Title	Bayesian analysis on spatial econometric models
Author(s)	高, 平
Degree Grantor	北海道大学
Degree Name	博士(経済学)
Dissertation Number	甲第14028号
Issue Date	2020-03-25
DOI	https://doi.org/10.14943/doctoral.k14028
Doc URL	https://hdl.handle.net/2115/94708
Type	doctoral thesis
File Information	Gao_Ping.pdf

HOKKAIDO UNIVERSITY

Bayesian analysis on spatial econometric models

A DISSERTATION PRESENTED
BY
PING GAO
TO
GRADUATE SCHOOL OF ECONOMICS AND BUSINESS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF ECONOMICS
IN THE FIELD OF
ECONOMETRICS AND ECONOMIC STATISTICS

HOKKAIDO UNIVERSITY
SAPPORO, HOKKAIDO
MARCH 2020

Thesis advisor: Professor Hikaru Hasegawa Ping Gao

Bayesian analysis on spatial econometric models

Abstract

For discrete dependent variables, the most commonly used models are logit models and probit models. Due to the widespread existence of spatial effects, researchers have applied spatial econometric models to discrete response variables. When spatial dimension is considered as an independent factor, a logistic structured additive regression (STAR) model may be used to explore the spatial effects. When the discrete dependent variables are ordinal and correlated, a multivariate ordered probit model can be applied to capture the interactions of ordinal responses. Currently, there are few researchers considering the spatial effects on multivariate ordinal responses, and the spatial dependencies have been confirmed to exist in many ordinal dependent variables. Therefore, a new multivariate probit model is needed to capture the interactions and the spatial dependencies of ordinal responses. Besides spatial dependencies, temporal dependencies also widely exist in ordinal dependent variables. Therefore, a more comprehensive model needs to be proposed to capture temporal and spatial dependencies simultaneously for the multivariate ordinal responses.

The traditional estimation method of discrete dependent variable models is the maximum likelihood (ML) method, but Beron and Vijverberg (2004) found it difficult to estimate. Albert and Chib (1993) introduced latent continuous variables into probit models and proposed the Bayesian inference for the estimation. Following Albert and Chib (1993), Bayesian inference have been widely applied to discrete response variable models (Chen and Dey, 2000; Hasegawa, 2010; Jeliazkov et al., 2008; LeSage, 1999). LeSage (1999) introduced Albert and Chib (1993)'s method in spatial econometrics and proposed the Markov Chain Monte Carlo (MCMC) method to estimate the spatial models. Because the Bayesian method can be applied to complex models that cannot or is difficult be estimated by traditional methods, it is gradually recognized and used in spatial econometrics.

In Chapter 2, the STAR models are applied to the chronic disease data in elderly Chinese people. Chronic diseases have become important factors affecting the health of elderly Chinese people. Because the prevalence of chronic disease varies among the provinces, it is necessary to understand the spatial effects on these diseases, as well as their relationships with potential risk factors. Because the structured additive regression (STAR) model can combine spatial effects, nonlinear effects of continuous factors, and the linear or fixed effects into a single model, it is applied to the data obtained from the

2000, 2006, and 2010 Chinese Urban and Rural Elderly Population Surveys. R2BayesX package is used to conduct the Bayesian analysis. The empirical results show that the province is a critical influencing factor, and the highest spatial effect usually appeared in two types of provinces: economically developed provinces, and economically backward provinces with complex terrain.

Spatial influences, besides being treated as an independent factor, have also been confirmed to exist in many dependent variables. Since many ordinal discrete responses are correlated and spatially dependent, it is better to study them together, rather than separately. Thus, a model that captures the interactions and the spatial relationships of multivariate ordinal outcomes in needed. Following Smith and LeSage (2004) and Jeliazkov et al. (2008), a multivariate spatial ordered probit (MSOP) model is proposed to address this need in Chapter 3. In applying this model, the parameters are calculated using the Bayesian inference based on the MCMC sampling. The validity and accuracy of the MSOP model is verified by simulated datasets, and the model performs very well with the simulated data. In addition, this study illustrates the model by applying it to two response variables, self-rated health and life satisfaction of elderly people in six provinces in East China. The empirical results show that the spatial dependencies are indispensable on the response variables.

Besides spatial dependencies, temporal dependencies also widely exist. In Chapter 4, the MSOP model is extended to the dynamic multivariate spatial ordered probit (DMSOP) model, wherein the spatial dependencies are explained using an additive error specification, as in Smith and LeSage (2004). The DMSOP model is the first attempt to capture temporal and spatial dependencies simultaneously for the multivariate ordinal responses. The parameters are still calculated using the Bayesian inference based on the MCMC sampling. A simulation study is implemented to demonstrate the validity and accuracy of the model. The DMSOP model performs effectively with the simulated data. It successfully explains the temporal and spatial dependencies, and the interaction of multivariate ordinal response variables as well as the effects of the explanatory variables. Additionally, the study applied the DMSOP model to the survey data from China Family Panel Studies (CFPS). The spatial coefficient and the temporal coefficient are considerable, and thus the spatial and temporal dependencies in the empirical study are critical.

Key Words: Bayesian inference, dynamic, Markov chain Monte Carlo (MCMC), multivariate response variables, ordered probit model, spatial dependency.

Acknowledgments

I WOULD LIKE TO express my sincere gratitude to my advisor Prof. Hikaru Hasegawa for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. At the same time, I would also like to thank my advisor Prof. Hikaru Hasegawa for his fatherly care in my life. Let me feel the warmth in a foreign country.

Besides my advisor, I would like to thank the rest of my deputy advisors: Prof. Shingo Takagi, and Prof. Yoshihide Kakizawa, for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

In addition, I would like to thank the Hokkaido University President's Fellowship for its financial support.

My sincere thanks also goes to my fellow labmate Lei Shi, who has given me a lot of help in study and life. She gave me a lot of convenience during the doctoral course. As an international student in Japan, it is really lucky to be able to get help from others. During the doctoral course, we did the research together and learned from each other.

Last but not the least, I would like to thank my family: my parents, my husband and my daughter for supporting me spiritually throughout writing this thesis and my life in general.

Contents

Listing of figures

List of Tables

1

Introduction

In econometrics, numerous dependent variables are continuous variables, but some dependent variables are discrete, binary, or polychotomous. For example, an analyst may wish to model the presence or absence of a disease, a choice of travel mode (walking, bus, rail, car, etc.), or shop location choice (city center, suburb, residential area, etc.). Sometimes, these discrete response variables are ordered, such as shop scale (small, medium, and large), grade (A, B, C, D, and F), or health (very unhealthy, unhealthy, general, healthy, and very healthy). When dealing with these discrete response variables or discrete ordinal response variables, methods for continuous dependent variables cannot be applied. Therefore, logit models and probit models are appeared to deal with discrete response variables. In the case of discrete ordinal response variables, the ordered logit (OL) models or ordered probit (OP) models can be applied.

The above choices, scales, or situations are sometimes affected by geographic location. For example, shopping malls are more likely be located in city center, while factories are often located in suburbs. Additionally, these choices may affected by neighbors or neighboring areas. For example, the economic development level of the surrounding provinces is high, and local economic development will also be driven. It can be obtained that the influence of spatial factors cannot be ignored. Therefore, the spatial framework is needed to deal with these types of discrete response variables.

The traditional estimation method of logit or probit models is the maximum likelihood (ML) method. Some researchers examined the ML method to obtain the parameter estimates (Anselin, 2013; Arbia, 2006; LeSage, 1999; Wang et al., 2013), but found it difficult to estimate (Beron and Vijverberg, 2004). In addition, the ML method is based on the asymptotic theory. This method is reliable when the sample size is large enough, while when the sample size is small, the method is either underestimated or overestimated. Gelman et al. (2013) give an example to calculate the mortality rate of kidney cancer in a country with only 1,000 people. They pointed out that when the sample size is very small, the mortality rate will far from its true value if someone accidentally dies from kidney cancer.

In contrast, we can apply another method to estimate: Bayesian inference. Bayesian inference can combine prior information and sample information together to calculate the parameters. Albert and Chib (1993) introduced latent continuous variables into probit models and proposed the Bayesian inference for the estimation. Following Albert and Chib (1993), Bayesian inference and latent continuous variables have been widely applied to discrete response variable models (Chen and Dey, 2000; Hasegawa, 2010; Jeliazkov et al., 2008; LeSage, 1999). LeSage (1999) introduced Albert and Chib (1993)'s method in spatial econometrics and proposed the Markov Chain Monte Carlo (MCMC) method to estimate the spatial models.

First, the background of a univariate spatial model called structured additive regression (STAR) model will be introduced. In discrete response variable models, the explanatory variables can be discrete, or continuous. In addition to these influencing factors, the spatial dimension is increasingly being considered as an independent factor, which can be examined using geostatistical methods. Kriging is a common geostatistical method that produces a map of a quantity of interest over a geographical region. In addition, as an extension to kriging, universal kriging takes into account the linearity of the covariate (Cressie, 2015, pp.151–172). However, when the effect of the covariate is nonlinear, the universal kriging method is not appropriate. In this case, a geoadditive model may be a better choice. The geoadditive model was introduced by Kammann and Wand (2003), and accounts for non-linear covariate effects under the assumption of additivity. Today, many researchers apply this model in their studies (Basile et al., 2013; Geniaux and Napoléone, 2008; Sauleau et al., 2007; Wand et al., 2011).

Another powerful model used in spatial analyses is the generalized additive model (GAM). The GAM is suitable for modeling nonlinear effects of continuous covariates in regression models with non-Gaussian responses. Furthermore, STAR models extend GAM models by including spatial effects, the nonlinear effects of continuous factors, and linear or fixed effects in one model (Kneib, 2006). STAR models include generalized linear models and generalized additive models as special cases, but also allow for a wider class of effects, such as geographical or spatial-temporal effects (Fahrmeir et al., 2004, 2013; Umlauf et al., 2015).

Next, the reasons for the establishment of a spatial model that deal with multivariate ordinal dependent variables will be explained. Spatial influences, besides being treated as an independent factor, have also been confirmed to exist in many dependent variables, such as the Japanese business cycle (Kakamu and Wago, 2007), the land use change (Verburg et al., 2004; Zhou and Kockelman, 2008), high-yielding variety rice adoption (Holloway et al., 2002), and even the presidential election (Smith and LeSage, 2004).

LeSage (1999) explained five specific spatial models: the first-order spatial autoregressive model, the mixed autoregressive model, the spatial autoregressive error model, the spatial Durbin model, and the general spatial model which includes both the spatial lag term and a spatially correlated error structure. However, all these models address continuous responses. McMillen (1992) first introduced the spatial probit models to address the need of discrete responses. Smith and LeSage (2004) extended the model by including an additive error specification. Furthermore, Kakamu and Wago (2007) extended that model to the spatial panel probit model by including panel data, and they applied it to the business cycle in Japan. All the response variables in the above spatial probit models are binary, and when ordinal, these models can be extended to spatial ordered probit (SOP) model.

All the above studies are univariate models. In the case of multivariate outcomes, Chen and Dey (2000), Chib and Greenberg (1998), Hasegawa (2010), and Jeliazkov et al. (2008) illustrated the multivariate probit models with binary and ordinal outcomes. Using such a model, Hasegawa (2010) analyzed tourists' satisfaction, and Lauer (2003) analyzed the impact of family background, cohort, and gender on educational attainment in France and Germany.

Currently, there are few researchers considering multivariate outcomes in spatial econometrics. Except that Baltagi et al. (2016) proposed a panel bivariate spatial probit model. Since many phenomena are correlated and spatially dependent, it is better to study them together, rather than separately. Thus, a model that captures the interactions of multivariate outcomes and spatial dependency is necessary.

Following the probit model with spatial dependencies in Smith and LeSage (2004) and models for multivariate ordinal outcomes in Jeliazkov et al. (2008), this study proposes a multivariate spatial ordered probit (MSOP) model to meet the above requirements. The MSOP model is the first to analyze the spatial dependencies of multivariate ordered response variables.

Next, the intention of a dynamic multivariate spatial ordered probit (DMSOP) model will be demonstrated. Besides spatial dependencies, temporal dependencies also widely exist. The dynamic or panel ordered OP model is needed to capture this characteristic. Hasegawa (2009) proposed a Bayesian dynamic panel OP model and applied it to subjective well-being. In spatial econometrics, Kakamu and Wago (2007) introduced a Bayesian panel spatial probit model, and then, Kakamu et al. (2010) used the above model to estimate the regional business cycle in Japan. In the case of ordered responses, Wang and Kockelman (2009c) extended Smith and LeSage (2004)'s model to the dynamic spatial ordered probit (DSOP) model.

Following the dynamic OP model with spatial dependencies in Wang and Kockelman (2009c) and models for multivariate ordinal outcomes in Jeliazkov et al. (2008), this study extends the MSOP model to the dynamic multivariate spatial ordered probit (DMSOP) model, wherein the spatial dependencies are explained as in Smith and LeSage (2004). The DMSOP model is the first attempt to capture temporal and spatial dependencies simultaneously for the multivariate ordinal responses.

The rest of this study is structured as follows. Chapter 2 use the STAR model to analyze chronic disease. In applying this model, a `R2BayesX` package is used to fit the STAR models. Chapter 3 proposes a MSOP model to capture the interactions of multivariate ordinal responses and spatial dependencies of dependent variables. The validity and accuracy of the MSOP model is verified by the simulation study. In addition, the MSOP model is applied to two subjective ordered discrete variables, self-rated health and life satisfaction of elderly people in six provinces in East China. Chapter 4 extends the MSOP model to the DMSOP model, wherein temporal and spatial dependencies can be captured simultaneously for the multivariate ordinal responses. A simulation study is implemented to demonstrate the validity and accuracy of the DMSOP model. Furthermore, the DMSOP model is illustrated using the panel data of adults in ten provinces in China. Chapter 5 presents the concluding remarks and future researches.

2

Bayesian analysis of a STAR model*

In this chapter, a structured additive regression (STAR) model is applied to the chronic disease data of elderly people in China. The medical definition of a chronic disease is a disease that persists for a long time. For example, the U.S. National Center for Health Statistics defines a chronic disease as one that lasts for three months or more. In China, more than 70% of elderly people struggle with a chronic disease. In addition, the

*This chapter is based on Gao and Hasegawa (2018).

majority of the elderly suffer from multiple chronic diseases (Thorpe and Howard, 2006; Vogeli et al., 2007; Wolff et al., 2002). The prevalence of chronic diseases is related to many potential risk factors, such as age, lifestyle habits (smoking, drinking), a lack of exercise, and so on.

With the rising prevalence of chronic diseases (Freedman and Martin, 2000; Thorpe and Howard, 2006) and the large elderly population (the population aged 60 and over in China was $2,308,600$ in 2016), the number of elderly Chinese people suffering from chronic diseases is very high. Because chronic diseases are essentially permanent, they introduce a heavy economic burden to families and society. In China, the most common cause of death is chronic diseases, rather than infectious diseases (He et al., 2005). For example, chronic disease-induced deaths accounted for 71.88% of all deaths among residents in Kunming (Yunnan provincial capital) in the period 2007–2010 (Li et al., 2012). The high number of elderly people with severe chronic conditions places a significant burden on medical care. Thus, the Chinese government faces enormous challenges in terms of medical investment. Numerous studies have examined chronic diseases, with many focusing on the factors influencing such diseases.

Fillingim et al. (2009) studied samples from different countries (China, France, Sweden, United States, etc.), and found that the prevalence of the most common forms of pain caused by chronic diseases is higher in women than in men. Furthermore, Zhen (2010) used data on Gansu province (in China), finding that the resistance of females to chronic disease pain is poor and that females' thresholds for discomfort are lower than those of men. As a result, women are more likely to visit a doctor and, thus, are more likely to be diagnosed with a chronic disease. Thus, gender is hypothesized to have a great influence on the reported prevalence of chronic diseases and of specific chronic diseases (e.g., hypertension and heart disease) in elderly Chinese people, and that elderly females are more likely to have chronic diseases and specific chronic diseases than are elderly males.

This chapter uses reported prevalence rather than prevalence, because the prevalence is not the true prevalence. Chronic diseases are usually non-fatal diseases and persist for a long time. Many elderly people may have chronic diseases, but may not be aware of this, in which case, they will report not having a disease, even though they do. Thus, the study can only obtain the reported prevalence.

Woolf et al. (2015) noted that people with a high economic status are more conscious of self-care, and that such individuals are more likely to be diagnosed with chronic diseases. Zhen (2010) found the reported prevalence of chronic diseases is significantly affected by access to health care. Income and medical security are greater among the urban elderly than among the rural elderly. Therefore, census register is hypothesized to be a critical factor related to chronic diseases and to specific chronic diseases (e.g., hypertension and heart disease), and that the reported prevalence of chronic diseases and specific chronic diseases is higher among the urban elderly than among the rural elderly.

Chen (2005), Ye (2013), and Zhao et al. (2015) found that marital status also has an affect on chronic diseases in China. Furthermore, Ye (2013) analyzed data on Jilin province, and found that those who are single have the lowest prevalence of chronic diseases, while the prevalence among divorced/widowed persons is the highest. Thus, the reported prevalence is hypothesized to be higher among divorced and widowed elderly people than among other types of elderly people.

The WHO (2005) reported that chronic diseases among the elderly are predominantly attributable to unhealthy habits during youth, such as excessive smoking and drinking. In addition, using data on China, Chen (2005), Jiao et al. (2002), and Zhao et al.

(2015) found similar results, namely, that cigarette smoking and alcohol usage are risk factors for chronic diseases. Therefore, (cigarette) smoking and drinking (alcohol) are hypothesized to have a considerable influence on the reported prevalence of chronic diseases and specific chronic diseases in China.

Numerous studies have confirmed that age has a considerable effect on chronic diseases in China (Chen, 2005; Jiao et al., 2002), with the prevalence increasing with age. Furthermore, Jiao et al. (2002) and Ye (2013) pointed out that education and exercise have non-negligible effects on chronic diseases, with the prevalence increasing for lower levels of education and less exercise. In this chapter, exercise is divided into two categories: sports activities and cultural activities.

Based on the above studies, I hypothesize that the prevalence increases in older people who get less exercise. However, people with higher levels of education are hypothesized more likely to report having chronic diseases because they learn more about the dangers of such diseases and pay more attention to their health.

The above studies are limited to a single province or city. Furthermore, with the exception of some statistical descriptive reports, few studies consider the spatial dimension as an independent factor in chronic disease research in China. There are tremendous differences in economic development levels, medical conditions, and living conditions among the provinces in China, all of which can affect the diagnosis and treatment of a chronic disease. As a result, the prevalence of chronic diseases is quite different among the provinces. Thus, province is hypothesized to be a critical factor affecting chronic diseases and specific chronic diseases. In addition, the reported prevalence is hypothesized to be similar to the case of the census register, in that it is higher in economically developed provinces.

In summary, based on past studies and on China's tremendous geographic differences, the following factors are hypothesized to be important factors affecting the prevalence of chronic diseases and specific chronic diseases in elderly Chinese people: gender, census register (urban or rural), marital status, smoking, drinking, age, education years, sports activities and cultural activities, and province. In addition, because most of the samples between surveys in 2006 and 2010 are the same, the effects of time is taken into account.

Previous studies usually only consider the linear effects of the continuous covariates (such as education years) on a chronic disease, even though they may have nonlinear effects. Because STAR models can include spatial effects, the nonlinear effects of continuous factors, and linear or fixed effects in a single model, a STAR model is applied to determine which covariates have considerable effects on chronic diseases in China.

The rest of this chapter is structured as follows. Section 2.1 introduces the STAR models, and their estimations. Section 2.2 explains the data. Section 2.3 presents the STAR model applied in this chapter, and describes the `R2BayesX` settings. Section 2.4 discusses the empirical results for chronic diseases, as well as for two specific chronic diseases (hypertension and heart disease) using the `R2BayesX` package. Section 2.5 presents the summary of this chapter.

2.1 Estimation Method

2.1.1 STAR Models

STAR models were first introduced by Fahrmeir et al. (2004), and not only contain generalized linear effects, but also allow for nonlinear effects of continuous covariates and spatial effects.

For generalized linear models, the mean μ of the response variable y is linked to a linear predictor η by

$$\mu = h^{-1}(\eta), \ \eta = \boldsymbol{x}'\boldsymbol{\gamma}, \tag{2.1}$$

where h is a known link function and $\boldsymbol{\gamma}$ denotes unknown regression coefficients.

Following Fahrmeir et al. (2004, p.734), in STAR models, the linear predictor is replaced by the following structured additive predictor:

$$\eta = f_1(x_1) + \cdots + f_p(x_q) + \boldsymbol{w}'\boldsymbol{\gamma}, \tag{2.2}$$

where $x_1, \ldots, x_q$ are nonlinear covariates, and f_j are smooth functions, which can represent potentially non-linear effects of continuous covariates or spatially structured and unstructured effects.

Furthermore, when the response variable is binary, the link function can be assumed to be a logit function, and the following logistic STAR model is considered:

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \eta = f_1(x_1) + \cdots + f_p(x_q) + \boldsymbol{w}'\boldsymbol{\gamma}, \tag{2.3}$$

where p denotes the probability of a specific event occurring (such as the probability of a person having a chronic disease).

2.1.2 Estimation of STAR Models

In this chapter, the STAR model is estimated using a Bayesian inference. For the Bayesian inference, all components of the STAR models must be supplemented with appropriate prior assumptions.

In the STAR model (2.2), $\boldsymbol{w}'\boldsymbol{\gamma}$ denotes the fixed effects. In general, a diffuse prior $p(\boldsymbol{\gamma}) \propto const$ is assumed for the fixed effects parameter $\boldsymbol{\gamma}$. At the same time, specific priors are given to the functions $f_j(\cdot)$, and depend on the type of the covariate.

For the nonlinear effects of continuous covariates $f_j(\cdot)$, Bayesian P-splines are utilized. P-splines are an improvement over B-splines, and introduce a penalty variable to prevent over fitting.

The basic idea behind P-splines is dividing the data interval into a relatively large number of sub-intervals, and an unknown smooth function f of a covariate x can be approximated by a linear combination of some basis functions. P-splines can be approximated by a polynomial spline of degree l, defined on a set of equally spaced knots $x^{\min} = \zeta_1 < \zeta_2 < \cdots < \zeta_m = x^{\max}$ within the domain of x. Following Fahrmeir et al. (2013, pp.426–431), a spline can be expressed by an adequate linear combination of $d = m + l - 1$ B-spline basis functions:

$$f(x) = \sum_{j=1}^{d} \beta_j B_j(x), \tag{2.4}$$

where $B_j(x)$ of degree l is defined as follows: for $j = 1, \ldots, d-1$,

$$\begin{cases} B_j^0(x) = I(\zeta_j \leq x < \zeta_{j+1}) & l = 0 \\ B_j^1(x) = \dfrac{x - \zeta_{j-1}}{\zeta_j - \zeta_{j-1}} I(\zeta_{j-1} \leq x < \zeta_j) + \dfrac{\zeta_{j+1} - x}{\zeta_{j+1} - \zeta_j} I(\zeta_j \leq x < \zeta_{j+1}) & l = 1 \\ B_j^l(x) = \dfrac{x - \zeta_{j-l}}{\zeta_j - \zeta_{j-l}} B_{j-1}^{l-1}(x) + \dfrac{\zeta_{j+1} - x}{\zeta_{j+1} - \zeta_{j+1-l}} B_j^{l-1}(x) & l \geq 2, \end{cases} \tag{2.5}$$

where $I(\cdot)$ is an indicator function.

The crucial choice for P-splines is the number of knots: too few knots may not be flexible enough, while choosing too many knots may over fit the data. To prevent over fitting, a penalty term is included. The penalty terms are expressed as in Eilers and

Marx (1996):

$$P(\lambda) = \frac{1}{2}\lambda \sum_{j=r+1}^{d} (\Delta^r \beta_j)^2, \tag{2.6}$$

where λ is the smoothing parameter and Δ^r denotes the rth-order differences. In most applications, second-order differences are chosen, which are defined as $\Delta^2\beta_j = \Delta^1\Delta^1\beta_j = \Delta^1\beta_j - \Delta^1\beta_{j-1} = \beta_j - 2\beta_{j-1} + \beta_{j-2}$. Furthermore, when $\lambda \to \infty$, the function estimate for $f(x)$ is close to linear in the case of second-order differences.

In applications, a common choice for P-splines is B-splines of degree $l = 3$, with $m = 20$ equidistant knots. These setting ensure that the estimated function is twice continuously differentiable, while allows sufficient flexibility to capture the typical non-linear mode.

The general form of the prior of $\boldsymbol{\beta}_j$ for a P-spline is given by the multivariate normal distribution:

$$p(\boldsymbol{\beta}_j|\tau_j^2) \propto \exp\left(-\frac{1}{2\tau_j^2}\boldsymbol{\beta}_j'\boldsymbol{K}_j\boldsymbol{\beta}_j\right), \tag{2.7}$$

where $\boldsymbol{K}_j$ is a penalty matrix, and τ_j^2 is a prior variance, which determines the impact of the prior distribution on the function estimates. For the full Bayesian inference, weakly informative inverse Gamma hyperpriors $\tau_j^2 \sim \mathrm{IG}(a_j, b_j)$ are assigned to τ_j^2, with $a_j = b_j = 0.001$ as a general setting. More detailed information about the Bayesian P-splines can be found in Lang and Brezger (2004).

As mentioned above, f_j denotes smooth functions that can be used to represent the potentially non-linear effects of continuous covariates or to represent a spatial effect. If f_j represents a spatial effect, it is expressed as $f_{spat}(\cdot)$.

The spatial effect in a STAR model is the effect of a spatial covariate. Usually, this is a proxy for unobserved influential factors, some of which may have a strong spatial correlation (structured), while others may be present only locally (unstructured). Thus, in order to distinguish between these two kinds of spatial effects, $f_{spat}(\cdot)$ is split into a spatially correlated (structured) part $f_{str}(\cdot)$ and spatially uncorrelated (unstructured) part $f_{unstr}(\cdot)$, i.e. $f_{spat}(\cdot) = f_{str}(\cdot) + f_{unstr}(\cdot)$. The structured spatial effects simply indicate that the spatial effects are correlated. There is no specific structure imposed on the spatial effects.

The spatially structured effect $f_{str}(\cdot)$ can be specified using stationary Gaussian random field (GRF) priors. When the place of residence is known exactly, given by geographical x-y coordinates, the spatial analysis can be conducted using a stationary GRF. The estimation of a GRF is based on the centroids of particular regions for geosplines and geokriging terms. More detailed information about stationary GRF priors can be found in Fahrmeir et al. (2013, pp.500–530). The spatially unstructured effect $f_{unstr}(\cdot)$ can be specified using simple Gaussian independent and identically distributed (i.i.d.) priors, and denotes the random effect of a covariate.

2.2 Empirical Data

A chronic disease is defined using the medical definition: a disease that lasts for a long time (more than three months) and cannot be cured.

The data for this chapter are taken from the 2000, 2006, and 2010 Chinese Urban and Rural Elderly Population Surveys, conducted by the China Research Center on Aging of the National Committee on Aging. The survey in 2000 only investigated whether people were suffering from chronic diseases, while those in 2006 and 2010 were expanded to include questions on specific chronic diseases. And most of the samples between surveys in 2006 and 2010 are the same. Thus, two specific chronic diseases (hypertension and heart disease, both of which are common in China, with a prevalence of more than 10%)

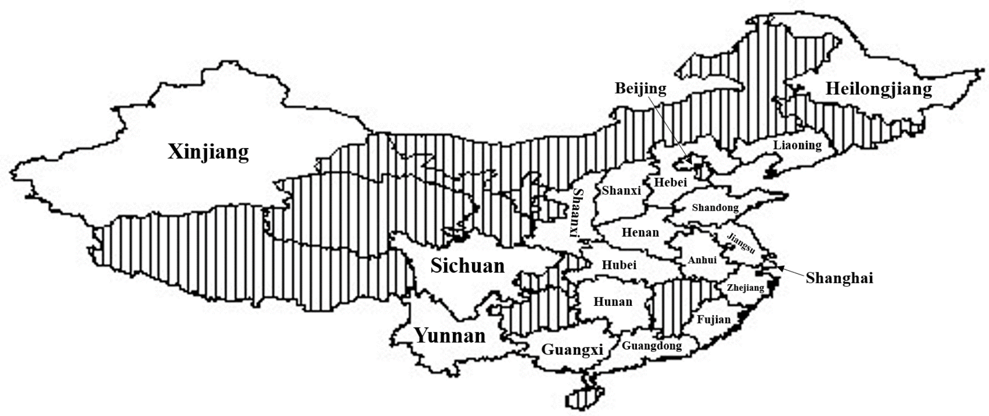


Figure 2.1: 20 selected provinces, municipalities and autonomous regions of China mainland

are analyzed in 2006 and 2010.

Moreover, the surveys focused on the following 20 representative provinces, municipalities, and autonomous regions: the eastern region - Beijing, Shanghai, Hebei, Liaoning, Jiangsu, Zhejiang, Fujian, Shandong, and Guangdong; the central region - Heilongjiang, Anhui, Henan, Shanxi, Hubei, and Hunan; and the western region - Sichuan, Yunnan, Shaanxi, Xinjiang, and Guangxi. The selected provinces, municipalities, and autonomous regions are shown in Figure 2.1. In China, the degree of economic development is closely related to geographical location. In general, provinces in the eastern region are almost economically developed provinces, provinces in the central region are moderately developed provinces, and provinces in the western region tend to be economically backward provinces.

The data sampling method used is the same as that of the Fifth Population Census;

based on the distribution of the population aged 60 years and older, a quota from each of the regions is determined. Then, stratified sampling is used to confirm that the survey results represent the total elderly population in China (Gao and Li, 2016).

After the survey samples were determined, the interviewers conducted household surveys. Here, interviews were conducted by interviewers, who then completed the questionnaires on behalf of the interviewees, based on their responses. No questionnaires were completed by the interviewees. Then, the interviewers checked and verified the responses after the investigation. As a result, the data accuracy is high. The three surveys generated $20,256$ responses, $19,947$ responses, and $19,986$ responses, respectively.

2.3 Empirical Models

2.3.1 Model Specification

Because the samples include data on whether people suffer from chronic diseases in 2000, 2006, and 2010, the first step is a Bayesian analysis of the geographic distribution of chronic diseases and their relationships with potential risk factors. Moreover, for 2006 and 2010, data are included on common chronic diseases (hypertension and heart disease). Thus, the second step is a Bayesian analysis of the geographic distribution of these two common chronic diseases and their relationships with potential risk factors. The second step is a refinement of the first step. Therefore, this paper describes how to implement the first step only.

Given a set of observations y_i, $1 \leq i \leq n$, y_i is a binary response for a chronic disease, such that

$$y_i = \begin{cases} 1 & \text{if one has a chronic disease} \\ 0 & \text{otherwise.} \end{cases}$$

Because the responses are binary, a logistic STAR model is considered to estimate the probability of an elderly person having a chronic disease ($y_i = 1$) versus the probability of an elderly person not having a chronic disease ($y_i = 0$):

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \eta_i = f_1(x_{i1}) + \cdots + f_p(x_{iq}) + \boldsymbol{w}_i'\boldsymbol{\gamma},\ 1 \leq i \leq n, \tag{2.8}$$

where $p_i = \Pr(y_i = 1)$, $x_{i1}, \ldots, x_{iq}$ are p continuous covariates, f_j are smooth functions, $\boldsymbol{w}_i = (w_{i1}, \ldots, w_{ir})'$ is a vector of r categorical covariates, and $\boldsymbol{\gamma}$ is an r-dimensional vector of unknown regression coefficients for the categorical covariates $\boldsymbol{w}_i$. The response is distributed as a Bernoulli random variable, such that $f(y_i|\eta_i) = {p_i}^{y_i}(1-p_i)^{(1-y_i)} = \exp[y_i\eta_i - \log(1+\exp(\eta_i))]$ for $y_i = 0, 1$, where $\eta_i = \text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right)$.

Based on the previous studies mentioned in the Introduction, the linear effects of the following categorical covariates are analyzed: gender of the elderly person (female or male), census register (urban or rural), marital status (live with spouse, live differently with spouse, widowed, divorce, and unmarried), smoking (smoked previously, currently smoke, and never smoke), and drinking (drank previously, currently drink, and never drink). In addition, the potential nonlinear effects of the following continuous covariates are investigated: the elderly's age (Age), education years (EY), number of sports activities (SA), and cultural activities (CA). Furthermore, because the observations on chronic diseases are associated with where a person lives, it is important to account for geographical/spatial differences in the analysis. Therefore, by taking spatial and nonlinear effects into account, the predictor shown in (2.8) is replaced by the following predictor:

$$\eta_i = f_1(Age_i) + f_2(EY_i) + f_3(SA_i) + f_4(CA_i) + f_{spat}(Province_i) + \boldsymbol{w}_i'\boldsymbol{\gamma}, \tag{2.9}$$

where $f_1(Age_i)$, $f_2(EY_i)$, $f_3(SA_i)$, and $f_4(CA_i)$ are nonlinear smooth effects of the continuous covariates, and $f_{spat}(Province)$ is the effect of the spatial covariate $Province_i$. Here, $Province_i \in \{1, \ldots, S\}$ is an integer, where S is the number of surveyed regions. Every integer indicates the province, municipality, or autonomous region in which the respondent is living. For example, $Province_i = 1$ means the respondent lives in Heilongjiang province, and $Province_i = 8$ denotes a respondent living in Beijing city. In this chapter, the number of surveyed provinces, municipalities, and autonomous regions is 20; thus $S = 20$ (i.e., $Province_i \in \{1, \ldots, 20\}$).

In China, the provinces are connected in that they usually have some similarity and correlation, and so they are spatially correlated. Therefore, the structured spatial effect $f_{str}(province)$ is considered rather than the unstructured spatial effect $f_{unstr}(province)$.

Finally, the following STAR model is estimated:

$$\eta_i = \beta_0 + f_1(Age_i) + f_2(EY_i) + f_3(SA_i) + f_4(CA_i) + f_{str}(province_i) + \boldsymbol{w}_i'\boldsymbol{\gamma}, \quad (2.10)$$

where $f_{str}(province_i)$ is a structured spatial effect, and $\boldsymbol{w}_i'\boldsymbol{\gamma}$ are the linear effects of the following categorical covariates: gender of the elderly, census register, marital status, smoking, and drinking.

In this chapter, the estimation of the above STAR model is obtained using a Bayesian inference. In the STAR model (2.10), for the fixed effects $\boldsymbol{w}_i'\boldsymbol{\gamma}$, a diffuse prior $p(\boldsymbol{\gamma}) \propto const$ is assumed for the parameter $\boldsymbol{\gamma}$. For the nonlinear effects of the continuous covariates $f_1(Age)$, $f_2(EY)$, $f_3(SA)$, and $f_4(CA)$, Bayesian P-splines are utilized. In addition, the structured spatial effect $f_{str}(province)$ is estimated using the stationary Gaussian random field (GRF) approach, because the geographical x-y coordinates of every surveyed province in this chapter are known exactly.

2.3.2 Model Implementation

For each data set (2000, 2006 and 2010), the STAR model (2.10) is fitted to a chronic disease.

All categorical covariates are changed into dummy variates. For example, marital status has five categories: live with spouse; live differently with spouse; widowed; divorce; and unmarried. Four dummy variates are used to represent this categorical covariate: marital statusA: live with spouse; marital statusB: live differently with spouse; marital statusC: widowed; marital statusD: divorce.

The model (2.10) can be implemented in `R2BayesX`, an open `R` package for STAR models. For this model, $25,000$ Markov chain Monte Carlo (MCMC) iterations were carried out after a burn in sample of $2,000$. In general, these random numbers are correlated. Thus, every 10th sampled parameter of the Markov chain is stored. The posterior mean, posterior standard deviation, posterior median, and 90% and 95% credible intervals for all parameters, estimated from the posterior distributions, are used to assess the model fit.

When fitting $f_{str}(Province)$ in `R2BayesX`, a "map" argument is needed. It can be an object of class "SpatialPolygonsDataFrame" or an object of class "`bnd`." Spatial polygon data in China can be downloaded as shape files. Furthermore, using the function `shp2bnd()` of the `R` package – shapefiles package, the shapefiles can be changed to "`bnd`" objects, – `Chinabnd`. The class "`bnd`" is a `list()` of polygon matrices, with x- and y-coordinates of the boundary points in the first and second columns, respectively, which can be used to calculate the centroids of polygons to estimate the smooth bivariate effects of the resulting coordinates.

2.4 Empirical Results

A logistic STAR model is used to estimate the probability of an elderly person having a chronic disease ($y_i = 1$) versus the probability of an elderly person not having a chronic disease ($y_i = 0$). It is well known that the coefficients in a logistic regression model do not represent marginal effects, but rather log odds. It is difficult to interpret the coefficients, and this problem becomes even more complex when considering non-linear effects. Therefore, I focus on which group has a greater impact, not the extent of the effect. For example, I examine whether the prevalence of a chronic disease in elderly females is higher than that of elderly males, not the marginal effects of the prevalence of a chronic disease in elderly females and males.

From (2.8), the structured additive predictor η_i and probability $p_i = \Pr(y_i = 1)$ are positively related. If the coefficient is positive, η_i of the experimental group is larger than η_i of the control group. Then, $p_i = \Pr(y_i = 1)$ of the experimental group is larger, which means the experimental group is more likely to have chronic diseases. Otherwise, if the coefficient is negative, the control group is more likely to have chronic diseases. Furthermore, when all other coefficients remain unchanged, a certain coefficient becomes larger than η_i becomes larger, and the probability $p_i = \Pr(y_i = 1)$ becomes larger, that is, a larger coefficient denotes a greater effect on a chronic disease.

Table 2.1 displays the variables used in the models and gives their meanings and values. Table 2.2 compares the hypotheses and the empirical results on the reported prevalence of a chronic disease, hypertension, and heart disease.

Table 2.1: Variables in the STAR models

Variable	Description	Value
Chronic Disease	Whether one has a chronic disease or not.	Having chronic disease is 1, else is 0.
Hypertension	Whether one has hypertension or not.	Having hypertension is 1, else is 0.
Heart Disease	Whether one has heart disease or not.	Having heart disease is 1, else is 0.
Time	Time of investigation.	2010 is 1; 2006 is 0.
Gender	Gender of the elderly person with categories 'male' and 'female'.	Female is 1; Male is 0.
Census Register	Census register with categories 'urban' and 'rural'.	Urban is 1; Rural is 0.
Marital Status	Marital status with categories 'live with spouse', 'live differently with spouse', 'widowed', 'divorce' and 'unmarried'	Using four dummy variables: Marital StatusA: live with spouse; Marital StatusB: live differently with spouse; Marital StatusC: widowed; Marital StatusD: divorce. Yes is 1, no is 0.
Smoking	Smoking condition with categories 'smoked previously', 'currently smoke' and 'never smoke'.	Using two dummy variables: SmokingA: smoked previously; SmokingB: currently smoke; Yes is 1, no is 0.
Drinking	Drinking condition with categories 'drank previously', 'currently drink' and 'never drink'.	Using two dummy variables: DrinkingA: drank previously; DrinkingB: currently drink; Yes is 1, no is 0.
Age	Age of the elderly people in years.	Continuous covariate, minimum value is 60.
Education Years	One's education years.	Continuous covariate, minimum value is 0.
Sports Activities	Number of sports activities one takes part in.	Continuous covariate, value changes from 0 to 5.
Cultural Activities	Number of cultural activities one takes part in.	Continuous covariate, value changes from 0 to 10.
Province	Province where one lives in.	Spatial covariate, integer, value changes from 1 to 20.

Table 2.2: Comparison of Hypotheses and Empirical Results on the Reported Prevalence

Variable	Hypotheses	Empirical Results (Posterior Mean)						
		Chronic Disease			Hypertension		Heart Disease	
		2000	2006	2010	2006	2010	2006	2010
Gender	considerable influence females > males	0.340**	0.344**	0.247**	0.158**	0.093**	0.489**	0.372**
Census Register	critical influence urban > rural	0.642**	0.631**	0.544**	0.599**	0.462**	0.872**	0.368**
Marital StatusA	considerable influence	-0.037	0.115	-0.100	-0.036	-0.033	0.342	0.111
Marital StatusB	divorced/widowed >	-0.101	0.258	-0.378	-0.045	-0.392	0.268	-0.024
Marital StatusC	other types	-0.034	0.098	-0.213	-0.028	-0.018	0.258	0.056
Marital StatusD		-0.031	0.181	-0.085	-0.024	-0.325	0.193	-0.058
SmokingA	considerable influence	0.430**	0.261**	0.267**	-0.051	-0.009	0.116*	0.117*
SmokingB	smoke > never smoke	0.037	-0.122**	-0.022	-0.237**	-0.178**	-0.003	-0.180**
DrinkingA	considerable influence	0.233**	0.274**	0.181**	0.173**	0.221**	0.018	0.057
DrinkingB	drink > never drink	-0.294**	-0.175**	-0.239**	-0.156**	-0.087**	-0.306**	-0.253**
Age	non-negligible influence older > younger	0.005**	0.005**	0.011**	0.006**	0.003**	0.010**	0.014**
Education Years	non-negligible influence longer > shorter	0.003**	0.005**	0.002**	0.004**	0.002**	0.003**	0.002**
Sports Activities	non-negligible influence more > less	0.003**	0.003**	0.003**	0.003**	0.003**	0.003**	0.004**
Cultural Activities	non-negligible influence more > less	0.004**	0.003**	0.004**	0.005**	0.005**	0.003**	0.003**
Province	critical influence developed > backward	0.416**	0.375**	0.054**	0.144**	0.201**	0.307**	1.975**

[1] "$**$" and "$*$" denote that zero is not included in the 95% and 90% credible intervals, respectively.

2.4.1 Empirical Results for Chronic Diseases

In Table 2.2, for the covariates of gender and census register, zero is not included in the 95% credible intervals in 2000, 2006, and 2010. Therefore, these covariates do affect chronic diseases. In addition, the posterior means of gender and census register are positive. This indicates that in comparing elderly females and elderly males (female is 1, male is 0), and urban elderly people and rural elderly people (urban is 1, rural is 0), the reported prevalence of the former groups is higher than that of the latter groups. These results are also consistent with the hypotheses.

In general, marital status may affect the health of elderly people (Kiecolt-Glaser and Newton, 2001). However, being married does not guarantee health benefits. A decline in the quality of marriage has a negative effect on mental and physical health (Wickrama et al., 1997). Therefore, marital status has a negligible effect on a chronic disease, because the 95% and 90% credible intervals of marital status include zero in all three years. Thus, the hypothesis of marital status is rejected.

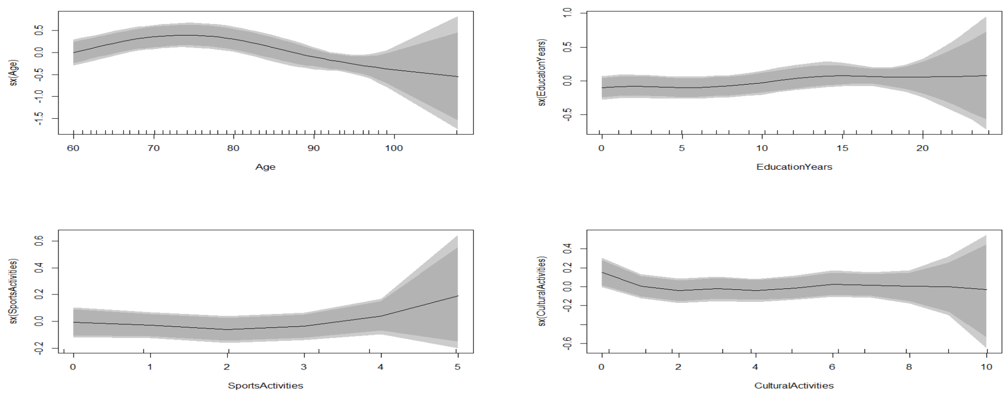
For the covariates smoking and drinking, zero was not included in the 95% credible intervals of smokingA (smoked previously), drinkingA (drank previously), and drinkingB (currently drink) in all three years. Therefore, the hypotheses is held that smoking and drinking do affect chronic diseases. However, smoking and drinking have two different kinds of effects on a chronic disease. The results of smokingA (smoked previously)/drinkingA (drank previously) are consistent with the following hypotheses: the reported prevalence in elderly people who smoked/drank previously, but no longer do so, is higher than that in elderly people who never smoke/drink. However, the results of smokingB (currently smoke) and drinkingB (currently drink) seem to be counter-intuitive: the reported prevalence in elderly people who currently smoke/drink is lower than that in those who never smoke/drink. One possible reason for this apparent contradiction in the case of drinking may be the following: moderate drinking has little effect on health, and some reports even show that moderate drinking is beneficial to our health (Fillmore et al., 2006).

In addition, as explained above, a larger coefficient denotes a greater effect on chronic diseases. From Table 2.2, for the fixed effects of the categorical covariates on chronic diseases, the census register has the greatest effect on chronic diseases, followed by gender, smoking, and drinking.

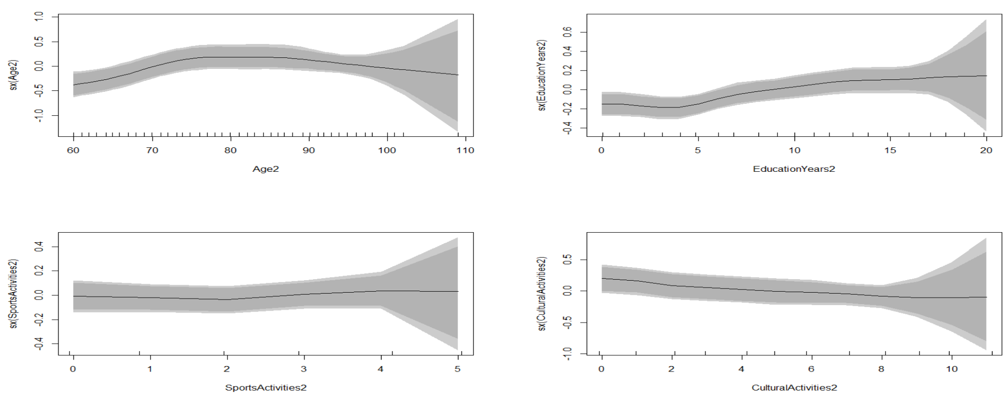
The nonlinear effects of the continuous covariates (age, years of education, sports activities, and cultural activities) are considerable, but they are all very small (mean values are less than 0.012) in the three years. Furthermore, the posterior means are all positive, which indicates a greater reported prevalence for older people with a higher level of education, and who do more sports activities and cultural activities. Age and education are consistent with the hypotheses, but sports activities and cultural activities are contrary to the hypotheses. One possible reason may be as follows: people who do more sports activities and cultural activities pay more attention to their health, and so are more likely to go to the hospital for an examination, and more likely to report having a chronic disease. Figure 2.2 gives the detailed nonlinear effects of these continuous covariates on a chronic disease, with 95% credible bands in all three years. The tails of all continuous covariates are wide because the numbers of observations in these parts are all very small.

Since a larger coefficient denotes a greater effect on chronic diseases, and the coefficient of province on chronic diseases is relatively large, thus, the spatial effect of province is critical.

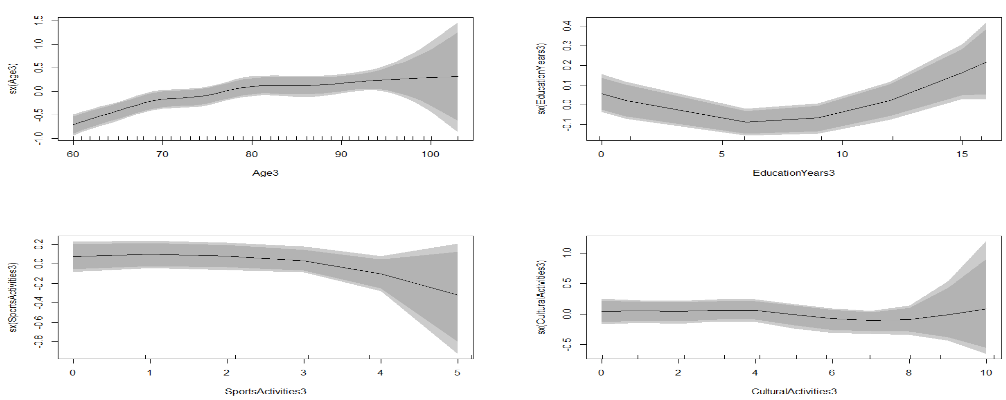
In 2000, Xinjiang province's structured spatial effect was the highest, followed by



(a) 2000



(b) 2006



(c) 2010

Figure 2.2: Effects of smooth terms on a chronic disease with 90% and 95% credible bands in 2000, 2006 and 2010

Anhui, Shaanxi, Sichuan, Guangxi, Beijing, and Shanxi provinces. Shandong province's structured spatial effect was the lowest (see Figure 2.3(a)). In 2006, Xinjiang province's structured spatial effect is still the highest, followed by Anhui, Hubei, and Hunan provinces. Guangdong province's structured spatial effect was the lowest, and Shandong, Guangxi and Fujian provinces' structured spatial effects are very low (see Figure 2.3(b)). In 2010, Zhejiang province's structured spatial effect is the highest, Hunan province's structured spatial effect is the lowest. In addition, Anhui, Fujian, and Xinjiang provinces' structured spatial effects are relatively high, and Shandong and Liaoning provinces' structured spatial effects are relatively low (see Figure 2.3(c)).

In conclusion, the province is a critical factor affecting chronic diseases, but that the high reported prevalence of chronic diseases is not only in economically developed provinces (such as Zhejiang, Beijing, Fujian), but also in the economically backward provinces with complex terrain (such as Xinjiang, Guangxi), as shown in Figure 2.3. In addition, special attention should be paid to Xinjiang and Anhui provinces, because their structured spatial effects are quite high in all three years.

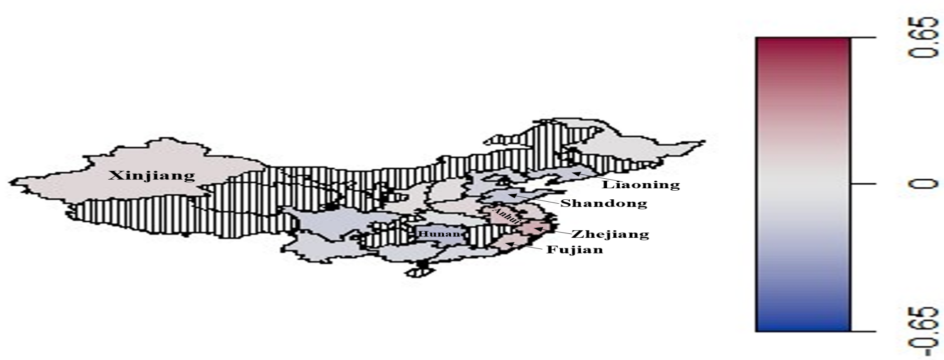
The high reported prevalence does not necessarily indicate that the health conditions in these provinces are bad. On the contrary, the high reported prevalence may indicate that elderly people pay more attention to their health. However, a high reported prevalence in economically backward provinces with complex terrain may indicate that the health conditions in these provinces are poor, for example, in Xinjiang province.

2.4.2 Empirical Results for Hypertension

As shown in Table 2.2, for the covariates of gender, census register, smokingB (currently smoke), drinkingA (drank previously), drinkingB (currently drink), age, education years, sports activities, cultural activities, and province, zero is not included in the

(a) 2000

(b) 2006



(c) 2010

Figure 2.3: Structured spatial effect on a chronic disease in 2000, 2006 and 2010

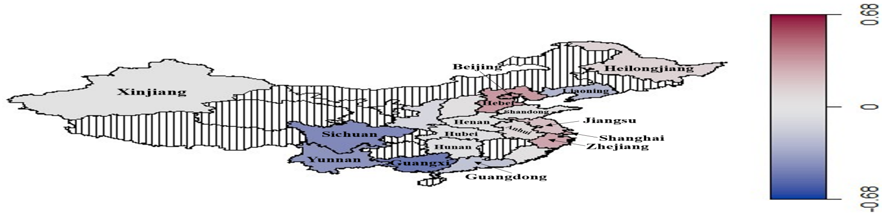
95% credible intervals in 2006 and 2010. Therefore, they have considerable effects on hypertension. But the hypothesis for marital status is rejected.

The empirical results for the fixed effects of the categorical covariates and for the nonlinear effects of the continuous covariates on hypertension are very similar to the results for chronic diseases. Thus, they are not repeated here. The differences between general chronic diseases and hypertension are mainly reflected in the spatial effects.

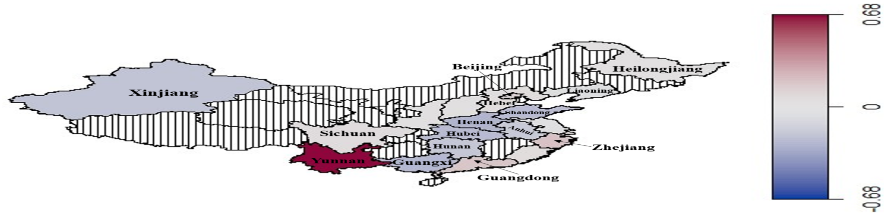
Since a larger coefficient denotes a greater effect, and the coefficient of province on hypertension is relatively large, thus, the spatial effect of province on hypertension is critical. In 2006, Hebei, Beijing, and Zhejiang provinces' structured spatial effects on hypertension are the highest, followed by Jiangsu, Shanghai and Heilongjiang. Guangxi, Yunnan, and Sichuan provinces' structured spatial effects are the lowest, although Guangdong and Liaoning provinces' structured spatial effects are also relatively low (see Figure 2.4(a)). In 2010, Yunnan province's structured spatial effect is the highest (see Figure 2.4(b)), showing a marked increase over the low level in 2006.

In conclusion, the province is a critical factor affecting hypertension, but that the highest reported prevalence occurs mainly in economically developed provinces (e.g., Zhejiang, Beijing, and Guangdong), and not in economically backward provinces, as shown in Figure 2.4. In addition, note that Zhejiang province's structured spatial effects are quite high in both years.

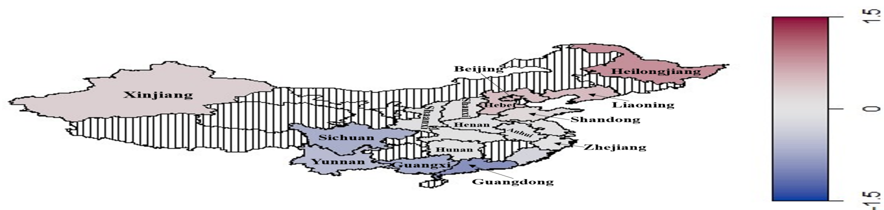
The high reported prevalence in economically developed provinces may indicate that the elderly pay more attention to their health. However, the high reported prevalence in Yunnan province indicates that the health conditions in this province are poor. Reports in 2013 revealed that for every 10 Kunming (Yunnan provincial capital) residents, two suffer from hypertension, but that about 70% of patients are not aware of this.



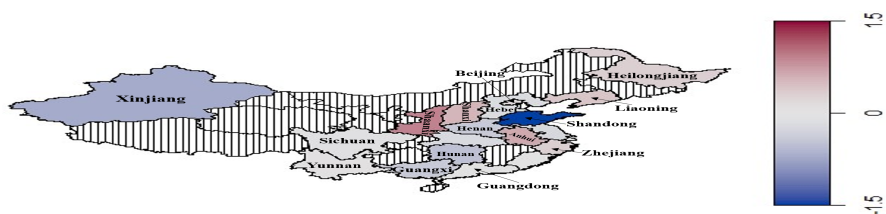
(a) Hypertension-2006



(b) Hypertension-2010



(c) Heart disease-2006



(d) Heart disease-2010

Figure 2.4: Structured spatial effect on hypertension and heart disease in 2006 and 2010

2.4.3 Empirical Results for Heart Disease

As shown in Table 2.2, for the covariates of gender, census register, smokingA (smoked previously), drinkingB (currently drink), age, education years, sports activities, cultural activities, and province, zero is not included in the 95% and 90% credible intervals in 2006 and 2010. Therefore, they have considerable effects on heart disease. But the hypothesis for marital status is rejected.

The empirical results for the fixed effects of the categorical covariates and the nonlinear effects of the continuous covariates on heart disease are very similar to the results for chronic diseases. Thus, they are not repeated here. The differences between general chronic diseases and heart disease are still mainly reflected in the spatial effects.

As explained above, a larger coefficient denotes a greater effect, and the coefficient of province on heart disease is relatively large, thus, the spatial effect of province on heart disease is critical. In 2006, Heilongjiang province's structured spatial effect was the highest, followed by Liaoning, Beijing, Hebei, and Xinjiang. Guangdong, Guangxi, Yunnan, and Sichuan provinces' structured spatial effects were the lowest (see Figure 2.4(c)). In 2010, Shaanxi province's structured spatial effect was the highest, followed by Shanxi and Anhui. Shandong province's structured spatial effect was the lowest, although Xinjiang, Hunan, and Guangxi were relatively low (see Figure 2.4(d)).

In conclusion, the province is a critical factor affecting heart disease. As shown in Figure 2.4, in 2006, the highest reported prevalence appeared not only in economically backward provinces (e.g., Xinjiang), but also in the economically developed provinces (e.g., Hebei, Beijing) and the moderately developed provinces (e.g., Heilongjiang, Liaoning). In contrast, in 2010, high spatial effects appeared in moderately developed provinces only (e.g., Shaanxi, Shanxi, and Anhui).

Table 2.3: Posterior Estimates of the Parameters for Chronic Disease (Add Time)

	Chronic Disease			Hypertension			Heart Disease		
Variable	Mean	Sd	Median	Mean	Sd	Median	Mean	Sd	Median
Fixed effects:									
(Intercept)	0.789	0.173	0.784	-0.911	0.191	-0.906	-2.209	0.271	-2.191
Time	0.129	0.027	0.129**	0.329	0.027	0.330**	0.088	0.030	0.088**
Gender	0.292	0.033	0.292**	0.129	0.034	0.129**	0.431	0.036	0.431**
CR	0.590	0.033	0.590**	0.549	0.031	0.549**	0.721	0.035	0.721**
MSA	0.015	0.105	0.017	-0.054	0.123	-0.056	0.193	0.162	0.192
MSB	0.006	0.150	0.003	-0.280	0.166	-0.281*	-0.031	0.211	-0.029
MSC	-0.048	0.107	-0.047	-0.051	0.125	-0.050	0.120	0.163	0.119
MSD	0.067	0.166	0.065	-0.204	0.174	-0.201	0.053	0.219	0.052
SmokingA	0.260	0.045	0.261**	-0.020	0.043	-0.021	0.113	0.047	0.114**
SmokingB	-0.076	0.035	-0.077**	-0.207	0.037	-0.206**	-0.090	0.042	-0.089**
DrinkingA	0.226	0.045	0.225**	0.181	0.042	0.181**	0.022	0.046	0.022
DrinkingB	-0.201	0.034	-0.201**	-0.125	0.036	-0.124**	-0.280	0.041	-0.280**
Nonlinear effects:									
sx(Age)	0.003	0.004	0.002**	0.004	0.005	0.002**	0.011	0.012	0.007**
sx(EY)	0.003	0.006	0.002**	0.002	0.003	0.001**	0.002	0.002	0.001**
sx(CA)	0.003	0.006	0.002**	0.002	0.003	0.002**	0.002	0.003	0.002**
sx(SA)	0.002	0.003	0.001**	0.003	0.007	0.001**	0.002	0.004	0.001**
sx(Province)	0.118	0.057	0.105**	0.106	0.052	0.094**	0.356	0.181	0.312**

[1] sx(·) corresponds to the smooth functions in STAR models.

[2] CR, MSA, MSB, MSC, MSD, CA, EY and SA are the abbreviation of Census Register, Marital StatusA, Marital StatusB, Marital StatusC, Marital StatusD, Cultural Activities, Education Years and Sports Activities, respectively.

[3] "**" and "*" denote that zero is not included in the 95% and 90% credible intervals, respectively. Mean, Sd and Median are posterior mean, posterior standard deviation and posterior median, respectively.

2.4.4 Empirical Results of Adding Time Factor

Because most of the samples between surveys in 2006 and 2010 are the same, the effects of time is taken into account. And the estimation results are shown in Table 2.3.

Table 2.3 gives the posterior means, posterior standard deviations, and posterior medians for all covariates (including time) of a chronic disease, hypertension, and heart disease. In Table 2.3, for the covariate of time, zero is not included in the 95% credible intervals. Therefore, time do affect chronic diseases. In addition, the posterior means of time are positive. This indicates that in comparing 2006 and 2010 (2010 is 1, 2006 is 0), the reported prevalence in 2010 is higher than that in 2006.

The higher reported prevalence in 2010 does not necessarily indicate that the health conditions of the elderly people are worse. On the contrary, the high reported prevalence may indicate that elderly people pay more attention to their health with the time going.

2.5 Chapter Summary

In this chapter, the following factors are hypothesized to be important factors affecting the prevalence of chronic diseases and specific chronic diseases in elderly Chinese people: gender, census register (urban or rural), marital status, smoking, drinking, age, education years, sports activities and cultural activities, and province. In addition, because most of the samples between surveys in 2006 and 2010 are the same, the effects of time is taken into account. To verify the hypotheses, a STAR model is applied to determine which covariates have considerable effects on chronic diseases and specific chronic diseases (hypertension and heart disease).

3

Bayesian Analysis of a MSOP Model

Currently, there is few researchers considering multivariate outcomes in spatial econometrics. Since many phenomena are correlated and spatially dependent, it is better to study them together, rather than separately. Thus, a model that captures the interactions of multivariate outcomes and spatial dependency is necessary. Based on the ordered probit (OP) models and the spatial ordered probit (SOP) models (Jeliazkov et al., 2008; Smith and LeSage, 2004), a multivariate spatial ordered probit (MSOP) model is proposed to capture the spatial dependencies and interaction of multivariate response variables as well as the effects of the explanatory variables.

The rest of this chapter is structured as follows. Section 3.1 introduces the model specification for the MSOP model. Section 3.2 explains details of the Bayesian inference for the model. Section 3.3 presents the model for the simulated dataset and its simulation results. Section 3.4 explains the empirical data and MCMC settings, and discusses the empirical results for the two multivariate outcomes: self-rated health and life satisfaction. Section 3.5 presents the summary of this chapter.

3.1 Model Specification

3.1.1 Ordered Probit Models

In a standard OP model, observed response variable y_i for latent response variable z_i is expressed as:

$$y_i = l \text{ if } \gamma_{l-1} < z_i \leq \gamma_l,\ l = 1, \ldots, L, \tag{3.1}$$

where l denotes the classifications of the response variable y, and γ_l $(l = 1, \ldots, L)$ are the cutpoints. Furthermore, they are specified as follows:

$$\gamma_0 = -\infty < \gamma_1 < \gamma_2 < \cdots < \gamma_{L-1} < \gamma_L = \infty. \tag{3.2}$$

The latent variable model is expressed as follows:

$$z_i = \boldsymbol{x}_i' \boldsymbol{\beta} + \xi_i,\ i = 1, \ldots, n, \tag{3.3}$$

where i denotes observations, and z_i is a latent response variable for individual i, which cannot be observed. $\boldsymbol{x}_i$ is a q-dimensional vector of explanatory variables, and $\boldsymbol{\beta}$ is the corresponding unknown coefficients. ξ_i is assumed to follow an i.i.d. normal distribution with common variance: $\xi_i \sim \mathrm{N}(0, \sigma^2)$.

3.1.2 Spatial Ordered Probit Models

Similar to the standard OP models, the relation between observed response variable y_{ik} and latent response variable z_{ik} in a SOP model is explained through the cutpoints γ_l as follows:

$$y_{ik} = l \text{ if } \gamma_{l-1} < z_{ik} \leq \gamma_l,\ \ l = 1, \ldots, L, \tag{3.4}$$

$$\gamma_0 = -\infty < \gamma_1 < \gamma_2 < \cdots < \gamma_{L-1} < \gamma_L = \infty, \tag{3.5}$$

where i now denotes regions ($i = 1, \ldots, m$), and k denotes individuals inside each region ($k = 1, \ldots, n_i$). Therefore, the total number of observations is $\sum_{i=1}^{m} n_i = n$.

Following Smith and LeSage (2004), the latent variable model including spatial dependencies is in the following form:

$$z_{ik} = \boldsymbol{x}_{ik}'\boldsymbol{\beta} + \delta_i + \xi_{ik},\ \ i = 1, \ldots, m, \text{and } k = 1, \ldots, n_i, \tag{3.6}$$

$$\delta_i = \rho \sum_{j=1}^{m} w_{ij}\delta_j + \epsilon_i, i = 1, \ldots, m, \tag{3.7}$$

where δ_i represents a regional effect, and ξ_{ik} represents a individualistic effect. z_{ik} is a latent response variable, and $\boldsymbol{x}_{ik}$ is a q-dimensional vector of the explanatory variables, and thus $\boldsymbol{x}_{ik} = (x_{ik1}, \ldots, x_{ikq})'$. ρ is called the spatial coefficient, which reflects the degree of spatial influence. Weights w_{ij} reflect the closeness between regions i and j, which can be quantified by using contiguity or distance. The residuals ϵ_i are assumed to follow $\mathrm{N}(0, \sigma^2)$. The individualistic effects ξ_{ik} in the SOP model are assumed conditionally i.i.d. normal variates with zero means and common variance v_i; that is, $\xi_{ik} \sim \mathrm{N}(0, v_i)$. The regional differences in v_i capture possible heteroscedasticity in the model.

3.1.3 Multivariate Spatial Ordered Probit Models

Following the probit model with spatial dependencies in Smith and LeSage (2004) and models for multivariate ordinal outcomes in Jeliazkov et al. (2008), this chapter proposes a MSOP model, wherein the observed response variable y_{iks} is a censored form of the latent response variable z_{iks}:

$$y_{iks} = l_s \text{ if } \gamma_{s,l_s-1} < z_{iks} \leq \gamma_{s,l_s}, \text{ for } s = 1, \ldots, S, \text{and } l_s = 1, \ldots, L_s, \tag{3.8}$$

where i denotes spatial regions, k denotes individuals inside each region ($k = 1, \ldots, n_i$), and s denotes the sequence of observed response variables. Therefore, y_{iks} denotes the sth observed response variable of individual k in region i. l_s denotes the classifications of response variable y_{iks}. In addition, the cutpoints γ_{s,l_s} are specified as in Chen and Dey (2000):

$$-\infty = \gamma_{s,0} < \gamma_{s,1} = 0 < \gamma_{s,2} < \cdots < \gamma_{s,L_s-1} = 1 < \gamma_{s,L_s} = \infty. \tag{3.9}$$

The latent response variable z_{iks} in the MSOP model is specified as follows:

$$z_{iks} = \boldsymbol{x}'_{ikst}\boldsymbol{\beta}_s + \delta_{is} + \xi_{iks}, \text{ for } i = 1, \ldots, m, k = 1, \ldots, n_i, \text{and } s = 1, \ldots, S, \tag{3.10}$$

where $\boldsymbol{x}_{iks} = (x_{iks1}, x_{iks2}, \ldots, x_{iksq_s})'$ is a q_s-dimensional vector of the explanatory variables. Correspondingly, $\boldsymbol{\beta}_s = (\beta_{s1}, \beta_{s2}, \ldots, \beta_{sq_s})'$ is a q_s-dimensional vector of the regression parameters for $\boldsymbol{x}_{iks}$. δ_{is} and ξ_{iks} represent regional effect and individualistic effect for the sth outcome, respectively.

The economic development level and the supporting infrastructure (such as medical, education, health environment, sanitation etc.) are quite different among regions. But these conditions are common in the same region, and thus observations in the same region have some common characteristics. In this chapter, the regional effect δ_{is} is used to capture all unobserved common effects for observations within region i for the sth outcome. Furthermore, the above conditions are more or less affected by neighboring regions. For example, the economic development levels of the surrounding regions are high, and local economic development will also be driven. Therefore, the regional effects may have spatial autocorrelation. Thus, a spatial autoregressive model can be used:

$$\delta_{is} = \rho \sum_{j=1}^{m} w_{ij}\delta_{js} + \epsilon_{is}, \text{ for } i = 1, \ldots, m, \text{and } s = 1, \ldots, S, \tag{3.11}$$

where ρ is the spatial coefficient, which reflects the degree of spatial influence. Weights w_{ij} reflect the closeness between regions i and j, and a commonly used definition is the queen contiguity, which defines $w_{ij} = 1$ for regions sharing a common edge or a common vertex; otherwise, $w_{ij} = 0$. In addition, $w_{ii} = 0$ and row sums $\sum_{j=1}^{m} w_{ij}$ are normalized to 1. ϵ_{is} are assumed to follow a normal distribution; that is,

$$\epsilon_{is} \sim \mathrm{N}(0, \sigma_s^2).$$

$\delta_{i1}, \ldots, \delta_{iS}$ are assumed to be conditionally independent given ρ and $\boldsymbol{\sigma}^2$, and thus

$$\boldsymbol{\epsilon}_i \sim \mathrm{N}(\mathbf{0}, \mathrm{diag}(\boldsymbol{\sigma}^2)),$$

where $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{is}, \ldots, \epsilon_{iS})'$, and $\boldsymbol{\sigma}^2 = \left(\sigma_1^2, \ldots, \sigma_s^2, \ldots, \sigma_S^2\right)'$.

Moreover, defining $\boldsymbol{\epsilon}_s = (\epsilon_{1s}, \ldots, \epsilon_{is}, \ldots, \epsilon_{ms})'$, and $\boldsymbol{\delta}_s = (\delta_{1s}, \ldots, \delta_{is}, \ldots, \delta_{ms})'$, the following equation is obtained:

$$\boldsymbol{\delta}_s = \rho W \boldsymbol{\delta}_s + \boldsymbol{\epsilon}_s, \ \text{for}\, s = 1, \ldots, S. \tag{3.12}$$

Additionally, $\boldsymbol{\delta}_1 = (\delta_{11}, \ldots, \delta_{1s}, \ldots, \delta_{1S})', \ldots, \boldsymbol{\delta}_m = (\delta_{m1}, \ldots, \delta_{ms}, \ldots, \delta_{mS})'$ are assumed to be conditionally i.i.d. variates, and thus $\delta_{1s}, \ldots, \delta_{is}, \ldots, \delta_{ms}$ are conditionally independent given ρ and σ_s^2, and thus

$$\boldsymbol{\epsilon}_s \sim \mathrm{N}(\mathbf{0}, \sigma_s^2 \boldsymbol{I}_m).$$

Stacking all S outcomes, (3.10) can be written as:

$$\boldsymbol{z}_{ik} = \boldsymbol{X}_{ik}\boldsymbol{\beta} + \boldsymbol{\delta}_i + \boldsymbol{\xi}_{ik}, \ \text{for}\, i = 1, \ldots, m, \text{and } k = 1, \ldots, n_i, \tag{3.13}$$

where $\boldsymbol{z}_{ik}$ is an S-dimensional vector of the latent response variables, and thus $\boldsymbol{z}_{ik} = (z_{ik1}, \ldots, z_{iks} \ldots, z_{ikS})'$. The matrix of covariates $\boldsymbol{X}_{ik}$ is organized as:

$$\boldsymbol{X}_{ik} = \begin{pmatrix} \boldsymbol{x}'_{ik1} & & & & \\ & \ddots & & & \\ & & \boldsymbol{x}'_{iks} & & \\ & & & \ddots & \\ & & & & \boldsymbol{x}'_{ikS} \end{pmatrix}_{S \times q},$$

where $\boldsymbol{x}_{iks}$ is a q_s-dimensional vector of the explanatory variables, and thus the number of columns of $\boldsymbol{X}_{ik}$ is $\sum_{s=1}^{S} q_s = q$. Correspondingly, $\boldsymbol{\beta} = \left(\boldsymbol{\beta}'_1, \ldots, \boldsymbol{\beta}'_s, \ldots, \boldsymbol{\beta}'_S\right)'$ is a

q-dimensional vector of the regression parameters. $\boldsymbol{\delta}_i = (\delta_{i1}, \ldots, \delta_{is}, \ldots, \delta_{iS})'$ is an S-dimensional vector of the regional effects, and $\boldsymbol{\xi}_{ik} = (\xi_{ik1}, \ldots, \xi_{iks}, \ldots, \xi_{ikS})'$ is an S-dimensional vector of the individualistic effects. The individualistic effects $\boldsymbol{\xi}_{ik}$ are assumed conditionally i.i.d. normal variates with zero means and common covariance matrix $\boldsymbol{V}$; that is,

$$\boldsymbol{\xi}_{ik} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{V}),$$

where $\boldsymbol{V}$ is an $S \times S$ symmetric positive definite covariance matrix.

Furthermore, I define that:

$$G_{iks} = (\gamma_{s,l_s-1}, \gamma_{s,l_s}] \text{ if } y_{iks} = l_s, \text{for } l_s = 1, \ldots, L_s, \text{and } s = 1, \ldots, S,$$

$$G_{ik} = G_{ik1} \times \cdots \times G_{ikS}, \text{ for } i = 1, \ldots, m, \text{and } k = 1, \ldots, n_i.$$

Defining $\boldsymbol{y}_{ik} = (y_{ik1}, \ldots, y_{ikS})'$ as an S-dimensional vector of the response variables, it can be derived that:

$$p(\boldsymbol{y}_{ik}|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}_{ik}) = p(\boldsymbol{y}_{ik}|\boldsymbol{\gamma}, \boldsymbol{z}_{ik}) = 1(\boldsymbol{z}_{ik} \in G_{ik}), \tag{3.14}$$

where $1(\cdot)$ is an indicator function.

Let $\boldsymbol{\xi}_i = \begin{pmatrix} \boldsymbol{\xi}_{i1} \\ \vdots \\ \boldsymbol{\xi}_{ik} \\ \vdots \\ \boldsymbol{\xi}_{in_i} \end{pmatrix}_{n_iS\times 1}$, $\boldsymbol{z}_i = \begin{pmatrix} \boldsymbol{z}_{i1} \\ \vdots \\ \boldsymbol{z}_{ik} \\ \vdots \\ \boldsymbol{z}_{in_i} \end{pmatrix}_{n_iS\times 1}$, and $\boldsymbol{X}_i = \begin{pmatrix} \boldsymbol{X}_{i1} \\ \vdots \\ \boldsymbol{X}_{ik} \\ \vdots \\ \boldsymbol{X}_{in_i} \end{pmatrix}_{n_iS\times q}$,

where $\boldsymbol{\xi}_i$ denotes an n_iS-dimensional vector of the individualistic effects for all the individuals in region i ($\boldsymbol{\xi}_{ik}$ is an S-dimensional vector, and thus the dimension of $\boldsymbol{\xi}_i$ is

$\sum_{k=1}^{n_i} S = n_i S$). $\boldsymbol{z}_i$ denotes an $n_i S$-dimensional vector of the latent response variables for all the individuals in region i, and $\boldsymbol{X}_i$ denotes an $n_i S \times q$ matrix of covariates for all the individuals in region i ($\boldsymbol{X}_{ik}$ is an $S \times q$ matrix of covariates, and thus the dimension of $\boldsymbol{X}_i$ is $\sum_{k=1}^{n_i} S \times q = n_i S \times q$). Based on the above definitions, (3.13) can be rewritten as:

$$\boldsymbol{z}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \mathbf{1}_i \otimes \boldsymbol{\delta}_i + \boldsymbol{\xi}_i, \text{ for } i = 1, \ldots, m, \tag{3.15}$$

where $\mathbf{1}_i$ denotes a $n_i \times 1$ vector of 1's, that is, $\mathbf{1}_i = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{n_i \times 1}$.

Moreover, stacking all regions, (3.15) can be summarized as:

$$\boldsymbol{z} = \boldsymbol{X}\boldsymbol{\beta} + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta} + \boldsymbol{\xi}, \tag{3.16}$$

$$\boldsymbol{\delta} = \rho(\boldsymbol{W} \otimes \boldsymbol{I}_S)\boldsymbol{\delta} + \boldsymbol{\epsilon}, \tag{3.17}$$

$$\boldsymbol{\xi} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{I}_n \otimes \boldsymbol{V}), \tag{3.18}$$

$$\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2)), \tag{3.19}$$

where $\boldsymbol{z} = \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_i \\ \vdots \\ \boldsymbol{z}_m \end{pmatrix}_{nS \times 1}$, $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_i \\ \vdots \\ \boldsymbol{X}_m \end{pmatrix}_{nS \times q}$, $\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_s \\ \vdots \\ \boldsymbol{\beta}_S \end{pmatrix}_{q \times 1}$, $\boldsymbol{\Delta} = \begin{pmatrix} \mathbf{1}_1 & & \\ & \ddots & \\ & & \mathbf{1}_m \end{pmatrix}_{n \times m}$,

$$\boldsymbol{\delta} = \begin{pmatrix} \boldsymbol{\delta}_1 \\ \vdots \\ \boldsymbol{\delta}_i \\ \vdots \\ \boldsymbol{\delta}_m \end{pmatrix}_{mS\times 1}, \boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_1 \\ \vdots \\ \boldsymbol{\xi}_i \\ \vdots \\ \boldsymbol{\xi}_m \end{pmatrix}_{nS\times 1}, \boldsymbol{W} = \begin{pmatrix} w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mm} \end{pmatrix}_{m\times m}, \boldsymbol{\epsilon} = \begin{pmatrix} \boldsymbol{\epsilon}_1 \\ \vdots \\ \boldsymbol{\epsilon}_i \\ \vdots \\ \boldsymbol{\epsilon}_m \end{pmatrix}_{mS\times 1},$$

$$\text{diag}(\boldsymbol{\sigma}^2) = \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_S^2 \end{pmatrix}_{S\times S}, \boldsymbol{\sigma}^2 = \begin{pmatrix} \sigma_1^2 \\ \vdots \\ \sigma_S^2 \end{pmatrix}_{S\times 1}, \text{ and } \boldsymbol{I}_S = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}_{S\times S}.$$

I define $\boldsymbol{B}_\rho = \boldsymbol{I}_{ms} - \rho(\boldsymbol{W} \otimes \boldsymbol{I}_S)$, and assume that $\boldsymbol{B}_\rho$ is nonsingular. Then, (3.17) can be expressed as:

$$\boldsymbol{\delta} = \boldsymbol{B}_\rho^{-1}\boldsymbol{\epsilon}. \tag{3.20}$$

Since $\boldsymbol{\epsilon} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_m \otimes \text{diag}(\boldsymbol{\sigma}^2))$, $\boldsymbol{\delta}$ has a multivariate normal distribution; that is, $\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2 \sim \mathrm{N}\left[\boldsymbol{0}, \boldsymbol{B}_\rho^{-1}(\boldsymbol{I}_m \otimes \text{diag}(\boldsymbol{\sigma}^2))(\boldsymbol{B}_\rho^{-1})'\right]$. Therefore, the density of $\boldsymbol{\delta}$ is as follows:

$$\begin{aligned} p(\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2) &\propto |\boldsymbol{B}_\rho^{-1}(\boldsymbol{I}_m \otimes \text{diag}(\boldsymbol{\sigma}^2))(\boldsymbol{B}_\rho^{-1})'|^{-1/2} \exp\left\{-\frac{1}{2}\boldsymbol{\delta}' \left[\boldsymbol{B}_\rho^{-1}(\boldsymbol{I}_m \otimes \text{diag}(\boldsymbol{\sigma}^2))(\boldsymbol{B}_\rho^{-1})'\right]^{-1} \boldsymbol{\delta}\right\} \\ &\propto |\,\text{diag}(\boldsymbol{\sigma}^2)|^{-m/2}|\boldsymbol{B}_\rho| \exp\left(-\frac{1}{2}\boldsymbol{\delta}'\boldsymbol{B}_\rho'(\boldsymbol{I}_m \otimes \text{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho\boldsymbol{\delta}\right). \end{aligned}$$

Similarly, based on (3.12) and (3.1.3), the density of $\boldsymbol{\delta}_s$ can be obtained. I define $\boldsymbol{A}_\rho = \boldsymbol{I}_m - \rho\boldsymbol{W}$, and assume that $\boldsymbol{A}_\rho$ is nonsingular. $\boldsymbol{\epsilon}_s \sim \mathrm{N}(\boldsymbol{0}, \sigma_s^2\boldsymbol{I}_m)$, and then $\boldsymbol{\delta}_s|\rho, \sigma_s^2 \sim \mathrm{N}\left[\boldsymbol{0}, \sigma_s^2(\boldsymbol{A}_\rho'\boldsymbol{A}_\rho)^{-1}\right]$. Therefore, the density of $\boldsymbol{\delta}_s$ is as follows:

$$\begin{aligned} p(\boldsymbol{\delta}_s|\rho, \sigma_s^2) &\propto |\sigma_s^2(\boldsymbol{A}_\rho{}'\boldsymbol{A}_\rho)^{-1}|^{-1/2} \exp\left\{-\frac{1}{2}\boldsymbol{\delta}_s' \left[\sigma_s^2(\boldsymbol{A}_\rho{}'\boldsymbol{A}_\rho)^{-1}\right]^{-1} \boldsymbol{\delta}_s\right\} \\ &\propto (\sigma_s^2)^{-m/2}|\boldsymbol{A}_\rho| \exp\left(-\frac{1}{2\sigma_s^2}\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s\right). \end{aligned}$$

Because $\boldsymbol{\xi} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{I}_n \otimes \boldsymbol{V})$, and $\boldsymbol{z} = \boldsymbol{X}\boldsymbol{\beta} + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta} + \boldsymbol{\xi}$, $\boldsymbol{z}$ has a multivariate normal distribution; that is, $\boldsymbol{z}|\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V} \sim \mathrm{N}\left[\boldsymbol{X}\boldsymbol{\beta} + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}, \boldsymbol{I}_n \otimes \boldsymbol{V}\right]$. As a result, the density of $\boldsymbol{z}$ is as follows:

$$\begin{aligned}
p(\boldsymbol{z}|\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}) &\propto |\boldsymbol{I}_n \otimes \boldsymbol{V}|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})'(\boldsymbol{I}_n \otimes \boldsymbol{V})^{-1}(\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})\right\} \\
&= |\boldsymbol{V}|^{-n/2} \exp\left\{-\frac{1}{2}(\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})\right\} \\
&= \prod_{i=1}^{m} \prod_{k=1}^{n_i} \left\{|\boldsymbol{V}|^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{z}_{ik} - \boldsymbol{X}_{ik}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik} - \boldsymbol{X}_{ik}\boldsymbol{\beta} - \boldsymbol{\delta}_i)\right]\right\}.
\end{aligned}$$

3.2 Parameter Estimation of MSOP Models

3.2.1 Prior Distributions of Parameters

To complete the Bayesian model, the following prior distributions of the parameters are introduced to the MSOP model:

$$\begin{aligned}
&\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{c}, \boldsymbol{T}), \\
&\boldsymbol{V} \sim \mathrm{IW}(n_0, \boldsymbol{Q}_0), \\
&\sigma_s^2 \sim \mathrm{IG}(e, g), \\
&\rho \sim \mathrm{U}(\lambda_{min}^{-1}, \lambda_{max}^{-1}),
\end{aligned}$$

where $\mathrm{IW}(n_0, \boldsymbol{Q}_0)$ denotes an Inverse-Wishart distribution with degree of freedom n_0 and $S \times S$ scale matrix $\boldsymbol{Q}_0$. $\mathrm{IG}(e, g)$ denotes an Inverse-Gamma distribution with shape e and scale g. λ_{min} and λ_{max} are the minimum and maximum eigenvalues of $\boldsymbol{W}$.

The prior density functions of the parameters are as follows:

$$\begin{aligned}
\pi(\boldsymbol{\beta}) &\propto \exp\left[-\frac{1}{2}(\boldsymbol{\beta}-\boldsymbol{c})'\boldsymbol{T}^{-1}(\boldsymbol{\beta}-\boldsymbol{c})\right], \\
\pi(\boldsymbol{V}) &\propto |\boldsymbol{V}|^{-\frac{n_0+S+1}{2}} \exp\left[-\frac{1}{2}\operatorname{tr}(\boldsymbol{Q}_0\boldsymbol{V}^{-1})\right], \\
\pi(\sigma_s^2) &\propto (\sigma_s^2)^{-(e+1)} \exp\left(-\frac{g}{\sigma_s^2}\right), \\
\pi(\rho) &\propto 1.
\end{aligned}$$

In addition, because $\boldsymbol{\gamma}$ has the limitation of $-\infty = \gamma_{s,0} < \gamma_{s,1} = 0 < \gamma_{s,2} < \cdots < \gamma_{s,L_s-1} = 1 < \gamma_{s,L_s} = \infty$, it is obvious that

$$\pi(\boldsymbol{\gamma}) \propto 1 \text{ for } \gamma_{s,2} < \cdots < \gamma_{s,L_s-2}, s = 1, \ldots, S.$$

3.2.2 MCMC Sampler of MSOP Models

The joint posterior density function of the parameters is as follows:

$$\begin{aligned}
p(\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}|\boldsymbol{y}) &\propto p(\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z})p(\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}) \\
&= p(\boldsymbol{y}|\boldsymbol{z}, \boldsymbol{\gamma})p(\boldsymbol{z}|\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V})p(\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2)\pi(\boldsymbol{\beta})\pi(\rho)\pi(\boldsymbol{\sigma}^2)\pi(\boldsymbol{V})\pi(\gamma),
\end{aligned}$$

where the probability of y_{iks} equals 1 when $\gamma_{s,l_s-1} < z_{iks} \leq \gamma_{s,l_s}$, and 0 otherwise. Given $\boldsymbol{z}$ and $\boldsymbol{\gamma}$, y_{ik} are conditionally i.i.d, and thus

$$p(\boldsymbol{y}|\boldsymbol{z}, \boldsymbol{\gamma}) = \prod_{i=1}^{m}\prod_{k=1}^{n_i} p(\boldsymbol{y}_{ik}|\gamma, \boldsymbol{z}_{ik}) = \prod_{i=1}^{m}\prod_{k=1}^{n_i} 1(\boldsymbol{z}_{ik} \in G_{ik}).$$

Before giving the parameter estimations, I must explain Cowles (1996)'s findings first. Cowles (1996) pointed out that the MCMC sampling of cutpoints $\boldsymbol{\gamma}$ conditioned on latent data $\boldsymbol{z}$ by using uniform distribution is not stable, and she suggested joint sampling of latent data $\boldsymbol{z}$ and cutpoints $\boldsymbol{\gamma}$. In this chapter, the conditional distributions of $\boldsymbol{\gamma}_s$ and $\boldsymbol{z}_s$ are obtained jointly.

Furthermore, the sample of $\boldsymbol{\gamma}$ can be obtained using a uniform distribution. Almost all the literature on the theme of SOP models used such an estimation method (Wang and Kockelman, 2009a,b,c). Following Cowles (1996), the subsequent studies (Albert and Chib, 2001; Chen and Dey, 2000) offered a better way to obtain parameters $\boldsymbol{\gamma}$ to remove the ordering constraint using the one-to-one map. In this chapter, the sampling of $\boldsymbol{\gamma}$ uses the transformation as in Chen and Dey (2000).

The parameters are estimated using a set of full conditional distributions (FCDs). Based on the joint posterior density function and prior distributions, the Appendix A.1 presents details of FCDs of all the parameters. Then, the parameters are estimated by MCMC sampling as follows:

1. Sample $\boldsymbol{\beta}|\boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{y} \sim \mathrm{N}(\boldsymbol{B}^{-1}\boldsymbol{b}, \boldsymbol{B}^{-1})$, where $\boldsymbol{B} = \boldsymbol{X}'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})\boldsymbol{X} + \boldsymbol{T}^{-1}$, and $\boldsymbol{b} = \boldsymbol{X}'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}) + \boldsymbol{T}^{-1}\boldsymbol{c}$.

2. Sample $\boldsymbol{\delta}|\boldsymbol{\beta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{y} \sim \mathrm{N}(\boldsymbol{\Omega}^{-1}\boldsymbol{\eta}, \boldsymbol{\Omega}^{-1})$, where $\boldsymbol{\Omega} = (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{\Delta} \otimes \boldsymbol{I}_S) + \boldsymbol{B}_\rho'(\boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho$, and $\boldsymbol{\eta} = (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta})$.

3. Sample $\rho|\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{z}, \boldsymbol{\gamma}, \boldsymbol{y}$ by using a Metropolis-Hastings (M-H) algorithm as explained in the Appendix A.1.3.

4. Sample $\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g}{\sigma_s^2}|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2_{-s}, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{y} \sim \chi^2(m + 2e)$, where $\boldsymbol{\sigma}^2_{-s} = (\sigma_1^2, \dots, \sigma_{s-1}^2, \sigma_{s+1}^2, \dots, \sigma_S^2)'$.

5. Sample $\boldsymbol{V}|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{y} \sim \mathrm{IW}(\tilde{n}, \tilde{\boldsymbol{Q}})$, where $\tilde{n} = n + n_0$, and $\tilde{\boldsymbol{Q}} = \sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ik} - \boldsymbol{X}_{ik}\boldsymbol{\beta} - \boldsymbol{\delta}_i)(\boldsymbol{z}_{ik} - \boldsymbol{X}_{ik}\boldsymbol{\beta} - \boldsymbol{\delta}_i)' + \boldsymbol{Q}_0$.

6. Let $\boldsymbol{z}_s = (z_{11s}, \ldots, z_{1n_is}, \ldots, z_{m1s}, \ldots, z_{mn_ms})'$ denotes the vector of the sth element z_{iks} from $\boldsymbol{z}_{ik}(i = 1, \ldots, m; k = 1, \ldots, n_i)$, and let $\boldsymbol{z}_{-s}$ denotes the vector removing $\boldsymbol{z}_s$ from $\boldsymbol{z}$. Similarly, let $z_{ik,-s}$ denotes the vector removing z_{iks} from $\boldsymbol{z}_{ik}$. Furthermore, defining $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_s, \ldots, \boldsymbol{\gamma}_S)'$, $\boldsymbol{\gamma}_s = (\gamma_{s,2}, \ldots, \gamma_{s,Ls-2})'$, and thus $\boldsymbol{\gamma}_{-s}$ denotes the vector removing $\boldsymbol{\gamma}_s$ from $\boldsymbol{\gamma}$. For $s = 1, \ldots, S$, sample $\boldsymbol{\gamma}_s, \boldsymbol{z}_s | \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}_{-s}, \boldsymbol{z}_{-s}, \boldsymbol{y}$ as follows:

    (a) Sample $z_{iks} | \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}_{-s}, \boldsymbol{y} \sim \mathrm{N}(\tilde{h}_{iks}, \tilde{v}_{iks}) 1(\boldsymbol{z}_{ik} \in G_{ik})$, with $\tilde{h}_{iks} = h_{iks} + \boldsymbol{v}_{s,-s} \boldsymbol{V}_{-s,-s}^{-1} (\boldsymbol{z}_{ik,-s} - \boldsymbol{h}_{ik,-s})$, and $\tilde{v}_{iks} = V_{s,s} - \boldsymbol{v}_{s,-s} \boldsymbol{V}_{-s,-s}^{-1} \boldsymbol{v}_{-s,s}$, where $\boldsymbol{v}_{s,-s} = (V_{s,1}, \ldots, V_{s,s-1}, V_{s,s+1}, \ldots, V_{s,S})$, $\boldsymbol{v}_{-s,s} = (V_{1,s}, \ldots, V_{s-1,s}, V_{s+1,s}, \ldots, V_{S,s})'$, and $\boldsymbol{V}_{-s,-s}$ denotes an $(S-1) \times (S-1)$ matrix removing the elements $v_{s,s}$, $\boldsymbol{v}_{-s,s}$ and $\boldsymbol{v}_{s,-s}$ from $\boldsymbol{V}$. I define $\boldsymbol{h}_{ik} = \boldsymbol{X}_{ik}\boldsymbol{\beta} + \boldsymbol{\delta}_i$, and thus $\boldsymbol{h}_{ik,-s}$ denotes the vector removing h_{iks} from $\boldsymbol{h}_{ik}$.

    (b) Sample γ_{s,l_s} by using the transformations as $\gamma_{s,l_s} = \frac{\gamma_{s,l_s-1} + \exp(\kappa_{s,l_s})}{1+\exp(\kappa_{s,l_s})}$ for $2 < l_s < L_s - 2$, where the samples of κ_{s,l_s} are generated by using a M-H algorithm as explained in the Appendix A.1.7.

3.3 Model Simulation of MSOP Models

3.3.1 Simulated Dataset

A simulation study is implemented to demonstrate the validity and accuracy of the model. The number of response variables is set as $S = 2$ (y_1 and y_2) for ease of understanding and explanation.

In the simulated dataset, bell-shaped frequency distributions are considered because the latent variables $\boldsymbol{z}$ are normally distributed; that is, the worst and best situations have lower frequencies, and the general situation has the highest frequency. I assumed that there are $L_1 = L_2 = 5$ ordered categories for the two response variables. Therefore, the frequencies for $y_1, y_2 = 1$ and $y_1, y_2 = 5$ should be the lowest, and those for $y_1, y_2 = 3$ should be the highest. In this chapter, the parameters are set to obtain such suitable frequency tables using the following reparameterization process:

$$y_{iks} = l_s \text{ if } \gamma^*_{s,l_s-1} < z^*_{iks} \leq \gamma^*_{s,l_s}, \text{ for } s = 1,2, \text{and } l_s = 1, \dots, 5, \tag{3.21}$$

where $-\infty = \gamma^*_{s,0} < \gamma^*_{s,1} = 0 < \gamma^*_{s,2} < \gamma^*_{s,3} < \gamma^*_{s,4} < \gamma^*_{s,5} = \infty$. I assume that

$$z^*_{ik} \sim \mathrm{N}\left(\begin{pmatrix} \boldsymbol{x}'_{ik1}\boldsymbol{\beta}^*_1 + \delta^*_{i1} \\ \boldsymbol{x}'_{ik2}\boldsymbol{\beta}^*_2 + \delta^*_{i2} \end{pmatrix}, \boldsymbol{V}^*\right), \tag{3.22}$$

where

$$\boldsymbol{V}^* = \begin{pmatrix} 1 & v^* \\ v^* & 1 \end{pmatrix}.$$

Further, the parameters are reparameterized as:

$$\tau_s = \frac{1}{\gamma^*_{s,4}}, \ \gamma_{s,l_s} = \tau_s \gamma^*_{s,l_s}, \ \beta_s = \tau_s \beta^*_s, \ \delta_s = \tau_s \delta^*_s, \ z_{iks} = \tau_s z^*_{iks}, l_s = 2,3; s = 1,2. \tag{3.23}$$

Based on the above model, the dataset is generated using the following steps:

1. Set the values of $\gamma^*_{s,2}$, $\gamma^*_{s,3}$, $\gamma^*_{s,4}$, v^*, δ^* and $\boldsymbol{\beta}_s$. From $\boldsymbol{\beta}_s$ and $\gamma^*_{s,4}$, $\boldsymbol{\beta}^*_s$ can be derived from (3.23).

2. Generate $\boldsymbol{z}^*_{ik}$ from (3.22).

3. Generate y_{iks} from $\boldsymbol{z}^*_{ik}$ using (3.21).

4. Calculate the frequency table of y_{iks}. If the frequency table is not suitable, try

procedures 1 to 3 again for other values of $\gamma^*_{s,2}$, $\gamma^*_{s,3}$, $\gamma^*_{s,4}$, v^*, δ^*, and $\boldsymbol{\beta}_s$. If the frequency table is suitable, derive the values of z_{iks}, $\gamma_{s,2}$ and $\gamma_{s,3}$ using (3.23). Further, derive

$$\boldsymbol{V} = \begin{pmatrix} \tau_1^2 & \tau_1\tau_2 v^* \\ \tau_1\tau_2 v^* & \tau_2^2 \end{pmatrix}.$$

In the simulated dataset, the number of region m is set as 6, 10, 15, and 20, and that of individuals in each region n_i is set as 100, 200, and 300 in different experiments. First, I generate a suitable dataset for the case of $m = 20$ and $n_i = 300$ (the largest case). Then, I proceed to generate the dataset for the case of small m and n_i. For example, for the case of $m = 20$ and $n_i = 100$, I randomly select the observations of size $n_i = 100$ from the observations of size $n_i = 300$ for each region. Further, for the case of $m = 6$ and $n_i = 100$, I first randomly select 6 regions among 20 regions, and then I randomly select the observations of size $n_i = 100$ from the observations of size $n_i = 300$ for each selected region.

The 20 regions are assumed to be located as shown in Figure 3.1. The weight matrix is generated based on queen contiguity. For example, region 15 is contiguous with regions 8, 9, 10, 14, 16, and 20. It is then row-standardized. For the case of $m = 6$, 10, and 15, the corresponding weight matrix is obtained based on the selected region.

There are $S = 2$ response variables, and each response variable has $q_s = 2$ explanatory variables. They are generated using a standard uniform distribution [$\mathrm{U}(0, 1)$]. Note that the simulation does not include the intercept. The explanatory variables are set to be different for the two response variables. The vector of individualistic effects $\boldsymbol{\delta}$ is generated using the multivariate normal distribution $\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2 \sim \mathrm{N}\left[\boldsymbol{0}, \boldsymbol{B}_\rho^{-1}(\boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2))(\boldsymbol{B}_\rho^{-1})'\right]$, where $\boldsymbol{B}_\rho = \boldsymbol{I}_{ms} - \rho(\boldsymbol{W} \otimes \boldsymbol{I}_S)$, and spatial coefficient ρ is set as 0.3, 0.5, and 0.7 in

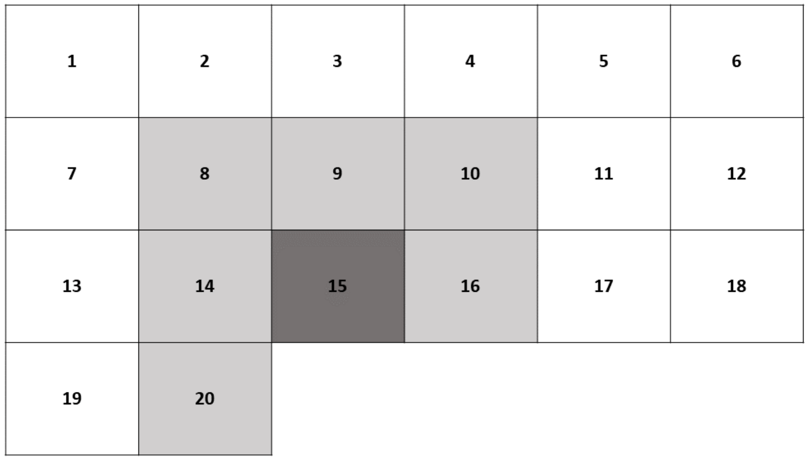


Figure 3.1: Location of regions in the simulated dataset for the MSOP model

different experiments. $\boldsymbol{\sigma}^2$ is set the same in different experiments as $(0.02, 0.01)'$.

To generate a suitable dataset, adjustment parameters τ_1 and τ_2 are finally set as $\tau_1 = \tau_2 = \sqrt{0.1}$. Since there are two response variables, the setting of covariance v^* is extremely important: the greater is the covariance, the stronger is the interaction of the response variables. To verify the stability of the model, v^* is set as 0.9 (strong interaction), 0.5 (moderate interaction), and 0.1 (weak interaction) in different experiments. Thus, covariance matrix $\boldsymbol{V}$ is set as $\boldsymbol{V} = \begin{pmatrix} 0.10 & 0.09 \\ 0.09 & 0.10 \end{pmatrix}$, $\boldsymbol{V} = \begin{pmatrix} 0.10 & 0.05 \\ 0.05 & 0.10 \end{pmatrix}$, and $\boldsymbol{V} = \begin{pmatrix} 0.10 & 0.01 \\ 0.01 & 0.10 \end{pmatrix}$ in different experiments.

For the setting of $v^* = 0.9$, the regression parameters are set as $\boldsymbol{\beta} = (0.55, 0.35, 0.41, 0.56)'$ and $\boldsymbol{\beta} = (0.40, 0.49, 0.45, 0.52)'$ in different experiments. For the setting of $v^* = 0.5$, the regression parameters are set as $\boldsymbol{\beta} = (0.56, 0.38, 0.46, 0.56)'$ and $\boldsymbol{\beta} = (0.46, 0.49, 0.49, 0.53)'$

in different experiments. For the setting of $v^* = 0.1$, the regression parameters are set as $\boldsymbol{\beta} = (0.58, 0.36, 0.45, 0.56)'$ and $\boldsymbol{\beta} = (0.45, 0.49, 0.49, 0.52)'$ in different experiments. In addition, the cutpoints are set as the same in different experiments as follows: $-\infty = \gamma_{1,0} < \gamma_{1,1} = 0 < \gamma_{1,2} = 0.31 < \gamma_{1,3} = 0.66 < \gamma_{1,4} = 1 < \gamma_{1,5} = \infty$, and $-\infty = \gamma_{2,0} < \gamma_{2,1} = 0 < \gamma_{2,2} = 0.33 < \gamma_{2,3} = 0.68 < \gamma_{2,4} = 1 < \gamma_{2,5} = \infty$.

For each setting of v^*, there are four settings for the number of region m, three settings for the individuals in each region n_i, two settings for regression parameters $\boldsymbol{\beta}$, and three settings for spatial coefficient ρ, and thus there are $4 \times 3 \times 2 \times 3 = 72$ experiments for each setting of v^*. The frequency tables for $v^* = 0.9, 0, 5$, and 0.1 in this simulation study are shown in Tables A.1, A.2, and A.3, respectively.

3.3.2 Simulation Results

In the simulation, the following diffuse priors are used:

$$
\begin{aligned}
&\boldsymbol{\beta} \sim \mathrm{N}(\mathbf{1}, 10^{12}\boldsymbol{I}),\\
&\boldsymbol{V} \sim \mathrm{IW}(1, \boldsymbol{Q}_0), \boldsymbol{Q}_0 = \begin{pmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{pmatrix},\\
&\sigma_1^2 = \sigma_2^2 \sim \mathrm{IG}(0, 0),\\
&\rho \sim \mathrm{U}(\lambda_{min}^{-1}, \lambda_{max}^{-1}),\\
&(\lambda_{min} \text{ and } \lambda_{max} \text{ are the minimum and maximum eigenvalues for different } \boldsymbol{W}),\\
&\boldsymbol{\kappa}_1 \sim \mathrm{N}(\mathbf{0}_{2\times 1}, 100\boldsymbol{I}_{2\times 2}),\\
&\boldsymbol{\kappa}_2 \sim \mathrm{N}(\mathbf{0}_{2\times 1}, 100\boldsymbol{I}_{2\times 2}),
\end{aligned}
$$

and the posterior results are calculated using MCMC sampling.

When using the MCMC method, the number of draws should be set as high as possible to get the convergent value of the parameters. As the number increases, the calculation time and memory requirements also increase. In addition, when the sample is large, the calculation speed must be a big problem. In this empirical study, there are $20 * 30 = 6,000$ observations in the simulated datasets. Therefore, the parameters are assumed convergent when sampled parameter distributions appear to stabilize.

The simulation results of $3,000$ and $11,000$ iterations have been compared, and the first $1,000$ samples are discarded as the burn-in period. The sampled parameter distributions appear to stabilize after $1,000$ iterations, and there is no obvious differences in simulation results between these two settings of iteration. Finally, each MCMC simulation is run for $3,000$ iterations, and the first $1,000$ samples are discarded as the burn-in period. Thus, the parameters are calculated based on the final $2,000$ runs.

Table 3.1 gives partial simulation results of posterior means and true values as well as the mean squared error (MSE) of $\boldsymbol{\beta}$ and ρ for the setting of $v^* = 0.9$, $v^* = 0.5$, and $v^* = 0.1$, respectively. The supplemental materials Tables A.4, A.5, and A.6 provide the complete results of 72 simulation experiments.

The MSE of $\boldsymbol{\beta}$ and ρ are calculated as follows:

$$MSE_{\boldsymbol{\beta}} = \frac{1}{2000} \sum_{t=1}^{2000} (\boldsymbol{\beta}^{(t)} - \boldsymbol{\beta}_{true})'(\boldsymbol{\beta}^{(t)} - \boldsymbol{\beta}_{true}), \tag{3.24}$$

$$MSE_{\rho} = \frac{1}{2000} \sum_{min}^{max} (\rho^{(t)} - \rho_{true})^2, \tag{3.25}$$

where $\boldsymbol{\beta}^{(t)}$ and $\rho^{(t)}$ represent the final tth simulation results.

From Table 3.1, for both settings of $\boldsymbol{\beta}_{true}$, no matter how large ρ_{true} is, the $MSE_{\boldsymbol{\beta}}$ are all very small. Furthermore, for all settings of $\boldsymbol{\beta}_{true}$ and ρ_{true}, the $MSE_{\boldsymbol{\beta}}$ and

Table 3.1: Partial simulation results of the MSOP models in the case of $v^* = 0.9$, $v^* = 0.5$ and $v^* = 0.1$

			simulation results						
ρ_{true}	m	n_i	β_{11}	β_{12}	β_{21}	β_{22}	MSE_β	ρ	MSE_ρ
			$v^* = 0.9$, $\boldsymbol{\beta}_{true} = (0.55, 0.35, 0.41, 0.56)'$						
0.3	6	100	0.5224	0.4074	0.4330	0.5464	0.0077	0.2850	0.0570
		200	0.5363	0.3848	0.4101	0.5544	0.0030	0.3379	0.0391
		300	0.5556	0.3698	0.4102	0.5700	0.0015	0.3549	0.0433
	10	100	0.5468	0.3806	0.4701	0.5422	0.0068	0.2805	0.0576
		200	0.5334	0.3570	0.4100	0.5543	0.0012	0.2476	0.0577
		300	0.5405	0.3485	0.4114	0.5593	0.0007	0.2541	0.0534
	15	100	0.5392	0.3560	0.4450	0.5489	0.0028	0.2835	0.0517
		200	0.5306	0.3426	0.4123	0.5615	0.0011	0.3528	0.0464
		300	0.5425	0.3459	0.4076	0.5733	0.0007	0.2976	0.0467
	20	100	0.5421	0.3650	0.4325	0.5524	0.0018	0.2606	0.0518
		200	0.5374	0.3518	0.4102	0.5577	0.0007	0.3625	0.0446
		300	0.5484	0.3495	0.4076	0.5688	0.0004	0.3408	0.0426
0.5	6	100	0.5173	0.3932	0.3902	0.5233	0.0079	0.3397	0.0758
0.7	6	100	0.5204	0.3725	0.3933	0.5699	0.0051	0.5020	0.0688
	20	100	0.5226	0.3401	0.4309	0.5747	0.0025	0.6561	0.0181
		300	0.5419	0.3357	0.4177	0.5784	0.0010	0.6819	0.0147
			$v^* = 0.9$, $\boldsymbol{\beta}_{true} = (0.40, 0.49, 0.45, 0.52)'$						
0.3	6	100	0.3774	0.5142	0.4674	0.4842	0.0059	0.2647	0.0532
0.5	6	100	0.3820	0.5134	0.4418	0.4992	0.0045	0.3141	0.0849
0.7	6	100	0.3595	0.5177	0.4399	0.5051	0.0059	0.4538	0.0952
	20	100	0.3699	0.4778	0.4675	0.5246	0.0023	0.6438	0.0200
		300	0.3878	0.4743	0.4564	0.5271	0.0008	0.6661	0.0165
			$v^* = 0.5$, $\boldsymbol{\beta}_{true} = (0.56, 0.38, 0.46, 0.56)'$						
0.3	6	100	0.5691	0.3423	0.4767	0.5191	0.0100	0.2808	0.0446
0.5	6	100	0.5609	0.3423	0.4740	0.5247	0.0097	0.3448	0.0671
0.7	6	100	0.5477	0.3543	0.4678	0.5462	0.0087	0.4361	0.1037
	20	100	0.5545	0.3909	0.4715	0.4886	0.0075	0.6905	0.0134
		300	0.5712	0.4003	0.4617	0.5314	0.0021	0.6503	0.0193
			$v^* = 0.5$, $\boldsymbol{\beta}_{true} = (0.46, 0.49, 0.49, 0.53)'$						
0.3	6	100	0.4354	0.4450	0.5231	0.4935	0.0119	0.2603	0.0481
0.5	6	100	0.4388	0.4711	0.4850	0.4935	0.0089	0.3436	0.0742
0.7	6	100	0.4339	0.4731	0.4960	0.5017	0.0087	0.4350	0.1122
	20	100	0.4455	0.5016	0.5031	0.4585	0.0078	0.6463	0.0236
		300	0.4613	0.5135	0.4912	0.5076	0.0018	0.6304	0.0269
			$v^* = 0.1$, $\boldsymbol{\beta}_{true} = (0.58, 0.36, 0.45, 0.56)'$						
0.3	6	100	0.6256	0.3614	0.5223	0.5989	0.0176	0.2930	0.0508
0.5	6	100	0.6047	0.3694	0.5343	0.6071	0.0194	0.3906	0.0560
0.7	6	100	0.5980	0.3598	0.5235	0.5998	0.0159	0.4736	0.0893
	20	100	0.5852	0.3357	0.4403	0.5660	0.0035	0.6380	0.0253
		300	0.5859	0.3485	0.4386	0.5731	0.0014	0.6543	0.0161
			$v^* = 0.1$, $\boldsymbol{\beta}_{true} = (0.45, 0.49, 0.49, 0.52)'$						
0.3	6	100	0.4772	0.5168	0.5606	0.5696	0.0178	0.3092	0.0559
0.5	6	100	0.4776	0.5123	0.5748	0.5762	0.0205	0.3652	0.0689
0.7	6	100	0.4890	0.4884	0.5521	0.5767	0.0172	0.4757	0.0867
	20	100	0.4403	0.4656	0.4778	0.5258	0.0036	0.6482	0.0243
		300	0.4421	0.4819	0.4733	0.5302	0.0014	0.6365	0.0211

MSE_ρ are the smallest for the case of $m = 20$. In addition, for the same m, the larger is the n_i, the smaller is the $MSE_{\boldsymbol{\beta}}$. For most cases, the MSE_ρ tend to be small with the increase in observations n_i. That is, the larger is the region and the more are the observations, the higher is the accuracy of the estimate.

But the posterior mean of ρ deviates from its true value by a large value when ρ equals 0.5 or 0.7; that is, an increase in the spatial coefficient ρ lends to greater deviation. Wang and Kockelman (2009c) explained for the deviation as follows: when positive spatial correlation exists, but is not fully recognized, the coefficients tend to be more biased because areas with higher response magnitudes will have a greater impact on model estimates .

All in all, considering the magnitudes of the parameter values, the estimation results are quite close to the true values. It satisfactorily explains the spatial dependencies and interaction of multivariate response variables as well as the effects of the explanatory variables.

3.4 Empirical Study

3.4.1 Empirical Data

The data for this empirical study are taken from the 2010 Chinese Urban and Rural Elderly Population Surveys, conducted by the China Research Center on Aging of the National Committee on Aging (Gao and Li, 2016). This survey is a household survey, and interviews were conducted by interviewers, who then completed the questionnaires on behalf of the interviewees based on their responses. No questionnaires were completed by the interviewees. Then, the interviewers checked and verified the responses after the investigation. As a result, the data accuracy is high.

Figure 3.2: The map of China mainland (31 provinces, municipalities and autonomous regions)

In this empirical study, the data of six provinces in East China (Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, and Shandong) is used, as shown in Figure 3.2. After deleting the missing data, there are $5,916$ valid survey samples. In addition, the weights w_{ij} in this empirical study are quantified based on contiguity, and the contiguity is defined by queen contiguity. From Figure 3.2, the normalized weight matrix can be easily obtained (row sum is 1) as follows:

$$\boldsymbol{W} = \begin{array}{cc} & \begin{array}{cccccc} Shanghai & Jiangsu & Zhejiang & Fujian & Anhui & Shandong \end{array} \\ \begin{array}{c} Shanghai \\ Jiangsu \\ Zhejiang \\ Fujian \\ Anhui \\ Shandong \end{array} & \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 1/4 & 0 & 1/4 & 0 & 1/4 & 1/4 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 0 & 0 & 1/3 \\ 0 & 1/2 & 0 & 0 & 1/2 & 0 \end{pmatrix} \end{array}.$$

Table 3.2 explains all the variables used in the MSOP model. In the model, the response variables are self-rated health and life satisfaction, and the explanatory variables include gender, age, education (education level), marriage (marital status), census register (urban and rural), chronic disease, and need for care. Here, age is a continuous variable, which has been changed as [age-min(age)] /[max(age)-min(age)] (in this empirical study, max(age)=100 and min(age)=60), and thus the range of age is from 0 to 1. The other explanatory variables are discrete variables.

Table 3.2: Variables in the empirical MSOP model

Variable	Description	Value
Dependent Variables		
Self-rated Health	Self-evaluation of the health condition.	Very poor is 1; poor is 2; general is 3; good is 4; and very good is 5.
Life Satisfaction	Satisfaction with life of the elderly people.	Very dissatisfied is 1; dissatisfied is 2; general is 3; satisfied is 4; and very satisfied is 5.
Explanatory variables		
Gender	Gender of the elderly person with categories 'male' and 'female'.	Male is 1; female is 0.
Age	Age of the elderly people in 2010.	Change as: (age-min(age)) /(max(age)-min(age))
Education	One's education level.	Junior high school and above is 1; else is 0.
Marriage	Marital status of the elderly people.	Live with a spouse is 1; else is 0.
Census Register	Census register with categories 'urban' and 'rural'.	Rural is 1; urban is 0.
Chronic Disease	Whether one has a chronic disease or not.	Having a chronic disease is 1, else is 0.
Need for Care	Whether one needs daily care from others or not.	Need daily care is 1; don't need is 0.

3.4.2 Posterior Results of Parameters

In the empirical study, the priors are assigned the same as in the simulation part, except $\rho \sim \mathrm{U}(-0.9003, 0.7982)$, because the minimum and maximum eigenvalues of above $\boldsymbol{W}$ have changed to -1.1107 and 1.2528. MCMC simulations are run for $5,000$ iterations, and the first $1,000$ samples are discarded as the burn-in period.

Table 3.3: Posterior Estimates of the empirical MSOP model

Parameter		Self-rated Health mean	sd	median		Life Satisfaction mean	sd	median
$\boldsymbol{\beta}$	β_{11} (Gender)	0.0207	0.0077	0.0206**	β_{12} (Gender)	-0.0139	0.0093	-0.0138
	β_{21} (Age)	-0.1254	0.0209	-0.1249**	β_{22} (Age)	0.1733	0.0256	0.1728**
	β_{31} (Education)	0.0389	0.0091	0.0391**	β_{32} (Education)	0.0606	0.0113	0.0606**
	β_{41} (Marriage)	0.0115	0.0082	0.0116	β_{42}(Marriage)	0.0647	0.0106	0.0647**
	β_{51} (CR)	-0.0436	0.0085	-0.0437**	β_{52} (CR)	-0.0830	0.0102	-0.0831**
	β_{61} (CD)	-0.1893	0.0085	-0.1892**	β_{62} (CD)	-0.0201	0.0108	-0.0201*
	β_{71} (NC)	-0.2443	0.0110	-0.2441**	β_{72} (NC)	-0.0509	0.0133	-0.0508**
$\boldsymbol{\delta}$	δ_{11} (Shanghai)	0.6992	0.0156	0.6990**	δ_{12} (Shanghai)	0.5005	0.0198	0.5010**
	δ_{21} (Jiangsu)	0.7284	0.0159	0.7280**	δ_{22} (Jiangsu)	0.5833	0.0202	0.5833**
	δ_{31} (Zhejiang)	0.7460	0.0159	0.7460**	δ_{32} (Zhejiang)	0.5543	0.0203	0.5546**
	δ_{41} (Anhui)	0.6247	0.0159	0.6244**	δ_{42} (Anhui)	0.4406	0.0198	0.4405**
	δ_{51} (Fujian)	0.6492	0.0157	0.6487**	δ_{52} (Fujian)	0.5224	0.0203	0.5226**
	δ_{61} (Shandong)	0.7157	0.0159	0.7155**	δ_{62} (Shandong)	0.5401	0.0201	0.5400**
ρ		0.7646	0.0399	0.7752**				
$\boldsymbol{\sigma}^2$	σ_1^2	0.0398	0.0395	0.0288**	σ_2^2	0.0223	0.0232	0.0160**
$\boldsymbol{V}$	$\boldsymbol{V}_{1,1}$	0.0622	0.0014	0.0621**	$\boldsymbol{V}_{2,2}$	0.0976	0.0022	0.0976**
	$\boldsymbol{V}_{1,2}$	0.0224	0.0012	0.0224**	$\boldsymbol{V}_{2,1}$	0.0224	0.0012	0.0224**
$\boldsymbol{\gamma}$	$\boldsymbol{\gamma}_{1,2}$	0.2718	0.0057	0.2716**	$\boldsymbol{\gamma}_{2,2}$	0.1980	0.0063	0.1979**
	$\boldsymbol{\gamma}_{1,3}$	0.6747	0.0058	0.6746**	$\boldsymbol{\gamma}_{2,3}$	0.5375	0.0061	0.5375**

[1] CR, CD and NC are the abbreviation of self-rated health, Census Register, Chronic Disease and Need for Care, respectively.
[2] "$**$" and "$*$" denote that zero is not included in the 95% and 90% credible intervals, respectively. Mean, Sd and Median are posterior mean, posterior standard deviation and posterior median, respectively.

Table 3.3 gives the posterior means, posterior standard deviations, and posterior medians of all estimated parameters. From Table 3.3, all the explanatory variables except marriage have evident effects on self-rated health. In detail, the regression parameters of gender and education are positive. Therefore, self-rated health in elderly males and people with high education level is more likely to be higher than that in elderly females and people with low education level, respectively. The regression parameters of age, census register, chronic disease, and need for care are negative. Therefore, self-rated health in younger elderly people, urban elderly people, people without chronic disease, and people without need for care is more likely to be higher than that in older elderly people, rural elderly people, people with chronic disease, and people with need for care, respectively.

However, for life satisfaction, all the explanatory variables except gender have evident effects on it. In detail, the regression parameters of age, education, and marriage are positive. Thus, life satisfaction in older elderly people is more likely to be higher than that in younger elderly people. Life satisfaction in people with high education level and people living with a spouse is more likely to be higher than that in people with low education level and people in other marital statuses, respectively. The regression parameters of census register, chronic disease, and need for care are negative. Thus, life satisfaction in urban elderly people, people without chronic disease, and people without the need for care is more likely to be higher than that in rural elderly people, people with chronic disease, and people with the need for care, respectively.

In addition, Table 3.3 shows that the five explanatory variables of age, education, census register, chronic disease, and need for care have evident effects on both response variables, but the directions are different among them. Age has a negative effect on self-rated health, but it has a positive effect on life satisfaction; education has a positive effect on both response variables; and census register, chronic disease, and need for care have negative effects on both response variables. In addition, gender only has an evident (positive) effect on self-rated health, and marriage only has an evident (positive) effect on life satisfaction.

In this empirical study, $\boldsymbol{\delta}$ is a 12 ($m = 6, S = 2$)-dimensional vector of the regional effects. It can be separated into two sub-vectors, $\boldsymbol{\delta}_{i1}$ and $\boldsymbol{\delta}_{i2}$. $\boldsymbol{\delta}_{i1}$ and $\boldsymbol{\delta}_{i2}$ represent all unobserved common regional effects for the observations within region $i\,(i = 1, \ldots, 6)$, that is, for self-rated health and life satisfaction, respectively. From Table 3.3, the values of vector $\boldsymbol{\delta}_{i1}$ and $\boldsymbol{\delta}_{i2}$ are all relatively large; that is, the regional effects are indispensable on the response variables. For self-rated health, Zhejiang province's regional effect is the greatest, while for life satisfaction, Jiangsu province's regional effect is the greatest. In addition, the regional effects of self-rated health are always greater than those of life satisfaction in the same region.

The posterior mean of spatial coefficient ρ is 0.7646. It is large and positive, and thus the spatial dependencies in this empirical study are critical. Furthermore, one province's regional effect has a positive effect on the other provinces' regional effects.

$\boldsymbol{V}_{1,1}$ denotes the variance of the individualistic effect for self-rated health; $\boldsymbol{V}_{1,2}$ and $\boldsymbol{V}_{1,2}$ denote the covariance of the individualistic effect for self-rated health and life satisfaction; and $\boldsymbol{V}_{2,2}$ denotes the variance of the individualistic effect for life satisfaction. Since $\boldsymbol{V}_{1,2}$ and $\boldsymbol{V}_{1,2}$ are evident and positive, the individualistic effects for self-rated health and life satisfaction are positively related.

$\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ represents the vector of estimated cutpoints for self-rated health and life satisfaction, respectively. For examples, if $0 < z_{ik1} \leq 0.2718$, self-rated health is "poor"; and if $0.5375 < z_{ik2} \leq 1$, life satisfaction is "satisfied". The cutpoints are critical to obtain accurate parameter estimates.

3.4.3 Posterior Results of Marginal Effects

Table 3.4 lists the marginal effects of all explanatory variables for self-rated health and life satisfaction. The derivations of marginal effects for continuous explanatory variables and discrete explanatory variables are given in Appendix A.2.

From Table 3.4, corresponding to the posterior results of self-rated health in Table 3.3, for self-rated health, the marginal effects of all the explanatory variables except marriage are evident; for life satisfaction, the marginal effects of all the explanatory variables except gender are evident. In detail, the marginal effects of gender are evident negative when "self-rated health $= 1, 2, 3$" and evident positive when "self-rated health $= 4, 5$." That is, self-rated health in elderly males is more likely to be "good" or "very good." The marginal effects of marriage are evident negative when "life satisfaction $= 1, 2, 3$" and evident positive when "life satisfaction $= 4, 5$." That is, life satisfaction in people living with a spouse is more likely to be "satisfied" or "very satisfied."

The marginal effects of age, education, census register, chronic disease, and need for care are evident on both response variables. Besides age, the direction of marginal effects of the other four variables on both response variables are consistent. In detail, the marginal effects of age are positive when "self-rated health $= 1, 2, 3$" & "life satisfaction $= 4, 5$," and negative when "self-rated health $= 4, 5$" & "life satisfaction $= 1, 2, 3$." Therefore, self-rated health in younger elderly people is more likely to be "good" or "very good," whereas life satisfaction in older elderly people is more likely to be "satisfied" or "very satisfied." The marginal effects of education are negative when "self-rated health $= 1, 2, 3$" & "life satisfaction $= 1, 2, 3$," and positive when "self-rated health $= 4, 5$" & "life satisfaction $= 4, 5$." That is, in people with high education level, self-rated health is more likely to be "good" or "very good," and life satisfaction is more likely to be

Table 3.4: Marginal effects of the explanatory variables in the empirical MSOP model

		Gender	Age	Education	Marriage	CR	CD	NC
	mean	-0.0182	0.0264	-0.0328	-0.0102	0.0383	0.1217	0.1961
SRH=1	sd	0.0069	0.0044	0.0074	0.0073	0.0075	0.0058	0.0121
	median	-0.0181**	0.0264**	-0.0329**	-0.0102	0.0383**	0.1216**	0.1959**
	mean	-0.0147	0.0213	-0.0283	-0.0082	0.0311	0.1388	0.1850
SRH=2	sd	0.0055	0.0036	0.0068	0.0059	0.0061	0.0072	0.0082
	median	-0.0146**	0.0212**	-0.0284**	-0.0083	0.0311**	0.1386**	0.1849**
	mean	-0.0021	0.0030	-0.0054	-0.0009	0.0036	0.0830	-0.1261
SRH=3	sd	0.0011	0.0009	0.0022	0.0008	0.0017	0.0083	0.0125
	median	-0.0019**	0.0029**	-0.0051**	-0.0008	0.0035**	0.0829**	-0.1258**
	mean	0.0195	-0.0281	0.0365	0.0109	-0.0407	-0.1842	-0.2132
SRH=4	sd	0.0073	0.0046	0.0086	0.0078	0.0079	0.0084	0.0092
	median	0.0194**	-0.0280**	0.0366**	0.0110	-0.0406**	-0.1840**	-0.2132**
	mean	0.0131	-0.0192	0.0255	0.0072	-0.0279	-0.1124	-0.0846
SRH=5	sd	0.0049	0.0035	0.0062	0.0051	0.0055	0.0074	0.0046
	median	0.0130**	-0.0191**	0.0254**	0.0073	-0.0279**	-0.1122**	-0.0844**
	mean	0.0046	-0.0172	-0.0186	-0.0222	0.0272	0.0064	0.0184
LS=1	sd	0.0031	0.0027	0.0034	0.0039	0.0035	0.0034	0.0053
	median	0.00454	-0.0171**	-0.0186**	-0.0222**	0.0272**	0.0065*	0.0182**
	mean	0.0061	-0.0228	-0.0262	-0.0286	0.0362	0.0088	0.0226
LS=2	sd	0.0041	0.0035	0.0048	0.0048	0.0047	0.0047	0.0060
	median	0.00604	-0.0228**	-0.0261**	-0.0285**	0.0362**	0.0088*	0.0227**
	mean	0.0075	-0.0280	-0.0351	-0.0334	0.0463	0.0111	0.0249
LS=3	sd	0.0050	0.0042	0.0070	0.0054	0.0059	0.0061	0.0060
	median	0.00754	-0.0279**	-0.0350**	-0.0334**	0.0464**	0.0111*	0.0250**
	mean	-0.0091	0.0342	0.0388	0.0453	-0.0571	-0.0128	-0.0364
LS=4	sd	0.0061	0.0052	0.0071	0.0078	0.0074	0.0068	0.0102
	median	-0.00904	0.0342**	0.0388**	0.0452**	-0.0571**	-0.0129*	-0.0363**
	mean	-0.0097	0.0361	0.0429	0.0431	-0.0550	-0.0143	-0.0332
LS=5	sd	0.0065	0.0054	0.0082	0.0068	0.0067	0.0078	0.0082
	median	-0.00964	0.0360**	0.0428**	0.0431**	-0.0550**	-0.0142*	-0.0332**

[1] SRH, LS, CR, CD and NC are the abbreviation of self-rated health, life satisfaction, Census Register, Chronic Disease and Need for Care, respectively.

[2] "$**$" and "$*$" denote that zero is not included in the 95% and 90% credible intervals, respectively. Mean, Sd and Median are posterior mean, posterior standard deviation and posterior median, respectively.

"satisfied" or "very satisfied." The marginal effects of census register, chronic disease and need for care are positive when "self-rated health $= 1, 2, 3$ or $1, 2$" & "life satisfaction $= 1, 2, 3$," and negative when "self-rated health $= 4, 5$ or $3, 4, 5$" & "life satisfaction $= 4, 5$." Therefore, in urban elderly people, in people without chronic disease, and in people without need for care self-rated health, self-rated health is more likely to be "good" or "very good," and life satisfaction is more likely to be "satisfied" or "very satisfied."

In conclusion, elderly males are more positive about their self-rated health, and people living with a spouse are more positive about their life satisfaction. People with high education level, urban elderly people, people without chronic disease, and people without need for care are more positive about self-rated health and life satisfaction. Older elderly people's self-rated health is more negative, but their life satisfaction is higher.

In this empirical study, the effect of age is opposite for the two response variables, and this result was also obtained by Rodgers et al. (2017). They found that self-rated health decreased with age, but life satisfaction increased with age. In addition, high life satisfaction is associated with good self-rated health (Enkvist et al., 2012; Palmore and Luikart, 1972). Therefore, life satisfaction may also be positively affected by the positive factors for self-rated health, such as urban (Ford et al., 2007), high education level (Damián et al., 2008), no existence of chronic disease (Damián et al., 2008; Lima-Costa et al., 2004), and no need for care (Ho et al., 1995). Moreover, Palmore and Luikart (1972) pointed out that gender has little or no relation with life satisfaction, and Enkvist et al. (2012) showed that marital status could predict life satisfaction significantly. These studies are all consistent with our finding.

Furthermore, the turning points of the positive signs ("good" and "very good" for self-rated health; "satisfied" and "very satisfied" for life satisfaction) and negative signs ("poor" and "very poor" for self-rated health; "dissatisfied" and "very dissatisfied" for life satisfaction) of the marginal effects for all the explanatory variables on both dependent variables are always “general.”

3.5 Chapter Summary

This chapter proposed a new algorithm for a MSOP model. In applying this model, the parameters are calculated using the Bayesian inference based on MCMC sampling. The parameters are estimated using a set of FCDs. To obtain the FCDs, this chapter assumed diffuse prior distributions for all the parameters. The prior distributions of all the other parameters are similar to those in former studies on Bayesian spatial econometrics (Kakamu and Wago, 2007; Smith and LeSage, 2004; Wang and Kockelman, 2009a,b,c) except for cutpoints $\boldsymbol{\gamma}$. The prior distributions of $\boldsymbol{\gamma}$ in this study are referred to the method in Chen and Dey (2000). This method removed the ordering constraint using the one-to-one map, and it has been proven to be more stable than using the uniform distribution.

The validity and accuracy of the model is verified by the simulated dataset. The estimation results are quite close to the true values. Therefore, the MSOP model performs very well with the simulated data. In addition, the study illustrated the model by applying it to the survey data of elderly people in six provinces in East China.

4

Bayesian Analysis of a DMSOP Model

Besides spatial dependencies, temporal dependencies also widely exist in the spatial econometrics. The dynamic ordered probit (DOP) model is needed to capture this characteristic. Following the DOP model with spatial dependencies in Wang and Kockelman (2009c) and models for multivariate ordinal outcomes in Jeliazkov et al. (2008), this chapter proposes a dynamic multivariate spatial ordered probit (DMSOP) model, wherein the spatial dependencies are explained using an additive error specification, as in Smith and LeSage (2004). The DMSOP model is the first attempt to capture temporal and spatial dependencies simultaneously for the multivariate ordinal responses.

The rest of this chapter is structured as follows. Section 4.1 introduces the model specification for the DMSOP model. Section 4.2 explains details of the Bayesian inference for the model. Section 4.3 presents the model for the simulated dataset and its simulation results. Section 4.4 explains the empirical data and MCMC settings, and discusses the empirical results for the two multivariate outcomes: life satisfaction and self-rated health. Section 4.5 presents the summary of this chapter.

4.1 Model Specification of DMSOP Models

This chapter proposes a DMSOP model, in which the observed response variable y_{ikst} is ordinal and has the following relationship with the latent response variable z_{ikst}:

$$y_{ikst} = l_s \text{ if } \gamma_{s,l_s-1} < z_{ikst} \leq \gamma_{s,l_s}, \text{ for } s = 1, \ldots, S, l_s = 1, \ldots, L_s, \text{and } t = 0, 1, \ldots, T, \tag{4.1}$$

where i indexes spatial regions ($i = 1, \ldots, m$), k indexes individuals inside each region ($k = 1, \ldots, n_i$). Therefore, the total number of observations is $\sum_{i=1}^{m} n_i = n$. s indexes the sequence of observed response variables and t indexes time periods. l_s indexes the classifications of response variable y_{ikst}. Additionally, the cutpoints γ_{s,l_s} are specified, as in Chen and Dey (2000):

$$-\infty = \gamma_{s,0} < \gamma_{s,1} = 0 < \gamma_{s,2} < \cdots < \gamma_{s,L_s-1} = 1 < \gamma_{s,L_s} = \infty. \tag{4.2}$$

The latent response variable z_{ikst} in the DMSOP model is specified as follows:

$$\begin{cases} z_{ikst} = \sum_{g=1}^{S} \lambda_{sg} z_{ikg(t-1)} + \boldsymbol{x}'_{ikst}\boldsymbol{\beta}_s + \delta_{ist} + \xi_{ikst}, \ \text{for}\, t = 1, \ldots, T, \\ z_{iks0} = \boldsymbol{x}'_{iks0}\boldsymbol{\beta}_{s0} + \delta_{is0} + \xi_{iks0}, \ \text{for}\, t = 0, \end{cases} \tag{4.3}$$

where λ_{sg} is the temporal coefficient between z_{ikst} and $z_{ikg(t-1)}$. $\sum_{g=1}^{S} \lambda_{sg} z_{ikg(t-1)}$ reflects the effects of all the variables in the last time period. Commonly, $\sum_{g=1}^{S} |\lambda_{sg}| \leq 1$ to guarantee the system's stability. $\boldsymbol{x}_{ikst} = (x_{iks1t}, x_{iks2t}, \ldots, x_{iksq_st})'$ is a q_s-dimensional vector of the explanatory variables. Correspondingly, $\boldsymbol{\beta}_s = (\beta_{s1}, \beta_{s2}, \ldots, \beta_{sq_s})'$ is a q_s-dimensional vector of the regression parameters for $\boldsymbol{x}_{ikst}$. δ_{ist} and ξ_{ikst} represent regional effects and individualistic effects for the sth outcome in period t, respectively. δ_{ist} is assumed to be time-invariant, as in Wang and Kockelman (2009c):

$$\delta_{ist} \equiv \delta_{is} \quad \text{for all } t = 0, 1, \ldots, T.$$

Therefore, (4.3) can be rewritten as:

$$\begin{cases} z_{ikst} = \sum_{g=1}^{S} \lambda_{sg} z_{ikg(t-1)} + \boldsymbol{x}'_{ikst}\boldsymbol{\beta}_s + \delta_{is} + \xi_{ikst}, \ \text{for}\, t = 1, \ldots, T, \\ z_{iks0} = \boldsymbol{x}'_{iks0}\boldsymbol{\beta}_{s0} + \delta_{is} + \xi_{iks0}, \ \text{for}\, t = 0, \end{cases} \tag{4.4}$$

where δ_{is} represents all unobserved common effects for observations within region i for the sth outcome. The regional effects usually have spatial autocorrelation. Thus, a spatial autoregressive model, as in Smith and LeSage (2004), can be used:

$$\delta_{is} = \rho \sum_{j=1}^{m} w_{ij}\delta_{js} + \epsilon_{is}, \ \text{for}\, i = 1, \ldots, m, \text{and } s = 1, \ldots, S, \tag{4.5}$$

where ρ is the spatial coefficient, which reflects the degree of spatial influence. Weights w_{ij} reflect the closeness between regions i and j, and the most commonly used definition is the queen contiguity, which defines $w_{ij} = 1$ for regions sharing a common edge or a common vertex; otherwise, $w_{ij} = 0$. Additionally, $w_{ii} = 0$ and row sums $\sum_{j=1}^{m} w_{ij}$ are normalized to 1. Note that if a given region i has no neighbors, and then $w_{ij} = 0$ for all $j = 1, 2, \ldots, m$. In this case, $\sum_{j=1}^{m} w_{ij}$ are normalized to 0. The disturbance $\boldsymbol{\epsilon}_{is}$ captures the regional effects that are not spatially distributed. $\boldsymbol{\epsilon}_{is}$ is assumed to follow an i.i.d. normal distribution as:

$$\boldsymbol{\epsilon}_{is} \sim \mathrm{N}(0, \boldsymbol{\sigma}_s^2).$$

Stacking all regions for the sth outcome, (4.5) can be written as:

$$\boldsymbol{\delta}_s = \rho W \boldsymbol{\delta}_s + \boldsymbol{\epsilon}_s, \text{ for } s = 1, \ldots, S, \tag{4.6}$$

where $\boldsymbol{\delta}_s = (\delta_{1s}, \ldots, \delta_{is}, \ldots, \delta_{ms})'$, and $\boldsymbol{\epsilon}_s = (\epsilon_{1s}, \ldots, \epsilon_{is}, \ldots, \epsilon_{ms})'$. $\delta_{1s}, \ldots, \delta_{is}, \ldots, \delta_{ms}$ are assumed to be conditionally independent given ρ and σ_s^2. Therefore,

$$\boldsymbol{\epsilon}_s \sim \mathrm{N}(\mathbf{0}, \sigma_s^2 \boldsymbol{I}_m). \tag{4.7}$$

I define $\boldsymbol{A}_\rho = \boldsymbol{I}_m - \rho \boldsymbol{W}$, and assume that $\boldsymbol{A}_\rho$ is nonsingular. Then, $\boldsymbol{\delta}_s | \rho, \sigma_s^2 \sim \mathrm{N}\left[\mathbf{0}, \sigma_s^2 (\boldsymbol{A}_\rho' \boldsymbol{A}_\rho)^{-1}\right]$. Therefore, the density of $\boldsymbol{\delta}_s$ is:

$$p(\boldsymbol{\delta}_s | \rho, \sigma_s^2) \propto (\sigma_s^2)^{-m/2} |\boldsymbol{A}_\rho| \exp\left(-\frac{1}{2\sigma_s^2} \boldsymbol{\delta}_s' \boldsymbol{A}_\rho' \boldsymbol{A}_\rho \boldsymbol{\delta}_s\right).$$

I define $\boldsymbol{z}_{ikt} = (z_{ik1t}, \ldots, z_{ikst} \ldots, z_{ikSt})'$ as an S-dimensional vector of the latent response variables, and organize the matrix of covariates $\boldsymbol{X}_{ikt}$ as follows:

$$\boldsymbol{X}_{ikt} = \begin{pmatrix} \boldsymbol{x}'_{ik1t} & & & & \\ & \ddots & & & \\ & & \boldsymbol{x}'_{ikst} & & \\ & & & \ddots & \\ & & & & \boldsymbol{x}'_{ikSt} \end{pmatrix}_{S\times q}.$$

Therefore, the number of columns of $\boldsymbol{X}_{ikt}$ is $\sum_{s=1}^{S} q_s = q$. Furthermore, I define

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1g} & \cdots & \lambda_{1S} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \lambda_{s1} & \cdots & \lambda_{sg} & \cdots & \lambda_{sS} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \lambda_{S1} & \cdots & \lambda_{Sg} & \cdots & \lambda_{SS} \end{pmatrix}_{S\times S},$$

where $\lambda_{sg} \neq 0$ for all $s \neq g$. That is, each latent variable in period t is affected by all the latent variables in the last time period $t-1$.

Then, (4.4) can be rewritten as:

$$\begin{cases} \boldsymbol{z}_{ikt} = \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} + \boldsymbol{X}_{ikt}\boldsymbol{\beta} + \boldsymbol{\delta}_i + \boldsymbol{\xi}_{ikt}, \text{ for } t = 1, \ldots, T, \\ z_{ik0} = \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 + \delta_i + \xi_{ik0}, \text{ for } t = 0, \end{cases} \tag{4.8}$$

where $\boldsymbol{\beta} = \left(\boldsymbol{\beta}'_1, \ldots, \boldsymbol{\beta}'_s, \ldots, \boldsymbol{\beta}'_S\right)'$ and $\boldsymbol{\beta}_0 = \left(\boldsymbol{\beta}'_{10}, \ldots, \boldsymbol{\beta}'_{s0}, \ldots, \boldsymbol{\beta}'_{S0}\right)'$ are q-dimensional vectors of the regression parameters. $\boldsymbol{\xi}_{ikt} = (\xi_{ik1t}, \ldots, \xi_{ikst}, \ldots, \xi_{ikSt})'$ is an S-dimensional vector of the individualistic effects. The individualistic effects $\boldsymbol{\xi}_{ikt}$ are assumed conditionally i.i.d. normal variates with zero means and a common covariance matrix $\boldsymbol{V}$;

that is,

$$\boldsymbol{\xi}_{ikt} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{V}), \text{ for } t = 0, 1, \ldots, T,$$

where $\boldsymbol{V}$ is an $S \times S$ symmetric positive definite covariance matrix. $\boldsymbol{\delta}_i = (\delta_{i1}, \ldots, \delta_{is}, \ldots, \delta_{iS})'$ is an S-dimensional vector of the regional effects. $\delta_{i1}, \ldots, \delta_{iS}$ are assumed to be conditionally independent given ρ and $\boldsymbol{\sigma}^2 = \left(\sigma_1^2, \ldots, \sigma_s^2, \ldots, \sigma_S^2\right)'$. Defining $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{is}, \ldots, \epsilon_{iS})'$, from (4.5), it is obvious that

$$\boldsymbol{\epsilon}_i \sim \mathrm{N}(\mathbf{0}, \mathrm{diag}(\boldsymbol{\sigma}^2)).$$

As is common in spatial econometrics (Kakamu and Wago, 2007; LeSage, 1999; Smith and LeSage, 2004; Wang and Kockelman, 2009c), for each $t \in \{0, 1, \ldots, T\}$, observations can be stacked by region and individual, and then the DMSOP model can be summarized as follows:

$$\begin{cases} \boldsymbol{z}_t = (\boldsymbol{I}_n \otimes \boldsymbol{\Lambda})\boldsymbol{z}_{t-1} + \boldsymbol{X}_t\boldsymbol{\beta} + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta} + \boldsymbol{\xi}_t, \text{ for } t = 1, \ldots, T, \\ z_0 = \boldsymbol{X}_0\boldsymbol{\beta}_0 + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta} + \boldsymbol{\xi}_0, \text{ for } t = 0, \end{cases} \tag{4.9}$$

$$\boldsymbol{\delta} = \rho(\boldsymbol{W} \otimes \boldsymbol{I}_S)\boldsymbol{\delta} + \boldsymbol{\epsilon}, \tag{4.10}$$

$$\boldsymbol{\xi}_t \sim \mathrm{N}(\mathbf{0}, \boldsymbol{I}_n \otimes \boldsymbol{V}), \text{ for } t = 0, 1, \ldots, T, \tag{4.11}$$

$$\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2)), \tag{4.12}$$

where $\boldsymbol{z}_t = \begin{pmatrix} \boldsymbol{z}_{1t} \\ \vdots \\ \boldsymbol{z}_{it} \\ \vdots \\ \boldsymbol{z}_{mt} \end{pmatrix}_{nS \times 1}$, with each $\boldsymbol{z}_{it} = \begin{pmatrix} \boldsymbol{z}_{i1t} \\ \vdots \\ \boldsymbol{z}_{ikt} \\ \vdots \\ \boldsymbol{z}_{in_it} \end{pmatrix}_{n_iS \times 1}$, $\boldsymbol{X}_t = \begin{pmatrix} \boldsymbol{X}_{1t} \\ \vdots \\ \boldsymbol{X}_{it} \\ \vdots \\ \boldsymbol{X}_{mt} \end{pmatrix}_{nS \times q}$, with each $\boldsymbol{X}_{it} = \begin{pmatrix} \boldsymbol{X}_{i1t} \\ \vdots \\ \boldsymbol{X}_{ikt} \\ \vdots \\ \boldsymbol{X}_{in_it} \end{pmatrix}_{n_iS\times q}$, $\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_s \\ \vdots \\ \boldsymbol{\beta}_S \end{pmatrix}_{q\times 1}$, $\boldsymbol{\Delta} = \begin{pmatrix} \mathbf{1}_1 & & & & \\ & \ddots & & & \\ & & \mathbf{1}_i & & \\ & & & \ddots & \\ & & & & \mathbf{1}_m \end{pmatrix}_{n\times m}$, with

each $\mathbf{1}_i = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{n_i\times 1}$, $\boldsymbol{\delta} = \begin{pmatrix} \boldsymbol{\delta}_1 \\ \vdots \\ \boldsymbol{\delta}_i \\ \vdots \\ \boldsymbol{\delta}_m \end{pmatrix}_{mS\times 1}$, $\boldsymbol{\xi}_t = \begin{pmatrix} \boldsymbol{\xi}_{1t} \\ \vdots \\ \boldsymbol{\xi}_{it} \\ \vdots \\ \boldsymbol{\xi}_{mt} \end{pmatrix}_{nS\times 1}$, with each $\boldsymbol{\xi}_{it} = \begin{pmatrix} \boldsymbol{\xi}_{i1t} \\ \vdots \\ \boldsymbol{\xi}_{ikt} \\ \vdots \\ \boldsymbol{\xi}_{in_it} \end{pmatrix}_{n_iS\times 1}$,

$\boldsymbol{W} = \begin{pmatrix} w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mm} \end{pmatrix}_{m\times m}$, $\boldsymbol{\epsilon} = \begin{pmatrix} \boldsymbol{\epsilon}_1 \\ \vdots \\ \boldsymbol{\epsilon}_i \\ \vdots \\ \boldsymbol{\epsilon}_m \end{pmatrix}_{mS\times 1}$, $\operatorname{diag}(\boldsymbol{\sigma}^2) = \begin{pmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_S^2 \end{pmatrix}_{S\times S}$.

Defining $\boldsymbol{B}_\rho = \boldsymbol{I}_{mS} - \rho(\boldsymbol{W} \otimes \boldsymbol{I}_S)$, from (4.10) and (4.12), $\boldsymbol{\delta}$ has a multivariate normal distribution; that is, $\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2 \sim \mathrm{N}\left[\mathbf{0}, \boldsymbol{B}_\rho^{-1}(\boldsymbol{I}_m \otimes \operatorname{diag}(\boldsymbol{\sigma}^2))(\boldsymbol{B}_\rho^{-1})'\right]$. Therefore, the density of $\boldsymbol{\delta}$ is:

$$p(\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2) \propto |\operatorname{diag}(\boldsymbol{\sigma}^2)|^{-m/2}|\boldsymbol{B}_\rho| \exp\left(-\frac{1}{2}\boldsymbol{\delta}'\boldsymbol{B}_\rho'(\boldsymbol{I}_m \otimes \operatorname{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho\boldsymbol{\delta}\right).$$

Furthermore, I define $\boldsymbol{z}_t^{\boldsymbol{\Lambda}} = \boldsymbol{z}_t - (\boldsymbol{I}_n \otimes \boldsymbol{\Lambda})\boldsymbol{z}_{t-1}$, and stack all the observations by time, then (4.9) can be expressed as:

$$\begin{cases} \boldsymbol{z}^{\boldsymbol{\Lambda}} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\Theta}\boldsymbol{\delta} + \boldsymbol{\xi}, \ \text{for } t = 1, \ldots, T, \\ z_0 = \boldsymbol{X}_0\boldsymbol{\beta}_0 + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta} + \xi_0, \ \text{for}\, t = 0, \end{cases} \tag{4.13}$$

where $\boldsymbol{z}^{\boldsymbol{\Lambda}} = \begin{pmatrix} \boldsymbol{z}_1^{\boldsymbol{\Lambda}} \\ \boldsymbol{z}_2^{\boldsymbol{\Lambda}} \\ \vdots \\ \boldsymbol{z}_T^{\boldsymbol{\Lambda}} \end{pmatrix}_{nST\times 1}$, $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \vdots \\ \boldsymbol{X}_T \end{pmatrix}_{nST\times q}$, $\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \\ \vdots \\ \boldsymbol{\xi}_T \end{pmatrix}_{nST\times q}$, and $\boldsymbol{\Theta} = \boldsymbol{1}_T \otimes (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)$, with $\boldsymbol{1}_T$ is a T-dimensional vector of ones.

From (4.9) and (4.11), it can be derived that $\boldsymbol{z}_t$ has a multivariate normal distribution; that is,

$$\begin{cases} \boldsymbol{z}_t|\boldsymbol{z}_{t-1}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda} \sim \mathrm{N}\left[(\boldsymbol{I}_n \otimes \boldsymbol{\Lambda})\boldsymbol{z}_{t-1} + \boldsymbol{X}_t\boldsymbol{\beta} + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}, \boldsymbol{I}_n \otimes \boldsymbol{V}\right], \ \text{for } t = 1, \ldots, T, \\ \boldsymbol{z}_0|\boldsymbol{\beta}_0, \boldsymbol{\delta}, \boldsymbol{V} \sim \mathrm{N}\left[\boldsymbol{X}_0\boldsymbol{\beta}_0 + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}, \boldsymbol{I}_n \otimes \boldsymbol{V}\right], \ \text{for}\, t = 0. \end{cases}$$

Based on the definition of $\boldsymbol{z}_t^{\boldsymbol{\Lambda}}$, the conditional density of $\boldsymbol{z}_t$ for $t = 1, \ldots, T$ can be expressed as:

$$\begin{aligned} &p(\boldsymbol{z}_t|\boldsymbol{z}_{t-1}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \\ &\propto |\boldsymbol{V}|^{-n/2} \exp\left\{-\frac{1}{2}(\boldsymbol{z}_t^{\boldsymbol{\Lambda}} - \boldsymbol{X}_t\boldsymbol{\beta} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_t^{\boldsymbol{\Lambda}} - \boldsymbol{X}_t\boldsymbol{\beta} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})\right\} \\ &= \prod_{i=1}^{m}\prod_{k=1}^{n_i}\left\{|\boldsymbol{V}|^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)\right]\right\}, \end{aligned}$$

and the conditional density of $\boldsymbol{z}_0$ can be expressed as:

$$
\begin{aligned}
& p(\boldsymbol{z}_0|\boldsymbol{\beta}_0, \boldsymbol{\delta}, \boldsymbol{V}) \\
\propto & |\boldsymbol{V}|^{-n/2} \exp\left\{-\frac{1}{2}(\boldsymbol{z}_0 - \boldsymbol{X}_0\boldsymbol{\beta}_0 - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - \boldsymbol{X}_0\boldsymbol{\beta}_0 - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})\right\} \\
= & \prod_{i=1}^{m}\prod_{k=1}^{n_i}\left\{|\boldsymbol{V}|^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{z}_{ik0} - \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik0} - \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 - \boldsymbol{\delta}_i)\right]\right\}.
\end{aligned}
$$

For $\boldsymbol{z} = (\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_T)'$, it is obvious that:

$$
\begin{aligned}
& p(\boldsymbol{z}|\boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) = p(\boldsymbol{z}_T, \boldsymbol{z}_{T-1}, \ldots, \boldsymbol{z}_1|\boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \\
= & p(\boldsymbol{z}_T|\boldsymbol{z}_{T-1}, \ldots, \boldsymbol{z}_1, \boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \times p(\boldsymbol{z}_{T-1}|\boldsymbol{z}_{T-2}, \ldots, \boldsymbol{z}_1, \boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \times \cdots \times p(\boldsymbol{z}_1|\boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \\
= & p(\boldsymbol{z}_T|\boldsymbol{z}_{T-1}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \times p(\boldsymbol{z}_{T-1}|\boldsymbol{z}_{T-2}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \times \cdots \times p(\boldsymbol{z}_1|\boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \\
= & \prod_{t=1}^{T} p(\boldsymbol{z}_t|\boldsymbol{z}_{t-1}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \propto |\boldsymbol{V}|^{-Tn/2} \exp\left\{-\frac{1}{2}(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{\Theta}\boldsymbol{\delta})'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{\Theta}\boldsymbol{\delta})\right\} \\
\propto & \prod_{t=1}^{T}\left\{|\boldsymbol{V}|^{-n/2} \exp\left[-\frac{1}{2}(\boldsymbol{z}_t^{\boldsymbol{\Lambda}} - \boldsymbol{X}_t\boldsymbol{\beta} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_t^{\boldsymbol{\Lambda}} - \boldsymbol{X}_t\boldsymbol{\beta} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})\right]\right\} \\
\propto & \prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i}\left\{|\boldsymbol{V}|^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)\right]\right\}.
\end{aligned}
$$

Moreover, I define that:

$$G_{iks} = (\gamma_{s,l_s-1}, \gamma_{s,l_s}] \ \text{ if } y_{iks} = l_s, \text{for } l_s = 1, \ldots, L_s, \text{and } s = 1, \ldots, S,$$

$$G_{ik} = G_{ik1} \times \cdots \times G_{ikS}, \ \text{for}\, i = 1, \ldots, m, \text{and } k = 1, \ldots, n_i.$$

Defining $\boldsymbol{y}_{ikt} = (y_{ik1t}, \ldots, y_{ikSt})'$ as an S-dimensional vector of the response variables, it can be derived that

$$p(\boldsymbol{y}_{ikt}|\boldsymbol{\gamma}, \boldsymbol{z}_{ikt}) = 1(\boldsymbol{z}_{ikt} \in G_{ik}), \text{ for } t = 0, 1, \ldots, T,$$

where $1(\cdot)$ is an indicator function.

4.2 Parameter Estimation of DMSOP Models

4.2.1 Prior Distributions of Parameters

In this chapter, the following prior distributions of the parameters are introduced to the DMSOP model:

$$\begin{aligned}
&\boldsymbol{\beta}_0 \sim \mathrm{N}(\boldsymbol{c}_0, \boldsymbol{\Gamma}_0), \\
&\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{c}, \boldsymbol{\Gamma}), \\
&\rho \sim \mathrm{Beta}(a_0, b_0) \text{ centered on zero with } |\rho| \leq 1, \\
&\sigma_s^2 \sim \mathrm{IG}(e_0, g_0), \\
&\boldsymbol{V} \sim \mathrm{IW}(n_0, \boldsymbol{Q}_0), \\
&\lambda_{sg} \sim \mathrm{U}(-1, 1) \text{ with } \sum_{g=1}^{S} |\lambda_{sg}| \leq 1,
\end{aligned}$$

where $\mathrm{IW}(n_0, \boldsymbol{Q}_0)$ denotes an Inverse-Wishart distribution with degree of freedom n_0 and $S \times S$ scale matrix $\boldsymbol{Q}_0$, and $\mathrm{IG}(e_0, g_0)$ denotes an Inverse-Gamma distribution with shape e_0 and scale g_0. The beta prior of ρ in this chapter is similar to LeSage and Parent (2007), as follows:

$$\pi(\rho) \propto (1+\rho)^{a_0-1}(1-\rho)^{b_0-1} \times 1(-1 \leq \rho \leq 1).$$

Moreover, when $a_0 = b_0 = 1$, the prior of ρ is a uniform distribution, and thus, this prior is more flexible than the uniform prior.

Because $\boldsymbol{\gamma}$ has the limitation of $-\infty = \gamma_{s,0} < \gamma_{s,1} = 0 < \gamma_{s,2} < \cdots < \gamma_{s,L_s-1} = 1 < \gamma_{s,L_s} = \infty$, it is obvious that

$$\pi(\boldsymbol{\gamma}) \propto 1 \text{ for } \gamma_{s,2} < \cdots < \gamma_{s,L_s-2}, s = 1, \ldots, S.$$

4.2.2 MCMC Sampler of DMSOP Models

The joint posterior density function of the parameters is as follows:

$$\begin{aligned}
&p(\boldsymbol{\beta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \gamma, \boldsymbol{z}, \boldsymbol{z}_0, \boldsymbol{\Lambda}|\boldsymbol{y}, \boldsymbol{y}_0) \\
&\propto p(\boldsymbol{y}, \boldsymbol{y}_0|\boldsymbol{\beta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \gamma, \boldsymbol{z}, \boldsymbol{z}_0, \boldsymbol{\Lambda}) p(\boldsymbol{\beta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \gamma, \boldsymbol{z}, \boldsymbol{z}_0, \boldsymbol{\Lambda}) \\
&= p(\boldsymbol{y}|\boldsymbol{z}, \boldsymbol{\gamma}) p(\boldsymbol{y}_0|\boldsymbol{z}_0, \gamma) p(\boldsymbol{z}|\boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) p(\boldsymbol{z}_0|\boldsymbol{\beta}_0, \boldsymbol{\delta}, \boldsymbol{V}) p(\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2) \pi(\boldsymbol{\beta}_0) \pi(\boldsymbol{\beta}) \pi(\rho) \pi(\boldsymbol{\sigma}^2) \pi(\boldsymbol{V}) \pi(\boldsymbol{\gamma}) \pi(\boldsymbol{\Lambda}),
\end{aligned}$$

where the probability of y_{iks} equals 1 when $\gamma_{s,l_s-1} < z_{iks} \leq \gamma_{s,l_s}$, and 0 otherwise. Given $\boldsymbol{z}$ and $\boldsymbol{\gamma}$, y_{ik} are conditionally i.i.d., and thus,

$$p(\boldsymbol{y}|\boldsymbol{z}, \gamma) = \prod_{t=1}^{T} \prod_{i=1}^{m} \prod_{k=1}^{n_i} p(\boldsymbol{y}_{ikt}|\gamma, \boldsymbol{z}_{ikt}) = \prod_{t=1}^{T} \prod_{i=1}^{m} \prod_{k=1}^{n_i} 1(\boldsymbol{z}_{ikt} \in G_{ik}),$$

$$p(\boldsymbol{y}_0|\boldsymbol{z}_0, \gamma) = \prod_{i=1}^{m} \prod_{k=1}^{n_i} p(\boldsymbol{y}_{ik0}|\gamma, \boldsymbol{z}_{ik0}) = \prod_{i=1}^{m} \prod_{k=1}^{n_i} 1(\boldsymbol{z}_{ik0} \in G_{ik}).$$

In this chapter, the parameters are estimated using a set of full conditional distributions. Note that the sample of $\boldsymbol{\gamma}$ can be obtained using a uniform distribution. Almost all the literature on the theme of spatial ordered probit (SOP) models used such an estimation method (Wang and Kockelman, 2009c). In this chapter, the sampling of $\boldsymbol{\gamma}$

uses transformations, as in Chen and Dey (2000).

Based on the joint posterior density function and prior distributions, the Appendix B.1 presents the details of full conditional distributions of all the parameters. Let $\star$ denote all the other parameters, $\boldsymbol{y}_0$ and $\boldsymbol{y}$ except the given parameters, then the parameters are estimated by MCMC sampling as follows:

1. Sample $\boldsymbol{\beta}_0|\star \sim \mathrm{N}(\boldsymbol{B}_{\boldsymbol{\beta}_0}^{-1}\boldsymbol{b}_{\boldsymbol{\beta}_0}, \boldsymbol{B}_{\boldsymbol{\beta}_0}^{-1})$, $\boldsymbol{B}_{\boldsymbol{\beta}_0} = \boldsymbol{X}_0'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})\boldsymbol{X}_0 + \boldsymbol{\Gamma}_0^{-1}$, and $\boldsymbol{b}_{\boldsymbol{\beta}_0} = \boldsymbol{X}_0'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}) + \boldsymbol{\Gamma}_0^{-1}\boldsymbol{c}_0$.

2. Sample $\boldsymbol{\beta}|\star \sim \mathrm{N}(\boldsymbol{B}_{\boldsymbol{\beta}}^{-1}\boldsymbol{b}_{\boldsymbol{\beta}}, \boldsymbol{B}_{\boldsymbol{\beta}}^{-1})$, where $\boldsymbol{B}_{\boldsymbol{\beta}} = \boldsymbol{X}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})\boldsymbol{X} + \boldsymbol{\Gamma}^{-1}$, and $\boldsymbol{b}_{\boldsymbol{\beta}} = \boldsymbol{X}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{\Theta}\boldsymbol{\delta}) + \boldsymbol{\Gamma}^{-1}\boldsymbol{c}$.

3. Sample $\boldsymbol{\delta}|\star \sim \mathrm{N}(\boldsymbol{\Omega}^{-1}\boldsymbol{\eta}, \boldsymbol{\Omega}^{-1})$, $\boldsymbol{\Omega} = \boldsymbol{\Theta}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})\boldsymbol{\Theta} + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{\Delta} \otimes \boldsymbol{I}_S) + \boldsymbol{B}_\rho'(\boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho$, and $\boldsymbol{\eta} = \boldsymbol{\Theta}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{X}\boldsymbol{\beta}) + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - \boldsymbol{X}_0\boldsymbol{\beta}_0)$.

4. Sample $\rho|\star$ using a Metropolis-Hastings (M-H) algorithm as explained in the Appendix B.1.4.

5. Sample $\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g_0}{\sigma_s^2}|\star \sim \chi^2(m + 2e_0)$, where $\boldsymbol{A}_\rho = \boldsymbol{I}_m - \rho\boldsymbol{W}$.

6. Sample $\boldsymbol{V}|\star \sim \mathrm{IW}(\tilde{n}, \tilde{\boldsymbol{Q}})$, where $\tilde{n} = (T+1)n + n_0$, and $\tilde{\boldsymbol{Q}} = \sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)' + \sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ik0} - \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 - \boldsymbol{\delta}_i)(\boldsymbol{z}_{ik0} - \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 - \boldsymbol{\delta}_i)' + \boldsymbol{Q}_0$.

7. Sample $\lambda_{sg}|\star \sim \mathrm{N}\left(\frac{\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i} A_{iksgt}z_{ikg(t-1)}}{\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i} z_{ikg(t-1)}^2}, \frac{\tilde{v}_s}{\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i} z_{ikg(t-1)}^2}\right)$, where $A_{iksgt} = z_{ikst} - (\boldsymbol{\lambda}_{s,-g}\boldsymbol{z}_{ik(-g)(t-1)} + + \boldsymbol{x}_{ikst}'\boldsymbol{\beta}_s + \delta_{is} + \boldsymbol{v}_{s,-s}\boldsymbol{V}_{-s,-s}^{-1}(\boldsymbol{z}_{ik(-s)t} - \boldsymbol{h}_{ik(-s)t}))$, and $\boldsymbol{h}_{ik(-s)t} = \boldsymbol{\lambda}_{-s,g}z_{ikg(t-1)} + \boldsymbol{\Lambda}_{-s,-g}\boldsymbol{z}_{ik(-g)(t-1)} + \boldsymbol{X}_{ik(-s)t}\boldsymbol{\beta}_{-s} + \boldsymbol{\delta}_{i(-s)}$. $\boldsymbol{z}_{ik(-s)t}$

denotes the vector removing z_{ikst} from $\boldsymbol{z}_{ikt}$, and $\boldsymbol{z}_{ik(-g)(t-1)}$ denotes the vector removing $z_{ikg(t-1)}$ from $\boldsymbol{z}_{ik(t-1)}$. $\boldsymbol{\lambda}_{s,-g} = (\lambda_{s1}, \ldots, \lambda_{s,g-1}, \lambda_{s,g+1}, \ldots, \lambda_{sS})$ and $\boldsymbol{v}_{s,-s} = (V_{s,1}, \ldots, V_{s,s-1}, V_{s,s+1}, \ldots, V_{s,S})$ are $1 \times (S-1)$ vectors, $\boldsymbol{\lambda}_{-s,g} = (\lambda_{1g}, \ldots, \lambda_{s-1,g}, \lambda_{s+1,g}, \ldots, \lambda_{Sg})'$ and $\boldsymbol{v}_{-s,s} = (V_{1,s}, \ldots, V_{s-1,s}, V_{s+1,s}, \ldots, V_{S,s})'$ are $(S-1) \times 1$ vectors. $\boldsymbol{\Lambda}_{-s,-g}$ and $\boldsymbol{V}_{-s,-s}$ denote the $(S-1) \times (S-1)$ matrices, removing $\boldsymbol{\lambda}_{s,-g}$ and $\boldsymbol{\lambda}_{-s,g}$ from $\boldsymbol{\Lambda}$, and $\boldsymbol{v}_{s,-s}$ and $\boldsymbol{v}_{-s,s}$ from $\boldsymbol{V}$, respectively.

8. Sample $\boldsymbol{z}_{ik0}|\star \sim \mathrm{N}(\boldsymbol{B}_{\boldsymbol{z}_0}^{-1}\boldsymbol{b}_{\boldsymbol{z}_0}, \boldsymbol{B}_{\boldsymbol{z}_0}^{-1})1(\boldsymbol{z}_{ik0} \in G_{ik})$, where $\boldsymbol{B}_{\boldsymbol{z}_0} = \boldsymbol{\Lambda}'\boldsymbol{V}^{-1}\boldsymbol{\Lambda} + \boldsymbol{V}^{-1}$, and $\boldsymbol{b}_{\boldsymbol{z}_0} = \boldsymbol{\Lambda}'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik1} - \boldsymbol{X}_{ik1}\boldsymbol{\beta} - \boldsymbol{\delta}_i) + \boldsymbol{V}^{-1}(\boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 + \boldsymbol{\delta}_i)$.

9. Sample $\boldsymbol{z}_{ikt}|\star \sim \mathrm{N}(\boldsymbol{B}_{\boldsymbol{z}_t}^{-1}\boldsymbol{b}_{\boldsymbol{z}_t}, \boldsymbol{B}_{\boldsymbol{z}_t}^{-1})1(\boldsymbol{z}_{ikt} \in G_{ik})$ for $t = 1, \ldots, T-1$, where $\boldsymbol{B}_{\boldsymbol{z}_t} = \boldsymbol{V}^{-1} + \boldsymbol{\Lambda}'\boldsymbol{V}^{-1}\boldsymbol{\Lambda}$, and $\boldsymbol{b}_{\boldsymbol{z}_t} = \boldsymbol{V}^{-1}(\boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} + \boldsymbol{X}_{ikt}\boldsymbol{\beta} + \boldsymbol{\delta}_i) + \boldsymbol{\Lambda}'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik(t+1)} - \boldsymbol{X}_{ik(t+1)}\boldsymbol{\beta} - \boldsymbol{\delta}_i)$.

10. Sample $\boldsymbol{z}_{ikT}|\star \sim \mathrm{N}(\boldsymbol{\Lambda}\boldsymbol{z}_{ik(T-1)} + \boldsymbol{X}_{ikT}\boldsymbol{\beta} + \boldsymbol{\delta}_i, \boldsymbol{V})1(\boldsymbol{z}_{ikT} \in G_{ik})$.

11. Sample γ_{s,l_s} using the transformations as $\gamma_{s,l_s} = \frac{\gamma_{s,l_s-1}+\exp(\kappa_{s,l_s})}{1+\exp(\kappa_{s,l_s})}$ for $2 < l_s < L_s-2$, where the samples of κ_{s,l_s} are generated using a M-H algorithm as explained in the Appendix B.1.10.

4.3 Model Simulation of DMSOP Models

4.3.1 Simulated Dataset

A simulation study is implemented to demonstrate the validity and accuracy of the DMSOP model. The number of response variables are set as $S = 2$ ($\boldsymbol{y}_1$ and $\boldsymbol{y}_2$) for ease of understanding and explanation. Additionally, I assumed that there are $L_1 = L_2 = 5$ ordered categories for the two response variables.

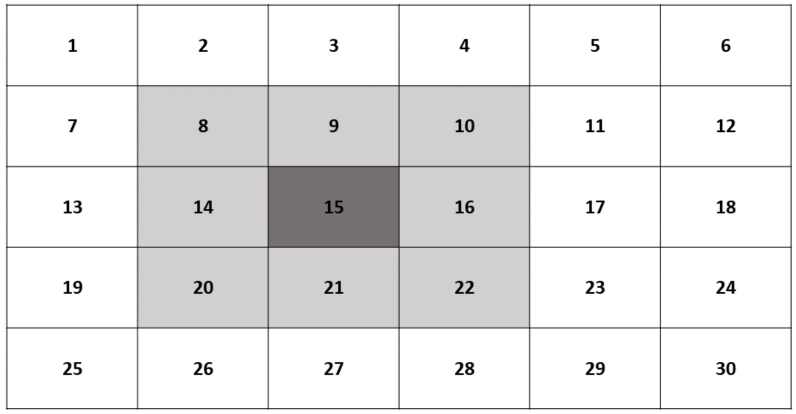


Figure 4.1: Location of regions in the simulated dataset for the DMSOP model

In the simulated dataset, the number of region m is set as 30, and that of individuals in each region n_i is set as 10 in different experiments. The time-series T is set as 7 (initial time is set as 0). The 30 regions are assumed to be located as shown in Figure 4.1. The weight matrix is generated based on queen contiguity. For example, region 15 is adjacent to regions 8, 9, 10, 14, 16, 20, 21 and 22. It is then row-standardized.

In the simulated dataset, each response variable has $q_s = 2$ explanatory variables. They are generated using a standard uniform distribution, $\mathrm{U}(0, 1)$. Note that the simulation does not include the intercept. The explanatory variables are set to be different for the two response variables. $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_0$ are set the same in different experiments as follows: $\boldsymbol{\beta} = (0.25, 0.51, 0.33, 0.46)'$, $\boldsymbol{\beta}_0 = (0.47, 0.50, 0.49, 0.53)'$.

The vector of regional effects $\boldsymbol{\delta}$ is generated using the multivariate normal distribution:

$$\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2 \sim \mathrm{N}\left[\mathbf{0}, \boldsymbol{B}_\rho^{-1}(\boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2))(\boldsymbol{B}_\rho^{-1})'\right],$$

where $\boldsymbol{B}_\rho = \boldsymbol{I}_{ms} - \rho(\boldsymbol{W} \otimes \boldsymbol{I}_S)$, and spatial coefficient ρ is set as 0.3, 0.5, and 0.7 in different experiments. $\boldsymbol{\sigma}^2$ is set the same in different experiments as $(0.02, 0.01)'$.

Since there are two response variables, the setting of covariance v_{12}/v_{21} is extremely important: the greater is the covariance, the stronger is the interaction of the response variables. To verify the stability of the model, v_{12}/v_{21} is set as 0.05 (weak interaction), 0.4 (moderate interaction), and 0.9 (strong interaction) in different experiments. Thus, covariance matrix $\boldsymbol{V}$ is set as $\boldsymbol{V} = \begin{pmatrix} 0.10 & 0.05 \\ 0.05 & 0.10 \end{pmatrix}$, $\boldsymbol{V} = \begin{pmatrix} 0.50 & 0.40 \\ 0.40 & 0.50 \end{pmatrix}$, and $\boldsymbol{V} = \begin{pmatrix} 1.0 & 0.9 \\ 0.9 & 1.0 \end{pmatrix}$ in different experiments. Based on the settings of $\boldsymbol{V}$, the vector of individualistic effects $\boldsymbol{\xi}_t$ is generated using the multivariate normal distribution $\boldsymbol{\xi}_t \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}_n \otimes \boldsymbol{V}),$ for $t = 0, 1, \ldots, T$.

Furthermore, the temporal coefficient matrix $\boldsymbol{\Lambda}$ also has three settings: $\boldsymbol{\Lambda} = \begin{pmatrix} 0.15 & 0.03 \\ 0.015 & 0.12 \end{pmatrix}$, $\boldsymbol{\Lambda} = \begin{pmatrix} 0.30 & 0.06 \\ 0.03 & 0.24 \end{pmatrix}$, and $\boldsymbol{\Lambda} = \begin{pmatrix} 0.45, 0 & 0.09 \\ 0.045 & 0.36 \end{pmatrix}$ in different experiments. Because the categories are set as $L_1 = L_2 = 5$ for the two response variables, the cutpoints to be estimated are 2 for each response variable. In the simulations, the cutpoints are set the same in different experiments as follows: $-\infty = \gamma_{1,0} < \gamma_{1,1} = 0 < \gamma_{1,2} = 0.31 < \gamma_{1,3} = 0.66 < \gamma_{1,4} = 1 < \gamma_{1,5} = \infty$, and $-\infty = \gamma_{2,0} < \gamma_{2,1} = 0 < \gamma_{2,2} = 0.33 < \gamma_{2,3} = 0.68 < \gamma_{2,4} = 1 < \gamma_{2,5} = \infty$.

Based on the above parameter values, the dataset is then generated using the following model assumptions:

$$\begin{cases} \boldsymbol{z}_t = (\boldsymbol{I}_n \otimes \boldsymbol{\Lambda})\boldsymbol{z}_{t-1} + \boldsymbol{X}_t\boldsymbol{\beta} + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta} + \boldsymbol{\xi}_t, \text{ for } t = 1, \ldots, 7, \\ z_0 = \boldsymbol{X}_0\boldsymbol{\beta}_0 + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta} + \boldsymbol{\xi}_0, \text{ for } t = 0, \end{cases}$$

and

$y_{ikst} = l_s$ if $\gamma_{s,l_s-1} < z_{ikst} \leq \gamma_{s,l_s}$, for $s = 1, 2, l_s = 1, \ldots, 5,$ and $t = 0, 1, \ldots, 7.$

4.3.2 Simulation Results

In the simulation, the following less informative priors are used: $\boldsymbol{\beta}_0 \,\&\, \boldsymbol{\beta} \sim \mathrm{N}(\mathbf{1}, 10^{12}\boldsymbol{I})$; $\rho \sim \mathrm{U}(-1,1)$; $\sigma_1^2 = \sigma_2^2 \sim \mathrm{IG}(0,0)$; $\boldsymbol{V} \sim \mathrm{IW}(1, \boldsymbol{Q}_0), \boldsymbol{Q}_0 = \begin{pmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{pmatrix}$; $l_{sg} \sim \mathrm{U}(-1,1)$, with $\sum_{g=1}^{S} |l_{sg}| \leq 1$; $\boldsymbol{\kappa}_1/\boldsymbol{\kappa}_2 \sim \mathrm{N}(\mathbf{0}_{2\times1}, 100\boldsymbol{I}_{2\times2})$. The posterior results are calculated using MCMC sampling.

Considering the large size of the simulated dataset, each MCMC simulation is run for $3,000$ iterations, and the first $1,000$ samples are discarded as the burn-in period. Thus, the parameters are calculated based on the final $2,000$ runs. In this chapter, the main parameters I focus on are $\boldsymbol{\beta}$, $\boldsymbol{\Lambda}$ and ρ. Mean squared error (MSE) is used to describe their estimation accuracy as follows:

$$MSE_{\boldsymbol{\beta}} = \frac{1}{2000}\sum_{\varphi=1}^{2000}(\boldsymbol{\beta}^{(\varphi)} - \boldsymbol{\beta}_{true})'(\boldsymbol{\beta}^{(\varphi)} - \boldsymbol{\beta}_{true}), \tag{4.14}$$

$$MSE_{\boldsymbol{\Lambda}} = \sum_{s=1}^{2}\sum_{g=1}^{2} MSE_{\lambda_{sg}} = \sum_{s=1}^{2}\sum_{g=1}^{2}\left\{\frac{1}{2000}\sum_{\varphi=1}^{2000}(\lambda_{sg}^{(\varphi)} - \lambda_{sg_{true}})^2\right\}, \tag{4.15}$$

$$MSE_{\rho} = \frac{1}{2000}\sum_{\varphi=1}^{2000}(\rho^{(\varphi)} - \rho_{true})^2, \tag{4.16}$$

where $\boldsymbol{\beta}^{(\varphi)}$, $\lambda_{sg}^{(\varphi)}$ and $\rho^{(\varphi)}$ represent the final φth simulation results.

Table 4.1 gives partial simulation results of posterior means and true values as well as MSE of $\boldsymbol{\beta}$, $\boldsymbol{\Lambda}$ and ρ. The Appendix Tables B.1, B.2, and B.3 provide the complete results of 27 simulation experiments. From Table 4.1, as the covariance matrix $\boldsymbol{V}$

increases, $MSE_{\boldsymbol{\beta}}$, $MSE_{\boldsymbol{\Lambda}}$ and MSE_{ρ} tend to be larger. That is, the magnitudes of coefficients tend to exhibit greater bias. One reason for this phenomenon is that, as the covariance matrix increases, the interaction of latent response variables is enhanced, thus adding uncertainty to the model. This uncertainty leads to greater bias. As the spatial coefficient ρ increases, $MSE_{\boldsymbol{\beta}}$ and $MSE_{\boldsymbol{\Lambda}}$ tend be larger. That is, magnitudes of coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\Lambda}$ tend to exhibit greater bias. However, a noteworthy tendency is apparent; as the spatial coefficient ρ increases, MSE_{ρ} tends to decrease. Therefore, the model can accurately capture high spatial dependency. As the temporal coefficient matrix $\boldsymbol{\Lambda}$ increases, $MSE_{\boldsymbol{\beta}}$ tends be larger, but $MSE_{\boldsymbol{\Lambda}}$ and MSE_{ρ} do not show a clear increasing trend, and sometimes they decrease. This indicates that the model can precisely capture temporal dependency.

From Table 4.1, the estimation results (posterior means) are quite close to the true values. Even in extreme situations ($\boldsymbol{V} = \begin{pmatrix} 1.0 & 0.9 \\ 0.9 & 1.0 \end{pmatrix}$, $\boldsymbol{\Lambda} = \begin{pmatrix} 0.45, 0 & 0.09 \\ 0.045 & 0.36 \end{pmatrix}$, and $\rho = 0.7$), the biases are very small. Thus, the DMSOP model performs effectively with the simulated data. It successfully explains the temporal and spatial dependencies, and the interaction of multivariate ordinal response variables as well as the effects of the explanatory variables.

4.4 Empirical Study

4.4.1 Empirical Data

In this empirical study, the survey data from China Family Panel Studies (CFPS) is used, which is conducted by the Institute of Social Science Survey of Peking University. The CFPS aims to collect data at three levels: individual, family, and community. The

Table 4.1: Partial simulation results of the DMSOP models

Parm	True	simulation results								
	$\boldsymbol{\Lambda}$	(0.15,0.03,0.015,0.12)			(0.3,0.06,0.03,0.24)			(0.45,0.09,0.045,0.36)		
	ρ	0.3	0.5	0.7	0.3	0.5	0.7	0.3	0.5	0.7
		$\boldsymbol{V}$=(0.1,0.05,0.05,0.1)								
β_{11}	0.25	0.2474	0.2506	0.2545	0.2435	0.2416	0.2348	0.2485	0.2654	0.2576
β_{12}	0.51	0.5063	0.5019	0.5083	0.5135	0.5087	0.5179	0.5066	0.5106	0.5032
β_{21}	0.33	0.3318	0.3269	0.3107	0.3315	0.3284	0.3429	0.3316	0.3307	0.3273
β_{22}	0.46	0.4428	0.4469	0.4470	0.4569	0.4452	0.4355	0.4510	0.4309	0.4269
λ_{11}	/	0.1639	0.1471	0.1587	0.3043	0.3107	0.3055	0.4453	0.4544	0.4405
λ_{12}	/	0.0436	0.0622	0.0589	0.0669	0.0613	0.0600	0.0982	0.0972	0.1173
λ_{21}	/	0.0165	0.0070	-0.0053	0.0110	0.0081	0.0049	0.0162	0.0181	0.0150
λ_{22}	/	0.1484	0.1496	0.1579	0.2798	0.2852	0.2817	0.4082	0.3984	0.4077
ρ	/	0.3125	0.5056	0.7219	0.3038	0.5079	0.7098	0.3577	0.5949	0.7258
$MSE_{\boldsymbol{\beta}}$	/	0.0020	0.0022	0.0024	0.0019	0.0023	0.0031	0.0020	0.0031	0.0034
$MSE_{\boldsymbol{\Lambda}}$	/	0.0016	0.0024	0.0032	0.0024	0.0030	0.0029	0.0036	0.0026	0.0044
MSE_{ρ}	/	0.0467	0.0285	0.0113	0.0432	0.0265	0.0121	0.0405	0.0262	0.0112
		$\boldsymbol{V}$=(0.5,0.4,0.4,0.5)								
β_{11}	0.25	0.2465	0.2341	0.2475	0.2489	0.2501	0.2740	0.2475	0.2579	0.2783
β_{12}	0.51	0.5015	0.5000	0.5070	0.4886	0.4827	0.4932	0.4788	0.4887	0.4942
β_{21}	0.33	0.3736	0.3680	0.3518	0.3453	0.3541	0.3707	0.3372	0.3231	0.3366
β_{22}	0.46	0.4395	0.4420	0.4635	0.4470	0.4364	0.4473	0.4434	0.4344	0.4376
λ_{11}	/	0.1250	0.1117	0.1212	0.2527	0.2527	0.2630	0.4317	0.4359	0.4160
λ_{12}	/	0.0707	0.0805	0.0764	0.1204	0.1235	0.1169	0.1208	0.1275	0.1420
λ_{21}	/	-0.0109	-0.0107	-0.0194	0.0015	-0.0093	0.0034	0.0521	0.0522	0.0356
λ_{22}	/	0.1662	0.1682	0.1740	0.2984	0.3087	0.2829	0.3902	0.3915	0.3971
ρ	/	0.1581	0.3476	0.6668	0.2237	0.3885	0.6779	0.2492	0.5269	0.7169
$MSE_{\boldsymbol{\beta}}$	/	0.0074	0.0069	0.0060	0.0057	0.0072	0.0087	0.0068	0.0077	0.0081
$MSE_{\boldsymbol{\Lambda}}$	/	0.0055	0.0074	0.0076	0.0106	0.0130	0.0077	0.0027	0.0031	0.0058
MSE_{ρ}	/	0.0928	0.0834	0.0193	0.0749	0.0650	0.0185	0.0620	0.0471	0.0147
		$\boldsymbol{V}$=(1,0.9,0.9,1)								
β_{11}	0.25	0.2967	0.2798	0.2874	0.2854	0.2715	0.2670	0.2260	0.2349	0.2687
β_{12}	0.51	0.4770	0.4719	0.4623	0.5208	0.4692	0.4581	0.4474	0.4233	0.4152
β_{21}	0.33	0.3847	0.3880	0.4020	0.3028	0.3916	0.3888	0.3731	0.3694	0.3813
β_{22}	0.46	0.4426	0.4607	0.4759	0.5295	0.4924	0.4937	0.4302	0.4496	0.4521
λ_{11}	/	0.1044	0.0813	0.1135	0.3164	0.2948	0.2750	0.4568	0.4742	0.4517
λ_{12}	/	0.0687	0.1003	0.0782	0.0181	0.0657	0.0897	0.0941	0.0765	0.0943
λ_{21}	/	-0.0150	-0.0586	-0.0466	0.0031	0.0078	-0.0144	0.0602	0.0726	0.0354
λ_{22}	/	0.1541	0.2005	0.1908	0.2507	0.2699	0.3030	0.3608	0.3534	0.3968
ρ	/	0.3275	0.5671	0.6924	0.3542	0.5341	0.6342	0.1564	0.2977	0.5662
$MSE_{\boldsymbol{\beta}}$	/	0.0129	0.0121	0.0165	0.0137	0.0135	0.0155	0.0144	0.0169	0.0203
$MSE_{\boldsymbol{\Lambda}}$	/	0.0068	0.0220	0.0135	0.0046	0.0025	0.0087	0.0009	0.0024	0.0023
MSE_{ρ}	/	0.0713	0.0513	0.0263	0.0713	0.0410	0.0355	0.1092	0.1077	0.0555

samples cover 25 provinces/cities/autonomous regions. In 2010, the project completed a baseline survey. All baseline family members and their future kinship/adopted children in 2010 are defined as permanently tracked members. The project completed three tracking surveys in 2012, 2014 and 2016. The statistical indicators of the 2010 survey is not consistent with those of the other three surveys. Considering consistency and comparability, I just use the survey data of 2012, 2014 and 2016. Additionally, I solely focus on traceable adult (aged 16 and above) individual data.

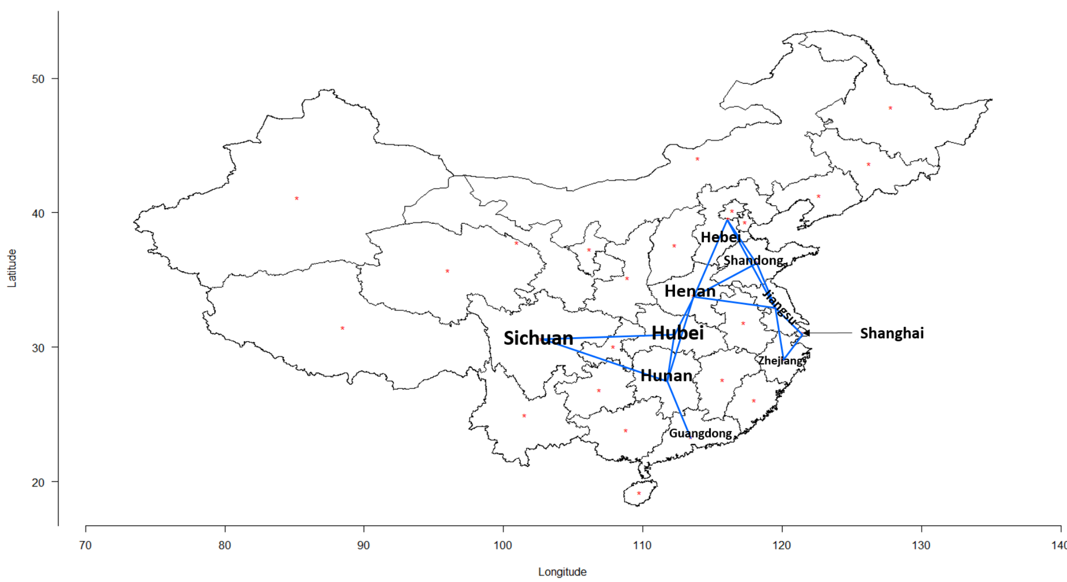


Figure 4.2: Ten selected provinces and adjacent provinces within 210km for the DMSOP model

The data of ten provinces in China (Hebei, Shanghai, Jiangsu, Zhejiang, Shandong, Henan, Hubei, Hunan, Guangdong, and Sichuan) is used, as shown in Figure 4.2. These ten provinces are selected based on the GDP ranking of provinces in 2018. After deleting the missing data, there are $2,520$ valid survey samples. Additionally, the weights w_{ij} in this empirical study are quantified based on distances, which are calculated from the latitude and longitude coordinates of the geometric center points. The greater the distance, the smaller the weight. All provinces within 210km are assumed to be adjacent, and define $w_{ij} = 1/\text{distance}$ for regions that are adjacent; otherwise, $w_{ij} = 0$. After that, the normalized weight matrix (row sum is 1) can be easily obtained as follows:

$$\boldsymbol{W} = \begin{pmatrix} 0 & 0 & 0.24 & 0 & 0.48 & 0.28 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.47 & 0.53 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.11 & 0.31 & 0 & 0.20 & 0.22 & 0.16 & 0 & 0 & 0 & 0 \\ 0 & 0.63 & 0.37 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.35 & 0 & 0.36 & 0 & 0 & 0.29 & 0 & 0 & 0 & 0 \\ 0.16 & 0 & 0.19 & 0 & 0.21 & 0 & 0.30 & 0.14 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.44 & 0 & 0.40 & 0 & 0.16 \\ 0 & 0 & 0 & 0 & 0 & 0.20 & 0.37 & 0 & 0.28 & 0.15 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.00 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.51 & 0.49 & 0 & 0 \end{pmatrix}.$$

Table 4.2 explains all the variables used in the DMSOP model. In the model, the response variables are life satisfaction and self-rated health, and the explanatory variables include age, gender, urban, education, employment and income. Here, age, education and income are continuous variables. The other explanatory variables are discrete variables.

Table 4.3 provides the summary statistics of these variables, and shows that the worst and best situations for life satisfaction and self-rated health have lower frequencies, while the general situation has the highest frequency. With time, the income and education levels have improved. Moreover, rural residents have gradually become urban residents, and unemployment has eased.

Table 4.2: Variables in the empirical DMSOP model

Variable	Description	Value
Dependent Variables		
Life Satisfaction	Satisfaction with life for the interviewees.	Very dissatisfied is 1; dissatisfied is 2; general is 3; satisfied is 4; and very satisfied is 5.
Self-rated Health	Self-evaluation of the health condition.	Very poor is 1; poor is 2; general is 3; good is 4; and very good is 5.
Explanatory variables		
Age	Age of the interviewees.	Change to log(age).
Gender	Gender of the interviewees.	Male is 1; female is 0.
Urban	Census register with categories 'urban' and 'rural'.	Urban is 1; rural is 0.
Education	One's education years.	From 0 to 19.
Employment	Employment status with categories 'employment' and 'Unemployment'.	Unemployment is 0; employment is 1.
Income	Annual income in last year.	Change to log(income).

4.4.2 Posterior Results of Parameters

In the empirical study, the priors are assigned the same as in the simulation part. MCMC simulations are run for $3,000$ iterations, and the first $1,000$ samples are discarded as the burn-in period.

From Table 4.4, the three explanatory variables of age, education and employment have evident effects on both response variables, but the directions are different among them. In terms of age, self-rated health decreased with age, but life satisfaction increased with age. This finding is compatible with Rodgers et al. (2017). In terms of education, one would believe that high education levels may have a positive impact, as in Fernández-Ballesteros et al. (2001). In this empirical study, high education levels illustrate the opposite effect. Similarly, Melin et al. (2003) also discovered this seemingly abnormal

Table 4.3: Summary statistics of empirical data

Dependent variables:					
	life satisfaction				
wave	1	2	3	4	5
1	120(0.05)	269(0.11)	1143(0.45)	653(0.26)	335(0.13)
2	45(0.02)	125(0.05)	749(0.30)	949(0.38)	652(0.26)
3	67(0.03)	171(0.07)	949(0.38)	777(0.31)	556(0.22)
	self-rated health				
wave	1	2	3	4	5
1	181(0.07)	448(0.18)	975(0.39)	626(0.25)	290(0.12)
2	164(0.07)	353(0.14)	1011(0.40)	606(0.24)	386(0.15)
3	184(0.07)	451(0.18)	1038(0.41)	470(0.19)	377(0.15)
Explanatory variables:					
Time-invariant	gender				
	0	1			
	1031	1489			
	log(age)				
wave	Mean	Sd	Min	Median	Max
1	1.57	0.14	1.20	1.59	1.89
2	1.59	0.13	1.26	1.61	1.90
3	1.61	0.13	1.30	1.63	1.91
	log(income)				
wave	Mean	Sd	Min	Median	Max
1	3.16	1.90	0.00	4.18	6.26
2	3.60	1.67	0.00	4.30	5.58
3	3.87	1.49	0.00	4.40	6.02
	education				
wave	Mean	Sd	Min	Median	Max
1	9.33	4.24	0.00	9.00	18.00
2	9.76	4.09	0.00	9.00	19.00
3	9.79	4.12	0.00	9.00	19.00
	urban			employment	
wave	0	1		0	1
1	1034	1486		346	2174
2	948	1572		104	2416
3	882	1638		52	2468

[1] The value in $(\cdot)$ denotes the frequency.

effect, and found that university education was negatively associated with sexual and family life satisfaction. In terms of employment, life satisfaction and self-rated health in employed people are more likely to be higher than in unemployed people.

Table 4.4: Posterior Estimates of the empirical DMSOP model

Parameter		Life Satisfaction				Self-rated Health		
		mean	sd	median		mean	sd	median
$\boldsymbol{\beta}$	β_{11} (Age)	0.0843	0.0383	0.0822**	β_{21} (Age)	-0.5589	0.0359	-0.5580**
	β_{12} (Gender)	-0.0004	0.0111	-0.0004	β_{22} (Gender)	0.0528	0.0108	0.0528**
	β_{13} (Urban)	-0.0045	0.0118	-0.0048	β_{23} (Urban)	0.0146	0.0118	0.0144
	β_{14} (Education)	-0.0034	0.0014	-0.0035**	β_{24} (Education)	-0.0045	0.0014	-0.0045**
	β_{15} (Employ)	0.0645	0.0301	0.0643**	β_{25} (Employ)	0.0595	0.0295	0.0599**
	β_{16} (Income)	-0.0010	0.0034	-0.0010	β_{26} (Income)	0.0040	0.0034	0.0040
$\boldsymbol{\delta}$	δ_{11}(Hebei)	0.3363	0.0681	0.3415**	δ_{12}(Hebei)	1.1739	0.0627	1.1750**
	δ_{21}(Shanghai)	0.3269	0.0711	0.3316**	δ_{22}(Shanghai)	1.1105	0.0642	1.1120**
	δ_{31}(Jiangsu)	0.3630	0.0706	0.3689**	δ_{32}(Jiangsu)	1.1258	0.0641	1.1265**
	δ_{41}(Zhejiang)	0.3102	0.0702	0.3157**	δ_{42}(Zhejiang)	1.1220	0.0643	1.1236**
	δ_{51}(Shandong)	0.3803	0.0704	0.3868**	δ_{52}(Shandong)	1.1899	0.0638	1.1928**
	δ_{61}(Henan)	0.3287	0.0691	0.3331**	δ_{62}(Henan)	1.1724	0.0625	1.1741**
	δ_{71}(Hubei)	0.3146	0.0727	0.3191**	δ_{72}(Hubei)	1.0854	0.0678	1.0856**
	δ_{81}(Hunan)	0.3003	0.0724	0.3067**	δ_{82}(Hunan)	1.1310	0.0653	1.1322**
	δ_{91}(Guangdong)	0.2543	0.0693	0.2610**	δ_{92}(Guangdong)	1.1363	0.0632	1.1391**
	δ_{101}(Sichuan)	0.3085	0.0703	0.3124**	δ_{102}(Sichuan)	1.0839	0.0632	1.0850**
$\boldsymbol{\Lambda}$	λ_{11}	0.3345	0.0105	0.3346**	λ_{12}	0.0672	0.0062	0.0670**
	λ_{21}	0.0373	0.0071	0.0370**	λ_{22}	0.4183	0.0093	0.4183**
ρ		0.9221	0.0425	0.9321**				
$\boldsymbol{\sigma}^2$	σ_1^2	0.0016	0.0013	0.0012**	σ_2^2	0.0128	0.0167	0.0075**
$\boldsymbol{V}$	v_{11}	0.1203	0.0028	0.1203**	v_{12}	0.0291	0.0017	0.0291**
	v_{21}	0.0291	0.0017	0.0291**	v_{22}	0.1232	0.0026	0.1232**
$\boldsymbol{\gamma}$	$\boldsymbol{\gamma}_{1,2}$	0.2265	0.0074	0.2268**	$\boldsymbol{\gamma}_{2,2}$	0.2959	0.0061	0.2956**
	$\boldsymbol{\gamma}_{1,3}$	0.6731	0.0057	0.6732**	$\boldsymbol{\gamma}_{2,3}$	0.7137	0.0052	0.7136**
$\boldsymbol{\beta}_0$	β_{110} (Age)	0.1583	0.0434	0.1548**	β_{210} (Age)	-0.4740	0.0393	-0.4742**
	β_{120} (Gender)	-0.0462	0.0152	-0.0462**	β_{220} (Gender)	0.0967	0.0152	0.0964**
	β_{130} (Urban)	0.0051	0.0159	0.0050	β_{230} (Urban)	-0.0015	0.0160	-0.0015
	β_{140} (Education)	0.0056	0.0019	0.0056**	β_{240} (Education)	0.0084	0.0019	0.0085**
	β_{150} (Employ)	-0.0439	0.0230	-0.0433*	β_{250} (Employ)	0.0306	0.0231	0.0300
	β_{160} (Income)	0.0092	0.0043	0.0091**	β_{260} (Income)	-0.0012	0.0042	-0.0012

[1] "$**$" and "$*$" denote that zero is not included in the 95% and 90% credible intervals, respectively. Mean, Sd and Median are posterior mean, posterior standard deviation and posterior median, respectively.

In this empirical study, $\boldsymbol{\delta}$ is a 20 ($m = 10,\ S = 2$)-dimensional vector of the regional effects. $\boldsymbol{\delta}_{i1}$ and $\boldsymbol{\delta}_{i2}$ represent all unobserved common regional effects for the observations within region $i\,(i = 1, \ldots, 10)$, that is, for self-rated health and life satisfaction, respectively. From Table 4.4, the values of vector $\boldsymbol{\delta}_{i1}$ and $\boldsymbol{\delta}_{i2}$ are all relatively large; that is, the regional effects are indispensable on the response variables. Additionally, the regional effects of self-rated health are always greater than those of life satisfaction in the same region.

The posterior mean of the temporal coefficients λ_{11} and λ_{21} are 0.3345 and 0.0373, respectively. That is, the temporal effect of life satisfaction in the last period $(z_{ik1(t-1)})$ on life satisfaction is 0.3345, and on self-rated health is 0.0373. Similarly, temporal effect of self-rated health in the last period $(z_{ik2(t-1)})$ on life satisfaction is 0.0672, and on self-rated health is 0.4183. The temporal coefficients λ_{11} and λ_{22} are large, and thus, the temporal dependencies in this empirical study are critical. Furthermore, temporal coefficients λ_{11} and λ_{22} are much larger than λ_{12} and λ_{21}, and thus, self-temporal effects are much stronger than mutual temporal effects.

The posterior mean of spatial coefficient ρ is 0.9221. It is large and positive, and thus, the spatial dependencies in this empirical study are critical; one province's regional effect has a positive effect on the other provinces' regional effects.

$\sigma_1^2 = 0.0016$ and $\sigma_2^2 = 0.0128$ are the variances of the disturbances of life satisfaction and self-rated health, respectively. The variances are very small, that is, the disturbances vary slightly.

$v_{11} = 0.1203$ and $v_{22} = 0.1232$ denote the variances of the individualistic effect for life satisfaction and self-rated health, respectively; $v_{12}/v_{21} = 0.0291$ denotes the covariance of the individualistic effects. Since v_{12}/v_{21} are evident and positive, the individualistic effects for life satisfaction and self-rated health are positively correlated.

4.5 Chapter Summary

This chapter proposed a new algorithm for a DMSOP model, wherein temporal and spatial interactions are captured simultaneously for the multivariate discrete ordered responses. The parameters are estimated using the Bayesian inference based on MCMC sampling. The core of Bayesian inference is to obtain the full conditional posterior distributions. In this chapter, I highlight the estimation method of $\boldsymbol{\Lambda}$ and $\boldsymbol{\gamma}$. We cannot obtain all the values in the temporal matrix $\boldsymbol{\Lambda}$ at once, and the λ_{sg} have to be estimated individually. Thus, conditional multivariate normal distribution of the latent response variable z_{ikst} is necessary. In previous studies on Bayesian spatial econometrics (Kakamu and Wago, 2007; Smith and LeSage, 2004; Wang and Kockelman, 2009c), $\boldsymbol{\gamma}$ is estimated using the uniform distribution. Here, I use a more stable method, as in Chen and Dey (2000), which removes the ordering constraint using the one-to-one map. The validity and accuracy of the model is verified by the simulated dataset. Additionally, the DMSOP model is applied to the survey data from CFPS.

5

Conclusions

This study mainly focuses on three types of models: an empirical STAR model, a newly proposed MSOP model, and an extended model of the MSOP model, the DMSOP model. Due to the excellent performance of Bayesian inference in spatial econometrics, all the parameters in the models are calculated using the Bayesian inference.

In Chapter 2, the following factors are hypothesized to be important factors affecting the prevalence of chronic diseases and specific chronic diseases in elderly Chinese people:

gender, census register (urban or rural), marital status, smoking, drinking, age, education years, sports activities and cultural activities, and province. In addition, because most of the samples between surveys in 2006 and 2010 are the same, the effect of time is taken into account. Because the STAR models can combine spatial effects, nonlinear effects of continuous factors, and the linear or fixed effects into a single model, a logistic STAR model is applied to verify the hypotheses. In applying this model, an `R` package called `R2BayesX` (Umlauf et al., 2015) is used to fit the STAR models. The empirical results show that except for the effect of marital status is negligible, the effects of other factors are considerable. Among them, the effect of province and time are particularly obvious.

In Chapter 3, this study proposes a MSOP model, which can captures the interactions of multivariate ordinal responses and spatial dependencies of dependent variables. In applying this model, the parameters are calculated using the Bayesian inference based on MCMC sampling. The prior distributions of all the other parameters are similar to those in former studies on Bayesian spatial econometrics (Kakamu and Wago, 2007; Smith and LeSage, 2004; Wang and Kockelman, 2009a,b,c) except for cutpoints. In previous studies, cutpoints are estimated using the uniform distributions. Here, a more stable method is used, as in Chen and Dey (2000), which removes the ordering constraint using the one-to-one map. To verify the validity and accuracy of the MSOP model, this study performs a simulation study. The estimation results are quite close to the true values. Therefore, the MSOP model performs very well with the simulated data. In addition, this study illustrates the MSOP model by applying it to two subjective ordered discrete variables, self-rated health and life satisfaction of elderly people in six provinces in East China. The empirical results show that the spatial dependencies in this empirical study are critical, and the individualistic effects for self-rated health and life satisfaction are positively related.

In Chapter 4, this study extends the MSOP model to DMSOP model, wherein temporal and spatial dependencies can be captured simultaneously for the multivariate ordinal responses. The parameters are estimated using the Bayesian inference. The core of Bayesian inference is to obtain the full conditional posterior distributions. Here, the estimation method of temporal coefficient matrix and cutpoints should be highlighted. The values in the temporal matrix cannot be obtained at once, and thus have to be estimated individually using the conditional multivariate normal distribution of the latent response variable. The estimation method of cutpoints is the same as in MSOP model, which removes the ordering constraint using the one-to-one map. The validity and accuracy of the model is verified by the simulated dataset. The DMSOP model performs effectively with the simulated data. It successfully explains the temporal and spatial dependencies, and the interaction of multivariate ordinal response variables as well as the effects of the explanatory variables. Additionally, the study applied the DMSOP model to the survey data from China Family Panel Studies. The spatial coefficient and the temporal coefficient are considerable, and thus the spatial and temporal dependencies in this empirical study are critical.

In this study, Bayesian method was used to explore the above three spatial econometric models. Nevertheless, there are still some remaining issues and future research about this study.

First, the nonlinear effects in empirical STAR model are considerable, but very small. Thus, I should build a better method to reconsider the nonlinear effects.

Next, The MSOP model includes spatial dependencies of dependent variables. Although the MSOP model performs very well with the simulated data, comparison should be implemented with the multivariate OP models that do not contain spatial dependencies. Therefore, a model comparison should be conducted in future work.

In addition, the MSOP model in this study only addresses inter-regional spatial correlation, and simplifies the intra-regional spatial correlation to a shared random effect. It is easy to think of reasons for intraregional interaction: people live close to each other and know each other, and thus a spatial model that formulates spatial dependence among intraregional agents may needed. In this case, the intra-regional interaction and inter-regional interaction could be captured by the following models:

$$\boldsymbol{z}_{iks} = \rho_1 \sum_{r=1}^{n_i} w_{kr}^1 z_{irs} + \boldsymbol{x}_{iks}' \boldsymbol{\beta}_s + \delta_{is} + \xi_{iks}, \tag{5.1}$$

$$\delta_{is} = \rho_2 \sum_{j=1}^{m} w_{ij}^2 \delta_{js} + \epsilon_{is}, \tag{5.2}$$

where ρ_1 is the intraregional spatial coefficient, weights w_{kr}^1 reflect the closeness between individuals k and r inside region i; ρ_2 is the interregional spatial coefficient, weights w_{ij}^2 reflect the closeness between regions i and j. z_{iks} is the latent response variable, δ_{is} represents all unobserved common effects for observations within region i for the sth outcome, and ξ_{iks} denotes the individualistic effect of individual k in region i for the sth outcome.

The above models (5.1) and (5.2) are more comprehensive than the MSOP model in this study. They can capture not only the interregional interaction, but also intraregional interaction. In multivariate spatial econometrics, this study is only an attempt. I will try to use the above models in future works to capture the interregional interaction and intraregional interaction simultaneously for the multivariate ordinal responses.

Furthermore, in DMSOP model, all unobserved common effects for observations within region i for the sth outcome in period t, the regional effects δ_{ist} is assumed to be time-invariant: $\delta_{ist} \equiv \delta_{is}$ for all $t = 0, 1, \ldots, T$ as in Wang and Kockelman (2009c). In practice, the regional effect δ_{ist} changed over time may be more common and reasonable. The regional effects changed over time can be described as follows:

$$\delta_{ist} = \rho \sum_{j=1}^{m} w_{ij} \delta_{jst} + \epsilon_{ist}, \ \ \epsilon_{ist} \sim \mathrm{N}(0, \boldsymbol{\sigma}_s^2), \tag{5.3}$$

where i denotes spatial regions ($i = 1, \ldots, m$), s denotes the sequence of observed response variables ($s = 1, \ldots, S$), and t denotes time periods ($t = 0, \ldots, T$). ρ is the spatial coefficient, which reflects the degree of spatial influence. Spatial weights w_{ij} reflect the closeness between regions i and j.

In this alternative model (5.3), the regional effects δ_{ist} is time variant. I will try to use Bayesian inference to estimate this time variant model, and compare the accuracy of the estimates of this DMSOP model and the DMSOP model proposed in this study.

A

Full Conditional Distribution and Marginal Effects of MSOP Model

A.1 Full Conditional Distribution of Parameters

A.1.1 The Full Conditional Distribution of $\boldsymbol{\beta}$

Let $\star$ denotes all the other parameters except the given parameters, and then the full conditional distribution (FCD) of $\boldsymbol{\beta}$ is expressed as follows:

$$
\begin{aligned}
&p(\boldsymbol{\beta}|\star) \propto p(\boldsymbol{z}|\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V})\pi(\boldsymbol{\beta}) \\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})\right\} \\
&\quad \times \exp\left[-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{c})'\boldsymbol{T}^{-1}(\boldsymbol{\beta} - \boldsymbol{c})\right] \\
&\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})\boldsymbol{X}\boldsymbol{\beta} - 2\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}) + \boldsymbol{\beta}'\boldsymbol{T}^{-1}\boldsymbol{\beta} - 2\boldsymbol{\beta}'\boldsymbol{T}^{-1}\boldsymbol{c}\right]\right\} \\
&= \exp\left\{-\frac{1}{2}\boldsymbol{\beta}'[\boldsymbol{X}'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})\boldsymbol{X} + \boldsymbol{T}^{-1}]\boldsymbol{\beta} + \boldsymbol{\beta}'\left[\boldsymbol{X}'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}) + \boldsymbol{T}^{-1}\boldsymbol{c}\right]\right\}.
\end{aligned}
$$

Defining $\boldsymbol{B} = \boldsymbol{X}'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})\boldsymbol{X} + \boldsymbol{T}^{-1}$, and $\boldsymbol{b} = \boldsymbol{X}'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z} - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}) + \boldsymbol{T}^{-1}\boldsymbol{c}$, the FCD of $\boldsymbol{\beta}$ can be expressed as:

$$
\boldsymbol{\beta}|\boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \gamma, \boldsymbol{z}, \boldsymbol{y} \sim \mathrm{N}(\boldsymbol{B}^{-1}\boldsymbol{b}, \boldsymbol{B}^{-1}).
$$

A.1.2 The Full Conditional Distribution of $\boldsymbol{\delta}$

The FCD of $\boldsymbol{\delta}$ is expressed as follows:

$$\begin{aligned}
p(\boldsymbol{\delta}|\star) &\propto p(\boldsymbol{z}|\boldsymbol{\beta},\boldsymbol{\delta},\boldsymbol{V})p(\boldsymbol{\delta}|\rho,\boldsymbol{\sigma}^2) \\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{z}-\boldsymbol{X\beta}-(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)\boldsymbol{\delta})'(\boldsymbol{I}_n\otimes\boldsymbol{V}^{-1})(\boldsymbol{z}-\boldsymbol{X\beta}-(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)\boldsymbol{\delta})\right\} \\
&\quad\times\exp\left(-\frac{1}{2}\boldsymbol{\delta}'\boldsymbol{B}_\rho'(\boldsymbol{I}_m\otimes\mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho\boldsymbol{\delta}\right) \\
&\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\delta}'(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)'(\boldsymbol{I}_n\otimes\boldsymbol{V}^{-1})(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)\boldsymbol{\delta}-2\boldsymbol{\delta}'(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)'(\boldsymbol{I}_n\otimes\boldsymbol{V}^{-1})(\boldsymbol{z}-\boldsymbol{X\beta})\right.\right. \\
&\qquad\qquad\left.\left.+\boldsymbol{\delta}'\boldsymbol{B}_\rho'(\boldsymbol{I}_m\otimes\mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho\boldsymbol{\delta}\right]\right\} \\
&= \exp\left\{-\frac{1}{2}\boldsymbol{\delta}'\left[(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)'(\boldsymbol{I}_n\otimes\boldsymbol{V}^{-1})(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)+\boldsymbol{B}_\rho'(\boldsymbol{I}_m\otimes\mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho\right]\boldsymbol{\delta}\right. \\
&\qquad\left.+\boldsymbol{\delta}'(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)'(\boldsymbol{I}_n\otimes\boldsymbol{V}^{-1})(\boldsymbol{z}-\boldsymbol{X\beta})\right\}.
\end{aligned}$$

Defining $\boldsymbol{\Omega}=(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)'(\boldsymbol{I}_n\otimes\boldsymbol{V}^{-1})(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)+\boldsymbol{B}_\rho'(\boldsymbol{I}_m\otimes\mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho$, and $\boldsymbol{\eta}=(\boldsymbol{\Delta}\otimes\boldsymbol{I}_S)'(\boldsymbol{I}_n\otimes\boldsymbol{V}^{-1})(\boldsymbol{z}-\boldsymbol{X\beta})$, the FCD of $\boldsymbol{\delta}$ can be expressed as:

$$\boldsymbol{\delta}|\boldsymbol{\beta},\rho,\boldsymbol{\sigma}^2,\boldsymbol{V},\gamma,\boldsymbol{z},\boldsymbol{y}\sim\mathrm{N}(\boldsymbol{\Omega}^{-1}\boldsymbol{\eta},\boldsymbol{\Omega}^{-1}),$$

where $\boldsymbol{\Omega}^{-1}$ depends on $\boldsymbol{B}_\rho$, which depends on ρ.

A.1.3 The Full Conditional Distribution of ρ

The FCD of ρ is expressed as follows:

$$\begin{aligned}
p(\rho|\star) &\propto p(\boldsymbol{\delta}|\rho,\boldsymbol{\sigma}^2)\pi(\rho) \\
&\propto |\boldsymbol{B}_\rho|\exp\left(-\frac{1}{2}\boldsymbol{\delta}'\boldsymbol{B}_\rho'(\boldsymbol{I}_m\otimes\mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho\boldsymbol{\delta}\right)1(\lambda_{min}^{-1}\leq\rho\leq\lambda_{max}^{-1}).
\end{aligned}$$

Since $p(\rho|\star)$ is not a standard distribution, we can use a M-H algorithm to obtain the samples.

1) Sample ρ^{new} from the candidate density: $\rho^{new} = \rho^{old} + aG$, where a is a known constant, and G can be a standard normal distribution or t distribution. In this study, we set $G \sim \mathrm{N}(0, 1)$.

2) Since $\rho \in [\lambda_{min}^{-1}, \lambda_{max}^{-1}]$, if proposed ρ^{new} is not within the interval, it is rejected with probability 1. Otherwise, we calculate the acceptance probability $\alpha(\rho^{new},\ \rho^{old}) = \min(\frac{p(\rho^{new}|\star)}{p(\rho^{old}|\star)}, 1)$.

3) Draw a uniform random deviate U. If $U < \alpha$, we set $\rho = \rho^{new}$, otherwise, $\rho = \rho^{old}$.

4) Calculate the updated ratio of ρ. Then, a is adjusted based on the updated ratio. If the updated ratio is less than 0.4, and then $a = a/1.1$; and if the updated ratio is less than 0.6, and then $a = a * 1.1$. Thus, the updated ratio is always close to 0.5.

A.1.4 The Full Conditional Distribution of $\boldsymbol{\sigma}^2$

The FCD of $\boldsymbol{\sigma}^2$ is expressed as follows:

$$p(\boldsymbol{\sigma}^2|\star) \propto p(\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2)\pi(\boldsymbol{\sigma}^2).$$

Given $\boldsymbol{\beta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}$ and $\boldsymbol{y}$, $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \ldots, \boldsymbol{\delta}_S$ are independent, and thus

$$\begin{aligned} p(\boldsymbol{\sigma}^2|\star) &\propto \prod_{s=1}^{S} \left\{ p(\boldsymbol{\delta}_s|\rho, \boldsymbol{\sigma}^2)\pi(\sigma_s^2) \right\} \\ &\propto \prod_{s=1}^{S} \left\{ (\sigma_s^2)^{-m/2} \exp\left(-\frac{1}{2\sigma_s^2}\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s \right) (\sigma_s^2)^{-(e+1)} \exp\left(-\frac{g}{\sigma_s^2} \right) \right\} \\ &\propto \prod_{s=1}^{S} \left\{ (\sigma_s^2)^{-(m/2+e+1)} \exp\left(-\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g}{2\sigma_s^2} \right) \right\}. \end{aligned}$$

Extract the element σ_s^2 from $\boldsymbol{\sigma}^2$, and then

$$p(\sigma_s^2|\star) \propto (\sigma_s^2)^{-(m/2+e+1)} \exp\left(-\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g}{2\sigma_s^2}\right).$$

Let $\theta_s = \frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s+2g}{\sigma_s^2}$, and thus $\sigma_s^2 = \frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s+2g}{\theta_s}$. Therefore, $\left|\frac{\mathrm{d}\,\sigma_s^2}{\mathrm{d}\,\theta_s}\right| = \frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s+2g}{\theta_s^2}$.

$$\begin{aligned}
p(\theta_s|\star) &= p(\sigma_s^2(\theta_s)|\star)\left|\frac{\mathrm{d}\,\sigma_s^2}{\mathrm{d}\,\theta_s}\right| \\
&\propto (\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g}{\theta_s})^{-(m/2+e+1)} \exp\left(-\frac{\theta_s}{2}\right)\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g}{\theta_s^2} \\
&\propto \theta_s^{\frac{m+2e}{2}-1}\exp\left(-\frac{\theta_s}{2}\right).
\end{aligned}$$

Therefore, the FCD of σ_s^2 can be expressed as:

$$\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g}{\sigma_s^2}|\boldsymbol{\beta},\boldsymbol{\delta},\rho,\boldsymbol{\sigma}_{-s}^2,\boldsymbol{V},\boldsymbol{\gamma},\boldsymbol{z},\boldsymbol{y} \sim \chi^2(m+2e).$$

More specifically, if $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_S^2 = \sigma_0^2$, and then

$$\begin{aligned}
p(\sigma_0^2|\star) &\propto p(\boldsymbol{\delta}|\rho,\sigma_0^2)\pi(\sigma_0^2) \\
&\propto (\sigma_0^2)^{-mS/2}\exp\left(-\frac{1}{2\sigma_0^2}\boldsymbol{\delta}_s'\boldsymbol{B}_\rho'\boldsymbol{B}_\rho\boldsymbol{\delta}_s\right)(\sigma_0^2)^{-(e+1)}\exp\left(-\frac{g}{\sigma_0^2}\right) \\
&\propto (\sigma_0^2)^{-(mS/2+e+1)}\exp\left(-\frac{\boldsymbol{\delta}'\boldsymbol{B}_\rho'\boldsymbol{B}_\rho\boldsymbol{\delta} + 2g}{2\sigma_0^2}\right).
\end{aligned}$$

Similarly, the FCD of σ_0^2 can be expressed as:

$$\frac{\boldsymbol{\delta}'\boldsymbol{B}_\rho'\boldsymbol{B}_\rho\boldsymbol{\delta} + 2g}{\sigma_0^2}|\boldsymbol{\beta},\boldsymbol{\delta},\rho,\boldsymbol{V},\boldsymbol{\gamma},\boldsymbol{z},\boldsymbol{y} \sim \chi^2(mS+2e).$$

A.1.5 The Full Conditional Distribution of $\boldsymbol{V}$

The FCD of $\boldsymbol{V}$ is expressed as follows:

$$
\begin{aligned}
p(\boldsymbol{V}|\star) &= p(\boldsymbol{z}|\boldsymbol{\beta},\boldsymbol{\delta},\boldsymbol{V})\pi(\boldsymbol{V}) \\
&\propto \prod_{i=1}^{m}\prod_{k=1}^{n_i}\left\{|\boldsymbol{V}|^{-1/2}\exp\left[-\frac{1}{2}(\boldsymbol{z}_{ik}-\boldsymbol{X}_{ik}\boldsymbol{\beta}-\boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik}-\boldsymbol{X}_{ik}\boldsymbol{\beta}-\boldsymbol{\delta}_i)\right]\right\} \\
&\quad \times |\boldsymbol{V}|^{-\frac{n_0+S+1}{2}}\exp\left[-\frac{1}{2}\operatorname{tr}(\boldsymbol{Q}_0\boldsymbol{V}^{-1})\right] \\
&= |\boldsymbol{V}|^{-\frac{n+n_0+S+1}{2}}\exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ik}-\boldsymbol{X}_{ik}\boldsymbol{\beta}-\boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik}-\boldsymbol{X}_{ik}\boldsymbol{\beta}-\boldsymbol{\delta}_i)+\operatorname{tr}(\boldsymbol{Q}_0\boldsymbol{V}^{-1})\right]\right\} \\
&= |\boldsymbol{V}|^{-\frac{n+n_0+S+1}{2}}\exp\left\{-\frac{1}{2}\operatorname{tr}\left[\left(\sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ik}-\boldsymbol{X}_{ik}\boldsymbol{\beta}-\boldsymbol{\delta}_i)(\boldsymbol{z}_{ik}-\boldsymbol{X}_{ik}\boldsymbol{\beta}-\boldsymbol{\delta}_i)'+\boldsymbol{Q}_0\right)\boldsymbol{V}^{-1}\right]\right\}.
\end{aligned}
$$

Defining $\tilde{n}=n+n_0$, and $\tilde{\boldsymbol{Q}}=\sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ik}-\boldsymbol{X}_{ik}\boldsymbol{\beta}-\boldsymbol{\delta}_i)(\boldsymbol{z}_{ik}-\boldsymbol{X}_{ik}\boldsymbol{\beta}-\boldsymbol{\delta}_i)'+\boldsymbol{Q}_0$, the FCD of $\boldsymbol{V}$ can be expressed as:

$$
\boldsymbol{V}|\boldsymbol{\beta},\boldsymbol{\delta},\rho,\boldsymbol{\sigma}^2,\boldsymbol{\gamma},\boldsymbol{z},\boldsymbol{y}\sim \mathrm{IW}(\tilde{n},\tilde{\boldsymbol{Q}}).
$$

A.1.6 The Full Conditional Distribution of $\boldsymbol{z}$

Let $\boldsymbol{z}_s=(z_{11s},\ldots,z_{1n_is},\ldots,z_{m1s},\ldots,z_{mn_ms})'$ denotes the vector of the sth element z_{iks} from $\boldsymbol{z}_{ik}(i=1,\ldots,m;k=1,\ldots,n_i)$, and let $\boldsymbol{z}_{-s}$ denotes the vector removing $\boldsymbol{z}_s$ from $\boldsymbol{z}$. Similarly, let $z_{ik,-s}$ denotes the vector removing z_{iks} from $\boldsymbol{z}_{ik}$. Furthermore, let $\boldsymbol{\gamma}=(\boldsymbol{\gamma}_1,\ldots,\boldsymbol{\gamma}_s,\ldots,\boldsymbol{\gamma}_S)'$ and $\boldsymbol{\gamma}_s=(\gamma_{s,2},\ldots,\gamma_{s,Ls-2})'$, and then $\boldsymbol{\gamma}_{-s}$ denotes the vector removing $\boldsymbol{\gamma}_s$ from $\boldsymbol{\gamma}$. We can get the joint conditional distribution of $\boldsymbol{\gamma}_s$ and $\boldsymbol{z}_s$:

$$
\begin{aligned}
& p(\boldsymbol{\gamma}_s, \boldsymbol{z}_s | \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \gamma_{-s}, \boldsymbol{z}_{-s}, \boldsymbol{y}) \\
& = p(\boldsymbol{z}_s | \gamma_s, \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}_{-s}, \boldsymbol{z}_{-s}, \boldsymbol{y}) \times p(\gamma_s | \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \gamma_{-s}, \boldsymbol{z}_{-s}, \boldsymbol{y}).
\end{aligned}
$$

Furthermore, given $\boldsymbol{\beta}, \boldsymbol{\delta}$ and $\boldsymbol{V}$, $\boldsymbol{z}_{ik}$ is independent of $\boldsymbol{z}_{-ik}$, and thus

$$
\begin{aligned}
p(\boldsymbol{z}_s | \gamma_s, \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \gamma_{-s}, \boldsymbol{z}_{-s}, \boldsymbol{y}) &= \prod_{i=1}^{m} \prod_{k=1}^{n_i} p(z_{iks} | \boldsymbol{\gamma}_s, \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}_{-s}, \boldsymbol{z}_{-s}, \boldsymbol{y}) \\
= \prod_{i=1}^{m} \prod_{k=1}^{n_i} p(\boldsymbol{z}_{iks} | \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \gamma, \boldsymbol{z}_{-s}, \boldsymbol{y}) &= \prod_{i=1}^{m} \prod_{k=1}^{n_i} p(\boldsymbol{y}_{ik} | \gamma, \boldsymbol{z}_{ik}) p(\boldsymbol{z}_{iks} | \boldsymbol{z}_{-s}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}).
\end{aligned}
$$

Defining $\boldsymbol{h}_{ik} = \boldsymbol{X}_{ik}\boldsymbol{\beta} + \boldsymbol{\delta}_i$, because $\boldsymbol{z}_{ik} \sim \mathrm{N}(\boldsymbol{X}_{ik}\boldsymbol{\beta} + \boldsymbol{\delta}_i, \boldsymbol{V})$, and then $\boldsymbol{z}_{ik} \sim \mathrm{N}(\boldsymbol{h}_{ik}, \boldsymbol{V})$. Furthermore, for the sake of convenience of expression, we reorganize $\boldsymbol{z}_{ik}$, $\boldsymbol{h}_{ik}$ and $\boldsymbol{V}$ as follows (the sth element become the first element):

$$
\boldsymbol{z}_{ik} = \begin{pmatrix} z_{iks} \\ \boldsymbol{z}_{ik,-s} \end{pmatrix}, \ \boldsymbol{h}_{ik} = \begin{pmatrix} h_{iks} \\ \boldsymbol{h}_{ik,-s} \end{pmatrix}, \ \text{and} \ \boldsymbol{V} = \begin{pmatrix} V_{s,s} & \boldsymbol{v}_{s,-s} \\ \boldsymbol{v}_{-s,s} & \boldsymbol{V}_{-s,-s} \end{pmatrix},
$$

where $\boldsymbol{v}_{s,-s} = (V_{s,1}, \ldots, V_{s,s-1}, V_{s,s+1}, \ldots, V_{s,S})$ is a $1 \times (S-1)$ vector, $\boldsymbol{v}_{-s,s} = (V_{1,s}, \ldots, V_{s-1,s}, V_{s+1,s}, \ldots, V_{S,s})'$ is an $(S-1) \times 1$ vector, and $\boldsymbol{V}_{-s,-s}$ denotes an $(S-1) \times (S-1)$ matrix removing the elements $V_{s,s}$, $\boldsymbol{v}_{-s,s}$ and $\boldsymbol{v}_{s,-s}$ from $\boldsymbol{V}$.

Based on the property of multivariate normal distribution, we can obtain the conditional probability distribution of z_{iks} given $\boldsymbol{z}_{-s}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}$; that is,

$$
z_{iks} | \boldsymbol{z}_{-s}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V} \sim \mathrm{N}(\tilde{h}_{iks}, \tilde{v}_{iks}), \ \text{for}\, i = 1, \ldots, m, \text{and } k = 1, \ldots, n_i,
$$

where

$$\tilde{h}_{iks} = h_{iks} + \boldsymbol{v}_{s,-s}\boldsymbol{V}_{-s,-s}^{-1}(\boldsymbol{z}_{ik,-s} - \boldsymbol{h}_{ik,-s}),$$
$$\tilde{v}_{iks} = V_{s,s} - \boldsymbol{v}_{s,-s}\boldsymbol{V}_{-s,-s}^{-1}\boldsymbol{v}_{-s,s}.$$

Thus,

$$\begin{aligned} p(z_{iks}|\boldsymbol{\gamma}_s, \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}_{-s}, \boldsymbol{z}_{-s}, \boldsymbol{y}) &\propto p(\boldsymbol{y}_{ik}|\gamma, \boldsymbol{z}_{ik})p(\boldsymbol{z}_{iks}|\boldsymbol{z}_{ik,-s}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}) \\ &\propto 1(\boldsymbol{z}_{ik} \in G_{ik}) \exp\left[-\frac{1}{2\tilde{v}_{iks}}(z_{iks} - \tilde{h}_{iks})^2\right]. \end{aligned}$$

Therefore, the FCD of z_{iks} can be expressed as:

$$z_{iks}|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}_{-s}, \boldsymbol{y} \sim \mathrm{N}(\tilde{h}_{iks}, \tilde{v}_{iks})1(\boldsymbol{z}_{ik} \in G_{ik}).$$

A.1.7 The Full Conditional Distribution of $\boldsymbol{\gamma}$

Given $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, ρ, $\boldsymbol{\sigma}^2$, $\boldsymbol{V}$ and $\boldsymbol{\gamma}_s$, $z_{11s}, z_{12s}, \ldots, z_{mn_ms}$ are independent, and thus

$$\begin{aligned} &p(\boldsymbol{\gamma}_s|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}_{-s}, \boldsymbol{z}_{-s}, \boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}_{-s})\pi(\boldsymbol{\gamma}_s) \\ &= \prod_{i=1}^{m}\prod_{k=1}^{n_i}\sum_{l_s=2}^{L_s-1} p(y_{iks} = l_s|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}_s, \boldsymbol{z}_{-s})1(\gamma_{s,2} < \cdots < \gamma_{s,L_s-1}) \\ &= \prod_{i=1}^{m}\prod_{k=1}^{n_i}\sum_{l_s=2}^{L_s-1} 1(y_{iks} = l_s)p(\gamma_{s,l_s-1} < z_{iks} \leq \gamma_{s,l_s}|\boldsymbol{z}_{-s}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V})1(\gamma_{s,2} < \cdots < \gamma_{s,L_s-1}). \end{aligned}$$

As explained before, $z_{iks}|\boldsymbol{z}_{-s}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V} \sim \mathrm{N}(\tilde{h}_{iks}, \tilde{v}_{iks})$. Therefore,

$$p(\gamma_{s,l_s-1} < z_{iks} \leq \gamma_{s,l_s}|\boldsymbol{z}_{-s}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}) = \Phi\left(\frac{\gamma_{s,l_s} - \tilde{h}_{iks}}{\tilde{v}_{iks}^{1/2}}\right) - \Phi\left(\frac{\gamma_{s,l_s-1} - \tilde{h}_{iks}}{\tilde{v}_{iks}^{1/2}}\right).$$

As a result,

$$
\begin{aligned}
&p(\boldsymbol{\gamma}_s|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}_{-s}, \boldsymbol{z}_{-s}, \boldsymbol{y}) \\
&\propto \prod_{i=1}^{m} \prod_{k=1}^{n_i} \sum_{l_s=2}^{L_s-1} 1(y_{iks} = l_s) \left[\Phi\left(\frac{\gamma_{s,l_s} - \tilde{h}_{iks}}{\tilde{v}_{iks}^{1/2}} \right) - \Phi\left(\frac{\gamma_{s,l_s-1} - \tilde{h}_{iks}}{\tilde{v}_{iks}^{1/2}} \right) \right] 1(\gamma_{s,2} < \cdots < \gamma_{s,L_s-1}).
\end{aligned}
$$

Based on this form of $p(\boldsymbol{\gamma}|\star)$, Chen and Dey (2000) offered a better way to get the parameters $\boldsymbol{\gamma}$. This new method removes the ordering constraint by using the one-to-one map:

$$
\gamma_{s,l_s} = \frac{\gamma_{s,l_s-1} + \exp(\kappa_{s,l_s})}{1 + \exp(\kappa_{s,l_s})}, 2 < l_s < L_s - 2.
$$

Defining $\boldsymbol{\kappa}_s = (\kappa_{s,2}, \ldots, \kappa_{s,Ls-2})'$, and we assume the prior for $\boldsymbol{\kappa}_s$ follows a normal distribution: $\boldsymbol{\kappa}_s \sim \mathrm{N}(\boldsymbol{\kappa}_{s0}, \boldsymbol{D}_{s0})$. That is,

$$
\pi(\boldsymbol{\kappa}_s) \propto \exp\left[-\frac{1}{2} (\boldsymbol{\kappa}_s - \boldsymbol{\kappa}_{s0})' \boldsymbol{D}_{s0}^{-1} (\boldsymbol{\kappa}_s - \boldsymbol{\kappa}_{s0}) \right].
$$

The Jacobian of transformation from $\boldsymbol{\gamma}_s$ to $\boldsymbol{\kappa}_s$ is as follows:

$$
\left| \frac{\partial \boldsymbol{\gamma}_s}{\partial \boldsymbol{\kappa}_s'} \right| = \prod_{l_s=2}^{L_s-2} \left| \frac{\partial \gamma_{s,l_s}}{\partial \kappa_{s,l_s}} \right| = \prod_{l_s=2}^{L_s-2} \frac{(1 - \gamma_{s,l_s-1}) \exp(\kappa_{s,l_s})}{(1 + \exp(\kappa_{s,l_s}))^2}.
$$

Therefore,

$$
\begin{aligned}
p(\boldsymbol{\kappa}_s|\star) &= p(\boldsymbol{\gamma}_s|\star) \left| \frac{\partial \boldsymbol{\gamma}_s}{\partial \boldsymbol{\kappa}_s'} \right| \pi(\boldsymbol{\kappa}_s) \\
&\propto \left\{ \prod_{i=1}^{m} \prod_{k=1}^{n_i} \sum_{l_s=2}^{L_s-1} 1(y_{iks} = l_s) \left[\Phi\left(\frac{\gamma_{s,l_s} - \tilde{h}_{iks}}{\tilde{v}_{iks}^{1/2}} \right) - \Phi\left(\frac{\gamma_{s,l_s-1} - \tilde{h}_{iks}}{\tilde{v}_{iks}^{1/2}} \right) \right] \right\} \\
&\times \left\{ \prod_{l_s=2}^{L_s-2} \frac{(1 - \gamma_{s,l_s-1}) \exp(\kappa_{s,l_s})}{(1 + \exp(\kappa_{s,l_s}))^2} \right\} \exp\left[-\frac{1}{2} (\boldsymbol{\kappa}_s - \boldsymbol{\kappa}_{s0})' \boldsymbol{D}_{s0}^{-1} (\boldsymbol{\kappa}_s - \boldsymbol{\kappa}_{s0}) \right].
\end{aligned}
$$

Since $p(\boldsymbol{\kappa}_s|\star)$ is not a standard distribution, we can use the a M-H algorithm to generate $\boldsymbol{\gamma}_s$ as follows:

1) Sample $\boldsymbol{\kappa}_s|\star$ marginally of $\boldsymbol{z}_s$ by drawing $\boldsymbol{\kappa}_s^{new}$ from the candidate student-t distribution $f_T(\boldsymbol{\kappa}_s|\hat{\boldsymbol{\kappa}}_s, \hat{\boldsymbol{D}}_s, v)$, where $\hat{\boldsymbol{\kappa}}_s = \text{argmax}p(\boldsymbol{\kappa}_s|\star), \hat{\boldsymbol{D}}_s = \left\{\left[-\frac{\partial^2 \log p(\boldsymbol{\kappa}_s|\star)}{\partial\boldsymbol{\kappa}_s\partial\boldsymbol{\kappa}_s'}\right]_{\boldsymbol{\kappa}_s=\hat{\boldsymbol{\kappa}}_s}\right\}^{-1}$, and v is the freedom.

2) Given the current $\boldsymbol{\kappa}_s^{old}$ and the proposed $\boldsymbol{\kappa}_s^{new}$, accept $\boldsymbol{\kappa}_s^{new}$ with the probability:

$$\alpha(\boldsymbol{\kappa}_s^{old}, \boldsymbol{\kappa}_s^{new}) = \min\left(\frac{p(\boldsymbol{\kappa}_s^{new}|\star) f_T(\boldsymbol{\kappa}_s^{old}|\hat{\boldsymbol{\kappa}}_s, \hat{\boldsymbol{D}}_s, v)}{p(\boldsymbol{\kappa}_s^{old}|\star) f_T(\boldsymbol{\kappa}_s^{new}|\hat{\boldsymbol{\kappa}}_s, \hat{\boldsymbol{D}}_s, v)}, 1\right).$$

3) Use the one-o-one map: $\gamma_{s,l_s} = \frac{\gamma_{s,l_s-1}+\exp(\kappa_{s,l_s})}{1+\exp(\kappa_{s,l_s})}, 2 < l_s < L_s - 2$ to obtain the cutpoints $\boldsymbol{\gamma}_s = (\gamma_{s,2}, \ldots, \gamma_{s,L_s-2})'$.

A.2 Marginal Effects

Learning from A.1.7, the probability of $y_{iks} = l_s$ can be written as:

$$\Pr(y_{iks} = l_s|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}_s, \boldsymbol{z}_{-s}) = \Phi\left(\frac{\gamma_{s,l_s} - \tilde{h}_{iks}}{\tilde{v}_{iks}^{1/2}}\right) - \Phi\left(\frac{\gamma_{s,l_s-1} - \tilde{h}_{iks}}{\tilde{v}_{iks}^{1/2}}\right),$$

where

$$\begin{aligned}\tilde{h}_{iks} &= h_{iks} + \boldsymbol{v}_{s,-s}\boldsymbol{V}_{-s,-s}^{-1}(\boldsymbol{z}_{ik,-s} - \boldsymbol{h}_{ik,-s}) = \boldsymbol{x}_{iks}'\boldsymbol{\beta}_s + \delta_{is} + \boldsymbol{v}_{s,-s}\boldsymbol{V}_{-s,-s}^{-1}(\boldsymbol{z}_{ik,-s} - \boldsymbol{h}_{ik,-s})\\ &= x_{iks1}\boldsymbol{\beta}_{s1} + \cdots + x_{iksc}\boldsymbol{\beta}_{sc} + \cdots + x_{iksq_s}\boldsymbol{\beta}_{sq_s} + \delta_{is} + \boldsymbol{v}_{s,-s}\boldsymbol{V}_{-s,-s}^{-1}(\boldsymbol{z}_{ik,-s} - \boldsymbol{h}_{ik,-s}),\\ \tilde{v}_{iks} &= V_{s,s} - \boldsymbol{v}_{s,-s}\boldsymbol{V}_{-s,-s}^{-1}\boldsymbol{v}_{-s,s}.\end{aligned}$$

If $x_{iksc}(c = 1, \ldots, q_s)$ is a continuous variable, the marginal effect can be obtained by the derivative as follows:

$$\frac{\partial \Pr(y_{iks} = l_s | \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}_s, \boldsymbol{z}_{-s})}{\partial x_{iksc}} = \left[\phi\left(\frac{\gamma_{s,l_s} - \tilde{h}_{iks}}{\tilde{v}_{iks}^{1/2}} \right) - \phi\left(\frac{\gamma_{s,l_s-1} - \tilde{h}_{iks}}{\tilde{v}_{iks}^{1/2}} \right) \right] \times (-\beta_{sc}).$$

If $x_{iksc}(c = 1, \ldots, q_s)$ is a discrete variable, the marginal effect of x_{iksc} going from l_c to $l_c + 1$ can be calculated by:

$$\left[\Phi\left(\frac{\gamma_{s,l_s} - \tilde{h}_{iks,l_c+1}}{\tilde{v}_{iks}^{1/2}} \right) - \Phi\left(\frac{\gamma_{s,l_s-1} - \tilde{h}_{iks,l_c+1}}{\tilde{v}_{iks}^{1/2}} \right) \right] - \left[\Phi\left(\frac{\gamma_{s,l_s} - \tilde{h}_{iks,l_c}}{\tilde{v}_{iks}^{1/2}} \right) - \Phi\left(\frac{\gamma_{s,l_s-1} - \tilde{h}_{iks,l_c}}{\tilde{v}_{iks}^{1/2}} \right) \right],$$

where

$$\begin{aligned}
\tilde{h}_{iks,l_c+1} &= x_{iks1}\boldsymbol{\beta}_{s1} + \cdots + (l_c+1)\boldsymbol{\beta}_{sc} + \cdots + x_{iksq_s}\boldsymbol{\beta}_{sq_s} + \delta_{is} + \boldsymbol{v}_{s,-s}\boldsymbol{V}_{-s,-s}^{-1}(\boldsymbol{z}_{ik,-s} - \boldsymbol{h}_{ik,-s}), \\
\tilde{h}_{iks,l_c} &= x_{iks1}\boldsymbol{\beta}_{s1} + \cdots + l_c\boldsymbol{\beta}_{sc} + \cdots + x_{iksq_s}\boldsymbol{\beta}_{sq_s} + \delta_{is} + \boldsymbol{v}_{s,-s}\boldsymbol{V}_{-s,-s}^{-1}(\boldsymbol{z}_{ik,-s} - \boldsymbol{h}_{ik,-s}).
\end{aligned}$$

A.3 Appendix Tables

Table A.1 provides the complete frequency table of 24 experiments for the setting of $v^* = 0.9$ in the simulation study of MSOP model.

Table A.1: Frequency table of the dependent variables ($v^* = 0.9$)

						y_1					y_2		
m	n_i	$\boldsymbol{\beta}_{true}$	ρ_{true}	1	2	3	4	5	1	2	3	4	5
6	100	(0.55,0.35,0.41,0.56)'	0.3	74	143	199	120	64	83	162	197	124	34
			0.5	75	142	194	125	64	82	172	190	122	34
			0.7	80	132	193	132	63	96	173	186	113	32
		(0.40,0.49,0.45,0.52)'	0.3	78	138	195	136	53	80	165	199	122	34
			0.5	79	134	196	133	58	82	169	194	121	34
			0.7	84	137	184	132	63	88	179	185	116	32
	200	(0.55,0.35,0.41,0.56)'	0.3	149	295	381	231	144	147	327	391	250	85
			0.5	152	297	370	236	145	149	345	379	244	83
			0.7	166	279	363	246	146	183	331	374	229	83
		(0.40,0.49,0.45,0.52)'	0.3	155	297	374	240	134	146	327	394	248	85
			0.5	158	291	370	242	139	149	338	391	241	81
			0.7	177	284	349	243	147	173	338	378	229	82
	300	(0.55,0.35,0.41,0.56)'	0.3	232	430	565	373	200	219	483	583	372	143
			0.5	240	431	550	374	205	227	493	579	355	146
			0.7	254	421	534	383	208	271	479	567	338	145
		(0.40,0.49,0.45,0.52)'	0.3	233	441	553	385	188	217	483	591	367	142
			0.5	240	435	543	388	194	227	491	583	356	143
			0.7	267	419	521	386	207	263	485	571	340	141
10	100	(0.55,0.35,0.41,0.56)'	0.3	118	230	372	199	81	123	270	338	184	85
			0.5	117	233	362	208	80	130	276	333	177	84
			0.7	121	223	359	215	82	142	287	322	168	81
		(0.40,0.49,0.45,0.52)'	0.3	116	233	372	206	73	123	270	335	191	81
			0.5	119	230	364	216	71	130	272	338	181	79
			0.7	123	229	353	213	82	141	290	324	163	82
	200	(0.55,0.35,0.41,0.56)'	0.3	231	486	731	397	155	221	527	692	386	174
			0.5	232	492	710	413	153	230	540	682	381	167
			0.7	243	475	699	421	162	265	548	667	360	160
		(0.40,0.49,0.45,0.52)'	0.3	236	486	735	399	144	220	520	691	399	170
			0.5	240	483	723	410	144	229	540	688	384	159
			0.7	248	476	696	422	158	263	552	671	350	164
	300	(0.55,0.35,0.41,0.56)'	0.3	370	738	1063	603	226	357	771	1031	595	246
			0.5	371	746	1037	623	223	372	790	1022	579	237
			0.7	383	722	1029	627	239	425	805	993	550	227

(to be continued)

(continued)

m	n_i	$\boldsymbol{\beta}_{true}$	ρ_{true}	y_1 1	2	3	4	5	y_2 1	2	3	4	5
10	300	(0.40,0.49,0.45,0.52)'	0.3	372	752	1055	606	215	355	774	1026	606	239
			0.5	378	748	1042	619	213	372	797	1020	584	227
			0.7	393	733	1008	640	226	421	813	993	543	230
15	100	(0.55,0.35,0.41,0.56)'	0.3	164	343	549	323	121	167	364	514	317	138
			0.5	156	338	539	345	122	172	359	518	309	142
			0.7	151	315	536	367	131	179	369	498	306	148
		(0.40,0.49,0.45,0.52)'	0.3	164	348	544	334	110	165	368	509	322	136
			0.5	158	344	537	349	112	170	363	515	316	136
			0.7	153	328	523	370	126	177	370	503	299	151
	200	(0.55,0.35,0.41,0.56)'	0.3	322	689	1100	641	248	317	710	1032	648	293
			0.5	307	689	1073	677	254	319	709	1033	644	295
			0.7	306	652	1053	709	280	344	700	1010	639	307
		(0.40,0.49,0.45,0.52)'	0.3	329	706	1079	651	235	311	710	1034	658	287
			0.5	320	691	1069	678	242	316	711	1039	647	287
			0.7	314	655	1041	723	267	341	705	1018	625	311
	300	(0.55,0.35,0.41,0.56)'	0.3	516	1042	1618	953	371	497	1078	1536	961	428
			0.5	499	1032	1589	998	382	504	1075	1538	949	434
			0.7	487	983	1564	1046	420	541	1059	1510	937	453
		(0.40,0.49,0.45,0.52)'	0.3	520	1076	1588	970	346	491	1084	1536	966	423
			0.5	507	1049	1577	1012	355	503	1080	1537	953	427
			0.7	497	994	1547	1067	395	537	1068	1513	929	453
20	100	(0.55,0.35,0.41,0.56)'	0.3	207	444	730	442	177	228	500	674	429	169
			0.5	196	439	714	471	180	230	504	673	418	175
			0.7	187	411	712	500	190	243	517	650	410	180
		(0.40,0.49,0.45,0.52)'	0.3	207	445	720	470	158	223	507	669	432	169
			0.5	197	442	711	484	166	226	506	675	422	171
			0.7	189	427	695	504	185	237	521	657	400	185
	200	(0.55,0.35,0.41,0.56)'	0.3	405	906	1439	875	375	425	980	1368	865	362
			0.5	387	903	1409	917	384	426	995	1357	855	367
			0.7	384	842	1403	961	410	463	991	1325	842	379
		(0.40,0.49,0.45,0.52)'	0.3	413	914	1427	891	355	418	979	1372	871	360
			0.5	398	897	1414	923	368	422	992	1374	852	360
			0.7	391	853	1384	973	399	456	995	1342	823	384
	300	(0.55,0.35,0.41,0.56)'	0.3	638	1370	2112	1329	551	667	1462	2050	1281	540
			0.5	618	1355	2081	1375	571	673	1474	2045	1256	552
			0.7	601	1280	2075	1434	610	725	1469	2003	1232	571
		(0.40,0.49,0.45,0.52)'	0.3	641	1408	2085	1348	518	658	1467	2057	1281	537
			0.5	622	1373	2073	1397	535	670	1478	2051	1257	544
			0.7	608	1298	2052	1458	584	715	1479	2014	1224	568

Table A.2 provides the complete frequency table of 24 experiments for the setting of $v^* = 0.5$ in the simulation study of MSOP model.

Table A.2: Frequency table of the dependent variables ($v^* = 0.5$)

				y_1					y_2				
m	n_i	$\boldsymbol{\beta}_{true}$	ρ_{true}	1	2	3	4	5	1	2	3	4	5
6	100	(0.56,0.38,0.46,0.56)'	0.3	76	144	180	150	50	72	150	196	127	55
			0.5	77	145	175	149	54	75	147	196	127	55
			0.7	79	132	181	148	60	85	145	190	125	55
		(0.46,0.49,0.49,0.53)'	0.3	76	142	179	155	48	72	148	194	131	55
			0.5	77	134	175	163	51	73	147	197	127	56
			0.7	77	131	180	154	58	85	146	191	123	55
	200	(0.56,0.38,0.46,0.56)'	0.3	167	269	365	280	119	140	292	392	246	130
			0.5	170	260	365	282	123	147	291	389	242	131
			0.7	175	236	363	290	136	161	291	380	240	128
		(0.46,0.49,0.49,0.53)'	0.3	168	267	361	286	118	139	287	393	252	129
			0.5	168	253	359	296	124	146	289	388	247	130
			0.7	166	250	355	293	136	162	290	381	240	127
	300	(0.56,0.38,0.46,0.56)'	0.3	241	392	558	424	185	195	438	592	379	196
			0.5	247	385	546	431	191	205	442	584	371	198
			0.7	257	351	538	443	211	225	443	578	358	196
		(0.46,0.49,0.49,0.53)'	0.3	245	384	551	437	183	194	433	596	382	195
			0.5	245	366	550	447	192	206	436	589	372	197
			0.7	247	362	531	453	207	226	438	582	361	193
10	100	(0.56,0.38,0.46,0.56)'	0.3	122	229	326	241	82	113	266	342	220	59
			0.5	125	224	322	240	89	120	264	345	211	60
			0.7	126	220	317	241	96	135	257	342	205	61
		(0.46,0.49,0.49,0.53)'	0.3	122	231	312	249	86	113	267	340	222	58
			0.5	123	223	309	257	88	118	263	347	213	59
			0.7	123	221	308	253	95	132	259	349	199	61
	200	(0.56,0.38,0.46,0.56)'	0.3	275	458	643	456	168	229	504	694	424	149
			0.5	275	457	633	458	177	244	504	694	410	148
			0.7	278	440	616	481	185	266	493	699	394	148
		(0.46,0.49,0.49,0.53)'	0.3	264	459	640	461	176	228	501	694	430	147
			0.5	269	450	626	477	178	240	499	699	414	148
			0.7	270	442	618	479	191	263	494	709	386	148
	300	(0.56,0.38,0.46,0.56)'	0.3	412	695	945	693	255	321	747	1064	640	228
			0.5	415	692	924	704	265	343	753	1058	623	223
			0.7	413	676	900	729	282	374	749	1056	597	224
		(0.46,0.49,0.49,0.53)'	0.3	405	688	948	697	262	319	745	1068	644	224
			0.5	409	684	925	714	268	342	746	1065	622	225
			0.7	409	673	910	716	292	372	751	1065	591	221

(to be continued)

(continued)

m	n_i	$\boldsymbol{\beta}_{true}$	ρ_{true}	y_1					y_2				
				1	2	3	4	5	1	2	3	4	5
15	100	(0.56,0.38,0.46,0.56)'	0.3	173	353	477	363	134	153	351	509	347	140
			0.5	172	347	475	360	146	157	346	510	344	143
			0.7	169	324	489	356	162	170	337	502	344	147
		(0.46,0.49,0.49,0.53)'	0.3	175	344	483	357	141	151	355	505	348	141
			0.5	171	332	471	381	145	156	347	509	344	144
			0.7	165	319	481	377	158	168	336	512	336	148
	200	(0.56,0.38,0.46,0.56)'	0.3	357	690	968	714	271	296	676	1017	676	335
			0.5	352	676	972	713	287	303	680	1010	669	338
			0.7	344	635	970	731	320	325	662	998	670	345
		(0.46,0.49,0.49,0.53)'	0.3	353	675	976	714	282	293	676	1014	684	333
			0.5	348	656	958	748	290	302	676	1009	676	337
			0.7	335	630	965	748	322	324	657	1019	654	346
	300	(0.56,0.38,0.46,0.56)'	0.3	541	1012	1479	1055	413	418	1026	1536	1026	494
			0.5	533	1001	1468	1065	433	431	1033	1525	1014	497
			0.7	518	947	1455	1090	490	462	1008	1513	1008	509
		(0.46,0.49,0.49,0.53)'	0.3	542	985	1489	1057	427	414	1021	1544	1030	491
			0.5	528	970	1464	1093	445	432	1023	1530	1018	497
			0.7	515	932	1456	1105	492	462	1002	1540	987	509
20	100	(0.56,0.38,0.46,0.56)'	0.3	207	453	656	497	187	207	493	690	438	172
			0.5	202	446	650	500	202	212	489	691	433	175
			0.7	196	420	660	501	223	229	476	682	435	178
		(0.46,0.49,0.49,0.53)'	0.3	208	444	652	499	197	207	492	687	442	172
			0.5	201	431	642	523	203	212	487	692	434	175
			0.7	193	413	651	523	220	229	476	691	424	180
	200	(0.56,0.38,0.46,0.56)'	0.3	444	877	1318	982	379	407	951	1374	868	400
			0.5	432	862	1316	992	398	422	950	1367	859	402
			0.7	417	815	1305	1025	438	448	933	1353	855	411
		(0.46,0.49,0.49,0.53)'	0.3	437	861	1317	993	392	404	949	1368	882	397
			0.5	430	839	1294	1034	403	415	952	1366	866	401
			0.7	409	808	1298	1043	442	448	932	1371	838	411
	300	(0.56,0.38,0.46,0.56)'	0.3	661	1289	1986	1478	586	593	1435	2059	1314	599
			0.5	649	1272	1967	1499	613	617	1436	2050	1298	599
			0.7	621	1212	1940	1547	680	654	1416	2036	1282	612
		(0.46,0.49,0.49,0.53)'	0.3	658	1262	1982	1498	600	589	1429	2065	1322	595
			0.5	643	1244	1954	1537	622	614	1432	2056	1298	600
			0.7	619	1196	1943	1559	683	654	1416	2059	1261	610

Table A.3 provides the complete frequency table of 24 experiments for the setting of $v^* = 0.1$ in the simulation study of MSOP model.

Table A.3: Frequency table of the dependent variables ($v^* = 0.1$)

						y_1					y_2		
m	n_i	$\boldsymbol{\beta}_{true}$	ρ_{true}	1	2	3	4	5	1	2	3	4	5
6	100	(0.58,0.36,0.45,0.56)'	0.3	70	117	213	135	65	62	145	193	127	73
			0.5	70	114	214	135	67	69	143	190	127	71
			0.7	74	114	200	136	76	77	151	182	122	68
		(0.45,0.49,0.49,0.52)'	0.3	72	114	214	137	63	66	143	192	128	71
			0.5	70	118	212	132	68	72	144	189	122	73
			0.7	71	111	207	138	73	77	153	178	123	69
	200	(0.58,0.36,0.45,0.56)'	0.3	168	246	393	259	134	137	285	389	259	130
			0.5	167	246	387	262	138	147	279	390	256	128
			0.7	171	241	366	266	156	159	294	373	247	127
		(0.45,0.49,0.49,0.52)'	0.3	170	246	390	271	123	140	284	390	258	128
			0.5	163	257	383	264	133	148	282	388	253	129
			0.7	167	244	373	271	145	160	290	372	250	128
	300	(0.58,0.36,0.45,0.56)'	0.3	239	397	578	393	193	208	445	593	366	188
			0.5	239	394	569	397	201	222	439	590	362	187
			0.7	242	386	538	408	226	240	463	557	347	193
		(0.45,0.49,0.49,0.52)'	0.3	243	390	579	406	182	208	442	596	368	186
			0.5	239	397	570	398	196	220	440	590	361	189
			0.7	239	380	555	411	215	242	457	556	351	194
10	100	(0.58,0.36,0.45,0.56)'	0.3	116	201	346	238	99	109	245	355	204	87
			0.5	119	195	348	237	101	120	240	354	199	87
			0.7	121	191	332	242	114	131	253	337	200	79
		(0.45,0.49,0.49,0.52)'	0.3	115	203	342	244	96	111	244	355	203	87
			0.5	115	201	344	238	102	121	243	351	201	84
			0.7	115	195	330	249	111	129	256	335	206	74
	200	(0.58,0.36,0.45,0.56)'	0.3	244	441	673	451	191	219	500	686	435	160
			0.5	251	432	662	459	196	237	493	687	426	157
			0.7	260	414	629	481	216	260	508	670	415	147
		(0.45,0.49,0.49,0.52)'	0.3	242	438	673	466	181	221	504	686	432	157
			0.5	242	435	666	471	186	238	497	687	427	151
			0.7	240	432	635	489	204	256	504	673	422	145
	300	(0.58,0.36,0.45,0.56)'	0.3	381	691	987	673	268	331	776	1039	610	244
			0.5	385	675	979	685	276	353	768	1037	603	239
			0.7	389	659	923	723	306	382	815	978	596	229
		(0.45,0.49,0.49,0.52)'	0.3	376	695	974	699	256	331	776	1043	608	242
			0.5	373	686	972	705	264	353	772	1038	602	235
			0.7	371	678	933	726	292	383	802	988	599	228

(to be continued)

(continued)

m	n_i	$\boldsymbol{\beta}_{true}$	ρ_{true}	y_1					y_2				
				1	2	3	4	5	1	2	3	4	5
15	100	(0.58,0.36,0.45,0.56)'	0.3	146	286	548	367	153	148	358	533	299	162
			0.5	140	275	550	375	160	149	357	525	305	164
			0.7	129	255	538	388	190	151	360	515	313	161
		(0.45,0.49,0.49,0.52)'	0.3	142	280	549	376	153	149	360	526	307	158
			0.5	138	274	551	374	163	154	353	527	308	158
			0.7	124	260	536	402	178	152	361	513	318	156
	200	(0.58,0.36,0.45,0.56)'	0.3	305	619	1049	715	312	311	734	1029	619	307
			0.5	295	603	1035	739	328	311	732	1020	628	309
			0.7	273	560	1014	774	379	313	738	1010	629	310
		(0.45,0.49,0.49,0.52)'	0.3	301	605	1064	731	299	309	741	1022	630	298
			0.5	286	601	1059	736	318	314	731	1025	631	299
			0.7	255	584	1020	781	360	310	739	1006	638	307
	300	(0.58,0.36,0.45,0.56)'	0.3	479	939	1569	1070	443	465	1122	1556	904	453
			0.5	458	912	1556	1107	467	465	1113	1544	920	458
			0.7	419	858	1503	1178	542	471	1135	1502	917	475
		(0.45,0.49,0.49,0.52)'	0.3	470	928	1571	1103	428	460	1125	1547	922	446
			0.5	447	917	1562	1120	454	467	1111	1553	919	450
			0.7	402	890	1501	1186	521	468	1132	1503	928	469
20	100	(0.58,0.36,0.45,0.56)'	0.3	197	377	731	499	196	202	469	676	434	219
			0.5	193	368	730	507	202	210	467	667	438	218
			0.7	187	348	709	524	232	224	470	656	441	209
		(0.45,0.49,0.49,0.52)'	0.3	193	382	726	508	191	205	469	672	441	213
			0.5	189	377	727	503	204	215	464	669	440	212
			0.7	177	360	706	538	219	222	474	654	447	203
	200	(0.58,0.36,0.45,0.56)'	0.3	417	818	1397	957	411	419	946	1351	866	418
			0.5	413	799	1379	984	425	431	943	1346	865	415
			0.7	403	748	1341	1032	476	457	952	1327	856	408
		(0.45,0.49,0.49,0.52)'	0.3	409	816	1405	979	391	417	957	1345	874	407
			0.5	396	814	1396	984	410	434	946	1346	868	406
			0.7	375	790	1345	1037	453	453	951	1328	863	405
	300	(0.58,0.36,0.45,0.56)'	0.3	646	1272	2075	1414	593	621	1450	2052	1252	625
			0.5	635	1238	2056	1457	614	638	1442	2041	1255	624
			0.7	608	1178	1982	1543	689	675	1482	1972	1242	629
		(0.45,0.49,0.49,0.52)'	0.3	636	1271	2068	1453	572	617	1455	2044	1268	616
			0.5	616	1263	2055	1472	594	642	1439	2046	1258	615
			0.7	583	1229	1977	1548	663	672	1472	1984	1248	624

Table 3.1 gives partial simulation results of posterior means and true values as well as the mean squared error (MSE) of $\boldsymbol{\beta}$ and ρ. Here, Table A.4 provides the complete results of 24 simulation experiments for the setting of $v^* = 0.9$.

Table A.4: Simulation results in the case of $v^* = 0.9$

			simulation results ($\boldsymbol{\beta}_{true} = (0.55, 0.35, 0.41, 0.56)'$)						
ρ_{true}	m	n_i	β_{11}	β_{12}	β_{21}	β_{22}	MSE_β	ρ	MSE_ρ
0.3	6	100	0.5224	0.4074	0.4330	0.5464	0.0077	0.2850	0.0570
		200	0.5363	0.3848	0.4101	0.5544	0.0030	0.3379	0.0391
		300	0.5556	0.3698	0.4102	0.5700	0.0015	0.3549	0.0433
	10	100	0.5468	0.3806	0.4701	0.5422	0.0068	0.2805	0.0576
		200	0.5334	0.3570	0.4100	0.5543	0.0012	0.2476	0.0577
		300	0.5405	0.3485	0.4114	0.5593	0.0007	0.2541	0.0534
	15	100	0.5392	0.3560	0.4450	0.5489	0.0028	0.2835	0.0517
		200	0.5306	0.3426	0.4123	0.5615	0.0011	0.3528	0.0464
		300	0.5425	0.3459	0.4076	0.5733	0.0007	0.2976	0.0467
	20	100	0.5421	0.3650	0.4325	0.5524	0.0018	0.2606	0.0518
		200	0.5374	0.3518	0.4102	0.5577	0.0007	0.3625	0.0446
		300	0.5484	0.3495	0.4076	0.5688	0.0004	0.3408	0.0426
0.5	6	100	0.5173	0.3932	0.3902	0.5233	0.0079	0.3397	0.0758
		200	0.5308	0.3715	0.3943	0.5361	0.0033	0.4303	0.0404
		300	0.5473	0.3600	0.4061	0.5638	0.0013	0.4116	0.0503
	10	100	0.5496	0.3652	0.4474	0.5552	0.0036	0.3887	0.0572
		200	0.5305	0.3520	0.3961	0.5584	0.0015	0.4203	0.0461
		300	0.5409	0.3452	0.4054	0.5626	0.0007	0.3904	0.0525
	15	100	0.5383	0.3456	0.4289	0.5573	0.0018	0.4574	0.0357
		200	0.5278	0.3388	0.4004	0.5640	0.0014	0.4939	0.0327
		300	0.5435	0.3408	0.4031	0.5768	0.0009	0.4468	0.0395
	20	100	0.5370	0.3569	0.4211	0.5605	0.0013	0.4629	0.0370
		200	0.5320	0.3458	0.4017	0.5599	0.0009	0.5165	0.0292
		300	0.5451	0.3433	0.4058	0.5719	0.0005	0.5076	0.0289
0.7	6	100	0.5204	0.3725	0.3933	0.5699	0.0051	0.5020	0.0688
		200	0.5292	0.3574	0.4205	0.5707	0.0023	0.5537	0.0451
		300	0.5427	0.3557	0.4331	0.5864	0.0023	0.5480	0.0444
	10	100	0.5314	0.3296	0.4643	0.5796	0.0061	0.5414	0.0523
		200	0.5221	0.3250	0.4207	0.5727	0.0027	0.5392	0.0550
		300	0.5407	0.3242	0.4255	0.5725	0.0018	0.5355	0.0534
	15	100	0.5225	0.3201	0.4441	0.5703	0.0043	0.6269	0.0265
		200	0.5260	0.3215	0.4092	0.5716	0.0022	0.6446	0.0228
		300	0.5374	0.3242	0.4144	0.5819	0.0017	0.6290	0.0249
	20	100	0.5226	0.3401	0.4309	0.5747	0.0025	0.6561	0.0181
		200	0.5274	0.3353	0.4131	0.5701	0.0013	0.6981	0.0123
		300	0.5419	0.3357	0.4177	0.5784	0.0010	0.6819	0.0147

(to be continued)

(continued)

ρ_{true}	m	n_i	simulation results ($\boldsymbol{\beta}_{true} = (0.40, 0.49, 0.45, 0.52)'$)						
			β_{11}	β_{12}	β_{21}	β_{22}	MSE_β	ρ	MSE_ρ
0.3	6	100	0.3774	0.5142	0.4674	0.4842	0.0059	0.2647	0.0532
		200	0.3864	0.5234	0.4539	0.5018	0.0034	0.3594	0.0486
		300	0.3906	0.5108	0.4536	0.5276	0.0016	0.3813	0.0509
	10	100	0.3682	0.4871	0.4939	0.4794	0.0065	0.2628	0.0608
		200	0.3692	0.4799	0.4551	0.5003	0.0024	0.2610	0.0557
		300	0.3836	0.4775	0.4549	0.5018	0.0014	0.2265	0.0642
	15	100	0.3785	0.4716	0.4807	0.4833	0.0043	0.2940	0.0543
		200	0.3858	0.4699	0.4524	0.4983	0.0016	0.3747	0.0533
		300	0.3936	0.4763	0.4480	0.5157	0.0007	0.2541	0.0528
	20	100	0.3793	0.4867	0.4690	0.4982	0.0022	0.2428	0.0549
		200	0.3851	0.4819	0.4543	0.5042	0.0010	0.3990	0.0415
		300	0.3939	0.4831	0.4506	0.5181	0.0004	0.3320	0.0434
0.5	6	100	0.3820	0.5134	0.4418	0.4992	0.0045	0.3141	0.0849
		200	0.3889	0.5016	0.4335	0.5101	0.0022	0.4298	0.0408
		300	0.3900	0.4912	0.4509	0.5318	0.0014	0.4171	0.0488
	10	100	0.3903	0.4928	0.4813	0.4955	0.0037	0.3654	0.0709
		200	0.3774	0.4841	0.4414	0.5097	0.0017	0.3928	0.0544
		300	0.3897	0.4749	0.4535	0.5141	0.0009	0.3634	0.0629
	15	100	0.3822	0.4761	0.4693	0.4979	0.0026	0.4407	0.0397
		200	0.3828	0.4682	0.4453	0.5098	0.0016	0.5157	0.0279
		300	0.3929	0.4700	0.4507	0.5214	0.0009	0.4421	0.0382
	20	100	0.3777	0.4879	0.4581	0.5100	0.0016	0.4471	0.0392
		200	0.3812	0.4785	0.4448	0.5118	0.0010	0.5513	0.0256
		300	0.3909	0.4777	0.4493	0.5214	0.0005	0.5189	0.0250
0.7	6	100	0.3595	0.5177	0.4399	0.5051	0.0059	0.4538	0.0952
		200	0.3800	0.5050	0.4480	0.5180	0.0021	0.5325	0.0552
		300	0.3978	0.5038	0.4484	0.5302	0.0013	0.5426	0.0512
	10	100	0.3788	0.4800	0.4986	0.5299	0.0051	0.5105	0.0645
		200	0.3688	0.4714	0.4638	0.5206	0.0024	0.5239	0.0572
		300	0.3888	0.4681	0.4652	0.5247	0.0015	0.5513	0.0468
	15	100	0.3800	0.4576	0.4830	0.5166	0.0039	0.6177	0.0314
		200	0.3793	0.4567	0.4571	0.5153	0.0022	0.6381	0.0242
		300	0.3880	0.4582	0.4600	0.5324	0.0018	0.6157	0.0266
	20	100	0.3699	0.4778	0.4675	0.5246	0.0023	0.6438	0.0200
		200	0.3735	0.4735	0.4570	0.5211	0.0015	0.6907	0.0124
		300	0.3878	0.4743	0.4564	0.5271	0.0008	0.6661	0.0165

Table 3.1 gives partial simulation results of posterior means and true values as well as the mean squared error (MSE) of $\boldsymbol{\beta}$ and ρ. Here, Table A.5 provides the complete results of 24 simulation experiments for the setting of $v^* = 0.5$.

Table A.5: Simulation results in the case of $v^* = 0.5$

			simulation results ($\boldsymbol{\beta}_{true} = (0.56, 0.38, 0.46, 0.56)'$)						
ρ_{true}	m	n_i	β_{11}	β_{12}	β_{21}	β_{22}	MSE_β	ρ	MSE_ρ
0.3	6	100	0.5691	0.3423	0.4767	0.5191	0.0100	0.2808	0.0446
		200	0.5249	0.3663	0.4715	0.5395	0.0056	0.2847	0.0412
		300	0.5555	0.3655	0.4492	0.5312	0.0037	0.2721	0.0466
	10	100	0.5467	0.3416	0.4512	0.4827	0.0114	0.3100	0.0608
		200	0.5561	0.3585	0.4635	0.5260	0.0039	0.3143	0.0470
		300	0.5734	0.3750	0.4489	0.5396	0.0022	0.3052	0.0468
	15	100	0.5770	0.3841	0.4821	0.4834	0.0094	0.3190	0.0409
		200	0.5652	0.3837	0.4760	0.5211	0.0032	0.3258	0.0385
		300	0.5722	0.4018	0.4506	0.5273	0.0028	0.3945	0.0424
	20	100	0.5534	0.3829	0.4741	0.4816	0.0084	0.3182	0.0529
		200	0.5483	0.3836	0.4713	0.5176	0.0031	0.2921	0.0442
		300	0.5673	0.3972	0.4515	0.5312	0.0020	0.3042	0.0410
0.5	6	100	0.5609	0.3423	0.4740	0.5247	0.0097	0.3448	0.0671
		200	0.5259	0.3713	0.4648	0.5484	0.0053	0.3510	0.0626
		300	0.5506	0.3677	0.4510	0.5343	0.0037	0.3400	0.0652
	10	100	0.5511	0.3559	0.4589	0.4898	0.0097	0.4870	0.0361
		200	0.5553	0.3680	0.4668	0.5277	0.0034	0.4457	0.0374
		300	0.5698	0.3805	0.4589	0.5385	0.0020	0.4684	0.0316
	15	100	0.5822	0.3971	0.4875	0.4871	0.0095	0.4709	0.0332
		200	0.5722	0.3925	0.4712	0.5156	0.0039	0.4742	0.0300
		300	0.5698	0.4055	0.4541	0.5250	0.0030	0.5008	0.0235
	20	100	0.5557	0.3868	0.4792	0.4864	0.0080	0.5228	0.0324
		200	0.5458	0.3794	0.4736	0.5234	0.0029	0.5041	0.0241
		300	0.5650	0.3951	0.4543	0.5307	0.0019	0.4780	0.0353
0.7	6	100	0.5477	0.3543	0.4678	0.5462	0.0087	0.4361	0.1037
		200	0.5301	0.3798	0.4649	0.5481	0.0051	0.4566	0.0928
		300	0.5589	0.3750	0.4566	0.5332	0.0034	0.4153	0.1175
	10	100	0.5538	0.3584	0.4589	0.4968	0.0087	0.6375	0.0286
		200	0.5610	0.3691	0.4763	0.5205	0.0041	0.5739	0.0411
		300	0.5788	0.3813	0.4688	0.5328	0.0027	0.5845	0.0352
	15	100	0.5787	0.4061	0.4742	0.4953	0.0084	0.6152	0.0241
		200	0.5725	0.3914	0.4779	0.5175	0.0038	0.6088	0.0258
		300	0.5780	0.4077	0.4617	0.5251	0.0033	0.6276	0.0224
	20	100	0.5545	0.3909	0.4715	0.4886	0.0075	0.6905	0.0134
		200	0.5501	0.3855	0.4753	0.5219	0.0028	0.6608	0.0152
		300	0.5712	0.4003	0.4617	0.5314	0.0021	0.6503	0.0193

(to be continued)

(continued)

ρ_{true}	m	n_i	simulation results ($\boldsymbol{\beta}_{true} = (0.46, 0.49, 0.49, 0.53)'$)						
			β_{11}	β_{12}	β_{21}	β_{22}	MSE_β	ρ	MSE_ρ
0.3	6	100	0.4354	0.4450	0.5231	0.4935	0.0119	0.2603	0.0481
		200	0.4075	0.4805	0.4929	0.5164	0.0067	0.3017	0.0464
		300	0.4466	0.4777	0.4719	0.5022	0.0040	0.2712	0.0469
	10	100	0.4479	0.4696	0.4862	0.4473	0.0113	0.3460	0.0483
		200	0.4556	0.4865	0.4905	0.4859	0.0041	0.3128	0.0511
		300	0.4701	0.4968	0.4820	0.5029	0.0025	0.3340	0.0424
	15	100	0.4787	0.4989	0.5110	0.4530	0.0094	0.3520	0.0365
		200	0.4680	0.5024	0.5004	0.4877	0.0035	0.3573	0.0397
		300	0.4697	0.5214	0.4796	0.4956	0.0034	0.4125	0.0420
	20	100	0.4489	0.5002	0.5038	0.4492	0.0089	0.3380	0.0477
		200	0.4393	0.5037	0.4952	0.4857	0.0037	0.3312	0.0464
		300	0.4610	0.5164	0.4795	0.5003	0.0024	0.3146	0.0436
0.5	6	100	0.4388	0.4711	0.4850	0.4935	0.0089	0.3436	0.0742
		200	0.4031	0.4924	0.4859	0.5267	0.0070	0.3692	0.0575
		300	0.4418	0.4832	0.4696	0.5112	0.0037	0.3639	0.0596
	10	100	0.4214	0.4803	0.4931	0.4604	0.0105	0.4920	0.0365
		200	0.4376	0.4853	0.5010	0.4960	0.0040	0.4494	0.0391
		300	0.4577	0.4974	0.4884	0.5112	0.0019	0.4789	0.0311
	15	100	0.4625	0.5043	0.5122	0.4608	0.0082	0.5031	0.0270
		200	0.4575	0.5013	0.5052	0.4919	0.0033	0.4755	0.0329
		300	0.4615	0.5181	0.4833	0.4976	0.0029	0.5177	0.0231
	20	100	0.4372	0.5059	0.5049	0.4550	0.0087	0.5207	0.0314
		200	0.4297	0.5006	0.4985	0.4908	0.0037	0.4929	0.0289
		300	0.4500	0.5133	0.4827	0.5027	0.0021	0.4998	0.0256
0.7	6	100	0.4339	0.4731	0.4960	0.5017	0.0087	0.4350	0.1122
		200	0.4003	0.5046	0.4902	0.5270	0.0076	0.4617	0.0889
		300	0.4402	0.4969	0.4775	0.5096	0.0036	0.4439	0.1027
	10	100	0.4416	0.4666	0.4805	0.4637	0.0097	0.6511	0.0276
		200	0.4497	0.4794	0.4977	0.4987	0.0036	0.5691	0.0453
		300	0.4770	0.4929	0.4899	0.5084	0.0023	0.5898	0.0351
	15	100	0.4745	0.5017	0.5108	0.4590	0.0087	0.6059	0.0276
		200	0.4649	0.5039	0.5031	0.4886	0.0035	0.6089	0.0249
		300	0.4707	0.5175	0.4881	0.4985	0.0029	0.6106	0.0273
	20	100	0.4455	0.5016	0.5031	0.4585	0.0078	0.6463	0.0236
		200	0.4392	0.5012	0.5031	0.4969	0.0029	0.6430	0.0201
		300	0.4613	0.5135	0.4912	0.5076	0.0018	0.6304	0.0269

Table 3.1 gives partial simulation results of posterior means and true values as well as the mean squared error (MSE) of $\boldsymbol{\beta}$ and ρ. Here, Table A.6 provides the complete results of 24 simulation experiments for the setting of $v^* = 0.1$.

Table A.6: Simulation results in the case of $v^* = 0.1$

			simulation results ($\boldsymbol{\beta}_{true} = (0.58, 0.36, 0.45, 0.56)'$)						
ρ_{true}	m	n_i	β_{11}	β_{12}	β_{21}	β_{22}	MSE_β	ρ	MSE_ρ
0.3	6	100	0.6256	0.3614	0.5223	0.5989	0.0176	0.2930	0.0508
		200	0.6152	0.3501	0.4922	0.5895	0.0086	0.2849	0.0485
		300	0.5854	0.3492	0.4697	0.5844	0.0042	0.3021	0.0474
	10	100	0.5939	0.3383	0.4198	0.5670	0.0064	0.2919	0.0550
		200	0.6032	0.3243	0.4227	0.5320	0.0059	0.3348	0.0447
		300	0.5925	0.3342	0.4297	0.5399	0.0033	0.2901	0.0456
	15	100	0.5923	0.3136	0.4626	0.5451	0.0061	0.3174	0.0425
		200	0.5885	0.3232	0.4503	0.5561	0.0031	0.3281	0.0389
		300	0.5812	0.3386	0.4416	0.5661	0.0017	0.2569	0.0448
	20	100	0.5859	0.3263	0.4611	0.5706	0.0041	0.2732	0.0533
		200	0.5944	0.3362	0.4587	0.5627	0.0021	0.3038	0.0431
		300	0.5816	0.3467	0.4441	0.5744	0.0013	0.2815	0.0470
0.5	6	100	0.6047	0.3694	0.5343	0.6071	0.0194	0.3906	0.0560
		200	0.6154	0.3565	0.4973	0.5874	0.0088	0.3674	0.0663
		300	0.5975	0.3563	0.4691	0.5848	0.0044	0.3719	0.0624
	10	100	0.5921	0.3559	0.4169	0.5603	0.0068	0.4331	0.0466
		200	0.6038	0.3349	0.4240	0.5331	0.0052	0.4515	0.0367
		300	0.5921	0.3354	0.4300	0.5360	0.0034	0.4114	0.0442
	15	100	0.5900	0.3233	0.4541	0.5455	0.0050	0.4111	0.0450
		200	0.5944	0.3294	0.4477	0.5567	0.0029	0.4079	0.0468
		300	0.5891	0.3424	0.4383	0.5616	0.0017	0.3846	0.0491
	20	100	0.5799	0.3343	0.4537	0.5704	0.0034	0.4506	0.0386
		200	0.5943	0.3406	0.4569	0.5633	0.0020	0.4778	0.0288
		300	0.5845	0.3480	0.4433	0.5728	0.0013	0.4804	0.0251
0.7	6	100	0.5980	0.3598	0.5235	0.5998	0.0159	0.4736	0.0893
		200	0.6214	0.3637	0.5042	0.5862	0.0103	0.4643	0.0870
		300	0.5914	0.3491	0.4806	0.5844	0.0051	0.4769	0.0814
	10	100	0.6006	0.3637	0.3895	0.5614	0.0094	0.6026	0.0338
		200	0.6016	0.3428	0.4132	0.5355	0.0054	0.5977	0.0339
		300	0.5881	0.3465	0.4244	0.5389	0.0031	0.5847	0.0342
	15	100	0.5868	0.3262	0.4416	0.5407	0.0052	0.5802	0.0386
		200	0.5910	0.3252	0.4446	0.5484	0.0033	0.5647	0.0489
		300	0.5821	0.3403	0.4366	0.5600	0.0017	0.5506	0.0480
	20	100	0.5852	0.3357	0.4403	0.5660	0.0035	0.6380	0.0253
		200	0.5983	0.3403	0.4508	0.5604	0.0021	0.6400	0.0208
		300	0.5859	0.3485	0.4386	0.5731	0.0014	0.6543	0.0161

(to be continued)

(continued)

ρ_{true}	m	n_i	simulation results ($\boldsymbol{\beta}_{true} = (0.45, 0.49, 0.49, 0.52)'$)						
			β_{11}	β_{12}	β_{21}	β_{22}	MSE_β	ρ	MSE_ρ
0.3	6	100	0.4772	0.5168	0.5606	0.5696	0.0178	0.3092	0.0559
		200	0.4728	0.4894	0.5278	0.5522	0.0076	0.3243	0.0490
		300	0.4579	0.4863	0.4994	0.5520	0.0042	0.3039	0.0453
	10	100	0.4480	0.4917	0.4539	0.5254	0.0062	0.3135	0.0494
		200	0.4646	0.4704	0.4591	0.4935	0.0047	0.3516	0.0471
		300	0.4385	0.4734	0.4655	0.5006	0.0031	0.3240	0.0397
	15	100	0.4523	0.4655	0.4907	0.5039	0.0042	0.3204	0.0480
		200	0.4525	0.4595	0.4839	0.5152	0.0027	0.3346	0.0424
		300	0.4407	0.4758	0.4737	0.5247	0.0017	0.2517	0.0480
	20	100	0.4443	0.4687	0.4970	0.5355	0.0033	0.3205	0.0482
		200	0.4559	0.4683	0.4945	0.5221	0.0019	0.3407	0.0442
		300	0.4428	0.4807	0.4793	0.5359	0.0013	0.3105	0.0448
0.5	6	100	0.4776	0.5123	0.5748	0.5762	0.0205	0.3652	0.0689
		200	0.4830	0.4936	0.5495	0.5505	0.0104	0.3674	0.0635
		300	0.4612	0.4821	0.5117	0.5436	0.0043	0.3641	0.0573
	10	100	0.4486	0.4955	0.4566	0.5207	0.0069	0.4466	0.0487
		200	0.4591	0.4741	0.4673	0.4880	0.0046	0.4929	0.0264
		300	0.4395	0.4767	0.4711	0.5007	0.0028	0.4474	0.0323
	15	100	0.4455	0.4637	0.4935	0.5040	0.0045	0.4096	0.0467
		200	0.4488	0.4611	0.4921	0.5107	0.0027	0.4272	0.0443
		300	0.4418	0.4749	0.4764	0.5180	0.0016	0.3950	0.0482
	20	100	0.4469	0.4738	0.5010	0.5282	0.0031	0.4580	0.0396
		200	0.4555	0.4718	0.5029	0.5171	0.0019	0.4766	0.0343
		300	0.4393	0.4779	0.4862	0.5331	0.0013	0.4779	0.0286
0.7	6	100	0.4890	0.4884	0.5521	0.5767	0.0172	0.4757	0.0867
		200	0.4921	0.4924	0.5355	0.5598	0.0099	0.4884	0.0747
		300	0.4494	0.4789	0.5101	0.5448	0.0044	0.4758	0.0860
	10	100	0.4502	0.5001	0.4220	0.5270	0.0101	0.5646	0.0481
		200	0.4580	0.4720	0.4523	0.4971	0.0051	0.5830	0.0350
		300	0.4432	0.4744	0.4636	0.5036	0.0030	0.5612	0.0407
	15	100	0.4491	0.4564	0.4744	0.5031	0.0052	0.5592	0.0460
		200	0.4468	0.4543	0.4776	0.5105	0.0032	0.5869	0.0334
		300	0.4413	0.4715	0.4710	0.5178	0.0020	0.5704	0.0406
	20	100	0.4403	0.4656	0.4778	0.5258	0.0036	0.6482	0.0243
		200	0.4551	0.4706	0.4899	0.5214	0.0019	0.6474	0.0204
		300	0.4421	0.4819	0.4733	0.5302	0.0014	0.6365	0.0211

B

Full Conditional Distribution of DMSOP Model

B.1 Full Conditional Distribution of Parameters

B.1.1 The Full Conditional Distribution of $\boldsymbol{\beta}_0$

The full conditional distribution (FCD) of $\boldsymbol{\beta}_0$ is expressed as follows:

$$
\begin{aligned}
&p(\boldsymbol{\beta}_0|\star) \propto p(\boldsymbol{z}_0|\boldsymbol{\beta}_0, \boldsymbol{\delta}, \boldsymbol{V})\pi(\boldsymbol{\beta}_0) \\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{z}_0 - \boldsymbol{X}_0\boldsymbol{\beta}_0 - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - \boldsymbol{X}_0\boldsymbol{\beta}_0 - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})\right\} \\
&\quad \times \exp\left[-\frac{1}{2}(\boldsymbol{\beta}_0 - \boldsymbol{c}_0)'\boldsymbol{\Gamma}_0^{-1}(\boldsymbol{\beta}_0 - \boldsymbol{c}_0)\right] \\
&\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\beta}_0'\boldsymbol{X}_0'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})\boldsymbol{X}_0\boldsymbol{\beta}_0 - 2\boldsymbol{\beta}_0'\boldsymbol{X}_0'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}) + \boldsymbol{\beta}_0'\boldsymbol{\Gamma}_0^{-1}\boldsymbol{\beta}_0 - 2\boldsymbol{\beta}_0'\boldsymbol{\Gamma}_0^{-1}\boldsymbol{c}_0\right]\right\} \\
&= \exp\left\{-\frac{1}{2}\boldsymbol{\beta}_0'[\boldsymbol{X}_0'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})\boldsymbol{X}_0 + \boldsymbol{\Gamma}_0^{-1}]\boldsymbol{\beta}_0 + \boldsymbol{\beta}_0'\left[\boldsymbol{X}_0'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}) + \boldsymbol{\Gamma}_0^{-1}\boldsymbol{c}_0\right]\right\}.
\end{aligned}
$$

Defining $\boldsymbol{B}_{\boldsymbol{\beta}_0} = \boldsymbol{X}_0'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})\boldsymbol{X}_0 + \boldsymbol{\Gamma}_0^{-1}$, and $\boldsymbol{b}_{\boldsymbol{\beta}_0} = \boldsymbol{X}_0'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}) + \boldsymbol{\Gamma}_0^{-1}\boldsymbol{c}_0$, the FCD of $\boldsymbol{\beta}$ can be expressed as:

$$\boldsymbol{\beta}_0|\star \sim \mathrm{N}(\boldsymbol{B}_{\boldsymbol{\beta}_0}^{-1}\boldsymbol{b}_{\boldsymbol{\beta}_0}, \boldsymbol{B}_{\boldsymbol{\beta}_0}^{-1}).$$

B.1.2 The Full Conditional Distribution of $\boldsymbol{\beta}$

The FCD of $\boldsymbol{\beta}$ is expressed as follows:

$$
\begin{aligned}
&p(\boldsymbol{\beta}|\star) \propto p(\boldsymbol{z}|\boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda})\pi(\boldsymbol{\beta}) \\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{\Theta}\boldsymbol{\delta})'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{\Theta}\boldsymbol{\delta})\right\} \exp\left[-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{c})'\boldsymbol{\Gamma}^{-1}(\boldsymbol{\beta} - \boldsymbol{c})\right] \\
&\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})\boldsymbol{X}\boldsymbol{\beta} - 2\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{\Theta}\boldsymbol{\delta}) + \boldsymbol{\beta}'\boldsymbol{\Gamma}^{-1}\boldsymbol{\beta} - 2\boldsymbol{\beta}'\boldsymbol{\Gamma}^{-1}\boldsymbol{c}\right]\right\} \\
&= \exp\left\{-\frac{1}{2}\boldsymbol{\beta}'[\boldsymbol{X}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})\boldsymbol{X} + \boldsymbol{\Gamma}^{-1}]\boldsymbol{\beta} + \boldsymbol{\beta}'\left[\boldsymbol{X}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{\Theta}\boldsymbol{\delta}) + \boldsymbol{\Gamma}^{-1}\boldsymbol{c}\right]\right\}.
\end{aligned}
$$

Defining $\boldsymbol{B}_{\boldsymbol{\beta}} = \boldsymbol{X}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})\boldsymbol{X} + \boldsymbol{\Gamma}^{-1}$, and $\boldsymbol{b}_{\boldsymbol{\beta}} = \boldsymbol{X}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{\Theta}\boldsymbol{\delta}) + \boldsymbol{\Gamma}^{-1}\boldsymbol{c}$, the FCD of $\boldsymbol{\beta}$ can be expressed as:

$$\boldsymbol{\beta}|\star \sim \mathrm{N}(\boldsymbol{B}_{\boldsymbol{\beta}}^{-1}\boldsymbol{b}_{\boldsymbol{\beta}}, \boldsymbol{B}_{\boldsymbol{\beta}}^{-1}).$$

B.1.3 The Full Conditional Distribution of $\boldsymbol{\delta}$

The FCD of $\boldsymbol{\delta}$ is expressed as follows:

$$
\begin{aligned}
&p(\boldsymbol{\delta}|\star) \propto p(\boldsymbol{z}|\boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) p(\boldsymbol{z}_0|\boldsymbol{\beta}_0, \boldsymbol{\delta}, \boldsymbol{V}) p(\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2) \\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{\Theta}\boldsymbol{\delta})'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{\Theta}\boldsymbol{\delta})\right\} \exp\left(-\frac{1}{2}\boldsymbol{\delta}'\boldsymbol{B}_\rho'(\boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho\boldsymbol{\delta}\right) \\
&\quad \times \exp\left\{-\frac{1}{2}(\boldsymbol{z}_0 - \boldsymbol{X}_0\boldsymbol{\beta}_0 - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - \boldsymbol{X}_0\boldsymbol{\beta}_0 - (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta})\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\delta}'\boldsymbol{\Theta}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})\boldsymbol{\Theta}\boldsymbol{\delta} - 2\boldsymbol{\delta}'\boldsymbol{\Theta}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{X}\boldsymbol{\beta}) + \boldsymbol{\delta}'(\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)\boldsymbol{\delta}\right.\right. \\
&\qquad\qquad \left.\left. -2\boldsymbol{\delta}'(\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - \boldsymbol{X}_0\boldsymbol{\beta}_0) + \boldsymbol{\delta}'\boldsymbol{B}_\rho'(\boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho\boldsymbol{\delta}\right]\right\} \\
&= \exp\left\{-\frac{1}{2}\boldsymbol{\delta}'\left[\boldsymbol{\Theta}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})\boldsymbol{\Theta} + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{\Delta} \otimes \boldsymbol{I}_S) + \boldsymbol{B}_\rho'(\boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho\right]\boldsymbol{\delta}\right. \\
&\qquad\qquad \left. +\boldsymbol{\delta}'[\boldsymbol{\Theta}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{X}\boldsymbol{\beta}) + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - \boldsymbol{X}_0\boldsymbol{\beta}_0)]\right\}.
\end{aligned}
$$

Defining $\boldsymbol{\Omega} = \boldsymbol{\Theta}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})\boldsymbol{\Theta} + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{\Delta} \otimes \boldsymbol{I}_S) + \boldsymbol{B}_\rho'(\boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho$, and $\boldsymbol{\eta} = \boldsymbol{\Theta}'(\boldsymbol{I}_{Tn} \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}^{\boldsymbol{\Lambda}} - \boldsymbol{X}\boldsymbol{\beta}) + (\boldsymbol{\Delta} \otimes \boldsymbol{I}_S)'(\boldsymbol{I}_n \otimes \boldsymbol{V}^{-1})(\boldsymbol{z}_0 - \boldsymbol{X}_0\boldsymbol{\beta}_0)$, the FCD of $\boldsymbol{\delta}$ can be expressed as:

$$\boldsymbol{\delta}|\star \sim \mathrm{N}(\boldsymbol{\Omega}^{-1}\boldsymbol{\eta}, \boldsymbol{\Omega}^{-1}),$$

where $\boldsymbol{\Omega}^{-1}$ depends on $\boldsymbol{B}_\rho$, which depends on ρ.

B.1.4 The Full Conditional Distribution of ρ

The FCD of ρ is expressed as follows:

$$
\begin{aligned}
p(\rho|\star) &\propto p(\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2)\pi(\rho) \\
&\propto |\boldsymbol{B}_\rho| \exp\left(-\frac{1}{2}\boldsymbol{\delta}'\boldsymbol{B}_\rho'(\boldsymbol{I}_m \otimes \mathrm{diag}(\boldsymbol{\sigma}^2))^{-1}\boldsymbol{B}_\rho\boldsymbol{\delta}\right)(1+\rho)^{a_0-1}(1-\rho)^{b_0-1} \times 1(-1 \leq \rho \leq 1).
\end{aligned}
$$

Since $p(\rho|\star)$ is not a standard distribution, we can use a M-H algorithm to obtain the samples.

1) Sample ρ^{new} from the candidate density: $\rho^{new} = \rho^{old} + aG$, where a is a known constant, and G can be a standard normal distribution or t distribution. In this study, we set $G \sim \mathrm{N}(0, 1)$.

2) Since $\rho \in [-1, 1]$, if proposed ρ^{new} is not within the interval, it is rejected with probability 1. Otherwise, we calculate the acceptance probability $\alpha(\rho^{new},\ \rho^{old}) = \min(\frac{p(\rho^{new}|\star)}{p(\rho^{old}|\star)}, 1)$.

3) Draw a uniform random deviate U. If $U < \alpha$, we set $\rho = \rho^{new}$, otherwise, $\rho = \rho^{old}$.

4) Calculate the updated ratio of ρ. Then, a is adjusted based on the updated ratio. If the updated ratio is less than 0.4, and then $a = a/1.1$; and if the updated ratio is less than 0.6, and then $a = a * 1.1$. Thus, the updated ratio is always close to 0.5.

B.1.5 The Full Conditional Distribution of $\boldsymbol{\sigma}^2$

The FCD of $\boldsymbol{\sigma}^2$ is expressed as follows:

$$p(\boldsymbol{\sigma}^2|\star) \propto p(\boldsymbol{\delta}|\rho, \boldsymbol{\sigma}^2)\pi(\boldsymbol{\sigma}^2).$$

Given $\boldsymbol{\beta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0, \boldsymbol{y}_0$ and $\boldsymbol{y}$, $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \ldots, \boldsymbol{\delta}_S$ are independent, and thus,

$$\begin{aligned}
p(\boldsymbol{\sigma}^2|\star) &\propto \prod_{s=1}^{S} \left\{p(\boldsymbol{\delta}_s|\rho, \boldsymbol{\sigma}^2)\pi(\sigma_s^2)\right\} \\
&\propto \prod_{s=1}^{S} \left\{(\sigma_s^2)^{-m/2} \exp\left(-\frac{1}{2\sigma_s^2}\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s\right)(\sigma_s^2)^{-(e_0+1)} \exp\left(-\frac{g_0}{\sigma_s^2}\right)\right\} \\
&\propto \prod_{s=1}^{S} \left\{(\sigma_s^2)^{-(m/2+e_0+1)} \exp\left(-\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g_0}{2\sigma_s^2}\right)\right\}.
\end{aligned}$$

Extract the element σ_s^2 from $\boldsymbol{\sigma}^2$, and then

$$p(\sigma_s^2|\star) \propto (\sigma_s^2)^{-(m/2+e_0+1)} \exp\left(-\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g_0}{2\sigma_s^2}\right).$$

Let $\theta_s = \frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s+2g_0}{\sigma_s^2}$, and thus, $\sigma_s^2 = \frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s+2g_0}{\theta_s}$. Therefore, $\left|\frac{\mathrm{d}\,\sigma_s^2}{\mathrm{d}\,\theta_s}\right| = \frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s+2g_0}{\theta_s^2}$. Then, the FCD of $\boldsymbol{\sigma}_s^2$ is expressed as follows:

$$\begin{aligned}
p(\theta_s|\star) &= p(\sigma_s^2(\theta_s)|\star)\left|\frac{\mathrm{d}\,\sigma_s^2}{\mathrm{d}\,\theta_s}\right| \\
&\propto \left(\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g_0}{\theta_s}\right)^{-(m/2+e_0+1)} \exp\left(-\frac{\theta_s}{2}\right)\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g_0}{\theta_s^2} \\
&\propto \theta_s^{\frac{m+2e_0}{2}-1} \exp\left(-\frac{\theta_s}{2}\right).
\end{aligned}$$

Therefore, the FCD of σ_s^2 can be expressed as:

$$\frac{\boldsymbol{\delta}_s'\boldsymbol{A}_\rho'\boldsymbol{A}_\rho\boldsymbol{\delta}_s + 2g_0}{\sigma_s^2}|\star \sim \chi^2(m + 2e_0).$$

More specifically, if $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_S^2 = \sigma_0^2$, and then

$$\begin{aligned}
p(\sigma_0^2|\star) &\propto p(\boldsymbol{\delta}|\rho, \sigma_0^2)\pi(\sigma_0^2) \\
&\propto (\sigma_0^2)^{-mS/2} \exp\left(-\frac{1}{2\sigma_0^2}\boldsymbol{\delta}_s'\boldsymbol{B}_\rho'\boldsymbol{B}_\rho\boldsymbol{\delta}_s\right)(\sigma_0^2)^{-(e_0+1)} \exp\left(-\frac{g_0}{\sigma_0^2}\right) \\
&\propto (\sigma_0^2)^{-(mS/2+e_0+1)} \exp\left(-\frac{\boldsymbol{\delta}'\boldsymbol{B}_\rho'\boldsymbol{B}_\rho\boldsymbol{\delta} + 2g_0}{2\sigma_0^2}\right).
\end{aligned}$$

Similarly, the FCD of σ_0^2 can be expressed as:

$$\frac{\boldsymbol{\delta}'\boldsymbol{B}_\rho'\boldsymbol{B}_\rho\boldsymbol{\delta} + 2g_0}{\sigma_0^2}|\star \sim \chi^2(mS + 2e_0).$$

B.1.6 The Full Conditional Distribution of $\boldsymbol{V}$

The FCD of $\boldsymbol{V}$ is expressed as follows:

$$
\begin{aligned}
&p(\boldsymbol{V}|\star) \propto p(\boldsymbol{z}|\boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda})p(\boldsymbol{z}_0|\boldsymbol{\beta}_0, \boldsymbol{\delta}, \boldsymbol{V})\pi(\boldsymbol{V}) \\
&\propto \left\{\prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i} |\boldsymbol{V}|^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)\right]\right\} \\
&\quad \times \left\{\prod_{i=1}^{m}\prod_{k=1}^{n_i} |\boldsymbol{V}|^{-1/2} \exp\left[-\frac{1}{2}(\boldsymbol{z}_{ik} - \boldsymbol{X}_{ik}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik} - \boldsymbol{X}_{ik}\boldsymbol{\beta} - \boldsymbol{\delta}_i)\right]\right\} |\boldsymbol{V}|^{-\frac{n_0+S+1}{2}} \exp\left[-\frac{1}{2}\operatorname{tr}(\boldsymbol{Q}_0\boldsymbol{V}^{-1})\right] \\
&= |\boldsymbol{V}|^{-\frac{(T+1)n+n_0+S+1}{2}} \exp\left\{-\frac{1}{2}\left[\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)\right.\right. \\
&\qquad\qquad \left.\left. + \sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ik0} - \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik0} - \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 - \boldsymbol{\delta}_i) + \operatorname{tr}(\boldsymbol{Q}_0\boldsymbol{V}^{-1})\right]\right\} \\
&= |\boldsymbol{V}|^{-\frac{(T+1)n+n_0+S+1}{2}} \exp\left\{-\frac{1}{2}\operatorname{tr}\left[\left(\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\right.\right.\right. \\
&\qquad\qquad \left.\left.\left. + \sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ik0} - \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 - \boldsymbol{\delta}_i)(\boldsymbol{z}_{ik0} - \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 - \boldsymbol{\delta}_i)' + \boldsymbol{Q}_0\right)\boldsymbol{V}^{-1}\right]\right\}.
\end{aligned}
$$

Defining $\tilde{n} = (T+1)n + n_0$, and $\tilde{\boldsymbol{Q}} = \sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)' + \sum_{i=1}^{m}\sum_{k=1}^{n_i}(\boldsymbol{z}_{ik0} - \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 - \boldsymbol{\delta}_i)(\boldsymbol{z}_{ik0} - \boldsymbol{X}_{ik0}\boldsymbol{\beta}_0 - \boldsymbol{\delta}_i)' + \boldsymbol{Q}_0$, the FCD of $\boldsymbol{V}$ can be expressed as:

$$
\boldsymbol{V}|\star \sim \mathrm{IW}(\tilde{n}, \tilde{\boldsymbol{Q}}).
$$

B.1.7 The Full Conditional Distribution of $\boldsymbol{\Lambda}$

Let $\boldsymbol{z}_{ik(-s)t}$ denotes the vector removing z_{ikst} from $\boldsymbol{z}_{ikt}$. For the sake of convenience of expression, we reorganize $\boldsymbol{z}_{ikt}$, $\boldsymbol{\Lambda}$, $\boldsymbol{X}_{ikt}$, $\boldsymbol{\beta}$, $\boldsymbol{\delta}_i$ and $\boldsymbol{V}$ as follows (the sth element become the first element):

$$\boldsymbol{z}_{ikt} = \begin{pmatrix} z_{ikst} \\ \boldsymbol{z}_{ik(-s)t} \end{pmatrix}, \ \boldsymbol{\Lambda} = \begin{pmatrix} \lambda_{sg} & \boldsymbol{\lambda}_{s,-g} \\ \boldsymbol{\lambda}_{-s,g} & \boldsymbol{\Lambda}_{-s,-g} \end{pmatrix}, \ \boldsymbol{X}_{ikt} = \begin{pmatrix} x'_{ikst} & \\ & \boldsymbol{X}_{ik(-s)t} \end{pmatrix},$$

$$\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta} s \\ \boldsymbol{\beta}_{-s} \end{pmatrix}, \ \boldsymbol{\delta}_i = \begin{pmatrix} \delta_{is} \\ \boldsymbol{\delta}_{i(-s)} \end{pmatrix}, \text{ and } \boldsymbol{V} = \begin{pmatrix} V_{s,s} & \boldsymbol{v}_{s,-s} \\ \boldsymbol{v}_{-s,s} & \boldsymbol{V}_{-s,-s} \end{pmatrix}, \ \boldsymbol{z}_{ik(t-1)} = \begin{pmatrix} z_{ikg(t-1)} \\ \boldsymbol{z}_{ik(-g)(t-1)} \end{pmatrix},$$

where $\boldsymbol{\lambda}_{s,-g} = (\lambda_{s1}, \dots, \lambda_{s,g-1}, \lambda_{s,g+1}, \dots, \lambda_{sS})$ and $\boldsymbol{v}_{s,-s} = (V_{s,1}, \dots, V_{s,s-1}, V_{s,s+1}, \dots, V_{s,S})$ are $1 \times (S-1)$ vectors, $\boldsymbol{\lambda}_{-s,g} = (\lambda_{1g}, \dots, \lambda_{s-1,g}, \lambda_{s+1,g}, \dots, \lambda_{Sg})'$ and $\boldsymbol{v}_{-s,s} = (V_{1,s}, \dots, V_{s-1,s}, V_{s+1,s}, \dots, V_{S,s})'$ are $(S-1) \times 1$ vector, $\boldsymbol{\Lambda}_{-s,-g}$ and $\boldsymbol{V}_{-s,-s}$ denote $(S-1) \times (S-1)$ matrices removing above elements from $\boldsymbol{\Lambda}$ and $\boldsymbol{V}$, respectively.

Defining $h_{ikst} = \lambda_{sg} z_{ikg(t-1)} + \boldsymbol{\lambda}_{s,-g} \boldsymbol{z}_{ik(-g)(t-1)} + \boldsymbol{x}'_{ikst} \boldsymbol{\beta}_s + \delta_{is}$, and $\boldsymbol{h}_{ik(-s)t} = \boldsymbol{\lambda}_{-s,g} z_{ikg(t-1)} + \boldsymbol{\Lambda}_{-s,-g} \boldsymbol{z}_{ik(-g)(t-1)} + \boldsymbol{X}_{ik(-s)t} \boldsymbol{\beta}_{-s} + \boldsymbol{\delta}_{i(-s)}$, $\boldsymbol{z}_{ikt}$ has the following multivariate normal distribution:

$$\boldsymbol{z}_{ikt} = \begin{pmatrix} z_{ikst} \\ \boldsymbol{z}_{ik(-s)t} \end{pmatrix} \sim \mathrm{N} \left[\begin{pmatrix} h_{ikst} \\ \boldsymbol{h}_{ik(-s)t} \end{pmatrix}, \begin{pmatrix} V_{s,s} & \boldsymbol{v}_{s,-s} \\ \boldsymbol{v}_{-s,s} & \boldsymbol{V}_{-s,-s} \end{pmatrix} \right].$$

Based on the property of multivariate normal distribution, the conditional distribution of z_{ikst} given $\boldsymbol{z}_{ik(-s)t}, \boldsymbol{z}_{ik(t-1)}, \boldsymbol{z}_{ik0}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}$ can be expressed as:

$$z_{ikst} | \boldsymbol{z}_{ik(-s)t}, \boldsymbol{z}_{ik(t-1)}, \boldsymbol{z}_{ik0}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda} \sim \mathrm{N}(\tilde{h}_{ikst}, \tilde{v}_s),$$

where

$$\tilde{h}_{ikst} = \lambda_{sg} z_{ikg(t-1)} + \boldsymbol{\lambda}_{s,-g} \boldsymbol{z}_{ik(-g)(t-1)} + \boldsymbol{x}'_{ikst} \boldsymbol{\beta}_s + \delta_{is} + \boldsymbol{v}_{s,-s} \boldsymbol{V}^{-1}_{-s,-s} (\boldsymbol{z}_{ik(-s)t} - \boldsymbol{h}_{ik(-s)t}),$$
$$\tilde{v}_s = V_{s,s} - \boldsymbol{v}_{s,-s} \boldsymbol{V}^{-1}_{-s,-s} \boldsymbol{v}_{-s,s}.$$

The FCD of $\boldsymbol{\Lambda}$ is expressed as follows:

$$
\begin{aligned}
&p(\boldsymbol{\Lambda}|\star) \propto p(\boldsymbol{z}|\boldsymbol{z}_0, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda})\pi(\boldsymbol{\Lambda}) \\
&\propto \prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i} p(\boldsymbol{z}_{ikt}|\boldsymbol{z}_{ik(t-1)}, \boldsymbol{z}_{ik0}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \prod_{s=1}^{S}\prod_{g=1}^{S}\pi(\lambda_{sg}) \\
&\propto \prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i} p(\boldsymbol{z}_{ikst}|\boldsymbol{z}_{ik(-s)t}, \boldsymbol{z}_{ik(t-1)}, \boldsymbol{z}_{ik0}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) p(\boldsymbol{z}_{ik(-s)t}|\boldsymbol{z}_{ik(t-1)}, \boldsymbol{z}_{ik0}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda})
\end{aligned}
$$

Extract the element λ_{sg} from $\boldsymbol{\Lambda}$, and then

$$
\begin{aligned}
p(\lambda_{sg}|\star) &\propto \prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i} p(\boldsymbol{z}_{ikst}|\boldsymbol{z}_{ik(-s)t}, \boldsymbol{z}_{ik(t-1)}, \boldsymbol{z}_{ik0}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) \\
&\propto \prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i} \exp\left[-\frac{1}{2\tilde{v}_s}(z_{ikst} - \tilde{h}_{ikst})^2\right] \\
&\propto \prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i} \exp\Big\{-\frac{1}{2\tilde{v}_s}\big[z_{ikst} - \big(\lambda_{sg} z_{ikg(t-1)} + \boldsymbol{\lambda}_{s,-g}\boldsymbol{z}_{ik(-g)(t-1)} + \boldsymbol{x}'_{ikst}\boldsymbol{\beta}_s + \delta_{is} \\
&\qquad + \boldsymbol{v}_{s,-s}\boldsymbol{V}^{-1}_{-s,-s}(\boldsymbol{z}_{ik(-s)t} - \boldsymbol{h}_{ik(-s)t})\big)\big]^2\Big\}.
\end{aligned}
$$

Defining $A_{iksgt} = z_{ikst} - (\boldsymbol{\lambda}_{s,-g}\boldsymbol{z}_{ik(-g)(t-1)} + \boldsymbol{x}'_{ikst}\boldsymbol{\beta}_s + \delta_{is} + \boldsymbol{v}_{s,-s}\boldsymbol{V}^{-1}_{-s,-s}(\boldsymbol{z}_{ik(-s)t} - \boldsymbol{h}_{ik(-s)t}))$,

$$
\begin{aligned}
p(\lambda_{sg}|\star) &\propto \prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i} \exp\left[-\frac{1}{2\tilde{v}_s}(A_{iksgt} - \lambda_{sg} z_{ikg(t-1)})^2\right] \\
&\propto \exp\left[-\frac{1}{2}\lambda_{sg}^2 \frac{\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i} z_{ikg(t-1)}^2}{\tilde{v}_s} + \lambda_{sg}\frac{\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i} A_{iksgt} z_{ikg(t-1)}}{\tilde{v}_s}\right].
\end{aligned}
$$

Therefore, the FCD of λ_{sg} can be expressed as:

$$\lambda_{sg}|\star \sim \mathrm{N}\left(\frac{\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i} A_{iksgt} z_{ikg(t-1)}}{\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i} z_{ikg(t-1)}^2}, \frac{\tilde{v}_s}{\sum_{t=1}^{T}\sum_{i=1}^{m}\sum_{k=1}^{n_i} z_{ikg(t-1)}^2}\right).$$

In each draw, the value of λ_{sg} needs to be limited to $(-1,1)$. Moreover, we only accept those λ_{sg} where $\sum_{g=1}^{S}|\lambda_{sg}| \leq 1$ for $s = 1,\ldots,S$.

B.1.8 The Full Conditional Distribution of $\boldsymbol{z}_0$

The FCD of $\boldsymbol{z}_0$ is expressed as follows:

$$\begin{aligned}
&p(\boldsymbol{z}_0|\star) \propto p(\boldsymbol{z}|\boldsymbol{z}_0,\boldsymbol{\beta},\boldsymbol{\delta},\boldsymbol{V},\boldsymbol{\Lambda})p(\boldsymbol{z}_0|\boldsymbol{\beta}_0,\boldsymbol{\delta},\boldsymbol{V})p(\boldsymbol{y}_0|\boldsymbol{z}_0,\boldsymbol{\gamma})\\
&\propto \left\{\prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i}\exp\left[-\frac{1}{2}(\boldsymbol{z}_{ikt}-\boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)}-\boldsymbol{X}_{ikt}\boldsymbol{\beta}-\boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ikt}-\boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)}-\boldsymbol{X}_{ikt}\boldsymbol{\beta}-\boldsymbol{\delta}_i)\right]\right\}\\
&\times\left\{\prod_{i=1}^{m}\prod_{k=1}^{n_i}|\boldsymbol{V}|^{-1/2}\exp\left[-\frac{1}{2}(\boldsymbol{z}_{ik0}-\boldsymbol{X}_{ik0}\boldsymbol{\beta}_0-\boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik0}-\boldsymbol{X}_{ik0}\boldsymbol{\beta}_0-\boldsymbol{\delta}_i)\right]\right\}\times\left\{\prod_{i=1}^{m}\prod_{k=1}^{n_i}1(\boldsymbol{z}_{ik0}\in G_{ik})\right\}\\
&\propto \prod_{i=1}^{m}\prod_{k=1}^{n_i}\exp\left\{-\frac{1}{2}\left[(\boldsymbol{z}_{ik1}-\boldsymbol{\Lambda}\boldsymbol{z}_{ik0}-\boldsymbol{X}_{ik1}\boldsymbol{\beta}-\boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik1}-\boldsymbol{\Lambda}\boldsymbol{z}_{ik0}-\boldsymbol{X}_{ik1}\boldsymbol{\beta}-\boldsymbol{\delta}_i)\right.\right.\\
&\qquad\qquad\left.\left.+(\boldsymbol{z}_{ik0}-\boldsymbol{X}_{ik0}\boldsymbol{\beta}_0-\boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik0}-\boldsymbol{X}_{ik0}\boldsymbol{\beta}_0-\boldsymbol{\delta}_i)\right]\right\}1(\boldsymbol{z}_{ik0}\in G_{ik})\\
&\propto \prod_{i=1}^{m}\prod_{k=1}^{n_i}\exp\left\{-\frac{1}{2}\boldsymbol{z}_{ik0}'(\boldsymbol{\Lambda}'\boldsymbol{V}^{-1}\boldsymbol{\Lambda}+\boldsymbol{V}^{-1})\boldsymbol{z}_{ik0}+\boldsymbol{z}_{ik0}'[\boldsymbol{\Lambda}'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik1}-\boldsymbol{X}_{ik1}\boldsymbol{\beta}-\boldsymbol{\delta}_i)+\boldsymbol{V}^{-1}(\boldsymbol{X}_{ik0}\boldsymbol{\beta}_0+\boldsymbol{\delta}_i)]\right\}1(\boldsymbol{z}_{ik0}\in G_{ik})
\end{aligned}$$

Defining $\boldsymbol{B}_{\boldsymbol{z}_0} = \boldsymbol{\Lambda}'\boldsymbol{V}^{-1}\boldsymbol{\Lambda}+\boldsymbol{V}^{-1}$, and $\boldsymbol{b}_{\boldsymbol{z}_0} = \boldsymbol{\Lambda}'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik1}-\boldsymbol{X}_{ik1}\boldsymbol{\beta}-\boldsymbol{\delta}_i)+\boldsymbol{V}^{-1}(\boldsymbol{X}_{ik0}\boldsymbol{\beta}_0+\boldsymbol{\delta}_i)$, the FCD of $\boldsymbol{z}_{ik0}$ can be expressed as:

$$\boldsymbol{z}_{ik0}|\star \sim \mathrm{N}(\boldsymbol{B}_{\boldsymbol{z}_0}^{-1}\boldsymbol{b}_{\boldsymbol{z}_0}, \boldsymbol{B}_{\boldsymbol{z}_0}^{-1})1(\boldsymbol{z}_{ik0}\in G_{ik}).$$

B.1.9 The Full Conditional Distribution of $\boldsymbol{z}$

The FCD of $\boldsymbol{z}$ is expressed as follows:

$$
\begin{aligned}
&p(\boldsymbol{z}|\star) \propto p(\boldsymbol{y}|\boldsymbol{z},\boldsymbol{\gamma})p(\boldsymbol{z}|\boldsymbol{z}_0,\boldsymbol{\beta},\boldsymbol{\delta},\boldsymbol{V},\boldsymbol{\Lambda}) \\
&\propto \prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i} 1(\boldsymbol{z}_{ikt} \in G_{ik}) \exp\left[-\frac{1}{2}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)\right].
\end{aligned}
$$

For $t = 1, \ldots, T-1$, z_{ikt} appears for both periods t and $t+1$. Extract $\boldsymbol{z}_{ikt}$ from $\boldsymbol{z}$,

$$
\begin{aligned}
p(\boldsymbol{z}_{ikt}|\star) &\propto 1(\boldsymbol{z}_{ikt} \in G_{ik}) \exp\left[-\frac{1}{2}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ikt} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} - \boldsymbol{X}_{ikt}\boldsymbol{\beta} - \boldsymbol{\delta}_i)\right] \\
&\quad \times \exp\left[-\frac{1}{2}(\boldsymbol{z}_{ik(t+1)} - \boldsymbol{\Lambda}\boldsymbol{z}_{ikt} - \boldsymbol{X}_{ik(t+1)}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik(t+1)} - \boldsymbol{\Lambda}\boldsymbol{z}_{ikt} - \boldsymbol{X}_{ik(t+1)}\boldsymbol{\beta} - \boldsymbol{\delta}_i)\right] \\
&\propto 1(\boldsymbol{z}_{ikt} \in G_{ik}) \exp\left\{-\frac{1}{2}\boldsymbol{z}_{ikt}'(\boldsymbol{V}^{-1} + \boldsymbol{\Lambda}'\boldsymbol{V}^{-1}\boldsymbol{\Lambda})\boldsymbol{z}_{ikt} + \boldsymbol{z}_{ikt}'[\boldsymbol{V}^{-1}(\boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} + \boldsymbol{X}_{ikt}\boldsymbol{\beta} + \boldsymbol{\delta}_i)\right. \\
&\qquad \left. + \boldsymbol{\Lambda}'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik(t+1)} - \boldsymbol{X}_{ik(t+1)}\boldsymbol{\beta} - \boldsymbol{\delta}_i)]\right\}.
\end{aligned}
$$

Defining $\boldsymbol{B}_{\boldsymbol{z}_t} = \boldsymbol{V}^{-1} + \boldsymbol{\Lambda}'\boldsymbol{V}^{-1}\boldsymbol{\Lambda}$, and $\boldsymbol{b}_{\boldsymbol{z}_t} = \boldsymbol{V}^{-1}(\boldsymbol{\Lambda}\boldsymbol{z}_{ik(t-1)} + \boldsymbol{X}_{ikt}\boldsymbol{\beta} + \boldsymbol{\delta}_i) + \boldsymbol{\Lambda}'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ik(t+1)} - \boldsymbol{X}_{ik(t+1)}\boldsymbol{\beta} - \boldsymbol{\delta}_i)$, the FCD of $\boldsymbol{z}_{ikt}$ can be expressed as:

$$
\boldsymbol{z}_{ikt}|\star \sim \mathrm{N}(\boldsymbol{B}_{\boldsymbol{z}_t}^{-1}\boldsymbol{b}_{\boldsymbol{z}_t}, \boldsymbol{B}_{\boldsymbol{z}_t}^{-1})1(\boldsymbol{z}_{ikt} \in G_{ik}),\ t = 1, \ldots, T-1.
$$

For $t = T$, z_{ikT} only appears for periods T. Extract $\boldsymbol{z}_{ikT}$ from $\boldsymbol{z}$,

$$
\begin{aligned}
&p(\boldsymbol{z}_{ikT}|\star) \propto 1(\boldsymbol{z}_{ikT} \in G_{ik}) \exp\left[-\frac{1}{2}(\boldsymbol{z}_{ikT} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(T-1)} - \boldsymbol{X}_{ikT}\boldsymbol{\beta} - \boldsymbol{\delta}_i)'\boldsymbol{V}^{-1}(\boldsymbol{z}_{ikT} - \boldsymbol{\Lambda}\boldsymbol{z}_{ik(T-1)} - \boldsymbol{X}_{ikT}\boldsymbol{\beta} - \boldsymbol{\delta}_i)\right] \\
&\propto 1(\boldsymbol{z}_{ikT} \in G_{ik}) \exp\left\{-\frac{1}{2}\boldsymbol{z}_{ikT}'\boldsymbol{V}^{-1}\boldsymbol{z}_{ikT} + \boldsymbol{z}_{ikT}'\boldsymbol{V}^{-1}(\boldsymbol{\Lambda}\boldsymbol{z}_{ik(T-1)} + \boldsymbol{X}_{ikT}\boldsymbol{\beta} + \boldsymbol{\delta}_i)\right\}.
\end{aligned}
$$

Therefore, the FCD of $\boldsymbol{z}_{ikT}$ can be expressed as:

$$\boldsymbol{z}_{ikT}|\star \sim \mathrm{N}(\boldsymbol{\Lambda}\boldsymbol{z}_{ik(T-1)} + \boldsymbol{X}_{ikT}\boldsymbol{\beta} + \boldsymbol{\delta}_i, \boldsymbol{V})1(\boldsymbol{z}_{ikT} \in G_{ik}).$$

B.1.10 The Full Conditional Distribution of $\boldsymbol{\gamma}$

The FCD of $\boldsymbol{\gamma}_s$ is expressed as follows:

$$\begin{aligned}
&p(\boldsymbol{\gamma}_s|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \gamma_{-s}, \boldsymbol{z}, \boldsymbol{z}_0, \boldsymbol{\Lambda}, \boldsymbol{y}, \boldsymbol{y}_0) \propto p(\boldsymbol{y}, \boldsymbol{y}_0|\gamma_s, \boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0, \boldsymbol{\Lambda})\pi(\boldsymbol{\gamma}_s) \\
&\propto \prod_{t=0}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i}\sum_{l_s=2}^{L_s-1} p(y_{ikst} = l_s|\boldsymbol{\beta}, \boldsymbol{\delta}, \rho, \boldsymbol{\sigma}^2, \boldsymbol{V}, \boldsymbol{\gamma}_s, \boldsymbol{z}, \boldsymbol{z}_0, \boldsymbol{\Lambda})1(\gamma_{s,2} < \cdots < \gamma_{s,L_s-1}) \\
&\propto \left\{\prod_{t=1}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i}\sum_{l_s=2}^{L_s-1} 1(y_{ikst} = l_s)p(\gamma_{s,l_s-1} < z_{ikst} \le \gamma_{s,l_s}|\boldsymbol{z}_{ik(-s)t}, \boldsymbol{z}_{ik(t-1)}, \boldsymbol{z}_{ik0}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda})\right\} \\
&\quad \times \left\{\prod_{i=1}^{m}\prod_{k=1}^{n_i}\sum_{l_s=2}^{L_s-1} 1(y_{iks0} = l_s)p(\gamma_{s,l_s-1} < z_{iks0} \le \gamma_{s,l_s}|\boldsymbol{\beta}_0, \boldsymbol{\delta}, \boldsymbol{V})\right\} 1(\gamma_{s,2} < \cdots < \gamma_{s,L_s-1}).
\end{aligned}$$

As explained before, $z_{ikst}|\boldsymbol{z}_{ik(-s)t}, \boldsymbol{z}_{ik(t-1)}, \boldsymbol{z}_{ik0}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda} \sim \mathrm{N}(\tilde{h}_{ikst}, \tilde{v}_s)$, therefore,

$$p(\gamma_{s,l_s-1} < z_{ikst} \le \gamma_{s,l_s}|\boldsymbol{z}_{ik(-s)t}, \boldsymbol{z}_{ik(t-1)}, \boldsymbol{z}_{ik0}, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{V}, \boldsymbol{\Lambda}) = \Phi\left(\frac{\gamma_{s,l_s} - \tilde{h}_{ikst}}{\tilde{v}_s^{1/2}}\right) - \Phi\left(\frac{\gamma_{s,l_s-1} - \tilde{h}_{ikst}}{\tilde{v}_s^{1/2}}\right),$$

and

$$p(\gamma_{s,l_s-1} < z_{iks0} \le \gamma_{s,l_s}|\boldsymbol{\beta}_0, \boldsymbol{\delta}, \boldsymbol{V}) = \Phi\left(\frac{\gamma_{s,l_s} - \tilde{h}_{iks0}}{\tilde{v}_s^{1/2}}\right) - \Phi\left(\frac{\gamma_{s,l_s-1} - \tilde{h}_{iks0}}{\tilde{v}_s^{1/2}}\right),$$

where

$$\tilde{h}_{iks0} = \boldsymbol{x}'_{iks0}\boldsymbol{\beta}_{s0} + \delta_{is} + \boldsymbol{v}_{s,-s}\boldsymbol{V}^{-1}_{-s,-s}(\boldsymbol{z}_{ik(-s)0} - \boldsymbol{X}_{ik(-s)0}\boldsymbol{\beta}_{(-s)0} + \boldsymbol{\delta}_{i(-s)}).$$

As a result,

$$
\begin{aligned}
&p(\boldsymbol{\gamma}_s|\boldsymbol{\beta},\boldsymbol{\delta},\rho,\boldsymbol{\sigma}^2,\boldsymbol{V},\boldsymbol{\gamma}_{-s},\boldsymbol{z},\boldsymbol{z}_0,\boldsymbol{\Lambda},\boldsymbol{y})\\
&\propto \prod_{t=0}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i}\sum_{l_s=2}^{L_s-1} 1(y_{ikst}=l_s)\left[\Phi\left(\frac{\gamma_{s,l_s}-\tilde{h}_{ikst}}{\tilde{v}_s^{1/2}}\right)-\Phi\left(\frac{\gamma_{s,l_s-1}-\tilde{h}_{ikst}}{\tilde{v}_s^{1/2}}\right)\right]1(\gamma_{s,2}<\cdots<\gamma_{s,L_s-1}).
\end{aligned}
$$

Based on this form of $p(\boldsymbol{\gamma}|\star)$, Chen and Dey (2000) offered a better way to get the parameters $\boldsymbol{\gamma}$. This new method removes the ordering constraint using the one-to-one map:

$$
\gamma_{s,l_s}=\frac{\gamma_{s,l_s-1}+\exp(\kappa_{s,l_s})}{1+\exp(\kappa_{s,l_s})}, 2<l_s<L_s-2.
$$

Defining $\boldsymbol{\kappa}_s=(\kappa_{s,2},\ldots,\kappa_{s,Ls-2})'$, and we assume the prior for $\boldsymbol{\kappa}_s$ follows a normal distribution: $\boldsymbol{\kappa}_s\sim \mathrm{N}(\boldsymbol{\kappa}_{s0},\boldsymbol{K}_{s0})$. That is,

$$
\pi(\boldsymbol{\kappa}_s)\propto \exp\left[-\frac{1}{2}(\boldsymbol{\kappa}_s-\boldsymbol{\kappa}_{s0})'\boldsymbol{K}_{s0}^{-1}(\boldsymbol{\kappa}_s-\boldsymbol{\kappa}_{s0})\right].
$$

The Jacobian of transformation from $\boldsymbol{\gamma}_s$ to $\boldsymbol{\kappa}_s$ is as follows:

$$
\left|\frac{\partial\boldsymbol{\gamma}_s}{\partial\boldsymbol{\kappa}_s'}\right|=\prod_{l_s=2}^{L_s-2}\left|\frac{\partial\gamma_{s,l_s}}{\partial\kappa_{s,l_s}}\right|=\prod_{l_s=2}^{L_s-2}\frac{(1-\gamma_{s,l_s-1})\exp(\kappa_{s,l_s})}{(1+\exp(\kappa_{s,l_s}))^2}.
$$

Therefore,

$$
\begin{aligned}
p(\boldsymbol{\kappa}_s|\star)&=p(\boldsymbol{\gamma}_s|\star)\left|\frac{\partial\boldsymbol{\gamma}_s}{\partial\boldsymbol{\kappa}_s'}\right|\pi(\boldsymbol{\kappa}_s)\\
&\propto\left\{\prod_{t=0}^{T}\prod_{i=1}^{m}\prod_{k=1}^{n_i}\sum_{l_s=2}^{L_s-1} 1(y_{ikst}=l_s)\left[\Phi\left(\frac{\gamma_{s,l_s}-\tilde{h}_{ikst}}{\tilde{v}_s^{1/2}}\right)-\Phi\left(\frac{\gamma_{s,l_s-1}-\tilde{h}_{ikst}}{\tilde{v}_s^{1/2}}\right)\right]\right\}\\
&\quad\times\left\{\prod_{l_s=2}^{L_s-2}\frac{(1-\gamma_{s,l_s-1})\exp(\kappa_{s,l_s})}{(1+\exp(\kappa_{s,l_s}))^2}\right\}\exp\left[-\frac{1}{2}(\boldsymbol{\kappa}_s-\boldsymbol{\kappa}_{s0})'\boldsymbol{K}_{s0}^{-1}(\boldsymbol{\kappa}_s-\boldsymbol{\kappa}_{s0})\right].
\end{aligned}
$$

Since $p(\boldsymbol{\kappa}_s|\star)$ is not a standard distribution, we can use the a M-H algorithm to generate $\boldsymbol{\gamma}_s$ as follows:

1) Sample $\boldsymbol{\kappa}_s|\star$ marginally of $\boldsymbol{z}_s$ by drawing $\boldsymbol{\kappa}_s^{new}$ from the candidate student-t distribution $f_T(\boldsymbol{\kappa}_s|\hat{\boldsymbol{\kappa}_s}, \hat{\boldsymbol{K}}_s, v)$, where $\hat{\boldsymbol{\kappa}_s} = \text{argmax} p(\boldsymbol{\kappa}_s|\star), \hat{\boldsymbol{K}}_s = \left\{ \left[-\frac{\partial^2 \log p(\boldsymbol{\kappa}_s|\star)}{\partial \boldsymbol{\kappa}_s \partial \boldsymbol{\kappa}_s'} \right]_{\boldsymbol{\kappa}_s = \hat{\boldsymbol{\kappa}_s}} \right\}^{-1}$, and v is the freedom.

2) Given the current $\boldsymbol{\kappa}_s^{old}$ and the proposed $\boldsymbol{\kappa}_s^{new}$, accept $\boldsymbol{\kappa}_s^{new}$ with the probability:

$$\alpha(\boldsymbol{\kappa}_s^{old}, \boldsymbol{\kappa}_s^{new}) = \min \left(\frac{p(\boldsymbol{\kappa}_s^{new}|\star) f_T(\boldsymbol{\kappa}_s^{old}|\hat{\boldsymbol{\kappa}_s}, \hat{\boldsymbol{K}}_s, v)}{p(\boldsymbol{\kappa}_s^{old}|\star) f_T(\boldsymbol{\kappa}_s^{new}|\hat{\boldsymbol{\kappa}_s}, \hat{\boldsymbol{K}}_s, v)}, 1 \right).$$

3) Use the one-to-one map: $\gamma_{s,l_s} = \frac{\gamma_{s,l_s-1} + \exp(\kappa_{s,l_s})}{1+\exp(\kappa_{s,l_s})}, 2 < l_s < L_s - 2$ to obtain the cutpoints $\boldsymbol{\gamma}_s = (\gamma_{s,2}, \ldots, \gamma_{s,L_s-2})'$.

B.2 Appendix Tables

Table 4.1 gives partial simulation results of posterior means and true values as well as MSE of $\boldsymbol{\beta}$, $\boldsymbol{\Lambda}$ and ρ. Here, Tables B.1, B.2, and B.3 provide the complete results of 27 simulation experiments of the DMSOP model.

Table B.1: Simulation results in the case of $\boldsymbol{V} = (0.1, 0.05, 0.05, 0.1)$

Parm	True	simulation results								
	$\boldsymbol{V}$	(0.1,0.05,0.05,0.1)								
	$\boldsymbol{\Lambda}$	(0.15,0.03,0.015,0.12)			(0.3,0.06,0.03,0.24)			(0.45,0.09,0.045,0.36)		
	ρ	0.3	0.5	0.7	0.3	0.5	0.7	0.3	0.5	0.7
β_{11}	0.25	0.2474	0.2506	0.2545	0.2435	0.2416	0.2348	0.2485	0.2654	0.2576
β_{12}	0.51	0.5063	0.5019	0.5083	0.5135	0.5087	0.5179	0.5066	0.5106	0.5032
β_{21}	0.33	0.3318	0.3269	0.3107	0.3315	0.3284	0.3429	0.3316	0.3307	0.3273
β_{22}	0.46	0.4428	0.4469	0.4470	0.4569	0.4452	0.4355	0.4510	0.4309	0.4269
λ_{11}	/	0.1639	0.1471	0.1587	0.3043	0.3107	0.3055	0.4453	0.4544	0.4405
λ_{12}	/	0.0436	0.0622	0.0589	0.0669	0.0613	0.0600	0.0982	0.0972	0.1173
λ_{21}	/	0.0165	0.0070	-0.0053	0.0110	0.0081	0.0049	0.0162	0.0181	0.0150
λ_{22}	/	0.1484	0.1496	0.1579	0.2798	0.2852	0.2817	0.4082	0.3984	0.4077
ρ	/	0.3125	0.5056	0.7219	0.3038	0.5079	0.7098	0.3577	0.5949	0.7258
σ_1	0.02	0.0104	0.0108	0.0104	0.0104	0.0104	0.0104	0.0101	0.0101	0.0103
σ_2	0.01	0.0107	0.0104	0.0101	0.0096	0.0097	0.0101	0.0093	0.0095	0.0100
v_{11}	/	0.0983	0.0980	0.0970	0.0988	0.0975	0.0999	0.0963	0.0970	0.1000
v_{12}/v_{21}	/	0.0488	0.0487	0.0479	0.0497	0.0490	0.0502	0.0488	0.0499	0.0494
v_{22}	/	0.1002	0.0989	0.0989	0.1023	0.1025	0.1057	0.1023	0.1046	0.1027
$\gamma_{1,2}$	0.31	0.3147	0.3140	0.3091	0.3046	0.3092	0.3089	0.2941	0.2784	0.2882
$\gamma_{1,3}$	0.66	0.6736	0.6716	0.6624	0.6709	0.6744	0.6758	0.6699	0.6619	0.6644
$\gamma_{2,2}$	0.33	0.3285	0.3295	0.3226	0.3269	0.3292	0.3195	0.3343	0.3297	0.3132
$\gamma_{2,3}$	0.68	0.6797	0.6742	0.6725	0.6784	0.6812	0.6806	0.6901	0.6879	0.6658
β_{110}	0.47	0.4272	0.4447	0.4296	0.4213	0.4357	0.4114	0.4135	0.4414	0.4201
β_{120}	0.5	0.5160	0.5018	0.5140	0.5111	0.4971	0.5047	0.5124	0.5068	0.5109
β_{210}	0.49	0.5191	0.5171	0.4887	0.5237	0.5251	0.5028	0.5243	0.5148	0.4892
β_{220}	0.53	0.4777	0.4812	0.4793	0.4821	0.4892	0.4882	0.4879	0.4799	0.4736
$MSE_{\boldsymbol{\beta}}$	/	0.0020	0.0022	0.0024	0.0019	0.0023	0.0031	0.0020	0.0031	0.0034
$MSE_{\boldsymbol{\Lambda}}$	/	0.0016	0.0024	0.0032	0.0024	0.0030	0.0029	0.0036	0.0026	0.0044
MSE_{ρ}	/	0.0467	0.0285	0.0113	0.0432	0.0265	0.0121	0.0405	0.0262	0.0112

Table B.2: Simulation results in the case of $\boldsymbol{V} = (0.5, 0.4, 0.4, 0.5)$

Parm	True	simulation results								
	$\boldsymbol{V}$	(0.5,0.4,0.4,0.5)								
	$\boldsymbol{\Lambda}$	(0.15,0.03,0.015,0.12)			(0.3,0.06,0.03,0.24)			(0.45,0.09,0.045,0.36)		
	ρ	0.3	0.5	0.7	0.3	0.5	0.7	0.3	0.5	0.7
β_{11}	0.25	0.2465	0.2341	0.2475	0.2489	0.2501	0.2740	0.2475	0.2579	0.2783
β_{12}	0.51	0.5015	0.5000	0.5070	0.4886	0.4827	0.4932	0.4788	0.4887	0.4942
β_{21}	0.33	0.3736	0.3680	0.3518	0.3453	0.3541	0.3707	0.3372	0.3231	0.3366
β_{22}	0.46	0.4395	0.4420	0.4635	0.4470	0.4364	0.4473	0.4434	0.4344	0.4376
λ_{11}	/	0.1250	0.1117	0.1212	0.2527	0.2527	0.2630	0.4317	0.4359	0.4160
λ_{12}	/	0.0707	0.0805	0.0764	0.1204	0.1235	0.1169	0.1208	0.1275	0.1420
λ_{21}	/	-0.0109	-0.0107	-0.0194	0.0015	-0.0093	0.0034	0.0521	0.0522	0.0356
λ_{22}	/	0.1662	0.1682	0.1740	0.2984	0.3087	0.2829	0.3902	0.3915	0.3971
ρ	/	0.1581	0.3476	0.6668	0.2237	0.3885	0.6779	0.2492	0.5269	0.7169
σ_1	0.02	0.0141	0.0163	0.0150	0.0144	0.0149	0.0148	0.0130	0.0121	0.0130
σ_2	0.01	0.0093	0.0095	0.0087	0.0080	0.0077	0.0073	0.0082	0.0093	0.0079
v_{11}	/	0.4798	0.4695	0.4770	0.4649	0.4681	0.4762	0.4660	0.4581	0.4660
v_{12}/v_{21}	/	0.3799	0.3735	0.3704	0.3630	0.3632	0.3681	0.3646	0.3582	0.3632
v_{22}	/	0.4764	0.4722	0.4576	0.4663	0.4658	0.4716	0.4745	0.4690	0.4710
$\gamma_{1,2}$	0.31	0.3228	0.3180	0.3127	0.3043	0.3124	0.3127	0.3125	0.3133	0.3213
$\gamma_{1,3}$	0.66	0.6884	0.6853	0.6762	0.6666	0.6833	0.6884	0.6503	0.6591	0.6556
$\gamma_{2,2}$	0.33	0.3408	0.3421	0.3494	0.3412	0.3468	0.3398	0.3314	0.3430	0.3349
$\gamma_{2,3}$	0.68	0.6648	0.6624	0.6651	0.6682	0.6693	0.6840	0.6823	0.6843	0.6786
β_{110}	0.47	0.4325	0.4390	0.4788	0.4340	0.4434	0.4801	0.4301	0.4513	0.4833
β_{120}	0.5	0.4590	0.4548	0.4432	0.4509	0.4516	0.4564	0.4457	0.4638	0.4627
β_{210}	0.49	0.4990	0.4581	0.4530	0.5064	0.4542	0.4596	0.5018	0.4505	0.4575
β_{220}	0.53	0.4945	0.5293	0.5252	0.4836	0.5295	0.5258	0.4849	0.5101	0.5061
$MSE_{\boldsymbol{\beta}}$	/	0.0074	0.0069	0.0060	0.0057	0.0072	0.0087	0.0068	0.0077	0.0081
$MSE_{\boldsymbol{\Lambda}}$	/	0.0055	0.0074	0.0076	0.0106	0.0130	0.0077	0.0027	0.0031	0.0058
MSE_{ρ}	/	0.0928	0.0834	0.0193	0.0749	0.0650	0.0185	0.0620	0.0471	0.0147

Table B.3: Simulation results in the case of $\boldsymbol{V} = (1, 0.9, 0.9, 1)$

Parm	True	simulation results								
	$\boldsymbol{V}$	(1,0.9,0.9,1)								
	$\boldsymbol{\Lambda}$	(0.15,0.03,0.015,0.12)			(0.3,0.06,0.03,0.24)			(0.45,0.09,0.045,0.36)		
	ρ	0.3	0.5	0.7	0.3	0.5	0.7	0.3	0.5	0.7
β_{11}	0.25	0.2967	0.2798	0.2874	0.2854	0.2715	0.2670	0.2260	0.2349	0.2687
β_{12}	0.51	0.4770	0.4719	0.4623	0.5208	0.4692	0.4581	0.4474	0.4233	0.4152
β_{21}	0.33	0.3847	0.3880	0.4020	0.3028	0.3916	0.3888	0.3731	0.3694	0.3813
β_{22}	0.46	0.4426	0.4607	0.4759	0.5295	0.4924	0.4937	0.4302	0.4496	0.4521
λ_{11}	/	0.1044	0.0813	0.1135	0.3164	0.2948	0.2750	0.4568	0.4742	0.4517
λ_{12}	/	0.0687	0.1003	0.0782	0.0181	0.0657	0.0897	0.0941	0.0765	0.0943
λ_{21}	/	-0.0150	-0.0586	-0.0466	0.0031	0.0078	-0.0144	0.0602	0.0726	0.0354
λ_{22}	/	0.1541	0.2005	0.1908	0.2507	0.2699	0.3030	0.3608	0.3534	0.3968
ρ	/	0.3275	0.5671	0.6924	0.3542	0.5341	0.6342	0.1564	0.2977	0.5662
σ_1	0.02	0.0173	0.0194	0.0175	0.0027	0.0165	0.0187	0.0136	0.0130	0.0144
σ_2	0.01	0.0056	0.0000	0.0015	0.0194	0.0018	0.0027	0.0090	0.0097	0.0086
v_{11}	/	0.9344	0.9444	0.9339	0.9578	0.8993	0.9261	0.9156	0.8976	0.8766
v_{12}/v_{21}	/	0.8521	0.8610	0.8730	0.9009	0.8375	0.8419	0.8162	0.8111	0.8035
v_{22}	/	0.9663	0.9770	1.0141	1.0400	0.9764	0.9607	0.9331	0.9384	0.9312
$\gamma_{1,2}$	0.31	0.3089	0.2970	0.3109	0.3230	0.2986	0.2992	0.2906	0.2890	0.2920
$\gamma_{1,3}$	0.66	0.6645	0.6604	0.6644	0.6547	0.6611	0.6631	0.6491	0.6444	0.6493
$\gamma_{2,2}$	0.33	0.3310	0.3423	0.3412	0.3420	0.3595	0.3478	0.3391	0.3365	0.3295
$\gamma_{2,3}$	0.68	0.6746	0.6872	0.6918	0.7027	0.6875	0.6732	0.6807	0.6836	0.6807
β_{110}	0.47	0.4316	0.3862	0.3940	0.4638	0.3647	0.3644	0.3965	0.3636	0.3706
β_{120}	0.5	0.4315	0.4598	0.4777	0.4555	0.4560	0.4739	0.4048	0.4142	0.4394
β_{210}	0.49	0.5128	0.5729	0.5735	0.5553	0.5899	0.5741	0.5153	0.5702	0.5420
β_{220}	0.53	0.4244	0.4173	0.4246	0.4490	0.4089	0.4232	0.4405	0.4087	0.4236
$MSE_{\boldsymbol{\beta}}$	/	0.0129	0.0121	0.0165	0.0137	0.0135	0.0155	0.0144	0.0169	0.0203
$MSE_{\boldsymbol{\Lambda}}$	/	0.0068	0.0220	0.0135	0.0046	0.0025	0.0087	0.0009	0.0024	0.0023
MSE_{ρ}	/	0.0713	0.0513	0.0263	0.0713	0.0410	0.0355	0.1092	0.1077	0.0555

References

Albert, J. H. and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association 88*(422), 669–679.

Albert, J. H. and S. Chib (2001). Sequential ordinal modeling with applications to survival data. *Biometrics 57*(3), 829–836.

Anselin, L. (2013). *Spatial Econometrics: Methods and Models*. Springer Science & Business Media.

Arbia, G. (2006). *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*. Springer Science & Business Media.

Baltagi, B. H., P. H. Egger, and M. Kesina (2016). Bayesian spatial bivariate panel probit estimation. In *Spatial Econometrics: Qualitative and Limited Dependent Variables*, pp. 119–144. Emerald Group Publishing Limited.

Basile, R., L. Benfratello, and D. Castellani (2013). Geoadditive models for regional count data: An application to industrial location. *Geographical Analysis 45*(1), 28–48.

Beron, K. J. and W. P. Vijverberg (2004). Probit in a spatial context: A Monte Carlo analysis. In *Advances in Spatial Econometrics*, pp. 169–195. Springer.

Chen, J. (2005). The prevalence study on main chronic diseases and the associated behavioral risk factors among the middle and old aged people in some areas of Hubei province. Master's thesis, Wuhan University. (In Chinese).

Chen, M. H. and D. K. Dey (2000). Bayesian analysis for correlated ordinal data models. In *Generalized Linear Models: A Bayesian Perspective*, pp. 133–157. New York: Marcel Dekker Inc.

Chib, S. and E. Greenberg (1998). Analysis of multivariate probit models. *Biometrika 85*(2), 347–361.

Cowles, M. K. (1996). Accelerating Monte Carlo Markov chain convergence for cumulative-link generalized linear models. *Statistics and Computing 6*(2), 101–111.

Cressie, N. (2015). *Statistics for Spatial Data*. New York: John Wiley & Sons.

Damián, J., R. Pastor-Barriuso, and E. Valderrama-Gama (2008). Factors associated with self-rated health in older people living in institutions. *BMC Geriatrics 8*(1), 5.

Eilers, P. H. and B. D. Marx (1996). Flexible smoothing with B-splines and penalties. *Statistical Science 11*(2), 89–102.

Enkvist, Å., H. Ekström, and S. Elmståhl (2012). What factors affect life satisfaction (LS) among the oldest-old? *Archives of Gerontology and Geriatrics 54*(1), 140–145.

Fahrmeir, L., T. Kneib, and S. Lang (2004). Penalized additive regression for space-time data: A Bayesian perspective. *Statistica Sinica 14*, 731–761.

Fahrmeir, L., T. Kneib, S. Lang, and B. Marx (2013). *Regression: Models, Methods and Applications*. Springer Science & Business Media.

Fernández-Ballesteros, R., M. D. Zamarrón, and M. Á. Ruiz (2001). The contribution of socio-demographic and psychosocial factors to life satisfaction. *Ageing and Society 21*(1), 25–43.

Fillingim, R. B., C. D. King, M. C. Ribeiro-Dasilva, B. Rahim-Williams, and J. L. Riley (2009). Sex, gender, and pain: A review of recent clinical and experimental findings. *The Journal of Pain 10*(5), 447 – 485.

Fillmore, K. M., W. C. Kerr, T. Stockwell, T. Chikritzhs, and A. Bostrom (2006). Moderate alcohol use and reduced mortality risk: Systematic error in prospective studies. *Addiction Research and Theory 14*(2), 101–132.

Ford, J., M. Spallek, and A. Dobson (2007). Self-rated health and a healthy lifestyle are the most important predictors of survival in elderly women. *Age and Ageing 37*(2), 194–200.

Freedman, V. A. and L. G. Martin (2000). Contribution of chronic conditions to aggregate changes in old-age functioning. *American Journal of Public Health 90*(11), 1755–1760.

Gao, P. and H. Hasegawa (2018). Bayesian spatial analysis of chronic diseases in elderly Chinese people using a STAR model. *Statistics in Transition new series 19*(4), 645–670.

Gao, P. and H. D. Li (2016). New characteristics of active life expectancy of the elderly in China. *Advances in Aging Research 5*(1), 27–39.

Gelman, A., H. S. Stern, J. B. Carlin, D. B. Dunson, A. Vehtari, and D. B. Rubin (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC.

Geniaux, G. and C. Napoléone (2008). Semi-parametric tools for spatial hedonic models: An introduction to mixed geographically weighted regression and geoadditive models. In *Hedonic Methods in Housing Markets*, pp. 101–127. Springer.

Hasegawa, H. (2009). Bayesian dynamic panel-ordered probit model and its application to subjective well-being. *Communications in Statistics-Simulation and Computation 38*(6), 1321–1347.

Hasegawa, H. (2010). Analyzing tourists' satisfaction: A multivariate ordered probit approach. *Tourism Management 31*(1), 86–97.

He, J., D. Gu, X. Wu, K. Reynolds, X. Duan, C. Yao, J. Wang, C.-S. Chen, J. Chen, R. P. Wildman, M. J. Klag, and P. K. Whelton (2005). Major causes of death among men and women in China. *The New England Journal of Medicine 353*(11), 1124–1134.

Ho, S., J. Woo, J. Lau, S. Chan, Y. Yuen, Y. Chan, and I. Chi (1995). Life satisfaction and associated factors in older Hong Kong Chinese. *Journal of the American Geriatrics Society 43*(3), 252–255.

Holloway, G., B. Shankar, and S. Rahmanb (2002). Bayesian spatial probit estimation: A primer and an application to HYV rice adoption. *Agricultural Economics 27*(3), 383–402.

Jeliazkov, I., J. Graves, and M. Kutzbach (2008). Fitting and comparison of models for multivariate ordinal outcomes. In *Bayesian Econometrics*, pp. 115–156. Emerald Group Publishing Limited.

Jiao, S., X. Yin, Y. Wang, J. Xie, and B. Guo (2002). Analysis on behaviors risk factors of chronic disease in residents of Beijing and study its countermeasure. *China Public Health 18*(2), 197–198. (In Chinese).

Kakamu, K. and H. Wago (2007). Bayesian spatial panel probit model with an application to business cycle in Japan. Technical report, Citeseer.

Kakamu, K., H. Wago, and H. Tanizaki (2010). Estimation of regional business cycle in Japan using Bayesian panel spatial autoregressive probit model. *Handbook of Regional Economics*, 555–571.

Kammann, E. and M. P. Wand (2003). Geoadditive models. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 52*(1), 1–18.

Kiecolt-Glaser, J. K. and T. L. Newton (2001). Marriage and health: His and hers. *Psychological Bulletin 127*(4), 472–503.

Kneib, T. (2006). *Mixed model based inference in structured additive regression.* Ph. D. thesis, Ludwig-Maximilians-Universität München.

Lang, S. and A. Brezger (2004). Bayesian P-splines. *Journal of Computational and Graphical Statistics 13*(1), 183–212.

Lauer, C. (2003). Family background, cohort and education: A French–German comparison based on a multivariate ordered probit model of educational attainment. *Labour Economics 10*(2), 231–251.

LeSage, J. P. (1999). The Theory and Practice of Spatial Econometrics. *http://www. spatial-econometrics. com/html/sbook. pdf*.

LeSage, J. P. and O. Parent (2007). Bayesian model averaging for spatial econometric models. *Geographical Analysis 39*(3), 241–267.

Li, Z. K., M. R. Zhang, Z. Yang, L. Shen, Y. Ma, H. Q. Gong, R. Tian, and F. Yan (2012). The epidemiological characteristics and PYLL analysis of death of chronic diseases among residents in Kunming city in 2007-2010. *Chinese Primary Health Care 26*(3), 84–86. (In Chinese).

Lima-Costa, M. F., J. O. Firmo, and E. Uchôa (2004). The structure of self-rated health among older adults: the Bambuí Health and Ageing Study (BHAS). *Revista De Saude Publica 38*(6), 827–834.

McMillen, D. P. (1992). Probit with spatial autocorrelation. *Journal of Regional Science 32*(3), 335–348.

Melin, R., K. S. Fugl-Meyer, and A. R. Fugl-Meyer (2003). Life satisfaction in 18-to 64-year-old Swedes: In relation to education, employment situation, health and physical activity. *Journal of Rehabilitation Medicine 35*(2), 84–90.

Palmore, E. and C. Luikart (1972). Health and social factors related to life satisfaction. *Journal of Health and Social Behavior 13*(1), 68–80.

Rodgers, V., S. Neville, and S. La Grow (2017). Health, functional ability and life satisfaction among older people 65 years and over: A cross-sectional study. *Contemporary Nurse 53*(3), 284–292.

Sauleau, E. A., A. Hennerfeind, A. Buemi, and L. Held (2007). Age, period and cohort effects in Bayesian smoothing of spatial cancer survival with geoadditive models. *Statistics in Medicine 26*(1), 212–229.

Smith, T. E. and J. P. LeSage (2004). A Bayesian probit model with spatial dependencies. In *Spatial and Spatiotemporal Econometrics*, pp. 127–160. Emerald Group Publishing Limited.

Thorpe, K. E. and D. H. Howard (2006). The rise in spending among medicare beneficiaries: The role of chronic disease prevalence and changes in treatment intensity. *Health Affairs 25*(5), w378–w388.

Umlauf, N., D. Adler, T. Kneib, S. Lang, and A. Zeileis (2015). Structured additive regression models: An R interface to BayesX. *Journal of Statistical Software 63*(21), 1–46.

Verburg, P. H., J. R. R. van Eck, T. C. de Nijs, M. J. Dijst, and P. Schot (2004). Determinants of land-use change patterns in the Netherlands. *Environment and Planning B: Planning and Design 31*(1), 125–150.

Vogeli, C., A. E. Shields, T. A. Lee, T. B. Gibson, W. D. Marder, K. B. Weiss, and D. Blumenthal (2007). Multiple chronic conditions: Prevalence, health consequences, and implications for quality, care management and costs. *Journal of General Internal Medicine 22*(3), 391–395.

Wand, H., C. Whitaker, and G. Ramjee (2011). Geoadditive models to assess spatial variation of HIV infections among women in local communities of Durban, South Africa. *International Journal of Health Geographics 10:28*(1), 1–9.

Wang, H., E. M. Iglesias, and J. M. Wooldridge (2013). Partial maximum likelihood estimation of spatial probit models. *Journal of Econometrics 172*(1), 77–89.

Wang, X. and K. M. Kockelman (2009a). Application of dynamic spatial ordered probit model: Patterns of ozone concentration in Austin, Texas. *Transportation Research Record 2136*(1), 45–56.

Wang, X. and K. M. Kockelman (2009b). Application of the dynamic spatial ordered probit model: Patterns of land development change in Austin, Texas. *Papers in Regional Science 88*(2), 345–365.

Wang, X. and K. M. Kockelman (2009c). Bayesian inference for ordered response data wit a dynamic spatial-ordered probit model. *Journal of Regional Science 49*(5), 877–913.

WHO (2005). Preventing chronic diseases: A vital investment. Technical report, World Health Organization.

Wickrama, K., F. O. Lorenz, R. D. Conger, and G. H. Elder Jr (1997). Marital quality and physical illness: A latent growth curve analysis. *Journal of Marriage and the Family 59*, 143–155.

Wolff, J. L., B. Starfield, and G. Anderson (2002). Prevalence, expenditures and complications of multiple chronic conditions in the elderly. *Archives of Internal Medicine 162*, 2269–2276.

Woolf, S. H., L. Aron, L. Dubay, S. M. Simon, E. Zimmerman, and K. X. Luk (2015). How are income and wealth linked to health and longevity? *Urban Institute*.

Ye, S. (2013). Analysis on non-communicable chronic diseases status and its influencing factors among adults in Gong Zhuling of Jilin province in 2012. Master's thesis, Jilin University. (In Chinese).

Zhao, J., R. Wan, Z. Liu, Q. Wan, Q. Zhang, and Y. Ruan (2015). Prevalence and influencing factors of chronic diseases in residents aged 15 years and above in urban and rural area of Yunnan province, 2010-2012. *Chinese Journal of Public Health 31*(7), 878–882. (In Chinese).

Zhen, S. (2010). Study of the current situation and influencing factors of chronic illnesses in Gansu. Master's thesis, Lanzhou University. (In Chinese).

Zhou, B. and K. M. Kockelman (2008). Neighborhood impacts on land use change: A multinomial logit model of spatial relationships. *The Annals of Regional Science 42*(2), 321–340.