A Bayesian Approach to Spatial Interaction Model with Spatial Random Effects for Origins and Destinations: A Case of Internal Migration within the City of Seoul

Yongwan Chun*

Abstract: This study presents a Bayesian hierarchical approach to spatial interaction model. Although spatial interaction model has been widely used in migration modeling, many researches point out the misspecification problem, which is due to spatial structure effects. However, the misspecification problem can also arise when variables for the characteristics of origins and destinations do not explain migration flows properly. A Bayesian hierarchical approach provides an efficient way to deal with the problem by taking spatial random effects into consideration. Spatially structured parameters for origins and destinations are capable of capturing latent or unobservable characteristics and then of revealing their pattern in Bayesian framework. An application of internal migration among 25 autonomous districts (gd) within the city of Seoul in 2002 demonstrates empirically that a Bayesian hierarchical model can be advantageous in applying spatial interaction model to migration flows.

Key Words: Bayesian hierarchical model, spatial interaction model, spatial random effects, migration.

요약: 본 연구는 공간 상호작용 모델에 대한 베이지안 계층모델 접근방법을 인구이동 연구의 핵심에서 제시한다. 인구이동 연구에서 공간 상호작용 모델이 별리 사용될수도 불구하고, 기존의 많은 연구들은 공간 구조 효과(spacial structure effects)에 기인한 섬점유류(misspecification)를 지적하고 있다. 그러나 이러한 섬점유류의 문제는 출발지와 도착지의 특성을 반영하는 변수들인 인구이동을 적절하게 설명하지 못하는 경우에도 발생할 수 있다. 베이지안 계층모델 접근방법은 공간 임의 효과(spacial random effects)를 통하여 이 문제를 효과적으로 처리할 수 있도록 한다. 각각의 출발지와 도착지에 발생된 공간적으로 구조화된 계층변수들은, 출발지와 도착지에 대한 잠재적이거나 또는 관측될 수 없는 특징들을 담아내어 그 결과를 도출내게 할 수 있다. 2002년에 발생한 서울 내부 구별 인구이동의 분석은 인구이동 연구에 공간 상호작용 모델을 적용하여 베이지안 계층모델 접근방법이 유효하다는 것을 보여준다.

주요어: 베이지안 계층 모델, 공간 상호작용 모델, 공간 임의 효과, 인구이동.

1. Introduction

Providing reasons for population migration and describing migration patterns have been important objectives of migration research. Governments and businesses have been interested in population migration because forecasting migration flows has the potential to influence socio-economic sectors such as employment opportunities and housing situations. This interest in migration motivated the migration modeling using formulated approaches. The most frequently utilized model in this context is the spatial interaction model (Fotheringham and O’Kelly, 1989). Spatial interaction model, which is based on the gravity model, focuses on two things: (1) geographical distances between origins and destinations, and (2) the influence

* Ph. D., Candidate, The Department of Geography, The Ohio State University (chun.49@osu.edu)
of origins and destinations in investigating migration flows. Geographical distances usually function as impedance for migration. This fact can provide a behavioral rationale for geographical movement. The influence of origins and destinations are expected to reflect the attractiveness of regions for movement. In the estimation of numerically formulated spatial interaction model, it is common practice to use Poisson regression, assuming that each migration flow is independent from the other migration flows (Baxter, 1983a; Congdon, 1992; Flowerdew, 1991; Flowerdew and Aitkin, 1982).

Many researchers have discussed misspecification problem in spatial interaction model. One of the primary issues is spatial structure effects. The spatial structure effects indicate that in migration system, the spatial arrangement of destinations has an influence on the choice for a destination in conjunction with the effects of destinations (Baxter, 1983b; Fotheringham, 1981; Fotheringham and Webber, 1980; Lo, 1991; Sheppard, 1979). That is, the heterogeneous distribution of regions in space leads to competition among close alternative destinations in the choice process. The spatial structure effects have been commonly discussed considering spatial locations and the effects of regions simultaneously. However, Tiefelsdorf (2003) separates "pure" spatial effects which are originated from spatial arrangement of destinations from the effects of regions. In this analytical perspective, he investigates the misspecification which is solely generated by spatial arrangement of regions in migration system.

This issue of this paper is misspecification associated with the effects of regions rather than on geographical distances between regions. This concern can be addressed by looking into existing spatial autocorrelation in the effects of origins, the effects of destinations, or both of them. LeSage and Pace (2005) present the problem of spatial autocorrelation in spatial interaction model. According to them, because variables for origin are collected from geographical units, they are likely to be spatially autocorrelated; likewise, variables for destination are not free from spatial autocorrelation. When this spatial autocorrelation is ignored, the estimated coefficient values become biased and statistical inference can be misled.

Many previous studies have utilized the random effects model to deal with spatial autocorrelation (Congdon, 1993, 2002a; Law and Haining, 2004; LeSage and Llano, 2006). Random effects model simultaneously incorporates region specific random variation with exogenous variables in a regression model. Specifically, with the specification that the random variation are spatially structured, spatial autocorrelation can be dealt with in random effects model (e.g., Cressie and Chan, 1989; Law and Haining, 2004). Regarding to the problem in context of spatial interaction model, region specific random effects can be applied to (1) origins, (2) destinations, and (3) both of origins and destinations simultaneously. Furthermore, random effects model is also used to handle overdispersion, which is defined as greater variability shown than expected in count data modeling. The violation of independent assumption may cause overdispersion (Congdon, 2002b; Wakefield et al., 2000).

This paper aims to illustrate spatial random effects for origins and destinations by using Bayesian hierarchical approach to formulate spatial interaction model. Recent interest of Bayesian approach has been achieved based on the development of Markov Chain Monte Carlo (MCMC) methods and software such as WinBUGS, in which MCMC is easily implemented (Charles and Louis, 2000). As MCMC methods such as Gibbs sampler and Metropolis-Hasting algorithm take advantage of a random sampling approach. These methods eliminate the need for complicated numerical calculation in the process of making fit models. Thus, Bayesian approaches can be utilized for spatial problems in which conventional statistics methods are intractable due to complicated spatial dependent structure. The spatial random effects can be also modeled easily in the framework of Bayesian approaches (Congdon, 2002a). There is much literature on the application of Bayesian spatial modeling in various fields (Banerjee et al., 2004; LeSage, 2000; Moller,