

연령별 인구이동 특성에 대한 탐색적 공간 데이터 분석(ESDA) : 대구시를 사례로*

김 감 영**

Exploratory Spatial Data Analysis (ESDA) for Age-Specific Migration Characteristics : A Case Study on Daegu Metropolitan City*

Kim, Kamyoung**

요약 : 본 연구의 목적은 연령별 인구이동의 구조와 공간적 패턴에 대한 이해를 향상시킬 수 있는 다양한 탐색적 공간 데이터 분석(Exploratory Spatial Data Analysis: ESDA) 기법들을 제시하는데 있다. ESDA 기법의 하나로 지역의 연령별 이동성향과 이동구조를 파악하기 위한 도구로 인구이동 피라미드를 고안하였다. 인구이동 피라미드는 연령별 전입, 전출, 순이동 정보를 이용하여 작성되는 그래픽 도구이다. 또한 연령별 인구이동의 공간적 패턴을 파악하기 위하여 다양한 국지적 연관성 측정 지표인 Local Moran's I_i , Getis-Ord's G_i^* 와 이를 응용한 AMOEBA 기법을 이용하였다. 사례 분석 결과, 제시한 ESDA 기법들을 이용하여 연령별 인구이동 구조, 연령별 이동의 연관성, 연령별 순이동률의 공간 군집 패턴 등을 확인할 수 있었으며, 연령별 공간 군집 사이의 상응관계를 탐색할 수 있었다. 이러한 정보는 인구이동 현상에 대한 상세한 이해뿐만 아니라 인구이동을 유발하는 지역의 특성을 유추하는데 도움을 준다.

주요어 : 탐색적 공간 데이터 분석, 연령별 인구이동, 인구이동 피라미드, 공간 군집, 순이동률

Abstract : The purpose of the study is to propose and evaluate Exploratory Spatial Data Analysis(ESDA) methods for examining age-specific population migration characteristics. First, population migration pyramid which is a pyramid-shaped graph designed with in-migration, out-migration, and net migration by age(or age group), was developed as a tool exploring age-specific migration propensities and structures. Second, various spatial statistics techniques based on local indicators of spatial association(LISA) such as Local Moran's I_i , Getis-Ord G_i^* , and AMOEBA were suggested as ways to detect spatial clusters of age-specific net migration rate. These ESDA techniques were applied to age-specific population migration of Daegu Metropolitan City. Application results demonstrated that suggested ESDA methods can effectively detect new information and patterns such as contribution of age-specific migration propensities to population changes in a given region, relationship among different age groups, hot and cold spot of age-specific net migration rate, and similarity between age-specific spatial clusters.

Key Words : Exploratory Spatial Data Analysis(ESDA), Age-Specific Migration, Population Migration Pyramid, Spatial Clusters, Net Migration Rate

1. 연구 배경 및 연구 목적

사회가 점차 고도화되면서 인구이동은 지역 및 공간 변화를 이해하기 위한 기초 현상으로 받아들여지고 있다. 사회의 고도화에 따른 핵가족화, 만혼의 증가, 여성 경제활동 참여 인구의 증가 등은 출산력의 감소를 가져왔다. 통계청의 발표에 따르면, 2009년 우리나라의 합계출산률은 1.15로 대체수준인 2.1을 한참 밑돌고 있으며, 전세계적으로도 가장 낮은 수치를 보이고 있을 만큼 심각한 저출

산 문제에 당면하고 있다. 저출산과 함께 근대화 및 의학기술의 발달에 의한 평균수명의 증가는 사회의 노령화 정도를 빠르게 진전시키고 있다. 이러한 저출산·고령화 사회에서 인구 변화에 대한 인구이동의 영향력은 더욱 커질 수밖에 없다. 이제 인구이동은 인구의 공간적 재분포에 결정적인 영향을 주고 있으며, 지역의 변화를 이해하는데 그 중요성이 증대되고 있다. 왜냐하면 인구이동은 기원지와 목적지 인구의 단순한 증감을 넘어 인구학적, 사회적, 경제적 구조를 변화시키기 때문이다

* 이 논문은 2009학년도 경북대학교 신입교수정착연구비에 의하여 연구되었음.

** 경북대학교 사범대학 지리교육과 전임강사(Full-Time Lecturer, Department of Geography Education, Kyungpook National University)(kamyoungkim@knu.ac.kr)

(Clark and Onaka, 1983). 따라서 인구 및 사회 변화를 유발하는 인구이동에 대한 체계적인 분석이 절실히 요구된다. 특히 인구의 절대감소를 경험하는 지역에 대한 정책 대안 수립과정에서 지역의 변화를 직접적으로 보여주는 인구이동 특성을 분석하는 것이 선행되어야 한다.

인구이동은 선택적이고 차별적인 과정으로 성, 연령, 교육수준 등에 따라 이동의 의사결정과정과 목적지의 선택이 달라진다(Peters and Larkin, 2005; 이희연, 2005). 따라서 지역의 인구이동을 연령별, 성별, 교육수준별로 구분하여 분석하면 지역 전체 인구이동을 분석할 때에는 파악할 수 없는 새로운 정보를 탐지하는 것이 가능하다. 우리나라의 경우, 최근 수도권 집중 현상과 지역 불균형 원인을 설명하기 위하여 선별적 인구이동에 대한 논의가 활발하게 이루어지고 있다(최은영, 2004; 권상철, 2005; 최진호, 2008; 최은영 등, 2010). 선별적 인구이동에 대한 이러한 논의들에서 공통적으로 다루어지고 있는 공간적 스케일은 광역시도 혹은 대도시권이다. 이러한 접근은 지역 전체적 관점에서 인구이동의 선별성을 파악하고 그 원인을 분석하여 대안을 수립하는데 도움을 준다. 그러나 광역시도 혹은 대도시권 내부의 하위 지역들이 선별적 인구이동과정에서 모두 동일한 특성을 보이는 것은 아니며, 그 내부에서도 다양한 공간적 변이가 존재한다. 따라서 지역의 특성과 여건에 맞는 대안을 수립하기 위해서는 전체로서 광역시도 혹은 대도시권과 외부와의 선별적 인구이동 특성뿐만 아니라 하위 공간단위 내에서의 선별적 인구이동의 공간적 차이를 파악하는 것이 요구된다. 한편 기존 연구들에서 선별적 인구이동을 이해하는 변수로 연령을 공통적으로 들고 있다. 연령은 지역의 이동성 수준을 이해하는 핵심변수라는 것이 많은 연구를 통하여 밝혀졌다.

이상을 종합해보면 선별적 인구이동을 파악하려면 상이한 공간 스케일에서의 분석이 요구되고 이 과정에서 연령 변수를 고려하는 것이 중요함을 알 수 있다. 즉 상이한 공간 스케일에서 인구이동 특성을 파악하기 위하여 연령별 인구이동의 구조와 공간적 패턴 분석이 요구된다. 그러나 기존 논의의 경우 이러한 요구를 충족시키지 못한다. 지역의 연령별 인구이동 구조와 성향을 파악하기 위하

여 제시된 방법인 인구이동 도표(migration schedule)의 경우(Pittenger, 1974; Rogers et al., 1978; Coll and Stillwell, 1999), 지역의 연령별 이동률의 상대적 차이는 파악할 수 있지만 여러 지역 사이의 연령별 인구이동의 절대적 차이를 파악할 수 없고 연령별 인구이동이 지역의 인구 증감에 어떻게 관여하는지를 파악하기 어렵다. 한편 최근 들어 연령별 순이동률의 지도화 및 공간 통계적 기법을 적용하여 연령별 인구이동의 공간적 특성을 파악하고자 하는 시도(Johnson et al., 2005)가 이루어지고 있지만 매우 제한적이다.

이러한 맥락에서 본 연구의 목적은 한 지역에 대한 연령별 인구이동 성향과 구조를 효과적으로 시각화하고 지역의 하위 공간 단위에서 나타나는 연령별 인구이동의 공간적 패턴을 탐색하는 방법을 제시하는데 있다. 이러한 연구 목적을 달성하기 위하여 본 연구는 방법론적 측면에서 탐색적 공간 데이터 분석(Exploratory Spatial Data Analysis: ESDA)에 기초한다. ESDA는 공간 데이터 자체가 갖고 있는 고유한 성격을 다양한 측면에서 심도 깊게 살펴보는 탐색적이고 기술적인(descriptive) 접근으로, 인구이동 데이터처럼 정보의 양이 많고 복잡한 데이터에서 새로운 패턴을 탐지하고, 이를 바탕으로 현상에 대한 가설과 모델을 수립하는데 유용하다. 연령별 인구이동의 구조 및 공간적 특성을 탐지할 수 있는 ESDA 기법들을 제시하고 구체적인 사례에 적용하여 그 유용성을 평가한다. 이러한 과정을 통하여 탐색된 새로운 정보들은 인구이동 현상에 대한 이해의 정도를 높여 주고 인구이동을 유발하는 지역적 특성을 유추하는데 도움을 준다. 나아가 이러한 정보를 바탕으로 인구이동에 대한 다양한 가설을 수립하고 인구이동의 원인을 설명하는 모델의 수립과정에서 유의미한 변수를 선정하는데 도움을 줄 수 있다.

2. 연령별 선별적 인구이동

19세기 말 Ravenstein(1885)이 인구이동 법칙을 정리하여 제시한 이후, 공간 상호작용 모델과 포아송 회귀 모델을 비롯한 다변량 통계분석 기법을 이용하여 인구이동을 모델링하고 원인을 설명하고자 하는 시도에서부터 선별적 인구이동의 공간적

특성을 밝히고자 하는 연구까지 인구이동에 대한 다양한 연구들이 수행되어 왔다. 인구이동은 선별적 이동으로, 경제, 사회 및 인구학적 특성, 즉 연령, 성, 교육정도, 직업, 생애주기 및 혼인관계 등에 따라 차별적으로 나타난다(이희연, 2005). 많은 연구들이 이러한 인구이동의 선택성과 차별성을 밝히고자 시도하였고 그러한 과정 속에서 교육기회, 고용기회, 주택공급과 같은 도시 혹은 지역의 경제 구조 및 공간 구조가 선별성을 유발시킨다는 결론에 도달하였다(최은영, 2004; 권상철, 2005; 최진호, 2008; 최은영 등, 2010). 여기에서는 선별적 인구이동, 특히 연령별 인구이동과 관련된 논의들을 중심으로 살펴본다.

먼저 최근 국내에서 이루어진 선별적 인구이동과 관련된 논의들을 살펴보자. 최은영(2004)은 수도권으로의 선택적 인구이동을 파악하기 위하여 연령별 순이동을 살펴 보았다. 그 결과 수도권 집중의 결정적 역할을 하는 연령층이 20대임을 보여주었고 이를 수도권의 교육기회와 관련지어 설명하였다. 비슷하게 최은영 등(2010)은 부산광역시와 다른 지역과의 연령별, 학력별 선택적 인구이동을 분석하여 부산광역권의 인구이동 특성을 파악하였다. 최진호(2008)는 수도권 인구이동에 초점을 맞추어 성별, 연령별, 교육수준별, 혼인상태별, 경제활동상태별, 직업별 인구이동 등 좀 더 다양한 수준에서 선별성을 기술하였다. 그 결과 연령, 혼인상태, 학력 등이 수도권으로의 인구이동에 큰 영향을 미치는 것으로 나타났다. 한편 이러한 연구들이 통계데이터를 요약하여 인구이동의 선별성을 기술하는데 그친 반면, 권상철(2005)은 통계 분석 기법을 이용하여 인구이동에 대한 선별성의 영향을 파악하고자 시도하였다. 그는 로지스틱 회귀분석을 이용하여 수도권의 인구집중과 관련하여 실업률, 제조업 비율, 노동직 비율, 농촌/도시구분과 같은 유출지역의 지역적 특성과 연령, 성별, 교육정도, 결혼상태 등 이주자의 선별성이 인구 유출에 미치는 영향 정도를 평가하였다. 그 결과 연령, 교육수준, 제조업 비율, 노동직 비율, 실업률 등의 순으로 인구이동에 대한 중요성이 나타나 수도권으로의 이동이 경제적인 고려를 배경에 두고 이루어지고 있음을 보여주었다.

이상의 논의들이 갖는 몇 가지 공통점이 있다.

하나는 선택적 인구이동의 평가가 광역시 혹은 대도시권과 같이 상대적으로 넓은 공간 단위를 대상으로 이루어졌다는 점이다. 선별성의 논의를 광역시 혹은 대도시권 단위로 논의할 경우 광역권 단위로 선별적 인구이동에 의해서 발생하는 문제점에 대한 대안을 수립하는데 유용한 정보를 제공할 수 있지만, 광역권이라는 공간 단위의 내적 다양성을 파악하는데 한계가 있다. 두 번째 공통점은 대부분의 연구에서 인구이동의 선별성의 중요한 요인으로 연령을 들고 있다는 점이다. 즉 인구이동은 연령의 측면에서 매우 선택적임을 보여주었다. 연령은 국가적 혹은 지역적 단위에서 이동성의 수준을 이해하는데 핵심변수라는 것이 많은 연구를 통하여 밝혀졌다(Plane and Heins, 2003).

선별적 인구이동에서 연령에 초점을 둔 연구들에 대한 검토는 연령별 인구이동에 대한 새로운 탐색적 방법을 고안하는데 도움을 준다. 먼저 Pittenger(1974)는 연령그룹을 x 축으로, 전입률, 전출률, 순이동률을 y 축으로 하는 그래프를 작성하고, 이러한 값들이 높은지 여부와 어느 연령그룹에서 가장 높게 나타나는지 여부에 따라 그래프의 유형을 범주화하였고 유형이 전형적으로 나타나는 지역의 특성을 기술하였다. Rogers et al. (1978)은 인구이동을 연령별로 구분하여 관측된 이동을 분포에 적합한 함수를 추정하여 도식화한 것을 '모형 인구이동 도표(model migration schedule)'라고 명명하였다. 이 도표를 통하여 어느 연령대에서 이동의 최고점이 나타나는지 혹은 몇 개의 정점이 나타나는지 등을 파악할 수 있는데, 이러한 정보는 지역의 경제적 여건을 이해하는데 유용하다. Coll and Stillwell(1999)은 한 지역의 연령별 전입률 및 전출률 인구이동 도표를 전체 지역 평균과 비교하여 몇 가지 유형으로 범주화하였다. 전체 평균에 비하여 20~30대의 인구이동이 높은 유형, 20~30대의 이동이 높으면서 65세 이상의 이동도 높은 유형, 평균과 유사하지만 이동의 최고점이 30~40대에서 나타나는 유형 등을 파악하여 이들 지역이 갖는 지역적 특성을 기술하고 유형의 공간적 패턴을 시각화하였다. 한편 최은영 등(2010)은 연령별 순이동지수를 그래프화하여 지역의 연령별 이동 특성을 파악하고자 하였다.

이와 같이 인구이동의 연령별 특성을 파악하기

위한 노력은 오래 전부터 수행되어 왔다. 그러나 연령별 인구이동 특성을 탐색하는 방법은 인구이동 도표(migration schedule)에 국한되어 있었다. 연령별 인구이동 특성을 파악하는데 인구이동 도표는 몇 가지 제한점을 갖는다. 첫째, 이동을 정보를 이용하기 때문에 지역 간의 연령별 이동률의 상대적 차이를 파악할 수 있지만 연령별 인구이동의 절대적 차이를 파악할 수 없다는 것이다. 예를 들어 한 지역에서 20대 인구 1,000명 중에 100명이 이동하는 것과 다른 지역에서 동일 연령대 인구 100명 중에 10명이 이동하는 것을 인구이동 도표 상에서 구분할 수 없다. 왜냐하면 해당 연령대에 대한 두 지역의 이동률은 동일하기 때문이다. 둘째, 인구이동 도표에서 연령별 이동을 구조를 통하여 지역 인구변화를 간접적으로 추정할 수 있지만, 연령별 인구이동이 지역의 인구증감에 어떻게 관여하는지 어느 연령층이 인구의 증감을 주도하는지 등을 파악하는 것은 어렵다.

한 지역의 연령그룹별 이동특성을 기술하기 위하여 연령별 이동을 그래프를 작성하는 것을 넘어 연령별 인구이동의 공간적 군집을 파악하고자 하는 시도들이 최근이 이루어지기 시작하였다. Johnson et al.(2005)는 연령그룹별 순이동의 시공간적 패턴, 특히 연령그룹별 순이동의 공간 군집 패턴을 파악하기 위하여 공간적 자기상관 분석을 이용하였다. 연령그룹별로 전역적 Moran's I를 산출하여 순이동률이 공간적 자기상관을 보임을 밝혔고, Moran 산점도(scatter plot)를 이용하여 양의 순이동률 혹은 음의 순이동률의 공간적 군집패턴을 시각화하였다. 그러나 이렇게 확인된 군집은 개별 공간단위를 평가하여 확인한 결과로 군집을 구성하는 공간단위 사이의 관련성을 파악할 수 없다는 단점을 지닌다(이상일 등, 2010).

이상에서 본 것처럼, 연령별 인구이동 특성을 파악하는데 연령별 인구이동 도표는 몇 가지 맹점을 갖고 있으며 연령별 인구이동의 공간적 군집패턴을 파악하고자 하는 시도는 매우 제한적으로 이루어 졌다. 따라서 본 연구는 하나의 공간 단위에서의 연령별 인구이동 특성뿐만 아니라 하위 공간 단위에서 나타나는 이동 패턴을 탐색할 수 있는 다양한 탐색적 기법들을 제시하고자 한다. 전입률, 전출률 혹은 이동률을 이용하여 연령별 이

동의 상대적 차이를 기술하는 인구이동 도표를 절대적 관점으로 확대하여, 전출, 전입, 총이동, 순이동을 도표화하는 새로운 시각화 방법을 고안한다. 또한 하위 공간 단위에서 나타나는 연령별 인구이동 특성을 탐색할 수 있는 다양한 탐색적 기법들을 제시한다. 구체적으로 연령별 인구이동의 공간적 군집을 파악할 수 있는 Local Moran's I, Getis-Ord's G^* 와 이를 응용한 AMOEBA 등을 소개하고 구체적인 사례에 적용하여 이 기법들의 유용성을 평가한다.

3. 연령별 인구이동의 탐색적 분석

연령별 인구이동 특성을 파악하기 위한 탐색적 방법은 크게 두 부분으로 나뉜다. 하나는 단일 지역의 연령별 인구이동 구조를 파악하기 위한 것이고, 다른 하나는 여러 지역에 걸쳐 나타나는 연령별 인구이동의 공간적 군집을 탐색하는 것이다. 먼저 Rogers et al.(1978)의 인구이동 도표를 개념적으로 발전시켜 연령별 전입, 전출, 총이동, 순이동 구조를 파악할 수 있는 새로운 탐색적 방법을 고안한다. 본 연구에서는 이를 '인구이동 피라미드(population migration pyramid)'라고 명명한다. 인구이동 피라미드는 인구피라미드와 유사한 방식으로 작성된다. 인구피라미드에서 x 축은 성별 구분을, y 축은 연령(혹은 연령그룹)을 표현하는데 반하여, 인구이동 피라미드에서 x 축은 전입량(x 축의 + 영역), 전출량(x 축의 -영역), 전입량과 전출량의 차이인 순이동량을 함께 표현하고 y 축은 연령(혹은 연령그룹)을 나타낸다. 전입량과 전출량에 의해서 표현되는 그래프 길이의 합은 총이동량을 나타내게 된다. 이러한 시각화 기법을 통하여 한 지역의 연령별 전출입, 총이동 및 순이동 구조를 파악할 수 있고 이를 통하여 이동성향뿐만 아니라 지역의 인구변화에 대한 연령별 기여 정도를 평가할 수 있다.

다음으로 연령별 인구이동의 공간적 군집을 탐색하기 위하여 국지적 공간 연관성 지표(local indicators of spatial association: LISA)를 활용한다. LISA의 활용은 방법론적 측면에서 ESDA에 뿌리를 두고 있다. ESDA는 데이터셋의 공간적 특성을 탐지하는 문제로의 탐색적 데이터 분석(Exploratory

Data Analysis: EDA)의 확장이다. EDA는 그래픽적인 방법들을 이용하여 데이터를 시각적으로 표현하고, 데이터 내에 존재하는 패턴과 관계들을 탐색한 후, 데이터로부터 가설을 도출하여 보다 적합한 통계적 모델을 찾아내는 분석 방법으로 정의된다. 이러한 EDA는 전통적인 가설 검정 방법에 비해 기술적(descriptive) 방법들을 중시하며 단순하고 직관적인 방법을 이용함으로써 원 데이터에 보다 충실하다는 특징을 갖는다. 이를 통하여 EDA는 데이터와 데이터의 기저에 놓인 구조에 대한 분석자의 통찰력을 극대화하는 것을 일차적인 목표로 한다(NIST Engineering Statistics Handbook). 이러한 EDA의 관점과 궤를 같이하여 ESDA는 공간데이터에 대한 다양한 과학적 지도학적 시각화 등의 탐색적 기법을 이용하여 데이터에서 공간적 패턴을 발견하고, 지리 데이터에 기반을 둔 가설을 수립하며, 나아가 공간 모델을 평가하는 것을 목적으로 한다(Haining et al., 1998:457).

인구이동 자료와 같은 대용량 데이터 분석을 위하여 EDA를 활용하는 것이 이상적이고 직관적인 것처럼 보이지만, 공간적 자기상관과 같은 공간 데이터가 갖는 특수한 성질 때문에 독립성 가정 하에서 개발된 통계학의 EDA 기법을 활용하는 것이 쉬운 일이 아니다(Anselin, 1999). 이러한 이유로 지리정보시스템(Geographic Information Systems: GIS)과 국지적 통계량(local statistics)을 결합하고 공간데이터의 특수한 성질을 고려한 다양한 ESDA 기법들이 개발되었다. GIS와 ESDA의 결합을 통한 공간데이터의 시각화는 ESDA의 기본 목적을 구현하는 최적의 환경을 제공하고 있다. 또한 다양한 국지적 통계량의 개발은 연구 지역 전체에 걸쳐 일반적으로 나타나는 보편적인 경향을 파악하는 것을 넘어 하위 지역에서 나타나는 국지적 특이성을 주목함으로써 공간적 이질성을 시각화하고 모델링하는 것을 가능하게 하였다(이상일, 2008). 특히 Local Moran's I_i (Anselin, 1995), Getis-Ord's G_i^* (Ord and Getis, 1995)와 같은 LISA는 시각화 기법을 통하여 공간적 군집, 공간적 특이점, 공간 체계의 확인을 가능하게 한다(Anselin, 1995:94).

Local Moran's I_i 는 공간 단위 i 와 주변 공간 단위들 간의 공변량을 계산하는 것으로 공간적 특이점 및 공간 군집의 탐지가 탁월하지만 핫스팟

(높은 값들의 군집)과 콜드스팟(낮은 값들의 군집)을 수치상으로 구분하지 못하는 한계를 갖는다(이상일 등, 2010). 한편 Getis-Ord's G_i^* 는 기준 공간 단위 i 와 이의 주변 공간 단위가 전체에서 차지하는 비중을 계산하는 것으로, 통계량의 수치로부터 핫스팟과 콜드스팟을 구분할 수 있지만, 공간적 특이점을 확인하는 것은 어렵다. 이와 같이 LISA는 관심 위치와 그 이웃의 유사성에 대한 통계량을 산출하여 유의한 위치들의 공간적 분포를 시각화함으로써 군집을 확인할 수 있는 특징을 갖지만, 이러한 과정을 통하여 산출된 공간 군집은 개별 공간 단위의 평가 결과의 집합으로, 공간 군집을 형성하는 단위 지역들 간에 어떠한 관계에 있는지 평가할 수 없는 한계를 지닌다(이상일 등, 2010).

공간 군집의 확인에서 이러한 제한점을 극복하고자 제시된 것이 AMOEBA(A Multidirectional Optimal Ecotope-Based Algorithm) 기법이다. AMOEBA는 Aldstadt and Getis(2006)가 제시한 불규칙적인 공간 군집의 지리적 형태를 식별하는 방법론으로, 기준 공간 단위에 Getis-Ord의 G_i^* 통계량이 극대화될 때까지 이웃한 공간 단위 및 이웃의 이웃을 단계적으로 포함시켜 공간 군집의 범역을 설정한다. 그러나 기존의 AMOEBA는 군집 분석을 수행하는 공간데이터에 양의 공간적 자기상관이 존재하면서 핫스팟(혹은 콜드스팟)들이 군집하여 분포하는 경우 더 높은(혹은 더 낮은) 값을 갖는 군집을 탐지하는데 어려움이 있다(이상일 등, 2010). 최근 이상일 등(2010)은 기존 알고리즘의 중복 군집 제거 논리가 갖는 제한점을 보완하여 수정 AMOEBA를 제시하였다. 이 알고리즘은 양의 공간적 자기상관을 가지면서, 핫스팟 혹은 콜드스팟의 분포가 서서히 변하거나 독립적인 경우 모두 효과적으로 군집을 탐색할 수 있다.

4. 탐색적 기법의 적용

이상에서 논의한 탐색적 공간데이터 분석 기법들을 대구광역시의 연령별 인구이동 데이터에 적용하였다. 분석에 사용한 인구이동 데이터는 주민등록 전입신고 자료를 바탕으로 작성된 인구이동 통계이다. 이 데이터는 가구 단위의 지역간 인구

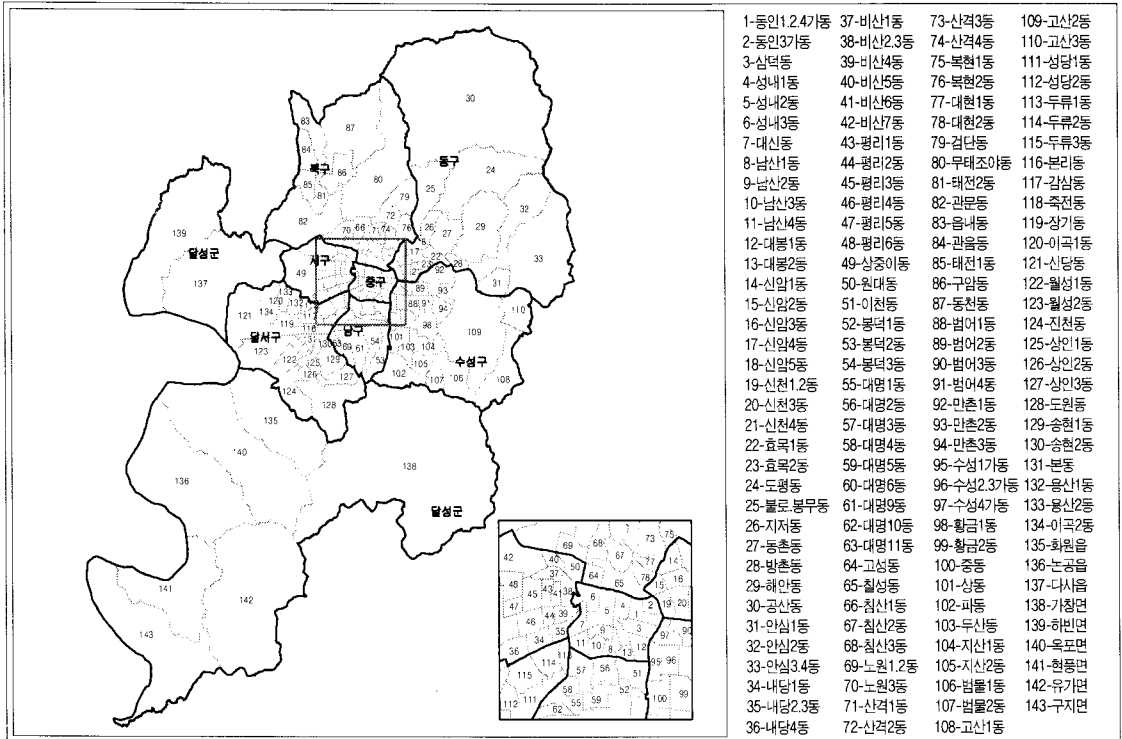


그림 1. 사례 연구 지역(2008년 기준)

이동 자료로, 이주자의 성별 정보뿐만 아니라 연령별 정보를 포함하고 있다. 2001년부터 제공되는 인구이동통계의 기본 공간 스케일은 읍면동이다. 공간 스케일이 읍면동 단위 수준까지 상세하기 때문에 대구와 같은 대도시 내의 하위 공간 단위에서 나타나는 인구이동의 공간 군집을 분석하는 것이 가능하다. 데이터의 분석 시점은 2008년이고 지도화를 위하여 이 시점의 대구시 행정구역을 활용하였다. <그림 1>은 본 연구의 공간적 범위에 해당하는 대구광역시 행정동의 분포를 보여주고 있다.

Microsoft Office Access를 이용하여 인구이동 원자료를 시도, 시군구, 읍면동 단위로 집계하였으며, Microsoft Office Excel을 이용하여 집계된 결과를 그래프화 하였다. 연령별 순이동률 사이의 상관관계를 계산하기 위하여 SPSS 17.0을 이용하였다. 지역별 인구이동 자료의 지도화를 위하여 ESRI의 ArcGIS 9.3을 이용하였다. 특히 Local Moran's I와 Getis-Ord's G_i^* 를 이용한 군집분석을 위하여 ArcGIS의 Spatial Statistics 툴을 활용하였다. 그러나 AMOEBA를 이용한 군집분석은 상

용 GIS 프로그램에서 지원하지 않기 때문에 이상 일 등(2010)이 ArcGIS 환경에서 개발한 모듈을 활용하였다.

1) 연령별 인구이동 특성

먼저 대구시에 대한 연령별 인구이동 도표를 작성하여 연령별 이동 성향을 파악하였다(그림 2). 연령별 이동률은 각 연령의 이동자수(전입자+전출자)를 해당 연령의 연앙인구(기준년도 인구와 이전 연도 인구의 평균)로 나누어 산출하였다. <그림 2>에서 보는 것처럼, 0세에서 높게 형성된 인구이동률은 점차 낮아져 17세에서 최저점을 형성한다. 이후 가파른 상승세를 보여 27세에서 최고점을 형성하고 이후 빠르게 감소한 후 40세 이후에는 안정기에 접어들어 이동률에 큰 변화가 발생하지 않는다. 다만 70세 이후 노년층에서의 이동률이 서서히 증가하는 패턴을 보인다. 어린 연령층에서 이동성이 높게 나타나는 것은 그 부모 세대의 높은 이동 성향 때문이다. 한편 중고등학교

연령별 인구이동 특성에 대한 탐색적 공간 데이터 분석(ESDA) : 대구시를 사례로

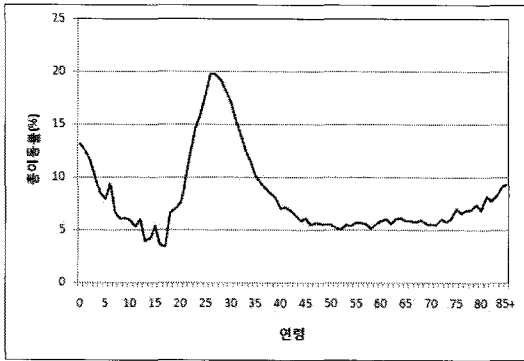


그림 2. 대구시 인구이동 도표(2008년, 자료: 통계청)

이 다른 두 지역에 비하여 다소 완만하게 감소하고 있음을 알 수 있다. 또 다른 특이한 사항은 모든 구군에서 6세, 12세, 15세에서 극지적인 이동률의 정점이 나타난다는 사실이다. 즉 초등학교, 중학교, 고등학교 진학 바로 전의 연령에서 이동률이 높게 나타나고 있다. 이러한 연령에서 교육 관련 요소가 인구이동의 중요한 요인임을 유추해 볼 수 있다.

이상에서 살펴본 연령별 인구이동 도표는 이동률을 기반으로 하고 있다. 따라서 도표를 해석하

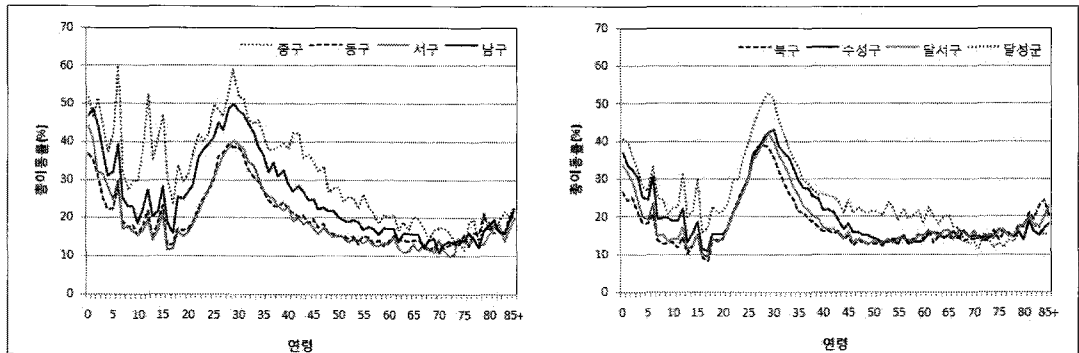


그림 3. 대구시 구별 인구이동 도표(2008년, 자료: 통계청)

제도교육에 의한 이동의 제약이 학생연령층의 이동성향을 낮게 한 것으로 볼 수 있다. 20~30대의 높은 이동성향은 경제적 기회와 관련되는 것으로 풀이할 수 있다.

연령별 인구이동 도표를 구별로 작성하여 보면 대구시의 내적 다양성을 파악할 수 있다(그림 3). <그림 3>을 <그림 2>와 비교하여 보았을 때, 기본적인 패턴은 대구시와 유사하지만 구별로 차이를 확인할 수 있다. 특히 중구의 연령별 이동률의 편차가 심하고 2008년의 경우 6세에서 이동률의 최정점이 나타났다. 또한 다른 지역에 비하여 20대 후반 이동률의 정점 이후 감소가 상대적으로 느리게 나타나고 있다. 중구, 달성군, 남구 등이 상대적으로 높은 이동성향을 보이는 반면 동구, 서구, 북구, 달서구는 상대적으로 낮은 이동성향을 보인다. 수성구의 경우 인구규모가 비슷한 북구, 달서구와 비교하였을 때, 유년층의 이동률이 높고, 최정점을 이루는 연령이 30세로 다른 두 지역보다 뒤에 나타난다. 또한 최정점을 형성한 후 이동률

는데 유의할 점이 한 가지 있다. 이동률은 상대적 개념으로 기저의 인구에 대한 이동자수의 비율이다. 따라서 인구가 적은 상태에서 약간의 이동은 비율 변화에 큰 영향을 줄 수 있다. 이와는 반대로 많은 인구 상황에서 인구이동이 상대적으로 많더라도 비율에 대한 영향이 작을 수 있다. <그림 3>에서 중구의 이동률이 높게 나타나는 것은 상대적으로 적은 인구에 기인한 것으로 볼 수 있다. 이와는 반대로 북구, 수성구, 달서구 등이 낮은 이동성향을 보이는 것은 상대적으로 많은 인구에 기인한 것으로 볼 수 있다. 인구이동의 절대량에서 북구, 수성구, 달서구 등은 중구보다 월등히 많다. 이를 통하여 알 수 있는 것은 연령별 인구이동 도표를 이용하여 연령별 인구이동의 절대적 특성을 파악할 수 없다는 점이다. 또한 이동률은 전입자수와 전출자수를 같이 고려하여 산출한 수치이기 때문에 연령별 이동성향이 지역의 인구 변화에 어떠한 영향을 주는지 파악할 수 없다.

이와 같은 연령별 인구이동 도표에 대한 대안으

로 인구이동 피라미드를 작성하였다. 인구이동 피라미드는 연령별 전입, 전출 및 순이동을 바탕으로 피라미드를 작성한 것으로 그 모양은 인구 피라미드와 유사하다(그림 4). 순이동이 전입 영역에 표시되면 양의 순이동을, 전출 영역에 표시되면 음의 순이동을 의미한다. 연령별 막대 그래프의 폭은 총이동을 의미한다. <그림 4>에서 보는 것처럼 연령별 인구이동 피라미드는 절대수치와 상대수치를 이용하여 작성하는 것이 가능하다. 이러한 인구이동 피라미드를 이용하여 연령별 인구이동 도표에서는 확인이 어려운 새로운 패턴을 탐색하는 것이 가능하다. 이동자수를 이용하여 작성한 대구시의 2008년 인구이동 피라미드를 보면(그림 4a) 전입자수에 상응하는 전출자수를 가지고 있어 피라미드가 대칭적인 나무의 모양을 하고 있다. 이는 Ravenstein(1885)의 인구이동 법칙 중 하나인 “모든 이동의 흐름은 이를 보상하는 흐름을 갖는다”는 것을 잘 보여준다고 하겠다. 연령별로 살펴보면 20~30대 연령층의 인구이동이 가장 많고 점차 연령이 증가하면서 이동량이 감소하는 패턴을 보인다. 젊은 연령층의 이동이 많은 것은 가족을 형성하지 않아 이동이 상대적으로 용이하고 졸업 후 취업 기회를 찾아 이동한 때문으로 것으로 풀이된다. 20세 이하의 경우 약 16~17세 정도에서 가장 적은 인구이동을 보여 고등학교 진학이 이동을 제약하는 것으로 풀이된다. 순이동의 경우

10세 이하 연령층을 제외한 대부분의 연령대에서 전출 영역에 표시되어 음의 순이동을 보이는 것을 알 수 있고, 이러한 현상은 25~30세 사이에서 최고에 달한다. 이 연령층의 유출이 대구의 사회적 증가에 의한 인구 감소의 주된 양상임을 알 수 있고, 나아가 이 연령층을 흡입할 수 있는 대구의 경제활동 기반이 취약함을 유추할 수 있다.

한편, <그림 4b>는 연령별 전입률, 전출률, 순이동률을 이용하여 인구이동 피라미드를 작성한 것이다. 이 피라미드에서 전입률과 전출률의 합을 의미하는 각 연령의 막대 그래프의 길이는 <그림 2>에서와 같이 이동률을 의미한다. 연령별 이동성향은 <그림 2>에서 관찰한 것과 유사하지만, <그림 4b>를 이용하여 어느 연령층의 이동성향이 지역의 인구 증감에 영향을 주었는지 파악할 수 있다. 전반적으로 전입률과 전출률은 대칭적인 모습을 보이지만 20~30세 사이에서 비대칭 정도가 강하게 나타나고 있다. 즉 음의 순이동률이 나타나면서 전입률보다 전출률이 두드러지게 나타나고 있다. 1~12세 사이에서는 양의 순이동률이, 42~71세 사이에서는 음의 순이동률이 나타나고 있지만 그 비율은 높지 않다.

인구이동 피라미드를 하위 지역별로 작성하여 살펴보면 좀 더 다양한 패턴을 탐색하는 것이 가능하다. <그림 5>는 2008년 대구시의 구별 연령별 인구이동 데이터를 토대로 작성한 인구이동 피라

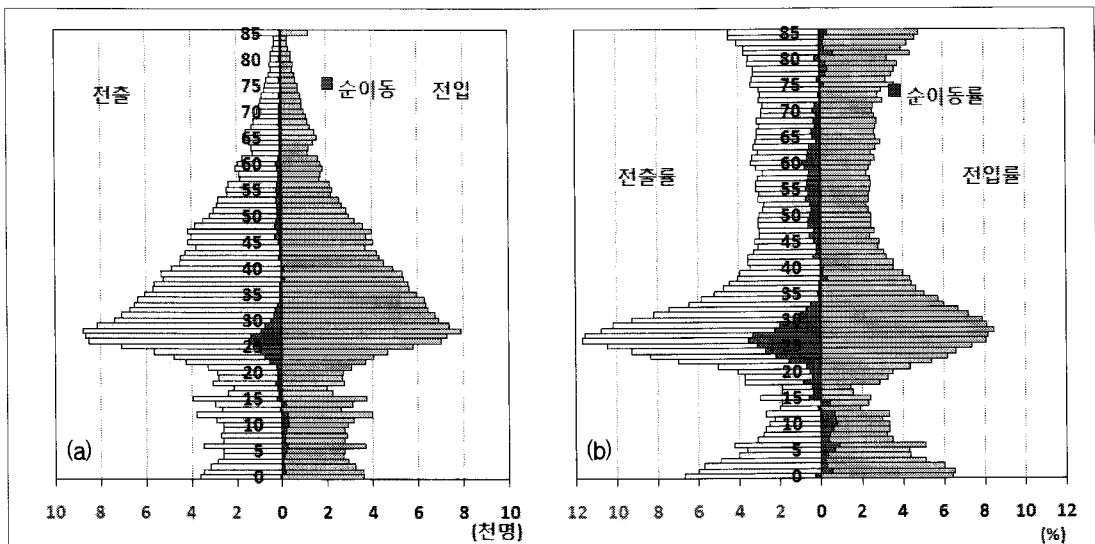


그림 4. 대구시 인구이동 피라미드: a-이동자수, b-이동률(2008년, 자료: 통계청)

연령별 인구이동 특성에 대한 탐색적 공간 데이터 분석(ESDA) : 대구시를 사례로

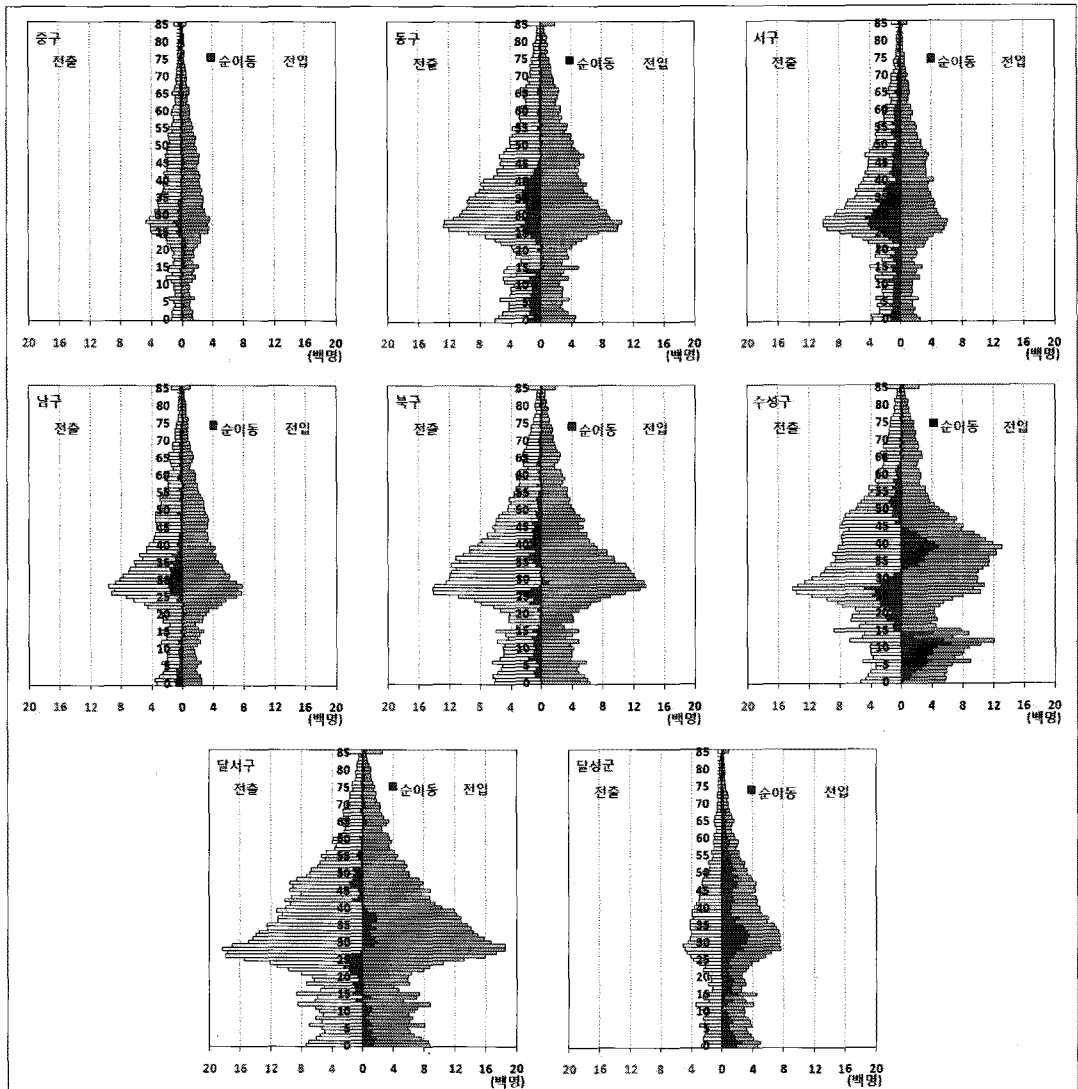


그림 5. 대구시 구별 인구이동 피라미드 (2008년, 자료: 통계청)

미드이다. 지역별 비교를 위하여 x축의 범위를 동일하게 설정하였다. 인구이동 피라미드의 모양에서 수성구를 제외한 모든 구들의 모양이 대구시와 비슷한 형태, 즉 전반적으로 대칭적이며, 유소년층의 인구이동이 상대적으로 적고 20~30대에서 이동량의 최고치가 나타나며 이후 연령에서는 점차적으로 줄어드는 패턴을 보인다. 중구의 경우 다른 지역에 비하여 이동량이 적으며, 20대 초반에서 양의 순이동이 나타나지만 그 정도는 미약하다. 동구, 서구, 남구, 북구 등은 대부분의 연령대

에서 음의 순이동이 나타나고 있어 심각한 인구유출을 경험하고 있음을 알 수 있다. 동구의 경우 0~15세, 20~45세 사이의 음의 순이동이 특징적이다. 이를 통하여 이 지역이 교육적 여건이나 경제적 여건에서 취약함을 유추해 볼 수 있다. 서구의 경우 음의 순이동 경향이 거의 모든 연령대에서 나타나고 있고 25~40세 사이에서 그 양상이 두드러진다. 남구의 경우도 25~40세에서 음의 순이동이 나타나지만 서구에 비하여 그 정도가 약함을 알 수 있다. 북구의 경우 대부분의 연령대에서 음의 순이동이 나타나지만 그 강도는 크지 않다.

수성구는 다른 구들에 비교하여 가장 역동적인 연령별 인구이동 패턴을 보여준다. 다른 구의 경우 보통 25~30세 연령그룹에서 전입, 전출, 순이동 모두에서 최정점을 이루지만 수성구의 경우 9~14세의 전입, 25~30세의 전출, 35~40세의 전입 등 여러 연령그룹에서 국지적 정점이 나타나 비대칭적인 인구이동 피라미드를 만들어 내고 있다. 순이동의 측면에서 보면 0~14세까지는 양의 순이동을 보이다가 15~31세까지는 음의 순이동을 보이고, 다시 32~44세 사이에서 양의 순이동을 보이다 46~60세 사이에서 음의 순이동을 보인다. 즉 일정한 간격을 두고 양과 음의 순이동이 반복됨을 알 수 있다. 이를 통하여 연령별 인구이동의 세대간 관련성을 유추할 수 있다. 보호를 필요로 하는 0~14세의 인구이동은 그 부모세대에 해당하는 30대 초반에서 40대 초반 연령의 인구이동과 강한 관련성을 보이고, 10대 후반에서 20대에 걸친 연령의 인구이동도 40~50대 연령에 해당하는 그 부모세대의 인구이동과 어느 정도 관련되어 있음을 알 수 있다. 이를 통하여 수성구는 학령인구과 그 부모세대를 흡입하는데 유리한 조건을 가지고 있는 반면, 20~30대를 유인할 수 있는 경제활동의 기반이 취약함을 유추할 수 있다. 달서구는 가장 많은 인구이동량을 보이고 있고 연령별 패턴은 수성구와 유사하게 양과 음의 순이동이 반복되는 패턴을 보이지만 그 정도는 약하다. 마지막으로 달성군의 경우 이동량은 많지 않지만 거의 모든 연령대에서 양의 순이동이 나타나는 것을 확인할 수 있다.

이상에서 살펴본 것처럼 하위 지역별로 작성된 인구이동 피라미드를 통하여 파악된 이러한 연령별 인구이동 패턴은 지역 전체로 살펴보았을 때에는 파악할 수 없었던 새로운 패턴으로, 이를 통하여 지역의 다양한 특성, 예를 들어 경제여건, 교육환경, 주거환경을 등을 유추할 수 있었다.

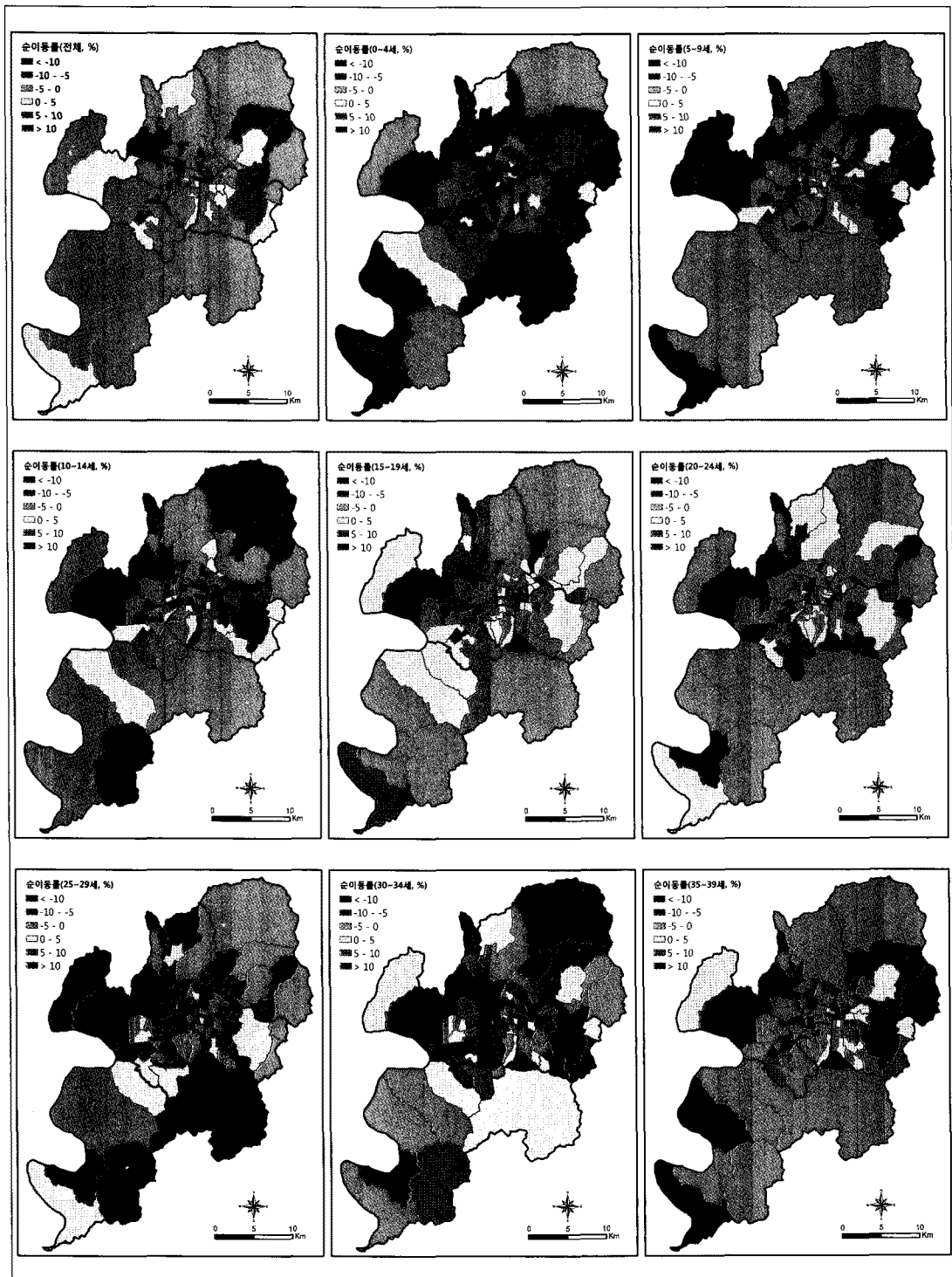
2) 연령그룹별 순이동률의 공간적 패턴

연령별 인구이동을 그 하위 단위별로 지도화하고 분석하면 상위 수준에서 탐지하지 못한 새로운 패턴을 탐지할 수 있다. 예를 들어 지역 내의 어느 부분으로 특정 연령층이 집중하고 어느 부분에

서는 빠져나가는지 그 공간적 패턴을 파악할 수 있다. <그림 6>은 대구시 동별 연령그룹별 순이동률의 공간적 패턴을 보여주고 있다. <그림 6>의 첫 번째 지도는 연령 구분 없이 전체를 대상으로 동별 순이동률을 지도화한 것이다. 순이동률은 전입자에서 전출자를 뺀 순이동자수를 연앙인구로 나누어 계산하며, 인구이동의 결과가 인구증가에 어떠한 영향을 미치고 있는지를 보여준다. 지도에서 보는 것처럼, 대구의 광범위한 지역에서 음의 순이동률을 보이고 있어 사회적 증가에 의한 인구 감소가 대구 전역에 걸쳐 나타나고 있음을 알 수 있다. 또한 음의 순이동률 경향은 도심지역에서 강하게 나타남을 알 수 있다. 이러한 전반적인 음의 순이동률에도 불구하고 수성구, 달서구, 북구의 일부지역에서 양의 순이동률이 나타나고 있다. 특히 중구 남산4동, 수성구 범어3동·수성4가동, 달서구 성당2동·본리동·월성1동, 북구 태전2동·관음동, 동구 안심1동은 10% 이상의 높은 양의 순이동률을 보인다.

0~4세 그룹의 경우 달성군과 북구 일부지역, 수성구와 동구의 일부지역, 달서구 일부지역에서 양의 순이동률이 나타나고 도심의 일부지역을 제외하고 대부분의 지역에서 음의 순이동률이 나타나고 있다. 5~9세와 10~14세의 순이동률 패턴도 0~4세와 매우 유사한 패턴을 보인다. 특히 수성구에서 양의 순이동률을 보이는 동들의 집합을 확인할 수 있다. 15~19세, 20~24세의 양 및 음의 순이동률 분포 패턴은 이전의 연령그룹과 비교하였을 때 그 경향이 뚜렷하지 않은데, 이 연령대가 고등학교 및 대학 재학에 의해서 이동에 제약을 받기 때문에 풀이된다. 이와 대조적으로 대학 졸업 후 취업 연령대에 속하는 25~29세는 공간적 패턴이 좀 더 명확하다. 산업시설이 밀집되어 있는 대구의 북동부지역에서 양의 순이동률 패턴이 나타나고 대부분 지역에서 음의 순이동률이 나타나고 있다. 이는 대구의 일부지역을 제외하고 고용을 창출할 수 있는 경제활동 기반이 취약함을 좀 더 직접적으로 보여준다. 30~34세, 35~39세, 40~44세의 패턴은 강도는 약하지만 각각 0~4세, 5~9세, 10~14세의 패턴과 유사함을 볼 수 있다. 이는 부모와 자녀 세대가 같이 이동한다는 것을 의미한다. 45세 이상의 연령그룹들의 순이동률 공

연령별 인구가동 특성에 대한 탐색적 공간 데이터 분석(ESDA): 대구시를 사례로



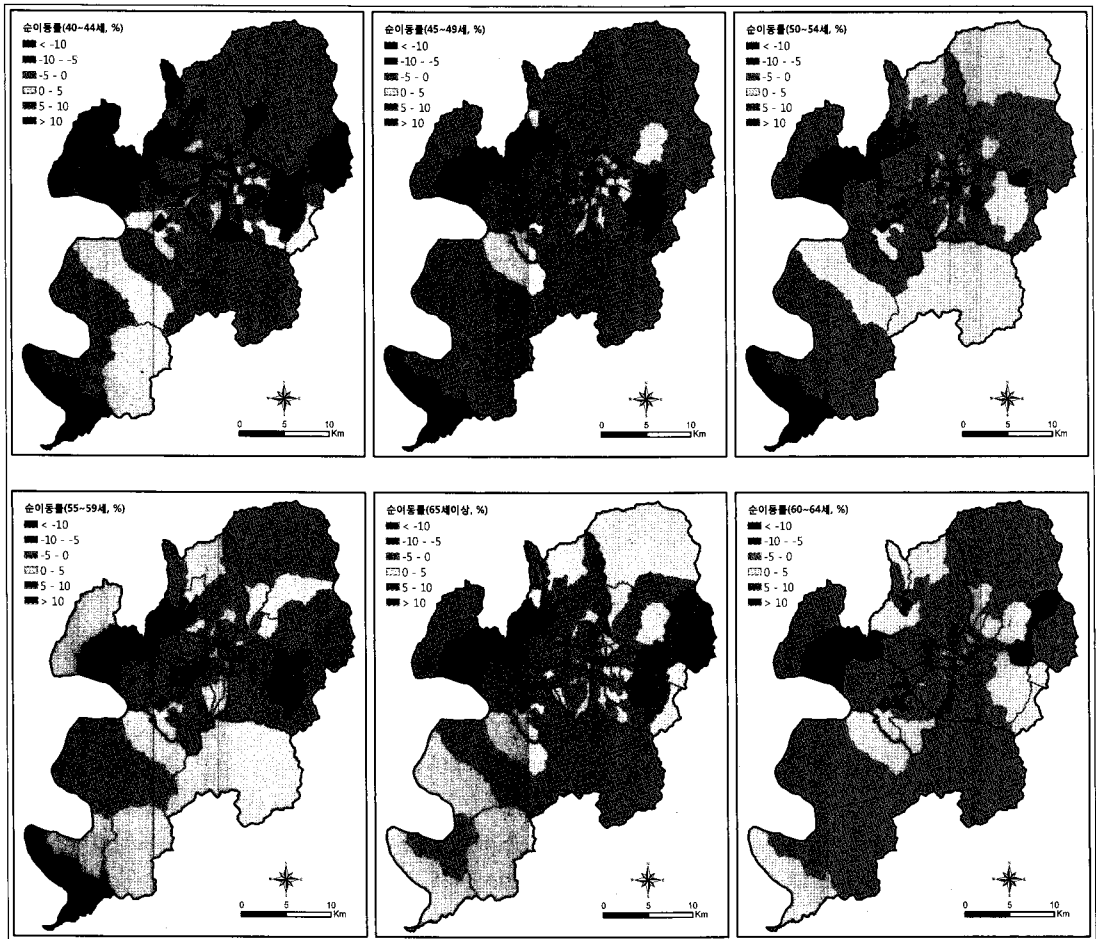


그림 6. 대구시 동별 연령그룹별 순이동률 분포 (2008년, 자료: 통계청)

간 패턴은 큰 차이를 보이지 않으며 일부 지역에서 양의 순이동률을 보이는 것을 제외하여 뚜렷한 패턴이 나타나지 않는다. 또한 음의 순이동률 경향도 강하게 나타나고 있지 않다.

지금까지의 관찰로부터 몇 가지 흥미로운 사실을 발견하였다. 하나는 연령별 순이동률 패턴이 30년(세대) 간격을 두고 매우 유사하게 나타나고 있다는 사실이다. 이를 통계적으로 파악하기 위하여 연령별 순이동률 사이의 상호 관련성 정도를 상관계수를 이용하여 측정하였다(표 1). <표 1>에 제시된 피어슨 상관계수 r 값은 모두 통계적으로 유의하고 전체적으로 연령그룹 사이에 높은 정적 상관관계를 가지고 있음을 알 수 있다. 그러나 각 연령그룹별로 가장 높은 상관관계를 보이는 연령 그룹은 상이함을 알 수 있다. 0~4세, 5~9세, 10

~14세, 15~19세, 20~24세 연령그룹은 각각 부모세대 연령그룹이 속하는 30~34세, 35~39세, 40~44세, 45~49세, 50~54세 연령그룹과 가장 높은 상관관계를 보인다. 이러한 세대간 연관성은 25~29세, 50세 이후 연령그룹에서는 나타나지 않는다. 이들 연령그룹에서는 오히려 세대간 관계보다는 전후 연령그룹과 높은 상관관계를 보인다.

또 다른 흥미로운 점은 연령별 순이동률의 분포가 공간적 자기상관을 보인다는 사실이다. 즉 양의 순이동률과 음의 순이동률이 공간적으로 밀려 있는 것을 확인하였다. 연령별 순이동률이 공간적 자기상관을 보이는지를 통계적으로 확인하기 위하여 전역적 공간 연관성 측정치인 Moran's I 를 계산하였다(표 2). <표 2>에서 보는 것처럼, 전체적으로 동별 순이동률의 군집은 유의수준 0.05에서

표 1. 연령그룹간의 순이동률 상관관계 (2008년, 자료: 통계청)

	0~4세	5~9세	10~14세	15~19세	20~24세	25~29세	30~34세	35~39세	40~44세	45~49세	50~54세	55~59세	60~64세	65세 이상
0~4세	1	.859	.765	.764	.688	.800	<u>.933</u>	.926	.818	.812	.775	.794	.818	<u>.842</u>
5~9세	.859	1	.860	.704	.641	.665	.851	<u>.926</u>	.869	.760	.672	.658	.728	.757
10~14세	.765	.860	1	.674	.562	.588	.727	.889	<u>.914</u>	.750	.610	.614	.674	.762
15~19세	.764	.704	.674	1	.849	.765	.766	.755	.791	<u>.910</u>	.855	.792	.763	.775
20~24세	.688	.641	.562	.849	1	.796	.746	.662	.691	.827	<u>.863</u>	.779	.766	.703
25~29세	.800	.665	.588	.765	.796	1	<u>.869</u>	.728	.695	.789	.839	.845	.845	.828
30~34세	<u>.933</u>	.851	.727	.766	.746	.869	1	.883	.798	.813	.803	.832	.861	.846
35~39세	<u>.926</u>	<u>.926</u>	.889	.755	.662	.728	.883	1	.901	.833	.748	.760	.799	.850
40~44세	.818	.869	<u>.914</u>	.791	.691	.695	.798	.901	1	.852	.750	.742	.775	.819
45~49세	.812	.760	.750	<u>.910</u>	.827	.789	.813	.833	.852	1	.894	.861	.858	.853
50~54세	.775	.672	.610	.855	.863	.839	.803	.748	.750	.894	1	<u>.898</u>	.869	.838
55~59세	.794	.658	.614	.792	.779	.845	.832	.760	.742	.861	<u>.898</u>	1	.897	.868
60~64세	.818	.728	.674	.763	.766	.845	.861	.799	.775	.858	.869	.897	1	<u>.902</u>
65세 이상	.842	.757	.762	.775	.703	.828	.846	.850	.819	.853	.838	.868	<u>.902</u>	1

통계적으로 유의하게 나타났다. 순이동률의 균집 여부를 연령그룹별로 살펴보면 그 결과가 상이함을 알 수 있다. 0~4세, 5~9세, 10~14세 연령

표 2. 연령그룹별 순이동률의 공간적 자기상관: Global Moran's I (2008년, 자료: 통계청)

연령그룹	Moran's I	Z-score	p-value
0~4세	0.13	2.78	0.005
5~9세	0.16	3.40	0.001
10~14세	0.25	5.16	0.000
15~19세	0.07	1.48	0.139
20~24세	0.04	0.90	0.368
25~29세	0.17	3.56	0.000
30~34세	0.18	3.78	0.000
35~39세	0.17	3.61	0.000
40~44세	0.11	2.39	0.017
45~49세	0.01	0.43	0.667
50~54세	0.01	0.32	0.749
55~59세	0.05	1.27	0.204
60~64세	0.04	0.94	0.347
65세 이상	0.07	1.50	0.134
전체	0.10	2.10	0.036

에서 순이동률의 공간적 균집이 통계적으로 유의하게 나타난 반면, 순이동률 패턴이 뚜렷하지 않았던 15~19세, 20~24세 연령그룹의 Moran's I 통계량은 유의하지 않은 것으로 나타났다. 특히 10~14세 연령그룹의 순이동률이 가장 강한 공간적 균집 경향을 보였다. 24~29세, 30~34세, 35~39세, 40~44세 연령그룹의 순이동률의 공간적 균집이 통계적으로 유의하게 나타났고, 이후 연령 그룹에서는 공간적 균집이 나타나지 않았다. 이러한 연령별 순이동률의 공간적 균집 경향은 <그림 6>의 연령별 순이동률의 공간적 패턴에서 관찰한 것과 매우 유사하다.

3) 연령그룹별 순이동률의 공간 균집 분석

앞에서 연령그룹별 순이동률이 공간적 자기상관을 보임을 확인하였다. 이 부분에서는 연령그룹별 공간 균집이 어디에서 나타나는지 확인하고자 한다. 이를 위하여 먼저 연령별 순이동률에 대한 Local Moran's I_i를 계산한 후 Z 값을 산출하여 유의수준 0.05에서 통계적으로 유의한 값들을 지도화하였고, <표 2>에서 전역적으로 공간적 자기상관을 보이는 연령그룹에 대한 지도만을 제시하

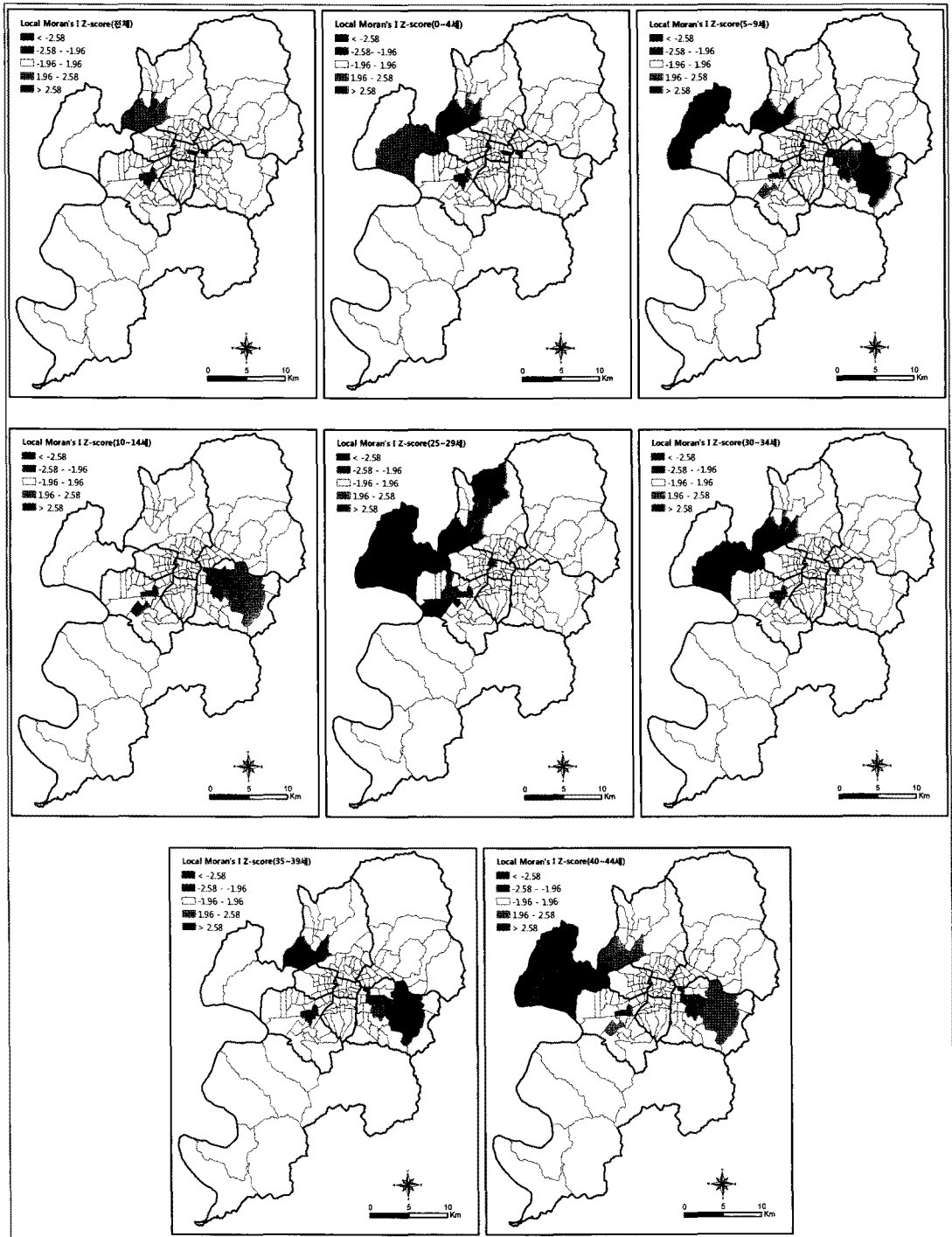


그림 7. 연령그룹별 순이동률에 대한 Local Moran's I 군집분석(2008년)

였다. 높은 양의 Moran's I_i 와 Z 값은 공간 단위가 높거나 혹은 낮은 유사한 값으로 둘러싸여 있는 경우로, 공간적 군집을 의미한다. 반면 공간 단위가 상이한 값(높은 경우 낮은 값, 낮은 경우 높

연령별 인구이동 특성에 대한 탐색적 공간 데이터 분석(ESDA): 대구시를 사례로

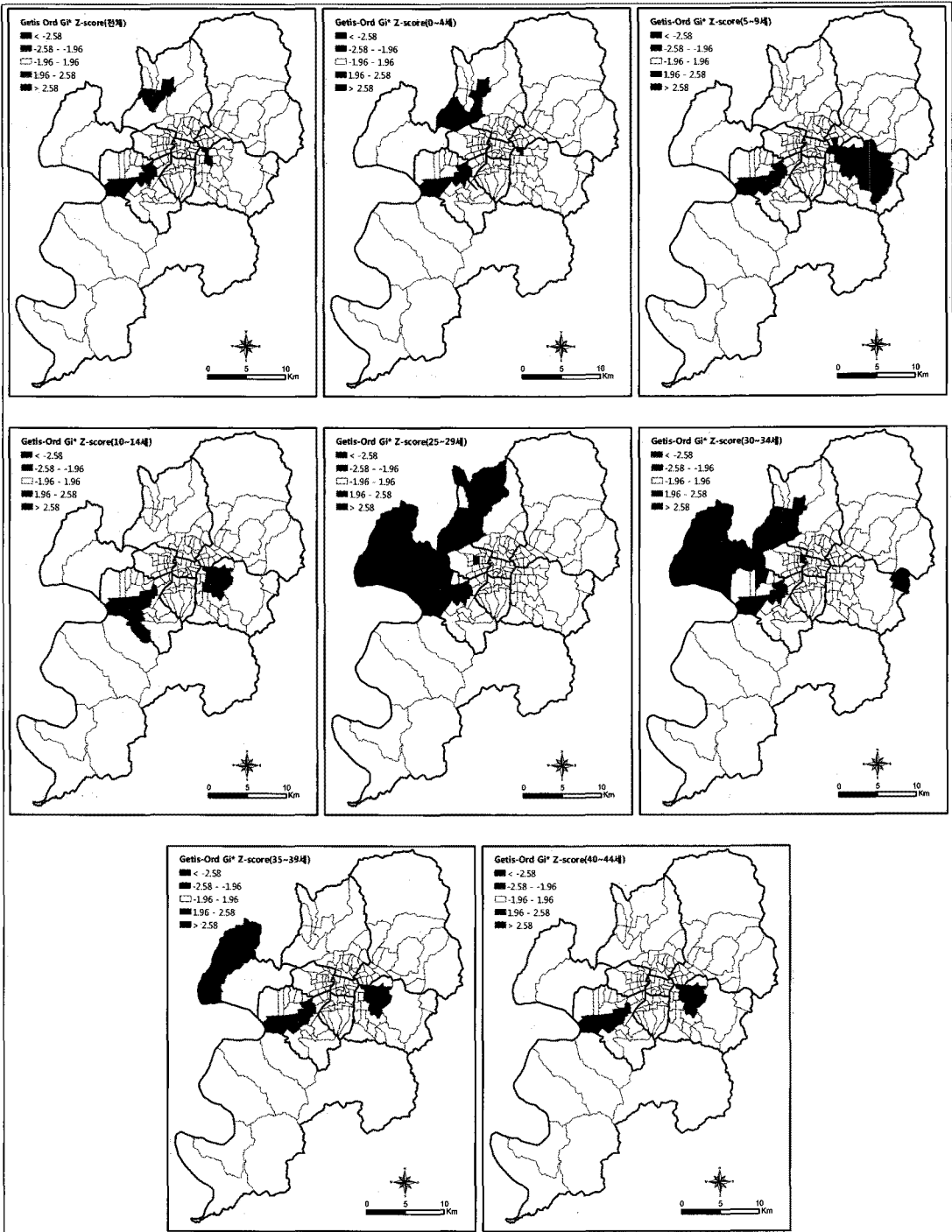


그림 8. 연령그룹별 순이동률에 대한 Getis-Ord G_i^* 군집분석(2008년)

은 값)으로 둘러싸여 있는 경우 Moran's I_i 와 Z 값은 음의 값을 갖고, 공간적 특이점을 의미한다.

<그림 7>에서 붉은 색으로 나타난 양의 Z 값 중 25~29세의 성내3동만이 낮은 값이 낮은 값으

로 둘러싸인 군집이고, 나머지는 모두 높은 값으로 둘러싸인 군집이다. 거의 모든 연령그룹에서 양의 순이동률의 군집을 보이는 곳은 달서구 성당2동과 본리동, 북구 관문동이다. 연령그룹별로 확인된 군집 패턴은 다음과 같다. 0~4세의 경우 달성군의 다사읍과 북구의 관문동·태전2동 지역에서 통계적으로 유의한 공간군집이 나타났다. 5~9세, 10~14세의 경우 수성구의 동들로 구성된 뚜렷한 공간 군집을 확인할 수 있다. 25~29세의 경우 대구의 북서부 지역에 높은 값들의 공간적 군집이 뚜렷하게 나타났다. 이러한 경향은 조금 약화되기는 했지만 30~34세에서도 지속된다. 한편 35~39세, 40~44세에서 다시 수성구 중심으로 양의 순이동률의 공간적 군집이 확인되었다. 이러한 군집 패턴은 이들의 자녀 세대인 5~9세, 10~14세의 패턴과 유사하지만 그 정도는 약함을 알 수 있다. 한편 지도에서 녹색은 각 연령그룹별로 공간 단위가 상이한 값으로 둘러싸인 공간적 특이점을 나타낸다.

Local Moran's I_i 는 공간적 군집과 공간적 특이점을 확인하는데 유용하지만, 군집 확인에 있어서 계산된 통계량만으로는 높은 값들로 둘러싸인 군집인 핫스팟과 낮은 값들로 둘러싸인 군집인 콜드스팟을 구분할 수 없다는 한계를 갖는다. 연령별 순이동률의 공간적 군집에 대한 핫스팟과 콜드스팟을 구분하기 위하여 Getis-Ord G_i^* 통계량과 이에 대한 Z 값을 계산하여 지도화하였다(그림 8). <그림 8> 역시 유의수준 0.05에서 통계적으로 유의한 공간 단위들만 시각화하였다. 양의 Z 값(높은 G_i^* 값)은 높은 값들의 공간적 군집을 나타내고, 음의 Z 값(낮은 G_i^* 값)은 낮은 값들의 공간적 군집을 나타낸다. 연령그룹별로 확인된 공간적 군집은 Moran's I_i 에서 확인된 공간적 군집과 매우 유사하지만 전체적으로 Getis-Ord G_i^* 를 이용하여 확인된 군집이 Moran's I_i 를 이용하여 확인된 군집보다 좀 더 연속적으로 나타나고 군집의 공간적 패턴이 좀 더 명확하다. 특징적인 패턴 중의 하나는 25~29세, 30~34세 연령그룹에서 나타나는 달성군, 달서구, 북구를 연결하는 높은 값들의 군집이다. Moran's I_i 와 같이 Getis-Ord G_i^* 또한 통계적으로 유의한 낮은 값들의 공간적 군집을 거의 확인하지 못하였다. 또한 음의 순이동률

을 갖는 공간 단위가 높은 값의 공간 군집에 포함되는 문제가 나타나기도 하였다. 예를 들어, 모든 연령그룹에서 음의 순이동률을 보이면서 Moran's I_i 의 분포에서 공간적 특이점으로 나타났던 달서구 본동(<그림 8>의 0~4세 지도에서 \blacktriangledown 표시 지역)이 높은 값들의 공간 군집에 포함되어 있다.

이상에서 살펴 본 것처럼, Local Moran's I_i 와 Getis-Ord G_i^* 모두 연령별 순이동률에 대한 높은 값의 공간적 군집을 확인하는 데에는 뛰어나지만, 낮은 값의 공간적 군집을 확인하는 데에는 취약하였다. 또한 이러한 국지적 연관성 측정 통계량은 개별 공간 단위별로 평가하여 시각화함으로써 군집의 공간적 패턴을 확인하는 것으로, 공간 상에서 연속적으로 나타나는 군집을 구성하는 단위 지역 사이에 어떠한 관련성이 있는지는 평가할 수 없는 한계를 지닌다. 한편 Getis-Ord G_i^* 를 이용하여 확인한 공간 군집의 경우 이질적인 값을 갖는 공간 단위를 포함하기도 하여, 유사한 값들을 갖는 공간 단위들의 집합이라는 '공간 군집'의 의미에 배치되는 결과를 보이기도 하였다. 공간 군집 확인에서 발견되는 이러한 문제점을 극복하고자 AMOEBA를 이용하였다. AMOEBA는 기준 공간 단위를 중심으로 순차적으로 이웃한 공간 단위들을 포함시켜가면서 포함된 공간 단위 전체를 대상으로 Getis-Ord G_i^* 를 극대화시켜 군집을 확인한다.

<그림 9>는 연령그룹별 순이동률에 AMOEBA를 적용하여 공간 군집을 확인한 결과이다. Getis-Ord G_i^* 의 통계량의 Z 값을 산출한 후 유의수준 0.05에서 통계적으로 유의한 것만 시각화하였다. 붉은 색으로 표현된 양의 Z 값은 높은 양의 순이동률을 갖는 공간 단위들의 군집을, 녹색으로 표현된 음의 Z 값은 음의 순이동률을 갖는 공간 단위들의 군집을 나타낸다. 분석 결과에서 가장 특징적인 사항은 AMOEBA가 모든 연령그룹에서 음의 순이동률을 갖는 공간 군집을 확인하였다는 점이다. 전체적으로 보면 북구의 관문동과 태전2동, 달서구의 본리동과 성당2동, 수성구의 범어3동과 수성4가동에서 강한 핫스팟이 나타나고 있고, 수성구의 고산2동과 동구의 안심1동을 연결하는 지역에서도 핫스팟을 확인할 수 있다. 이와는 반대로 대구 전역에 걸쳐 광범위한 콜드스팟이 나타나고

연령별 인구이동 특성에 대한 탐색적 공간 데이터 분석(ESDA): 대구시를 사례로

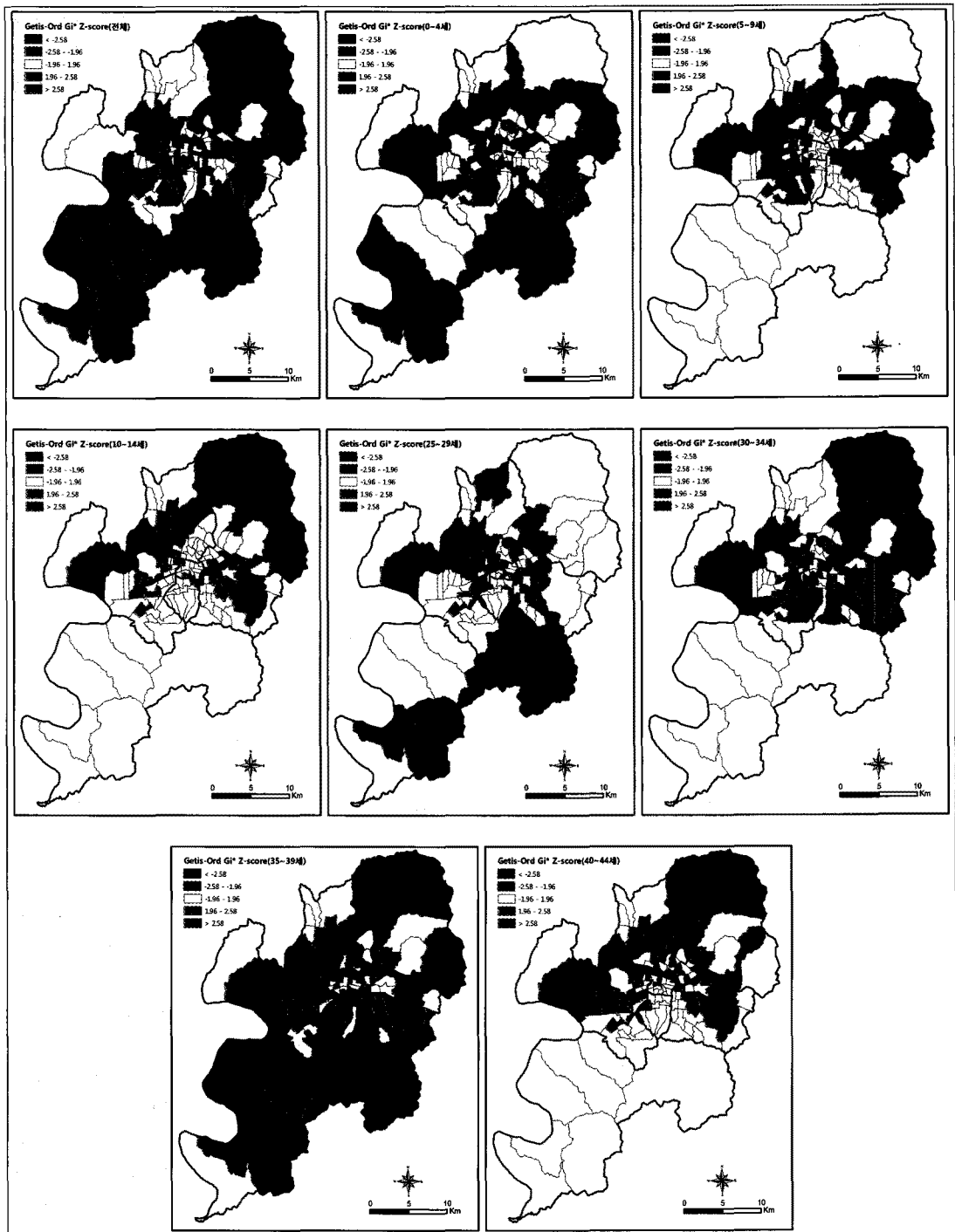


그림 9. 연령그룹별 순이동률에 대한 AMOEBA 군집 분석(2008년)

있다. 연령그룹별로 확인된 군집을 살펴보면 다음과 같다. 0~4세의 경우 달성군의 다사읍, 북구의

관문동과 태전2동을 포함하는 지역, 달서구의 본리동과 성당2동, 동구의 신천3동과 수성구의 범어

3동에서 강한 핫스팟이 나타나고, 수성구의 고산1·2동, 동구의 안심1동을 포함하는 지역에서도 어느 정도 높은 값의 군집이 나타난다. 핫스팟이 국지적으로 나타나는 반면 콜드스팟은 대구시 전역에 걸쳐 광범위하게 분포하고 있다. 5~9세 연령 그룹에서 확인된 특징적인 패턴은 0~4세에서는 미약했던 수성구를 중심으로 양의 순이동률의 강한 군집이다. 또한 음의 순이동률의 공간적 군집이 대구시 중심부에서 북쪽으로 축소되어 나타난다. 10~14세의 군집 패턴은 5~9세의 군집 패턴과 매우 유사하게 나타난다. 25~29세의 경우, 대구의 북서부 지역에서 강한 양의 순이동률의 군집을 확인할 수 있고, 대구의 중심부에서 수성구를 연결하는 지역에서 음의 순이동률의 군집이 광범위하게 나타나고 있다. 30~34세의 양의 순이동률 군집 패턴은 0~4세와 동일하며, 음의 순이동률 군집은 대구 중심부에서 북쪽에 걸쳐 나타나고 있다. 35~39세의 양의 순이동률 군집 또한 5~9세의 군집과 유사하며, 음의 순이동률 군집은 대구시의 대부분의 지역에 걸쳐 나타난다. 40~44세의 군집 패턴 또한 10~14세의 패턴과 매우 유사하다.

전체적으로 양의 순이동률의 공간 군집은 매우 제한적인 공간 범위에서 나타나는 반면, 음의 순이동률의 공간 군집은 광범위하게 나타나는 것을 볼 수 있다. 그중에서 달성군의 다사읍과 북구의 관문동·태전2동을 포함하는 지역, 달서구의 본리동과 성당2동을 포함하는 지역 두 곳이 거의 모든 연령그룹에서 양의 순이동률의 공간 군집으로 확인되어 대구 내에서 인구이동의 주된 목적지가 되고 있음을 알 수 있다. 수성구의 경우 연령그룹별로 양의 순이동률 군집과 음의 순이동률 군집이 반복하여 나타나는 흥미로운 경향을 보여주었다. 또한 0~4세, 5~9세, 10~14세의 양의 순이동률 군집 패턴은 그 부모세대라 할 수 있는 30~34세, 35~39세, 40~44세의 패턴과 유사하게 나타나, 세대 간 군집 패턴에서의 상응관계를 확인할 수 있다. 마지막으로 <그림 6>과 군집 분석의 결과인 <그림 7·8·9>를 비교해볼 때, AMOEBA가 연령별 순이동률의 공간적 군집을 가장 효과적으로 탐지해 낼 수 있다. 즉 Local Moran's I_i , Getis-Ord G_i^* 는 연령별 순이동률에 대한 높은 값의 공간적 군집을 확인하는데 뛰어났지만 낮은 값의

공간적 군집을 탐색해내지 못하였다. 뿐만 아니라 Local Moran's I_i 는 핫스팟과 콜드스팟을 지도상으로 구분할 수 없었으며, Getis-Ord G_i^* 는 낮은 값을 갖는 공간 단위를 포함하는 높은 값의 공간 군집을 확인하는 데에 문제점을 보였다. 이와는 달리 AMOEBA는 높은 값들의 공간 군집과 낮은 값들의 공간적 군집을 효과적으로 탐지하였다.

사례 분석을 통하여 본 연구에서 제시한 ESDA 기법들을 이용하여 연령별 인구이동 구조, 연령별 관련성, 연령별 순이동률의 공간 군집 패턴 등을 확인할 수 있었으며, 연령별 공간 군집 사이의 상응관계를 탐색할 수 있었다. ESDA를 이용하여 탐색된 이러한 정보들을 바탕으로 인구이동 현상에 대한 이해 정도를 높일 수 있을 뿐만 아니라 인구이동을 유발하는 지역적 특성, 예를 들어 경제여건, 교육환경, 주거환경 등을 유추할 수 있었다. 이러한 정보는 인구이동에 대한 다양한 가설을 수립하고 인구이동의 원인을 설명하는 모델의 수립과정에서 유의미한 변수를 선정하는데 도움을 준다.

5. 요약 및 결론

저출산·고령화를 특징으로 하는 현대사회에서 인구이동은 지역의 인구학적, 사회적, 경제적 변화를 이해하는 중요한 현상으로 받아들여지고 있다. 이러한 인구이동은 연령, 성, 교육수준, 직업 등에 따라 선택적이고 차별적으로 나타나는 현상이다. 특히 이주자들의 선택적 혹은 차별적 이동성향에 영향을 주는 중요한 요인 중 하나는 연령이라는 것이 많은 연구들에 의해서 확인되었다. 이러한 배경 하에서 본 연구는 연령별 인구이동의 구조와 공간적 패턴에 대한 이해를 향상시킬 수 있는 다양한 탐색적 공간 데이터 분석 기법들을 제시하였고, 사례연구를 통하여 기법들의 효용성을 보여주었다.

먼저 지역의 연령별 인구이동 구조를 파악하기 위하여 인구이동 지표의 개념을 발전시켜 인구이동 피라미드를 고안하였다. 인구이동 피라미드는 연령별(혹은 연령그룹별) 전입, 전출, 순이동 정보를 바탕으로 작성한 것으로, 전입량, 전출량, 순이동, 총이동을 모두 표현할 수 있다. 이를 이용하여 한 지역에서 어느 연령의 이동성향이 높으며, 연

연령별 이동성향이 지역의 인구 증감에 어떻게 기여하는지 파악할 수 있었고 연령별 인구이동 구조의 지역적 차이를 파악하는 것이 가능하였다. 또한 인구이동 피라미드로부터 지역의 경제여건, 교육 환경, 주거환경 등 지역의 특성을 유추할 수 있었다.

다음으로 연령별 인구이동의 공간적 구조를 파악하기 위하여 연령별 순이동률을 이용하여 공간 군집 분석을 수행하였다. 공간 군집 분석을 위하여 국지적 연관성 측정 지표인 Local Moran's I_i , Getis-Ord G_i^* 와 이를 응용한 AMOEBA 기법을 이용하였다. 사례연구를 통한 분석 결과 Local Moran's I_i , Getis-Ord G_i^* 는 연령별 순이동률에 대한 높은 값의 공간적 군집을 확인하는데 뛰어났지만, 낮은 값의 공간적 군집을 탐색해내지 못하였다. 뿐만 아니라 Local Moran's I_i 는 핫스팟과 콜드스팟을 지도상으로 구분할 수 없었으며, Getis-Ord G_i^* 는 낮은 값을 갖는 공간 단위를 포함하는 높은 값의 공간 군집을 확인하는데 어려움이 있어 유사한 값을 갖는 공간 단위들의 집합이라는 '공간 군집' 개념에 위배되는 결과를 보여주었다. 이러한 기법들과는 달리 AMOEBA는 높은 값들의 공간 군집과 낮은 값들의 공간 군집을 효과적으로 탐지하여 '공간 군집'의 본연의 의미에 맞는 결과를 도출하였다.

대구의 사례연구를 통하여 파악된 연령별 인구이동 특성 및 공간적 패턴을 살펴보면 다음과 같다. 연령별 이동성향에서 20~30대의 이동성향이 높게 나타나지만, 초·중·고등학교 취학 직전 연령에서 높은 이동성향이 탐지되어 이들 연령대에서 교육 관련 요소가 인구이동을 유발시키는 주된 요인임을 알 수 있었다. 구별 인구이동 피라미드의 경우 수성구가 가장 특이한 모양을 보여주었는데, 이를 통하여 수성구에서 인구이동에서 세대간의 관련성이 다른 지역에 비하여 더 뚜렷하게 나타남을 확인하였다. 순이동률의 공간적 패턴은 연령그룹별로 상이한 패턴을 보였고, 30년 간격을 두고 유사한 패턴이 나타남을 확인하였다. 세대간의 연결성은 연령그룹별 순이동률에 대한 상관관계 분석에서도 확인되었다. 전체적으로 순이동률은 공간적 자기상관을 보이지만, 이러한 경향은 연령그룹별로 상이하게 나타났다. 0~4세, 5~9세, 10~14세, 25~29세, 30~34세, 45~39세, 40~44

세 연령그룹의 공간적 자기상관 측정치는 통계적으로 유의한 반면 나머지 연령그룹은 그렇지 않았다. AMOEBA를 이용한 공간 군집 분석 결과, 연령그룹별 인구이동에서의 상용관계를 공간적으로도 확인하였다. 군집 분석 결과 흥미로운 패턴은 수성구를 중심으로 5~9세, 10~14세에서 높은 값의 공간 군집이 형성되었다가 이후 연령 그룹에서는 낮은 값의 공간 군집으로 바뀌었다. 이 지역에서 높은 값의 공간 군집은 35~39세, 40~44세 연령그룹에서 다시 나타나 연령별 인구이동의 연계성이 공간적으로도 나타남을 확인하였다.

본 연구에서 제시한 연령별 인구이동에 대한 탐색적 기법들은 인구이동 현상에 대한 이해를 풍부하게 하는데 기여할 것으로 기대된다. 제시된 분석기법을 이용하여 탐색된 새로운 패턴은 인구이동을 설명하는 모델링 과정에 그리고 지역의 인구이동과 관련된 구체적인 대안을 수립하는데 유용한 정보를 제공할 것이다. 그러나 본 연구는 ESDA에 기초하여 연령별 인구이동 구조와 공간적 패턴을 탐색하고 기술하는데 초점을 두었다. 따라서 ESDA가 인구이동에 대한 가설을 수립하고 인구이동의 원인을 설명하는 모델을 만드는 데 어떻게 기여하는지를 보여주지 못한 한계를 갖는다.

문헌

- 권상철, 2005, 우리나라 수도권으로의 인구이동: 시기별 유출지역 특성과 이주자 선별성의 상대적 중요도 평가, 한국지역지리학회지, 11(6), 571-584.
- 이상일, 2008, 거주지 분화에 대한 공간통계학적 접근(II): 국지적 공간 분리성 측도를 이용한 탐색적 공간데이터 분석, 대한지리학회지, 43(1), 134-153.
- 이상일·조대현·손학기·채미옥, 2010, 공간 클러스터의 범역 설정을 위한 GIS-기반 방법론 연구 - 수정 AMOEBA 기법-, 대한지리학회지, 45(4), 502-520.
- 이희연, 2005, 인구학: 인구의 지리학적 이해, 법문사, 서울.
- 최은영, 2004, 선택적 인구이동과 공간적 불평등의 심화: 수도권을 중심으로, 한국도시지리학회

- 지, 7(2), 57-69.
- 최은영·구동화·조순기, 2010, 부산 대도시권의 인구이동(II): 이동 가구의 선택성, 한국지역지리학회지, 16(2), 123-136.
- 최진호, 2008, 한국 지역 간 인구이동의 선별성과 이동 이유: 수도권을 중심으로, 한국인구학, 31(3), 159-178.
- Aldstadt, J. and Getis, A., 2006, Using AMOEBA to create a spatial weights matrix and identify spatial clusters, *Geographical Analysis*, 38(4), 327-343.
- Anselin, L., 1995, Local indicators of spatial association-LISA, *Geographical Analysis*, 27(2), 93-115.
- Anselin, L., 1999, Interactive techniques and exploratory spatial data analysis, In P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind (eds.), *Geographical Information Systems: Principles, Techniques, Management and Applications*, Wiley, New York, 251-64.
- Clark, W. A. V. and Onaka, J., 1983, Life cycle and housing adjustment as explanations of residential mobility, *Urban Studies*, 20(1), 47-57.
- Coll, A. G. and Stillwell, J., 1999, Inter-Provincial migration in Spain: temporal trends and age-specific patterns, *International Journal of Population Geography*, 5, 97-115.
- Haining, R., Wise, S., and Ma, J., 1998, Exploratory spatial data analysis in a Geographic Information System environment, *Journal of the Royal Statistical Society. Series D*, 47(3), 457-469.
- Johnson, K. M., Voss, P. R., Hammer, R. B., Fuguitt, G. V., and McNiven, S., 2005, Temporal and spatial variation in age-specific net migration in the United States, *Demography*, 42(4), 791-812.
- NIST Engineering Statistics Handbook <http://www.itl.nist.gov/div898/handbook/>
- Ord, J. K. and Getis, A., 1995, Local spatial autocorrelation statistics: distributional issues and an application, *Geographical Analysis*, 27(4), 286-306.
- Peters, G. L. and Larkin, R. P., 2005, *Population Geography: Problems, Concepts, and Prospects*, 8th Edition, Kendall/Hunt, Dubuque, Iowa.
- Pittenger, D. B., 1974, A Typology of age-specific net migration rate distributions, *Journal of the American Planning Association*, 40(4), 278-283.
- Plane, D. A. and Heins, F., 2003, Age articulation of U. S. inter-metropolitan migration flows, *The Annals of Regional Science*, 37, 107-130.
- Ravenstein, E. G., 1885, The Laws of Migration, *Journal of the Statistical Society of London*, 48(2), 167-235.
- Rogers, A., Raquillet, R., and Castro, L. J., 1978, Model migration schedules and their applications, *Environment and Planning A*, 10(5), 475-502.

(접수: 2010.8.11, 수정: 2010.9.14, 채택: 2010.10.10)