지역거주 의향에 미치는 영향요인들의 조건결합*
- 의사결정나무분석 활용 -

Identifying Combinations of the Conditions Affecting the Residents' Intention to Reside in Their Local Area
- Application of Decision Tree Analysis Method -

최 영 출 (충북대학교 행정학과 교수)

Abstract

Young-Chool Choi

This paper aims to identify combinations of the variables associated with local residents' intention to reside in their local area, to discover which domain of the variables affects their intentions most, and to analyze the characteristics underlying the higher intention groups and the lower intention groups. To this end, this paper employs decision tree analysis method which has rarely been used in the field of population mobility studies. The conclusions of this study are summarized as follows. First, it is found that among three domains of variables associated with local residents' intention to reside in, the domain of policy-related variables has been proved to be the strongest. Second, not only central government but local governments as well need to know multiple combinations of the variables affecting residents' intention to reside in their local area and to reflect them into policy-making processes concerned.

Keywords: residents' intention to reside in, decision tree analysis, combinations of the conditions affecting residents' intention to reside in

I. 서 론

우리나라 뿐만 아니라 많은 국가들의 경우, 오늘날 겪고 있는 지역문제 중의 하나는 지역내의 인구감소라고 할 수 있다. 저 출산문제가 날로 심각해 지고 있고, 동시에 지역 내의 인구가 타 지역으로 이주함에 따라 인구가 감소되는 지역이 늘어나고 있다. 인구가 감소되면 지역 내 산업이 쇠퇴할 뿐만 아니라 지역을 유지시켜 주는 자원이 감소되고 이는 각종 주민 지원 서비스의 감소로 이어지게 된다. 이러한 지원서비스 감소는 주민들의 거주의향을 감소시키고, 타

* 이 논문은 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2016S1A3A2924563).

지역으로의 이주를 촉진시키게 되고 이는 다시 지역 내 인구감소를 가중시키게 되는 악순환을 가져오게 된다. 지역내의 인구감소는 다시 자치단체로서의 존립문제에까지 영향을 미치는 것으로서 실로 중요한 문제가 되고 있다. 따라서 오늘날 중앙정부나, 광역 및 기초 등 모든 지방정부에서는 국가전체적 차원에서 인구감소를 우려하며 저 출산대책에 많은 정책적 관심을 가지고 있다. 이와 동시에 지역내에서는 저 출산 문제에 더하여 지역 내의 인구를 지키고 타 지역 인구를 유입시키는데 더욱 많은 관심을 가지고 있다. 지역 내의 인구를 지키고, 타 지역 인구를 유입시키기 위해서는 단순한 접근방법으로는 어렵다. 지역에의 거주지 선택에는 다양한 요인들이 결합되어 있다. 지역 내 주민들의 거주의향이 무엇인지를 파악하고 이러한 거주의향을 파악하여 상응하는 정책방향을 제시하는 것은 무엇보다도 중요한 과제가 되고 있다. 그러나 주민들의 거주의향에 관련된 기존의 연구들은 대부분 개인적 배경들이 지역 내 거주의향에 어느 정도 영향을 미치는 것인지에 초점을 두고 있어서 설령 이러한 정보를 얻는다고 하여도 중앙정부나 지방자치단체 입장에서는 제한적인 용도로 밖에 활용할 수밖에 없다. 이러한 문제인식하에서 본 연구는 우리나라 지역 주민들 2,700명을 대상으로 한 설문지 조사결과를 바탕으로 하여 지역내 거주의향에 영향을 미치는 영향요인들의 다양한 결합조건들을 도출해 냄으로써 거주의향의 핵심적 요소들을 파악하고, 이를 통하여 중앙정부나 지방자치단체에게 지역 내 인구유출 방지 및 나아가서는 인구유입 대책을 수립하는 데 필요한 정보를 제공하는 것을 목적으로 한다.

Ⅱ. 이론적 논의와 연구문제

1. 이론적 논의

주민들이 자기가 살고 있는 지역에 계속 거주하고자 하는 거주의향에는 많은 요인들이 영향을 미친다. 기존의 선행연구들이 (이석원외, 2017; 정연택. 김재태. 2014; Zhao et al., 2019; Kim et al., 2005; Pagliara and Preston, 2003; Brasington, 2002; Cirillo, 2000; Champion, 2002; Cheshire and Sheppard, 2004; Knox and Pinch. 2000; Meen and Meen, 2003) 거주의향 요인으로 제시하고 있는 영향요인들을 보면 다양하다. 연령, 소득수준, 결혼상태, 학력, 지역의 수준, 직장까지의 교통소요 시간, 직장까지의 교통 비용, 슈퍼마켓까지의 교통 비용, 인구밀도, 주거지역, 주거지 인근 학교 수준, 지역의 인구규모, 주택가격, 성별, 함께 사는 어린이 수 등 다양하다. 이러한 요인들은 크게 보아서 개인적 요인들과 주거지역의 요인으로 구분된다. 성별이나, 연령, 소득수준, 결혼상태, 학력과 같은 것들은 개인적 요인으로 구분될 수 있고, 직장까지의 거리나 교통비용, 주택가격, 지역의 인구규모 등과 같은 것들은 주거지역 요인으로

구분될 수 있다. 이러한 연구들은 주로 개인적 배경 요인들과 해당 지역의 특성 요인들이 주민들의 거주의향과 직접, 간접으로 영향관계를 맺고 있는 것으로 가정한다. 그러나 최근에 들어와서 삶의 만족이나 행복을 연구하는 학자들은 (이승종외, 2013; 이지혜. 최영출, 2014; 노영숙. 장지현, 2018; 서재호외, 2012; 이승종외, 2014; 이상엽. 정건섭, 2011; 주상현, 2012; 최영출, 2017; 최영출, 2019; Diaz-Morales et al., 2013; Etzioni, 2018; Magee et al., 2013; Naci & Ioannidis, 2015; Lachmann et al., 2018; Ehrhardt, 2000; Western & Tomaszewski, 2016; Michalos, 2008; Graafland & Compen, 2015) 삶의 만족요인이나 행복결정요인들을 설명하면서, 이러한 요인들이 잘 갖추어진 지역들의 경우 거주의향이 높고 그렇지 않은 지역일수록 주민들의 거주의향이 낮다는 연구결과를 많이 제시하고 있다. 이러한 연구들은 인간이 궁극적으로 추구하는 가치라고 할 수 있는 행복이나 웰빙, 삶의 만족과 가치들에 영향을 미치는 요인들을 다차원적으로 제시하고 이러한 요인들이 주민들의 지역 거주의향에 영향을 미치는 것을 강조하고 있다. 이러한 연구들은 정책적 관점에서 주민들의 거주의향이나 삶의 만족을 검토하는 연구라고 할 수 있다. 즉, 주민들의 삶의 분야별로 자원배분이나 정책적 고려를 통하여 주민들의 삶의 만족도를 충족시켜 줌으로써 행복도 증진 및 거주의향정도를 높여주고자 하는데 관심을 두고 있는 것이 특징이다. 이러한 부류의 연구를 하는 국내 연구로는 대표적으로 이승종외 (2013),[1] 서재호 외 (2012), 이승종외 (2016) 등의 연구를 들 수 있다. 이들은 커뮤니티웰빙이라는 용어를 사용하여 지역사회 수준의 웰빙에 영향을 미치는 요인들로서 자본 (capital)개념을 들어서 설명하고 있다. 이들의 연구내용을 보면 (이승종외, 2013; 이승종외 2016) 커뮤니티웰빙에 영향을 미치는 자본으로서 인본적 자본 (human capital), 경제적 자본 (economic capital), 자연자본 (natural capital), 인프라 자본 (infra capital), 문화자본 (cultural capital), 사회자본 (social capital) 등 6개 자본으로 구분하고 있다. 각 자본에는 3개에서 8개까지의 지표를 포함한다. 가령, 인본적 자본에는 교육, 복지등 분야의 8개 지표, 경제자본에는 지역의 재정자립도, 고용수준 등 6개지표, 문화자본에는 지역의 문화프로그램 등 3개지표, 자연자본에는 공기오염도 등 6개지표, 인프라 자본에는 도로상태 등 6개 지표, 사회자본에는 지역사회 주민들의 유대관계 등 6개 지표가 포함되어 있다. 자본개념에 바탕을 두고 커뮤니티 웰빙 수준을 측정하고자 하는 이러한 연구들의 특징은 지역사회 수준의 웰빙수준을 측정하는 데 있어서 개인적, 지역적 요인 뿐만 아니라 정책적 요인들을 고려하여 다차원적으로 상대적 중요성 등을 확인하는 장점이 있다고 하겠다. 이러한 연구의 범주에 들어가는 또 다른 연구로는 오미영 (2008), 이경영외 (2018), 이승종외 (2016)의 연구들이 있다. 이들의 연구도 기본적으로는 주민들의 거주의향에는 정주여건, 지방자치단체와의 관계, 사회적 자본 등 정책적 요소들이 영향을 미친다고 보는 관점들이다. 본 연구에서는 주민들의 거주의향에는 개인적 요인들, 지역적 요인들 뿐만 아니라 주민들의 삶의 수준에 영향을 미치는 여러 영역들에 대한 정책적 요인들도 중

[1] 이 연구팀은 한국연구재단의 SSK연구사업을 실시하고 있으며, 매 2년마다 전국단위 서베이를 하고 있음. 본 논문은 이 연구팀에서 2017년에 실시한 원자료를 활용하도록 허용받았음.

요하다고 생각한다. 특히, 지방자치가 실시된 이후 복지, 교육, 지역경제, 자연재해, 사회자본 등 여러 영역들에 걸쳐서 자치단체간 차별적 노력들이 나타나고 있다. 그 결과, 자치단체장의 성향이나 지방의회 의원들의 정치적 성향 등에 따라서 특정 분야에 대한 재정적, 제도적 자원 배분이 강화되거나 축소되는 등의 가시적 결과들이 나타나고 있다. 이러한 정책적 요인들이 주민 삶에 미치는 정도가 크다고 볼 수 있기 때문에 이러한 요인들도 거주의향에 크게 영향을 미치는 것으로 가정할 수 있다. 따라서 본 연구에서는 기존의 연구들과는 달리, 거주의향에 영향을 미치는 요인들로서, 주민들의 개인적 배경요인, 지역적 특성 요인, 정책적 요인들을 모두 포함시키기로 한다. 이렇게 함으로써 주민들의 지역내 거주의향에 영향을 미치는 요인들을 다차원적으로 파악할 수 있다고 생각한다.

2. 연구문제

본 연구에서는 위에서 논의한 이론적 배경과 문제인식을 바탕으로 하여 아래와 같은 연구문제를 선정한다.

〈연구문제 1〉: 우리나라 지역주민들이 현재 살고 있는 지역의 거주의향에 미치는 중요한 독립변수들은 무엇인가? 개인적 특성요인, 지역적 특성요인, 정책적 특성요인 중 어느 특성요인중 가장 크게 영향을 미치는 요인은 무엇인가?

〈연구문제 2〉: 우리나라 지역주민들이 현재 살고 있는 지역의 높은 거주의향에 미치는 결합조건들은 무엇인가?

〈연구문제 3〉: 우리나라 지역주민들이 현재 살고 있는 지역의 낮은 거주의향에 미치는 결합조건들은 무엇인가? 등이다.

Ⅲ. 조사설계

1. 연구대상

본 연구에서의 응답 대상은 우리나라 17개 시도의 27개 시군구 지역에 거주하는 지역의 주민들 2,700명으로 하였다. 전국 지역 주민들을 대량으로 샘플링 하여 주민들의 만족도나 거주의향과 같은 연구를 정기적으로 수행하기는 용이하지 않다. 서울대학교 행정대학원의 커뮤니티웰빙 연구센터에서는 한국연구재단의 SSK 사업을 수행하는 과정에서 한국갤럽에 의뢰하여 우리나라 시도의 시군구 27개 지역을 선정하여 설문지를 수집하여 매 2년마다 데이터베이스화하고 있다.[2] 샘플링 방법은 층화추출과 집락추출을 결합하였다. 이 과정에서 인구규모, 지역특성 등을 기준으로 27개 시군구가 선정되었고, 각 시군구별로 100명씩의 주민을 추출하여 이

들을 대상으로 설문지를 배포하고 수집한바 있다. 가장 최근의 자료 수집은 2017년 10월부터 11월까지 실시되었으며, 여기서 수집된 자료를 연구진의 동의를 얻어서 활용한다.

2. 연구방법

본 연구의 목적을 달성하기 위해서 가능한 연구방법은 다양하다. 본 연구에서는 위에서 설정한 연구의 목적을 효과적으로 달성하기 위하여 의사결정나무 분석 (decision tree analysis)을 활용한다. 본 연구의 연구문제들은 3가지로서 현재 주민들이 살고 있는 지역에 계속 거주의사가 있는지 즉 거주의향과 관련된다. 거주의향을 1점에서 10점까지 나누어서 거주의향이 낮은 항목부터 높은 항목까지 10개로 등급화한 항목에 표시를 하게 하고 있다. 이 거주의향이 종속변수가 된다. 이 종속변수에 영향을 미치는 다양한 독립변수들을 선정하여 의사결정나무 분석을 실시한다. 이는 지역에의 거주의향에 영향을 미치는 다양한 결합조건들을 도출하여 중앙정부나 지방정부가 자원배분 및 정책결정과정에서 주민들의 거주의향을 높이고 타 지역으로의 이주를 예방하는 데 필요한 의미있는 정보를 제공하기 위한 것이다. 물론 이러한 결합조건 분석에 있어서 QCA (Qualitative Comparative Analysis)나 fsQCA (fuzzy-set Qualitative Comparative Analysis) 방법을 사용할 수 있다. 그러나 QCA나 fsQCA의 경우에는 사례 수가 적은 경우 (5<=사례수<=55)에 일반적으로 사용된다.[3] 아울러, QCA 계열 방법의 경우, 다양한 독립변수들의 결합조건을 제시해 주기는 하나 독립변수들의 중요도를 제시해 주지 못한다. 그 반면에 의사결정나무 분석은 다양한 결합조건들 뿐만 아니라 분석에 포함된 변수들의 상대적인 중요도를 제시해 주기 때문에 더 많은 정보를 제공해 준다. 의사결정나무 분석은 데이터마이닝을 위한 방법 중의 하나로서 바로 이러한 장점 때문에 본 연구목적 달성을 위해서 효과적인 방법이라고 생각한다. 의사결정 나무분석의 특징을 정리해 본다면 첫째, 의사결정나무 분석은 개체들을 유사한 형태를 보이는 소그룹의 일원으로 분류해 주며, 일정한 기준에 따른 층으로 나누어 준다 (허명회, 2013). 이렇게 분류된 소그룹의 특성을 파악하게 되면 중앙정부나 지방정부는 이러한 정보를 바탕으로 삶의 만족수준 향상을 위한 중점지역 선정이나 차등적 관리를 하는데 도움을 줄 수 있다. 둘째, 많은 설명변수들 중에서 활용가치가 있는 소수의 변수들을 용이한 방법으로 단계적으로 선별해내는데 도움을 준다 (손지은. 김성범, 2013; 김옥숙, 2011; 김희진, 2002; 최영출, 2015). 회귀분석과 같은 방법들도 이러한 목적 달성에 유용하나, 의사결정나무 분석은 성격이 서로 다른 변수들도 한 번에 고려하여 다양한 결합조건들을 파악하게 해 줌으로써, 유의미한 변수들의 분류에 대단히 유용하다. 셋째, 회귀분석과 같은 방법과는 달리 독립변수들의 특정조합이 갖는 효과를 식별해 냄으로서 모수적 모형

---

2) 전술한 바와 같이, 본 연구는 서울대 SSK 연구단의 자료 활용에 대한 동의를 얻어 원자료를 통하여 수행하였음.

3) 사례수가 적은 경우 질적 방법으로는 근거이론 같은 방법을 사용하기도 한다 (유호웅외, 2018). 또 최근에는 시뮬레이션 방법들을 통하여 질적 측면을 양화하기 위한 시도도 다양하게 이루어지고 있다 (장지현, 2017).

에서 어떤 변수 간 상호작용을 고려해야 하는지에 대한 지침을 제공해 준다. 즉, 회귀분석에서는 독립변수들 중에서 통계적으로 유의미한 변수들만 회귀식에서 이용될 수 있으나 (김순귀외, 2000) 의사결정나무 분석에서는 통계적으로 유의미한 변수가 아니더라도 다른 변수와 결합이 되면 종속변수에 유의미한 영향을 미칠 수 있다는 점을 반영해 준다. 이러한 이유로 본 연구에서는 지역에의 계속 거주의향을 종속변수로 하고 여기에 영향을 미칠 수 있는 다양한 독립변수들의 특정조합이 갖는 효과를 식별해 보기 위하여 의사결정나무 분석을 활용하고자 한다.

3. 변수선정 및 측정방법

본 연구에서는 종속변수를 지역에의 계속 거주의향 수준으로 하고, 독립변수는 선행연구들의 결과를 바탕으로 하여 다음과 같이 개인적 배경요인, 지역적 특성요인, 그리고 정책적 특성요인 등으로 구분하여 선정한다. 따라서, 성별, 연령, 소득수준, 학력, 일자리 형태, 결혼상태 등 개인적 배경요인들이 포함된다. 지역적 특성으로는 지역유형 (대도시, 중소도시, 농촌), 지역의 인구규모, 지역의 재정상태가 포함된다. 지역관련 정책적 요인[4]으로서는 인본적 자본 (8개 지표), 경제적 자본 (4개 지표), 문화적 자본 (3개 지표), 자연 자본 (6개 지표), 인프라 자본 (6개 지표), 그리고 사회자본 (6개 지표) 등 6개 자본 영역으로 구분하고 각 영역별 지표들의 평균값을 활용한다. 다만, 본 연구에서는 선행연구에서 일부 발견되는 언론의 자유, 사회의 불균형 정도와 같은 거버넌스적 변수들과, 개인들의 정서적 측면에 관한 변수들은 제외한다. 이는 중앙정부나 지방정부 입장에서 볼 때, 지역주민들의 정서 상태를 파악하기 쉽지 않다는 점 및 언론의 자유 같은 거버넌스적 측면은 지역별 차등화가 용이하지 않다는 점 때문이다. 중앙이나 지방의 정책결정자 입장에서 필요한 것은 오히려 지역적 특성과 지역주민들의 배경 및 주민삶에 영향을 미치는 주요 정책분야별 정보들이 오히려 정책과정에서 용이하게 활용할 수 있다는 장점이 있다. 아래와 같은 변수들을 선정하고 일정한 기준에 따라 측정하였다.

〈표 1〉 변수 및 측정방법

영역	변수/영역	측정방법	비고
지역	지역유형	특,광역시의 자치구, 중소도시, 군 등 3개 분류	특광역시의 자치구 1, 중소도시 2, 농촌군 3
	인구규모	인구수(명)	2017년 기준
	재정상태	재정자립도(%)	2017년 기준
개인	성별	남자, 여자	
	연령	연령	비율변수로 처리
	소득	9개 카테고리로 구분	서열변수로 처리

4) 본 연구에서는 설문지에서 표현된 용어들과의 일관성을 감안하여 기 사용된 용어들을 그대로 사용하기로 함.

	결혼상태	기혼, 미혼, 사별 또는 이혼	명목변수로 처리
	학력	중졸이하, 고졸이하, 대졸, 대학원 이상	서열변수로 처리
	일자리 형태	정규직, 비정규직, 프리랜스, 자영업	명목변수로 처리
정책 분야	인본적 자본 (human capital)	공공도서관, 평생교육, 지역내 학교수준, 건강수준, 의료, 노인, 장애인, 보육 등 8개	매우 불만 1, 매우 만족 10 (평균점수)
	경제 자본 (economics capital)	지역내 일자리, 자치단체 살림살이, 지역경제활성화, 생활비 등 4개	매우 불만 1, 매우 만족 10 (평균점수)
	문화자본 (cultural capital)	문화활동 수준, 지역문화시설, 지역고유의 전통/문화/가치 등 3개	매우 불만 1, 매우 만족 10 (평균점수)
	자연자본 (natural capital)	공기오염도, 공원녹지공간, 지역내 하천, 재활용 및 쓰레기수거, 거리청결도, 주변소음 등 6개	매우 불만 1, 매우 만족 10 (평균점수)
	인프라 자본 (infra capital)	대중교통, 도로유지수준, 교통체증, 재난대비, 인도의 보행자편의, 치안 및 경찰 등 6개	매우 불만 1, 매우 만족 10 (평균점수)
	사회자본 (social capital)	지역사회단체활동, 지역사회 협력활동, 지역주민 신뢰, 자치단체신뢰, 중앙정부신뢰, 지역신문/TV 신뢰 등 6개	매우 불만 1, 매우 만족 10 (평균점수)
종속 변수	거주의향	지역내에서의 계속 거주의향	최하 1점, 최상 10점 척도로 하여 평균 점수 이하를 낮음, 평균점수 이상을 높음으로 2등분

4. 모형의 평가

과학적 연구절차에 의해 양적 분석을 하는 경우, 분석을 위한 모형에 대한 평가가 이루어져야 한다. 모형에 대한 평가는 적용한 방법론에 따라 다소 다르겠지만 일반적으로는 모형의 유의성 평가 -> 독립변수의 유의성 평가-> 모형의 적합도 평가->모형 해석 등의 순서로 수행된다 (조용준, 2018). 먼저, 모형의 유의성 검증부터 살펴 본다. 회귀분석, 판별분석, 군집분석, 로지스틱 회귀분석 등의 경우에는 산출된 모형 (함수식 등)이 통계적으로 유의미한지 일정한 기준에 의하여 검증하고 유의하지 않으면 이후의 단계는 진행하지 않고 새로운 변수 투입이나 제거 과정을 거쳐서 새로운 유의미한 모형을 찾는 것이 일반적이다. 그러나 의사결정나무 분석과 같은 데이터마이닝 알고리즘에서는 이러한 방법으로 모형의 유의성을 검증하지 않는다 (조용준, 2018). 의사결정나무 분석의 경우에는 의사결정나무 모형을 통하여 노드를 분리시키는 노드분할 기준과 범주합치기 기준을 0.05 또는 0.01로 사전에 설정하여 이 기준을 통과하여 나무모형이 구축되는 경우 모형의 유의성이 있다고 할 수 있다. 따라서 의사결정 나무분석 모형의 경우에는 회귀분석 모형과 같은 유의성 검증절차와는 다르다. 두 번째는 독립변수의 유의성 평가문제이다. 이는 종속변수에 영향을 미친다고 가정하는 영향변수가 통계적으로 유의한지를 개별적으로 파악하는 단계이다. 이 과정도 일반적인 통계분석과 데이터마이닝의 방법은 다르다 (조용준, 2018). 의사결정나무 분석의 경우에는 투입된 독립변수들 중에서 나무모형에 포함된 변

수들은 통계적으로 유의성이 있다고 가정하기 때문에 일반적 통계분석과는 달리 산출된 독립변수의 중요도를 검토하게 된다. 세 번째는 모형의 적합도 평가 문제이다. 모형의 적합도 평가는 산출된 모형이 얼마나 정확한지 즉 예측정확도를 평가하는 것이다. 예측하고자 하는 종속변수가 범주형일 때에는 분류나 판별이 목적이기 때문에 설정한 모형이 실제값을 정확히 예측하고 실제값과 같이 분류하였는 지 (예: 양호, 불량 또는 성공, 실패 등)명확히 셀 수 있으므로 예측정확도로 평가가 가능하다. 본 연구의 경우, 주민들의 거주의향을 10점 척도로 하여 (계속 거주의향이 매우 낮음은 1, 매우 높음은 10) 구한 다음 전체 평균 점수를 (7.22점) 구하여 평균점수 이하는 낮음으로 하고, 평균점수 이상은 높음으로 하여 2개로 구분하였다. 즉 범주형 변수로 처리하였다. 따라서 모형의 적합도 평가는 카이제곱 통계량에 의하여 수행된다.

의사결정나무분석에서 나무를 키우는 방법, 즉 분리기준은 CHAID, 포괄적 CHAID, CRT, QUEST 등이 있다 (허명회, 2013; 조용준, 2018). 이 중에서 본 연구는 CRT 기준을 채택한다. CRT는 종속변수가 범주형일 때 주로 사용하며, 다른 기준에 비하여 자식노드내 종속변수값들이 최대한 동질적이 되도록 부모노드를 분리한다. 이를 통하여 노드내 동질성을 극대화해 주는 장점이 있다. 본 연구에서 거주의향은 '낮음', '높음' 2개 범주형으로 구분하였고 노드내의 동질성 극대화를 위하여 CRT를 활용한다.

Ⅳ. 분석결과

1. 응답자의 일반적 배경

본 연구에서 활용한 2,700명 응답자들의 일반적 배경을 요약해 보면 아래 표에서 보는 바와 같다. 응답자 배경 (2)의 경우에는 변수들이 명목척도가 아니기 때문에 기술통계값으로 별도 처리하도록 하였다. 전체 응답자 중에서 지역특성별로 보면, 대도시 주민들이 1100명으로 약 40%를 차지하고, 중소도시는 692명으로서 약 26%, 그리고 농촌지역은 908명으로서 약 34%를 차지한다. 성별로는 남성이 1339명으로서 49.6%이고, 여성이 1361명으로서 50.4%로서 비슷한 비율로 나타나고 있다.

〈표 2〉 응답자의 일반적 배경(1)

		N	퍼센트	유효퍼센트	누적퍼센트/(표준편차)
지역특성	대도시	1100	40.7	40.7	40.7
	중소도시	692	25.6	25.6	66.4
	농촌지역	908	33.6	33.6	100

	합계	2700	100	100	
성별	남성	1339	49.6	49.6	49.6
	여성	1361	50.4	50.4	100
	합계	2700	100	100	
나이	합계	19	88	48.81	(15.236)
혼인 상태	미혼	522	19.3	19.4	19.4
	기혼 (사실혼, 주말부부, 배우자의 요양)	2063	76.4	76.6	96
	이혼,별거,사별 등	109	4.0	4.0	100
	합계	2694	99.8	100	
	무응답	6	2		
	합계	2700	100		
학력	초등학교 졸업 이하	202	7.5	7.5	7.5
	중학교 졸업	292	10.8	10.9	18.4
	고등학교 졸업	1144	42.4	42.5	60.9
	전문대학 및 대학교 졸업	1011	37.4	37.6	98.5
	대학원 이상	40	1.5	1.5	100
	합계	2689	99.6	100	
	무응답	11	4		
	합계	2700	100		
경제 활동	예	1931	71.5	71.5	71.5
	아니오	769	28.5	100	
	합계	2700	100.0		
일자리 형태	정규직 임금근로자	918	34	47.5	47.5
	비정규직 임금근로자	310	11.5	16.1	63.6
	자영업자	633	23.4	32.8	96.4
	무급가족 종사자	42	1.6	2.2	98.5
	기타 종사자 (프리랜서, 보험설계사 등)	28	1.0	1.5	100.0
	합계	1931	71.5	100	
	결측값	769	28.5		
	합계	2700	100.0		
월 가구 소득	0-99만원	108	4.0	4.0	4.0
	100-199만원	317	11.7	11.8	15.8
	200-299만원	441	16.3	16.4	32.2
	300-399만원	625	23.1	23.2	55.4
	400-499만원	540	20.0	20.1	75.5
	500-599만원	394	14.6	14.7	90.2
	600-699만원	171	6.3	6.4	96.5
	700-799만원	48	1.8	1.8	98.3
	800-899만원	18	0.7	0.7	99.0
	900-999만원	15	0.6	0.6	99.6
	1000만원이상	12	0.4	0.4	100.0.
	합계	2689	99.6	100	
	무응답	11	4		
	합계	2700	100		

〈표 3〉 응답자의 일반적 배경(2)

	N	최솟값	최대값	평균	표준편차
재정자립도	2700	12.05	67.65	29.6174	13.87015
인구	2700	38200	555622	191453.28	145273.029
개인 human	2635	2.13	10.00	6.9105	1.13565
개인 economic	2700	1.00	9.75	6.7566	1.26733
개인 cultural	2700	1.00	10.00	6.8147	1.32827
개인 natural	2700	1.67	9.83	7.0590	0.98216
개인 infra	2700	2.00	10.00	6.9262	1.10091
개인 social	2700	1.33	10.00	6.8581	1.16977
계속적 지역거주의향	2700	1	10	7.22	1.371
유효수(목록별)	2635				

주: 개인human 등 '개인'명칭이 붙은 것은 서울대 연구진이 구성한 설문응답에서의 집단human과 구분하기 위한 것임. 코딩과정에서 '개인'명칭이 붙어서 처리되었기 때문에 그대로 사용하기로 함.

정책적 요소라고 할 수 있는 인본적 자본의 경우에는 평균값이 6.91이며, 경제적 자본은 6.75 수준이다. 문화적 자본은 6.81이고 자연적 자본은 가장 높은 7.05이다. 인프라 자본과 사회적 자본은 각각 6.92와 6.85이다. 종속변수인 계속적 거주의향의 평균 점수는 7.22로 나타나고 있다.

2. 의사결정 나무분석 결과

1) 전반적 경향

우리나라 지역 주민 2,700명을 대상으로 하여 현재 살고 있는 지역에 대한 계속 거주의향 영향요인을 확인하기 위하여 의사결정나무 분석을 수행하였다. 전체 나무구조 중 상위마디의 결정 변수일수록 분리하는데 있어서 중요하다는 의미이다. 아래 그림은 지역주민들의 거주의향을 결정짓는 전체 나무구조이다. 여기에서 거주의향의 구분에 있어서 전체 평균값 7.22를 기준으로 하여 평균값 보다 낮으면 '낮음'으로 높으면 '높음'으로 구분된 것이다.

〈그림 1〉 거주의향 수준 의사결정나무 모형

위의 그림을 보면 노드 0(뿌리노드, root node)는 총 2,700명의 사례로 구성되어 있으며 이 중 거주의향 낮음이 1,482 명으로서 54.9%이고 높음이 1,218 명이며 45.1%를 차지하고 있다. 뿌리노드는 인본적 자본 (개인human)에 의하여 노드 1과 노드 2로 분리된다. 노드 1에는 인본적 자본이 6.938보다 작거나 같은 사람들로 구성되어 있고, 노드 2에는 인본적 자본이 6.938보다 큰 사람들로 구성되어 있다. 노드 1에 포함되어 있는 1,210명 중에서 거주의향이 낮음이 856명으로서 70.7%이고 높음이 364명으로서 29.3%를 차지하고 있다. 노드 1은 다시 자연적

자본(개인natural) 에 의하여 노드 3과 노드 4로 구분되는 데 노드 3인 경우에는 자연적 자본이 5.750보다 작거나 같은 사람들 229명으로 구성되어 있다. 이 노드 3은 더 이상 분리되지 않고 최종 거주의향이 '낮음'노드로 확정되어 있다. 노드 4는 자연적 자본이 5.750 이상인 사람들로서 다시 지역특성에 의하여 분리되어 지고 있다. 이러한 과정을 거쳐서 전체모형 중 최종 끝마디 노드는 19개로 나타나 있다.

한편, 전체 모형의 유의성 평가를 보면, 노드를 분리시키는 노드분할 기준과 범주합치기 기준을 본 연구에서는 0.05 로 사전에 설정하여 이 기준을 통과하여 나무모형이 구축되었기 때문에 유의수준 0.05에서 나무모형의 유의성이 있다고 할 수 있다. 또 독립변수의 유의성은 회귀분석과는 달리 상대적 중요도가 제시된다. 이는 후술하는 바와 같다. 모형의 예측력 즉 적합도를 보면, 69.3% 수준으로서 낮지 않은 수준이라고 할 수 있다. 아래 표에서 보는 바와 같이, 리스크 추정치는 30.9%이다.

〈표 4〉 모형의 위험도 (오분류율)

추정값	표준오차
.309	.009

주: 성장방법:CRT
종속변수: 평균 7.22 이하는 0, 이상은 1

위의 표의 리스크 산출 근거가 되는 교차분류표를 보면, 아래 표에서 보는 바와 같다. 행은 종속변수에 대하여 실제 관측된 변수이고 열은 예측 범주이다. 아래 표는 총 2,700명의 사례를 총 4가지 (2*2) 로 분류한 표이다. 실제 '낮음' 1,482 (1108+374) 중에서 1108 사례를 옳게 분류하여 74.8%의 정분류율 (correct percent)을 기록하고 있다. 실제 '높음'인 1218명 사례에 대하여는 758명을 옳게 분류하여 62.2%의 정분류율을 기록하고 있다. 이에 따라 총 오류율은 30.9% 수준이다.

〈표 5〉 교차분류표

감시됨	예측		
	낮음	높음	정확도(%)
낮음	1108	374	74.8%
높음	460	758	62.2%
전체 퍼센트	58.1%	41.9%	69.1%

성장방법: CRT
종속변수: 평균 7.22이하는 0, 이상은 1

2) 독립변수들의 중요도

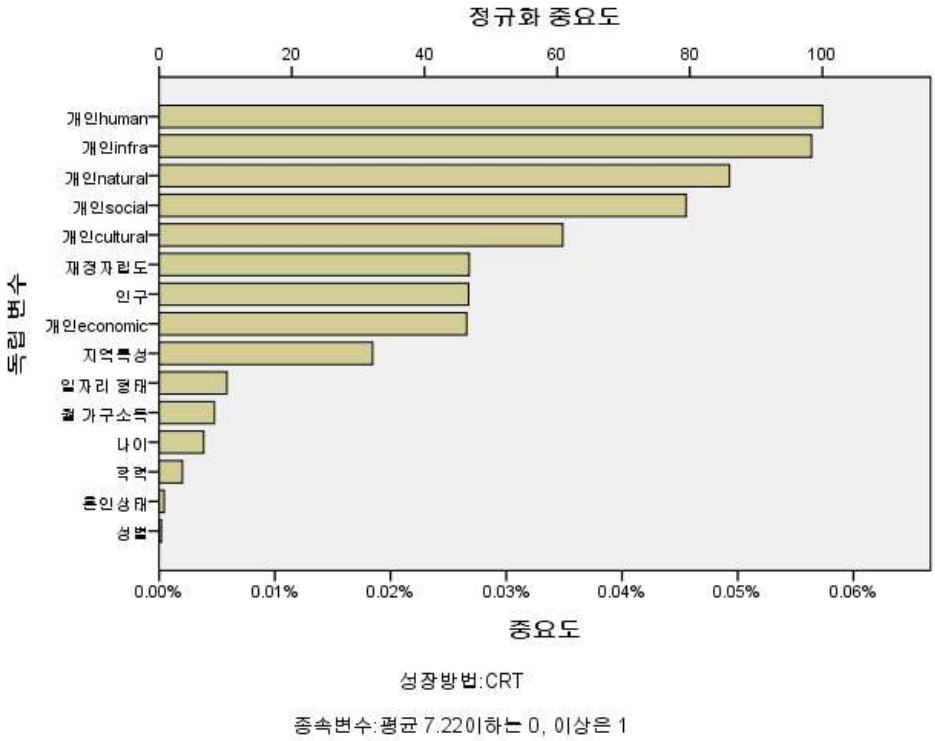
지역주민들의 거주의향에 영향을 미치는 독립변수들의 중요도를 놓고 볼 때, 인본적 자본(개인human) 변수가 가장 크게 영향을 미치는 것을 알 수 있다. 여기서 중요도값은 첫 번째 마디를 형성하는 개인human 변수의 중요도를 1로 둘 때, 나머지 변수들의 상대적인 중요도를 나타내고 있는 것이다. 표에서 보는 바와 같이, 지역주민들의 거주의향에 영향을 미치는 가장 중요한 변수는 인본적 자본 (공공도서관, 평생교육, 지역내 학교수준, 건강수준, 의료, 노인, 장애인, 보육 등 8개)으로 나타나고 있고, 두 번째 중요한 것은 인프라 자본 (개인infra)으로 나타나고 있다. 아래 표에서 보는 바와 같이, 상대적 중요도로 보면, 개인적 배경 변수나 지역적 변수들 (인구, 재정규모, 지역특성)보다도 정책적 변수들 (자본변수)이 거주의향에 더 큰 영향을 미치는 것으로 나타나고 있다.

〈표 6〉 독립변수들의 상대적 중요도

독립변수	중요도	정규화 중요도
개인human	0.057	100.0%
개인infra	0.056	98.3%
개인natural	0.049	86.0%
개인social	0.046	79.4%
개인cultural	0.035	60.9%
재정자립도	0.027	46.8%
인구	0.027	46.7%
개인economic	0.027	46.4%
지역특성	0.018	32.2%
일자리형태	0.006	10.2%
월가구소득	0.005	8.3%
나이	0.004	6.7%
학력	0.002	3.5%
혼인상태	0.000	0.8%
성별	0.000	0.3%

아래 그림은 지역주민들의 거주의향에 영향을 미치는 변수들의 상대적 중요도를 그래프로 보여 주고 있다. 아래 그림에서 보듯이 정책적 변수 6개 중에서 5개가 가장 중요한 순위에 포함되어 있음을 알 수 있다. 그 다음은 지역적 변수들로서 재정자립도, 인구, 지역특성 순이다. 개인적 배경 요인들은 상대적으로 거주의향에 덜 영향을 미치는 것으로 나타나고 있다.

〈그림 2〉 거주의향에 대한 변수들의 상대적인 중요도



주민들의 거주의향에 대한 의사결정나무 분석의 경우, 분류기의 효율성을 나타내는 개념인 이득 (gain)과 이득도표 (gain chart)를 살펴 볼 필요가 있다. 여기서 이득이란 최종노드에서 나타난 실제 '높음' 또는 '낮음'의 사례 수를 말한다. 아래 표에서 보면 노드 18에서 가장 많은 이득 (gain)이 있다. 노드 18에 포함된 사례 수는 133이며 이는 전체 응답자 2,700명 중 4.9%에 해당된다. 노드 18에서의 '높음'에 대한 이득은 124이며 이 124는 전체 '높음'에 응답한 케이스 중에서 10.2%가 노드 18에서 나온다는 의미이다. 그리고 이 노드 18에 있는 133명 중에서 '높음'에 응답한 비율은 93.2%라는 의미이다. 노드 36에서 그 다음으로 많은 이득이 있고, 그 다음은 노드 28이다. 이러한 순서로 이득의 순위가 매겨져 있다. 한편 지수값은 평균을 100으로 보았을 때의 비율값을 의미한다.

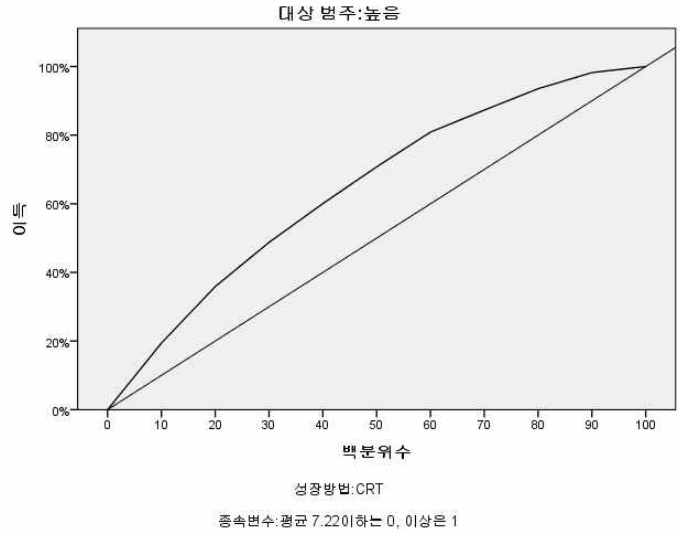
〈표 7〉 노드에 대한 이익

노드	노드별					
	노드		이득		응답	지수
	N	퍼센트	N	퍼센트		
18	133	4.9	124	10.2	93.2	206.7
36	31	1.1	26	2.1	83.9	185.9
28	188	7.0	153	12.6	81.4	180.4
31	23	0.9	18	1.5	78.3	173.5
34	168	6.2	118	9.7	70.2	155.7
25	120	4.4	75	6.2	62.5	138.5
27	57	2.1	34	2.8	59.6	132.2
35	412	15.3	210	17.2	51.0	113.0
33	29	1.1	14	1.1	48.3	107.0
29	67	2.5	32	2.6	47.8	105.9
32	122	4.5	58	4.8	47.5	105.4
26	189	7.0	89	7.3	47.1	104.4
9	56	2.1	25	2.1	44.6	99.0
24	52	1.9	19	1.6	36.5	81.0
14	505	18.7	142	11.7	28.1	62.3
30	204	7.6	52	4.3	25.5	56.5
3	229	8.5	23	1.9	10.0	22.3
23	33	1.2	3	0.2	9.1	20.2
13	82	3.0	3	0.2	3.7	8.1

주: 성장방법: CRT, 종속변수: 평균 7.22이하는 0, 이상은 1

아래 그림의 이득도표는 이득을 그래프로 설명해 주고 있다. 이 그래프에서 y 좌표값은 전체 사례의 x 퍼센트에 해당하는 사례 중 의사결정나무에 의하여 확보가 가능한 '높음'사례의 퍼센트를 말한다. 예를 들어서 이득표의 노드 18에서 전체의 4.9% 사례중 10.2의 '높음' 거주의향 사례가 나오므로 좌표점이 (4.9, 10.2)에 찍히게 된다. 다음 노드 36이 추가되어 (4.9+1.1, 10.2, 2.1), 즉 (6.0, 12.3)에 좌표점이 찍힌다. 그 다음에는 다시 노드 28이 추가되어 좌표점이 (6.0+7.0, 12.3+12.6), 즉 (13.0, 24.9)가 될 것이다. 이런 식으로 진행되며, 좌표점과 좌표점 사이는 연결된다 (조용준, 2018; 허명회, 2013). 이것은 전체 케이스의 13%에 대해서 거주의향 '높음' 판정을 함으로써 전체 '높음'케이스의 24.9%를 포착할 수 있음을 의미한다. 이런 방법으로 계속 확대해 나갈 수 있다. 이득 함수가 (0,0)과 (100, 100)을 통과하는 직선보다 위쪽에 나타날수록 거주의향 '높음'사례를 정확히 탐지해 내는 분류규칙이다. 아래 그림에서 보는 바와 같이, 지역 주민들의 거주의향 영향요인의 이득함수는 (0,0)과 (100, 100)을 통과하는 직선보다 위쪽에 나타나 있음을 알 수 있다.

〈그림 3〉 이득도표 (gain chart)



3) 거주의향이 높은 그룹

위의 〈그림 3〉 거주의향 수준 의사결정나무 모형에 나타난 정보를 바탕으로 하여 거주의향이 높은 상위그룹의 주요 특성을 살펴보기로 한다. 최종 끝 노드19개 중에서 '높음' 순위별로 상위 3개 노드의 특성을 살펴본다.

〈표 8〉 거주의향이 높은 상위 3개 그룹

노드	노드별						노드 특성
	노드		이득		응답	지수	
	N	퍼센트	N	퍼센트			
18	133	4.9	124	10.2	93.2	206.7	인본적 자본의 값이 6.938보다 높으며, 인구가 71,481 명 보다 작은 자치단체에 살고, 인프라 자본이 6.750 점보다 높고, 월가구 소득이 500 만원 이상 되는 주민 그룹
36	31	1.1	26	2.1	83.9	185.9	인본적 자본의 값이 6.938 보다 높고, 자치단체의 인구가 71,481 명 보다 높으며, 다시 인본적 자본 점수가 7.562 보다 높으나 재정자립도가 25.73% 보다 높으며, 인본적 자본 점수가 8.56보다 낮은 범위에 있는 주민 그룹
28	188	7.0	153	12.6	81.4	180.4	인본적 자본의 값이 6.938 보다 높고, 인구가 71,481 보다 작은 자치단체에 살며, 인프라 자본 점수가 6.750 보다 높고, 월 가구 소득이 400만원 보다 낮으나 문화자본 점수가 6.5보다 높은 주민 그룹

거주의향이 높은 상위 3개 그룹의 특성을 살펴 보면, 상위 3개 그룹중 가장 거주의향이 높은 그룹은 표에서 보는 바와 같이 노드 18이다. 노드 18의 특성을 보면, 133명이 포함되어 있는데 인본적 자본의 값이 6.938보다 높으며, 인구가 71,481 보다 작은 자치단체에 살고, 인프라 자본이 6.750 점보다 높고, 월가구소득이 500만원 이상 되는 주민 그룹이다. 노드 36은 인본적 자본의 값이 6.938 보다 높고, 자치단체의 인구가 71,481 명 보다 높으며, 다시 인본적 자본 점수가 7.562 보다 높으나 재정자립도가 25.73% 보다 높으며, 인본적 자본 점수가 8.56보다 낮은 범위에 있는 주민 그룹이다. 인본적 자본이 2번 이상 기준으로 포함되어 있는 것이 특징이다. 다음 노드 28은 인본적 자본의 값이 6.938 보다 높고, 인구가 71,481 보다 작은 자치단체에 살며, 인프라 자본 점수가 6.750 보다 높고, 월 가구 소득이 400만원 보다 낮으나 문화자본 점수가 6.5보다 높은 주민 그룹이다.

4) 거주의향이 낮은 그룹

지역주민들의 거주의향이 특히 낮은 하위 3개 노드의 특성을 살펴 보면 아래 표에서 보는 바와 같다. 하위 3개 그룹은 노드, 3, 23, 그리고 13이다.

〈표 9〉 거주의향이 낮은 하위 3개 그룹

노드	노드별						노드 특성
	노드		이득		응답	지수	
	N	퍼센트	N	퍼센트			
3	229	8.5	23	1.9	10.0	22.3	인본적 자본의 값이 6.938 보다 작거나 같으며, 자연 자본값이 5.750 보다 낮은 주민 그룹
23	33	1.2	3	0.2	9.1	20.2	인본적 자본의 값이 6.938 보다 작거나 같으며, 자연 자본값이 5.750보다 높으나, 지역특성상 농촌지역이며, 인본적 자본 값이 5.438 보다 낮고, 자연 자본값이 6.417 보다 낮은 주민 그룹
13	82	3.0	3	0.2	3.7	8.1	인본적 자본의 값이 6.938 보다 작거나 같으며, 자연 자본 값이 5.750 보다 높고, 지역특성상 대도시 또는 중소도시에 해당되며, 재정자립도 면에서 20.975% 보다 낮은 자치단체에 사는 주민 그룹

노드 3은 229명의 주민이 포함되어 있다. 이 그룹은 인본적 자본의 값이 6.938 보다 작거나 같으며, 자연 자본값이 5.750 보다 낮은 주민 그룹이다. 이 노드의 경우 229 명중 90%에 해당하는 206명이 거주의향이 낮음에 포함되어 있다. 다른 말로 한다면, 인본적 자본의 값이 6.94 보다 낮고, 자연 자본값이 5.75보다 낮은 의식을 하고 있는 주민들의 경우, 거주의향이 대단히 낮다는 것을 의미한다. 노드 23은 33명이 포함되어 있다. 이 노드의 특징은 인본적 자본의 값이 6.938 보다 작거나 같으며, 자연 자본값이 5.750보다 높으나, 지역특성상 농촌지역이며, 인본적

자본 값이 5.438 보다 낮고, 자연 자본값이 6.417 보다 낮은 주민 그룹이다. 여기에는 인본적 자본의 기준이 두 번 포함되어 있다. 다음 노드 13은 82명이 포함되어 있다. 이 그룹의 특성은 인본적 자본의 값이 6.938 보다 작거나 같으며, 자연 자본 값이 5.750 보다 높고, 지역특성상 대도시 또는 중소도시에 해당되며, 재정자립도 면에서 20.975% 보다 낮은 자치단체에 사는 주민 그룹이다. 이는 인본적 자본이나 자연 자본 면에서 낮은 만족감을 가지고 있으면서 대도시 자치구 또는 중소 도시에 살고, 동시에 재정자립도 면에서 특히 낮은 자치단체에 사는 주민이 포함된다.

V. 결 론

본 연구는 우리나라 지역 주민들이 현재 살고 있는 지역에 대한 거주의향에 영향을 미치는 요인들의 다양한 결합조건들을 확인하는데 연구의 목적을 두었다. 이를 통하여 중앙정부는 총괄적인 측면에서 지방자치단체에 대한 인구유입 등 문제에 있어서 각종 보조금 지원정책이나 자원배분을 효율적으로 하는데 도움이 될 수 있다. 아울러 개별 지방자치단체는 해당 자치단체의 실정에 맞게 인구를 유입시키거나 유출을 방지하는데 필요한 정책적 정보를 확보할 수 있게 된다. 이러한 필요성에서 기존에도 유사한 연구목적을 가진 연구들이 있어 왔다. 그러나 대부분의 국내외 연구들은 주민들의 거주의향에 영향을 미치는 요인들을 회귀분석과 같은 방법으로 도출하고자 한 경우가 많았다. 이러한 연구결과들도 유의미한 정보를 제공해 준다. 그러나, 개별 자치단체들이 처한 조건들이 상이하기 때문에 획일적인 정보만 가지고는 정책적으로 큰 의미를 제공하기 어렵다. 이러한 문제점을 해소해 주는데 도움을 줄 수 있는 방법이 데이터마이닝 기법이며, 본 연구가 선택한 의사결정나무 분석 방법이 여기에 해당한다. 의사결정나무 분석은 개체들을 유사한 행태를 보이는 소그룹의 일원으로 분류해 주며, 개체들을 일정한 기준에 따른 층으로 분류해 준다. 이러한 분류는 지방자치단체의 행정을 수행하는 데 유용한 기준이 될 수 있다. 아울러 독립변수들의 특정 조합이 가지는 효과를 식별해 냄으로써 모수적 모형에서 어떤 변수간 상호작용을 고려해야 하는 지를 알 수 있다.

이러한 배경 하에서 본 연구는 우리나라 27개 시군구의 2,700명 주민들을 대상으로 하여 주민들의 거주의향에 영향을 미치는 변수들을 설정하고 -개인적 배경 변수, 지역특성 변수, 정책적 변수-이러한 변수들 중에서 어떠한 변수들이 영향을 미치는 데 있어서 중요한 변수인지, 그리고 거주의향이 높은 그룹의 특성과 낮은 그룹의 특성이 무엇인지를 파악해 내는 것을 연구목적으로 하였다. 본 연구의 분석 결과 얻어진 결론은 다음과 같다. 첫째 개인적 배경변수, 지역특성 변수, 그리고 정책적 변수들 중에서 지역 주민들의 거주의향에 대해서는 정책적 변수들이 크게 영향을 미치는 것으로 나타났다. 성별이나 소득수준과 같은 개인적 배경과 지역특성과 같

은 변수들 보다도 오히려 정책적 변수들이 영향을 크게 미치는 것으로 나타났다는 점이다. 정책적 변수들 중에서도 인본적 자본변수 (공공도서관, 평생교육, 지역내 학교수준, 건강수준, 의료, 노인, 장애인, 보육 등 8개)들이 다른 정책변수들 보다도 상대적으로 더 중요한 영향을 미치고 있다는 점이다. 본 연구에서 사용한 인프라 자본, 경제적 자본, 문화적 자본, 자연 자본 및 사회자본 보다도 인본적 자본 변수들이 크게 거주의향에 영향을 미치고 있다는 점이다. 인본적 자본 변수들은 교육이나 사회복지에 관련되는 변수들이다. 이는 상대적 중요도 분석에서 나타난 바 있다. 교육과 복지 분야가 주민들의 거주의향을 높이는데 가장 크게 영향을 미치고 있다는 점은 지방자치단체들이 크게 관심을 가져야 하는 점이라고 생각된다. 둘째 거주의향이 높은 주민들의 상위그룹 특성을 살펴 볼 때, 인본적 자본 값이 높은 경우가 공통적으로 해당되나 자치단체 규모면에서 볼 때는 인구규모가 71,481 명보다 작은 경우나 큰 경우 모두 다른 조건들만 갖추어지면 거주의향을 높일 수 있다는 점이다. 이러한 점이 바로 회귀분석과 같은 방법을 통해서 나온 분석 결과와 다른 점이다. 노드 18의 경우, 인본적 자본의 값이 6.938보다 높으며, 인구가 71,481 명 보다 작은 자치단체에 살고, 인프라 자본은 6.750 점보다 높고, 월가구소득이 500만원 이상 되는 주민 그룹이 거주의향이 높은 주민들이다. 한편, 또 거주의향이 높은 노드는 36이다. 이 노드의 특성을 보면, 인본적 자본의 값이 6.938 보다 높고, 자치단체의 인구는 노드 18과는 달리 71,481 명 보다 높다는 점이다. 다만, 이 조건에 더하여, 다시 인본적 자본 점수가 7.562 보다 높으나 재정자립도가 25.73% 보다 높으며, 인본적 자본 점수가 8.56보다 낮은 범위에 있는 주민 그룹이다. 즉, 인구규모가 높은 자치단체이지만 재정자립도 면에서 높은 상태가 되면서, 일정한 정도 수준의 인본적 자본 점수를 가지게 되면 높은 거주의향을 가지게 된다는 점이다. 이러한 세부적인 조건들에 대해서 자치단체들은 정밀하게 고려해 볼 필요가 있다. 셋째 거주의향이 낮은 주민 그룹들의 경우도 다양한 조건들이 포함되어 있어서 자치단체들은 조건들에 대한 정책적 고려를 할 필요성이 있다는 점이다. 거주의향이 낮은 그룹에 포함되는 노드 3의 경우 인본적 자본의 값이 6.938 보다 작거나 같으며, 자연 자본값이 5.750 보다 낮은 주민들이 포함되고 있다. 이는 교육 및 복지 분야 및 자연 자본 면에서 낮은 수준이 되면 주민들은 거주의향이 낮아져서 타 지역으로의 유출 가능성이 높아진다는 점이다. 인구유출이 심한 자치단체의 경우, 이러한 분야에 대한 적극적인 노력이 필요하다는 점을 시사해 준다. 아울러, 인본적 자본의 값이 낮으면서 지역특성상 대도시 또는 중소도시에 해당되며, 재정자립도 면에서 20.975% 보다 낮은 자치단체에 사는 주민들은 특히 거주의향이 낮다는 점이다. 대도시의 자치구나 중소도시이면서 재정자립도가 낮은 지역인 경우 특히 인본적 자본 값이 낮으면 주민들의 거주의향이 낮아진다는 점도 자원배분이나 정책결정과정에서 고려해야 할 사항이라고 생각한다.

〈참고문헌〉

김순귀·이영훈·정동빈. (2000). 이차 자기회귀오차 구조를 갖는 선형회귀모형의 자료영향도 평가.「품질경영학회지」, 28(2): 57-69.

노영숙·장지현. (2018). 지역사회에 기반한 사회적 기업의 딜레마.「사회적경제와 정책연구」, 8(1): 27-52.

배승종·김대식·은상규. (2016). 정주여건을 고려한 의사결정 나무 기법 활용 농촌지역 유형화.「한국농공학회논문집」, 58(6): 79-92.

서인석·이승종·기영화. (2014). 지방행정서비스와 커뮤니티 웰빙 구성요소 간 관계: Kee 모형의 6개 자본 요인을 중심으로.「한국행정연구」, 23(2):243-269.

서재호·김혜정·이승종. (2012). 지방정부의 가치지향으로서 커뮤니티 웰빙에 대한 연구.「한국행정학보」, 46(1): 33-56.

오미영. (2008). 지방자치단체-지역주민 관계가 지역민의 지역사회 유대, 지역사회만족 그리고 거주의향에 미치는 영향.「정책분석평가획보」, 18(1): 327-350.

유호웅·이정주·최외출. (2018). 근거이론에 기반한 새마을교육 특성분석.「사회적경제와 정책연구」, 8(3): 183-213.

이경영·정문기·정예은. (2018). 지속거주의향의 영향요인에 관한 연구.「한국행정연구」, 27(2): 179-209.

이상엽·정건섭. (2011). 커뮤니티 비즈니스의 발전방향 연구.「한국비교정부학보」, 15(2): 221-252.

이석원·정승화·문현승. (2017). 전북혁신도시 거주민의 주거만족 및 거주의향에 관한 연구.「지역사회논문집」, 85: 43-53.

이승종·기영화·김윤지·김남숙. (2013). 커뮤니티웰빙 지표에 대한 공무원과 전문가의 평가비교 연구: AHP 기법의 적용.「한국행정학보」, 47(2): 313-337.

이지혜·최영출. (2014). 청소년의 행복도에 영향을 미치는 교육적 요인분석.「한국비교정부학보」, 18(2): 67-84.

장지현. (2017). 교육경쟁력 강화를 위한 피이드백 루프 탐색과 정책시뮬레이션.「사회적경제와 정책연구」, 7(3): 109-144.

장지현·최영출. (2017). 주민참여 영향요인이 도시발전 수준에 미치는 인과관계 분석: 한국자치행정학보」, 31(3): 25-52.

조용준. (2018). 빅데이타 SPSS 최신분석기법. 서울: 한나래.

주상현. (2012). 지방자치단체 커뮤니티 비즈니스 활성화 방안. 사회적경제와 정책연구」, 2(1): 1-39.

정연택·김재태. (2014). 소형주택 임차인의 주거만족도 및 주거가치, 계속 거주의향에 관한 연구.

「주거환경」, 12(3): 267-285.

최영출. (2019). 지역주민들의 삶의 만족에 대한 영향요인 결합조건 탐색.「한국자치행정학보」, 33(1): 1-26.

최영출. (2015). 커뮤니티 웰빙수준에 미치는 영향요인들의 인과적 관계.「한국비교정부학보」, 19(2): 71-91.

허명회. (2013). SPSS Statistics. 서울: 데이터솔루션.

Brasington, D. M. (2002). The demand for local public goods: The case of public school quality. *Public Finance Review*, 30(3), 163-187.

Champion, T. (2002). *Migration, residential preferences and the changing environment of cities: ESRC Cities Programme project end of award report*. University of Newcastle upon Tyne.

Cheshire, P., & Sheppard, S. (2004). Capitalising the value of free schools: the impact of supply characteristics and uncertainty. *The Economic Journal*, *114*(499), F397-F424.

Cirillo, C., Daly, A., & Lindveld, K. (1998). Eliminating bias due to the repeated measurements problem in stated preference data. In *Operations Research and Decision Aid Methodologies in Traffic and Transportation Management* (pp. 286-301). Springer, Berlin, Heidelberg.

Díaz-Morales, J. F., Jankowski, K. S., Vollmer, C., & Randler, C. (2013). Morningness and life satisfaction: further evidence from Spain. *Chronobiology international*, 30(10): 1283-1285. 3.

Diener, E., & Suh, E. M. (2000). Measuring subjective well-being to compare the quality of life of cultures. *Culture and subjective wellbeing*, MIT Press: Cambridge. MA.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all?. *Journal of Economic Behavior & Organization*, 27(1): 35-47.

Ehrhardt, J. J., Saris, W. E., & Veenhoven, R. (2000). Stability of life-satisfaction over time. *Journal of Happiness Studies*, 1(2): 177-205.

Etzioni, A. (2018). Happiness Is the Wrong Metric. In Happiness is the Wrong Metric (pp. 3-40). *Springer*, Cham.

Graafland, J. & Compen, B. (2015). Economic freedom and life satisfaction: mediation by income per capita annd generalized trust. *Journal of Happiness Studies*, 16: 789-810.

Guttman, L. & Louis, S. (1982). On the definition and varieties of attitude and wellbeing. *Social Indicators Research*, 10(2): 159-174.

Jankowski, K. (2012). Morningness/Eveningness and Satisfaction with Life inn a Polish Sample. *Chrononobiology International*, 29: 780-785.

Kahneman, D & Riss, Jason. (2005). Living and thinking about it: two perspectives on life. In Huppert, Felicia, Baylis, N., Keveme, B. *The science of wellbeing: Oxford*: Oxford University Press.

Kim, J. H., Pagliara, F., Preston, J. (2005). The intention to move and residential location choice behavior. *Urban Studies*, 42(9): 1621-1636.

Knox, P., &Pinch, S. (2014). *Urban social geography: an introduction*. Routledge.

Maria Kockelman, K. (1997). Effects of location elements on home purchase prices and rents in San Francisco Bay Area. *Transportation Research Record*, 1606(1), 40-50.

Kocur, G., Adler, T., Hyman, W., &Aunet, B. (1982). *Guide to forecasting travel demand with direct utility assessment. United States Department of Transportation, Urban Mass Transportation Administration*. Report UMTA-NH-11-0001-82-1, Washington DC.

Lachmann, B., Sariyska, R., Kannen, C., Błaszkiewicz, K., Trendafilov, B., Andone, I., ... &Montag, C. (2018). Contributing to overall life satisfaction: personality traits versus life satisfaction variables revisited—Is replication impossible?. *Behavioral Sciences,* 8(1): 1.

Leech, D., & Campos, E. (2003). Is comprehensive education really free?: a case-study of the effects of secondary school admissions policies on house prices in one local area. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 166(1): 135-154.

LI, S. and WU, F. (2004) Contextualizing residential mobility and housing choice: evidence from urban China, *Environment and Planning A*, 36: 1-6.

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: analysis and applications*. Cambridge university press.

Magee, C. A., Miller, L. M., & Heaven, P. C. (2013). Personality trait change and life satisfaction in adults: The roles of age and hedonic balance. *Personality and Individual differences.* 55(6): 694-698.

McMahon, D. M. (2004). From the happiness of virtue to the virtue of happiness: 400 BC–AD 1780. *Daedalus*, 133(2): 5-17.

McNutt, M. K., & Judge, A. V. (1990). The superswell and mantle dynamics beneath the South Pacific. *Science*. 248(4958): 969-975.

Meen, D., & Meen, G. (2003). Social behaviour as a basis for modelling the urban housing market: a review. *Urban studies*, 40(5-6): 917-935.

Michalos, A. (2008). Education, happiness and well-being. *Social Indicator*s, 87: 347-366.

Michalos, A. C. (2008). Ancient observations on business ethics: Middle East meets West. *Journal of Business Ethics*. 79(1-2): 9-19.

Morikawa, T. (1989). *Incorporating stated preference data in travel demand analysis* (Doctoral dissertation, Massachusetts Institute of Technology).

Murphy, J. J., & Stevens, T. H. (2004). Contingent valuation, hypothetical bias, and experimental economics. *Agricultural and Resource Economics Review*, 33(2): 182-192.

Naci, H. & Ioannidis, J. (2015). Evaluation of wellness determinants and interventions by citizen scientists. *JAMA*, 314(2): 121-2.

de D. Ortuzar, J., Martínez, F. J., &Varela, F. J. (2000). Stated preferences in modelling accessibility. *International Planning Studies*, 5(1): 65-85.

Pagliaro, F., & Preston, J. (2003). The Impact of Transport on Residential Location, Final Report. *Transport Studies Unit, University of Oxford, Oxford,* UK.

Pearmain, D., Swnan son Armain, D., Swanson, J BRADEYWW tated Preference Techniques: A Guide to Practice. London: Steer Davies Gleave and Hague Consulting Group.

Ryff, C. & Keyes, C. (1995). The structure of psychological wellbeing revisited. *Journal of Personality and Social Psychology,* 69(4): 719-727.

United Nations. (2008). *A Community-based Approach in UNHCR Operations*. UN: UN Refugee Agency.

Walker, B., Marsh, A., Wardman, M., & Niner, P. (2002). Modelling tenants' choices in the public rented sector: a stated preference approach. *Urban Studies*, 39(4), 665-688.

Wang, D., & Li, S. M. (2004). Housing preferences in a transitional housing system: the case of Beijing, China. *Environment and Planning* A, 36(1): 69-87.

Western, M. & Tomaszewski, W. (2016). PLOS ONE|DPO: 10.1371/journal.pone.0163345 October 3.

Zhao, R., Zhou, H. & Tu, T. (2019). Reginal Determinants Of Residential Intention Of Migrants In China: Evidence From The Chinese National Migrants Dynamic Monitoring Survey In 2015. *Modern China Studies*, 26(1): 1-49.

접수일(2019년 04월 06일)
수정일(2019년 04월 22일)
게재확정일(2019년 04월 25일)

〈국문초록〉

지역거주 의향에 미치는 영향요인들의 조건결합: 의사결정나무분석 활용

본 연구는 우리나라의 주민들이 현재 살고 있는 지역에 거주하고자 하는 거주의향에 어떠한 변수들이 상대적으로 강한 영향을 미치는 지, 거주의향이 높은 그룹의 특성은 무엇인지 및 거주의향이 낮은 그룹의 특성은 무엇인지를 확인하고자 하는 것이 목적이다. 이를 통하여 중앙정부나 지방자치단체들에게 인구유입 강화 및 인구유출 방지를 위한 시사점을 제공해 주고자 한다. 이를 위해서 현재까지 주민들의 거주의향 분석에는 거의 활용되고 있지 않았던 데이터마이닝 기법 중의 하나인 의사결정나무 분석 방법을 활용하였다. 분석을 위해서는 전국 시군구 27개 주민들 각각 100명을 샘플링하여 2,700 명을 대상으로 수집한 설문지 자료를 활용하였다. 그 결과 19개의 최종 노드가 발생되었다. 거주의향에 영향을 미치는 요인은 개인적 배경이나 지역적 특성 변수 보다도 정책적 변수들, 특히 교육, 복지와 같은 인본적 자본 변수들이 영향을 크게 미치는 것으로 나타났다. 거주의향이 높은 상위그룹 3개와 낮은 하위 그룹 특성을 살펴볼 때, 개별 자치단체들이 처한 상황을 고려하여 거주의향 향상을 위한 맞춤형 접근을 할 필요가 있다는 점을 시사해 준다.

주제어: 거주의향, 의사결정나무 분석, 거주의향 결합조건