ESTIMATION OF SPATIAL AUTOREGRESSIVE MODELS

WITH DYADIC OBSERVATIONS AND LIMITED

DEPENDENT VARIABLES

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

SHALI LUO

Dr. J. Isaac Miller, Dissertation Supervisor

DECEMBER 2012

UMI Number: 3537893

UMI®
Dissertation Publishing

UMI 3537893

ProQuest®

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

المنارة للاستشارات

www.manaraa.com

The undersigned, appointed by the Dean of the Graduate School, have examined the

dissertation entitled:

ESTIMATION OF SPATIAL AUTOREGRESSIVE MODELS

WITH DYADIC OBSERVATIONS AND LIMITED

DEPENDENT VARIABLES

presented by Shali Luo

a candidate for the degree of Doctor of Philosophy and hereby certify that, in their

opinion, it is worthy of acceptance.

_____

Dr. J. Isaac Miller

_____

Dr. Shawn Ni

_____

Dr. Vitor Trindade

_____

Dr. Chong He

I dedicate this dissertation to
my grandma, my parents, and my husband.

ACKNOWLEDGMENTS

I am deeply grateful to my advisor, Dr. Zack Miller. Without his guidance, encouragement, and occasionally much needed pushes, I might have been less adventurous in exploring the new econometric techniques that I have developed in the dissertation. During this academic journey, I have learned so much from Dr. Miller. His intellectual passion and seriousness has taught me what makes a good researcher. My heart-felt gratitude also goes to Dr. Shawn Ni. During my graduate years, I have greatly benefited from numerous discussions with Dr. Ni regarding my coursework and research projects. I highly look up to his ingenuity in research ideas as well as all-round academic expertise. I also want to express my thanks to Dr. Vitor Trindade for the insightful suggestions and comments that he provided me at various stages of my dissertation. His questions and advice are always thought-provoking and have made me to think deeper and further. Being an outside committee member, Dr. Chong He has been a very resourceful expert to consult with. I really appreciate her making time to talk over the various technical issues that I encountered in the process of completing the dissertation. She helped me to bridge the gap between textbook knowledge and real-life applications. In addition, I would like to extend my thankfulness to Dr. Chris Wikle. When I approached Dr. Wikle with modeling problems, he kindly agreed to read my rough draft and arrange extra meetings to further discuss my questions. Not only did he provide helpful comments, but also directed me to valuable references. His care for students and intellectuality has encouraged me to confront difficult research moments with a more positive attitude.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

viii

Chapter 1

INTRODUCTION

When data are collected with reference to points or regions in space, it is unrealistic to assume independence among individual observations since values observed at one location or over one region may depend on those at nearby locations. To appropriately handle spatially correlated data, statisticians design spatial statistical models and their applications appear in a wide range of issue areas such as geography, meteorology, resource conservation, disease control, and crime studies. The popularity of spatial statistical models also has led theoretical econometricians and applied economists to explore spatial/strategic interaction as an important feature of social interaction models and spillover effect or network models. However, contrary to statisticians who treat the spatial process in the error term, econometricians tend to specify the spatial dependence in the mean component of the regression model. This difference may be attributed to the fact that spatial analysis in statistics is mainly conducted for mapping and prediction, whereas economists are more interested in drawing reliable inferences about a causal relationship in the presence of spatial correlation and identifying the impact of spatial dependence on the phenomenon of interest. In a historical point of view, the publication of Anselin's (1988) influential book, *Spatial Econometrics: Methods and Models*, stimulated econometricians to have worked towards developing spatial models that are better tailored to the specific features of economics data. One of such efforts results in advancing spatial econometric methods for constructing spatial interaction models.

1

Spatial interaction models are referred to as "models that focus on flows between origins and destinations" (LeSage and Pace, 2008, p. 942; also see Sen and Smith, 1995). This type of flow data is often found in the studies of international trade flows, population migration, transportation, communications and information flow, network, and regional and interregional economic analysis. Conventionally, spatial interaction models include a distance variable as an attempt to remove the spatial correlation among sampled origin-destination (OD) pairs. However, spatial dependence may not properly be controlled for with a distance variable that measures the bilateral distance between each origin and destination (e.g., Porojan 2001; Lee and Pace, 2005). Upon recognizing the inadequacy of the distance variable, LeSage and Pace (2008) argue that flow data involves both an origin and a destination at the same time so that spatial dependence may arise from multiple sources. As a better tuned approach to handle spatial correlation among data that features a directional flow, LeSage and Pace extend the traditional spatial interaction model by including three spatial connectivity matrices that explicitly capture origin-based, destination-based, and origin-to-destination dependence among flow data.

Nonetheless, LeSage and Pace's methodological advance is directed at continuous dependent variables, and it is not readily applicable to binary or censored data. This dissertation aims to fill the gap by extending the spatial OD modeling technique to the limited dependent variable cases. It should be noted that, due to the non-linearity of regression models for limited dependent variables with spatial correlation, the commonly used maximum likelihood (ML) estimator is no longer appropriate. Therefore, this research adopts Bayesian procedures for model estimation.

In Chapter 2, I extend the spatial OD model in order to accommodate flow data that represent binary choice outcomes. Data on interregional flows or interactions such as firms'

2

investment decisions, interregional collaboration, and contracts are often collected as dichotomous and they may exhibit spatial dependence. Chapter 2 proposes a spatial OD probit which incorporates the three spatial connectivity matrices advanced by LeSage and Pace in the regression model of the latent dependent variable. However, as McMillen (1992) observes, it is difficult to adapt ML estimators to probit models with spatial correlation. In order to avoid the inconsistency issue associated with an ML estimator, I show that a spatial OD probit model may be estimated with a Bayesian method that is similar to what LeSage and Pace (2009) discuss in their study. In addition, I suggest an effective way to deal with self-directed OD pairs. Due to the structure of weight matrices for OD flow samples, intra-regional or self-directed pairs are included in the estimation by construction, even though they are of no interest to the researcher. For example, in studies of international conflict, it would be not sensible to include observations on civil conflict because the causes of the former are different from those of the latter. This research introduces an elimination approach that entirely removes self-directed pairs from the estimation procedure, while preserving the correlation structure embodied by the weight matrices.

As an application, the spatial OD probit model is employed to investigate militarized interstate dispute (MID) initiations, observations of which can be viewed as a "directional flow" from the initiator to the target. Using a cross-section of 26 European countries drawn from the period leading up to WWII, I find empirical evidence for two types of spatial correlation: target-based and initiator-to-target based. Compared to a benchmark model that ignores spatial correlation, the effect estimates of the explanatory variables noticeably change under the spatial OD probit model, suggesting the inadequacy of models that do not account for spatial correlation in this context.

3

In Chapter 3, I focus on another type of limited dependent variable measured as censored data. When flow data are generated with an underlying spatial process but censored, a standard spatial OD model is less useful. To tackle this issue, I propose a Tobit version of the spatial OD model. For the purposes of this chapter, I use bilateral trade data where censoring occurs at zero and where trade flows between origins and destinations are expected to be correlated with those of neighboring OD pairs. To be consistent with the previous studies of international trade, I use the conventional log-linear specification of the gravity model but augment it with spatial OD modeling technique.

Due to its simple estimation and interpretation, the gravity equation in its log-linearized form has served as the workhorse in applied trade studies. However, trade data often contain zero flows, so they are not suitable for direct log transformation as required by log-linearized gravity models. Accordingly, handling zero-valued observations is an important issue in empirical trade studies. Following Eaton and Tamura (1994)'s approach to zero trade values, my proposed model is based on an assumption that there exists a threshold level for trade data. When trade volume reaches the threshold, the model records the observed trade value. Otherwise, it is treated as unobserved and latent.

I term this improved technique a spatial OD threshold Tobit model. It avoids leaving out zero-valued observations while utilizing the log-linearized gravity model. In addition, it addresses multiple forms of spatial correlation embedded in "directional" trade flows in a similar fashion as LeSage and Pace's spatial OD modeling. Not surprisingly, the addition of the threshold parameter as well as the spatial process in model specification complicates the estimation of the model. This complication is resolved by employing a Bayesian approach. It is

4

also worth noting that this model can accommodate other situations where the censored point is positive or where the dependent variable does not need to be log-transformed.

As an illustration, the proposed spatial OD threshold Tobit model is applied to export flows among 32 Asian countries in 1990. The empirical results point to all three types of spatial dependence: exporter-based, importer-based, and exporter-to-importer based. After considering the multiple sources of spatial correlation in bilateral trade flows, I find that the impact of conventional trade variables changes in a noticeable way, calling into question the assumption of spatial independence when modeling trade flows.

In sum, given that flow data or data on spatial or strategic interactions often represent binary outcomes or censored values, this dissertation has extended the existing spatial OD modeling technique and proposed models and estimation strategies to deal with the aforementioned types of limited dependent variables in the context of spatial correlation among the observations. The findings of this dissertation confirm Anselin's (1988) argument that, like temporal dependence, spatial dependence may generate inconsistent estimates if not properly handled.

References

Anselin, L. 1988. *Spatial Econometrics: Methods and Model*s. Boston: Kluwer Academic.

Eaton, J. and A. Tamura. 1994. "Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns," *Journal of the Japanese and International Economies*, 8, 478-510.

Lee, M.L. and R.K. Pace. 2005. "Spatial Distribution of Retail Sales," *Journal of Real Estate Finance and Economics*, 31(1), 53-69.

LeSage, J.P. and R.K. Pace. 2008. "Spatial Econometric Modeling of Origin-Destination Flows," *Journal of Regional Science*, 48(5), 941-967.

LeSage, J.P. and R.K. Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton: Chapman & Hall/CRC.

McMillen, D.P. 1992. "Probit with Spatial Autocorrelation," *Journal of Regional Science*, 32(3), 335-348.

Porojan, A. 2001. "Trade Flows and Spatial Effects: The Gravity Model Revisited*," Open Economies Review*, 12, 265-280.

Sen, A. and T.E. Smith. 1995. *Gravity Models of Spatial Interaction Behavior*. Heidelberg: Springer-Verlag.

Chapter 2

A SPATIAL ORIGIN-DESTINATION MODEL OF CORRELATES OF WORLD WAR II IN EUROPE

2.1 Abstract

This study examines cases in which interregional flows are measured as binary choice outcomes and proposes a spatial origin-destination (OD) probit. The use of Bayesian modeling avoids inconsistent estimates, which result from maximum likelihood estimation when spatially lagged terms of the dependent variable are included on the right-hand side of the equation. Drawing on three spatial dependence matrices, which are similar to those used by LeSage and Pace (2008), capturing origin-based, destination-based, and origin-to-destination-based dependence, the spatial OD probit model corrects for the spatial autocorrelation among binary observations of directional flows. The model is used to analyze militarized interstate dispute (MID) initiation, which features directed dyads as the unit of analysis and a binary measure as the dependent variable. Upon fitting the spatial OD probit model to the conflict initiation data for 26 European countries for the years 1933-1941, this study finds empirical evidence for two types of spatial dependence: target-based and initiator-to-target based. After further taking the dyadic dependence into consideration, this study reveals that the effects of conventional political variables such as capabilities change in a notable way, suggesting that models that do not account for spatial dependence are inadequate.

2.2 Introduction

Spatial econometrics, as a subfield of econometrics, focuses on "the treatment of spatial interaction (spatial autocorrelation) and spatial structure (spatial heterogeneity) in regression models for cross-sectional and panel data" (Anselin, 2003). In particular, the focus on spatial interaction has generated much appreciation among both applied and theoretical econometricians, as manifested by applications of spatial econometric models in an increasingly wide range of fields in economics. At an early stage of this development, econometric models that explicitly incorporated spatial interaction were limited to regional studies. However, recent empirical studies employing spatial econometric methods have been found in such diverse areas as international trade, public economics, labor economics and agricultural economics. As the scope of spatial econometrics widens, advances are being made towards developing more sophisticated model specifications, which are tuned to particular data structures.

The progress of spatial econometrics also provides innovative ideas and useful methodological tools for researchers in other social sciences where the potential problem of spatial dependence looms large. Over the past ten years, political scientists, among others, have employed various technical and statistical methods to address the spatial dependence embedded in their data structure (e.g., Gleditsch and Ward, 2000; Ward and Gleditsch, 2002; Gartzke and Gleditsch, 2008). Nonetheless, most of the methods they use suffer from an inability to directly model the spatial dependence process, and consequently are incapable of specifying the distinctive impact of spatial dependence on the phenomenon under scrutiny and the effect of each of the causal factors apart from spatial interaction.

8

Ward and Gleditsch (2008), for example, introduce two spatial econometric models in the areas of democracy and voting: the spatial autoregressive (SAR) model and the spatial error (SER) model. Since their spatial econometric models reduce estimation bias by allowing observations that are spatially connected to borrow useful information from each other, they are instrumental in finding new empirical evidence related to democracy and voting behavior. Despite this important advance over previous strategies for handling spatial dependence in these particular areas, no spatial econometric model has ever been applied to the study of interstate conflict, a prominent research area in political science (Neumayer and Plumper, 2010, p. 147).

This research gap may be attributed to two methodological issues. First, the conventional SAR model assumes that the dependent variable is continuous and relies mainly on maximum likelihood for estimation; this assumption creates a problem for international conflict scholars who view their outcome variables as a binary choice (e.g., war versus peace). Second, it is challenging to model spatial dependence in an analysis where the units are pairs of countries (or dyads in the terminology of political scientists), given that spatial dependence could come from either side of a dyad and probably even go beyond that (Gartzke and Gleditsch, 2008). For the majority of quantitative studies on interstate conflict, the dyad is the favored unit of analysis due to the belief that this level of analysis better reflects the strategic interactions between two states (e.g., Bremer, 1992; Bennett and Stam, 2000). On the other hand, as King (2001, p. 498) aptly points out, "dyadic observations in international conflict data have complex dependence structures."

Not surprisingly, economists have long sought not only to examine the causes, economic or not, of international conflict and its consequences, but also to develop economic measures and

9

policy tools to contain conflicts. Early works by Pigou (1921) and Robbins (1937) looked into international economic interdependence and the risk of war. Isard (1994) and Polachek (1994) first formalized "peace economics" as a research field. In recent years, war and peace have become a more active research area; its relationships with economic factors, such as international trade (e.g., Polachek, 1997), business cycles (e.g., Hess and Orphanides, 2001a; Blomberg and Hess, 2002), growth (e.g., Koubi, 2005), and state fiscal capacity and taxation (e.g., Besley and Persson, 2008), have been vigorously investigated by economists.

This study extends the spatial origin-destination modeling (hereafter referred to as spatial OD model) proposed by LeSage and Pace (2008) to probit regression, and introduces a Bayesian approach to its estimation. This extended model, after necessary tuning to accommodate the unique features of dyadic interstate conflict data, is then used to explore the impact of spatial dependence on the propensity for militarized interstate dispute (MID) initiation among dyads. Also, this approach to conflict data enables us to examine the performance of conflict-inducing factors in the presence of spatial effect terms in a probit regression model. In doing so, this research attempts to increase the knowledge of dyadic conflict studies in the disciplines of economics and political science.

The rest of this study is organized as follows. Section 2.3 reviews the original spatial OD model. Section 2.4 discusses the technical difficulties in applying the spatial OD model to binary dependent variables. A Bayesian approach to probit spatial OD regression model is presented in Section 2.5. In Section 2.6, the proposed spatial OD probit model is applied to investigate conflict initiation among European countries during the WWII period. Section 2.7 reports the

10

empirical results and discusses their implications, and Section 2.8 offers some concluding remarks.

2.3 Spatial Econometric Modeling of Origin-Destination Flows: the Continuous Case

2.3.1 Structure of Origin-Destination Models

Since OD flows are directional, one pair of regions will yield two observations where their origin and destination status can be reversed. Therefore, if n regions are considered under a spatial OD model, the number of observations becomes $n^2 = N$. Here, an n by n square matrix $\mathbf{Y}$ is used to denote interregional flows from each of the n origin regions to each of the n destination regions, with each column recording a specific origin's outflows to each of the n potential destination regions and each row corresponding to the inflows toward a given destination from each of the n potential origins. Specifically, the OD flow matrix is organized as follows:

$$
\begin{pmatrix}
o_1 \to d_1 & o_2 \to d_1 & \cdots & o_n \to d_1 \\
o_1 \to d_2 & o_2 \to d_2 & \cdots & o_n \to d_2 \\
\vdots & \vdots & \ddots & \vdots \\
o_1 \to d_n & o_2 \to d_n & \cdots & o_n \to d_n
\end{pmatrix}
$$

It is worth noting that the diagonal elements of this OD flow matrix represent intra-regional rather than inter-regional flows. They are often set to zero when the objective is to model inter-regional flows (e.g., Fischer et al., 2006).[1] To reflect an "*origin-centric* ordering" of OD flows

---

[1] In the study of state-level population migration flows, LeSage and Pace (2008, p. 960) create a separate model for flows on the main diagonal of the flow matrix by setting all elements of the covariates corresponding to the diagonal of the flow matrix equal to zero, thus "prevent[ing] these variables from entering the interregional migration flow model" while including additional explanatory variables for the intra-regional model.

11

(LeSage and Pace, 2008, p. 944), the matrix $\mathbf{Y}$ is then vectorized into an N by 1 matrix y (i.e., $y = vec(\mathbf{Y})$).[2]

In a typical spatial interaction model, where each region is treated as an observation, explanatory variables that represent k characteristics for each of the n regions are expressed in an n by k matrix, X. In keeping with the *origin-centric* arrangement of y, this X matrix is stacked n times in a spatial OD model to form an N by k matrix which tallies destination characteristics and is thereby denoted $\mathbf{X_d}$. LeSage and Pace note that $\mathbf{X_d}$ and X are related via a Kronecker product operation so that $\mathbf{X_d} = \iota_n \otimes X$, where ι_n is an n by 1 vector of ones (p. 945). Similarly, $X \otimes \iota_n$ would produce another N by k matrix, $\mathbf{X_o}$ which contains origin characteristics. In the matrix $\mathbf{X_o}$, the characteristics of each region are repeated n times before being stacked together, thus resulting in a matrix with dimensions of N by k.

In the same fashion as would occur by stacking the matrix $\mathbf{Y}$ to form the vector y, an N by 1 vector g recording the distances from origins to destinations is formed by stacking the n columns of an n by n OD distance matrix $\mathbf{G}$ into a variable vector.

So far, the resulting regression model assumes the following form:

$$y = \alpha \iota_N + \mathbf{X_d}\beta_d + \mathbf{X_o}\beta_o + \gamma g + \varepsilon, \qquad (2.1)$$

where α is the parameter associated with the constant term vector ι_N; β_d and β_o are the k by 1 parameter vectors associated with the covariate matrices $\mathbf{X_d}$ and $\mathbf{X_o}$, respectively; γ is the scalar parameter measuring the effect of the distance variable g; and ε is an N by 1 vector of errors under the assumption that $\varepsilon \sim N(0, \sigma^2 \mathbf{I_N})$.

---

[2] The *vec* operator converts a matrix into a column vector by stacking its columns in sequence.

Equation (2.1) represents a spatial OD model under the assumption of independent observations. It should be noted that although the distance between regions in each dyad is still included as an explanatory variable in a spatial OD model, it is considered to be inadequate for capturing spatial dependence between dyads. The inadequacy has urged econometricians to review the potential sources of dependence and thus explore more appropriate ways to model them. LeSage and Pace's (2008) study sheds new light on this exploration. Based on an omitted variables argument, LeSage and Pace demonstrate that if unobserved forces or missing covariates exert a similar impact on "neighboring" observations, including spatial lags of the dependent variable would be useful in capturing dependence among OD pairs. Furthermore, they argue that when it comes to the forms of dependence in the case of OD flows, "neighboring regions include neighbors to the origin, neighbors to the destination, and perhaps a link between neighbors of the origin and neighbors of the destination" (p. 947).

2.3.1.1 Multiple Sources of Spatial Dependence in Origin-Destination Models

To bring the idea of capturing spatial dependence using spatial lags of the dependent variable into line with the possible types of dependence that could exist among OD pairs, LeSage and Pace (2008) extend the spatial OD model above by introducing spatial lags defined through three spatial connectivity structures. To be exact, their spatial OD model is expressed as:

$$y = \rho_d \mathbf{W}_d y + \rho_o \mathbf{W}_o y + \rho_w \mathbf{W}_w y + \alpha \iota_N + \mathbf{X_d} \beta_d + \mathbf{X_o} \beta_o + \gamma g + \varepsilon . \qquad (2.2)$$

Equation (2.2) includes three spatial weight matrices: $\mathbf{W}_d$, $\mathbf{W}_o$, and $\mathbf{W}_w$. They are built upon the more familiar row-standardized, n by n first-order contiguity weight matrix $\mathbf{W}$, but are adapted to the neighboring relationships unique to OD flows.

13

First, $\mathbf{W}_d$ equals $\mathbf{I}_n \otimes \mathbf{W}$. This N by N weight matrix embodies the notion that factors causing flows from an origin to a destination may bring about similar flows to nearby destinations; accordingly, the spatial lag $\mathbf{W}_d y$ attempts to pick up this type of destination-based dependence by the use of average flows from one origin to the neighbors of a given destination. Using similar reasoning, a second N by N weight matrix $\mathbf{W}_o = \mathbf{W} \otimes \mathbf{I}_n$ is developed in order to reflect origin-based dependence and the spatial lag $\mathbf{W}_o y$ measures an average of flows into one destination from the neighborhood of an origin.

Furthermore, the idea of removing origin-based and destination-based dependence in sequence via the product $(\mathbf{I}_N - \rho_d \mathbf{W}_d)(\mathbf{I}_N - \rho_o \mathbf{W}_o)$ motivates the introduction of a third dependence structure.[3] Termed as origin-to-destination dependence, this third type is modeled by the spatial weight matrix $\mathbf{W}_w = \mathbf{W} \otimes \mathbf{W}$. Since $\mathbf{W}_w$ represents a second-order connectivity between the neighborhood of an origin and the neighborhood of a destination, the spatial lag $\mathbf{W}_w y$ indicates an average of flows from the neighborhood of an origin to the neighborhood of a destination. In light of the relationship among the three spatial parameters implied by the product term described above, LeSage and Pace (2008) consider both restricted and unrestricted models. With the restriction imposed that $\rho_w = -\rho_d \rho_o$, Equation (2.2) yields a "successive spatial filtering" (p. 955). However, when this restriction is relaxed, Equation (2.2) represents a more generalized model specification for spatial OD modeling. The model comparison in LeSage and Pace (2008) indicates that the generalized model specification outperforms all other alternative models, including the one subject to the restriction of $\rho_w = -\rho_d \rho_o$.

2.3.1.2 Estimation of Spatial Origin-Destination Models

---

[3] On p. 954, LeSage and Pace (2008) demonstrate the third type of spatial connectivity by expanding the product term, $(\mathbf{I}_N - \rho_d \mathbf{W}_d)(\mathbf{I}_N - \rho_o \mathbf{W}_o) = \mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o + \rho_d \rho_o \mathbf{W}_d \mathbf{W}_o = \mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w$. The authors also note that the order of this operation does not affect the result.

14

It is important to note that Equation (2.2) is subject to endogeneity problems due to the inclusion of spatial lags. To make this point, it will suffice to show that in general, $Cov(\mathbf{W}_d y, \varepsilon | \mathbf{X}) = \sigma^2 (\mathbf{W}_d (\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)^{-1}) \neq 0.$[4] This endogeneity problem is potentially worse with multiple spatial lags. Hence, OLS estimates of the Equation (2.2) would be inconsistent in most cases.[5] A commonly used estimation method is maximum likelihood.

The spatial OD model in Equation (2.2) implies the reduced form equation:

$$y = (\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)^{-1}(\alpha \iota_N + \mathbf{X_d} \beta_d + \mathbf{X_o} \beta_o + \gamma g)$$

$$+ (\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)^{-1} \varepsilon . \tag{2.3}$$

Letting $\mathbf{X} = (\iota_N, \mathbf{X_d}, \mathbf{X_o}, g)$, $\boldsymbol{\beta} = (\alpha, \beta_d', \beta_o', \gamma)'$ and $\boldsymbol{A} = (\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)$, the notation can be simplified:

$$y = \boldsymbol{A}^{-1} \mathbf{X} \boldsymbol{\beta} + \boldsymbol{A}^{-1} \varepsilon . \tag{2.4}$$

The log likelihood function for this model up to an irrelevant constant is:

$$\ln L(\rho_d, \rho_o, \rho_w, \boldsymbol{\beta}, \sigma^2) = -\frac{N}{2} \ln \sigma^2 + \ln |\boldsymbol{A}| - \frac{1}{2\sigma^2} \varepsilon(\theta)' \varepsilon(\theta) \tag{2.5}$$

where $\varepsilon(\theta) = \boldsymbol{A} y - \mathbf{X} \boldsymbol{\beta}$ and $|\boldsymbol{A}|$ refers to the determinant of the matrix $\boldsymbol{A}$.

2.3.2 Strengths and Limitations of the Original Spatial Origin-Destination Model

---

[4] $Cov(\mathbf{W}_d y, \varepsilon | \mathbf{X}) = E\big((\mathbf{W}_d (\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)^{-1} \varepsilon \cdot \varepsilon')\big)$
$= \mathbf{W}_d (\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)^{-1} E(\varepsilon \cdot \varepsilon') = \sigma^2 (\mathbf{W}_d (\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)^{-1})$.
[5] Lee (2002) shows that under certain regularity conditions, the OLS estimates of a SAR model can be consistent in large group interactions.

15

The validity of the conventional spatial interaction models, which include a geographical distance variable as a way of treating spatial dependence in interregional flows, has long been questioned. Furthermore, the fact that each observation of OD flows involves two regions (or locations) may also engender dependence among the observations in multiple ways. Indeed, LeSage and Pace (2008, p. 942) concur with Griffith and Jones' insight (1980, p. 190) that flows stemming from an origin are "enhanced or diminished in accordance with the propensity of emissiveness of its neighboring origin locations" whereas flows towards a destination are "enhanced or diminished in accordance with the propensity of attractiveness of its neighboring destination locations." In line with this reasoning, LeSage and Pace (2008) propose a revised model with three spatial weight matrices (as shown in (2.2) above), which are designed to capture origin-based, destination-based, and origin-to-destination-based dependence among OD flow pairs, and their modeling is in a fashion compliant with the standard SAR model. They also provide an illustration of the spatial OD model, using data on state-level population migration flows, and find evidence for origin-, destination- and origin-to-destination-based dependence in the population migration flows between U.S. states.[6]

Since LeSage and Pace's spatial OD model effectively deals with different types of potential dependence in data featuring a directional flow, it can be applied to the realm of international conflict where spatial dependence among dyadic pairs of countries is prevalent. In a dyadic analysis of interstate war,[7] an incidence of war can be viewed as a directional "flow" from the initiator country (i.e., the origin) to the target country (i.e., the destination). However, in practice, the spatial OD model is not readily applicable to a dyadic analysis of international conflict owing to the binary nature of the data. Originally, spatial OD models were set up to cope

---

[6] In their sample, LeSage and Pace include only the 48 contiguous states and the District of Columbia.

[7] In this study, the terms "war" and "conflict" are used interchangeably.

with continuous outcome variables that are generated by spatially dependent processes; accordingly, maximum likelihood estimation (MLE) was considered the most appropriate method for estimating the model parameters. Unfortunately, the binary nature of dependent variables in conflict data renders MLE computationally ineffective. When dealing with binary dependent variables, researchers normally follow a latent variable approach by transforming the nonlinear relationship into a linear one with a link function.

In a similar manner, this study turns to a probit link, and extends LeSage and Pace's spatial OD model by incorporating the spatial weight structures into a standard probit model. However, an ML estimator becomes unfit, once terms of "lagged" spatial effects are introduced into the probit model. For this reason, a Bayesian approach is employed instead to estimate the new model using a probit link. The next two sections offer the technical details used to bridge the spatial OD modeling with a probit regression model, as well as the Bayesian procedure exploited to estimate parameters of interest.

2.4 Building a Spatial OD Probit Model

When investigating interregional flows or interactions, the researcher often needs to model dependent variables that represent a binary choice outcome. For instance, interregional collaboration behavior, mobility, and firms' investment decisions are all binary choices that may exhibit spatial dependence. In these scenarios, it would be impractical to apply the spatial OD model designed for continuous outcome variables.

In applied research, probit models are frequently used for handling binary dependent variables. However, standard probit models do not account for spatial dependence. Given its

17

potential applications, efforts have been attempted to develop a probit estimator for spatially correlated binary data. For example, McMillen (1992) observes that "[p]robit and other limited dependent variable models are neglected in the spatial autocorrelation literature"; he ascribes this neglect to "the difficulty of adapting maximum-likelihood estimators to models with dependent observations." In his study, McMillen presents a probit version of the SAR model and proposes an EM algorithm for estimation, but he also notes that this estimation procedure "is not easy." One of the drawbacks of McMillen's EM algorithm is that it does not readily provide consistent covariance matrix estimates (p. 347). Drawing on work by Albert and Chib (1993), LeSage (2000) proposes a Bayesian approach to SAR probit models which demonstrates several advantages over the method set forth by McMillen.

Like a standard probit, the spatial OD probit model proposed in this study relies on the use of a latent variable y^* to model the binary dependent variable of interest. The unobserved, latent variable y^* is then modeled using the spatial OD flow regression relation:[8]

$$y^* = \rho_d \mathbf{W}_d y^* + \rho_o \mathbf{W}_o y^* + \rho_w \mathbf{W}_w y^* + \alpha \iota_N + \mathbf{X_d} \beta_d + \mathbf{X_o} \beta_o + \gamma g + \varepsilon \qquad (2.6)$$

Furthermore, without loss of generality, σ^2 is assumed to be equal to one (i.e., $\varepsilon \sim N(0, \mathbf{I_N})$) for identification. The only difference between the setup of this spatial OD probit and a standard probit is the addition of spatial lag terms to the regression of the latent variable.

Equation (2.6) can be written in its reduced form as follows:

$$y^* = \boldsymbol{A}^{-1} \mathbf{X} \boldsymbol{\beta} + \boldsymbol{A}^{-1} \varepsilon \qquad (2.7)$$

---

[8] Similar ideas of model building based on lags of the latent variable can also be found in dynamic probit models (e.g., Davutyan and Parke, 1995; Dueker, 1999).

where $\mathbf{X\beta}$ and $\mathbf{A}$ are defined in the same way as in Equation (2.4) above. The latent variable y^* is linked to the observed binary variable y through the following measurement equation:

$$y_i = 1 \text{, if } y_i^* \geq 0$$

$$y_i = 0 \text{, if } y_i^* < 0 \tag{2.8}$$

The covariance matrix $E\big((A^{-1}\varepsilon)(A^{-1}\varepsilon)'\big) = (A'A)^{-1}$ implies that the error terms are heteroskedastic as well as autocorrelated. As McMillen points out, "[h]eteroskedasticity can be relatively benign in a model with a continuous dependent variable, but it is a serious problem in a discrete dependent variable model (p. 339)." Due to the heteroskedasticity innate to model specification, an ML estimator for a spatial probit suffers from inconsistency. This problem becomes clear when one inspects the marginal probabilities of a spatial probit model. Specifically, in a spatial OD probit model, the marginal probabilities are:

$$\Pr(y_i = 1|\mathbf{X}) = \Pr([A^{-1}]_i\mathbf{X\beta} + [A^{-1}]_i\varepsilon \geq 0)$$

where $[A^{-1}]_i$ is defined to be the ith row of the inverse matrix A^{-1}, or

$$\Pr(y_i = 1|\mathbf{X}) = \Pr\left(u_i \leq \frac{[A^{-1}]_i\mathbf{X\beta}}{\sigma_i}\right) \tag{2.9}$$

using the marginal distribution from a multivariate normal with variance-covariance matrix $(A'A)^{-1}$.

In expression (2.9), σ_i stands for the square root of the ith diagonal element of the variance-covariance matrix and u_i follows a standard normal. As $\sigma_i \neq \sigma_0$ for all i, this structurally built-in heteroskedasticity implies that the ML estimator for standard probit would be

inconsistent in the spatial model case. Moreover, the fact that the error terms $\boldsymbol{A}^{-1}\varepsilon$ are correlated aggravates the inappropriateness of applying the standard probit estimator in a spatial probit.

To circumvent the aforementioned difficulties, LeSage (2000) proposes a Bayesian MCMC procedure to estimate the SAR probit model. In order to generate the posterior distributions for the parameters of interest, the MCMC sampler combines Albert and Chib's (1993) data augmentation approach, which introduces N latent variables as parameters that can be estimated using Gibbs sampling, with the Metropolis-Hastings-within-Gibbs method. Gelfand and Smith (1990) show that Gibbs sampling produces a Markov Chain whose stationary distribution is the true joint distribution of the parameters. For this reason, this Bayesian estimation procedure is regarded as more desirable. LeSage's approach provides a general framework for estimating probit models that explicitly incorporate spatial dependence in model specification.

This study adapts the Bayesian procedure suggested by LeSage to a family of models involving spatial OD flows, among which the SAR probit model can be viewed, from a technical standpoint, as a special case. Technically speaking, this adaption rests on the inclusion of two additional spatial lag parameters into LeSage's Bayesian SAR probit model, similar to the OD model of LeSage and Pace (2008), although the three weight matrices present in the spatial OD probit model have quite different denotations from that of the weight matrix applied in the SAR probit model. In principle, once the conditional distribution of the latent observations is derived, the problem at hand reduces to a spatial regression model which is of a more typical form for simulation, and from there, the conditional distributions for all model parameters can be derived sequentially.

2.5 Bayesian Modeling

In a spatial OD probit model, spatial interdependence induces a truncated multivariate normal distribution (TMVN) for the latent parameters y^* as shown in (2.10).

$$y^* \sim TMVN(A^{-1}X\beta, (A'A)^{-1}) \tag{2.10}$$

Hence, the conditional prior density of y^* takes the form:

$$\pi(y^*|\beta, \rho_d, \rho_o, \rho_w) \propto |A| \, exp\left(-\frac{1}{2}(Ay^* - X\beta)'(Ay^* - X\beta)\right) \tag{2.11}$$

where, as noted earlier, $|A|$ is defined to be the determinant of the matrix A.

To conduct a Bayesian analysis, one starts by assigning reasonable prior distributions for the parameters, and then, given the data, derives the corresponding posterior distribution for each parameter. This study follows the common practice in Bayesian spatial modeling by allowing β to assume the form of a multivariate normal distribution, and then applying diffuse priors on the spatial lag parameters, ρ_d, ρ_o, and ρ_w (LeSage and Pace 2009, p. 221). To be specific,

$$\beta \sim MVN(0, T), \quad T = I_k \cdot 10^4$$

$$\rho_i \sim U(-1, 1), \quad i = d, o, w \tag{2.12}$$

where k denotes the number of coefficient parameters.

Given the prior densities specified above, it follows from Bayes' Theorem that,

$$p(\beta, \rho_d, \rho_o, \rho_w, y^*|y) \propto p(y|\beta, \rho_d, \rho_o, \rho_w, y^*) \cdot \pi(\beta, \rho_d, \rho_o, \rho_w, y^*) \tag{2.13}$$

21

Due to the fact that given y^*, $\boldsymbol{\beta}$, ρ_d, ρ_o, and ρ_w become redundant in defining the conditional density of y, along with the assumed independence among $\boldsymbol{\beta}$, ρ_d, ρ_o, and ρ_w, we have

$$p(\boldsymbol{\beta}, \rho_d, \rho_o, \rho_w, y^*|y) \propto p(y|y^*) \cdot \pi(y^*|\boldsymbol{\beta}, \rho_d, \rho_o, \rho_w) \cdot \pi(\boldsymbol{\beta}) \cdot \pi(\rho_d) \cdot \pi(\rho_o) \cdot \pi(\rho_w) \qquad (2.14)$$

Here, $\pi(\cdot)$ and $p(\cdot)$ are used to distinguish between prior and posterior densities.

2.5.1 The Conditional Posterior Distribution of $\boldsymbol{\beta}$

The conditional posterior density of $\boldsymbol{\beta}$, given ρ_d, ρ_o, ρ_w, and y^*, is proportional to the joint posterior density (2.14) with ρ_d, ρ_o, ρ_w, and y^* held constant.

$$p(\boldsymbol{\beta}| \rho_d, \rho_o, \rho_w, y^*, y) = \frac{p(\boldsymbol{\beta}, \rho_d, \rho_o, \rho_w, y^*|y)}{p(\rho_d, \rho_o, \rho_w, y^*|y)} \propto p(\boldsymbol{\beta}, \rho_d, \rho_o, \rho_w, y^*|y)$$

$$\propto \pi(y^*|\boldsymbol{\beta}, \rho_d, \rho_o, \rho_w) \cdot \pi(\boldsymbol{\beta})$$

$$\propto |\boldsymbol{A}| \, exp\left(-\frac{1}{2}(\boldsymbol{A}y^* - \mathbf{X}\boldsymbol{\beta})'(\boldsymbol{A}y^* - \mathbf{X}\boldsymbol{\beta})\right) \cdot |T|^{-\frac{1}{2}} exp\left(-\frac{1}{2}\boldsymbol{\beta}'T^{-1}\boldsymbol{\beta}\right) \qquad (2.15)$$

where the last step follows by substituting in (2.11) and (2.12).

After some algebraic manipulation, this can be simplified to read as follows:

$$\propto |\boldsymbol{A}| \cdot |T|^{-\frac{1}{2}} \cdot exp\left\{-\frac{1}{2}[(\boldsymbol{\beta} - \Xi\mathbf{X}'\boldsymbol{A}y^*)'\Xi^{-1}(\boldsymbol{\beta} - \Xi\mathbf{X}'\boldsymbol{A}y^*) + b]\right\}$$

$$\Xi = (\mathbf{X}'\mathbf{X} + T^{-1})^{-1}, \text{ and } b = y^{*'}\boldsymbol{A}'\boldsymbol{A}y^* - (\Xi\mathbf{X}'\boldsymbol{A}y^*)'\Xi^{-1}(\Xi\mathbf{X}'\boldsymbol{A}y^*)$$

and is proportional to

$$\propto exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \Xi\mathbf{X}'\boldsymbol{A}y^*)'\Xi^{-1}(\boldsymbol{\beta} - \Xi\mathbf{X}'\boldsymbol{A}y^*)\right\} \qquad (2.16)$$

which is the kernel of a multivariate normal distribution. Thus, we can infer that the conditional posterior distribution of $\boldsymbol{\beta}$ is:

$$\boldsymbol{\beta}|\rho_d, \rho_o, \rho_w, y^*, y \sim MVN(\Xi\mathbf{X}'\boldsymbol{A}y^*, \Xi),$$

and this is equivalent to

$$\boldsymbol{\beta}|\rho_d, \rho_o, \rho_w, y^* \sim MVN(\Xi\mathbf{X}'\boldsymbol{A}y^*, \Xi). \qquad (2.17)$$

2.5.2 The Conditional Posterior Distributions of the ρ's

Similarly, the conditional posteriors for ρ's can be derived. Using the same argument as in (2.15), together with (2.11) and (2.12), we can write,

$$p(\rho_d|\boldsymbol{\beta}, \rho_o, \rho_w, y^*, y) \propto \pi(y^*|\boldsymbol{\beta}, \rho_d, \rho_o, \rho_w) \cdot \pi(\rho_d)$$

and this implies that

$$p(\rho_d|\boldsymbol{\beta}, \rho_o, \rho_w, y^*) \propto \boldsymbol{I}(-1 < \rho_d < 1) \cdot |\boldsymbol{A}| \, exp\left(-\frac{1}{2}(\boldsymbol{A}y^* - \mathbf{X}\boldsymbol{\beta})'(\boldsymbol{A}y^* - \mathbf{X}\boldsymbol{\beta})\right)$$

$$(2.18)$$

where $\boldsymbol{I}(-1 < \rho_d < 1)$ denotes an indicator function that takes the value 1 if ρ_d is in the open interval $(-1,1)$. This restriction is imposed because of the model assumptions and computational

feasibility.[9] As independence is presumed among $\boldsymbol{\beta}$, ρ_d, ρ_o, and ρ_w, the conditional posterior distributions for ρ_o and ρ_w take the same form as (2.18). However, unlike the conditional posterior for $\boldsymbol{\beta}$ derived above, the conditional posterior distributions for the ρ's do not have a known form, and thus cannot be sampled directly. For this reason, this study turns to a Metropolis-Hastings algorithm for sampling.

2.5.3 Sampling of y^*

For brevity, this study will, from this point on, express the latent parameters as $y^* \sim TMVN(\mu(\mathbf{X}), \Omega)$, with $\mu(\mathbf{X}) = \boldsymbol{A}^{-1}\mathbf{X}\boldsymbol{\beta}$ denoting the mean and $\Omega = (\boldsymbol{A}'\boldsymbol{A})^{-1}$ representing the variance-covariance matrix. As illustrated earlier, $y^* \sim TMVN(\mu(\mathbf{X}), \Omega)$ subject to a vector of inequality restrictions $a \leq y^* \leq b$. Here, a and b depend on the observed values of 0 and 1 for the elements of y; thus, individual elements of a may be $-\infty$ and individual elements of b may be $+\infty$ given that y_i^* is only truncated on one side. The fact that the covariance matrix Ω is not diagonal implies that the marginal distributions for individual elements of y^* are not univariate truncated normal, thus ruling out the option of sampling from a sequence of univariate truncated normal distributions in order to obtain simulations for the individual elements y_i^*.

To circumvent this problem, LeSage and Pace (2009) adopt the approach set forth by Geweke (1991) and generate simulations for individual elements y_i^* relying on the conditional distribution of y_i^* given all other elements of y^* (denoted as y_{-i}^*). Showing that sampling from

---

[9] Being an autoregressive coefficient, $|\rho_d|$, is required to be less than 1 for the model to be stable. Moreover, LeSage and Pace (2009, p. 128) note that a feasible range for the spatial parameter is $(1/\lambda_{\min}, 1/\lambda_{\max})$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalues of the spatial weight matrix. In this application, the three spatial weight matrices all have a minimum eigenvalue of -1 and a maximum eigenvalue of 1.

$y^* \sim TMVN(\mu(\mathbf{X}), \Omega)$ subject to inequality restrictions $a \leq y^* \leq b$ is equivalent to

constructing samples from an N-variate normal distribution $z \sim N(0, \Omega)$ subject to the inequality

restrictions $\underline{c} \leq z \leq \overline{c}$, where $\underline{c} = a - \mu(\mathbf{X})$ and $\overline{c} = b - \mu(\mathbf{X})$, Geweke works with the precision

matrix Ψ (i.e., the inverse of the variance-covariance matrix) of the truncated multivariate

normal distribution from which samples of y^* need to be drawn. Geweke establishes that

$E(z_i|z_{-i}) = \gamma_{-i} z_{-i}$, where $\gamma_{-i} = -\Psi_{-i}/\Psi_{i,i}$, with Ψ_{-i} being the ith row of Ψ excluding the ith

element and $\Psi_{i,i}$ the scalar in the ith row and column. For the spatial OD probit model, the

precision matrix is $\Psi = \Omega^{-1} = \mathbf{A}'\mathbf{A}$.

LeSage and Pace (2009) apply the result stated above to sample individual z_i from a

normal conditional distribution, and then acquire samples of y^* using the transformation $y^* = \mu(\mathbf{X}) + z$. More formally, the normal conditional distribution of z_i takes the following form:

$$z_i|z_{-i} = \gamma_{-i} z_{-i} + h_i v_i$$

$$h_i = (\Psi_{i,i})^{-1/2} \tag{2.19}$$

and v_i is sampled from a standard normal subject to the following truncation criteria:

$$(\underline{c_i} - \gamma_{-i} z_{-i})/h_i \leq v_i \leq (\overline{c_i} - \gamma_{-i} z_{-i})/h_i$$

$$\underline{c_i} = -\infty \text{ and } \overline{c_i} = -\mu(\mathbf{X})_i \text{ for } y_i = 0$$

$$\underline{c_i} = -\mu(\mathbf{X})_i \text{ and } \overline{c_i} = +\infty \text{ for } y_i = 1 \tag{2.20}$$

A Gibbs sampling procedure is taken to produce the vector z, where in each iteration the

most recent samples of $z_1, z_2, \ldots, z_{i-1}, z_{i+1}, z_{i+2}, \ldots, z_N$ are used for simulating z_i. Constructed

this way, the vector z, in conjunction with the relation $y^* = \mu(\mathbf{X}) + z$, will give the vector y^*,

which is then used to simulate draws from the conditional posterior distributions of the other

model parameters. It should be noted that Geweke's Gibbs sampler implementation for the truncated multivariate normal distribution sometimes shows slow convergence (Rodriguez-Yam, Davis and Scharf, 2004).[10] And this problem becomes more acute when sampling a high-dimensional truncated multivariate normal with non-diagonal covariance matrix.

To improve convergence, this paper employs the Geweke-Hajivassiliou-Keane (GHK) multivariate normal simulator (Geweke, 1991; Hajivassiliou, 1990; and Keane, 1994), a technique that samples recursively from truncated univariate normals after a Cholesky transformation. The GHK algorithm works as follows.[11]

Instead of drawing from the original distribution of the latent variable,

$$y^* \sim N(\mu(\mathbf{X}), \Omega) \qquad s.t. \qquad a \leq y^* \leq b,$$

I draw a random vector

$$\epsilon \sim N(\mathbf{0}, \mathbf{I}) \qquad s.t. \qquad a^* \equiv a - \mu(\mathbf{X}) \leq C\epsilon \leq b - \mu(\mathbf{X}) \equiv b^*$$

where C denotes the lower triangular Cholesky factor of Ω, $CC' = \Omega$.

Due to the triangular structure of C, the restrictions on ϵ are recursive. Specifically,

$$\epsilon_1 \sim N(0,1) \qquad s.t. \qquad a_1^*/C_{11} \leq \epsilon_1 \leq b_1^*/C_{11} ,$$

$$\epsilon_2 \sim N(0,1) \qquad s.t. \qquad (a_2^* - C_{21} \cdot \epsilon_1)/C_{22} \leq \epsilon_2 \leq (b_2^* - C_{21} \cdot \epsilon_1)/C_{22} ,$$

---

[10] Rodriguez-Yam, Davis and Scharf (2004) propose an efficient Gibbs sampler to the truncated multivariate normal distribution; but, their implementation is developed with a view to application to constrained linear regressions and therefore is not easily adopted in cases of high-dimensional truncated multivariate normals.
[11] Börsch-Supan and Hajivassiliou (1993) provide a detailed description of this recursive conditioning method.

and for $i = 2, \ldots, N$,

$$\epsilon_i \sim N(0,1) \qquad s.t. \qquad (a_i^* - \sum_{j=1}^{i-1} C_{ij}\, \epsilon_j)/C_{ii} \leq \epsilon_i \leq (b_i^* - \sum_{j=1}^{i-1} C_{ij}\, \epsilon_j)/C_{ii}\ .$$

Thus ϵ_i can be sampled sequentially from univariate truncated normals. The simulated vector ϵ and the relation $y^* = \mu(\mathbf{X}) + C\epsilon$ will give the desired truncated random vector y^*.

2.5.4 Implementing MCMC Sampling

Assigning arbitrary initial values for the parameters (denoted by the superscript (0)), this study sequentially samples the conditional distributions for the model parameters following the steps sketched below:

1. $p\left(\boldsymbol{\beta} \middle| \rho_d^{(0)}, \rho_o^{(0)}, \rho_w^{(0)}, y^{*(0)}\right)$, which is a multivariate normal distribution with mean and variance defined in (2.16). Label the sampled vector $\boldsymbol{\beta}$ as $\boldsymbol{\beta}^{(1)}$.

2. $p\left(\rho_d \middle| \boldsymbol{\beta}^{(1)}, \rho_o^{(0)}, \rho_w^{(0)}, y^{*(0)}\right)$, which can be acquired by means of a Metropolis-Hastings sampler based on a normal jumping density, along with rejection sampling in order to confine ρ_d to the $(-1, 1)$ interval. Label this updated value $\rho_d^{(1)}$.

3. $p\left(\rho_o \middle| \boldsymbol{\beta}^{(1)}, \rho_d^{(1)}, \rho_w^{(0)}, y^{*(0)}\right)$, which applies the same Metropolis-Hastings algorithm as in step (2). The newly updated value for ρ_d is used when making a draw for ρ_o.

4. $p\left(\rho_w \middle| \boldsymbol{\beta}^{(1)}, \rho_d^{(1)}, \rho_o^{(1)}, y^{*(0)}\right)$, which is similar to steps (2) and (3), except that now the updated values for both ρ_d and ρ_o are employed.

5. $p\left(y^* \middle| \boldsymbol{\beta}^{(1)}, \rho_d^{(1)}, \rho_o^{(1)}, \rho_w^{(1)}\right)$, which requires draws from the left- or right-truncated normal distributions defined in section 2.5.3.

27

Then, return to step (1) and replace the initial values $\boldsymbol{\beta}^{(0)}, \rho_d^{(0)}, \rho_o^{(0)}, \rho_w^{(0)}, y^{*(0)}$ with the updated values $\boldsymbol{\beta}^{(1)}, \rho_d^{(1)} \rho_o^{(1)}, \rho_w^{(1)}, y^{*(1)}$. This process is repeated to obtain a large sample of draws that can be used to make valid inferences with respect to the model parameters.

2.6 Spatial OD Probit Model of the Correlates of World War II in Europe

As a way of illustrating the utility of the proposed spatial OD probit model, this section looks into the causes of interstate conflict in Europe during the period 1933-1941.

2.6.1 Methodological Rationale for Spatial Modeling of Interstate Conflict

Due to its devastating destructiveness to human lives and socio-economic development, interstate war has been a very active research topic for both political scientists and economists. Not surprisingly, the former group tends to focus more on the causal factors of war (e.g., Russett and Oneal, 2001; Choi and James, 2005; Gartzke and Gleditsch, 2008), while the latter is more interested in examining the relationships between war and economic fundamentals (e.g., Besley and Persson, 2007 and 2008, on state fiscal policy and taxation; Hess and Orphanides, 1995 and 2001a, and Blomberg and Hess, 2002, on business cycles). Still, some economists are also inclined to investigate issues such as the impact of democratic institutions and election cycles on the propensity for war (e.g., Hess and Orphanides, 2001b).

It is important to note that most conflict scholars tend to treat each incident of interstate war as a separate and isolated event without giving due consideration to the interdependence that may exist among incidents of war. However, the assumption of independence among a series of war incidents is unfounded. Arguably, once a country is already at war, its ability and willingness to get involved in another one will be more or less affected. More generally, as a

28

directional development between two countries, a country's war decisions are conditional on the conflict behavior of other "connected" pairs of countries (Gartzke and Gleditsch, 2008). Gartzke and Gleditsch (2008) treat alliance ties as one of the important linkages for conflicts within dyads and concede that "many other sources of extra-dyadic ties and dependence other than alliance links are possible" (p. 32). Clearly, of the possible sources of dependence between interstate conflicts, the most fundamental link should arise from geographical proximity. After all, many alliances were also formed out of geo-strategic concerns.

Interstate war involves a strategic interaction between at least two states. For this reason, many political scientists prefer to use the dyad (i.e., a pair of states) as the unit of analysis. This level of analysis is considered to better reflect the mutual influence of states within the setting of war, and therefore to produce a more balanced assessment of the causes of interstate conflict. Furthermore, conflict scholars began to notice the importance of examining the initiation of war over involvement, since the former is a more useful concept in pinpointing the origin of interstate conflict. Hence, the notion of directed dyads, which differentiates between aggressor and victim, was introduced, allowing researchers to test directional hypotheses pertaining to the differential characteristics of potential initiator and target (Bennett and Stam, 2000). For instance, the use of directed dyads enables researchers to test contending theories such as democratic peace, economic interdependence, and policy preference similarity. The advantage of directed dyadic level analyses is well substantiated by Ray (2001, p. 374): "virtually the entire difference between the conflict proneness of jointly democratic pairs of states and other pairs might be accounted for by the aggressiveness or assertiveness of democratic states against autocratic states. Or, that difference could be produced entirely by the aggressiveness of autocratic states.

www.manaraa.com

Only directed dyadic level analyses can reveal where in between those extremes the 'real world' is located."

The directed dyadic research design has contributed to advancing the scientific knowledge of interstate conflict: characteristics of *identified* initiator and target are both essential for explaining conflict behavior. More importantly, this research design brings to light several potential sources of spatial dependence among conflict observations. Although conflict scholars include a geographical distance variable as a way of capturing the geopolitical nature of interstate war, it may, to some extent, even be considered a control for spatial dependence.[12] The effectiveness of this treatment of spatial dependence has, however, long been challenged, urging researchers to design better spatial econometric models (Anselin 1988).

Just like temporal dependence in time series data, spatial dependence will, if not handled properly, cause an estimator to be biased, and consequently lead to incorrect inferences with respect to model parameters. The consequences of ignoring this dimension of dependence in conflict studies are potentially more detrimental, given that geographical proximity is regarded as an explanatory variable of interest. In fact, political scientists have become increasingly aware of the neglect of spatial dependence in conflict studies. For example, Gleditsch and Ward (2000) introduce the concept of spatial dependence in to the study of war and examine the impact of spatial clustering on the relationship between democratization and war. Ward and Gleditsch (2002) incorporate the advances made in the implementation of conditional models for categorical variables in the statistical analysis of spatial patterns; they estimate an autologistic

---

[12] Hegre, Oneal and Russett (2009, p. 766) assert the necessity of including geographic distance besides contiguity in the conflict equation, acknowledging that the logic of including distance in the gravity model of trade, where distance is "a broad concept that may include anything that either facilitates or hinders the movement of goods and services," also applies to the conflict equation.

30

model of democracy and international war using pseudolikelihood and MCMC techniques. Beck, Gleditsch and Beardsley (2006) illustrate the SAR model using a simple social requisites model that explains democracy by GDP per capita and the SER model using an example that probes the political determinants of dyadic trade. They argue that the SAR model usually should be preferred to the SER model, and experiment with two different ways of designing the weight matrix **W** based on political economy notions of distance. In their recent book, *An Introduction to Spatial Regression Models in the Social Sciences*, Ward and Gleditsch (2008) utilize a SAR model to examine the distribution of democracy, voting outcomes, and the political determinants of trade flows. These aforementioned studies represent some notable efforts by political scientists to bring advances in spatial dependence modeling to bear on international studies and war studies. However, the simplicity of these models renders them less useful in modeling spatial interdependence in dyadic conflict data.

Gartzke and Gleditsch (2008) were the first to consider the application of spatial dependence in international conflict in a dyadic setting. They utilized "third order" and "fourth order" alliances, as well as a measure of a third state's location between two disputants in a conflict, to model the dependence-generating process. Their empirical results highlight the presence of potentially complex spatial dependency relations in dyadic conflict data, although the adequacy of the two types of linkages they use to measure spatial dependence warrants further scrutiny. It is worth noting that their construct of "inbetweenness" does not distinguish between a dyad having two or more large inbetweenness ratios in relation to other dyads that are involved in disputes, and a dyad with only one large inbetweeness score; this is because they only "consider the maximum inbetweenness score for each dyad" (p. 21). From a policy point of view, identifying the causal process underlying the initiation of war should be more informative

31

than just that of war involvement. Also, the approach in Gartzke and Gleditsch (2008) neglects to consider the intricacy of the types of spatial dependence involved in war initiation.

The solutions suggested for addressing spatial dependence in existing conflict studies share a common drawback. The spatial connectivity they depict does not conceptualize war as a directional strategic behavior between two states; it thereby fails to take into account all the possible sources of spatial dependence underlying the behavioral decision to wage war. It is plausible that the propensity for war initiation between a dyad is influenced by their geographic proximity, either through one party or both, to an ongoing nearby conflict and that this connectivity can be further shaped by a disputant's initiator or target status in a potential militarized conflict.

If we treat an observation of directed dyadic war data as a directional flow from the aggressor to the victim, then recent developments in spatial econometrics should help shed light on tackling the spatial interdependence among observations of war initiation. The rationale for spatial OD modeling of interregional flows can be applied to the study of international conflict. The likelihood of the initiation of war by a potential aggressor may be enhanced or diminished by its aggressive behavior towards the neighboring states of the potential target (i.e., target-based spatial dependence). The likelihood of being attacked by a potential aggressor may be enhanced or diminished by the belligerent behavior of the aggressor's neighboring states towards the potential victim (i.e., initiator-based dependence). Quite likely, the propensity for initiating a conflict is also affected by the aggressive behavior exhibited by the neighbors of the potential attacker towards the neighbors of the potential target (i.e., initiator-to-target-based dependence).

32

Therefore, a more comprehensive treatment of spatial dependence in directed dyads of conflict initiation requires modeling the linkages that exist between dyads in relation to the status of each dyad member in a conflict. In keeping with this argument, this study applies a spatial OD probit model to analyze the initiation of MIDs among 26 European countries during the WWII period. The empirical findings indicate that by directly modeling multiple sources of spatial dependence among dyads, as represented by the neighboring relationships that are constructed through the side of initiator, the side of target, as well as both sides, the spatial OD probit model reveals how war initiations are correlated in space; as a consequence, this study obtains estimates which more reliably reflect the impact of political factors contributing to the likelihood of war initiation.

2.6.2 Research Design

The initiation of conflict among 26 European countries during the WWII period is analyzed in the context of a spatial OD probit model. The choice of the sample data is based on the following two considerations: 1) conflict initiations were relatively more frequent and concentrated in this region during the time span under scrutiny; 2) this data selection is consistent with Beck, Gleditsch and Beardsley's (2006, p. 37) remark that a European sample includes "'high-quality' observations that we are relatively confident in." A list of sampled countries is reported in Appendix I.

2.6.2.1 Data and Modeling

In this subsection, this study discusses data sources and describes how the data are structured to fit the specification of a spatial OD probit model; this model embodies spatial

connectivity terms representing initiator (i.e., origin)-centric and target (i.e., destination)-centric dependence as well as a spatial link highlighting concurrent neighboring relationships across dyads between both initiators and targets.

The illustrative application under consideration focuses on the militarized interstate disputes among European countries during the period from 1933 to 1941, and investigates how spatial dependence involved in the dispute initiations in Europe affects the estimation of classical war regression models. The cutoffs for the study period are selected with a view towards two historical events, which define the study of the nine year period as being more homogenous and relatively plagued with warfare. To be precise, on January 30, 1933, Adolf Hitler was appointed the Chancellor of Germany, which signified a turning point in European diplomatic history, and on December 11, 1941, Hitler declared war on the United States. The involvement of the U.S. marked a new phase in the War and made it truly a world war in scale. In addition, data collection with respect to the explanatory variables used in the analysis is more available and reliable for European countries. In a word, this study believes that the period chosen is quite suitable for analyzing spatial dependence among dispute initiations.

All the data used in this analysis are retrieved from the Expected Utility Generation and Data Management Program (EUGene) Version 3.204.[13] With the directed dyad-year being the unit of analysis, the time frame examined in this study, and the region specified to be Europe, EUGene outputs 6,030 cases from the Correlates of War (COW) Militarized Interstate Disputes (MID) data with Side A noted as initiator and Side B noted as target. The MID data is used as the basis for creating the dependent variable. EUGene is also utilized to generate data for four

---

[13] EUGene is a free software, but copyrighted. It can be downloaded at http://www.eugenesoftware.org/. EUGene generates datasets for quantitative analysis of international relations, with country-year, directed-dyad-year, non-directed-dyad-year, and directed-dispute-dyad-year as the unit of analysis.

independent variables: (1) democracy and (2) national material capabilities for the initiator and the target states, respectively, (3) a measure of geographical distance within each dyad, and (4) alliance ties within dyads. Like many studies in spatial econometrics, this study aggregates the data over the time span in question to obtain a cross-national dataset, yielding 756 observations in total. Among the 756 directed dyads contained in the sample, many did not experience a MID during the study period.

To avoid the issue of possible reverse causality, values of the year 1932 are recorded for the explanatory variables.[14] As can be seen from the OD flow matrix presented in section 2.3.1, a self-directed pair is conventionally included for each region in a sample of interregional flows (i.e., those on the diagonal of the flow matrix) in order to utilize the model structure associated with spatial OD modeling. Since this study is only concerned with militarized interstate disputes, it is natural to code the initiation variable as zero for the self-directed dyads.[15] In addition, it should be noted that data is unavailable for Demark and Norway for the year 1941. After removing these two countries from the sample, the cross-sectional data covers 26 European countries and consists of 676 directed dyads (including 26 self-directed pairs).

2.6.2.2 Definition of Variables and Weight Matrices

The dependent variable, *cwinit*, measures whether there was an initiation of a militarized interstate dispute (MID) by the first state (i.e., State A) in a directed-dyad. A MID is "a set of

---

[14] As a robustness check, I also tried two other measures of the explanatory variables. One measure sets the variables at their respective mid-year value of the time span studied (i.e., year 1937's value). The other measure records the average value of each variable over the nine year interval for the conflict-free pairs and takes the values of the year previous to the first conflict initiation for each conflict dyad. ML estimation results are quite robust across all three measures of the covariates. Since the lagged value approach is more justifiable from a technical point of view, it is chosen for the empirical analysis in the paper.

[15] These dyads will be handled such that this choice is irrelevant. A more detailed discussion on operationalization issues regarding these dyads will be presented in Section 2.6.2.3.

interactions between or among states involving threats to use military force, displays of military force, or actual uses of military force" (Gochman and Maoz, 1984, p. 587; see also Jones, Bremer and Singer, 1996). It is coded as "1" if the first country (i.e., initiator) of a dyad initiated at least one MID against the second country (i.e., target) over the nine year period; otherwise, it is coded "0".[16]

This study employs four independent variables: democracy, national material capabilities, geographic distance, and alliance. Because the main purpose of this study is to illustrate modeling spatial dependence among binary observations of directional flows rather than to present an exhaustive model of interstate conflict and given the specific historical period under study, other potential explanatory variables such as economic interdependence are not included, which also helps save computation time. *polity1* and *polity2* record, respectively, the democracy score for the two members of a dyad (i.e., *polity1* for initiator and *polity2* for target). The data were originally compiled by the Polity IV project, and are now available as a user dataset in EUGene. The Polity dataset measures and aggregates five different aspects of democratic institutions in each country: (1) competitiveness of participation, (2) regulation of participation, (3) competitiveness of executive recruitment, (4) openness of executive recruitment and (5) constraints on the executive. The averaged democracy score runs from -10 to 10, ranging from the most autocratic (the least democratic) to the most democratic. Missing Polity IV values are replaced by Polity III values whenever the latter are available. The democratic peace proposition maintains that democratic countries in a dyad are less war prone due to the high levels of institutional constraints on the war decisions of their political leaders (see Russett and Oneal, 2001; Choi and James, 2005). In this study, the directed dyadic data structure allows us to

---

[16] In Appendix II, dyads that experienced more than one initiation are listed.

analyze separately how the regime type of a state affects its war behavior with respect to its role as a MID initiator or target.

cap1 and *cap2* measure the national material capabilities of each state in a dyad. These indicators are introduced to capture how an increase or a decrease in the national capabilities of the initiator or target state influences the likelihood of a MID initiation. Each state's national capabilities are assessed over six components as follows: total population, urban population, iron and steel production, energy consumption, military personnel, and military expenditure. More specifically, a composite index is computed by summing all the scores for each of the six capability components in a given year, converting each state's absolute component to a share of the international system, and then averaging across the 6 components (see Singer, Bremer, and Stuckey, 1972; Singer, 1987).

The *distance* variable measures the distance in miles between capitals, adjusted for contiguity. The default option in the data management program of EUGene relies on the distance calculation method proposed in Bueno de Mesquita's (1981) book, *The War Trap*. Conventionally, *distance* is included to test the hypothesis of geopolitics that the geographic location of countries is a key factor in determining foreign policy. Later on, it is suggested that the distance measure may also help capture the spatial dependence among war data (Hegre, Oneal and Russett, 2009), although its sufficiency has been called into question.

In the COW dataset, *alliance* is originally coded as a categorical variable, with 1=defense pact, 2=neutrality, 3=entente, and 4=no agreement. In this study, I follow the convention in the literature and convert it into a binary measure by combining the first three categories. Therefore, "1" denotes that dyad partners were allied and "0" otherwise.

37

In spatial econometrics, the physical distance between observations serves as a basis for constructing the weight matrix $\mathbf{W}$ which directly models the spatial connectivity structure. Although socio-economic distances are sometimes considered in specifying the spatial interaction structure in social studies, Gleditsch (2007) acutely points out that "[a]lthough states other than geographic neighbors can be important, we can identify the most important relationships between states by examining dependence determined by geographical proximity." Here, neighboring states are first identified, with $\mathbf{W}_{ij}$ ($i \neq j$) being coded as "1" if the distance between the two members of a dyad is zero; it is coded as "0" otherwise.[17] Based on this "neighbors" information, the weight matrix $\mathbf{W}$ is then row standardized to have row sums of one. Through the Kronecker product operations which were described earlier, three particular weight matrices $\mathbf{W}_d$, $\mathbf{W}_o$, and $\mathbf{W}_w$ are produced to reflect initiator-, target- and initiator-to-target-based dependence among dyads. Specifying the weight matrices based on geographic distances is also sensible when examining the spatial connectivity of conflict dyads. As Gartzke and Gleditsch (2008, p. 18) assert, being a determinant of dyadic interaction, "distance should also matter for the degree of dependence on other dyads." Nonetheless, it is worth noting that in the spatial econometrics literature, a number of ways other than contiguity have been used, in reference to the researcher's knowledge of, or theory about, the diffusiveness of spatial interaction, for defining neighboring relations and thus the weight matrix (e.g., LeSage and Pace, 2004).[18] Of

---

[17] Beck, Gleditsch and Beardsley (2006) define countries as connected if they are within 500 km of one another, a distance equivalent to 311 miles. Given that this study focuses on European countries, a shorter cutoff criterion of 212 miles was also applied in determining neighbors. The empirical results were similar to what are reported here.

[18] To alleviate the concern that the observed spatial connectivity among conflict initiation may only be a reflection of underlying alliance ties, I also constructed weight matrices based on alliance relationships and ran two other models: one model replaces the distance-based weight matrices with alliance-based weight matrices and the other includes both sets of weight matrices. It turns out that the model with alliance-based weight matrices provides a worse fit whereas the model employed in the paper provides the same fit as does the more inclusive model. These estimation results support the notion that strategic concerns trump political concerns.

course, different designs of the weight matrix affect estimation results differently. The row-standardized weight matrix **W** used in the present application is included in Appendix III.

In addition, the data for the explanatory variables are further arranged so as to comply with the notations developed in Section 2.3. Given that the explanatory variables contain both initiator and target characteristics, the distance between a pair, and the alliance ties between dyad partners, a 676 by 2 matrix that has *polity2* and *cap2* as its columns is partitioned out and labeled **X$_d$** to represent attributes of target countries only. Similarly, a second 676 by 2 matrix that includes *polity1* and *cap1* is denoted **X$_o$**, standing for initiator countries' characteristics.

2.6.2.3 Handling the Self-Directed Dyads

2.6.2.3.1 A Dummy Variable Approach

In keeping with the structure of the weight matrix **W**, self-directed pairs are included in the data by construction. However, the dependent variable for these self-directed pairs assumes a value of zero because only interstate conflicts are considered. Accordingly, the zero values of these observations are different in nature from those observed for non-conflict dyads. If we ignore this difference and treat the self-directed pairs the same as other dyads, the estimation results would be biased.

To resolve this problem, a dummy variable is utilized to keep the observations for self-directed dyads from biasing the estimation without changing the structure of the model. To be exact,

$$y^* = (\iota_N - d) \odot \rho_d \mathbf{W}_d y^* + (\iota_N - d) \odot \rho_o \mathbf{W}_o y^* + (\iota_N - d) \odot \rho_w \mathbf{W}_w y^* +$$

www.manaraa.com

$$(\iota_N - d) \odot \alpha \iota_N + (\iota_N - d) \odot \mathbf{X_d}\beta_d + (\iota_N - d) \odot \mathbf{X_o}\beta_o + \gamma g + \theta D + \varepsilon, \qquad (2.21)$$

where d is an N by 1 binary vector with its elements taking the value "1" for self-directed dyads and "0" otherwise. The term $(\iota_N - d)$ does not appear before the distance variable g because distances for self-directed pairs are recorded as zero in the first place. And it is the same with the alliance dummy D, which is coded as zero for self-directed dyads. Equation (2.19) can be simplified as

$$y^* = (\mathbf{I}_N - (\iota_N - d) \odot \rho_d \mathbf{W}_d - (\iota_N - d) \odot \rho_o \mathbf{W}_o - (\iota_N - d) \odot \rho_w \mathbf{W}_w)^{-1}$$
$$(\iota_N - d) \odot \mathbf{X}\boldsymbol{\beta} + \epsilon \qquad (2.22)$$

where $\epsilon = (\mathbf{I}_N - (\iota_N - d) \odot \rho_d \mathbf{W}_d - (\iota_N - d) \odot \rho_o \mathbf{W}_o - (\iota_N - d) \odot \rho_w \mathbf{W}_w)^{-1}\varepsilon$.

The above Hadamard products perform the operations of setting the values of all regressors, both conventional and spatial lags, to zero for the self-directed observations while maintaining intact the spatial structures.

2.6.2.3.2 An Elimination Approach

Admittedly, assigning zero values to the regressors for self-directed pairs is a double-edged tactic. It avoids the downside associated with the often-used method of only setting intra-regional flows to zero, while keeping the observed values of explanatory variables, a practice which might bias the estimated effects of the explanatory variables. However, with both the dependent and independent variables being recorded as zero for the self-directed pairs, this approach invites the possibility of underestimating the standard errors of the model.

Given that this study focuses on the spatial dependence among interstate conflicts and given that civil (intrastate) conflicts and interstate conflicts are believed to operate under different mechanisms, a more sensible way to deal with self-directed dyads in this context is to pull them out of the estimation procedure entirely, while preserving the primary dependence structure embodied in the weight matrices. To this end, this study makes use of a 'selection' matrix to take out of the weight matrices, $\mathbf{W}_d$, $\mathbf{W}_o$, and $\mathbf{W}_w$, respectively, the rows and columns corresponding to self-directed pairs, so the altered weight matrices are conformable to the data structure after those self-directed observations are extracted.

Before proceeding with the empirical analysis, this study first illustrates this "selection" process with a relatively simplified example which involves only 4 countries. Suppose the locations of the 4 countries are graphically represented as follows:

	Country1	
Country3	Country2	
		Country4

Based on the first-order contiguity relationship, the row-standardized weight matrix $\mathbf{W}$ for this simple example becomes:

$$\begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

41

and accordingly the origin-centric weight matrix $\mathbf{W}_o = \mathbf{W} \otimes \mathbf{I}_4$ is identified as[19]

$$\begin{pmatrix}
0 & & & & 1/2 & & & & 1/2 & & & & 0 & & & \\
& 0 & & & & 1/2 & & & & 1/2 & & & & 0 & & \\
& & 0 & & & & 1/2 & & & & 1/2 & & & & 0 & \\
& & & 0 & & & & 1/2 & & & & 1/2 & & & & 0 \\
1/3 & & & & 0 & & & & 1/3 & & & & 1/3 & & & \\
& 1/3 & & & & 0 & & & & 1/3 & & & & 1/3 & & \\
& & 1/3 & & & & 0 & & & & 1/3 & & & & 1/3 & \\
& & & 1/3 & & & & 0 & & & & 1/3 & & & & 1/3 \\
1/2 & & & & 1/2 & & & & 0 & & & & 0 & & & \\
& 1/2 & & & & 1/2 & & & & 0 & & & & 0 & & \\
& & 1/2 & & & & 1/2 & & & & 0 & & & & 0 & \\
& & & 1/2 & & & & 1/2 & & & & 0 & & & & 0 \\
0 & & & & 1 & & & & 0 & & & & 0 & & & \\
& 0 & & & & 1 & & & & 0 & & & & 0 & & \\
& & 0 & & & & 1 & & & & 0 & & & & 0 & \\
& & & 0 & & & & 1 & & & & 0 & & & & 0
\end{pmatrix}$$

where the blank cells are all zeros.

To eliminate the role of the self-directed dyads from the estimation, the 1st, 6th, 11th and 16th rows and columns of $\mathbf{W}_o$ need to be removed. This can be accomplished by pre-multiplying $\mathbf{W}_o$ by a "selection" matrix $\boldsymbol{B}$ and post-multiplying $\mathbf{W}_o$ by the transpose of $\boldsymbol{B}$, with $\boldsymbol{B}$ denoting a 12 by 16 sparse matrix as shown below:

---

[19] To save space, this study demonstrates only the operations on the $\mathbf{W}_o$ matrix. But the same operations can be applied to the $\mathbf{W}_d$ and $\mathbf{W}_w$ matrices as well.

42

$$B = \begin{pmatrix} 0 & 1 & & & & & & & & & & \\ & & 1 & & & & & & & & & \\ & & & 1 & & & & & & & & \\ & & & & 1 & & & & & & & \\ & & & & 0 & 1 & & & & & & \\ & & & & & & 1 & & & & & \\ & & & & & & & 1 & & & & \\ & & & & & & & & 1 & & & \\ & & & & & & & & 0 & 1 & & \\ & & & & & & & & & & 1 & \\ & & & & & & & & & & & 1 \\ & & & & & & & & & & 1 & 0 \end{pmatrix}$$

This selection operation is performed on $\mathbf{W}_d$ and $\mathbf{W}_w$ as well. Naturally, the altered weight matrices need to be row standardized once again to reflect the exclusion of neighboring relationships with self-directed pairs.

This alteration of the weight matrices sustains the rationale for origin-centric, destination-centric, and origin-to-destination dependence structures; however, it relieves the clumsiness of the original weight matrices, $\mathbf{W}_o$, $\mathbf{W}_d$ and $\mathbf{W}_w$ when self-directed dyads are inadvertently included due to construction rather than because they are of interest to the researcher. It should be noted that the elimination approach is preferred to the dummy variable approach, since the latter often leads to underestimating standard errors. Therefore, this study applies the elimination approach to the self-directed dyads in the illustrative example of international conflict and makes Bayesian inferences accordingly. The new model is given by

$$\boldsymbol{B}y^* = \rho_d\boldsymbol{B}\mathbf{W}_d^R\boldsymbol{B}'\boldsymbol{B}y^* + \rho_o\boldsymbol{B}\mathbf{W}_o^R\boldsymbol{B}'\boldsymbol{B}y^* + \rho_w\boldsymbol{B}\mathbf{W}_w^R\boldsymbol{B}'\boldsymbol{B}y^* + \boldsymbol{B}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{B}\varepsilon \qquad (2.23)$$

with renormalized matrices $\mathbf{W}_d^R$, $\mathbf{W}_o^R$, and $\mathbf{W}_w^R$.

2.6.3 Modified Moran *I* test and Spatial Dependence in the Limited Variable Case

In spatial studies, it is common practice to perform a spatial dependence test before implementing a spatial regression analysis. One widely used test is the Moran *I* test. To determine if a variable is distributed in a nonrandom spatial pattern, the Moran *I* statistic examines the spatial correlation of residuals. Nonetheless, a caveat of this statistic is the assumption of a continuous dependent variable, which is not satisfied in the case considered in this study.

Lin and Zhang (2007) propose a modified Moran *I* test for non-continuous cases wherein deviance residuals, rather than regression residuals, are examined for checking spatial autocorrelation among observations. Based on established statistical literature, which demonstrates that deviance residuals of loglinear models are asymptotically normal (see Agresti, 1990), Lin and Zhang show that the Moran *I* based on the log-likelihood (deviance) residuals of generalized linear models is analogous to the Moran *I* based on linear regression residuals. Their deviance residual Moran *I* test is implemented in four steps as follows: 1) in the model-fitting process, estimate the parameters of explanatory variables, 2) compute the deviance residual Moran *I* statistic,

$$I_{DR} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij}(r_{i,d} - \bar{r}_d)(r_{j,d} - \bar{r}_d)}{\left[\sum_{i=1}^{N}(r_{i,d} - \bar{r}_d)^2 / N\right]\left[\sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij}\right]}$$

where the deviance residual $r_{i,d}$ takes the place of x_i in the conventional Moran *I*; 3) compute the mean and variance of I_{DR}; and 4) calculate the *p* value of I_{DR} relying on the asymptotic normality assumption (p. 297). This study employs this variation of the Moran *I* test only as an exploratory

44

tool. Significant test statistics would further motivate the necessity of using a spatial OD probit model.

The modified Moran I test is applied to the deviance residuals from a standard probit with the weight matrix considered, in turn, being $\mathbf{W}_d$, $\mathbf{W}_o$ and $\mathbf{W}_w$ so as to flag any of the three possible forms of spatial dependence. Specifically, when the weight matrix is specified as $\mathbf{W}_d$, the test statistic has a p-value of 5.128e-16, which signals target (destination)-based dependence. While the weight matrix takes the format of $\mathbf{W}_o$, the test statistic reports a p-value of 0.2275. And the Moran I statistic holds a p-value of 0.5998 when $\mathbf{W}_w$ is alternatively considered to be the weight matrix in the test. Although the test results only suggest the existence of target-centric dependency relations, it should be kept in mind that Moran I test normally checks for spatial correlation among singletons and it tends to over count the average linkages represented by the $\mathbf{W}_w$ matrix. Hence, the test may not easily pick up the spatial connectivity among dyadic observations arising from this type of dual neighboring relationship.

2.6.4 Further Exploratory Data Analysis

As an additional exploratory step, this study adds all the three spatial lag terms, $\mathbf{W}_d y$, $\mathbf{W}_o y$ and $\mathbf{W}_w y$, simultaneously to the standard probit, but still relies on the standard ML estimator for estimation. While this estimator suffers from an inconsistency problem as discussed earlier, preliminary estimation results from this regression may be somewhat informative. But it should be noted that the model estimated this way is no longer the same as (2.6) in that spatial lags in y rather than y^* are used. The coefficient on $\mathbf{W}_d y$ is, as hinted at by the Moran I tests above, significant at the 0.001 level; however the coefficient of $\mathbf{W}_w y$ turns out to be significant

45

at the 0.01 level as well. This result should not be too surprising in light of Moran I's potential limitations in handling the unique structure of $\mathbf{W}_w$.

Columns 2 and 3 in TABLE 2.1 present estimated results using standard probit estimators. While Column 2 does not include spatial lags, Column 3 does. The spatial coefficient estimates displayed in Column 3 are indicative of the presence of spatial dependence in the conflict data examined in this study.[20] The positive coefficient on $\mathbf{W}_d y$ seems to suggest that a state's decision to attack a target is positively related to its conflict initiation behavior relative to the neighboring countries of the intended target. This relationship in conflict behavior may arise out of logistic needs, strategic considerations, or the mere intention of reinforcing its position in the region by bringing more territory under its control.

Insignificant as it is, the negative sign of the coefficient for $\mathbf{W}_o y$ hints that states are less likely to assault their potential target when their neighboring state is engaged in a conflict with the same target. A disincentive to getting into other states' spheres of influence appears to offer a plausible explanation for this negative association among conflict initiations. The negative coefficient associated with $\mathbf{W}_w y$ alludes to a reverse correlation between a state's conflict initiation and its neighbors' assaults on the neighbors of its latent target. This dampening effect on the outbreak of conflict may be accounted for by the potential originator's concern with its own domestic stability and a reluctance to get involved in an already unstable situation. From the perspective of a potential initiator, waging a war against a target which shares a border with the state(s) assaulted by the neighbor(s) of the initiator may mean further complicating the existing chaos. For instance, Gleditsch (2007) finds that conflicts, civil or interstate, in neighboring

---

[20] It should be noted that we should read these estimation results with caution because they are still biased. That is, the estimator is a standard probit model instead of a spatial OD probit model.

countries tend to increase a country's proneness to civil war. In line with this reasoning, concerns about maintaining political stability within the home territory should lower the incentive for the potential originator to set off the conflict. For one thing, handling refugee flows created by an ongoing conflict in a neighboring country may deplete a potential initiator's available resources. More importantly, it is unwise for a state to drag itself into a new war if that state is not ready to pool all its resources for the war effort. In a similar vein, Levy (1982) contends that war-weariness may inhibit war initiation behavior by other countries. This effect may be more salient for states within the vicinity of either an ongoing war or one concluded not long ago. Geographical proximity allows neighboring states to be aware of the devastating forces of war. Overall, the preliminary tests discussed so far seem to lend some support for the possibility of spatial dependence in conflict initiation.

2.7 Spatial OD Probit Estimated with a MCMC Sampler

As discussed in Section 2.3, in order to obtain consistent estimates for the proposed spatial OD probit model, a Bayesian approach is more appropriate. Since the rationale for the Bayesian method and the sampling procedure have already been explained in earlier sections, this study now focuses on the specifics pertaining to sampling the spatial lag parameters ρ_i ($i = d, o, w$), and then presents the empirical findings.

For MCMC simulation, initial values need to be assigned to parameters in order to start the simulation process. This study uses zeros as the starting values for parameters of the explanatory variables and the spatial lagged terms. As ρ_i is supposed to be within the interval of $(-1,1)$, this study follows LeSage and Pace's (2009) approach by employing a normal distribution for the proposal density along with rejection sampling to confine the simulated

47

values of ρ_i to the desired range. The proposed value of ρ_i^p is deemed as a random deviate from the current value of ρ_i^c with the random deviation drawn from a standard normal distribution and adjusted by a tuning parameter m. Precisely, this procedure can be expressed as follows:

$$\rho_i^p = \rho_i^c + m \cdot N(0, 1). \tag{2.24}$$

The reason for tuning the proposed values is to allow the Metropolis-Hastings sampling procedure to cover the entire conditional distribution rather than to get stuck in a low probability area. By monitoring the acceptance rates of the Metropolis-Hastings procedure, this study sets m at 1/20 and the resultant acceptance rates for all ρ_i's are around 40%.[21]

2.7.1 Estimation Results from the MCMC Sampler

(TABLE 2.1 about here)

60,000 iterations were run and the first 10,000 were discarded as burn-in. Following the conventional practice in MCMC simulations, a 95% credibility interval together with the posterior mean and standard deviation are reported in Appendix IV. Bayesian estimates including a p-value (which evaluates whether a coefficient is significantly different from zero (LeSage and Pace, 2009, p. 292)) are presented in Column 4 of TABLE 2.1 alongside with the ML estimates in Columns 2 and 3. It should be noted that the Bayesian p-value is held as being comparable to the conventional p-value associated with the asymptotic t-statistic (LeSage and Pace, 2009; Gelman et al., 1995). Column 4 displays estimated results with self-directed dyads removed (i.e., the elimination approach is applied). As noted earlier, the elimination approach

---

[21] LeSage and Pace (2009, p. 137) suggest the use of a tuning parameter that results in an acceptance rate between 40 and 60 percent.

should be more preferable in handling unnecessary self-directed pairs. The MCMC simulation results for the spatial OD probit show that in the context of conflict initiation, the propensity for starting a dispute in one dyad is influenced by the dispute onsets in other related dyads, even when the conventional political factors of democracy, national capabilities, geographic distance and alliance are controlled in a model of directed dyadic conflict.

To be specific, the spatial coefficients ρ_d and ρ_w are both statistically significant, with the former taking a positive sign and the latter being negative; while ρ_o turns out to be insignificant though negative (see Column 4). The positive sign of ρ_d suggests that states are more likely to attack the neighbors of their intended target as well. This positive correlation may be attributed to logistic and strategic concerns. The famous Maginot Line provides a good case in point. Due to the construction of this defense line, Germany was dissuaded from a direct attack on France, but invaded Belgium first in order to conquer France. As explained earlier, a negative ρ_w means that instigation of a conflict by the neighbor of a potential aggressor, with the neighbor of the intended target, tends to discourage the conflict behavior of the latent initiator.

Inclusion of the three terms for dyadic dependence notably affects the estimated coefficients of political variables. Although neither of the two democracy variables shows statistical significance, it is interesting to note the change in the magnitude of estimated coefficient for *polity1*. Under a standard probit (i.e., Column 2), the coefficient that was estimated with the unrealistic assumption of independence between observations is -0.0045 for the initiator, *polity1*. However, once the spatial dependence is properly taken into account, *polity1* shows no substantive effect as well and virtually assumes a coefficient of zero. In the context of the time span covered in this study, it is plausible that the regime type of a potential

49

initiator should not be relevant to its conflict initiation behavior, as for many countries security concerns unquestionably overwhelmed political concerns during that historical period.

Both the initiator and target's national capabilities are found to have a significant, positive influence on the likelihood of dispute initiation, although the power status of the initiator appears to have a relatively larger bearing than that of the target. In contrast to the standard probit case where independence is assumed among observations (i.e., Column 2), the discrepancy in the coefficient sizes of national capabilities is narrowed considerably in the spatial OD probit (see Column 4), with the estimate for *cap1* being reduced by 35% and that of *cap2* increased by 13%. These results imply that no matter whether it is from the potential aggressor or the target, a change of the power status quo tends to encourage an interstate dispute. It seems that conflict initiation behavior is either a result of strengthened power for the latent initiator, who is now better prepared to overwhelm or challenge the intended target, or due to an urge to contain the power growth of the target.

With the addition of the spatial lag terms, *distance* still appears to be negatively correlated to dispute proneness. This is consistent with the existing literature that when countries are far apart, they are less likely to fight each other (Boulding,1963; Gleditsch, 2002). It is plausible that due to geopolitical or logistic reasons, two countries separated by a large distance are less subject to conflicts of interest. In terms of magnitude, the estimated coefficient for *distance* is considerably larger in the spatial OD probit (-0.0011) than in a non-spatial probit (-0.0009). And *alliance* exhibits no statistical significance though it takes on a positive sign, indicating that alliance ties do not have an appreciable impact on the propensity for conflict initiation within dyads. This finding is consistent with previous dyadic analyses of the

50

relationship between alliance dyads and war dyads (Ostrom and Hoole, 1978; Bremer, 1992). If as the more traditional view sees it, alliances normally grow "out of expediency and reflect nothing deeper than a temporary need of two or more states to coordinate their actions against one or more other states" , then war between allies may not be expected to be either more or less frequent than between non-allied states (Bremer, 1992, p. 315).

Overall, the multifaceted spatial dependence captured by the spatial OD probit model implies that the effects of explanatory variables also work through different channels.

2.7.2 Marginal Effects in the Spatial OD Probit Model

While attaining statistical significance is an essential factor in determining the relevance of independent variables, passing this milestone does not ensure that these variables have a meaningful influence over the dependent variable in a substantive sense. In a probit model, to specifically determine the extent to which such explanatory variables influence the dependent variable, researchers should report the marginal effect that each independent variable exerts on the outcome variable in a percentage change term. Due to the non-linearity of the normal probability density in probit models, the marginal effect of an explanatory variable varies with the level of the variable itself. Thus, the mean value, or a particular observation of each regressor is often used when interpreting model estimates. Accordingly, the marginal effect is inferred as being the change in the probability of a given event occurring, which is associated with a change in the average or a particular level of an explanatory variable.

In the spatial OD probit model, the marginal effects are more complicated, because a change in the value of an explanatory variable for the ith observation will not only affect the

current dyad, but its effect may also diffuse to other dyads. Part of the emitted effect will ultimately be routed back to the starting dyad, depending on the types of the dyadic dependency relations as well as on the strength of such relations. This implies that a full treatment of marginal effects should consider the direct impact on the current dyad in question and the indirect or spatial spillover impact on neighboring dyads via the three forms of dependence specified in the model.

Moreover, the country-specific regressors have an origin-destination structure, defined above by Kronecker products with the common n by k matrix, X. Thus, a change in one country's regressor immediately affects all dyads in which that country is either an origin or a destination, and then the effects are propagated through the spatial diffusion mechanism to other dyads. The only columns of X specific to each dyad rather than each country are the constant term, distance between countries in the dyad, and alliance.[22]

Consider the linear model in (2.23). We redefine $\boldsymbol{A} = (\mathbf{I}_{N-n} - \rho_d \boldsymbol{B}\mathbf{W}_d^R\boldsymbol{B}' - \rho_o\boldsymbol{B}\mathbf{W}_o^R\boldsymbol{B}' - \rho_w\boldsymbol{B}\mathbf{W}_w^R\boldsymbol{B}')$, in order to rewrite the model as $\boldsymbol{B}y^* = \boldsymbol{A}^{-1}\boldsymbol{B}\mathbf{X}\boldsymbol{\beta} + \boldsymbol{A}^{-1}\boldsymbol{B}\varepsilon$, similarly to (2.7). Let e_i' denote the binary row vector that selects the ith row of a subsequent conformable matrix or vector – i.e., e_i is a vector with a unit in the ith element and zeros elsewhere. The ith observation of $\boldsymbol{B}y^*$ is thus $e_i'\boldsymbol{B}y^*$, which we simplify hereafter as y_i^* with the convention that self-directed dyads are omitted. Recalling that $\mathbf{X_d} = \iota_n \otimes \mathrm{X}$ and $\mathbf{X_o} = \mathrm{X} \otimes \iota_n$, we wish to examine the derivative

---

[22] Marginal effects of a change in distance are complicated by the fact that we adjust for contiguity. Such an adjustment means that distance is not a differentiable function of latitude and longitude. We do not consider these marginal effects.

$$\frac{\partial E[y_i^* | e_i' \boldsymbol{A}^{-1} \boldsymbol{B} \mathrm{X}]}{\partial \mathrm{X}} = \frac{\partial e_i' \boldsymbol{A}^{-1} \boldsymbol{B}(\iota_n \otimes \mathrm{X}) \beta_d}{\partial \mathrm{X}} + \frac{\partial e_i' \boldsymbol{A}^{-1} \boldsymbol{B}(\mathrm{X} \otimes \iota_n) \beta_o}{\partial \mathrm{X}} \qquad (2.25)$$

To simplify notation, let $e_i^* = (\boldsymbol{A}^{-1}\boldsymbol{B})' e_i$, which is the ith column of the n^2 by $n(n-1)$ matrix $(\boldsymbol{A}^{-1}\boldsymbol{B})'$. Let $\mathbf{F}_i^* = \left(e_{i,1}^*, \dots, e_{i,n}^*\right)$, where $e_{i,1}^*$ denotes the first n by 1 subvector of e_i^*, etc. Thus, $e_i^* = vec(\mathbf{F}_i^*)$ and we may write

$$\frac{\partial E[y_i^* | e_i' \boldsymbol{A}^{-1} \boldsymbol{B} \mathrm{X}]}{\partial \mathrm{X}} = \frac{\partial vec(\mathbf{F}_i^*)'(\iota_n \otimes \mathrm{X}) \beta_d}{\partial \mathrm{X}} + \frac{\partial vec(\mathbf{F}_i^*)'(\mathrm{X} \otimes \iota_n) \beta_o}{\partial \mathrm{X}}$$

$$= \frac{\partial tr(\beta_d \iota_n' \mathbf{F}_i^{*\prime} \mathrm{X})}{\partial \mathrm{X}} + \frac{\partial tr(\mathbf{F}_i^{*\prime} \iota_n \beta_o' \mathrm{X}')}{\partial \mathrm{X}}$$

$$= \mathbf{F}_i^* \iota_n \beta_d' + \mathbf{F}_i^{*\prime} \iota_n \beta_o' \qquad (2.26)$$

The last line follows from the derivative of a trace (see, e.g., Lütkepohl, 1996). This expression is an n by K matrix of partial derivatives, and the effect on y_i^* of changing a certain country-specific characteristic for a certain country is given by the element in the corresponding column and row, respectively. Such an effect could be written as

$$\frac{\partial E[y_i^* | e_i' \boldsymbol{A}^{-1} \boldsymbol{B} \mathrm{X}]}{\partial x_{jk}} = [\mathbf{F}_i^* \iota_n \beta_d' + \mathbf{F}_i^{*\prime} \iota_n \beta_o']_{jk} \qquad (2.27)$$

for country $j = 1, \dots, n$ and characteristic $k = 1, \dots, K$.

For the model given jointly by (2.8) and (2.23) with a probit link function, the marginal effect on the probability of conflict initiation in dyad i given a change in characteristic k of country j is given by

53

$$\frac{\partial E[y_i | e_i' \boldsymbol{A}^{-1} \boldsymbol{B} \mathrm{X}]}{\partial x_{jk}} = \phi \left(\frac{e_i' \boldsymbol{A}^{-1} \boldsymbol{B} \mathrm{X} \beta}{\sqrt{e_i' (\boldsymbol{A}'\boldsymbol{A})^{-1} e_i}} \right) \frac{[\mathbf{F}_i^* \iota_n \beta_d' + \mathbf{F}_i^{*\prime} \iota_n \beta_o']_{jk}}{\sqrt{e_i' (\boldsymbol{A}'\boldsymbol{A})^{-1} e_i}} \tag{2.28}$$

using the usual chain rule. These marginal effects are of course estimated by inserting estimates of $\boldsymbol{A}$ and β, of which β_d and β_o are subvectors.

From these derivative matrices, scalar summaries for the whole system, along the lines of those proposed recently by LeSage and Thomas-Agnan (2012), may be created by averaging across all i and j. Such scalar summaries allow the calculation of both a direct effect (i.e., the effect of changing any country's regressor on dyads involving that country) and a network or indirect effect (i.e., the effect on pairs not involving that country). Total effect is the sum of these two effects. For the national capabilities variable, which records a country's proportion of the total system capabilities, this study reports the average marginal effects associated with a 1% increase of this regressor. The marginal impacts of *polity* are computed when the variable increases by one unit. TABLE 2.2 presents the marginal effect estimates for the country-specific regressors of the spatial OD probit model.

(TABLE 2.2 about here)

As shown in TABLE 2.2, enhanced power status will have a positive direct impact. Specifically, a 1% increase in the capability of a country (either potential initiator or target) increases the probability of conflict between that country and any other country by an average of 2.148%. However, the network effect is estimated to be negative though smaller, providing evidence for the notion that two countries may become less prone to mutual conflict when facing a common threat. This subtlety cannot be picked up by a model with no spatial correlation. Overall, a 1% upsurge in capabilities will enhance the propensity for dispute initiation by

1.599%. This positive total effect reaffirms the realist perspective on balance of power that increased power status is likely to fuel interstate conflict, either by allowing a state to become more assertive or prompting neighboring countries to perceive it as an imminent danger. Consistent with the coefficient estimates, *polity* has an insignificant marginal impact on conflict initiation. A unit increase in the *polity* measure of a country reduces the probability of conflict for a dyad involving that country by 0.108% while this dampening effect is only about 0.105% across the whole system.

2.7.3 Individual Marginal Effects in the Spatial OD Probit Model

An interesting question may arise, "What would have been the implications if a certain condition had been changed for a particular country?" For example, suppose that Germany had not taken the road of militarization during the WWII period, but instead had maintained military capabilities similar to that of Sweden (as the two countries have about the same size of territory and Sweden's *cap* level is around the median level of the sample). Then what would have been the impacts on all the sampled dyads and in particular, on those dyads that involve Germany as the potential initiator? This question can be answered by inspecting scalar summaries for individual origins which average the derivative matrices specified in (2.28) across i with j and k fixed.

TABLE 2.3 shows the baseline probabilities estimated by the spatial OD probit model of a dispute initiated by Germany against each of the other sample countries. For illustration, we report the marginal impact of a small (1%) reduction in Germany's capability using the formula derived in (2.28). We also report the forecast change in initiation propensity from lowering

55

Germany's capability from the recorded value of 0.0700 at the time to the sample median level of 0.0061.

The baseline probabilities range from a 5.4% probability that Germany would target Portugal to a 50.8% probability of targeting France. The marginal effects of a 1% decrease in Germany's capability range from a moderate decrease of 1.3% (vs. Portugal) to a more substantial decrease of 4.9% (vs. the Soviet Union) in the initiation probability. On average, the decrease in the likelihood of dispute initiation is 3.7% for all pairs with conflict potentially originating with Germany, 2.23% for all pairs involving Germany, and 0.13% for the whole system. Under the counterfactual of lowering Germany's capability to the level of sample median, the likelihood of dispute initiation by Germany would have fallen by 4.51% to 28.33%. To be more specific, a less power-thirsty Germany during WWII would have reduced its propensity for attacking France by 27.91% and for targeting Russia by 28.33%. On average, the decrease in initiation probability is 17.93% when Germany is the potential aggressor, 11.24% when Germany is involved at all, and 0.60% for the whole system. This reinforces the previous conclusion pertaining to the war-inducing effects of military power.

(TABLE 2.3 about here)

2.8 Conclusion

Interstate conflict is a widely-studied subject among political scientists and economists. However, most existing quantitative research was conducted under the unrealistic assumption that each observation is spatially-independent. Fortunately, recent innovations in spatial econometrics can help researchers to investigate the possibility of spatial dependence in conflict

data. In fact, some recent studies employ spatial econometrics modeling, such as the SAR model,[23] when looking into the causes of international conflict. Nevertheless, due to the constraints of model structures, existing research that explores the spatial dependence in conflict data is limited to certain cases where the unit of analysis is monadic, consisting of individual states. This state of affairs is less than satisfying because an interstate conflict is a phenomenon in which at least two parties are, by definition, involved. When leaving out one conflict participant or ignoring interaction between participants, empirical analysis is likely to produce biased estimates, and consequently to distort the understanding of the nature of interstate conflict. For these reasons, conflict scholars prefer the dyad as the unit of analysis to the monad. There are two kinds of dyadic analysis: non-directed dyads and directed dyads. While the former does not distinguish between aggressor and victim, the latter does. However, this level of analysis demands a more complex structure of connectivity to reflect the spatial dependence; this goes beyond what the traditional SAR model can handle. Due to this methodological difficulty, no previous study has attempted to directly model the spatial interdependence in dyadic conflict data.

This study has extended LeSage and Pace's (2008) spatial origin-destination modeling to cases with a binary dependent variable, and then applied the spatial OD probit model to dispute initiations among 26 European countries during the WWII period. As noted, by using a combination of three spatial connectivity matrices for origin-, destination-, and origin-to-destination-based dependence, LeSage and Pace's spatial OD model extended the traditional spatial model and addressed the spatial dependence in interregional flows. Their spatial OD model relied on MLE for estimation because the model's dependent variable was a continuous

---

[23] The SAR model is preferred to the SER model because it directly models the dependence-generating process instead of treating it as a nuisance term.

measure. This study proposes a probit version of LeSage and Pace's spatial OD model where the dependent variable is dichotomous. Because a traditional ML estimator for the spatial OD probit model would suffer from inconsistency due to the inclusion of lagged terms of the dependent variable on the right-hand side of the equation, this study adopts a Bayesian simulation procedure for model estimation. When applied to conflict initiation data, the constructs of the three spatial connectivity matrices in the spatial OD probit model are intended to capture the dependence among conflict initiations arising from the initiator side, the target side and the initiator-to-target link.

The empirical results from the spatial OD probit model of interstate dispute initiation indicate that spatial dependence exists between conflict initiations, and that this spatial relationship is more complex than the one specified in the existing literature. The positive, significant coefficient associated with target-centric dependence signifies that aggressors tend to attack the neighbors of their intended victim as well. This inviting effect of spatial dependence on dispute initiations may be accounted for by strategic and logistic needs. On the other hand, the initiator-to-target-based dependence exerts a negative impact on the propensity for a dispute. A potential originator tends to be discouraged from taking actions against its intended target if there is a conflict between the neighbor of the originator and the neighbor of the target. This negative association between conflict initiation behaviors may be explained by concerns about maintaining domestic stability or by a war-weariness effect. The coefficient estimate for the initiator-centric dependence assumes a negative sign but is statistically insignificant. The negative sign seems to suggest that countries restrain themselves from starting a dispute against the same target that their neighboring state preys on. This could be due to a tacit understanding of the existence of spheres of influence, which may not have been uncommon in WWII Europe.

However, the data used in this study offer insufficient empirical evidence to draw such a conclusion.

In addition, the introduction of spatial correlation enables researchers to examine effects from spillovers of conflict throughout the system. For instance, the spillovers of national capabilities indicate that two countries in a dyad may be less bound to engaging in a conflict when facing a threat from a third country. Overall, the effect estimates of the spatial OD probit model suggest that increased power, whether it is related to an initiator or a target, is likely to heighten the chances of conflict in a dyad. However, geographical distance appears to suppress the outbreak of conflict, and alliances have no bearing on conflict initiation within dyads even after control for spatial dependence.

In sum, the proposed spatial OD probit model takes into account the interdependence among directed dyads, and therefore is instrumental in producing more reliable estimates of conflict-inducing factors as well as a better understanding of the dynamics of interstate conflict behavior.

References

Agresti, A. 1990. *Categorical Data Analysis*. New York: Wiley.

Albert, J.H. and S. Chib. 1993. "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88, 669-679.

Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Boston: Kluwer Academic.

Anselin, L. 2003. "Spatial Econometrics," in B. Baltagi (ed.) *A Companion to Theoretical Econometrics*, Oxford: Blackwell Publishing Ltd, 310-330.

Beck, N.L., K.S. Gleditsch and K.C. Beardsley. 2006. "Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy," *International Studies Quarterly*, 50(1), 27-44.

Bennett, S.D. and A.C. Stam. 2000. "Research Design and Estimator Choices in the Analysis of Interstate Dyads: When Decisions Matter," *Journal of Conflict Resolution*, 44(5), 653-685.

Besley, T.J., and T. Persson. 2008. "Wars and State Capacity." *Journal of the European Economic Association*, 6(2–3), 522-530.

Bivand, R.S., E.J. Pebesma, and V. Gómez-Rubio. 2008. *Applied Spatial Data Analysis with R*. New York: Springer.

Blomberg, B.S., and G.D. Hess. 2002. "The Temporal Links between Conflict and Economic Activity," *Journal of Conflict Resolution*, 46(1), 74-90.

Börsch-Supan, Axel and Vassilis A. Hajivassiliou. 1993. "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models," *Journal of Econometrics*, 58, 347-368.

Bremer, S.A. 1992. "Dangerous Dyads: Conditions Affecting the Likelihood of Interstate War, 1816-1965," *Journal of Conflict Resolution*, 36(2), 309-341.

Bueno de Mesquita, B. 1981. *The War Trap*. New Haven: Yale University Press.

60

Bueno de Mesquita, B. 1988. "The Contribution of Expected Utility Theory to the Study of International Conflict," *The Journal of Interdisciplinary History*, 18(4), 629-652.

Choi, S.-W. and P. James. 2005. *Civil-Military Dynamics, Democracy, and International Conflict: A New Quest for International Peace*. New York: Palgrave.

Cramer, C. 2002. "Homo Economicus Goes to War: Methodological Individualism, Rational Choice and the Political Economy of War," *World Development*, 30(11), 1845-1864.

Davutyan, N. and W.R. Parke. 1995. "The Operations of the Bank of England, 1890-1908: A Dynamic Probit Approach," *Journal of Money, Credit and Banking*, 27(4), 1099-1112.

Dueker, M. 1999. "Measuring Monetary Policy Inertia in Target Fed Funds Rate Changes," *Federal Reserve Bank of St Louis Review*, September/October, 3-9.

Dulles, E. 1942. "War and Investment Opportunities: A Historical Analysis," *American Economic Review*, 32(1), 112-122.

Fischer, M.M., T. Scherngell, and E. Jansenberger. 2006. "The Geography of Knowledge Spillovers between High-Technology Firms in Europe Evidence from a Spatial Interaction Modelling Perspective," *Geographical Analysis*, 38(3), 288-309.

Garfinkel, M.R. 1994. "Domestic Politics and International Conflict," *American Economic Review*, 84(5), 1294-1309.

Garfinkel, M.R. 1990. "Arming as a Strategic Investment in a Cooperative Equilibrium," *American Economic Review*, 80(1), 50-68.

Gartzke, E.A. and K.S. Gleditsch. 2008. Online manuscript. "The Ties that Bias: Specifying and Operationalizing Components of Dyadic Dependence in International Conflict".

Gelfand, A.E. and A.F.M. Smith. 1990. "Sampling-Based Approach to Calculating Marginal Densities," *Journal of American Statistical Association*, 85, 398-409.

Geweke, J. 1991. "Efficient Simulation from the Multivariate Normal and Student-t Distributions Subject to Linear Constraints and the Evaluation of Constraint Probabilities," in *Proceedings of the 23rd Symposium on the Interface Between Computing Science and Statistics*, E. Kermanidas, (ed.), 571-578.

Gleditsch, K.S. 2007. "Transnational Dimensions of Civil War," *Journal of Peace Research*, 44(3), 293-309.

Grossman, H. and M. Kim. 1995. "Swords or Plowshares? A Theory of the Security of Claims to Property," *Journal of Political Economy*, 103(6), 1275-1288.

Hajivassiliou, V. 1990. "Smooth Simulation Estimation of Panel Data LDV Models," Discussion paper, Yale University.

Hegre, H., J.R. Oneal and B. Russett. 2009. Online Manuscript. "Trade Does Promote Peace: New Simultaneous Estimates of the Reciprocal Effects of Trade and Conflict".

Hess, G.D. and A. Orphanides. 1995. "War Economics: An Economic Rational Voter Framework," *American Economic Review*, 85(4), 828-846.

Hess, G.D. and A. Orphanides. 2001a. "Economic Conditions, Elections, and the Magnitude of Foreign Conflicts," *Journal of Public Economics*, 80, 121-140.

Hess, G.D. and A. Orphanides. 2001b. "War and Democracy," *Journal of Political Economy*, 109(4), 776-810.

Isard, W. 1994. "Peace Economics: A Topical Perspective," *Peace Economics, Peace Science and Public Policy*, 1(2), 11-13.

Keane, M. 1994. "A Computationally Practical Simulation Estimator for Panel Data," *Econometrica*, 62(1), 95-116.

King, G. 2001. "Proper Nouns and Methodological Propriety: Pooling Dyads in International Relations Data," *International Organization*, 55, 497-507.

Koubi, V. 2005. "War and Economic Performance," *Journal of Peace Research*, 42(1), 67-82.

Lee, L.F. 2002. "Consistency and Efficiency of Least Squares Estimation for Mixed Regressive, Spatial Autoregressive Models," *Econometric Theory*, 18, 252-277.

LeSage, J.P. 2000. "Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models," *Geographical Analysis*, 32(1), 19-35.

LeSage, J.P. and R.K. Pace. 2004. "Introduction," *Advances in Econometrics: Volume 18: Spatial and Spatiotemporal Econometrics*. Oxford: Elsevier Ltd, 1-32.

LeSage, J.P. and R.K. Pace. 2008. "Spatial Econometric Modeling of Origin-Destination Flows," *Journal of Regional Science*, 48(5), 941-967.

LeSage, J.P. and R.K. Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton: Chapman & Hall/CRC.

LeSage, James P. and Christine Thomas-Agnan. 2012. "Interpreting Spatial Econometric Origin-Destination Flow Models," Department of Finance and Economics, Texas State University – San Marcos, mimeo.

Lin, G. and T.L. Zhang. 2007. "Loglinear Residual Tests of Moran's *I* Autocorrelation and Their Applications to Kentucky Breast Cancer Data," *Geographical Analysis*, 39, 293-310.

Lütkepohl, Helmut. 1996. *Handbook of Matrices*. New York: John Wiley & Sons.

Lynch, S.M. 2007. *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New York: Springer.

McMillen, D.P. 1992. "Probit with Spatial Autocorrelation," *Journal of Regional Science*, 32(3), 335-348.

Neumayer, E. and T. Plümper. 2010. "Spatial effects in dyadic data," *International Organization*, 64(1), 145-166.

Ostrom, C. W. and F. Hoole. 1978. "Alliances and war revisited," *International Studies Quarterly*, 22, 215-236.

Pigou, A.C. 1921. *The Political Economy of War*. London: Macmillan.

Polachek, S.W. 1994. 'Peace Economics: A Trade Theory Perspective," *Peace Economics, Peace Science and Public Policy*, 1(2), 12-15.

Polachek, S.W. 1997. "Why Democracies Cooperate More and Fight Less: The Relationship between International Trade and Cooperation," *Review of International Economics*, 5(3), 295-309.

Ray, J.L. 2001. "Integrating Levels of Analysis in World Politics," *Journal of Theoretical Politics*, 13(4), 355-388.

Robbins, L. 1939. *The Economic Causes of War*. London: Jonathan Cape Ltd.

Russett, B. and J.R. Oneal. 2001. *Triangulating Peace: Democracy, Interdependence, and International Organizations*. New York and London: W.W. Norton & Company.

Singer, J. D., S. Bremer, and J. Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965," in Bruce Russett (ed.) *Peace, War, and Numbers*, Beverly Hills: Sage, 19-48.

Singer, J. D. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985, *International Interactions*, 14, 115-132.

Spiegelhalter, D.J., N.G. Best, B.P. Carlin and A. van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 583-639.

Ward, M. and K.S. Gleditsch. Book manuscript. *An Introduction to Spatial Regression Models in the Social Sciences*. 2008 preprint, Sage Publishers.

Appendix I

List of Sample Countries

U.K., Ireland, Netherlands, Belgium, Luxembourg, France, Switzerland, Spain, Portugal, Germany, Poland, Austria, Hungary, Czech, Italy, Albania, Yugoslavia, Greece, Bulgaria, Romania, Russia, Estonia, Latvia, Lithuania, Finland, Sweden.

Appendix II

List of Directed Dyads with More Than One Initiation Over 1933-1941

Initiator	Target	Year of Initiation
Switzerland	Germany	1939
Switzerland	Germany	1940
Germany	Belgium	1936
Germany	Belgium	1939
Germany	Portugal	1940
Germany	Portugal	1941
Germany	Czech	1938
Germany	Czech	1939
Germany	Bulgaria	1940
Germany	Bulgaria	1941
Germany	Russia	1939
Germany	Russia	1940
Germany	Sweden	1939
Germany	Sweden	1940

Germany	Sweden	1941
Hungary	Czech	1938
Hungary	Czech	1939
Italy	UK	1937
Italy	UK	1939
Italy	France	1937
Italy	France	1939
Italy	France	1940
Italy	Spain	1936
Italy	Spain	1940
Italy	Germany	1934
Italy	Germany	1939
Italy	Albania	1934
Italy	Albania	1939
Romania	Hungary	1939
Romania	Hungary	1941
Russia	Poland	1938
Russia	Poland	1939
Russia	Bulgaria	1940
Russia	Bulgaria	1941
Russia	Estonia	1939
Russia	Estonia	1940
Russia	Latvia	1939
Russia	Latvia	1940

Weight Matrix **W**

	GBR	IRL	NED	BEL	LUX	FRA	SUI	ESP	POR	GER	POL	AUT	HUN	CZE	ITA	ALB	YUG	GRE	BUL	ROU	RUS	EST	LAT	LTU	FIN	SWE
GBR	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IRL	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NED	0	0	0	0.5	0	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BEL	0	0	0.25	0	0.25	0.25	0	0	0	0.25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LUX	0	0	0	0.333	0	0.333	0	0	0	0.333	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FRA	0	0	0	0.167	0.167	0	0.167	0.167	0	0.167	0	0	0	0	0.167	0	0	0	0	0	0	0	0	0	0	0
SUI	0	0	0	0	0.25	0	0	0	0	0.25	0	0.25	0	0	0.25	0	0	0	0	0	0	0	0	0	0	0
ESP	0	0	0	0	0	0.5	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
POR	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GER	0	0	0.111	0.111	0.111	0.111	0.111	0	0	0	0.111	0.111	0	0.111	0	0	0	0	0	0	0	0	0	0.111	0	0
POL	0	0	0	0	0	0	0	0	0	0.167	0	0	0	0.167	0	0	0	0	0	0.167	0.167	0	0.167	0.167	0	0
AUT	0	0	0	0	0	0	0.167	0	0	0.167	0	0	0.167	0.167	0.167	0	0.167	0	0	0	0	0	0	0	0	0
HUN	0	0	0	0	0	0	0	0	0	0	0	0.25	0	0.25	0	0	0.25	0	0	0.25	0	0	0	0	0	0
CZE	0	0	0	0	0	0	0	0	0	0.2	0.2	0.2	0.2	0	0	0	0	0	0	0.2	0	0	0	0	0	0
ITA	0	0	0	0	0	0.25	0.25	0	0	0	0	0.25	0	0	0	0	0.25	0	0	0	0	0	0	0	0	0
ALB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.5	0	0	0	0	0	0	0	0
YUG	0	0	0	0	0	0	0	0	0	0	0	0.143	0.143	0	0.143	0.143	0	0.143	0.143	0.143	0	0	0	0	0	0
GRE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.333	0.333	0	0.333	0	0	0	0	0	0	0
BUL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.333	0.333	0	0.333	0	0	0	0	0	0
ROU	0	0	0	0	0	0	0	0	0	0	0.167	0	0.167	0.167	0	0	0.167	0	0.167	0	0.167	0	0	0	0	0
RUS	0	0	0	0	0	0	0	0	0	0	0.2	0	0	0	0	0	0	0	0	0.2	0	0.2	0.2	0	0.2	0
EST	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0.5	0	0	0
LAT	0	0	0	0	0	0	0	0	0	0	0.25	0	0	0	0	0	0	0	0	0	0.25	0.25	0	0.25	0	0
LTU	0	0	0	0	0	0	0	0	0	0.333	0.333	0	0	0	0	0	0	0	0	0	0	0	0.333	0	0	0
FIN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0.5
SWE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0

**Bayesian Estimates of the Spatial OD Modeling of MID Initiation
Europe, 1933-1941**

	Mean	**S.D.**	**2.5%**	**Median**	**97.5%**	**Sample**
p_d	0.2148	0.0765	0.0693	0.2140	0.3668	50000
p_o	-0.0915	0.0958	-0.2865	-0.0898	0.0923	50000
p_w	-0.2626	0.1202	-0.4875	-0.2685	-0.0091	50000
intcpt.	-1.3138	0.2128	-1.7389	-1.3118	-0.9044	50000
polity1	0.0001	0.0102	-0.0197	0.0000	0.0203	50000
polity2	-0.0096	0.0113	-0.0319	-0.0096	0.0125	50000
cap1	9.5507	1.9963	5.6703	9.5367	13.5138	50000
cap2	5.1845	2.2338	0.8498	5.1693	9.6239	50000
distance	-0.0011	0.0002	-0.0016	-0.0011	-0.0007	50000
alliance	0.0252	0.2192	-0.4095	0.0267	0.4515	50000

TABLE 2.1 Probit Estimates of MID Initiations, Europe, 1933-1941

Variable	Standard Probit Estimator		Spatial OD Probit MCMC Sampling
	w/o Spatial Lags	w/ Spatial Lags	w/o Self-directed Dyads
p_d^a		1.6271***	
		(0.3076)	
p_o^a		-0.0927	
		(0.5230)	
p_w^a		-3.3015***	
		(0.9170)	
p_d			0.2148***
			(0.0765)
p_o			-0.0915
			(0.0958)
p_w			-0.2626**
			(0.1202)
intcpt.	-1.2801***	-0.8539***	-1.3138***
	(0.1624)	(0.2464)	(0.2128)
polity1	-0.0045	-0.0025	0.0001
	(0.0104)	(0.0113)	(0.0102)
polity2	-0.0096	-0.0098	-0.0096
	(0.0103)	(0.0114)	(0.0113)
cap1	14.7576***	7.1875***	9.5507***
	(1.8614)	(2.2136)	(1.9963)
cap2	4.5841**	6.9784***	5.1845**
	(2.0082)	(2.2348)	(2.2338)
distance	-0.0009***	-0.0013***	-0.0011***
	(0.0002)	(0.0002)	(0.0002)
alliance	0.0533	-0.0344	0.0252
	(0.2204)	(0.2372)	(0.2192)
N	650	650	650

Significance levels: '***' indicates that p-value < 0.01, '**' indicates that $0.01 < p$-value < 0.05, and '*' indicates that $0.05 < p$-value < 0.1.

Standard errors are in parentheses (except in Column 4 where standard deviations of the simulations are reported) .

Note: As the model specification in Bayesian modeling relies on $y*$ instead of y, the spatial parameters are distinguished from the ones estimated in the standard probit by denoting the latter as ρ_d^a, ρ_o^a, and ρ_w^a.

TABLE 2.2 Marginal Effects from the Spatial OD Modeling
of MID Initiations, Europe, 1933-1941

Regressor	Direct Impacts	Network Impacts	Total Impacts
polity	-0.108%	0.003%	-0.105%
cap	2.148%	-0.549%	1.599%

*Note***:** The marginal effects of *polity* are calculated from a unit increase; and the marginal effects of *cap* are based on a 1% increase.

TABLE 2.3

**Impacts from Changing Germany's Capabilities
on Its Likelihood of Initiation**

Directed Dyads			Baseline	Impact of a 1% Decrease in Germany's Capabilities	Impact of a Decrease to the Sample Median
Germany	→	France	50.779%	-4.756%	-27.911%
Germany	→	Soviet Union	49.857%	-4.893%	-28.326%
Germany	→	Poland	46.821%	-4.794%	-27.082%
Germany	→	Lithuania	44.983%	-4.662%	-26.030%
Germany	→	Belgium	42.719%	-4.535%	-24.920%
Germany	→	Czech	40.560%	-4.680%	-24.960%
Germany	→	Luxembourg	39.633%	-4.492%	-23.970%
Germany	→	Austria	38.352%	-4.616%	-24.111%
Germany	→	Netherlands	38.241%	-4.292%	-22.826%
Germany	→	Switzerland	37.495%	-4.526%	-23.539%
Germany	→	U.K.	30.220%	-3.979%	-19.614%
Germany	→	Hungary	26.253%	-4.025%	-18.550%
Germany	→	Italy	24.187%	-3.839%	-17.334%
Germany	→	Sweden	23.552%	-3.651%	-16.607%
Germany	→	Latvia	22.873%	-3.683%	-16.465%
Germany	→	Yugoslavia	22.179%	-3.661%	-16.162%
Germany	→	Finland	18.337%	-3.176%	-13.570%
Germany	→	Estonia	17.246%	-3.129%	-13.049%
Germany	→	Romania	16.358%	-3.051%	-12.526%
Germany	→	Albania	15.731%	-2.946%	-12.049%
Germany	→	Bulgaria	14.586%	-2.823%	-11.323%
Germany	→	Ireland	12.836%	-2.392%	-9.719%
Germany	→	Spain	8.635%	-1.877%	-6.960%
Germany	→	Greece	7.453%	-1.722%	-6.147%
Germany	→	Portugal	5.411%	-1.302%	-4.509%
Mean: Germany as origin			27.81%	-3.66%	-17.93%
Mean: All dyads with Germany			22.09%	-2.23%	-11.24%
Mean: Whole system			10.18%	-0.13%	-0.60%

Note: Marginal Effects are calculated based on Germany's *cap* value being 0.0700 for the sample period.

Chapter 3

DO TRADE FLOWS INTERACT IN SPACE?

SPATIAL ORIGIN-DESTINATION MODELING OF GRAVITY

3.1 Abstract

Since the gravity model is widely used as an empirical tool for investigating bilateral trade flows, trade economists strive to improve it on two fronts: theoretical underpinnings and methodological soundness. Unfortunately, most empirical studies are conducted under an unrealistic assumption of independence among trade flows, even when a remoteness variable is included in the equation. By extending the spatial origin-destination (OD) techniques proposed by LeSage and Pace (2008), this study explores spatial interdependence among bilateral trade flows and proposes a spatial OD threshold Tobit as an improved estimation technique for the gravity model. This spatial threshold Tobit is designed to capture multiple forms of spatial autocorrelation embedded in "directional" trade flows while accounting for the corner solution where trade volumes are recorded as zero. This newly proposed model is then applied to export flows among 32 Asian countries in 1990. The empirical results provide evidence for the presence of all three types of spatial dependence: exporter-based, importer-based and exporter-to-importer-based. After further taking into account the multifaceted spatial correlation in bilateral trade flows, this study finds that the effect of conventional trade variables changes in a noticeable way. This finding implies that gravity models of trade may produce biased estimates if they fail to properly address spatial dependence.

72

3.2 Introduction

Since Tinbergen (1962)'s seminal work that introduced the gravity equation as an empirical specification of bilateral trade flows, the gravity model "has dominated empirical research in international trade" (Helpman et al., 2008, p. 442). In its basic form, the gravity model explains trade volumes based on the economic sizes (often measured by real GDP) of two trading partners and the distance separating them through a functional form analogous to Newton's Law of Gravity with stochastic features. Over time, this basic model has been extended by including other explanatory variables to help better understand the mechanism of trade (e.g., border effect, McCallum, 1995) and to evaluate institutional or policy impacts on trade flows (e.g., preferential trade agreement and membership in WTO, Feenstra, 2004). As Anderson (1979, p. 106) succinctly summarizes it, "[A]pplied to a wide variety of goods and factors moving over regional and national borders under differing circumstances, [the gravity model] usually produces a good fit." Given its strong explanatory power, along with the simplicity in formulation and ease of interpretation when expressed in a log-linear form, the gravity model has been widely viewed as a successful empirical tool for investigating trade flows (Anderson and van Wincoop, 2003).

Despite its popularity in empirical studies, the gravity model has been subject to several criticisms and recalibrations by trade economists. In the past forty years, trade scholars have endeavored to buttress two aspects of the gravity model: theoretical justification and econometric soundness. The fact that this model was not originally derived from any received theory of trade economics has urged researchers to question its validity and to probe into its theoretical underpinnings. To develop theoretical guidance for the applications of the gravity model, trade

73

economists have set forth trade models under different assumptions regarding market characteristics and types of trade goods, and verified the link between the gravity model and various trade theories (Anderson 1979; Bergstrand, 1985, 1989; Deardorff 1988; Anderson and van Wincoop, 2003; Helpman et al., 2008).

On the empirical front, several issues challenge the conventional econometric specification of the gravity model. First, the gravity model has usually been transformed into a log-linear form and then estimated with OLS due to computational convenience in empirical analysis. This simplicity in estimation does not come without a price. The log transformed dependent variable requires that all trade volumes under scrutiny be positive, because logarithm of zero is mathematically undefined. Hence, how to handle zero trade values has been an active research topic for applied trade economists. Second, heteroskedasticity has recently arisen at the center of the discussions on the econometric specification of the gravity model (e.g., Santos Silva and Tenreyro, 2006; Martin and Pham 2008). The potential for biased estimates when employing OLS estimation in the presence of heteroskedasticity leads to an important question: Which functional form of the gravity model is more appropriate, multiplicative or log-linear?

A large volume of literature has been devoted to the aforementioned two econometric issues. However, an equally important but largely neglected problem with the empirical modeling of international trade is how to deal with the interdependence of trade flows. Though spatial dependence between trade flows has long been suspected (Anderson and van Wincoop, 2003), only a very few empirical studies attempt to incorporate the spatial features of bilateral trade flows into the econometric specification of the gravity model. Porojan (2001) is the first to apply spatial econometric techniques to the gravity model and his research provides evidence for

74

the existence of spatial correlation among bilateral trade flows. Later, in responding to the growing interest in operationalizing "multilateral resistance" (Anderson and van Wincoop, 2003), Behrens et al. (2012) propose an econometric specification of the gravity model which is akin to spatial econometric modeling in structure. Notwithstanding the structural resemblance, the authors emphasize that their model is motivated by the theory-driven "multilateral resistance" terms rather than by an urge to directly model the spatial patterns of trade data. Admittedly, the notion of "multilateral resistance" perceives trade flows as interdependent. Moreover, recognizing the empirical difficulty arising from a lack of reliable data on price indices, which reflect multilateral resistance to trade and hence are expected to exhibit spatial dependence, LeSage and Pace (2009, p. 217) suggest the inclusion of spatial lags of the dependent variable to accommodate spatially dependent, unobserved variables, and provide further motivation for spatial modeling of international trade. All in all, these researches shed light on the significance of taking spatial dependence into account in applied international trade studies.

The methodological contribution of this study addresses modeling dependent variables that combine three empirically relevant, nonstandard statistical features: censored, dyadic, and spatially correlated data. By extending the recent advances of spatial econometrics in modeling spatial autocorrelation for data featuring origin-destination flows (see LeSage and Pace, 2008, 2009), this study proposes a spatial origin-destination (OD) threshold Tobit model to allow a censored and directed dyadic dependent variable, of which trade data represents a typical example. This newly proposed methodology is applied to a cross-section of bilateral export flows to concurrently deal with the potential multiple sources of spatial dependence in trade data as well as the problem of zero trade values. Since maximum likelihood estimation becomes infeasible in this case, this study develops a Bayesian procedure to estimate the model.

The empirical results of this study provide support for exporter-based (origin-based), importer-based (destination-based), as well as exporter-to-importer-based (origin-to-destination-based) spatial correlation in export flows. Economic sizes and geographical distance are statistically meaningful determinants of trade flows, although magnitudes of their impacts are not close to unity after taking spatial dependence into account. Besides, sizable network effects are detected for GDP, the omission of which obscures the mechanism through which economic sizes affect trade flows.

The rest of the study is organized as follows. Section 3.3 reviews the main theoretical developments of the gravity model and the technical issues involved in the econometric specification of the model. Section 3.4 discusses the technical difficulties in applying the spatial origin-destination (OD) model to censored data and introduces a spatial OD threshold gravity model. Section 3.5 presents a Bayesian procedure for model estimation. In section 3.6, the proposed spatial OD threshold gravity model is applied to examine bilateral trade flows among 32 Asian countries in the year of 1990, and the empirical results are discussed. Section 3.7 concludes.

3.3 Theoretical and Methodological Developments of the Gravity Model

3.3.1 Theoretical Contributions to the Legitimacy of the Gravity Model

As the earliest applications of the gravity equation to international trade were not derived from formal theoretical modeling, the validity of this model was often questioned (e.g., Polak, 1996). Since the early 1980s, efforts have been made towards the development of theoretical

underpinnings for the gravity model and as a result, theory-bounded studies have demonstrated that the gravity equation can be derived from very different trade frameworks (Deardorff 1989).

Anderson (1979) was the first researcher to develop a theory-guided gravity model. Assuming product differentiation by country of origin and employing distance as a proxy for transport costs, he derived the gravity equation based on a demand function with Constant Elasticity of Substitution (CES). Applying the same preferences, Bergstrand (1985, p. 474) demonstrated that the gravity equation can be derived as "a reduced form from a partial equilibrium subsystem of a general equilibrium model." Bergstrand (1989) advanced the microeconomic foundations for the gravity equation with the additional assumption of monopolistic competition, although he still used existing price indexes to approximate the complex price terms induced by the theoretical model (p.147). Helpman and Krugman (1985) and Helpman (1987) found that the gravity equation is consistent with trade theories based on imperfect competition. Deardorff (1998) reconciled the gravity equation with a classical theory of trade — the Heckscher-Ohlin framework, showing how the equation can be motivated from a factor endowment model.

Built upon Anderson (1979), Anderson and van Wincoop (2003) refined the theoretical derivation of the gravity model and highlighted the need to account for the endogeneity of prices, embodied in the multilateral resistance terms for exporting and importing regions. More recent theoretical explorations induce the gravity equation for models of international trade with heterogeneous firms (Melitz, 2003 and Helpman et al., 2008).

3.3.2 Econometric Issues of the Log-linear Gravity Model

While commenting on Deardorff's work, Bergstrand (1998, p. 23) accredited the fascination with the gravity equation to "the consistently strong empirical explanatory power of the model." Using the log-linearized gravity model allows researchers to obtain high R^2 values even in its "basic" specification where variation in trade flows is explained by GDPs of the country pairs in the sample and their bilateral distance (Bergstrand and Egger, 2011, p. 3). However, the log-linear gravity model has been criticized due to several econometric issues, such as zero trade values, heteroskedasticity, and spatial correlation among trade flows.

3.3.2.1 Presence of Zero Trade Values

Despite its appeal of simplicity in estimation, the log-linear form of the gravity model imposes the additional restriction that sampled trade data can have only positive values in order for the log-transformed dependent variable to be well defined. This restriction excludes zero-valued observations from estimation, obviously overlooking the fact that some country pairs do not trade with each other due to economic and/or geographical reasons. In this case, zero observations are not the result of rounding or recording errors and thus should be included in the estimation. In fact, zero trade flows are not uncommon. Noted as early as Linnemann (1966), half of the world's bilateral trade flows are zeros. Santos Silva and Tenreyro (2006) have 48% of their observations recorded as zeros while Helpman et al. (2008, p. 442) observe that "about half of the country pairs [in their sample] do not trade with one another."

Moreover, recent theoretical developments in trade literature rationalize the presence of zero trade flows in their theoretical models. Haveman and Hummels (2004, p. 213) derive the

gravity equation based on an assumption of incomplete specialization that permits the occurrences of zero values in trade data, "a fact that is difficult to reconcile with the complete specialization framework used to derive gravity equations." Helpman et al. (2008) develops a gravity model that takes into account firm heterogeneity and fixed costs. Their model predicts zero trade volume between two countries when "no firm in country i finds it profitable to export to country j" (p. 451) and therefore accounts for the self-selection of firms into the export markets. Hence, appropriately handling observations of zero trade volumes has emerged as an important empirical issue for trade economists.

Several strategies have been suggested in empirical analysis of bilateral trade to cope with the "zero problem". While some are proposed on an ad hoc basis, others are more theory motivated. The most commonly used approaches are either to drop all zero observations (e.g., Baier and Bergstrand, 2007) or arbitrarily add a very small positive number to all trade flows (e.g., Felbermayr and Kohler, 2006; Linders and de Groot, 2006). Although both approaches are simple to implement, they are theoretically unjustifiable and statistically unsound. Disregarding zero trade flows discards information and can lead to biased estimates, especially given that these zero-valued observations are not randomly distributed (Helpman et al., 2008; Burger et al., 2009). Likewise, the approach of replacing zero values with a small positive constant is less than advisable. The choice of the positive constant is arbitrary and regression estimates have been found to be sensitive to the selected constant (Flowerdew and Aitkin, 1982; Burger et al., 2009).

Helpman et al. (2008) suggest that there is a selection process hidden in the observed trade data that involves two-stages. The first stage concerns the decision of trade or no trade with another country while the second stage determines the volume of trade once the choice of trade is

made in the first stage. Their theoretical reasoning is appealing; however, identification of the model parameters requires different sets of explanatory variables for the two separate estimation stages. Since it is difficult to justify a proxy for the fixed cost of exports that is not also a proxy for the variable cost of trade (e.g., Anderson, 2011), the selection model has been viewed as less straightforward to implement.

Several studies use the standard Tobit model to estimate the gravity equation with zero flows (e.g., Rose, 2004; Anderson and Marcouiller, 2002). Eaton and Tamura (1994), in particular, put forward an innovative way to deal with zero trade volumes. In their modified gravity model, the volume of trade between a pair of countries records a positive value only if the potential trade exceeds some minimum amount. When a trade volume reaches a certain threshold, trading with another country becomes profitable and will thus occur. Rauch and Trindade (2002) follow their approach when estimating the impact of business and social networks on international trade. Later, Ranjan and Tobias (2007) adopt and name this method a threshold Tobit model. The threshold Tobit incorporates a self-selection process and zero flows are the outcomes of economic decision-making based on the potential profitability of engaging in bilateral trade.

3.3.2.2 Heteroskedasticity

Another extensively researched econometric issue in the empirical literature of the gravity model is heteroskedasticity.

Early exploration into this issue centered around omitted variables. It is plausible that one country exports different quantities to two countries even though the two export markets have the

80

same GDP and are equally distanced from the exporter due to historical, cultural, ethnic, or political factors. If these factors are correlated with the gravity variables but not controlled for, OLS estimation would yield biased estimates. To alleviate this problem, researchers extend the basic gravity model by including such variables as common language, colonial history, and trading blocs (e.g., Helpman et al., 2008; Porojan, 2001). An important insight from the recent theoretical development of the gravity equation is that the traditional specification suffers from omitted variable bias because it does not consider relative prices. Anderson and van Wincoop (2003) illustrate that the flow of bilateral trade is not only influenced by trade obstacles at the bilateral level, but also by the obstacles relative to all other countries, which they term as "multilateral resistance". Thus, omission of the "multilateral resistance" terms is thought to be a serious source of heteroskedasticity. However, since it is not easy to quantify "multilateral resistance", country-specific fixed-effects are suggested as a simple alternative to represent the essence of these terms.

More recently, Santos Silver and Tenreyro's influential study (2006) revives the attention to the heteroskedasticity problem but in a somewhat different context. The authors point out that the log-linear transformation of the gravity equation tends to cause the error terms to be correlated with the explanatory variables and thus strongly recommend estimating the gravity model in its original multiplicative form with a Poisson Pseudo Maximum Likelihood (PPML) estimator. Their work inspires intense discussions about the appropriate estimator for the gravity model. Martin and Pham (2008) contend that the PPML estimator is less subject to bias if heteroskedasticity is the only problem, but this estimator does not appear to be robust to the joint problems of heteroskedasticity and zero trade flows. Moreover, as the Poisson model assumes equidispersion (i.e., the conditional mean and the conditional variance of the dependent variable

are assumed to be equal), which is argued to be too restrictive, other nonlinear estimators are explored (e.g., Heckman Maximum Likelihood, Martin and Pham, 2008; Zero-Inflated Negative Binomial, Burger et al., 2009; Gamma Pseudo Maximum Likelihood (GPML), Martinez-Zarzoso, 2013).

Certainly, discussions on the appropriate estimator for the gravity model have enhanced researchers' awareness of the potential bias associated with the log-linear gravity formulation. Nonetheless, simulation studies that compare the performance of various estimators against different patterns of heteroskedasticity cannot set one estimator apart from the others. In many cases, the newly proposed estimation techniques actually produce less desirable results (e.g., Martin and Pham, 2008; Martinez-Zarzoso, 2013). It should also be noted that as the underlying data generating process is unknown, performance of the estimators can only be assessed with simulated data and the data simulation process, in particular the way in which zero-valued observations are generated could affect the performance of specific estimators (Martin and Pham, 2008, p. 14-15).

3.3.2.3 Spatial Interdependence among Trade Flows

Since Krugman (1991), trade scholars have begun to appreciate how geography matters to trade. However, most of the previous empirical studies of the gravity model of trade fail to explicitly account for the role of location (Porojan, 2001). It is reasonable to argue that trade flows are not isolated events, but interact with one another due to the geographical location of the members of trading pairs. Conventionally, a bilateral distance variable is included in the gravity model to proxy for trade barriers, especially trade costs; however, it is not sufficient for capturing spatial dependence across trade flows.

82

Using a relative distance measure distance defined as the actual distance divided by the average distance of the importing country from its trading partners, Polak (1996) brings up the idea of controlling for the location effect in the gravity model. His distance measure is in nature close to the use of standardized weight matrix in spatial econometric models (Porojan, 2001, p. 277). Similarly, Hamilton and Winters (1992) call for "more differentiated measures of 'distance'". Frankel and Wei (1998) echo Polak's insight and include both the distance variable and an additional "remoteness indicator" (calculated as the average of a country's distances to its trading partners, weighted by the partners' income) in their model specification.

On the other hand, Anderson and van Wincoop (2003, p. 170) critique the use of an "atheoretic 'remoteness' variable", stressing that "the remoteness index does not capture any of the other trade barriers." Based on their own theoretical reasoning of the gravity model, the authors advocate the inclusion of "multilateral resistance" variables in the empirical specification of the gravity model. The mathematical expression of the "multilateral resistance" terms elucidates the fact that bilateral trade depends not only on the trade barriers between a pair of regions but also on the average trade barriers that both regions face with all their trading partners, alluding to the need to control for interdependence when estimating the gravity equation systems. Since it is not straightforward to operationalize the notion of "multilateral resistance", they suggest the use of country-specific dummies as an alternative to capture such correlation in empirical analysis. Some applied works choose to specify more specific forms for the alluded correlation in their cross-country regressions. For instance, Baier and Bergstrand (2009) introduce income-weighted average bilateral distance and income-weighted average border variables for both the exporting and importing regions to construct the unobservable theoretically-motivated "multilateral resistance" terms.

Most recently, in their derivation of the gravity model, Behrens et al. (2012) adopt a 'dual' approach that hinges only on *observable* trade flows and accordingly propose an econometric specification of the gravity model that explicitly deals with "cross-sectional correlations among trade flows" (p. 774). The interdependence structure embedded in their gravity model indicates that a given trade flow depends on all other trade flows to the same destination. The authors purposely interpret their econometric specification as revealing cross-sectional correlation, not spatial correlation, among bilateral trade flows. By doing so, they try to stress the point that the autoregressive structure is originated from their theoretical model and not constructed out of the econometric concern to control for spatial effects, thus providing a better approach to quantify the notion of "multilateral resistance" highlighted by Anderson and van Wincoop (2003). At the same time, they also concede that a number of other measures (e.g., economic distance and socio-economic distance) have been employed to define the connectivity structure in spatial econometrics. In that sense, the formulation of their empirical model *de facto* attests to the presence of spatial effect in trade data. Their "interaction matrix" relates the trade flow X_{ij} from region i to j to all the trade flows from the other regions k to region j (p. 779). Essentially this matrix reflects an origin-based type of dependence, one of the three potential forms of dependence among flow data explored in LeSage and Pace (2008)'s spatial origin-destination model.

Overall, it appears that most empirical gravity models implicitly assume independence. Of the few studies that do allow for interdependence among trade flows, interdependence tends to be construed only in the narrow sense of relative trade barriers. Trade scholars rarely pay close attention to the potential spatial correlation among bilateral trade flows that is induced by geographical/spatial location, though spatial effects, including spatial spillovers, have been more

84

carefully investigated and documented in other issue areas such as growth and FDI (e.g., Weinhold, 2002; Blonigen et al., 2007). Arguably, similar spatial patterns may also exist in bilateral trade, since neighboring countries could have an impact on each other's trading behavior as a result of geographical proximity. This may be explained by technology spillover, infrastructure layout, or dissemination of ideas and policy orientations. Blum and Goldfarb (2006) show that the gravity model even holds in the case of digital goods consumed over the Internet, which do not have trading costs. Their finding that Americans are more likely to visit websites from nearby countries, even after controlling for language, income and immigrant stock, suggests that the spatial correlation existing among trade flows should be more wide-ranging than the mere manifestation of relative trade barriers. As cautioned by Anselin (1988), when such spatial effects are present, OLS estimator will be biased and spatial econometric techniques should be employed.

3.3.2.3.1 Explicitly Modeling Spatial Effects in the Gravity Model

Porojan (2001) takes note of the neglect of spatial effects in the international trade literature. Following Anselin's (1988) advice on how to handle data in the presence of spatial correlation, Porojan's research is the first to use spatial econometric models to estimate the gravity equation. Employing both import and export data for 15 EU member states and 7 OECD countries in 1995, he compares the performance of several spatial econometric specifications: the spatial error (SEM) model, the spatial autoregressive (SAR) model, and the spatial autocorrelation (SAC) model that contains spatial dependence in both the dependence variable and the error term with or without controlling for heteroskedasticity. The test results indicate that each spatial specification exhibits better model performance than OLS (p. 275). More

importantly, after explicitly taking into account the inherent spatial effects, Porojan finds that "the magnitude of the estimated parameters changes considerably and, with it, the measures on the predicted trade flows" (p. 266), thereby calling attention to the proper handling of spatial dependence in trade data.[24]

While their model specification is not motivated by spatial econometric modeling per se, Behrens et al. (2012) adapts spatial econometric techniques to deal with cross-sectional interdependence among trade flows and thus advances empirical analysis of spatial correlation in bilateral trade data. In their SARMA model, which includes an autoregressive term in the mean component of the model and a first-order moving average process for the errors, the "interaction matrix" that defines the autoregressive structure only takes care of exporter (origin)-based dependence as termed by LeSage and Pace (2008). It is reasonable to suspect that a similar dependency relationship may also exist among trade flows, which arises from the importing side.

A common drawback of the previous efforts with spatial modeling is that they do not perceive bilateral trade in the context of a directional flow involving both an exporter (or an origin) and an importer (or a destination) at the same time. Rather, their model specifications only concern dependence between trade flows coming from a selected side of the trading pairs, depending on whether import or export data is examined. As these approaches fail to consider the latent multiple sources of spatial correlation among bilateral trade flows, they are inadequate in capturing the impact of spatial dependence on flows of trade. In their study, Behrens et al. (2012) recognize the distinctive feature of origin-destination data and ascribe the absence of

---

[24] However, it should be noted that LeSage and Pace (2012) indicate that the non-linear relationship between y and X in the SAR and other spatial lag models makes it inappropriate to interpret the coefficient estimates as if they reflect linear regression slope estimates. For correct interpretation, they propose scalar summary measures to calculate the direct, indirect and total effects associated with changing explanatory variables in various types of spatial regression models.

applications of spatial econometric modeling in trade literature to a lack of well-developed origin-destination based models in spatial econometrics (P. 774). Although they briefly refer to LeSage and Pace's (2008) work on spatial econometric modeling of origin-destination flows as an exception, their model specification does not reflect any consideration of correlation between trade flows that might stem from the destination side. Furthermore, both Porojan (2001) and Behrens et al. (2012) treat the remaining spatial correlation in the error term. It is more desirable to model the spatial dependence process in the mean component of the model rather than push it to the disturbance term, especially when learning the specific source and impact of spatial correlation is informative to trade policy-making. Also, in their recent work, LeSage and Pace (2012) contend that the choice between the spatial autoregressive model and other types of spatial models, including the spatial error model and its variants, should be guided by the nature of spillovers. They define "global" spillovers as feedback or self-reinforcing effects and "local" spillovers as impacting only nearby or immediate neighbors. In this sense, international trade reflects "global" spillovers. Trade flows give rise to technology dissemination and economies of scale, which will lead to enhanced efficiencies and reduced production costs, and subsequently economic growth. Growth allows more trade opportunities and encourages a new round of diffusion of technology and economies of scale. Given the feedback effects of international trade, a spatial autoregressive type of model appears to be more appropriate in handling the spatial correlation that exists in trade flows.

As a way to justify spatial origin-destination modeling of flow data, LeSage and Pace (2008) invoke Griffith and Jones (1980)'s insight: flows stemming from an origin are "enhanced or diminished in accordance with the propensity of emissiveness of its neighboring origin locations," whereas flows into a destination are "enhanced or diminished in accordance with the

propensity of attractiveness of its neighboring destination locations." This observation should apply to international trade which also features a directional flow from an origin (the exporting country) to a destination (the importing country). Specifically, the attractiveness of an importer could be enhanced or diminished by that of its neighboring importers while the emissiveness of an exporter could be enhanced or diminished by that of its nearby exporters. Such spatial correlation may be due to technology and innovation spillovers among exporting countries that are in close proximity to each other; may arise out of the adoption of similar trade policy orientations among exporting countries that are neighbors; may be brought about by the established trading infrastructure which makes it economical to access a cluster of export markets that are geographically close to one another; or may be better explained by the competition between neighboring exporting countries or importing countries.

Hence, it will be less productive to model spatial correlation in bilateral trade flows without giving due consideration to the directional and dyadic features of trade data. Only by modeling spatial correlation in terms of the specific role assumed by each member in a trading pair (i.e., being in the status of an exporter or an importer), will a researcher be able to "zero in" the multiple sources of spatial dependence among trade flows. Accordingly, the spatial OD model proposed by LeSage and Pace (2008) offers a more promising solution to controlling for spatial effects when investigating international trade. The core idea of LeSage and Pace's modeling strategy is to extend the conventional spatial autoregressive (SAR) model by constructing three spatial lags of the dependent variable based on three different connectivity structures, which are meant to capture respectively spatial dependence arising from neighboring relationships among origin regions, among destination regions as well as dual neighboring relationships between both origins and destinations across trading pairs. The use of spatial lags is

88

recommended as a more efficient way to deal with omitted variables (LeSage and Pace, 2008; Porojan, 2001; Behrens et al., 2012).

It is also worth noting that both Porojan (2001) and Behrens et al. (2012) use the log-linear formulation of the gravity model while illustrating the need to account for interdependence when studying trade issues. As discussed earlier, this functional form of the gravity model requires that all sampled trade volumes to be positive. The inability to accommodate zero trade flows in the estimation unfortunately limits the applicability of their models, because it is quite common to observe zero values in trade data. Lebreton and Roi (2009) is the first attempt to apply LeSage and Pace's spatial OD modeling technique to trade data and it examines the effect of exchange rate volatility on import flows. By fitting the model to trade flows between Oceanian countries, their empirical results indicate the presence of both positive origin-based and destination-based dependence. However, the authors also point out a limitation that the spatial OD model is designed for continuous dependent variable and thus cannot handle censored data with zero trade values (Lebreton and Roi, 2009, p. 5).

This study extends the spatial OD model to accommodate the limited dependent variable nature of bilateral trade data.[25] The newly proposed spatial gravity model is designed to capture the multiple types of spatial dependence embedded in trade data while allowing for zero-valued flows in the sample. Thus, this model precludes the inconsistency problem suffered by OLS due to the presence of spatial correlation and accordingly makes full use of the trade flow data.

---

[25] As LeSage and Pace (2009) point out, gravity models have often been employed to explain OD flows which are "fundamentally spatial in nature". Their spatial OD modeling extends the traditional gravity model to handle spatial dependence. Moreover, acknowledging the possible presence of a large number of zero flows in flow data such as population migration flows and international trade flows, they mention the potential of extending the spatial OD modeling to treat the zero flows problem.

3.4 Model Specification

LeSage and Pace's (2008) spatial OD model represents a more tailored approach to dealing with spatial correlation in data featuring an origin-to-destination flow such as bilateral trade flows. However, their model specification is not readily applicable to the gravity model of trade if the data contain zero trade values, as their spatial OD model is proposed for continuous dependent variables. When zero trade flows are present in a data set, it becomes an issue of a limited dependent variable and it normally requires a limited dependent variable approach, such as a Tobit-type (spatial) model. Moreover, zero trade values pose another technical challenge when the log-linear form of the gravity model is employed for the sake of the computational ease. To handle zero trade observations and spatial autocorrelation concurrently, this study employs a spatial OD model in conjunction with the threshold gravity model first introduced by Eaton and Tamura (1994).[26]

According to the threshold gravity model, the volume of trade between a pair of countries records a positive value only if the potential trade exceeds a certain minimum amount (i.e., threshold). As Ranjan and Tobias (2007) point out, the threshold Tobit allows us to "remain true to the mixed discrete-continuous nature of trade data" by "assign[ing] meaningful probabilities to the event of no trade" and also helps to "avoid the problem of taking the log of zero" (p. 818).

Following Eaton and Tamura (1994)'s framework, the trade flow from country j to country k is modeled as:

---

[26] Applying a Bayesian procedure to the threshold gravity model proposed by Eaton and Tamura (1994), Ranjan and Tobias (2007) examine the impact of contract enforcement on bilateral trade flows. LeSage and Pace (2009) briefly mention the potential to combine their spatial Tobit model with the idea of a threshold value of trade initiated by Eaton and Tamura and later adopted by Ranjan and Tobias.

$$\ln\left(y_{jk}^* + a\right) = X_{jk}\beta + \varepsilon_{jk}, \quad \varepsilon_{jk} \sim N\left(0,\sigma^2\right) \tag{3.1}$$

where
$$y_{jk} = \begin{cases} y_{jk}^*, & if \ \ y_{jk}^* > 0 \\ 0 \ \ , & if \ \ y_{jk}^* \leq 0 \end{cases} \quad and \quad a > max[0, -y_{jk}^*]$$

Or equivalently,

$$\underbrace{y_{jk}^*}_{\text{desired trade}} = \underbrace{- \ a}_{\text{fixed cost}} + \underbrace{exp\left(X_{jk}\beta + \varepsilon_{jk}\right)}_{\text{potential trade}} \tag{3.2}$$

From an economic point of view, the threshold parameter a can be interpreted as a fixed or average cost of international trade,[27] and y_{jk}^* represents the *desired* amount of bilateral trade. The actual *observed* trade volume, y_{jk} , equals y_{jk}^* if the potential trade more than covers the fixed cost. On the contrary, if the potential trade falls below the fixed cost and bilateral trade becomes undesirable or unprofitable (i.e., desired trade is negative), then the *observed* trade volume y_{jk} shows up as zero. That is to say, trade will occur only when trade is desired (when $y_{jk}^* > 0$). In this sense, the practice of adding an arbitrary "1" to all sampled trade data to have their logged term defined is not convincing, as it imposes an arbitrary one-unit trade cost.

Also, it is worth noting that technically speaking, the log-linear formulation of the model may help alleviate possible heteroskedasticity, as it is known that homoskedasticity in logs allows a reasonable heteroskedasticity in levels (e.g., LeSage and Thomas-Agnan, 2012).

Using matrix notation, equation (3.1) can be rewritten as:

---

[27] Similarly, Rauch and Trindade (2002, p. 119) think of a as "an amount of 'melting' that occurs as soon as the trip starts, independent of the distance travelled." See also Ranjan and Tobias (2007).

$$\ln(y^* + a \cdot \iota_N) = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I) \tag{3.3}$$

LeSage and Pace (2008)'s spatial OD model employs three spatial lag terms of the dependent variable defined through the weight matrices $\mathbf{W}_d$, $\mathbf{W}_o$, and $\mathbf{W}_w$ to model respectively spatial correlation stemming from neighboring relationships among the exporting countries, among the importing countries, as well as concurrent neighboring relationships across trading pairs. The construction of weight matrices W_d, W_o, and W_w is explained in detail in Chapter 2 (also see LeSage and Pace, 2008). As explained above, the left hand side of equation (3.3) represents the logged value of potential trade. If spatial effects are present, they are expected to influence the latent trade volumes. Taking the log of the dependent variable attempts to correct for potential heteroskedasticity. Therefore, the spatial OD modeling of the threshold Tobit is specified as follows:

$$\ln(y^* + a \cdot \iota_N) = \rho_d \mathbf{W}_d \cdot \ln(y^* + a \cdot \iota_N) + \rho_o \mathbf{W}_o \cdot \ln(y^* + a \cdot \iota_N)$$

$$+ \rho_w \mathbf{W}_w \cdot \ln(y^* + a \cdot \iota_N) + X\beta + \varepsilon \tag{3.4}$$

For brevity, we use v^* to denote $\ln(y^* + a \cdot \iota_N)$ and accordingly, (3.4) can be rearranged as:

$$(\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w) \cdot v^* = X\beta + \varepsilon \tag{3.5}$$

Setting $\boldsymbol{A} = (\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)$, we can simplify equation (3.5):

$$\boldsymbol{A} \cdot v^* = X\beta + \varepsilon \tag{3.6}$$

or equivalently,

92

$$v^* = A^{-1}X\beta + A^{-1}\varepsilon \qquad (3.7)$$

Eaton and Tamura (1994) rely on maximum likelihood for the estimation of their threshold Tobit model. However, once spatial lags are introduced into the model, maximum likelihood estimation (MLE) becomes infeasible, as explained below. Hence, this study develops a Bayesian estimation algorithm which properly accounts for the discrete-continuous feature of trade data while avoiding the computational difficulty associated with an ML estimator.

3.4.1 Computational Difficulty with MLE

When spatial interaction is taken into consideration, an ML estimator is not an appropriate choice in dealing with the involved relationships among observations. This issue becomes clear when we inspect the likelihood function for positive observed values of y.

For notational convenience, in this subsection, we use the subscript "i" to denote observations instead of "j, k" which indicates the exporting and importing countries for each observation of bilateral trade volume. From expression (3.6), we get

$$[A]_i \cdot v^* = [X]_i\beta + \varepsilon_i \qquad (3.8)$$

where $[A]_i$ and $[X]_i$ designate the ith row of the A and X matrices, respectively.

In keeping with Eaton and Tamura (1994)'s approach, by means of a Jacobian transformation, the likelihood function for positive y_i is written as

$$Pr(y_i > 0) \cdot f(y_i | y_i > 0) = f(\varepsilon_i) \cdot \frac{A_{i,i}}{y_i + a}$$

93

$$= \frac{A_{i,i}}{y_i + a} \cdot \frac{1}{(2\pi\sigma^2)^{1/2}} exp\left\{-\frac{([A]_i \cdot v^* - [X]_i\beta)^2}{2\sigma^2}\right\} \qquad (3.9)$$

with $A_{i,i}$ denoting the element in the ith row and column of A.

As can be seen from (3.9), vector y^* includes both observed values and unobserved values, which would greatly increase the computational difficulty of ML estimation.

3.4.2 A Bayesian Approach to the Spatial OD Threshold Tobit Model

With the assumption that $\varepsilon \sim N(0,\ \sigma^2 I)$, spatial interdependence induces a multivariate normal distribution for $v^* = \ln(y^* + a \cdot \iota_N)$ as shown in (3.10).

$$v^* = \ln(y^* + a \cdot \iota_N) \sim N(A^{-1}X\beta,\ \sigma^2(A'A)^{-1}) \qquad (3.10)$$

Accordingly, $y^* + a \cdot \iota_N \sim Lognorm(A^{-1}X\beta,\ \sigma^2(A'A)^{-1})$ and y^* follows a shifted multivariate log-normal distribution with conditional prior density

$$\pi(y^*|\beta,\ a,\ \rho_d,\ \rho_o,\ \rho_w,\ \sigma^2) = \frac{|A| \cdot (\sigma^2)^{-\frac{N}{2}}}{(y_1^*+a) \cdot (y_2^*+a) \cdots (y_N^*+a)(\sqrt{2\pi})^{\frac{N}{2}}}$$

$$\cdot exp\left\{-\frac{1}{2\sigma^2}(A \cdot v^* - X\beta)'(A \cdot v^* - X\beta)\right\} \qquad (3.11)$$

Following the common practice of Bayesian analysis, this study assigns a multivariate normal prior for β, an inverse-gamma prior for σ^2,[28] and diffuse priors for the threshold parameter a as well as the spatial lag parameters, ρ_d, ρ_o, and ρ_w. Specifically,

$$\beta \sim N\left(\mathbf{0}, \boldsymbol{T}\right), \qquad \mathbf{T} = \boldsymbol{I}_k \cdot 10^6$$

$$\rho_i \sim U\left(-1, 1\right), \quad i = d, o, w$$

$$\pi(\sigma^2) \propto (\sigma^2)^{-1}$$

$$a \sim U\left(0, \boldsymbol{D}\right),^{29} \quad \boldsymbol{D} = 0.005$$

Given the priors described above, we can derive the conditional posterior distributions of model parameters as follows by applying Bayes' Theorem:

$$p(\beta \mid \sigma^2, \rho_d, \rho_o, \rho_w, a, y^*)$$

$$\propto exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{A} \cdot v^* - \boldsymbol{X}\beta)'(\boldsymbol{A} \cdot v^* - \boldsymbol{X}\beta)\right\} \cdot exp\left(-\frac{1}{2}\beta'\boldsymbol{T}^{-1}\beta\right)$$

$$\propto exp\left\{-\frac{1}{2\sigma^2}[\beta - (\boldsymbol{X}'\boldsymbol{X} + \sigma^2\boldsymbol{T}^{-1})^{-1}\boldsymbol{X}'\boldsymbol{A}v^*]'(\boldsymbol{X}'\boldsymbol{X} + \sigma^2\boldsymbol{T}^{-1})[\beta - (\boldsymbol{X}'\boldsymbol{X} + \sigma^2\boldsymbol{T}^{-1})^{-1}\boldsymbol{X}'\boldsymbol{A}v^*]\right\}$$

which is the kernel of a multivariate normal distribution. Therefore, the conditional posterior of β is:

---

[28] When we let both the shape parameter and the scale parameter of an inverse-gamma distribution infinitely approach zero, then its density is $\pi(\sigma^2) \propto (\sigma^2)^{-1}$.
[29] The specified upper bound for a is based on the threshold estimate obtained using Eaton and Tamura's threshold Tobit model and is also compatible with the Bayesian estimation results reported in Ranjan and Tobias (2007) after adjusting for the differences in the measurement of trade data. Moreover, it is noteworthy that adjusting the upper bound of a does not much affect the spatial parameters, which indicates the robustness of spatial effects and therefore the need to control for spatial correlation.

$$\beta \,|\, \sigma^2, \, \rho_d, \, \rho_o, \, \rho_w, \, a, \, y^* \sim N\left((X'X + \sigma^2 T^{-1})^{-1} X'Av^*, \, \left(\frac{X'X}{\sigma^2} + T^{-1}\right)^{-1}\right). \qquad (3.12)$$

And,

$$p(\sigma^2 | \beta, \rho_d, \rho_o, \rho_w, a, y^*)$$

$$\propto (\sigma^2)^{-\frac{N}{2}} \cdot exp\left\{-\frac{1}{2\sigma^2}(A \cdot v^* - X\beta)'(A \cdot v^* - X\beta)\right\} \cdot (\sigma^2)^{-1}$$

$$= (\sigma^2)^{-\left(\frac{N}{2}\right)-1} \cdot exp\left\{-\frac{\frac{1}{2}(A \cdot v^* - X\beta)'(A \cdot v^* - X\beta)}{\sigma^2}\right\}$$

Due to the use of a conjugate prior for σ^2, its conditional posterior also follows an inverse-gamma distribution:

$$\sigma^2 | \beta, \rho_d, \rho_o, \rho_w, a, y^* \sim IG\left(\frac{N}{2}, \frac{1}{2}(A \cdot v^* - X\beta)'(A \cdot v^* - X\beta)\right). \qquad (3.13)$$

With presumed independence among the spatial lag parameters, the posterior conditionals for ρ's all take the same form as (3.14):

$$p(\rho_d | \beta, \rho_o, \rho_w, \sigma^2, a, y^*) \propto I(-1 < \rho_i < 1) \cdot |A|$$

$$\cdot exp\left\{-\frac{1}{2\sigma^2}(A \cdot v^* - X\beta)'(A \cdot v^* - X\beta)\right\} \qquad (3.14)$$

Here $I(-1 < \rho_i < 1)$ denotes an indicator function which assumes the value of 1 if ρ_d is in the open interval $(-1, 1)$.[30] As the conditional posterior distributions of ρ's do not have a standard

---

[30] This restriction is imposed both due to model assumptions (i.e., ρ's are autoregressive parameters) and computational feasibility (see LeSage and Pace, 2009). LeSage and Pace further suggest imposing the stability restriction that $|\sum_i \rho_i| < 1, i = d, o, w$ (p. 221).

www.manaraa.com

form and thus cannot be sampled directly, this study employs a Metropolis-Hastings algorithm to obtain samples for ρ's.

The conditional posterior distribution for the threshold parameter, a, is obtained as:

$$p(a|\beta, \sigma^2, \rho_d, \rho_o, \rho_w, y^*) \propto I(0 < a < D) \cdot |A| \cdot \frac{1}{(y_1^* + a) \cdot (y_2^* + a) \cdots (y_N^* + a)}$$

$$\cdot exp\left\{-\frac{1}{2\sigma^2}(A \cdot v^* - X\beta)'(A \cdot v^* - X\beta)\right\} \qquad (3.15)$$

This distribution has an unknown form as well. Again a Metropolis-Hastings sampling scheme is used to simulate draws from (3.15).[31]

3.4.2.1 Sampling of Latent Trade Values y^*

For notational convenience, we introduce $\mu = A^{-1}X\beta$ and $\Omega = \sigma^2(A'A)^{-1}$ and express (3.10) as

$$v^* \sim MVN(\mu, \Omega) \qquad (3.16)$$

According to the setup of the threshold Tobit model, the latent variable y_i^* equals observed trade flow y_i whenever the former turns out to be positive. Otherwise, the latent trade flow is not observed. Therefore, with a predetermined a, the vector of v^* is comprised of both known and unknown values.

By utilizing the properties of the multivariate normal distribution, LeSage and Pace (2009) design an elegant way to obtain latent y_i^* for censored observations. To this end, trade

---

[31] Based on generated data experiments, Ranjan and Tobias (2007, p. 826) choose to sample this distribution via grid approximation.

97

data needs to be rearranged and then have the zero values stacked on top of the non-censored ones. This way, the resultant v^* vector encompasses two sub-vectors v_1^* and v_2, with v_1^* corresponding to the n_1 censored observations and v_2 containing the n_2 non-censored observations only. Mathematically, $v^* = \left(v_1^{*\prime}, v_2^{\prime}\right)^{\prime}$, where v^*, v_1^* and v_2 are three column vectors with dimensions N by 1, n_1 by 1 and n_2 by 1, respectively; and $N = n_1 + n_2$.

As LeSage and Pace (2009) astutely point out, we can treat the censored observations contained in v_1^* as random variables, and construct the mean vector and covariance matrix for this random vector conditional on the v_2 vector of uncensored observations. In so doing, we can generate latent y_i^*'s from the derived distribution for the random vector v_1^*. Specifically, the conditional posterior distribution for the n_1 censored observations can be expressed as a multivariate truncated normal distribution $v_1^* \sim TMVN(\mu_1^*, \Omega_{11}^*)$ subject to the constraints that $v_1^* \leq \ln(a) \cdot \iota_{n_1}$, drawing upon the fact that v_1^* and v_2 are the two partitioned vectors of a multivariate normal.

Analogous to the partitioning of the multivariate normal v^* into two sub-vectors v_1^* and v_2, the mean vector and covariance matrix of v^* are correspondingly partitioned as $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ with sizes $\begin{bmatrix} n_1 \\ n_2 \end{bmatrix}$ and $\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$ with sizes $\begin{bmatrix} n_1 \times n_1 & n_1 \times n_2 \\ n_2 \times n_1 & n_2 \times n_2 \end{bmatrix}$. Furthermore, the conditional distribution of v_1^* given v_2 is a (n_1-variate) multivariate normal with the mean vector and covariance matrix defined as follows:

$$\mu_1^* = E(v_1^*|v_2) = \mu_1 + \Omega_{12}\Omega_{22}^{-1}(v_2 - \mu_2) \tag{3.17}$$

$$\Omega_{11}^* = Var(v_1^*|v_2) = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21} \tag{3.18}$$

With the precision matrix $\Psi = \Omega^{-1} = (1/\sigma^2)(A'A)$, Pace and LeSage (2007) point out that by applying Corollary 8.5.12 in Harville (1997), the mean vector of v_1^* can be rewritten as

$$\mu_1^* = \mu_1 - \Psi_{11}^{-1} \Psi_{12}(v_2 - \mu_2)$$

to save computational time.

Given (3.17) and (3.18), we have

$$\ln\left(y_1^* + a \cdot \iota_{n_1}\right) = v_1^* \sim TMVN(\mu_1^*, \ \Omega_{11}^*) \qquad s.t. \quad v_1^* \leq \ln(a) \cdot \iota_{n_1} \tag{3.19}$$

where ι_{n_1} is a column vector of ones with n_1 elements.

To generate the random vector v_1^*, this study turns to the Geweke-Hajivassiliou-Keane (GHK) multivariate normal simulator (Geweke, 1991; Hajivassiliou, 1990; and Keane, 1994), a technique that samples recursively from truncated univariate normals after a Cholesky transformation. The GHK algorithm is described below.

First, draw a random vector

$$\epsilon \sim N(0, \ I) \qquad s.t. \qquad a^* \equiv -\infty - \mu_1^* \leq C\epsilon \leq \ln(a) \cdot \iota_{n_1} - \mu_1^* \equiv b^*$$

where C denotes the lower triangular Cholesky factor of Ω_{11}^*, $CC' = \Omega_{11}^*$.

Due to the triangular structure of C, the restrictions on ϵ are recursive. To be exact,

$$\epsilon_1 \sim N(0,1) \qquad s.t. \qquad a_1^*/C_{11} \leq \epsilon_1 \leq b_1^*/C_{11} \ ,$$

$$\epsilon_2 \sim N(0,1) \qquad s.t. \qquad (a_2^* - C_{21}\epsilon_1)/C_{22} \leq \epsilon_2 \leq (b_2^* - C_{21}\epsilon_1)/C_{22} \ ,$$

99

and for $i = 2, \ldots, N$,

$$\epsilon_i \sim N(0, 1) \qquad s.t. \qquad (a_i^* - \sum_{j=1}^{i-1} C_{ij}\, \epsilon_j)/C_{ii} \le \epsilon_i \le (b_i^* - \sum_{j=1}^{i-1} C_{ij}\, \epsilon_j)/C_{ii} \, .$$

Thus ϵ_i can be sampled sequentially from univariate truncated normals. The simulated vector ϵ

and the relation $v_1^* = \mu_1^* + C\epsilon$ will give the desired truncated random vector v_1^*.

By transforming the equation $\ln\!\left(y_1^* + a \cdot \iota_{n_1}\right) = v_1^*$, we can solve for y_1^* and obtain

$$y_1^* = exp(v_1^*) - a \cdot \iota_{n_1} = exp(\mu_1^* + C\epsilon) - a \cdot \iota_{n_1}$$

in order to simulate values of y^* corresponding to observed zero trade flows.

3.4.2.2 Bayesian Simulation Procedure for the Spatial Threshold Tobit Model

This study fits the spatial OD threshold Tobit model using a Gibbs sampler, which

sequentially samples the conditional distributions of the model parameters following the steps

sketched below:

1. $p\left(\beta \,\middle|\, \sigma^{2\,(0)},\, \rho_d^{(0)},\, \rho_o^{(0)},\, \rho_w^{(0)},\, a^{(0)},\, y^{*\,(0)}\right)$, which is a multivariate normal distribution

    with mean and variance defined in (3.12). Label the sampled vector β as $\beta^{(1)}$.

2. $p\left(\sigma^2 \,\middle|\, \beta^{(1)},\, \rho_d^{(0)},\, \rho_o^{(0)},\, \rho_w^{(0)},\, a^{(0)},\, y^{*\,(0)}\right)$, which has an inverse-gamma distribution with

    the shape and scale parameters specified in (3.13). Label this sampled σ^2 as $\sigma^{2\,(1)}$.

3. $p\left(\rho_d \,\middle|\, \beta^{(1)},\, \sigma^{2\,(1)},\, \rho_o^{(0)},\, \rho_w^{(0)},\, a^{(0)},\, y^{*\,(0)}\right)$, which can be acquired by means of a

    Metropolis-Hastings sampler based on a normal jumping density, along with rejection

    sampling in order to confine ρ_d to the $(-1, 1)$ interval. Label this updated value $\rho_d^{(1)}$.

100

4. $p\left(\rho_o \big| \beta^{(1)}, \sigma^{2(1)}, \rho_d^{(1)}, \rho_w^{(0)}, a^{(0)}, y^{*(0)}\right)$, which applies the same Metropolis-Hastings algorithm as in step (3). The newly updated value for ρ_d is used when making a draw for ρ_o.

5. $p\left(\rho_w \big| \beta^{(1)}, \sigma^{2(1)}, \rho_d^{(1)}, \rho_o^{(1)}, a^{(0)}, y^{*(0)}\right)$, which is similar to steps (3) and (4), except that now the updated values for both ρ_d and ρ_o are employed.

6. $p\left(a \big| \beta^{(1)}, \sigma^{2(1)}, \rho_d^{(1)}, \rho_o^{(1)}, \rho_w^{(1)}, y^{*(0)}\right)$, which also takes a Metropolis-Hastings substep along with rejection sampling to simulate draws limited to the interval $(0, \boldsymbol{D})$.

7. Sampled values for the latent y_1^* can be obtained through a transformation of the draws generated from the right-truncated multivariate normal v_1^*, as described in Section 3.4.2.1.

This process is repeated to collect a large sample of simulated values that can be used to make valid inferences with respect to the model parameters.

3.5   Data, Construction of Weight Matrices, and Effects Estimates

3.5.1   The Data

To illustrate the spatial effects among bilateral trade flows, this study employs the basic setup of the gravity equation and fits the spatial OD threshold Tobit to a sample of 32 Asian countries in 1990 (see Appendix I). The explanatory variables, $\boldsymbol{X}$, include real GDP of the two economies in a pair, the distance within each pair, as well as a contiguity dummy.[32] The dependent variable is bilateral exports. Data on export volumes, GDP, and bilateral distance are

---

[32] The explanatory variables are all log transformed except the contiguity variable.

collected from Santos Silva and Tenreyro (2006)'s dataset.[33] In the original data, trade volumes are measured in thousands of US dollars and GDP in level term. For computational manageability, this study records export volumes in billions of US dollars and GDP in millions of US dollars.

When a contiguity variable is included in Santos Silva and Tenreyro's dataset, it is based only on land borders. This operationalization may be too restrictive because if two countries are separated only by a small body of water, they are *de facto* neighbors. For this reason, this study adopts a broader definition of contiguity which acknowledges not only a land border, but also a water border. The contiguity data are retrieved from the Expected Utility Generation and Data Management Program (EUGene) <http://www.eugenesoftware.org/> Version 3.204, which allows the user to choose from several different distances of water body under which two countries are to be considered as contiguous. This study uses a more conservative separation distance of less than 25 miles of water body as an alternative criterion for determining contiguity.

3.5.2 Handling Weight Matrices

3.5.2.1 Choice of the Weight Matrices

In spatial analysis, the weight matrices are used to capture the inherent spatial correlation in the data. Therefore, these matrices should be constructed in view of the characteristics of the data under study to make them more relevant to the embedded spatial patterns.

---

[33] Santos Silva and Tenreyro post their data and definition of variables at http://privatewww.essex.ac.uk/~jmcss/LGW.html. It needs to be noted that when comparing different estimators using the Anderson-van Wincoop (2003) gravity model, which controls for multilateral resistance by including exporter- and importer-specific effects, Santos Silva and Tenreyro (2006) do not use countries' GDPs as explanatory variables as only bilateral variables can be identified given the cross-sectional data employed.

Porojan (2001, p. 271) utilizes a contiguity weight matrix wherein countries that share a land border or separated by a small body of water are coded as contiguous. This choice is better than using a predetermined capital distance as the cutoff for denoting neighbors, especially for trade participants with a large territory, such as China and Russia, which might require the cutoff distance to be quite large in order to capture possible spatial interactions. On the other hand, Porojan's weight matrix does not reflect the dyadic and directional features of flow data by differentiating the sources of spatial correlation.

As pointed out by Behrens et al. (2012), the weight matrix may be defined in different ways. Behrens et al.'s weight matrix is called the "interaction matrix" to be distinguished from the more common, distance-based formulation, because it assigns weights based on the population ratio of the exporting country in a trading pair to the total population of countries in the sample. However, this way of defining the interdependence structure of trade flows is not compelling, because it rigidly sets the influence of one trade flow invariable with respect to all other relevant trade flows. For example, if two trading pairs share the same importing country and have a similar population size in the respective exporting country (i.e., same ratio to the total population of the sample) but are differentiated by the bilateral distance within each pair, then all other bilateral trade flows involving the same importing country are supposed to exert the exactly same effect on the trade volumes of the stated two pairs according to their interaction matrix. Further, though Behrens et al. acknowledge that trade data features an origin-to-destination flow, their model is incapable of handling the complex connectivity structure embodied in such directional flows.

It can be argued that geographical distance matters to trade behavior in different ways. Admittedly, transportation cost is an important factor when countries decide with whom to trade. Yet, it is important to note that physical proximity allows countries to benefit from the spread of technologies, ideas, and policies, which are all conducive to the promotion of trade (e.g., technology spillover, LeSage et al., 2007). Therefore, this study will still rely on a first-order contiguity weight matrix $\mathbf{W}$, where $\mathbf{W}_{ij}$ ($i \neq j$) is coded as "1" if the two members of a pair are contiguous and coded as "0" otherwise. A row-standardized weight matrix $\mathbf{W}$ based on the stated definition of contiguity is included in Appendix II. This $\mathbf{W}$ matrix is then adapted to the neighboring relationships unique to OD flows to build the three spatial weight matrices $\mathbf{W}_d$, $\mathbf{W}_o$, and $\mathbf{W}_w$ as proposed by LeSage and Pace (2008) to specifically model origin-centric, destination-centric and origin-to-destination-centric dependence. Through Kronecker product operations, $\mathbf{W}_d$, which equals $\mathbf{I}_n \otimes \mathbf{W}$, embodies the notion that factors causing trade flows from an exporting country to an importing country may induce or dampen similar flows to nearby importers; $\mathbf{W}_o = \mathbf{W} \otimes \mathbf{I}_n$ is intended to reflect origin-based dependence wherein a country's exporting behavior may simulate or impede similar flows of trade from its neighboring exporters to the same destination; and $\mathbf{W}_w = \mathbf{W} \otimes \mathbf{W}$ represents a second-order connectivity between the neighborhood of an exporting country and the neighborhood of an importing country.[34]

3.5.2.2 Elimination of Self-Directed Pairs

In keeping with the structure of the weight matrix $\mathbf{W}$, self-directed pairs are included in the data by construction. However, when the values of the dependent variable are set to zero for

---

[34] In this study, lower case n denotes the number of countries in the sample, whereas the upper case N stands for the total number of observations, which equals n^2 by construction according to the spatial OD modeling set forth by LeSage and Pace (2008), and equals $n^2 - n = n(n-1)$ after self-directed pairs are removed from the data using the elimination step proposed in the next section.

the self-directed pairs, it fails to distinguish the zero values for these observations from those observed for no-trade pairs, thus leading to biased estimation. To avoid this bias, several researchers include an additional set of explanatory variables to estimate a model of self-directed pairs in tandem with a model of bilateral flows (see LeSage and Pace, 2008).

Behrens et al. (2012) include internal absorptions in their data; however, their construction of the "interaction matrix" assigns no weights to self-directed pairs (i.e., own trade flows), thus excluding the role of self-directed pairs from the interactive network. Santos Silva and Tenreyro (2006) consider only non-self-directed trade pairs in their investigation of the proper functional form for the gravity equation. Accordingly, it is difficult to assume that bilateral trade flows and internal flows operate under the same mechanism and thus need to be estimated jointly, especially for the cross-section considered in the application of this chapter. In 1990, external trade volumes and internal trade flows were not on a comparable scale for many of the sample countries, either because they were not yet capable of participating in international trade or because they focused on an inward looking trade policy and international trade only accounted for a very small part of their national income. In this sense, it may not be necessary to include self-directed pairs in the model. However, to accommodate diverging perspectives, this study will fit the spatial OD threshold Tobit model to the data augmented with internal flows as an empirical exercise.[35] The latter case does not require the elimination process that will be

---

[35] As a crude measure, this study calculates internal trade flows as the difference between GDP and trade balance as suggested by Lebreton and Roi (2009, p.5). To maintain data consistency, the GDP data is multiplied by external balance on goods and services (% of GDP) to back calculate trade balance. Data on external balance on goods and services (% of GDP) is retrieved from World Development Indicators (WDI) online version. For Cambodia and United Arab Emirates, this data is not available for the year 1990, we instead use a later year (1993 for Cambodia and 2001 for UAE) when this information first becomes available. Data on internal distance is taken from the *GeoDist* database compiled by Thierry Mayer and Soledad Zignago, which can be downloaded at http://www.cepii.fr/anglaisgraph/bdd/distances.htm. The internal distance of a country is computed as $d_{ii} = .67\sqrt{area/\pi}$, where country area is measured in square kilometers.

described momentarily, which removes the structural rigidity of the weight matrices associated with self-directed dyads.

To eliminate the role of self-directed pairs from the estimation, the $(n(m-1)+m)^{\text{th}}$ rows and columns of the weight matrices need to be removed, with n denoting the number of countries in the sample and $m = 1, 2, \cdots, n$. This can be accomplished by pre-multiplying each weight matrix by a "selection" matrix $\boldsymbol{B}$ and post-multiplying the resulting matrix by the transpose of $\boldsymbol{B}$. Here $\boldsymbol{B}$ represents a $n(n-1)$ by n^2 sparse matrix as defined in Chapter 2. The elimination procedure that removes self-directed pairs frees the researcher from the inadvertent constraints innate to the construction of spatial weight matrices. Using matrix notation, this process can be expressed as

$$\boldsymbol{B}v^* = \rho_d \boldsymbol{B}\mathbf{W}_d^R \boldsymbol{B}'\boldsymbol{B}v^* + \rho_o \boldsymbol{B}\mathbf{W}_o^R \boldsymbol{B}'\boldsymbol{B}v^* + \rho_w \boldsymbol{B}\mathbf{W}_w^R \boldsymbol{B}'\boldsymbol{B}v^* + \boldsymbol{B}X\boldsymbol{\beta} + \boldsymbol{B}\varepsilon, \qquad (3.20)$$

where v^* is defined as in Section 3.4.4. $\mathbf{W}_d^R$, $\mathbf{W}_o^R$, and $\mathbf{W}_w^R$ signify a renormalization of the modified weight matrices which exclude neighboring relationships with self-directed pairs, though renormalization takes place after the elimination step.

3.5.2.3 Weight Matrices Compliant with the Restructured Data

As explained in Section 3.4, we need to rearrange the data to stack zero-valued observations above non-zero ones in order to take advantage of the properties of a multivariate normal distribution and thus derive the conditional density of latent v_1^* given v_2. This data restructuring necessitates additional operations on the abovementioned modified weight matrices, $\mathbf{W}_d^R$, $\mathbf{W}_o^R$, and $\mathbf{W}_w^R$ so as to keep the inherent connectivity structure intact. However, this procedure is more data-specific, depending on the positions of zero observations in the data.

106

Let $[D]_i$ represent the rows of an N by N identity matrix D with the row numbers, i, corresponding to the positions of zero observations in the data that already have self-directed pairs removed. Then place the block of $[D]_i$ above the remaining rows of D, denoted as $[D]_{i-}$, to create a new matrix M. That is $M = \left([D]_i{'},\ [D]_{i-}{'}\right)'$. Pre-multiplying each of the matrices, $\mathbf{W}_d^R$, $\mathbf{W}_o^R$, and $\mathbf{W}_w^R$, by M and then post-multiplying each by the transpose of M will produce revised weight matrices that comply with the restructured data (i.e., zero values placed on top of non-zero observations). Hence, the new model is given by

$$M(Bv^*) = \rho_d M(B\mathbf{W}_d^R B')M'[M(Bv^*)] + \rho_o M(B\mathbf{W}_o^R B')M'[M(Bv^*)]$$

$$+ \rho_w M(B\mathbf{W}_w^R B')M'[M(Bv^*)] + M(BX\boldsymbol{\beta}) + M(B\varepsilon) \qquad (3.21)$$

Given that M is an orthogonal matrix, (3.21) can be simplified as:[36]

$$M(Bv^*) = \rho_d M(B\mathbf{W}_d^R B')(Bv^*) + \rho_o M(B\mathbf{W}_o^R B')(Bv^*)$$

$$+ \rho_w M(B\mathbf{W}_w^R B')(Bv^*) + M(BX\boldsymbol{\beta}) + M(B\varepsilon) \qquad (3.22)$$

3.5.3 Marginal Effects in Spatial OD Tobit Models

It is not straightforward to interpret the estimated coefficients of a Tobit model due to its inherent nonlinearity. The spatial autoregressive structure of the spatial OD threshold Tobit model further complicates the interpretation. Using marginal effects, which are partial derivatives reflecting how changes in an explanatory variable affect the expected value of y_i, is considered to be useful. As in other spatial OD models, the origin-centric ordering of the

---

[36] Since M is a square matrix with orthonormal column (and row) vectors, we know M is orthogonal and $M'M = M^{-1}M = I$.

variables in a spatial OD Tobit implies that for non-bilateral regressors (e.g., GDP in this application), marginal effects should be calculated as country-specific rather than dyad (or observation)-specific, because a change in one country's regressor immediately affects all dyads in which that country is either an origin or a destination and then the effects are propagated through the spatial spillover mechanism to other dyads.

Given that coefficient estimates of a conventional regression model are interpreted as averaging over impacts on all observations arising from changes in explanatory variables, LeSage and Thomas-Agnan (2012) propose the use of scalar summary measures which can provide interpretation of spatial autoregressive interaction models in a consistent manner. By averaging over the relevant marginal effects associated with changing a given characteristic for all regions, $i = 1, \ldots, n$, these scalar summaries allow the calculation of direct effects – i.e., origin and destination effects arising from changing a typical country's regressor on pairs involving that country, distinguished by the origin or destination status of the said country in those pairs, network (indirect) effects – i.e., the effects on pairs not involving the said country, as well as intraregional effects on self-directed pairs. Total effects are the sum of these four types of effects.

In matrix notation, the partial derivatives measuring total effects on the latent variable (represented by the flow matrix Y^*) from changing X_i^k (region $i = 1, \ldots, n$ and characteristic $k = 1, \ldots, K$) are given by

$$\text{TE} = \begin{pmatrix} \partial Y^*/\partial X_1^k \\ \partial Y^*/\partial X_2^k \\ \vdots \\ \partial Y^*/\partial X_n^k \end{pmatrix} = (\mathbf{I}_N - \rho_d \mathbf{W}_d - \rho_o \mathbf{W}_o - \rho_w \mathbf{W}_w)^{-1} \begin{pmatrix} Q d_1 \beta_d^k + Q o_1 \beta_o^k \\ Q d_2 \beta_d^k + Q o_2 \beta_o^k \\ \vdots \\ Q d_N \beta_d^k + Q o_N \beta_o^k \end{pmatrix} \qquad (3.23)$$

where Qd_i is an $n \times n$ matrix of zeros with the ith row adjusted to be a vector of ones, and Qo_i is an $n \times n$ zero matrix with the ith column replaced by a vector of ones. A scalar summary measuring the total effects of a change in the typical region's kth characteristic can be obtained by averaging across all the elements of the $N \times n$ matrix TE in (3.23) and thus takes the form:

$$te = (1/N)l'_N \cdot TE \cdot l_n.$$

A scalar summary of the destination effects can be calculated by averaging across the elements in the matrix TE that correspond to the partial derivatives for pairs in which the country with changed characteristic is the destination. Mathematically, this scalar measure can be expressed as $de = (1/N)l'_N \cdot DE \cdot l_n$, where DE is an $N \times n$ matrix that retains the $[1 + n \cdot (r - 1)]$th rows $(r = 1, \ldots, n)$ of TE matrix while having the remaining rows set to zero.

Similarly, a scalar summary of the origin effects can be constructed by averaging across the elements in the matrix TE which correspond to the partial derivatives for pairs in which the country with changed characteristic is the origin. This can be expressed as $oe = (1/N)l'_N \cdot OE \cdot l_n$, where OE is an $N \times n$ matrix that retains the $[(1 + n \cdot (r - 1)): nr, r]$ elements of TE (i.e., the rth n elements of the rth column of the matrix TE with $r = 1, \ldots, n$) while setting the remaining elements to zero.

Further, a scalar summary of the intraregional effects can be created using $ie = (1/N)l'_N \cdot IE \cdot l_n$, with IE being an $N \times n$ zero matrix adjusted to pass over, in the corresponding row and column positions, the elements of the matrix TE that represent partial derivatives for self-directed pairs.

Consequently a scalar summary for the network effects can be obtained using $ne = te - de - oe - ie$.

In order to interpret the newly proposed spatial OD threshold Tobit model, this study adopts LeSage and Thomas-Agnan's approach with a modification that sets the intraregional effects to 'zero' when internal flows are not included in the model estimation. Not surprisingly, the nonlinear nature of Tobit means that the β_d^k and β_o^k in (3.23) should be replaced by derivatives which are no longer constant scalars as would be the case of a linear regression model. Instead, they are two $n \times n$ matrices with varying elements depending on the specific values of explanatory variables observed for each OD flow. This study uses bold letters to distinguish them from the coefficient estimates. By organizing the column vector $\boldsymbol{X\beta}$ into an $n \times n$ origin-centric matrix $\boldsymbol{Z}$, we write,[37]

$$\boldsymbol{\beta_d^k} = \Phi\left(\frac{\boldsymbol{Z} - \boldsymbol{H} \cdot \tau}{\boldsymbol{U}}\right) \cdot \beta_d^k + \phi\left(\frac{\boldsymbol{H} \cdot \tau - \boldsymbol{Z}}{\boldsymbol{U}}\right) \cdot \left(-\frac{\beta_d^k}{\boldsymbol{U}}\right) \cdot \tau$$

and

$$\boldsymbol{\beta_o^k} = \Phi\left(\frac{\boldsymbol{Z} - \boldsymbol{H} \cdot \tau}{\boldsymbol{U}}\right) \cdot \beta_o^k + \phi\left(\frac{\boldsymbol{H} \cdot \tau - \boldsymbol{Z}}{\boldsymbol{U}}\right) \cdot \left(-\frac{\beta_o^k}{\boldsymbol{U}}\right) \cdot \tau \qquad (3.24)$$

where $\boldsymbol{H}$ stands for an $n \times n$ matrix of ones, $\boldsymbol{\tau}$ represents the censoring point, and $\boldsymbol{U}$ designates an $n \times n$ diagonal matrix with the diagonals set equal to the square root of the diagonals in the covariance matrix Ω defined in (3.16).

---

[37] To comply with the origin-centric ordering, the diagonal elements of the matrix $\boldsymbol{Z}$ in this application are set to zero as place holders, since we do not consider sample countries' internal trade flows. And subsequently the diagonals of $\boldsymbol{\beta_d^k}$ and $\boldsymbol{\beta_o^k}$ are replaced by zeros to reflect the exclusion of interregional flows.

3.6   Empirical Results and Interpretation

Following the algorithm described in Section 3.4.2.2, this study runs 35,000 iterations. Inspection of the trace plots for all model parameters indicates quick convergence to a steady state. Thus this research uses a burn-in period of 5,000 iterations and draws inferences based on the remaining 30,000 iterations. As with the conventional practice in Bayesian analysis, a 95% credibility interval together with posterior mean and standard deviation that are associated with each model parameter are reported in TABLE 3.1.

TABLE 3.2 displays the results of several other techniques commonly used for the estimation of the gravity equation alongside those from the spatial OD threshold Tobit.[38] Column 1 presents the OLS estimates using the logarithm of exports as the dependent variable. As noted earlier, this requires dropping all the observations of zero bilateral trade flow. Only 612 country pairs, or 61.7% of the current sample, record positive export flows. Column 2 shows the OLS estimates with $\ln(y_i + 1)$ being the dependent variable, and in Column 3, OLS results are presented using $\ln(y_i + 0.0049)$ as the dependent variable, where the added positive constant is opted in light of the threshold estimate from the threshold Tobit model (see Column 5). Column 4 exhibits results of standard Tobit and Column 5 lists threshold Tobit estimates based on Eaton and Tamura (1994). The spatial OD threshold Tobit results are presented in Column 6 in a compatible format.

The signs of all of the parameter estimates are remarkably stable across all models except for the contiguity variable, which seems not substantially different from zero in these models.

---

[38] For non-Bayesian estimations, a 95% confidence interval is reported in parentheses below each point estimate, while for the Bayesian estimation a 95% credible interval is presented.

Inspecting the first three columns, we find that different approaches to log transforming the dependent variable lead to noticeable changes in OLS estimates. As shown in Column 1, the coefficients for exporter's GDP and distance are almost equal to positive one and negative one, respectively, while the GDP coefficient for importer is also on a comparable scale. However, these results are obtained using positive export flows only. When the zero observations are included for estimation, the magnitude of conventional trade variables decreases noticeably. As illustrated in Column 3, when we set the added positive constant (i.e., fix the threshold parameter) to be 0.0049, the sizes of the two income elasticities decrease by more than half, with exporter income-elasticity declining from 0.9970 to 0.4404 and importer income-elasticity decreasing from 0.8813 to 0.4013. As for the distance parameter, its magnitude changes from -1.0081 to -0.2006. Further, when an arbitrary constant of "1" is added to the export flow data before log transforming them as in some previous trade studies, the sizes of all parameters except for contiguity decrease to about one twentieth of those estimated only with positive observations.[39] It is interesting to note that once a threshold parameter is included, the sign of the coefficient on contiguity changes from negative to positive as shown in Columns 5 and 6 as compared to the others, which is consistent with what trade theory on bilateral trade costs would predict, though this coefficient seems not significantly different from zero in almost all the models. Also, it is quite consistent across these models that the magnitude of coefficient on exporter's income turns out to be larger than that of importer's, though this coefficient estimate itself is not directly comparable across the models, which we will discuss further a bit later.

---

[39] Although in this case, the value "1" is quite large given that export flows are measured in billions of US dollars, this exercise illustrates that the choice of the positive constant to be added to trade data in order to make use of the log-linearized gravity equation does affect the estimation results and therefore should not be made on an ad hoc basis. For instance, Behrens et al. (2012) augment zero trade flows by adding 1, which might have exerted an unduly impact on their estimates given that the Canada-US exports dataset is measured in million US dollars for the year 1993.

Estimates from the spatial OD threshold Tobit suggest that bilateral trade flows are indeed correlated in space and the interdependence among export flows arises from multiple sources. Specifically, none of the 95% credible intervals for the spatial coefficients contains the value zero, with ρ_d and ρ_o both showing a positive sign whereas ρ_w turning out to be negative. The positive sign of ρ_d suggests that when a country exports to another country, it is likely to export to the neighbors of its destination as well. This spillover effect may be partly due to potential economies of scale. Exporting to countries that are clustered geographically allows an exporter to take advantage of the established trade route and existing infrastructures geared to export activities. On the demand side, countries that are located in close geographic proximity, especially those of a smaller size, are predisposed to possess a similar endowment of resources, which may lead them to import the same types of goods. A positive ρ_o signals that one country's exports to a given destination tend to be positively related to the trade volumes from its neighboring countries to the same importer, alluding to a different type of spillover effect. This resemblance in trading behavior among exporters who are geographically proximate may be attributable to the easiness of the dissemination of technologies and innovations, relocation of skilled labor, and even policy imitation to occur among neighboring countries. In this sense, proximity appears to provide trade-promoting opportunities rather than create market competition. Moreover, similarities in resource endowment among neighboring exporting countries may lead them to specialize in the production of same or similar types of goods.

On the other hand, the negative sign of ρ_w indicates a competitive relationship across trading pairs when a 'dual' neighboring relationship exists both at the origins and the destinations. When both exporter countries are eyeing the same export markets while both importer countries are looking to the same suppliers, stronger trade ties within one pair may

cause concerns on the part of the other exporter (or importer) about disadvantaged trade position with the importer (or exporter) in the said pair. In the context of concurrent neighboring relationships across two pairs between both exporters and importers, this competition is likely to induce a negative impact on the trade flows within the other pair of traders.

The estimate of the threshold parameter is around 0.0049 and zero falls outside the 95% credible intervals. This implies that on average, the potential trade volumes need to be at least 4.9 million for an exporter country to be willing (i.e., for it to be profitable) to carry out trade transactions.

It appears that geographical distance negatively affects trade volumes. Its coefficient estimate from the spatial OD threshold Tobit is -0.1842. However, this estimate is quite different from the ones obtained under the standard OLS and Tobit models (columns 1 and 4), which are very close to unity, being -1.0081 and -1.0079, respectively. It is also slightly smaller than the distance coefficient produced by Eaton and Tamura's threshold Tobit. This is consistent with LeSage and Thomas-Agnan (2012)'s observation that diminished importance of distance after accounting for spatial dependence "often occurs for the spatial variants of gravity models" (p. 23). In a similar vein, Fotheringham and Webber (1980) note that in the presence of spatial autocorrelation, the estimated parameter on the distance variable captures both "a 'true' friction of distance effect" and a measure of the map pattern (p. 34). Joining their insight, Porojan (2001, p. 275) further explicates that the spatial lag in his model captures an important part of the spatial effect, which the traditional formulation of the gravity model partially picked up through the distance variable. Since spatial origin-destination modeling is better tailored to flow data in capturing spatial effects, it is not unexpected that the estimated impact of distance weakens.

114

Contiguity shows no discernible impact on export flows. Although the respective coefficient takes a positive sign for both the spatial and non-spatial threshold Tobits, zero falls near the center of the intervals for this coefficient in all but one models. This is consistent with Ranjan and Tobias (2007)'s finding. While the authors do not offer a formal explanation for the insignificance of the contiguity effect, they draw attention to the difference in their new model specification which accounts for the discrete-continuous nature of bilateral trade data (p. 830). More importantly, in competition with spatial terms built on contiguity relationship, a bilateral contiguity dummy may prove inadequate in distinguishing the involved effects of contiguity on trade flows.

Whether the spatial OD threshold Tobit model is estimated with or without the observations on internal trade flows (i.e., columns 6 and 7), the estimation results are quite consistent in terms of the sign and significance of coefficients. It should be noted that according to the design of spatial weight matrix $\mathbf{W}$, all diagonal elements (i.e., weights assigned to intra-regional flows or in this application, internal flows) are set to zero. This structure is transferred to $\mathbf{W}_d$, $\mathbf{W}_o$ and $\mathbf{W}_w$, leading to different treatment of neighboring relationships with self-directed pairs and non-self-directed pairs. Thus it is not surprising that the inclusion of own trade flows may attenuate spatial effects. Nonetheless, positive exporter- and importer-based dependence still emerge for this extended data, though the magnitude of the coefficient estimates turns out to be smaller. Besides, the coefficient size on *distance* increases as expected. Compared to bilateral trade flows, internal flows record much larger trade volumes but occur within relatively shorter distances, tilting towards a resisting effect of distance on trade. Moreover, both GDP coefficients are estimated to be larger with the addition of internal flow data, which is unsurprising given that internal flows account for a much higher proportion of GDP.

115

As LeSage and Thomas-Agnan (2012) correctly point out, estimates for non-bilateral variables in spatial OD models (i.e., exporter-GDP and importer-GDP in the current application) are not directly comparable to those from OLS. Hence, it is more appropriate to calculate scalar summary effect estimates that reflect marginal effects associated with changes in regional characteristics on average flows so as to provide interpretation in a fashion consistent with that of conventional linear regression models. TABLE 3.3 shows summary effect estimates of GDP for the spatial OD threshold Tobit model. The first column displays the (averaged) marginal effects on the latent y_i^*, while the second column presents the (averaged) marginal effects on y_i. The scalar summary estimates for the latent and observed regressands turn out to be similar. As far as the effect estimates for y_i are concerned, a one percent increase in GDP of the typical exporter (i.e., origin) country is likely to lead to a 0.5941 percent increase in export flows, while a one percent increase in GDP of the typical importer (i.e., destination) country is likely to lead to a 0.3351 percent increase in export flows. These results imply that the impact of exporter's GDP is larger than that of importer's GDP. The network effects are also estimated to be positive, suggesting that a one percent increase in the GDP of the typical country is likely to lead to a 0.2190 percent increase in export flows due to spatial spillover (separately from the flows that are already captured by the origin and destination effects). In this example, total effects reflect the sum of the origin, destination and network effects on trade flows. The total effects of increasing GDP by one percent are a 1.1482 increase in export flows.

As an exploratory step, TABLE 3.4 compares the marginal effects of GDP for the latent variable y^* in the spatial Tobit as well as for the three OLS models considered in this study. This comparison purports to address the question of whether the different coefficients in these models

116

are compensated by the different a's to produce similar marginal effects. For illustration, we focus on the application that considers bilateral flows only (i.e., columns 1-6),

For models with no spatial correlation (i.e., the first three columns), the origin effects and destination effects are the same as the coefficient estimates for exporter-GDP and importer-GDP and the network effects are all zeros. As shown in TABLE 3.4, the marginal effects from these different models are very dissimilar, though all four types of summary measures consistently identify positive impacts of GDP on export flows across these models as expected. This implies that the different impacts estimated of GDP by the spatial OD threshold model are not statistical artifacts emerging from the choice of a. By allowing for spatial correlation, the spatial OD threshold Tobit detects sizeable network effects, which are estimated to be 0.1993 when the effects on the latent variable are examined. The asymmetric income impacts between exporter and importer are more distinct under the spatial model, with the origin effects (0.5180) standing about 1.8 times the magnitude of the destination effects (0.2877). Nonetheless, both the origin and destination effects are quite smaller than unity, revealing a much reduced influence of GDPs on export volumes once spatial dependence is appropriately controlled for. This result should not be surprising. In fact, several previous studies discuss this issue. For instance, Grossman (1998) questions the unrealistic large magnitude of coefficients on GDPs. Porojan (2001) reports considerable changes in the size of estimated parameters when applying alternative spatial econometric models to both import and export data, though he does not calculate the marginal effects, which should provide the more appropriate interpretation given the nonlinearity introduced by spatial correlation. Moreover, applied trade economists have always been aware of potential omitted variables in the empirical specification of the gravity model. If there are variables left out of the model which are correlated with the GDP measures and positively affect

117

export activities, estimates of GDP coefficients will be inflated. In fact, several studies have recommended the use of spatial lags as a more efficient way to deal with the issue of omitted variables (LeSage and Pace, 2008; Behrens et al., 2012; Porojan, 2001). For instance, Behrens et al. (2012) argue that the use of lagged terms is "more robust to potential misspecification concerning the form of interdependence" (p. 775).

3.7 The Issue of Model Fit

As the econometric models shown in TABLE 3.2 have different assumptions for the underlying data distribution and are estimated using quite different techniques, it is not straightforward to make model comparison. Thus, this study employs several measures of model fit as exploratory tools, which would be helpful in evaluating how well the model represents the data. As noted, both the standard OLS and Tobit regression models drop observations of zero flows in the estimation due to the log transformation of the dependent variable. For comparison purposes, we focus on the spatial OD threshold Tobit and Eaton and Tamura's threshold Tobit, the two models that fully utilize the sample data.

As an equivalent to R^2, several pseudo R^2 statistics have been developed for non OLS regression models. Efron's pseudo R^2 is an extension to the "percent variance explained" interpretation of R^2 in linear regressions. This measure was directed at binary-outcome models and writes as:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{\pi}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{3.25}$$

where $\hat{\pi}_i$ denotes model predicted probabilities. However, it could be used for continuous models and sometimes is called the sum of squares R^2. It should be noted that neither the spatial

118

OD threshold Tobit nor the threshold Tobit is estimated by minimizing variance. We employ the Efron's pseudo R^2 as an exploratory tool.

Considering the right skewedness of our sample trade data as illustrated below in FIGURE 3.1, the mean value (0.2599) does not reasonably represent the average export flows among the sample countries. Thus, we replace the sample mean with the median (0.001) in the pseudo R^2 statistic in (3.25). Based on this modified measure, the pseudo R^2 is 0.139 for the spatial threshold Tobit and 0.127 for the non-spatial threshold Tobit.

A more commonly used pseudo R^2 statistic is McFadden's likelihood-ratio index,

$$R^2 = 1 - \frac{\mathcal{L}(M_{\text{Full}})}{\mathcal{L}(M_{\text{Intcpt}})} .$$
(3.26)

According to the log likelihood function for the Tobit model (Wooldridge, 2002), the log likelihood value for the spatial OD threshold Tobit is -1646.413 and that for the non-spatial threshold Tobit is -1735.047, while the log likelihood for an intercept only model is -1775.2. Thus the McFadden pseudo R^2 for the threshold Tobit with and without spatial correlation are 0.073 and 0.023, respectively.

With respect to the pseudo R^2 measures reported above, the spatial threshold Tobit model appears to provide a moderately better fit than the non-spatial threshold Tobit. The relatively small pseudo R^2 values may be due to the estimation difficulty in accommodating simultaneously both the substantial numbers of zero values and the highly right skewedness of the data.
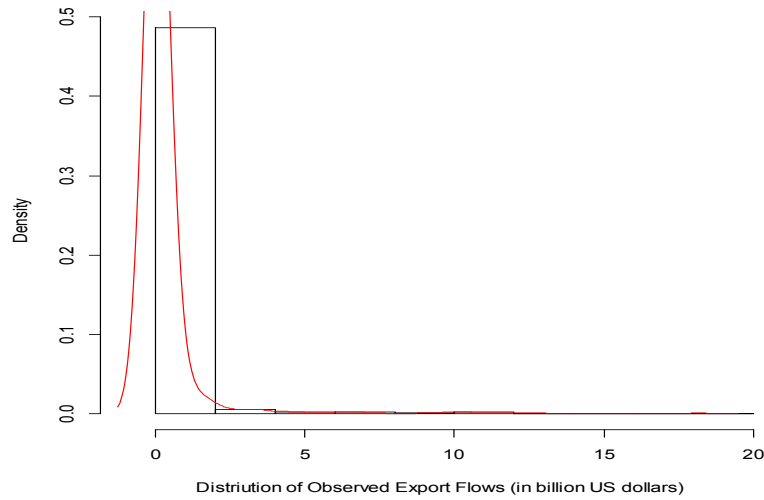
FIGURE 3.1 Distribution of the Observed Values of Export Flows

In addition to comparing different models , it is also useful to check how a specific model performs against the data. As the spatial OD threshold Tobit is estimated using a Bayesian procedure, we also implement Bayesian posterior predictive checks for the model.

For the 380 observations of zero flows, the spatial OD threshold Tobit correctly predicts 194 cases, a classification rate of 51.05%. As for the other 612 positive observations, the spatial model tends to have a better predictive capability for comparatively small trade values than for large ones. When the six observations that have a trade volume greater than 10 (in billion US dollars) are excluded from the sample data, the median of the remaining data is only 0.0267 (in billon US dollars). For instance, based on 3,000 simulations, Figure 3.2 shows the posterior predictive frequency for observation 465, $y_i = 0.0265$. It records the export flows from China to Nepal, and this trade volume is close to the median value after the exclusion of those particularly large observations. The red vertical line in the plot represents the observed data for this trade pair. The data appears quite plausible under the model.

120
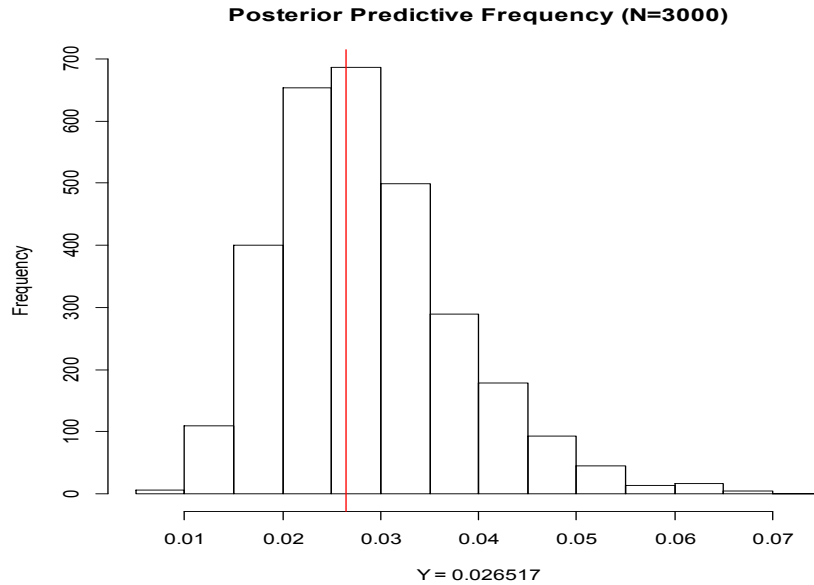
FIGURE 3.2 Posterior Predictive Frequency for $y_i = 0.0265$

However, the posterior predicted value for large observations tends to be too small. For instance, FIGURE 3.3 shows that the posterior predictions of observation 590 cluster around 0.5227, which is quite far from the actual data $y_i = 1.0282$ (indicated by the red vertical line). The discrepancy between the predicted value and the observed value gets larger when we move further away from the median of the data. In 256 cases of the 612 positive observations, the observed value is within 1.95 standard deviations of the predicted value.

Posterior Predictive Frequency (N=3000)

$Y = 1.028$
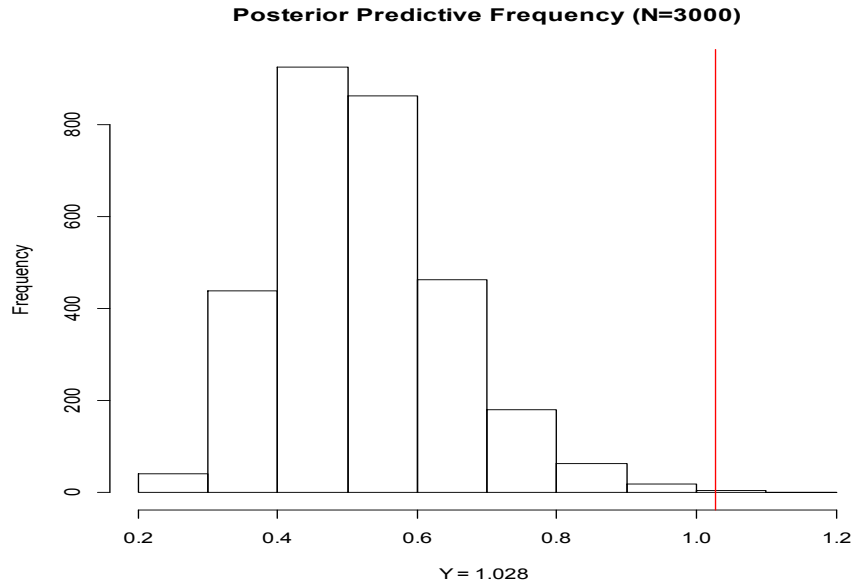
FIGURE 3.3 Posterior Predictive Frequency for $y_i = 1.0282$

3.8   Conclusion

Since the introduction of the gravity model by Tinbergen (1962) and Linnemann (1966), trade scholars have sought to advance the model in two aspects. On the one hand, trade economists aim to establish a close link between the gravity model and trade theories. On the other hand, empirical researchers deal with several important econometric issues, such as the handling of zero trade flows and heteroskedasticity. However, applied trade economists rarely tap the question of correlation among trade flows, especially in the context of geographic location, although spatial dependence has been more heavily studied in some other related areas such as growth and FDI. So far, only a very few empirical studies have attempted to explicitly incorporate the interdependence of bilateral trade flows into the gravity model. Yet, the methods they employ to model the dependence structure overlook the fact that bilateral trade data are dyadic in nature and characterize a directional flow from the exporter to the importer. In this

122

context, previous methods fail to adequately account for the interdependence between trade flows.[40]

Recent development of spatial econometric modeling provides a useful tool for researchers to analyze spatial dependence in data featuring a directional flow, such as population migration and bilateral trade flows. This study extends the spatial origin-destination modeling set forth by LeSage and Pace (2008) and develops a specification of the gravity model that can both effectively handle the multiple forms of spatial dependence existing among bilateral trade flows as well as the zero trade flow problem whose presence defies the log-linear formulation of the gravity model. Since the proposed model introduces spatial connectivity structures that render an ML estimator infeasible, this study relies on a Bayesian approach for model estimation.

When the multiple sources of spatial dependence in bilateral trade data are explicitly taken into account, the magnitude of the effect estimated for conventional trade variables changes substantially. To be specific, the traditional specification of the gravity model overestimates the income elasticity of both home country (i.e., exporter country in the case studied here) and host country (i.e., importer country in this study), as well as the trade friction effect of distance. Contiguity exerts no appreciable impact on bilateral trade, especially when measures of spatial correlation built upon contiguity are included in the model. Besides, sizable network effects are identified for GDP, revealing the additional channels through which economic size affects trade flows.

---

[40] As noted earlier, Lebreton and Roi (2009)'s online manuscript is an exception. However, as a straight application of LeSage and Pace's (2008) spatial origin-destination modeling technique, this paper does not tackle zero trade flow problem.

The empirical results of this study point to the presence of all three types of spatial dependence which LeSage and Pace (2008) argue as plausible in data characterized by a bilateral directional flow. The exporter (origin)-based dependence shows a positive effect on bilateral trade flows, suggesting that exporters who are located close to one another geographically tend to export to the same markets. This positive correlation among trade flows which arises from the end of exporters may find its explanation in a spatial spillover of technologies, ideas and polices among neighboring exporting countries. Or similar resource endowments may lead nearby countries to specialize in the production of same exported goods. Positive importer (destination)-based dependence implies that an exporter's trade activities with a given importer are likely to bring about similar export flows to the neighbors of the importer. This spatial effect may be attributed to economies of scale and similarities in import demands. The third type of spatial dependence appears among trading pairs that are featured by a dual neighboring relationship both between exporters and importers. The negative effect of this exporter (origin)-to-importer (destination) dependence reflects a competitive link among such pairs of traders.

Relying on spatial connectivity structures that manifest multidirectional spatial dependence in bilateral trade flows, the proposed spatial OD threshold Tobit allows the researcher to evaluate in a more balanced fashion the impact of geographic distance on trade flows and alleviates the estimation bias in the covariates of the gravity model due to spatial correlation as well as possible omitted variables.

The contribution of this study is twofold. Methodologically, it advances an econometric model by considering the complexity of spatial autocorrelation embedded in "directional" trade flows while dealing with the corner solution where trade volumes are recorded as zero.

124

Empirically, it provides evidence that bilateral trade flows are indeed correlated in space and that conventional trade variables have much lesser impacts than previously reported, which actually work through multiple channels due to the multifaceted spatial dependence of trade flows. On the other hand, in fitting the sample data that contain a sizable amount of zeros as well as some particularly large values, the spatial OD threshold Tobit model performs slightly better than the non-spatial threshold Tobit model. Future research should try to improve the model by accounting for this problem.

Although it is sensible to specify weight matrices based on geographic distance when examining the spatial correlation of bilateral trade flows, there may be other ways than contiguity for defining neighboring relations and thereby the weight matrix given the researcher's knowledge of, or theory about, the diffusiveness of spatial interaction (e.g., LeSage and Pace, 2004). And of course, different designs of the weight matrix may affect estimation results differently.

References

Anderson, J. E. 2011. "The Gravity Model," online manuscript.

Anderson, J.E. and E. van Wincoop. 2004. "Trade Costs," *Journal of Economic Literature*, 42, 691-751.

Anderson, J.E. and van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, 170-192.

Anselin, L. 1988. *Spatial Econometrics: Methods and Model*s. Boston: Kluwer Academic.

Baier, S.L. and J.H. Bergstrand. 2007. "Do Free Trade Agreements Actually Increase Members' International Trade?" *Journal of International Economics*, 71, 72-95.

Baier, S.L. and J.H. Bergstrand. 2009. "*Bonus vetus* OLS: A Simple Method for Approximating International Trade-Cost Effects Using the Gravity Equation," *Journal of International Economics*, 77, 77-85.

Behrens, K., C. Ertur, and W. Koch. 2012. "'Dual' Gravity: Using Spatial Econometrics to Control for Multilateral Resistance," *Journal of Applied Econometrics*, 27(5), 773-794.

Bergstrand, J. H. 1989. "The Generalized Gravity Equation, Monopolistic Competition, and the Factor–Proportions Theory in International Trade", *Review of Economics and Statistics*, 71(1), 143-153.

Bergstrand, J.H. 1985. "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence", *Review of Economics and Statistics*, 67(3), 474-481.

Bergstrand, J.H. 1998. "Comment," in *The Regionalization of the World Economy*, J.A. Frankel, (ed.), 23-28.

Blonigen, B.A., R.B. Davies, G.R. Waddell, and H.T. Naughton. 2007. "FDI in Space: Spatial Autoregressive Relationships in Foreign Direct Investment," *European Economic Review*, 51, p. 1303-1325.

126

Blum, B.S. and A. Goldfarb. 2006. "Does the Internet Defy the Law of Gravity?" *Journal of International Economics*, 70(2), 384-405.

Burger, M., F. van Oort and G. Linders. 2009. "On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation," *Spatial Economic Analysis*, 4(2), 167-190.

Conley, T.G. and E. Ligon. 2002. "Economic Distance and Cross-Country Spillovers," *Journal of Economic Growth*, 7, 157-187.

Eaton, J. and A. Tamura. 1994. "Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns," *Journal of the Japanese and International Economies*, 8, 478-510.

Feenstra, R.C, J.R. Markusen and A.K. Rose. 2001. "Using the Gravity Equation to Differentiate among Alternative Theories of Trade," *Canadian Journal of Economics*, 34(2), 430-447.

Feenstra, R.C. 2004. *Advanced International Trade: Theory and Evidence*. Princeton: Princeton University Press.

Felbermayr, G.J. and W. Kohler. 2006. "Exploring the Intensive and Extensive Margins of World Trade," *Review of World Economics*, 142(4), 642-674.

Fotheringham, A.S. and M.J. Webber. 1980. "Spatial Structure and the Parameters of Spatial Interaction Models," *Geographical Analysis*, 12(1), 33-46.

Frankel, J.A. and S.-J. Wei. 1998. "Regionalization of World Trade and Currencies: Economics and Politics," In *The Regionalization of the World Economy*, J.A. Frankel (ed.), London: Chicago University Press, 189-226.

Grossman, G.M. 1998. "Comment," In *The Regionalization of the World Economy*, J.A. Frankel (ed.), London: Chicago University Press, 29-31.

Hamilton, C.B. and A.L. Winters. 1992. "Opening up International Trade with Eastern Europe," *Economic Policy*, 14, 77-116.

127

Haveman J. and D. Hummels. 2004. "Alternative Hypotheses and the Volume of Trade: The Gravity Equation and the Extent of Specialization," *Canadian Journal of Economics*, 37(1), 199-218.

Harville, D.A. 1997. *Matrix Algebra From a Statistician's Perspective*. New York: Springer-Verlag.

Helpman, E. and P. Krugman. 1985. *Market Structure and Foreign Trade*. Cambridge: MIT Press.

Helpman, E., M. Melitz and Y. Rubinstein. 2008. "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123(2), 441-487.

Helpman. E. 1987. "Imperfect Competition and International Trade: Evidence from Fourteen Industrial countries," *Journal of the Japanese and International Economies*, 1, 62-81.

Lebreton, M. and L. Roi. 2009. "A Spatial Interaction Model with Spatial Dependence for Trade Flows in Oceania: A Preliminary Analysis," online manuscript.

LeSage, J.P., M.M. Fischer and T. Scherngell. 2007. "Knowledge Spillovers across Europe: Evidence from a Poisson Spatial Interaction Model with Spatial Effects," *Papers in Regional Science*, 86(3), 393-421.

LeSage, J.P. and R.K. Pace. 2004. "Introduction," *Advances in Econometrics: Vol. 18: Spatial and Spatiotemporal Econometrics*. Oxford: Elsevier Ltd, 1-32.

LeSage, J.P. and R.K. Pace. 2008. "Spatial Econometric Modeling of Origin-Destination Flows," *Journal of Regional Science*, 48(5), 941-967.

LeSage, J.P. and R.K. Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton: Chapman & Hall/CRC.

LeSage, J.P. and R.K. Pace. 2012. "Interpreting Spatial Econometric Models," Department of Finance and Economics, Texas State University – San Marcos, mimeo.

LeSage, J.P. and C. Thomas-Agnan. 2012. "Interpreting Spatial Econometric Origin-Destination Flow Models," Department of Finance and Economics, Texas State University – San Marcos, mimeo.

128

Linders, G.M. and H.L.F. de Groot. "Estimation of the Gravity Equation in the Presence of Zero Flows," online manuscript.

Martin, W. and C.S. Pham. 2008. "Estimating the Gravity Model When Zero Trade Flows Are Frequent," online manuscript.

Martinez-Zarzoso, I. 2013. "The Log of Gravity revisited." *Applied Economics*, 45(3), 311-327.

Mayer, T. and S. Zignago. 2011. "Notes on CEPII's Distance Measures: The *GeoDist* Database," CEPII Working Paper 201-25.

Melitz, M.J. 2003. "The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695-1725.

Pace, R.K. and J.P. LeSage. 2007. "Spatial Econometric Models, Prediction," in *Encyclopedia of Geographical Information Science*, S. Shekhar and H. Xiong (eds.). New York: Springer-Verlag

Polak, J.J. 1996. "Is APEC a Natural Regional Trading Bloc? A Critique of the 'Gravity Model' of International Trade," *World Economy*, 19(5), 533-543.

Porojan, A. 2001. "Trade Flows and Spatial Effects: The Gravity Model Revisited*," Open Economies Review*, 12, 265-280.

Ranjan, P. and J.L. Tobias. 2007. "Bayesian Inference for the Gravity Model," *Journal of Applied Econometrics*, 22, 817-838.

Rauch, J.E. and V. Trindade. 2002. "Ethnic Chinese Networks in International Trade," *Review of Economics and Statistics*, 84(1), 116-130.

Rose, A.K. 2004. "Do We Really Know That the WTO Increases Trade?" *American Economic Review*, 94(1), 98-114.

Santos Silva, J.M.C. and S. Tenreyro. 2006. "The Log of Gravity," *Review of Economics and Statistics*, 88(4), 641-658.

Smarzynska, B.K. 2001. "Does Relative Location Matter for Bilateral Trade Flows? An Extension of the Gravity Model," *Journal of Economic Integration*, 16(3), 379-398.

Tinbergen, J. 1962. *Shaping the World Economy*. New York: The Twentieth Century Fund.

Weinhold, D. 2002. "The Importance of Trade and Geography in the Pattern of Spatial Dependence of Growth Rates," *Review of Development Economics*, 6(3), 369-382.

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

Appendix V

List of Sample Countries (Abbreviation in Parentheses)

Bahrain (BAH)	Jordan (JOR)	Russian Federation (RUS)
Bangladesh (BNG)	Korea, Rep. (ROK)	Saudi Arabia (SAU)
Bhutan (BHU)	Lao PDR (LAO)	Singapore (SIN)
Brunei (BRU)	Lebanon (LEB)	Sri Lanka (SRI)
Cambodia (CAM)	Malaysia (MAL)	Syrian Arab Rep. (SYR)
China (CHN)	Maldives (MAD)	Thailand (THI)
India (IND)	Mongolia (MON)	Turkey (TUR)
Indonesia (INS)	Nepal (NEP)	United Arab Emirates (UAE)
Iran (IRN)	Oman (OMA)	Vietnam (DRV)
Israel (ISR)	Pakistan (PAK)	Yemen (YEM)
Japan (JPN)	Philippines (PHI)	

Appendix VI

First-order Contiguity Matrix **W**

	RUS	IRN	TUR	SYR	LEB	JOR	ISR	SAU	YEM	BAH	UAE	OMA	CHN	MON	ROK	JPN	IND	BHU	PAK	BNG	SRI	MAD	NEP	THI	CAM	LAO	DRV	MAL	SIN	BRU	PHI	INS	SUM
RUS	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
IRN	1	0	1	0	0	0	0	1	0	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	7
TUR	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
SYR	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
LEB	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
JOR	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
ISR	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
SAU	0	1	0	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
YEM	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
BAH	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
UAE	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
OMA	0	1	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
CHN	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0	0	1	0	0	1	1	0	0	0	0	0	9
MON	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
ROK	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
JPN	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
IND	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	8
BHU	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
PAK	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
BNG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
SRI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MAD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
NEP	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
THI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	4
CAM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	3
LAO	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	4
DRV	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	3
MAL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	5
SIN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2
BRU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
PHI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	2
INS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	0	1	0	5

132

TABLE 3.1 **Bayesian Estimates of the Spatial OD Threshold Tobit Model of Export Flows, Asia, 1990**

	Mean	S.D.	2.5%	Median	97.5%	Sample
intcpt.	-6.7752	0.6666	-8.0859	-6.7740	-5.4725	30000
Log exporter's GDP	0.3286	0.0270	0.2767	0.3285	0.3816	30000
Log importer's GDP	0.2877	0.0268	0.2359	0.2876	0.3410	30000
Log distance	-0.1842	0.0660	-0.3125	-0.1839	-0.0558	30000
Contiguity	0.0331	0.1876	-0.3324	0.0333	0.4005	30000
ρ_d	0.3498	0.0299	0.2905	0.3499	0.4078	30000
ρ_o	0.3418	0.0323	0.2793	0.3420	0.4069	30000
ρ_w	-0.1473	0.0371	-0.2221	-0.1476	-0.0731	30000
α	0.0049	0.0001	0.0047	0.0049	0.0050	30000

133

TABLE 3.2   Regression Estimates of the Traditional Gravity Equation

Estimator Dep. Var.	OLS $\ln(y_i)$	OLS $\ln(y_i + 1)$	OLS $\ln(y_i + 0.0049)$	Tobit $\ln(y_i)$	Threshold Tobit $\ln(y_i + a)$	Spatial OD Threshold Tobit[a] $\ln(y_i + a)$	Spatial OD Threshold Tobit[b] $\ln(y_i + a)$
intcpt.	-15.8456 * (-18.2169, -13.4744)	-0.6689 * (-0.9225, -0.4153)	-10.7401* (-11.8838, -9.5964)	-15.8567 * (-18.2218, -13.4916)	-15.6865 * (-17.4427, -13.9303)	-6.7752 * (-8.10859, -5.4725)	-8.8987 * (-10.4636, -7.4026)
Log exp-GDP	0.9970 * (0.9084, 1.0856)	0.0559 * (0.0477, 0.0641)	0.4404 * (0.4030, 0.4778)	0.9979 * (0.9095, 1.0863)	0.6473 * (0.5885, 0.7061)	0.3286 * (0.2767, 0.3816)	0.4920 * (0.4278, 0.5586)
Log imp-GDP	0.8813 * (0.7956, 0.9670)	0.0594 * (0.0512, 0.0676)	0.4013 * (0.3639, 0.4387)	0.8812 * (0.7957, 0.9667)	0.5740 * (0.5187, 0.6293)	0.2877 * (0.2359, 0.3410)	0.4638 * (0.4017, 0.5310)
Log distance	-1.0081 * (-1.2925, -0.7237)	-0.0442 * (-0.0740, -0.0144)	-0.2006 * (-0.3351, -0.0661)	-1.0079 * (-1.2917, -0.7241)	-0.1907 * (-0.3536, -0.0278)	-0.1842 * (-0.3125, -0.0558)	-0.9395 * (-1.0535, -0.8297)
Contiguity	-0.0982 (-0.8308, 0.6344)	-0.0931 * (-0.1786, -0.0076)	-0.0479 (-0.4334, 0.3376)	-0.0967 (-0.8274, 0.6340)	0.1517 (-0.3069, 0.6103)	0.0331 (-0.3324, 0.4005)	-0.4162 (-0.8763, 0.0427)
ρ_d						0.3498 * (0.2905, 0.4078)	0.1514 * (0.0917, 0.2132)
ρ_o						0.3418 * (0.2793, 0.4069)	0.1131 * (0.0472, 0.1748)
ρ_w						-0.1473 * (-0.2221, -0.0731)	0.0512 (-0.0209, 0.1153)
α					0.0049 * (0.0033, 0.0065)	0.0049 * (0.0047, 0.0050)	0.0049 * (0.0048, 0.0050)

95% confidence intervals for OLS and ML-based estimates, and 95% credible intervals for Bayesian estimates (sample size of 30,000).

* denotes zero not in interval.

[a] considers bilateral trade flows only ($N = 992$), i.e., excluding self-directed pairs .

[b] includes both bilateral and internal trade flows ($N = 1024$).

TABLE 3.3   Marginal Estimates of GDP for the Spatial OD Threshold Tobit

	Marginal Effects on y^*	Marginal Effects on y
Origin Effects	0.5180	0.5941
Destination Effects	0.2877	0.3351
Network Effects	0.1993	0.2190
Total Effects	1.0050	1.1482

TABLE 3.4   Effect Estimates of GDP for the OLS Models and for the Latent Variable in the Spatial OD Threshold Tobit

Estimator	OLS	OLS	OLS	Spatial OD Threshold Tobit
Dep. Var.	$\ln(y_i)$	$\ln(y_i + 1)$	$\ln(y_i + 0.0049)$	$\ln(y_i^* + a)$
Origin Effects	0.9970	0.0559	0.4404	0.5180
Destination Effects	0.8813	0.0594	0.4013	0.2877
Network Effects	0	0	0	0.1993
Total Effects	1.8783	0.1153	0.8417	1.0050

Chapter 4

CONCLUSION

The potential problem of inconsistent estimates due to spatial correlation has led researchers to design spatial modeling for improved estimation. This dissertation further considers modeling spatial correlation in flow data that are recorded as binary or censored values. Each observation of flow data, by definition, involves an origin and a destination at the same time, so spatial dependence is expected to arise in flow data in a more complicated manner. Accordingly, three spatial lag terms are constructed to specifically capture spatial correlation between observations on OD flows induced by a neighboring relationship between origins, between destinations, as well as a dual neighboring relationship both at the origin and the destination. This approach is similar to the spatial OD modeling suggested by LeSage and Pace (2008), and the three spatial lags are incorporated into regression models with binary and censored dependent variables, respectively. However, the non-linearity of the limited dependent variable models in the presence of spatial lags makes an ML estimator inconsistent. To circumvent the inconsistent estimation, this study develops Bayesian estimation procedures for the newly proposed spatial models.

In Chapter 2, a spatial OD probit model is proposed to tackle with spatial dependence embedded in binary flow data. By incorporating the spatial lags into the

137

latent regression model and taking advantage of the data augmentation approach of Bayesian analysis, I develop an estimation strategy to avoid the inconsistency problem associated with MLE in this context. To illustrate, the spatial OD probit is applied to a cross-sectional data on militarized interstate dispute initiations among European countries. The empirical results indicate that spatial correlation exists between conflict initiations, and that this correlation is more complex than that discussed in the existing literature. This chapter finds evidence for two types of spatial correlation: target-based and initiator-to-target based. The positive target-centric correlation suggests that initiators tend to attack the neighbors of their intended targets as well, possibly reflecting strategic or logistic needs. And the negative coefficient on the initiator-to-target based correlation signifies that a potential initiator tends to be discouraged from taking actions against its intended target if there is a conflict between the neighbor of the initiator and the neighbor of the target. This negative association between conflict initiation behaviors may have arisen from initiators' desire for maintaining domestic stability or by a war-weariness effect. Moreover, compared to base models that do not control for spatial correlation, the estimated effects of explanatory variables in the spatial OD probit model change noticeably. For example, after considering spatial correlation, the national capabilities variable is found to have a larger impact on conflict initiation. Moreover, the spatial model is capable of detecting the spillover effects of regressors on conflict. The negative spillovers associated with national capabilities imply that two countries within a dyad are less prone to militarized dispute when both perceive an external threat due to the power increase of a third country. Since the proposed spatial OD probit model accounts for

138

spatial correlation among directional flows measured as binary outcomes, it is instrumental in producing more reliable estimates of conflict-inducing factors as well as a better understanding of the dynamics of interstate conflict behavior.

In Chapter 3, I design a spatial OD threshold Tobit as a way to model flow data that are censored. As is well-known, international trade data are often left-censored at zero due to the lack of trade activities and they are likely to exhibit spatial correlation probably as a result of technology diffusion and labor mobility. Built off of Eaton and Tamura's (1994) threshold gravity model and LeSage and Pace's spatial OD modeling technique, this chapter purports to address in tandem the zero problem that challenges the log-linear formulation of the gravity model and the multiple sources of spatial correlation in trade data. The use of a threshold parameter as well as assigning spatial processes to the latent trade variable make it possible to incorporate the spatial structures in the commonly employed log-linear gravity model, while allowing for the inclusion of zero-valued observations in model estimation. In this context, the three spatial lags are meant to reflect exporter-centric, importer-centric, and exporter-to-importer based dependence among trade flows. Using data on export flows among European countries in 1990, the spatial OD threshold Tobit model indicates the presence of all three types of spatial dependence. The positive coefficient on the exporter-based dependence implies that exporters who are located in geographic proximity tend to export to the same markets. This positive correlation among export flows may be explained by a spatial spillover of technologies, ideas, and policy orientations among neighboring exporting countries. On the other hand, positive importer-based dependence denotes that trade activities with an

importer are likely to promote similar export flows to the neighbors of the importer due to economies of scale in structuring transport as well as similarities in import demands. Furthermore, the exporter-to-importer correlation is negative, reflecting a competitive link among trading pairs that feature a dual neighboring relationship both between exporters and importers. When taking into account the multiple types of spatial dependence in bilateral trade flows, the spatial OD threshold Tobit estimates moderated income elasticity of both exporter country and importer country. Distance still appears to exert a trade friction effect, and the contiguity variable no longer shows any impact on export flows once spatial dependence is controlled for.

In short, this dissertation has advanced spatial OD modeling for two types of limited dependent variables: binary and censored. Since the spatial connectivity structures included in the newly proposed models consider the complexity of spatial correlation among observations on "directional" flows, these models are not only instrumental in revealing how various types of flow data interact in space, but also more effective in alleviating the estimation bias in the covariates of spatial interaction models due to spatial correlation as well as possible omitted variables.

VITA

Shali Luo received a B.A. in English and an M.A. in American Studies from Foreign Affairs College (now China Foreign Affairs University) in 1995 and 2002, an M.A. in Political Science and an M.A. in Statistics from the University of Missouri-Columbia in 2005 and 2010, respectively, and a Ph.D. in Economics from the University of Missouri-Columbia in 2012. She specializes in applied econometrics, Bayesian analysis, spatial modeling, international political economy, and empirical industrial organization. She has a joint paper with Seung-Whan Choi that is forthcoming in *International Interactions*. The first essay of her dissertation (Chapter 2) is awarded the 2012 Missouri Valley Economic Association (MVEA) Distinguished Student Paper. She also has several working papers some of which are under review.