

지역주민들의 삶의 만족에 대한 영향요인 결합조건 탐색 의사결정 나무분석 (decision tree analysis)의 적용 *

최영출

본 연구의 목적은 우리나라 27개 시군구 지역 주민 2,700명을 대상으로 한 설문조사 결과를 통하여 삶의 만족도에 대한 영향요인들의 결합조건을 의사결정나무 분석 모형으로 분석해 보고, 주민들의 삶의 만족수준을 제고시킬 수 있는 정책적 시사점을 얻고자 하는 데 있다. 본 연구로부터 얻은 시사점을 정리하면 다음과 같다. 첫째, 전국적 차원에서 볼 때, 끝마디로 조건들이 결정되는 삶의 만족 영향요인들의 조합은 모두 25개이며 가장 중요한 영향변수는 인구이고, 재정자립도 등 다양한 조건들의 결합들이 주민들의 삶의 만족수준에 영향을 미친다는 것을 볼 수 있다. 이러한 사실은 다양하게 결합되는 조건들에 대한 차별적 정책 조합을 통하여 중앙정부나 지방정부가 실질적으로 주민들의 만족수준을 높일 수 있는 정보로 활용할 수 있다는 점이다. 둘째, 대도시는 16개 노드, 중소도시는 14개 노드, 그리고 농촌군의 경우에도 14개 노드로 구분되어 지역유형별로 주민들을 노드별 특성을 고려하여 분류하고 이 분류에 맞게 차등적 삶의 만족 수준 제고 정책을 수행할 수 있다는 점이다. 이를 통하여 주민들의 삶의 만족수준을 제고시킬 수 있는 차등적 관리방안을 제시할 수 있다는 점에서 본 연구의 시사점을 찾을 수 있다.

[주제어: 삶의 만족, 의사결정 나무분석, 결합조건]

I. 서론

삶의 만족은 개인이 일상생활을 하면서 일상의 상황에 대해서 가지고 있는 좋은 감정의 정도라고 할 수 있다. 삶의 만족은 후술하는 바와 같이 주관적인 웰빙과 거의 같은 뜻으로 사용된다. 노벨경제학상을 수상한 바 있는 미국 프린스턴 대학교의 Kahneman 교수는 삶의 만족은 개인이 자기 생활을 해 나가는데 있어서 자기가 설정한 목표를 달성하게 하는 데 있어서나 개인적 기대를 충족시키는데 있어서 상당한 정도로 영향을 미친다고 하면서 (<https://www.haaretz.com>) 삶의 만족도의 중요성을 강조한다. 삶의 만족은 인간이 궁극적으로 추구하는 이상적인 상태임과 동시에 중앙정부나 지방정부도 궁극적인 목적이 주민들의 삶의 만족에 있는 만큼 자원배분과정이나 정책과정을 통해서 주민들의 삶의 만족을 향상시키기 위하여 노력해야 할 책무를 지닌다. 삶의 만족에 대해서 국가간, 지역간 수준을 측정하여 비교하거나, 또는 삶의 만족에 미치는 영향요인을 발견하기 위한 노력 등 오늘날에는 삶의 만족에 대한 관심이 점점 커지고

* 이 연구는 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2016S1A3A2024563).

있다. 특히, 객관적인 여건이 비슷한 지역에서도 삶의 만족수준이 개인별로 다를 수 있고, 지역 단위에서도 타 지역단위와 삶의 만족 수준이 다를 수 있다. 그러나 지금까지는 삶의 만족수준에 관련된 연구들이 주로 삶의 만족수준을 일정한 설문에 의하여 측정하는 수준 측정 연구와 이러한 삶의 만족 수준에 영향을 미치는 영향요인들을 발견하는데 초점을 둔 연구가 대부분이다. 특히, 후자의 경우에는 삶의 만족수준에 영향을 미칠 것으로 가정되는 몇 가지 독립변수들을 선정하고 이러한 독립변수들이 통계적으로 유의성이 있는지를 찾는 이론적 탐색적 수준의 연구들이 영향요인과 관련된 연구들이었다. 그러나, 이러한 연구들의 경우, 중앙정부나 지방정부의 정책결정시 실질적으로 도움이 되는 유용한 정보를 제공해 주는 데에는 한계가 있다. 왜냐하면, 설령 몇 가지 독립변수들의 영향요인들을 파악한다고 하더라도 그러한 독립변수들을 정책결정 과정에서 반영하기가 용이하지 않을 뿐 만 아니라 유사한 집단별로 군집화를 형성하기 어렵기 때문이다. 예를 들어서 성별과 소득수준이 삶의 수준에 영향을 미친다고 하더라도 정책과정 속에서 자원배분이나 제도적 혜택 부여 과정에서 특정지역에 대한 추가적인 배분이나 차등적 관리를 통한 반영이 용이하지 않기 때문이다. 이러한 문제점을 해소하기 위해서는 삶의 만족수준이 높은 지역이나 낮은 지역은 어떠한 변수들의 조합으로 구성되어 있는지를 다양한 집단으로 나누어 설명할 수 있어야 한다. 이렇게 되면, 삶의 만족수준이 높은 지역유형과 낮은 지역유형, 중간 지역 유형 등 다양한 특성을 가진 지역유형을 파악할 수 있게 되고, 이러한 지역유형에는 상응하는 자원배분과 같은 제도적 고려를 통하여 주민들의 삶의 수준을 높일 수 있는 구체적인 방안들을 찾을 수 있게 하기 때문이다. 본 연구는 이러한 배경하에서 기존의 지역 주민들의 삶의 만족수준에 미치는 영향요인 연구를 좀 더 진전시키고, 중앙정부나 지방정부의 정책과정에서 실질적으로 도움이 될 수 있는 정보를 제공하는 것을 연구의 목적으로 한다.

II. 이론적 논의

1. 삶의 만족의 의미와 유사개념

삶의 만족개념은 상호 유사한 다른 개념과도 같이 살펴 볼 필요가 있다. 먼저, 삶의 만족(life satisfaction)은 사람들이 그들의 일상생활에서 자기들이 가지고 있는 일정한 기준에 비추어, 현재의 상황에 대해서 어느 정도 좋은 감정을 가지고 있는가 하는 점과 관련된다(Jankowski, 2012; Diaz-Morales et al., 2013; Bailey et al., 2007). 즉, 사람들이 현재의 모습에 대하여 자기들이 가지고 있는 일정한 기준에 비교하여 일상생활이 어디에 위치하고 있는지를 제시해 주는 정도라고 할 수 있겠다. 삶의 만족은 현재의 느낌에 대한 평가라기보다는 현재의 생활에 대한 긍정적인 태도와 관계된다. 따라서 삶의 만족은 경제적 지위, 교육정도, 경험, 주거 등과 관련하여 측정되어져 왔다. 삶의 만족은 후술하는 주관적 웰빙(subjective wellbeing)의 중요한 부분이라고 할 수 있다. 삶의 만족을 측정하기 위한 지표나 지수와 같은 측정도구의 개

발이나 방법론적 노력도 많이 진행되어 왔다. 삶의 만족을 측정하기 위한 지수로서는 Happy Planet Index (HPI), Living Planet Index, OECD Better Life Index, Where-to-be-born Index, World Happiness Report, World Values Survey 등이 있다. 물론 이러한 지수들은 후술하는 웰빙이나 행복과 같은 개념을 측정하기 위해서도 활용되기 때문에 삶의 만족 개념에만 국한되는 것은 아니다.

다음은 행복 (happiness)에 대해서 알아본다. 행복은 전반적으로 즐거움과 기쁨과 같은 정서적 감정에 대한 좋은 느낌 또는 긍정적이고 즐거운 감정상태, 기쁨과 즐거움과 같은 감정에 대한 현재의 좋은 느낌, 삶의 만족수준에 대한 평가 등 다양하게 정의되고 있다 (Anand, 2016; Deci & Ryan, 2006, Kahneman et al., 2005). 행복은 1960년대 이래로 사회심리학, 임상 및 의학 분야 그리고 행복경제학 등 학계에서 많은 연구가 수행되어져 왔다. 행복과 관계되는 이론들은 매슬로우의 욕구단계설, 자기결정 이론 (self-determination theory), 선택의 자유이론 (freedom of choice), 긍정심리학 (positive psychology) 등이 포함된다. 특히 학계에서는 행복을 측정하기 위한 많은 척도들이 개발되어져 왔다. 주관적 행복 척도 (SHS, Subjective Happiness Scale), 긍정 및 부정 정서척도 (PANAS, Positive and Negative Affect Schedule), 그리고 삶의 만족 척도 (SWLS, Satisfaction with Life Scale) 등이 포함된다. 영국은 2012년에 국가 웰빙 (national wellbeing)에 대한 측정을 시작했는데 이는 국민전체행복 (Gross National Happiness)를 측정해 오고 있는 부탄의 영향을 받은 것이다. 행복의 정의에 대해서 2012년 세계행복보고서 (World Happiness Report)는 인지적인 삶에 대한 평가 (cognitive life evaluation)에 바탕을 두고 있다. 따라서 행복을 측정하기 위한 질문은 일반적으로 간단하다. 가령, 'How happy are you with your life as a whole' 또는 'How happy are you now'와 같은 식이다. 물론 이러한 개념정의와 측정방법에 대해서 비판도 제기된다. 대표적인 예로서 Etzioni (2018) 및 Diener and Suh (2000) 등은 행복개념이라는 것이 주관적 웰빙 (subjective wellbeing)과 거의 구분되지 않으며, 다양한 측면을 포괄하지 못하기 때문에 잘못된 개념이라고 하기도 한다.

삶의 만족과 유사한 개념으로서 웰빙(wellbeing)도 있다. 웰빙 역시 개념 자체가 명확하지 않다. 웰빙은 개인 또는 집단이 긍정적인 (positive) 상태에 있는 것을 말한다 (Naci and Ioannidis, 2015; Ryff & Keyes, 1995; Guttman and Louis, 1982; Magee & Miller, 2013). 웰빙은 단순히 건강을 넘어서서 물리적, 정신적, 사회적, 개인적 측면에서 성취나 만족, 활력 등 다차원적으로 만족하는 상태를 의미한다. 웰빙을 연구하는 접근방법도 다양한데, 발전심리학 (development psychology), 성격심리학 (personality psychology), 그리고 임상심리학 (clinical psychology) 등으로 대별할 수 있다. 발전심리학은 심리적 웰빙이 인생에 걸쳐서 일어나는 성장패턴 측면에서 분석되는 것을 의미한다. 성격심리학은 Maslow의 자기실현 (self-actualization)개념, Jung의 개별화 (individuation), Allport의 성숙 (maturity)과 같은 개념을 응용하여 분석하는 것이 특징이다. 그리고 임상심리학은 정신적 질병이 없는 것이 심리적 웰빙을 구성한다고 보는 입장이라고 할 수 있겠다.

위에서는 삶의 만족, 행복, 웰빙이라는 유사하면서도 상호 구분되는 개념들에서 살펴 보았다.

다양한 학자들의 의견을 반영하여 개념적 정리를 해 보았지만 명확히 개념적으로 구분된다고 보기 어렵다. 이러한 이유로 Lachmann et al. (2018)은 삶의 만족, 행복, 주관적 웰빙은 상호교호적으로 (interchangeably) 사용되고 있다고 지적한다. 그들은 이러한 3가지 개념들은 상당한 정도로 상호 중복되고 있다고 주장한다. 물론 세계적으로 볼 때 각종 국제기구에서는 World Happiness Report 처럼 ‘행복’ (happiness)이라는 용어를 많이 사용하여 국제적으로는 행복 개념이 많이 사용되고 있기는 하나, 실생활에서는 오히려 ‘삶의 만족’ 수준을 측정하는 것이 간단하고 명료하기 때문에 삶의 만족이 더욱 용이한 측정지표가 될 수 있다고 Lachmann et al. (2018)은 주장하고 있다. 이들의 의견을 빌리면 삶의 만족은 삶의 만족도 척도 (Satisfaction with Life Scale, SWLS)를 사용하여 용이하게 측정할 수 있다고 한다. 가령, ‘전반적으로 볼 때, 당신의 삶에 어느 정도 만족하십니까?’ (How satisfied are you with your life overall?)와 같은 단일의 질문이 그 예이다. 이는 독일의 사회경제 패널 (German Socio-Economic Panel, SOEP)에서 이용되는 질문이기도 하다. 이러한 주장외에도 현실적으로 삶의 만족이나, 행복, 웰빙과 같은 개념들이 구분의 실익이 크게 없다고 할 수 있다 (최영출, 2015). 이 유사개념들이 구분되어 별도로 다루어진다면 실제 정책현실에서 차별성으로 인한 구분 실익이 있어야 하나, 그러한 구분 실익이 현실에서는 없다는 것이다. 그 이유는 설령 3가지 개념들의 하위개념을 설정하여 측정지표를 다르게 한다고 하더라도 총론적인 면에서 유사개념들은 상관관계가 대단히 높을 수밖에 없다. 동시에, 설령 미세한 차이가 각 개념마다 발생한다고 하더라도 중앙정부나 지방정부의 정책결정가들 입장에서 자원배분 등 면에서 구분의 실익이 없다고 생각한다. 이러한 관점에서 본 연구에서는 삶의 만족, 행복, 웰빙 등의 유사 개념 중에서 Lachmann et al. (2018)의 견해처럼 삶의 만족이 이해하기 쉽고 측정하기 쉽다는 이유로 이를 핵심개념으로 활용하고자 한다. 여기서 삶의 만족은 여러 학자들의 의견을 종합하고, 특히 Lachmann et al. (2018)의 견해를 반영하여, ‘자기들이 가지고 있는 개인적 기준을 바탕으로 하여 전반적인 관점에서 현재의 생활에 만족하는 정도’로 정의하고자 한다. 아울러, 이러한 삶의 만족은 유사개념인 웰빙, 특히 주관적 웰빙 및 행복과도 상호 교호적으로 사용가능하다고 전제한다.

2. 삶의 만족에 영향을 미치는 영향요인

삶의 만족에 미치는 영향요인들은 무엇인가? 이 역시 웰빙의 영향요인, 행복의 영향요인과도 중복된다고 할 수 있다. 넓게 보면 삶의 만족에 대해서 개인의 퍼스넬리티에 관련된 요인을 강조하는 연구자도 있고, 개인이 처한 상태 및 환경을 강조하는 연구자들도 있다. 전자의 경우에는 소위 5개 성격특성 모델 (big five personality) 또는 5개 성격특성의 약자를 따서 만든 OCEAN 모델이라고 하는 것에 기초하고 있다. 이 5개 성격특성은 경험에 대한 적극성 (openness to experience), 성실성 (conscientiousness), 외향성 (extraversion), 친화성 (agreeableness), 신경증 (neuroticism) 이라고 할 수 있다. 사람들이 이러한 성격특성에서 어떤 정도인가에 따라 삶의 만족정도는 달라진다고 보는 입장이다. 그러한 입장은 지나치게 개인적 특성을 강조하고, 외부적 환경 조건에 대해서는 덜 강조하는 경향이 있다. 이와는 다소 다르게,

개인적 성격 보다는 개인이 처한 개인적 상황이나 조건들을 강조하는 주장도 있다 (Wolfers, 2003). 이러한 주장들은 개인의 연령이나 생활의 경험, 가족상태, 개인의 직업, 결혼유무, 직업 등과 같은 조건들이 삶의 만족에 중요한 영향요인이라고 주장한다. 이러한 개인적 조건들과는 달리, 중앙정부나 지방정부의 정책이나 정치적 안정성, 자유정도 등 거버넌스적 요소를 강조하는 주장도 있다. 이와 같이, 삶의 만족에 미치는 영향요인들도 실로 다양하다. 이를 좀 더 세부적으로 정리해 보면 다음과 같다.

먼저 개인적 차원의 영향요인을 중시하는 연구들이다. 가령, Ehrhardt et al. (2000)은 삶의 만족에 미치는 영향요인으로서 심리적 건강, 사회적 능력, 지적능력 등을 들며, Sastre (1999)는 인구학적 특성, 경제적 특성, 사회적 특성 등을 든다. 또, 도시적 특성, 소득 수준, 학력 수준, 일자리 상태 등을 들기도 한다 (Bailey et al., 2007; Guttman & Louis, 1982). 이와 대조적으로 사회적 차원의 영향요인들을 강조하는 학자들도(McMahon, 2004; Naci & Ioannidis, 2015; Ryff & Keyes, 1995; Easterlin, 1995) 있다. 이들은 사회의 경제적 수준, 정치적 자유, 민주주의 수준, 사회의 불균등 정도, 언론의 자유, 사회적 통합 정도 등을 든다. 우리나라의 많은 연구들은 (박순주, 2001; 최영출, 2018; 임희섭, 1996; 권중돈, 조주연, 2000; 고재욱, 이동열, 2011; 최영출, 2016; 최영출, 2015; 김철주, 이상엽, 2013; 이지혜, 최영출, 2014; 이지희·이은희, 2017; 구혜경, 2018)은 개인적 조건으로서 성별, 연령, 학력, 일자리 형태, 소득 수준, 결혼상태, 지역적 특성 등을 공통적으로 많이 활용하고 있다. 사회적 조건, 개인적 조건 중에서 어떤 조건들이 더욱 중요성을 가지는 가? 이는 연구의 목적에 따라 다를 것이다. 이론적 연구에 치중할수록 관련되는 모든 조건들을 연구에 포함시킬려고 할 것이다. 그러나, 정책적 관점에서 볼 때 지역 주민들 개개인의 심리적 상태나 정신적 상태를 측정하여 정책에 반영하기는 용이하지 않다. 이보다는 오히려 지역적 조건이나 개인의 경제상태나 개인적 조건과 같은 것들이 더욱 중요할 수 있다. 삶의 만족 수준에 미치는 이러한 조건들의 특징이나 영향요인들을 알 수 있다면, 중앙정부나 지방정부가 이러한 조건들을 반영하여 중장기적 또는 단기적으로 삶의 만족도를 개선시키는 데 활용할 수 있을 것이다.

III. 조사설계

1) 연구대상

본 연구의 대상은 우리나라 지역주민 2,700명이다. 서울대학교 행정대학원의 커뮤니티웰빙 연구센터에서는 한국갤럽에 의뢰하여 우리나라 시도의 시군구 27개 지역을 선정하여 설문지를 수집한 바 있다. 이 과정에서 인구규모, 지역 특성 등을 기준으로 27개 시군구가 선정되었고, 각 시군구별로 100명씩의 주민을 유의 추출하여 (purposive sampling) 이 들을 대상으로 설문지를 배포하고 수집한바 있다. 이 자료 수집은 2017년 10월부터 11월까지 실시되었으며, 여기서 수집

된 자료를 활용한다.

2) 연구방법

본 연구에서는 연구의 목적을 효과적으로 달성하기 위하여 의사결정 나무분석 (decision tree analysis)을 활용한다. 본 연구의 목적은 전술한 바와 같이 삶의 만족수준에 영향을 미치는 다양한 결합조건들을 도출하여 중앙정부나 지방정부가 자원배분 및 정책결정과정에 삶의 만족 수준 향상에 관한 의미있는 정보를 제공하기 위한 것이다. 의사결정 나무분석은 바로 이같은 연구 목적 달성을 위해서 효과적인 방법이라고 생각한다. 그 이유는 첫째, 의사결정나무분석은 개체들을 유사한 형태를 보이는 소그룹의 일원으로 분류해 주며, 일정한 기준에 따른 층으로 나누어 준다 (허명희, 2013; 이희연, 노승철, 2017). 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). 회귀분석과 같은 방법들도 이러한 목적 달성에 유용하나, 의사결정나무분석은 성격이 서로 다른 변수들도 한번에 고려하여 다양한 결합조건들을 파악하게 해 줌으로써, 유의미한 변수들의 분류에 대단히 유용하다. 셋째, 회귀분석과 같은 방법과는 달리 독립변수들의 특정조합이 갖는 효과를 식별해 냄으로서 모수적 모형에서 어떤 변수 간 상호작용을 고려해야 하는지에 대한 지침을 제공해 준다. 즉, 회귀분석에서는 독립변수들 중에서 통계적으로 유의미한 변수들만 회귀식에서 이용될 수 있으나 (김순귀외, 2012; 임은정, 정순희, 2015) 의사결정 나무분석에서는 통계적으로 유의미한 변수가 아니더라도 다른 변수와 결합이 되면 종속변수에 유의미한 영향을 미칠 수 있다는 점을 반영해 준다. 이러한 이유로 본 연구에서는 삶의 만족수준을 종속변수로 하고 여기에 영향을 미칠 수 있는 다양한 독립변수들의 특정조합이 갖는 효과를 식별해 보기 위하여 의사결정나무분석을 활용하고자 한다.

3) 변수선정

본 연구에서는 종속변수를 삶의 만족수준으로 하고, 독립변수는 선행연구들의 결과를 바탕으로 하여 다음과 같은 9 개의 변수를 선정한다. 독립변수에는 지역주민들이 거주하는 해당 지역의 특성과 아울러 개별 지역주민들의 개인적 특성을 포함한다. 따라서, 성별, 연령, 지역유형, 지역의 인구규모, 지역의 재정상태, 소득수준, 학력, 일자리 형태, 결혼상태 등이 포함된다. 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 정서상태를 파악하기 쉽지 않다는 점 및 언론의 자유같은 거버넌스적 측면은 지역별 차등화가 용이하지 않다는 점 때문이다. 중앙이나 지방의 정책결정자 입장에서 필요한 것은 오히려 지역적 특성과 지역주민들의 구성형태 정보들이 오히려 정책과정에서 용이하게 활용할 수 있다는 장점이 있다. 이러한 관점에서 본 연구에서는 우선 전국적 차원에서 응답자들에 대한 의사결정 나무분석을 실시하고, 각 지역유형별 (대도시, 중소도시, 농촌군)로 구

분하여 분석을 실시하고자 한다.

<표 1> 변수 및 측정방법

영역	변수	측정방법	비고
지역	지역유형	특,광역시외의 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
	결혼상태	기혼, 미혼, 사별 또는 이혼	명목변수로 처리
	학력	중졸이하, 고졸이하, 대졸, 대학원 이상	서열변수로 처리
	일자리 형태	정규직, 비정규직, 프리랜스, 자영업	명목변수로 처리
삶의 만족도	지역내의 전반적인 삶의 만족도	지역내에서의 전반적인 삶의 만족도	비율변수로 처리 (최하 1점, 최상 10점)

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 경우 산출된 모형 (함수식 등)이 통계적으로 유의미한지 일정한 기준에 의하여 검증하고 유의하지 않으면 이후의 단계는 진행하지 않고 새로운 변수 투입이나 제거과정을 거쳐서 새로운 유의미한 모형을 찾는 것이 일반적이다. 그러나 의사결정 나무분석과 같은 데이터마이닝 알고리즘에서는 이러한 방법으로 모형의 유의성을 검증하지 않는다 (조용준, 2018). 의사결정 나무분석의 경우에는 의사결정나무모형을 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). 의사결정 나무분석의 경우에는 투입된 독립변수들 중에서 나무모형에 포함된 변수들은 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 분류

하였는 지 (예: 양호, 불량 또는 성공, 실패 등) 명확히 셀 수 있으므로 예측정확도로 평가가 가능하다. 그러나 종속변수가 연속형이 되면 실제값과 모형에 의한 예측값과의 괴리가 어느 정도 되는지를 통해서 정확도를 추정할 수 있다. 이를 측정하기 위해서는 여러 가지 방법이 있으나 뿌리노드(=노드 0)의 표준편차 값을 이용한 분산과 회귀나무 모형에 의하여 설명된 분산비율을 이용하여 결정계수를 (R^2) 추정한다 (허명희, 2013). 마지막으로 모형해석인데, 모형해석의 경우 산출된 모형이 함수식인지 (예: 회귀모형), 규칙기반인지 (예: 의사결정 나무분석), 아니면 모형산출의 근거가 없는 블랙박스인 경우 (예: 신경망 분석, 휴리스틱 시뮬레이션)인지에 따라 다르다 (조용준, 2018). 본 연구와 같은 의사결정 나무분석의 경우에는 여러 가지 변수들의 다양한 결합 조건에 (인구규모, 소득규모, 지역유형 등) 의하여 종속변수 결과를 설명하고 해석할 수 있다.

IV. 분석결과

1) 조사대상자의 일반적 현황

아래 표에서 보는 바와 같이, 전체 응답자 수는 2,700명이며, 각 항목별 결측치가 일부 있기는 하나 대체적으로 잘 응답한 것으로 나타나고 있다. 27개 시군구에서 각각 100명씩 추출되었고, 27개 시군구의 인구, 재정자립도 및 지역유형은 지역적 특성을 반영하기 위하여 포함되었다.

<표 2> 조사대상자의 일반적 현황

구분	사례 수/평균	비율(%)/평균 점수	비고
인구	205772		2017년 기준
재정자립도	29.61		2017년 기준
지역유형	대도시(자치구)	1,100	40.7
	중소도시	692	25.6
	농촌군	908	33.6
월가구소득	0-99만원	108	4.0
	100-199만원	317	11.8
	200-299만원	441	16.4
	300-399만원	625	23.2
	400-499만원	540	20.1
	500-599만원	394	14.7
	600-699만원	171	6.4
	700-799만원	48	1.8
	800-899만원	18	0.7
	900-999만원	15	0.6
일자리형태	정규직임금근로자	918	47.5
	비정규직임금근로자	310	16.1

	자영업자	633	32.8	
	무급가족종사자	42	2.2	
	기타종사자	28	1.5	
경제활동	경제활동합	1931	71.5	
	경제활동안함	769	28.5	
주거형태	자가	2078	77.2	
	전세	488	18.1	
	반전세	27	1.0	
	월세	78	2.9	
	공동주택 (기숙사, 고시원등)	18	0.7	
	기타	3	0.1	
학력	초등졸업 이하	202	7.5	
	중학교 졸업	292	10.9	
	고등학교 졸업	1144	42.5	
	전문대학 또는 대학교졸업	1011	37.6	
	대학원이상	40	1.5	
나이	나이	48.8		
성별	남성	1339	49.6	
	여성	1361	50.4	
혼인상태	미혼	522	19.4	
	기혼	2063	76.6	
	이혼, 별거, 사별등	109	4.0	
종교유무	있음	960	35.6	
	없음	1740	64.4	
지역에 대한 만족도	전반적으로 볼 때 귀하가 사는 지역의 삶에 어느 정도 만족하십니까?	7.025		1점-10점

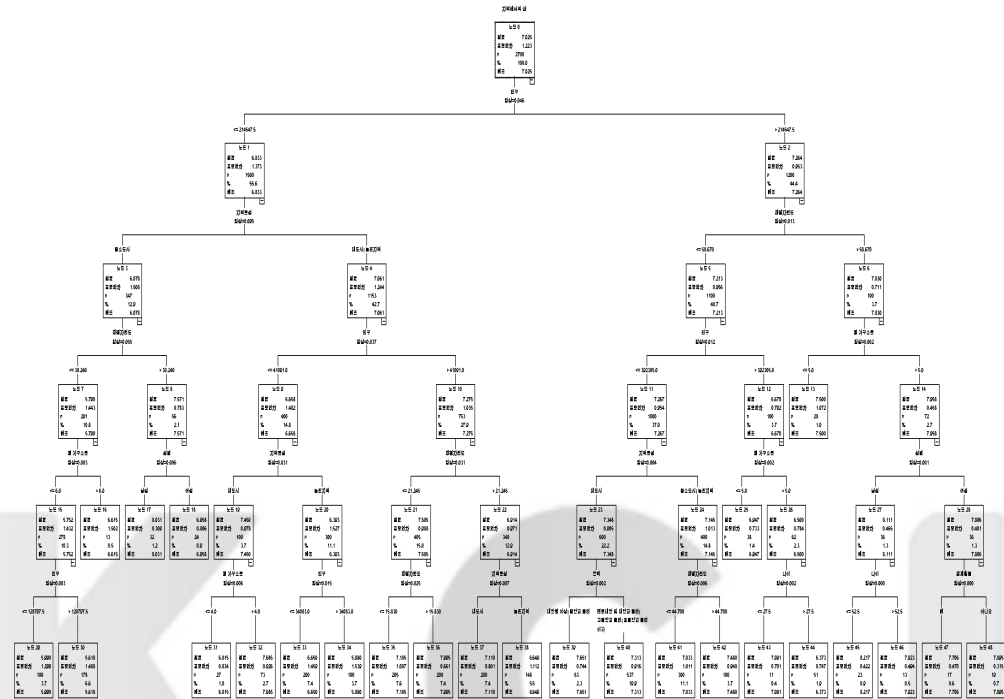
주: 결측치는 제외하였음.

한편, 종속변수인 삶의 만족도는 전반적으로 볼 때의 지역의 삶에 대한 만족도를 묻는 질문으로 구성되어 있으며, 평균은 7.025점으로 나타나고 있다.

2) 전국적 차원의 의사결정 나무분석

전국 응답자들 (2,700명)의 지역 내에서의 삶의 만족도에 영향을 미치는 요인을 확인하기 위하여 의사결정 나무분석을 수행하였다. 먼저 형성된 상위마디의 변수일수록 영향요인 면에서 중요하다라는 것을 의미한다. 계속 분리되어 자식마디의 특성이 누적되어 가며, 자동적으로 지역내 삶의 만족도의 평균이 변화되게 된다. 아래 그림은 전국 응답자 전체의 지역내에서의 삶의 만족도 수준을 나타내고 있는 의사결정 나무 모형이다. 아래 그림을 보면, 전체 응답자의 지역내에서의 삶의 만족도 점수는 7.025 점이며, 이는 여러 변수들의 상호작용을 거쳐서 8.217점 (노드 45)까지 높아지거나 5.62 (노드 30) 점까지 낮아지고 있다.

<그림 1> 전국적 차원의 의사결정나무 모형



끝마디로 조건들이 결정되는 조합은 모두 25개로 나타나고 있다. 이러한 조건결합들의 평균 점수 및 특징을 표로 나타내 보면 아래 표와 같다. 표에서 보는 바와 같이, 삶의 만족 수준이 가장 높은 집단은 노드 45이며, 이 집단은 인구가 21457명 이상 되는 지역에 살며, 지역의 재정자립도가 54.67% 인 지역에 살고, 월 가구소득이 500만원 이상이 되며, 성별은 남성이고, 나이가 52.5세 보다 낮은 사람들의 그룹으로서 평균점수는 8.22로서 높다. 그리고, 이 집단에 포함되는 사례 수는 23명으로 나타나고 있다. 즉 전체 응답자 2700명 중에서는 약 1% 수준이라고 할 수 있다. 이를 만약 우리나라 전체 국민으로 확대한다면, 약 50만명 수준이 된다고도 할 수 있다. 이 그룹에 속한 사람들은 다른 사람들 보다도 삶의 만족 수준이 상대적으로 높은 집단에 속하고 있다고 할 수 있겠다.

<표 3> 전국 응답자들의 의사결정나무 모형 노드의 일반적 현황

노드	N	퍼센트	평균
45	23	0.9%	8.22
17	32	1.2%	8.3
36	200	7.4%	7.99
46	13	0.5%	7.92
48	19	0.7%	7.89

47	17	0.6%	7.71
32	73	2.7%	7.68
39	63	2.3%	7.65
13	28	1.0%	7.50
42	100	3.7%	7.48
40	537	19.9%	7.31
35	205	7.6%	7.19
37	200	7.4%	7.11
43	11	0.4%	7.09
41	300	11.1%	7.03
18	24	0.9%	6.96
25	38	1.4%	6.95
31	27	1.0%	6.81
33	200	7.4%	6.65
38	148	5.5%	6.65
16	13	0.5%	6.62
44	51	1.9%	6.37
29	100	3.7%	5.99
34	100	3.7%	5.88
30	178	6.6%	5.62

성장방법 : CRT

종속변수 : 지역에서의 삶

이러한 노드 중에서 편의상 삶의 만족수준이 높은 상위 3개 노드와 가장 낮은 하위 노드의 특성을 설명해 보면 아래 표에서 보는 바와 같다. 중앙정부나 지방정부 입장에서는 삶의 만족수준이 높은 그룹과 하위 그룹에 대해서는 별도의 관심을 가질 필요성이 있기 때문에 본 연구에서는 상위 3개 그룹 및 하위 3개 그룹에 대해서만 특징을 살펴 보기로 한다.

<표 4> 전국 응답자들의 상위 3개 노드 특성

노드	N	퍼센트	평균	특징
45	23	0.9%	8.22	인구가 비교적 많은 214547명 이상 되는 지역에 살며, 지역의 재정자립도가 54.67% 인 지역에 살고, 월 가구소득이 500만원 이상이며, 성별은 남성이고, 나이가 52.5세 보다 낮은 사람들의 그룹
17	32	1.2%	8.3	인구가 214547명 보다 작은 지역이며, 지역적으로는 중소도시에 살고, 재정자립도가 30.24% 보다 높은 지역에 살고, 성별로는 남성인 사람들의 그룹
36	200	7.4%	7.99	인구가 214574명보다 작은 지역이며, 지역유형으로는 대도시 또는 농촌군지역이고, 그러나 인구가 41000명 보다는 많으며, 재정자립도가 21.3% 보다 낮으나 15.8% 보다는 높은 지역에 사는 주민그룹

<표 5> 전국 응답자들의 하위 3개 노드 특성

노드	N	퍼센트	평균	특징
29	100	3.7%	5.99	인구가 214547명 보다 작고, 중소도시에 속하며, 재정자립도가 30.24% 보다 낮고, 월가구소득이 600만원 보다 낮으면서, 인구가 128797명보다 작은 지역에 사는 주민그룹
34	100	3.7%	5.88	인구가 214547명 보다 작고, 대도시 또는 농촌지역이며, 인구가 41091명 보다 작고, 지역유형이 농촌지역이며, 인구가 34043 명 보다 큰 지역의 주민 그룹
30	178	6.6%	5.62	인구가 214547명 보다 작고, 중소도시에 속하며, 재정자립도가 30.24% 보다 낮고, 월가구소득이 600만원보다 낮으며, 인구가 128797명 보다 작은 지역에 사는 주민 그룹

아래 표는 전국 응답자들의 삶의 만족 수준을 구성하는 변수들의 상대적인 중요도를 나타낸 것이다. 상대적 중요도는 첫 번째 마디를 형성하는 인구의 중요도를 1로 볼 때, 나머지 변인들의 상대적인 중요도를 나타낸다. 전국적 차원에서의 응답자들의 삶의 만족 수준에서 가장 중요한 변수는 인구규모였고, 두 번째 중요한 변수는 지역특성, 그리고 재정자립도 순이었다.

<표 6> 독립변수 중요도 (전체)

독립 변수	중요도	정규화 중요도
인구	0.231	100.0%
지역특성	0.190	82.2%
재정자립도	0.186	80.5%
월 가구소득	0.065	28.2%
일자리 형태	0.025	10.8%
주거형태	0.014	6.0%
학력	0.014	5.9%
나이	0.010	4.3%
성별	0.006	2.8%
경제활동	0.003	1.4%
혼인 상태	0.002	0.8%
종교 유무	1.808E-5	0.0%

성장방법 : CRT

종속변수 : 지역에서의 삶

한편, 전국 응답자를 대상으로 한 전국적 차원의 모형의 유의성은 노드 분할 및 범주합치기 유의수준을 0.05로 하여 의사결정 나무모형이 구축된 것이므로 모형의 통계적 유의성은 있다고 하겠다. 아울러, 뿌리 노드 (노드=0)에서 표준편차가 1.223이고, 분산은 1.4957이며, 위험도 추정

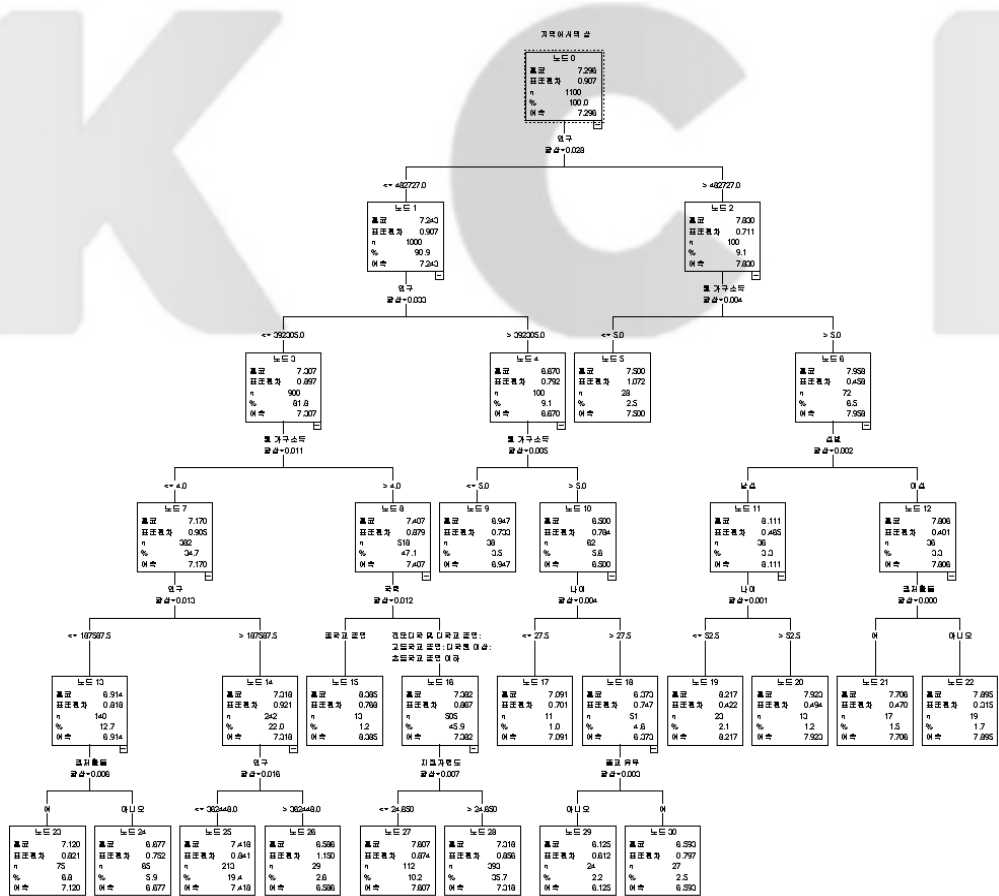
값이 0.78로 나타남에 따라, 회귀나무에 의하여 설명된 분산비율, 즉 결정계수 (R^2)는 $\{1-(0.78/1.223)\}$ 이 되어 0.37 수준이다. 따라서, 모형의 유의성과 적합도 면에서는 문제가 없는 것으로 나타나고 있다.

3) 지역유형별 의사결정나무분석

(1) 대도시

대도시 자치구들 응답자들의 삶의 만족에 대한 의사결정나무 분석결과는 아래 그림과 같다. 대도시 지역 주민 1100 명들의 경우, 삶의 만족 평균은 7.296으로 나타났다. 16개의 끝마디가 나타나 있으며, 가장 높은 노드는 노드 15로서 평균이 8.38이고, 가장 낮은 노드는 29로서 평균이 6.12로 나타나고 있다.

<그림 2> 대도시 지역의 의사결정나무 모형



아래 표는 대도시 지역 주민들의 노드에 대한 이익표가 나타나 있다.

<표 7> 대도시 지역주민들의 의사결정나무 모형 노드의 일반적 현황

노드	N	퍼센트	평균
15	13	1.2%	8.38
19	23	2.1%	8.22
20	13	1.2%	7.92
22	19	1.7%	7.89
21	17	1.5%	7.71
27	112	10.2%	7.61
5	28	2.5%	7.50
25	213	19.4%	7.42
28	393	35.7%	7.32
23	75	6.8%	7.12
17	11	1.0%	7.09
9	38	3.5%	6.95
24	65	5.9%	6.68
30	27	2.5%	6.59
26	29	2.6%	6.59
29	24	2.2%	6.12

성장방법 : CRT

종속변수 : 지역에서의 삶

대도시 지역 주민들의 경우 가장 높은 노드는 15로서, 인구가 482727명 보다 작은 지역 중에서, 다시 인구가 392305보다 적으며, 월가구소득이 400만원 이상이고, 학력면에서는 초중고 졸업인 주민 그룹이라고 할 수 있다. 이 그룹에 포함되는 주민 사례수는 13명이고 대도시 응답자 1100명 중에서 약 1.2%를 차지하는 비율이다. 대도시 지역주민들의 경우에도 가장 높은 상위 노드 3개와 하위 노드 3개의 특성을 살펴 보기로 한다.

<표 8> 대도시 지역주민들의 상위 3개 노드 특성

노드	N	퍼센트	평균	특성
15	13	1.2%	8.38	인구가 482727명 보다 작은 지역 중에서, 다시 인구가 392305보다 적으며, 월가구소득이 400만원 이상이고, 학력면에서는 초중고 졸업인 주민 그룹
19	23	2.1%	8.22	인구가 482727명 보다 많은 지역에 살고, 월 가구소득이 500만원 이상이며 성별이 남자이고 나이가 52.5세보다 작은 주민 그룹
20	13	1.2%	7.92	인구가 482727명 보다 많은 지역에 살고, 월가구소득이 500만원 이상이며 성별이 남자이고, 나이가 52.5세 보다 많은 주민 그룹

아래 표는 대도시 지역주민들의 삶의 만족 하위 3개 노드 그룹의 특성을 살펴 본다.

<표 9> 대도시 하위 3개 노드 특성

노드	N	퍼센트	평균	특성
30	27	2.5%	6.59	인구가 482727명 보다 작은 지역이며, 인구가 392305명 보다는 많은 지역으로서 월가구소득이 500만원 이상이나 나이가 27.5세 보다 많고, 종교를 가지고 있는 주민 그룹
26	29	2.6%	6.59	인구가 482727명 보다 작은 지역이며, 다시 392305명 보다 작고, 월 가구소득이 400만원 이하이며, 인구가 187597명 보다 많은 지역이고, 인구가 382449명 보다 많은 지역 주민 그룹
29	24	2.2%	6.12	인구가 482727명 보다 작고, 다시 인구가 392305명보다 많으며 월가구소득이 500만원 이상이며, 나이가 27.5세보다 많고, 종교를 가지고 있지 않은 주민 그룹

아래 표에서는 대도시 지역주민들의 삶의 만족 수준의 노드 분리에 영향을 미치는 변수들의 중요도를 보여 주고 있다. 대도시 지역주민들의 경우 인구, 재정자립도, 월가구소득 순으로 노드 분리에 영향을 미치고 있다.

<표 10> 독립변수 중요도 (대도시)

독립 변수	중요도	정규화 중요도
인구	0.095	100.0%
재정자립도	0.040	42.2%
월가구소득	0.023	24.1%
나이	0.015	15.8%
경제활동	0.014	15.0%
학력	0.013	13.9%
주거형태	0.008	8.5%
혼인상태	0.005	5.7%
일자리형태	0.005	5.5%
종교유무	0.003	3.2%
성별	0.002	1.7%

성장방법 : CRT

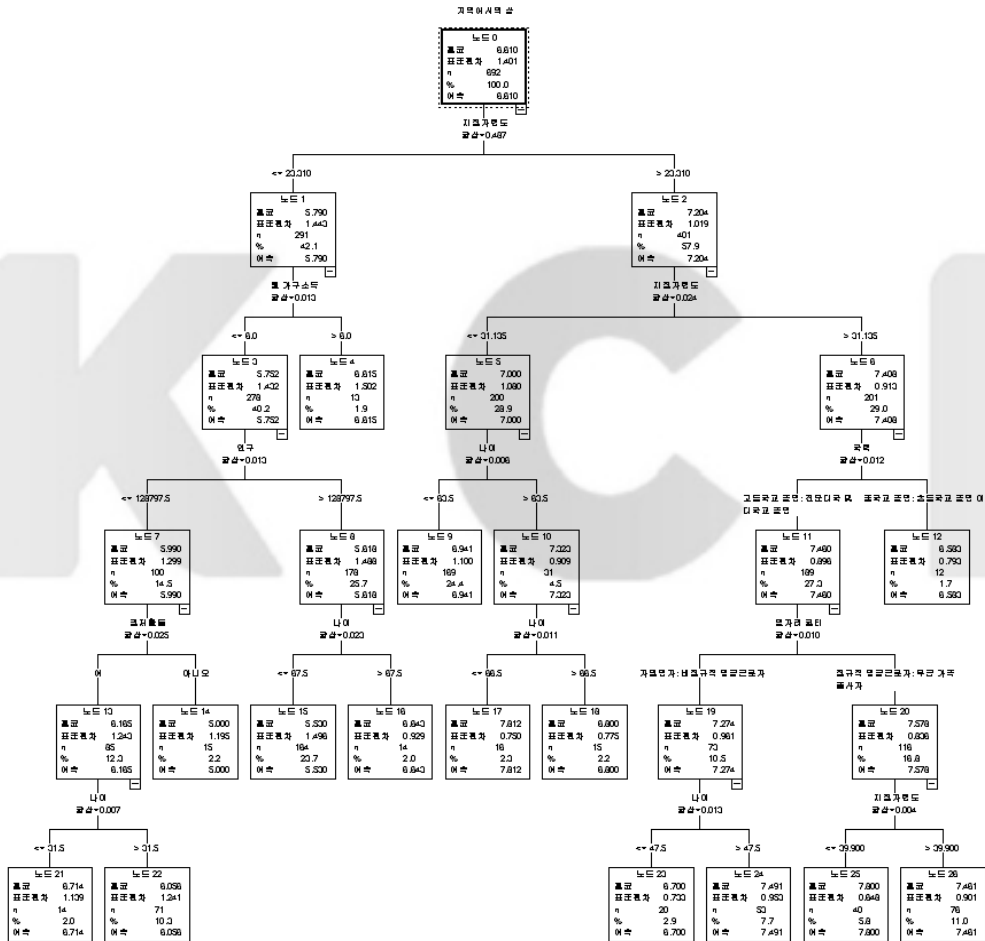
종속변수 : 지역에서의 삶

한편, 대도시 지역 응답자를 대상으로 한 의사결정 나무모형은 전국모형과 마찬가지로 노드 분할 및 범주합치기 유의수준을 0.05로 하여 구축된 것이므로 모형의 통계적 유의성은 있다고 하겠다. 아울러, 뿌리 노드 (노드=0)에서 표준편차가 0.907이고, 분산은 0.8226이며, 위험도 추정 값이 0.477로 나타남에 따라, 대도시 지역표본에 대하여, 회귀나무에 의하여 설명된 분산비율, 즉 결정계수 (R^2) 는 $\{1-(0.477/0.8226)\}$ 이 되어 0.43 수준이다. 대도시 지역의 나무모형의 경우에도 모형의 유의성과 적합도는 문제가 없다고 할 수 있다.

(2) 중소도시

중소도시 지역주민들의 삶의 만족에 대한 의사결정나무 분석결과는 아래 그림과 같다. 중소도시 지역 주민 692명들의 경우, 삶의 만족 평균은 6.810으로 나타났다. 14개의 끝마디가 나타났으며, 가장 높은 노드는 노드 17로서 평균이 7.81이고, 가장 낮은 노드는 14로서 평균이 5.00으로 나타나고 있다.

<그림 3> 중소도시 지역의 의사결정나무 모형



중소도시 지역 주민들의 삶의 만족수준을 보면 가장 높은 노드는 17로서 재정자립도가 23.3% 보다 높은 지역에 살며, 동시에 재정자립도가 31.13% 보다는 낮은 지역이고 나이가 63.5세보다 높고 66.5세보다는 적은 주민 그룹이다. 이 그룹에 포함되는 사례 수는 16명으로서 전체 응답자 중에서 약 2%에 해당된다.

<표 11> 중소도시 지역주민들의 의사결정나무 모형 노드의 일반적 현황

노드	N	퍼센트	평균
17	16	2.3%	7.81
25	40	5.8%	7.80
24	53	7.7%	7.49
26	76	11.0%	7.46
9	169	24.4%	6.94
18	15	2.2%	6.80
21	14	2.0%	6.71
23	20	2.9%	6.70
16	14	2.0%	6.64
4	13	1.9%	6.62
12	12	1.7%	6.58
22	71	10.3%	6.06
15	164	23.7%	5.53
14	15	2.2%	5.00

성장방법 : CRT

종속변수 : 지역에서의 삶

중소도시 지역주민들의 경우에도 가장 높은 상위 노드 3개와 하위 노드 3개의 특성을 살펴보기로 한다. 아래 표에서는 중소도시 지역주민들의 상위 3개 노드의 특성을 보여 주고 있다.

<표 12> 중소도시 지역주민들의 상위 3개 노드 특성

노드	N	퍼센트	평균	특성
17	16	2.3%	7.81	재정자립도가 23.3% 보다 높으나 31.13% 보다는 낮은 지역이고, 나이가 63.5세보다 높고 66.5세보다는 적은 주민 그룹
25	40	5.8%	7.80	재정자립도가 31.13% 보다 높은 지역이면서 학력이 고등학교, 전문대학, 대학교 졸업중 하나이고, 일자리형태가 정규직 근로자 또는 무급가족 종사자 중 하나이면서 재정자립도가 39.9% 보다 낮은 지역에 사는 주민 그룹
24	53	7.7%	7.49	재정자립도가 31.13% 보다 높은 지역이며 학력면에서 고등학교, 전문대 또는 대학 졸업자중 하나이고, 일자리 형태가 자영업자 또는 비정규직 임금근로자면서 나이가 47.5세보다 많은 주민 그룹

아래 표는 중소도시 지역주민들의 삶의 만족 하위 3개 노드 그룹의 특성을 살펴 본다.

<표 13> 중소도시 하위 3개 노드 특성

노드	N	퍼센트	평균	특성
22	71	10.3%	6.06	재정자립도가 23.3% 보다 낮은 지역에 살고, 월 가구소득이 600만원 이하이면서 인구가 128797명 이하인 지역에 살며 경제활동은 하면서 나이가 31.5세보다 많은 주민 그룹
15	164	23.7%	5.53	재정자립도가 23.3% 보다 낮은 지역에 살고, 월가구소득이 600만원 이하이면서 인구가 128797명 보다 높은 지역이나 나이가 67.5세보다 낮은 주민 그룹

14	15	2.2%	5.00	재정자립도가 23.3% 보다 낮은 지역에 살고, 월가구소득이 600만원 이하이면서 인구가 128797명보다 낮은 지역이며 경제활동을 하지 않은 주민 그룹
----	----	------	------	---

아래 표에서는 중소도시 지역주민들의 삶의 만족 수준의 노드 분리에 영향을 미치는 변수들의 중요도를 보여 주고 있다. 중소도시 지역주민들의 경우 재정자립도, 인구, 월가구소득 순으로 노드분리에 영향을 미치고 있다.

<표 14> 독립변수 중요도 (중소도시)

독립 변수	중요도	정규화 중요도
재정자립도	0.519	100.0%
인구	0.311	60.0%
월 가구소득	0.099	19.2%
일자리 형○태	0.076	14.6%
나이	0.067	12.9%
주거형태	0.032	6.1%
경제활동	0.029	5.5%
학력	0.026	5.0%
혼인상태	0.011	2.2%
성별	0.004	0.9%
종교유무	0.002	0.3%

성장방법 : CRT

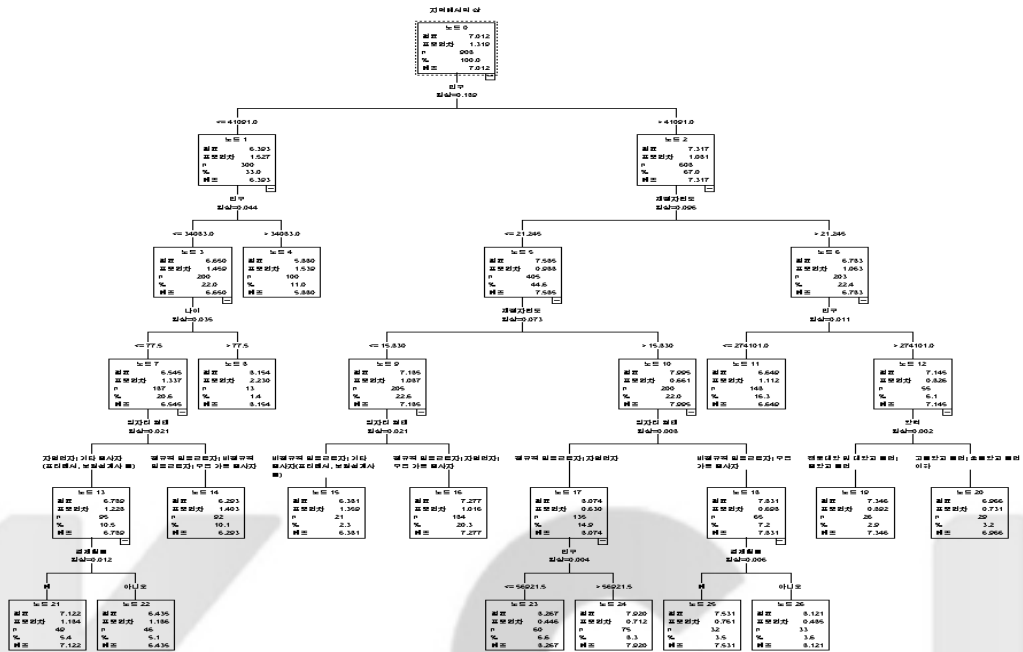
종속변수 : 지역에서의 삶

한편, 중소도시 응답자를 대상으로 나무모형의 경우에도 노드 분할 및 범주합치기 기준을 0.05로 하여 구축된 것이므로 모형의 통계적 유의성은 있다고 하겠다. 아울러, 뿌리 노드 (노드 =0)에서 표준편차가 1.401 이고, 분산은 1.555이며, 위험도 추정값이 0.917로 나타남에 따라, 회귀나무에 의하여 설명된 분산비율, 즉 결정계수 (R^2) 는 $\{1-(0.817/1.555)\}$ 이 되어 0.42 수준이다. 따라서 중소도시 응답자를 대상으로 한 의사결정 나무모형의 경우에도 모형의 유의성과 적합도는 문제가 없다고 할 수 있다.

(3) 농촌 군

농촌 군 지역주민들의 삶의 만족에 대한 의사결정나무 분석결과는 아래 그림과 같다. 농촌지역 908 명들의 경우, 삶의 만족 평균은 7.012로 나타났다. 14개의 끝마디가 나타났으며, 가장 높은 노드는 노드 23 으로서 평균이 8.27 이고, 가장 낮은 노드는 4로서 평균이 5.88로 나타나고 있다.

<그림 4> 농촌 군 지역의 의사결정나무 모형



아래 표는 농촌 군 지역 주민들의 노드별 특성이 나타나 있다. 농촌 군지역 주민들에 대한 특성을 보면, 노드 23의 경우 삶의 만족 수준이 가장 높다. 이 그룹은 인구가 41091명 보다 많은 지역에 살고, 재정자립도가 21.24% 보다 낮으나 15.83% 보다는 높은 지역에 살며, 일자리 형태가 정규직임금근로자 또는 자영업자이면서 인구가 56921명 보다 낮은 지역에 사는 주민 그룹이다. 이 그룹에 포함된 사례 수는 60명이며 전체 응답자의 약 6.6%를 차지하고 있다.

<표 15> 농촌 군 지역주민들의 의사결정나무 모형 노드의 일반적 현황

노드	N	퍼센트	평균
23	60	6.6%	8.27
8	13	1.4%	8.15
26	33	3.6%	8.12
24	75	8.3%	7.92
25	32	3.5%	7.53
19	26	2.9%	7.35
16	184	20.3%	7.28
21	49	5.4%	7.12
20	29	3.2%	6.97
11	148	16.3%	6.65
22	26	5.1%	6.43
15	21	2.3%	6.38
14	92	10.1%	6.29
4	100	11.0%	5.88

성장방법 : CRT

종속변수 : 지역에서의 삶

아래 표에서는 농촌 군 지역 주민들의 상위 3개 노드의 일반적 특성을 보여 주고 있다.

<표 16> 농촌 군 지역주민들의 상위 3개 노드 특성

노드	N	퍼센트	평균	특성
23	60	6.6%	8.27	인구가 41091명 보다 많은 지역에 살고, 재정자립도가 21.24% 보다 낮으나 15.83% 보다는 높은 지역에 살며, 일자리 형태가 정규직임금근로자 또는 자영업자이면서 인구가 56921명 보다 낮은 지역에 사는 주민 그룹
8	13	1.4%	8.15	인구가 41091 명보다 낮으며, 다시 인구면에서 34083 보다는 낮은 지역에 살고 나이가 77.5세 보다 많은 주민 그룹
26	33	3.6%	8.12	재정자립도가 21.24% 보다 낮으나 15.83% 보다는 높은 지역에 살고, 일자리형태가 비정규직 임금근로자 또는 무급가족종사자이며, 경제활동을 하지 않는 주민 그룹

<표 17> 농촌 군 지역주민들의 하위 3개 노드 특성

노드	N	퍼센트	평균	특성
15	21	2.3%	6.38	인구가 41091 보다 많으며, 재정자립도가 21.24% 보다 낮고, 다시 재정자립도 면에서 15.83% 보다 낮으며 일자리형태가 비정규직 임금근로자 혹은 기타종사자 주민 그룹
14	92	10.1%	6.29	인구가 41091 보다 낮으며, 다시 인구가 34083 보다 낮고, 나이가 77.5세 보다 작으며, 일자리 형태가 정규직 임금근로자, 비정규직 임금근로자 혹은 무급가족종사자 중 하나인 주민 그룹
4	100	11.0%	5.88	인구가 41091 보다 낮은 지역이면서 다시 인구가 34083 보다는 많은 지역 주민 그룹

아래 표는 농촌 군 지역주민들의 삶의 만족 수준을 구성하는 변수들의 상대적인 중요도를 나타낸 것이다. 상대적 중요도는 첫 번째 마디를 형성하는 인구의 중요도를 1로 볼 때, 나머지 변인들의 상대적인 중요도를 나타낸다. 농촌 지역 주민들의 경우 삶의 만족 수준에서 가장 중요한 변수는 인구규모였고, 두 번째 중요한 변수는 재정자립도, 그리고 일자리 형태 순이었다.

<표 17> 독립변수 중요도 (농촌)

독립 변수	중요도	정규화 중요도
인구	0.328	100.0%
재정자립도	0.246	75.1%
일자리형태	0.067	20.5%
나이	0.052	15.8%
주거형태	0.035	10.7%
월 구조득	0.035	10.7%
학력	0.027	8.1%
경제활동	0.019	5.9%
혼인상태	0.016	4.8%
종교 유무	0.003	1.1%
성별	0.001	0.4%

성장방법 : CRT 종속변수 : 지역에서의 삶

한편, 농촌 군 응답자를 대상으로 한 모형의 경우에도 노드 분할 및 범주합치기 유의수준을 0.05로 하여 의사결정 나무모형이 구축된 것이므로 모형의 통계적 유의성은 있다고 하겠다. 아울러, 뿌리 노드 (노드=0)에서 표준편차가 1.310이고, 분산은 1.72 이며, 위험도 추정값이 0.88로 나타남에 따라, 회귀나무에 의하여 설명된 분산비율, 즉 결정계수 (R^2) 는 $\{1-(0.88/1.72)\}$ 이 되어 0.49 수준이다. 농촌 군 지역의 경우에도 나무모형의 유의성과 적합도는 큰 문제가 없다고 할 수 있다.

3) 시사점

본 연구에서는 우리나라 27개 시군구 지역 주민 2,700명을 대상으로 하여 삶의 만족도에 대한 영향요인들의 결합조건을 의사결정나무 분석 모형으로 분석해 보았다. 그 분석결과 얻은 시사점을 정리하면 다음과 같다. 첫째, 전국적 차원에서 볼 때, 끝마디로 조건들이 결정되는 조합은 모두 25개로서 다양한 조건들의 결합들이 주민들의 삶의 만족수준에 영향을 미친다는 것을 볼 수 있다는 점이다. 이러한 사실은 다양하게 결합되는 조건들에 대한 차별적 정책 조합을 통하여 중앙정부나 지방정부가 실질적으로 주민들의 만족수준을 높일 수 있는 정보로 활용할 수 있다는 점이다. 둘째, 전국적 차원에서 볼 때, 지역 주민들의 삶의 만족수준을 결정짓는 가장 중요한 조건은 인구특성 변수로 나타나고 있다. 즉, 인구규모가 전국 시군구의 평균인 205,772 명 수준이 되고 (약 21만명), 재정자립도가 일정수준 이상이 되며 (약 54.7%), 월 가구소득이 500만원 이상이고 성별이 남성이며, 나이가 52.5세 보다 낮은 주민 그룹이 가장 만족도가 높은 주민들로 나타나고 있다. 전국적 관점에서 이러한 주민들의 경우에는 이러한 높은 삶의 만족수준을 지속적으로 높일 수 있는 대책이 필요하다. 반면, 인구가 일정규모 보다 작고 (약 21만명), 중소도시에 속하며, 재정자립도가 30.24%보다 낮고, 월 가구소득이 600만원 보다 낮으며, 인구가 약 12만명 보다 작은 지역에 사는 중소도시 지역 주민들은 삶의 만족수준이 가장 낮은 주민그룹이다. 전국적으로 이러한 특성을 가지고 있는 지역 주민들에 대해서는 삶의 만족수준을 높일 수 있는 특단의 정책들이 필요하다고 할 수 있다. 이 지역 주민들의 비율을 전국으로 확산해 볼 때, 약 330만명이 된다. 이 그룹에 속하는 전국의 약 330만 명 되는 주민들의 삶의 수준을 전략적으로 높일 수 있는 방안이 필요할 것이다. 셋째, 지역유형 별로 볼 때 대도시의 구의 경우에도 일정한 인구수준 규모가 주민삶의 만족도를 가르는 기준이 되고 있으며 그 다음에는 재정자립도와 월가구소득이 중요한 기준이 되고 있음을 알 수 있다. 아울러 대도시 구의 경우에는 16개의 끝 마디가 생성되고 있어서 주민들을 16개 특성별 그룹으로 삶의 만족수준 관리를 할 수 있다는 점이다. 중소도시의 경우에는 재정자립도, 인구수, 월가구 소득변수들이 주민 삶의 만족 수준을 구분짓는 중요한 영향변수로 작용하고 있다는 점을 알 수 있다. 특히 중소도시의 경우에는 14개의 끝 마디가 생성되었는 바 14개의 특성을 가진 주민 그룹으로 구분할 수 있다는 점을 제시한다. 중소도시들의 경우 각 주민들이 14개 특성 그룹으로 구분되기 때문에 각 중소도시 지방정부들은 이러한 14개 특성 그룹으로 주민 삶의 만족수준을 관리할 수 있다. 농촌군의 경우에는 인구, 재정자립도 및 일자리 형태가 주민 삶의 만족수준을 가르는 기준으로 작용하고 있다. 중소도시의 경우와 같이 모두 14개의 끝 마디가 나타나고 있어서 14개 주민 그룹으로 구분하여

삶의 만족도 관리를 할 수 있을 것이다. 특히, 농촌군의 경우에는 일자리형태가 삶의 만족수준을 구분짓는 기준으로 되고 있다는 점이 다른 지역유형과 차별성을 가지고 있다고 할 수 있다. 넷째, 지역 주민들의 삶의 만족수준을 높이기 위하여 중앙정부 관점에서는 전국적 차원에서 분석한 25개 노드별 관리를 할 필요가 있으며, 지방정부 차원에서는 각 지역 유형별 노드 수에 따라 특성을 고려하여 삶의 만족수준 관리를 할 필요가 있다는 점이다. 즉, 대도시 지역의 경우에는 각 구청은 16개의 유형을 고려하여 자기 구 지역 주민들을 특성별로 차별적 관리를 할 필요가 있다는 점이다. 마찬가지로 중소도시는 14개로 끝 마디가 구분되기 때문에 이러한 특성을 반영하여 해당 구 지역 주민들에 대한 차별적 서비스 제공과 삶의 만족도 수준 관리를 할 필요성을 제기해 준다 (노영숙, 장지현, 2018). 농촌군의 경우에도 14개 끝 마디가 생성되고 있기 때문에 이러한 특성별로 14개 정도로 주민들을 구분하여 차별적 행정을 수행해 준다면 주민들의 삶의 만족도는 지속적으로 제고될 수 있을 것이라고 생각한다. 다만, 이러한 과정에서 자기 지역들이 가지고 있는 지역자원들을 지역유형에 부합하게 활용할 필요성은 있다고 하겠다. 다섯째, 주민들의 삶의 만족도 분석을 하는 데 있어서 기존에 많이 수행하던 회귀분석과 같은 양적 방법 대신에 본 연구에서 수행한 의사결정 나무분석 방법은 다양한 독립변수들의 결합조건에 의하여 만족수준을 형성하는 여러 가지 조건들의 결합을 제시해 주기 때문에 정책적인 측면에서 이전 분석 방법들보다 훨씬 실용적인 대안들을 모색가능하게 해 준다는 장점이 있다. 이러한 측면에서 실질적, 정책적 시사점 뿐만 아니라 이론적 시사점이 있다고 할 수 있겠다.

V. 결 론

주민들의 삶의 만족 수준 제고는 중앙정부나 지방정부가 궁극적으로 달성하고자 하는 정책의 목표라고 할 수 있다. 동서고금을 막론하고 국가나 지방자치단체는 지역의 주민들이 현재보다 높은 삶의 만족수준을 누릴 수 있도록 노력해야 할 법적 도덕적 의무를 가지고 있다. 이러한 목적에서 삶의 만족 수준에 관하여 외국이나 국내에서 이 분야에 대한 지속적인 연구들이 수행되어 왔다. 그러나 대부분의 연구들은 주민들의 삶의 만족수준에 미치는 영향요인들을 회귀분석과 같은 방법으로 분석한 결과 통계적으로 유의미한 제한된 독립변수들의 중요성을 제시하는데 그치고 있다. 그러나 현실세계에서는 삶의 만족수준이라는 것에 다양한 독립변수들의 결합조건들이 존재할 수 밖에 없다. 일정한 조건들이 높아야 반드시 어떠한 결과가 나타나는 것이 아니라 일정한 조건들이 낮더라도 다른 조건들이 결합되는 경우 유사한 결과가 나타날 수 있다는 점이다. 이러한 점을 감안해 볼 때, 본 연구에서 사용한 의사결정 나무분석은 그 결과가 제시해주는 점이 정책적으로 더욱 의미가 있다고 할 수 있다고 생각한다. 전국적 응답자들의 경우 25개의 끝마디가 생성된 바 있다. 이는 전국적 차원에서 중앙정부가 삶의 만족 수준을 제고하는 정책을 수립하고 집행하는데 있어서 25개 특성별로 그룹핑하여 차등적 자원배분을 할 수 있다는 점을 시사해 준다. 아울러 지방정부 차원에서는 각 지역유형별로 - 시군구- 유형별 특성을

고려하여 가령 대도시 구의 경우에는 16개 그룹, 중소도시의 경우에는 14개 그룹, 그리고 농촌군의 경우에는 14개 그룹으로 주민들을 유형화하여 각 만족수준에 상응하는 차등적 정책을 수립할 수 있을 것이다. 이러한 측면에서 본 연구의 의의가 있다고 할 수 있다.

참고문헌

- 강현철 · 한상태 · 최종후 · 이성건 · 김은석 · 엄익현 · 김미경. (2006). 빅데이터분석을 위한 데이터마이닝 방법론. 서울: 자유아카데미.
- 고재욱 · 이동열. (2011). 노인 일자리사업 참여자의 삶의 만족도와 정책효과에 관한 연구. 「한국노년학」. 31(4): 1209-1228.
- 구혜경. (2018). 소비자의 자기결정성이 윤리적 소비행동과 소비자 행복에 미치는 영향. 「사회적경제와 정책연구」. 8(1): 113-142.
- 권중돈 · 조주연. (2000). 노년기의 삶의 만족도에 영향을 미치는 요인. 「한국노년학」. 20(3): 87-95.
- 김순귀 · 정동빈 · 박영술. (2012). 「SPSS를 이용한 로지스틱 회귀모형의 이해와 응용」. 서울: 한나래 아카데미.
- 김옥숙. (2011). 청소년의 희망적 사고의 결정요인에 관한 연구. 경기대학교 대학원 박사학위논문.
- 김철주 · 이상엽. (2013). 공공기관의 사회적 책임 강화방안. 「사회적경제와 정책연구」. 2(1): 41-80.
- 김희진. (2002). E-MINER와 CLEMENTIME을 이용한 의사결정나무 사례 분석. 이화여자대학교 석사학위논문.
- 노영숙 · 장지현. (2018). 지역사회에 기반한 사회적기업의 딜레마. 「사회적경제와 정책연구」. 8(1): 27-52.
- 박순주. (2001). 만성 폐질환자의 삶의 질 설명모형. 연세대학교 박사논문.
- 배정수. (2014). 의사결정나무 분석기법을 활용한 고등학생 진로결정수준 및 진로준비 행동결정요인 우선순위 탐구. 순천향대학교 대학원 박사학위 논문.
- 손지은 · 김성범. (2013). 의사결정나무 모델에서의 중요 룰 선택기법. 2013년 대한산업공학회 추계학술대회 논문집.
- 이지희 · 이은희. (2017). 저소득층의 삶의 질에 영향을 미치는 요인분석: 심리적 안녕감의 매개효과를 중심으로. 「한국비교정부학보」. 21(2): 85-109.
- 이지혜 · 최영출. (2014). 청소년의 행복도에 미치는 교육적 요인분석. 「한국비교정부학보」. 18(2): 67-84.
- 이희연 · 노승철. (2017). 「고급통계분석론」. 서울: 문우사.

- 임은정·정순희. (2015). 의사결정나무 분석을 통한 한국 중고령자의 점진적 은퇴의사 결정에 관한 연구. *Financial Planning Review*. 8(4): 167-193.
- 임희섭. (1996). 삶의 질의 개념적 연구. 「한국행정연구」. 1: 49-60.
- 장지현. (2017). 교육경쟁력 강화를 위한 피이드백 루프 탐색과 정책시뮬레이션. 「사회적경제와 정책연구」. 7(3): 109-144.
- 조용준. (2018). 빅데이터 SPSS 최신분석 기법. 서울: 한나래.
- 최영출. (2015). 커뮤니티웰빙수준에 미치는 영향요인들의 인과적 관계. 「한국비교정부학보」. 19(2): 71-91.
- 최영출. (2016). 지방분권과 국가경쟁력 및 국민행복도와의 인과관계 분석. 「한국자치행정학보」. 30(1): 1-30.
- 최영출. (2018). 소득수준과 삶의 만족도와의 관계 분석 : 이항로짓분석의 활용, 「한국자치행정학보」. 32(3): 1-18.
- 허명희. (2013). SPSS Statistics 분류분석. 데이터솔루션.
- 허명희·이용구. (2008). 데이터마이닝 모델링과 사례. 서울: 한나래.
- Anand, P. (2016). *Happiness Explained*. Oxford: Oxford University Press.
- Bailey, T., Eng, W., Frisch, M., Snyder, C. (2007). Hope and optimism as related to life satisfaction, *Journal of Positive Psychology*. 2(3): 168-169.
- Deci, E. · Ryan, R. (2006). Hedonia, eudaimonia, and wellbeing: an introduction. *Journal of Happiness Studies*. 9(1): 1-11.
- Diaz-Morales, J. · Jankowski, K. · Vollmer, C. · Randler, C. (2013). Morningness and life satisfaction; further evidence from Spain. *Chronobiology International*, 30: 1283-1285.
- Diener, E., Suh, E. M. (2000). *Measuring subjective wellbeing to compare the quality of life of cultures*. In 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.
- 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 in a Polish Sample. *Chronobiology 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.
- Lachmann, B., Sariyska, R., Kannen, C., Błaszkiwicz, 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.
- 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. (2004). From the happiness of virtue to the virtue of happiness: 400 B.C.-A.D., *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.
- Michalos, A. C. (2008). Ancient observations on business ethics: Middle East meets West. *Journal of Business Ethics*. 79(1-2): 9-19.
- Moksnes, U. & Espnes, G. (2013). Self-esteem and life satisfaction in adolescents—Gender and age as potential moderators. *Qual. Life Res*. 22: 2921-2928.
- Naci, H. & Ioannidis, J. (2015). Evaluation of wellness determinants and interventions by citizen scientists. *JAMA*, 314(2): 121-2.
- Ryff, C. & Keyes, C. (1995). The structure of psychological wellbeing revisited. *Journal of Personality and Social Psychology*. 69(4): 719-727.
- Sastre, M. (1999). Lay conceptions of wellbeing and rules used in wellbeing judgement among young, middle-aged, and elderly adults. *Social Indicators Research*. 46: 137-155.
- Wolfers, J. (2003). Is business cycle volatility costly? Evidence from surveys of subjective well being. *International finance*. 6(1), 1-26.
- “Why Nobel Prize Winner Daniel Kahneman Gave Up on Happiness”, www.haaretz.com, last modified Dec 7.2018, accessed date 20190110, <https://www.haaretz.com/israel-news/.premium.MAGAZINE-why-nobel-prize-winner-daniel-kahneman-gave-up-on-happiness-1.6528513>

최영출(崔永出): 영국 뉴캐슬대학교 (Newcastle University)에서 정책학 박사학위를 받고 (논문제목: Privatisation of Local Government Services and the Politics of Transaction Cost in the UK) 현재 충북대학교 행정학과 교수로 재직 중이다. 주요 관심분야는 정책분석 및 평가, 지역정책, 지방자치, 교육정책 등이다. 영문저서인 *The Dynamics of Public Service Contracting: The British Experience* (Policy Press, UK, 1999)가 있으며 *The Impact of the Daejeon City Knowledge Network* (2010), *Conditions associated with adolescents' happiness in OECD countries* (2016) 등의 논문이 있다. 한국정책분석평가학회 회장을 역임하고, 현재는 한국비교정부학회 회장을 수행하고 있다 <E-mail: ycchoi@cbu.ac.kr/ 논문접수일: 2019. 2. 22/ 논문수정일: 2019. 3. 13/ 게재확정일: 2019. 3. 18>.



Analysing multiple combinations of the conditions affecting local residents' life satisfaction: Application of decision tree analysis method

Choi, Young-Chool

This article aims at identifying multiple combinations of the factors affecting local residents' satisfaction with their daily lives in their local area by using decision-tree analysis method, to examine which domain of the factors affects their life satisfaction most, and to look at the characteristics underlying the more satisfied groups and the less satisfied groups. For this purpose, this article utilizes decision tree analysis method which has seldom been used in the field of happiness studies. The conclusions of this study are summarized as follows. First, it is discovered that the total number of significant nodes explaining residents' life satisfaction at national level is 25, and the most important factor affecting life satisfaction is population size of local government through which local residents are served. Second, it is also found that metropolitan cities, small and medium sized cities, and rural areas are classified into 16 nodes, 14 nodes and 14 nodes, respectively. This paper puts forward policy implications for central government and local government to be able to reflect the research results into their policy processes.

[Keywords: Residents' Life Satisfaction, Decision Tree Analysis, Combinations of the Conditions Affecting Life Satisfaction]