 0.97]
6	Colima	6.34	6.14	[3.74, 10.16]	0.90	0.92	[0.67, 1.00]	0.78	0.80	[0.50, 0.96]
7	Chiapas	10.21	9.89	[6.45, 15.77]	0.89	0.92	[0.64, 1.00]	0.79	0.81	[0.50, 0.97]
8	Chihuahua	5.06	4.95	[2.47, 8.28]	0.90	0.92	[0.69, 1.00]	0.77	0.79	[0.47, 0.97]
9	Distrito Federal	1.26	1.19	[0.72, 2.16]	0.94	0.95	[0.81, 0.99]	0.75	0.76	[0.46, 0.95]
10	Durango	1.67	1.62	[0.99, 2.66]	0.90	0.93	[0.66, 1.00]	0.79	0.81	[0.51, 0.97]
11	Guanajuato	3.90	3.76	[2.38, 6.22]	0.90	0.93	[0.65, 1.00]	0.79	0.81	[0.51, 0.97]
12	Guerrero	1.44	1.39	[0.81, 2.36]	0.90	0.92	[0.68, 1.00]	0.79	0.81	[0.52, 0.97]
13	Hidalgo	5.01	4.83	[2.95, 8.16]	0.91	0.94	[0.70, 1.00]	0.77	0.79	[0.49, 0.96]
14	Jalisco	2.78	2.68	[1.45, 4.68]	0.92	0.94	[0.75, 1.00]	0.75	0.77	[0.45, 0.96]
15	México	1.45	1.36	[0.79, 2.71]	0.93	0.94	[0.81, 0.99]	0.73	0.74	[0.46, 0.93]
16	Michoacán	3.02	2.93	[1.79, 4.77]	0.91	0.94	[0.69, 1.00]	0.77	0.79	[0.48, 0.97]
17	Morelos	2.14	2.07	[1.34, 3.34]	0.91	0.94	[0.68, 1.00]	0.79	0.81	[0.51, 0.97]
18	Nayarit	11.23	10.88	[7.14, 17.51]	0.90	0.93	[0.67, 1.00]	0.79	0.81	[0.51, 0.97]
19	Nuevo León	3.06	2.94	[1.69, 5.19]	0.93	0.94	[0.77, 1.00]	0.75	0.77	[0.47, 0.95]
20	Oaxaca	2.38	2.31	[1.49, 3.69]	0.90	0.93	[0.64, 1.00]	0.81	0.83	[0.53, 0.98]
21	Puebla	6.76	6.55	[4.12, 10.75]	0.92	0.94	[0.71, 1.00]	0.78	0.79	[0.50, 0.96]
22	Querétaro	2.45	2.33	[1.48, 4.09]	0.94	0.95	[0.83, 0.99]	0.77	0.78	[0.50, 0.94]
23	Quintana Roo	9.72	9.54	[4.69, 15.94]	0.91	0.93	[0.71, 1.00]	0.76	0.78	[0.48, 0.96]
24	San Luis Potosí	3.34	3.25	[1.72, 5.55]	0.91	0.93	[0.70, 1.00]	0.75	0.77	[0.48, 0.96]
25	Sinaloa	6.48	6.26	[4.06, 10.18]	0.91	0.93	[0.67, 1.00]	0.78	0.80	[0.50, 0.97]
26	Sonora	6.48	6.27	[4.05, 10.10]	0.91	0.94	[0.70, 1.00]	0.78	0.80	[0.49, 0.97]
27	Tabasco	5.56	5.38	[3.53, 8.66]	0.93	0.96	[0.71, 1.00]	0.78	0.80	[0.48, 0.97]
28	Tamaulipas	3.69	3.57	[2.16, 5.95]	0.89	0.91	[0.65, 1.00]	0.80	0.82	[0.53, 0.97]
29	Tlaxcala	2.94	2.85	[1.65, 4.80]	0.90	0.93	[0.69, 1.00]	0.76	0.78	[0.48, 0.96]
30	Veracruz	3.30	3.20	[2.08, 5.13]	0.91	0.94	[0.67, 1.00]	0.79	0.81	[0.51, 0.97]
31	Yucatán	2.44	2.36	[1.46, 3.89]	0.93	0.95	[0.73, 1.00]	0.77	0.79	[0.49, 0.96]
32	Zacatecas	5.85	5.66	[3.71, 9.11]	0.91	0.94	[0.67, 1.00]	0.78	0.80	[0.48, 0.97]

Notes: 95% CI indicates 95% credible interval.



(1) Aguascalientes (2) Baja California (3) Baja California Sur (4) Campeche (5) Coahuila (6) Colima (7) Chiapas (8) Chihuahua (9) Distrito Federal (10) Durango (11) Guanajuato (12) Guerrero (13) Hidalgo (14) Jalisco (15) Mexico (16) Michoacan

Figure 2.8: Recession Probabilities for MS-SAR



(17) Morelos (18) Nayarit (19) Nuevo Leon (20) Oaxaca
(21) Puebla (22) Queretaro (23) Quintana Roo (24) San Luis Potosi
(25) Sinaloa (26) Sonora (27) Tabasco (28) Tamaulipas
(29) Tlaxcala (30) Veracruz (31) Yucatan (32) Zacatecas

Figure 2.8: Recession Probabilities for MS-SAR (Continued)

Table 2.8: Estimated Parameters for MS

Code	State	μ_0			μ_1		
		Mean	Median	95% CI	Mean	Median	95% CI
1	Aguascalientes	−1.73	−1.85	[−3.54, 0.55]	1.39	1.39	[0.69, 2.07]
2	Baja California	−2.07	−2.15	[−3.43, −0.18]	1.03	1.04	[0.40, 1.60]
3	Baja California Sur	−0.21	−0.08	[−2.18, 1.12]	1.19	1.18	[0.47, 2.04]
4	Campeche	−1.18	−1.13	[−2.36, −0.30]	0.00	−0.13	[−1.06, 1.72]
5	Coahuila	−0.73	−0.63	[−2.75, 0.79]	0.94	0.93	[−0.23, 2.19]
6	Colima	−0.55	−0.47	[−2.39, 0.89]	1.28	1.23	[0.24, 2.56]
7	Chiapas	−0.54	−0.45	[−2.27, 0.73]	0.77	0.72	[−0.23, 2.05]
8	Chihuahua	−1.20	−1.06	[−3.54, 0.52]	0.85	0.85	[−0.06, 1.81]
9	Distrito Federal	−1.75	−1.77	[−3.57, 0.20]	0.80	0.80	[0.36, 1.27]
10	Durango	−0.71	−0.69	[−2.09, 0.41]	0.71	0.69	[0.11, 1.41]
11	Guanajuato	−0.68	−0.60	[−2.53, 0.73]	1.02	0.99	[0.17, 2.00]
12	Guerrero	−1.05	−1.08	[−2.35, 0.30]	0.79	0.79	[0.23, 1.36]
13	Hidalgo	−1.29	−1.28	[−3.43, 0.67]	1.10	1.11	[0.28, 1.92]
14	Jalisco	−1.54	−1.49	[−4.02, 0.52]	0.91	0.91	[0.23, 1.60]
15	México	−1.97	−2.04	[−3.31, −0.13]	1.24	1.25	[0.73, 1.70]
16	Michoacán	−1.28	−1.30	[−3.28, 0.54]	0.84	0.84	[0.15, 1.55]
17	Morelos	−0.44	−0.36	[−2.01, 0.66]	0.83	0.80	[0.19, 1.64]
18	Nayarit	−0.46	−0.39	[−2.18, 0.87]	0.97	0.93	[−0.09, 2.27]
19	Nuevo León	−1.89	−1.96	[−3.80, 0.34]	1.29	1.29	[0.51, 2.00]
20	Oaxaca	−0.36	−0.19	[−2.14, 0.67]	0.74	0.69	[0.07, 1.73]
21	Puebla	−0.89	−0.82	[−2.92, 0.82]	1.15	1.14	[0.21, 2.16]
22	Querétaro	−2.05	−2.10	[−3.55, −0.30]	1.56	1.57	[0.92, 2.17]
23	Quintana Roo	−0.94	−0.80	[−3.40, 0.91]	1.40	1.37	[0.19, 2.69]
24	San Luis Potosí	−1.35	−1.42	[−3.17, 0.61]	1.20	1.21	[0.38, 1.99]
25	Sinaloa	−0.48	−0.37	[−2.33, 0.82]	0.91	0.88	[0.01, 1.99]
26	Sonora	−0.54	−0.44	[−2.48, 0.93]	1.10	1.08	[0.18, 2.11]
27	Tabasco	−0.32	−0.19	[−2.47, 1.15]	1.21	1.20	[0.50, 1.97]
28	Tamaulipas	−1.25	−1.29	[−2.85, 0.40]	1.03	1.03	[0.22, 1.85]
29	Tlaxcala	−1.11	−1.18	[−2.54, 0.42]	0.96	0.97	[0.15, 1.76]
30	Veracruz	−0.27	−0.14	[−1.98, 0.81]	0.89	0.83	[0.19, 1.96]
31	Yucatán	−0.39	−0.34	[−2.02, 0.86]	1.10	1.09	[0.45, 1.86]
32	Zacatecas	−0.26	−0.14	[−2.15, 1.07]	1.10	1.08	[0.30, 2.02]

Notes: 95% CI indicates 95% credible interval.

Table 2.8: Estimated Parameters for MS (Continued)

		σ^2			p_{11}			p_{00}		
Code	State	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI
1	Aguascalientes	3.34	3.14	[1.91, 5.87]	0.93	0.94	[0.78, 0.99]	0.75	0.76	[0.48, 0.94]
2	Baja California	2.47	2.32	[1.37, 4.42]	0.94	0.95	[0.81, 0.99]	0.75	0.76	[0.50, 0.94]
3	Baja California Sur	4.67	4.53	[2.96, 7.33]	0.93	0.96	[0.72, 1.00]	0.78	0.80	[0.49, 0.97]
4	Campeche	4.90	4.73	[3.02, 7.68]	0.86	0.89	[0.60, 0.99]	0.84	0.87	[0.56, 0.98]
5	Coahuila	15.54	15.03	[9.62, 24.27]	0.91	0.93	[0.68, 1.00]	0.78	0.80	[0.49, 0.96]
6	Colima	7.04	6.80	[4.10, 11.35]	0.90	0.92	[0.69, 1.00]	0.78	0.79	[0.51, 0.96]
7	Chiapas	9.50	9.22	[5.87, 14.80]	0.90	0.93	[0.66, 1.00]	0.78	0.80	[0.50, 0.97]
8	Chihuahua	5.76	5.61	[2.91, 9.61]	0.91	0.93	[0.70, 1.00]	0.76	0.78	[0.47, 0.96]
9	Distrito Federal	1.60	1.52	[0.97, 2.68]	0.94	0.95	[0.83, 0.99]	0.76	0.77	[0.49, 0.94]
10	Durango	1.63	1.57	[0.88, 2.72]	0.91	0.93	[0.71, 1.00]	0.78	0.79	[0.51, 0.96]
11	Guanajuato	4.94	4.78	[2.93, 7.93]	0.91	0.93	[0.69, 1.00]	0.78	0.79	[0.50, 0.96]
12	Guerrero	1.64	1.56	[0.95, 2.80]	0.92	0.93	[0.74, 0.99]	0.78	0.80	[0.53, 0.95]
13	Hidalgo	5.08	4.88	[2.89, 8.45]	0.93	0.94	[0.75, 1.00]	0.76	0.78	[0.49, 0.95]
14	Jalisco	3.47	3.32	[1.90, 5.91]	0.93	0.94	[0.76, 0.99]	0.75	0.77	[0.46, 0.95]
15	México	1.74	1.64	[0.98, 3.14]	0.94	0.94	[0.83, 0.99]	0.74	0.75	[0.47, 0.94]
16	Michoacán	3.13	3.00	[1.70, 5.27]	0.92	0.94	[0.74, 1.00]	0.75	0.76	[0.47, 0.95]
17	Morelos	2.41	2.32	[1.47, 3.85]	0.91	0.94	[0.69, 1.00]	0.79	0.81	[0.52, 0.97]
18	Nayarit	12.05	11.66	[7.61, 18.91]	0.91	0.93	[0.68, 1.00]	0.79	0.81	[0.51, 0.97]
19	Nuevo León	4.35	4.12	[2.51, 7.54]	0.94	0.95	[0.82, 0.99]	0.75	0.76	[0.48, 0.94]
20	Oaxaca	2.75	2.65	[1.73, 4.33]	0.90	0.93	[0.65, 1.00]	0.79	0.81	[0.50, 0.97]
21	Puebla	6.97	6.71	[4.14, 11.24]	0.92	0.94	[0.72, 1.00]	0.77	0.78	[0.49, 0.95]
22	Querétaro	3.16	3.02	[1.94, 5.31]	0.95	0.95	[0.85, 0.99]	0.77	0.79	[0.52, 0.95]
23	Quintana Roo	10.10	9.97	[4.24, 16.97]	0.90	0.92	[0.69, 1.00]	0.76	0.78	[0.48, 0.96]
24	San Luis Potosí	3.91	3.74	[2.02, 6.77]	0.92	0.94	[0.75, 0.99]	0.74	0.76	[0.47, 0.95]
25	Sinaloa	6.64	6.41	[4.13, 10.50]	0.91	0.93	[0.67, 1.00]	0.78	0.80	[0.50, 0.97]
26	Sonora	7.51	7.26	[4.57, 11.76]	0.92	0.94	[0.71, 1.00]	0.77	0.79	[0.49, 0.96]
27	Tabasco	4.46	4.32	[2.72, 7.00]	0.93	0.96	[0.73, 1.00]	0.77	0.79	[0.47, 0.97]
28	Tamaulipas	3.66	3.53	[2.05, 6.12]	0.91	0.93	[0.73, 0.99]	0.78	0.79	[0.53, 0.95]
29	Tlaxcala	2.62	2.49	[1.29, 4.67]	0.89	0.91	[0.70, 0.99]	0.74	0.75	[0.48, 0.94]
30	Veracruz	3.44	3.33	[2.17, 5.38]	0.91	0.94	[0.68, 1.00]	0.80	0.81	[0.51, 0.97]
31	Yucatán	2.63	2.54	[1.56, 4.18]	0.92	0.94	[0.73, 1.00]	0.78	0.80	[0.51, 0.96]
32	Zacatecas	5.89	5.70	[3.69, 9.13]	0.92	0.95	[0.68, 1.00]	0.78	0.80	[0.47, 0.97]

Notes: 95% CI indicates 95% credible interval.

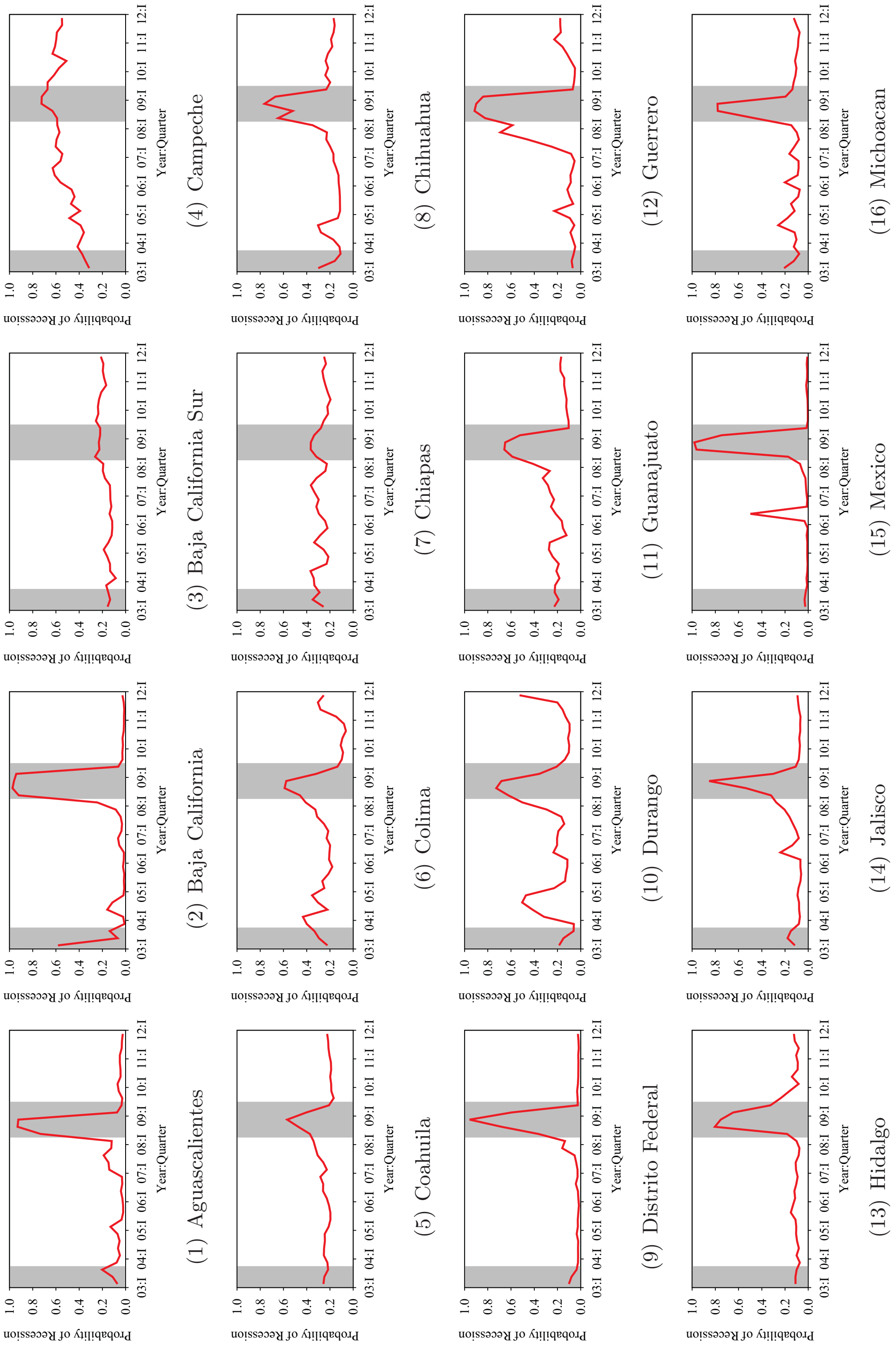
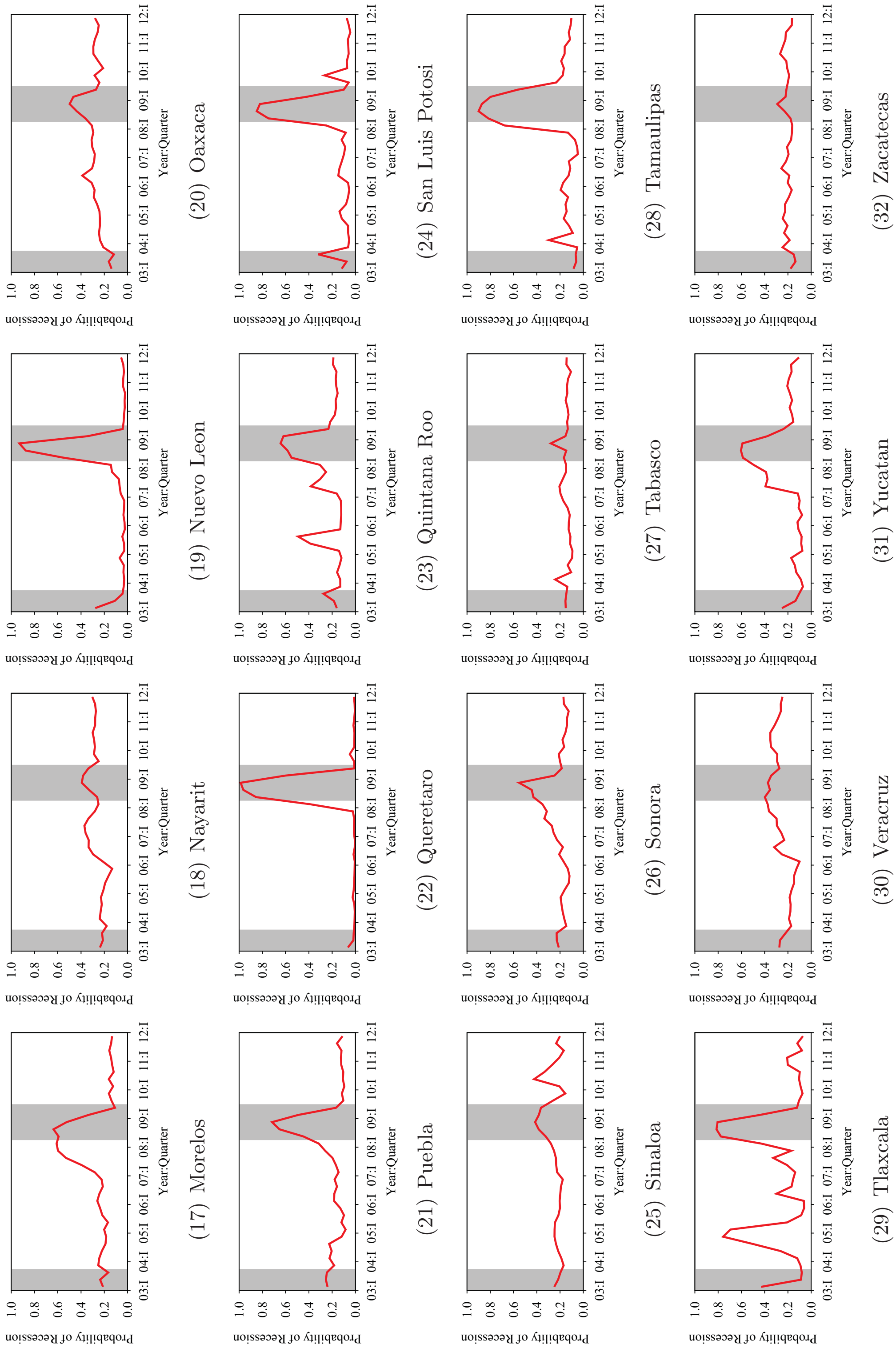


Figure 2.9: Recession Probabilities for MS



(17) Morelos

(18) Nayarit

(19) Nuevo Leon

(20) Oaxaca

(21) Puebla

(22) Queretaro

(23) Quintana Roo

(24) San Luis Potosi

(25) Sinaloa

(26) Sonora

(27) Tabasco

(28) Tamaulipas

(29) Tlaxcala

(30) Veracruz

(31) Yucatan

(32) Zacatecas

Figure 2.9: Recession Probabilities for MS (Continued)

Table 2.9: Estimated Parameters for MS-AR(1)

		μ_0			μ_1			ϕ		
Code	State	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI
1	Aguascalientes	−1.29	−1.33	[−3.27, 0.62]	1.16	1.17	[0.29, 1.99]	0.20	0.20	[−0.11, 0.53]
2	Baja California	−2.14	−2.25	[−3.67, 0.11]	1.11	1.13	[0.40, 1.72]	−0.04	−0.05	[−0.32, 0.32]
3	Baja California Sur	−0.02	0.10	[−2.07, 1.44]	1.69	1.66	[0.84, 2.75]	−0.40	−0.40	[−0.69, −0.09]
4	Campeche	−1.43	−1.38	[−2.63, −0.50]	−0.16	−0.28	[−1.28, 1.59]	−0.22	−0.22	[−0.53, 0.09]
5	Coahuila	−0.69	−0.61	[−2.57, 0.76]	0.89	0.87	[−0.30, 2.20]	0.13	0.14	[−0.18, 0.45]
6	Colima	−0.56	−0.48	[−2.34, 0.78]	1.04	0.99	[0.02, 2.37]	0.16	0.16	[−0.15, 0.48]
7	Chiapas	−0.64	−0.57	[−2.33, 0.63]	0.75	0.69	[−0.27, 2.10]	−0.10	−0.10	[−0.42, 0.21]
8	Chihuahua	−1.53	−1.45	[−4.04, 0.48]	1.02	1.04	[0.04, 1.93]	−0.05	−0.06	[−0.37, 0.29]
9	Distrito Federal	−1.89	−1.95	[−3.73, 0.25]	0.88	0.89	[0.28, 1.43]	−0.05	−0.05	[−0.45, 0.37]
10	Durango	−0.65	−0.55	[−2.20, 0.36]	0.51	0.46	[−0.10, 1.38]	0.32	0.32	[−0.05, 0.67]
11	Guanajuato	−0.82	−0.73	[−2.78, 0.72]	1.07	1.06	[0.12, 2.11]	−0.02	−0.02	[−0.36, 0.34]
12	Guerrero	−1.06	−1.09	[−2.44, 0.31]	0.76	0.77	[0.08, 1.42]	−0.02	−0.02	[−0.38, 0.39]
13	Hidalgo	−1.89	−1.99	[−3.99, 0.49]	1.32	1.34	[0.37, 2.14]	−0.18	−0.19	[−0.49, 0.18]
14	Jalisco	−1.42	−1.36	[−3.84, 0.51]	0.88	0.88	[0.11, 1.68]	0.03	0.03	[−0.32, 0.35]
15	México	−1.61	−1.75	[−3.31, 0.41]	1.02	1.02	[0.31, 1.67]	0.20	0.21	[−0.16, 0.56]
16	Michoacán	−1.39	−1.44	[−3.43, 0.53]	0.93	0.93	[0.19, 1.66]	−0.04	−0.04	[−0.36, 0.29]
17	Morelos	−0.77	−0.79	[−2.08, 0.53]	1.03	1.05	[0.27, 1.77]	−0.21	−0.22	[−0.58, 0.19]
18	Nayarit	−0.45	−0.38	[−2.20, 0.91]	1.03	0.99	[−0.08, 2.41]	−0.05	−0.05	[−0.37, 0.27]
19	Nuevo León	−1.88	−1.96	[−3.93, 0.43]	1.27	1.28	[0.40, 2.05]	0.10	0.10	[−0.25, 0.41]
20	Oaxaca	−0.53	−0.36	[−2.40, 0.67]	0.79	0.76	[0.11, 1.68]	−0.21	−0.21	[−0.53, 0.11]
21	Puebla	−0.85	−0.78	[−2.92, 0.76]	1.13	1.12	[0.12, 2.18]	0.06	0.05	[−0.29, 0.40]
22	Querétaro	−1.89	−1.98	[−3.65, 0.32]	1.44	1.44	[0.53, 2.29]	0.13	0.14	[−0.27, 0.49]
23	Quintana Roo	−0.98	−0.83	[−3.40, 0.88]	1.40	1.36	[0.10, 2.84]	−0.06	−0.06	[−0.37, 0.27]
24	San Luis Potosí	−1.37	−1.43	[−3.26, 0.62]	1.24	1.26	[0.27, 2.09]	−0.02	−0.02	[−0.37, 0.35]
25	Sinaloa	−0.47	−0.36	[−2.33, 0.84]	0.97	0.93	[0.05, 2.16]	−0.03	−0.03	[−0.36, 0.30]
26	Sonora	−0.65	−0.58	[−2.67, 0.98]	1.33	1.31	[0.29, 2.42]	−0.24	−0.24	[−0.56, 0.08]
27	Tabasco	−0.25	−0.24	[−2.22, 1.47]	1.85	1.84	[1.13, 2.65]	−0.55	−0.55	[−0.82, −0.28]
28	Tamaulipas	−1.25	−1.27	[−2.97, 0.42]	1.01	1.02	[0.09, 1.90]	−0.01	−0.02	[−0.37, 0.36]
29	Tlaxcala	−0.99	−1.00	[−2.60, 0.46]	0.99	1.00	[0.11, 1.83]	0.11	0.11	[−0.27, 0.50]
30	Veracruz	−0.27	−0.14	[−2.04, 0.86]	0.98	0.92	[0.20, 2.13]	−0.10	−0.10	[−0.44, 0.23]
31	Yucatán	−0.41	−0.40	[−1.88, 0.90]	1.38	1.38	[0.55, 2.27]	−0.14	−0.15	[−0.50, 0.22]
32	Zacatecas	−0.15	−0.05	[−2.12, 1.32]	1.52	1.48	[0.65, 2.61]	−0.40	−0.40	[−0.70, −0.10]

Notes: 95% CI indicates 95% credible interval.

Table 2.9: Estimated Parameters for MS-AR(1) (Continued)

		σ^2			p_{11}			p_{00}		
Code	State	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI
1	Aguascalientes	3.69	3.50	[1.95, 6.52]	0.92	0.94	[0.73, 1.00]	0.75	0.77	[0.48, 0.95]
2	Baja California	2.37	2.20	[1.34, 4.40]	0.94	0.95	[0.81, 0.99]	0.77	0.79	[0.52, 0.95]
3	Baja California Sur	3.97	3.83	[2.38, 6.35]	0.93	0.95	[0.73, 1.00]	0.78	0.80	[0.49, 0.97]
4	Campeche	4.52	4.37	[2.72, 7.20]	0.87	0.89	[0.62, 0.99]	0.84	0.87	[0.57, 0.98]
5	Coahuila	16.15	15.56	[10.04, 25.45]	0.90	0.93	[0.68, 1.00]	0.78	0.80	[0.50, 0.97]
6	Colima	7.21	6.99	[4.28, 11.40]	0.91	0.93	[0.69, 1.00]	0.78	0.80	[0.50, 0.97]
7	Chiapas	9.28	9.01	[5.53, 14.67]	0.90	0.93	[0.67, 1.00]	0.78	0.79	[0.50, 0.97]
8	Chihuahua	5.33	5.16	[2.32, 9.30]	0.91	0.93	[0.73, 0.99]	0.76	0.77	[0.48, 0.96]
9	Distrito Federal	1.60	1.50	[0.93, 2.81]	0.94	0.95	[0.82, 0.99]	0.76	0.78	[0.49, 0.95]
10	Durango	1.68	1.63	[0.93, 2.74]	0.91	0.93	[0.69, 1.00]	0.78	0.79	[0.50, 0.97]
11	Guanajuato	5.03	4.84	[2.84, 8.40]	0.91	0.93	[0.71, 1.00]	0.78	0.79	[0.51, 0.96]
12	Guerrero	1.74	1.65	[0.98, 2.98]	0.92	0.93	[0.73, 0.99]	0.78	0.80	[0.53, 0.96]
13	Hidalgo	4.62	4.33	[2.52, 8.25]	0.94	0.95	[0.80, 0.99]	0.77	0.78	[0.50, 0.95]
14	Jalisco	3.65	3.50	[1.94, 6.31]	0.93	0.94	[0.75, 1.00]	0.75	0.77	[0.47, 0.95]
15	México	1.96	1.84	[1.02, 3.57]	0.93	0.94	[0.78, 0.99]	0.74	0.75	[0.46, 0.94]
16	Michoacán	3.13	2.99	[1.66, 5.37]	0.92	0.94	[0.73, 1.00]	0.75	0.76	[0.47, 0.95]
17	Morelos	2.25	2.16	[1.31, 3.78]	0.92	0.94	[0.75, 1.00]	0.79	0.81	[0.54, 0.96]
18	Nayarit	12.62	12.17	[7.88, 19.93]	0.91	0.93	[0.68, 1.00]	0.79	0.81	[0.51, 0.97]
19	Nuevo León	4.23	3.99	[2.39, 7.51]	0.94	0.95	[0.81, 0.99]	0.76	0.77	[0.49, 0.95]
20	Oaxaca	2.48	2.40	[1.44, 4.03]	0.91	0.93	[0.67, 1.00]	0.78	0.79	[0.49, 0.97]
21	Puebla	7.34	7.06	[4.33, 11.92]	0.92	0.94	[0.73, 1.00]	0.77	0.79	[0.49, 0.96]
22	Querétaro	3.34	3.12	[1.92, 5.93]	0.94	0.95	[0.83, 0.99]	0.77	0.79	[0.51, 0.95]
23	Quintana Roo	10.46	10.32	[4.18, 17.97]	0.91	0.92	[0.71, 1.00]	0.76	0.78	[0.48, 0.96]
24	San Luis Potosí	4.14	3.94	[2.07, 7.30]	0.92	0.94	[0.75, 0.99]	0.75	0.76	[0.48, 0.95]
25	Sinaloa	6.92	6.69	[4.18, 10.94]	0.90	0.93	[0.68, 1.00]	0.78	0.80	[0.49, 0.96]
26	Sonora	7.17	6.90	[4.22, 11.68]	0.92	0.94	[0.72, 1.00]	0.77	0.79	[0.50, 0.96]
27	Tabasco	2.99	2.90	[1.61, 4.88]	0.94	0.96	[0.77, 1.00]	0.75	0.77	[0.45, 0.96]
28	Tamaulipas	3.92	3.73	[2.18, 6.72]	0.91	0.93	[0.72, 0.99]	0.78	0.80	[0.52, 0.96]
29	Tlaxcala	2.61	2.48	[1.29, 4.70]	0.89	0.90	[0.68, 0.99]	0.76	0.77	[0.50, 0.96]
30	Veracruz	3.56	3.43	[2.21, 5.67]	0.91	0.94	[0.69, 1.00]	0.79	0.82	[0.50, 0.97]
31	Yucatán	2.44	2.33	[1.41, 4.07]	0.92	0.94	[0.73, 1.00]	0.79	0.81	[0.54, 0.96]
32	Zacatecas	5.00	4.83	[3.01, 7.97]	0.92	0.95	[0.69, 1.00]	0.77	0.79	[0.47, 0.96]

Notes: 95% CI indicates 95% credible interval.

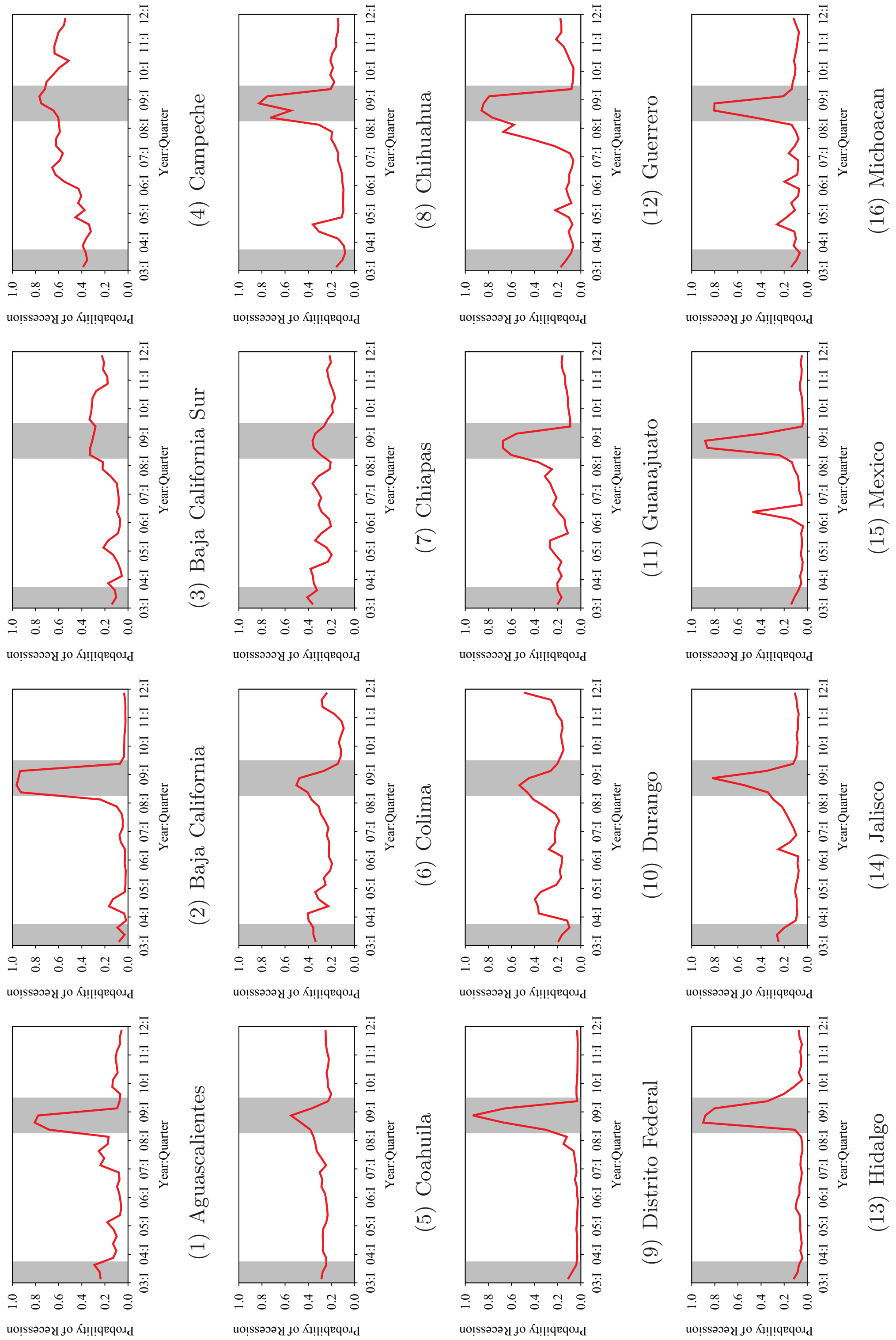


Figure 2.10: Recession Probabilities for MS-AR(1)

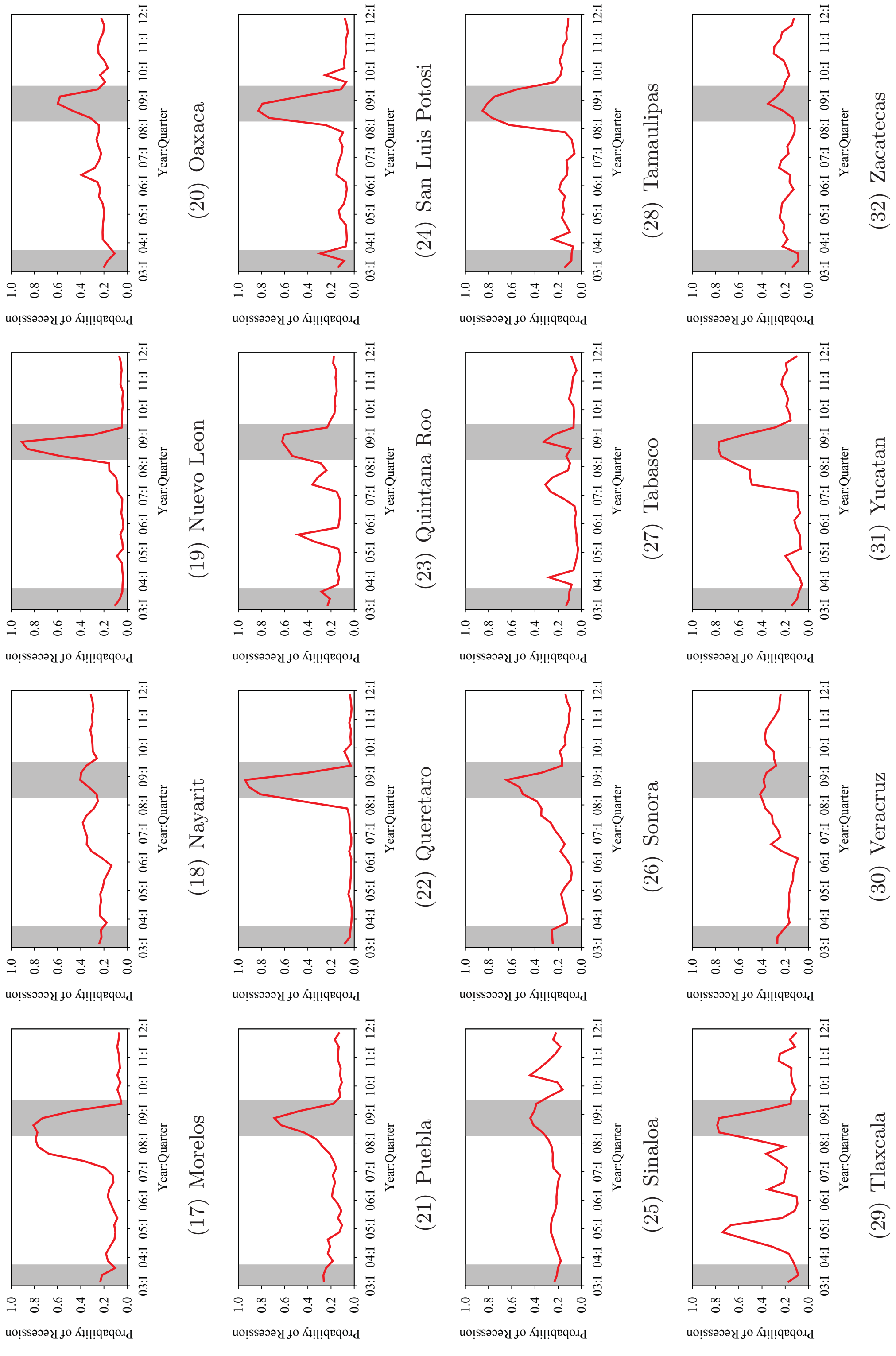


Figure 2.10: Recession Probabilities for MS-AR(1) (Continued)

Table 2.10: Estimated Parameters for MS-SAR-AR(1)

	ρ		
	Mean	Median	95% CI
Spatial Dependence	0.29	0.29	[0.24, 0.33]

		μ_0			μ_1			ϕ		
Code	State	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI
1	Aguascalientes	−1.16	−1.15	[−3.03, 0.54]	1.03	1.02	[0.18, 1.91]	0.10	0.10	[−0.20, 0.41]
2	Baja California	−1.41	−1.46	[−3.23, 0.32]	0.77	0.77	[0.01, 1.50]	−0.03	−0.03	[−0.36, 0.32]
3	Baja California Sur	−0.13	−0.02	[−2.10, 1.30]	1.49	1.46	[0.68, 2.49]	−0.40	−0.40	[−0.70, −0.09]
4	Campeche	−1.61	−1.56	[−2.76, −0.73]	−0.31	−0.46	[−1.49, 1.64]	−0.19	−0.19	[−0.48, 0.11]
5	Coahuila	−0.71	−0.62	[−2.54, 0.65]	0.69	0.65	[−0.40, 1.96]	0.11	0.11	[−0.17, 0.40]
6	Colima	−0.52	−0.44	[−2.20, 0.73]	0.94	0.88	[−0.06, 2.30]	0.09	0.09	[−0.20, 0.39]
7	Chiapas	−0.73	−0.65	[−2.35, 0.51]	0.54	0.48	[−0.52, 1.93]	−0.15	−0.14	[−0.47, 0.18]
8	Chihuahua	−1.38	−1.16	[−3.92, 0.43]	0.79	0.79	[−0.13, 1.72]	−0.05	−0.05	[−0.34, 0.26]
9	Distrito Federal	−2.06	−2.21	[−3.87, 0.23]	0.65	0.66	[0.14, 1.13]	−0.11	−0.12	[−0.47, 0.27]
10	Durango	−0.68	−0.54	[−2.25, 0.23]	0.34	0.28	[−0.26, 1.40]	0.16	0.16	[−0.20, 0.51]
11	Guanajuato	−0.54	−0.43	[−2.24, 0.61]	0.81	0.75	[−0.04, 2.04]	−0.03	−0.03	[−0.33, 0.26]
12	Guerrero	−0.67	−0.59	[−2.05, 0.28]	0.48	0.45	[−0.12, 1.28]	0.02	0.03	[−0.31, 0.36]
13	Hidalgo	−1.49	−1.55	[−3.52, 0.54]	1.13	1.15	[0.19, 2.00]	−0.20	−0.21	[−0.52, 0.15]
14	Jalisco	−1.39	−1.26	[−3.79, 0.40]	0.63	0.62	[−0.06, 1.37]	−0.01	−0.01	[−0.33, 0.28]
15	México	−1.24	−1.32	[−2.95, 0.47]	0.81	0.81	[0.15, 1.46]	0.20	0.20	[−0.13, 0.51]
16	Michoacán	−0.80	−0.65	[−2.81, 0.55]	0.68	0.66	[−0.02, 1.47]	−0.14	−0.14	[−0.45, 0.16]
17	Morelos	−0.63	−0.60	[−1.97, 0.50]	0.78	0.77	[0.10, 1.53]	−0.27	−0.27	[−0.60, 0.06]
18	Nayarit	−0.51	−0.44	[−2.23, 0.80]	0.90	0.86	[−0.20, 2.24]	−0.06	−0.06	[−0.37, 0.26]
19	Nuevo León	−1.36	−1.43	[−3.15, 0.58]	1.22	1.23	[0.48, 1.93]	−0.10	−0.09	[−0.40, 0.19]
20	Oaxaca	−0.40	−0.23	[−2.13, 0.55]	0.61	0.54	[−0.03, 1.77]	−0.22	−0.22	[−0.51, 0.07]
21	Puebla	−0.70	−0.59	[−2.71, 0.75]	0.98	0.95	[0.00, 2.07]	0.03	0.03	[−0.31, 0.37]
22	Querétaro	−1.78	−1.85	[−3.44, 0.28]	1.26	1.26	[0.46, 2.02]	0.06	0.07	[−0.28, 0.38]
23	Quintana Roo	−0.91	−0.79	[−3.24, 0.89]	1.38	1.34	[0.12, 2.75]	−0.05	−0.05	[−0.34, 0.27]
24	San Luis Potosí	−1.17	−1.19	[−2.98, 0.57]	1.03	1.04	[0.19, 1.82]	−0.10	−0.10	[−0.41, 0.23]
25	Sinaloa	−0.52	−0.41	[−2.31, 0.72]	0.82	0.78	[−0.09, 2.04]	−0.06	−0.06	[−0.39, 0.27]
26	Sonora	−0.61	−0.52	[−2.60, 0.88]	1.16	1.14	[0.21, 2.20]	−0.27	−0.27	[−0.56, 0.02]
27	Tabasco	−0.14	−0.09	[−2.18, 1.55]	1.78	1.77	[1.05, 2.54]	−0.65	−0.65	[−0.94, −0.36]
28	Tamaulipas	−0.90	−0.80	[−2.53, 0.39]	0.85	0.81	[−0.11, 2.16]	−0.08	−0.08	[−0.42, 0.26]
29	Tlaxcala	−0.77	−0.62	[−2.65, 0.42]	0.61	0.56	[−0.14, 1.69]	0.12	0.12	[−0.21, 0.45]
30	Veracruz	−0.35	−0.21	[−2.07, 0.73]	0.82	0.76	[0.04, 1.96]	−0.11	−0.12	[−0.44, 0.22]
31	Yucatán	−0.16	−0.09	[−1.92, 1.12]	1.46	1.43	[0.71, 2.39]	−0.08	−0.08	[−0.41, 0.26]
32	Zacatecas	−0.20	−0.09	[−2.09, 1.16]	1.30	1.25	[0.43, 2.52]	−0.38	−0.38	[−0.67, −0.08]

Notes: 95% CI indicates 95% credible interval.

Table 2.10: Estimated Parameters for MS-SAR-AR(1) (Continued)

		σ^2			p_{11}			p_{00}		
Code	State	Mean	Median	95% CI	Mean	Median	95% CI	Mean	Median	95% CI
1	Aguascalientes	3.31	3.15	[1.72, 5.74]	0.91	0.93	[0.71, 0.99]	0.75	0.77	[0.48, 0.95]
2	Baja California	2.72	2.60	[1.50, 4.64]	0.92	0.94	[0.74, 0.99]	0.77	0.79	[0.50, 0.96]
3	Baja California Sur	4.22	4.07	[2.56, 6.70]	0.93	0.95	[0.73, 1.00]	0.78	0.80	[0.50, 0.97]
4	Campeche	4.12	3.98	[2.42, 6.66]	0.86	0.88	[0.59, 0.99]	0.84	0.87	[0.55, 0.98]
5	Coahuila	12.99	12.52	[8.08, 20.40]	0.90	0.93	[0.68, 1.00]	0.79	0.81	[0.50, 0.97]
6	Colima	6.55	6.34	[3.95, 10.30]	0.90	0.93	[0.67, 1.00]	0.79	0.80	[0.51, 0.96]
7	Chiapas	10.05	9.73	[6.12, 15.91]	0.90	0.93	[0.65, 1.00]	0.79	0.81	[0.51, 0.97]
8	Chihuahua	4.68	4.59	[1.92, 8.12]	0.90	0.92	[0.71, 0.99]	0.76	0.77	[0.46, 0.96]
9	Distrito Federal	1.18	1.09	[0.65, 2.13]	0.94	0.96	[0.82, 0.99]	0.75	0.76	[0.47, 0.95]
10	Durango	1.74	1.68	[1.03, 2.82]	0.91	0.94	[0.68, 1.00]	0.79	0.81	[0.50, 0.97]
11	Guanajuato	4.03	3.88	[2.38, 6.50]	0.90	0.93	[0.67, 1.00]	0.79	0.81	[0.52, 0.97]
12	Guerrero	1.50	1.45	[0.84, 2.44]	0.90	0.93	[0.69, 1.00]	0.79	0.81	[0.52, 0.97]
13	Hidalgo	4.49	4.26	[2.50, 7.78]	0.93	0.95	[0.76, 1.00]	0.77	0.79	[0.51, 0.95]
14	Jalisco	2.87	2.76	[1.53, 4.85]	0.92	0.94	[0.74, 1.00]	0.75	0.76	[0.46, 0.96]
15	México	1.66	1.58	[0.83, 2.96]	0.92	0.94	[0.75, 0.99]	0.74	0.75	[0.46, 0.95]
16	Michoacán	2.93	2.84	[1.62, 4.82]	0.91	0.93	[0.70, 1.00]	0.77	0.78	[0.47, 0.97]
17	Morelos	1.94	1.86	[1.13, 3.18]	0.92	0.94	[0.73, 1.00]	0.79	0.81	[0.54, 0.96]
18	Nayarit	11.68	11.30	[7.32, 18.31]	0.90	0.93	[0.66, 1.00]	0.79	0.81	[0.51, 0.97]
19	Nuevo León	2.75	2.60	[1.52, 4.77]	0.93	0.94	[0.78, 0.99]	0.76	0.77	[0.49, 0.95]
20	Oaxaca	2.15	2.07	[1.31, 3.41]	0.91	0.94	[0.67, 1.00]	0.79	0.81	[0.49, 0.97]
21	Puebla	7.12	6.87	[4.25, 11.53]	0.91	0.94	[0.70, 1.00]	0.78	0.79	[0.51, 0.96]
22	Querétaro	2.48	2.34	[1.44, 4.34]	0.94	0.95	[0.82, 0.99]	0.77	0.78	[0.51, 0.95]
23	Quintana Roo	9.83	9.64	[4.33, 16.73]	0.91	0.92	[0.71, 1.00]	0.76	0.77	[0.47, 0.96]
24	San Luis Potosí	3.28	3.14	[1.72, 5.63]	0.92	0.94	[0.73, 1.00]	0.76	0.77	[0.48, 0.95]
25	Sinaloa	6.68	6.47	[4.09, 10.67]	0.91	0.93	[0.67, 1.00]	0.78	0.80	[0.50, 0.97]
26	Sonora	5.97	5.75	[3.51, 9.71]	0.92	0.94	[0.73, 1.00]	0.77	0.79	[0.49, 0.96]
27	Tabasco	3.64	3.53	[2.26, 5.75]	0.95	0.97	[0.77, 1.00]	0.78	0.80	[0.49, 0.97]
28	Tamaulipas	3.71	3.55	[2.09, 6.15]	0.90	0.92	[0.67, 0.99]	0.80	0.81	[0.53, 0.97]
29	Tlaxcala	2.95	2.86	[1.68, 4.84]	0.89	0.92	[0.66, 1.00]	0.78	0.79	[0.50, 0.97]
30	Veracruz	3.41	3.30	[2.11, 5.34]	0.91	0.94	[0.69, 1.00]	0.80	0.82	[0.51, 0.97]
31	Yucatán	2.22	2.14	[1.28, 3.59]	0.92	0.94	[0.70, 1.00]	0.78	0.80	[0.51, 0.96]
32	Zacatecas	5.06	4.91	[2.99, 8.03]	0.91	0.94	[0.66, 1.00]	0.78	0.80	[0.49, 0.97]

Notes: 95% CI indicates 95% credible interval.

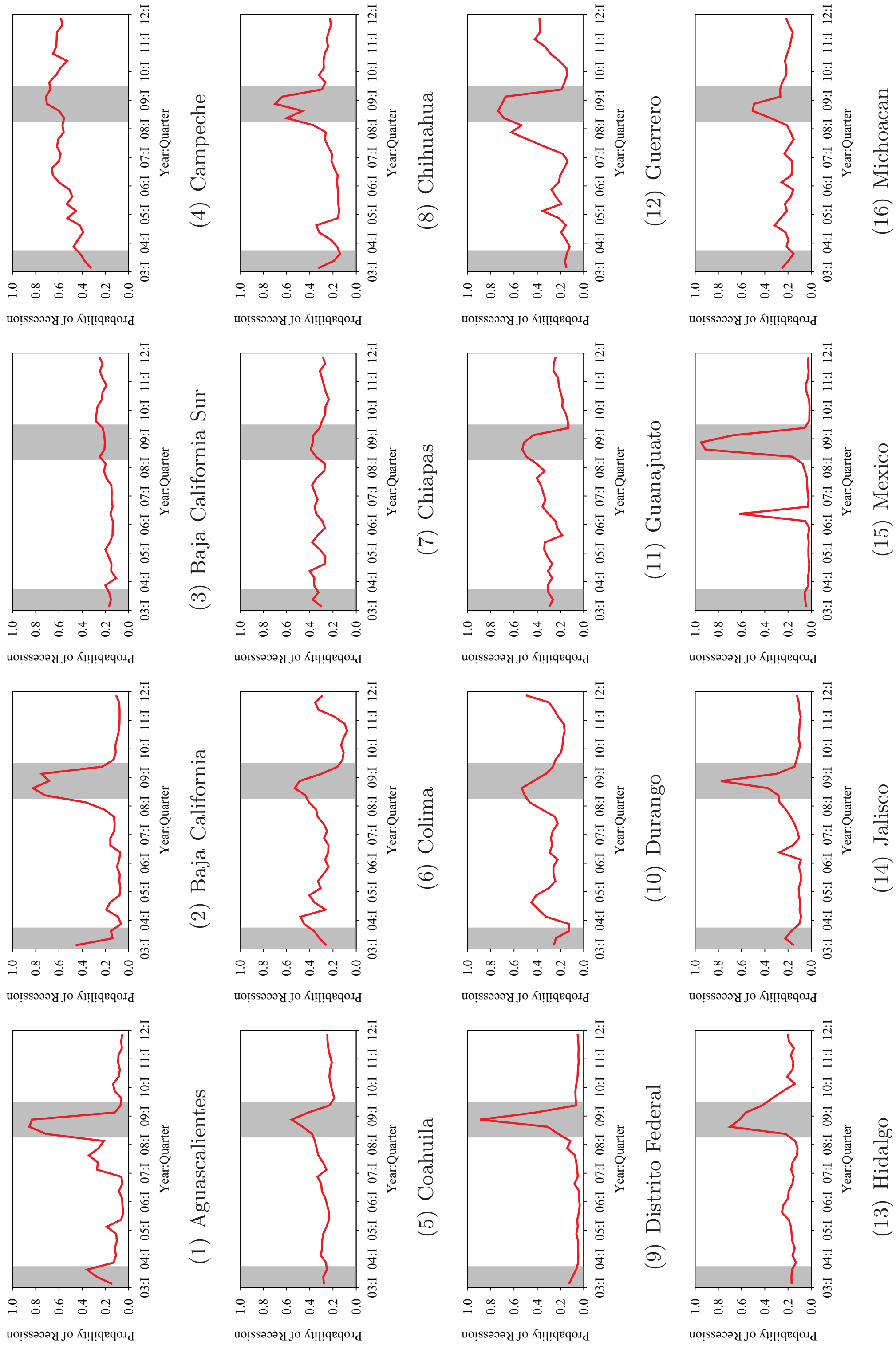


Figure 2.11: Recession Probabilities for MS-SAR-AR(1)

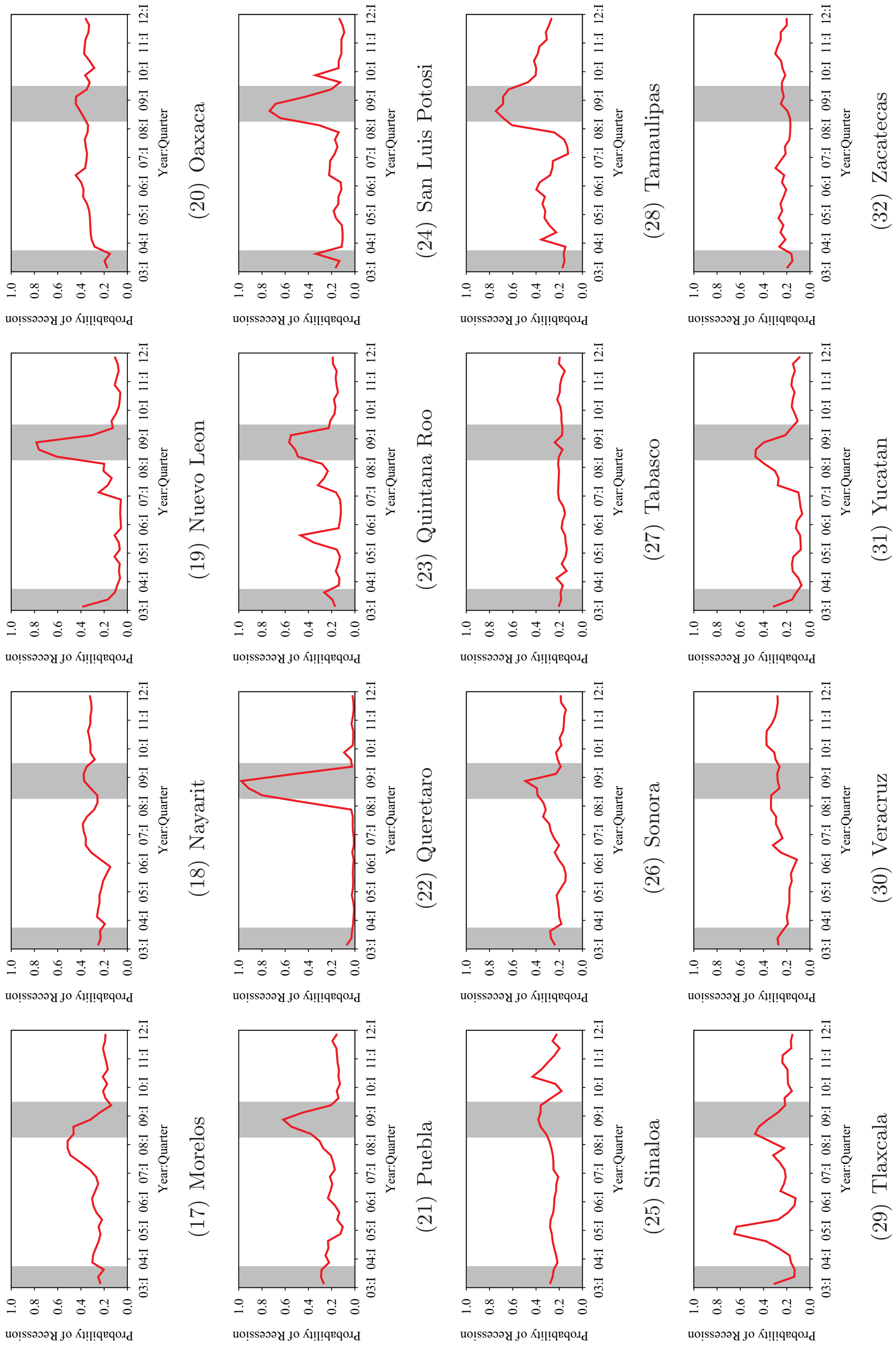


Figure 2.11: Recession Probabilities for MS-SAR-AR(1) (Continued)

Chapter 3

Do Workers Really Benefit from Agglomeration? The Case of the Mexican Banking Sector*

3.1 Introduction

It has been often said that our economic activities benefit from agglomeration arising from interactions between job seekers and firms in labor markets, linkages between intermediate and final goods suppliers, and knowledge creation and spillover.[1] Duranton and Puga (2004) studied the theoretical mechanisms that underlie the micro-foundations of agglomeration economies. Rosenthal and Strange (2004) surveyed empirical findings to identify the factors that enhance agglomeration of economic activities. The objective of these studies was to uncover the source and nature of agglomeration economies. In recent years, there has been an expanding literature on agglomeration economies that attempts to answer why workers living in denser areas earn more than those living in less dense areas. The aforementioned studies attempt to identify what factors explain wage premiums arising from agglomeration and examine whether agglomeration benefits persist even after controlling for possible factors.

The existing literature clarified the existence of an agglomeration effect on wages in denser areas, resulting from positive externalities of agglomeration. Recent studies have

*I would like to specially thank Kentaro Nakajima and Kensuke Teshima for their helpful comments and suggestions. I also thank Nobuaki Hamaguchi, Yoichi Matsubayashi, Akio Namba, Komei Sasaki, and all the participants in the 2013 annual meeting of the Applied Regional Science Conference at Kyoto University and a seminar at the Research Institute of Economy, Trade and Industry for their useful comments and suggestions. Naturally, any remaining errors are my own. I am grateful for the benefits accorded to me during my stay at the Instituto Tecnológico Autónomo de México. This research was conducted under a scholarship granted by the Government of Mexico, through the Ministry of Foreign Affairs of Mexico.

[1] Marshall (1890) was among the first to observe benefits from agglomeration of economic activities.

investigated the possible explanations for this phenomenon from static and dynamic perspectives. Notably, two important ideas may reveal why wages are higher in bigger cities. First, sorting of skills by city size explains agglomeration economies; inherently, high-skilled workers tend to locate in big cities. This phenomenon has been explored by Combes et al. (2008a), Mion and Naticchioni (2009), Combes et al. (2012b), and Matano and Naticchioni (2012). Combes et al. (2010) noted that about half the agglomeration effects on wages estimated without control variables are explained by spatial sorting of inherent workers' skills.[2] Sorting of skills is classified as a static aspect of an agglomeration economy. Second, differences in city sizes give rise to differences in learning speeds. Workers in bigger cities easily accumulate human capital by taking advantage of their meaningful experiences there. de la Roca and Puga (2012) found that experiences accumulated in bigger cites make the slope of the wage profile steeper, and this effect does not disappear even after migration to smaller areas.[3] Gould (2007) also found that human capital gains acquired from working in big cities remain valuable for white-collar workers even when they relocate to rural areas, although the same cannot be said for blue-collar workers.[4] Learning by working in big cities is classified as a dynamic aspect of an agglomeration economy.

In this paper, we explore the idea of agglomeration effects on wages, especially when firms have branch networks. The existing literature has proved the existence of an agglomeration effect on wages even after controlling for spatial sorting of skills. However, agglomeration affects both workers and firms simultaneously. In fact, Combes et al. (2010, 2012a) showed evidence that total factor productivities of firms in denser areas are also higher, due to agglomeration economies. Thus, the manner in which workers directly benefit from local agglomeration in bigger areas remains unclear when firms' branches are located in less dense areas as well. Specifically, this paper tries to clarify agglomeration effects on wages from two perspectives. One effect is that workers directly benefit from agglomeration in denser areas, that is, from a local perspective. Another effect is that workers indirectly benefit from agglomeration through firms' branch networks, that is, from a global perspective. In the latter case, the branch network would be beneficial for workers employed in less dense areas through transfer mechanism.

---

[2]Combes et al. (2010, 2011) referred to spatial sorting of skills as endogenous quality of labor. They also examined endogenous quantity of labor, or in other words, the simultaneous causality of wage and density. That is, more productive areas attract more workers, resulting in denser areas. However, the effect of endogenous quantity of labor is not as strong as that of endogenous quality of labor.

[3]See also Puga (2010) for an explanation about learning by working in big cities.

[4]Glaeser and Maré (2001) and Glaeser and Resseger (2010) also found evidence of static urban wage premiums, that human capital accumulation is faster in bigger areas, and that learning effects are stronger in skill-intensive areas.

For this purpose, we focus on banks' branches networks. If workers directly benefit from an agglomeration economy, higher wages would be paid in branches located in denser areas after controlling for individual characteristics and occupation. If the indirect effect through banks' branch network is significant, bank fixed effects would be positively correlated with agglomeration (e.g., average population density of places where each bank is located), thus suggesting that higher wages might be explained by higher bank fixed effects.

For our empirical analysis, we need workers' micro-data, including their firm-level information.[5] The Mexican workers' micro-data set from the National Survey of Occupation and Employment (*Encuesta Nacional de Ocupación y Empleo*) also provides information regarding the names of the workers' firms. Thus, we examine direct and indirect effects of agglomeration on wages by making use of firm information. Notably, this paper focuses on the banking sector, which suits the objective of this paper; in terms of population density, the banks have a large number of branches across different areas. Furthermore, since banking services are offered locally, the effects of agglomeration economies are striking. Conversely, for the manufacturing sector, factories might be located in less dense regions, depending on transport costs. Thus, we consider that banks' branches located in denser areas strongly would affect workers' and firms' productivities, from a local perspective.

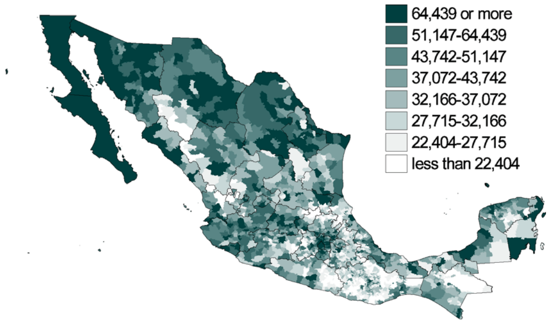
Our empirical framework follows the simplest framework proposed by Combes et al. (2011), which integrates the agglomeration variable (e.g., population or employment density) into a standard Mincerian equation. This basic framework was also followed by Mion and Naticchioni (2009).[6] The coefficient estimate of the agglomeration variable measures the extent to which the agglomeration economy leads to a kind of wage premium.[7]

We further assess what types of workers can enjoy more benefits from agglomeration. The existing literature shows that depending on their education level, workers are differently affected by the agglomeration economy. For example, Hering and Poncet (2010) examined the micro-data of Chinese workers and found that high-skilled workers tend to benefit from geographical externality whereas low-skilled workers do not. This difference is considered to be remarkable in developing countries. We explore these possibilities by dividing the

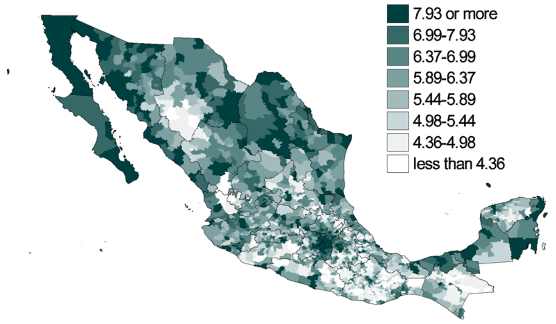
[5] Mion and Naticchioni (2009) and Matano and Naticchioni (2012) used employee-employer matched data. However, their motivations differ from ours. Our interest lies in identifying the direct and indirect agglomeration effects on wages in branch networks.

[6] See also Combes et al. (2008a, Chap. 11).

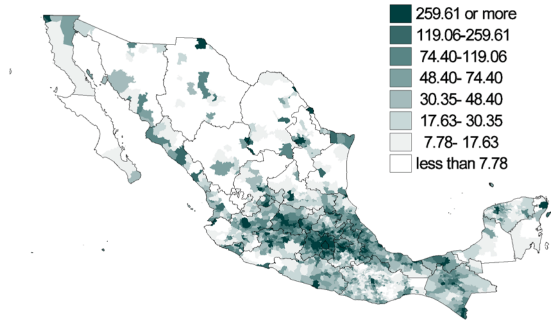
[7] The endogeneity problem of population density gives rise to a biased estimator. Following Ciccone and Hall (1996), we also use lagged population density as an instrumental variable. Although long-lagged data is preferred, we use the population density in 1990. This approach is quite common in the literature and has been used by Combes et al. (2008a), Mion and Naticchioni (2009), Combes et al. (2010), and Matano and Naticchioni (2012).



(a) GDP per Capita



(b) Years of Schooling



(c) Population Density

Figure 3.1: Geographic Distributions in 2005

Notes: The spatially smoothed population densities are used. See Section 3.4.
Data source: SNIM 2005

sample by bank and workers' education level.

First, we describe the Mexican municipal data. Figure 3.1 shows maps of gross domestic product (GDP) per capita in Panel (a), average years of schooling in Panel (b), and density of population aged 15 and over in Panel (c) in 2005. We can see that GDP per capita and years of schooling are highly correlated. In addition, population density also shows positive relationships with GDP per capita and years of schooling around the central region of Mexico.[8] Although years of schooling cannot capture unobservable skills, the positive relationship between years of schooling and population density implies sorting of skills by area size. In the empirical study, therefore, we analyze whether there is an agglomeration effect on wages even after controlling for years of schooling and individual observable characteristics.

This paper contributes to the existing literature on agglomeration economies by un-

[8]The administrative areas in the northern states are comparatively bigger than those in the central and southern states and their population densities tend to be smaller.

covering a route for agglomeration effects on wages in branch networks. Our empirical findings show that agglomeration effects on wages continue to remain even after controlling for individual characteristics, occupation, time effects, firm size, and bank fixed effects in the pooled data. At the same time, workers indirectly benefit from agglomeration through banks' branch networks, suggesting that branch networks extend additional agglomeration effects, especially to workers employed in less dense areas. Furthermore, we find that within the banks' branch networks, high-skilled workers are likely to benefit from agglomeration, whereas low-skilled workers are not the case; the basic wage tends to be the same between workers with regard to occupation, but the wage premium is likely to be paid based on individual skills.[9] To make matters worse, low-skilled workers can be negatively affected by agglomeration as a congestion effect.

The rest of the paper is organized as follows. In Section 2, we briefly document Mexico's banking sector. In Section 3, we explain the empirical framework used in this paper to identify the effect of density. Section 4 presents the data. Section 5 presents the estimation results and our interpretations for them. Finally, Section 6 concludes.

3.2 Mexican Banking Sector

In this paper, we focus on the banking sector in Mexico. Table 3.1 presents the list of commercial banks licensed under Mexican legislation through the National Banking and Securities Commission (*Comisión Nacional Bancaria y de Valores*, CNBV). None of the workers in our data set works in 11 of the 42 commercial banks considered in this study. According to Banco de México (2011), the large-sized banks in Mexico are BBVA Bancomer, Banamex, Santander, Banorte, HSBC, Inbursa, and Scotiabank, which are part of financial groups.

A characteristic feature of the Mexican banking sector is its considerably high share of foreign capital. After the reprivatization of banks around 1991, the negotiation of the North American Free Trade Agreement (NAFTA) raised expectations about the entry of the U.S. and Canadian banks into the Mexican market. The NAFTA has indeed played an important role in reducing regulations for foreign banks within the Mexican banking system. Simultaneously, the peso crisis in 1994 also greatly accelerated deregulation on foreign investment in the Mexican banking sector, to attract more capital and keep Mexican banks from collapsing (Lubrano, 1998; Minushkin, 2000; Sigmond, 2011). Then, regulations on equity stakes were finally eliminated in December 1998, and the entry of foreign banks

[9] This view comes from the commission system and performance-related payment.

Table 3.1: Commercial Banks in Mexico

Bank Name	Assets	Bank Name	Assets
Large-Sized	4679.9	Small Subsidiaries of Foreign Banks	425.2
BBVA Bancomer	1263.5	Deutsche Bank México*	226.3
Banco Nacional de México (Banamex)	1112.1	Bank of America Mrrill Lynch*	51.9
Banco Santander	747.8	Banco Credit Suisse (México)*	47.8
Banco Mercantil del Norte (Banorte)	603.3	Banco JP Morgan*	40.3
HSBC México	485.5	Barclays Bank México*	18.5
Banco Inbursa	243.5	Bank of Tokyo-Mitsubishi UFJ (México)*	14.6
Scotiabank	224.2	American Express Bank México	12.6
Medium-Sized	767.5	Volkswagen Bank	5.2
Banca Afirme	104.3	The Royal Bank of Scotland México*	3.8
Banco del Bajío	102.8	UBS Bank México*	2.3
Banco Interacciones	101.8	ING Bank México*	1.3
Ixe Banco	100.1	The Bank of New York Mellon*	0.8
Banco Regional de Monterrey	72.2	Associated with Commercial Chains	306.3
Banco Invex	43.5	Banco Fácil	174.9
Banco Monex	42.2	Banco Azteca	92.5
Banca Mifel	37.6	Bancoppel	18.0
Banco Multiva	31.8	Banco Ahorro Famsa	14.5
CIBanco	25.9	Banco Wal-Mart de México Adelante	6.3
Banco Ve por Más	24.3		
InterBanco*	17.8		
Banco Compartamos	17.3		
BANSÍ	13.5		
Banco Actinver	11.7		
Banco Base	10.4		
ABC Capital	5.6		
Banco Autofin México	4.8		

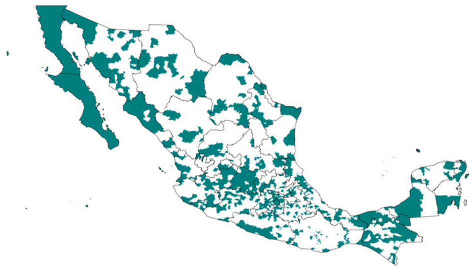
Notes: As of December 2012. Assets are expressed in billions of pesos. * indicates no sample in our data set.
Source: *Comisión Nacional Bancaria y de Valores*

finally got underway. The main transactions of acquisitions during the early 2000s included: Santander (Spain) buying Serfin in May 2000, BBVA (Spain) buying Bancomer in July 2000, Scotiabank (Canada) buying Inverlat in November 2000, Citibank (US) buying Banamex in August 2001, and HSBC (UK) buying Bital in November 2002.[10]

Figure 3.2 presents the geographical distribution of the branches of the large-sized banks at the municipal level. BBVA Bancomer has the largest number of branches in Mexico, followed by Banamex. These banks cover the majority of Mexican municipalities and have nation-wide banking networks. The locations of their branches, by and large, correspond

[10] See Minushkin (2000) and Turrent (2011) for a historical overview of the Mexican banking sector.

(a) BBVA Bancomer (547; 1796)



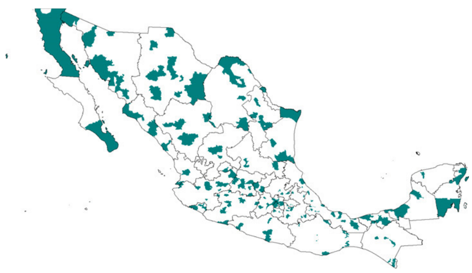
(b) Banamex (528; 1703)

(c) Santander (303; 1142)

(d) Banorte (345; 1142)

(e) HSBC (328; 1040)

(f) Inbursa (114; 294)



(g) Scotiabank (208; 639)

Figure 3.2: Banks' Branches at the Municipality Level in December 2012

Notes: Colored municipalities have at least one bank branch. The numbers on the left- and right-hand sides within the parentheses indicate the numbers of municipalities and branches, respectively.

Data source: CNBV

to municipalities with higher GDP per capita (Figure 3.1).

3.3 Empirical Framework

3.3.1 Theoretical Background

We slightly modify the micro-economic foundation suggested by Combes et al. (2008a,b).[11] The profit of firm j operating in area a at time t is given by

$$\pi_{jt} = p_t y_{jt} - \sum_{i \in (j,a,t)} w_{it} \ell_{it} - r_t k_{jt},$$

where p_t is the market price of the product, and y_{jt} is the output. For any worker i, w_{it} is the wage rate, and ℓ_{it} is the amount of labor supply. In addition, k_{jt} represents the other factors of production, and r_t is their market price.

We assume that the production function takes a Cobb-Douglas form with constant returns to scale:

$$y_{jt} = A_{jat} \left(\sum_{i \in (j,a,t)} s_{it} \ell_{it} \right)^{\xi} (k_{jt})^{1-\xi}, \quad 0 < \xi \leq 1,$$

where A_{jat} is the total factor productivity, and s_{it} represents skills of worker i. It is assumed that A_{jat} differs in terms of the population (or employment) density of area a. This assumption reflects the empirical findings of Combes et al. (2010) and Combes et al. (2012a) regarding agglomeration economies. Solving profit maximization with respect to labor yields the following condition on the wage rate:

$$w_{it} = \xi p_t A_{j(it)at} \left(\frac{k_{j(it)t}}{\sum_{i \in (j,a,t)} s_{it} \ell_{it}} \right)^{1-\xi} s_{it}.$$

Matching it with the first-order condition for profit maximization with respect to other factors, we have the following wage equation:

$$w_{it} = B_{j(it)at} s_{it}, \quad \text{where} \quad B_{j(it)at} \equiv \xi (1-\xi)^{(1-\xi)/\xi} \left(\frac{p_t}{(r_t)^{1-\xi}} \right)^{1/\xi} (A_{j(it)at})^{1/\xi}. \qquad (3.1)$$

We can see that wage of worker i differs in terms of total factor productivity $A_{j(it)at}$ and individual skills $s_{i,t}$ in equation (3.1).[12]

[11] Note that for simplification, we do not explicitly consider how branch networks affect the wage rate between firms' branches in this model.

[12] Combes et al. (2008a) also considered the case wherein inputs and output markets are segmented between

3.3.2 Specification and Interpretation

To derive a regression model, we need to specify the firm productivity term $B_{j(it)t}$ and the individual skill term s_{it}, respectively. First, assuming that firm productivity depends on the population density of area a and firm-year fixed effect ψ_{jt}, we have

$$\log(B_{j(it)at}) = \alpha^j \log(\mathrm{Dens}_{a(j(it)t)t}) + \psi_{jt}, \tag{3.2}$$

where α^j is a parameter measuring how much agglomeration affects firm productivity. It is considered that α^j differs at the firm level.

Second, the individual skill of worker i is assumed to take the following form:

$$\log(s_{it}) = \alpha^i \log(\mathrm{Dens}_{a(it)t}) + \boldsymbol{X}_{it}\boldsymbol{\beta} + \mu_i, \tag{3.3}$$

where α^i is a parameter measuring how much agglomeration affects labor efficiency regarding individual skills, $\boldsymbol{X}_{it}$ is a vector of time-varying workers characteristics, $\boldsymbol{\beta}$ is a vector of parameters, and μ_i is a worker fixed effect. In the spirit of human capital externalities noted by Rauch (1993), Acemoglu and Angrist (2001), and Moretti (2004), we largely interpret the idea of this externality such that the denser area affects labor (in)efficiency through the exchange of ideas, imitation, learning by doing, and fierce competition in equation (3.3). It is considered that α^i differs at the individual level.

Taking the logarithm of equation (3.1) and inserting equations (3.2) and (3.3) into it, we obtain the following regression model:

$$\log(w_{it}) = \alpha \log(\mathrm{Dens}_{at}) + \psi_{jt} + \boldsymbol{X}_{it}\boldsymbol{\beta} + \mu_i + u_{ijat}, \tag{3.4}$$

where $\alpha \equiv \alpha^j + \alpha^i$, and u_{ijat} is an error term. For convenience of explanation, our notations for firm and area indices j and a, respectively, are now simplified. Our empirical framework may be interpreted as an extended approach that integrates the agglomeration effect into a Mincerian equation.

The key point for empirical analysis is that the coefficient of population density α captures effects of local agglomeration on wages through total factor productivity as well as individual skills in area a. The coefficient estimate of population density measures the overall effect among the effects from total factor productivity and individual skills, which represents the benefits from local agglomeration that workers directly receive. Note that in our empirical analysis, we interpret ψ_{jt} as bank fixed effects, not banks' branch fixed

areas. In that case, wage differences across areas can occur through total factor productivity, price of the output, or price of the other factor.

effects.

Focusing on banks' branches, we separately identify local agglomeration effects on wages and the effect of global banks' branch networks on wages. That is, $\alpha \log(\text{Dens}_{at})$ captures whether workers employed in area a directly benefit from local agglomeration effects of wage, and simultaneously, ψ_{jt} captures whether they indirectly benefit from bank' branch networks' effects.

3.3.3 Estimation Strategy

The first objective of this paper is to examine whether the wage is higher in denser areas and if so by how much. The wage of worker i employed in bank j and working in area a at time t is given by

$$\log(w_{ijat}) = \alpha \log(\text{Dens}^s_{at}) + \boldsymbol{X}_{it}\boldsymbol{\beta} + \psi_j + \tau_t + u_{ijat}, \tag{3.5}$$

where α is a parameter of our interest, Dens^s_{at} is the spatially smoothed population density of area a, $\boldsymbol{X}_{it}$ is a row vector of the worker's characteristics, ψ_j is the bank fixed effect, τ_t is a time effect, and u_{ijat} is the error term.[13] Due to data limitations, we deal with banks' branches at the municipality level. Owing to collinearity in the data, we cannot separately estimate bank and individual fixed effects. Thus, individual fixed effect μ_i is dropped in regression model (3.5). From coefficient α, we assess whether workers directly benefit from working in denser areas. Furthermore, we investigate whether workers indirectly benefit from banks' branch networks by focusing on bank fixed effects ψ_j.

The second objective of this paper is to investigate whether a local agglomeration effect on wages still exists among branches of the same firms. Previous studies on agglomeration economies did not distinguish how workers receive the benefit from agglomeration within the framework of a branch network. The existing literature showed that wages are higher in denser areas. However, as agglomeration itself affects both workers and firms, this finding does not indicate whether workers really enjoy the direct benefits arising from local agglomeration. If the branch network plays an important role, workers will receive agglomeration benefits from a global perspective, and not a local perspective.

The third objective of this paper is to analyze heterogeneous effects of agglomeration

[13] Combes et al. (2011) prefer a two-step procedure for identifying the effect of density. In the first step, the wage is regressed on with area dummies. In the second step, the estimates of area dummies are regressed on by population density. If we follow the same methodology, a problem arises. In our case, the sample size is inadequate to allow consistent estimates of municipality dummies; some municipality dummies might be estimated by using only a few workers. To avoid this problem, we introduce the population density directly in the Mincerian equation.

on wages between high- and low-skilled workers. The existing literature already mentions that workers tend to be differently affected by the agglomeration economy depending on skill levels. As stated earlier, Hering and Poncet (2010) analyzed the micro-data of Chinese workers and found that high-skilled workers tend to benefit from geographical externality, whereas low-skilled workers do not. This difference is considered to be remarkable in developing countries. We explore these possibilities by dividing the sample by bank and workers' education level.

3.3.4 Estimation Issues

There are some important estimation issues to be controlled for. If we conduct a regression analysis of model (3.5) by using ordinary least squares (OLS), the estimate of α might be biased due to the endogeneity problem arising from several reasons. Combes et al. (2010, 2011) referred to two main sources of endogeneity: endogenous quantity and quality of labor.[14] As for endogenous quantity of labor, more productive areas attract more workers, and consequently, the density also rises to a greater extent. To solve this endogeneity problem, they suggested the instrumental variable (IV) method using long lagged density variable.[15] We also rely on the same estimation method by using lagged population density by several decades.

The endogenous quality of labor underscores the workers' skill differences across areas. It is often said that skilled workers tend to prefer living in denser areas for several reasons, such as better amenities in such areas. As a result, higher wages are observed in denser areas. As seen in Figure 3.1, high-skilled people are likely to locate in higher wage areas and in denser areas. In the literature of spatial sorting, individual fixed effects are included to take into account unobservable individual effects. In this paper, instead of including individual effects, we deal with this issue by including observable individual characteristics. As explained previously, this is because area dummies and banks' fixed effects cannot be identified by collinearity.[16]

Furthermore, considering bank fixed effects is also of great importance. As noted before, more productive firms are likely to be located in denser areas, and thus, the high correlation

[14] See also Combes et al. (2008b, Chap. 11) for a discussion on endogeneity problems.

[15] Ciccone and Hall (1996) originally used this method.

[16] In the literature for spatial sorting, individual fixed effects have been included to take into account unobservable individual effects. In this paper, instead of including individual effects, we deal with this issue by including observable individual characteristics. This is because municipality dummies cannot be identified by collinearity unless there are movers between municipalities. Similarly, banks' fixed effects cannot be identified by collinearity unless workers have changed jobs between banks.

between population density and firm fixed effect is expected. This point has not been emphasized in most previous studies. We assess the importance of firm fixed effects in the empirics of agglomeration economies by studying the correlation between estimates of ψ_j and population density.[17]

3.4 Data

We use workers' micro-data from the National Survey of Occupation and Employment (*Encuesta Nacional de Ocupación y Empleo*, ENOE) in Mexico. The ENOE has been conducted by the National Institute of Statistics and Geography (*Instituto Nacional de Estadística y Geografía*,, INEGI) every quarter from 2005.[18] We use data for the time span 2005:Q1–2012:Q4. The ENOE contains data of individual characteristics (e.g., age, gender, marriage status, years of schooling, and living state and municipality) and employment information (e.g., hourly wage, occupation, firm name, industry classification, and firm size).

Our dependent variable, hourly wage, is adjusted by using the consumer price index (CPI) published by the INEGI. The base period is 2005:Q1. Spatial differences in cost-of-living are adjusted by using area differences in minimum wages. In Mexico, there are three categories of minimum wage depending on the degree of urbanization. The base area is that of the lowest minimum wage of the three categories. In line with the literature of pertaining to the Mincerian equation, we calculate potential years of experience by subtracting six years and years of schooling from age.[19] Besides, we control for gender, marriage status, occupation, and firm size. We use the two-digit level occupation codes. Firm size dummies are created for 1–10 workers, 11–20 workers, 21–50 workers, 51–100

[17] The estimates of bank fixed effects are calculated using `areg` in Stata. Consider a regression model:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{D}\boldsymbol{\psi} + \boldsymbol{u},$$

where $\boldsymbol{D}$ is a matrix of banks' fixed effects, and $\boldsymbol{X}$ does not contain constant term. From the Frisch–Waugh–Lovell theorem, we can obtain

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}\boldsymbol{M_D}\boldsymbol{X})^{-1}\boldsymbol{X}\boldsymbol{M_D}\boldsymbol{y},$$

where $\boldsymbol{M_D} \equiv \mathbf{I} - \boldsymbol{D}(\boldsymbol{D}^\top\boldsymbol{D})^{-1}\boldsymbol{D}^\top$. We use estimates of bank fixed effects obtained from the regression model

$$(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) - (\bar{\boldsymbol{y}} - \bar{\boldsymbol{X}}\hat{\boldsymbol{\beta}}) = \boldsymbol{D}\boldsymbol{\psi}^m + \boldsymbol{v},$$

where $\bar{\boldsymbol{y}}$ and $\bar{\boldsymbol{X}}$ indicate average values of dependent and explanatory variables, respectively. Note that the estimates of $\boldsymbol{\psi}^m$ are centered around the mean of $\hat{\boldsymbol{\psi}}$.

[18] The previous version of the ENOE, the National Survey of Employment (*Encuesta Nacional de Empleo*, ENE) had been conducted until 2004:Q4. We do not use this database in this paper.

[19] When years of experience take on a negative value, we set it to zero.

workers, 101–250 workers, 251–500 workers, and 501 workers and over.[20]

This paper focuses on the banking sector, especially commercial banks. The ENOE offers the four-digit code of industry classification and also reports the names of banks. We choose commercial banks licensed under Mexican legislation in August 2012. Finally, our data set includes 31 of the 42 commercial banks.

Population density is available at the municipal level from the population census conducted every five years. Based on the censuses, the National System of Municipal Information (*Sistema Nacional de Información Municipal*, SNIM) provides its summarized municipal data on population and area.[21] We use the 1990, 2005, and 2010 population censuses, in which population aged 15 and over is used to calculate population density.[22] Then, we match the two data sets of the ENOE and SNIM by using the municipal code. The key problem is that the geographical units correspond to administrative units, instead of economic or employment areas. To mitigate this problem, we use spatially smoothed population density.

Let $z_a^s = \sum_{m=1}^{R} \mathbf{1}_{am}(d) z_m$ denote the spatially local sum data of municipality a, where R stands for the number of municipalities; z_m, the raw data of municipality m; and $\mathbf{1}_{am}(d)$, the amth element of the indicator matrix, in which the amth element takes the value of 1 if the distance between municipalities a and m is less than d km and 0 otherwise.[23] We set $d = 40$ km. Thus, the spatially smoothed population density is calculated as $\text{Dens}_a^s = \text{Pop}_a^s / \text{Area}_a^s$, where Pop_a^s and Area_a^s are spatially local sums of population and area, respectively, of municipality a. The descriptive statistics of the variables are shown in Table 3.2.

3.5 Estimation Results

3.5.1 Agglomeration Effects on Wages in the Banking Sector

Table 3.3 presents the OLS estimation results of regression model (3.5). The coefficient estimates of population density in Columns (1)–(4) show significantly positive values. In

[20] As noted in Brown and Medoff (1989), large-sized firms tend to pay more than small-sized firms. We also control for firm size in the regression.

[21] The data are available in `http://www.snim.rami.gob.mx/`.

[22] There is no information regarding municipal area in the 1990 population census. Therefore, we complement municipal areas in 1990 with the corresponding data from the 2000 population census. In doing so, we add the complemented data to that of the original municipalities.

[23] SNIM also offers the latitudes and longitudes of municipalities, from which the bilateral distances between any two municipalities can be calculated using the formula suggested by Vincenty (1975).

Table 3.2: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
Hourly Wage (pesos, Price in 2005:Q1)	39.02	27.33	7.06	203.49
Population Density (Population/km^2)	337.55	649.32	1.29	2789.27
Population Density in 1990 (Population/km^2)	231.89	488.99	0.76	2050.07
Years of Schooling	14.26	2.80	0	21
Experience	12.75	9.95	0	67.25
Femal Dummy (= 1 if a person is female)	0.50	0.50	0	1
Marriage Dummy (= 1 if a person is married)	0.47	0.50	0	1
Firm Size Dummy (1–10 Workers)	0.17	0.38	0	1
Firm Size Dummy (11–20 Workers)	0.30	0.46	0	1
Firm Size Dummy (21–50 Workers)	0.27	0.45	0	1
Firm Size Dummy (51–100 Workers)	0.10	0.30	0	1
Firm Size Dummy (101–250 Workers)	0.05	0.22	0	1
Firm Size Dummy (251–500 Workers)	0.03	0.16	0	1
Firm Size Dummy (501 Workers and over)	0.08	0.27	0	1

Notes: The number of observations is 12875. Hourly wage is adjusted using CPI. In addition, spatial differences in cost-of-living are also adjusted by area differences in minimum wage, and the base area is that of the lowest minimum wage of three categories. Population aged 15 and over is used to calculate of population density. Dummy variables of bank, occupation, and time are not shown due to space limitations.

Column (1), the density elasticity of wage is 0.030. In this regression, firm size and bank fixed effects are not controlled for. After controlling for firm size, the result does not change much. The corresponding density elasticity of wage is 0.021. However, controlling for bank fixed effects changes the coefficient estimate of population density considerably. In Column (4), after controlling for both firm size and bank fixed effects, the density elasticity of wage is 0.019, which is approximately 63% smaller than that in Column (1). These estimation results imply that bank fixed effects are correlated with population density. Therefore, omitting bank fixed effects leads to bias in the coefficient estimate of population density.[24]

Table 3.4 shows the IV estimation results. The coefficient estimates of population density in Columns (1)–(4) are still significantly positive at the 10% at least. The F statistic for weak identification (Weak IV) exceeds thresholds proposed by Stock and Yogo (2005) for the maximal relative bias and maximal size. The results of the Dubin–Wu–Hansuman (DWH) test indicate that endogeneity should be controlled for. The density

[24] The coefficient estimate of years of schooling shows a significantly positive value (0.050) in Column (4). While the period of our data set spans from 2005:Q1–2012:Q4, the estimate takes a value very close to that obtained in the existing literature. For example, Chiquiar (2008) estimated the Mincerian wage equation by using Mexican micro-data gathered from the 1990 and 2000 population censuses. According to his estimation results with full control variables, the estimates are 0.040 for 1990 and 0.051 for 2000.

Table 3.3: OLS Estimation Results for Agglomeration Effects on Wages

	Dependent Variable: $\log(w_{ijat})$			
Explanatory Variable	(1)	(2)	(3)	(4)
Population Density	0.030***	0.027***	0.021**	0.019**
	(0.009)	(0.009)	(0.008)	(0.008)
Years of Schooling	0.061***	0.061***	0.050***	0.050***
	(0.003)	(0.003)	(0.003)	(0.003)
Experience	0.020***	0.020***	0.019***	0.019***
	(0.002)	(0.002)	(0.002)	(0.002)
Experience Squared	−0.018***	−0.018***	−0.020***	−0.019***
	(0.005)	(0.005)	(0.005)	(0.005)
Female Dummy	−0.024**	−0.023**	−0.031***	−0.031***
	(0.011)	(0.011)	(0.010)	(0.010)
Marriage Dummy	0.075***	0.074***	0.075***	0.074***
	(0.015)	(0.015)	(0.015)	(0.015)
Occupation Dummy	Yes	Yes	Yes	Yes
State Dummy	Yes	Yes	Yes	Yes
Time Dummy	Yes	Yes	Yes	Yes
Firm Size Dummy	No	Yes	No	Yes
Bank Fixed Effect Dummy	No	No	Yes	Yes
Number of Observations	12875	12875	12875	12875
Adjusted R^2	0.299	0.301	0.333	0.335

Notes: Heteroskedasticity-robust standard errors clustered by municipality are in the parentheses. Population density is expressed in logarithm. Experience squared is divided by 100. * denotes statistical significance at the 10% level, ** at the 5% level and *** at the 1% level.

elasticity of wage in Column (4) is 0.016 after controlling for both firm size and bank fixed effects. To sum up, the population density positively affects individual wage in the pooled data set of the banking sector. Note, however, that the magnitude of population density on wages decreases by introducing bank fixed effects and the significance level also falls, which implies the existence of indirect agglomeration effects through banks' branch networks.

Figure 3.3 shows the relationship between the bank fixed effects estimated in Column (4) of Table 3.3 and the average population density of places where each bank is located. Some banks are excluded due to small sample sizes. We select banks with more than 15 branches in our data set. Clearly, bank fixed effects are positively correlated with population density.[25] The robustness check for endogeneity is also done by IV. Table 3.5 shows the

[25]This positive relationship might reflect differences in the retail and wholesale banking between banks.

Table 3.4: IV Estimation Results for Agglomeration Effects on Wages

	Dependent Variable: $\log(w_{ijat})$			
Explanatory Variable	(1)	(2)	(3)	(4)
Population Density	0.025***	0.023***	0.018**	0.016*
	(0.010)	(0.007)	(0.009)	(0.009)
Years of Schooling	0.061***	0.061***	0.050***	0.050***
	(0.003)	(0.002)	(0.003)	(0.003)
Experience	0.020***	0.020***	0.019***	0.019***
	(0.002)	(0.002)	(0.002)	(0.002)
Experience Squared	−0.018***	−0.018***	−0.020***	−0.019***
	(0.005)	(0.004)	(0.005)	(0.005)
Female Dummy	−0.024**	−0.023**	−0.032***	−0.031***
	(0.011)	(0.009)	(0.010)	(0.010)
Marriage Dummy	0.074***	0.074***	0.075***	0.074***
	(0.015)	(0.010)	(0.015)	(0.015)
Occupation Dummy	Yes	Yes	Yes	Yes
State Dummy	Yes	Yes	Yes	Yes
Time Dummy	Yes	Yes	Yes	Yes
Firm Size Dummy	No	Yes	No	Yes
Bank Fixed Effect Dummy	No	No	Yes	Yes
Number of Observations	12875	12875	12875	12875
F Statistic (Weak IV)	2707.268	72725.203	2769.943	2743.229
DWH Test (p-value)	0.014	0.015	0.120	0.132

Notes: Heteroskedasticity-robust standard errors clustered by municipality are in the parentheses. Population density is expressed in logarithm. Experience squared is divided by 100. The IV of the population density is the population density in 1990. F Statistic (Weak IV) is robust Kleinbergen–Paap rk Wald F statistic for test of weak instruments. DWH Test is a Dubin–Wu–Hausman test for endogeneity. * denotes statistical significance at the 10% level, ** at the 5% level and *** at the 1% level.

OLS and IV estimation results of the following regression model:

$$\hat{\psi}_j = \delta + \eta \log(\overline{\text{Dens}}^{s}_{c(j)}) + v_j,$$

where $\hat{\psi}_j$ is a fixed effect of bank j estimated in Column (4) of Table 3.3, δ is a constant term, η is a parameter of our interest, $\overline{\text{Dens}}^{s}_{c(j)}$ is the average population density of places where bank j is located, and v_j is the error term. We can see that IV estimation also shows a statistically significant coefficient estimate of population density at the 10% level. Therefore, thus far, our results imply that workers benefit from agglomeration directly as well as indirectly through banks' branch networks, that is, from both the local and the global perspectives. The latter effects would be beneficial for workers employed in less

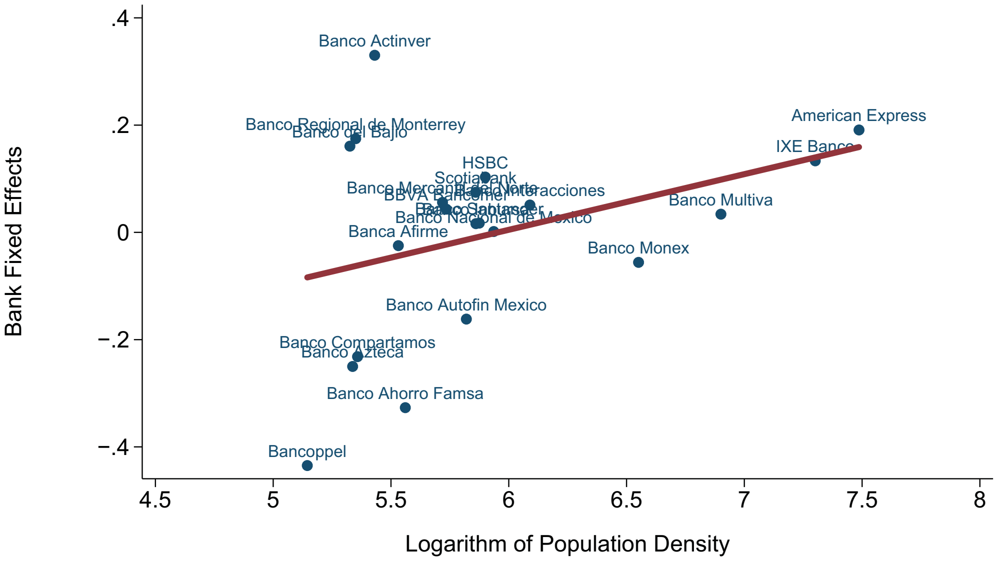


Figure 3.3: Bank Fixed Effects and Population Density

Notes: The estimated regression line is

$$\hat{\psi}_j = -\underset{(0.354)^*}{0.618} + \underset{(0.054)^*}{0.104} \log(\overline{\text{Dens}}_{c(j)}).$$

where heteroskedasticity-consistent standard errors are in the parentheses, and * denotes statistical significance at the 10% level. See Table 3.5 for detailed estimation results.

dense areas.

3.5.2 Agglomeration Effects on Wages between Banks' Branches

We further analyze whether the local agglomeration effect on wages still exists among branches of the same banks by exploiting our data set. Table 3.6 presents the IV estimation results of regression model (3.5) by bank. We select banks with sample sizes exceeding 500. The F statistic for weak identification (Weak IV) exceeds the thresholds proposed by Stock and Yogo (2005) for the maximal relative bias and maximal size. The density elasticity of wage in this case is significantly positive at the 10% level for Santander only. However, the other banks do not show significant estimates. The magnitudes of population density are also small for Azteca, Banamex, Bancomer, and HSBC. These results constitute great

Table 3.5: Relationship between Bank Fixed Effects and Population Density

	Dependent Variable: $\hat{\psi}_j$	
	OLS	IV
Explanatory Variable	(1)	(2)
Population Density	0.104*	0.101*
	(0.054)	(0.053)
Constant	−0.618*	−0.602*
	(0.354)	(0.346)
Number of Observations	21	21
Adjusted R^2	0.130	
F-Test (Weak IV)		14325.636
DWH Test (p-value)		0.454

Notes: Heteroskedasticity-consistent standard errors are in the parentheses. Population density is expressed in logarithm. Population density is average population density of banks' locating municipalities. The IV of the population density is the population density in 1990. F Statistic (Weak IV) is robust Kleinbergen–Paap rk Wald F statistic for test of weak instruments. DWH Test is a Dubin–Wu–Hausman test for endogeneity. * denotes statistical significance at the 10% level.

significance in the empirics of agglomeration economies; the effects of local agglomeration on wages received directly by workers, on average, might cease to exist between branches of the same banks.

Our interpretations of these estimation results are as follows. As shown in Figure 3.3, banks located in comparatively dense areas benefit from agglomeration. As a result, higher wages are offered to workers indirectly through banks' branch networks. Thus, workers might not benefit from agglomeration directly in the places of banks' branches. We further explore whether workers' skill levels play an important role in benefits from agglomeration in Section 3.5.3.

3.5.3 Heterogeneous Agglomeration Effects on Wages between High- and Low-Skilled Workers

Table 3.7 presents the IV estimation results of regression model (3.5) by education level. We divide the sample into high- and low-skilled workers. High- and low-skilled workers are classified as those who have studied at the university level or higher and those who have studied in a high school or at a lower level, respectively. As noted in Hering and Poncet (2010), high-skilled workers are likely to benefit from agglomeration economies, unlike their low-skilled counterparts. We further analyze the heterogeneous effects of local

Table 3.6: IV Estimation Results for Agglomeration Effects on Wages by Bank

	Dependent Variable: $\log(w_{ijat})$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Explanatory Variable	Azteca	Compartamos	Banorte	Inbursa	Banamex	Santander	Bancomer	HSBC	Scotiabank
Population Density	0.004	0.016	0.059	0.010	0.007	0.059^{*}	0.008	−0.009	0.054
	(0.024)	(0.034)	(0.043)	(0.022)	(0.018)	(0.031)	(0.019)	(0.028)	(0.053)
Years of Schooling	0.013^{**}	0.030^{**}	0.046^{***}	0.053^{***}	0.057^{***}	0.050^{***}	0.055^{***}	0.052^{***}	0.052^{***}
	(0.006)	(0.012)	(0.010)	(0.007)	(0.005)	(0.008)	(0.006)	(0.009)	(0.014)
Experience	0.016^{**}	0.005	0.009	0.021^{***}	0.024^{***}	0.019^{***}	0.025^{***}	0.020^{***}	0.021^{**}
	(0.007)	(0.008)	(0.006)	(0.006)	(0.004)	(0.005)	(0.004)	(0.007)	(0.009)
Experience Squared	-0.037^{*}	−0.005	−0.001	-0.027^{*}	-0.022^{*}	-0.035^{***}	-0.034^{***}	−0.013	−0.035
	(0.021)	(0.028)	(0.012)	(0.015)	(0.012)	(0.012)	(0.011)	(0.020)	(0.022)
Female Dummy	-0.094^{**}	−0.017	−0.033	−0.004	-0.073^{***}	−0.019	-0.049^{**}	0.005	−0.085
	(0.039)	(0.053)	(0.049)	(0.034)	(0.024)	(0.032)	(0.025)	(0.033)	(0.064)
Marriage Dummy	0.057	0.018	0.020	0.084^{**}	0.025	0.155^{***}	0.049	0.123^{***}	0.156^{***}
	(0.037)	(0.059)	(0.054)	(0.036)	(0.028)	(0.042)	(0.033)	(0.040)	(0.054)
Occupation Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Size Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	1066	503	1104	1301	2435	1134	2322	1407	524
F Statistic (Weak IV)	1419.564	2560.780	987.969	2039.736	1733.645	495.271	2200.662	1685.509	698.610
DWH Test (p-value)	0.240	0.326	0.308	0.850	0.564	0.158	0.091	0.037	0.228

Notes: Heteroskedasticity-consistent standard errors clustered by municipality are in the parentheses. Population density is expressed in logarithm. Experience squared is divided by 100. The IV of the population density is the population density in 1990. F Statistic (Weak IV) is robust Kleinbergen–Paap rk Wald F statistic for test of weak instruments. DWH Test is a Dubin–Wu–Hausman test for endogeneity. * denotes statistical significance at the 10% level, ** at the 5% level and *** at the 1% level.

Table 3.7: Heterogeneous Agglomeration Effects on Wages by Education Level

	Dependent Variable: $\log(w_{ijat})$			
	High-Skilled	Low-Skilled	High-Skilled	Low-Skilled
Explanatory Variable	(1)	(2)	(3)	(4)
Population Density	0.049***	−0.017	0.037***	−0.022
	(0.012)	(0.017)	(0.011)	(0.016)
Years of Schooling	0.058***	0.052***	0.049***	0.043***
	(0.004)	(0.008)	(0.004)	(0.008)
Experience	0.022***	0.018***	0.020***	0.015***
	(0.003)	(0.004)	(0.003)	(0.003)
Experience Squared	−0.025***	−0.011	−0.026***	−0.010
	(0.009)	(0.008)	(0.009)	(0.007)
Female Dummy	−0.020	−0.039*	−0.027**	−0.040**
	(0.013)	(0.020)	(0.012)	(0.018)
Marriage Dummy	0.077***	0.064**	0.073***	0.074***
	(0.018)	(0.025)	(0.018)	(0.025)
Occupation Dummy	Yes	Yes	Yes	Yes
State Dummy	Yes	Yes	Yes	Yes
Time Dummy	Yes	Yes	Yes	Yes
Firm Size Dummy	No	No	Yes	Yes
Bank Fixed Effect Dummy	No	No	Yes	Yes
Number of Observations	9112	3763	9112	3763
F Statistic (Weak IV)	2405.624	2634.349	2467.517	2676.069
DWH Test (p-value)	0.446	0.005	0.762	0.015

Notes: Heteroskedasticity-consistent standard errors clustered by municipality are in the parentheses. Population density is expressed in logarithm. Experience squared is divided by 100. The IV of the population density is the population density in 1990. F Statistic (Weak IV) is robust Kleinbergen–Paap rk Wald F statistic for test of weak instruments. DWH Test is a Dubin–Wu–Hausman test for endogeneity. * denotes statistical significance at the 10% level, ** at the 5% level and *** at the 1% level.

agglomeration on wages by bank.

Tables 3.8 and 3.9 present the IV estimation results of regression model (3.5) by bank and skill level, respectively. Unlike Table 3.6, Table 3.8 shows that the population densities for Banorte, Banamex, and Santander have significantly positive effects on wages at the 5% level. High-skilled workers are likely to directly enjoy benefits arising from local agglomeration even within the branch network. On the contrary, Table 3.9 shows that the population densities for Compartamos, Banamex, HSBC, and Scotiabank have significantly negative effects on wages (a positive effect exists only for Inbursa). Thus, low-skilled workers are rather likely to suffer from local agglomeration in the branch networks.

Table 3.8: Heterogeneous Agglomeration Effects on Wages for High-Skilled Workers by Bank

	Dependent Variable: $\log(w_{ijat})$								
Explanatory Variable	(1) Azteca	(2) Compartamos	(3) Banorte	(4) Inbursa	(5) Banamex	(6) Santander	(7) Bancomer	(8) HSBC	(9) Scotiabank
Population Density	0.025	0.029	0.094**	−0.013	0.052**	0.087**	0.018	0.024	−0.002
	(0.036)	(0.041)	(0.046)	(0.039)	(0.022)	(0.035)	(0.024)	(0.022)	(0.064)
Years of Schooling	0.006	0.040**	0.029	0.062***	0.060***	0.045***	0.059***	0.046***	0.064***
	(0.013)	(0.019)	(0.018)	(0.010)	(0.008)	(0.011)	(0.008)	(0.012)	(0.020)
Experience	0.024***	0.043***	0.013*	0.021**	0.028***	0.023**	0.027***	0.022***	0.024**
	(0.007)	(0.015)	(0.007)	(0.009)	(0.005)	(0.010)	(0.005)	(0.008)	(0.012)
Experience Squared	−0.081***	−0.185***	−0.008	−0.025	−0.038**	−0.057*	−0.045***	−0.022	−0.061*
	(0.023)	(0.057)	(0.016)	(0.022)	(0.016)	(0.032)	(0.015)	(0.026)	(0.035)
Female Dummy	−0.096*	0.012	−0.016	−0.009	−0.057*	−0.045	−0.049	0.010	−0.063
	(0.057)	(0.065)	(0.063)	(0.035)	(0.030)	(0.036)	(0.030)	(0.036)	(0.072)
Marriage Dummy	−0.010	−0.009	0.082	0.094**	0.025	0.156***	0.050	0.112**	0.205***
	(0.051)	(0.079)	(0.063)	(0.041)	(0.031)	(0.055)	(0.038)	(0.047)	(0.073)
Occupation Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Size Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	531	320	757	986	1643	809	1718	1144	394
F Statistic (Weak IV)	3284.652	2349.525	1011.268	1438.799	1458.763	416.603	1411.721	1482.010	616.642
DWH Test (p-value)	0.597	0.592	0.428	0.331	0.777	0.035	0.446	0.129	0.910

Notes: Heteroskedasticity-consistent standard errors clustered by municipality are in the parentheses. Population density is expressed in logarithm. Experience squared is divided by 100. The IV of the population density is the population density in 1990. F Statistic (Weak IV) is robust Kleinbergen–Paap rk Wald F statistic for test of weak instruments. DWH Test is a Dubin–Wu–Hausman test for endogeneity. * denotes statistical significance at the 10% level, ** at the 5% level and *** at the 1% level.

Table 3.9: Heterogeneous Agglomeration Effects on Wages for Low-Skilled Workers by Bank

	Dependent Variable: $\log(w_{ijat})$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Explanatory Variable	Azteca	Compartamos	Banorte	Inbursa	Banamex	Santander	Bancomer	HSBC	Scotiabank
Population Density	0.005	-0.106^{*}	-0.106	0.064^{**}	-0.064^{***}	-0.043	0.028	-0.094^{*}	-0.518^{***}
	(0.037)	(0.059)	(0.098)	(0.028)	(0.023)	(0.075)	(0.026)	(0.054)	(0.120)
Years of Schooling	0.015	0.053	-0.005	0.084^{***}	0.062^{***}	0.053^{***}	0.025	0.096^{**}	0.171^{***}
	(0.014)	(0.051)	(0.023)	(0.027)	(0.018)	(0.020)	(0.016)	(0.041)	(0.040)
Experience	0.015	0.006	0.007	0.025^{**}	0.018^{**}	0.022^{*}	0.028^{***}	0.017	-0.019
	(0.011)	(0.013)	(0.009)	(0.010)	(0.008)	(0.011)	(0.007)	(0.016)	(0.016)
Experience Squared	-0.026	0.013	-0.008	-0.034	0.001	-0.034	-0.037^{***}	0.011	0.047
	(0.032)	(0.030)	(0.017)	(0.027)	(0.020)	(0.030)	(0.013)	(0.034)	(0.032)
Female Dummy	-0.121^{**}	0.057	-0.027	-0.041	-0.076^{*}	0.074^{*}	-0.064^{*}	0.001	-0.098
	(0.050)	(0.093)	(0.075)	(0.074)	(0.041)	(0.043)	(0.038)	(0.080)	(0.092)
Marriage Dummy	0.093^{*}	0.072	0.027	0.069	0.009	0.144^{**}	0.075	0.212^{**}	0.446^{***}
	(0.049)	(0.095)	(0.073)	(0.092)	(0.055)	(0.068)	(0.054)	(0.086)	(0.095)
Occupation Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Size Dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	535	183	347	315	792	325	604	263	130
F Statistic (Weak IV)	512.707	683.411	453.098	3003.726	1246.710	282.212	2078.073	223.609	440.274
DWH Test (p-value)	0.483	0.045	0.568	0.777	0.076	0.648	0.057	0.784	0.409

Notes: Heteroskedasticity-consistent standard errors clustered by municipality are in the parentheses. Population density is expressed in logarithm. Experience squared is divided by 100. The IV of the population density is the population density in 1990. F Statistic (Weak IV) is robust Kleinbergen–Paap rk Wald F statistic for test of weak instruments. DWH Test is a Dubin–Wu–Hausman test for endogeneity. * denotes statistical significance at the 10% level, ** at the 5% level and *** at the 1% level.

The manner in which workers directly benefit from agglomeration economies in terms of individual skills is highly heterogeneous. More importantly and interestingly, high-skilled workers are likely to benefit from agglomeration, unlike low-skilled workers. To make matters worse, low-skilled workers can be negatively affected by agglomeration as a congestion effect in the branch network, wherein while the basic wage tends to be the same between workers, it is likely to be paid based on individual skills through the commission system and performance-related payment. Fierce competition in agglomerated areas might give rise to negative externalities.

3.6 Concluding Remarks

In this paper, we examined whether workers really earn higher wages in agglomerated areas. Previous studies on the empirics of agglomeration economies have shown that higher wages tend to be paid in denser areas. However, agglomeration itself affects both workers and firms simultaneously, and as a result, the route of the agglomeration effects on wages received directly by workers was still unclear, notably when firms have branch networks. Our objective was to study two possibilities: (1) wage is higher in denser areas even across branches within the same firms, suggesting that workers directly benefit from agglomeration, from a local perspective, and (2) branches of the same firms are, on average, located in denser areas, and thus, their overall wages, on average, are higher throughout the branch network, suggesting that workers indirectly benefit from branch networks, from a global perspective. Furthermore, we analyzed heterogeneous effects of agglomeration on wages between high- and low-skilled workers. Using Mexican workers' micro-data sets, which also includes their firm-level information, we analyzed agglomeration effects on wages by focusing on the banking sector as it has a number of branches spread across regions throughout the country.

We found that population density has a significantly positive effect on wages in the pooled data across banks. These results are consistent with the findings of previous studies. Furthermore, we found that banks locating branches in denser areas tend to pay workers better, suggesting that workers indirectly benefit from agglomeration through banks' branch networks. Thus, branch networks would provide additional agglomeration effects, especially for workers employed in less dense areas. More importantly and interestingly, we found that high-skilled workers are likely to benefit from agglomeration, unlike their low-skilled counterparts. Within the banks' branch network, the basic wage tends to be the same between workers in terms of occupation, but it is likely to be paid based on individual skills through the commission system and performance-related payment. Moreover, low-skilled workers can be negatively affected by agglomeration as a congestion effect, and fierce competition in denser areas might give rise to negative externalities. Our estimation results suggest that all types of workers cannot necessarily enjoy direct agglomeration benefits from a local perspective.

A limitation of our paper is that the dynamic aspect of the agglomeration economy has not been examined. Even if there is no static agglomeration effect on wages, more valuable experience in big cities might make the wage profile steeper. Another limitation is that we did not theoretically explore how the wage system is determined within branch networks in Section 3.3. It should be worthwhile to construct a more rigorous theoretical model to uncover the underlying mechanism of branch locations and wage distribution in further studies.

Bibliography

Acemoglu, Daron and Joshua Angrist (2001) "How large are human-capital externalities? Evidence from compulsory-schooling laws," in Bernanke, Ben S. and Kenneth Rogoff eds. *NBER Macroeconomics Annual 2000* Vol. 15: MIT Press, Chap. 1, pp. 9–74.

Albert, James H. and Siddhartha Chib (1993) "Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts," *Journal of Business & Economic Statistics* 11(1), pp. 1–15.

Anselin, Luc (1988) *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer Academic Press.

——— (2003) "Spatial externalities, spatial multipliers, and spatial econometrics," *International Regional Science Review* 26(2), pp. 153–166.

——— (2006) "Spatial econometrics," in Mills, Terence C. and Kerry Patterson eds. *Palgrave Handbook of Econometrics* Vol. 1 Econometric Theory, Basingstoke: Palgrave Macmillan, Chap. 26, pp. 901–969.

——— (2010) "Thirty years of spatial econometrics," *Papers in Regional Science* 89(1), pp. 3–25.

Anselin, Luc and Anil Kumar Bera (1998) "Spatial dependence in linear regression models with an introduction to spatial econometrics," in Ullah, Aman and David E.A. Giles eds. *Handbook of Applied Economic Statistics*, New York: Marcel Dekkar, Chap. 7, pp. 237–289.

Anselin, Luc, Anil K. Bera, Raymond Florax, and Mann J. Yoon (1996) "Simple diagnostic tests for spatial dependence," *Regional Science and Urban Economics* 26(1), pp. 77–104.

Baldwin, Richard, Ricard Forslid, Philippe Martin, Gianmarco Ottaviano, and Frederic Robert-Nicoud eds. (2003) *Economic Geography and Public Policy*, New Jersey: Princeton University Press.

Banco de México (2011) "Financial System Report: September 2011," Banco de México. `http://www.banxico.org.mx/` (Last Checked: November 16, 2013).

Blanchflower, David G. and Andrew J. Oswald (1994) *The Wage Curve*, Cambridge, MA: MIT Press.

Brown, Charles and James Medoff (1989) "The employer size-wage effect," *Journal of Political Economy* 97(5), pp. 1027–1059.

Burns, Arthur F. and Wesley C. Mitchell (1946) *Measuring Business Cycles*, New York: National Bureau of Economic Research.

Card, David (1995) "The wage curve: A review," *Journal of Economic Literature* 33(2), pp. 785–799.

Chiang, Shu-hen (2009) "The effects of industrial diversification on regional unemployment in Taiwan: Is the portfolio theory applicable?" *Annals of Regional Science* 4(43), pp. 947–962.

Chib, Siddhartha (1995) "Marginal likelihood from the Gibbs output," *Journal of the American Statistical Association* 90(432), pp. 1313–1321.

——— (1996) "Calculating posterior distributions and modal estimates in Markov mixture models," *Journal of Econometrics* 75(1), pp. 79–97.

——— (2001) "Markov chain Monte Carlo methods: Computation and inference," in Heckman, James J. and Edward E. Leamer eds. *Handbook of Econometrics* Vol. 5, Amsterdam: Elsevier, Chap. 57, pp. 3569–3649.

Chib, Siddhartha and Ivan Jeliazkov (2001) "Marginal likelihood from the Metropolis-Hastings output," *Journal of the American Statistical Association* 96(453), pp. 270–281.

Chiquiar, Daniel (2008) "Globalization, regional wage differentials and the Stolper-Samuelson theorem: Evidence from Mexico," *Journal of International Economics* 74(1), pp. 70–93.

Chiquiar, Daniel and Manuel Ramos-Francia (2005) "Trade and business-cycle synchronization: Evidence from Mexican and U.S. manufacturing industries," *North American Journal of Economics and Finance* 16(2), pp. 187–216.

Ciccone, Antonio and Robert E. Hall (1996) "Productivity and the density of economic activity," *American Economic Review* 86(1), pp. 54–70.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon (2008a) "Spatial wage disparities: Sorting matters!," *Journal of Urban Economics* 63(2), pp. 723–742.

Combes, Pierre-Philippe, Thierry Mayer, and Jacques-François Thisse (2008b) *Economic Geography: The Integration of Regions and Nations*, New Jersey: Princeton University Press.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux (2010) "Estimating agglomeration economies with history, geology, and worker effects," in

Glaeser, Edward L. ed. *Agglomeration Economics*, Chicago: University of Chicago Press, Chap. 1, pp. 15–66.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon (2011) "The identification of agglomeration economies," *Journal of Economic Geography* 11(2), pp. 253–266.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux (2012a) "The productivity advantages of large cities: Distinguishing agglomeration from firm selection," *Econometrica* 80(6), pp. 2543–2594.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux (2012b) "Sorting and local wage and skill distributions in France," *Regional Science and Urban Economics* 42(6), pp. 913–930.

de la Roca, Jorge and Diego Puga (2012) "Learning by working in big cities," CEPR Discussion Papers No. 9243.

Delajara, Marcelo (2011) "Comovimiento regional del empleo durante el ciclo económico en México," *El Trimestre Económico* LXXVIII(3), pp. 613–642.

Doornik, Jurgen A. and Marius Ooms (2006) *Introduction to Ox*, London: Timberlake Consultants Press. `http://www.doornik.com/`.

Duranton, Gilles and Diego Puga (2004) "Micro-foundations of urban agglomeration economies," in Henderson, J. Vernon and Jacques-François Thisse eds. *Handbook of Regional and Urban Economics* Vol. 4, Amsterdam: Elsevier, Chap. 48, pp. 2063–2117.

Epifani, Paolo and Gino A. Gancia (2005) "Trade, migration and regional unemployment," *Regional Science and Urban Economics* 35(6), pp. 625–644.

Erquizio-Espinal, Alfredo (2010) "Gran Recesión 2008–2009 en EE.UU. y México: Un enfoque regional," *Paradigma económico* 2(2), pp. 5–40.

Francis, John (2009) "Agglomeration, job flows and unemployment," *Annals of Regional Science* 43(1), pp. 181–198.

Fujita, Masahisa (2010) "The evolution of spatial economics: From Thünen to the new economic geography," *Japanese Economic Review* 61(1), pp. 1–32.

Fujita, Masahisa and Jacques-François Thisse (2013) *Economics of Agglomeration: Cities, Industrial Location, and Globalization*, New York: Cambridge University Press, 2nd edition.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables (1999) *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge, MA: MIT Press.

Glaeser, Edward L. and David C. Maré (2001) "Cities and skills," *Journal of Labor Economics* 19(2), pp. 316–342.

Glaeser, Edward L. and Matthew G. Resseger (2010) "The complementarity between cities and skills," *Journal of Regional Science* 50(1), pp. 221–244.

Gould, E. D. (2007) "Cities, workers, and wages: A structural analysis of the urban wage premium," *Review of Economic Studies* 74(2), pp. 477–506.

Hamilton, James D. (1989) "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica* 57(2), pp. 357–384.

——— (2008) "Regime switching models," in Durlauf, Steven N. and Lawrence E. Blume eds. *The New Palgrave Dictionary of Economics* Vol. 7, New York: Palgrave Macmillan, 2nd edition, pp. 53–57.

Hamilton, James D. and Michael T. Owyang (2012) "The propagation of regional recessions," *Review of Economics and Statistics* 94(4), pp. 935–947.

Hanson, Gordon H. (1998) "North American economic integration and industry location," *Oxford Review of Economic Policy* 14(2), pp. 30–44.

——— (2005) "Market potential, increasing returns and geographic concentration," *Journal of International Economics* 67(1), pp. 1–24.

Harris, John R. and Michael P. Todaro (1970) "Migration, unemployment and development: A two-sector analysis," *American Economic Review* 60(1), pp. 126–142.

Helpman, Elhanan (1998) "The size of regions," in Pines, David, Efraim Sadka, and Itzhak Zilcha eds. *Topics in Public Economics: Theoretical and Applied Analysis*, Cambridge: Cambridge University Press, Chap. 2, pp. 33–54.

Helpman, Elhanan and Oleg Itskhoki (2010) "Labour market rigidities, trade and unemployment," *Review of Economic Studies* 77(3), pp. 1100-1137.

Hering, Laura and Sandra Poncet (2010) "Market access impact on individual wages: Evidence from China," *Review of Economics and Statistics* 92(1), pp. 145–159.

Hirata, Hideaki, M. Ayhan Kose, and Christopher Otrok (2013) "Regionalization vs. globalization," IMF Working Papers 13/19.

Holloway, Garth, Bhavani Shankar, and Sanzidur Rahman (2002) "Bayesian spatial probit estimation: A primer and an application to HYV rice adoption," *Agricultural Economics* 27(3), pp. 383–402.

Hynninen, Sanna-Mari and Jukka Lahtonen (2007) "Does population density matter in the process of matching heterogeneous job seekers and vacancies?" *Empirica* 34(5), pp. 397–410.

Izraeli, Oded and Kevin J. Murphy (2003) "The effect of industrial diversity on state unemployment rate and per capita income," *Annals of Regional Science* 37(1), pp. 1–14.

Kano, Shigeki and Makoto Ohta (2005) "Estimating a matching function and regional matching efficiencies: Japanese panel data for 1973–1999," *Japan and the World Economy*

17(1), pp. 25–41.

Kelejian, Harry H. and Ingmar R. Prucha (1998) “A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances,” *Journal of Real Estate Finance and Economics* 17(1), pp. 99–121.

Kim, Chang-Jin and Charles R. Nelson (1998) “Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching,” *Review of Economics and Statistics* 80(2), pp. 188–201.

——— (1999a) “Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle,” *Review of Economics and Statistics* 81(4), pp. 608–616.

——— (1999b) *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, Cambridge, MA: MIT Press.

Kose, M. Ayhan, Christopher Otrok, and Charles H. Whiteman (2003) “International business cycles: World, region, and country-specific factors,” *American Economic Review* 93(4), pp. 1216–1239.

Krugman, Paul (1991a) *Geography and Trade*, Cambridge, MA: MIT Press.

——— (1991b) “Increasing returns and economic geography,” *Journal of Political Economy* 99(3), pp. 483–499.

——— (1995) *Development, Geography, and Economic Theory*, Cambridge, MA: MIT Press.

Krugman, Paul and Raul Livas-Elizondo (1996) “Trade policy and the Third World metropolis,” *Journal of Development Economics* 49(1), pp. 137–150.

LeSage, James P. and R. Kelley Pace (2009) *Introduction to Spatial Econometrics*, Boca Raton: CRC Press.

Lubrano, Mike (1998) “Foreign investment in the financial sector of Mexico,” *United States–Mexico Law Journal* 6, pp. 81–87.

Marshall, Alfred (1890) *Principles of Economics*, London: Macmillan.

Matano, Alessia and Paolo Naticchioni (2012) “Wage distribution and the spatial sorting of workers,” *Journal of Economic Geography* 12(2), pp. 379–408.

Mejía-Reyes, Pablo and Jeanett Campos-Chávez (2011) “Are the Mexican states and the United States business cycles synchronized? Evidence from the manufacturing production,” *economía mexicana NUEVA ÉPOCA* XX(1), pp. 79–112.

Mejía-Reyes, Pablo and Alfredo Erquizio-Espinal (2012) *Expansiones y Recesiones en los Estados de México*, Hermosillo: Pearson Educación–Universidad de Sonora.

Minushkin, Susan (2000) “*Banqueros* and *bolseros*: Structural change and financial market liberalisation in Mexico," *Journal of Latin American Studies* 34(4), pp. 915–944.

Mion, Giordano and Paolo Naticchioni (2009) "The spatial sorting and matching of skills and firms," *Canadian Journal of Economics* 42(1), pp. 28–55.

Moretti, Enrico (2004) "Human capital externalities in cities," in Henderson, J. Vernon and Jacques-François Thisse eds. *Handbook of Regional and Urban Economics* Vol. 4, Amsterdam: Elsevier, Chap. 51, pp. 2243–2291.

Ohtsuka, Yoshihiro (2010) "Estimation of regional bussiness cycle in Japan with Markov switching spatial autoregressive-AR model," *Nihon Tōkei Gakkai shi* 40(2), pp. 89–109. (in Japanese).

Ord, Keith (1975) "Estimation methods for models of spatial interaction," *Journal of the American Statistical Association* 70(349), pp. 120–126.

Owyang, Michael T., Jeremy Piger, and Howard J. Wall (2005) "Business cycle phases in U.S. states," *Review of Economics and Statistics* 87(4), pp. 604–616.

Owyang, Michael T., Jeremy M. Piger, Howard J. Wall, and Christopher H. Wheeler (2008) "The Economic performance of cities: A Markov-switching approach," *Journal of Urban Economics* 64(3), pp. 538–550.

Owyang, Michael T., David E. Rapach, and Howard J. Wall (2009) "States and the business cycle," *Journal of Urban Economics* 65(2), pp. 181–194.

Petrongolo, Barbara and Christopher A. Pissarides (2001) "Looking into the black box: A survey of the matching function," *Journal of Economic Literature* 39(2), pp. 390–431.

Pflüger, Michael and Takatoshi Tabuchi (2010) "The size of regions with land use for production," *Regional Science and Urban Economics* 40(6), pp. 481–489.

Pissarides, Christopher A. (2000) *Equilibrium Unemployment Theory*, Cambridge, MA: MIT Press, 2nd edition.

Puga, Diego (2010) "The magnitude and causes of agglomeration economies," *Journal of Regional Science* 50(1), pp. 203–219.

Rauch, James E. (1993) "Productivity gains from geographic concentration of human capital: Evidence from the cities," *Journal of Urban Economics* 34(3), pp. 380–400.

Redding, Stephen and Anthony J. Venables (2004) "Economic geography and international inequality," *Journal of International Economics* 62(1), pp. 53–82.

Rogerson, Richard, Robert Shimer, and Randall Wright (2005) "Search-theoretic models of the labor market: A survey," *Journal of Economic Literature* 43(4), pp. 959–988.

Rosenthal, Stuart S. and William C. Strange (2001) "The determinants of agglomeration," *Journal of Urban Economics* 50(2), pp. 191–229.

——— (2004) "Evidence on the nature and sources of agglomeration economies," in Henderson, J. Vernon and Jacques-François Thisse eds. *Handbook of Regional and Urban Economics* Vol. 4, Amsterdam: Elsevier, Chap. 49, pp. 2119–2171.

Sato, Yasuhiro (2000) "Search theory and the wage curve," *Economics Letters* 66(1), pp. 93–98.

Sigmond, Karen B. (2011) "Banking regulation in Mexico: Lessons from financial crises," *Mexican Law Review* IV(1), pp. 3–33.

Simon, Curtis J. (1988) "Frictional unemployment and the role of industrial diversity," *Quarterly Journal of Economics* 103(4), pp. 715–728.

Sosa, Sebastian (2008) "External shocks and business cycle fluctuations in Mexico: How important are U.S. factors?," IMF Working Papers 08/100.

Stock, James H. and Mark W. Watson (1989) "New indexes of coincident and leading economic indicators," in Blanchard, Olivier Jean and Stanley Fischer eds. *NBER Macroeconomics Annual 1989* Vol. 4, Cambridge, MA: MIT Press, Chap. 7, pp. 351–409.

Stock, James H. and Motohiro Yogo (2005) "Testing for weak instruments in linear IV regression," in Andrews, Donald W. K. and James H. Stock eds. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge: Cambridge University Press, Chap. 5, pp. 80–108.

Stole, Lars A. and Jeffrey Zwiebel (1996a) "Intra-firm bargaining under non-binding contracts," *Review of Economic Studies* 63(3), pp. 375–410.

——— (1996b) "Organizational design and technology choice under intrafirm bargaining," *American Economic Review* 86(1), pp. 195–222.

Suedekum, Jens (2005) "Increasing returns and spatial unemployment disparities," *Papers in Regional Science* 84(2), pp. 159–181.

Torres, Alberto and Oscar Vela (2003) "Trade integration and synchronization between the business cycles of Mexico and the United States," *North American Journal of Economics and Finance* 14(3), pp. 319–342.

Turrent, Eduardo (2011) "A brief summary of banking in Mexico," Banco de México. `http://www.banxico.org.mx/` (Last Checked: November 16, 2013).

Vincenty, Thaddeus (1975) "Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations," *Survey Review* 23(176), pp. 88–93.

vom Berge, Philipp (forthcoming) "Search unemployment and new economic geography," *Annals of Regional Science* Online First.

World Bank (2009) *World Development Report 2009: Reshaping Economic Geography*, Washington, D.C.: World Bank.

Zierahn, Ulrich Theodor (forthcoming) "Agglomeration, congestion, and regional unemployment disparities," *Annals of Regional Science* Online First.

Ziesemer, Thomas (2005) "Monopolistic competition and search unemployment," *Metroeconomica* 56(3), pp. 334–359.