Modelling spatial weight matrices and lags in spatial panel models

Tan, Chang

DOI:
10.33612/diss.630685888

Modelling Spatial Weight Matrices and Lags in Spatial Panel Models

Chang Tan

Publisher: University of Groningen, Groningen, The Netherlands

Printed by: Ipskamp Printing
P.O. Box 333
7500 AH Enschede
The Netherlands

university of groningen

Modelling spatial weight matrices and lags in spatial panel models

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. C. Wijmenga
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Thursday 28 September 2023 at 14.30 hours

by

Chang Tan

born on 1 April 1992
in Liaoning, China

Supervisor
Prof. J.P. Elhorst

Co-supervisor
Dr. M. Kesina

Assessment Committee
Prof. R.J.M. Alessie
Prof. L.F. Lee
Prof. F. van Oort

Acknowledgements

The journey to a PhD has been the best experience in my life ever.

I am so lucky that I could have the best supervisors in the world: Paul Elhorst and Michaela Kesina. All your patience, kindness, and conscientiousness helped me grow and moved me along. It was you who let me know the happiness of doing research and encouraged me not to give up. Besides research, you also helped me prepare for the job market interviews, listened to my nonsense, and tolerated my sloppiness. Whenever I had questions, you were always there.

I would like to extend my appreciation to the members of the reading committee, Lung-Fei Lee, Rob Alessie, and Frank van Oort. Thank you for reading this manuscript and valuable comments.

I am deeply grateful to Arthur de Boer, Max Groneck, Tristan Kohl, Helena Lei, Ellen Nienhuis, Kristian Peters, Hanneke Tamling, who helped me solve all kinds of problems in daily life. Without your help, I could not survive from the tough weather in the Netherlands. Special thanks go to my friends Ioanna Tziolas, Cheng Chen, Wout de Vries, Ruchen Hu, Huan Liu, Zhaoxin Liu, Chenming Peng, Ziwei Rao, Lennard Rekker, Nannette Stoffers, Lotte Swank, Mark van Oldeniel, Jasmira Wiersma, Suqing Wu, Jiahua Zong, and Xiying Zhang; thank you for your generous support.

I truly appreciate my parents. You encouraged me to chase my dream, spoke words of wisdom, and let me choose my way of life. No words can fully express the gratitude to my love, Xin Jiang, thank you for your support, understanding, and care when I got lost. You feel like home. You are my home.

A new journey is about to start. Let's keep going!

Chang Tan,
Groningen, 31 January 2023

Contents

Chapter 1

Introduction

1.1 General introduction

"Everything in the world is connected with something else" and "Every qualitatively defined system has a special type of interaction" (Spirkin, 1983, p. 84). In social science, social and economic interactions play an indispensable role in the spread of diseases, the personal employment options, the trade flows of many goods and services, etc. Spatial econometric models provide possibilities to explain the spatial, economic, and social interactions among individual units in a system. Three different types of interaction effects are considered in the spatial econometric literature: endogenous interaction effects among the dependent variable ($\boldsymbol{y}_t$), exogenous interaction effects among the regressors ($\boldsymbol{X}_t$), and interaction effects among the error term ($\boldsymbol{v}_t$) for individual units at time t where $t = 1, \ldots, T$. Correspondingly, a spatial econometric model may contain one spatial lag in the dependent variable ($\boldsymbol{W}\boldsymbol{y}_t$), K spatial lags in the regressors ($\boldsymbol{W}\boldsymbol{X}_t$), and one in the error term, where $\boldsymbol{W}$ is the $N \times N$ spatial weight matrix capturing the connection between units, which is an essential part of each spatial econometric model.

Spatial interactions, which is the main focus of spatial econometrics, obviously depend on the specification of the spatial weight matrix $\boldsymbol{W}$. However, an often-criticized aspect of using spatial econometric models in empirical research is that $\boldsymbol{W}$ is specified in advance rather than being estimated along with the parameters in the model (Corrado and Fingleton, 2012). Question (I) in the spatial econometric literature is therefore the specification of the corresponding spatial weight

matrix. Issues about the specification of $\boldsymbol{W}$ are the following. First, whether the spatial weight matrix should take a sparse binary contiguity form or a dense distance decay form and whether the spatial weight matrix is pre-specified, selected according to posterior test statistics, or directly estimated. Second, how to normalize the spatial weight matrix. Common approaches are normalization by the row-sums, column-sums, the maximal or minimal row-sum, or the largest eigenvalue of $\boldsymbol{W}$. Among them, normalization by the row-sums or the largest eigenvalue are the most popular ones and they are also known as row normalization and scalar normalization, respectively. Row normalization can be achieved by dividing each element by the corresponding row-sum and this normalization ensures that the weighting operation can be interpreted as an averaging of the values of other individual units. Scalar normalization can be realized by dividing each element by the largest eigenvalue of $\boldsymbol{W}$ and this normalization ensures that economic interpretation in terms of distance decay is valid when the spatial weight matrix has a distance decay structure. Third, whether different spatial lags should share the same or different spatial weight matrices. In theory, each of the $K + 2$ spatial lags, if included, may have its own $\boldsymbol{W}$. The reality is that almost every empirical study and each Monte Carlo simulation experiment in an econometric-theoretical study investigating the finite sample properties of a newly proposed spatial econometric model estimator or test statistic, adopts one common specification of $\boldsymbol{W}$ for all spatial lags. One common pre-specified spatial weight matrix for all spatial lags may cause identification problems. We will discuss the identification problem later in the second following paragraph. The parameterization of the spatial weight matrix turns out to be a simple and effective solution.

Question (II) emerging from the spatial econometric literature is how to specify the spatial lags. Specifically, this question focuses on the specification of the spatial lag in the error term. This lag captures unobserved shocks following a spatial pattern other than those covered by time fixed effects and thus improves the efficiency of the estimator if accounted for. The representative model is the spatial error model (SEM) in the spatial autoregressive (AR) form, while econometric-theoretical and empirical research on models with spatial moving average (MA) errors have received limited attention. These two forms of spatial errors can reflect different types of shocks. Shocks captured by AR errors are global and those

captured by MA errors are local (Fingleton, 2008b). To distinguish global and local shocks, sparse weight matrices which contain many zero elements are generally taken as examples. Shocks in the AR form are global because they rebound and diffuse multiple rounds to all units even for the units which are not directly connected, while shocks in the MA form are local since they only diffuse one round to the units which are connected (Fingleton and LeGallo, 2007). The latter implies that for spatial weight matrices which are dense, the shocks spread likewise to all units when captured by the MA form. Therefore, the interpretations of the AR and MA errors should be considered together with the specification of the spatial weight matrices.

Question (III) standing in the spatial econometric literature is the identification problem. It needs to be noted that there are different identification problems in spatial econometrics. The one that we pay attention to in this thesis is that the parameters to be estimated of a model cannot be recovered from its reduced form. This identification problem has been discussed in section 3 of Gibbons and Overman (2012) and also mentioned as one of the identification problems by Halleck Vega and Elhorst (2015). The inclusion of $\boldsymbol{W}\boldsymbol{X}_t$, which are also known as Durbin term, has the advantage that it generates fully flexible indirect effects. The indirect effects are often the main interest of applied spatial econometric studies since they measure the marginal impact of changes in the explanatory variables of other units on the dependent variable of the focal unit. Attempts to identify the parameters of models with Durbin term have been made. Bramoullé et al. (2009) investigate the identification of the parameters of the linear-in-means model, which is also known as the spatial Durbin (SD) model. Lee, Liu, and Lin (2010) propose a specific network structure providing information on how units are connected to each other. This structure helps to solve the identification problem of various interaction effects in the general nesting spatial (GNS) model. Similarly, specific functional forms of the spatial weight matrix can help solving the identification problem. By imposing exponential or inverse distance decay functional forms of the spatial weight matrix, which exploits the geographical arrangement of the units in the sample, the distance decay parameter of the spatial weight matrix of each spatial lag can be estimated, thereby avoiding that this matrix is completely pre-specified and the same for all spatial lags. Using this simple parametric method, each regressor can

interact with a different spatial weight matrix and exhibit a different distance decay pattern. This extension helps to identify the unknown parameters.

In spatial panel models, the inclusion of fixed effects, for both individual units and time periods, is important. The individual fixed effects are often incorporated in the model to control for time invariant unobserved individual characteristics which can be correlated with the existing regressors. As for the time fixed effects, they can control for time-varying unobservables that affect all units in the sample. However, if the fixed effects are estimated jointly with the other parameters, we may encounter incidental parameter problems. Two transformations can be used to concentrate out these fixed effects. The within transformation uses $\boldsymbol{J}_N = (\boldsymbol{I}_N - \frac{1}{N}\boldsymbol{\iota}_N\boldsymbol{\iota}_N^{'})$ and $\boldsymbol{J}_T = (\boldsymbol{I}_T - \frac{1}{T}\boldsymbol{\iota}_T\boldsymbol{\iota}_T^{'})$ to eliminate both fixed effects where $\boldsymbol{\iota}_N$ and $\boldsymbol{\iota}_T$ are the $N \times 1$ and $T \times 1$ vectors of ones, respectively. The orthogonal transformation, on the other hand, uses $\boldsymbol{F}_{N,N-1}$ and $\boldsymbol{F}_{T,T-1}$ to eliminate the fixed effects, where $[\boldsymbol{F}_{N,N-1}, \frac{1}{\sqrt{N}}\boldsymbol{\iota}_N]$ and $[\boldsymbol{F}_{T,T-1}, \frac{1}{\sqrt{T}}\boldsymbol{\iota}_T]$ are the orthonormal eigenvector matrices of $\boldsymbol{J}_N$ and $\boldsymbol{J}_T$. The within and orthogonal transformation are called direct and transformation approach in Lee and Yu (2010a). To be consistent with the panel data literature, we adopt the nomenclature of within and orthogonal transformation. The transformed equations can be subsequently estimated by quasi-maximum likelihood (QML). Lee and Yu (2010a) prove the consistency and derive the asymptotic distribution of the QML estimators based on the within and orthogonal transformation for the SAR model with AR errors, which is also called the SARAR model. They also derive the estimation procedures for two panel data settings: individual fixed effects only and both types of fixed effects. In the former case, the variance parameter needs to be bias-corrected when T is finite and the QML estimator based on the within transformation is applied, while the QML estimator based on the orthogonal transformation is consistent and properly centred. For the latter case, the QML estimator based on the within transformation is consistent only when both N and T are large, whereas the QML estimator based on the orthogonal transformation is consistent when either N or T is large. Bias correction is needed to remove the non-centrality of the QML estimator based on the within approach. One limitation of the orthogonal transformation is that it requires $\boldsymbol{W}$ to be row-normalized when both fixed effects are included and concentrated out. The QML estimators based on the within and orthogonal transformation are labelled

the QML within and QML orthogonal estimator for simplicity.

The three studies in this dissertation contribute to the literature by providing solutions to the three questions mentioned above: (I) the specification of the spatial weight matrix of each spatial lag, (II) the specification of the spatial lag itself, and (III) the issue of identification. In Chapter 2 we depart from the spatial Durbin model (SDM), which contains spatial lags in both dependent variable and the regressors, and we allow the dependent variable and each regressor to interact with different spatial weight matrices by parameterizing them with a distance decay parameter. These distance decay parameters are estimated together with the response parameters of the model. This study extends the literature by answering Question (I), the specification of $\boldsymbol{W}$. In Chapter 3 we focus on the estimation of the spatial autoregressive (SAR) model with spatial MA errors, which is also known as the SARMA model, with both fixed effects. Different non-parameterized spatial weight matrices are allowed for the spatial lags in the dependent variable and the errors. Possible solutions to Question (II), the specification of the spatial lag, are provided in this study. In the last study of Chapter 4, we consider the estimation of the general nesting spatial (GNS) model with AR or MA errors, which are known as GNSAR and GNSMA model, respectively. These two models contain all three types of spatial lags. The parameterized spatial weight matrices that are different for each spatial lag under various model settings are also considered in this chapter. As an extension of the previous two studies, this study not only answers the first two questions but it also provides a solution to Question (III), the identification problem. An overview of the content of each chapter is given in Table 1.1.

Table 1.1. Description of the Chapters in the Thesis

Chapter	**Model**	**Fixed Effects**	**Setting of N and T**	**Parameterization of W**	**Normalization of W**	**Method of Concentrating out Fixed Effect**
Chapter 2	SDM	Both	Large N, finite T	Yes	Row normalization Scalar normalization	Within transformation
Chapter 3	SARMA	Both	Large N, finite T	No	Row normalization Scalar normalization	Within transformation
Chapter 4	GNSAR GNSMA	Both	Large N, large T.	Yes	Row normalization	Orthogonal transformation

1.2 Outline of the thesis

This dissertation is organized as follows. Chapter 2 provides the QML within estimator of the SDM with parameterized spatial weight matrices. Chapter 3 focuses on the QML within estimator of the panel SARMA model with pre-specified though different spatial weight matrices. Chapter 4 sets out the GNSAR and GNSMA models and derives the QML orthogonal estimator of the parameters when the spatial weight matrices are parameterized and it discusses the identification problem. Chapter 5 concludes.

We start with the prominent spatial Durbin model (SDM), which contains spatial lags in the dependent variable and the regressors. In Chapter 2, we estimate the common parameters and the distance decay parameters which are imposed for each spatial weight matrix within the model setting. Although each variable can have its own spatial weight matrix, practitioners generally adopt one common pre-specified spatial weight matrix for all of them. This chapter breaks this practice by parameterizing each spatial weight matrix with a distance decay parameter. We propose a maximum likelihood approach that estimates the decay parameters along with the other parameters in the model. We consider negative exponential and inverse distance decay matrices, and two types of normalization, row normalization and normalization of the weight matrix by its largest eigenvalue. We investigate the performance of the proposed estimation approach in a Monte Carlo simulation experiment and present the results of an empirical application on military expenditures. An important finding is that the indirect effects, which are important for many empirical studies, can be severely biased when one common pre-specified spatial weight matrix is used.

Chapter 3 advocates the wider use of the spatial autoregressive (AR) model with spatial moving average (MA) errors, individual and time fixed effects, and different spatial weight matrices for each spatial lag. We derive and investigate the asymptotic properties of the QML within estimator of the model when N is large and T is finite. We also argue and demonstrate that the interpretation of the spatial MA errors in various empirical applications is more obvious than the spatial AR errors since most shocks are local rather than global. A Monte Carlo simulation and an empirical example based on military expenditure data are used to illustrate

this. The findings also indicate that each spatial lag should interact with different spatial weight matrices.

Chapter 4 sets out GNS models with individual and time fixed effects, a spatial AR or spatial MA error process, and distance-based parameterized spatial weight matrices that are different for each spatial lag in the model. The QML orthogonal estimator for these two models is developed based on the QML within estimators in Chapters 2 and 3. The asymptotic distribution of the proposed estimator is provided and its finite sample properties are investigated by means of Monte Carlo simulation. To illustrate the benefits of the proposed models in an empirical setting, the Yang (2021) US house price data set is used.

The final chapter summarizes the main findings, provides discussions, and mentions directions for future research.

Chapter 2

Parameterizing spatial weight matrices in the spatial Durbin panel data model*

* This chapter was written together with M. Kesina and J.P. Elhorst.

2.1 Introduction

In a spatial econometric model the behaviour of one cross-sectional unit is co-determined by spatial lags observed on other cross-sectional units. Spatial lags can be taken with respect to the dependent variable, the explanatory variables, and/or the error term, known as the spatially lagged dependent variable, spatially lagged explanatory variables, and the spatially lagged error term, respectively. The degree of co-determination further depends on the mutual relationships among the cross-sectional units formalized by the spatial weight matrix. In the literature, there exists a variety of spatial econometric models that contain one or multiple forms of spatial lags (for an overview see Elhorst, 2014).

One prominent model is the spatial Durbin model (SDM) that contains a spatial lag in the dependent variable and each of the explanatory variables, enabling estimation of endogenous interaction effects and contextual effects. This model is advocated by LeSage and Pace (2009), and has been widely applied to study different topics in economics, e.g. energy economics (Feng and Chen, 2018; Li and Li, 2020; Chen et al., 2021, to mention a few), innovations (Tientao et al., 2016; Pan et al., 2021; Xiao and Mao, 2021, to mention a few), regional and growth economics (Beer and Riedl, 2012; Koroglu and Sun, 2016; Sabater and Graham, 2019, to mention a few). Furthermore, it is widely used in the peer effects literature, but then under the label linear-in-means model (Bramoullé et al., 2009, 2020; De Paula et al., 2020; among others).

One advantage of the SDM in empirical research is its flexibility in modelling spillovers. The main interest of spatial econometric practitioners are not the point estimates of the parameters but the marginal impact of changes that explanatory variables have on the dependent variable. Two marginal effects are of interest: the direct effect of changing the explanatory variable of one unit on the dependent variable of that unit itself, and the cumulative effect of changing the explanatory variable of one unit on the dependent variable of other units (LeSage and Pace, 2009). This cumulative effect is known as the indirect effect, the spatial spillover effect, the indirect spillover effect or the spillover effect. Elhorst (2010) and Halleck Vega and Elhorst (2015) demonstrate that only models that at least include spatially lagged regressors are able to produce spillover effects that can take any empirical value.

In contrast, models that include a spatially lagged dependent variable and/or a spatially lagged error only are less flexible since they impose restrictions on the magnitude of spillover effects in advance.

Although the econometric theoretical literature distinguishes different spatial weight matrices in models with multiple types of spatial lags, empirical applications and Monte Carlo simulations tend to use one common specification of the spatial weight matrix for all spatial lags. Especially in the SDM the same spatial weight matrix is used for all regressors. This is a very restrictive assumption as the effect and the spatial range of spatial lags are likely to be different for different regressors. One potential solution is to pre-specify different spatial weight matrices for the spatially lagged dependent variable and each of the spatially lagged regressors, estimate the model for different combinations of spatial weight matrices, and then select the combination that produces the highest log-likelihood function value (Leenders, 2002), the highest Bayesian posterior model probability (LeSage and Pace, 2009, Ch.6), or the best spatial J-test result (Kelejian, 2008). However, this approach poses several challenges and thus is hardly used. First, the empirical researcher does not know ex-ante which spatial weight matrix to use for the different types of spatial lags. Second, when multiple spatial weight matrices are selected, this procedure becomes demanding and time-consuming. Conversely, if the true spatial weight matrix is not in the choice set because this set is kept small, the researcher runs the risk of picking a wrong spatial weight matrix.

Another solution is to estimate the spatial weight matrix along with the other parameters in the model (Ahrens and Bhattacharjee, 2015; Lam and Souza, 2019; De Paula et al., 2020; Lewbel et al., 2021, to mention a few). However, these studies either focus on one form of spatial lag only, usually the spatially lagged dependent variable (the first two cited studies), or use the SDM with one common spatial weight matrix (the last two cited studies). Furthermore, many studies are based on the assumption that the number of time units (T) exceeds the number of cross-sectional units (N), whereas the majority of empirical studies is based on $N >> T$. Some approaches limit the number of unknown elements of the spatial weight matrix, which leads to a sparse spatial weight matrix.

Our study contributes to the literature in the following ways. First, we let the spatial lag of each regressor have its own spatial weight matrix to allow for differ-

ent spatial ranges. Furthermore, we parameterize each spatial weight matrix with a different distance decay parameter and estimate those together with the other parameters of the model. We consider two frequently used functional forms of distance decay, the negative exponential and the inverse distance form, and two types of normalization, row normalization and normalization of the weight matrix by its largest eigenvalue.[1]

The proposed maximum likelihood (ML) estimator extends previous work of Lee and Yu (2010a) by allowing for a parameterization of all spatial weight matrices in the model without imposing sparsity in advance or requiring a large time dimension. Our approach is applicable to the common large N and finite T framework.

We also provide mathematical expressions for the direct and indirect spillover effect. We conduct Monte Carlo experiments and find that our approach works well in estimating the parameters and direct and indirect spillover effects. Especially the latter are more sensitive to the choice of the spatial weight matrix.

A Matlab function of the proposed ML estimator will be made available to give other researchers the opportunity to apply it in their own empirical setting. An empirical application on military expenditures from Yesilyurt and Elhorst (2017) is utilized to show the usefulness of the proposed parameterization approach for applied researchers. We find empirical evidence in favor of both the spatial Durbin model and spillover effects, and show that this is because we do not adopt the same pre-specified spatial weight matrix for all regressors in the model.

The remainder of this chapter is organized as follows. In Section 2.2, we review previous studies that parameterized the spatial weight matrix or suggested alternative estimation approaches. In Section 2.3, we set out the spatial Durbin model with parameterized weight matrices for each spatial lag and explain how the parameters and direct and indirect spillover effects can be estimated. In Section 2.4, we present the results of a Monte Carlo experiment. In Section 2.5 we discuss the results of an empirical analysis. In Section 2.6, we draw conclusions.

[1] Whether or not to row normalize distance decay matrices has been an issue of discussion in the spatial econometric literature. Furthermore, there is no unifying consensus about which normalization to use. Therefore, we provide estimators and mathematical expressions for the different types of distance decay and normalization.

2.2 Parameterizing spatial weight matrix and alternative approaches

Specifying the spatial weight matrix has been recognized as difficult and controversial (Bavaud, 1998; Corrado and Fingleton, 2012, among others). Researchers generally agree that spatial weights should decrease with some generalized distance, but a univocal specification is still missing. According to Corrado and Fingleton (2012), the spatial econometric model and the specification of the spatial weight matrix, commonly symbolized by $\boldsymbol{W}$, should be theory-driven. While several studies develop an economic-theoretical model that leads to a certain type of spatial econometric model supporting the inclusion of spatially lagged dependent variables, regressors and/or errors[2], research on an economic-theoretical derivation of the spatial weight matrix is scarce. Bavaud (1998) provides a systematic overview of the general theoretical properties of spatial weight models, but does not explain how to apply the theory in practice. The author even apologizes for this omission (p. 154). Instead, practitioners tend to adopt one of the many popular specifications and use it for all spatial lags in the model.[3]

Nevertheless, there are studies that specify a functional form for the elements of $\boldsymbol{W}$ that depends on one or several measures of distance (e.g. geographical distance, or economic distance, or combinations thereof) and one or more distance decay parameters. The two most popular parameterized spatial weight matrices are the inverse distance matrix and the exponential distance decay matrix based on geographical distance only.[4] However, many studies utilizing parameterized

[2] Studies with derivations leading to SDM or the linear-in-means model are of Ertur and Koch (2009) on economic growth, of Yesilyurt and Elhorst (2017) on military expenditures, of Heijnen and Elhorst (2018) on the diffusion of waste disposal taxes, and of Szumilo (2020) on housing prices. Blume et al. (2015) set up a theoretical model, a social interaction game, in which utility depends on an individual's own action and that of member's actions, from which the linear-in-means model can be derived. Xu and Lee (2019) provide a similar theoretical foundation for spatial econometric models.

[3] Popular choices for the spatial weight matrix are (i) p-order binary contiguity matrices (if $p = 1$ only first-order neighbours are included, if $p = 2$ first and second order neighbors are considered, and so on); (ii) inverse or exponential distance decay matrices (with or without a cut-off point); (iii) q-nearest neighbour matrices (where q is a positive integer); (iv) block diagonal or group interaction matrices where each block represents a group of units that interact with each other but not with observation in other groups; and (v) weight matrices based on socio-economic variables rather than distance.

[4] Examples of more complex forms of parameterizations are Cliff and Ord (1973) who use a $\boldsymbol{W}$ matrix whose elements are written as a Cobb-Douglas function of the physical distance and the length of the common border between each pair of units; or Elhorst and Halleck Vega (2017) who use a gravity type of function based on the distance, and the economic sizes (measured by population or gross product) of both units.

spatial weight matrices use pre-specified values for the distance decay parameters (Murdoch et al., 1993; Fingleton and LeGallo, 2008, among others). The reason is that parameterizing $\boldsymbol{W}$ leads to an econometric model that is non-linear in its parameters and thus requires non-linear estimation techniques enhancing the complexity of the analysis.[5] There are also studies that really estimate the distance decay parameter (Fischer et al., 2009; Dubin, 1988; Kakamu, 2005; and Halleck Vega and Elhorst, 2015, to mention a few). However, these approaches also have limitations. First, some of them focus on one spatial lag only. Second, the main focus lies either on the empirical application or the spatial econometric regression model, but less on the estimation of the distance decay parameter, whose estimation remains vague.

A more recent strand of literature infers $\boldsymbol{W}$ from the data using various (geo)statistical modelling techniques, provided that the number of observations over time (T) exceeds the number of observations across space (N) (Ahrens and Bhattacharjee, 2015; Lam and Souza, 2020; also see chapter 4 of Beenstock and Felsenstein, 2019, for an overview). However, this approach does not address the majority of empirical studies that are based on $N >> T$.[6] Even if $\boldsymbol{W}$ is assumed to be symmetric and the leading diagonal of $\boldsymbol{W}$ is set to zero (since units cannot affect themselves), the number of unknown elements amounts to $\frac{1}{2}N(N+1)$ and are under-identified when T is not sufficiently large. Recent studies in the social interactions literature propose to estimate the unknown elements of $\boldsymbol{W}$, but instead of assuming that T is sufficiently large and they assume that the number of unknown elements of $\boldsymbol{W}$ is sufficiently small (De Paula et al. 2020; Lewbel et al., 2021). However, whether this assumption makes sense cannot be specified in advance and only be determined by first estimating the strength of the distance decay effect.

One final approach is of Debarsy and LeSage (2018) who investigate a model with a spatially lagged dependent variable and specify the spatial weight matrix as a convex combination of several underlying matrices. The contribution of each submatrix to the overall matrix is estimated together with the other model parameters using Bayesian estimation. One limitation is that only one spatial lag with such

[5] This argument is also mentioned in Corrado and Fingleton (2012, p.217): "These could be estimated alongside other model parameters, but because of the difficulty this would entail, it makes practical sense to assign values to these coefficients a priori."

[6] Lam and Souza (2020) do mention some exceptions in their paper where N is not necessarily small.

an overall spatial weight matrix is considered. If the number of sub-matrices to construct the overall matrix is M and the number of spatial lags is extended to $K+1$, as in a spatial Durbin model, then $M(K+1)$ additional parameters need to be estimated. In contrast, only $K+1$ additional (distance decay) parameters need to be estimated when adopting a parameterized inverse distance or exponential distance matrix for each spatial lag in the model.

To the best of our knowledge, there is no paper that proposes an estimation approach that (i) allows for different spatial weight matrices for all spatial lags of the model and that (ii) enables estimating all those spatial weight matrices in a unified framework. We aim to fill this gap with this chapter, where we allow for different parameterizations (e.g. exponential and inverse distance decay) and normalizations (row normalization and scalar normalization).

2.3 The methodology of parameterization

2.3.1 The parameterized spatial Durbin model

We focus on the following parameterized spatial Durbin model (SDM)

$$y_{it} = \rho_0 \sum_{j=1}^{N} w_{ij}(\delta_0) y_{jt} + \sum_{k=1}^{K} x_{kit}\beta_{0k} + \sum_{k=1}^{K}\sum_{j=1}^{N} w_{kij}(\alpha_{0k}) x_{kjt}\phi_{0k} + c_{0i} + \varpi_{0t} + v_{it}, \quad (2.1)$$

where $i = 1,\ldots,N$, $t = 1,\ldots,T$, N is the number of cross-sectional units, and T is the number of time periods. We assume a large N, finite T and a balanced panel. y_{it} represents the dependent variable of unit i at time t, and x_{kit} the kth non-stochastic explanatory variable with coefficient β_{0k}. The term $\sum_{j=1}^{N} w_{ij}(\delta_0) y_{jt}$ denotes the spatial lag of the dependent variable of other units than i, and the accompanying coefficient ρ_0 the impact of this spatial lag. Similarly, the regressors $\sum_{j=1}^{N} w_{kij}(\alpha_{0k}) x_{kjt}$ ($k = 1,\ldots,K$) denote the spatial lags of the explanatory variables, whose impacts are measured by the coefficients ϕ_{0k}. The elements $w_{ij}(\delta_0)$ and $w_{kij}(\alpha_{0k})$ measure the relationships between unit i and j. These elements are heterogeneous for the different types of regressors and are parameterized e.g. depend on additional parameters δ_0 and α_{0k}. The individual fixed effects c_{0i} ($i = 1,\ldots,N$) control for unobserved individual-specific, time-invariant effects. Similarly, the time period fixed

effects ϖ_{0t} $(t = 1, \ldots, T-1)$[7] control for time-specific, unit-invariant effects. As T is finite, ϖ_{0t} may just be incorporated in the set of explanatory variables. We write them out explicitly to make clear that they should not be spatially lagged. We allow c_{0i} and ϖ_{0t} to be correlated with x_{kit} and $\sum_{j=1}^{N} w_{kij}(\alpha_{0k})x_{kjt}$. When the subscript 0 is added to a parameter or a vector of parameters, we denote the true value of this parameter or vector. Finally, v_{it} is an error term whose properties are specified in Section 2.3.3.

The main differences of Equation (2.1) compared to the standard SDM in the literature are twofold. First, we allow each spatial lag to have a different spatial weight matrix. This provides more flexibility, because the spatial range, e.g. the distance at which explanatory variables still have an effect on other units, will no longer be assumed to be the same but different from one explanatory variable to another. This setup generalizes most previous empirical spatial econometric studies that employ one common weight matrix for all spatial lags.[8] Second, instead of pre-specifying spatial weight matrices, we parameterize them by distance decay parameters δ_0 and α_{0k} $(k = 1, .., K)$, and estimate these decay parameters jointly with the other parameters of the model.

Model (2.1) can be rewritten in matrix notation. If we stack the individual observations for each time period t, the model reads as

$$\begin{aligned} \boldsymbol{y}_t &= \rho_0 \boldsymbol{W}(\delta_0)\boldsymbol{y}_t + \sum_{k=1}^{K} \boldsymbol{x}_{kt}\beta_{0k} + \sum_{k=1}^{K} \boldsymbol{W}(\alpha_{0k})\boldsymbol{x}_{kt}\phi_{0k} + \boldsymbol{c}_0 + \varpi_{0t}\boldsymbol{\iota}_N + \boldsymbol{v}_t, \\ & t = 1, \ldots, T, \end{aligned} \tag{2.2}$$

where $\boldsymbol{y}_t$, $\boldsymbol{x}_{kt}$, $\boldsymbol{c}_0$, $\boldsymbol{v}_t$ are $N \times 1$ vectors, and $\boldsymbol{\iota}_N$ is an $N \times 1$ vector of ones. $\boldsymbol{W}(\delta_0)$ and $\boldsymbol{W}(\alpha_{0k})$ are $N \times N$ matrices describing the connectivity of all N cross-sectional units in the sample. The corresponding reduced form of Equation (2.2) reads as

$$\begin{aligned} \boldsymbol{y}_t &= (\boldsymbol{I}_N - \rho_0 \boldsymbol{W}(\delta_0))^{-1} \left(\sum_{k=1}^{K} \boldsymbol{x}_{kt}\beta_{0k} + \sum_{k=1}^{K} \boldsymbol{W}(\alpha_{0k})\boldsymbol{x}_{kt}\phi_{0k} + \boldsymbol{c}_0 + \varpi_{0t}\boldsymbol{\iota}_N + \boldsymbol{v}_t \right), \\ & t = 1, \ldots, T. \end{aligned} \tag{2.3}$$

[7] One time dummy is left aside to avoid perfect multicollinearity with the individual fixed effects.

[8] For the same reason, it extends the classic linear-in-means model characteristic of the social interactions literature (De Paula et al., 2020; Lewbel et al., 2021), where ρ_0 captures endogenous interaction effects and ϕ_{0k} contextual effects.

2.3.2 The parameterized spatial weight matrix and its normalizations

Two commonly used parameterized specifications of the spatial weight matrix are the negative exponential and the inverse distance matrix. For $i \neq j$, the ijth element of the negative exponential matrix in its raw, e.g. unnormalized, form is given by

$$w_{ij}(\varrho_0)^r = e^{-d_{ij}\varrho_0}, \tag{2.4}$$

and the ijth element of the inverse distance matrix is

$$w_{ij}(\varrho_0)^r = 1/d_{ij}^{\varrho_0}, \tag{2.5}$$

where d_{ij} denotes the distance between units i and j and ϱ_0 the distance decay parameter of the corresponding spatial lag. Importantly, we depart from a heterogeneous distance decay structure, i.e., each spatial lag has its own distance decay parameter ϱ_0 . Both functional forms have the property that $w_{ij}(\varrho_0)^r \overset{\varrho_0 \to 0}{\longrightarrow} 1$ and $w_{ij}(\varrho_0)^r \overset{\varrho_0 \to \infty}{\longrightarrow} 0$. This implies that, depending on the value ϱ_0, the full weight matrix may be sparse or dense, i.e., for large values of ϱ_0 it has many zeros and hardly any non-zero elements, while small values of ϱ_0 lead to a dense matrix without or hardly any zeros. Note that the diagonal elements of the spatial weight matrix are assumed to be zero to prevent units from predicting themselves.

In general, elements of the spatial weight matrix are normalized. We consider two frequently used normalizations: normalization by row-sums and scalar normalization by the largest eigenvalue. Let $\boldsymbol{W}^r$ denote the matrix in raw form before normalization. Row normalization is achieved by $w_{ij} = w_{ij}^r / \sum_{j=1}^N w_{ij}^r$ and has the effect that the elements of the resulting weight matrix in every row sum to one. Row normalization is generally applied for the two following reasons. It facilitates the interpretation of operations with the weight matrix as an averaging of neighboring values (Anselin and Bera, 1998), and the spatial autoregressive parameter ρ_0 takes values in the parameter space $(1/\kappa_{min}, 1)$, where κ_{min} is the smallest negative eigenvalue of $\boldsymbol{W}$ (Ord, 1981). However, row normalization has also been criticized. Kelejian and Prucha (2010) demonstrate that normalization of the elements of the spatial weight matrix by a different factor for each row as opposed to a single factor is likely to lead to misspecification problems. Elhorst (2014, Section 2.4) explains

that this especially concerns distance decay matrices, because its economic interpretation in terms of distance decay will then no longer be valid (Anselin 1988, pp. 23-24). The reasons for this are twofold. First, due to row normalization the spatial weight matrix might lose its property of being symmetric, as a result of which the impact of unit i on unit j is not the same as that of unit j on unit i. Second, remote and centrally located units may end up having the same impact, i.e., independent on their relative location.[9] To avoid such misspecification problems, Kelejian and Prucha (2010) propose a normalization procedure where each element of $\boldsymbol{W}^r$ is divided by its largest eigenvalue κ^r_{max}, to get $\boldsymbol{W} = (1/\kappa^r_{max})\boldsymbol{W}^r$. When applying this normalization, ρ_0 again takes values in the parameter space $(1/\kappa_{min}, 1)$, with the difference that the smallest negative eigenvalue of this matrix takes another value. Since some rows will sum up to values greater than one and other to values smaller than one, a scalar normalized weight matrix can no longer be interpreted as an averaging of neighboring values. In empirical work both types of normalizations are applied. Therefore we consider both row normalized and scalar normalized (by its largest eigenvalue) forms of the negative exponential and inverse distance matrices.

2.3.3 Assumptions

Lee and Yu (2010a) set out the assumptions under which the ML estimator of the model parameters in Equation (2.1) are identified, consistent, and asymptotically normal, both when assuming that the error terms are normally distributed or not, and provided that the $\boldsymbol{W}$ matrices are non-stochastic, i.e., not parameterized. The regular rate of convergence is $\sqrt{N}$, provided that T is finite or fixed. Details are set out in Appendix 2.A.1. In this section we discuss the assumptions which need to be adapted, such that the proof by Lee and Yu (2010a) carries over to the model in this study.

Assumption 2.1. *The idiosyncratic errors v_{it}, $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$ are normally distributed or are i.i.d across i and t with mean zero, variance σ_0^2 and finite fourth*

[9] The following example illustrates this. Consider a centrally located unit and a remote unit that both have two neighbors. The distance of the first unit to its neighbours is d, while the distance of the second unit to its neighbours is a multiple of d. Despite this difference in location, the entries in the inverse distance matrix describing the spatial arrangement of the units in the sample will be $1/2$ in both cases, provided that the spatial weight matrix is row normalized.

moment.

Assumption 2.2. *The matrices* $\boldsymbol{W}(\delta_0)$ *and* $\boldsymbol{W}(\alpha_{0k})(k = 1,...K)$ *are exogenous, but stochastic with diagonal elements equal to zero. Before normalization, the row and column sums of each spatial weight matrix are uniformly bounded or, alternatively, diverge to infinity though at a rate slower than* N*.*

Assumption 2.3. $\boldsymbol{S} = \boldsymbol{I}_N - \rho_0 \boldsymbol{W}(\delta_0)$ *is invertible and uniformly bounded in row and column sums in absolute value.* ρ_0 *is in the interior of* $\mathbb{P}$*, where* $\mathbb{P}$ *is an open but bounded interval.* ρ_0 *is bounded away from zero.*

Assumption 2.4. ϕ_{0k} *for* $k = 1, \ldots, K$ *is bounded away from zero.*

Assumption 2.5. *For all* k $\rho_0\beta_{0k}\boldsymbol{W}(\delta_0) + \phi_{0k}\boldsymbol{W}(\alpha_{0k}) \neq \boldsymbol{0}_N$*, and* $[\boldsymbol{I}_N, \boldsymbol{W}(\delta_0), \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0k})]$ *are linearly independent.*

Assumption 2.6. *The regressors* $\boldsymbol{x}_{kt}$ *are nonstochastic and bounded uniformly. There is no multicollinearity among the regressors and their spatially lagged counterparts.*

If the error terms are assumed to be normally distributed, the parameters can be estimated by ML. If they are not truly normally distributed, they can be estimated by quasi (Q)ML based on the properties specified in Assumption 2.1. Although we consider both estimators, we only use the term ML. Assumption 2.2 allows for weak divergence of the spatial weight matrix. In the case of a non-stochastic $\boldsymbol{W}$, $\boldsymbol{W}$ is assumed to be uniformly bounded in both row and column sums in absolute value (Lee and Yu, 2010, Kelejian and Prucha, 2010, among others). We are dealing with spatial weight matrices that depend on distance decay parameters δ_0 and α_{0k} $(k = , 1, ..., K)$, which makes the spatial weight matrices stochastic. Gupta (2019) shows that many established estimation methods also work with an exogenous stochastic spatial weight matrix.[10] However, the fact that the spatial weight matrix is stochastic, needs to be formalized in the weak divergence assumption, which requires that the row and column sums of the stochastic spatial weight matrix may diverge to infinity as long as the sample size N diverges to infinity faster. This is

[10] There are also studies that allow for endogenous spatial weight matrices (Qu and Lee, 2015). Endogeneity can arise due to feedback effects, e.g. if the spatial weight matrix is based on economic variables, which also depend on the dependent variable. However, if the spatial weight matrix depends on geographic distance between countries, counties or cities, such feedback effects do not occur, since distance is exogenous. This property does not change when distance is parameterized.

in line with Assumption 2 in Gupta (2019). Assumption 2.2 can be fulfilled when the spatial weight matrix has a distance decay function form. For an exponential distance decay matrix the condition of uniformly boundedness corresponds to a distance decay parameter of $\varrho_0 > 0$, and for an inverse distance matrix to $\varrho_0 > 1$ (Elhorst et al., 2021).[11] Lee (2004) shows that the row and column sums of the spatial weight matrix in a cross-sectional setting may also diverge to infinity as long as the sample size N diverges to infinity faster.[12] For this reason, we also consider and experiment with an inverse distance matrix using a lower bound in the interval $(0, 1]$.

Assumption 2.3 is a modification of Assumption 3 in Lee and Yu (2010a). The invertibility of $\boldsymbol{S}$ guarantees that Equation (2.3) is valid. For ρ_0 we use the interval $(-1, +1)$, with the exception of $\rho_0 = 0$. If $\boldsymbol{W}(\delta_0)$ is scalar or row normalized the upper bound is 1 by construction. The lower bound might be smaller than -1 when scalar normalization is applied, but negative values for ρ_0 in models with only one spatial lag in the dependent variable are considerably less common than positive ones, let alone negative values smaller than -1 (Elhorst, 2014, Section 2.5). Therefore, imposing this lower bound is hardly restrictive. ρ_0 should be bounded away from zero. If ρ_0 equals zero, then the distance decay parameter δ_0 is not identified. This follows from the information matrix of the parameters derived in Appendix 2.A.4.[13] The same non-identification problem occurs for α_{0k} if $\phi_{0k} = 0$ ($k = 1, \ldots, K$), which is excluded in Assumption 2.4.

Assumption 2.5 states further identification assumptions. The first part is different and less restrictive than the condition $\rho_0\beta_{0k} + \phi_{0k} \neq 0$ for all k in Bramoullé et al. (2009, Proposition 1) and Lee and Yu (2016, Lemmas 2 and 3). This is due to the fact that we have not one common but different weight matrices for each spatial lag. Re-

[11] When $\boldsymbol{W}$ is a parameterized exponential distance matrix, the corresponding row or column sum of this series in a continuous space (to ease calculations) can be calculated as the integral $\int_1^N e^{-\varrho_0 x} dx = (1/\varrho_0)e^{-\varrho_0}(1 - e^{-N})$. The row or column sums represented by this integral are upper bounded for $\varrho_0 > 0$ if N goes to infinity. Similarly, the integral $\int_1^N \frac{1}{x^{\varrho_0}} dx = 1/(1 - \varrho_0)(N^{1-\varrho_0} - 1)$ is upper bounded for $\varrho_0 > 1$ if N goes to infinity.

[12] Although not part of the formal proof in Lee and Yu (2010a), they state that this result also carries over to a panel data setting. This point occurs at $1/(1 - \varrho_0)(N^{1-\varrho_0} - 1)/N = 1/(1 - \varrho_0)(1/N_0^{\varrho} - 1/N)$, which converges to zero and thus is upper bounded for $\varrho_0 > 0$ if N goes to infinity.

[13] If ρ_0 equals zero, then the elements in the information matrix that are based on second order derivatives involving δ_0 equal zero (see equations A.35,A.40,A.44,A.47,A.48,A.49). Consequently one row and column contains zeros only. Thus the information matrix is not invertible and the variance covariance matrix is not defined.

writing the reduced form in Equation (2.3) using the fact that $(\boldsymbol{I}_N - \rho_0 \boldsymbol{W}(\delta_0))^{-1} = \boldsymbol{I}_N + \rho_0 \boldsymbol{W}(\delta_0) + \rho_0^2 \boldsymbol{W}^2(\delta_0) + \ldots$ yields

$$\begin{aligned}\boldsymbol{y}_t &= (\boldsymbol{I}_N - \rho_0 \boldsymbol{W}(\delta_0))^{-1} \left[\sum_{k=1}^{K} (\boldsymbol{x}_{kt} \beta_{0k} + \boldsymbol{W}(\alpha_{0k}) \boldsymbol{x}_{kt} \phi_{0k}) + \boldsymbol{c} + \varpi_{0t} \boldsymbol{\iota}_N + \boldsymbol{W}_t \right] \\ &= \sum_{k=1}^{K} \left[\boldsymbol{x}_{kt} \beta_{0k} + [\rho_0 \beta_{0k} \boldsymbol{W}(\delta_0) + \phi_{0k} \boldsymbol{W}(\alpha_{0k})][\boldsymbol{I}_N + \sum_{g=2}^{\infty} \rho_0^{g-1} \boldsymbol{W}^{g-1}(\delta_0) \boldsymbol{x}_{kt}] \right] \\ &+ (\boldsymbol{I}_N - \rho_0 \boldsymbol{W}(\delta_0))^{-1} [\boldsymbol{c}_0 + \varpi_{0t} \boldsymbol{\iota}_N + \boldsymbol{v}_t]. \end{aligned} \tag{2.6}$$

This expression shows that the spatial lags $\boldsymbol{W}(\delta_0)\boldsymbol{y}_t$ and $\boldsymbol{W}(\alpha_{0k})\boldsymbol{x}_{kt}$ $(k = 1, \ldots, K)$ cancel each other out if $\rho_0 \beta_{0k} + \phi_{0k} = 0$ for all k and $\delta_0 = \alpha_{0,1} = \ldots = \alpha_{0K}$. Under these restrictions, the spatial Durbin model (SDM) reduces to the spatial error model (SEM) (Burridge, 1981), as a result of which the coefficients of the spatial lags $\boldsymbol{W}(\delta_0)\boldsymbol{y}_t$ and $\boldsymbol{W}(\alpha_{0k})\boldsymbol{x}_{kt}$ $(k = 1, \ldots, K)$ are not identified.[14] However, in empirical work we do not expect the estimates for the decay parameters δ_0 and α_{0k} to be identical. Thus this assumption is not very restrictive.

Bramoullé et al. (2009) and Lee and Yu (2016) show that $[\boldsymbol{I}_N, \boldsymbol{W}(\delta_0), \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0k})]$ should also be linearly independent, as stated in the second part of Assumption 2.5. Since the diagonal elements of $\boldsymbol{W}(\delta_0)$ and $\boldsymbol{W}(\alpha_{0k})$ are assumed to be zero for all k, the identity matrix and $\boldsymbol{W}(\delta_0)$ are independent from each other by construction. If δ_0 goes to infinity, the elements of $\boldsymbol{W}(\delta_0)$ go to zero, such that it seems that we end up with a null matrix that is dependent on the identity matrix. However, the elements of distant units go to zero much faster than nearby units, as a result of which the distance decay matrix converges to a binary contiguity matrix, which is independent of the identity matrix.[15] For the same reasons, the product of two distance decay matrices will not be equal to one of them by construction: $\boldsymbol{W}(\delta_0) \neq \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0k})$.

As in Lee and Yu (2010a), we assume exogenous and uniformly bounded regressors. Given the exogeneity of the explanatory variables and the weak divergence of the corresponding spatial weight matrices, the spatially lagged regressors

[14] We also validated these findings as part of our Monte Carlo simulation experiment in the next section.

[15] Suppose two elements 1/2 and 1/4 and that both are taken to the power 10. Then the first element becomes 1024 times as large as the second element. When the matrix is row or scalar normalized subsequently, the first element will dominate the second, i.e., the last element can be set to zero equally well without changing the structure of the weight matrix.

are also bounded uniformly. Assumption 2.6 also rules out multicollinearity among the regressors.

2.3.4 Estimation

Generally, individual fixed effects are concentrated out by demeaning the variables by their individual-specific means (Baltagi, 2005).[16] The dependent variable of this demeaned model reads as $\widetilde{y}_{it} = y_{it} - \frac{1}{T}\sum_{t=1}^{T} y_{it}$. Similar transformations are applied to right-hand side elements of the regression equation. The demeaned model for time period t is

$$\widetilde{y}_{it} = \rho_0 \sum_{j=1}^{N} w_{ij}(\delta_0)\widetilde{y}_{jt} + \sum_{k=1}^{K} \widetilde{x}_{kit}\beta_{0k} + \sum_{k=1}^{K}\sum_{j=1}^{N} w_{ij}(\alpha_{0k})\widetilde{x}_{kjt}\phi_{0k} + \widetilde{\omega}_{0t} + \widetilde{v}_{it}. \quad (2.7)$$

Although this transformation does eliminate c_{0i}, it induces linear dependence of the transformed errors $\widetilde{v}_{it}$ over time. Consequently $\widehat{\sigma}^{2w}$ will be biased when T is small or fixed. To get an unbiased estimate, Lee and Yu (2010a) propose the following bias-correction (bc) $\widehat{\sigma}^{2w-be} = \frac{T}{T-1}\widehat{\sigma}^{2w}$. This correction can be easily carried out after the parameters of the model have been estimated. When stacking the individual observations in each time period t, the demeaned spatial Durbin model of Equation (2.7) can be rewritten as

$$\widetilde{\boldsymbol{y}}_t = \rho_0 \boldsymbol{W}(\delta_0)\widetilde{\boldsymbol{y}}_t + \sum_{k=1}^{K} \widetilde{\boldsymbol{x}}_{kt}\beta_{0k} + \sum_{k=1}^{K} \boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k} + \widetilde{\omega}_{0t}\boldsymbol{\iota}_N + \widetilde{\boldsymbol{v}}_t, \quad t = 1, \ldots, T, \quad (2.8)$$

where $\widetilde{\boldsymbol{y}}_t$, $\widetilde{\boldsymbol{x}}_{kt}$, and $\widetilde{\boldsymbol{v}}_t$ are $N \times 1$ vectors, and $\boldsymbol{\iota}_N$ is an $N \times 1$ vector of ones. When also $\boldsymbol{x}_{kt}$ and $\boldsymbol{W}(\alpha_{0k})\boldsymbol{x}_{kt}$ for each k, the model reads as

$$\widetilde{\boldsymbol{y}}_t = \rho_0 \boldsymbol{W}(\delta_0)\widetilde{\boldsymbol{y}}_t + \widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta}_0 + \widetilde{\boldsymbol{v}}_t, \quad (2.9)$$

where the $N \times (2K+1)$ matrix $\widetilde{\boldsymbol{X}}_t^{*}$ is the corresponding sorted matrix of all explanatory variables, their spatial lags and dummy variables for time periods, and

[16] This is because these fixed effects are not of interest (or not reported), cannot be estimated consistently when T is small, and might affect the accuracy of parameter estimates when taking them up as part of the regressors if N grows large.

is given by

$$\widetilde{\boldsymbol{X}}_t^* = [\widetilde{\boldsymbol{x}}_{1t}, \ldots, \widetilde{\boldsymbol{x}}_{Kt}, \boldsymbol{W}(\alpha_{0,1})\widetilde{\boldsymbol{x}}_{1t}, \ldots, \boldsymbol{W}(\alpha_{0K})\widetilde{\boldsymbol{x}}_{Kt}, \boldsymbol{\iota}_N]. \tag{2.10}$$

To avoid multicollinearity, $T-1$ time period dummies are included. Then $\boldsymbol{\zeta}_0 = (\boldsymbol{\beta}_0^{'}, \boldsymbol{\phi}_0^{'}, \widetilde{\omega}_{0t})^{'}$. The remaining parameters to be estimated are the spatial autoregressive parameter ρ_0, the distance decay parameters for all spatial lags $\boldsymbol{\alpha}_0 = (\delta_0, \alpha_{0,1}, \ldots., \alpha_{0K})'$ and σ_0^2. Defining the full set of parameters as $\boldsymbol{\theta} = (\boldsymbol{\zeta}^{'}, \rho, \sigma^2, \boldsymbol{\alpha}^{'})^{'}$, the log-likelihood function of the model is given by

$$\ln L^w(\boldsymbol{\theta}) = -\frac{NT}{2}\log(2\pi\sigma^2) + T\log|\boldsymbol{I}_N - \rho\boldsymbol{W}(\delta)| - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t^{'}(\boldsymbol{\zeta})\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}), \tag{2.11}$$

where

$$\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}) = \widetilde{\boldsymbol{y}}_t - \rho\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}. \tag{2.12}$$

The ML estimator can be obtained by maximizing the log-likelihood function with respect to $\boldsymbol{\theta}$.[17] The parameter vector $\boldsymbol{\zeta}$ and the scalar σ^2 can be solved analytically from their first order conditions, given ρ and $\boldsymbol{\alpha} = (\delta, \alpha_1, \ldots, \alpha_K)^{'}$, which yields

$$\widehat{\boldsymbol{\zeta}}^w(\rho, \boldsymbol{\alpha}) = [\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}\widetilde{\boldsymbol{X}}_t^*]^{-1}[\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{S}(\rho, \delta)\widetilde{\boldsymbol{y}}_t], \tag{2.13}$$

$$\widehat{\sigma}^{2w}(\rho, \boldsymbol{\alpha}) = \frac{1}{NT}\sum_{t=1}^{T}(\boldsymbol{S}(\rho, \delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t^*\widehat{\boldsymbol{\zeta}}^w)^{'}(\boldsymbol{S}(\rho, \delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t^*\widehat{\boldsymbol{\zeta}}^w), \tag{2.14}$$

where $\boldsymbol{S}(\rho, \delta) = \boldsymbol{I}_N - \rho\boldsymbol{W}(\delta)$. By substituting these solutions for $\widehat{\boldsymbol{\zeta}}^w$ and $\widehat{\sigma}^{2w}$ into Equation (2.11), the concentrated log-likelihood function with respect to the $K+2$ remaining estimated parameters $\widehat{\rho}^w$ and $\widehat{\boldsymbol{\alpha}}^w$ are obtained. Given the solution of this maximization problem for $\widehat{\rho}^w$ and $\widehat{\boldsymbol{\alpha}}^w$, we can subsequently determine the unconditional ML estimates of $\widehat{\boldsymbol{\zeta}}^w$ and $\widehat{\sigma}^{2w}$, as well as the bias-corrected outcome $\widehat{\sigma}^{2w-bc} = \frac{T}{T-1}\widehat{\sigma}^{2w}$.

We explored two ways to find the ML estimates of ρ and $\boldsymbol{\alpha}$. The first is to set their first order derivatives equal to zero, given $\widehat{\boldsymbol{\zeta}}^w$ and $\widehat{\sigma}^{2w}$, and to solve this sys-

[17] Mathematical expressions for the first and second order conditions and the information matrix, which will be used to determine the variance-covariance matrix of the parameters, are reported in the appendix.

tem of equations numerically. Technically, this is complicated not only because the parameters appear in multiple equations, but also because we ended up with multiple solutions, including unbounded solutions. The second way is to maximize the concentrated log-likelihood function of ρ and $\boldsymbol{\alpha}$. The concentrated log-likelihood within this setup is programmed as

$$\log L^w(\rho, \boldsymbol{\alpha} | \widehat{\boldsymbol{\zeta}}^w, \widehat{\sigma}^{2w}) = -\frac{NT}{2}\log(2\pi\widehat{\sigma}^{2w}) + T\log|\boldsymbol{I}_N - \rho\boldsymbol{W}(\delta)| - \frac{NT}{2}, \quad (2.15)$$

where $\widehat{\sigma}^{2w}$ is programmed as in (2.14), and $\widehat{\boldsymbol{\zeta}}^w$ as part of $\widehat{\sigma}^{2w}$ in (2.14) is programmed as in (2.13). This has the effect that if one or more values of ρ and $\boldsymbol{\alpha}$ change, the estimates for $\widehat{\boldsymbol{\zeta}}^w(\rho, \boldsymbol{\alpha})$ and $\widehat{\sigma}^{2w}(\rho, \boldsymbol{\alpha})$ change accordingly.

One final issue is the estimation of ρ and δ if $\boldsymbol{W}(\delta)$ is not row but scalar normalized. The issue is that $\rho\boldsymbol{W}(\delta) = \rho\boldsymbol{W}^r(\delta)/\kappa^r_{max}$ is equivalent with $\rho^*\boldsymbol{W}^*(\delta)$, where $\rho^* = c\rho$ and $\boldsymbol{W}^*(\delta) = \boldsymbol{W}(\delta)/c$ and c denotes any scalar factor (Kelejian and Prucha, 2010, p.55). For this reason, Ahrens and Bhattacharjee (2015) normalize by $\rho = 1$ and then try to estimate the elements of the spatial weight matrix. We apply the following approach. First, ρ is estimated setting $\delta = 1$ (the default value) and $\boldsymbol{W}(\delta)$ is scalar normalized, which yields $\rho = \rho_{initial}$. This $\rho_{initial}$ is then kept fixed when maximizing the concentrated log-likelihood function for δ and the other α_k ($k = 1, \ldots, K$). During this iterative process $\boldsymbol{W}(\delta)$ is scalar normalized every time δ changes. Only after the optimal values for δ, α_k ($k = 1, \ldots, K$) are found, among which $\widehat{\delta}_{ML}$, ρ is estimated again to get $\rho = \widehat{\rho}_{ML}$. We developed an estimation routine, which enables estimating the proposed model for different normalizations and distance decay functions. A description of the routine and the expressions for the variance-covariance matrix of the parameter estimates can be found in the appendix of this chapter.

2.3.5 Interpretation: direct and indirect spillover effects

A quantitative interpretation of the coefficient estimates in Equation (2.1) is not advisable, because they do not represent the marginal effects of the explanatory variables (LeSage and Pace, 2009). These marginal effects are obtained by taking the first-order derivatives of the reduced form of the SDM in Equation (2.3). This

yields an $N \times N$ matrix $\mathbf{\Psi}_k$ of the form

$$\mathbf{\Psi}_k = (\boldsymbol{I}_N - \rho \boldsymbol{W}(\delta))^{-1}(\boldsymbol{I}_N \cdot \beta_k + \boldsymbol{W}(\alpha_k) \cdot \phi_k) \qquad \text{for} \quad k = 1, ..., K. \tag{2.16}$$

Since the weight matrices do not change over time, this expression is the same for each time period t. Every diagonal element of $\mathbf{\Psi}_k$ reflects the impact of changing the kth explanatory variable of one observational unit i on the dependent variable of i. Every off-diagonal element reflects the impact of changing the kth explanatory variable of one observational unit i on the dependent variable of another unit j, $j \neq i$. To suppress the amount of output, LeSage and Pace (2009) suggest two summary indicators: the direct effect DE_k measured as the average of all diagonal elements of $\mathbf{\Psi}_k$, and the indirect spillover effect IE_k measured as the average row or column sum of the off-diagonal elements excluding the respective diagonal element. Both effects are calculated as

$$DE_k = \frac{1}{N} tr(\mathbf{\Psi}_k) = \frac{1}{N} tr\{(\boldsymbol{I}_N - \rho \boldsymbol{W}(\delta))^{-1}(\boldsymbol{I}_N \cdot \beta_k + \boldsymbol{W}(\alpha_k) \cdot \phi_k)\}, \tag{2.17}$$

$$\begin{aligned} IE_k &= \frac{1}{N} \boldsymbol{\iota}_N^{'} \mathbf{\Psi}_k \boldsymbol{\iota}_N - \frac{1}{N} tr(\mathbf{\Psi}_k) \\ &= \frac{1}{N} \boldsymbol{\iota}_N^{'} \{(\boldsymbol{I}_N - \rho \boldsymbol{W}(\delta))^{-1}(\boldsymbol{I}_N \cdot \beta_k + \boldsymbol{W}(\alpha_k) \cdot \phi_k)\} \boldsymbol{\iota}_N - \\ &\quad \frac{1}{N} tr\{(\boldsymbol{I}_N - \rho \boldsymbol{W}(\delta))^{-1}(\boldsymbol{I}_N \cdot \beta_k + \boldsymbol{W}(\alpha_k) \cdot \phi_k)\}. \end{aligned} \tag{2.18}$$

These expressions show two facts. First, the parameters ϕ_k of the spatial lags of the explanatory variables contribute to the required flexibility of the indirect spillover effects, as claimed in the introduction of this paper. This can be seen by imposing the restriction $\phi_k = 0$ to obtain the SAR model without spatial lags in the explanatory variables. Then the matrix $\mathbf{\Psi}_k$ simplifies to $\mathbf{\Psi}_k^{SAR} = (\boldsymbol{I}_N - \rho \boldsymbol{W}(\delta))^{-1}(\boldsymbol{I}_N \cdot \beta_k) = \beta_k (\boldsymbol{I}_N - \rho \boldsymbol{W}(\delta))^{-1} \equiv \beta_k \boldsymbol{S}_I$, and as a result the direct effect to $\frac{\beta_k}{N} tr(\boldsymbol{S}_I)$ and the indirect spillover effect to $\frac{\beta_k}{N} \boldsymbol{\iota}_N^{'} \boldsymbol{S}_I \boldsymbol{\iota}_N - \frac{\beta_k}{N} tr(\boldsymbol{S}_I)$. Consequently, the ratio between the indirect spillover effect and the direct effect becomes

$$\frac{IE_k^{SAR}}{DE_k^{SAR}} = \frac{\frac{\beta_k}{N} \boldsymbol{\iota}_N^{'} \boldsymbol{S}_I \boldsymbol{\iota}_N - \frac{\beta_k}{N} tr(\boldsymbol{S}_I)}{\frac{\beta_k}{N} tr(\boldsymbol{S}_I)} = \boldsymbol{\iota}_N^{'} \boldsymbol{S}_I \boldsymbol{\iota}_N / tr(\boldsymbol{S}_I) - 1, \tag{2.19}$$

which is independent of the parameter β_k. This independence implies that this ratio is the same for every explanatory variable in the model.[18]

The second fact Equations (2.17) and (2.18) show is that parameterizing the spatial weight matrix of every explanatory variable enhances this flexibility by the decay parameters δ and α_k. This is because the distance at which explanatory variables may still have effect on other units, i.e., the spatial range, may also be different from one variable to another.

To draw statistical inferences on the direct and indirect spillover effects, we use the delta method since it saves on computation time. Mathematical expressions are provided in Appendix 2.A.6.

2.4 Monte Carlo simulations

We conduct Monte Carlo experiments to evaluate the performance of our proposed estimator with the parameterized exponential and inverse distance matrix in row normalized and scalar normalized form. Therefore we generate separate experiments for each type of spatial weight matrix and normalization and investigate the performance of our estimator and compare it with other existing approaches.

Our data generating process contains two explanatory variables: $x_{1it} \sim \mathcal{N}(1, 2)$ and $x_{2it} \sim \mathcal{N}(-1.5, 3.5)$. The coefficients of the first variable and its spatial lag are $\beta_1 = -1$ and $\phi_1 = 1.5$, and of the second variable and its spatial lag are $\beta_2 = 0.2$ and $\phi_2 = -0.3$. The unobserved individual fixed effects and the error terms are both generated from a normal distribution with mean 0 and variance 1. We set $N = \{200, 800\}$ and $T = 5$ and the number of iterations is 1000. To construct distance matrices between the cross-sectional units, we use the coordinates of N data points evenly set in a rectangle of 10×20 for $N = 200$ and 20×40 for $N = 800$. By using a rectangle, the row and scalar normalized matrices are different by construction. These distances are used to calculate the inverse distance and the exponential distance decay matrix. The decay parameters are $\delta = 2$, $\alpha_1 = 1.5$, and $\alpha_2 = 3$. Finally, the parameter of the spatial lag of the dependent variable is set to $\rho = 0.5$.

[18] Elhorst (2010) is the first who has pointed this out, though only based on a simple numerical example, and for this reason advocated the inclusion of spatial lags in the explanatory variables to avoid unnecessary restrictions on the sign and magnitude of spatial spillovers in empirical research.

The simulation design described above serves as our basic case and is labelled Case I. Additionally, we carry out further experiments, where we vary either the spatial lag parameter ρ or one of the distance decay parameters.

We compare the performance of three estimators. We label our proposed approach where all spatial weight matrices are parameterized as PWFE, where P indicates parameterized, W indicates spatial weight matrix, and FE indicates fixed effects.

In addition we investigate what would happen if a practitioner replaces all spatial weight matrices by one common spatial weight matrix with a decay parameter of 1 for all spatial lags. This reflects the most widely used approach applied by practitioners. Since researchers do not know the true decay parameters and the true specification of the spatial weight matrices, this investigation may throw more light on how harmful this can be and is referred to as WFE (one common $\boldsymbol{W}$ and fixed effects). Finally, we evaluate the performance of an estimator assuming that the true spatial weight matrices are known to the researcher. Note that this reflects an ideal, but hypothetical situation. However, this may throw more light on the relative performance of our approach and the associated costs in terms of estimation errors compared to this ideal estimator which we refer to as TWFE (true T spatial weight matrices $\boldsymbol{W}$ and fixed effects).

In our analysis we investigate both the parameter estimates and the estimates of the direct and indirect spillover effects.[19] To judge the performance of the estimators, we consider several statistics: the average bias (Bias), the root mean square error (RMSE), the median bias (Mbias), and the median absolute value of the bias (Mabias). The latter two are used as they are more robust to outliers. Furthermore, we test for each parameter whether it equals its true value and report mean and standard deviations of the p-values. If the underlying asymptotic distribution is true, then under the null the p-values should follow a $U(0,1)$ distribution, and thus should have a mean p-value of 0.5 and a standard deviation of approximately 0.29.

Table 2.1 reports the Bias, RMSE, Mbias, and Mabias of the parameter estimates; Table 2.2 reports the same statistics of the direct and indirect spillover effects; and

[19] LeSage and Pace (2018) demonstrate that past studies' focus exclusively on point estimates may not provide useful information regarding the statistical properties of the marginal effects, i.e., the direct and indirect spillover effects obtained from these point estimates.

Table 2.3 reports mean and standard deviation of the p-values of both parameter estimates and direct and indirect effects. To save space, we focus on the results of the negative exponential distance matrix. Results for the inverse distance matrix can be found in Appendix 2.A.7. When focusing on the proposed estimator PWFE

Table 2.1. Simulation results for Case I: $\rho = 0.5$, $\delta = 2$, $\alpha_1 = 1.5$, $\alpha_2 = 3$

Settings				Exponential: row normalized									Exponential: scalar normalized								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)
WFE	**5**	**200**	**Bias**	0.037	-0.011	0.327	-0.140	-0.209	0.099				0.032	-0.010	0.309	-0.156	-0.237	0.082			
			RMSE	0.038	0.015	0.329	0.153	0.214	0.116				0.033	0.015	0.312	0.171	0.245	0.102			
			Mbias	0.037	-0.011	0.327	-0.138	-0.207	0.098				0.032	-0.010	0.308	-0.156	-0.236	0.081			
			Mabias	0.037	0.011	0.327	0.138	0.207	0.098				0.032	0.011	0.308	0.156	0.236	0.081			
TWFE	**5**	**200**	**Bias**	-0.000	0.000	-0.001	0.000	-0.002	-0.011				0.000	0.000	-0.002	0.000	-0.003	-0.011			
			RMSE	0.008	0.010	0.024	0.025	0.025	0.053				0.008	0.010	0.027	0.027	0.029	0.053			
			Mbias	0.000	0.000	-0.002	-0.000	-0.002	-0.013				0.000	0.000	-0.002	0.000	-0.004	-0.012			
			Mabias	0.005	0.007	0.016	0.016	0.017	0.037				0.005	0.007	0.019	0.017	0.019	0.037			
PWFE	**5**	**200**	**Bias**	0.001	-0.000	0.011	-0.003	-0.014	-0.013	0.036	-0.014	0.156	0.001	0.000	0.002	-0.003	-0.006	-0.013	0.012	-0.004	0.189
			RMSE	0.010	0.010	0.055	0.036	0.053	0.054	0.268	0.109	0.858	0.008	0.010	0.044	0.038	0.032	0.054	0.283	0.095	0.930
			Mbias	0.001	-0.000	0.007	0.000	-0.012	-0.016	0.033	-0.016	-0.003	0.001	-0.000	0.001	-0.000	-0.005	-0.017	0.013	-0.004	0.029
			Mabias	0.006	0.007	0.034	0.024	0.033	0.038	0.172	0.070	0.490	0.005	0.007	0.028	0.026	0.022	0.038	0.194	0.064	0.516
WFE	**5**	**800**	**Bias**	0.033	-0.010	0.359	-0.185	-0.200	0.104				0.029	-0.010	0.368	-0.196	-0.196	0.091			
			RMSE	0.033	0.011	0.360	0.188	0.202	0.108				0.029	0.011	0.368	0.199	0.198	0.096			
			Mbias	0.033	-0.010	0.358	-0.183	-0.200	0.104				0.029	-0.010	0.367	-0.194	-0.1951	0.091			
			Mabias	0.033	0.010	0.358	0.183	0.200	0.104				0.029	0.010	0.367	0.194	0.195	0.091			
TWFE	**5**	**800**	**Bias**	0.000	0.000	-0.001	-0.001	-0.0010	0.000				0.000	0.000	-0.001	-0.001	-0.001	0.000			
			RMSE	0.004	0.005	0.014	0.014	0.015	0.026				0.004	0.005	0.015	0.014	0.016	0.026			
			Mbias	0.000	-0.000	-0.001	-0.002	-0.001	-0.001				0.000	-0.000	-0.001	-0.002	-0.000	-0.001			
			Mabias	0.003	0.003	0.009	0.009	0.010	0.017				0.002	0.003	0.011	0.009	0.011	0.017			
PWFE	**5**	**800**	**Bias**	0.000	0.000	-0.000	-0.002	-0.002	-0.001	0.013	0.002	0.058	0.000	0.000	-0.001	-0.002	-0.008	-0.001	0.010	0.003	0.064
			RMSE	0.005	0.005	0.030	0.020	0.027	0.026	0.135	0.057	0.399	0.004	0.005	0.025	0.021	0.017	0.026	0.137	0.050	0.413
			Mbias	0.000	-0.000	-0.001	-0.000	-0.002	-0.002	0.011	0.000	0.006	0.0000	-0.000	-0.002	-0.000	-0.001	-0.001	0.004	0.000	0.015
			Mabias	0.003	0.003	0.021	0.013	0.019	0.016	0.090	0.036	0.261	0.003	0.003	0.017	0.014	0.012	0.017	0.096	0.034	0.266

Table 2.2. Direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 for Case I using the parameter estimates in Table 2.1

Settings				Exponential: row normalized				Exponential: scalar normalized			
	T	N		DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Bias**	0.001	0.002	0.224	-0.156	0.000	0.001	0.165	-0.150
			RMSE	0.007	0.011	0.242	0.177	0.007	0.010	0.191	0.170
			Mbias	0.001	0.002	0.225	-0.155	0.000	0.001	0.162	-0.150
			Mabias	0.005	0.007	0.225	0.155	0.005	0.007	0.162	0.150
TWFE	**5**	**200**	**Bias**	0.000	0.000	-0.005	0.001	0.000	0.000	-0.006	0.001
			RMSE	0.007	0.010	0.069	0.049	0.007	0.010	0.072	0.047
			Mbias	-0.001	0.000	-0.004	0.001	-0.001	0.000	-0.008	0.001
			Mabias	0.005	0.007	0.048	0.032	0.005	0.007	0.051	0.030
PWFE	**5**	**200**	**Bias**	0.000	0.000	-0.001	-0.002	0.000	0.000	-0.002	-0.003
			RMSE	0.007	0.010	0.088	0.069	0.007	0.010	0.095	0.067
			Mbias	-0.001	0.000	-0.005	0.002	-0.001	0.000	-0.008	0.001
			Mabias	0.005	0.007	0.058	0.047	0.005	0.007	0.062	0.044
WFE	**5**	**800**	**Bias**	0.001	0.001	0.275	-0.223	0.001	0.001	0.271	-0.216
			RMSE	0.004	0.006	0.282	0.228	0.004	0.006	0.278	0.221
			Mbias	0.001	0.001	0.275	-0.220	0.001	0.000	0.268	-0.215
			Mabias	0.002	0.004	0.275	0.220	0.002	0.004	0.268	0.215
TWFE	**5**	**800**	**Bias**	0.000	0.000	-0.002	-0.002	0.000	0.000	-0.001	-0.002
			RMSE	0.004	0.006	0.044	0.027	0.004	0.005	0.043	0.027
			Mbias	0.000	0.000	-0.003	-0.002	0.000	0.000	-0.002	-0.003
			Mabias	0.002	0.004	0.031	0.018	0.002	0.004	0.030	0.018
PWFE	**5**	**800**	**Bias**	0.000	0.000	-0.003	-0.003	0.000	0.000	-0.002	-0.003
			RMSE	0.004	0.006	0.057	0.039	0.004	0.005	0.056	0.039
			Mbias	0.000	0.000	-0.007	0.001	0.000	0.000	-0.003	0.000
			Mabias	0.002	0.004	0.037	0.026	0.002	0.004	0.038	0.025

Table 2.3. Mean and standard deviation of the p-values of the parameters and the direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2

Settings				Negative exponential: row normalized											
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.001	0.338	0.000	0.058	0.000				0.496	0.486	0.023	0.113
			Std	0.009	0.295	0.000	0.152	0.005				0.285	0.291	0.092	0.213
TWFE	**5**	**200**	**Mean**	0.496	0.491	0.495	0.494	0.478				0.489	0.493	0.492	0.497
			Std	0.288	0.293	0.289	0.295	0.288				0.279	0.292	0.290	0.295
PWFE	**5**	**200**	**Mean**	0.488	0.487	0.483	0.482	0.466	0.496	0.473	0.485	0.489	0.487	0.505	0.486
			Std	0.302	0.291	0.288	0.286	0.291	0.289	0.301	0.284	0.279	0.294	0.295	0.291
WFE	**5**	**800**	**Mean**	0.000	0.167	0.000	0.000	0.000				0.501	0.496	0.000	0.000
			Std	0.000	0.233	0.000	0.000	0.000				0.298	0.294	0.000	0.005
TWFE	**5**	**800**	**Mean**	0.504	0.502	0.502	0.499	0.492				0.505	0.495	0.488	0.490
			Std	0.292	0.289	0.288	0.293	0.294				0.301	0.290	0.291	0.294
PWFE	**5**	**800**	**Mean**	0.499	0.503	0.482	0.498	0.480	0.492	0.479	0.487	0.503	0.491	0.497	0.493
			Std	0.298	0.289	0.290	0.290	0.297	0.294	0.300	0.291	0.302	0.291	0.289	0.289
				Negative exponential: scalar normalized											
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.006	0.353	0.000	0.059	0.001				0.496	0.489	0.090	0.115
			Std	0.027	0.296	0.000	0.152	0.007				0.280	0.290	0.205	0.209
TWFE	**5**	**200**	**Mean**	0.491	0.493	0.489	0.498	0.487				0.490	0.493	0.479	0.500
			Std	0.286	0.293	0.288	0.294	0.293				0.279	0.292	0.286	0.296
PWFE	**5**	**200**	**Mean**	0.520	0.494	0.581	0.494	0.643	0.491	0.538	0.503	0.491	0.489	0.501	0.489
			Std	0.284	0.292	0.264	0.284	0.238	0.285	0.281	0.280	0.279	0.293	0.290	0.288
WFE	**5**	**800**	**Mean**	0.000	0.184	0.000	0.000	0.000				0.496	0.498	0.000	0.001
			Std	0.000	0.243	0.000	0.001	0.000				0.298	0.291	0.000	0.009
TWFE	**5**	**800**	**Mean**	0.510	0.502	0.497	0.497	0.489				0.503	0.495	0.483	0.486
			Std	0.296	0.289	0.289	0.290	0.287				0.300	0.290	0.289	0.289
PWFE	**5**	**800**	**Mean**	0.528	0.504	0.566	0.502	0.644	0.500	0.526	0.490	0.502	0.492	0.491	0.500
			Std	0.289	0.289	0.273	0.290	0.233	0.281	0.273	0.288	0.301	0.290	0.292	0.292

first, we observe the following patterns. First, the biases are small and acceptable.[20] Generally, they are smaller for the coefficients of the variables than for the decay parameters, and smaller for $N = 800$ than for $N = 200$. Similarly, increasing the sample size leads to a decrease in the RMSE and the median absolute value of the bias.

In line with the findings from Table 2.1, bias, RMSE, Mbias, and Mabias of the direct and indirect effects are very low in Table 2.2. While the biases in the coefficients and decay parameters are already small in Table 2.1, the biases in the direct and indirect spillover effects in Table 2.2 appear to be even smaller.[21] Increasing the sample size leads to a decrease of bias, RMSE, Mbias, and Mabias.

Table 2.3 shows that for PWFE average p-values and their standard deviation closely fluctuate around 0.5 and 0.29, respectively. Only for the coefficient ρ differences are slightly higher when scalar normalization is used (for the negative expo-

[20] One exception is the average bias in the decay parameter α_2. However, this is driven by a few outliers as can be seen when comparing with the median bias.

[21] The explanation could be that biases in these effects are a nonlinear function of five parameters ($\rho, \beta_k, \phi_k, \delta, \alpha_k$, see Equations (2.17) and (2.18)). Consequently, a bias in one of these parameter may be compensated by another opposite bias in one of the other parameters, as a result of which the bias in the indirect spillover effect, the main focus of applied practitioners, remains small.

nential matrix mean and standard deviation are 0.65 and 0.23, respectively, and for the inverse distance matrix we have an average of 0.69 and a standard deviation of 0.21). However, this has no adverse effect on the p-values of the direct and indirect spillover effects derived from ρ and other coefficients.

As expected, the performance of the model deteriorates when using the WFE estimator which assumes the same spatial weight matrix with a decay parameter of 1 for all spatial lags, while the true but unknown values are 2, 1.5 and 3. This is pronounced in the higher values of bias, RMSE, Mbias, and Mabias in Table 2.1. Especially, the estimates on ϕ_1, ϕ_2, and ρ are severely biased. When investigating the effect estimates in Table 2.2, we see that the bias in the direct effect estimates is small, not more that 0.0002 at the maximum, but the bias in the indirect spillover effects is large. We find biases that exceed 0.5 (in absolute value). This is a cause for concern and supports the approach advocated in this chapter whereby the decay parameters of the spatial lags are estimated rather than set at one particular value for all spatial lags in advance. These biases are also reflected in the p-values in Table 2.3. Except for the direct effect estimates, the p-values of indirect effects and the parameter estimates are far off the desired values.

In the exceptional case that researchers would know the true specification of all spatial weight matrices, the model parameters can be estimated with greater accuracy than in the situation where the decay parameters also need to be estimated (see Table 2.1). This is also evident when comparing the median absolute value of the bias (Mabias) of the indirect spillover effects of TWFE and PWFE reported in Table 2.2. As expected, in Table 2.3 TWFE's measures of the p-values are closer to the desired levels. However, TWFE only reflects a hypothetical situation as practitioners do not know the true values of the decay parameters. In general, the differences in the results of TWFE and PWFE are rather small compared to the differences and negative consequences of using wrong pre-specified weight matrices as with WFE.

We conduct similar experiments where we vary ρ, α_2, or α_1. An overview of the different parameter configurations and the results can be found in Appendix 2.A.7. We summarize the main findings here and focus on the estimates of the direct and indirect effects in particular.

Changing the ρ parameter from rather strong (0.5 in Case I) to mild spatial dependence (Case II, $\rho = 0.25$), or even to a negative value (Case III, $\rho = -0.25$), does

not affect the pattern of results. By contrast, it does change when setting $\rho = 0.01$ (Case IV), which simulates the case where a spatial lag hardly matters and its decay parameter is difficult to identify, in line with Assumption 2.3. This is confirmed in the Monte Carlo simulation results: bias and RMSE of δ increase substantially and to unacceptable levels. In case of the exponential distance matrix, the bias in ρ remains more or less the same, while in case of the inverse distance matrix, both the bias and the RMSE of this parameter increase more. Importantly, this appears to have no effect on the bias and RMSE of the direct and indirect spillover effects. This is reassuring news for practitioners mainly interested in direct and indirect spillover effects estimates of the explanatory variables.

In the remaining experiments, we aim at investigating the parameter space of distance decay parameter in greater detail. Recall, uniformly boundedness in Assumption 2.2 implies the following lower bounds: 0 for the exponential distance decay matrix and 1 for the inverse distance matrix. Weak divergence in Assumption 2.2 implies a lower bound of 0 for the inverse distance matrix. Thus we investigate the consequences when the distance decay parameter is close to these lower bounds, e.g. when $\alpha_2 = 1$ (Case V) and $\alpha_2 = 0.5$ (Case VI). For the exponential distance decay matrix the bias decreases, especially if $\alpha_2 = 1$. At $\alpha_2 = 0.5$ the bias reduction becomes smaller. On the other hand, the median bias always remains close to zero. For the inverse distance matrix we see no improvement, but a deterioration, although the median bias remains close to zero, especially for the larger value of $N = 800$. This implies that the distance decay parameter of an exponential distance matrix can be estimated with greater accuracy than that of an inverse distance matrix. This is because of the property that $\boldsymbol{W}(\varrho_0)$ of the exponential decay matrix is uniformly bounded in both row and column sums in absolute value for $\varrho_0 > 0$, whereas this is not the case for the inverse distance matrix. Then this property only holds if $\varrho_0 > 1$. This implies that values of $\varrho_0 < 1$ need to be handled with care when employing inverse distance matrices.

Another pattern occurs when increasing the decay parameter to an extremely high value, i.e., $\alpha_1 = 10$ (Case VII). Note this value resembles the case that $w_{ij}(\varrho_0)^r \overset{\varrho_0 \to \infty}{\to} 0$, as a result of which the spatial weight matrix boils down to a first-order binary contiguity matrix. Bias and RMSE of α_1 increase substantially, while bias and RMSE of the corresponding spatial lag parameter ϕ_1 remain largely the same. The

increased bias and RMSE in the decay parameter α_1 appears to have no effect on the bias and RMSE of the direct and indirect spillover effects. This finding implies that a practitioner who finds a high value for one of the decay parameters, perhaps one equal to the upper bound, can still value the direct and indirect spillover effects of the corresponding explanatory variable or might adopt a first-order binary contiguity matrix equally well.

2.5 Empirical analysis

The empirical analysis in this section is based on Yesilyurt and Elhorst (YE) (2017), who investigate military spending measured as a ratio of GDP, also known as the defense burden, in 144 countries over the period 1993 to 2007. Explanatory variables are GDP, population, international war, civil war, and political regime. The dependent variable and the first two explanatory variables are measured in logs, while the latter three are measured as scores. The scores on the variables international war and civil war range from 0 (no war) to 10 (greatest). The variable political regime ranges from -10 to +10, where -10 indicates strongly autocratic and +10 strongly democratic countries.[22]

In their paper, YE compare several spatial econometric models and eight potential spatial weight matrices. However, exponential distance decay and inverse distance matrices have not been investigated. Using Bayesian comparison methods developed by LeSage (2014, 2015), they find that the SAR model, e.g. the specification that only includes a spatial lag in the dependent variable, in combination with a spatial weight matrix that is specified as a first-order binary contiguity matrix based on maritime borders produces the highest Bayesian posterior model probability. YE do not find any (or hardly any) significant indirect spillover effects, also not when estimating the SDM instead of the SAR model. This finding is typical of many empirical studies applying SDM: often none or only some of the spatial lags in the explanatory variables and/or indirect spillover effects appear to be significant. One potential problem could be that these studies adopt one common spatial weight matrix for all spatial lags in the model. In the Monte Carlo simulations, it has been found that the direct effects are less sensitive for selecting the wrong

[22] For more details on the data and the variables we refer the reader to Yesilyurt and Elhorst (2017).

spatial weight matrix. By contrast, the indirect spillover effects, which tend to be the main focus of empirical studies, appeared to be extremely sensitive to a wrong choice of the spatial weight matrix.

Therefore we investigate whether using parameterized spatial weight matrices based on distances alter the results. Specifically, we distinguish different matrices for each spatial lag in the model. The results of several estimations are reported in Table 2.4. First column reports the results from an SDM using the preferred binary contiguity matrix from YE. The coefficient estimate of the spatial autoregressive parameter ρ is 0.241 and significant at the 1% level. Three of the five explanatory variables in the model appear to have coefficients and direct effects that are significant at the 1% level and one at the 10% level. However, only one of their corresponding spatial lags and indirect spillover effects, the political regime, appears to be significant at the 5% level. We also test whether there is empirical evidence in favor of the SDM, e.g. we conduct a Wald test testing $H_0 : \phi = 0$. The test statistic is 7.85 with a p-value of 0.16. Consequently, no empirical evidence in favor of the SDM is found. The next four columns of Table 2.4 consider two generalizations of YE.

First, we depart from the binary contiguity matrix and investigate row and scalar normalized negative exponential and inverse distance matrices when estimating one common decay parameter for all spatial lags in the model. This is the first step a practitioner can undertake to determine the best performing spatial weight matrix empirically. The estimated distance decay parameters amount to 2.022 and 2.305 for the row and scalar normalized exponential and to 2.113 and 0.766 for the row and scalar normalized inverse distance matrices, and are all significant. However, when comparing the performance of the SDM for these matrices with that of the binary contiguity matrix, measured in terms of the log-likelihood function value, these parameterized spatial weight matrices mean no improvement. Just as for the binary contiguity matrix, no empirical evidence in favor of SDM is found: at most one spatial lag of the explanatory variables appears to have a coefficient and one indirect spillover effect that is (weakly) significant; and also the Wald test does not reject.

In the remainder of the analysis we relax the assumption that all spatial weight matrices are the same for each spatial lag. We apply the proposed PWFE estimator

that allows for different distance decay parameters for each spatial weight matrix. Due to the fact that the row normalized exponential distance decay matrix produces the highest log-likelihood function value of the previous four options being considered, the practitioner may use this specification to see whether the performance of the model can be improved by estimating the decay parameter for each single spatial lag.[23]

Column 6 reports the estimation results of PWFE. The log-likelihood function value is -1311.0, which is slightly better than the log-likelihood function value of -1311.4 for the binary contiguity matrix, even though the estimated decay parameters for the spatial lags of internal war and political regime reach the upper bound. For comparison, we run another regression, where we use the estimated spatial weight matrices based on the estimated distance decay parameters from Column 6. We label this estimator as EWFE and the results are reported in Column 7. When the decay parameter equals the upper bound of 10, we found out that most of the elements of the row normalized exponential distance decay matrix are rather small (smaller than 0.005) and can be replaced by zeros. The results reported in Column 7 confirm that this has hardly any effect on the outcomes. Finally, Column 8 reports the estimation results when the spatial lags for which an upper bound of 10 was found are replaced by the original binary contiguity matrix used by YE (labeled EWFEBC). In this case the log-likelihood function value increases further to -1309.23.

The coefficient estimates and the direct and indirect effects do not change when the spatial weight matrices are pre-specified using the estimated values obtained for the decay parameters or when the spatial weight matrices for the spatial lags of internal war and political regime are specified as a binary contiguity matrix. By contrast, the t-values of the spatial lags and of the indirect spillover effects tend to go up slightly. This is because the decay parameters are no longer part of the variance-covariance matrix that is used to determine their significance levels. It reflects the common approach applied by practitioners who pre-specify the spatial weight matrices.

[23] The reason to limit this to one matrix is because of the necessity to consider different starting values, which complicates the analysis. By comparing several experimental designs, we find out that the best option to find the global optimum is to use both a low and a high starting value (0.5 and 3.5) for the decay parameter of each spatial lag. Since we have $K+1$ decay parameters $(\delta, \alpha_1, \ldots, \alpha_K)$, this design is accompanied with 2^{K+1} different sets of starting values, which is 64 for our empirical illustration.

When comparing the direct effects of YE in Column 1 with those in Columns 6 to 8 we see, in line with the Monte Carlo simulation experiments, that they are not sensitive for the choice of weight matrix. In contrast, when comparing the indirect spillover effects, we see different results. For GDP we find a significant indirect spillover effect in Columns 7 and 8 of about -1.4. The direct effect of this variable indicates that if the level of GDP in a country increases, the defense burden in that country also increases though less than proportionally. The spillover effect shows that it also has a diminishing effect on the defense burden of its neighbours. The indirect spillover effect of civil war is also (weakly) significant in Columns 7 and 8 and fluctuates around -0.100. This means that a country sharing a border with a country involved in a civil war does not consider this as a potential threat. Finally, the indirect spillover effect of the political regime is significant and takes values that range from -0.012 to -0.035. A similar and significant effect of -0.030 was found when adopting one common binary contiguity matrix. This outcome demonstrates that the defense burden does not only increase with a higher level of autocracy in the own country, but also with higher levels of autocracy in neighbouring countries.

In sum, three indirect spillover effects appear to be significant. This is in contrast to the model of YE based on one common binary contiguity matrix, where only one is significant. The magnitudes of the indirect spillover effects also differ, in line with the Monte Carlo simulation results in Section 4. The same applies to the coefficients obtained for the spatially lagged regressors, as a result of which the hypothesis $H_0 : \phi = 0$ now also needs to be rejected. We find Wald test statistics with p-values of 0.02.

The overall conclusion of this empirical illustration is that allowing for different spatial ranges of the explanatory variables affects parameter estimates and indirect spillover effects. Stronger and significant empirical evidence is found in favour of SDM and indirect spillover effects.

2.6 Conclusion

The purpose of this chapter is to break the practice of employing one common and pre-specified spatial weight matrix for all spatial lags by parameterizing each spatial weight matrix with a distance decay parameter. We investigate two func-

tional forms of distance decay, the negative exponential and the inverse distance matrices, and two types of normalization, row normalization and normalization of the weight matrix by its largest eigenvalue. The spatial Durbin model is taken as point of departure due to its popularity in empirical research and because its indirect spillover effects are more flexible compared to models that do not contain spatial lags in the explanatory variables. It is found that the extension with distance decay parameters helps to further increase this flexibility, because the distance at which explanatory variables still have an effect on other units, i.e., the spatial range, will no longer be the same but be different from one explanatory variable to another.

The difficulty of parameterizing spatial weight matrices is that the econometric model is no longer linear in its parameters. Our literature overview shows that previous studies were not able to deal with this adequately. We extend the work of Lee and Yu (2010a) and propose a maximum likelihood approach where the response parameters of the model and the distance decay parameters are estimated jointly.

Our Monte Carlo simulation experiment shows that our estimation procedure performs well in terms of bias (average, median, and median absolute value), root mean square error, and p-values both of the parameters and the direct and indirect spillover effects that can be derived using these parameters. The main finding of our Monte Carlo simulation experiment and also of our empirical application is that the performance of an estimator that erroneously assumes a common spatial weight matrix for all spatial lags performs worse. This especially holds for the indirect spillover effects, which are often the main focus of applied spatial econometric research. This is a cause for concern since it reflects the most widely used approach applied by practitioners. By contrast, it supports the approach advocated in this chapter whereby the decay parameters of the spatial lags are estimated rather than set at one particular value for all spatial lags in advance.

In follow-up research we aim at extending the model and consider a spatial lag in the error term and and different specification of the spatial errors.

Table 2.4. Military expenditures according to the binary contiguity matrix (YE), one common and different parameterized weight matrices

		One common decay parameter				Proposed model: ED_rn		
Determinants/W	YE	ED_rn	ED_sn	ID_rn	ID_sn	PWFE	EWFE	EWFEBC
ρ	0.241***	0.257***	0.555***	0.231***	0.454***	0.252***	0.251***	0.247***
	(8.69)	(5.67)	(8.41)	(3.74)	(5.03)	(5.29)	(7.93)	(7.80)
GDP	-0.530***	-0.514***	-0.506***	-0.534***	-0.553***	-0.487***	-0.487***	-0.486***
	(-5.57)	(-5.33)	(-5.30)	(-5.55)	(-5.71)	(-5.04)	(-5.07)	(-5.08)
Population	1.184***	1.213***	0.552	0.816**	0.875**	1.204***	1.204***	1.171***
	(3.18)	(2.93)	(1.40)	(2.07)	(2.27)	(2.86)	(2.98)	(2.90)
International War	0.073*	0.085**	0.079*	0.074*	0.068	0.084**	0.084**	0.076*
	(1.72)	(1.99)	(1.83)	(1.74)	(1.59)	(1.97)	(1.99)	(1.81)
Civil War	0.006	0.007	0.010	0.006	0.002	0.010	0.010	0.007
	(0.37)	(0.47)	(0.65)	(0.37)	(0.14)	(0.66)	(0.67)	(0.47)
Political Regime	-0.016***	-0.018***	-0.017***	-0.017***	-0.018***	-0.018***	-0.018***	-0.018***
	(-3.28)	(-3.54)	(-3.42)	(-3.50)	(-3.46)	(-3.58)	(-3.61)	(-3.50)
W(α_1)*GDP	0.085	-0.041	0.116	0.219	0.712***	-0.944	-0.944*	-0.950*
	(0.53)	(-0.20)	(0.53)	(1.09)	(2.67)	(-0.76)	(-1.77)	(-1.78)
W(α_2)*Population	-0.517	-1.170	3.930*	0.544	-1.399	-1.466	-1.467*	-1.389**
	(-0.84)	(-1.55)	(1.75)	(0.63)	(-0.31)	(-1.43)	(-1.81)	(-1.71)
W(α_3)*International War	-0.055	-0.110	-0.289	-0.080	-0.221	-0.066	-0.066	-0.032
	(-0.70)	(-1.13)	(-1.15)	(-0.93)	(-0.43)	(-1.17)	(-1.46)	(-0.40)
W(α_4)*Civil War	-0.029	-0.081*	-0.097	-0.050	0.045	-0.078	-0.078**	-0.091**
	(-1.03)	(-1.81)	(-0.89)	(-1.39)	(0.47)	(-1.27)	(-2.06)	(-2.42)
W(α_5)*Political Regime	-0.019**	0.000	-0.011	-0.016	-0.022	-0.005	-0.005	-0.023**
	(-2.06)	(0.02)	(-0.36)	(-1.37)	(-0.63)	(-0.58)	(-0.76)	(-2.44)
δ (Spatially lagged dependent variable)		2.022***	2.305***	2.113***	0.766***	2.003***		
		(3.94)	(5.36)	(4.44)	(7.71)	(3.56)		
α_1						0.410		
						(0.74)		
α_2						1,484		
						(1.07)		
α_3						10.00		
						(0.45)		
α_4						2,321		
						(1.12)		
α_5						10.00		
						(0.25)		
DE_GDP	-0.527***	-0.522***	-0.510***	-0.529***	-0.547***	-0.505***	-0.504***	-0.509***
	(-5.59)	(-5.41)	(-5.29)	(-5.50)	(-5.63)	(-5.23)	(-5.11)	(-5.33)
DE_Population	1.170***	1.172***	0.649*	0.852**	0.863**	1.160***	1.180***	1.124***
	(3.11)	(2.91)	(1.69)	(2.22)	(2.30)	(2.83)	(2.92)	(2.76)
DE_Internationl War	0.071*	0.080*	0.073*	0.071*	0.066	0.080*	0.079*	0.077*
	(1.64)	(1.89)	(1.72)	(1.66)	(1.54)	(1.89)	(1.85)	(1.83)
DE_Civil War	0.005	0.003	0.008	0.003	0.003	0.006	0.006	0.003
	(0.34)	(0.21)	(0.50)	(0.21)	(0.17)	(0.40)	(0.39)	(0.18)
DE_Political Regime	-0.018***	-0.018***	-0.018***	-0.018***	-0.018***	-0.019***	-0.018***	-0.019***
	(-3.41)	(-3.52)	(-3.43)	(-3.62)	(-3.48)	(-3.61)	(-3.51)	(-3.85)
IE_GDP	-0.053	-0.224	-0.119	0.120	0.729*	-1.406	-1.414**	-1.387**
	(-0.26)	(-0.86)	(-0.72)	(0.49)	(1.59)	(-0.85)	(-1.96)	(-2.00)
IE_Population	-0.305	-1.113	3.064*	0.916	-1.585	-1.510	-1.565	-1.446
	(-0.40)	(-1.21)	(1.89)	(0.86)	(-0.23)	(-1.14)	(-1.61)	(-1.40)
IE_International War	-0.046	-0.113	-0.177	-0.078	-0.300	-0.057	-0.060	-0.012
	(-0.48)	(-0.92)	(-0.98)	(-0.73)	(-0.37)	(-0.80)	(-1.04)	(-0.13)
IE_Civil War	-0.034	-0.103*	-0.066	-0.061	0.073	-0.097	-0.096*	-0.115**
	(-0.95)	(-1.72)	(-0.83)	(-1.31)	(0.46)	(-1.20)	(-1.94)	(-2.37)
IE_Political Regime	-0.030**	-0.006	-0.015	-0.025*	-0.047	-0.012	-0.012*	-0.035***
	(-2.57)	(-0.35)	(-0.66)	(-1.56)	(-0.80)	(-1.10)	(-1.64)	(-2.68)
Observations	2160	2160	2160	2160	2160	2160	2160	2160
Log-likelihood function value	-1311.39	-1313.66	-1316.37	-1319.19	-1333.49	-1311.02	-1311.02	-1309.23
R-squared	0.702	0.702	0.701	0.700	0.694	0.702	0.702	0.702
H0: $\phi = 0$ (p-value)	7.85 (0.16)	8.39 (0.14)	6.81 (0.23)	6.57 (0.25)	8.85 (0.12)	13.21 (0.02)	13.21 (0.02)	16.52 (0.00)

Notes: YE=Yesilurt and Elhorst (2017), PWFE=Parameterized W matrices and fixed effects, EWFE=Estimated W matrices and fixed effects, EWFEBC=Estimated W or binary contiguity (BC) matrices and fixed effects, ED=Exponential distance decay, ID=Inverse distance decay, rn=row normalized, sn=scalar normalized; *,**,*** significant at respectively 10%, 5% and 1%.

2.A Appendix: Derivatives, information matrix, asymptotic normality and delta method

2.A.1 First-order derivatives of the spatial weight matrix

To generalize the mathematical expressions in the subsequent sections, i.e., to avoid that they all need to be repeated for both types of spatial weight matrices and both types of normalizations, we first introduce the symbol $\boldsymbol{Z}(\varrho)$ representing the first order derivative of the row or scalar normalized spatial weight matrix with respect to ϱ, $\boldsymbol{Z}(\varrho) = \frac{\partial \boldsymbol{W}(\varrho)}{\partial \varrho}$. $\boldsymbol{W}^r$ represents the weight matrix in raw form before it is normalized. Consequently, we have $w_{ij}^r(\varrho) = e^{-d_{ij}\varrho}$ and $w_{ij}(\varrho) = \frac{e^{-d_{ij}\varrho}}{\sum_j e^{-d_{ij}\varrho}}$ for the exponential decay matrix before and after row normalization, respectively. Similarly, we have $w_{ij}^r(\varrho) = {d_{ij}}^{-\varrho}$ and $w_{ij}(\varrho) = \frac{{d_{ij}}^{-\varrho}}{\sum_j {d_{ij}}^{-\varrho}}$ for the inverse distance decay matrix before and after row normalization, respectively.

Under row normalization, the elements of the matrix of first order derivatives for the exponential and inverse distance matrix are

$$z_{ij}(\varrho) = \frac{(\sum_j d_{ij} e^{-d_{ij}\varrho}) e^{-d_{ij}\varrho} - (\sum_j e^{-d_{ij}\varrho}) d_{ij} e^{-d_{ij}\varrho}}{(\sum_j e^{-d_{ij}\varrho})^2}, \tag{2.A.1}$$

$$z_{ij}(\varrho) = \frac{(\sum_j ln(d_{ij}) d_{ij}^{-\varrho}) {d_{ij}}^{-\varrho} - (\sum_j {d_{ij}}^{-\varrho}) ln(d_{ij}) {d_{ij}}^{-\varrho}}{(\sum_j {d_{ij}}^{-\varrho})^2}, \tag{2.A.2}$$

respectively.

If the matrices are scalar normalized by the largest eigenvalue, we respectively have $w_{ij} = e^{-d_{ij}\varrho} / \kappa_{max}^r(\varrho)$ and $w_{ij} = {d_{ij}}^{-\varrho} / \kappa_{max}^r(\varrho)$. In this case, the matrix $\boldsymbol{Z}(\varrho)$ of first order derivatives is easier expressed in full form.

Since $\boldsymbol{W}^r(\varrho)$ is real and symmetric, it has distinct eigenvalues κ_i^r and eigenvectors $\boldsymbol{\tau}_i^r$ $(i = 1, \ldots, N)$ with properties $\boldsymbol{\tau}_i^{r'} \boldsymbol{\tau}_i^r = 1$ and $\partial \kappa_i = \boldsymbol{\tau}_i^{r'} \partial \boldsymbol{W}^r(\varrho) \boldsymbol{\tau}_i$ (Magnus, 1985, Equation 4). Using these properties for the largest eigenvalue and corresponding eigenvector of $\boldsymbol{W}^r(\varrho)$, the matrix of first order derivatives for the exponential and inverse distance decay matrices are

$$\boldsymbol{Z}(\varrho) = \frac{\frac{\partial \boldsymbol{W}^r(\varrho)}{\partial \varrho} \kappa_{max}^r(\varrho) - \boldsymbol{\tau}^{r'}(\varrho) \frac{\partial \boldsymbol{W}^r(\varrho)}{\partial \varrho} \boldsymbol{\tau}(\varrho) \boldsymbol{W}^r(\varrho)}{[\kappa_{max}^r(\varrho)]^2}, \tag{2.A.3}$$

$$\boldsymbol{Z}(\varrho) = \frac{\frac{\partial \boldsymbol{W}^r(\varrho)}{\partial \varrho}\kappa^r_{max}(\varrho) - \boldsymbol{\tau}^{r'}(\varrho)\frac{\partial \boldsymbol{W}^r(\varrho)}{\partial \varrho}\boldsymbol{\tau}^r(\varrho)\boldsymbol{W}^r(\varrho)}{[\kappa^r_{max}(\varrho)]^2}, \tag{2.A.4}$$

respectively, where the typical element of $\frac{\partial \boldsymbol{W}^r(\varrho)}{\partial \varrho} = -d_{ij}e^{-\varrho d_{ij}}$ for the exponential and $\frac{\partial \boldsymbol{W}^r(\varrho)}{\partial \varrho} = -ln(d_{ij})d_{ij}^{-\varrho}$ for the inverse distance decay matrix.

2.A.2 First-order derivatives of the parameters

Without loss of generality, the presentation with respect to the distance decay parameters is limited to δ, α_1 and α_2. This also implies that only two explanatory variables are considered ($K = 2$). For the same reason, time dummies are left aside. The first-order derivatives are

$$\frac{\partial lnL^w(\boldsymbol{\theta})}{\partial \boldsymbol{\zeta}} = \frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}\widetilde{\boldsymbol{v}}_t = \frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}[\widetilde{\boldsymbol{y}}_t - \rho\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta}], \tag{2.A.5}$$

$$\frac{\partial lnL^w(\boldsymbol{\theta})}{\partial \rho} = -T\text{tr}(\boldsymbol{G}) + \frac{1}{\sigma^2}\sum_{t=1}^{T}[(\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta})'\widetilde{\boldsymbol{v}}_t + \widetilde{\boldsymbol{v}}_t'\boldsymbol{G}\widetilde{\boldsymbol{v}}_t], \tag{2.A.6}$$

$$\frac{\partial lnL^w(\boldsymbol{\theta})}{\partial \sigma^2} = -\frac{NT}{2\sigma^2} + \sum_{t=1}^{T}\frac{\widetilde{\boldsymbol{v}}_t'\widetilde{\boldsymbol{v}}_t}{2\sigma^4}, \tag{2.A.7}$$

$$\frac{\partial lnL^w(\boldsymbol{\theta})}{\partial \delta} = -\rho T\text{tr}(\boldsymbol{\Delta}(\delta)) + \frac{\rho}{\sigma^2}\sum_{t=1}^{T}[(\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta})'\widetilde{\boldsymbol{v}}_t + \widetilde{\boldsymbol{v}}_t'\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{v}}_t], \tag{2.A.8}$$

$$\frac{\partial lnL^w(\boldsymbol{\theta})}{\partial \alpha_1} = \sum_{t=1}^{T}\frac{\widetilde{\boldsymbol{v}}_t'\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1}{\sigma^2} = \sum_{t=1}^{T}\frac{(\widetilde{\boldsymbol{y}}_t - \rho\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta})'\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1}{\sigma^2}, \tag{2.A.9}$$

$$\frac{\partial lnL^w(\boldsymbol{\theta})}{\partial \alpha_2} = \sum_{t=1}^{T}\frac{\widetilde{\boldsymbol{v}}_t'\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2}{\sigma^2} = \sum_{t=1}^{T}\frac{(\widetilde{\boldsymbol{y}}_t - \rho\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta})'\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2}{\sigma^2}, \tag{2.A.10}$$

where $\widetilde{\boldsymbol{v}}_t = \widetilde{\boldsymbol{y}}_t - \rho\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta}$. To shorten notation, we further use $\boldsymbol{G} = \boldsymbol{W}(\delta)\boldsymbol{S}^{-1}(\rho,\delta)$ and $\boldsymbol{\Delta}(\delta) = \boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}(\rho,\delta)$.

2.A.3 Second-order derivatives of the parameters

In line with $\boldsymbol{\theta} = (\boldsymbol{\zeta}', \rho, \sigma^2, \boldsymbol{\alpha}')'$, the Hessian matrix can be partitioned into a block-matrix consisting of six rows and column groups with respect to $\boldsymbol{\zeta}'$, ρ, σ^2, δ, α_1, and α_2. Since the Hessian matrix is symmetric, we provide the expressions for the

diagonal and upper-diagonal blocks of this matrix only.

The submatrices in the first row group of the Hessian matrix are

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'} = -\sum_{t=1}^{T}\frac{\widetilde{\boldsymbol{X}}_t^{*'}\widetilde{\boldsymbol{X}}_t^{*}}{\sigma^2}, \tag{2.A.11}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\rho} = -\sum_{t=1}^{T}\frac{\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t}{\sigma^2} = -\sum_{t=1}^{T}\frac{\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{G}\widetilde{\boldsymbol{v}}_t}{\sigma^2}, \tag{2.A.12}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\sigma^2} = -\sum_{t=1}^{T}\frac{\widetilde{\boldsymbol{X}}_t^{*'}\widetilde{\boldsymbol{v}}_t}{\sigma^4}, \tag{2.A.13}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\delta} = -\sum_{t=1}^{T}\frac{\rho\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t}{\sigma^2} = -\sum_{t=1}^{T}\frac{\rho\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \rho\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{v}}_t}{\sigma^2}, \tag{2.A.14}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\alpha_1} = \sum_{t=1}^{T}\frac{(\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1})'\widetilde{\boldsymbol{v}}_t - \widetilde{\boldsymbol{X}}^{*'}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1}{\sigma^2}, \tag{2.A.15}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\alpha_2} = \sum_{t=1}^{T}\frac{(\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2})'\widetilde{\boldsymbol{v}}_t - \widetilde{\boldsymbol{X}}^{*'}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2}{\sigma^2},. \tag{2.A.16}$$

The submatrices in the second row group are

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\rho^2} = -T\mathrm{tr}(\boldsymbol{G}^2) - \sum_{t=1}^{T}\frac{(\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \widetilde{\boldsymbol{v}}_t)'\boldsymbol{G}'\boldsymbol{G}(\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \widetilde{\boldsymbol{v}}_t)}{\sigma^2}, \tag{2.A.17}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\rho\partial\sigma^2} = -\sum_{t=1}^{T}\frac{(\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\widetilde{\boldsymbol{v}}_t}{\sigma^4} = -\sum_{t=1}^{T}\frac{(\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta})'\widetilde{\boldsymbol{v}}_t + \widetilde{\boldsymbol{v}}_t'\boldsymbol{G}\widetilde{\boldsymbol{v}}_t}{\sigma^4}, \tag{2.A.18}$$

$$\begin{aligned}\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\rho\partial\delta} &= -T\mathrm{tr}(\boldsymbol{\Delta}(\delta)) - \rho T\mathrm{tr}(\boldsymbol{G}\boldsymbol{\Delta}(\delta)) \\ &\quad + \sum_{t=1}^{T}\frac{(\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{v}}_t)'\widetilde{\boldsymbol{v}}_t - \rho(\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{G}\widetilde{\boldsymbol{v}}_t)'(\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{v}}_t)}{\sigma^2},\end{aligned} \tag{2.A.19}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\rho\partial\alpha_1} = -\sum_{t=1}^{T}\frac{(\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1}{\sigma^2} = -\sum_{t=1}^{T}\frac{(\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{G}\widetilde{\boldsymbol{v}}_t)'\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1}{\sigma^2}, \tag{2.A.20}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\rho\partial\alpha_2} = -\sum_{t=1}^{T}\frac{(\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2}{\sigma^2} = -\sum_{t=1}^{T}\frac{(\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{G}\widetilde{\boldsymbol{v}}_t)'\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2}{\sigma^2}. \tag{2.A.21}$$

The submatrices in the third row group are

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial(\sigma^2)^2} = \frac{NT}{2\sigma^4} - \sum_{t=1}^{T} \frac{\widetilde{\boldsymbol{v}}_t^{'}\widetilde{\boldsymbol{v}}_t}{\sigma^6}, \tag{2.A.22}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\sigma^2\partial\delta} = \frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\delta\partial\sigma^2} = -\sum_{t=1}^{T} \frac{\rho(\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t)^{'}\widetilde{\boldsymbol{v}}_t}{\sigma^4} = -\sum_{t=1}^{T} \frac{\rho(\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{v}}_t)^{'}\widetilde{\boldsymbol{v}}_t}{\sigma^4}, \tag{2.A.23}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\sigma^2\partial\alpha_1} = \frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\alpha_1\partial\sigma^2} = -\sum_{t=1}^{T} \frac{\widetilde{\boldsymbol{v}}_t^{'}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1}{\sigma^4}, \tag{2.A.24}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\sigma^2\partial\alpha_2} = \frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\alpha_2\partial\sigma^2} = -\sum_{t=1}^{T} \frac{\widetilde{\boldsymbol{v}}_t^{'}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2}{\sigma^4}. \tag{2.A.25}$$

The submatrices in the fourth row group are

$$\begin{aligned}\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\delta^2} &= -\rho T\mathrm{tr}(\boldsymbol{G}) - \rho^2 T\mathrm{tr}(\boldsymbol{\Delta}(\delta)\boldsymbol{\Delta}(\delta)) \\ &\quad + \sum_{t=1}^{T} \frac{\rho(\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{G}\widetilde{\boldsymbol{v}}_t)^{'}\widetilde{\boldsymbol{v}}_t - \rho^2(\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{v}}_t)^{'}(\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{v}}_t)}{\sigma^2},\end{aligned} \tag{2.A.26}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\delta\partial\alpha_1} = -\sum_{t=1}^{T} \frac{\rho(\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{v}}_t)^{'}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1}{\sigma^2}, \tag{2.A.27}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\delta\partial\alpha_2} = -\sum_{t=1}^{T} \frac{\rho(\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta} + \boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{v}}_t)^{'}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t2}\phi_2}{\sigma^2}. \tag{2.A.28}$$

The submatrices in the fifth row group are

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\alpha_1^2} = \sum_{t=1}^{T} \frac{\widetilde{\boldsymbol{v}}_t^{'}\frac{\partial^2\boldsymbol{W}(\alpha_1)}{\partial\alpha_1^2}\widetilde{\boldsymbol{x}}_{t1}\phi_1 - (\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1)^{'}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1}{\sigma^2}, \tag{2.A.29}$$

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\alpha_1\partial\alpha_2} = -\frac{(\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2)^{'}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1}{\sigma^2}. \tag{2.A.30}$$

Finally, the submatrix in the sixth row group is

$$\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\alpha_2^2} = \sum_{t=1}^{T} \frac{\widetilde{\boldsymbol{v}}_t^{'}\frac{\partial^2\boldsymbol{W}(\alpha_2)}{\partial\alpha_2^2}\widetilde{\boldsymbol{x}}_{t2}\phi_2 - (\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2)^{'}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2}{\sigma^2}. \tag{2.A.31}$$

2.A.4 Information matrix

The submatrices in the first row group of the information matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_T}^{w}$ are

$$-E\left[\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}\right] = \frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}\widetilde{\boldsymbol{X}}_t^{*}, \tag{2.A.32}$$

$$-E\left[\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\rho}\right] = \frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta}, \tag{2.A.33}$$

$$-E\left[\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\sigma^2}\right] = \mathbf{0}, \tag{2.A.34}$$

$$-E\left[\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\delta}\right] = \frac{\rho}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta}, \tag{2.A.35}$$

$$-E\left[\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\alpha_1}\right] = \frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1, \tag{2.A.36}$$

$$-E\left[\frac{\partial^2 lnL^w(\boldsymbol{\theta})}{\partial\boldsymbol{\zeta}\partial\alpha_2}\right] = \frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2. \tag{2.A.37}$$

The submatrices in the second row group of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_T}^{w}$ are

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\rho^2}\right] = T\mathrm{tr}[\boldsymbol{G}(\boldsymbol{G}+\boldsymbol{G}')] + \frac{1}{\sigma^2}\sum_{t=1}^{T}(\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta})'\boldsymbol{G}'\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta}, \tag{2.A.38}$$

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\rho\partial\sigma^2}\right] = \frac{T}{\sigma^2}\mathrm{tr}(\boldsymbol{G}), \tag{2.A.39}$$

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\rho\partial\delta}\right] = \rho T\mathrm{tr}[\boldsymbol{\Delta}(\delta)(\boldsymbol{G}+\boldsymbol{G}')] + \frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta})'\boldsymbol{G}'\boldsymbol{\Delta}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta}, \tag{2.A.40}$$

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\rho\partial\alpha_1}\right] = \frac{1}{\sigma^2}\sum_{t=1}^{T}(\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta})'\boldsymbol{G}'\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1, \tag{2.A.41}$$

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\rho\partial\alpha_2}\right] = \frac{1}{\sigma^2}\sum_{t=1}^{T}(\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta})'\boldsymbol{G}'\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2. \tag{2.A.42}$$

The submatrices in the third row group of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_T}^{w}$ are

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial(\sigma^2)^2}\right] = \frac{NT}{2\sigma^4}, \tag{2.A.43}$$

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\sigma^2\partial\delta}\right] = \frac{\rho T}{\sigma^2}\mathrm{tr}(\boldsymbol{\Delta}(\delta)), \tag{2.A.44}$$

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\sigma^2\partial\alpha_1}\right] = \mathbf{0}, \tag{2.A.45}$$

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\sigma^2\partial\alpha_2}\right] = \mathbf{0}. \tag{2.A.46}$$

The submatrices in the fourth row group of $\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T}$ are

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\delta^2}\right] = \rho^2 T\mathrm{tr}[\boldsymbol{\Delta}(\delta)(\boldsymbol{\Delta}(\delta)+\boldsymbol{\Delta}(\delta)')] + \frac{\rho^2}{\sigma^2}\sum_{t=1}^{T}(\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta})'\boldsymbol{\Delta}(\delta)'\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}, \tag{2.A.47}$$

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\delta\partial\alpha_1}\right] = \frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta})'\boldsymbol{\Delta}(\delta)'\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1, \tag{2.A.48}$$

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\delta\partial\alpha_2}\right] = \frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta})'\boldsymbol{\Delta}(\delta)'\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2. \tag{2.A.49}$$

The submatrices in the fifth row group of $\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T}$ are

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\alpha_1^2}\right] = \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1)'\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1, \tag{2.A.50}$$

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\alpha_1\alpha_2}\right] = \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2)'\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{t1}\phi_1. \tag{2.A.51}$$

Finally, the submatrix in the sixth row group is

$$-E\left[\frac{\partial^2 L^w(\boldsymbol{\theta})}{\partial\alpha_2^2}\right] = \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2)'\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{t2}\phi_2. \tag{2.A.52}$$

2.A.5 Asymptotic normality

Lee and Yu (2010a) show in Theorem 2(1) and Lemma A.4 that the asymptotic distribution of the ML estimator (see Section 2.3.4) is given by

$$\sqrt{NT}(\hat{\boldsymbol{\theta}}^w - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Upsilon}^w_{\boldsymbol{\theta}_T}), \tag{2.A.53}$$

where $\boldsymbol{\Upsilon}^w_{\boldsymbol{\theta}_T} = \lim\frac{T}{T-1}(\frac{1}{NT}\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T})^{-1}(\frac{1}{NT}\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T} + \boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T})(\frac{1}{NT}\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T})^{-1}$, and $\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T}$ is specified in Appendix 2.A.4. The matrix $\boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T}$ reads as

$$\boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T} = \frac{(T-1)}{T}\frac{\mu_4 - 3\sigma_0^4}{\sigma_0^4}$$

$$= \begin{pmatrix} \mathbf{0}_{2K\times 2K} & * & * & * & * \\ \mathbf{0}_{1\times 2K} & \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{G})_{ii}^2 & * & * & * \\ \mathbf{0}_{1\times 2K} & \frac{1}{2\sigma^2 N}\mathrm{tr}(\boldsymbol{G}) & \frac{1}{4\sigma^4} & * & * \\ \mathbf{0}_{1\times 2K} & \frac{\rho}{N}\sum_{i=1}^{N}(\boldsymbol{G})_{ii}(\boldsymbol{\Delta}(\delta))_{ii} & \frac{\rho}{2\sigma^2 N}\mathrm{tr}(\boldsymbol{\Delta}(\delta)) & \frac{\rho^2}{N}\sum_{i=1}^{N}(\boldsymbol{\Delta}(\delta))_{ii}^2 & * \\ \mathbf{0}_{2\times 2K} & \mathbf{0}_{2\times 1} & \mathbf{0}_{2\times 1} & \mathbf{0}_{2\times 1} & \mathbf{0}_{2\times 2} \end{pmatrix}, \tag{2.A.54}$$

where μ_4 is the fourth moment of the error term. If the error terms are assumed to be normally distributed, the matrix $\boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T}$ cancels out since $\mu_4 - 3\sigma_0^4 = 0$ under this circumstance. This yields $\mathbf{Y}^w_{\boldsymbol{\theta}_T} = \lim \frac{T}{T-1}(\frac{1}{NT}\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T})^{-1}$.

2.A.6 Direct and indirect spillover effects: the delta method

To draw statistical inferences on the direct and indirect spillover effects, expressions for standard errors are needed. Two methods can be used: bootrapping or the delta method. To save computation time, we use the delta method which is an extension of the method described in Arbia et al. (2020). We depart from $\sqrt{NT}(\hat{\boldsymbol{\theta}}^w - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}(0, \mathbf{Y}^w_{\boldsymbol{\theta}_T})$, derived in Appendix A.5, but instead of $\boldsymbol{\theta}$ and $\mathbf{Y}^w_{\boldsymbol{\theta}_T}$, we consider $\boldsymbol{\varphi} = (\boldsymbol{\beta}', \boldsymbol{\phi}', \rho, \boldsymbol{\alpha}')'$ and $\mathbf{Y}^w_{\boldsymbol{\varphi}_T}$, i.e., after rows and columns for the time dummies and σ^2 have been removed since they are not needed to determine the direct and indirect spillover effects. Applying propositions 1 and 2 and remark 2 in Arbia et al. (2020), we get

$$\sqrt{N}[DE_k(\widehat{\boldsymbol{\varphi}}^w) - DE_k(\boldsymbol{\varphi})] \xrightarrow{d} N[0, \boldsymbol{A}_k^{DE}(\boldsymbol{\varphi})(\mathbf{Y}^w_{\boldsymbol{\varphi}_T})^{-1}\boldsymbol{A}_k^{DE}(\boldsymbol{\varphi})'], \tag{2.A.55}$$

$$\sqrt{N}[IE_k(\widehat{\boldsymbol{\varphi}}^w) - IE_k(\boldsymbol{\varphi})] \xrightarrow{d} N[0, \boldsymbol{A}_k^{IE}(\boldsymbol{\varphi})(\mathbf{Y}^w_{\boldsymbol{\varphi}_T})^{-1}\boldsymbol{A}_k^{IE}(\boldsymbol{\varphi})'], \tag{2.A.56}$$

where $\boldsymbol{A}_k^{DE}(\boldsymbol{\varphi}) = \frac{\partial DE_k(\boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}'}$ and $\boldsymbol{A}_k^{IE}(\boldsymbol{\varphi}) = \frac{\partial IE_k(\boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}'}$. These first order derivatives for the direct effects take the form

$$\frac{\partial DE_k(\boldsymbol{\varphi})}{\partial \beta_k} = \frac{1}{N} tr(\boldsymbol{S}^{-1}(\rho, \delta)), \tag{2.A.57}$$

$$\frac{\partial DE_k(\boldsymbol{\varphi})}{\partial \phi_k} = \frac{1}{N} tr\{\boldsymbol{S}^{-1}(\rho, \delta)\boldsymbol{W}(\alpha_k)\}, \tag{2.A.58}$$

$$\frac{\partial DE_k(\boldsymbol{\varphi})}{\partial \rho} = \frac{1}{N} tr\{\boldsymbol{S}^{-1}(\rho, \delta)\boldsymbol{W}(\delta)\boldsymbol{S}^{-1}(\rho, \delta)\boldsymbol{\Pi}\}, \tag{2.A.59}$$

$$\frac{\partial DE_k(\boldsymbol{\varphi})}{\partial \delta} = \frac{1}{N} tr\{\rho\boldsymbol{S}^{-1}(\rho, \delta)\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}(\rho, \delta)\boldsymbol{\Pi}\}, \tag{2.A.60}$$

$$\frac{\partial DE_k(\boldsymbol{\varphi})}{\partial \alpha_k} = \frac{1}{N} tr\{\boldsymbol{S}^{-1}(\rho, \delta)\boldsymbol{Z}(\alpha_k)\phi_k\}, \tag{2.A.61}$$

and for the indirect effects take the form

$$\frac{\partial IE_k(\boldsymbol{\varphi})}{\partial \beta_k} = \frac{1}{N}\boldsymbol{\iota}_N' \boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{\iota}_N - \frac{1}{N} tr(\boldsymbol{S}^{-1}(\rho,\delta)), \tag{2.A.62}$$

$$\frac{\partial IE_k(\boldsymbol{\varphi})}{\partial \gamma_k} = \frac{1}{N}\boldsymbol{\iota}_N' \{\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{W}(\alpha_k)\}\boldsymbol{\iota}_N - \frac{1}{N} tr\{\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{W}(\alpha_k)\}, \tag{2.A.63}$$

$$\frac{\partial IE_k(\boldsymbol{\varphi})}{\partial \rho} = \frac{1}{N}\boldsymbol{\iota}_N' \{\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{W}(\delta)\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{\Pi}\}\boldsymbol{\iota}_N - \frac{1}{N} tr\{\boldsymbol{S}^{-1}\boldsymbol{W}(\delta)\boldsymbol{S}^{-1}\boldsymbol{\Pi}\}, \tag{2.A.64}$$

$$\frac{\partial IE_k(\boldsymbol{\varphi})}{\partial \delta} = \frac{1}{N}\boldsymbol{\iota}_N' \{\rho\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{\Pi}\}\boldsymbol{\iota}_N - \frac{1}{N} tr\{\rho\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{\Pi}\}, \tag{2.A.65}$$

$$\frac{\partial IE_k(\boldsymbol{\varphi})}{\partial \alpha_k} = \frac{1}{N}\boldsymbol{\iota}_N' \{\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{Z}(\alpha_k)\phi_k\}\boldsymbol{\iota}_N' - \frac{1}{N} tr\{\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{Z}(\alpha_k)\phi_k\}, \tag{2.A.66}$$

where $\boldsymbol{\Pi} = \boldsymbol{I}_N \cdot \beta_k + \boldsymbol{W}(\alpha_k) \cdot \phi_k$. Contrary to Arbia et al. (2020), our effect estimates contain the terms $\boldsymbol{Z}(\alpha_k)$, which represent the first order derivative of the spatial weight matrices with respect to α_k. The mathematical expressions of this derivative for the inverse distance and exponential decay matrix and for both normalizations can be found in Appendix 2.A.1.

2.A.7 Matlab routines and additional simulation results

Appendix 2.A.7 contains a description of the developed estimation routine and additional simulation results.

PWFE: A routine for practitioners

To be able to estimate the parameters of the model set out in this paper, we developed a Matlab routine entitled PWFE with the following options:

1. Type of spatial weight matrix: exponential distance decay matrix (Edist) or inverse distance matrix (Idist).

2. Type of normalization of the spatial weight matrix: row normalization (rsn) or scalar normalization by the largest eigenvalue (men).

3. Type of model: different distance decay parameters for each spatial lag (multi), one common distance decay parameter for all spatial lags (same), and one

common distance decay parameter for all spatial lags in the explanatory variables, but not the dependent variable (one).

4. Method to determine significance levels of the direct and indirect spillover effects: the recommended delta method (del) or the bootstrap method based on LeSage and Pace (2009) (bt).

5. gamma0, lb and ub are respectively the starting values, the lower bounds and the upper bounds on ρ and the distance decay parameters when using row normalization, and on the distance decay parameters when using scalar normalization by the largest eigenvalue.

Additional simulation results

Case I represents the basic run: $\beta_1 = -1, \beta_2 = 0.2, \phi_1 = 1.5, \phi_2 = -0.3, \rho = 0.5, \sigma^2 = 1, \delta = 2, \alpha_1 = 1.5, \alpha_2 = 3$. The main text contains the results for the exponential distance decay matrix. The section also reports the results for the inverse distance matrix. In general, the results are similar to the results of the exponential distance decay matrix.

Table A.1. Simulation results for Case I: $\rho = 0.5$, $\delta = 2$, $\alpha_1 = 1.5$, $\alpha_2 = 3$

				Inverse distance: row normalized										Inverse distance: scalar normalized									
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)
WFE	**5**	**200**	**Bias**	0.0120	-0.0060	0.4470	-0.2655	-0.3788	0.0663				0.0113	-0.0069	0.3639	-0.3393	-0.4143	0.0517			
			RMSE	0.0141	0.0118	0.4778	0.3322	0.4292	0.0880				0.0135	0.0123	0.4067	0.4033	0.4657	0.0770			
			Mbias	0.0117	-0.0057	0.4565	-0.2559	-0.3657	0.0662				0.0109	-0.0066	0.3790	-0.3245	-0.3922	0.0506			
			Mabias	0.0117	0.0082	0.4565	0.2567	0.3657	0.0667				0.0109	0.0086	0.3798	0.3245	0.3922	0.0544			
TWFE	**5**	**200**	**Bias**	-0.0002	0.0001	-0.0091	0.0004	-0.0087	-0.0107				-0.0002	0.0001	-0.0100	-0.0001	-0.0099	-0.0107			
			RMSE	0.0071	0.0098	0.0700	0.0335	0.0546	0.0532				0.0071	0.0098	0.0731	0.0361	0.0573	0.0533			
			Mbias	-0.0003	0.0000	-0.0064	0.0001	-0.0062	-0.0131				-0.0003	0.0000	-0.0092	0.0010	-0.0079	-0.0127			
			Mabias	0.0049	0.0066	0.0474	0.0214	0.0362	0.0375				0.0049	0.0066	0.0486	0.0237	0.0368	0.0378			
PWFE	**5**	**200**	**Bias**	0.0013	-0.0005	0.0614	-0.0163	-0.0694	-0.0136	0.1488	-0.0456	0.1735	0.0003	0.0000	0.0170	-0.0048	-0.0192	-0.0136	-0.0182	-0.0139	0.2566
			RMSE	0.0080	0.0099	0.2597	0.0874	0.1549	0.0540	0.5797	0.2049	1.0035	0.0074	0.0099	0.1845	0.0718	0.0642	0.0538	0.3142	0.1847	1.0528
			Mbias	0.0010	-0.0006	0.0374	-0.0050	-0.0532	-0.0159	0.0745	-0.0513	-0.0580	0.0002	-0.0004	0.0141	0.0011	-0.0169	-0.0158	-0.0224	-0.0196	0.0092
			Mabias	0.0052	0.0067	0.1400	0.0547	0.0855	0.0384	0.2401	0.1331	0.4735	0.0050	0.0066	0.1240	0.0502	0.0412	0.0380	0.2043	0.1363	0.4358
WFE	**5**	**800**	**Bias**	0.0051	-0.0048	0.3512	-0.6169	-0.4439	0.0646				0.0045	-0.0043	0.3095	-0.5893	-0.4402	0.0545			
			RMSE	0.0063	0.0072	0.3748	0.6427	0.4707	0.0706				0.0058	0.0068	0.3375	0.6152	0.4682	0.0612			
			Mbias	0.0049	-0.0051	0.3517	-0.6127	-0.4385	0.0644				0.0044	-0.0046	0.3090	-0.5847	-0.4341	0.0541			
			Mabias	0.0049	0.0055	0.3517	0.6127	0.4385	0.0644				0.0044	0.0050	0.3090	0.5847	0.4341	0.0541			
TWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0070	-0.0017	-0.0042	0.0002				0.0000	0.0000	-0.0075	-0.0018	-0.0047	0.0002			
			RMSE	0.0036	0.0052	0.0555	0.0189	0.0348	0.0261				0.0035	0.0052	0.0589	0.0201	0.0376	0.0261			
			Mbias	0.0000	-0.0003	-0.0067	-0.0016	-0.0031	-0.0009				0.0000	-0.0003	-0.0073	-0.0020	-0.0028	-0.0010			
			Mabias	0.0023	0.0034	0.0366	0.0122	0.0232	0.0166				0.0023	0.0034	0.0399	0.0127	0.0247	0.0165			
PWFE	**5**	**800**	**Bias**	0.0002	-0.0001	0.0089	-0.0071	-0.0258	-0.0007	0.0490	-0.0082	0.0361	0.0000	0.0000	-0.0100	-0.0047	-0.0074	-0.0007	0.0062	0.0042	0.0572
			RMSE	0.0037	0.0052	0.1318	0.0468	0.0898	0.0262	0.1901	0.0992	0.3615	0.0036	0.0052	0.1027	0.0467	0.0411	0.0262	0.1342	0.0884	0.3765
			Mbias	0.0001	-0.0004	-0.0003	-0.0020	-0.0229	-0.0019	0.0373	-0.0087	-0.0011	0.0000	-0.0003	-0.0101	-0.0023	-0.0079	-0.0016	0.0096	0.0019	0.0201
			Mabias	0.0024	0.0034	0.0840	0.0314	0.0578	0.0168	0.1106	0.0659	0.2369	0.0023	0.0034	0.0698	0.0318	0.0282	0.0167	0.0893	0.0594	0.2391

Table A.2. Direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 for Case I using the parameter estimates in Table A.1

Settings				Inverse distance: row normalized				Inverse distance: scalar normalized			
	T	N		DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Bias**	0.0003	0.0017	0.1813	-0.2062	0.0001	-0.0001	0.0911	-0.2898
			RMSE	0.0074	0.0102	0.4750	0.2716	0.0073	0.0099	0.4091	0.3377
			Mbias	-0.0001	0.0018	0.1251	-0.2176	-0.0002	-0.0001	0.0463	-0.3015
			Mabias	0.0050	0.0066	0.2714	0.2214	0.0050	0.0066	0.2583	0.3017
TWFE	**5**	**200**	**Bias**	-0.0004	0.0003	-0.0153	0.0028	-0.0004	0.0002	-0.0168	0.0023
			RMSE	0.0071	0.0099	0.2134	0.0668	0.0071	0.0099	0.2094	0.0649
			Mbias	-0.0007	0.0003	-0.0281	0.0032	-0.0008	0.0001	-0.0310	0.0024
			Mabias	0.0049	0.0066	0.1434	0.0442	0.0049	0.0065	0.1387	0.0442
PWFE	**5**	**200**	**Bias**	-0.0003	0.0004	0.0122	-0.0033	-0.0002	0.0004	0.0229	-0.0037
			RMSE	0.0073	0.0099	0.4117	0.1402	0.0073	0.0099	0.3950	0.1295
			Mbias	-0.0005	0.0003	-0.0061	0.0117	-0.0004	0.0003	-0.0020	0.0050
			Mabias	0.0051	0.0065	0.2695	0.0957	0.0051	0.0064	0.2606	0.0925
WFE	**5**	**800**	**Bias**	-0.0005	0.0008	-0.0591	-0.5646	-0.0005	0.0006	-0.0455	-0.5470
			RMSE	0.0037	0.0053	0.2479	0.5820	0.0037	0.0053	0.2393	0.5649
			Mbias	-0.0005	0.0005	-0.0712	-0.5649	-0.0005	0.0003	-0.0555	-0.5480
			Mabias	0.0024	0.0035	0.1786	0.5649	0.0024	0.0035	0.1671	0.5480
TWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0126	-0.0024	0.0000	0.0000	-0.0124	-0.0025
			RMSE	0.0036	0.0052	0.1615	0.0392	0.0036	0.0052	0.1617	0.0389
			Mbias	0.0000	-0.0003	-0.0169	-0.0010	0.0000	-0.0002	-0.0175	-0.0027
			Mabias	0.0024	0.0035	0.1077	0.0271	0.0024	0.0035	0.1111	0.0274
PWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0217	-0.0060	0.0000	0.0000	-0.0208	-0.0073
			RMSE	0.0036	0.0053	0.2365	0.0873	0.0036	0.0052	0.2314	0.0867
			Mbias	0.0000	-0.0002	-0.0275	-0.0025	0.0000	-0.0002	-0.0225	-0.0037
			Mabias	0.0024	0.0036	0.1624	0.0580	0.0024	0.0035	0.1559	0.0586

Table A.3. Case I: Mean and standard deviation of the p-values of the parameters (Table A.1) and the direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 (Table A.2)

Settings				Inverse distance: row normalized											
	T	N		β_1**(-1)**	β_2**(0.2)**	ϕ_1**(1.5)**	ϕ_2**(-0.3)**	ρ**(0.5)**	δ**(2)**	α_1**(1.5)**	α_2**(3)**	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.2133	0.4387	0.0308	0.1919	0.0708				0.4917	0.4908	0.2685	0.2430
			Std	0.2483	0.3019	0.1153	0.2767	0.1740				0.2866	0.2877	0.2968	0.2947
TWFE	**5**	**200**	**Mean**	0.4918	0.4938	0.4845	0.4896	0.4800				0.4961	0.4935	0.4945	0.5010
			Std	0.2814	0.2925	0.2906	0.2920	0.2938				0.2807	0.2926	0.2939	0.2929
PWFE	**5**	**200**	**Mean**	0.4905	0.4887	0.4620	0.4494	0.4348	0.4785	0.4382	0.4759	0.4961	0.4941	0.4593	0.4428
			Std	0.2960	0.2916	0.2970	0.3022	0.3109	0.2839	0.2984	0.2915	0.2897	0.2942	0.3092	0.3056
WFE	**5**	**800**	**Mean**	0.2788	0.3707	0.0208	0.0005	0.0031				0.5069	0.5073	0.2514	0.0015
			Std	0.2867	0.2987	0.0903	0.0079	0.0218				0.2997	0.2910	0.2975	0.0170
TWFE	**5**	**800**	**Mean**	0.5092	0.5015	0.4953	0.4996	0.4881				0.5045	0.4990	0.4951	0.4932
			Std	0.2979	0.2903	0.2917	0.2909	0.2877				0.3001	0.2902	0.2943	0.2853
PWFE	**5**	**800**	**Mean**	0.4996	0.5008	0.4686	0.4834	0.4518	0.4792	0.4587	0.4848	0.5026	0.4983	0.4731	0.4837
			Std	0.2953	0.2893	0.2978	0.2924	0.3016	0.2873	0.3025	0.2911	0.3009	0.2900	0.3018	0.2973
				Inverse distance: scalar normalized											
	T	**N**		β_1**(-1)**	β_2**(0.2)**	ϕ_1**(1.5)**	ϕ_2**(-0.3)**	ρ**(0.5)**	δ**(2)**	α_1**(1.5)**	α_2**(3)**	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.2308	0.4224	0.0796	0.1372	0.0548				0.4909	0.4961	0.2698	0.1579
			Std	0.2560	0.3013	0.1868	0.2382	0.1476				0.2848	0.2876	0.2983	0.2520
TWFE	**5**	**200**	**Mean**	0.4920	0.4955	0.4836	0.4917	0.4819				0.4948	0.4936	0.4814	0.4996
			Std	0.2817	0.2927	0.2893	0.2914	0.2963				0.2806	0.2911	0.2926	0.2936
PWFE	**5**	**200**	**Mean**	0.4990	0.4928	0.5032	0.5060	0.6553	0.5232	0.4724	0.5393	0.4996	0.4923	0.4858	0.4813
			Std	0.2853	0.2906	0.2879	0.2904	0.2248	0.2745	0.2907	0.2637	0.2903	0.2900	0.3020	0.2984
WFE	**5**	**800**	**Mean**	0.3138	0.3923	0.0431	0.0011	0.0039				0.5052	0.5059	0.2546	0.0022
			Std	0.2888	0.2999	0.1344	0.0165	0.0296				0.2998	0.2895	0.2955	0.0250
TWFE	**5**	**800**	**Mean**	0.5093	0.5011	0.4871	0.5006	0.4859				0.5042	0.4990	0.4848	0.4909
			Std	0.2980	0.2908	0.2920	0.2911	0.2916				0.3000	0.2900	0.2991	0.2832
PWFE	**5**	**800**	**Mean**	0.5065	0.5024	0.5430	0.4966	0.6897	0.5617	0.5189	0.4989	0.5025	0.4988	0.4823	0.4859
			Std	0.2946	0.2901	0.2785	0.2893	0.2065	0.2694	0.2880	0.2835	0.3007	0.2896	0.3029	0.2893

We also run several simulations to investigate the parameter spaces of the spatial autocorrelation parameter and the distance decay parameters in greater detail.

Specifically, we modify ρ, δ, α_1, and α_2. This yields 6 additional parameter configurations, which are summarized in Table A.4. Changes with respect to Case I are Case II: $\rho = 0.25$, Case III: $\rho = -0.25$, Case IV: $\rho = -0.01$, Case V: $\alpha_2 = 1$, Case VI: $\alpha_2 = 0.5$, and Case VII: $\alpha_1 = 10$.

Table A.4. Summary of cases

	ρ	δ	α_1	α_2
Case I	0.5	2	1.5	3
Case II	0.25	2	1.5	3
Case III	-0.25	2	1.5	3
Case IV	0.01	2	1.5	3
Case V	0.5	2	1.5	1
Case VI	0.5	2	1.5	0.5
Case VII	0.5	2	10	3

Table A.5. Simulation results for Case II: $\rho = 0.25$, $\delta = 2$, $\alpha_1 = 1.5$, $\alpha_2 = 3$

				Negative exponential: row normalized									Negative exponential: scalar normalized								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.25)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.25)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)
WFE	**5**	**200**	**Bias**	0.0499	-0.0133	0.3540	-0.1470	-0.4342	0.0767				0.0432	-0.0130	0.3342	-0.1764	-0.4829	0.0611			
			RMSE	0.0507	0.0168	0.3564	0.1597	0.4386	0.0972				0.0441	0.0165	0.3376	0.1898	0.4892	0.0850			
			Mbias	0.0498	-0.0133	0.3541	-0.1456	-0.4325	0.0748				0.0429	-0.0130	0.3340	-0.1741	-0.4807	0.0594			
			Mabias	0.0498	0.0134	0.3541	0.1456	0.4325	0.0750				0.0429	0.0131	0.3340	0.1741	0.4807	0.0617			
TWFE	**5**	**200**	**Bias**	-0.0001	0.0002	-0.0014	0.0003	-0.0021	-0.0109				0.0000	0.0002	-0.0020	0.0002	-0.0031	-0.0108			
			RMSE	0.0080	0.0098	0.0246	0.0250	0.0288	0.0532				0.0079	0.0098	0.0282	0.0266	0.0336	0.0532			
			Mbias	-0.0002	0.0001	-0.0012	-0.0004	-0.0018	-0.0133				0.0002	0.0000	-0.0022	0.0001	-0.0035	-0.0133			
			Mabias	0.0052	0.0066	0.0171	0.0160	0.0202	0.0378				0.0055	0.0066	0.0202	0.0170	0.0229	0.0379			
PWFE	**5**	**200**	**Bias**	0.0016	-0.0001	0.0070	-0.0020	-0.0113	-0.0145	0.2183	-0.0126	0.1667	0.0009	0.0001	0.0021	-0.0021	-0.0048	-0.0138	0.0151	-0.0072	0.1930
			RMSE	0.0110	0.0100	0.0504	0.0358	0.0679	0.0541	0.9893	0.1061	0.8656	0.0090	0.0099	0.0447	0.0384	0.0383	0.0538	0.5926	0.0961	0.9393
			Mbias	0.0011	-0.0004	0.0058	0.0006	-0.0096	-0.0168	0.0876	-0.0140	0.0063	0.0007	-0.0002	0.0000	-0.0001	-0.0041	-0.0165	0.0151	-0.0012	0.0371
			Mabias	0.0071	0.0067	0.0316	0.0241	0.0452	0.0381	0.3696	0.0698	0.4919	0.0060	0.0067	0.0289	0.0260	0.0250	0.0379	0.3736	0.0620	0.5235
WFE	**5**	**800**	**Bias**	0.0445	-0.0112	0.3908	-0.1956	-0.4499	0.0820				0.0399	-0.0110	0.4028	-0.2073	-0.4649	0.0720			
			RMSE	0.0448	0.0124	0.3915	0.1987	0.4515	0.0872				0.0402	0.0122	0.4037	0.2107	0.4669	0.0776			
			Mbias	0.0446	-0.0115	0.3897	-0.1931	-0.4503	0.0817				0.0398	-0.0113	0.4017	-0.2059	-0.4643	0.0721			
			Mabias	0.0446	0.0115	0.3897	0.1931	0.4503	0.0817				0.0398	0.0113	0.4017	0.2059	0.4643	0.0721			
TWFE	**5**	**800**	**Bias**	0.0001	0.0000	-0.0007	-0.0012	-0.0009	0.0001				0.0001	0.0000	-0.0007	-0.0013	-0.0007	0.0001			
			RMSE	0.0040	0.0052	0.0145	0.0136	0.0171	0.0260				0.0040	0.0052	0.0158	0.0143	0.0186	0.0261			
			Mbias	0.0001	-0.0003	-0.0010	-0.0016	-0.0006	-0.0012				0.0000	-0.0003	-0.0007	-0.0017	0.0000	-0.0012			
			Mabias	0.0026	0.0034	0.0099	0.0088	0.0115	0.0166				0.0025	0.0034	0.0111	0.0092	0.0126	0.0165			
PWFE	**5**	**800**	**Bias**	0.0002	0.0000	-0.0006	-0.0014	-0.0007	-0.0010	0.0331	0.0018	0.0592	0.0001	0.0000	-0.0009	-0.0015	-0.0003	-0.0007	0.0181	0.0019	0.0633
			RMSE	0.0051	0.0052	0.0286	0.0198	0.0348	0.0261	0.2771	0.0559	0.3992	0.0043	0.0052	0.0252	0.0210	0.0196	0.0260	0.2599	0.0475	0.4132
			Mbias	0.0001	-0.0003	-0.0015	0.0000	-0.0010	-0.0021	0.0194	0.0002	0.0052	0.0000	-0.0003	-0.0012	0.0000	-0.0005	-0.0017	0.0146	0.0015	0.0168
			Mabias	0.0034	0.0033	0.0189	0.0133	0.0230	0.0165	0.1736	0.0359	0.2595	0.0027	0.0034	0.0168	0.0138	0.0130	0.0166	0.1752	0.0327	0.2666
				Inverse distance: row normalized									**Inverse distance: scalar normalized**								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.25)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.25)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)
WFE	**5**	**200**	**Bias**	0.0176	-0.0064	0.5422	-0.3069	-0.8186	0.0489				0.0161	-0.0073	0.4524	-0.3880	-0.8044	0.0365			
			RMSE	0.0192	0.0120	0.5699	0.3654	0.8570	0.0757				0.0179	0.0125	0.4907	0.4421	0.8468	0.0677			
			Mbias	0.0175	-0.0063	0.5501	-0.3034	-0.8200	0.0491				0.0159	-0.0073	0.4600	-0.3815	-0.7955	0.0358			
			Mabias	0.0175	0.0085	0.5501	0.3034	0.8200	0.0531				0.0159	0.0090	0.4600	0.3815	0.7955	0.0451			
TWFE	**5**	**200**	**Bias**	-0.0002	0.0001	-0.0078	0.0005	-0.0077	-0.0109				-0.0002	0.0001	-0.0089	0.0001	-0.0090	-0.0109			
			RMSE	0.0072	0.0098	0.0710	0.0333	0.0607	0.0532				0.0072	0.0098	0.0750	0.0358	0.0646	0.0532			
			Mbias	-0.0003	0.0001	-0.0065	-0.0003	-0.0068	-0.0132				-0.0003	0.0001	-0.0079	0.0012	-0.0062	-0.0128			
			Mabias	0.0050	0.0066	0.0482	0.0216	0.0406	0.0375				0.0049	0.0066	0.0498	0.0237	0.0418	0.0378			
PWFE	**5**	**200**	**Bias**	0.0033	-0.0011	0.1137	-0.0165	-0.1514	-0.0140	0.1407	-0.0946	0.1581	0.0004	0.0000	0.0122	-0.0047	-0.0179	-0.0141	-0.0602	-0.0096	0.2617
			RMSE	0.0099	0.0100	0.3008	0.0916	0.3355	0.0538	1.9500	0.2376	0.9937	0.0077	0.0099	0.1825	0.0721	0.0789	0.0538	0.7821	0.1788	1.0627
			Mbias	0.0028	-0.0014	0.0706	-0.0020	-0.0578	-0.0166	-0.1722	-0.0927	-0.0727	0.0002	-0.0003	0.0105	0.0008	-0.0134	-0.0156	-0.0402	-0.0215	0.0167
			Mabias	0.0062	0.0067	0.1540	0.0544	0.1156	0.0384	0.7285	0.1576	0.4888	0.0053	0.0067	0.1211	0.0511	0.0470	0.0380	0.4374	0.1372	0.4373
WFE	**5**	**800**	**Bias**	0.0084	-0.0046	0.5957	-0.6687	-1.0413	0.0515				0.0075	-0.0040	0.5523	-0.6444	-0.9970	0.0441			
			RMSE	0.0092	0.0070	0.6142	0.6956	1.0575	0.0587				0.0083	0.0066	0.5733	0.6716	1.0158	0.0522			
			Mbias	0.0082	-0.0050	0.5964	-0.6751	-1.0742	0.0512				0.0074	-0.0043	0.5521	-0.6452	-1.0109	0.0444			
			Mabias	0.0082	0.0053	0.5964	0.6751	1.0742	0.0512				0.0074	0.0049	0.5521	0.6452	1.0109	0.0444			
TWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0063	-0.0016	-0.0039	0.0001				0.0000	0.0000	-0.0066	-0.0018	-0.0041	0.0001			
			RMSE	0.0036	0.0052	0.0570	0.0189	0.0389	0.0261				0.0036	0.0052	0.0605	0.0201	0.0419	0.0261			
			Mbias	0.0000	-0.0003	-0.0063	-0.0017	-0.0028	-0.0011				0.0000	-0.0003	-0.0065	-0.0020	-0.0028	-0.0011			
			Mabias	0.0023	0.0034	0.0373	0.0121	0.0253	0.0164				0.0023	0.0034	0.0397	0.0127	0.0283	0.0164			
PWFE	**5**	**800**	**Bias**	0.0005	-0.0002	0.0263	-0.0078	-0.0502	-0.0007	0.0635	-0.0220	0.0300	0.0001	0.0000	-0.0099	-0.0047	-0.0057	-0.0007	-0.0022	0.0043	0.0572
			RMSE	0.0040	0.0053	0.1736	0.0484	0.1964	0.0262	0.8617	0.1243	0.3683	0.0037	0.0052	0.1057	0.0467	0.0462	0.0261	0.3100	0.0883	0.3773
			Mbias	0.0003	-0.0004	-0.0007	-0.0025	-0.0243	-0.0018	0.0294	-0.0103	-0.0025	0.0000	-0.0003	-0.0091	-0.0027	-0.0053	-0.0018	0.0156	0.0029	0.0183
			Mabias	0.0025	0.0035	0.0872	0.0313	0.0776	0.0165	0.2624	0.0675	0.2428	0.0024	0.0034	0.0708	0.0316	0.0303	0.0163	0.1749	0.0582	0.2395

Table A.6. Simulation results for Case III: $\rho = -0.25$, $\delta = 2$, $\alpha_1 = 1.5$, $\alpha_2 = 3$

				Negative exponential: row normalized									Negative exponential: scalar normalized								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(-0.25)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(-0.25)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)
WFE	5	200	**Bias**	0.0584	-0.0135	0.3890	-0.1408	-0.7500	0.1019				0.0456	-0.0116	0.3850	-0.1634	-0.7500	0.0893			
			RMSE	0.0588	0.0166	0.3911	0.1533	0.7500	0.1183				0.0461	0.0151	0.3875	0.1759	0.7500	0.1074			
			Mbias	0.0583	-0.0135	0.3891	-0.1411	-0.7500	0.0999				0.0454	-0.0116	0.3842	-0.1631	-0.7500	0.0887			
			Mabias	0.0583	0.0135	0.3891	0.1411	0.7500	0.0999				0.0454	0.0117	0.3842	0.1631	0.7500	0.0887			
TWFE	5	200	**Bias**	-0.0002	0.0002	-0.0012	0.0003	-0.0012	-0.0110				-0.0001	0.0002	-0.0018	0.0003	-0.0020	-0.0110			
			RMSE	0.0085	0.0098	0.0261	0.0250	0.0310	0.0531				0.0085	0.0099	0.0299	0.0266	0.0368	0.0531			
			Mbias	-0.0002	0.0000	-0.0011	-0.0005	-0.0008	-0.0141				0.0000	0.0000	-0.0022	-0.0001	-0.0025	-0.0135			
			Mabias	0.0057	0.0066	0.0181	0.0159	0.0216	0.0376				0.0059	0.0067	0.0212	0.0173	0.0248	0.0379			
PWFE	5	200	**Bias**	0.0015	0.0000	0.0065	-0.0027	-0.0298	-0.0158	0.0283	-0.0055	0.1572	-0.0010	0.0006	-0.0040	-0.0034	-0.0064	-0.0138	-0.0776	0.0174	0.1967
			RMSE	0.0130	0.0101	0.0444	0.0355	0.0993	0.0543	0.7737	0.0973	0.8392	0.0105	0.0100	0.0441	0.0381	0.0443	0.0536	0.5214	0.1060	0.9163
			Mbias	0.0008	-0.0004	0.0054	-0.0007	-0.0227	-0.0178	-0.0896	-0.0096	0.0218	-0.0003	0.0003	-0.0042	-0.0010	-0.0078	-0.0161	-0.0438	-0.0013	0.0367
			Mabias	0.0086	0.0070	0.0284	0.0239	0.0601	0.0391	0.3574	0.0628	0.4777	0.0069	0.0065	0.0277	0.0254	0.0283	0.0381	0.3027	0.0587	0.5199
WFE	5	800	**Bias**	0.0487	-0.0088	0.4447	-0.2002	-0.7500	0.1130				0.0415	-0.0085	0.4706	-0.2144	-0.7500	0.1044			
			RMSE	0.0488	0.0102	0.4453	0.2031	0.7500	0.1169				0.0417	0.0100	0.4712	0.2174	0.7500	0.1085			
			Mbias	0.0487	-0.0090	0.4434	-0.1981	-0.7500	0.1122				0.0415	-0.0087	0.4700	-0.2130	-0.7500	0.1043			
			Mabias	0.0487	0.0090	0.4434	0.1981	0.7500	0.1122				0.0415	0.0088	0.4700	0.2130	0.7500	0.1043			
TWFE	5	800	**Bias**	0.0001	0.0000	-0.0007	-0.0012	-0.0009	0.0000				0.0001	0.0000	-0.0007	-0.0013	-0.0006	0.0000			
			RMSE	0.0043	0.0052	0.0154	0.0136	0.0187	0.0261				0.0043	0.0052	0.0168	0.0143	0.0205	0.0261			
			Mbias	0.0001	-0.0004	-0.0009	-0.0015	-0.0006	-0.0009				0.0000	-0.0004	-0.0004	-0.0019	-0.0004	-0.0009			
			Mabias	0.0029	0.0034	0.0105	0.0088	0.0125	0.0162				0.0027	0.0034	0.0116	0.0092	0.0138	0.0162			
PWFE	5	800	**Bias**	0.0002	0.0000	-0.0002	-0.0017	-0.0059	-0.0011	0.0030	0.0026	0.0544	-0.0002	0.0001	-0.0016	-0.0017	-0.0010	-0.0007	-0.0151	0.0051	0.0631
			RMSE	0.0062	0.0053	0.0266	0.0198	0.0485	0.0262	0.2964	0.0536	0.3957	0.0049	0.0052	0.0259	0.0210	0.0214	0.0261	0.2388	0.0489	0.4130
			Mbias	0.0001	-0.0001	-0.0005	-0.0001	-0.0047	-0.0015	-0.0201	0.0018	0.0041	-0.0001	-0.0003	-0.0014	-0.0001	-0.0013	-0.0016	-0.0109	0.0007	0.0153
			Mabias	0.0042	0.0034	0.0177	0.0132	0.0312	0.0164	0.1857	0.0350	0.2551	0.0030	0.0034	0.0172	0.0138	0.0144	0.0162	0.1519	0.0305	0.2694
				Inverse distance: row normalized									**Inverse distance: scalar normalized**								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(-0.25)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(-0.25)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)
WFE	5	200	**Bias**	0.0116	0.0018	0.9930	-0.0725	-0.7500	0.1027				0.0120	0.0000	0.8859	-0.1475	-0.7500	0.0802			
			RMSE	0.0136	0.0099	0.9988	0.1441	0.7500	0.1188				0.0139	0.0097	0.8932	0.1947	0.7500	0.0993			
			Mbias	0.0116	0.0017	0.9954	-0.0754	-0.7500	0.1025				0.0120	-0.0001	0.8907	-0.1518	-0.7500	0.0793			
			Mabias	0.0116	0.0066	0.9954	0.1017	0.7500	0.1025				0.0120	0.0063	0.8907	0.1539	0.7500	0.0796			
TWFE	5	200	**Bias**	-0.0001	0.0002	-0.0058	0.0006	-0.0058	-0.0111				-0.0001	0.0001	-0.0070	0.0003	-0.0072	-0.0111			
			RMSE	0.0075	0.0098	0.0707	0.0333	0.0661	0.0531				0.0074	0.0098	0.0756	0.0356	0.0716	0.0531			
			Mbias	-0.0001	0.0001	-0.0047	-0.0001	-0.0049	-0.0136				-0.0002	0.0001	-0.0070	0.0008	-0.0049	-0.0136			
			Mabias	0.0051	0.0066	0.0471	0.0218	0.0444	0.0378				0.0050	0.0067	0.0501	0.0235	0.0460	0.0377			
PWFE	5	200	**Bias**	0.0023	-0.0005	0.0467	-0.0201	-0.2044	-0.0171	-0.0872	-0.0256	0.1346	0.0002	0.0002	0.0127	-0.0048	-0.0305	-0.0143	-0.0459	-0.0031	0.2519
			RMSE	0.0102	0.0100	0.2288	0.0888	0.3794	0.0547	1.5094	0.1968	0.9522	0.0080	0.0099	0.1777	0.0716	0.0978	0.0538	0.5865	0.1615	1.0385
			Mbias	0.0018	-0.0007	0.0241	-0.0076	-0.1391	-0.0194	-0.3630	-0.0291	-0.0779	0.0002	-0.0001	0.0094	0.0010	-0.0261	-0.0165	0.0060	-0.0117	0.0181
			Mabias	0.0063	0.0068	0.1391	0.0537	0.1993	0.0387	0.5628	0.1361	0.4718	0.0053	0.0067	0.1191	0.0510	0.0569	0.0383	0.3304	0.1149	0.4356
WFE	5	800	**Bias**	0.0034	0.0010	1.1639	-0.3267	-0.7500	0.1147				0.0033	0.0012	1.0920	-0.3479	-0.7500	0.0984			
			RMSE	0.0049	0.0052	1.1682	0.3465	0.7500	0.1183				0.0049	0.0053	1.0970	0.3673	0.7500	0.1024			
			Mbias	0.0034	0.0008	1.1631	-0.3268	-0.7500	0.1143				0.0033	0.0010	1.0892	-0.3495	-0.7500	0.0982			
			Mabias	0.0036	0.0035	1.1631	0.3268	0.7500	0.1143				0.0035	0.0035	1.0892	0.3495	0.7500	0.0982			
TWFE	5	800	**Bias**	0.0001	0.0000	-0.0055	-0.0016	-0.0035	0.0000				0.0001	0.0000	-0.0055	-0.0018	-0.0034	0.0000			
			RMSE	0.0036	0.0052	0.0575	0.0189	0.0433	0.0261				0.0036	0.0052	0.0611	0.0201	0.0466	0.0261			
			Mbias	0.0001	-0.0003	-0.0058	-0.0018	-0.0015	-0.0013				0.0000	-0.0003	-0.0065	-0.0021	-0.0019	-0.0012			
			Mabias	0.0023	0.0034	0.0378	0.0121	0.0286	0.0163				0.0023	0.0035	0.0415	0.0131	0.0305	0.0162			
PWFE	5	800	**Bias**	0.0003	-0.0001	0.0081	-0.0075	-0.0750	-0.0012	0.0403	-0.0005	0.0357	0.0000	0.0001	-0.0070	-0.0042	-0.0084	-0.0007	-0.0501	0.0061	0.0601
			RMSE	0.0043	0.0053	0.1341	0.0470	0.2418	0.0262	1.1611	0.1093	0.3622	0.0038	0.0052	0.1097	0.0464	0.0578	0.0261	0.3269	0.0872	0.3772
			Mbias	0.0003	-0.0003	0.0027	-0.0028	-0.0520	-0.0022	-0.1414	-0.0057	-0.0033	0.0000	-0.0003	-0.0076	-0.0031	-0.0075	-0.0018	-0.0206	0.0008	0.0197
			Mabias	0.0029	0.0035	0.0866	0.0318	0.1209	0.0160	0.3352	0.0674	0.2411	0.0025	0.0034	0.0713	0.0313	0.0341	0.0160	0.1709	0.0581	0.2343

Table A.7. Simulation results for Case IV: $\rho = 0.01$, $\delta = 2$, $\alpha_1 = 1.5$, $\alpha_2 = 3$

				Negative exponential: row normalized									Negative exponential: scalar normalized								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.01)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.01)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)
WFE	**5**	**200**	**Bias**	0.0630	-0.0159	0.3667	-0.1497	-0.6785	0.0680				0.0556	-0.0161	0.3439	-0.1915	-0.7589	0.0492			
			RMSE	0.0637	0.0189	0.3692	0.1629	0.6820	0.0904				0.0565	0.0191	0.3477	0.2046	0.7641	0.0770			
			Mbias	0.0630	-0.0159	0.3674	-0.1487	-0.6778	0.0665				0.0554	-0.0162	0.3432	-0.1906	-0.7561	0.0472			
			Mabias	0.0630	0.0159	0.3674	0.1487	0.6778	0.0682				0.0554	0.0162	0.3432	0.1906	0.7561	0.0529			
TWFE	**5**	**200**	**Bias**	-0.0001	0.0002	-0.0013	0.0003	-0.0016	-0.0110				-0.0001	0.0002	-0.0019	0.0003	-0.0026	-0.0109			
			RMSE	0.0082	0.0098	0.0254	0.0250	0.0306	0.0531				0.0082	0.0098	0.0291	0.0266	0.0358	0.0531			
			Mbias	-0.0002	0.0000	-0.0012	-0.0007	-0.0013	-0.0137				0.0001	0.0000	-0.0021	0.0000	-0.0030	-0.0139			
			Mabias	0.0055	0.0066	0.0174	0.0160	0.0215	0.0375				0.0056	0.0067	0.0208	0.0173	0.0246	0.0373			
PWFE	**5**	**200**	**Bias**	0.0010	0.0001	0.0058	-0.0023	-0.0182	-0.0152	1.2907	-0.0043	0.1704	0.0003	0.0002	-0.0004	-0.0026	-0.0076	-0.0135	1.8134	-0.0019	0.1854
			RMSE	0.0129	0.0102	0.0530	0.0356	0.1231	0.0542	3.5977	0.1068	0.8599	0.0087	0.0099	0.0362	0.0384	0.0570	0.0537	4.0955	0.0466	0.9068
			Mbias	0.0004	-0.0004	0.0030	0.0002	0.0070	-0.0178	-0.1789	-0.0020	0.0224	0.0004	0.0001	-0.0004	0.0002	0.0002	-0.0160	0.1197	-0.0034	0.0310
			Mabias	0.0092	0.0069	0.0352	0.0241	0.0674	0.0389	1.3517	0.0757	0.4909	0.0059	0.0066	0.0232	0.0260	0.0338	0.0386	1.5000	0.0317	0.5039
WFE	**5**	**800**	**Bias**	0.0564	-0.0156	0.4066	-0.2241	-0.7207	0.0700				0.0509	-0.0141	0.4148	-0.2367	-0.7587	0.0593			
			RMSE	0.0566	0.0164	0.4074	0.2268	0.7221	0.0762				0.0512	0.0150	0.4157	0.2400	0.7605	0.0663			
			Mbias	0.0565	-0.0157	0.4065	-0.2225	-0.7206	0.0697				0.0510	-0.0141	0.4152	-0.2365	-0.7597	0.0598			
			Mabias	0.0565	0.0157	0.4065	0.2225	0.7206	0.0697				0.0510	0.0141	0.4152	0.2365	0.7597	0.0598			
TWFE	**5**	**800**	**Bias**	0.0000	-0.0002	0.0001	-0.0006	-0.0005	-0.0010				0.0000	-0.0002	0.0000	-0.0006	-0.0006	-0.0010			
			RMSE	0.0042	0.0050	0.0142	0.0136	0.0170	0.0261				0.0042	0.0050	0.0157	0.0143	0.0187	0.0261			
			Mbias	0.0002	-0.0003	0.0006	-0.0005	-0.0001	-0.0011				0.0001	-0.0003	0.0001	-0.0007	-0.0001	-0.0012			
			Mabias	0.0027	0.0035	0.0101	0.0091	0.0116	0.0173				0.0028	0.0035	0.0110	0.0099	0.0126	0.0172			
PWFE	**5**	**800**	**Bias**	0.0002	-0.0002	0.0012	-0.0018	-0.0015	-0.0020	3.5224	-0.0001	0.0294	0.0001	-0.0002	-0.0002	-0.0019	-0.0012	-0.0017	1.5770	0.0014	0.0360
			RMSE	0.0057	0.0052	0.0269	0.0196	0.0519	0.0261	5.2387	0.0541	0.3803	0.0044	0.0051	0.0193	0.0212	0.0345	0.0261	3.8587	0.0240	0.3845
			Mbias	0.0000	-0.0003	0.0009	-0.0004	-0.0005	-0.0022	2.1432	-0.0001	-0.0020	0.0002	-0.0003	-0.0009	-0.0015	0.0026	-0.0019	-0.0800	0.0026	-0.0036
			Mabias	0.0037	0.0035	0.0180	0.0135	0.0225	0.0172	2.1432	0.0386	0.2384	0.0029	0.0034	0.0130	0.0146	0.0191	0.0170	1.4975	0.0171	0.2489
				Inverse distance: row normalized									**Inverse distance: scalar normalized**								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.01)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.01)	σ^2(1)	δ(2)	α_1(1.5)	α_2(3)
WFE	**5**	**200**	**Bias**	0.0186	-0.0042	0.7048	-0.2392	-0.9913	0.0513				0.0175	-0.0055	0.6026	-0.3296	-0.9748	0.0386			
			RMSE	0.0199	0.0106	0.7163	0.2853	0.9934	0.0773				0.0189	0.0112	0.6182	0.3655	0.9791	0.0689			
			Mbias	0.0183	-0.0044	0.7046	-0.2417	-1.0100	0.0518				0.0173	-0.0056	0.6051	-0.3340	-1.0100	0.0374			
			Mabias	0.0183	0.0074	0.7046	0.2423	1.0100	0.0561				0.0173	0.0081	0.6051	0.3340	1.0100	0.0482			
TWFE	**5**	**200**	**Bias**	-0.0001	0.0002	-0.0067	0.0005	-0.0068	-0.0110				-0.0001	0.0001	-0.0079	0.0002	-0.0081	-0.0110			
			RMSE	0.0073	0.0098	0.0711	0.0333	0.0640	0.0531				0.0073	0.0098	0.0756	0.0357	0.0688	0.0531			
			Mbias	-0.0002	0.0001	-0.0048	-0.0004	-0.0059	-0.0134				-0.0002	0.0000	-0.0080	0.0009	-0.0056	-0.0136			
			Mabias	0.0050	0.0066	0.0478	0.0216	0.0427	0.0377				0.0049	0.0066	0.0503	0.0235	0.0447	0.0376			
PWFE	**5**	**200**	**Bias**	0.0028	-0.0008	0.0683	-0.0219	-0.1991	-0.0161	0.0877	-0.0559	0.1344	0.0003	0.0001	0.0028	-0.0063	-0.0330	-0.0139	0.9040	-0.0040	0.2481
			RMSE	0.0097	0.0099	0.2309	0.0956	0.4047	0.0543	2.6426	0.1819	0.9833	0.0076	0.0098	0.1445	0.0725	0.1302	0.0538	3.1464	0.1120	1.0462
			Mbias	0.0027	-0.0008	0.0368	-0.0065	-0.1428	-0.0184	-0.9271	-0.0529	-0.0809	0.0001	0.0000	-0.0004	-0.0022	-0.0213	-0.0160	-0.1111	-0.0093	0.0166
			Mabias	0.0063	0.0066	0.1325	0.0547	0.2197	0.0387	1.3611	0.1185	0.4887	0.0050	0.0066	0.0985	0.0504	0.0692	0.0381	1.4997	0.0820	0.4363
WFE	**5**	**800**	**Bias**	0.0075	-0.0045	1.0509	-1.4989	-1.0081	0.0386				0.0069	-0.0019	0.8052	-0.5164	-1.0093	0.0535			
			RMSE	0.0083	0.0068	1.0578	1.5081	1.0082	0.0475				0.0077	0.0055	0.8150	0.5363	1.0094	0.0602			
			Mbias	0.0076	-0.0048	1.0498	-1.4985	-1.0100	0.0385				0.0069	-0.0021	0.8012	-0.5193	-1.0100	0.0531			
			Mabias	0.0076	0.0051	1.0498	1.4985	1.0100	0.0386				0.0069	0.0038	0.8012	0.5193	1.0100	0.0531			
TWFE	**5**	**800**	**Bias**	0.0000	-0.0002	-0.0033	-0.0007	-0.0038	-0.0011				0.0001	0.0000	-0.0059	-0.0018	-0.0037	0.0001			
			RMSE	0.0036	0.0050	0.0570	0.0189	0.0407	0.0261				0.0036	0.0052	0.0610	0.0201	0.0446	0.0261			
			Mbias	0.0001	-0.0004	-0.0021	-0.0006	-0.0033	-0.0010				0.0000	-0.0003	-0.0063	-0.0020	-0.0025	-0.0011			
			Mabias	0.0024	0.0034	0.0373	0.0126	0.0278	0.0171				0.0023	0.0034	0.0412	0.0129	0.0291	0.0163			
PWFE	**5**	**800**	**Bias**	0.0005	-0.0004	0.0271	-0.0078	-0.0867	-0.0022	0.1072	-0.0157	0.0284	0.0001	0.0000	-0.0081	-0.0047	-0.0157	-0.0007	0.7582	0.0034	0.0569
			RMSE	0.0040	0.0051	0.1411	0.0499	0.2797	0.0262	2.4456	0.0972	0.3771	0.0037	0.0052	0.0915	0.0467	0.0979	0.0261	2.8403	0.0586	0.3760
			Mbias	0.0005	-0.0006	0.0132	-0.0007	-0.0195	-0.0021	-0.6966	-0.0142	-0.0039	0.0001	-0.0003	-0.0077	-0.0032	-0.0026	-0.0018	-0.0382	0.0013	0.0125
			Mabias	0.0027	0.0035	0.0973	0.0312	0.1519	0.0173	1.0631	0.0702	0.2444	0.0024	0.0034	0.0586	0.0316	0.0454	0.0163	1.2141	0.0388	0.2387

Table A.8. Simulation results for Case V: $\rho = 0.5$, $\delta = 2$, $\alpha_1 = 1.5$, $\alpha_2 = 1$

				Negative exponential: row normalized									Negative exponential: scalar normalized								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(1)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(1)
WFE	**5**	**200**	**Bias**	0.0332	-0.0014	0.3244	0.0591	-0.1561	0.0670				0.0268	-0.0006	0.3295	0.0667	-0.1462	0.0526			
			RMSE	0.0343	0.0102	0.3267	0.0818	0.1629	0.0886				0.0280	0.0101	0.3322	0.0904	0.1562	0.0777			
			Mbias	0.0331	-0.0015	0.3233	0.0605	-0.1550	0.0638				0.0266	-0.0006	0.3284	0.0681	-0.1435	0.0507			
			Mabias	0.0331	0.0068	0.3233	0.0622	0.1550	0.0662				0.0266	0.0068	0.3284	0.0695	0.1435	0.0554			
TWFE	**5**	**200**	**Bias**	0.0000	0.0002	-0.0013	0.0010	-0.0027	-0.0106				0.0000	0.0002	-0.0021	0.0002	-0.0038	-0.0105			
			RMSE	0.0077	0.0097	0.0240	0.0448	0.0250	0.0534				0.0077	0.0097	0.0278	0.0488	0.0296	0.0534			
			Mbias	0.0000	0.0001	-0.0012	0.0024	-0.0024	-0.0128				0.0002	0.0001	-0.0016	0.0018	-0.0034	-0.0124			
			Mabias	0.0051	0.0064	0.0163	0.0303	0.0173	0.0375				0.0053	0.0064	0.0190	0.0339	0.0194	0.0378			
PWFE	**5**	**200**	**Bias**	0.0017	0.0000	0.0126	-0.0070	-0.0175	-0.0131	0.0462	-0.0175	0.0526	0.0005	0.0002	0.0020	-0.0050	-0.0074	-0.0133	0.0141	-0.0036	0.0754
			RMSE	0.0099	0.0098	0.0565	0.0716	0.0557	0.0543	0.2632	0.1091	0.3963	0.0082	0.0097	0.0473	0.0733	0.0330	0.0540	0.2751	0.0973	0.4596
			Mbias	0.0010	-0.0002	0.0089	-0.0030	-0.0156	-0.0161	0.0364	-0.0209	-0.0130	0.0005	0.0001	0.0003	-0.0019	-0.0059	-0.0157	0.0076	-0.0044	0.0027
			Mabias	0.0064	0.0065	0.0366	0.0495	0.0353	0.0386	0.1729	0.0719	0.1831	0.0054	0.0066	0.0304	0.0540	0.0222	0.0378	0.1869	0.0673	0.1826
WFE	**5**	**800**	**Bias**	0.0304	-0.0025	0.3588	0.0370	-0.1628	0.0793				0.0266	-0.0025	0.3691	0.0412	-0.1567	0.0688			
			RMSE	0.0307	0.0059	0.3595	0.0488	0.1652	0.0845				0.0269	0.0059	0.3699	0.0534	0.1598	0.0746			
			Mbias	0.0305	-0.0028	0.3574	0.0387	-0.1622	0.0788				0.0265	-0.0027	0.3680	0.0422	-0.1562	0.0686			
			Mabias	0.0305	0.0043	0.3574	0.0395	0.1622	0.0788				0.0265	0.0042	0.3680	0.0427	0.1562	0.0686			
TWFE	**5**	**800**	**Bias**	0.0001	0.0001	-0.0007	-0.0016	-0.0011	0.0002				0.0001	0.0001	-0.0008	-0.0018	-0.0010	0.0002			
			RMSE	0.0038	0.0052	0.0138	0.0249	0.0148	0.0261				0.0038	0.0052	0.0150	0.0269	0.0160	0.0261			
			Mbias	0.0002	-0.0002	-0.0011	-0.0010	-0.0009	-0.0007				0.0001	-0.0002	-0.0009	-0.0016	-0.0003	-0.0010			
			Mabias	0.0025	0.0034	0.0095	0.0166	0.0100	0.0171				0.0025	0.0034	0.0106	0.0177	0.0111	0.0171			
PWFE	**5**	**800**	**Bias**	0.0002	0.0000	0.0002	-0.0032	-0.0027	-0.0007	0.0136	0.0010	0.0186	0.0001	0.0001	-0.0012	-0.0031	-0.0011	-0.0007	0.0094	0.0026	0.0221
			RMSE	0.0045	0.0052	0.0298	0.0397	0.0268	0.0263	0.1334	0.0558	0.1553	0.0041	0.0052	0.0252	0.0426	0.0171	0.0262	0.1341	0.0487	0.1651
			Mbias	0.0002	-0.0002	-0.0002	-0.0016	-0.0023	-0.0019	0.0122	0.0010	0.0051	0.0000	-0.0002	-0.0015	-0.0025	-0.0013	-0.0015	0.0056	0.0014	0.0048
			Mabias	0.0029	0.0034	0.0199	0.0276	0.0182	0.0165	0.0884	0.0365	0.0985	0.0025	0.0034	0.0169	0.0296	0.0115	0.0168	0.0916	0.0333	0.1017
				Inverse distance: row normalized									**Inverse distance: scalar normalized**								
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(1)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(1)
WFE	**5**	**200**	**Bias**	0.0083	0.0020	0.6580	0.2156	0.0323	0.0083				0.0079	0.0013	0.5945	0.1678	-0.0064	0.0058			
			RMSE	0.0110	0.0100	0.6699	0.2435	0.1164	0.0542				0.0107	0.0099	0.6092	0.2095	0.1169	0.0537			
			Mbias	0.0080	0.0019	0.6609	0.2262	0.0412	0.0056				0.0075	0.0012	0.5966	0.1797	0.0034	0.0051			
			Mabias	0.0082	0.0065	0.6609	0.2264	0.0826	0.0359				0.0078	0.0065	0.5966	0.1808	0.0800	0.0356			
TWFE	**5**	**200**	**Bias**	-0.0001	0.0001	-0.0131	-0.0072	-0.0122	-0.0106				-0.0001	0.0001	-0.0160	-0.0106	-0.0147	-0.0106			
			RMSE	0.0071	0.0097	0.0780	0.0925	0.0590	0.0533				0.0071	0.0097	0.0869	0.0995	0.0655	0.0533			
			Mbias	-0.0003	0.0000	-0.0089	-0.0024	-0.0096	-0.0130				-0.0003	-0.0001	-0.0104	-0.0058	-0.0104	-0.0129			
			Mabias	0.0049	0.0064	0.0503	0.0646	0.0389	0.0374				0.0049	0.0064	0.0557	0.0709	0.0436	0.0373			
PWFE	**5**	**200**	**Bias**	0.0015	0.0000	0.0705	-0.0212	-0.0729	-0.0135	0.2150	-0.0443	0.2993	0.0005	0.0002	0.0394	-0.0043	-0.0257	-0.0135	-0.0112	-0.0218	0.3485
			RMSE	0.0080	0.0097	0.2966	0.1547	0.1454	0.0539	0.7113	0.2245	1.7386	0.0076	0.0097	0.2469	0.1342	0.0745	0.0539	0.3233	0.2175	1.6814
			Mbias	0.0012	-0.0002	0.0281	-0.0220	-0.0639	-0.0158	0.1353	-0.0417	-0.0959	0.0002	0.0000	0.0090	-0.0075	-0.0235	-0.0166	-0.0029	-0.0234	-0.0089
			Mabias	0.0050	0.0065	0.1708	0.0986	0.0959	0.0379	0.2817	0.1481	0.3474	0.0051	0.0064	0.1543	0.0918	0.0491	0.0378	0.2076	0.1538	0.3066
WFE	**5**	**800**	**Bias**	0.0038	0.0005	0.5479	0.1743	-0.1033	0.0209				0.0035	0.0004	0.5040	0.1611	-0.1123	0.0175			
			RMSE	0.0053	0.0052	0.5574	0.2042	0.1371	0.0340				0.0051	0.0052	0.5149	0.1933	0.1452	0.0319			
			Mbias	0.0037	0.0002	0.5418	0.1762	-0.0977	0.0199				0.0034	0.0001	0.4999	0.1615	-0.1084	0.0169			
			Mabias	0.0038	0.0034	0.5418	0.1762	0.0992	0.0232				0.0036	0.0035	0.4999	0.1618	0.1084	0.0208			
TWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0093	-0.0053	-0.0061	0.0002				0.0000	0.0000	-0.0098	-0.0057	-0.0066	0.0002			
			RMSE	0.0036	0.0052	0.0582	0.0830	0.0358	0.0261				0.0035	0.0052	0.0615	0.0846	0.0382	0.0261			
			Mbias	0.0000	-0.0002	-0.0100	-0.0046	-0.0058	-0.0011				0.0000	-0.0002	-0.0093	-0.0075	-0.0059	-0.0010			
			Mabias	0.0023	0.0034	0.0400	0.0555	0.0239	0.0166				0.0023	0.0034	0.0421	0.0566	0.0252	0.0166			
PWFE	**5**	**800**	**Bias**	0.0003	0.0000	0.0100	-0.0176	-0.0378	-0.0006	0.0712	-0.0101	0.1265	0.0001	0.0000	-0.0097	-0.0057	-0.0110	-0.0007	0.0110	0.0047	0.1233
			RMSE	0.0037	0.0052	0.1438	0.1290	0.0961	0.0262	0.2084	0.1038	1.0973	0.0036	0.0052	0.1207	0.1153	0.0420	0.0262	0.1340	0.0952	0.8804
			Mbias	0.0002	-0.0002	0.0009	-0.0130	-0.0369	-0.0017	0.0519	-0.0104	-0.0344	0.0000	-0.0002	-0.0085	-0.0059	-0.0098	-0.0017	0.0127	0.0017	0.0080
			Mabias	0.0024	0.0034	0.0944	0.0893	0.0641	0.0167	0.1209	0.0694	0.2505	0.0023	0.0034	0.0824	0.0799	0.0295	0.0168	0.0898	0.0627	0.2210

Table A.9. Simulation results for Case VI: $\rho = 0.5$, $\delta = 2$, $\alpha_1 = 1.5$, $\alpha_2 = 0.5$

				Negative exponential: row normalized									Negative exponential: scalar normalized								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(0.5)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(0.5)
WFE	**5**	**200**	**Bias**	0.0311	0.0019	0.3261	0.1985	-0.1284	0.0732				0.0243	0.0024	0.3400	0.2054	-0.1042	0.0553			
			RMSE	0.0323	0.0102	0.3283	0.2060	0.1365	0.0934				0.0257	0.0103	0.3425	0.2134	0.1171	0.0794			
			Mbias	0.0310	0.0017	0.3257	0.1996	-0.1259	0.0710				0.0241	0.0023	0.3386	0.2067	-0.1019	0.0541			
			Mabias	0.0310	0.0067	0.3257	0.1996	0.1259	0.0718				0.0241	0.0066	0.3386	0.2067	0.1019	0.0570			
TWFE	**5**	**200**	**Bias**	0.0000	0.0002	-0.0014	0.0004	-0.0029	-0.0106				0.0000	0.0002	-0.0023	-0.0014	-0.0041	-0.0105			
			RMSE	0.0077	0.0097	0.0241	0.0663	0.0251	0.0535				0.0077	0.0097	0.0285	0.0717	0.0302	0.0534			
			Mbias	-0.0001	0.0001	-0.0016	0.0017	-0.0019	-0.0129				0.0001	0.0000	-0.0021	0.0010	-0.0026	-0.0127			
			Mabias	0.0051	0.0064	0.0162	0.0450	0.0169	0.0375				0.0052	0.0064	0.0199	0.0506	0.0204	0.0377			
PWFE	**5**	**200**	**Bias**	0.0007	0.0001	0.0077	0.0164	-0.0086	-0.0134	0.0381	-0.0075	0.1035	0.0003	0.0001	0.0051	0.0171	-0.0050	-0.0131	0.0218	-0.0049	0.1518
			RMSE	0.0095	0.0097	0.0542	0.0738	0.0507	0.0544	0.2566	0.1060	0.3216	0.0082	0.0097	0.0482	0.0807	0.0332	0.0540	0.2689	0.0985	0.6306
			Mbias	0.0002	-0.0001	0.0037	0.0152	-0.0064	-0.0168	0.0300	-0.0103	0.0000	0.0001	0.0000	0.0013	0.0163	-0.0034	-0.0156	0.0163	-0.0088	0.0000
			Mabias	0.0061	0.0065	0.0347	0.0497	0.0315	0.0384	0.1695	0.0713	0.0000	0.0054	0.0064	0.0305	0.0541	0.0224	0.0377	0.1851	0.0676	0.0000
WFE	**5**	**800**	**Bias**	0.0295	0.0008	0.3613	0.1961	-0.1483	0.0853				0.0258	0.0004	0.3725	0.2012	-0.1427	0.0738			
			RMSE	0.0298	0.0054	0.3619	0.1986	0.1510	0.0902				0.0262	0.0053	0.3732	0.2040	0.1461	0.0792			
			Mbias	0.0295	0.0006	0.3599	0.1978	-0.1478	0.0850				0.0257	0.0001	0.3714	0.2020	-0.1418	0.0744			
			Mabias	0.0295	0.0035	0.3599	0.1978	0.1478	0.0850				0.0257	0.0036	0.3714	0.2020	0.1418	0.0744			
TWFE	**5**	**800**	**Bias**	0.0001	0.0001	-0.0008	-0.0016	-0.0013	0.0002				0.0001	0.0001	-0.0009	-0.0019	-0.0012	0.0002			
			RMSE	0.0038	0.0052	0.0138	0.0408	0.0147	0.0261				0.0038	0.0052	0.0150	0.0442	0.0159	0.0261			
			Mbias	0.0001	-0.0002	-0.0011	-0.0009	-0.0010	-0.0007				0.0001	-0.0002	-0.0010	-0.0010	-0.0008	-0.0010			
			Mabias	0.0024	0.0034	0.0094	0.0277	0.0099	0.0171				0.0024	0.0034	0.0107	0.0296	0.0109	0.0169			
PWFE	**5**	**800**	**Bias**	-0.0001	0.0000	-0.0009	0.0160	0.0001	-0.0010	0.0137	0.0044	0.0602	0.0000	0.0000	-0.0004	0.0174	-0.0004	-0.0007	0.0147	0.0032	0.0670
			RMSE	0.0045	0.0052	0.0295	0.0497	0.0260	0.0263	0.1311	0.0550	0.1131	0.0041	0.0052	0.0253	0.0533	0.0169	0.0262	0.1305	0.0483	0.1262
			Mbias	-0.0001	-0.0002	-0.0013	0.0136	-0.0005	-0.0024	0.0123	0.0028	0.0056	-0.0001	-0.0002	-0.0003	0.0147	-0.0002	-0.0016	0.0108	0.0025	0.0118
			Mabias	0.0029	0.0035	0.0203	0.0319	0.0181	0.0167	0.0863	0.0355	0.0056	0.0026	0.0034	0.0165	0.0354	0.0116	0.0169	0.0887	0.0329	0.0118
				Inverse distance: row normalized									**Inverse distance: scalar normalized**								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(0.5)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(1.5)	α_2(0.5)
WFE	**5**	**200**	**Bias**	0.0077	0.0027	0.6857	0.3009	0.0944	0.0074				0.0072	0.0021	0.6191	0.2591	0.0589	0.0046			
			RMSE	0.0105	0.0101	0.6962	0.3177	0.1382	0.0539				0.0101	0.0100	0.6320	0.2825	0.1203	0.0534			
			Mbias	0.0074	0.0025	0.6885	0.3080	0.1008	0.0047				0.0068	0.0021	0.6210	0.2687	0.0668	0.0041			
			Mabias	0.0078	0.0066	0.6885	0.3080	0.1103	0.0363				0.0075	0.0065	0.6210	0.2687	0.0901	0.0358			
TWFE	**5**	**200**	**Bias**	-0.0001	0.0002	-0.0155	-0.0140	-0.0143	-0.0105				-0.0001	0.0001	-0.0192	-0.0184	-0.0176	-0.0105			
			RMSE	0.0071	0.0097	0.0803	0.1083	0.0617	0.0532				0.0071	0.0097	0.0899	0.1169	0.0695	0.0532			
			Mbias	-0.0003	-0.0001	-0.0096	-0.0121	-0.0111	-0.0130				-0.0003	-0.0001	-0.0135	-0.0148	-0.0148	-0.0126			
			Mabias	0.0049	0.0063	0.0515	0.0754	0.0421	0.0372				0.0048	0.0064	0.0575	0.0811	0.0465	0.0374			
PWFE	**5**	**200**	**Bias**	0.0011	0.0001	0.0759	0.0098	-0.0496	-0.0133	0.1502	-0.0394	0.7293	0.0005	0.0001	0.0500	0.0110	-0.0234	-0.0131	-0.0043	-0.0240	0.6982
			RMSE	0.0080	0.0097	0.3172	0.1658	0.1336	0.0539	0.6288	0.2341	2.1167	0.0077	0.0097	0.2760	0.1549	0.0782	0.0538	0.3348	0.2350	2.0298
			Mbias	0.0007	-0.0001	0.0224	-0.0035	-0.0400	-0.0155	0.1020	-0.0286	0.0000	0.0002	0.0000	0.0091	0.0009	-0.0205	-0.0160	0.0106	-0.0230	0.0610
			Mabias	0.0052	0.0065	0.1750	0.0958	0.0878	0.0381	0.2872	0.1568	0.0000	0.0052	0.0064	0.1690	0.0949	0.0515	0.0377	0.2132	0.1653	0.0610
WFE	**5**	**800**	**Bias**	0.0036	0.0010	0.5760	0.2912	-0.0472	0.0218				0.0034	0.0008	0.5319	0.2740	-0.0564	0.0184			
			RMSE	0.0051	0.0053	0.5844	0.3077	0.0998	0.0346				0.0049	0.0053	0.5415	0.2916	0.1046	0.0324			
			Mbias	0.0035	0.0007	0.5700	0.2934	-0.0421	0.0208				0.0032	0.0006	0.5265	0.2741	-0.0502	0.0178			
			Mabias	0.0037	0.0035	0.5700	0.2934	0.0647	0.0238				0.0035	0.0035	0.5265	0.2741	0.0673	0.0216			
TWFE	**5**	**800**	**Bias**	0.0001	0.0001	-0.0101	-0.0076	-0.0067	0.0002				0.0001	0.0001	-0.0107	-0.0081	-0.0073	0.0002			
			RMSE	0.0036	0.0052	0.0602	0.1040	0.0368	0.0261				0.0035	0.0052	0.0638	0.1050	0.0394	0.0261			
			Mbias	0.0000	-0.0002	-0.0105	-0.0095	-0.0064	-0.0011				0.0000	-0.0002	-0.0094	-0.0110	-0.0062	-0.0011			
			Mabias	0.0023	0.0034	0.0421	0.0727	0.0245	0.0166				0.0023	0.0034	0.0431	0.0723	0.0257	0.0166			
PWFE	**5**	**800**	**Bias**	0.0002	0.0000	0.0101	0.0075	-0.0223	-0.0006	0.0522	-0.0045	0.5545	0.0000	0.0000	-0.0028	0.0067	-0.0089	-0.0006	0.0180	0.0050	0.4617
			RMSE	0.0037	0.0052	0.1443	0.1345	0.0876	0.0262	0.1949	0.1025	1.7598	0.0036	0.0052	0.1219	0.1251	0.0424	0.0262	0.1319	0.0962	1.5249
			Mbias	0.0001	-0.0002	-0.0001	0.0027	-0.0181	-0.0017	0.0331	-0.0052	0.0094	0.0000	-0.0001	-0.0020	0.0004	-0.0082	-0.0017	0.0191	0.0032	0.0363
			Mabias	0.0024	0.0035	0.0928	0.0860	0.0573	0.0167	0.1194	0.0672	0.0094	0.0023	0.0035	0.0827	0.0797	0.0298	0.0167	0.0862	0.0626	0.0363

Table A.10. Simulation results for Case VII: $\rho = 0.5$, $\delta = 2$, $\alpha_1 = 10$, $\alpha_2 = 3$

				Negative exponential: row normalized									Negative exponential: scalar normalized								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(10)	α_2(3)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(10)	α_2(3)
WFE	**5**	**200**	**Bias**	0.2534	-0.0450	1.5971	-0.1596	-1.5000	2.2829				0.2085	-0.0402	1.6986	-0.2018	-1.5000	2.1663			
			RMSE	0.2535	0.0466	1.5994	0.2073	1.5000	2.2899				0.2087	0.0419	1.7009	0.2451	1.5000	2.1729			
			Mbias	0.2530	-0.0451	1.5979	-0.1563	-1.5000	2.2866				0.2082	-0.0401	1.6969	-0.2021	-1.5000	2.1682			
			Mabias	0.2530	0.0451	1.5979	0.1572	1.5000	2.2866				0.2082	0.0401	1.6969	0.2022	1.5000	2.1682			
TWFE	**5**	**200**	**Bias**	0.0001	0.0002	-0.0012	0.0002	-0.0021	-0.0107				0.0002	0.0002	-0.0016	0.0001	-0.0031	-0.0106			
			RMSE	0.0093	0.0098	0.0145	0.0249	0.0255	0.0534				0.0093	0.0098	0.0156	0.0266	0.0297	0.0533			
			Mbias	0.0001	0.0000	-0.0011	-0.0004	-0.0018	-0.0137				0.0004	0.0001	-0.0018	0.0003	-0.0028	-0.0131			
			Mabias	0.0063	0.0065	0.0097	0.0159	0.0169	0.0369				0.0062	0.0066	0.0109	0.0171	0.0204	0.0369			
PWFE	**5**	**200**	**Bias**	0.0009	0.0002	0.0001	-0.0024	-0.0045	-0.0140	0.0293	1.4186	0.1635	0.0028	-0.0001	0.0001	-0.0024	-0.0041	-0.0122	-0.0674	-0.6678	0.1529
			RMSE	0.0108	0.0098	0.0213	0.0356	0.0397	0.0542	0.2826	4.1010	0.8462	0.0107	0.0098	0.0179	0.0385	0.0329	0.0536	0.1627	1.0703	0.9090
			Mbias	0.0010	-0.0002	0.0004	0.0002	-0.0049	-0.0162	0.0074	-0.0205	0.0279	0.0025	-0.0003	0.0007	0.0008	-0.0036	-0.0144	-0.0560	-0.1061	-0.0090
			Mabias	0.0069	0.0065	0.0151	0.0240	0.0272	0.0384	0.1882	1.4112	0.4803	0.0073	0.0065	0.0129	0.0260	0.0220	0.0376	0.1002	0.1061	0.5126
WFE	**5**	**800**	**Bias**	0.2328	-0.0352	1.7976	-0.2622	-1.5000	2.4829				0.1985	-0.0321	1.8405	-0.3343	-1.5000	2.5047			
			RMSE	0.2329	0.0357	1.7982	0.2724	1.5000	2.4846				0.1985	0.0327	1.8412	0.3423	1.5000	2.5062			
			Mbias	0.2327	-0.0356	1.7952	-0.2576	-1.5000	2.4780				0.1987	-0.0324	1.8406	-0.3338	-1.5000	2.4995			
			Mabias	0.2327	0.0356	1.7952	0.2576	1.5000	2.4780				0.1987	0.0324	1.8406	0.3338	1.5000	2.4995			
TWFE	**5**	**800**	**Bias**	0.0002	0.0000	-0.0002	-0.0012	-0.0009	0.0002				0.0001	-0.0002	0.0000	-0.0006	-0.0006	-0.0010			
			RMSE	0.0047	0.0052	0.0078	0.0136	0.0150	0.0261				0.0047	0.0050	0.0081	0.0142	0.0156	0.0262			
			Mbias	0.0001	-0.0003	-0.0002	-0.0014	-0.0011	-0.0010				0.0003	-0.0004	-0.0001	-0.0007	-0.0001	-0.0011			
			Mabias	0.0030	0.0034	0.0054	0.0088	0.0102	0.0172				0.0033	0.0035	0.0054	0.0098	0.0106	0.0178			
PWFE	**5**	**800**	**Bias**	0.0000	0.0000	0.0000	-0.0015	-0.0015	-0.0008	0.0188	0.5285	0.0590	0.0012	-0.0003	0.0008	-0.0024	-0.0003	-0.0011	-0.0339	-0.3645	0.0074
			RMSE	0.0055	0.0052	0.0109	0.0198	0.0235	0.0261	0.1569	2.0145	0.3924	0.0052	0.0050	0.0089	0.0215	0.0176	0.0262	0.0759	0.6059	0.3820
			Mbias	-0.0001	-0.0003	-0.0006	-0.0002	-0.0019	-0.0022	0.0123	0.0830	0.0112	0.0013	-0.0004	0.0008	-0.0017	-0.0005	-0.0010	-0.0279	-0.0268	-0.0358
			Mabias	0.0038	0.0034	0.0072	0.0132	0.0156	0.0168	0.1056	0.7129	0.2544	0.0035	0.0035	0.0060	0.0145	0.0117	0.0178	0.0473	0.0268	0.2475
				Inverse distance: row normalized									**Inverse distance scalar normalization**								
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(10)	α_2(3)	β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	σ^2(1)	δ(2)	α_1(10)	α_2(3)
WFE	**5**	**200**	**Bias**	0.1335	0.0031	6.4894	1.1504	-1.5000	5.7710				0.1229	0.0003	6.2172	0.9328	-1.5000	5.1923			
			RMSE	0.1337	0.0107	6.4923	1.1872	1.5000	5.7764				0.1231	0.0102	6.2203	0.9771	1.5000	5.1981			
			Mbias	0.1332	0.0030	6.4876	1.1451	-1.5000	5.7699				0.1227	0.0002	6.2201	0.9256	-1.5000	5.1900			
			Mabias	0.1332	0.0070	6.4876	1.1451	1.5000	5.7699				0.1227	0.0068	6.2201	0.9256	1.5000	5.1900			
TWFE	**5**	**200**	**Bias**	0.0001	0.0002	-0.0013	0.0004	-0.0043	-0.0108				0.0003	0.0002	-0.0016	0.0000	-0.0063	-0.0108			
			RMSE	0.0091	0.0097	0.0145	0.0330	0.0460	0.0532				0.0090	0.0098	0.0155	0.0355	0.0512	0.0533			
			Mbias	0.0005	0.0001	-0.0016	-0.0002	-0.0016	-0.0134				0.0005	0.0001	-0.0016	0.0011	-0.0044	-0.0127			
			Mabias	0.0062	0.0064	0.0097	0.0217	0.0292	0.0372				0.0063	0.0065	0.0108	0.0234	0.0346	0.0374			
PWFE	**5**	**200**	**Bias**	0.0004	0.0002	0.0026	-0.0093	-0.0222	-0.0146	0.0942	0.3420	0.1795	0.0035	-0.0001	0.0030	-0.0054	-0.0091	-0.0126	-0.0771	-0.4902	0.1736
			RMSE	0.0115	0.0098	0.0200	0.0714	0.0937	0.0541	0.3705	1.8538	0.8862	0.0112	0.0098	0.0188	0.0651	0.0558	0.0536	0.1781	0.7939	0.9235
			Mbias	0.0003	-0.0002	0.0032	-0.0013	-0.0221	-0.0167	0.0546	-0.0015	0.0047	0.0032	-0.0004	0.0028	-0.0032	-0.0089	-0.0149	-0.0552	-0.1115	-0.0394
			Mabias	0.0076	0.0066	0.0138	0.0475	0.0666	0.0385	0.2186	0.8331	0.4412	0.0073	0.0066	0.0136	0.0469	0.0350	0.0377	0.1027	0.1115	0.4167
WFE	**5**	**800**	**Bias**	0.0743	-0.0136	7.6844	-3.2441	-1.5000	8.5477				0.0671	-0.0137	7.6766	-3.4183	-1.5000	7.9037			
			RMSE	0.0744	0.0145	7.6875	3.2542	1.5000	8.5495				0.0672	0.0146	7.6797	3.4284	1.5000	7.9054			
			Mbias	0.0745	-0.0137	7.6815	-3.2525	-1.5000	8.5417				0.0672	-0.0139	7.6799	-3.4204	-1.5000	7.8985			
			Mabias	0.0745	0.0137	7.6815	3.2525	1.5000	8.5417				0.0672	0.0139	7.6799	3.4204	1.5000	7.8985			
TWFE	**5**	**800**	**Bias**	0.0001	-0.0002	-0.0001	-0.0008	-0.0018	-0.0010				0.0001	-0.0002	-0.0001	-0.0009	-0.0022	-0.0010			
			RMSE	0.0047	0.0050	0.0078	0.0189	0.0325	0.0261				0.0047	0.0050	0.0082	0.0204	0.0364	0.0261			
			Mbias	0.0003	-0.0003	-0.0001	-0.0011	-0.0002	-0.0009				0.0003	-0.0003	-0.0001	-0.0008	-0.0008	-0.0012			
			Mabias	0.0031	0.0035	0.0054	0.0124	0.0221	0.0173				0.0032	0.0035	0.0056	0.0140	0.0251	0.0173			
PWFE	**5**	**800**	**Bias**	0.0002	-0.0002	0.0009	-0.0060	-0.0082	-0.0019	0.0266	0.0501	0.0419	0.0015	-0.0003	0.0022	-0.0093	-0.0017	-0.0013	-0.0326	-0.2436	0.0272
			RMSE	0.0058	0.0050	0.0103	0.0479	0.0667	0.0261	0.1744	0.6379	0.3743	0.0053	0.0050	0.0092	0.0554	0.0405	0.0261	0.0668	0.4043	0.4071
			Mbias	0.0002	-0.0003	0.0009	-0.0002	-0.0049	-0.0018	0.0026	-0.0180	-0.0004	0.0016	-0.0004	0.0019	-0.0025	-0.0001	-0.0014	-0.0266	-0.0489	-0.0167
			Mabias	0.0036	0.0035	0.0068	0.0303	0.0444	0.0174	0.1099	0.4109	0.2275	0.0035	0.0035	0.0062	0.0345	0.0273	0.0174	0.0366	0.0489	0.2555

Table A.11. Direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 for Case II and III using the parameter estimates in Table A.5 and A.6

$\rho = 0.25$				Exponential: row normalized				Exponential: scalar normalized				Inverse-dis: row normalized				Inverse-dis: scalar normalized			
	T	N		DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_2	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Bias**	0.0012	0.0001	0.0975	-0.0867	0.0001	-0.0008	0.0723	-0.0953	-0.0003	0.0005	0.0411	-0.1244	0.0000	-0.0008	0.0292	-0.1919
			RMSE	0.0071	0.0099	0.1109	0.1012	0.0070	0.0099	0.0911	0.1082	0.0070	0.0098	0.2203	0.1653	0.0070	0.0098	0.2219	0.2209
			Mbias	0.0010	0.0000	0.0999	-0.0855	-0.0001	-0.0010	0.0714	-0.0954	-0.0006	0.0003	0.0136	-0.1305	-0.0003	-0.0010	0.0047	-0.1976
			Mabias	0.0048	0.0066	0.0999	0.0856	0.0049	0.0065	0.0717	0.0955	0.0049	0.0065	0.1422	0.1327	0.0049	0.0066	0.1533	0.1977
TWFE	**5**	**200**	**Bias**	-0.0004	0.0003	-0.0022	0.0006	-0.0004	0.0003	-0.0029	0.0005	-0.0004	0.0003	-0.0081	0.0014	-0.0004	0.0002	-0.0092	0.0011
			RMSE	0.0070	0.0098	0.0435	0.0327	0.0070	0.0098	0.0454	0.0321	0.0070	0.0097	0.1305	0.0441	0.0070	0.0097	0.1308	0.0434
			Mbias	-0.0006	0.0001	-0.0019	0.0002	-0.0005	0.0001	-0.0047	-0.0001	-0.0006	0.0003	-0.0132	0.0013	-0.0006	0.0002	-0.0160	0.0011
			Mabias	0.0049	0.0065	0.0300	0.0212	0.0049	0.0066	0.0323	0.0207	0.0049	0.0065	0.0884	0.0304	0.0048	0.0064	0.0879	0.0299
PWFE	**5**	**200**	**Bias**	-0.0004	0.0004	0.0030	-0.0016	-0.0004	0.0004	0.0021	-0.0022	-0.0002	0.0005	0.0319	0.0015	-0.0003	0.0004	0.0143	-0.0041
			RMSE	0.0070	0.0099	0.0592	0.0466	0.0070	0.0098	0.0640	0.0459	0.0070	0.0097	0.2601	0.0931	0.0071	0.0097	0.2605	0.0884
			Mbias	-0.0005	0.0002	-0.0009	-0.0001	-0.0006	0.0003	-0.0009	0.0008	-0.0004	0.0003	0.0187	0.0122	-0.0004	0.0002	0.0000	0.0032
			Mabias	0.0049	0.0065	0.0379	0.0312	0.0049	0.0066	0.0411	0.0305	0.0050	0.0064	0.1738	0.0654	0.0050	0.0063	0.1717	0.0633
WFE	**5**	**800**	**Bias**	0.0005	0.0014	0.1132	-0.1240	0.0008	0.0004	0.1158	-0.1200	0.0001	0.0000	-0.0385	-0.2970	0.0001	0.0001	-0.0069	-0.3020
			RMSE	0.0036	0.0054	0.1182	0.1274	0.0037	0.0053	0.1207	0.1236	0.0036	0.0052	0.1409	0.3109	0.0036	0.0052	0.1397	0.3156
			Mbias	0.0006	0.0012	0.1129	-0.1225	0.0008	0.0002	0.1147	-0.1198	0.0001	-0.0003	-0.0518	-0.2989	0.0001	-0.0001	-0.0198	-0.3046
			Mabias	0.0023	0.0035	0.1129	0.1225	0.0024	0.0035	0.1147	0.1198	0.0024	0.0034	0.1090	0.2989	0.0023	0.0035	0.1022	0.3046
TWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0011	-0.0015	0.0000	0.0000	-0.0008	-0.0016	0.0000	0.0000	-0.0074	-0.0018	0.0000	0.0000	-0.0069	-0.0019
			RMSE	0.0036	0.0053	0.0273	0.0181	0.0036	0.0052	0.0272	0.0181	0.0036	0.0052	0.1009	0.0255	0.0036	0.0052	0.1018	0.0257
			Mbias	0.0000	-0.0002	-0.0017	-0.0015	0.0000	-0.0002	-0.0005	-0.0020	0.0000	-0.0002	-0.0095	-0.0015	0.0000	-0.0002	-0.0081	-0.0022
			Mabias	0.0023	0.0035	0.0189	0.0122	0.0023	0.0035	0.0190	0.0121	0.0024	0.0034	0.0657	0.0175	0.0024	0.0034	0.0686	0.0178
PWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0007	-0.0019	0.0000	0.0000	-0.0008	-0.0019	0.0000	0.0000	-0.0057	-0.0041	0.0000	0.0000	-0.0118	-0.0056
			RMSE	0.0036	0.0053	0.0369	0.0263	0.0036	0.0052	0.0367	0.0262	0.0036	0.0052	0.1549	0.0587	0.0036	0.0052	0.1546	0.0585
			Mbias	0.0000	-0.0002	-0.0024	0.0003	0.0000	-0.0002	-0.0017	0.0001	0.0000	-0.0002	-0.0085	-0.0021	0.0000	-0.0002	-0.0127	-0.0040
			Mabias	0.0024	0.0035	0.0240	0.0176	0.0023	0.0035	0.0248	0.0171	0.0024	0.0035	0.1064	0.0391	0.0024	0.0035	0.1033	0.0389
$\rho = -0.25$				Exponential: row normalized				Exponential: scalar normalized				Inverse-dis: row normalized				Inverse-dis: scalar normalized			
	T	N		DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_2	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Bias**	0.0035	-0.0034	0.0702	-0.0438	0.0017	-0.0031	0.0777	-0.0519	0.0020	0.0009	0.3503	-0.0062	0.0026	-0.0001	0.3153	-0.0540
			RMSE	0.0079	0.0103	0.0733	0.0540	0.0072	0.0102	0.0808	0.0610	0.0073	0.0097	0.3545	0.0630	0.0075	0.0097	0.3203	0.0830
			Mbias	0.0034	-0.0037	0.0702	-0.0436	0.0016	-0.0032	0.0778	-0.0516	0.0018	0.0007	0.3522	-0.0081	0.0025	-0.0003	0.3179	-0.0558
			Mabias	0.0052	0.0071	0.0702	0.0436	0.0049	0.0070	0.0778	0.0517	0.0049	0.0065	0.3522	0.0453	0.0049	0.0064	0.3179	0.0609
TWFE	**5**	**200**	**Bias**	-0.0004	0.0003	-0.0007	0.0002	-0.0004	0.0003	-0.0011	0.0001	-0.0004	0.0003	-0.0033	0.0005	-0.0004	0.0003	-0.0040	0.0004
			RMSE	0.0070	0.0098	0.0247	0.0209	0.0070	0.0098	0.0260	0.0209	0.0070	0.0097	0.0698	0.0271	0.0070	0.0097	0.0718	0.0270
			Mbias	-0.0004	0.0002	-0.0010	-0.0005	-0.0004	0.0001	-0.0018	-0.0002	-0.0005	0.0001	-0.0053	0.0000	-0.0005	0.0002	-0.0063	0.0004
			Mabias	0.0050	0.0064	0.0169	0.0136	0.0049	0.0064	0.0183	0.0135	0.0049	0.0064	0.0467	0.0179	0.0049	0.0064	0.0495	0.0186
PWFE	**5**	**200**	**Bias**	-0.0003	0.0004	-0.0023	-0.0009	-0.0004	0.0004	-0.0044	-0.0021	-0.0004	0.0003	-0.0114	-0.0025	-0.0003	0.0004	0.0083	-0.0027
			RMSE	0.0070	0.0098	0.0336	0.0284	0.0070	0.0098	0.0373	0.0284	0.0070	0.0097	0.1551	0.0549	0.0070	0.0097	0.1529	0.0535
			Mbias	-0.0004	0.0004	-0.0023	0.0006	-0.0004	0.0003	-0.0050	-0.0002	-0.0005	0.0001	-0.0201	0.0023	-0.0005	0.0003	0.0001	0.0020
			Mabias	0.0048	0.0064	0.0212	0.0188	0.0049	0.0064	0.0237	0.0190	0.0050	0.0063	0.1031	0.0382	0.0049	0.0063	0.0991	0.0372
WFE	**5**	**800**	**Bias**	0.0028	0.0007	0.0939	-0.0753	0.0029	-0.0005	0.1073	-0.0760	0.0022	-0.0004	0.4315	-0.1324	0.0019	0.0001	0.4114	-0.1483
			RMSE	0.0045	0.0052	0.0947	0.0772	0.0046	0.0052	0.1080	0.0780	0.0042	0.0052	0.4344	0.1445	0.0040	0.0052	0.4146	0.1592
			Mbias	0.0028	0.0006	0.0936	-0.0743	0.0029	-0.0008	0.1067	-0.0757	0.0022	-0.0006	0.4313	-0.1326	0.0019	-0.0001	0.4106	-0.1505
			Mabias	0.0032	0.0035	0.0936	0.0743	0.0032	0.0035	0.1067	0.0757	0.0028	0.0035	0.4313	0.1326	0.0027	0.0034	0.4106	0.1505
TWFE	**5**	**800**	**Bias**	0.0000	0.0001	-0.0006	-0.0010	0.0000	0.0001	-0.0004	-0.0011	0.0000	0.0001	-0.0039	-0.0012	0.0000	0.0001	-0.0035	-0.0013
			RMSE	0.0036	0.0052	0.0150	0.0113	0.0036	0.0052	0.0153	0.0115	0.0036	0.0051	0.0555	0.0153	0.0036	0.0051	0.0566	0.0156
			Mbias	0.0000	-0.0001	-0.0010	-0.0013	0.0000	-0.0001	-0.0004	-0.0012	0.0000	-0.0001	-0.0068	-0.0014	0.0000	-0.0002	-0.0045	-0.0015
			Mabias	0.0024	0.0035	0.0104	0.0073	0.0024	0.0034	0.0107	0.0075	0.0024	0.0034	0.0364	0.0101	0.0023	0.0034	0.0389	0.0104
PWFE	**5**	**800**	**Bias**	0.0000	0.0001	-0.0017	-0.0011	0.0000	0.0001	-0.0014	-0.0013	0.0000	0.0000	-0.0132	-0.0021	0.0000	0.0001	-0.0058	-0.0030
			RMSE	0.0036	0.0052	0.0214	0.0161	0.0036	0.0052	0.0219	0.0163	0.0036	0.0052	0.0931	0.0347	0.0036	0.0052	0.0938	0.0352
			Mbias	-0.0001	-0.0001	-0.0021	0.0002	0.0000	-0.0001	-0.0010	-0.0004	0.0000	-0.0002	-0.0127	-0.0003	0.0000	-0.0002	-0.0064	-0.0017
			Mabias	0.0024	0.0034	0.0141	0.0108	0.0024	0.0035	0.0147	0.0109	0.0023	0.0034	0.0629	0.0234	0.0024	0.0034	0.0618	0.0238

Table A.12. Direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 for Case V and VI using the parameter estimates in Table A.8 and A.9

$\alpha_2 = 1$				Exponential: row normalized				Exponential: scalar normalized				Inverse-dis: row normalized				Inverse-dis: scalar normalized			
	T	N		DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_2	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Bias**	0.0018	0.0027	0.3107	0.1328	0.0020	0.0027	0.3206	0.1274	0.0081	0.0036	1.6614	0.5012	0.0067	0.0025	1.3320	0.3707
			RMSE	0.0075	0.0107	0.3254	0.1582	0.0075	0.0107	0.3376	0.1531	0.0114	0.0107	1.8511	0.5890	0.0103	0.0103	1.4972	0.4660
			Mbias	0.0015	0.0028	0.3078	0.1337	0.0019	0.0028	0.3182	0.1273	0.0080	0.0039	1.5487	0.4663	0.0065	0.0027	1.2286	0.3482
			Mabias	0.0051	0.0070	0.3078	0.1337	0.0049	0.0069	0.3182	0.1273	0.0085	0.0069	1.5487	0.4663	0.0073	0.0066	1.2286	0.3482
TWFE	**5**	**200**	**Bias**	-0.0004	0.0003	-0.0051	0.0031	-0.0004	0.0003	-0.0061	0.0020	-0.0005	0.0001	-0.0253	-0.0040	-0.0005	0.0001	-0.0288	-0.0066
			RMSE	0.0073	0.0104	0.0700	0.0870	0.0073	0.0102	0.0736	0.0839	0.0071	0.0099	0.2369	0.1715	0.0071	0.0099	0.2479	0.1737
			Mbias	-0.0006	0.0004	-0.0047	0.0059	-0.0006	0.0004	-0.0069	0.0022	-0.0006	0.0000	-0.0323	-0.0035	-0.0007	0.0000	-0.0407	-0.0082
			Mabias	0.0051	0.0069	0.0468	0.0588	0.0050	0.0069	0.0523	0.0566	0.0050	0.0065	0.1595	0.1234	0.0049	0.0065	0.1727	0.1259
PWFE	**5**	**200**	**Bias**	-0.0004	0.0004	-0.0033	-0.0057	-0.0004	0.0003	-0.0035	-0.0067	-0.0001	0.0002	0.0418	0.0032	0.0000	0.0002	0.0641	0.0113
			RMSE	0.0074	0.0104	0.0890	0.1331	0.0073	0.0103	0.1006	0.1282	0.0076	0.0099	0.5503	0.2594	0.0076	0.0099	0.5253	0.2432
			Mbias	-0.0006	0.0004	-0.0046	0.0020	-0.0005	0.0003	-0.0074	-0.0017	-0.0002	0.0001	-0.0585	-0.0219	-0.0003	0.0001	-0.0182	0.0007
			Mabias	0.0051	0.0069	0.0579	0.0984	0.0050	0.0068	0.0680	0.0951	0.0051	0.0065	0.3336	0.1686	0.0051	0.0064	0.3253	0.1676
WFE	**5**	**800**	**Bias**	0.0013	0.0006	0.3426	0.1004	0.0013	0.0000	0.3350	0.0934	0.0004	0.0004	0.7803	0.3333	0.0006	0.0003	0.7146	0.2993
			RMSE	0.0040	0.0056	0.3482	0.1115	0.0040	0.0055	0.3407	0.1051	0.0037	0.0053	0.8379	0.3809	0.0037	0.0052	0.7701	0.3473
			Mbias	0.0012	0.0003	0.3417	0.1026	0.0013	-0.0003	0.3348	0.0952	0.0005	0.0001	0.7634	0.3310	0.0006	0.0001	0.6994	0.2951
			Mabias	0.0026	0.0036	0.3417	0.1026	0.0026	0.0037	0.3348	0.0952	0.0024	0.0034	0.7634	0.3310	0.0023	0.0035	0.6994	0.2951
TWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0023	-0.0029	0.0000	0.0000	-0.0018	-0.0032	0.0000	0.0000	-0.0201	-0.0064	0.0000	0.0000	-0.0197	-0.0070
			RMSE	0.0038	0.0055	0.0439	0.0496	0.0038	0.0055	0.0430	0.0492	0.0036	0.0052	0.1690	0.1609	0.0036	0.0052	0.1676	0.1574
			Mbias	-0.0001	-0.0003	-0.0031	-0.0005	-0.0001	-0.0003	-0.0015	-0.0010	0.0000	-0.0002	-0.0338	-0.0082	0.0000	-0.0002	-0.0323	-0.0105
			Mabias	0.0025	0.0037	0.0305	0.0335	0.0024	0.0036	0.0295	0.0332	0.0024	0.0035	0.1184	0.1089	0.0024	0.0035	0.1168	0.1044
PWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0033	-0.0051	0.0000	0.0000	-0.0028	-0.0054	0.0000	0.0000	-0.0361	-0.0125	0.0000	0.0000	-0.0259	-0.0051
			RMSE	0.0038	0.0055	0.0569	0.0776	0.0038	0.0055	0.0561	0.0764	0.0036	0.0052	0.2754	0.2307	0.0036	0.0052	0.2633	0.2172
			Mbias	0.0000	-0.0003	-0.0070	-0.0024	-0.0001	-0.0003	-0.0049	-0.0027	0.0000	-0.0002	-0.0463	-0.0182	0.0000	-0.0003	-0.0243	-0.0099
			Mabias	0.0024	0.0037	0.0364	0.0529	0.0024	0.0037	0.0375	0.0541	0.0024	0.0035	0.1846	0.1594	0.0024	0.0035	0.1769	0.1512
$\alpha_2 = 0.5$				**Exponential: row normalized**				**Exponential: scalar normalized**				**Inverse-dis: row normalized**				**Inverse-dis: scalar normalized**			
	T	N		DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_2	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Bias**	0.0024	0.0021	0.3676	0.3582	0.0030	0.0027	0.4103	0.3438	0.0106	0.0041	2.1487	0.7698	0.0086	0.0031	1.7259	0.6133
			RMSE	0.0077	0.0106	0.3812	0.3687	0.0078	0.0107	0.4259	0.3552	0.0135	0.0110	2.3566	0.8572	0.0118	0.0105	1.9019	0.6997
			Mbias	0.0021	0.0021	0.3631	0.3599	0.0029	0.0029	0.4064	0.3410	0.0105	0.0044	2.0128	0.7199	0.0085	0.0032	1.5957	0.5745
			Mabias	0.0050	0.0070	0.3631	0.3599	0.0051	0.0070	0.4064	0.3410	0.0105	0.0073	2.0128	0.7199	0.0088	0.0068	1.5957	0.5745
TWFE	**5**	**200**	**Bias**	-0.0004	0.0003	-0.0054	0.0026	-0.0004	0.0002	-0.0068	0.0004	-0.0005	0.0001	-0.0319	-0.0131	-0.0006	0.0001	-0.0367	-0.0166
			RMSE	0.0074	0.0103	0.0709	0.1288	0.0073	0.0102	0.0764	0.1247	0.0072	0.0099	0.2458	0.1978	0.0072	0.0099	0.2588	0.2025
			Mbias	-0.0006	0.0002	-0.0052	0.0015	-0.0006	0.0002	-0.0076	0.0007	-0.0007	-0.0001	-0.0386	-0.0165	-0.0007	-0.0001	-0.0443	-0.0255
			Mabias	0.0051	0.0069	0.0479	0.0896	0.0050	0.0068	0.0544	0.0869	0.0050	0.0066	0.1682	0.1443	0.0049	0.0065	0.1786	0.1438
PWFE	**5**	**200**	**Bias**	-0.0003	-0.0004	0.0023	0.0359	-0.0003	-0.0003	0.0053	0.0355	0.0002	0.0000	0.1082	0.0573	0.0001	0.0000	0.0972	0.0469
			RMSE	0.0074	0.0103	0.0902	0.1448	0.0073	0.0102	0.1032	0.1438	0.0078	0.0099	0.6525	0.3157	0.0077	0.0099	0.6049	0.2905
			Mbias	-0.0005	-0.0005	-0.0004	0.0322	-0.0004	-0.0003	-0.0003	0.0283	0.0000	-0.0002	-0.0266	0.0072	-0.0002	-0.0001	-0.0162	0.0089
			Mabias	0.0051	0.0068	0.0570	0.0959	0.0050	0.0069	0.0683	0.0943	0.0053	0.0067	0.3642	0.1703	0.0052	0.0066	0.3519	0.1724
WFE	**5**	**800**	**Bias**	0.0015	0.0000	0.3749	0.3495	0.0015	-0.0006	0.3638	0.3196	0.0007	0.0004	1.0202	0.5669	0.0008	0.0004	0.9309	0.5185
			RMSE	0.0041	0.0055	0.3804	0.3528	0.0041	0.0055	0.3694	0.3232	0.0037	0.0053	1.0802	0.6036	0.0037	0.0053	0.9848	0.5535
			Mbias	0.0015	-0.0003	0.3742	0.3518	0.0015	-0.0009	0.3626	0.3210	0.0008	0.0001	1.0096	0.5582	0.0008	0.0002	0.9185	0.5117
			Mabias	0.0027	0.0037	0.3742	0.3518	0.0027	0.0037	0.3626	0.3210	0.0023	0.0035	1.0096	0.5582	0.0024	0.0034	0.9185	0.5117
TWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0028	-0.0028	0.0000	0.0000	-0.0023	-0.0032	0.0000	0.0000	-0.0223	-0.0085	0.0000	0.0000	-0.0218	-0.0089
			RMSE	0.0038	0.0055	0.0436	0.0810	0.0038	0.0054	0.0427	0.0797	0.0036	0.0052	0.1754	0.2000	0.0036	0.0052	0.1741	0.1953
			Mbias	-0.0001	-0.0003	-0.0036	-0.0015	-0.0001	-0.0003	-0.0022	-0.0024	0.0000	-0.0002	-0.0343	-0.0149	-0.0001	-0.0002	-0.0318	-0.0184
			Mabias	0.0024	0.0036	0.0301	0.0553	0.0024	0.0036	0.0287	0.0540	0.0024	0.0035	0.1158	0.1381	0.0024	0.0035	0.1194	0.1382
PWFE	**5**	**800**	**Bias**	0.0000	-0.0003	-0.0006	0.0318	0.0000	-0.0003	-0.0003	0.0310	0.0000	-0.0001	-0.0079	0.0281	0.0000	-0.0001	-0.0086	0.0220
			RMSE	0.0038	0.0055	0.0568	0.0994	0.0038	0.0055	0.0559	0.0960	0.0036	0.0052	0.2831	0.2557	0.0036	0.0052	0.2701	0.2367
			Mbias	0.0000	-0.0005	-0.0043	0.0286	0.0000	-0.0005	-0.0011	0.0272	0.0000	-0.0004	-0.0276	0.0092	0.0000	-0.0003	-0.0178	0.0101
			Mabias	0.0024	0.0037	0.0364	0.0627	0.0024	0.0036	0.0373	0.0640	0.0024	0.0035	0.1919	0.1615	0.0024	0.0035	0.1806	0.1499

Table A.13. Direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 for Case IV and VII using the parameter estimates in Table A.7 and A.10

$\rho = 0.01$				Exponential: row normalized				Exponential: scalar normalized				Inverse-dis: row normalized				Inverse-dis: scalar normalized			
	T	N		DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Bias**	0.0018	-0.0018	0.0514	-0.0565	0.0002	-0.0024	0.0364	-0.0702	0.0003	0.0001	0.1134	-0.0721	0.0008	-0.0010	0.0963	-0.1309
			RMSE	0.0072	0.0099	0.0633	0.0686	0.0070	0.0100	0.0534	0.0801	0.0070	0.0097	0.1359	0.1062	0.0070	0.0097	0.1293	0.1523
			Mbias	0.0016	-0.0018	0.0520	-0.0556	0.0001	-0.0026	0.0353	-0.0704	0.0002	-0.0001	0.1093	-0.0740	0.0006	-0.0012	0.0884	-0.1338
			Mabias	0.0048	0.0066	0.0523	0.0561	0.0048	0.0067	0.0380	0.0706	0.0048	0.0063	0.1093	0.0794	0.0048	0.0064	0.0889	0.1339
TWFE	**5**	**200**	**Bias**	-0.0004	0.0003	-0.0013	0.0003	-0.0004	0.0003	-0.0017	0.0003	-0.0004	0.0003	-0.0051	0.0008	-0.0004	0.0003	-0.0059	0.0006
			RMSE	0.0070	0.0097	0.0318	0.0253	0.0070	0.0097	0.0334	0.0251	0.0070	0.0097	0.0929	0.0336	0.0070	0.0097	0.0944	0.0333
			Mbias	-0.0004	0.0002	-0.0019	-0.0004	-0.0004	0.0002	-0.0029	-0.0001	-0.0005	0.0002	-0.0065	0.0001	-0.0005	0.0001	-0.0120	0.0002
			Mabias	0.0049	0.0064	0.0218	0.0163	0.0048	0.0064	0.0239	0.0164	0.0048	0.0064	0.0609	0.0228	0.0048	0.0064	0.0638	0.0229
PWFE	**5**	**200**	**Bias**	-0.0004	0.0004	-0.0002	-0.0015	-0.0004	0.0004	-0.0013	-0.0021	-0.0004	0.0004	-0.0052	-0.0015	-0.0004	0.0004	-0.0013	-0.0045
			RMSE	0.0070	0.0097	0.0443	0.0359	0.0070	0.0097	0.0423	0.0353	0.0070	0.0097	0.1811	0.0693	0.0070	0.0097	0.1604	0.0680
			Mbias	-0.0005	0.0003	-0.0015	0.0007	-0.0004	0.0003	-0.0028	0.0006	-0.0005	0.0001	-0.0117	0.0057	-0.0005	0.0003	-0.0034	0.0013
			Mabias	0.0048	0.0063	0.0285	0.0245	0.0048	0.0063	0.0276	0.0237	0.0049	0.0063	0.1198	0.0468	0.0048	0.0063	0.1034	0.0479
WFE	**5**	**800**	**Bias**	0.0005	-0.0014	0.0578	-0.0963	0.0005	-0.0010	0.0596	-0.0934	0.0008	-0.0007	0.2741	-0.7008	0.0009	0.0000	0.1774	-0.2144
			RMSE	0.0036	0.0052	0.0619	0.0987	0.0036	0.0051	0.0640	0.0960	0.0036	0.0051	0.2809	0.7057	0.0037	0.0052	0.1882	0.2259
			Mbias	0.0006	-0.0016	0.0577	-0.0955	0.0006	-0.0013	0.0591	-0.0928	0.0010	-0.0010	0.2727	-0.6990	0.0009	-0.0002	0.1755	-0.2165
			Mabias	0.0023	0.0036	0.0577	0.0955	0.0023	0.0036	0.0591	0.0928	0.0024	0.0035	0.2727	0.6990	0.0023	0.0034	0.1755	0.2165
TWFE	**5**	**800**	**Bias**	0.0000	-0.0002	0.0001	-0.0006	0.0000	-0.0002	0.0001	-0.0006	0.0000	-0.0002	-0.0028	-0.0005	0.0000	0.0001	-0.0048	-0.0016
			RMSE	0.0035	0.0050	0.0182	0.0139	0.0035	0.0050	0.0186	0.0140	0.0035	0.0050	0.0721	0.0197	0.0036	0.0052	0.0740	0.0195
			Mbias	0.0001	-0.0004	0.0004	-0.0005	0.0001	-0.0004	0.0002	-0.0006	0.0001	-0.0004	-0.0018	-0.0001	0.0000	-0.0002	-0.0063	-0.0018
			Mabias	0.0024	0.0035	0.0127	0.0093	0.0024	0.0035	0.0130	0.0096	0.0024	0.0035	0.0485	0.0134	0.0023	0.0034	0.0507	0.0134
PWFE	**5**	**800**	**Bias**	0.0000	-0.0002	0.0011	-0.0019	0.0000	-0.0002	0.0000	-0.0018	0.0000	-0.0002	-0.0079	-0.0039	0.0000	0.0001	-0.0092	-0.0041
			RMSE	0.0035	0.0050	0.0253	0.0200	0.0035	0.0050	0.0240	0.0203	0.0035	0.0050	0.0998	0.0528	0.0036	0.0052	0.1058	0.0449
			Mbias	0.0001	-0.0004	0.0020	-0.0002	0.0001	-0.0004	-0.0008	-0.0010	0.0001	-0.0004	-0.0135	0.0030	0.0000	-0.0002	-0.0100	-0.0024
			Mabias	0.0024	0.0035	0.0166	0.0135	0.0024	0.0035	0.0153	0.0136	0.0024	0.0035	0.0689	0.0327	0.0024	0.0034	0.0681	0.0305
$\alpha_1 = 10$				**Exponential: row normalized**				**Exponential: scalar normalized**				**Inverse-dis: row normalized**				**Inverse-dis: scalar normalized**			
	T	**N**		DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Bias**	0.0118	-0.0058	0.1635	0.0534	0.0044	-0.0068	0.2449	0.0178	0.0205	0.0075	2.5410	0.7192	0.0216	0.0049	2.4742	0.5747
			RMSE	0.0139	0.0119	0.1691	0.0854	0.0085	0.0123	0.2486	0.0686	0.0217	0.0124	2.5429	0.7340	0.0227	0.0110	2.4761	0.5922
			Mbias	0.0116	-0.0056	0.1640	0.0530	0.0041	-0.0068	0.2441	0.0179	0.0202	0.0075	2.5397	0.7156	0.0214	0.0049	2.4754	0.5726
.			**Mabias**	0.0116	0.0084	0.1640	0.0624	0.0059	0.0086	0.2441	0.0481	0.0202	0.0087	2.5397	0.7156	0.0214	0.0074	2.4754	0.5726
TWFE	**5**	**200**	**Bias**	-0.0005	0.0003	-0.0034	0.0010	-0.0005	0.0003	-0.0042	0.0010	-0.0004	0.0003	-0.0024	0.0018	-0.0004	0.0003	-0.0042	0.0020
			RMSE	0.0073	0.0103	0.0555	0.0486	0.0072	0.0102	0.0551	0.0466	0.0071	0.0099	0.0927	0.0656	0.0070	0.0099	0.0895	0.0633
			Mbias	-0.0007	0.0003	-0.0044	0.0015	-0.0007	0.0002	-0.0069	0.0010	-0.0006	0.0002	-0.0042	0.0036	-0.0006	0.0002	-0.0106	0.0020
			Mabias	0.0051	0.0067	0.0387	0.0301	0.0051	0.0067	0.0413	0.0294	0.0049	0.0066	0.0617	0.0412	0.0049	0.0066	0.0624	0.0413
PWFE	**5**	**200**	**Bias**	-0.0004	0.0004	-0.0020	-0.0034	-0.0004	0.0004	0.0017	-0.0035	-0.0004	0.0004	-0.0070	-0.0067	-0.0004	0.0004	0.0070	-0.0084
			RMSE	0.0073	0.0104	0.0661	0.0695	0.0073	0.0102	0.0577	0.0672	0.0071	0.0099	0.1776	0.1278	0.0071	0.0099	0.0989	0.1191
			Mbias	-0.0007	0.0003	-0.0056	0.0010	-0.0006	0.0005	-0.0003	0.0009	-0.0006	0.0002	-0.0362	0.0028	-0.0005	0.0002	0.0009	-0.0019
			Mabias	0.0051	0.0069	0.0445	0.0460	0.0051	0.0069	0.0409	0.0445	0.0050	0.0067	0.1225	0.0903	0.0050	0.0065	0.0630	0.0872
WFE	**5**	**800**	**Bias**	0.0138	0.0024	0.2514	-0.0010	0.0049	0.0022	0.2829	-0.0435	0.0095	-0.0010	3.1200	-1.4779	0.0080	-0.0019	3.1323	-1.5591
			RMSE	0.0143	0.0060	0.2526	0.0370	0.0062	0.0057	0.2840	0.0562	0.0101	0.0052	3.1219	1.4835	0.0088	0.0054	3.1342	1.5645
			Mbias	0.0138	0.0022	0.2502	0.0005	0.0051	0.0021	0.2827	-0.0432	0.0096	-0.0011	3.1185	-1.4803	0.0081	-0.0020	3.1334	-1.5607
			Mabias	0.0138	0.0038	0.2502	0.0252	0.0051	0.0038	0.2827	0.0440	0.0096	0.0036	3.1185	1.4803	0.0081	0.0037	3.1334	1.5607
TWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0008	-0.0021	0.0000	-0.0002	0.0000	-0.0013	0.0000	-0.0002	0.0007	-0.0021	0.0000	-0.0002	0.0009	-0.0022
			RMSE	0.0038	0.0055	0.0333	0.0275	0.0037	0.0053	0.0313	0.0267	0.0036	0.0051	0.0668	0.0410	0.0036	0.0051	0.0664	0.0410
			Mbias	0.0000	-0.0003	-0.0015	-0.0018	0.0001	-0.0003	0.0001	-0.0012	0.0001	-0.0004	-0.0017	-0.0023	0.0001	-0.0003	-0.0009	-0.0010
			Mabias	0.0024	0.0038	0.0223	0.0179	0.0024	0.0035	0.0211	0.0186	0.0024	0.0035	0.0450	0.0261	0.0024	0.0035	0.0443	0.0265
PWFE	**5**	**800**	**Bias**	0.0000	0.0000	-0.0011	-0.0025	0.0000	-0.0002	0.0040	-0.0048	0.0000	-0.0002	0.0023	-0.0131	0.0000	-0.0002	0.0099	-0.0188
			RMSE	0.0038	0.0055	0.0417	0.0395	0.0037	0.0053	0.0334	0.0392	0.0036	0.0051	0.1281	0.1028	0.0036	0.0051	0.0750	0.1052
			Mbias	0.0000	-0.0002	-0.0040	0.0012	0.0002	-0.0003	0.0043	-0.0027	0.0001	-0.0004	-0.0079	0.0026	0.0001	-0.0003	0.0049	-0.0035
			Mabias	0.0024	0.0037	0.0283	0.0262	0.0024	0.0035	0.0228	0.0264	0.0024	0.0035	0.0881	0.0656	0.0024	0.0035	0.0499	0.0665

Table A.14. Case II: Mean and standard deviation of the p-values of the parameters (Table A.5) and the direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 (Table A.11, upper panel)

Settings				Negative exponential: row normalized											
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.25)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0000	0.2838	0.0000	0.0508	0.0000				0.5071	0.5026	0.0834	0.1383
			Std	0.0002	0.2849	0.0000	0.1438	0.0000				0.2854	0.2899	0.1988	0.2260
TWFE	**5**	**200**	**Mean**	0.4960	0.4909	0.4951	0.4945	0.4780				0.4920	0.4944	0.4901	0.4952
			Std	0.2881	0.2929	0.2919	0.2950	0.2861				0.2806	0.2935	0.2874	0.2940
PWFE	**5**	**200**	**Mean**	0.4805	0.4868	0.4866	0.4837	0.4644	0.5106	0.4746	0.4869	0.4920	0.4917	0.5022	0.4855
			Std	0.3035	0.2909	0.2825	0.2864	0.2965	0.2864	0.2998	0.2830	0.2807	0.2926	0.2925	0.2895
WFE	**5**	**800**	**Mean**	0.0000	0.1271	0.0000	0.0000	0.0000				0.5124	0.5062	0.0022	0.0009
			Std	0.0000	0.2023	0.0000	0.0001	0.0000				0.2939	0.2943	0.0339	0.0124
TWFE	**5**	**800**	**Mean**	0.5009	0.5017	0.5007	0.4983	0.4912				0.5069	0.4991	0.4889	0.4925
			Std	0.2887	0.2883	0.2881	0.2924	0.2940				0.3000	0.2907	0.2914	0.2920
PWFE	**5**	**800**	**Mean**	0.4949	0.5028	0.4857	0.4983	0.4807	0.5016	0.4794	0.4870	0.5054	0.4982	0.4994	0.4952
			Std	0.2971	0.2890	0.2900	0.2902	0.2966	0.2918	0.2981	0.2906	0.3003	0.2909	0.2917	0.2896
				Negative exponential: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.25)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0003	0.2902	0.0000	0.0353	0.0000				0.5058	0.5000	0.1564	0.1088
			Std	0.0026	0.2880	0.0000	0.1144	0.0000				0.2790	0.2901	0.2450	0.2055
TWFE	**5**	**200**	**Mean**	0.4918	0.4925	0.4863	0.4983	0.4897				0.4924	0.4937	0.4806	0.5003
			Std	0.2874	0.2937	0.2867	0.2948	0.2913				0.2808	0.2926	0.2867	0.2969
PWFE	**5**	**200**	**Mean**	0.5304	0.4933	0.5668	0.4931	0.6880	0.5123	0.5431	0.5018	0.4925	0.4922	0.5009	0.4887
			Std	0.2820	0.2906	0.2687	0.2851	0.2126	0.2848	0.2870	0.2804	0.2804	0.2910	0.2863	0.2884
WFE	**5**	**800**	**Mean**	0.0000	0.1307	0.0000	0.0000	0.0000				0.5063	0.5097	0.0010	0.0017
			Std	0.0000	0.2051	0.0000	0.0004	0.0000				0.2940	0.2887	0.0149	0.0215
TWFE	**5**	**800**	**Mean**	0.5076	0.5015	0.4949	0.4959	0.4894				0.5065	0.4993	0.4851	0.4890
			Std	0.2947	0.2881	0.2898	0.2896	0.2887				0.2996	0.2905	0.2901	0.2884
PWFE	**5**	**800**	**Mean**	0.5451	0.5055	0.5615	0.5018	0.6856	0.5241	0.5372	0.4900	0.5050	0.4987	0.4945	0.5010
			Std	0.2839	0.2883	0.2792	0.2903	0.2134	0.2753	0.2702	0.2879	0.2999	0.2906	0.2944	0.2923
				Inverse distance: row normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.25)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0969	0.4300	0.0125	0.1316	0.0022				0.5006	0.5018	0.2903	0.2185
			Std	0.1831	0.3025	0.0673	0.2356	0.0132				0.2786	0.2887	0.3146	0.2902
TWFE	**5**	**200**	**Mean**	0.4912	0.4930	0.4874	0.4900	0.4853				0.4951	0.4937	0.5074	0.4976
			Std	0.2811	0.2925	0.2911	0.2927	0.2955				0.2806	0.2921	0.2900	0.2948
PWFE	**5**	**200**	**Mean**	0.4487	0.4856	0.4323	0.4511	0.4100	0.4739	0.3696	0.4727	0.4976	0.4957	0.4627	0.4398
			Std	0.3062	0.2925	0.3039	0.3055	0.3378	0.3062	0.3161	0.2954	0.2873	0.2917	0.2996	0.3114
WFE	**5**	**800**	**Mean**	0.0975	0.3784	0.0003	0.0007	0.0000				0.5130	0.5074	0.2054	0.0058
			Std	0.1717	0.2996	0.0054	0.0111	0.0000				0.2961	0.2882	0.2789	0.0521
TWFE	**5**	**800**	**Mean**	0.5103	0.5016	0.4980	0.4992	0.4939				0.5063	0.5011	0.5014	0.4938
			Std	0.2981	0.2900	0.2931	0.2904	0.2888				0.2995	0.2909	0.2930	0.2843
PWFE	**5**	**800**	**Mean**	0.4817	0.4951	0.4514	0.4808	0.4421	0.4993	0.4399	0.4788	0.5047	0.5007	0.4724	0.4791
			Std	0.2957	0.2931	0.3070	0.2945	0.3138	0.2829	0.3133	0.2930	0.3000	0.2912	0.2977	0.2988
				Inverse distance: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.25)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.1202	0.4106	0.0478	0.0866	0.0031				0.5000	0.4996	0.2646	0.1238
			Std	0.2045	0.3001	0.1464	0.1933	0.0181				0.2804	0.2888	0.2989	0.2277
TWFE	**5**	**200**	**Mean**	0.4909	0.4945	0.4846	0.4928	0.4884				0.4940	0.4941	0.4848	0.4976
			Std	0.2807	0.2923	0.2892	0.2921	0.2981				0.2809	0.2915	0.2915	0.2952
PWFE	**5**	**200**	**Mean**	0.5037	0.4931	0.4984	0.5014	0.6917	0.5846	0.4611	0.5351	0.4971	0.4953	0.4844	0.4801
			Std	0.2886	0.2908	0.2881	0.2924	0.2220	0.2477	0.2918	0.2664	0.2871	0.2903	0.3015	0.2985
WFE	**5**	**800**	**Mean**	0.1363	0.4035	0.0016	0.0013	0.0000				0.5117	0.5065	0.2216	0.0053
			Std	0.2078	0.3001	0.0185	0.0200	0.0000				0.2966	0.2888	0.2951	0.0459
TWFE	**5**	**800**	**Mean**	0.5101	0.5010	0.4890	0.4999	0.4921				0.5062	0.5010	0.4881	0.4925
			Std	0.2977	0.2904	0.2921	0.2901	0.2938				0.2996	0.2908	0.2972	0.2835
PWFE	**5**	**800**	**Mean**	0.5078	0.5026	0.5325	0.4955	0.7349	0.6102	0.5183	0.4983	0.5045	0.5007	0.4831	0.4856
			Std	0.2927	0.2895	0.2800	0.2886	0.1862	0.2485	0.2860	0.2838	0.2998	0.2910	0.3005	0.2885

Table A.15. Case III: Mean and standard deviation of the p-values of the parameters (Table A.6) and the direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 (Table A.12, lower panel)

Settings				Negative exponential: row normalized											
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(-0.25)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0000	0.2803	0.0000	0.0580	0.0000				0.4947	0.4973	0.0131	0.2076
			Std	0.0000	0.2801	0.0000	0.1549	0.0000				0.2980	0.2832	0.0593	0.2657
TWFE	**5**	**200**	**Mean**	0.4962	0.4911	0.4937	0.4941	0.4855				0.4966	0.4955	0.4934	0.4947
			Std	0.2869	0.2935	0.2910	0.2944	0.2925				0.2852	0.2924	0.2888	0.2952
PWFE	**5**	**200**	**Mean**	0.4832	0.4858	0.4951	0.4863	0.4582	0.4930	0.4856	0.4902	0.4976	0.4939	0.5016	0.4871
			Std	0.3060	0.2919	0.2882	0.2874	0.2937	0.2945	0.2907	0.2844	0.2856	0.2905	0.2878	0.2889
WFE	**5**	**800**	**Mean**	0.0000	0.2101	0.0000	0.0000	0.0000				0.4386	0.5309	0.0000	0.0008
			Std	0.0000	0.2534	0.0000	0.0001	0.0000				0.2953	0.2864	0.0000	0.0094
TWFE	**5**	**800**	**Mean**	0.4980	0.5018	0.5002	0.4976	0.4886				0.5084	0.5056	0.4936	0.5007
			Std	0.2879	0.2880	0.2899	0.2922	0.2936				0.2976	0.2926	0.2893	0.2935
PWFE	**5**	**800**	**Mean**	0.4909	0.5040	0.4904	0.4982	0.4806	0.4961	0.4826	0.4861	0.5070	0.5042	0.4934	0.4963
			Std	0.2938	0.2895	0.2910	0.2901	0.2943	0.2882	0.2967	0.2886	0.2973	0.2925	0.2893	0.2879
				Negative exponential: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(-0.25)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0001	0.3240	0.0000	0.0431	0.0000				0.5155	0.4979	0.0097	0.1595
			Std	0.0005	0.2871	0.0000	0.1306	0.0000				0.2860	0.2845	0.0497	0.2372
TWFE	**5**	**200**	**Mean**	0.4929	0.4927	0.4835	0.4987	0.4948				0.4959	0.4955	0.4818	0.4981
			Std	0.2879	0.2942	0.2852	0.2955	0.2917				0.2842	0.2925	0.2869	0.2942
PWFE	**5**	**200**	**Mean**	0.5568	0.4988	0.5477	0.4942	0.7449	0.5834	0.5631	0.5055	0.4954	0.4934	0.5036	0.4933
			Std	0.2775	0.2921	0.2760	0.2827	0.1762	0.2623	0.2810	0.2795	0.2842	0.2893	0.2838	0.2856
WFE	**5**	**800**	**Mean**	0.0000	0.2202	0.0000	0.0000	0.0000				0.4313	0.5229	0.0000	0.0012
			Std	0.0000	0.2586	0.0000	0.0001	0.0000				0.2954	0.2853	0.0000	0.0148
TWFE	**5**	**800**	**Mean**	0.5038	0.5017	0.4942	0.4946	0.4898				0.5085	0.5054	0.4875	0.4971
			Std	0.2938	0.2883	0.2913	0.2889	0.2904				0.2972	0.2924	0.2878	0.2907
PWFE	**5**	**800**	**Mean**	0.5787	0.5083	0.5386	0.5017	0.7505	0.5843	0.5529	0.4900	0.5072	0.5047	0.4923	0.5014
			Std	0.2741	0.2882	0.2852	0.2906	0.1781	0.2546	0.2693	0.2888	0.2973	0.2928	0.2947	0.2895
				Inverse distance: row normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(-0.25)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.2354	0.5114	0.0000	0.3808	0.0000				0.5073	0.5152	0.0003	0.4129
			Std	0.2482	0.2903	0.0000	0.2983	0.0000				0.2899	0.2886	0.0038	0.2959
TWFE	**5**	**200**	**Mean**	0.4920	0.4922	0.4907	0.4902	0.4880				0.4957	0.4948	0.4955	0.4953
			Std	0.2839	0.2925	0.2911	0.2940	0.2961				0.2837	0.2925	0.2885	0.2981
PWFE	**5**	**200**	**Mean**	0.4726	0.4863	0.4547	0.4589	0.4080	0.4282	0.4195	0.4714	0.4965	0.4989	0.4564	0.4514
			Std	0.3058	0.2903	0.2979	0.3047	0.3075	0.3254	0.3065	0.2963	0.2858	0.2912	0.3149	0.3101
WFE	**5**	**800**	**Mean**	0.3968	0.5207	0.0000	0.0267	0.0000				0.4623	0.5171	0.0000	0.0636
			Std	0.2923	0.2890	0.0000	0.0983	0.0001				0.2946	0.2860	0.0000	0.1542
TWFE	**5**	**800**	**Mean**	0.5117	0.5016	0.4982	0.4984	0.5000				0.5077	0.5040	0.4981	0.4977
			Std	0.2980	0.2896	0.2915	0.2893	0.2917				0.2983	0.2920	0.2909	0.2846
PWFE	**5**	**800**	**Mean**	0.4765	0.4965	0.4611	0.4821	0.4197	0.4572	0.4421	0.4796	0.5059	0.5028	0.4741	0.4842
			Std	0.2954	0.2886	0.3001	0.2909	0.3120	0.3089	0.3111	0.2883	0.2983	0.2923	0.3044	0.2978
				Inverse distance: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(-0.25)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.2183	0.5113	0.0000	0.2956	0.0001				0.4974	0.5121	0.0011	0.3463
			Std	0.2421	0.2861	0.0000	0.3013	0.0000				0.2976	0.2880	0.0076	0.3074
TWFE	**5**	**200**	**Mean**	0.4908	0.4936	0.4877	0.4935	0.4948				0.4952	0.4951	0.4882	0.4942
			Std	0.2826	0.2922	0.2906	0.2932	0.2981				0.2832	0.2923	0.2902	0.2933
PWFE	**5**	**200**	**Mean**	0.5329	0.4967	0.5108	0.5031	0.7507	0.6634	0.5091	0.5342	0.4962	0.4975	0.4902	0.4832
			Std	0.2830	0.2905	0.2901	0.2931	0.1830	0.2326	0.2881	0.2708	0.2845	0.2904	0.3028	0.2999
WFE	**5**	**800**	**Mean**	0.3977	0.5165	0.0000	0.0210	0.0000				0.4731	0.5194	0.0000	0.0410
			Std	0.2930	0.2910	0.0000	0.0884	0.0000				0.2942	0.2884	0.0000	0.1267
TWFE	**5**	**800**	**Mean**	0.5114	0.5009	0.4891	0.4991	0.4952				0.5079	0.5039	0.4897	0.4973
			Std	0.2980	0.2896	0.2897	0.2893	0.2910				0.2983	0.2921	0.2936	0.2852
PWFE	**5**	**800**	**Mean**	0.5159	0.5036	0.5178	0.4957	0.7922	0.6830	0.5245	0.4993	0.5060	0.5031	0.4816	0.4882
			Std	0.2893	0.2889	0.2901	0.2877	0.1552	0.2093	0.2878	0.2834	0.2982	0.2926	0.3013	0.2890

Table A.16. Case IV: Mean and standard deviation of the p-values of the parameters (Table A.7) and the direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 (Table A.14, upper panel)

Settings				Negative exponential: row normalized										
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.01)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0000	0.2272	0.0000	0.0503	0.0000				0.5113	0.5045	0.1522	0.1736
			Std	0.0000	0.2680	0.0000	0.1456	0.0000				0.2886	0.2878	0.2498	0.2471
TWFE	**5**	**200**	**Mean**	0.4964	0.4910	0.4944	0.4944	0.4809				0.4948	0.4947	0.4901	0.4947
			Std	0.2880	0.2932	0.2919	0.2949	0.2882				0.2835	0.2924	0.2862	0.2944
PWFE	**5**	**200**	**Mean**	0.4308	0.4824	0.4433	0.4862	0.4053	0.5976	0.4258	0.4864	0.4948	0.4927	0.5282	0.4876
			Std	0.3027	0.2922	0.2898	0.2864	0.2557	0.3575	0.2990	0.2810	0.2837	0.2906	0.2786	0.2915
WFE	**5**	**800**	**Mean**	0.0000	0.0343	0.0000	0.0000	0.0000				0.5113	0.5067	0.0195	0.0002
			Std	0.0000	0.0914	0.0000	0.0000	0.0000				0.2820	0.2806	0.0890	0.0023
TWFE	**5**	**800**	**Mean**	0.4955	0.5053	0.4963	0.4929	0.4964				0.4991	0.5010	0.4963	0.4945
			Std	0.2927	0.2820	0.2902	0.2934	0.2891				0.2863	0.2804	0.2910	0.2949
PWFE	**5**	**800**	**Mean**	0.4684	0.5027	0.4692	0.4970	0.4809	0.7440	0.4521	0.5005	0.4975	0.5010	0.5040	0.5056
			Std	0.2947	0.2859	0.2888	0.2817	0.2875	0.2948	0.2966	0.2922	0.2868	0.2802	0.2852	0.2859
				Negative exponential: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.01)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0000	0.2206	0.0000	0.0244	0.0000				0.5138	0.4970	0.2378	0.1108
			Std	0.0001	0.2652	0.0000	0.0917	0.0000				0.2797	0.2883	0.2966	0.2073
TWFE	**5**	**200**	**Mean**	0.4924	0.4926	0.4847	0.4987	0.4917				0.4946	0.4943	0.4816	0.4994
			Std	0.2880	0.2940	0.2860	0.2955	0.2906				0.2832	0.2921	0.2876	0.2958
PWFE	**5**	**200**	**Mean**	0.5529	0.4948	0.6175	0.4913	0.6534	0.7695	0.7071	0.5051	0.4941	0.4939	0.5579	0.4898
			Std	0.2706	0.2884	0.2545	0.2822	0.2126	0.2461	0.2146	0.2761	0.2831	0.2896	0.2652	0.2860
WFE	**5**	**800**	**Mean**	0.0000	0.0558	0.0000	0.0000	0.0000				0.5094	0.5090	0.0200	0.0007
			Std	0.0000	0.1238	0.0000	0.0000	0.0000				0.2823	0.2787	0.0947	0.0072
TWFE	**5**	**800**	**Mean**	0.4992	0.5044	0.4899	0.4964	0.4994				0.4990	0.5011	0.4876	0.4981
			Std	0.2956	0.2819	0.2842	0.2953	0.2952				0.2862	0.2804	0.2805	0.2963
PWFE	**5**	**800**	**Mean**	0.5525	0.5093	0.6048	0.5030	0.6667	0.7423	0.6908	0.5068	0.4975	0.5009	0.5705	0.5142
			Std	0.2742	0.2819	0.2511	0.2841	0.2151	0.2576	0.2193	0.2809	0.2866	0.2801	0.2596	0.2826
				Inverse distance: row normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.01)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0704	0.4715	0.0001	0.1569	0.0000				0.5073	0.5084	0.2439	0.2680
			Std	0.1414	0.2933	0.0011	0.2465	0.0000				0.2802	0.2888	0.2676	0.3029
TWFE	**5**	**200**	**Mean**	0.4917	0.4925	0.4895	0.4900	0.4869				0.4951	0.4942	0.5000	0.4972
			Std	0.2827	0.2924	0.2916	0.2933	0.2953				0.2826	0.2921	0.2905	0.2972
PWFE	**5**	**200**	**Mean**	0.4569	0.4911	0.4600	0.4508	0.3345	0.4362	0.4324	0.4649	0.4948	0.4969	0.4755	0.4438
			Std	0.2975	0.2942	0.2921	0.3025	0.2360	0.3423	0.2948	0.2979	0.2842	0.2913	0.3084	0.3090
WFE	**5**	**800**	**Mean**	0.1235	0.3991	0.0000	0.0000	0.0000				0.4928	0.5038	0.0002	0.0000
			Std	0.1985	0.3044	0.0000	0.0000	0.0000				0.2866	0.2788	0.0027	0.0000
TWFE	**5**	**800**	**Mean**	0.4993	0.5035	0.4801	0.4936	0.4866				0.4990	0.5009	0.4862	0.5021
			Std	0.2908	0.2818	0.2983	0.2899	0.2931				0.2862	0.2803	0.2980	0.2949
PWFE	**5**	**800**	**Mean**	0.4789	0.5013	0.3992	0.4997	0.3828	0.4601	0.3754	0.5015	0.4969	0.5007	0.4667	0.4933
			Std	0.2967	0.2820	0.2914	0.2833	0.2595	0.3394	0.2905	0.2905	0.2866	0.2805	0.2921	0.2979
				Inverse distance: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.01)	δ(2)	α_1(1.5)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0876	0.4496	0.0015	0.0793	0.0000				0.5049	0.5042	0.3256	0.1394
			Std	0.1631	0.2956	0.0133	0.1760	0.0004				0.2828	0.2893	0.3031	0.2317
TWFE	**5**	**200**	**Mean**	0.4911	0.4940	0.4860	0.4933	0.4922				0.4946	0.4947	0.4856	0.4963
			Std	0.2821	0.2922	0.2895	0.2926	0.2983				0.2824	0.2921	0.2910	0.2952
PWFE	**5**	**200**	**Mean**	0.5189	0.4954	0.5622	0.4969	0.6057	0.7405	0.5927	0.5292	0.4954	0.4961	0.5346	0.4748
			Std	0.2792	0.2889	0.2744	0.2939	0.2213	0.2247	0.2480	0.2711	0.2831	0.2896	0.2908	0.2998
WFE	**5**	**800**	**Mean**	0.1666	0.4761	0.0000	0.0032	0.0000				0.5033	0.5112	0.0250	0.0113
			Std	0.2305	0.2860	0.0000	0.0335	0.0000				0.2927	0.2893	0.0817	0.0707
TWFE	**5**	**800**	**Mean**	0.5107	0.5010	0.4891	0.4995	0.4946				0.5071	0.5026	0.4893	0.4947
			Std	0.2977	0.2899	0.2906	0.2896	0.2929				0.2988	0.2916	0.2949	0.2843
PWFE	**5**	**800**	**Mean**	0.5162	0.5016	0.5816	0.4935	0.6249	0.7361	0.6021	0.4984	0.5053	0.5018	0.5103	0.4808
			Std	0.2924	0.2901	0.2681	0.2878	0.2257	0.2185	0.2563	0.2838	0.2988	0.2918	0.2971	0.2870

Table A.17. Case V: Mean and standard deviation of the p-values of the parameters (Table A.8) and the direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 (Table A.13, upper panel)

Settings				Negative exponential: row normalized											
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(1.5)	α_2(1)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0046	0.4891	0.0000	0.2841	0.0054				0.4899	0.4755	0.0032	0.1836
			Std	0.0224	0.2926	0.0000	0.2975	0.0277				0.2901	0.2938	0.0318	0.2616
TWFE	**5**	**200**	**Mean**	0.4952	0.4923	0.4934	0.4820	0.4811				0.4893	0.4912	0.4926	0.4863
			Std	0.2887	0.2920	0.2905	0.2936	0.2897				0.2797	0.2914	0.2918	0.2942
PWFE	**5**	**200**	**Mean**	0.4878	0.4924	0.4779	0.4443	0.4571	0.4983	0.4643	0.4692	0.4883	0.4962	0.5058	0.4443
			Std	0.3058	0.2937	0.2913	0.2934	0.2944	0.2869	0.3000	0.3012	0.2801	0.2923	0.3002	0.2937
WFE	**5**	**800**	**Mean**	0.0000	0.4556	0.0000	0.2594	0.0000				0.4839	0.4953	0.0000	0.0983
			Std	0.0000	0.2909	0.0000	0.2872	0.0000				0.2977	0.2926	0.0000	0.1956
TWFE	**5**	**800**	**Mean**	0.5054	0.5024	0.5026	0.4983	0.4910				0.5049	0.4969	0.4885	0.4937
			Std	0.2927	0.2909	0.2880	0.2858	0.2927				0.3011	0.2898	0.2906	0.2859
PWFE	**5**	**800**	**Mean**	0.4981	0.5027	0.4840	0.4812	0.4783	0.4932	0.4827	0.4933	0.5039	0.4990	0.4976	0.4825
			Std	0.2963	0.2917	0.2899	0.2915	0.2941	0.2923	0.3018	0.2941	0.3023	0.2897	0.2892	0.2940
				Negative exponential: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(1.5)	α_2(1)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0197	0.4927	0.0000	0.2838	0.0292				0.4898	0.4766	0.0042	0.1957
			Std	0.0726	0.2925	0.0000	0.3028	0.0903				0.2930	0.2932	0.0356	0.2670
TWFE	**5**	**200**	**Mean**	0.4912	0.4930	0.4870	0.4811	0.4868				0.4904	0.4906	0.4765	0.4835
			Std	0.2868	0.2922	0.2909	0.2902	0.2937				0.2791	0.2899	0.2867	0.2884
PWFE	**5**	**200**	**Mean**	0.5218	0.4937	0.5617	0.4790	0.6493	0.5010	0.5285	0.5102	0.4914	0.4977	0.4990	0.4697
			Std	0.2855	0.2914	0.2716	0.2920	0.2338	0.2805	0.2826	0.2899	0.2802	0.2917	0.2916	0.2956
WFE	**5**	**800**	**Mean**	0.0000	0.4549	0.0000	0.2503	0.0000				0.4823	0.4934	0.0000	0.1154
			Std	0.0001	0.2917	0.0000	0.2827	0.0004				0.2983	0.2893	0.0000	0.2110
TWFE	**5**	**800**	**Mean**	0.5098	0.5025	0.4969	0.5031	0.4877				0.5034	0.4965	0.4821	0.4980
			Std	0.2963	0.2909	0.2897	0.2898	0.2873				0.2996	0.2883	0.2869	0.2891
PWFE	**5**	**800**	**Mean**	0.5273	0.5025	0.5677	0.4855	0.6477	0.5017	0.5309	0.5031	0.5024	0.4985	0.4943	0.4811
			Std	0.2875	0.2913	0.2735	0.2865	0.2304	0.2794	0.2745	0.2906	0.3008	0.2890	0.2938	0.2889
				Inverse distance: row normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(1.5)	α_2(1)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.3237	0.4890	0.0002	0.1379	0.4117				0.3332	0.4681	0.0179	0.1219
			Std	0.2915	0.2941	0.0037	0.2272	0.3143				0.2957	0.2966	0.0795	0.2080
TWFE	**5**	**200**	**Mean**	0.4935	0.4943	0.4772	0.4722	0.4736				0.4964	0.4910	0.4862	0.4730
			Std	0.2828	0.2924	0.2906	0.2957	0.2913				0.2814	0.2905	0.2974	0.2913
PWFE	**5**	**200**	**Mean**	0.4891	0.4974	0.4473	0.4579	0.4258	0.4742	0.4329	0.5445	0.4921	0.5005	0.4428	0.4545
			Std	0.2947	0.2919	0.2975	0.3080	0.3171	0.2769	0.3026	0.2528	0.2886	0.2916	0.3093	0.3117
WFE	**5**	**800**	**Mean**	0.3505	0.5021	0.0000	0.1760	0.3167				0.4991	0.5003	0.0139	0.1276
			Std	0.2917	0.2913	0.0003	0.2554	0.2983				0.2959	0.2907	0.0692	0.2149
TWFE	**5**	**800**	**Mean**	0.5088	0.5027	0.4872	0.4962	0.4894				0.5055	0.5011	0.4842	0.4964
			Std	0.2977	0.2919	0.2921	0.2926	0.2936				0.3007	0.2905	0.2908	0.2928
PWFE	**5**	**800**	**Mean**	0.4992	0.5018	0.4543	0.4534	0.4363	0.4690	0.4439	0.5071	0.5029	0.5001	0.4533	0.4494
			Std	0.2955	0.2923	0.2996	0.2999	0.2987	0.2877	0.3017	0.2735	0.3017	0.2901	0.3043	0.3043
				Inverse distance: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(1.5)	α_2(1)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.3386	0.4927	0.0012	0.2349	0.4318				0.3734	0.4797	0.0330	0.2146
			Std	0.2966	0.2923	0.0119	0.2874	0.3012				0.3031	0.2922	0.1189	0.2761
TWFE	**5**	**200**	**Mean**	0.4928	0.4949	0.4708	0.4684	0.4668				0.4957	0.4920	0.4732	0.4726
			Std	0.2826	0.2918	0.2905	0.2936	0.2942				0.2814	0.2895	0.2987	0.2935
PWFE	**5**	**200**	**Mean**	0.4950	0.4964	0.4721	0.5164	0.6459	0.5690	0.4584	0.5873	0.4949	0.4991	0.4794	0.4946
			Std	0.2881	0.2898	0.2937	0.2954	0.2286	0.2662	0.2990	0.2392	0.2886	0.2886	0.2955	0.3020
WFE	**5**	**800**	**Mean**	0.3664	0.5013	0.0001	0.2013	0.2871				0.4968	0.4996	0.0201	0.1511
			Std	0.2933	0.2905	0.0016	0.2678	0.2870				0.2962	0.2899	0.0901	0.2319
TWFE	**5**	**800**	**Mean**	0.5092	0.5025	0.4814	0.4963	0.4861				0.5049	0.5010	0.4687	0.4972
			Std	0.2979	0.2921	0.2953	0.2901	0.2949				0.3004	0.2906	0.2885	0.2914
PWFE	**5**	**800**	**Mean**	0.5067	0.5017	0.5062	0.4872	0.6935	0.5738	0.4964	0.5415	0.5029	0.5001	0.4719	0.4720
			Std	0.2950	0.2921	0.2898	0.2917	0.2007	0.2654	0.2916	0.2648	0.3010	0.2895	0.3013	0.2974

Table A.18. Case VI: Mean and standard deviation of the p-values of the parameters (Table A.9) and the direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 (Table A.13, lower panel)

Settings				Negative exponential: row normalized											
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(1.5)	α_2(0.5)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0081	0.4917	0.0000	0.0041	0.0234				0.4876	0.4807	0.0010	0.0009
			Std	0.0360	0.2930	0.0000	0.0248	0.0762				0.2947	0.2921	0.0172	0.0062
TWFE	**5**	**200**	**Mean**	0.4960	0.4938	0.4920	0.4697	0.4824				0.4890	0.4899	0.4896	0.4739
			Std	0.2892	0.2923	0.2904	0.2906	0.2893				0.2799	0.2910	0.2932	0.2927
PWFE	**5**	**200**	**Mean**	0.4907	0.4935	0.4860	0.5077	0.4893	0.4967	0.4795	0.7983	0.4891	0.5095	0.5096	0.5023
			Std	0.3000	0.2928	0.2892	0.2874	0.2978	0.2846	0.2999	0.2802	0.2800	0.2909	0.2954	0.2875
WFE	**5**	**800**	**Mean**	0.0000	0.5046	0.0000	0.0000	0.0000				0.4755	0.4950	0.0000	0.0000
			Std	0.0000	0.2921	0.0000	0.0000	0.0001				0.2966	0.2905	0.0000	0.0000
TWFE	**5**	**800**	**Mean**	0.5062	0.5028	0.5022	0.4955	0.4897				0.5051	0.4999	0.4882	0.4930
			Std	0.2933	0.2920	0.2877	0.2931	0.2920				0.3013	0.2898	0.2908	0.2922
PWFE	**5**	**800**	**Mean**	0.4997	0.5017	0.4879	0.5553	0.4835	0.4960	0.4847	0.7459	0.5044	0.5178	0.5059	0.5509
			Std	0.2959	0.2918	0.2958	0.2870	0.2916	0.2923	0.3008	0.3105	0.3020	0.2832	0.2912	0.2862
				Negative exponential: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(1.5)	α_2(0.5)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0333	0.4894	0.0000	0.0063	0.1120				0.4814	0.4771	0.0004	0.0018
			Std	0.1037	0.2955	0.0000	0.0404	0.1988				0.3018	0.2934	0.0039	0.0140
TWFE	**5**	**200**	**Mean**	0.4929	0.4939	0.4836	0.4698	0.4842				0.4902	0.4903	0.4744	0.4736
			Std	0.2885	0.2922	0.2905	0.2889	0.2952				0.2795	0.2904	0.2863	0.2893
PWFE	**5**	**200**	**Mean**	0.5203	0.4940	0.5645	0.5113	0.6447	0.5062	0.5231	0.7955	0.4913	0.5087	0.5020	0.5124
			Std	0.2856	0.2920	0.2749	0.2858	0.2348	0.2783	0.2836	0.2709	0.2796	0.2907	0.2891	0.2870
WFE	**5**	**800**	**Mean**	0.0000	0.5039	0.0000	0.0000	0.0002				0.4751	0.4894	0.0000	0.0000
			Std	0.0001	0.2903	0.0000	0.0001	0.0016				0.2975	0.2891	0.0000	0.0000
TWFE	**5**	**800**	**Mean**	0.5100	0.5029	0.4964	0.4955	0.4892				0.5035	0.5001	0.4817	0.4928
			Std	0.2964	0.2921	0.2899	0.2907	0.2892				0.2997	0.2891	0.2848	0.2898
PWFE	**5**	**800**	**Mean**	0.5264	0.5016	0.5663	0.5613	0.6504	0.5121	0.5301	0.7461	0.5028	0.5149	0.5014	0.5568
			Std	0.2874	0.2921	0.2746	0.2880	0.2297	0.2806	0.2757	0.3061	0.3007	0.2828	0.2957	0.2886
				Inverse distance: row normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(1.5)	α_2(0.5)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.3462	0.4854	0.0001	0.0366	0.3014				0.2658	0.4615	0.0073	0.0363
			Std	0.2986	0.2955	0.0012	0.1112	0.3035				0.2787	0.2982	0.0368	0.1020
TWFE	**5**	**200**	**Mean**	0.4937	0.4950	0.4810	0.4701	0.4717				0.4968	0.4915	0.4861	0.4775
			Std	0.2826	0.2925	0.2974	0.2889	0.2947				0.2818	0.2905	0.3026	0.2921
PWFE	**5**	**200**	**Mean**	0.4874	0.4980	0.4310	0.4586	0.4417	0.4558	0.4220	0.7921	0.4822	0.4990	0.4294	0.4539
			Std	0.2947	0.2926	0.3023	0.3117	0.3072	0.2841	0.3081	0.2904	0.2892	0.2911	0.3143	0.3148
WFE	**5**	**800**	**Mean**	0.3634	0.5017	0.0000	0.0301	0.4385				0.4945	0.5004	0.0042	0.0210
			Std	0.2933	0.2939	0.0000	0.0972	0.3071				0.2958	0.2911	0.0339	0.0769
TWFE	**5**	**800**	**Mean**	0.5084	0.5027	0.4893	0.4874	0.4899				0.5062	0.5018	0.4881	0.4878
			Std	0.2977	0.2921	0.2940	0.2892	0.2928				0.3013	0.2910	0.2931	0.2884
PWFE	**5**	**800**	**Mean**	0.5019	0.5017	0.4535	0.4668	0.4673	0.4755	0.4482	0.7494	0.5023	0.5024	0.4594	0.4613
			Std	0.2967	0.2919	0.2962	0.2982	0.2973	0.2906	0.3016	0.3301	0.3010	0.2902	0.3005	0.3009
				Inverse distance: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(1.5)	α_2(0.5)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.3646	0.4904	0.0005	0.0884	0.3804				0.3161	0.4744	0.0137	0.0783
			Std	0.3025	0.2952	0.0049	0.1811	0.3200				0.2889	0.2951	0.0575	0.1626
TWFE	**5**	**200**	**Mean**	0.4930	0.4953	0.4718	0.4654	0.4615				0.4964	0.4926	0.4715	0.4703
			Std	0.2826	0.2921	0.2951	0.2888	0.2945				0.2822	0.2902	0.3024	0.2864
PWFE	**5**	**200**	**Mean**	0.4913	0.4965	0.4483	0.5001	0.6305	0.5581	0.4347	0.7604	0.4887	0.4988	0.4561	0.4832
			Std	0.2882	0.2903	0.2995	0.3080	0.2365	0.2743	0.3025	0.2990	0.2895	0.2885	0.3026	0.3119
WFE	**5**	**800**	**Mean**	0.3785	0.5016	0.0000	0.0412	0.4241				0.4921	0.4997	0.0050	0.0287
			Std	0.2947	0.2929	0.0002	0.1170	0.3043				0.2965	0.2903	0.0333	0.0975
TWFE	**5**	**800**	**Mean**	0.5088	0.5026	0.4845	0.4891	0.4876				0.5055	0.5016	0.4765	0.4894
			Std	0.2978	0.2922	0.2971	0.2883	0.2949				0.3008	0.2912	0.2916	0.2879
PWFE	**5**	**800**	**Mean**	0.5064	0.5018	0.5052	0.4866	0.6909	0.5793	0.4929	0.7311	0.5020	0.5025	0.4690	0.4757
			Std	0.2942	0.2921	0.2882	0.2904	0.2060	0.2667	0.2952	0.3332	0.3000	0.2902	0.2970	0.2935

Table A.19. Case VII: Mean and standard deviation of the p-values of the parameters (Table A.10) and the direct (DE) and indirect (IE) spillover effects of variables x_1 and x_2 (Table A.13, lower panel)

Settings				Negative exponential: row normalized											
	T	N		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(10)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0000	0.0331	0.0000	0.1792	0.0000				0.4215	0.6258	0.0021	0.2719
			Std	0.0000	0.0642	0.0000	0.2559	0.0000				0.2522	0.2407	0.0221	0.2955
TWFE	**5**	**200**	**Mean**	0.5008	0.4914	0.4919	0.4951	0.4852				0.4893	0.4934	0.4878	0.5015
			Std	0.2958	0.2918	0.2840	0.2945	0.2913				0.2792	0.2923	0.2921	0.2967
PWFE	**5**	**200**	**Mean**	0.4990	0.4882	0.4780	0.4837	0.4823	0.5114	0.6105	0.4903	0.4881	0.4912	0.4971	0.4871
			Std	0.2866	0.2889	0.2880	0.2872	0.2936	0.2944	0.3225	0.2825	0.2792	0.2917	0.2935	0.2896
WFE	**5**	**800**	**Mean**	0.0000	0.0027	0.0000	0.0029	0.0000				0.0804	0.6600	0.0000	0.3796
			Std	0.0000	0.0103	0.0000	0.0333	0.0000				0.0960	0.2337	0.0000	0.3084
TWFE	**5**	**800**	**Mean**	0.5081	0.5027	0.4957	0.4996	0.4884				0.5042	0.4946	0.4898	0.4932
			Std	0.2926	0.2899	0.2860	0.2934	0.2927				0.2999	0.2902	0.2920	0.2964
PWFE	**5**	**800**	**Mean**	0.4924	0.5042	0.5044	0.4986	0.4956	0.5038	0.5530	0.4854	0.5031	0.4954	0.4928	0.4937
			Std	0.2888	0.2894	0.2855	0.2892	0.2944	0.2892	0.2702	0.2867	0.3010	0.2901	0.2881	0.2896
				Negative exponential: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(10)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0000	0.0511	0.0000	0.1382	0.0000				0.6342	0.6050	0.0000	0.3511
			Std	0.0000	0.0866	0.0000	0.2242	0.0000				0.2458	0.2463	0.0000	0.3067
TWFE	**5**	**200**	**Mean**	0.4964	0.4927	0.4852	0.4981	0.4743				0.4903	0.4927	0.4723	0.5040
			Std	0.2968	0.2933	0.2836	0.2944	0.2894				0.2794	0.2922	0.2889	0.2963
PWFE	**5**	**200**	**Mean**	0.5140	0.4908	0.5620	0.4894	0.6074	0.7000	0.6866	0.4985	0.4895	0.4909	0.5346	0.4869
			Std	0.2929	0.2886	0.2628	0.2802	0.2516	0.2298	0.3791	0.2853	0.2797	0.2892	0.2688	0.2897
WFE	**5**	**800**	**Mean**	0.0000	0.0046	0.0000	0.0003	0.0000				0.4872	0.6715	0.0000	0.2423
			Std	0.0000	0.0123	0.0000	0.0061	0.0000				0.2538	0.2269	0.0000	0.2866
TWFE	**5**	**800**	**Mean**	0.4974	0.5029	0.4964	0.4957	0.4941				0.4958	0.5068	0.4913	0.5000
			Std	0.2947	0.2806	0.2912	0.2946	0.2896				0.2881	0.2842	0.2867	0.2958
PWFE	**5**	**800**	**Mean**	0.5140	0.5035	0.5928	0.5014	0.6186	0.7314	0.7161	0.5010	0.4939	0.5086	0.5459	0.5143
			Std	0.2860	0.2814	0.2593	0.2842	0.2448	0.2087	0.3574	0.2859	0.2880	0.2853	0.2687	0.2875
				Inverse distance: row normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(10)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0000	0.7403	0.0000	0.0013	0.0000				0.3045	0.7016	0.0000	0.0001
			Std	0.0000	0.1839	0.0000	0.0114	0.0000				0.1721	0.2014	0.0000	0.0008
TWFE	**5**	**200**	**Mean**	0.4934	0.4923	0.4897	0.4910	0.4933				0.4915	0.4944	0.4950	0.5069
			Std	0.2880	0.2918	0.2818	0.2946	0.2961				0.2799	0.2937	0.2957	0.2942
PWFE	**5**	**200**	**Mean**	0.4895	0.4902	0.4822	0.4636	0.4680	0.4903	0.5192	0.4862	0.4912	0.4934	0.4650	0.4529
			Std	0.2906	0.2919	0.2819	0.3068	0.3022	0.2832	0.2788	0.2895	0.2802	0.2925	0.3039	0.3060
WFE	**5**	**800**	**Mean**	0.0078	0.4196	0.0007	0.0000	0.0585				0.5073	0.7985	0.0004	0.0000
			Std	0.0504	0.1798	0.0113	0.0000	0.1219				0.1752	0.1418	0.0071	0.0012
TWFE	**5**	**800**	**Mean**	0.4911	0.5026	0.4929	0.4916	0.4905				0.4982	0.5030	0.4915	0.5013
			Std	0.2929	0.2804	0.2923	0.2879	0.2954				0.2882	0.2809	0.2967	0.2913
PWFE	**5**	**800**	**Mean**	0.5003	0.5041	0.5035	0.5059	0.4876	0.4999	0.4949	0.5095	0.4959	0.5028	0.4862	0.5050
			Std	0.3015	0.2810	0.2881	0.2889	0.2975	0.2813	0.2844	0.2878	0.2883	0.2803	0.2980	0.2918
				Inverse distance: scalar normalized											
	T	**N**		β_1(-1)	β_2(0.2)	ϕ_1(1.5)	ϕ_2(-0.3)	ρ(0.5)	δ(2)	α_1(10)	α_2(3)	DE_x_1	DE_x_2	IE_x_1	IE_x_2
WFE	**5**	**200**	**Mean**	0.0000	0.7411	0.0000	0.0108	0.0000				0.2601	0.7262	0.0000	0.0014
			Std	0.0000	0.1816	0.0000	0.0505	0.0000				0.1632	0.1935	0.0000	0.0119
TWFE	**5**	**200**	**Mean**	0.4909	0.4936	0.4851	0.4916	0.4821				0.4921	0.4935	0.4785	0.5051
			Std	0.2910	0.2929	0.2827	0.2940	0.2933				0.2800	0.2926	0.2884	0.2934
PWFE	**5**	**200**	**Mean**	0.5149	0.4932	0.5323	0.5221	0.6780	0.7172	0.6879	0.5385	0.4939	0.4946	0.6589	0.4852
			Std	0.2901	0.2908	0.2719	0.2775	0.2207	0.2202	0.3720	0.2702	0.2799	0.2906	0.2279	0.2891
WFE	**5**	**800**	**Mean**	0.0036	0.3993	0.0002	0.0001	0.0361				0.5395	0.7829	0.0001	0.0003
			Std	0.0329	0.1788	0.0055	0.0024	0.0888				0.1827	0.1510	0.0026	0.0062
TWFE	**5**	**800**	**Mean**	0.4909	0.5025	0.4914	0.4988	0.4858				0.4980	0.5034	0.4849	0.5076
			Std	0.2909	0.2806	0.2905	0.2958	0.2903				0.2882	0.2816	0.2869	0.2938
PWFE	**5**	**800**	**Mean**	0.5276	0.5031	0.5621	0.5103	0.6663	0.7880	0.7144	0.5093	0.4960	0.5043	0.6509	0.5251
			Std	0.2776	0.2807	0.2725	0.2895	0.2240	0.1804	0.3479	0.2854	0.2887	0.2818	0.2266	0.2831

Chapter 3

The spatial autoregressive panel data model with spatial moving average errors*

* This chapter was written with J.P. Elhorst, except for the derivations in the appendix.

3.1 Introduction

The spatial autoregressive (SAR) model, the spatial error (SE) model, and the so-called SARAR model combining these two specifications are the most widely used spatial econometric models to introduce new methods of estimation. Leading examples focusing on cross-sectional data are Ord (1975), who introduces the maximum likelihood (ML) estimator of SAR and SE models; Anselin (1988, pp.82-86) and Kelejian and Prucha (1998, 1999), who develop instrumental variables (IV) and generalized method-of-moments (GMM) estimators of SAR, SE, and SARAR models; and Lee (2004) who derives the quasi maximum likelihood (QML) estimator of the SAR model. Leading examples focusing on spatial panel data are Elhorst (2003), who introduces the ML estimator of several SAR and SE models with fixed and random effects; Lee and Yu (2010a) who derive the QML estimators of the SARAR model with fixed effects; and Lee and Yu (2012) who derive the QML estimator of a general spatial panel model with fixed and random effects, which nests various spatial panel specifications existing in the literature. In addition to these econometric-theoretical studies, the SAR, SE, and SARAR models are also widely used in all kinds of empirical applications.

A spatial econometric model that received limited attention in both the econometric-theoretical and empirical literature is the SAR model with spatial moving average (MA) errors, hereinafter abbreviated to SARMA. The main reason to draw attention to this model is that in many empirical applications it is not likely that an unobserved shock in the error term of one unit is able to reach every other unit of the study area. Most shocks are local and only have a limited spatial range. Several spatial econometric researchers (Anselin, 2003; Fingleton and LeGallo, 2008; Elhorst, 2010) have pointed out that shocks captured by an AR error process are global since they diffuse among all the units, also when they are not connected according to the specification of the spatial weight matrix. Conversely, shocks captured by an MA error process are local since units can only affect each other when they are connected according to the spatial weight matrix. This finding indicates that the distinction between local and global shocks, and therefore the difference between spatial AR and MA errors and their relationship with the spatial weight matrix describing the spatial arrangement of the units in the sample deserves more

attention. Two important properties of spatial weight matrices in this regard have emerged from the existing literature (Kelejian and Prucha, 1998, 1999; Lee, 2004). The proportion of elements that is zero and whether the row and column sums are upper bounded in absolute value or diverge to infinity though slower than the sample size N. To investigate the relevance of these properties, we consider three different spatial weight matrices in this chapter. A first-order binary contiguity matrix which is sparse and whose row and column sums are uniformly bounded, an inverse distance decay matrix which is dense and whose row and columns sums are unbounded but go to infinity slower than the sample size N, and an exponential distance matrix which is dense but whose row and column sums are again uniformly bounded. A detailed overview of previous studies, which considered spatial MA errors, the difference between local and global shocks, the specification and properties of the three proposed matrices, and the relationship between local or global shocks and the sparsity or density of these matrices is the topic of the next section in this chapter. This detailed overview is also the first contribution of this chapter to the existing literature. Although spatial weight matrices of different spatial lags can be different, they are generally treated to be the same in empirical applications. The downside of this simplification will also be examined in this chapter.

The second contribution of this chapter to the existing literature is the derivation of a QML within estimator of the SARMA model including fixed effects in both the cross-sectional and time domain, and a proof that this estimator is consistent and asymptotically normal for the dominating panel data setting in applied spatial econometric research: N is large and T is finite. We show that time fixed effects, if included, are better treated as regular regressors, since the QML within estimator then also offers the opportunity to consider spatial weight matrices normalized by a scalar, such as the largest eigenvalue, as an alternative to row normalization. This is an advantage because row normalization is not always harmless (see Kelejian and Prucha, 2010), especially when working with inverse or exponential distance decay matrices. Our QML within estimator also explains why the popularity of the SARMA model has lagged behind the SARAR model. It is computationally more demanding and software for spatial MA error models have not or hardly

been made available to practitioners.[1]

This chapter is organized as follows. In Section 3.2 we introduce the SARMA model and its main competitor the SARAR model, give an overview of previous studies that focused on spatial moving average errors, provide a detailed explanation of global versus local interpretations, and explain the link with the two mentioned properties of spatial weight matrices. In Section 3.3 we set out the QML within estimator of the SARMA model. In Section 3.4 we report and discuss the results of a Monte Carlo (MC) simulation experiment evaluating the performance of the proposed estimator. We also show which biases occur when data generated by a SARMA model with different spatial weight matrices for each spatial lag is estimated by the SARAR model, or when one common spatial weight matrix is used for all spatial lags in the SARMA model. Section 3.5 contains an empirical application on military expenditures taken from Yesilyurt and Elhorst (2017) in which the SARMA and SARAR models are compared to each other. Finally, we draw conclusions and suggest topics for further research in Section 3.6.

3.2 Spatial MA and AR error processes

3.2.1 Specifications

The SARMA model for panel data reads as

$$\boldsymbol{y}_t = \rho_0 \boldsymbol{W}\boldsymbol{y}_t + \boldsymbol{X}_t\boldsymbol{\beta}_0 + \boldsymbol{c}_0 + \varpi_{0t}\boldsymbol{\iota}_N + \boldsymbol{u}_t,\ \boldsymbol{u}_t = \boldsymbol{v}_t - \lambda_0 \boldsymbol{M}\boldsymbol{v}_t, \tag{3.1}$$

and the SARAR model as

$$\boldsymbol{y}_t = \rho_0 \boldsymbol{W}\boldsymbol{y}_t + \boldsymbol{X}_t\boldsymbol{\beta}_0 + \boldsymbol{c}_0 + \varpi_{0t}\boldsymbol{\iota}_N + \boldsymbol{u}_t,,\ \boldsymbol{u}_t = \lambda_0 \boldsymbol{M}\boldsymbol{u}_t + \boldsymbol{v}_t, \tag{3.2}$$

where the $N \times 1$ vector $\boldsymbol{y}_t = (y_{1t}, y_{2t}, ..., y_{Nt})'$ reflects the dependent variable of units 1 to N at time t $(t = 1, \ldots, T)$. $\boldsymbol{W}\boldsymbol{y}_t$ represents the spatial lag of $\boldsymbol{y}_t$ and ρ_0 the spatial autoregressive coefficient of this variable. $\boldsymbol{X}_t$ is an $N \times K$ matrix of non-stochastic time-varying regressors and $\boldsymbol{\beta}_0$ the associated $K \times 1$ vector of coefficients. $\boldsymbol{u}_t = (u_{1t}, u_{2t}, ..., u_{Nt})'$ is an $N \times 1$ vector of error terms, which follows a spatial moving average (MA) error process in (3.1) or a spatial autoregressive error process

[1] The routines developed for the purpose of this chapter will be made available upon request.

in (3.2) with parameter λ_0. The elements v_{it} of the disturbance term vector $\boldsymbol{v}_t$ are *i.i.d* across i and t with zero mean, variance σ_0^2, and finite fourth moment. $\boldsymbol{W}$ and $\boldsymbol{M}$ are $N \times N$ spatial weight matrices generating spatial dependence among the cross-sectional units in the dependent variable and the error terms, respectively. Although $\boldsymbol{W}$ and $\boldsymbol{M}$ can be different, they are generally treated to be the same in empirical applications. This simplification will be expressly examined in this chapter. $\boldsymbol{c}_0$ is an $N \times 1$ vector of individual fixed effects, ϕ_{t0} is a time fixed effect, and $\boldsymbol{\iota}_N$ is an $N \times 1$ vector of ones.

3.2.2 Previous studies on spatial MA errors

Previous studies focusing on spatial moving average errors in spatial econometric models providing econometric-theoretical proofs that their estimators are consistent or that implement these type of errors in their empirical application are scarce. Table 3.1 provides an overview. Haining (1978) proposes a spatial MA model and investigates its stationarity, assuming that the spatial weight matrix takes the form of a square lattice. Huang (1984) derives the ML estimator and its asymptotic properties of a SARMA(p,q) model for cross-sectional data. Hepple (2003) sets out the ML estimator of the spatial MA model for cross-sectional data. In addition, she briefly pays attention to the Bayesian estimator of this model and the ML estimator when this model is extended with a spatial lag in the dependent variable, yielding the SARMA model. Fingleton and LeGallo (2007, p.44) modify the feasible generalized spatial two-stage least squares (FGS2SLS) estimator developed by Kelejian and Prucha (1998) for the cross-sectional SARAR model, such that it can be used to estimate a linear regression model not only with exogenous but also with endogenous explanatory variables and a spatial MA error term. In a series of follow-up studies, Fingleton (2008a,2008b) and Fingleton and LeGallo (2008) further explore the finite sample properties of this estimator, extend this model to spatial panels with unobserved unit-specific random effects, and illustrate their models and estimators with the results of some empirical studies. Baltagi and Liu (2011) improve Fingleton's (2008a) FGS2SLS by taking into account that the moment conditions that are based on OLS residuals rather than the true disturbances. Instead of the FGS2SLS estimator developed by Kelejian and Prucha (1998), Doğan and Taspinar (2013) modify the one-step GMM estimators developed by Lee (2007) and Liu et al. (2010) for the

Table 3.1. Studies on spatial moving average errors

Study	Spatial lags	Data	Matrix form	Estimation method
Haining (1978)	$\boldsymbol{Mv}$	CS	Squared lattice	MLE
Huang (1984)	$\boldsymbol{Wy}, \boldsymbol{Mv}$	CS	Not mentioned	MLE
Hepple (2003)	$\boldsymbol{Wy}, \boldsymbol{Mv}$	CS	Not mentioned	MLE, Bayesian
Fingleton and LeGallo (2007)	$\boldsymbol{Mv}$	CS	Contiguity	FGS2SLS
Fingleton (2008a)	$\boldsymbol{Mv}$	CS	Contiguity	FGS2SLS
Fingleton (2008b)	$\boldsymbol{Wy}_t, \boldsymbol{Mv}_t$	PD	Contiguity	FGS2SLS
Fingleton and LeGallo (2008)	$\boldsymbol{Wy}, \boldsymbol{Mv}$	CS	Negative exponential	FGS2SLS
Baltagi and Liu (2011)	$\boldsymbol{Mv}$	CS	Contiguity	GMM (OLS residuals)
Lee and Yu (2012)	$\boldsymbol{Wy}_t, \boldsymbol{Mv}_t$	PD	Not mentioned	QMLE-FD
Doğan and Taşpınar (2013)	$\boldsymbol{Wy}, \boldsymbol{Mv}$	CS	Group interaction	GMM
Fingleton et al. (2018)	$\boldsymbol{Wy}_t, \boldsymbol{Mv}_t$	PD	Group interaction	FGS3SLS
Baltagi et al. (2019)	$\boldsymbol{Wy}_t, \boldsymbol{Mv}_t$	PD	Contiguity	FGS4SLS
This study	$\boldsymbol{Wy}_t, \boldsymbol{Mv}_t$	PD	Contiguity, negative exponential, inverse distance, row and scalar normalized	QMLE-within

CS: cross-sectional data; PD: panel data. (Q)MLE=(Quasi) maximum likelihood estimator, FGS2SLS=Feasible generalized spatial two-stage least-squares, FGS3SLS=Extension of FGS2SLS, FGS4SLS=Extension of FGS3SLS.

cross-sectional SARAR model, such that it can be used to estimate the spatial MA and the SARMA models. The FGS2SLS estimator developed by Fingleton and others is also used in later work of Fingleton et al. (2018) to estimate a spatial panel SARMA model with hierarchical random effects in the spatial domain, i.e., for individual units and individual units within groups. They also rename their estimator FGS3SLS because the parameters are determined in three stages. In their empirical application, they specifically focus on regions across EU member countries and a spatial weight matrix which is specified as a group interaction matrix. Baltagi et al. (2019) further extends this model with dynamic effects in space and time, i.e., capturing a spatial, a time, and a space-time lag in the dependent variable, as well as a spatial lag, a time lag, and a spatial MA process in the error term, and to a FGS4SLS estimator to account for these additional dynamic effects. Lee and Yu (2012) derive the QML estimator of a general spatial panel model with fixed and random effects, which nests various spatial panel specifications, among which the SARMA model in Equation (3.1) of this chapter. However, due to its general character and the focus on the properties of their derived estimators, this model specification and its properties relevant for empirical studies, the subject of this study, do not get the attention they deserve. Another issue is that they take first-differences to concentrate out the individual effects rather than the commonly used within transformation. In addition to these journal papers, the spatial MA error specification is briefly discussed in several overview chapters (Anselin and Bera, 1998) or spatial econometric textbooks (Haining, 1990; LeSage and Pace, 2009).

The conclusion from this overview of previous studies is that estimation methods, which are first developed to estimate models with spatial autoregressive errors, are subsequently also developed for models with spatial moving average errors. The focus has been on ML, IV/GMM, and Bayesian estimators of the cross-sectional SARMA model, and on the feasible generalized spatial two-stage least squares (FGS2SLS) estimator of its panel data counterpart with at most random effects for the cross-sectional units. Fixed effects in space and especially time have not been considered in these studies, while the degree of local spatial dependence in the dependent variable and/or the error term when time fixed effects are left aside tends to be overestimated, an issue that is overlooked in all studies reported in Table 3.1. One exception is Lee and Yu (2012), but they do use first-differences

rather than the standard within approach to concentrate out the individual effects. For these reasons, we provide the quasi ML (QML) within estimator of the panel data SARMA model with individual and time fixed effects. The term quasi is used since this estimator does not require any specification of the distribution function of the error terms, except that they have zero mean, variance σ^2, and finite fourth moment. We focus on the large N and finite T panel data setting since it is representative for most empirical studies in spatial econometrics. The problem of N being large is the incidental parameter problem; if N is large so will be the number of individual fixed effects, which furthermore cannot be estimated consistently if T is finite. Generally, they are concentrated out by the demeaning procedure $\boldsymbol{I}_T - \frac{1}{T}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'$, which is the standard within transformation in the panel data literature (Arellano, 2003, Section 2.2; Baltagi, 2005).

3.2.3 Local versus global shocks

The spatial MA error of the SARMA model in (3.1) can be rewritten as

$$\boldsymbol{u}_t = (\boldsymbol{I}_N - \lambda_0\boldsymbol{M})\boldsymbol{v}_t = \boldsymbol{v}_t - \lambda_0\boldsymbol{M}\boldsymbol{v}_t, \tag{3.3}$$

and the spatial AR error of the SARAR model in (3.2) as

$$\begin{aligned}\boldsymbol{u}_t &= (\boldsymbol{I}_N - \lambda_0\boldsymbol{M})^{-1}\boldsymbol{v}_t = \left(\sum_{i=1}^{\infty}\lambda_0^i\boldsymbol{M}^i\right)\boldsymbol{v}_t \\ &= \boldsymbol{v}_t + \lambda_0\boldsymbol{M}\boldsymbol{v}_t + \lambda_0^2\boldsymbol{M}^2\boldsymbol{v}_t + \lambda_0^3\boldsymbol{M}^3\boldsymbol{v}_t + ...,\end{aligned} \tag{3.4}$$

where we used the property that $\boldsymbol{M}^0 = \boldsymbol{I}_N$. These two spatial error specifications have different economic interpretations. Spatial interaction effects among the error terms may occur if unobserved shocks follow a spatial pattern other than those covered by time fixed effects, which may cause the parameters of the model to be inefficient. This inefficiency may be avoided or reduced by accounting for these spatial interaction effects through estimating λ_0.

Several studies paid attention to the economic interpretation of spatial AR versus MA errors (Anselin, 2003; Fingleton[2] and LeGallo, 2008; Elhorst, 2010). If a shock occurs at a particular unit, a spatial MA error term will have the effect that this

[2] See also several follow-up studies of this author.

shock is transmitted to those units which belong to its neighbourhood set according to the specification of $\boldsymbol{M}$. If this matrix is sparse, it will only be transmitted to a limited number of units, mainly nearby units. If this matrix is dense, it will be transmitted to more units, whereby units located nearby, in accordance with Tobler's first law, will be more strongly influenced than units farther away. In other words, the sparser (denser) the spatial weight matrix, the smaller (greater) will be the spatial range of the area that is affected by this shock. Since the right-hand side of (3.3) only contains the first-order effect $\boldsymbol{M}\boldsymbol{v}_t$, the impact of the shock may be termed a local effect.

By contrast, if a shock occurs at a particular unit and the error term follows a spatial AR process, then the right-hand side of (3.4) not only contains the first-order effect $\boldsymbol{M}\boldsymbol{v}_t$ but also higher-order effects. Just as the spatial MA error term, the first-order effect is limited to those units of which the elements of $\boldsymbol{M}$ are non-zero. However, due to the second and higher-order effects, units that do not belong to the neighbourhood set of a unit that instigates the shock will also be affected. Such higher-order effects may also arise as a result of feedback effects, i.e. impacts passing through neighbouring units and back to the unit itself (e.g. $1 \rightarrow 2 \rightarrow 1$ and $1 \rightarrow 2 \rightarrow 3 \rightarrow 2 \rightarrow 1$). Since this multi-round diffusion effect ultimately affects all units by the shock that occurred at one location, such a shock may be termed a global effect. To explain this principle, all three cited studies (Anselin, 2003; Fingleton and LeGallo, 2008; Elhorst, 2010) adopt a sparse matrix, i.e., the first-order binary contiguity (BC) matrix whose elements are 1 if two units share a common border, and 0 otherwise. This is understandable because it is not plausible that the spatial matrix is dense when adopting a spatial AR process. If this matrix would be dense, it is not necessary to consider higher-order terms. Then it would be sufficient and more efficient to specify a spatial MA error term in combination with a spatial matrix that is dense. In many empirical applications it is also not likely to assume that an unobserved shock in one unit is able to reach any corner of the study area. Most shocks are local and only have a limited spatial range. Furthermore, external shocks that are shared by all units in the sample, such as business cycle effects, can be modelled by time fixed effects. Besides, the SARMA model also contains a spatial lag in the dependent variable ($\boldsymbol{W}\boldsymbol{y}_t$). Consequently, the impact of a local shock on the dependent variable of a particular unit may also spread further

due to endogenous interaction effects among the dependent variables of the units across the study area.

Finally, it should be emphasized that AR and MA error processes are not nested. Equation (3.4) seems to simplify to Equation (3.3) when the second and higher-order terms in (3.4) are zero, i.e., when $\lambda_0 = 0$. However, when this parameter is zero, the first-order term $\lambda_0 \boldsymbol{M}\boldsymbol{v}_t$ is also dropped, resulting in the SAR instead of the SARMA model.

3.2.4 Spatial weight matrices

The previous section has made clear that the choice between a spatial AR or MA process is also related to the spatial range of the spatial weight matrix. The spatial econometric literature has pointed out that this range depends on the properties which percentage of the elements is zero and whether or not their row and column sums are upper bounded in absolute value or diverge to infinity though slower than the sample size N. Assuming that $\boldsymbol{M}$ (note: the same holds for $\boldsymbol{W}$) is non-negative and in raw form a non-normalized matrix of known constants with zero diagonal elements (to prevent units from explaining themselves) either of the two additional assumptions is needed for the consistency of the QML/IV/GMM estimators set out in Kelejian and Prucha (1998, 1999), Lee (2004), and Lee and Yu (2010a):

$$I : 0 < \lim_{N \to +\infty} \sum_{j=1}^{N} |m_{ij}| \leq c, \tag{3.5}$$

where m_{ij} denotes an element of $\boldsymbol{M}$ and c is a constant. This is known as the boundedness condition and will be satisfied when the spatial weight matrix takes the form of a BC matrix, since the number of neighbours of each unit is less than the unit with the largest number of neighbours. This condition is not satisfied when the spatial weight matrix takes the form of an inverse distance (ID) matrix: $m_{ij} = 1/d_{ij}$, where d_{ij} denotes the distance between units i and j. Not only are all off-diagonal elements of this matrix non-zero, the row and column sums of this matrix also diverge to infinity. For example, if the distance to the first neighbour is 1, to the second is 2, to the third is 3, and so on, the corresponding row sum of this unit amounts to $1/1 + 1/2 + 1/3 + \ldots$, representing a series which is not upper

bounded. By contrast, the second assumption

$$II : \lim_{N\to+\infty} \frac{\sum_{j=1}^{N} m_{ij}}{N} \to 0, \tag{3.6}$$

is satisfied since $\lim_{N\to+\infty}(1/1 + 1/2 + 1/3 + \ldots)/N \to 0$. This is known as the weak convergence condition. A more detailed mathematical explanation is available in Elhorst et al. (2021). The exponential distance (ED) matrix, whose elements are defined as $m_{ij} = exp(-d_{ij})$, takes a position in between. Just as the ID matrix, all off-diagonal elements of this matrix are non-zero, but in contrast to the ID matrix it does satisfy condition I that its row and column sums are upper bounded. Using the same example as for the ID matrix, the row sum amounts $1/e + 1/e^2 + 1/e^3 + \ldots$, which is upper bounded by $1/(e-1)$.

In sum, while the off-diagonal elements of both the ED and the ID matrix are non-zero, the correlation between any two units of the ED matrix converges so much faster to 0 with distance in contrast to that of the ID matrix that its row and column sums are upper bounded. We will utilize these properties when testing the SARMA and SARAR models against each other.

3.3 Estimation of the SARMA panel data model

In this section, we derive the ML and QML within estimator of the SARMA model when using the within transformation to concentrate out the individual fixed effects, and we investigate the asymptotic properties of the estimator when N is large and T is finite.

The SARMA model in (3.1) contains both individual and time fixed effects. Importantly, when T is finite, these time fixed effects do not have to be concentrated out; they can also be taken up as part of the regressors $\boldsymbol{X}_t$. For this reason, we depart from the model

$$\boldsymbol{y}_t = \rho_0 \boldsymbol{W}\boldsymbol{y}_t + \boldsymbol{X}_t\boldsymbol{\beta}_0 + \boldsymbol{c}_0 + \boldsymbol{u}_t, \;\; \boldsymbol{u}_t = \boldsymbol{v}_t - \lambda_0 \boldsymbol{M}\boldsymbol{v}_t, \tag{3.7}$$

in this subsection and do no specify the time fixed effects separately. This model therefore covers both the possibility that a researcher includes time fixed effects or does not want to include time fixed effects, for example, because some regressors

do not vary over time.

In general, elements of the spatial weight matrix are row-normalized because it facilitates the interpretation of operations with the weight matrix as an averaging of neighbouring values (Anselin and Bera, 1998). However, row-normalization has also been criticized. Kelejian and Prucha (2010) demonstrate that normalization of the elements of the spatial weight matrix by a different factor for each row as opposed to a single factor is likely to lead to misspecification problems. To avoid such misspecification problems, they propose an alternative normalization procedure where each element of the spatial weight matrix in raw form is divided by its largest eigenvalue. Important, the QML within estimator set out in this section also covers the possibility to consider this kind of normalized matrices, which is useful especially when working with inverse or exponential distance decay matrices. In sum, since the majority of empirical studies is characterized by this panel data setting and the spatial weight matrix does not necessarily have to be row-normalized, the findings in this section are relevant for a broad range of empirical studies.

Let $\boldsymbol{S}(\rho) = \boldsymbol{I}_N - \rho\boldsymbol{W}$, $\boldsymbol{R}(\lambda) = \boldsymbol{I}_N - \lambda\boldsymbol{M}$, $\boldsymbol{\beta} = (\boldsymbol{\beta}^{'}, \rho, \lambda, \sigma^2)^{'}$, and $\boldsymbol{\zeta} = (\boldsymbol{\beta}^{'}, \rho, \lambda)^{'}$. When the subscript 0 is added to a parameter or a vector of parameters below, we denote the true value of this parameter or vector. $\boldsymbol{J}_T = \boldsymbol{I}_T - \frac{1}{T}\boldsymbol{\iota}_T\boldsymbol{\iota}_T^{'}$ is used to concentrate out the individual fixed effects, which is known as the within transformation in the panel data literature (see Arellano, 2003; Baltagi, 2005). This technique reads as $\widetilde{\boldsymbol{y}}_t = \boldsymbol{y}_t - \overline{\boldsymbol{y}}_t$ for $t = 1, 2, \ldots, T$, where the elements of $\overline{\boldsymbol{y}}_t$ are defined by $\overline{y}_{it} = \frac{1}{T}\sum_{t=1}^{T} y_{it}$. Similar transformations apply to the explanatory variables and the error term of the model, yielding $\widetilde{\boldsymbol{X}}_t$, $\widetilde{\boldsymbol{u}}_t$, and $\widetilde{\boldsymbol{v}}_t$.

Assuming as if the error terms are normally distributed, we will change this later when investigating the asymptotic properties based on the variance-covariance matrix of the proposed estimator, the log likelihood after $\boldsymbol{c}_0$ is concentrated out using the within transformation (indicated by the superscript w) is given by

$$\ln L^{w}(\boldsymbol{\theta}) = -\frac{NT}{2}\ln(2\pi\sigma^2) + T[\ln|\boldsymbol{S}(\rho)| - \ln|\boldsymbol{R}(\lambda)|] - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t^{'}(\boldsymbol{\zeta})\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}), \tag{3.8}$$

where $\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}) = \boldsymbol{R}^{-1}(\lambda)[\boldsymbol{S}(\rho)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}]$. Two important differences between the SARMA and SARAR model are that the matrix $\boldsymbol{R}(\lambda)$ needs to be inverted when

calculating $\widetilde{v}_t(\zeta)$ and that the Jacobian term $\ln|\boldsymbol{R}(\lambda)|$ does have a negative rather than a positive sign. The first and second order derivatives of (3.8) are given in (3.A.1) and (3.A.3) of the Appendix 3.A.1, respectively.

Next, we also concentrate out $\boldsymbol{\beta}$ and σ^2 to focus on (ρ, λ), whose QML within estimators, given λ and ρ, are

$$\widehat{\boldsymbol{\beta}}^{w}(\rho,\lambda) = [\sum_{t=1}^{T} \widetilde{\boldsymbol{X}}_t^{'} \boldsymbol{R}^{-1'}(\lambda)\boldsymbol{R}^{-1}(\lambda)\widetilde{\boldsymbol{X}}_t]^{-1}[\sum_{t=1}^{T} \widetilde{\boldsymbol{X}}_t^{'} \boldsymbol{R}^{-1'}(\lambda)\boldsymbol{R}^{-1}(\lambda)\boldsymbol{S}(\rho)\widetilde{\boldsymbol{y}}_t], \tag{3.9}$$

$$\widehat{\sigma}^{2w}(\rho,\lambda) = \frac{1}{NT}\sum_{t=1}^{T}[\boldsymbol{S}(\rho)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t\widehat{\boldsymbol{\beta}}^{w}(\rho,\lambda)]^{'}\boldsymbol{R}^{-1'}(\lambda)\boldsymbol{R}^{-1}(\lambda)[\boldsymbol{S}(\rho)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t\widehat{\boldsymbol{\beta}}^{w}(\rho,\lambda)], \tag{3.10}$$

as a result of which the corresponding concentrated log likelihood function of (ρ, λ) is

$$\ln L^{w}(\rho,\lambda) = -\frac{NT}{2}(\ln(2\pi)+1) - \frac{NT}{2}\ln\widehat{\sigma}^{2w}(\rho,\lambda) + T[\ln|\boldsymbol{S}(\rho)| - \ln|\boldsymbol{R}(\lambda)|]. \tag{3.11}$$

One problem of the within transformation is that the resulting error terms are linearly dependent over the time dimension. Consequently, the estimate $\widehat{\sigma}^{2w}$ of σ_0^2 using the within transformation will be consistent only when T is large. When T is finite, $\widehat{\sigma}^{2w}$ should be bias-corrected by $\frac{T}{T-1}\widehat{\sigma}^{2w}$ to obtain a consistent estimate of σ_0^2. This is demonstrated by Arellano (2003, p.24) for non-spatial panel models and by Lee and Yu (2010a) for a SARAR panel data model. This can be effected by replacing T in (3.11) with $T-1$.

To investigate the asymptotic properties of the proposed QML within estimator of the SARMA model, we make eight assumptions and prove two theorems. Assumptions similar to Lee and Yu (2010a) are labelled by **(LY)** and those that are different due to considering a spatial MA error process are labelled by **(MA)**. The assumptions are

Assumption 3.1 (LY). *The spatial weight matrices* $\boldsymbol{W}$ *and* $\boldsymbol{M}$ *are non-stochastic and have zero diagonal values.*

Assumption 3.2 (LY). *The disturbances* v_{it} *are i.i.d for* $i = 1, 2, ..., N$ *and* $t = 1, 2, ..., T$*, and have zero mean, variance* σ_0^2*, and* $E|v_{it}|^{4+\eta} < \infty$ *for some* η*.*

Assumption 3.3 (LY). $\boldsymbol{S}(\rho)$ *and* $\boldsymbol{R}(\lambda)$ *are invertible for all* ρ *and* λ *in compact intervals* $\mathbb{P}$ *and* Λ*, respectively.* ρ_0 *is the interior of* $\mathbb{P}$ *and* λ_0 *is the interior of* Λ*.*

Assumption 3.4 (MA). *The elements of $\boldsymbol{X}_t$ are non-stochastic and bounded uniformly in N and T in absolute value and the limit of $\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{'}\boldsymbol{R}^{-1'}\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t$ exists and is non-singular.*

Assumption 3.5 (LY). *$\boldsymbol{W}$ and $\boldsymbol{M}$ are uniformly bounded in both row and column sums in absolute value, and so are $\boldsymbol{S}^{-1}(\rho)$ and $\boldsymbol{R}^{-1}(\lambda)$ for $\rho \in \mathbb{P}$ and $\lambda \in \Lambda$.*

Assumption 3.6 (MA). *N is large and T is finite.*

Assumption 3.7 (MA). *Either (a) the limit of*

$$(\frac{1}{N}|\sigma_0^2\boldsymbol{R}^{'}\boldsymbol{S}^{-1'}\boldsymbol{S}^{-1}\boldsymbol{R}| - \frac{1}{N}|\sigma^{2\dagger}(\rho,\lambda)\boldsymbol{R}^{'}(\lambda)\boldsymbol{S}^{-1'}(\rho)\boldsymbol{S}^{-1}(\rho)\boldsymbol{R}(\lambda)|)$$

is not zero for $(\rho,\lambda) \neq (\rho_0,\lambda_0)$, as N tends to infinity; or (b) the limit of $\mathcal{H}(\lambda)$ is non-singular for each possible λ in Λ, and the limit of $\frac{1}{N}|\sigma_0^2\boldsymbol{R}^{'}\boldsymbol{R}| - \frac{1}{N}|\sigma^{2\dagger}(\lambda)\boldsymbol{R}^{'}(\lambda)\boldsymbol{R}(\lambda)|$ is not zero for $\lambda \neq \lambda_0$.

Assumption 3.8 (MA). *The limit of $\frac{1}{N^2}[tr(\boldsymbol{C}^s\boldsymbol{C}^s)tr(\boldsymbol{D}^s\boldsymbol{D}^s) - tr^2(\boldsymbol{C}^s\boldsymbol{D}^s)]$ is strictly positive as N tends to go infinity.*

We denote $\boldsymbol{G} = \boldsymbol{W}\boldsymbol{S}^{-1}$, $\ddot{\boldsymbol{G}} = \boldsymbol{R}^{-1}\boldsymbol{G}\boldsymbol{R}$, $\boldsymbol{K} = \boldsymbol{R}^{-1}\boldsymbol{M}$, $\boldsymbol{C} = \ddot{\boldsymbol{G}} - \frac{\mathrm{tr}\ddot{\boldsymbol{G}}}{N}\boldsymbol{I}_N$, and $\boldsymbol{D} = \boldsymbol{K} - \frac{\mathrm{tr}\boldsymbol{K}}{N}\boldsymbol{I}_N$. We further have

$$\mathcal{H}(\lambda) = \frac{1}{N(T-1)}\sum_{t=1}^{T}(\widetilde{\boldsymbol{X}}_t, \boldsymbol{G}\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0)^{'}\boldsymbol{R}^{-1'}(\lambda)\boldsymbol{R}^{-1}(\lambda)(\widetilde{\boldsymbol{X}}_t, \boldsymbol{G}\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0), \tag{3.12}$$

$$\sigma^{2\dagger}(\lambda) = \frac{\sigma_0^2}{N}\mathrm{tr}[(\boldsymbol{R}^{-1}(\lambda)\boldsymbol{R})^{'}(\boldsymbol{R}^{-1}(\lambda)\boldsymbol{R})],$$

$$\sigma^{2\dagger}(\rho,\lambda) = \frac{\sigma_0^2}{N}\mathrm{tr}[(\boldsymbol{R}^{-1}(\rho)\boldsymbol{S}(\rho)\boldsymbol{S}^{-1}\boldsymbol{R})^{'}(\boldsymbol{R}^{-1}(\lambda)\boldsymbol{S}(\rho)\boldsymbol{S}^{-1}\boldsymbol{R})].$$

Assumptions 3.1, 3.2, 3.3, and 3.5 have been discussed in Lee and Yu (2010a), although a few comments from a practitioner's point of view are in order. Assumption 3.1 can be relaxed by allowing the spatial weight matrices to contain non-zero diagonal elements. Generally, their diagonal elements are set to zero to avoid that units can affect themselves. However, there are also applications in which this assumption is not made, for example, when explaining self-flow data in trade flow models (Balaszi et al., 2018). Even though the elements of spatial weight matrices are generally assumed to be non-negative from an economic viewpoint, negative elements do not have to be excluded either (Bailey et al., 2016). This explains the

addition in absolute value in Assumption 3.5. Furthermore, even if all elements are non-negative, this does not guarantee that all elements of $\boldsymbol{S}^{-1}(\rho)$ and $\boldsymbol{R}^{-1}(\lambda)$ are also non-negative, which explains the terminology uniformly bounded in absolute value in general. Assumption 3.5 might also be relaxed; the row and columns sum may also diverge to infinity, but only to a lesser extent than the sample size N (see Lee, 2004; Lee and Yu, 2010a, footnote 14). Commonly it is assumed that ρ_0 and λ_0 take values in the interval $(-1, +1)$ to fulfill Assumption 3.3. Theoretically it is possible that ρ_0 is smaller than -1, since the lower bound is defined by 1 divided by the smallest eigenvalue of $\boldsymbol{W}$. The result will be smaller than -1 since this smallest eigenvalue will be in the interval (-1,0) when $\boldsymbol{W}$ is non-negative and normalized (Ord, 1978). However, negative values for ρ_0 in models with a spatial lag in the dependent variable are considerably less common than positive ones, let alone negative values smaller than -1 (Elhorst, 2014, Section 2.5). Therefore, imposing this lower bound is hardly restrictive. A similar argument applies to the coefficient of the MA error term.

Assumptions 3.4, 3.6, 3.7, and 3.8 are modified for the model with MA errors; if the inverse of the matrix $\boldsymbol{R}(\lambda)$ is taken when considering a spatial AR error process, it is not when considering a spatial MA process, or vice versa. Assumption 3.6 focuses on the large N and finite T panel data setting, representative for the majority of spatial econometric studies. This assumption allows researchers the opportunity to include time fixed effects as regular regressors without facing an incidental parameter problem due to T being large, and the opportunity to consider spatial weight matrices that are not row normalized but normalized by a scalar. Assumption 3.7 provides the identification conditions of the model with spatial MA errors. Assumption 3.7(a) globally identifies (ρ_0, λ_0), provided that $\boldsymbol{W} \neq \boldsymbol{M}$. In case $\boldsymbol{W} = \boldsymbol{M}$, (ρ_0, λ_0) and (λ_0, ρ_0) cannot be distinguished from each other based on 3.7(a). Instead, identification relies on Assumption 3.7(b). The non-singularity of $\mathcal{H}(\lambda)$ for each λ in the first part of Assumption 3.7(b) guarantees that ρ_0 and $\boldsymbol{\beta}_0$ are identified. The second part of Assumption 3.7(b) provides the identification of λ_0 for $\lambda \neq \lambda_0$.[3] These identification assumptions have been discussed in Lee and Yu (2010a) for the SARAR model and have been transformed here for the SARMA

[3] In Assumption 3.7, the equations in (3.12), and the proofs of Theorems 1 and 2 in the appendix, the notations $\boldsymbol{R}(\lambda)$ and $\boldsymbol{R}$ are used to denote this matrix evaluated at respectively the estimate of λ and the true value of $\lambda = \lambda_0$.

model.

Theorem 1. *Under Assumptions 3.1-3.7, $\boldsymbol{\beta}_0$ is identified and for $\widehat{\boldsymbol{\theta}}^{w}$ based on (3.9)-(3.11) of the within transformation, $\widehat{\boldsymbol{\theta}}^{w} - \boldsymbol{\theta}_T \xrightarrow{p} \mathbf{0}$, where $\boldsymbol{\theta}_T = \boldsymbol{\theta}_0 - (\mathbf{0}_{1\times(K+2)}, \frac{1}{T}\sigma_0^2)'$.*

Proof. See Appendix 3.A.2.

From this theorem it follows that $\sigma_T^2 = \sigma_0^2 - \frac{1}{T}\sigma_0^2$, which implies that the true value of $\sigma_0^2 = \frac{T}{T-1}\sigma_T^2$.

Theorem 2. *Under Assumptions 1-6 and 7(a) ($\boldsymbol{W} \neq \boldsymbol{M}$); or Assumption 1-6, 7(b), and 8 ($\boldsymbol{W} = \boldsymbol{M}$),*

$$\sqrt{NT}(\widehat{\boldsymbol{\theta}}^{w} - \boldsymbol{\theta}_T) \xrightarrow{w} \mathcal{N}\left(\mathbf{0}, lim\frac{T}{T-1}(\boldsymbol{\Sigma}^{w}_{\boldsymbol{\theta}_T})^{-1}(\boldsymbol{\Sigma}^{w}_{\boldsymbol{\theta}_T} + \boldsymbol{\Omega}^{w}_{\boldsymbol{\theta}_T})(\boldsymbol{\Sigma}^{w}_{\boldsymbol{\theta}_T})^{-1}\right), \tag{3.13}$$

where $\boldsymbol{\Sigma}^{w}_{\boldsymbol{\theta}_T}$ and $\boldsymbol{\Omega}^{w}_{\boldsymbol{\theta}_T}$ are given in (3.A.4) and (3.A.5) of Appendix 3.A.1;

Proof. See Appendix 3.A.3.

Importantly, the addition of the matrix $\boldsymbol{\Omega}^{w}_{\boldsymbol{\theta}_T}$ in (3.13) of this theorem, see Equation (3.A.5) for its detailed specification, covers the correction that is needed when the error terms are not assumed to be normally distributed. Only if they are, this matrix reduces to the null matrix and drops out. This correction affects the efficiency of the parameter estimates but not their consistency.

To determine the parameters of the SARMA model numerically based on the proposed QML within estimator, a routine has been developed where the researcher needs to specify the $N \times 1$ vector $\boldsymbol{y}_t$, the $N \times K$ matrix $\boldsymbol{X}_t$, and the two $N \times N$ spatial weight matrices $\boldsymbol{W}$ and $\boldsymbol{M}$. The researcher also has the opportunity to specify whether or not to include time fixed effects and whether or not to depart from normally distributed error terms. The routine first concentrates out the individual fixed effects and then maximizes the concentrated log-likelihood function in Equation (3.11), such that if the values of ρ and/or λ change, the estimates for $\widehat{\boldsymbol{\beta}}^{w}(\rho, \lambda)$ and $\widehat{\sigma}^{2w}(\rho, \lambda)$ in Equations (3.9) and (3.10) change accordingly. When this maximum is found, the estimate for $\widehat{\sigma}^{2w}$ is bias-corrected. A similar routine has been developed for the SARAR model. By comparing these two routines, it appeared that

the SARMA model is computationally more demanding than the SARAR model and thus requires more computation time. This is because the matrix $\boldsymbol{R}$ when determining $\widehat{\boldsymbol{\beta}}^{w}(\rho,\lambda)$ and $\widehat{\sigma}^{2w}(\rho,\lambda)$ needs to be inverted every time the concentrated log-likelihood function is determined for a particular pair of values (ρ,λ) in the maximization process, which is not needed when estimating the parameters of the SARAR model. This may have been an obstacle in the past when computing facilities were relatively limited and might explain why the SARAR model received more attention than the SARMA model. Another explanation might be that researchers may find it more difficult to work with different structures for the spatial lag in the dependent variable and in the error term. For the same reason and due to lack of economic-theoretical knowledge how to specify the spatial weight matrix, they also often consider one common spatial weight matrix for all spatial lags in the model. This chapter attempts to break this practice by considering both the SARMA and the SARAR model, by considering the opportunity to consider different spatial weight matrices for each spatial lag, and by making both routines available.

The routines are based on Assumption 3.6 that N is large and T is finite. Given that T is finite in most empirical spatial econometric studies, there is no obstacle to treat time fixed effects, if included, as regular regressors. When T grows large, it is nonetheless tempting not only to concentrate out the individual effects but also the time fixed effects. However, this requires more complicated transformations than the within approach and spatial weight matrices that are row normalized, or alternatively a more complicated bias correction procedure of the within estimator. Details are available in Lee and Yu (2010a, Section 3) and Elhorst (2014, Section 3.3.3), though only for the SARAR model. To illustrate the minor significance of this extension, we refer to the xsmle command in Stata. This routine offers the opportunity to estimate the SARAR model with individual and/or time fixed effects (ind, time or both), as well as to activate the "leeyu" suboption to concentrate out the fixed effects using the transformations developed by Lee and Yu (2010a). However, when both individual and time fixed effects are included and the "leeyu" suboption is activated, the following warning is given "Suboption -both- will be replaced with -ind-".

3.4 Monte Carlo simulations

We conduct a Monte Carlo experiment to examine the finite sample properties of the proposed QML within estimator of the SARMA model under different N and T settings. We consider $N = (200, 800)$ and $T = (5, 10)$ because these panel data settings are representative for applied studies in spatial econometrics. We generate the data using the SARMA model with a binary contiguity (BC) matrix for the dependent variable and a negative exponential (ED) distance decay matrix for the error term. To construct these matrices, we use the coordinates of the N data points evenly set in a rectangle of 10×20 for $N = 200$ and 20×40 for $N = 800$. The BC matrix is based on the rook principle. Row-normalization is performed for both matrices. The data generating process further contains one explanatory variable $\boldsymbol{x}_t$ with coefficient $\beta_0 = 1$, while $\boldsymbol{c}_0$, $\boldsymbol{\omega}_{0t} = (\omega_{0,1}, \omega_{0,2}, ..., \omega_{0T})$, and $\boldsymbol{v}_t$ are all generated from independent standard normal distributions, where $\sigma_0^2 = 1$. Two settings of spatial dependence in the dependent variable and spatial dependence in the error are considered, which are $(\rho_0, \lambda_0) = (0.2, 0.5)$ and $(\rho_0, \lambda_0) = (0.5, 0.2)$. The degree of spatial dependence in the dependent variable in the first setting is relatively mild (0.2) and the spatial dependence in the error term relatively large (0.5), while in the second setting this is turned around. The number of iterations is set to 1000. The R^2 of the DGP ranges from 30% to 50%.

We first estimate the model "correctly"; the SARMA model with BC matrix for the dependent variable and ED matrix for the errors are used both in the DGP and estimation of this model. We consider the QML within estimator before and after bias correction (bc), respectively. When bias correction is performed, σ^2 is labelled as σ_{bc}^2. Next, we investigate respectively what happens if a practitioner would estimate the SARMA model by one common spatial weight matrix, the BC matrix, for the two spatial lags ("wrong" matrix), would estimate the SARAR model with the BC matrix for the dependent variable and the ED matrix for the error terms ("wrong" model), or would estimate the SARAR model by one common spatial weight matrix, the BC matrix, for the two spatial lags ("wrong" matrix and "wrong" model). In each of these additional cases, we only consider the bias corrected σ_{bc}^2. Time fixed effects are included and treated as regular explanatory variable in all cases, but their coefficient estimates are not reported for reasons of space

and simplicity. For each panel data setting and estimated model, the empirical bias (Bias), the empirical standard deviation (E-SD), the empirical root mean square error (RMSE), and the empirical median bias (Mbias) are reported. In addition, we report the direct and indirect spillover effect of the explanatory variable in the model (denoted by DE_x and IE_x). The direct effect measures the impact of changing the explanatory variable of one unit on the dependent variable of that unit itself, and the indirect spillover effect the cumulative effect of changing the explanatory variable of one unit on the dependent variable of other units (LeSage and Pace, 2009). LeSage and Pace (2018) demonstrate that past studies' focus exclusively on point estimates may not provide useful information regarding the statistical properties of the marginal effects, i.e., the direct and indirect spillover effect of each variable obtained from these point estimates.

Tables 3.2 and 3.3 report the simulation results for $(\rho_0, \lambda_0) = (0.2, 0.5)$ and $(\rho_0, \lambda_0) = (0.5, 0.2)$, respectively. Both cases show similar patterns. The simulation results in the first panel "SARMA BC_ED" of these tables unequivocally show that the within approach without bias correction should be rejected. As expected, the estimate for σ^2 is severely biased in small samples, even though it halves when T increases from 5 to 10. The biases in β, ρ, λ, σ^2_{bc}, and the direct and indirect spillover effects are all small and acceptable and tend to fall as T increases but above all when N increases. Largely similar results are found for the E-SD, RMSE, and Mbias. Overall, the simulation results appear to be consistent with the theoretical predictions.

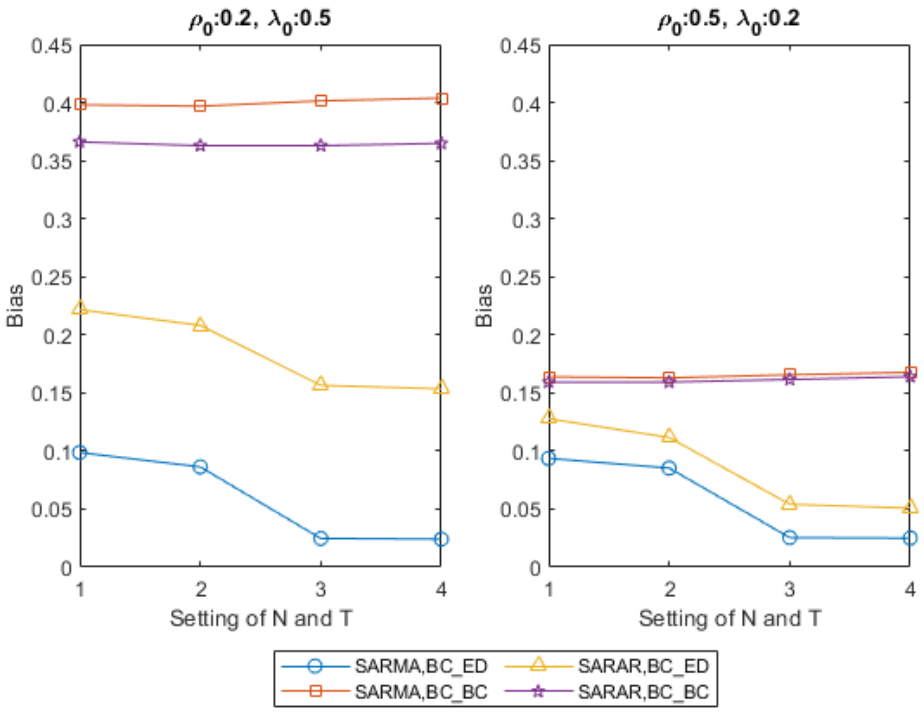
Table 3.2. Simulation results for the SARMA: ρ_0:0.2, λ_0:0.5

Setting				SARMA BC_ED							SARMA BC_BC						SARAR BC_ED						SARAR BC_BC					
	N	T		β	ρ	λ	σ^2	σ^2_{bc}	DE_x	IE_x	β	ρ	λ	σ^2_{bc}	DE_x	IE_x	β	ρ	λ	σ^2_{bc}	DE_x	IE_x	β	ρ	λ	σ^2_{bc}	DE_x	IE_x
(1)	200	5	Bias	-0.001	0.009	0.099	-0.214	-0.017	0.001	0.018	-0.002	0.027	0.398	-0.019	0.003	0.048	-0.001	0.025	-0.222	-0.031	0.002	0.041	-0.003	0.004	-0.366	-0.030	0.000	0.016
			E-SD	0.034	0.056	0.212	0.043	0.054	0.035	0.085	0.035	0.067	0.078	0.053	0.036	0.105	0.034	0.049	0.128	0.051	0.035	0.077	0.035	0.081	0.101	0.052	0.036	0.121
			RMSE	0.034	0.056	0.234	0.218	0.057	0.035	0.087	0.035	0.072	0.405	0.057	0.036	0.115	0.034	0.056	0.234	0.057	0.035	0.087	0.035	0.072	0.405	0.057	0.036	0.115
			Mbias	-0.001	0.006	0.103	-0.216	-0.020	0.002	0.010	-0.001	0.027	0.398	-0.022	0.003	0.040	-0.001	0.022	-0.208	-0.033	0.003	0.034	-0.001	0.003	-0.364	-0.031	0.001	0.005
(2)	200	10	Bias	-0.001	0.008	0.086	-0.110	-0.011	0.001	0.014	-0.001	0.029	0.397	-0.013	0.003	0.047	-0.001	0.024	-0.208	-0.025	0.002	0.038	-0.001	0.005	-0.363	-0.022	0.000	0.011
			E-SD	0.022	0.036	0.144	0.031	0.035	0.023	0.054	0.023	0.043	0.051	0.035	0.023	0.067	0.022	0.032	0.084	0.033	0.023	0.049	0.023	0.054	0.070	0.034	0.024	0.080
			RMSE	0.022	0.037	0.167	0.115	0.037	0.023	0.056	0.023	0.052	0.400	0.037	0.023	0.081	0.022	0.037	0.167	0.037	0.023	0.056	0.023	0.052	0.400	0.037	0.023	0.081
			Mbias	-0.001	0.008	0.087	-0.111	-0.012	0.001	0.011	-0.001	0.030	0.397	-0.014	0.004	0.045	-0.001	0.024	-0.203	-0.025	0.003	0.035	-0.001	0.006	-0.363	-0.022	0.001	0.009
(3)	800	5	Bias	-0.001	0.002	0.025	-0.202	-0.003	0.000	0.005	-0.001	0.033	0.401	-0.005	0.003	0.052	-0.001	0.017	-0.156	-0.017	0.001	0.026	-0.001	0.006	-0.363	-0.012	0.000	0.012
			E-SD	0.017	0.029	0.108	0.021	0.027	0.018	0.042	0.018	0.035	0.039	0.027	0.018	0.056	0.017	0.025	0.061	0.026	0.018	0.038	0.018	0.046	0.055	0.026	0.019	0.068
			RMSE	0.017	0.029	0.110	0.203	0.027	0.018	0.042	0.018	0.049	0.403	0.027	0.018	0.076	0.017	0.029	0.110	0.027	0.018	0.042	0.018	0.049	0.403	0.027	0.018	0.076
			Mbias	-0.001	0.003	0.024	-0.202	-0.003	0.000	0.003	-0.001	0.034	0.403	-0.006	0.003	0.050	-0.001	0.016	-0.154	-0.017	0.001	0.024	-0.001	0.006	-0.366	-0.012	0.000	0.009
(4)	800	10	Bias	0.000	0.003	0.024	-0.102	-0.002	0.000	0.005	-0.001	0.036	0.404	-0.005	0.004	0.056	0.000	0.017	-0.153	-0.017	0.002	0.026	-0.001	0.008	-0.365	-0.010	0.001	0.014
			E-SD	0.012	0.019	0.074	0.015	0.017	0.012	0.028	0.012	0.024	0.027	0.017	0.012	0.037	0.012	0.016	0.042	0.016	0.012	0.025	0.012	0.031	0.039	0.016	0.012	0.046
			RMSE	0.012	0.019	0.078	0.103	0.017	0.012	0.028	0.012	0.043	0.405	0.018	0.013	0.067	0.012	0.019	0.078	0.017	0.012	0.028	0.012	0.043	0.405	0.018	0.013	0.067
			Mbias	0.000	0.002	0.025	-0.102	-0.002	0.000	0.004	0.000	0.036	0.404	-0.005	0.005	0.055	0.000	0.016	-0.152	-0.016	0.002	0.024	0.000	0.009	-0.367	-0.010	0.002	0.013

Table 3.3. Simulation results for the SARMA: ρ_0:0.5, λ_0:0.2

Setting				SARMA BC_ED							SARMA BC_BC						SARAR BC_ED						SARAR BC_BC					
	N	T		β	ρ	λ	σ^2	σ^2_{bc}	DE_x	IE_x	β	ρ	λ	σ^2_{bc}	DE_x	IE_x	β	ρ	λ	σ^2_{bc}	DE_x	IE_x	β	ρ	λ	σ^2_{bc}	DE_x	IE_x
(1)	200	5	Bias	-0.002	0.006	0.094	-0.213	-0.016	0.001	0.036	-0.003	0.008	0.164	-0.017	0.002	0.044	-0.002	0.010	-0.128	-0.021	0.002	0.048	-0.002	0.004	-0.159	-0.020	0.001	0.034
			E-SD	0.035	0.048	0.200	0.044	0.055	0.039	0.180	0.036	0.050	0.075	0.054	0.039	0.185	0.035	0.044	0.161	0.054	0.039	0.169	0.036	0.054	0.081	0.054	0.039	0.193
			RMSE	0.035	0.049	0.221	0.217	0.057	0.039	0.184	0.036	0.051	0.180	0.057	0.039	0.190	0.035	0.049	0.221	0.057	0.039	0.184	0.036	0.051	0.180	0.057	0.039	0.190
			Mbias	-0.002	0.005	0.101	-0.215	-0.019	0.002	0.022	-0.002	0.008	0.167	-0.019	0.002	0.033	-0.002	0.009	-0.115	-0.022	0.003	0.034	-0.002	0.005	-0.164	-0.022	0.001	0.020
(2)	200	10	Bias	-0.002	0.007	0.085	-0.110	-0.011	0.001	0.030	-0.002	0.010	0.163	-0.012	0.002	0.042	-0.002	0.010	-0.112	-0.014	0.002	0.041	-0.002	0.008	-0.159	-0.013	0.001	0.034
			E-SD	0.023	0.032	0.136	0.032	0.036	0.026	0.117	0.023	0.033	0.050	0.035	0.025	0.120	0.023	0.029	0.108	0.035	0.025	0.108	0.023	0.035	0.054	0.035	0.026	0.126
			RMSE	0.023	0.032	0.161	0.114	0.037	0.026	0.120	0.023	0.034	0.170	0.037	0.025	0.127	0.023	0.032	0.161	0.037	0.026	0.120	0.023	0.034	0.170	0.037	0.025	0.127
			Mbias	-0.001	0.006	0.088	-0.110	-0.011	0.002	0.023	-0.002	0.011	0.162	-0.012	0.003	0.038	-0.002	0.010	-0.104	-0.014	0.002	0.032	-0.002	0.009	-0.159	-0.014	0.002	0.034
(3)	800	5	Bias	-0.001	0.002	0.025	-0.202	-0.003	0.000	0.011	-0.003	0.016	0.166	-0.007	0.003	0.064	-0.001	0.005	-0.054	-0.006	0.000	0.020	-0.003	0.014	-0.162	-0.008	0.003	0.055
			E-SD	0.018	0.025	0.103	0.022	0.027	0.020	0.091	0.018	0.026	0.037	0.027	0.020	0.100	0.018	0.024	0.081	0.026	0.020	0.087	0.018	0.028	0.042	0.027	0.020	0.107
			RMSE	0.018	0.025	0.106	0.203	0.027	0.020	0.092	0.018	0.031	0.170	0.028	0.020	0.119	0.018	0.025	0.106	0.027	0.020	0.092	0.018	0.031	0.170	0.028	0.020	0.119
			Mbias	-0.002	0.002	0.024	-0.203	-0.003	0.000	0.010	-0.003	0.017	0.166	-0.007	0.003	0.064	-0.002	0.005	-0.048	-0.007	0.001	0.017	-0.003	0.015	-0.163	-0.008	0.003	0.056
(4)	800	10	Bias	-0.001	0.003	0.025	-0.102	-0.002	0.000	0.012	-0.003	0.019	0.168	-0.007	0.004	0.070	-0.001	0.005	-0.051	-0.005	0.001	0.021	-0.003	0.017	-0.164	-0.007	0.003	0.062
			E-SD	0.012	0.017	0.071	0.016	0.018	0.013	0.061	0.012	0.018	0.026	0.018	0.013	0.066	0.012	0.016	0.055	0.017	0.013	0.057	0.012	0.019	0.029	0.017	0.013	0.071
			RMSE	0.012	0.017	0.076	0.103	0.018	0.013	0.062	0.013	0.026	0.170	0.019	0.013	0.096	0.012	0.017	0.076	0.018	0.013	0.062	0.013	0.026	0.170	0.019	0.013	0.096
			Mbias	-0.001	0.002	0.027	-0.102	-0.002	0.001	0.009	-0.003	0.019	0.167	-0.007	0.005	0.069	-0.001	0.005	-0.048	-0.005	0.001	0.019	-0.003	0.017	-0.164	-0.007	0.004	0.062

Figure 3.1. Bias of λ



Since the spatial moving average parameter λ is our point of our interest, Figure 3.1 graphs the bias that occurs in this parameter when adopting a wrong spatial weight matrix, a wrong model, or both. These biases are taken from Tables 3.2 and 3.3. The numbers 1 to 4 on the x-axis indicate the panel data settings of N and T, respectively $(N, T) = (200, 5), (200, 10), (800, 5), (800, 10)$, and the values on the y-axis the magnitude of the biases. When the model is estimated correctly, the bias of ρ decreases as N, T, or both increase, and especially if N increases. Furthermore, the bias is smallest compared to the other cases. Nevertheless, the bias is relatively large compared to that of the other parameters and marginal effects in the panels "SARMA BC_ED" of Tables 3.2 and 3.3. When the wrong model is estimated, namely the SARAR model, Figure 3.1 shows that the bias in λ increases, but that the downward pattern relative to N and T remains. By contrast, when one common spatial weight matrix is adopted for all spatial lags, both in the right model (SARMA) and in the wrong model (SARAR), not only the bias in λ further increases, but also the downward pattern relative to N and T completely disappears.

This finding demonstrates that the selection of the right spatial weight matrix and considering different spatial weight matrices for different spatial lags in the model is of crucial importance and perhaps even more important than the selection of the right spatial econometric model.

3.5 Empirical application: Military expenditures

The empirical analysis in this section is based on Yesilyurt and Elhorst (2017) (YE). YE investigate (the log of the level of) military spending measured as a ratio of GDP, also known as the defense burden, in 144 countries over the period 1993 to 2007. Consequently, $T = 15$ may be considered finite. Explanatory variables are GDP, population, international war, civil war, and political regime.[4] The first two explanatory variables are measured in logs, while the latter three are measured as scores, of which the scores on international war and civil war range from 0 (no war) to 10 (greatest, and on political regime from -10 (strongly autocratic) to +10 (strongly democratic).[5]

YE compare the performance of different spatial econometric models, among which the SAR model, but not the SARMA and the SARAR models, and eight potential specifications of the spatial weight matrix, among which a binary contiguity matrix, but not an exponential or distance decay matrix. Using Bayesian comparison methods developed by LeSage (2014, 2015), they find that the SAR model, the specification that only includes a spatial lag in the dependent variable, in combination with a spatial weight matrix specified as a row-normalized binary contiguity (BC) matrix based on maritime borders produces the highest Bayesian posterior model probability. It makes this study interesting for further research because the SARMA and SARAR models are not considered and LeSage's Bayesian comparison approach does not give any other option than to depart from one common spatial weight matrix for all spatial lags in the model. We investigate whether extending the SAR model with a spatial MA or AR error process and whether utilizing the

[4] Due to linking several data sources, YE were able to collect data on many countries from multiple continents (Europe, Asia, Africa, North and South America, and Australia). The number of 144 countries is also greater than any other study on military expenditures, and also covers countries frequently engaged in tensions (Cuba, Israel, Lebanon, Iran, Syrian, Kuwait, Vietnam, Mali, Ethiopia, both Congos, and Columbia).

[5] For more details on the data and the variables we refer the reader to Yesilyurt and Elhorst (2017).

opportunity to use different spatial weight matrices for each spatial lag improve the results. We consider row-normalized BC, ED, and ID spatial weight matrices, which differ in their degree of sparsity and boundedness, but for reasons explained in Section 3.2.3, we exclude the ID matrix when modelling AR spatial lags and the BC matrix when modelling MA spatial lags. We also consider scalar normalization by the largest eigenvalue of the distance-based spatial weight matrices.

Table 3.4 reports the estimation results using the QML within estimator with bias correction. The individual fixed effects are included but concentrated out and the time fixed effects are captured as year dummies and estimated as regular explanatory variables. One year dummy is dropped to avoid multicollinearity. Just as YE, the direct and indirect spillover effects of each variable are also determined. The delta method introduced in Arbia et al. (2020) is used to calculate the significance levels of these effects. Columns (1) and (2) report the estimation results of the SAR model using either the row-normalized BC or the row-normalized ED spatial weight matrix, respectively. Although the ED matrix was not part of the eight matrices considered by YE, it is clear that when this matrix were added to their choice set, the results would not have changed. The log likelihood function value when adopting the BC matrix is higher than that of the ED matrix. The coefficient estimate of the spatial autoregressive parameter ρ is 0.2368 and significant at the 1% level. Four of the five explanatory variables in the model appear to have coefficients, direct and indirect spillover effects that are significant at the 1%, 5% or 10% level. When adopting the row-normalized ED matrix, only three explanatory variables have this property. For this reason and the fact that its log likelihood function value is lower, we reject the ED matrix and stick to the BC matrix when modelling the spatial lag in the dependent variable in the next steps.

Columns (3) and (4) report the results when the SAR model using the row-normalized BC matrix is extended with a spatial AR error process based on either the row-normalized BC or the row-normalized ED matrix. The results demonstrate that the common practice to use one common spatial weight matrix for the spatial lag in the dependent variable and the spatial lag in the error term needs to be rejected. The log likelihood function value when adopting the row-normalized BC matrix also in the error term is lower than that of the ED matrix. Furthermore, while the spatial autoregressive parameter is close to zero and insignificant (-0.0162, t-value

Table 3.4. Military expenditures according to different combinations of spatial weight matrices

Model / Var \ W	SAR		SARAR		SARMA		
	(1)BC	(2)ED	(3)BC_BC	(4)BC_ED	(5)BC_ED	(6)BC_ID	(7)BC_ID*
GDP	-0.4900***	-0.4982***	-0.4770***	-0.4845***	-0.4838***	-0.5003***	-0.4812***
	(-5.27)	(-5.34)	(-5.05)	(-5.15)	(-5.14)	(-5.32)	(-4.95)
Pop	-0.3437**	-0.2804	-0.4667**	-0.3537*	-0.3596*	-0.3396*	-0.3604**
	(-2.02)	(-1.64)	(-2.32)	(-1.91)	(-1.92)	(-1.83)	(-1.96)
Int. War	0.0733*	0.0725*	0.0785*	0.0751*	0.0758*	0.0749*	0.0759*
	(1.82)	(1.80)	(1.93)	(1.85)	(1.86)	(1.85)	(1.81)
Civil War	0.0092	0.0085	0.0101	0.0098	0.0098	0.0072	0.0069
	(0.63)	(0.58)	(0.70)	(0.68)	(0.68)	(0.50)	(0.46)
PR	-0.0147***	-0.0156***	-0.0137***	-0.0152***	-0.0153***	-0.0143***	-0.0145***
	(-3.07)	(-3.26)	(-2.89)	(-3.18)	(-3.20)	(-3.03)	(-2.96)
ρ	0.2368***	0.3389***	-0.0162	0.1649***	0.1530***	0.1762***	0.2880***
	(8.42)	(8.35)	(-0.51)	(5.54)	(5.11)	(5.97)	(6.44)
λ			0.2587***	0.1744***	-0.2235***	-0.3426***	-0.2918***
			(9.01)	(3.50)	(-3.50)	(-3.40)	(-2.74)
dir_GDP	-0.4962***	-0.5035***	-0.4770***	-0.4874***	-0.4863***	-0.5037***	-0.4853***
	(-5.27)	(-5.34)	(-5.05)	(-5.15)	(-5.14)	(-5.32)	(-4.95)
dir_Pop	-0.3481**	-0.2835	-0.4667**	-0.3557*	-0.3614*	-0.3419*	-0.3635**
	(-2.02)	(-1.64)	(-2.32)	(-1.91)	(-1.92)	(-1.83)	(-1.96)
dir_Int. War	0.0742*	0.0733*	0.0785*	0.0755*	0.0762*	0.0754*	0.0765*
	(1.82)	(1.80)	(1.93)	(1.85)	(1.86)	(1.85)	(1.81)
dir_Civil War	0.0093	0.0086	0.0101	0.0099	0.0099	0.0072	0.0070
	(0.63)	(0.58)	(0.70)	(0.68)	(0.68)	(0.50)	(0.46)
dir_PR	-0.0148***	-0.0158***	-0.0137***	-0.0153***	-0.0154***	-0.0144***	-0.0147***
	(-3.07)	(-3.26)	(-2.89)	(-3.18)	(-3.2)	(-3.03)	(-2.96)
ind_GDP	-0.1448***	-0.2500***	0.0076	-0.0921***	-0.0843***	-0.1029***	-0.1172***
	(-4.12)	(-3.86)	(0.52)	(-3.49)	(-3.35)	(-3.66)	(-3.49)
ind_Pop	-0.1016*	-0.1407	0.0074	-0.0673*	-0.0627*	-0.0698*	-0.0878*
	(-1.93)	(-1.58)	(0.51)	(-1.77)	(-1.76)	(-1.72)	(-1.82)
ind_Int. War	0.0217*	0.0364*	-0.0012	0.0143*	0.0132*	0.0154*	0.0185*
	(1.75)	(1.71)	(-0.50)	(1.72)	(1.72)	(1.73)	(1.70)
ind_Civil War	0.0027	0.0043	-0.0002	0.0019	0.0017	0.0015	0.0017
	(0.63)	(0.58)	(-0.42)	(0.67)	(0.67)	(0.49)	(0.46)
ind_PR	-0.0043***	-0.0078***	0.0002	-0.0029***	-0.0027***	-0.0029***	-0.0035***
	(-2.79)	(-2.82)	(0.51)	(-2.64)	(-2.60)	(-2.60)	(-2.53)
logL	-1390.2946	-1394.6256	-1389.5789	-1387.5085	-1387.1566	-1387.0925	-1394.7790
rsqr	0.7003	0.6984	0.6861	0.6975	0.6970	0.6980	0.6970

***significant at 1% level **significant at 5% level *significant at 10% level

-0.51) when using the BC matrix, it is positive and significant (0.1649, t-value 5.54) when using the ED matrix. The estimates of λ indicating the spatial dependence in the error terms are significant in both settings. This makes sense. It is likely that the spatial pattern of unobserved shocks in the error term go beyond the first-order neighbours of each country. When modelling the AR error process by the same BC matrix as for the spatial lag in the dependent variable, it finally appears that none of the indirect spillover effects is significant any more.

Columns (5) and (6) report the results when estimating the SARMA model, in particular the SAR model using the row-normalized BC matrix extended with a spatial MA error process based on spatial weight matrices that are denser than the binary contiguity matrix, either the row-normalized ED or the row-normalized ID matrix. In line with the purpose of this chapter, it first appears that the log likelihood function value of the SARMA model compared to the SARAR model when using the same two spatial weight matrices, the row-normalized BC matrix for the spatial lag in the dependent variable and the row-normalized ED matrix for the spatial lag in the error term, appears to be higher. The difference is nonetheless small: -1387.1566 versus -1387.5085. This is because they yield largely similar results in terms of magnitude and significance. Second, it appears that the performance of the SARMA model in terms of the log likelihood function value can be improved further to -1387.0925 by replacing the row-normalized ED matrix by the row-normalized ID matrix. This finding supports the hypotheses that shocks are local because the spatial pattern of unobserved shocks in the error term go beyond the first-order neighbours of each country, that these shocks for this reason are better modelled by a spatial MA error process than by a spatial AR process, and that a spatial weight matrix denser than the binary contiguity matrix used to model the spatial lag in the dependent variables is the appropriate matrix to model this MA error process. The estimated spatial moving average parameter of the error term is -0.3426 and significant at the 1% level; it turns out to have a downward effect on the estimated autoregressive parameter in that it decreases from 0.2368 in the SAR model to 0.1762 in the SARMA model. This finding implies that if a particular country increases its military expenditure as a ratio of GDP (say by 1% point), neighbouring countries will follow (by 0.18% points), though to a lesser extent than is found when estimating the SAR model without a spatial lag in the error term.

Column (7) of Table 3.4 finally reports the estimation results when the row normalized inverse distance matrix is replaced by its scalar normalized counterpart. Since the log likelihood value does not increase, this alternative specification of the inverse distance matrix appears not to be fruitful.[6]

When comparing the parameter estimates, the direct effects, and the indirect spillover effects in the different columns with each other, it initially seems as if the differences are rather small. Especially, the signs, magnitudes, and significance levels of the parameters estimates and the direct effects do not differ to any great extent from each other, no matter which model is used (SAR, SARAR, or SARMA) and which specification of the spatial weight matrix is used (BC, ED, or ID). However, a different conclusion emerges considering the indirect spillover effects when a model is estimated with the "wrong" spatial weight matrix, such as the SAR model with an ED matrix in Column (2), or when one common spatial weight matrix is used for all spatial lags in the model, such as the SARAR model with a BC matrix in Column (3). Then both the signs and significance levels tend to be different. This is an important finding since indirect spillover effects are often the main focus of applied practitioners. They want to know whether changing a variable in one country has an effect on other counties. Our search for the most likely model and the most likely spatial weight matrix for the different spatial lags in the model show that this can really make a difference. The results of the SARMA models are more robust when different spatial weight matrices are used. The results found for the best performing model, the SARMA model with a spatial lag modelled by the BC matrix and a spatial MA error process modelled by the ID matrix, show that a focal country surrounded by countries with higher levels of GDP, countries with larger population, countries involved in an international war, and countries with an autocratic regime have a significant upward effect on the military expenditures of that country.

3.6 Conclusion

The SARMA model received considerably less attention in the theoretical and applied econometric literature than the SARAR model. We identify two reasons for

[6] The estimated results of the full set of models with scalar normalized spatial weight matrices can be obtained upon request.

this. First, the estimation of the spatial MA autocorrelation parameter is computationally more demanding than of spatial AR autocorrelation parameter since the matrix $\boldsymbol{R}(\lambda)$ needs to be inverted multiple times when estimating the SARMA model, while this inversion is not needed when estimating the SARAR model. This may have been an obstacle in the past when computing facilities were relatively limited. Second, estimation routines of the SARMA model have not been made available to practitioners. This chapter attempts to break this practice along four ways.

First, we develop the QML within estimator with bias correction to estimate the parameters of the SARMA model when N is large and T is finite. We investigate its asymptotic sample properties from an econometric-theoretical viewpoint and their finite sample properties from an empirical viewpoint using Monte Carlo simulations. The QML within estimator appears to be consistent, asymptotically normal, and due to the bias correction properly centred, while its finite sample properties appear to be acceptable and to improve as T increases but above all when N increases. The QML within estimator without bias correction needs to be rejected. We also investigate what happens when the data is generated by a SARMA model with different spatial weight matrices for each spatial lag but estimated by the SARMA model with one common spatial weight matrix, the SARAR model with similar spatial weight matrices, or the SARAR model with one common spatial weight matrix. This simulation experiment shows that the selection of the right spatial weight matrix and considering different spatial weight matrices for different spatial lags in the model is of crucial importance.

Second, the proposed estimator offers the opportunity to consider different spatial weight matrices for the spatial lags in the model, as well as spatial weight matrices that are normalized by its largest eigenvalue rather than row normalized. We have shown that this improves empirical applications.

Third, a Matlab routine of the proposed QML estimators will be made available to give other researchers the opportunity to apply it to their own data sets.

Fourth, we describe and compare the econometric and economic properties of spatial moving average errors and spatial autoregressive errors with each other in relation to the degree of sparsity and boundedness of three spatial weight matrices, and demonstrate that the interpretation of spatial moving average errors in various

empirical applications, among which our empirical application based on military expenditures, is more obvious than spatial autoregressive errors since most shocks are local rather than global.

In follow-up research we aim at extending the SARMA and SARAR models further with spatial lags in the explanatory variables, leading to so-called general nesting spatial models, and with parameterized specifications of distance decay spatial weight matrices to further increase the flexibility with which these matrices can be specified.

3.A Appendix: QML within estimator with bias correction

3.A.1 First and second order derivatives of the log likelihood in (3.8)

The first order derivatives are

$$\frac{1}{\sqrt{NT}}\frac{\partial \ln L^w(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{1}{\sigma^2}\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}(\lambda)\widetilde{\boldsymbol{X}}_t)'\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}) \\ \frac{1}{\sigma^2}\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}[(\boldsymbol{R}^{-1}(\lambda)\boldsymbol{W}\widetilde{\boldsymbol{y}}_t)'\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}) - \mathrm{tr}\boldsymbol{G}(\rho)] \\ \frac{1}{\sigma^2}\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}[-(\boldsymbol{K}(\lambda)\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}))'\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}) + \mathrm{tr}\boldsymbol{K}(\lambda)] \\ \frac{1}{2\sigma^4}\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}(\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\zeta})\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}) - N\sigma^2) \end{pmatrix}, \quad (3.A.1)$$

where $\boldsymbol{S}(\rho) = \boldsymbol{I}_N - \rho\boldsymbol{W}$, $\boldsymbol{G}(\rho) = \boldsymbol{W}\boldsymbol{S}^{-1}(\rho)$, $\boldsymbol{R}(\lambda) = \boldsymbol{I}_N - \lambda\boldsymbol{M}$, $\boldsymbol{H}(\lambda) = \boldsymbol{M}\boldsymbol{R}^{-1}(\lambda)$, and $\boldsymbol{K}(\lambda) = \boldsymbol{R}^{-1}(\lambda)\boldsymbol{M}$.

The score of the log likelihood function evaluated at $\boldsymbol{\theta}_T$, where $\boldsymbol{\theta}_T = \boldsymbol{\theta}_0 - (\boldsymbol{0}_{1\times(K+2)}, \frac{1}{T}\sigma_0^2)'$ is defined in Theorem 1, is

$$\frac{1}{\sqrt{NT}}\frac{\partial \ln L^w(\boldsymbol{\theta}_T)}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{1}{\sigma_T^2}\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\ddot{\boldsymbol{X}}_t'\widetilde{\boldsymbol{v}}_t \\ \frac{1}{\sigma_T^2}\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}(\ddot{\boldsymbol{G}}\ddot{\boldsymbol{X}}_t\boldsymbol{\beta}_0)'\widetilde{\boldsymbol{v}}_t + \frac{1}{\sigma_T^2}\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}(\widetilde{\boldsymbol{v}}_t'\ddot{\boldsymbol{G}}'\widetilde{\boldsymbol{v}}_t - \sigma_T^2\mathrm{tr}\ddot{\boldsymbol{G}}) \\ \frac{1}{\sigma_T^2}\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}(-\widetilde{\boldsymbol{v}}_t'\boldsymbol{K}'\widetilde{\boldsymbol{v}}_t + \sigma_T^2\mathrm{tr}\boldsymbol{K}) \\ \frac{1}{2\sigma_T^4}\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}(\widetilde{\boldsymbol{v}}_t'\widetilde{\boldsymbol{v}}_t - N\sigma_T^2) \end{pmatrix}, \quad (3.A.2)$$

where $\ddot{\boldsymbol{W}} = \boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{R}$, $\ddot{\boldsymbol{G}} = \ddot{\boldsymbol{W}}(\boldsymbol{I}_N - \rho\ddot{\boldsymbol{W}})^{-1} = \boldsymbol{R}^{-1}\boldsymbol{G}\boldsymbol{R}$, and $\ddot{\boldsymbol{X}}_t = \boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t$.

The second order derivatives are

$$
\begin{aligned}
\frac{1}{NT}\frac{\partial^2 \ln L^w(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} = \\
-\frac{1}{NT}
&\begin{pmatrix}
\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}(\lambda)\widetilde{\boldsymbol{X}}_t)'\boldsymbol{R}^{-1}(\lambda)\widetilde{\boldsymbol{X}}_t & * & * \\
\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}(\lambda)\boldsymbol{W}\widetilde{\boldsymbol{y}}_t)'\boldsymbol{R}^{-1}(\lambda)\widetilde{\boldsymbol{X}}_t & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}(\lambda)\boldsymbol{W}\widetilde{\boldsymbol{y}}_t)'\boldsymbol{R}^{-1}(\lambda)\boldsymbol{W}\widetilde{\boldsymbol{y}}_t + T\mathrm{tr}[\boldsymbol{G}^2(\rho)] & * \\
-\frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\zeta})\boldsymbol{R}^{-1}(\lambda)\boldsymbol{H}(\lambda)\widetilde{\boldsymbol{X}}_t - \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{K}(\lambda)\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}))'\boldsymbol{R}^{-1}(\lambda)\widetilde{\boldsymbol{X}}_t & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}(\lambda)\boldsymbol{H}(\lambda)\boldsymbol{W}\widetilde{\boldsymbol{y}}_t)'\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}) - \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{K}(\lambda)\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}))'\boldsymbol{R}^{-1}(\lambda)\boldsymbol{W}\widetilde{\boldsymbol{y}}_t & 0 & 0 \\
\frac{1}{\sigma^4}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\zeta})\boldsymbol{R}^{-1}(\lambda)\widetilde{\boldsymbol{X}}_t & \frac{1}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}(\lambda)\boldsymbol{W}\widetilde{\boldsymbol{y}}_t)'\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}) & 0 & 0
\end{pmatrix} \\
-\frac{1}{NT}
&\begin{pmatrix}
\boldsymbol{0}_{K\times K} & \boldsymbol{0}_{K\times 1} & \boldsymbol{0}_{K\times 1} & \boldsymbol{0}_{K\times 1} \\
\boldsymbol{0}_{1\times k} & 0 & 0 & 0 \\
\boldsymbol{0}_{1\times k} & 0 & \frac{1}{\sigma^2}\sum_{t=1}^{T}(2\boldsymbol{K}(\lambda)\boldsymbol{K}(\lambda)\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}))'\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}) + \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{K}(\lambda)\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}))'\boldsymbol{K}(\lambda)\widetilde{\boldsymbol{v}}_t - T\mathrm{tr}[\boldsymbol{H}^2(\lambda)] & * \\
\boldsymbol{0}_{1\times K} & 0 & -\frac{1}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{K}(\lambda)\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}))'\widetilde{\boldsymbol{v}}_t & -\frac{NT}{2\sigma^4}+\frac{1}{\sigma^6}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\zeta})\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta})
\end{pmatrix}.
\end{aligned}
\tag{3.A.3}
$$

Given the second order derivatives, the score of the log likelihood at $\boldsymbol{\theta}_T$, and Lemma A.4 of Lee and Yu (2010a), the information matrix $\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T}$ takes the form

$$\begin{aligned}\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T} &= -\mathrm{E}\frac{1}{NT}\frac{\partial^2 \ln L^w(\boldsymbol{\theta}_T)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \\ &= \frac{1}{\sigma_T^2}\begin{pmatrix} \mathbb{H} & * & * \\ \mathbf{0}_{1\times(K+1)} & 0 & * \\ \mathbf{0}_{1\times(K+1)} & 0 & 0 \end{pmatrix} \\ &+ \begin{pmatrix} \mathbf{0}_{K\times K} & * & * & * \\ \mathbf{0}_{1\times K} & \frac{1}{N}\mathrm{tr}(\ddot{\boldsymbol{G}}^s\ddot{\boldsymbol{G}}) & * & * \\ \mathbf{0}_{1\times K} & -\frac{1}{N}\mathrm{tr}(\boldsymbol{K}^s\ddot{\boldsymbol{G}}) & \frac{1}{N}\mathrm{tr}(\boldsymbol{K}^s\boldsymbol{K}) & * \\ \mathbf{0}_{1\times K} & \frac{1}{\sigma_T^2 N}\mathrm{tr}(\ddot{\boldsymbol{G}}) & -\frac{1}{\sigma_T^2 N}\mathrm{tr}(\boldsymbol{K}) & \frac{1}{2\sigma_T^4} \end{pmatrix},\end{aligned} \tag{3.A.4}$$

where $\mathbb{H} = \frac{1}{NT}\sum_{t=1}^{T}(\ddot{\boldsymbol{X}}_t, \ddot{\boldsymbol{G}}\ddot{\boldsymbol{X}}_t\boldsymbol{\beta}_0)'(\ddot{\boldsymbol{X}}_t, \ddot{\boldsymbol{G}}\ddot{\boldsymbol{X}}_t\boldsymbol{\beta}_0)$. A matrix with superscript s is defined by $\boldsymbol{A}^s = \boldsymbol{A}' + \boldsymbol{A}$ ($\boldsymbol{A} = \ddot{\boldsymbol{G}}^s, \boldsymbol{K}^s$). The matrix $\boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T}$ in Theorem 2 reads as

$$\begin{aligned}\boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T} &= \frac{T-1}{T}\frac{\mu_4 - 3\sigma_0^4}{\sigma_0^4}\times \\ &\begin{pmatrix} \mathbf{0}_{K\times K} & * & * & * \\ \mathbf{0}_{1\times K} & \frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{G}}_{ii}^2 & * & * \\ \mathbf{0}_{1\times K} & -\frac{1}{N}\sum_{i=1}^{N}\ddot{\boldsymbol{G}}_{ii}\boldsymbol{K}_{ii} & \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{K}_{ii}^2 & * \\ \mathbf{0}_{1\times K} & \frac{1}{2\sigma_T^2 N}\mathrm{tr}\ddot{\boldsymbol{G}} & -\frac{1}{2\sigma_T^2 N}\mathrm{tr}\boldsymbol{K} & \frac{1}{4\sigma_T^4} \end{pmatrix}.\end{aligned} \tag{3.A.5}$$

This matrix captures the correction that is needed to the information matrix when the error terms are not assumed to be normally distributed.

3.A.2 Proof of Theorem 1 (Consistency)

The proofs closely follow the steps and notations used in Lee and Yu (2010a), but we do make modifications for using the SARMA rather than the SARAR model. We present the analysis under the asymptotic setting that N tends to infinity and T is finite.

As pointed out in Section 3.3, one problem of the within transformation is that the resulting error terms are linearly dependent over the time dimension. Consequently, the estimate $\widehat{\sigma}^{2w}$ of σ_0^2 using the within transformation will be consistent

only when T is large. When T is finite, $\widehat{\sigma}^{2w}$ should be bias-corrected by $\frac{T}{T-1}\widehat{\sigma}^{2w}$ to obtain a consistent estimate of σ_0^2. This is demonstrated by Arellano (2003, Section 2.2) for non-spatial panel models, by Lee and Yu (2012) for spatial panel data models, and by Lee and Yu (2010a) for the SARAR model. One major difference between the last two studies is that Lee and Yu (2012) concentrate out individual specific effects by taking first-differences, while Lee and Yu (2010a) concentrate them out by an orthogonal transformation. However, all three studies conclude that the required bias correction can be effected by replacing T in (3.11) with $T-1$,[7] to get the adjusted log likelihood (not indicated by a symbol)

$$\begin{aligned}\ln L(\boldsymbol{\theta}) = &-\frac{N(T-1)}{2}\ln(2\pi\sigma^2) + (T-1)[\ln|\boldsymbol{S}(\rho)| - \ln|\boldsymbol{R}(\lambda)|] \\ &-\frac{1}{2\sigma^2}\sum_{t=1}^{T-1}\boldsymbol{v}_t^{*'}(\boldsymbol{\zeta})\boldsymbol{v}_t^*(\boldsymbol{\zeta}),\end{aligned} \tag{3.A.6}$$

where $\boldsymbol{v}_t^*(\boldsymbol{\zeta}) = \boldsymbol{R}^{-1}(\lambda)[\boldsymbol{S}(\rho)\boldsymbol{y}_t^* - \boldsymbol{X}_t^*\boldsymbol{\beta}]$. To be able to compare this log likelihood with that of the within transformation in Equation (3.8), we use the property that $\sum_{t=1}^{T-1}\boldsymbol{v}_t^{*'}\boldsymbol{v}_t^* = (\boldsymbol{v}_1', ..., \boldsymbol{v}_T')\,(\boldsymbol{J}_T \otimes \boldsymbol{I}_N)(\boldsymbol{v}_1, ..., \boldsymbol{v}_T) = \sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t'\widetilde{\boldsymbol{v}}_t$, where $\boldsymbol{J}_T = \boldsymbol{I}_T - \frac{1}{T}\boldsymbol{\iota}_T\boldsymbol{\iota}_T'$, and rewrite (3.A.6) as

$$\begin{aligned}\ln L(\boldsymbol{\theta}) = &-\frac{N(T-1)}{2}\ln(2\pi\sigma^2) + (T-1)[\ln|\boldsymbol{S}(\rho)| - \ln|\boldsymbol{R}(\lambda)|] \\ &-\frac{1}{2\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\zeta})\widetilde{\boldsymbol{v}}_t(\boldsymbol{\zeta}).\end{aligned} \tag{3.A.7}$$

The QML estimators of $\boldsymbol{\beta}$ and σ^2, given ρ and λ, starting from this log likelihood are

$$\widehat{\boldsymbol{\beta}}(\rho,\lambda) = [\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t'\boldsymbol{R}^{-1'}(\lambda)\boldsymbol{R}^{-1}(\lambda)\widetilde{\boldsymbol{X}}_t]^{-1}[\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t'\boldsymbol{R}^{-1'}(\lambda)\boldsymbol{R}^{-1}\boldsymbol{S}(\rho)\widetilde{\boldsymbol{y}}_t], \tag{3.A.8}$$

$$\begin{aligned}\widehat{\sigma}^2(\rho,\lambda) = &\frac{1}{N(T-1)}\sum_{t=1}^{T}[\boldsymbol{S}(\rho)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t\widehat{\boldsymbol{\beta}}(\rho,\lambda)]' \\ &\times \boldsymbol{R}^{-1'}(\lambda)\boldsymbol{R}^{-1}(\lambda)[\boldsymbol{S}(\rho)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t\widehat{\boldsymbol{\beta}}(\rho,\lambda)],\end{aligned} \tag{3.A.9}$$

[7] A detailed explanation is provided by Lee and Yu (2012, p.1373).

as a result of which the concentrated log likelihood function of (ρ, λ) is

$$\begin{aligned} \ln L(\rho,\lambda) = &-\frac{N(T-1)}{2}(\ln(2\pi)+1) - \frac{N(T-1)}{2}\ln\widehat{\sigma}^2(\rho,\lambda) \\ &+ (T-1)[\ln|\boldsymbol{S}(\rho)| - \ln|\boldsymbol{R}(\lambda)|]. \end{aligned} \tag{3.A.10}$$

By comparing (3.9)-(3.10) with (3.A.8)-(3.A.9), it appears that $\widehat{\boldsymbol{\beta}}(\rho,\lambda) = \widehat{\boldsymbol{\beta}}^{w}(\rho,\lambda)$ and $\widehat{\sigma}^2(\rho,\lambda) = \frac{T}{T-1}\widehat{\sigma}^{2w}(\rho,\lambda)$. Hence, after plugging in the estimates from the adjusted log likelihood into the concentrated log likelihood of the within transformation, (3.11) can be rewritten as

$$\begin{aligned} \ln L^{w}(\rho,\lambda) = &-\frac{NT}{2}(\ln(2\pi) + \ln(\frac{T-1}{T}) + 1) - \frac{NT}{2}\ln\widehat{\sigma}^2(\rho,\lambda) \\ &+ T[\ln|\boldsymbol{S}(\rho)| - \ln|\boldsymbol{R}(\lambda)|]. \end{aligned} \tag{3.A.11}$$

The expressions in (3.A.7) and (3.A.11) show that the difference between $T-1$ and T has no effect and that the maximum for $\widehat{\rho}$ and $\widehat{\lambda}$ of both transformations will be the same. Therefore, $\widehat{\boldsymbol{\beta}}^{w}(\rho,\lambda)$ and $\widehat{\boldsymbol{\beta}}(\rho,\lambda)$ using the log likelihood of the within transformation and its adjustment will also be the same. By contrast, the estimate of σ_0^2 using the within transformation will be different and consistent only when T is large. However, departing from the expressions for $\widehat{\sigma}^{2w}(\rho,\lambda)$ and $\widehat{\sigma}^2(\rho,\lambda)$, $\widehat{\sigma}^{2w}(\rho,\lambda)$ can be easily be bias-corrected by $\frac{T}{T-1}\widehat{\sigma}^{2w}$, which is numerically equal to $\widehat{\sigma}^2(\rho,\lambda)$.

Global identification of (ρ_0, λ_0)

Corresponding to $\frac{1}{N(T-1)}\ln L(\rho,\lambda)$ in (3.A.7), define $Q(\rho,\lambda) = \max_{\boldsymbol{\beta},\sigma^2} \mathrm{E}\frac{1}{N(T-1)}\ln L(\boldsymbol{\theta})$ and let $\mathcal{H}_{\rho_0}(\lambda) = \mathcal{H}_3(\lambda) - \mathcal{H}_2'(\lambda)\mathcal{H}_1^{-1}(\lambda)\mathcal{H}_2(\lambda)$. where $\mathcal{H}_i(\lambda)$ for $i = 1,2,3$ are the corresponding components of $\mathcal{H}(\lambda)$ in (3.12). We have

$$Q(\rho,\lambda) = -(\ln(2\pi)+1) - \frac{1}{2}\sigma^{*2}(\rho,\lambda) + \frac{1}{N}[\ln\boldsymbol{S}(\rho) - \ln\boldsymbol{R}(\lambda)], \tag{3.A.12}$$

where

$$\begin{aligned} \sigma^{*2}(\rho,\lambda) &= (\rho_0-\rho)^2\mathcal{H}_{\rho_0}(\lambda) + \sigma^{2\dagger}(\rho,\lambda) \\ &= (\rho_0-\rho)\mathcal{H}_{\rho_0}(\lambda) + \sigma_0^2\frac{1}{N}\mathrm{tr}(\boldsymbol{R}'\boldsymbol{S}'^{-1}\boldsymbol{S}'(\rho)\boldsymbol{R}'^{-1}(\lambda)\boldsymbol{R}^{-1}(\lambda)\boldsymbol{S}(\rho)\boldsymbol{S}^{-1}\boldsymbol{R}). \end{aligned} \tag{3.A.13}$$

At the true parameters, $Q(\rho_0,\lambda_0) = -\frac{1}{2}(\ln 2\pi + 1) - \frac{1}{2}\ln\sigma_0^2 + \frac{1}{N}\ln|\boldsymbol{S}(\rho_0)| - \frac{1}{n}\ln|\boldsymbol{R}(\lambda_0)|$. We need to prove that $\lim Q(\rho,\lambda) < \lim Q(\rho_0,\lambda_0)$ for any $(\rho,\lambda) \neq (\rho_0,\lambda_0)$. We have

$$\begin{aligned} &Q(\rho,\lambda) - Q(\rho_0,\lambda_0) = \\ &-\frac{1}{2}[\ln\sigma^{*2}(\rho,\lambda) - \ln\sigma_0^2] + \frac{1}{N}\ln|\boldsymbol{S}(\rho)| - \frac{1}{N}\ln|\boldsymbol{S}(\rho_0)| + [-(\frac{1}{N}\ln|\boldsymbol{R}(\lambda)| - \frac{1}{N}\ln|\boldsymbol{R}(\lambda_0)|)] \\ &= T_1(\rho,\lambda) - T_2(\rho,\lambda), \end{aligned}$$

where

$$T_1(\rho,\lambda) = -\frac{1}{2}[\ln\sigma^2(\rho,\lambda) - \ln\sigma_0^2] + \frac{1}{N}\ln|\boldsymbol{S}(\rho)| - \frac{1}{N}\ln|\boldsymbol{S}(\rho_0)| + [-(\frac{1}{N}\ln|\boldsymbol{R}(\lambda)| - \frac{1}{N}\ln|\boldsymbol{R}(\lambda_0)|)],$$

and $T_2(\rho,\lambda) = \frac{1}{2}\ln(1 + \frac{(\rho_0-\rho)^2(\mathcal{H}_3(\lambda) - \mathcal{H}_2'(\lambda)\mathcal{H}_1^{-1}(\lambda)\mathcal{H}_2(\lambda))}{\sigma^{2\dagger}(\rho,\lambda)})$.

Consider the spatial process $\boldsymbol{y}_t = \rho_0\boldsymbol{W}\boldsymbol{y}_t + \boldsymbol{u}_t$ without explanatory variables and the spatial MA error process $\boldsymbol{u}_t = \boldsymbol{v}_t - \lambda_0\boldsymbol{M}\boldsymbol{v}_t$ in period t. The log likelihood of this process is

$$\ln L_p(\rho,\lambda,\sigma^2) = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\ln\sigma^2 + \ln|\boldsymbol{S}(\rho)| - \ln|\boldsymbol{R}(\lambda)| - \frac{1}{\sigma^2}\boldsymbol{v}_t'(\rho,\lambda)\boldsymbol{J}_N\boldsymbol{v}_t(\rho,\lambda),$$

where $\boldsymbol{v}_t = \boldsymbol{R}^{-1}(\lambda)\boldsymbol{S}(\rho)\boldsymbol{y}_t$. Let $Q_p(\rho,\lambda) = \max_{\sigma^2}\mathrm{E}\frac{1}{N}\ln L_p(\rho,\lambda,\sigma^2)$ and $Q_p(\rho_0,\lambda_0)$ be $Q_p(\rho,\lambda)$ evaluated at (ρ_0,λ_0). It follows that $Q_p(\rho,\lambda) - Q_p(\rho_0,\lambda_0) = T_1(\rho,\lambda)$. Due to the information inequality $Q_p(\rho,\lambda) - Q_p(\rho_0,\lambda_0) \leq 0$, we have $T_1(\rho,\lambda) \leq 0$ for any (ρ,λ). By contrast, as $(\rho_0-\rho)^2(\mathcal{H}_3(\lambda) - \mathcal{H}_2'(\lambda)\mathcal{H}_1^{-1}(\lambda)\mathcal{H}_2(\lambda))$ is a quadratic function of ρ, given λ, and $\sigma^{2\dagger}(\rho,\lambda)$ is bounded away from zero, $T_2(\rho,\lambda) \geq 0$.

The finding that $Q_p(\rho,\lambda) - Q_p(\rho_0,\lambda_0) \leq 0$ for any (ρ,λ) implies that $-\frac{1}{2}\ln\sigma^{2\dagger}(\rho,\lambda) \leq -\frac{1}{2}\ln\sigma_0^2 + \frac{1}{N}\ln|\boldsymbol{S}(\rho)| - \frac{1}{N}\ln|\boldsymbol{S}(\rho_0)| + [-(\frac{1}{N}\ln|\boldsymbol{R}(\lambda)| - \frac{1}{N}\ln|\boldsymbol{R}(\lambda_0)|)]$. As $\frac{1}{N}\ln|\boldsymbol{S}(\rho)| - \frac{1}{N}\ln|\boldsymbol{S}(\rho_0)|$ and $\frac{1}{N}\ln|\boldsymbol{R}(\lambda)| - \frac{1}{N}\ln|\boldsymbol{R}(\lambda_0)|$ and thus $-(\frac{1}{N}\ln|\boldsymbol{R}(\lambda)| - \frac{1}{N}\ln|\boldsymbol{R}(\lambda_0)|)$ are all O(1) uniformly in (ρ,λ), $-\ln\sigma^{2\dagger}(\rho,\lambda)$ is bounded from above as σ_0^2 is bounded away from 0. Hence, $\sigma^{2\dagger}(\rho,\lambda)$ is is bounded away from 0.

Under Assumption 3.7(b), $\mathcal{H}_3(\lambda) - \mathcal{H}_2'(\lambda)\mathcal{H}_1^{-1}(\lambda)\mathcal{H}_2(\lambda)$ is positive, so that $T_2(\rho,\lambda) > 0$ for $\rho \neq \rho_0$ for any given λ. Given ρ_0,λ_0 is the unique maximizer of $\lim T_1(\rho,\lambda)$ under

$$\lim(\frac{1}{N}|\sigma_0^2\boldsymbol{R}'\boldsymbol{R}| - \frac{1}{N}|\sigma^{2\dagger}(\rho,\lambda)\boldsymbol{R}'(\lambda)\boldsymbol{R}(\lambda)|) \neq 0,$$

for $\lambda \neq \lambda_0$. Hence, both ρ_0 and λ_0 are identified. When Assumption 3.7(b) does not

hold, identification requires that $T_1(\rho,\lambda) < 0$. Under Assumption 3.7(a), we will have $T_1(\rho,\lambda) < 0$ whenever $(\rho,\lambda) \neq (\rho_0,\lambda_0)$. Hence, $\lim[Q(\rho_0,\lambda_0) - Q(\rho,\lambda)] > 0$ if $(\rho,\lambda) \neq (\rho_0,\lambda_0)$. This proves global identification of (ρ_0,λ_0).

Uniform convergence of $\frac{1}{N(T-1)}\ln L(\rho,\lambda) - Q(\rho,\lambda)$
As $\frac{1}{NT}\ln L^w(\rho,\lambda) - Q^w(\rho,\lambda) = \frac{1}{N(T-1)}\ln L(\rho,\lambda) - Q(\rho,\lambda)$, where $Q^w(\rho,\lambda) = \max_{\boldsymbol{\beta},\sigma^2} \mathrm{E}\frac{1}{NT}\ln L^w(\boldsymbol{\theta})$, we also prove the latter for simplicity. Denote

$$\widetilde{\mathcal{X}}_{\chi,t}(\lambda) = \boldsymbol{R}^{-1}(\lambda)(\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{'}\mathcal{H}_1^{-1}(\lambda)\mathcal{H}_2(\lambda))$$

and

$$\mathcal{V}_{\chi,t}(\lambda) = \frac{1}{N(T-1)}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{'}\boldsymbol{R}^{-1'}(\lambda)\boldsymbol{R}^{-1}(\lambda)\boldsymbol{S}(\rho)\boldsymbol{S}^{-1}\boldsymbol{R}\widetilde{\boldsymbol{v}}_t.$$

We have

$$\begin{aligned}\widehat{\sigma}^2(\rho,\lambda) &= (\rho_0-\rho)^2\mathcal{H}_{\rho_0}(\lambda)\\ &+ \frac{1}{N(T-1)}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t^{'}\boldsymbol{R}^{'}\boldsymbol{S}^{'-1}\boldsymbol{S}^{'}(\rho)\boldsymbol{R}^{'-1}(\lambda)\boldsymbol{R}^{-1}(\lambda)\boldsymbol{S}(\rho)\boldsymbol{S}^{-1}\boldsymbol{R}\widetilde{\boldsymbol{v}}_t\\ &+ 2(\rho_0-\rho)\frac{1}{N(T-1)}\sum_{t=1}^{T}\widetilde{\mathcal{X}}_{\chi,t}(\lambda)\boldsymbol{R}^{'-1}(\lambda)\boldsymbol{R}^{-1}(\lambda)\boldsymbol{S}(\rho)\boldsymbol{S}^{-1}\boldsymbol{R}\widetilde{\boldsymbol{v}}_t\\ &- \mathcal{V}_{\chi,t}^{'}(\lambda)\mathcal{H}_1^{-1}(\lambda)\mathcal{V}_{\chi,t}(\lambda), \qquad (3.A.14)\end{aligned}$$

as a result of which

$$\widehat{\sigma}^2(\rho_0,\lambda_0) = \sigma_0^2 + O_p(\frac{1}{\sqrt{NT}}) \text{ and } \widehat{\sigma}^{*2}(\rho_0,\lambda_0) = \sigma_0^2. \qquad (3.A.15)$$

From (3.A.10) and (3.A.12), $\frac{1}{N(T-1)}\ln L(\rho,\lambda) - Q(\rho,\lambda) = \frac{1}{2}\ln\sigma^{*2}(\rho,\lambda) - \frac{1}{2}\ln\widehat{\sigma}^2(\rho,\lambda)$. By the mean value theorem, $\frac{1}{N(T-1)}\ln L(\rho,\lambda) - Q(\rho,\lambda) = -\frac{1}{2}\frac{1}{\widetilde{\sigma}^2(\rho,\lambda)}(\widehat{\sigma}^2(\rho,\lambda) - \sigma^{*2}(\rho,\lambda))$, where $\widetilde{\sigma}^2(\rho,\lambda)$ lies between $\widehat{\sigma}^2(\rho,\lambda)$ and $\sigma^{*2}(\rho,\lambda)$. We need to show that (1) $\widehat{\sigma}^2(\rho,\lambda) - \sigma^{*2}(\rho,\lambda) \xrightarrow{p} 0$ uniformly in ρ and λ and (2) $\widetilde{\sigma}^2(\rho,\lambda)$ is bounded away from zero uniformly in ρ and λ in probability.

To prove Theorem 1: We have $\widehat{\sigma}^2(\rho,\lambda)$ and $\sigma^{*2}(\rho,\lambda)$ in (3.A.14) and (3.A.15). When N is large and T is finite, using Lemma A.12 in Lee (2004), $\widehat{\sigma}^2(\rho,\lambda) - \sigma^{*2}(\rho,\lambda) \xrightarrow{p}$

0 uniformly in ρ and λ. As $\widehat{\sigma}^2(\rho,\lambda) - \sigma^{*2}(\rho,\lambda) \xrightarrow{p} 0$ uniformly in ρ and λ, we have $\frac{1}{N(T-1)}\ln L(\rho,\lambda) - Q(\rho,\lambda) \xrightarrow{p} 0$ uniformly in ρ and λ.

Uniform equicontinuity of $Q(\rho,\lambda)$

From (3.A.12) and (3.A.13), $Q(\rho,\lambda)$ is uniformly equicontinuous in ρ and λ due to the facts: (1) $\frac{1}{N}\ln|\boldsymbol{S}(\rho)|$ and $\frac{1}{N}\ln|\boldsymbol{R}(\lambda)|$ uniformly equicontinuous in ρ and λ; (2)$(\rho-\rho_0)^2\mathcal{H}_{\rho_0}(\lambda)$ is uniformly equicontinuous in ρ and λ; (3) $\sigma^{2\dagger}(\rho,\lambda)$ is uniformly equicontinuous in ρ and λ.

Combining global identification, uniform convergence and equicontinuity, the consistency of $(\widehat{\rho}^w, \widehat{\lambda}^w)$, which is equivalent to $(\widehat{\rho}, \widehat{\lambda})$, follows.

Consistency of other parameters

From (3.9) and our comparison of QML estimators based on those two transformations, we know that $(\widehat{\boldsymbol{\beta}}^w, \widehat{\rho}^w, \widehat{\lambda}^w)$ is numerically the same as $(\widehat{\boldsymbol{\beta}}, \widehat{\rho}, \widehat{\lambda})$. From (3.10) and (3.A.9), we also know that $\frac{T}{T-1}\widehat{\sigma}^{2w}(\widehat{\rho}^w, \widehat{\lambda}^w) - \sigma_0^2 \xrightarrow{p} 0$ and $\widehat{\sigma}^2(\widehat{\rho}^w, \widehat{\lambda}^w) - \sigma_0^2 \xrightarrow{p} 0$. Hence, $\widehat{\boldsymbol{\beta}}^w(\widehat{\rho}^w, \widehat{\lambda}^w)$ is consistent.

3.A.3 Proof of Theorem 2 (Asymptotic Distribution)

For the within transformation, according to the Taylor expansion,

$$\sqrt{NT}(\widehat{\boldsymbol{\theta}}^w - \boldsymbol{\theta}_T) = (-\frac{1}{NT}\frac{\partial^2 \ln L^w(\bar{\boldsymbol{\theta}}^w)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'})^{-1} \times (\frac{1}{\sqrt{NT}}\frac{\partial \ln L^w(\boldsymbol{\theta}_T)}{\partial\boldsymbol{\theta}}),$$

where $\bar{\boldsymbol{\theta}}^w$ lies between $\boldsymbol{\theta}_T$ and $\widehat{\boldsymbol{\theta}}^w$. As we have

$$\begin{aligned}-\frac{1}{NT}\frac{\partial^2 \ln L^w(\bar{\boldsymbol{\theta}}^w)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} = &(-\frac{1}{NT}\frac{\partial^2 \ln L^w(\bar{\boldsymbol{\theta}}^w)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} - (-\frac{1}{NT}\frac{\partial^2 \ln L^w(\boldsymbol{\theta}_T^w)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'})) \\ &+ (-\frac{1}{NT}\frac{\partial^2 \ln L^w(\boldsymbol{\theta}_T^w)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} - \boldsymbol{\Sigma}_{\boldsymbol{\theta}_T}^w) + \boldsymbol{\Sigma}_{\boldsymbol{\theta}_T}^w,\end{aligned}$$

where the first term is $||\bar{\boldsymbol{\theta}}^w - \boldsymbol{\theta}_T|| \cdot O_p(1)$ and the second term, by using Lee and Yu (2010a, Lemma A.3), is $O_p(\frac{1}{\sqrt{NT}})$, $-\frac{1}{NT}\frac{\partial^2 \ln L^w(\bar{\boldsymbol{\theta}}^w)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} = ||\bar{\boldsymbol{\theta}}^w - \boldsymbol{\theta}_T|| \cdot O_p(1) + O_p(\frac{1}{\sqrt{NT}}) + \boldsymbol{\Sigma}_{\boldsymbol{\theta}_T}^w$.

Under Assumptions 3.7 and 3.8, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_T}^w$ in (3.A.4) is non-singular. We can prove the non-singularity of the limiting information matrix by using an argument by

contradiction. We need to prove that $\boldsymbol{\Sigma}^{w}_{\boldsymbol{\theta}_T}\boldsymbol{c} = 0$ implies that $\boldsymbol{c} = 0$, where $\boldsymbol{c} = (\boldsymbol{c}_1^{'}, c_2, c_3, c_4)^{'}$, c_2, c_3, c_4 are scalars and $\boldsymbol{c}_1$ is $K \times 1$ vector.

With $\boldsymbol{C} = \ddot{\boldsymbol{G}} - \frac{\text{tr}\ddot{\boldsymbol{G}}}{N}\boldsymbol{I}_N$ and $\boldsymbol{D} = \boldsymbol{K} - \frac{\text{tr}\boldsymbol{K}}{N}\boldsymbol{I}_N$, we have $\frac{1}{N}\text{tr}(\ddot{\boldsymbol{G}}^s\ddot{\boldsymbol{G}}) - 2(\frac{\text{tr}\ddot{\boldsymbol{G}}}{N})^2 = \frac{1}{2N}\text{tr}(\boldsymbol{C}^s\boldsymbol{C}^s)$, $\frac{1}{N}\text{tr}(\boldsymbol{K}^s\boldsymbol{K}) - 2(\frac{\text{tr}\boldsymbol{K}}{N})^2 = \frac{1}{2N}\text{tr}(\boldsymbol{D}^s\boldsymbol{D}^s)$, and $\frac{1}{N}\text{tr}(\boldsymbol{K}^s\ddot{\boldsymbol{G}}) - 2(\frac{\text{tr}\boldsymbol{K}}{N}\frac{\text{tr}\ddot{\boldsymbol{G}}}{N}) = \frac{1}{2N}\text{tr}(\boldsymbol{C}^s\boldsymbol{D}^s)$. Furthermore, we have $\mathcal{H}_{\boldsymbol{\beta}} = \frac{1}{NT}\sum_{t=1}^{T}\ddot{\boldsymbol{X}}_t^{'}\ddot{\boldsymbol{X}}_t$, $\mathcal{H}_{\boldsymbol{\beta}\rho} = \frac{1}{NT}\sum_{t=1}^{T}\ddot{\boldsymbol{X}}_t^{'}\ddot{\boldsymbol{G}}\ddot{\boldsymbol{X}}_t\boldsymbol{\beta}_0$, $\mathcal{H}_{\rho} = \frac{1}{NT}\sum_{t=1}^{T}(\ddot{\boldsymbol{G}}\ddot{\boldsymbol{X}}_t\boldsymbol{\beta}_0)^{'}\ddot{\boldsymbol{G}}\ddot{\boldsymbol{X}}_t\boldsymbol{\beta}_0$, and $\mathcal{H}_{\rho\boldsymbol{\beta}} = \mathcal{H}_{\boldsymbol{\beta}\rho}^{'}$. The detailed proof of $\boldsymbol{\Sigma}^{w}_{\boldsymbol{\theta}_T}\boldsymbol{c} = 0$ is as follows:

$$
\begin{aligned}
\frac{1}{\sigma_T^2}\mathcal{H}_{\boldsymbol{\beta}}\boldsymbol{c}_1 + \frac{1}{\sigma_T^2}\mathcal{H}_{\boldsymbol{\beta}\rho}c_2 &= 0,\\
\frac{1}{\sigma_T^2}\mathcal{H}_{\rho\boldsymbol{\beta}}\boldsymbol{c}_1 + (\frac{1}{\sigma_T^2}\mathcal{H}_{\rho} + \frac{1}{n}\text{tr}\ddot{\boldsymbol{G}}^s\ddot{\boldsymbol{G}})c_2 - \frac{1}{N}\text{tr}(\boldsymbol{K}^s\ddot{\boldsymbol{G}})c_3 + \frac{1}{\sigma_T^2 N}\text{tr}\ddot{\boldsymbol{G}}c_4 &= 0,\\
-\frac{1}{N}\text{tr}(\boldsymbol{K}^s\ddot{\boldsymbol{G}})c_2 + \frac{1}{N}\text{tr}(\boldsymbol{K}^s\boldsymbol{K})c_3 - \frac{1}{\sigma_T^2 N}\text{tr}\boldsymbol{K}c_4 &= 0,\\
\frac{1}{\sigma_T^2 N}\text{tr}\ddot{\boldsymbol{G}}c_2 - \frac{1}{\sigma_T^2 N}\text{tr}\boldsymbol{K}c_3 + \frac{1}{2\sigma_T^4}c_4 &= 0.
\end{aligned}
$$

By solving the third and fourth equations, we get

$$
\begin{aligned}
c_3 &= \frac{\text{tr}(\boldsymbol{C}^s\boldsymbol{D}^s)}{\text{tr}(\boldsymbol{D}^s\boldsymbol{D}^s)}c_2,\\
\frac{1}{\sigma_T^2 N}c_4 &= \frac{2}{N^2}\frac{\text{tr}\boldsymbol{K}\text{tr}(\boldsymbol{C}^s\boldsymbol{D}^s)}{\text{tr}(\boldsymbol{D}^s\boldsymbol{D}^s)}c_2 - \frac{2}{N^2}\text{tr}\ddot{\boldsymbol{G}}c_2,
\end{aligned}
$$

as a result of which $\boldsymbol{c}_1$, c_3, and c_4 can all be expressed in terms of c_2. Substituting these results in the second equation yields

$$
\frac{1}{\sigma_T^2}(\mathcal{H}_{\rho} - \mathcal{H}_{\rho\boldsymbol{\theta}}(\mathcal{H}_{\boldsymbol{\theta}})^{-1}\mathcal{H}_{\boldsymbol{\theta}\rho})c_2 + \frac{1}{2N}\text{tr}(\boldsymbol{C}^s\boldsymbol{C}^s)c_2 - \frac{1}{2N}\frac{\text{tr}^2(\boldsymbol{C}^s\boldsymbol{D}^s)}{\text{tr}(\boldsymbol{D}^s\boldsymbol{D}^s)}c_2 = 0,
$$

as a result of which $\boldsymbol{\Sigma}^{w}_{\boldsymbol{\theta}_T}\boldsymbol{c} = 0$ implies that

$$
\{\lim(\frac{1}{\sigma_T^2}\frac{1}{N}\text{tr}(\boldsymbol{D}^s\boldsymbol{D}^s)(\mathcal{H}_{\rho} - \mathcal{H}_{\rho\boldsymbol{\beta}}(\mathcal{H}_{\boldsymbol{\beta}})^{-1}\mathcal{H}_{\boldsymbol{\beta}\rho}) + \boldsymbol{\Phi})\},
$$

where $\boldsymbol{\Phi} = \frac{1}{2N^2}[\text{tr}(\boldsymbol{C}^s\boldsymbol{C}^s)\text{tr}(\boldsymbol{D}^s\boldsymbol{D}^s) - \text{tr}^2(\boldsymbol{C}^s\boldsymbol{D}^s)]$ and $\mathcal{H}_{\rho} - \mathcal{H}_{\rho\boldsymbol{\beta}}(\mathcal{H}_{\boldsymbol{\beta}})^{-1}\mathcal{H}_{\boldsymbol{\beta}\rho}$ are non-negative by the Cauchy-Schwarz inequality. Hence, the non-singularity of $\lim\boldsymbol{\Sigma}^{w}_{\boldsymbol{\theta}_T}$ follows from Assumption 3.7.

Since the score of the log likelihood function evaluated at $\boldsymbol{\theta}_0$ is a linear and

quadratic form of $\widetilde{v}_t$, which has zero mean and $\mathrm{E}\widetilde{v}_t^{'}\widetilde{v}_t = \frac{T-1}{T}N\sigma_0^2 = N\sigma_T^2$, and $\ddot{\boldsymbol{X}}_t$ is uncorrelated with $\widetilde{v}_t$, using Lee and Yu (2010a, Lemma A.4), we have

$$\begin{aligned}\mathrm{E}\Big(\frac{1}{\sqrt{NT}}\frac{\partial \ln L^w(\boldsymbol{\theta}_T)}{\partial \boldsymbol{\theta}} \cdot \frac{1}{\sqrt{NT}}\frac{\partial \ln L^w(\boldsymbol{\theta}_T)}{\partial \boldsymbol{\theta}'}\Big) &= \mathrm{cov}\Big(\frac{1}{\sqrt{NT}}\frac{\partial \ln L^w(\boldsymbol{\theta}_T)}{\partial \boldsymbol{\theta}}, \frac{1}{\sqrt{NT}}\frac{\partial \ln L^w(\boldsymbol{\theta}_T)}{\partial \boldsymbol{\theta}'}\Big), \\ &= \frac{T-1}{T}(\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T} + \boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T}),\end{aligned}$$

where $\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T}$ is given in (3.A.4) and $\boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T}$ is given in (3.A.5). When $\boldsymbol{v}_t$ are normally distributed, $\boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T} = \mathbf{0}_{(K+3)\times(K+3)}$ because $\mu_4 - 3\sigma_0^4 = 0$. By using the central limit theorem in Lee and Yu (2010a, Lemma A.1), $\frac{1}{\sqrt{NT}}\frac{\partial \ln L^w(\boldsymbol{\theta}_T)}{\partial \boldsymbol{\theta}} \xrightarrow{w} \mathcal{N}(0, \lim\frac{T-1}{T}(\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T} + \boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T}))$.

As $||\bar{\boldsymbol{\theta}}^w - \boldsymbol{\theta}_T|| = o_p(1)$ and $\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T}$ is non-singular in the limit, $(-\frac{1}{NT}\frac{\partial^2 \ln L^w(\bar{\boldsymbol{\theta}}^w)}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'})^{-1}$ is $O_p(1)$. It follows that $\widehat{\boldsymbol{\theta}}^w - \theta_T = O_p(\frac{1}{NT})$ and thus

$$\sqrt{NT}(\widehat{\boldsymbol{\theta}}^w - \boldsymbol{\theta}_T) = (\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T} + O_p(\frac{1}{\sqrt{NT}}))^{-1} \times (\frac{1}{\sqrt{NT}}\frac{\partial \ln L^w(\boldsymbol{\theta}_T)}{\partial \boldsymbol{\theta}}).$$

Using the fact that

$$(\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T} + O_p(\frac{1}{\sqrt{NT}}))^{-1} = (\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T})^{-1} + O_p(\frac{1}{\sqrt{NT}}),$$

we eventually have

$$\sqrt{NT}(\widehat{\boldsymbol{\theta}}^w - \boldsymbol{\theta}_T) \xrightarrow{w} \mathcal{N}(0, \lim\frac{T-1}{T}(\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T})^{-1}(\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T} + \boldsymbol{\Omega}^w_{\boldsymbol{\theta}_T})(\boldsymbol{\Sigma}^w_{\boldsymbol{\theta}_T})^{-1}).$$

Chapter 4

General nesting spatial panel models with spatial autoregressive or moving average errors and parameterized distance-based decay matrices

4.1 Introduction

Spatial econometric models study whether and to which extent cross-sectional units affect each other. The kind of interactions that are considered can take the form of endogenous spatial interactions in the dependent variable ($\boldsymbol{W}\boldsymbol{y}_t$), exogenous interactions in the regressors ($\boldsymbol{W}\boldsymbol{X}_t$), and interaction in the errors, which in turn can take the form of a spatial autoregressive (AR) structure, given by $\boldsymbol{u}_t = \lambda \boldsymbol{M}\boldsymbol{u}_t + \boldsymbol{v}_t$, or a moving average (MA) structure, given by $\boldsymbol{u}_t = \boldsymbol{v}_t - \lambda \boldsymbol{M}\boldsymbol{v}_t$. Although there exists a variety of spatial econometric models that contain one or multiple forms of spatial interactions (for an overview see Elhorst, 2014), the general nesting spatial (GNS) model that includes all three forms of spatial interactions, either with AR or MA errors, hardly received any attention. For example, the GNS model with AR errors is mentioned on page 53 in the introductory spatial econometric textbook of LeSage and Pace (2009), but it is not taken seriously given the fact that the specification of this model in contrast to other models is not numbered. The GNS model with AR errors is more extensively discussed in the spatial econometric textbook of Kelejian and Piras (2017). In Chapter 2 they briefly set out the instrumental variables (IV) and the maximum likelihood (ML) estimators of this model in a cross-sectional setting, specify the assumptions to be made for these estimators to be consistent and asymptotically normal, and take up a proof for this. However, in spite of many empirical illustrations, none of them provides the estimation results of this model. This lack of applications also characterizes the empirical literature. Lin (2015) adopts the GNS model to investigate peer effects among adolescents, though based on cross-sectional data only. Li and Li (2020) depart from a GNS model with fixed effects in space and time to investigate energy investment, economic growth, and carbon emissions, but eventually they only estimate the simpler spatial Durbin (SD) model. Huang et al. (2021) initially estimate a GNS model with fixed effects in space and time to explain green innovations for sustainable development in China. However, because the simpler SD model is not inferior in performance, they interpret all results based on this model. Further below we also discuss studies which have developed estimators for GNS models.

One major barrier hindering the wide application of GNS models is the identification problem, the issue that the coefficients of various spatial interactions can-

not be separately identified from the reduced form of the model. This identification problem is discussed in several studies. Manski (1993) demonstrates that peer effects and contextual effects in linear-in-means models, known as spatial lags in the dependent and independent variables in a spatial econometric model, cannot be distinguished from each other if the network or spatial weight matrix is specified as a group interaction matrix in which the diagonal elements just as the off-diagonal elements are specified as $1/N_g$, where N_g represents the number of people in each group g. This identification problem is known as the reflection problem. Bramoullé et al. (2009) show that identification fails when the linear-in-means model, the counterpart of a spatial Durbin model in spatial econometrics, simplifies to a model with correlated effects only, the counterpart of a spatial error model in spatial econometrics. Gibbons and Overman (2012, Section 3) argue that a weak instrument problem may occur when a linear-in-means model with peer effects and contextual effects is estimated by instrumental variables based on second and higher-order contextual effects. Although Halleck Vega and Elhorst (2015) show that these problems can be reduced by setting the diagonal elements of spatial weight matrix to zero, based on the argument that a unit cannot affect itself, by parameterizing the spatial weight matrix, and by estimating the model by maximum likelihood, they find that the GNS model might still suffer from identification problems because the spatial autoregressive parameter of the spatial lag in the dependent variable and the spatial autocorrelation coefficient of the spatial lag in the error term are easily interchanged. To break the curse of identification in a GNS model, Lee, Liu, and Lin (2010) assume a block diagonal network matrix for social interaction models, where each block represents a group of units that interact with each other but not with itself and not with units in other groups. However, just as all the other cited studies, this study departs from one uniform spatial weight matrix for all spatial lags in the model. In this chapter, we propose a distance-based parameterized spatial weight matrix that is different for each spatial lag in the model.

The spatial weight matrix is first specified as a functional form of distance with a distance decay parameter and then the corresponding distance decay parameter of each individual spatial lag is assumed to be different and estimated jointly with the response parameters of the model. The distance decay concept has been well applied in different fields of economics such as regional science, spatial econom-

ics, urban economics, and economic geography. Fotheringham (1981) describes the distance decay parameter as a measure of the relationship between distance and interaction. Using a simple parametric approach, such as the negative exponential form $e^{-d_{ij}\varrho_0}$, where d_{ij} is the geographical distance between units i and j, the degree of interaction between units of observations diminishes with the distance decay parameter ϱ_0. The advantage of allowing the distance decay parameters to be different for different spatial lags is that the problem of weak identification further diminishes. Generally, one pre-specified spatial weight matrix is used in empirical studies for all spatial lags in the model, while in this study different distance decay parameters are estimated. Consequently, more information is extracted from the data, as the distance decay parameters of the spatial lags of the regressors are likely to be different from one common distance decay parameter for all of them.

While spatial AR errors dominate the spatial econometric literature focusing on the GNS model, the errors might also take the form of spatial MA process. An overview of previous studies in this field is provided in Table 3.1 of Chapter 3. Fingleton and LeGallo (2007) distinguish AR and MA errors from each other in terms of the shock diffusion process among spatial units. Spatial AR errors represent global shocks since a shock that occurs in one unit not only spreads to nearby units to which this unit is connected, but due to its multiple rounds structure also to all other units farther away. Conversely, spatial MA errors reflect local shocks, as a shock in one unit only spreads to nearby units in one round and cannot spread further. Unlike GNSAR models, GNS models with the MA errors (GNSMA) have, to our knowledge, not been discussed.

Three types of estimators have been developed to estimate GNSAR models for cross-sectional data or spatial panels, with and without fixed effects in space and time: maximum likelihood (ML) (Burridge et al., 2016), quasi-maximum likelihood (QML) (Lee et al., 2010), two-stage-least-squares (2SLS) or IV (Lee and Yu, 2016; Kelejian and Piras, 2017), and Bayesian estimators (Hassan, 2017). For spatial econometric models in a panel data setting, the inclusion of fixed effects, for both spatial units and time periods, is important. Individual fixed effects are incorporated to control for time-invariant unobserved individual characteristics which can be correlated with the existing regressors[1], while time fixed effects control for unobserved

[1] Instead of fixed effects, it also possible to consider random effects (Millo, 2022).

characteristics varying from one time period to another. Lee and Yu (2010a) point out that time fixed effects might be important for space-time researches such as economic growth, convergence, and regional economics to capture business cycle effects typical for the time-series and macroeconomic literature. Neglecting fixed effects may thence bias the estimation results (Elhorst, 2003). However, the increased number of parameters due to the inclusion of fixed effects may cause an incidental parameter problem, as discussed in Neyman and Scott (1948). When spatial econometric models contain both fixed effects in space and time, two panel data settings may be distinguished: large N finite T and large N large T. In the former case, time fixed effects do not have to be included or if included they can be treated as normal regressors. In the latter case, the inclusion of time effects may cause an incidental parameter problem when T goes to infinity in addition to the incidental parameter problem caused by the individual fixed effects when N goes to infinity.

To estimate the spatial panel models with fixed effects, Lee and Yu (2010a) propose a QML estimator which uses an orthogonal transformation to concentrate out both fixed effects in the spatial autoregressive model with AR errors (SARAR model). Later on Lee and Yu (2012) apply this QML orthogonal estimator to estimate a general spatial econometric model which nests the spatial autoregressive model with MA errors (SARMA model) with both fixed effects. These QML estimators are consistent and properly centered requiring no additional bias correction procedures. In this study, these QML estimators are modified further to estimate panel GNS models with AR or MA errors by adding Durbin terms with parameterized spatial weight matrices to the SARAR and SARMA models.

This chapter contributes to the literature in the following ways. First, we are among the first to consider GNS panel data models with AR or MA errors, individual fixed effects and time fixed effects in which the spatial weight matrices are parameterized. Although GNS models are characterized by identification problems, these problems are reduced by parameterizing the spatial weight matrices with a different distance decay matrix for each lag in the model. Second, we derive QML estimators for these kind of models. Third, as an extension of the work of Lee et al. (2010) in which the social interaction model with group interaction matrices and individual fixed effects is estimated, we first show that identification problems

in the estimation of the GNS models are the same when either spatial AR or MA errors are used. Then we introduce the rank conditions for the identification when the spatial weight matrices are parameterized for the GNS models with these two different types of errors.

This chapter is organized as follows. In Section 4.2 we specify the GNSAR and GNSMA panel models, introduce the functional form of the parameterized spatial weight matrices, set out the corresponding QML estimator based on the orthogonal transformation to concentrate out the fixed effects, and show the rank conditions for identification. In Section 4.3 we conduct a Monte Carlo experiment to explore the finite sample properties of the proposed QML estimators. In Section 4.4 we illustrate the benefits of the GNSAR and GNSMA models by an empirical analysis using US house price data of Yang (2021). In the last section of this chapter, we draw conclusions.

4.2 Methodology

4.2.1 The GNSAR and GNSMA models

The GNSAR panel model with parameterized spatial weight matrices and individual and time effects consists of a combination of the spatial Durbin (SD) model in (4.1) and a spatial AR error process in (4.2),

$$\boldsymbol{y}_t = \rho_0 \boldsymbol{W}(\delta_0)\boldsymbol{y}_t + \boldsymbol{X}_t\boldsymbol{\beta}_0 + \sum_{k=1}^{K} \boldsymbol{W}(\alpha_{0k})\boldsymbol{x}_{kt}\phi_{0k} + \boldsymbol{c}_0 + \varpi_{0t}\boldsymbol{\iota}_N + \boldsymbol{u}_t, \tag{4.1}$$

$$\boldsymbol{u}_t = \lambda_0 \boldsymbol{W}(\gamma_0)\boldsymbol{u}_t + \boldsymbol{v}_t,\ t = 1, 2, ...T. \tag{4.2}$$

The spatial MA error process is given by

$$\boldsymbol{u}_t = \boldsymbol{v}_t - \lambda_0 \boldsymbol{W}(\gamma_0)\boldsymbol{v}_t,\ t = 1, 2, ...T. \tag{4.3}$$

The SD model in (4.1) combined with the spatial MA error process in (4.3) is the GNSMA panel model with parameterized spatial weight matrices and both fixed effects. In both models, ϖ_{0t} represents a time fixed effect and $\boldsymbol{c}_0$ an $N \times 1$ vector of individual fixed effects. $\boldsymbol{\iota}_N$ is an $N \times 1$ vector of ones. The dependent variable $\boldsymbol{y}_t = (y_{1t}, y_{2t}, ..., y_{Nt})'$ and the disturbances $\boldsymbol{v}_t = (v_{1t}, v_{2t}, ..., v_{Nt})'$ are $N \times 1$ vectors,

where v_{it} is *i.i.d* across i and t with zero mean and variance σ_0^2. $\boldsymbol{X}_t$ is an $N \times K$ matrix of non-stochastic time varying regressors whose kth column is denoted by $\boldsymbol{x}_{kt}$. $\boldsymbol{W}(\delta_0)$, $\boldsymbol{W}(\alpha_{0k})$ $(k = 1, ..., K)$, and $\boldsymbol{W}(\gamma_0)$ are $N \times N$ parameterized spatial weights matrices with zero diagonals, which capture spatial dependence in respectively the dependent variable, the regressors, and the error terms among the cross sectional units at each moment in time. δ_0, α_{0k}, and γ_0 are the corresponding distance decay parameters. ρ_0 is the spatial autoregressive parameter for the dependent variable and λ_0 is the spatial autocorrelation parameter for the spatial AR or MA error. For any parameter, the subscript 0 denotes that parameter evaluated at its true value.

The typical element of the parameterized spatial weight matrix, before it is row-normalized, is $e^{-d_{ij}\varrho_0}$, where d_{ij} is the distance between unit i and j and ϱ_0 is the distance decay parameter. This matrix has three important properties. First, the degree of interaction between two units diminishes if the distance decay parameter increases. The elements converge to 1 when $\varrho_0 \to 0$ and to 0 when $\varrho_0 \to \infty$. Second, by allowing the distance decay parameters to be different, the spatial lag of each regressor will have its own spatial weight matrix. Third, even though the off-diagonal elements of the exponential distance decay matrix are nonzero, this matrix satisfies the boundedness condition. The boundedness condition requires that both row and column sums before the spatial weight matrix is row-normalized is upper bounded when N goes to infinity, which it is as long as the distance decay parameter ϱ_0 is positive (see Chapter 2). Row normalization of the spatial weight matrices is popular in empirical studies on spatial interactions. Row normalization is required to concentrate out time fixed effects by the orthogonal transformation.

4.2.2 QML estimation of the GNSAR and GNSMA models

We take the estimation of the GNSAR model as point of departure since this model receives more attention in the literature and the method of concentrating out the fixed effects is basically the same for both the GNSAR and GNSMA models. The main difference is that the corresponding log-likelihood functions and notations are different. To estimate the GNSAR model with both fixed effects, we modify the QML estimator based on the orthogonal transformation set out in Lee and Yu (2010a). To concentrate out the individual fixed effects as a first step, the matrix $\boldsymbol{F}_{T,T-1}$ is used, where $[\boldsymbol{F}_{T,T-1}, \frac{1}{\sqrt{T}}\boldsymbol{\iota}_T]$ is the orthonormal eigenvector matrix of the

time mean operator $\boldsymbol{J}_T = \boldsymbol{I}_T - \frac{1}{T}\boldsymbol{\iota}_T\boldsymbol{\iota}_T^{'}$. After reshaping the dependent variable of dimension $NT \times 1$ into the matrix $[\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_T]$ of dimension $N \times T$, it can be post-multiplied by $\boldsymbol{F}_{T,T-1}$, to get $[\boldsymbol{y}_1^{\bullet}, \boldsymbol{y}_2^{\bullet}, ..., \boldsymbol{y}_{T-1}^{\bullet}] = [\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_T]\boldsymbol{F}_{T,T-1}$. Similar transformations are used to obtain $\boldsymbol{X}_t^{*\bullet}$ and $\boldsymbol{v}_t^{\bullet}$, where $\boldsymbol{X}_t^* = (\boldsymbol{x}_{1t}, ..., \boldsymbol{x}_{Kt}, \boldsymbol{W}(\alpha_{0,1})\boldsymbol{x}_{1t}, ..., \boldsymbol{W}(\alpha_{0K})\boldsymbol{x}_{Kt})$ before it is post-multiplied by $\boldsymbol{F}_{T,T-1}$. Due to this transformation, the GNSAR model given in (4.1) and (4.2) can be re-written as

$$\begin{aligned}
\boldsymbol{y}_t^{\bullet} &= \rho_0 \boldsymbol{W}(\delta_0)\boldsymbol{y}_t^{\bullet} + \boldsymbol{X}_t^{*\bullet}\boldsymbol{\zeta}_0 + \omega_{0t}^{\bullet}\boldsymbol{\iota}_N + \boldsymbol{u}_t^{\bullet}, \\
\boldsymbol{u}_t^{\bullet} &= \lambda_0 \boldsymbol{W}(\gamma_0)\boldsymbol{u}_t^{\bullet} + \boldsymbol{v}_t^{\bullet}, \quad t = 1, 2, ...T-1,
\end{aligned} \tag{4.4}$$

where $\boldsymbol{\zeta}_0 = (\boldsymbol{\beta}_0^{'}, \boldsymbol{\phi}_0^{'})^{'}$ and the actual sample size has become $N(T-1)$ instead of NT.

In the second step, the time fixed effects are concentrated out by the matrix $\boldsymbol{F}_{N,N-1}$, where $[\boldsymbol{F}_{N,N-1}, \frac{1}{\sqrt{N}}\boldsymbol{\iota}_N]$ is the orthonormal eigenvector matrix of the cross-sectional mean operator $\boldsymbol{J}_N = \boldsymbol{I}_N - \frac{1}{N}\boldsymbol{\iota}_N\boldsymbol{\iota}_N^{'}$. After concentrating out both fixed effects using the orthogonal transformation, (4.4) can be re-written as

$$\begin{aligned}
\boldsymbol{y}_t^{\bullet\bullet} &= \rho_0 (\boldsymbol{F}_{N,N-1}^{'}\boldsymbol{W}(\delta_0)\boldsymbol{F}_{N,N-1})\boldsymbol{y}_t^{\bullet\bullet} + \boldsymbol{X}_t^{*\bullet\bullet}\boldsymbol{\zeta}_0 + \boldsymbol{u}_t^{\bullet\bullet}, \\
\boldsymbol{u}_t^{\bullet\bullet} &= \lambda_0 (\boldsymbol{F}_{N,N-1}^{'}\boldsymbol{W}(\gamma_0)\boldsymbol{F}_{N,N-1})\boldsymbol{u}_t^{\bullet\bullet} + \boldsymbol{v}_t^{\bullet\bullet}, \quad t = 1, 2, ...T-1,
\end{aligned} \tag{4.5}$$

where $\boldsymbol{y}_t^{\bullet\bullet} = \boldsymbol{F}_{N,N-1}^{'}\boldsymbol{y}_t^{\bullet}$, $\boldsymbol{X}_t^{*\bullet\bullet} = \boldsymbol{F}_{N,N-1}^{'}\boldsymbol{X}_t^{*\bullet}$, and $\boldsymbol{v}_t^{\bullet\bullet} = \boldsymbol{F}_{N,N-1}^{'}\boldsymbol{v}_t^{\bullet}$. After this second-step transformation, the actual sample size decreases to $(N-1)(T-1)$. As we also have

$$\begin{aligned}
&(\boldsymbol{v}_1^{\bullet\bullet'}, ..., \boldsymbol{v}_{T-1}^{\bullet\bullet'})^{'} = (\boldsymbol{I}_{T-1} \otimes \boldsymbol{F}_{N,N-1}^{'})(\boldsymbol{v}_1^{\bullet'}, ..., \boldsymbol{v}_{T-1}^{\bullet'})^{'} = (\boldsymbol{F}_{T,T-1}^{'} \otimes \boldsymbol{F}_{N,N-1}^{'})(\boldsymbol{v}_1^{'}, ..., \boldsymbol{v}_T^{'})^{'}, \\
&\mathrm{E}(\boldsymbol{v}_1^{\bullet\bullet'}, ..., \boldsymbol{v}_{T-1}^{\bullet\bullet'})(\boldsymbol{v}_1^{\bullet\bullet'}, ..., \boldsymbol{v}_{T-1}^{\bullet\bullet'})^{'} = \sigma_0^2(\boldsymbol{F}_{T,T-1}^{'} \otimes \boldsymbol{F}_{N,N-1}^{'})(\boldsymbol{F}_{T,T-1} \otimes \boldsymbol{F}_{N,N-1}) \\
&= \sigma_0^2(\boldsymbol{I}_{T-1} \otimes \boldsymbol{I}_{N-1}),
\end{aligned}$$

the log-likelihood function of (4.5) reads as

$$\begin{aligned}
\ln L(\boldsymbol{\theta}) = &- \frac{(N-1)(T-1)}{2}\ln(2\pi\sigma^2) + (T-1)\ln|\boldsymbol{I}_{N-1} - \rho\boldsymbol{F}_{N,N-1}^{'}\boldsymbol{W}(\delta)\boldsymbol{F}_{N,N-1}| \\
&+ (T-1)\ln|\boldsymbol{I}_{N-1} - \lambda\boldsymbol{F}_{N,N-1}^{'}\boldsymbol{W}(\gamma)\boldsymbol{F}_{N,N-1}| - \frac{1}{2\sigma^2}\sum_{t=1}^{T-1}\boldsymbol{v}_t^{\bullet\bullet'}(\xi)\boldsymbol{v}_t^{\bullet\bullet}(\xi).
\end{aligned} \tag{4.6}$$

Using Lemma A.2 in Lee and Yu (2010a), which requires the spatial weight matrices to be row-normalized, the corresponding log determinant and inverse of $\boldsymbol{I}_{N-1} - \rho \boldsymbol{F}_{N,N-1}^{'} \boldsymbol{W}(\delta) \boldsymbol{F}_{N,N-1}$ are

$$|\boldsymbol{I}_{N-1} - \rho \boldsymbol{F}_{N,N-1}^{'} \boldsymbol{W}(\delta) \boldsymbol{F}_{N,N-1}| = \frac{1}{1-\rho} |\boldsymbol{I}_{N-1} - \rho \boldsymbol{W}(\delta)|,$$
$$(\boldsymbol{I}_{N-1} - \rho \boldsymbol{F}_{N,N-1}^{'} \boldsymbol{W}(\delta) \boldsymbol{F}_{N,N-1})^{-1} = \boldsymbol{F}_{N,N-1}^{'} (\boldsymbol{I}_N - \rho \boldsymbol{W}(\delta))^{-1} \boldsymbol{F}_{N,N-1}.$$

The same holds for $\boldsymbol{I}_{N-1} - \lambda \boldsymbol{F}_{N,N-1}^{'} \boldsymbol{W}(\gamma) \boldsymbol{F}_{N,N-1}$. For every two N-dimensional column vectors $\boldsymbol{p}_t$ and $\boldsymbol{q}_t$ we further have

$$\sum_{t=1}^{T-1} \boldsymbol{p}_t^{\bullet\bullet'} \boldsymbol{q}_t^{\bullet\bullet} = (\boldsymbol{p}_1^{'}, ..., \boldsymbol{p}_T^{'})(\boldsymbol{F}_{T,T-1} \otimes \boldsymbol{F}_{N,N-1}) \times (\boldsymbol{F}_{T,T-1}^{'} \otimes \boldsymbol{F}_{N,N-1}^{'})(\boldsymbol{q}_1^{'}, ..., \boldsymbol{q}_T^{'})^{'}$$
$$= (\boldsymbol{p}_1^{'}, ..., \boldsymbol{p}_T^{'})(\boldsymbol{J}_T \otimes \boldsymbol{J}_N)(\boldsymbol{q}_1^{'}, ..., \boldsymbol{q}_T^{'})^{'} = \sum_{t=1}^{T} \widetilde{\boldsymbol{p}}_t^{'} \boldsymbol{J}_N \widetilde{\boldsymbol{q}}_t,$$

where the element of $\widetilde{\boldsymbol{p}}_t$ is $\widetilde{p}_{it} = p_{it} - \frac{1}{T} \sum_{t=1}^{T} p_{it}$ for $t = 1, 2, ..., T$ and $\widetilde{q}_{it} = q_{it} - \frac{1}{T} \sum_{t=1}^{T} q_{it}$ is the element of $\widetilde{\boldsymbol{q}}_t$ for $t = 1, 2, ..., T$. Consequently, the corresponding log-likelihood of the GNSAR model after both fixed effects have been concentrated out is

$$\ln L(\boldsymbol{\theta}) = -\frac{(N-1)(T-1)}{2} \ln(2\pi\sigma^2) - (T-1)[\ln(1-\rho) + \ln(1-\lambda)]$$
$$+ (T-1)[\ln|\boldsymbol{S}(\rho, \delta)| + \ln|\boldsymbol{R}(\lambda, \gamma)|] - \frac{1}{2\sigma^2} \sum_{t=1}^{T} \widetilde{\boldsymbol{v}}_t^{'}(\boldsymbol{\xi}) \boldsymbol{J}_N \widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}), \qquad (4.7)$$

where $\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) = \boldsymbol{R}(\lambda, \gamma)[\boldsymbol{S}(\rho, \delta) \widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t^{*'} \boldsymbol{\zeta}]$ using the following notations: $\boldsymbol{S}(\rho, \delta) = \boldsymbol{I}_N - \rho \boldsymbol{W}(\delta)$, $\boldsymbol{R}(\lambda, \gamma) = \boldsymbol{I}_N - \lambda \boldsymbol{W}(\gamma)$, $\widetilde{\boldsymbol{y}}_t$ contains the element $\widetilde{y}_{it} = y_{it} - \frac{1}{T} \sum_{t=1}^{T} y_{it}$, $\widetilde{\boldsymbol{X}}_t^{*}$ has the element $\widetilde{x}_{kit}^{*} = x_{kit}^{*} - \frac{1}{T} \sum_{t=1}^{T} x_{kit}^{*}$, $\widetilde{\boldsymbol{v}}_t$ has the element $\widetilde{v}_{it} = v_{it} - \frac{1}{T} \sum_{t=1}^{T} v_{it}$, $\boldsymbol{\theta} = (\boldsymbol{\zeta}^{'}, \rho, \lambda, \sigma^2, \delta, \boldsymbol{\alpha}^{'}, \gamma)^{'}$, and $\boldsymbol{\xi} = (\boldsymbol{\zeta}^{'}, \rho, \lambda, \delta, \boldsymbol{\alpha}^{'}, \gamma)^{'}$. $\boldsymbol{\alpha}^{'} = (\alpha_1, \alpha_2, ..., \alpha_K)$. The corresponding ML estimates for the transformed GNSAR model are

$$\widehat{\boldsymbol{\zeta}} = \left[\sum_{t=1}^{T} \widetilde{\boldsymbol{X}^{*}}_t^{'} \boldsymbol{R}^{'}(\lambda, \gamma) \boldsymbol{J}_N \boldsymbol{R}(\lambda, \gamma) \widetilde{\boldsymbol{X}}_t^{*} \right]^{'} \left[\sum_{t=1}^{T} \widetilde{\boldsymbol{X}}_t^{*'} \boldsymbol{R}^{'}(\lambda, \gamma) \boldsymbol{J}_N \boldsymbol{R}(\lambda, \gamma) \boldsymbol{S}(\rho, \delta) \widetilde{\boldsymbol{y}}_t \right], \qquad (4.8)$$

$$\widehat{\sigma}^2 = \frac{1}{(N-1)(T-1)} \sum_{t=1}^{T} [\boldsymbol{S}(\rho,\delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}^*}'_t\boldsymbol{\zeta}]' \boldsymbol{R}'(\lambda,\gamma)\boldsymbol{J}_N\boldsymbol{R}(\lambda,\gamma)[\boldsymbol{S}(\rho,\delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}^{*'}_t\boldsymbol{\zeta}]. \tag{4.9}$$

Turning to the GNSMA model, the log-likelihood after both fixed effects have been concentrated out is

$$\begin{aligned} \ln L(\boldsymbol{\theta}) = &- \frac{(N-1)(T-1)}{2}\ln(2\pi\sigma^2) - (T-1)[\ln(1-\rho) - \ln(1-\lambda)] \\ &+ (T-1)[\ln|\boldsymbol{S}(\rho,\delta)| - \ln|\boldsymbol{R}(\lambda,\gamma)|] - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}'_t(\boldsymbol{\xi})\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}), \end{aligned} \tag{4.10}$$

where $\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) = \boldsymbol{R}^{-1}(\lambda,\gamma)[\boldsymbol{S}(\rho,\delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}^*_t\boldsymbol{\zeta}]$. The corresponding ML estimates for the transformed GNSMA model are

$$\widehat{\boldsymbol{\zeta}} = \left[\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}^{*'}_t\boldsymbol{R}^{-1'}(\lambda,\gamma)\boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)\widetilde{\boldsymbol{X}}^*_t\right]' \left[\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}^{*'}_t\boldsymbol{R}^{-1'}(\lambda,\gamma)\boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{S}(\rho,\delta)\widetilde{\boldsymbol{y}}_t\right], \tag{4.11}$$

$$\widehat{\sigma}^2 = \frac{1}{(N-1)(T-1)}\sum_{t=1}^{T}[\boldsymbol{S}(\rho,\delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}^*_t\boldsymbol{\zeta}]'\boldsymbol{R}^{-1'}(\lambda,\gamma)\boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)[\boldsymbol{S}(\rho,\delta)\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}^*_t\boldsymbol{\zeta}]. \tag{4.12}$$

To analyse the asymptotic properties and identification of the proposed estimators, the following assumptions are needed. These assumptions are extended based on those in Lee and Yu (2010a) and Chapters 2 and 3 of this thesis. It concerns assumptions for the SARAR model with non-parameterized spatial weight matrices, the SD model with parameterized spatial weight matrices, and the SARMA model with non-parameterized spatial weight matrices, respectively.

Assumption 4.1. *The spatial weight matrices $\boldsymbol{W}(\delta_0)$, $\boldsymbol{W}(\alpha_{0k})$, and $\boldsymbol{W}(\gamma_0)$ are row-normalized and have zero diagonals.*

Assumption 4.2. *The idiosyncratic errors v_{it} for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$ are i.i.d across i and t with zero mean, variance σ_0^2, and finite fourth moment such that $E|v_{it}|^{4+h} < \infty$ for some $h > 0$.*

Assumption 4.3. *$\boldsymbol{S}(\rho,\delta)$ and $\boldsymbol{R}(\lambda,\gamma)$ are invertible for all ρ, δ, λ, and γ in compact intervals $\mathbb{P}_1$, $\mathbb{P}_2$, $\mathbb{P}_2$, and $\mathbb{P}_4$, respectively. ρ_0, δ_0, λ_0, and γ_0 should be interior of $\mathbb{P}_1$, $\mathbb{P}_2$, $\mathbb{P}_2$, and $\mathbb{P}_4$, respectively. Furthermore, ρ_0 and λ_0 are bounded away from zero.*

Assumption 4.4 (for AR errors)**.** *The elements of $\boldsymbol{X}_t$ are non-stochastic and bounded, uniformly in N and T. The limit of $\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{R}'\boldsymbol{J}_N\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^{*}$ exists and is non-singular.*

Assumption 4.4′ (for MA errors)**.** *The elements of $\boldsymbol{X}_t$ are non-stochastic and bounded, uniformly in N and T. The limit of the limit of $\frac{1}{NT}\sum_{t=1}^{T}\widetilde{\boldsymbol{X}}_t^{*'}\boldsymbol{R}^{-1'}\boldsymbol{J}_N\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^{*}$ exists and is non-singular.*

Assumption 4.5. *Before row normalization, the row and column sums of the spatial weight matrices $\boldsymbol{W}(\delta_0)$, $\boldsymbol{W}(\alpha_{0k})$, and $\boldsymbol{W}(\gamma_0)$ are uniformly bounded in both row and column sums in absolute value. $\boldsymbol{S}^{-1}(\rho,\delta)$ and $\boldsymbol{R}^{-1}(\lambda,\gamma)$ are uniformly bounded in both row and column sums in absolute value for ρ, δ, λ, and γ in compact intervals $\mathbb{P}_1$, $\mathbb{P}_2$, $\mathbb{P}_2$, and $\mathbb{P}_4$.*

Assumption 4.6. *ϕ_{0k} for $k = 1, ..., K$ is bounded away from zero.*

Assumption 4.7. *N is large.*

Assumption 4.8 (for AR errors)**.** *Given*

$$\sigma^{2\dagger}(\lambda,\gamma) = \frac{\sigma_0^2}{N-1}tr[(\boldsymbol{R}(\lambda,\gamma)\boldsymbol{R}^{-1})'\boldsymbol{J}_N(\boldsymbol{R}(\lambda,\gamma)\boldsymbol{R}^{-1})]$$
$$\sigma^{2\dagger}(\rho,\delta,\lambda,\gamma) = \frac{\sigma_0^2}{N-1}tr[(\boldsymbol{R}(\lambda,\gamma)\boldsymbol{S}(\rho,\delta)\boldsymbol{S}^{-1}\boldsymbol{R}^{-1})'\boldsymbol{J}_N(\boldsymbol{R}(\lambda,\gamma)\boldsymbol{S}(\rho,\delta)\boldsymbol{S}^{-1}\boldsymbol{R}^{-1})]$$

(i) the limit of

$$(a)\ \ (\lambda,\gamma) = \frac{1}{(N-1)(T-1)}\sum_{t=1}^{T}(\widetilde{\boldsymbol{X}}_t^{*},\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta}_0)'\boldsymbol{R}'(\lambda,\gamma)\boldsymbol{J}_N \times\boldsymbol{R}(\lambda,\gamma)(\widetilde{\boldsymbol{X}}_t^{*},\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^{*}\boldsymbol{\zeta}_0)$$

is nonsingular for each λ and the limit of

$$(b)\ \ \frac{1}{N-1}ln|\sigma_0^2\boldsymbol{R}^{-1'}\boldsymbol{J}_N\boldsymbol{R}^{-1}| - \frac{1}{N-1}ln|\sigma^{2\dagger}(\lambda,\gamma)\boldsymbol{R}^{-1'}(\lambda,\gamma)\boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)|$$

is not zero for $\lambda \neq \lambda_0$ and $\gamma \neq \gamma_0$;
(ii) the limit of $\frac{1}{N-1}ln|\sigma_0^2\boldsymbol{R}^{-1'}\boldsymbol{S}^{-1'}\boldsymbol{J}_N\boldsymbol{S}^{-1}\boldsymbol{R}^{-1}| - \frac{1}{N-1}ln|\sigma_N^2(\rho,\delta,\lambda,\gamma)\boldsymbol{R}^{-1'}(\lambda,\gamma)$ $\boldsymbol{S}^{-1'}(\rho,\delta)\boldsymbol{J}_N\boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{R}^{-1}(\lambda,\gamma)|$ is not zero for $(\rho,\delta,\lambda,\gamma) \neq (\rho_0,\delta_0,\lambda_0,\gamma_0)$, as N tends to infinity.

Assumption 4.8′ (for MA errors). *Given*

$$\sigma^{2\dagger}(\lambda,\gamma) = \frac{\sigma_0^2}{N-1} tr[(\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{R})' \boldsymbol{J}_N (\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{R})]$$
$$\sigma^{2\dagger}(\rho,\delta,\lambda,\gamma) = \frac{\sigma_0^2}{N-1} tr[(\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{S}(\rho,\delta)\boldsymbol{S}^{-1}\boldsymbol{R})' \boldsymbol{J}_N (\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{S}(\rho,\delta)\boldsymbol{S}^{-1}\boldsymbol{R})]$$

(i) the limit of

$$(a)\ \ \mathcal{H}(\lambda,\gamma) = \frac{1}{(N-1)(T-1)} \sum_{t=1}^{T} (\widetilde{\boldsymbol{X}}_t^*, \boldsymbol{G}\widetilde{\boldsymbol{X}}_t^* \boldsymbol{\zeta}_0)' \boldsymbol{R}^{-1'}(\lambda,\gamma) \boldsymbol{J}_N \times \boldsymbol{R}^{-1}(\lambda,\gamma)(\widetilde{\boldsymbol{X}}_t^*, \boldsymbol{G}\widetilde{\boldsymbol{X}}_t^* \boldsymbol{\zeta}_0)$$

is nonsingular for each λ and the limit of

$$(b)\ \ \frac{1}{N-1} ln|\sigma_0^2 \boldsymbol{R}' \boldsymbol{J}_N \boldsymbol{R}| - \frac{1}{N-1} ln|\sigma^{2\dagger}(\lambda,\gamma)\boldsymbol{R}'(\lambda,\gamma)\boldsymbol{J}_N \boldsymbol{R}(\lambda,\gamma)|$$

is not zero for $\lambda \neq \lambda_0$ and $\gamma \neq \gamma_0$;
(ii) the limit of $\frac{1}{N-1} ln|\sigma_0^2 \boldsymbol{R}' \boldsymbol{S}^{-1'} \boldsymbol{J}_N \boldsymbol{S}^{-1} \boldsymbol{R}| - \frac{1}{N-1} ln|\sigma_N^2(\rho,\delta,\lambda,\gamma)\boldsymbol{R}'(\lambda,\gamma)\boldsymbol{S}^{-1'}(\rho,\delta)$ $\boldsymbol{J}_N \boldsymbol{S}^{-1}(\rho,\delta)\boldsymbol{R}(\lambda,\gamma)|$ is not zero for $(\rho,\delta,\lambda,\gamma) \neq (\rho_0,\delta_0,\lambda_0,\gamma_0)$, as N tends to infinity.

For the matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ below and in the appendices, we define $\boldsymbol{A}^s = \boldsymbol{A} + \boldsymbol{A}'$. We further use the truncated notations $\boldsymbol{G} = \boldsymbol{W}(\delta_0)\boldsymbol{S}^{-1}$, $\boldsymbol{H} = \boldsymbol{W}(\gamma_0)\boldsymbol{R}^{-1}$, and $\boldsymbol{Q} = \boldsymbol{R}^{-1}\boldsymbol{W}(\gamma_0)$.

Assumption 4.9 (for AR errors). *The limit of $\frac{1}{(N-1)^2}[tr(\boldsymbol{C}^s\boldsymbol{C}^s)tr(\boldsymbol{D}^s\boldsymbol{D}^s) - tr^2(\boldsymbol{C}^s\boldsymbol{D}^s)]$ is strictly positive, where $\boldsymbol{C} = \boldsymbol{J}_N\ddot{\boldsymbol{G}} - \frac{tr\boldsymbol{J}_N\ddot{\boldsymbol{G}}}{N-1}\boldsymbol{J}_N$, $\boldsymbol{D} = \boldsymbol{J}_N\boldsymbol{H} - \frac{tr\boldsymbol{J}_N\boldsymbol{H}}{N-1}\boldsymbol{J}_N$, and $\ddot{\boldsymbol{G}} = \boldsymbol{R}\boldsymbol{G}\boldsymbol{R}^{-1}$.*

Assumption 4.9′ (for MA errors). *The limit of $\frac{1}{(N-1)^2}[tr(\boldsymbol{C}^s\boldsymbol{C}^s)tr(\boldsymbol{D}^s\boldsymbol{D}^s) - tr^2(\boldsymbol{C}^s\boldsymbol{D}^s)]$ is strictly positive, where $\boldsymbol{C} = \boldsymbol{J}_N\ddot{\boldsymbol{G}} - \frac{tr\boldsymbol{J}\ddot{\boldsymbol{G}}}{N-1}\boldsymbol{J}_N$, $\boldsymbol{D} = \boldsymbol{J}_N\boldsymbol{Q} - \frac{tr\boldsymbol{J}_N\boldsymbol{Q}}{N-1}\boldsymbol{J}_N$, and $\ddot{\boldsymbol{G}} = \boldsymbol{R}^{-1}\boldsymbol{G}\boldsymbol{R}$.*

The row normalization of the spatial weight matrices $\boldsymbol{W}(\delta_0)$ and $\boldsymbol{W}(\gamma_0)$ in Assumption 4.1 ensures that the orthogonal transformation may be implemented to concentrate out the time fixed effects. It also causes the parameters ρ and λ to take values in the interval $(-1,1)$, in line with Assumption 4.3. Note that the spatial weight matrices $\boldsymbol{W}(\alpha_{0k})$ do not necessarily need to be row normalized, but that this makes most sense given the row-normalization of $\boldsymbol{W}(\delta_0)$ and $\boldsymbol{W}(\gamma_0)$. Assumption 4.2 states the standard regularity assumptions for the disturbance term

v_{it}. Assumption 4.3 ensures that the reduced form of the combination of (4.1) and (4.2) and the combination of (4.1) and (4.3) are valid. The first two parts are extensions of Assumption 3 in Lee and Yu (2010a). The last part is a modification of Assumption 2.3 in Chapter 2 to ensure that the distance decay parameters δ_0 and γ_0 are identified. Assumption 4.4 is for the GNSAR model and Assumption 4.4′ for the GNSMA model. They not only assume that all regressors are exogenous and uniformly bounded, but also rule out perfect multicollinearity between any pair of regressors, including the regressors $\boldsymbol{x}_{kt}$ and $\boldsymbol{W}(\alpha_{0k})\boldsymbol{x}_{kt}$. They extend Assumption 4 in Lee and Yu (2010a) and Assumption 3.4 in Chapter 3 with spatial Durbin terms. Assumption 4.5 extends Assumption 5 of Lee and Yu (2010a) and modifies Assumption 2.5 of Chapter 2 in which the boundedness condition also relies on the distance decay parameters. If $\phi_{0k} = 0$ then $\alpha_{0k} = 0$ is not identified for each $k = 1, ..., K$, which is excluded by Assumption 4.6 and taken from Assumption 2.5 in Chapter 2. Assumption 4.7, taken from Lee and Yu (2010a), states that the QML estimator based on the orthogonal transformation is consistent no matter whether T is large or finite. Assumption 4.8 and Assumption 4.8′ are the identification conditions of the QML estimators for the GNSAR and GNSMA models with both fixed effects, to be discussed below. Assumption 4.9 and 4.9′ ensure that the limits of the information matrices of the GNSAR and GNSMA models are non-singular.[2] This is because we exclude that ρ_0, λ_{0k} and ϕ_{0k} ($k = 1, ..., K$) are equal to 0. Assumptions 4.8, 4.8′, 4.9, and 4.9′ are adapted to GNSAR and GNSMA models in which the spatial weight matrices are parameterized.

According to Theorem 3 (2) in Lee and Yu (2010a), the asymptotic distribution of the QML estimator of the GNSAR model when both fixed effects are included, under assumptions 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, and 4.8(i) or 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.8(ii), and 4.9, is

$$\sqrt{(N-1)(T-1)}(\boldsymbol{\theta} - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, \mathbf{\Upsilon}(\boldsymbol{\theta}_0)). \tag{4.13}$$

where $\mathbf{\Upsilon}(\boldsymbol{\theta}_0) = \lim \mathbf{\Sigma}^{-1}(\boldsymbol{\theta}_0)(\mathbf{\Sigma}(\boldsymbol{\theta}_0) + \mathbf{\Omega}(\boldsymbol{\theta}_0))\mathbf{\Sigma}^{-1}(\boldsymbol{\theta}_0)$.

The same holds for the asymptotic distribution of the QML estimator of the

[2] Assumptions 4.9 and 4.9′ also ensure that the limits of the information matrices of the models are non-singular when the QML estimator based on the within transformation is applied.

GNSMA model when both fixed effects are included.[3] The first-order derivatives, second-order derivatives, $\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)$, and $\boldsymbol{\Omega}(\boldsymbol{\theta}_0)$ with respect to the log-likelihood function of the GNSAR model in the form of (4.7) and the GSNMA model in (4.10) are given in Appendix 4.A and Appendix 4.B.

4.2.3 Identification

We first show that the rank conditions for the identification of the fixed effects GNSAR and GNSMA models are the same, provided that the spatial weight matrices are not parameterized. This extends the work of Lee et al. (2010) in which the GNSAR model with network matrix only contains group fixed effects.[4] Then we generalize the rank conditions for identification when the spatial weight matrices are parameterized.

Consider the fixed effects GNSAR model given by

$$\begin{aligned} \boldsymbol{y}_t &= \rho_0 \boldsymbol{W}\boldsymbol{y}_t + \boldsymbol{X}_t\boldsymbol{\beta}_0 + \boldsymbol{W}\boldsymbol{X}_t\boldsymbol{\phi}_0 + \boldsymbol{c}_0 + \varpi_{0t}\boldsymbol{\iota}_N + \boldsymbol{u}_t, \\ \boldsymbol{u}_t &= \lambda_0 \boldsymbol{M}\boldsymbol{u}_t + \boldsymbol{v}_t, t = 1, ..., T, \end{aligned} \tag{4.14}$$

where $\boldsymbol{W}$ and $\boldsymbol{M}$ are not parameterized. The identification problem arises if $\rho_0\boldsymbol{\beta}_0 + \boldsymbol{\phi}_0 = \boldsymbol{0}$ and $\boldsymbol{W} = \boldsymbol{M}$. Under these circumstances, $\boldsymbol{\beta}_0$, $\boldsymbol{\phi}_0$, and ρ_0 can no longer be separately identified. To reveal the identification problems when $\boldsymbol{W} = \boldsymbol{M}$, we take the reduced form of the GNSAR model with both fixed effects as point of departure

$$\boldsymbol{y}_t^{\bullet\bullet} = \boldsymbol{S}^{\bullet -1}(\boldsymbol{X}_t^{\bullet\bullet}\boldsymbol{\beta}_0 + \boldsymbol{W}^{\bullet}\boldsymbol{X}_t^{\bullet\bullet}\boldsymbol{\phi}_0) + \boldsymbol{S}^{\bullet -1}\boldsymbol{R}^{\bullet -1}\boldsymbol{v}_t^{\bullet\bullet}, \ \ t = 1, 2, ..., T-1, \tag{4.15}$$

where $\boldsymbol{W}^{\bullet} = \boldsymbol{F}_{N,N-1}^{'}\boldsymbol{W}\boldsymbol{F}_{N,N-1}$, $\boldsymbol{M}^{\bullet} = \boldsymbol{F}_{N,N-1}^{'}\boldsymbol{M}\boldsymbol{F}_{N,N-1}$ and $\boldsymbol{S}^{\bullet -1} = \boldsymbol{I}_{N-1} - \rho_0\boldsymbol{W}^{\bullet}$, and $\boldsymbol{R}^{\bullet -1} = \boldsymbol{I}_{N-1} - \lambda_0\boldsymbol{M}^{\bullet}$. Lee et al. (2010) show that ρ_0 and λ_0 are identified when $\boldsymbol{W} \neq \boldsymbol{M}$ and thus $\boldsymbol{W}^{\bullet} \neq \boldsymbol{M}^{\bullet}$, since $\boldsymbol{\epsilon}_t = \lambda_0\boldsymbol{M}^{\bullet}\boldsymbol{\epsilon}_t + \rho_0\boldsymbol{W}^{\bullet}\boldsymbol{\epsilon}_t - \rho_0\lambda_0\boldsymbol{M}^{\bullet}\boldsymbol{W}^{\bullet}\boldsymbol{\epsilon}_t + \boldsymbol{v}_t^{\bullet\bullet}$ then follows a higher-order SAR process.[5] Conversely, ρ_0 and λ_0 cannot be identified when $\boldsymbol{W} = \boldsymbol{M}$ and thus $\boldsymbol{W}^{\bullet} = \boldsymbol{M}^{\bullet}$, since $\boldsymbol{\epsilon}_t$ then degenerates to $\boldsymbol{\epsilon}_t =$

[3] Details are provided in a working paper of Lee and Yu (2010c).

[4] Group fixed effects should not be confused with individual fixed effects. In principle, they might be interpreted as time fixed effects.

[5] Lee and Liu (2010) discuss the identification conditions for higher-order SAR and SARAR models in detail. Lee (2004) dicusses the identification conditions for higher-order SAR model. If $\rho_0\boldsymbol{\beta}_0 + \boldsymbol{\phi}_0 = \boldsymbol{0}$ in a SD model, then it simplifies to a spatial error model (Bramoulle et al., 2009, equation (6)). If this condition holds in a GNS model with spatial AR or MA errors, one obtains a higher-order spatial error model.

$(\rho_0 + \lambda_0)\boldsymbol{W}^\bullet\boldsymbol{\epsilon}_t - \rho_0\lambda_0\boldsymbol{W}^{\bullet 2}\boldsymbol{\epsilon}_t + \boldsymbol{v}_t^{\bullet\bullet}$ in which ρ_0 and λ_0 can be switched around without observing any difference.

The reduced form of the GNSMA model, which is not part of the work of Lee et al. (2010), reads as

$$\boldsymbol{y}_t^{\bullet\bullet} = \boldsymbol{S}^{\bullet -1}(\boldsymbol{X}_t^{\bullet\bullet}\boldsymbol{\beta}_0 + \boldsymbol{W}^\bullet\boldsymbol{X}_t^{\bullet\bullet}\boldsymbol{\phi}_0) + \boldsymbol{S}^{\bullet -1}\boldsymbol{R}^\bullet\boldsymbol{v}_t^{\bullet\bullet}, \quad t = 1, 2, ..., T-1, \tag{4.16}$$

Given that $(\boldsymbol{I}_{N-1} - \lambda_0\boldsymbol{M}^\bullet)^{-1} = \sum_{i=1}^{\infty}\lambda_0^i\boldsymbol{M}^{\bullet i}$ for the GNSMA model, $\boldsymbol{\epsilon}_t = -\lambda_0\boldsymbol{M}^\bullet\boldsymbol{\epsilon}_t + \rho_0\boldsymbol{W}^\bullet\boldsymbol{\epsilon}_t + \rho_0\lambda_0\boldsymbol{M}^\bullet\boldsymbol{W}^\bullet\boldsymbol{\epsilon}_t - (\sum_{i=2}^{\infty}\lambda_0^i\boldsymbol{M}^{\bullet i})\boldsymbol{\epsilon}_t + (\sum_{i=2}^{\infty}\lambda_0^i\rho_0\boldsymbol{M}^{\bullet i}\boldsymbol{W}^\bullet)\boldsymbol{\epsilon}_t + \boldsymbol{v}_t^{\bullet\bullet}$ also follows a higher-order SAR process. This process shows that ρ_0 and λ_0 cannot be separately identified when $\boldsymbol{W}^\bullet = \boldsymbol{M}^\bullet$.

Three conclusions can be drawn from this overview. First, the identification problems of the GNSAR and GNSMA model are similar to those of the SD model. Second, the identification conditions for ρ_0 and λ_0 within the GNSAR and GNSMA models are the same when the spatial weight matrices are not parameterized. Only when $\boldsymbol{W} \neq \boldsymbol{M}$ are ρ_0 and λ_0 identified, while $\boldsymbol{\phi}_0$ is identified via $\boldsymbol{\phi}_0 = \rho_0\boldsymbol{\beta}_0$. Third, when the spatial weight matrices are parameterized, then the identification conditions of the GNSAR and GNSMA models are different because the matrix $\boldsymbol{R}(\lambda, \gamma)$ of a spatial MA error process is inverted and of a spatial AR error process is not.

Given that the disturbances $\boldsymbol{v}_t$ of the GNSAR and GNSMA models follow different higher-order SAR processes, the rank conditions for the identification of the GNSAR and GNSMA models with both fixed effects are the same when the spatial weight matrices are not parameterized. Lemma 4.1 provides a sufficient condition for the rank condition under Assumption 4.8(i) for GNSAR and Assumption 4.8'(i) for GNSMA models with both fixed effects when the spatial weight matrices are not parameterized:

Lemma 4.1. *If* $\rho_0\boldsymbol{\beta}_0 + \boldsymbol{\phi}_0 \neq \boldsymbol{0}$ *and* $[\widetilde{\boldsymbol{X}}_t, \boldsymbol{W}\widetilde{\boldsymbol{X}}_t, \boldsymbol{W}^2\widetilde{\boldsymbol{X}}_t, \boldsymbol{\iota}_N]$ *has full column rank, then* $\boldsymbol{J}_N\boldsymbol{R}(\lambda)(\widetilde{\boldsymbol{X}}_t^*, \boldsymbol{G}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}_0)$ *has full column rank for the GNSAR model and* $\boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda)(\widetilde{\boldsymbol{X}}_t^*, \boldsymbol{G}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}_0)$ *has full column rank for the GNSMA model when both fixed effects are included.*

Expression (a) of Assumptions 4.8(i) and 4.8'(i) ensures that ρ_0, $\boldsymbol{\beta}_0$, and $\boldsymbol{\phi}_0$ in the GNSAR and GNSMA models are locally identified, respectively. Expression (b) of 4.8(i) and 4.8'(i) ensures that λ_0 is identified through the $\boldsymbol{S}\boldsymbol{R}^{-1}\boldsymbol{v}_t$ process for the

GNSAR model and $\boldsymbol{SRv}_t$ for the GNSMA model. When Expression (a) of Assumptions 4.8(i) and 4.8′(i) does not hold, but Assumption 4.8(ii) or 4.8′(ii), then (ρ_0, λ_0) is globally identified.

Lemma 4.2 provides a sufficient condition for the rank condition under Assumption 4.8(i) for GNSAR and Assumption 4.8′(i) for GNSMA models with both fixed effects when the spatial weight matrices are parameterized:

Lemma 4.2. *If* $\rho_0\beta_{0k}\boldsymbol{W}(\delta_0) + \phi_{0k}\boldsymbol{W}(\alpha_{0k}) \neq \boldsymbol{0}$ *and* $[\widetilde{\boldsymbol{X}}_t, \boldsymbol{W}(\delta_0)\widetilde{\boldsymbol{X}}_t, \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0,1})\widetilde{\boldsymbol{x}}_{1t}, ..., \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0K})\widetilde{\boldsymbol{x}}_{Kt}, \boldsymbol{\iota}_N]$ *has full column rank, then* $\boldsymbol{J}_N\boldsymbol{R}(\lambda,\gamma)(\widetilde{\boldsymbol{X}}_t^*, \boldsymbol{G}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}_0)$ *has full column rank for the GNSAR model and* $\boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)(\widetilde{\boldsymbol{X}}_t^*, \boldsymbol{G}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}_0)$ *has full column rank for the GNSMA model when both fixed effects are included.*

When the spatial weight matrices are parameterized, Expression (a) of 4.8(i) and 4.8′(i) ensures that ρ_0, δ_0, $\boldsymbol{\beta}_0$, and $\boldsymbol{\phi}_0$ are locally identified for the GNSAR and GNSMA models, respectively. Expression (b) of 4.8(i) and 4.8′(i) ensures that λ_0 and γ_0 are identified through the $\boldsymbol{SR}^{-1}\boldsymbol{v}_t$ process for the GNSAR model and $\boldsymbol{SRv}_t$ for the GNSMA model. When Expression (a) of Assumptions 4.8(i) and 4.8′(i) does not hold, but Assumption 4.8(ii) or 4.8′(ii), then $(\rho_0, \delta_0, \lambda_0, \gamma_0)$ are globally identified.

The first part of Lemma 4.2 is less restrictive than the first part in Lemma 4.1 due to employing a different weight matrix for each spatial lag. Lemma 4.2 rules out two cases: (i) $\rho_0\beta_k + \phi_k = 0$ and $\delta_0 = \alpha_{0k}$ for all k; and (ii) $\boldsymbol{W}(\delta_0) = \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0k})$. Under case (i), the GNS models will reduce to a spatial error model which only contain spatial interactions among the errors, which is shown in Chapter 3. In empirical studies, however, the estimated distance decay parameters are expected to be different since individual regressors tend to follow different distance decay patterns. Under case (ii), the full column rank conditions in Lemma 4.2 are not met. However, endogenous and exogenous spatial interactions, captured by $\boldsymbol{W}(\delta_0)\boldsymbol{y}_t$ and $\boldsymbol{W}(\alpha_{0k})\boldsymbol{x}_{kt}$, also tend to follow different distance decay patterns. Furthermore, they do not tend to diverge to infinity,[6] as a result of which $\boldsymbol{W}(\delta_0) = \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0k})$ does not hold, which is the second condition stated in Lemma 4.2. Overall, it can be concluded that Lemma 4.2 will not be restrictive in empirical applications. A formal proof of Lemma 4.2, and of Lemma 4.1 as special case of Lemma 4.2, is given in Appendix 4.C.

[6] If the distance decay parameter grows large, the spatial weight matrices will converge to binary contiguity matrix as the elements of distant units go to zero faster than nearby units.

4.3 Monte Carlo simulations

A Monte Carlo experiment is conducted to evaluate the performance of our proposed estimator for the fixed effects GNSAR and GNSMA models in which the spatial weight matrix is specified as a row-normalized negative exponential (ED) distance decay matrix.

Data are generated either from the process given in (4.1) and (4.2) for the GNSAR model or (4.1) and (4.3) for the GNSMA model. Two regressors are included: $x_{1it} \sim \mathcal{N}(-2,5)$ and $x_{2it} \sim \mathcal{N}(3,3)$. The coefficients of x_{1it} and its spatial lag are $\beta_1 = 1$ and $\phi_1 = -0.2$ and those of x_{2it} and its spatial lag are $\beta_2 = -1.5$ and $\phi_2 = 0.3$. $\boldsymbol{c}_0$, $\boldsymbol{\omega}_{0t}$, and $\boldsymbol{v}_t$ are generated from independent standard normal distributions. We set $N = (50, 200)$, $T = (50, 100)$, and the number of iterations equal to 1000. To construct distance matrices between the cross-sectional units, we use the coordinates of N data points evenly set in a rectangle of 5×10 for $N = 50$ and 10×20 for $N = 200$. These distances are used to calculate the exponential distance decay matrix. The distance decay parameters for the spatial weight matrices interacting with different spatial lags are set as $(\delta, \alpha_1, \alpha_2, \gamma) = (2, 4, 2.5, 3)$. Row-normalization of the spatial weight matrices is required. Two settings of the (ρ, λ) are considered. In the first setting $(\rho, \lambda) = (0.2, 0.5)$, such that the degree of spatial dependence in the dependent variable is relatively mild (0.2) and the spatial dependence in the error term relatively large (0.5). In the second setting $(\rho, \lambda) = (0.5, 0.2)$, such that these two degrees of spatial dependence are turned around.

In our experiments we focus on both the parameter estimates and the estimates of the direct and indirect effects. The formulas for the direct and indirect effects are taken from Chapter 2. They do not change when spatial AR or MA errors are added to the SD model with parameterized spatial weight matrices. To evaluate the performance of the estimators, four statistics are considered: the empirical bias (Bias), the empirical root mean square error (RMSE), the median bias (Mbias), and the median absolute value of the bias (Mabias). Among them, Mbias and Mabias are used since they are robust to outliers. Besides, the asymptotic distribution of the estimators is evaluated. Specifically, we test the null hypothesis whether each parameter equals its true value and report the mean and the standard deviations

of the p-values. If the underlying asymptotic distribution is true, then the p-values should follow a $U(0,1)$ distribution, and thus should have a mean p-value of 0.5 and a standard deviation of approximately 0.29.

Table 4.1 reports the simulations results when $\rho = 0.2$ and $\lambda = 0.5$ and Table 4.2 when $\rho = 0.5$ and $\lambda = 0.2$. These tables show that the simulation results for both cases exhibit similar patterns. As N, T, or both increase, the biases of the coefficient estimates, direct and indirect effects, and distance decay parameters tend to go down. The same holds for the RMSE, Mbias, and Mabias. However, there are some striking differences. First, the biases for the estimates of the GNSMA model are smaller than those of the GNSAR model for both settings, except for the parameter λ in the second setting. Second, the biases of the distance decay parameter estimates are generally larger than the biases of the response parameters. Among the distance decay parameters, the biases in γ in turn are larger than those of the other distance decay parameters. Third, when comparing the parameters ρ and λ, it turns out that the biases in λ are larger. Fourth, in general, the biases in the coefficients of the spatial lags of the explanatory variables are larger than those of the explanatory variables themselves. Similarly, the biases of the indirect effect estimates are larger than those of the direct effect estimates of the explanatory variables. Overall, the biases, RMSE, Mbias, and Mabias of the estimates are nevertheless relatively small, except for γ in the second case. Only if both N and T are large, do these four statistics of γ take acceptable small values. Table 4.3 reports the average p-values and standard deviations of the p-values. The simulation results show that all the estimated parameters and direct and indirect effects are properly centered for both models and both panel data settings of N and T.

4.4 Empirical application

The empirical application is based on a recent study of Yang (2021), which contains data on house prices, population density, income, and spatial information of 377 Metropolitan Statistical Areas (MSAs) over the period 1975Q1 to 2014Q4 (total number of time periods is 160). According to the United States Office of Management and Budget, the MSA is defined as the core area with a population density of 50,000 people or more including surrounding territory that has high commuting

Table 4.1. Simulation results for $\rho = 0.2$ and $\lambda = 0.5$

	T	N		β_1	β_2	ϕ_1	ϕ_2	ρ	λ	σ^2	δ	α_1	α_2	γ	DE_x_1	DE_x_2	IE_x_1	IE_x_2
				GNSAR														
(1)	**50**	**50**	**Bias**	0.0015	0.0003	0.0026	0.0041	0.0098	-0.0025	-0.0042	0.3669	0.0736	0.0967	0.1832	0.0000	0.0006	-0.0058	0.0071
			RMSE	0.0089	0.0080	0.0184	0.0428	0.0546	0.0466	0.0299	1.3542	0.2500	0.5674	0.9116	0.0047	0.0084	0.0509	0.0552
			Mbias	0.0018	0.0002	0.0020	0.0010	0.0028	-0.0032	-0.0038	0.0144	0.0218	0.0065	0.0432	0.0000	0.0005	-0.0009	0.0031
			Mabias	0.0060	0.0054	0.0125	0.0289	0.0357	0.0313	0.0203	0.4822	0.1268	0.3302	0.4961	0.0032	0.0056	0.0319	0.0380
(2)	**100**	**50**	**Bias**	0.0006	0.0000	0.0018	0.0023	0.0046	0.0004	-0.0020	0.1884	0.0446	0.0400	0.0804	-0.0001	0.0002	-0.0023	0.0039
			RMSE	0.0065	0.0053	0.0144	0.0293	0.0401	0.0341	0.0211	0.8158	0.1839	0.3479	0.5707	0.0036	0.0056	0.0380	0.0381
			Mbias	0.0005	0.0001	0.0017	0.0012	0.0008	0.0009	-0.0018	0.0235	0.0053	0.0052	0.0193	-0.0002	0.0002	0.0003	0.0026
			Mabias	0.0045	0.0036	0.0098	0.0192	0.0275	0.0237	0.0140	0.3849	0.0978	0.2282	0.3518	0.0025	0.0038	0.0269	0.0250
(3)	**50**	**200**	**Bias**	0.0004	0.0000	0.0011	0.0021	0.0022	-0.0006	-0.0007	0.0993	0.0256	0.0184	0.0414	0.0000	0.0001	-0.0007	0.0031
			RMSE	0.0047	0.0037	0.0101	0.0224	0.0289	0.0241	0.0150	0.4942	0.1164	0.2404	0.3807	0.0024	0.0038	0.0259	0.0287
			Mbias	0.0004	0.0001	0.0008	0.0001	0.0002	-0.0002	-0.0014	0.0052	0.0144	0.0030	0.0102	0.0000	0.0002	0.0002	0.0008
			Mabias	0.0032	0.0025	0.0067	0.0150	0.0186	0.0156	0.0102	0.2680	0.0714	0.1576	0.2461	0.0016	0.0026	0.0168	0.0185
(4)	**100**	**200**	**Bias**	0.0001	-0.0001	0.0003	0.0007	0.0020	-0.0011	-0.0008	0.0233	0.0054	0.0032	0.0335	0.0000	-0.0001	-0.0014	0.0013
			RMSE	0.0031	0.0026	0.0068	0.0146	0.0205	0.0171	0.0104	0.2976	0.0717	0.1592	0.2534	0.0016	0.0027	0.0176	0.0188
			Mbias	0.0001	-0.0001	0.0003	0.0005	0.0013	-0.0015	-0.0011	-0.0014	0.0036	-0.0054	0.0207	-0.0001	-0.0001	-0.0007	0.0005
			Mabias	0.0020	0.0018	0.0046	0.0096	0.0145	0.0115	0.0071	0.1934	0.0453	0.1036	0.1603	0.0011	0.0019	0.0117	0.0125
				GNSMA														
(5)	**50**	**50**	**Bias**	0.0008	0.0003	0.0007	0.0010	0.0037	0.0085	-0.0035	0.1656	0.0470	0.0582	0.1512	0.0000	0.0004	-0.0016	0.0019
			RMSE	0.0090	0.0073	0.0090	0.0196	0.0303	0.0431	0.0298	0.8746	0.2405	0.4186	0.8930	0.0040	0.0070	0.0203	0.0243
			Mbias	0.0004	0.0000	0.0006	0.0004	0.0033	0.0071	-0.0037	0.0117	0.0123	-0.0080	0.0007	0.0000	0.0003	-0.0003	0.0014
			Mabias	0.0058	0.0049	0.0064	0.0139	0.0199	0.0276	0.0200	0.3353	0.1371	0.2649	0.4832	0.0028	0.0049	0.0126	0.0166
(6)	**100**	**50**	**Bias**	0.0002	0.0000	0.0005	0.0005	0.0013	0.0027	-0.0017	0.0841	0.0269	0.0286	0.0672	-0.0001	0.0000	-0.0003	0.0008
			RMSE	0.0064	0.0048	0.0068	0.0131	0.0220	0.0287	0.0210	0.4883	0.1643	0.2663	0.5228	0.0029	0.0047	0.0152	0.0164
			Mbias	0.0001	-0.0002	0.0006	-0.0003	0.0009	0.0038	-0.0015	0.0066	0.0143	0.0119	-0.0138	-0.0002	0.0000	0.0008	0.0006
			Mabias	0.0043	0.0033	0.0048	0.0091	0.0158	0.0199	0.0140	0.2701	0.1052	0.1691	0.3081	0.0020	0.0033	0.0098	0.0115
(7)	**50**	**200**	**Bias**	0.0002	-0.0001	0.0003	0.0007	0.0010	0.0023	-0.0005	0.0385	0.0141	0.0129	0.0200	0.0000	0.0000	-0.0003	0.0010
			RMSE	0.0047	0.0035	0.0053	0.0098	0.0159	0.0201	0.0150	0.3164	0.1092	0.1793	0.3382	0.0021	0.0034	0.0098	0.0121
			Mbias	0.0004	-0.0001	0.0001	0.0009	0.0010	0.0024	-0.0008	0.0006	0.0047	0.0012	0.0041	0.0000	0.0000	0.0000	0.0011
			Mabias	0.0031	0.0022	0.0034	0.0066	0.0114	0.0132	0.0102	0.1909	0.0709	0.1152	0.2140	0.0014	0.0022	0.0067	0.0080
(8)	**100**	**200**	**Bias**	-0.0001	-0.0001	0.0001	0.0003	0.0005	0.0012	-0.0007	0.0055	0.0012	0.0019	0.0009	0.0000	-0.0001	-0.0004	0.0004
			RMSE	0.0032	0.0025	0.0035	0.0063	0.0107	0.0144	0.0103	0.1996	0.0721	0.1186	0.2189	0.0014	0.0024	0.0064	0.0078
			Mbias	-0.0002	-0.0001	0.0001	0.0004	0.0007	0.0019	-0.0008	-0.0079	-0.0014	-0.0084	-0.0068	0.0000	0.0000	-0.0004	0.0008
			Mabias	0.0020	0.0017	0.0025	0.0041	0.0071	0.0097	0.0071	0.1345	0.0473	0.0794	0.1420	0.0009	0.0016	0.0044	0.0052

Table 4.2. Simulation results for $\rho = 0.5$ and $\lambda = 0.2$

	T	N		β_1	β_2	ϕ_1	ϕ_2	ρ	λ	σ^2	δ	α_1	α_2	γ	DE_x_1	DE_x_2	IE_x_1	IE_x_2
				GNSAR														
(1)	**50**	**50**	**Bias**	-0.0001	0.0001	0.0009	0.0036	-0.0033	0.0135	-0.0038	0.0570	0.0349	0.0538	1.0644	0.0000	0.0002	0.0042	0.0064
			RMSE	0.0077	0.0072	0.0135	0.0340	0.0340	0.0634	0.0297	0.2778	0.1964	0.4851	2.9620	0.0048	0.0083	0.0636	0.0683
			Mbias	-0.0003	0.0001	0.0006	0.0005	-0.0031	0.0050	-0.0032	0.0320	0.0152	0.0035	0.0563	0.0000	0.0000	0.0080	0.0033
			Mabias	0.0055	0.0049	0.0090	0.0211	0.0233	0.0382	0.0194	0.1693	0.1280	0.2969	1.1885	0.0032	0.0057	0.0448	0.0421
(2)	**100**	**50**	**Bias**	-0.0001	0.0001	0.0001	0.0003	-0.0026	0.0061	-0.0029	0.0379	0.0195	0.0485	0.6605	0.0001	0.0000	0.0026	0.0002
			RMSE	0.0059	0.0048	0.0101	0.0225	0.0273	0.0464	0.0203	0.2116	0.1452	0.3159	2.2097	0.0036	0.0054	0.0524	0.0454
			Mbias	-0.0002	0.0001	0.0003	-0.0004	-0.0016	0.0025	-0.0031	0.0119	0.0090	0.0226	0.0128	0.0001	0.0000	0.0047	-0.0011
			Mabias	0.0038	0.0032	0.0066	0.0155	0.0185	0.0296	0.0136	0.1338	0.0866	0.1967	0.8314	0.0024	0.0036	0.0357	0.0305
(3)	**50**	**200**	**Bias**	0.0004	-0.0001	0.0001	0.0000	0.0002	0.0005	-0.0013	0.0190	0.0158	0.0240	0.2605	0.0000	-0.0001	-0.0005	0.0001
			RMSE	0.0042	0.0035	0.0074	0.0167	0.0176	0.0313	0.0144	0.1415	0.1004	0.2235	1.1759	0.0023	0.0039	0.0323	0.0341
			Mbias	0.0005	-0.0001	0.0003	-0.0006	0.0005	-0.0014	-0.0013	0.0047	0.0118	0.0100	0.0951	0.0000	-0.0001	0.0010	-0.0007
			Mabias	0.0028	0.0024	0.0051	0.0109	0.0122	0.0203	0.0097	0.0966	0.0664	0.1406	0.6055	0.0016	0.0027	0.0211	0.0225
(4)	**100**	**200**	**Bias**	0.0000	0.0000	-0.0002	0.0001	-0.0005	0.0001	-0.0008	0.0077	0.0038	0.0043	0.1369	0.0000	0.0000	0.0000	0.0001
			RMSE	0.0029	0.0025	0.0051	0.0111	0.0131	0.0219	0.0100	0.0957	0.0687	0.1441	0.7205	0.0016	0.0027	0.0245	0.0227
			Mbias	0.0001	0.0000	-0.0002	-0.0002	-0.0002	-0.0007	-0.0013	0.0005	0.0022	-0.0007	0.0294	0.0000	0.0000	0.0004	-0.0001
			Mabias	0.0020	0.0017	0.0035	0.0072	0.0086	0.0143	0.0071	0.0616	0.0480	0.0949	0.4303	0.0011	0.0018	0.0163	0.0144
				GNSMA														
(5)	**50**	**50**	**Bias**	-0.0002	0.0001	0.0010	0.0024	0.0020	0.0228	-0.0039	0.0058	0.0126	0.0401	0.8152	-0.0002	0.0001	-0.0048	0.0060
			RMSE	0.0078	0.0070	0.0112	0.0263	0.0280	0.0774	0.0298	0.2398	0.1998	0.4501	2.7562	0.0043	0.0073	0.0494	0.0522
			Mbias	-0.0004	0.0002	0.0014	0.0008	0.0018	0.0112	-0.0034	-0.0117	-0.0075	0.0002	-0.0664	-0.0002	0.0000	-0.0026	0.0030
			Mabias	0.0055	0.0047	0.0072	0.0163	0.0190	0.0397	0.0193	0.1587	0.1277	0.2835	1.2383	0.0028	0.0051	0.0338	0.0329
(6)	**100**	**50**	**Bias**	-0.0001	0.0000	0.0001	0.0000	0.0002	0.0105	-0.0028	0.0096	0.0073	0.0380	0.4953	0.0000	0.0000	-0.0020	0.0005
			RMSE	0.0058	0.0048	0.0084	0.0174	0.0217	0.0464	0.0203	0.1808	0.1441	0.2924	2.0059	0.0031	0.0049	0.0393	0.0345
			Mbias	0.0000	0.0001	0.0003	-0.0003	0.0003	0.0076	-0.0030	-0.0036	0.0028	0.0129	-0.0227	0.0000	0.0002	0.0007	0.0002
			Mabias	0.0038	0.0032	0.0055	0.0120	0.0149	0.0278	0.0137	0.1189	0.0919	0.1884	0.8049	0.0021	0.0033	0.0262	0.0233
(7)	**50**	**200**	**Bias**	0.0002	-0.0001	0.0004	-0.0001	0.0014	0.0066	-0.0010	0.0036	0.0085	0.0181	0.1999	0.0000	-0.0001	-0.0023	0.0004
			RMSE	0.0042	0.0034	0.0063	0.0129	0.0141	0.0311	0.0144	0.1283	0.1019	0.2085	1.1978	0.0021	0.0035	0.0238	0.0259
			Mbias	0.0004	-0.0001	0.0005	-0.0002	0.0010	0.0058	-0.0009	-0.0007	0.0067	0.0007	-0.0443	0.0000	0.0000	-0.0011	-0.0006
			Mabias	0.0027	0.0024	0.0043	0.0087	0.0097	0.0199	0.0096	0.0878	0.0681	0.1322	0.5430	0.0015	0.0024	0.0158	0.0170
(8)	**100**	**200**	**Bias**	-0.0001	0.0000	-0.0001	0.0001	0.0002	0.0032	-0.0006	-0.0002	-0.0002	0.0003	0.0599	0.0000	0.0000	-0.0010	0.0004
			RMSE	0.0029	0.0024	0.0043	0.0084	0.0103	0.0210	0.0100	0.0864	0.0703	0.1329	0.6639	0.0015	0.0025	0.0178	0.0168
			Mbias	0.0000	0.0000	-0.0001	-0.0003	0.0004	0.0028	-0.0011	-0.0051	-0.0013	-0.0029	-0.0409	-0.0001	0.0001	-0.0011	0.0005
			Mabias	0.0019	0.0016	0.0030	0.0054	0.0072	0.0135	0.0071	0.0568	0.0491	0.0900	0.4168	0.0010	0.0017	0.0116	0.0108

Table 4.3. Simulation results of the means and the standard deviations of the p-values

	T	N		β_1	β_2	ϕ_1	ϕ_2	ρ	λ	δ	α_1	α_2	γ	DE_x_1	DE_x_2	IE_x_1	IE_x_2
	ρ: 0.2. λ: 0.5			**GNSAR**													
(1)	**50**	**50**	**Mean**	0.5060	0.4883	0.5015	0.4893	0.5012	0.5293	0.5093	0.5029	0.4916	0.5282	0.5014	0.4864	0.5064	0.4885
			Std	0.2810	0.2931	0.2881	0.2935	0.3034	0.2886	0.2880	0.2876	0.2806	0.2784	0.2882	0.2934	0.2952	0.2922
(2)	**100**	**50**	**Mean**	0.5029	0.5046	0.5006	0.5021	0.4967	0.5287	0.5003	0.4922	0.5045	0.5317	0.4955	0.4954	0.4966	0.5026
			Std	0.2893	0.2842	0.2885	0.2849	0.2938	0.2793	0.2959	0.2880	0.2872	0.2912	0.2897	0.2828	0.2848	0.2884
(3)	**50**	**200**	**Mean**	0.4891	0.5108	0.5075	0.4945	0.5289	0.5553	0.5189	0.4959	0.5045	0.5409	0.5031	0.5012	0.5207	0.4988
			Std	0.2978	0.2873	0.2933	0.2917	0.2901	0.2785	0.2939	0.2886	0.2887	0.2854	0.2914	0.2832	0.2836	0.2936
(4)	**100**	**200**	**Mean**	0.5057	0.4950	0.4988	0.5144	0.5236	0.5510	0.5217	0.5233	0.5081	0.5598	0.5081	0.4955	0.5268	0.5105
			Std	0.2883	0.2854	0.2935	0.2979	0.2793	0.2706	0.2850	0.2882	0.2857	0.2729	0.2992	0.2832	0.2801	0.2960
	ρ: 0.2. λ: 0.5			**GNSMA**													
(5)	**50**	**50**	**Mean**	0.5074	0.4891	0.5021	0.4849	0.5101	0.4966	0.5217	0.5039	0.4923	0.4901	0.4979	0.4908	0.5086	0.4863
			Std	0.2913	0.2894	0.2888	0.2864	0.2855	0.2980	0.2857	0.2905	0.2779	0.2850	0.2908	0.2948	0.2936	0.2884
(6)	**100**	**50**	**Mean**	0.5015	0.5012	0.4931	0.5072	0.5157	0.4944	0.5039	0.4908	0.5074	0.5015	0.4990	0.4991	0.5011	0.4988
			Std	0.2894	0.2802	0.2803	0.2859	0.2831	0.2765	0.2857	0.2857	0.2867	0.2850	0.2919	0.2817	0.2886	0.2827
(7)	**50**	**200**	**Mean**	0.4836	0.5093	0.5070	0.4956	0.4866	0.5051	0.4966	0.4953	0.5061	0.4971	0.4961	0.5054	0.4934	0.4934
			Std	0.2931	0.2868	0.2992	0.2923	0.2862	0.2927	0.2867	0.2951	0.2849	0.2900	0.2902	0.2861	0.2885	0.2907
(8)	**100**	**200**	**Mean**	0.5068	0.4980	0.5018	0.5095	0.5036	0.4930	0.5053	0.5100	0.5079	0.5139	0.5019	0.4999	0.4983	0.5020
			Std	0.2908	0.2888	0.2858	0.2890	0.2875	0.2868	0.2875	0.2864	0.2846	0.2837	0.2922	0.2891	0.2876	0.2868
	ρ: 0.5. λ: 0.2			**GNSAR**													
(9)	**50**	**50**	**Mean**	0.5140	0.4972	0.5094	0.5093	0.5110	0.5293	0.5119	0.5190	0.5030	0.5546	0.5059	0.4922	0.5058	0.5096
			Std	0.2841	0.2957	0.2900	0.2877	0.2858	0.2839	0.2775	0.2919	0.2895	0.2915	0.2966	0.2909	0.2838	0.2869
(10)	**100**	**50**	**Mean**	0.5060	0.5101	0.5100	0.4976	0.4948	0.5100	0.4951	0.5038	0.4989	0.5558	0.5065	0.5171	0.4883	0.4948
			Std	0.2903	0.2853	0.2918	0.2922	0.2905	0.2867	0.2923	0.2948	0.2818	0.2782	0.2860	0.2842	0.2890	0.2911
(11)	**50**	**200**	**Mean**	0.5062	0.5007	0.4968	0.5098	0.5066	0.5239	0.4904	0.4922	0.5040	0.5305	0.4892	0.5012	0.5069	0.5126
			Std	0.2920	0.2873	0.2869	0.2850	0.2883	0.2919	0.2881	0.2889	0.2889	0.2794	0.2902	0.2888	0.2885	0.2893
(12)	**100**	**200**	**Mean**	0.5062	0.4966	0.5109	0.5149	0.5003	0.5169	0.5161	0.4993	0.5130	0.5142	0.4967	0.4968	0.5005	0.5146
			Std	0.2822	0.2934	0.2861	0.2904	0.2862	0.2842	0.2893	0.2805	0.2833	0.2749	0.2846	0.2936	0.2911	0.2927
	ρ: 0.5. λ: 0.2			**GNSMA**													
(13)	**50**	**50**	**Mean**	0.5058	0.4979	0.5098	0.5016	0.5057	0.4998	0.5038	0.5097	0.4988	0.5207	0.4994	0.4930	0.4986	0.5019
			Std	0.2828	0.2910	0.2899	0.2907	0.2912	0.2897	0.2870	0.2944	0.2943	0.3042	0.2897	0.2906	0.2835	0.2921
(14)	**100**	**50**	**Mean**	0.5017	0.5034	0.5097	0.4908	0.4968	0.4973	0.4884	0.4979	0.4972	0.5411	0.5025	0.5141	0.4895	0.4899
			Std	0.2890	0.2838	0.2950	0.2892	0.2905	0.2836	0.2906	0.2909	0.2792	0.2889	0.2861	0.2835	0.2927	0.2905
(15)	**50**	**200**	**Mean**	0.5006	0.4997	0.4938	0.5007	0.4976	0.4940	0.4887	0.4934	0.4944	0.5171	0.4883	0.4968	0.4976	0.5004
			Std	0.2920	0.2862	0.2849	0.2827	0.2898	0.2910	0.2949	0.2928	0.2876	0.2896	0.2903	0.2852	0.2926	0.2829
(16)	**100**	**200**	**Mean**	0.5075	0.4937	0.5102	0.5141	0.4901	0.5026	0.5106	0.4974	0.5131	0.4999	0.4939	0.4916	0.4935	0.5159
			Std	0.2868	0.2916	0.2892	0.2869	0.2876	0.2879	0.2890	0.2827	0.2840	0.2885	0.2841	0.2897	0.2917	0.2887

ties with the core. The dependent variable, the real house price change, is constructed as $log(P_{it}/CPI_{it}) - log(P_{i,t-1}/CPI_{i,t-1})$, where P_{it} is the Freddie Mac House Price Index (FMHPI) and CPI_{it} is Consumer Price Index (CPI). Explanatory variables are the percentage changes in population (pop) and real per capita income (income). Both variables are calculated as first-differences of these variables expressed in natural logarithms and thus can be interpreted as growth rates. Due to taking first-differences, the actual number of time periods reduces to 159. The CD statistic of Pesaran (2015) on the residuals of several specifications investigated by Yang (2021) suggest that there might still be some remaining spatial dependence in the residuals, indicating the need to examine other spatial weight matrices, parameterize them, and to include spatial errors. The CD statistics is -6.365 when the SD model is applied in combination with the regional national common factors, individual fixed effects, seasonal dummies, and the row-normalized binary contiguity matrix with cut-off point at 100 miles (W100), which is the best performing spatial weight matrix in her study.

For this purpose, we first estimate the GNSAR and GNSMA models when each spatial lag in the model is pre-specified by the same spatial weight matrix: the W100 used in Yang (2021) and a non-parameterized negative exponential matrix (ED), respectively. We further estimate the GNSAR and GNSMA models when the negative exponential matrices are parameterized. Three cases are considered: (i) all spatial lags share the same estimated distance decay parameter (ED-same); (ii) the estimated distance decay parameters are different for the spatial lag in the dependent variable and the error tern, but not for the spatial lags in the explanatory variables (ED-one); and (iii) the estimated distance decay parameters are different for all spatial lags (ED-multi). All models include spatial and time fixed effects and all spatial weight matrices are row-normalized. The results are reported in Table 4.4 by which Columns (1) - (5) capture the results of GNSAR models and Columns (6) - (10) of the GNSMA models.

The estimated values of the spatial autoregressive parameter ρ are all positive and highly significant across the different models; they take values in the range (0.6, 0.8). This result indicates that house prices in each individual MSA are highly dependent on those in neighbouring MSAs. Population and income growth rates are found to have an upward effect on house prices. The magnitudes of the estim-

ated coefficients and the corresponding direct effects, which fluctuate around 0.3 (coefficient) and 0.4 (direct effects) for population, and below or around 0.1 (coefficient and direct effect) for income, are similar to those in Yang (2021). By contrast, the estimated coefficients of the spatial lags and the indirect effects of these two explanatory variables differ across the equations. The coefficient of the spatial lag of population tends to be positive and significant in most cases, except when adopting Yang's W100 matrix, while its indirect effect is positive and significant in all cases, even though the magnitude of this effect is quite different across the equations. The indirect effect of income is found to be positive and significant in all cases, while its magnitude seems to increase as more distance decay parameters are estimated separately (Columns (9) and (10)). Although the estimated coefficients of the spatial errors λ are all highly significant, they also differ in terms of signs and magnitudes. Note that due to the specification of spatial AR and MA errors in Equations (4.2) and (4.3), a positive spatial autocorrelation coefficient in a spatial AR error process corresponds with a negative spatial autocorrelation coefficient in a spatial MA error process. The estimated λ's take large negative values in the GNSMA models the moment that more distance decay parameters are estimated separately (Columns (9) and (10)). Importantly, the distance decay parameter of the spatial AR errors appears to be large and of the spatial MA errors to be small. This is line with the Fingleton and LeGallo's (2007) statement that spatial AR errors represent global shocks and spatial MA errors represent local shocks. Due to its multiple rounds structure, a shock reaches every corner of the research area even if the spatial range of AR errors is small, while a shock in MA errors due to its one round structure can only reach areas located farther away if the spatial range of MA errors is large. Overall, the estimated distance decay parameters appear to be significant, except for the distance decay parameter of income (α_2) of the GNSMA model in Column (1). Moreover, since the estimated values of the distance decay parameters are quite different, the conclusion must be that each spatial lag, as expected, exhibits a different distance decay pattern.

For both the GNSAR and GNSMA models, the log-likelihood values increase when the non-parameterized ED matrix replaces the binary contiguity matrices with a cut-off point (W100) used by Yang (2021). Importantly, the log-likelihood values further increase when the ED matrix is parameterized, also if all spatial lags

share the same parameterized ED matrix. However, the log-likelihood values reach their maximum when the estimated distance decay parameters are assumed to be different for each spatial lag. Eventually, the log-likelihood value of the GNSAR model (193194.3107), reported in Column (6), turns out to be the largest across all models being considered. For this reason, the GNSAR model with different spatial weight matrices for each spatial lag may be labelled as the best performing model. Population shows a strong distance decay effect within this model ($\alpha_1 = 4.6185$). This finding implies that the spatial range of population, the distance at which it still has effect, is limited. The spatial AR error shows a milder distance decay effect ($\lambda = 2.2365$), while the distance decay effects exhibited by income and the house price change in neighbouring units are relatively small ($\alpha_2 = 0.9608$ and $\delta = 0.4893$). According to the estimated direct and indirect effects reported in Column (5), a one percentage point increase in population growth in one particular MSA is predicted to lead to an average 0.3479 percentage point increase in house prices in the own MSA, and a 2.2652 percentage point increase in house prices in neighbouring MSAs. By contrast, a one percentage point increase in income growth has much smaller direct and indirect effects on house prices, respectively 0.0912 and 0.7179 percentage points. The estimated indirect effects are larger than those in Yang (2021). There are two possible explanations for this finding. The larger indirect effects in our study might be due to modelling strong cross-sectional dependence by time fixed effects rather than the more general common factors considered in Yang (2021). On the other hand, there is evidence of mild (or weak) cross-sectional dependence captured by the spatial AR error in that the spatial autocorrelation coefficient is strongly significant ($\lambda = 0.5286$ is significant at the 1% significance level). This has been ignored in Yang (2021). The inclusion of the spatial error is important since an improperly specified error term may result in misleading conclusions concerning the statistical significance of the estimated parameters. The results in this study favour a spatial AR error process, which is in line with the complexity of the US housing market. This market has gone through nationwide economic shocks and housing-related policy changes, as a result of which geographic interactions and spillovers of shocks between cities influence the co-movement of US house price changes (Choi and Hansz, 2021). Since such shocks are global and affect all MSAs in the US, they are more likely to diffuse according

to a multiple rounds structure among the units.

Table 4.4. Estimation results explaining US house price changes

	GNSAR					GNSMA				
	W100	ED	ED_same	ED_one	ED_multi	W100	ED	ED_same	ED_one	ED_multi
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ρ	0.7891***	0.6607***	0.8179***	0.8120***	0.8104***	0.8127***	0.8060***	0.8280***	0.6353***	0.6348***
	(299.61)	(26.88)	(26.45)	(110.43)	(108.96)	(190.68)	(160.30)	(182.41)	(60.55)	(63.66)
pop	0.3061***	0.3015***	0.2749***	0.3145***	0.3126***	0.3014***	0.2838***	0.2727***	0.2839***	0.2839***
	(32.31)	(33.16)	(23.61)	(32.46)	(32.08)	(31.8)	(31.10)	(30.51)	(31.17)	(31.68)
income	0.0793***	0.0785***	0.0743***	0.0824***	0.0809***	0.0783***	0.0752***	0.0739***	0.0742***	0.0767***
	(19.16)	(19.94)	(18.48)	(20.49)	(19.51)	(18.95)	(19.25)	(19.12)	(19.50)	(29.45)
W*pop	-0.0011	0.3640***	0.1118**	0.2116***	0.1828***	-0.0244*	0.2095***	0.0933***	1.5353***	1.3106***
	(-0.08)	(14.14)	(1.68)	(6.02)	(9.07)	(-1.68)	(7.91)	(4.63)	(5.67)	(5.71)
W*income	-0.0115**	0.0503***	0.0001	0.0349***	0.0725***	-0.0170***	0.0185**	-0.0038	0.2982***	0.9815
	(-2.12)	(4.67)	(0.00)	(4.38)	(3.39)	(-3.22)	(1.99)	(-0.53)	(3.83)	(1.29)
λ	-0.2728***	0.5702***	0.1129	0.5262***	0.5286***	0.3052***	-0.3921***	-0.0676***	-2.5645***	-2.5173***
	(-36.32)	(19.73)	(0.72)	(42.58)	(42.08)	(26.68)	(-30.01)	(-3.23)	(-15.37)	(-16.64)
δ/α			1.3387***	0.4911***	0.4893***			1.3744***	1.9141***	1.9139***
			(9.00)	(60.59)	(60.49)			(54.60)	(53.21)	(379.32)
α_1/α				3.4803***	4.6185***				0.1599***	0.1930***
				(2.83)	(4.94)				(5.95)	(7.77)
α_2					0.9608***					0.0546
					(2.68)					(1.43)
γ				2.2455***	2.2365***				0.1990***	0.2019***
				(54.87)	(53.4)				(19.15)	(26.14)
di_pop	0.4213***	0.3591***	0.3625***	0.3530***	0.3479***	0.4198***	0.3587***	0.3624***	0.3565***	0.3575***
	(36.22)	(32.20)	(32.52)	(31.10)	(31.53)	(36.10)	(32.22)	(32.52)	(32.53)	(32.72)
di_income	0.1037***	0.0889***	0.0901***	0.0903***	0.0912***	0.1034***	0.0887***	0.0901***	0.0908***	0.0958***
	(22.52)	(20.17)	(20.51)	(20.33)	(20.5)	(22.48)	(20.22)	(20.52)	(20.53)	(40.62)
in_pop	0.9790***	1.6024***	1.7615***	2.4452***	2.2652***	1.0126***	2.1840***	1.7652***	4.6320***	4.0089***
	(18.69)	(26.49)	(16.08)	(10.86)	(13.79)	(18.70)	(16.21)	(16.74)	(6.28)	(7.26)
in_income	0.2080***	0.2906***	0.3182***	0.5332***	0.7179***	0.2143***	0.3944***	0.3177***	0.9303***	2.8022
	(11.98)	(9.37)	(8.75)	(9.91)	(6.38)	(12.01)	(8.68)	(8.96)	(4.38)	(1.42)
logL	188129.3224	192625.4966	192739.5662	193191.73	193194.3107	188094.7209	192553.6504	192736.1782	193065.7505	193069.7571
rsqr	0.7794	0.7543	0.7974	0.7280	0.7274	0.7810	0.7808	0.7998	0.7783	0.7786

* , **, and *** denote significance at the 10%, 5%, and 1% level. t-statistics are reported in parentheses.

4.5 Conclusion and discussion

We consider GNS models with individual and time fixed effects, a spatial autoregressive or spatial moving average error process, and parameterized spatial weight matrices that are different for each spatial lag in the model. These spatial weight matrices, which are specified as exponential distance decay matrices, add more flexibility to the coefficients of the spatial lags of the explanatory variables in the model and the direct and indirect effect estimates derived from them. They also help to reduce the identification problems that plagued the empirical literature estimating GNS models.

Consistent and properly centred QML estimators, based on the orthogonal transformation to concentrate out the individual and time fixed effects, are derived to estimate the distance decay parameters jointly with the response parameters in the model. The assumptions necessary for asymptotic normality, identification and consistency have been taken from previous studies, and adapted and extended to the model at hand. In our Monte Carlo experiment we examine the performance of the estimators of the response parameters, distance decay parameters, and direct and indirect effects. An empirical application using US house data taken from Yang (2021) shows that the GNS model with spatial AR errors and different distance decay parameters for different spatial lags outperforms all the other empirical models considered in this chapter.

Literature on the selection of spatial econometric models with either AR or MA errors is scarce. The reasons are that these two models are not nested and models with MA errors are computationally more difficult to estimate, as shown in Chapter 3. Bayesian posterior model probabilities (LeSage, 2014) or entropy measures (Herrera et al., 2019), in addition to log-likelihood values, may be a useful tool to compare the performance of these kind of non-nested models in future research.

4.A Appendix: The GNSAR model with both fixed effects

Without loss of generality, the presentation with respect to the distance decay parameters in Appendix 4.A and 4.B is limited to two explanatory variables ($K = 2$) with distance decay parameters α_1 and α_2.

$$\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^*)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) \\ \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) - (T-1)\mathrm{tr}\boldsymbol{J}_N\boldsymbol{G} \\ \frac{1}{\sigma^2}\sum_{t=1}^{T}[\boldsymbol{H}\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi})]'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) - (T-1)\mathrm{tr}\boldsymbol{J}_N\boldsymbol{H} \\ \frac{1}{2\sigma^4}\sum_{t=1}^{T}(\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\xi})\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) - (N-1)\frac{T-1}{T}\sigma^2) \\ -\rho(T-1)\mathrm{tr}[\boldsymbol{\Delta}(\delta)\boldsymbol{S}^{-1}] + \frac{\rho}{\sigma^2}\sum_{t-1}^{T}(\boldsymbol{R}\boldsymbol{\Delta}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) \\ \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{\Delta}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) \\ \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{\Delta}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) \\ -\lambda(T-1)\mathrm{tr}[\boldsymbol{\Delta}(\gamma)\boldsymbol{R}^{-1}] + \frac{\lambda}{\sigma^2}\sum_{t=1}^{T}[\boldsymbol{\Delta}(\gamma)(\boldsymbol{S}\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta})]'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) \end{pmatrix}, \tag{4.A.1}$$

where $\boldsymbol{G} = \boldsymbol{W}(\delta)\boldsymbol{S}^{-1}$ and $\boldsymbol{H} = \boldsymbol{W}(\gamma)\boldsymbol{R}^{-1}$. We also denote $\ddot{\boldsymbol{G}} = \boldsymbol{R}\boldsymbol{G}\boldsymbol{R}^{-1}$, $\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1} = \boldsymbol{\Delta}$, $\ddot{\boldsymbol{\Delta}} = \boldsymbol{R}\boldsymbol{\Delta}\boldsymbol{R}^{-1}$, and $\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1} = \boldsymbol{\Lambda}$. $\boldsymbol{Z}(\varrho) = \frac{\partial \boldsymbol{W}(\varrho)}{\partial \varrho}$ and $\boldsymbol{\Gamma}(\varrho) = \frac{\partial^2 \boldsymbol{W}(\varrho)}{\partial \varrho^2}$.

$$
-\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} =
$$

$$
\begin{pmatrix}
\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^*)'\boldsymbol{J}_N\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^* & * & * & * & * & * & * \\
\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^* & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t & * & * & * & * & * \\
 & +(T-1)\mathrm{tr}[\boldsymbol{J}_N\boldsymbol{G}^2] & & & & & \\
\frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\xi})\boldsymbol{J}_N\boldsymbol{W}(\gamma)\widetilde{\boldsymbol{X}}_t^* & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{H}\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{H}\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}))'\boldsymbol{J}_N\boldsymbol{H}\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) & 0 & * & * & * \\
+\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{H}\widetilde{\boldsymbol{v}}(\boldsymbol{\xi}))'\boldsymbol{J}_N\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^* & +\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{W}(\gamma)\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) & +(T-1)\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{H}^2) & & & & \\
\frac{1}{\sigma^4}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\xi})\boldsymbol{J}_N\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^* & \frac{1}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) & \frac{1}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{H}\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}))'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) & 0 & 0 & 0 & 0 \\
\frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^* & (T-1)\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1})+\rho(T-1)\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{G}\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}) & \frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{W}(\gamma)\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t & 0 & 0 & 0 & 0 \\
 & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) & \frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{H}\widetilde{\boldsymbol{v}}_t)'\boldsymbol{J}_N\boldsymbol{R}\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t & & & & \\
\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1)'\boldsymbol{J}_N\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^* & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1 & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{W}(\gamma)\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t & 0 & 0 & 0 & 0 \\
-\frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\xi})\boldsymbol{J}_N\boldsymbol{R}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t} & & +\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{H}\widetilde{\boldsymbol{v}}_t)'\boldsymbol{J}_N\boldsymbol{R}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1 & & & & \\
\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^* & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2 & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{W}(\gamma)\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t & 0 & 0 & 0 & 0 \\
-\frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\xi})\boldsymbol{J}_N\boldsymbol{R}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t} & & +\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{W}(\gamma)\boldsymbol{R}^{-1}\widetilde{\boldsymbol{v}}_t)'\boldsymbol{J}_N\boldsymbol{R}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2 & & & & \\
\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\xi})\boldsymbol{J}_N\boldsymbol{Z}(\gamma)\widetilde{\boldsymbol{X}}_t^* & \frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Z}(\gamma)\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t & (T-1)\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1})+\lambda(T-1)\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{H}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}) & 0 & 0 & 0 & 0 \\
+\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}))'\boldsymbol{J}_N\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^* & +\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\widetilde{\boldsymbol{v}}_t & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\widetilde{\boldsymbol{v}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t & & & & \\
\end{pmatrix}
$$

$$
\begin{pmatrix}
\mathbf{0}_{K\times K} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & -\frac{(N-1)(T-1)}{2\sigma^4}+\frac{1}{\sigma^6}\tilde{v}_t'(\xi)I_N\tilde{v}_t & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & +\frac{\rho}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{RW}(\delta)\tilde{y}_t)'I_N\boldsymbol{RZ}(\delta)\tilde{y}_t & 0 & \frac{\rho}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{RZ}(\delta)\tilde{y}_t)'I_N\tilde{v}_t & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & \frac{1}{\sigma^4}\sum_{t=1}^{T}\tilde{v}_t'I_N\boldsymbol{RZ}(\alpha_1)\tilde{x}_{1t}\phi_1 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & \frac{1}{\sigma^4}\sum_{t=1}^{T}\tilde{v}_t'I_N\boldsymbol{RZ}(\alpha_2)\tilde{x}_{2t}\phi_2 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & +\frac{\lambda}{\sigma^2}(\boldsymbol{W}(\gamma)\boldsymbol{R}^{-1}\tilde{v}_t)'I_N\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\tilde{v}_t & \frac{\lambda}{\sigma^4}\sum_{t=1}^{T}\tilde{v}_t'I_N\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\tilde{v}_t & 0 & 0 & 0
\end{pmatrix}
+
\begin{pmatrix}
\mathbf{0}_{K\times K} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & \rho(T-1)\mathrm{tr}(\boldsymbol{\Gamma}(\delta)\boldsymbol{S}^{-1})+\rho^2(T-1)\mathrm{tr}(\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}) & 0 & 0 & 0 \\
 & & & & -\frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R\Gamma}(\delta)\tilde{y}_t)'I_N\tilde{v}_t & & & \\
 & & & & +\frac{\rho^2}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{RZ}(\delta)\tilde{y}_t)'I_N\boldsymbol{RZ}(\delta)\tilde{y}_t & & & \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & \frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{RZ}(\delta)\tilde{y}_t)'I_N\boldsymbol{RZ}(\alpha_1)\tilde{x}_{1t}\phi_1 & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R\Gamma}(\alpha_1)\tilde{x}_{1t}\phi_1)'I_N\tilde{v}_t & 0 & 0 \\
 & & & & & +\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{RZ}(\alpha_1)\tilde{x}_{1t}\phi_1)'I_N\boldsymbol{RZ}(\alpha_1)\tilde{x}_{1t}\phi_1 & & \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & \frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{RZ}(\delta)\tilde{y}_t)'I_N\boldsymbol{RZ}(\alpha_2)\tilde{x}_{2t}\phi_2 & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{RZ}(\alpha_1)\tilde{x}_{1t}\phi_1)'I_N\boldsymbol{RZ}(\alpha_2)\tilde{x}_{2t}\phi_2 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & \frac{\rho\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Z}(\gamma)\boldsymbol{Z}(\delta)\tilde{y}_t)'I_N\tilde{v}_t & \frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Z}(\gamma)\boldsymbol{Z}(\alpha_1)\tilde{x}_{1t}\phi_1)'I_N\tilde{v}_t & 0 & 0 \\
 & & & & +\frac{\rho\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{RZ}(\delta)\tilde{y}_t)'I_N\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\tilde{v}_t & +\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{RZ}(\alpha_1)\tilde{x}_{1t}\phi_1)'I_N\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\tilde{v}_t & &
\end{pmatrix}
$$

$$
+\begin{pmatrix}
\mathbf{0}_{K\times K} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & \begin{matrix} -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{\Gamma}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t \\ +\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\boldsymbol{R}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2 \end{matrix} & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & \begin{matrix} \frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Z}(\gamma)\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t \\ +\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\widetilde{\boldsymbol{v}}_t \end{matrix} & \begin{matrix} \lambda(T-1)\mathrm{tr}(\boldsymbol{\Gamma}(\gamma)\boldsymbol{R}^{-1}) + \lambda^2(T-1)\mathrm{tr}(\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}) \\ -\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{\Gamma}(\gamma)\boldsymbol{R}^{-1}\widetilde{\boldsymbol{v}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t \\ +\frac{\lambda^2}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\widetilde{\boldsymbol{v}}_t)'\boldsymbol{J}_N\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\widetilde{\boldsymbol{v}}_t \end{matrix}
\end{pmatrix}. \tag{4.A.2}
$$

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \mathcal{H} + \begin{pmatrix} \mathbf{0}_{2K\times 2K} & * & * & * & * & * & \\ \mathbf{0}_{1\times 2K} & \frac{1}{N-1}\mathrm{tr}\ddot{\boldsymbol{G}}^s\boldsymbol{J}_N\ddot{\boldsymbol{G}} & * & * & * & * & * \\ \mathbf{0}_{1\times 2K} & \frac{1}{N-1}\mathrm{tr}\boldsymbol{H}^s\boldsymbol{J}_N\ddot{\boldsymbol{G}} & \frac{1}{N-1}\mathrm{tr}\boldsymbol{H}^s\boldsymbol{J}_N\boldsymbol{H} & * & * & * & * \\ \mathbf{0}_{1\times K} & \frac{1}{\sigma^2(N-1)}\mathrm{tr}(\boldsymbol{J}_N\ddot{\boldsymbol{G}}) & \frac{1}{\sigma^2(N-1)}\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{H}) & \frac{1}{2\sigma^4} & * & * & * \\ \mathbf{0}_{1\times 2K} & \frac{\rho}{N-1}\mathrm{tr}\ddot{\boldsymbol{\Delta}}^s\boldsymbol{J}_N\ddot{\boldsymbol{G}} & \frac{\rho}{N-1}\mathrm{tr}\ddot{\boldsymbol{\Delta}}^s\boldsymbol{J}_N\boldsymbol{H} & \frac{\rho}{\sigma^2(N-1)}\mathrm{tr}(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}}) & \frac{\rho^2}{N-1}\mathrm{tr}\ddot{\boldsymbol{\Delta}}^s\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}} & * & * \\ \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times 1} & 0 \\ \mathbf{0}_{1\times 2K} & \frac{\lambda}{N-1}\mathrm{tr}\boldsymbol{\Lambda}^s\boldsymbol{J}_N\ddot{\boldsymbol{G}} & \frac{\lambda}{N-1}\mathrm{tr}\boldsymbol{\Lambda}^s\boldsymbol{J}_N\boldsymbol{H} & \frac{\lambda}{\sigma^2(N-1)}\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{\Lambda}) & \frac{\rho\lambda}{N-1}\mathrm{tr}\boldsymbol{\Lambda}^s\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}} & \mathbf{0}_{K\times 1} & \frac{\lambda^2}{N-1}\mathrm{tr}\boldsymbol{\Lambda}^s\boldsymbol{J}_N\boldsymbol{\Lambda} \end{pmatrix}, \tag{4.A.3}$$

where $\mathcal{H} = \frac{1}{(N-1)(T-1)}\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^*, \boldsymbol{R}\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}, \mathbf{0}_{K\times 1}, \mathbf{0}_{K\times 1}, \rho\boldsymbol{R}\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}, \boldsymbol{R}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1, \boldsymbol{R}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2, \mathbf{0}_{K\times 1})'\boldsymbol{J}_N$
$\times(\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^*, \boldsymbol{R}\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}, \mathbf{0}_{K\times 1}, \mathbf{0}_{K\times 1}, \rho\boldsymbol{R}\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}, \boldsymbol{R}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1, \boldsymbol{R}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2, \mathbf{0}_{K\times 1})$

$$\boldsymbol{\Omega}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{0}_{2K\times 2K} & * & * & * & * & * & * \\ \mathbf{0}_{1\times 2K} & \frac{1}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{G}})_{ii}]^2 & * & * & * & ** & \\ \mathbf{0}_{1\times 2K} & \frac{1}{n}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{G}})_{ii}][(\boldsymbol{J}_N\boldsymbol{H})_{ii}] & \frac{1}{n}\sum_{i=1}^{N}[(\boldsymbol{J}_N\boldsymbol{H})_{ii}]^2 & * & * & * & \\ \mathbf{0}_{1\times 2K} & \frac{1}{2\sigma^2 n}\mathrm{tr}(\boldsymbol{J}_N\ddot{\boldsymbol{G}}) & \frac{1}{2\sigma^2 n}\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{H}) & \frac{1}{4\sigma^4} & * & ** & \\ \mathbf{0}_{1\times 2K} & \frac{\rho}{n}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{G}})_{ii}][(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}})_{ii}] & \frac{\rho}{n}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}})_{ii}][(\boldsymbol{J}_N\boldsymbol{H})_{ii}] & \frac{\rho}{2\sigma^2 n}\mathrm{tr}(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}}) & \frac{\rho^2}{n}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}})_{ii}]^2 & ** & \\ \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{0}_{K\times 1} \\ \mathbf{0}_{1\times 2K} & \frac{\lambda}{n}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{G}})_{ii}][(\boldsymbol{J}_N\boldsymbol{\Lambda})_{ii}] & \frac{\rho}{n}\sum_{i=1}^{N}[(\boldsymbol{J}_N\boldsymbol{\Lambda})_{ii}][(\boldsymbol{J}_N\boldsymbol{H})_{ii}] & \frac{\lambda}{2\sigma^2 n}\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{\Lambda}) & \frac{\rho\lambda}{n}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}})_{ii}][(\boldsymbol{J}_N\boldsymbol{\Lambda})_{ii}] & \mathbf{0}_{K\times 1} & \frac{\lambda^2}{n}\sum_{i=1}^{N}[(\boldsymbol{J}_N\boldsymbol{\Lambda})_{ii}]^2 \end{pmatrix}. \tag{4.A.4}$$

4.B Appendix: The GNSMA model with both fixed effects

$$\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^*)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) \\ \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) - (T-1)\mathrm{tr}\boldsymbol{J}_N\boldsymbol{G} \\ -\frac{1}{\sigma^2}\sum_{t=1}^{T}[\boldsymbol{Q}\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi})]'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) + (T-1)\mathrm{tr}\boldsymbol{J}_N\boldsymbol{H} \\ \frac{1}{2\sigma^4}\sum_{t=1}^{T}(\widetilde{\boldsymbol{v}}_t'(\boldsymbol{\xi})\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) - (N-1)\frac{T-1}{T}\sigma^2) \\ -\rho(T-1)\mathrm{tr}[\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}] + \frac{\rho}{\sigma^2}\sum_{t-1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) \\ \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) \\ \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) \\ \lambda(T-1)\mathrm{tr}[\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}] - \frac{\lambda}{\sigma^2}\sum_{t=1}^{T}[\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}(\boldsymbol{S}\widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta})]'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t(\boldsymbol{\xi}) \end{pmatrix}, \tag{4.B.1}$$

where $\boldsymbol{G} = \boldsymbol{W}(\delta)\boldsymbol{S}^{-1}$ and $\boldsymbol{Q} = \boldsymbol{R}^{-1}\boldsymbol{W}(\gamma)$. We also denote $\ddot{\boldsymbol{G}} = \boldsymbol{R}^{-1}\boldsymbol{G}\boldsymbol{R}$, $\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1} = \boldsymbol{\Delta}$, $\ddot{\boldsymbol{\Delta}} = \boldsymbol{R}^{-1}\boldsymbol{\Delta}\boldsymbol{R}$, and $\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma) = \boldsymbol{\Lambda}$.

$$
-\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} =
$$

$$
\begin{pmatrix}
\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^*)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^* & * & * & * & * \\
\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}\widetilde{\boldsymbol{X}}_t^* & \begin{matrix}\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t \\ +(T-1)\mathrm{tr}[\boldsymbol{J}_N\boldsymbol{G}^2]\end{matrix} & * & * & * \\
-\frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{v}_t'(\xi)\boldsymbol{J}_N\boldsymbol{Q}\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^* & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{Q}\widetilde{v}_t(\xi) & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Q}\widetilde{v}_t(\xi))'\boldsymbol{J}_N\boldsymbol{Q}\widetilde{v}_t(\xi) & 0 & * & * \\
-\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Q}\widetilde{v}(\xi))'\boldsymbol{J}_N\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^* & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{v}_t(\xi) & \frac{2}{\sigma^2}\sum_{t=1}^{T-1}(\boldsymbol{Q}\boldsymbol{Q}\widetilde{v}_t)'\boldsymbol{J}_N\widetilde{v}_t-(T-1)\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{H}^2) & & & \\
\frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{v}_t'(\xi)\boldsymbol{J}_N\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^* & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{v}_t(\xi) & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Q}\widetilde{v}_t(\xi))'\boldsymbol{J}_N\widetilde{v}_t(\xi) & 0 & 0 & 0 \\
\frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^* & \begin{matrix}(T-1)\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1})+\rho(T-1)\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{G}\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}) \\ -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{v}_t(\xi)\end{matrix} & -\frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{v}_t & 0 & 0 & 0 \\
 & & -\frac{\rho}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Q}\widetilde{v}_t)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\widetilde{\boldsymbol{y}}_t & & & \\
\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^* & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1 & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1)'\boldsymbol{J}_N\widetilde{v}_t & 0 & 0 & 0 \\
-\frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{v}_t'(\xi)\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t} & & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Q}\widetilde{v}_t)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1 & & & \\
\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^* & \frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2 & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\widetilde{v}_t & 0 & 0 & 0 \\
-\frac{1}{\sigma^2}\sum_{t=1}^{T}\widetilde{v}_t'(\xi)\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t} & & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{Q}\widetilde{v}_t)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2 & & & \\
-\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}\widetilde{v}_t'(\xi)\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^* & -\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\widetilde{v}_t & -(T-1)\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1})-\lambda(T-1)\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{H}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}) & 0 & 0 & 0 \\
-\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\widetilde{v}_t(\xi))'\boldsymbol{J}_N\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^* & -\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{W}(\delta)\widetilde{\boldsymbol{y}}_t)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\widetilde{v}_t & \begin{matrix}+\frac{1}{\sigma^2}\sum_{t=1}^{T}[(\lambda\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\boldsymbol{Q}+\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma))\widetilde{v}_t]'\boldsymbol{J}_N\widetilde{v}_t \\ \end{matrix} & & & \\
\end{pmatrix}
$$

$$
\begin{pmatrix}
\mathbf{0}_{K\times K} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & -\frac{(N-1)(T-1)}{2\sigma^4} + \frac{1}{\sigma^6}\tilde{v}_t'(\xi)I_N\tilde{v}_t & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & +\frac{\rho}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{W}(\phi)\tilde{y}_t)'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\tilde{y}_t & 0 & \frac{\rho}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\tilde{y}_t)'I_N\tilde{v}_t & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & \frac{1}{\sigma^4}\sum_{t=1}^{T}\tilde{v}_t'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\tilde{x}_{1t}\phi_1 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & \frac{1}{\sigma^4}\sum_{t=1}^{T}\tilde{v}_t'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\tilde{x}_{2t}\phi_2 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & +\frac{1}{\sigma^6}\sum_{t=1}^{T}\tilde{v}_t'(\lambda\boldsymbol{Q}\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\tilde{v}_t)'I_N\tilde{v}_t & -\frac{1}{\sigma^4}\sum_{t=1}^{T}\tilde{v}_t'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\tilde{v}_t & 0 & 0 & 0 & 0 \\
 & & +\frac{1}{\sigma^6}(\boldsymbol{R}^{-1}\boldsymbol{W}(\gamma)\tilde{v}_t)'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\tilde{v}_t & & & & & \\
+ & & & & & & & \\
\begin{pmatrix}
\mathbf{0}_{K\times K} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & \rho(T-1)\mathrm{tr}(\boldsymbol{\Gamma}(\phi)S^{-1}) + \rho^2(T-1)\mathrm{tr}(\boldsymbol{Z}(\delta)S^{-1}\boldsymbol{Z}(\delta)S^{-1}) & 0 & 0 & 0 \\
 & & & & -\frac{\rho}{\sigma^6}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{\Gamma}(\phi)\tilde{y}_t)'I_N\tilde{v}_t & & & \\
 & & & & +\frac{\rho^2}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\tilde{y}_t)'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\tilde{y}_t & & & \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & \frac{\rho}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\tilde{y}_t)'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\tilde{x}_{1t}\phi_1 & -\frac{1}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{\Gamma}(\alpha_1)\tilde{x}_{1t}\phi_1)'I_N\tilde{v}_t & 0 & 0 \\
 & & & & & +\frac{1}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\tilde{x}_{1t}\phi_1)'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\tilde{x}_{1t}\phi_1 & & \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & -\frac{\rho}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\tilde{y}_t)'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\tilde{x}_{2t}\phi_2 & \frac{1}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\tilde{x}_{1t}\phi_1)'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\tilde{x}_{2t}\phi_2 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & -\frac{\rho\lambda}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\tilde{y}_t)'I_N\tilde{v}_t & -\frac{\lambda}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\tilde{x}_{1t}\phi_1)'I_N\tilde{v}_t & 0 & 0 \\
 & & & & -\frac{\rho\lambda}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\tilde{y}_t)'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\tilde{v}_t & -\frac{\lambda}{\sigma^4}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\tilde{x}_{1t}\phi_1)'I_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\tilde{v}_t & & \\
\end{pmatrix}
\end{pmatrix}
$$

$$+\begin{pmatrix}
\mathbf{0}_{K\times K} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & \begin{matrix} -\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{\Gamma}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t \\ +\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2 \end{matrix} & 0 \\
\mathbf{0}_{1\times K} & 0 & 0 & 0 & 0 & 0 & \begin{matrix} -\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t \\ -\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\widetilde{\boldsymbol{v}}_t \end{matrix} & \begin{matrix} -\lambda(T-1)\mathrm{tr}(\boldsymbol{\Gamma}(\gamma)\boldsymbol{R}^{-1}) - \lambda^2(T-1)\mathrm{tr}(\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}) \\ +\frac{\lambda}{\sigma^2}\sum_{t=1}^{T}(2\lambda\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\widetilde{\boldsymbol{v}}_t + \boldsymbol{R}^{-1}\boldsymbol{\Gamma}(\gamma)\widetilde{\boldsymbol{v}}_t)'\boldsymbol{J}_N\widetilde{\boldsymbol{v}}_t \\ +\frac{\lambda^2}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\widetilde{\boldsymbol{v}}_t)'\boldsymbol{J}_N\boldsymbol{R}^{-1}\boldsymbol{Z}(\gamma)\widetilde{\boldsymbol{v}}_t \end{matrix}
\end{pmatrix}. \tag{4.B.2}$$

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \mathcal{H} + \begin{pmatrix} \mathbf{0}_{2K\times 2K} & * & * & * & * & * & \\ \mathbf{0}_{1\times 2K} & \frac{1}{N-1}\mathrm{tr}\ddot{\boldsymbol{G}}^s\boldsymbol{J}_N\ddot{\boldsymbol{G}} & * & * & * & * & * \\ \mathbf{0}_{1\times 2K} & -\frac{1}{N-1}\mathrm{tr}\boldsymbol{Q}^s\boldsymbol{J}_N\ddot{\boldsymbol{G}} & \frac{1}{N-1}\mathrm{tr}\boldsymbol{Q}^s\boldsymbol{J}_N\boldsymbol{Q} & * & * & * & * \\ \mathbf{0}_{1\times 2K} & \frac{1}{\sigma^2(N-1)}\mathrm{tr}(\boldsymbol{J}_N\ddot{\boldsymbol{G}}) & -\frac{1}{\sigma^2(N-1)}\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{Q}) & \frac{1}{2\sigma^4} & * & * & * \\ \mathbf{0}_{1\times 2K} & \frac{\rho}{N-1}\mathrm{tr}\ddot{\boldsymbol{\Delta}}^s\boldsymbol{J}_N\ddot{\boldsymbol{G}} & -\frac{\rho}{N-1}\mathrm{tr}\ddot{\boldsymbol{\Delta}}^s\boldsymbol{J}_N\boldsymbol{Q} & \frac{\rho}{\sigma^2(N-1)}\mathrm{tr}(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}}) & \frac{\rho^2}{N-1}\mathrm{tr}\ddot{\boldsymbol{\Delta}}^s\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}} & * & * \\ \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times K} & \mathbf{0}_{K\times 1} & 0 \\ \mathbf{0}_{1\times 2K} & -\frac{\lambda}{N-1}\mathrm{tr}\boldsymbol{\Lambda}^s\boldsymbol{J}_N\ddot{\boldsymbol{G}} & \frac{\lambda}{N-1}\mathrm{tr}\boldsymbol{\Lambda}^s\boldsymbol{J}_N\boldsymbol{Q} & -\frac{\lambda}{\sigma^2(N-1)}\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{\Lambda}) & -\frac{\rho\lambda}{N-1}\mathrm{tr}\boldsymbol{\Lambda}^s\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}} & \mathbf{0}_{K\times 1} & \frac{\lambda^2}{N-1}\mathrm{tr}\boldsymbol{\Lambda}^s\boldsymbol{J}_N\boldsymbol{\Lambda} \end{pmatrix}, \tag{4.B.3}$$

where $\mathcal{H}_{n,T} = \frac{1}{(N-1)(T-1)}\frac{1}{\sigma^2}\sum_{t=1}^{T}(\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^*, \boldsymbol{R}^{-1}\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}, \mathbf{0}_{K\times 1}, \mathbf{0}_{K\times 1}, \rho\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}, \boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1, \boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2, \mathbf{0}_{K\times 1})'\boldsymbol{J}_N$
$\times(\boldsymbol{R}^{-1}\widetilde{\boldsymbol{X}}_t^*, \boldsymbol{R}^{-1}\boldsymbol{G}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}, \mathbf{0}_{K\times 1}, \mathbf{0}_{K\times 1}, \rho\boldsymbol{R}^{-1}\boldsymbol{Z}(\delta)\boldsymbol{S}^{-1}\widetilde{\boldsymbol{X}}_t^*\boldsymbol{\zeta}, \boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_1)\widetilde{\boldsymbol{x}}_{1t}\phi_1, \boldsymbol{R}^{-1}\boldsymbol{Z}(\alpha_2)\widetilde{\boldsymbol{x}}_{2t}\phi_2, \mathbf{0}_{K\times 1})$

$$\boldsymbol{\Omega}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{0}_{2K\times 2K} & * & * & * & * & * & * \\ \mathbf{0}_{1\times 2K} & \frac{1}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{G}})_{ii}]^2 & * & * & * & ** & \\ \mathbf{0}_{1\times 2K} & -\frac{1}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{G}})_{ii}][(\boldsymbol{J}_N\boldsymbol{Q})_{ii}] & \frac{1}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\boldsymbol{Q})_{ii}]^2 & * & * & * & \\ \mathbf{0}_{1\times 2K} & \frac{1}{2\sigma^2 n}\mathrm{tr}(\boldsymbol{J}_N\ddot{\boldsymbol{G}}) & -\frac{1}{2\sigma^2 n}\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{Q}) & \frac{1}{4\sigma^4} & * & ** & \\ \mathbf{0}_{1\times 2K} & \frac{\rho}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{G}})_{ii}][(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}})_{ii}] & -\frac{\rho}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}})_{ii}][(\boldsymbol{J}_N\boldsymbol{Q})_{ii}] & \frac{\rho}{2\sigma^2 n}\mathrm{tr}(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}}) & \frac{\rho^2}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}})_{ii}]^2 & ** & \\ \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times 1} & \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{0}_{K\times 1} \\ \mathbf{0}_{1\times 2K} & -\frac{\lambda}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{G}})_{ii}][(\boldsymbol{J}_N\boldsymbol{\Lambda})_{ii}] & \frac{\rho}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\boldsymbol{\Lambda})_{ii}][(\boldsymbol{J}_N\boldsymbol{H})_{ii}] & -\frac{\lambda}{2\sigma^2 n}\mathrm{tr}(\boldsymbol{J}_N\boldsymbol{\Lambda}) & -\frac{\rho\lambda}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\ddot{\boldsymbol{\Delta}})_{ii}][(\boldsymbol{J}_N\boldsymbol{\Lambda})_{ii}] & \mathbf{0}_{K\times 1} & \frac{\lambda^2}{N}\sum_{i=1}^{N}[(\boldsymbol{J}_N\boldsymbol{\Lambda})_{ii}]^2 \end{pmatrix}. \tag{4.B.4}$$

4.C Appendix: Proofs of Lemma 4.2

We provide proofs of Lemma 4.2, since Lemma 4.1 is a special case of Lemma 4.2 with $\boldsymbol{W}(\delta_0) = \boldsymbol{W}(\alpha_{0k}) = \boldsymbol{W}$ for each k.

For the GNSAR model, we are interested in the sufficient condition that

$$\begin{aligned}\boldsymbol{J}_N\boldsymbol{R}(\lambda,\gamma)\boldsymbol{G}(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0 + \sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k})c_1 + \boldsymbol{J}_N\boldsymbol{R}(\lambda,\gamma)\widetilde{\boldsymbol{X}}_t c_2 \\ + \boldsymbol{J}_N\boldsymbol{R}(\lambda,\gamma)(\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt})c_3 = 0, \qquad (4.C.1)\end{aligned}$$

for $c_1 = c_2 = c_3 = 0$, where c_1, c_2, and c_3 are conformable scalars. This is analogous to Equation (G.1) in the proof of Lemma 5.2 in Lee et al. (2010) for the social interaction models. Denote $\kappa_1 = \frac{1}{N}\boldsymbol{\iota}_N'\boldsymbol{R}(\lambda,\gamma)\boldsymbol{G}(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0 + \sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k})$, $\kappa_2 = \frac{1}{N}\boldsymbol{\iota}_N'\boldsymbol{R}(\lambda,\gamma)\widetilde{\boldsymbol{X}}_t$, and $\kappa_3 = \frac{1}{N}\boldsymbol{\iota}_N'\boldsymbol{R}(\lambda,\gamma)[\boldsymbol{W}(\alpha_{0,1})\widetilde{\boldsymbol{x}}_{1t},...,\boldsymbol{W}(\alpha_{0K})\widetilde{\boldsymbol{x}}_{Kt}]$. Exploiting $\boldsymbol{W}(\delta_0)\boldsymbol{S}^{-1} = \boldsymbol{S}^{-1}\boldsymbol{W}(\delta_0)$[7], yields

$$\begin{aligned}&[\boldsymbol{J}_N\boldsymbol{R}(\lambda,\gamma)\boldsymbol{G}(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0 + \sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k}), \boldsymbol{J}_N\boldsymbol{R}(\lambda,\gamma)\widetilde{\boldsymbol{X}}_t, \boldsymbol{J}_N\boldsymbol{R}(\lambda,\gamma)(\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt})] \\ &= \boldsymbol{R}(\lambda,\gamma)\boldsymbol{S}^{-1}\{[\boldsymbol{W}(\delta_0)(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0 + \sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k}), \boldsymbol{S}\widetilde{\boldsymbol{X}}_t, \boldsymbol{S}(\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt})] \\ &- \boldsymbol{S}\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{\iota}_N(\kappa_1,\kappa_2,\kappa_3)\} \\ &= \boldsymbol{R}(\lambda,\gamma)\boldsymbol{S}^{-1}\{[\boldsymbol{W}(\delta_0)(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0 + \sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k}), \boldsymbol{S}\widetilde{\boldsymbol{X}}_t, \boldsymbol{S}(\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt})] \\ &- \boldsymbol{\iota}_N(\kappa_1^*,\kappa_2^*,\kappa_3^*)\}, \qquad (4.C.2)\end{aligned}$$

where $\kappa_m^* = \frac{1-\rho_0}{1-\lambda_0}\kappa_m$ for $m = 1,2,3$, and $\boldsymbol{S}\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{\iota}_N = \frac{1-\rho_0}{1-\lambda_0}\boldsymbol{\iota}_N$. Therefore, (4.C.1) is equivalent to

$$\begin{aligned}&\boldsymbol{W}(\delta_0)(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0 + \sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k})c_1 + \boldsymbol{S}\widetilde{\boldsymbol{X}}_t c_2 + \boldsymbol{S}(\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt})c_3 \\ &- \boldsymbol{\iota}_N(\kappa_1^* c_1 + \kappa_2^* c_2 + \kappa_3^* c_3) \\ &= \widetilde{\boldsymbol{X}}_t c_2 + \boldsymbol{W}(\delta_0)[\boldsymbol{W}(\alpha_{0,1})\widetilde{\boldsymbol{x}}_{1t}(\phi_{0,1}c_1 - \rho_0 c_3),...,\boldsymbol{W}(\alpha_{0K})\widetilde{\boldsymbol{x}}_{Kt}(\phi_{0K}c_1 - \rho_0 c_3)]\end{aligned}$$

[7] $\boldsymbol{W}(\delta_0)\boldsymbol{S}^{-1} = \boldsymbol{W}(\delta_0)(\boldsymbol{I}_N + \rho_0\boldsymbol{W}(\delta_0) + \rho_0^2\boldsymbol{W}^2(\delta_0) + ...) = (\boldsymbol{I}_N + \rho_0\boldsymbol{W}(\delta_0) + \rho_0^2\boldsymbol{W}^2(\delta_0) + ...)\boldsymbol{W}(\delta_0) = \boldsymbol{S}^{-1}\boldsymbol{W}(\delta_0)$.

$$+\sum_{k=1}^{K}(\boldsymbol{W}(\delta_0)\widetilde{\boldsymbol{x}}_{kt}\beta_{0k}c_1-\rho_0\boldsymbol{W}(\delta_0)\widetilde{\boldsymbol{x}}_{kt}c_2+\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}c_3)-\boldsymbol{\iota}_N(\kappa_1^*,\kappa_2^*\kappa_3^*)=0. \tag{4.C.3}$$

As $[\widetilde{\boldsymbol{X}}_t, \boldsymbol{W}(\delta_0)\widetilde{\boldsymbol{X}}_t, \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0,1})\widetilde{\boldsymbol{x}}_{1t}, ..., \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0K})\widetilde{\boldsymbol{x}}_{Kt}, \boldsymbol{\iota}_N]$ has assumed to have full column rank for each k, it follows that $c_2 = 0, \phi_{0k}c_1 - \rho_0 c_3 = 0$, and $\boldsymbol{W}(\delta_0)\beta_{0k}c_1 + \boldsymbol{W}(\alpha_{0k})c_3 = 0$. This in turn implies that $c_1 = 0$ and $c_3 = -\phi_{0k}c_1/\rho_0 = 0$ given the assumptions that $\rho_0\beta_{0k}\boldsymbol{W}(\delta_0) + \phi_{0k}\boldsymbol{W}(\alpha_{0k}) \neq 0$ and $\rho_0 \neq 0$. This completes the proof for the GNSAR model.

For the GNSMA model, we similarly have

$$\begin{aligned}\boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{G}(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0+\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k})c_1+\boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)\widetilde{\boldsymbol{X}}_t c_2\\ +\boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)(\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt})c_3=0.\end{aligned} \tag{4.C.4}$$

Denote $\kappa_1 = \frac{1}{N}\boldsymbol{\iota}_N'\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{G}(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0 + \sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k})$, $\kappa_2 = \frac{1}{N}\boldsymbol{\iota}_N'\boldsymbol{R}^{-1}(\lambda,\gamma)\widetilde{\boldsymbol{X}}_t$, and $\kappa_3 = \frac{1}{N}\boldsymbol{\iota}_N'\boldsymbol{R}^{-1}(\lambda,\gamma)[\boldsymbol{W}(\alpha_{0,1})\widetilde{\boldsymbol{x}}_{1t}, ..., \boldsymbol{W}(\alpha_{0K})\widetilde{\boldsymbol{x}}_{Kt}]$, to get

$$\begin{aligned}&[\boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{G}(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0+\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k}), \boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)\widetilde{\boldsymbol{X}}_t, \boldsymbol{J}_N\boldsymbol{R}^{-1}(\lambda,\gamma)(\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt})]\\ &=\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{S}^{-1}\{[\boldsymbol{W}(\delta_0)(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0+\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k}), \boldsymbol{S}\widetilde{\boldsymbol{X}}_t, \boldsymbol{S}(\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt})]\\ &-\boldsymbol{S}\boldsymbol{R}(\lambda,\gamma)\boldsymbol{\iota}_N(\kappa_1,\kappa_2,\kappa_3)\}\\ &=\boldsymbol{R}^{-1}(\lambda,\gamma)\boldsymbol{S}^{-1}\{[\boldsymbol{W}(\delta_0)(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0+\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k}), \boldsymbol{S}\widetilde{\boldsymbol{X}}_t, \boldsymbol{S}(\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt})]\\ &-\boldsymbol{\iota}_N(\kappa_1^*,\kappa_2^*,\kappa_3^*)\},\end{aligned} \tag{4.C.5}$$

where $\kappa_m^* = (1-\rho_0)(1-\lambda_0)\kappa_m$ for $m = 1,2,3$, and $\boldsymbol{S}\boldsymbol{R}(\lambda,\gamma)\boldsymbol{\iota}_N = (1-\rho_0)(1-\lambda_0)\boldsymbol{\iota}_N$. Therefore, (4.C.4) is equivalent to

$$\begin{aligned}&\boldsymbol{W}(\delta_0)(\widetilde{\boldsymbol{X}}_t\boldsymbol{\beta}_0+\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}\phi_{0k})c_1+\boldsymbol{S}\widetilde{\boldsymbol{X}}_t c_2+\boldsymbol{S}(\sum_{k=1}^{K}\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt})c_3\\ &-\boldsymbol{\iota}_N(\kappa_1^*c_1+\kappa_2^*c_2+\kappa_3^*c_3)\\ &=\widetilde{\boldsymbol{X}}_t c_2+\boldsymbol{W}(\delta_0)[\boldsymbol{W}(\alpha_{0,1})\widetilde{\boldsymbol{x}}_{1t}(\phi_{0,1}c_1-\rho_0 c_3), ..., \boldsymbol{W}(\alpha_{0K})\widetilde{\boldsymbol{x}}_{Kt}(\phi_{0K}c_1-\rho_0 c_3)]\end{aligned}$$

$$+\sum_{k=1}^{K}(\boldsymbol{W}(\delta_0)\widetilde{\boldsymbol{x}}_{kt}\beta_{0k}c_1-\rho_0\boldsymbol{W}(\delta_0)\widetilde{\boldsymbol{x}}_{kt}c_2+\boldsymbol{W}(\alpha_{0k})\widetilde{\boldsymbol{x}}_{kt}c_3)-\boldsymbol{\iota}_N(\kappa_1^*,\kappa_2^*\kappa_3^*)=0. \tag{4.C.6}$$

As $[\widetilde{\boldsymbol{X}}_t, \boldsymbol{W}(\delta_0)\widetilde{\boldsymbol{X}}_t, \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0,1})\widetilde{\boldsymbol{x}}_{1t}, ..., \boldsymbol{W}(\delta_0)\boldsymbol{W}(\alpha_{0K})\widetilde{\boldsymbol{x}}_{Kt}, \boldsymbol{\iota}_N]$ is assumed to have full column rank for each k, it follows that $c_2 = 0$, $\phi_{0k}c_1 - \rho_0 c_3 = 0$, and $\boldsymbol{W}(\delta_0)\beta_{0k}c_1 + \boldsymbol{W}(\alpha_{0k})c_3 = 0$. This in turn implies that $c_1 = 0$ and $c_3 = -\phi_{0k}c_1/\rho_0 = 0$ given the assumptions that $\rho_0\beta_{0k}\boldsymbol{W}(\delta_0) + \phi_{0k}\boldsymbol{W}(\alpha_{0k}) \neq 0$ and $\rho_0 \neq 0$. The desired result for the GNSMA model follows.

Chapter 5

Conclusion

This dissertation contains three chapters which are linked through their focus on the specification and estimation of several spatial econometric models and their spatial weight matrices. The central question of this thesis is threefold: (I) how to specify the spatial weight matrix or matrices, (II) how to specify the spatial lags, and (III) how to solve or diminish any identification problems of the model parameters. To address this central question, models from simple to challenging are considered. The general context and the more detailed context of each chapter is discussed in the previous chapters. This chapter summarizes and reiterates the main conclusions of each chapter, mentions some limitations, and provides potential directions for future research.

In Chapter 2 the prominent spatial Durbin model, which contains spatial lags in the dependent variable and the regressors, is taken as point of departure. This model generates more flexible indirect effects than their counterparts without Durbin terms. We further let each spatial lag and regressor have its own spatial weight matrix by parameterizing each of them with a different distance decay parameter. In the panel setting, this model can also contain individual and time fixed effects when N is large and T is finite. The QML estimator based on the within transformation to concentrate out the individual fixed effects is used to estimate the distance decay parameters together with the response parameters. In addition to the two functional forms of distance decay structure, the negative exponential and the inverse distance, two types of normalization, row normalization and scalar normalization, are considered. The simulation and empirical application results show that, instead

of adopting one common pre-specified matrix, each spatial lag should interact with a different spatial weight matrix. Therefore, this chapter answers Question (I), the specification of the spatial weight matrix or matrices, by estimating them within the SDM framework.

Chapter 3 considers the QML within estimator of the SARMA model with individual and time fixed effects, which includes spatial lags in the dependent variable and the errors. This model has gained less attention in the literature compared with the widely applied SARAR model. We focus on the popular panel setting in which N is large and T is finite. We extend the interpretation of local and global shocks by considering the spatial range of the spatial weight matrices: a spatial MA error process representing local shocks which tend to go together with a dense matrix and the spatial AR error process representing global shocks which tend to go together with a sparse matrix. The simulation results show that the model parameters, the direct effects, and the indirect effects are biased when wrong spatial weight matrices are used. This chapter therefore answers Question (II) regarding the specification of the spatial lags. Question (I) about the specification of the spatial weight matrix is also partly answered since the empirical results show that different types of spatial lags may interact with different spatial weight matrices in terms of sparsity or density.

As an extension of Chapter 2 and Chapter 3, Chapter 4 considers the estimation of panel GNS models with spatial AR or MA errors, individual and time fixed effects, and parameterized spatial weight matrices within the large N and large T setting. We exploit and modify the QML estimator based on the orthogonal transformation to estimate the distance decay parameters of the spatial weight matrices and the response parameters of these models. Monte Carlo simulations are conducted to show the acceptable performance of the QML estimators. The empirical application using US house data taken from Yang (2021) confirms the existence of global shocks in the US housing market, since the spatial AR error process is found to outperform its spatial MA counterpart. This chapter answers all three questions posed in this thesis. It answers Question (I) regarding the specification of the spatial weight matrix since it shows that distance-based decay matrices can be estimated along the response parameters of the model. It answers Question (II) regarding the specification of spatial lags since both MA and AR errors are considered. Finally,

it answers Question (III) by taking into account the rank conditions of the GNSAR and GNSMA models and by showing that the identification of its parameters is alleviated by parameterizing the spatial weight matrices.

Matlab routines will be made available for the proposed QML estimators for all the mentioned models; these routines provide opportunities to practitioners for their own research.

One issue for further research is how to select between spatial AR and MA errors. Besides the economic theory set out in Chapter 3, Bayesian posterior model probabilities are a powerful and straightforward tool for model selection since the model with the highest probability should be selected (LeSage, 2014). However, prior and posterior probabilities need to be derived not only for each model specification but also for different spatial weight matrices of each spatial lag in these specifications. These derivations are not yet available and left for future research. Attempts have also been made for selecting from non-nested spatial econometric models. Liu and Lee (2019) propose a non-degenerate likelihood-ratio test for model selection between two non-nested spatial econometric models. This test is based on QML estimators of the matrix exponential spatial specification (MESS) model and the SARAR model. However, the power of the test for model selection between models with spatial AR and MA errors is yet unknown. One of the possibilities is to explore whether AR versus MA tests in the time-series literature (e.g., King, 1983) are also applicable to spatial econometric models.

Another step forward of this thesis is to add "dynamic" elements to the spatial panel models. Elhorst (2021) categorizes spatial econometric models into four generations: the first generation contains models based on cross-sectional data, the second generation comprises models based on pooled data but with controls for individual and time fixed or random effects, the third generation captures dynamic spatial panel data models, and the current fourth generation, from 2021, includes spatial panel data models with all spatial lags, both fixed or random effects, and common factors. These models are quite difficult to apply. Parameterizing the spatial weight matrices in these advanced models will make the estimation far more difficult. Furthermore, they may suffer from additional identification problems. Lee and Yu (2016) propose identification conditions for dynamic spatial panel models with Durbin terms. The role of parameterized spatial weight matrix in the iden-

tification of the parameters in these dynamic models is left for future research.

References

Ahrens, A., & Bhattarchajee, A. (2015). Two-step Lasso estimation of the spatial weight matrix. *Econometrics*, 3, 128-155.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht, the Netherlands: Kluwer.

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26(2), 153-166.

Anselin, L., & Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In Ullah A., Giles D. (eds) *Handbook of Applied Economics Statistics*, 237-289. New York: Marcel Dekker.

Arbia, G., Bera, A. K., Doğan, O., & Taşpınar, S. (2020). Testing impact measures in spatial autoregressive models. *International Regional Science Review*, 43(1-2), 40-75.

Arellano, M. (2003). *Panel data econometrics*. Oxford University Press.

Bailey, N., Holly, S., & Pesaran, M. H. (2016). A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence. *Journal of Applied Econometrics*, 31(1), 249-280.

Balazsi, L., Matyas, L., & Wansbeek, T. (2018). The estimation of multidimensional fixed effects panel data models. *Econometric Reviews*, 37(3), 212-227.

Baltagi, B. H. (2005). *Econometric Analysis of Panel Data*, 3rd edn. Chichester: Wiley.

Baltagi, B. H., & Liu, L. (2011). An improved generalized moments estimator for a spatial moving average error model. *Economics Letters*, 113(3), 282-284.

Baltagi, B. H., Fingleton, B., & Pirotte, A. (2019). A time-space dynamic panel data model with spatial moving average errors. *Regional Science and Urban Economics*, 76, 13-31.

Bavaud, F. (1998). Models for spatial weights: a systematic look. *Geographical analysis*, 30(2), 153–171.

Beenstock, M., & Felsenstein, D. (2019). *The Econometric Analysis of Non-stationary Spatial Panel Data*. Heidelberg: Springer.

Beer, C., & Riedl, A. (2012). Modelling spatial externalities in panel data: The spatial Durbin model revisited. *Papers in Regional Science*, 91(2), 299-318.

Blume, L.E., Brock, W.A., Durlauf, S.N., & Jayaraman, R. (2015) Linear social interactions models. *Journal of Political Economy*, 123(2), 444-496.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1), 41-55.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2020). Peer effects in networks: A survey. *Annual Review of Economics*, 12, 603-629.

Burridge, P. (1981). Testing for a common factor in a spatial autoregression model. *Environment and Planning A*, 13(7), 795-400.

Burridge, P., Elhorst, J. P., & Zigova, K. (2016). *Group interaction in research and the use of general nesting spatial models*. Emerald Group Publishing Limited.

Chen, D., Lu, X., Hu, W., Zhang, C., & Lin, Y. (2021). How urban sprawl influences eco-environmental quality: Empirical research in China by using the spatial Durbin model. *Ecological Indicators*, 131, 108-113.

Choi, C. Y., & Hansz, J. A. (2021). From banking integration to housing market integration-Evidence from the comovement of US Metropolitan House Prices. *Journal of Financial Stability*, 54, 100883.

Cliff, A.,& Ord, J. (1973). *Spatial Autocorrelation*. London: Pion.

Corrado, L., & Fingleton, B. (2012). Where is the economics in spatial econometrics? *Journal of Regional Science*, 52(2), 210-239.

Davidson, R., & MacKinnon, J.G. (2004). *Econometric Theory and Methods*. Oxford University Press, New York.

Debarsy, N., & LeSage, J.P. (2018). Flexible dependence modeling using convex combinations of different types of connectivity structures. *Regional Science and Urban Economics*, 69, 48-68.

De Paula, A., Rasul, I., & Souza, P.C.L. (2020). *Identifying network ties from panel data: Theory and an application to tax competition*. ArXiv. https://arxiv.org/abs/1910.07452v2.

Doğan, O., & Taşpınar S. (2013). GMM estimation of spatial autoregressive models with moving average disturbances. *Regional Science and Urban Economics*, 43(6), 903-926.

Dubin, R. A. (1988). Estimation of regression coefficients in the presence of spatially autocorrelated error terms. *The Review of Economics and Statistics*, 70(3), 466-474.

Elhorst, J. P. (2003). Specification and estimation of spatial panel data models. *International Regional Science Review*, 26(3), 244-268.

Elhorst, J. P. (2010). Applied spatial econometrics: raising the bar. *Spatial Economic Analysis*, 5(1), 9-28.

Elhorst, J. P. (2014). *Spatial Econometrics from Cross-sectional Data to Spatial Panels*. Heidelberg: Springer.

Elhorst, J. P., & Halleck Vega, S. (2017). The SLX model: Extensions and the sensitivity of spatial spillovers to W. *Papeles de Economía Española*, 152, 34-50.

Elhorst, J. P. (2021). The dynamic general nesting spatial econometric model for spatial panels with common factors: Further raising the bar. *Review of Regional Research*, 1-19.

Elhorst, J. P., Gross, M., & Tereanu, E. (2021). Spillovers in space and time: Where spatial econometrics and global VAR models meet. *Journal of Economic Surveys*, 35(1), 192-226.

Ertur, C., & Koch, W. (2007). Growth, technological interdependence and spatial xxternalities: theory and evidence. *Journal of Applied Econometrics*, 22(6):1033-1062.

Feng, Z., & Chen, W. (2018). Environmental regulation, green innovation, and industrial green development: An empirical analysis based on the spatial Durbin model. *Sustainability*, 10(1), 223.

Fingleton, B. (2008a). A generalized method of moments estimator for a spatial model with moving average errors, with application to real estate prices. *Empirical Economics*, 34(1), 35-57.

Fingleton, B. (2008b). A generalized method of moments estimator for a spatial panel model with an endogenous spatial lag and spatial moving average errors. *Spatial Economic Analysis*, 3(1), 27-44.

Fingleton, B., & LeGallo, J. (2007). Finite sample properties of estimators of spatial models with autoregressive, or moving average, disturbances and system feedback. *Annales d'Économie et de Statistique*, 39-62.

Fingleton, B., & LeGallo, J. (2008). Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: Finite sample properties.*Papers in Regional Science*, 87(3), 319–339.

Fingleton, B., Le Gallo, J., & Pirotte, A. (2018). A multidimensional spatial lag panel data model with spatial moving average nested random effects errors. *Empirical Economics*, 55(1), 113-146.

Fischer, M.M., Scherngell, T., & Reismann, M. (2009). Knowledge spillovers and total factor productivity: Evidence using a spatial panel data model. *Geographical Analysis*, 41(2), 204–220.

Fotheringham, A. S. (1981). Spatial structure and distance-decay parameters. *Annals of the Association of American Geographers*, 71(3), 425-436.

Gibbons, S., & Overman, H. G. (2012). Mostly pointless spatial econometrics? *Journal of Regional Science*, 52(2), 172-191.

Gupta, A. (2019). Estimation of spatial autoregressions with stochastic weight matrices. *Econometric Theory*, 35, 417-463.

Haining, R. P. (1978). The moving average model for spatial interaction. *Transactions of the Institute of British Geographers*, 3(2), 202-225.

Haining, R. (1990). *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press, Cambridge.

Halleck Vega, S., & Elhorst, J. P. (2015). The SLX model. *Journal of Regional Science*, 55(3), 339-363.

Hassan, A. R. (2017). The interplay between the Bayesian and frequentist approaches: a general nesting spatial panel data model. *Spatial Economic Analysis*, 12(1), 92-112.

Heijnen, P., & Elhorst, J. P. (2018) The diffusion of local differentiated waste disposal taxes in the Netherlands. *De Economist*, 166, 239–258.

Hepple, L. W. (2003). Bayesian and maximum likelihood estimation of the linear model with spatial moving average disturbances. *School of Geographical Sciences, University of Bristol, Working Papers Series, University of Bristol, School of Geographical Sciences*.

Herrera, M., Mur, J., & Ruiz, M. (2019). A comparison study on criteria to select the most adequate weighting matrix. *Entropy*, 21(2), 160.

Huang, J. S. (1984). The autoregressive moving average model for spatial analysis. *Australian Journal of Statistics*, 26(2), 169-178.

Huang, H., Wang, F., Song, M., Balezentis, T., & Streimikiene, D. (2021). Green innovations for sustainable development of China: Analysis based on the nested spatial panel models. *Technology in Society*, 65, 101593.

Kakamu, K. (2005). Bayesian estimation of a distance functional weight matrix model. *Economics Bulletin*, 3(57), 1–6.

Kelejian, H. H. (2008). A spatial J-test for model specification against a single or a set of non-nested alternatives. *Letters in Spatial and Resource Sciences*, 1(1), 3-11.

Kelejian, H. H., & Piras, G. (2017). *Spatial econometrics*. Elsevier Academic Press.

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

Kelejian, H. H., & Prucha, I.R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40 (2), 509-533.

Kelejian, H. H., & Prucha, I.R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157(1), 53–67.

King, M. L. (1983). Testing for autoregressive against moving average errors in the linear regression model. *Journal of Econometrics*, 21(1), 35-51.

Koroglu, M., & Sun, Y. (2016). Functional-coefficient spatial Durbin models with nonparametric spatial weights: An application to economic growth. *Econometrics*, 4(1), 6.

Lam, C., & Souza, P. C. (2020). Estimation and selection of spatial weight matrix in a spatial lag model. *Journal of Business & Economic Statistics*, 38(3), 693-710.

Lee, L. F. (2003). Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econometric Reviews*, 22(4), 307-335.

Lee, L. F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, 72(6), 1899-1925.

Lee, L. F. (2007). GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *Journal of Econometrics*, 137(2), 489-514.

Lee, L. F., & Liu, X. (2010). Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances. *Econometric Theory*, 26(1), 187-230.

Lee, L. F., Liu, X., & Lin, X. (2010). Specification and estimation of social interaction models with network structures. *The Econometrics Journal*, 13(2), 145-176.

Lee, L.F., & Yu, J. (2010a). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154(2), 165-185.

Lee, L.F., & Yu, J. (2010b). Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, 40, 255-271.

Lee, L.F., & Yu, J. (2010c). *Spatial Panels: Random Components vs. Fixed Effects*. Unpublished manuscript. Downloaded from CiteSeerX, The Pennsylvania State University.

Lee, L.F., & Yu, J. (2012). Spatial panels: Random components versus fixed effects. *International Economic Review*, 53(4), 1369-1412.

Lee, L.F., & Yu, J. (2016). Identification of spatial Durbin panel models. *Journal of Applied Econometrics*, 31, 133–162.

Leenders, R. T. A. (2002). Modeling social influence through network autocorrelation: Constructing the weight matrix. *Social Networks*, 24(1), 21-47.

LeSage, J. P. (2014). Spatial econometric panel data model specification: a Bayesian approach. *Spatial Statistics*, 9, 122-145.

LeSage, J. P. (2015). Software for Bayesian cross section and panel spatial model comparison. *Journal of Geographical Systems*, 17, 297–310.

LeSage, J. P., & Pace, R.K. (2009). *Introduction to Spatial Econometrics*. Taylor & Francis CRC Press: Boca Raton.

LeSage, J. P., & Pace, R.K. (2018). Spatial econometric Monte Carlo studies: Raising the bar. *Empirical Economics*, 55(1), 17–34.

Lewbel, A., Qu, X., & Tang, X. (2021). *Social networks with unobserved links (revised 01 Feb 2021)*. Boston College Working Papers in Economics 1004, Boston College Department of Economics.

Li, J., & Li, S. (2020). Energy investment, economic growth and carbon emissions in China—Empirical analysis based on spatial Durbin model. *Energy Policy*, 140, 111425.

Lin, X. (2015). Utilizing spatial autoregressive models to identify peer effects among adolescents. *Empirical Economics*, 49(3), 929-960.

Liu, T., & Lee, L. F. (2019). A likelihood ratio test for spatial model selection. *Journal of Econometrics*, 213(2), 434-458.

Liu, X., Lee, L. F., & Bollinger, C. R. (2010). An efficient GMM estimator of spatial autoregressive models. *Journal of Econometrics*, 159(2), 303-319.

Magnus, J. R. (1985). On differentiating eigenvalues and eigenvectors. *Econometric Theory*, 1(2), 179-191.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3), 531-542.

Millo, G. (2022). The generalized spatial random effects model in R. *Journal of Spatial Econometrics*, 3(1), 1-18.

Murdoch, J. C., Rahmatian, M., & Thayer, M.A. (1993). A spatially autoregressive median voter model of recreation expenditures. *Public Finance Quarterly*, 21(3), 334-350.

Neyman, J., & Scott, E. L. (1948). Consistent estimates based on partially consistent observations. *Econometrica: Journal of the Econometric Society*, 1-32.

Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70, 120-126.

Ord J.K. (1981). Towards a theory of spatial statistics: A comment. *Geographical Analysis*, 13, 91-93.

Pan, X., Guo, S., Li, M., & Song, J. (2021). The effect of technology infrastructure investment on technological innovation——A study based on spatial Durbin model. *Technovation*, 107, 102315.

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. *Econometric reviews*, 34(6-10), 1089-1117.

Qu, X., & Lee, L.F. (2015). Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *Journal of Econometrics*, 184, 209-232.

Sabater, A., & Graham, E. (2019). International migration and fertility variation in Spain during the economic recession: A spatial Durbin approach. *Applied Spatial Analysis and Policy*, 12(3), 515-546.

Spirkin, A. G. (1983). *Dialectical materialism*. Progress Publishers.

Szumilo, N. (2021). Prices of peers: identifying endogenous price effects in the housing market. *The Economic Journal*, 131(639), 3041-3070.

Tientao, A., Legros, D., & Pichery, M. C. (2016). Technology spillover and TFP growth: A spatial Durbin model. *International Economics*, 145, 21-31.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2), 307-333.

Xiao, H., & Mao, J. (2021). Effects of postgraduate education on technological innovation: a study based on the spatial Durbin model. *Asia Pacific Education Review*, 22(1), 89-99.

Xu, X., & Lee, L. F. (2019). Theoretical foundations for spatial econometric research. *Regional Science and Urban Economics*, 76, 2–12.

Yang, C. F. (2021). Common factors and spatial dependence: An application to US house prices. *Econometric Reviews*, 40(1), 14-50.

Yesilyurt, M. E., & Elhorst, J. P. (2017). Impact of neighboring countries on military expenditures: A dynamic spatial panel approach. *Journal of Peace Research*, 54(6), 777-790.

Samenvatting

Ruimtelijk econometrische modellen bieden mogelijkheden om ruimtelijke, economische en sociale interacties tussen transversale waarnemingseenheden te verklaren door verschillende ruimtelijke vertragingen op te nemen. Ruimtelijke gewichtenmatrices zijn daarin een essentieel onderdeel, omdat ze weergeven hoe deze eenheden met elkaar verbonden zijn. Om ruimtelijk econometrische modellen te kunnen toepassen, dienen drie vragen te worden beantwoord. Het betreft achtereenvolgens (I) hoe de ruimtelijke gewichtenmatrix of -matrices te modelleren, (II) hoe ruimtelijke vertragingen te modelleren, en (III) langs welke weg de parameters van het model zijn geïdentificeerd. Dit proefschrift onderzoekt oplossingen en geeft antwoord op deze drie vragen. De panel structuur en schattingsmethode spelen hierbij een belangrijke rol. In ruimtelijke panel data modellen met individuele en tijdgebonden vaste effecten, worden deze vaste effecten in het algemeen namelijk eerst uit het model getransformeerd, middels ofwel de zogenoemde "within" transformatie of de orthogonale transformatie, waarna het getransformeerde model wordt geschat, in dit proefschrift op basis van quasi maximale waarschijnlijkheid (QML).

Hoofdstuk 2 gaat in op eerste vraag door te kijken naar de schatting van het ruimtelijke Durbin model (spatial Durbin (SD) model) met geparameteriseerde ruimtelijke gewichtenmatrices. Dit model bevat ruimtelijke vertragingen in de te verklaren variabele en elk van de verklarende variabelen. De kracht van dit model in empirisch onderzoek is dat het flexibele indirecte spillover effecten genereert, een maat voor de marginale effecten van veranderingen in één van de verklarende variabelen van een waarnemingseenheid op de te verklaren variabele van andere waarnemingseenheden. Deze flexibiliteit wordt opgevoerd door elke ruimtelijke

gewichtenmatrix te parameteriseren met een andere afstandsvervalparameter en deze parameters tezamen met de andere parameters van het model te schatten. De QML-schatter op basis van de within-transformatie wordt gebruikt om het SD-model te schatten met individuele en tijdgebonden effecten onder de veronderstelling dat N groot en T eindig is. Vanwege deze veronderstelling behoeven alleen de individuele effecten uit het model te worden getransformeerd, terwijl de tijdgebonden effecten kunnen worden behandeld als reguliere regressoren. De Monte Carlo-simulatieresultaten en empirische toepassing laten zien dat de gangbare praktijk om met één gemeenschappelijke ruimtelijke gewichtenmatrix te werken voor alle ruimtelijke vertragingen in het model moet worden verworpen en dat de parameterisering van de gewichtenmatrices de flexibiliteit waarmee indirecte spillover effecten worden gemeten inderdaad verbetert.

Hoofdstuk 3 gaat in op de tweede vraag met betrekking tot de specificatie van ruimtelijke vertragingen door ruimtelijke moving average (MA) en ruimtelijke autoregressieve (AR) processen in de storingstermen met elkaar te vergelijken. We breiden de interpretatie van lokale en globale schokken uit door rekening te houden met het ruimtelijke bereik van de gewichtenmatrices: het ruimtelijke MA-proces modelleert lokale schokken en correspondeert met goed gevulde ruimtelijke gewichtenmatrices, terwijl het ruimtelijke AR-proces globale schokken vastlegt en correspondeert met schaars gevulde ruimtelijke gewichtenmatrices. De QML-schatter op basis van de within-transformatie wordt gebruikt om een ruimtelijk econometrisch model te schatten met ruimtelijke vertragingen in de te verklaren variabele en de storingsterm met vooraf gespecificeerde maar verschillende ruimtelijke gewichtenmatrices. Dit model, beter bekend als het ruimtelijke autoregressieve (SAR) model met ruimtelijke MA-storingstermen, of kortweg het SARMA-model, bevat voorts individuele en tijdgebonden effecten en wordt wederom geschat onder de veronderstelling dat N groot en T eindig is. De Monte Carlo-simulatieresultaten laten zien dat de selectie van de juiste ruimtelijke gewichtenmatrix en het overwegen van verschillende ruimtelijke gewichtenmatrices voor verschillende ruimtelijke vertragingen in het model van cruciaal belang zijn en misschien zelfs belangrijker dan de selectie van het juiste ruimtelijk econometrische model. De empirische toepassing op basis van data ontleend aan Yesilyurt en Elhorst (2017), waarin militaire uitgaven worden verklaard, laat echter zien dat

het ruimtelijke proces in de storingsterm van dit model het beste gemodelleerd kan worden volgens een MA structuur en een goed gevulde ruimtelijke gewichtenmatrix, omdat de schokken een lokaal karakter vertonen. Met deze bevindingen beantwoordt dit hoofdstuk zowel niet alleen de tweede vraag gesteld in dit proefschrift, maar ook de eerste.

Hoofdstuk 4 gaat aanvankelijk in op de derde vraag betreffende identificatie van de modelparameters middels de schatting van het general nesting spatial (GNS) model met ruimtelijke AR- of MA-storingstermen en geparameteriseerde ruimtelijke gewichtenmatrices. Deze modellen bevatten ruimtelijke vertragingen in de te verklaren variabele, elk van de verklarende variabelen en de storingsterm. De QML-schatter gebaseerd op de orthogonale transformatie wordt niet alleen gebruikt maar ook afgeleid om deze GNS-modellen te schatten met individuele en tijdgebonden vaste effecten, in dit hoofdstuk onder de veronderstelling dat zowel N en T groot zijn. Langs deze weg wordt bewezen dat het parameteriseren van de ruimtelijke gewichtenmatrices het in de literatuur genoemde identificatieprobleem verlicht. De Monte Carlo simulatieresultaten laten zien dat de QML orthogonale schatter acceptabel presteert. De empirische toepassing op basis van een data set ontleend aan Yang (2021), waarin de stijging van huizenprijzen in Amerikaanse metropolen wordt verklaard middels veranderingen in inkomens- en bevolkingsgroei, laat zien dat elke van de vier ruimtelijke vertragingen in het model wordt gekenmerkt door een andere gewichtenmatrix en dat het ruimtelijke proces in de storingsterm in dit model het beste gemodelleerd kan worden volgens een autoregressief (AR) proces en een schaars gevulde ruimtelijke gewichtenmatrix, omdat de Amerikaanse huizenmarkt wordt gekenmerkt door schokken die een globaal karakter vertonen. Met deze bevindingen beantwoordt dit hoofdstuk tenslotte alle vragen die in dit proefschrift gesteld zijn.