

퍼지셋 질적비교분석의 사회과학적 활용
: Stata를 활용한 Y-검증과
N-검증을 중심으로

민 기 채 (Min, Ki-chaе)*

(E-mail : minkichae@daum.net)

논문접수일 : 2014년 7 월 1 일
논문심사일 : 2014년 7 월 24 일
게재확정일 : 2014년 8 월 27 일

* 학위취득대학 : 서울대학교
현직: 서울대학교 사회복지연구소 선임연구원

퍼지셋 질적비교분석의 사회과학적 활용*
: Stata를 활용한 Y-검증과
N-검증을 중심으로

<국문요약>

그 동안 사회과학에서 양적 연구방법과 질적 연구방법은 서로의 단점을 보완하고 장점을 취하고자 노력하여 왔다. 그 중 라긴(Ragin)의 퍼지셋 질적비교분석 방법은 양적 및 질적 방법 사이에서 연구방법론상의 획기적인 기여를 하였다. 퍼지셋 이상형 분석(fuzzy-set ideal type analysis)과 퍼지셋 다중 결합원인 분석(fuzzy-set multiple conjunctural causations analysis)은 사례의 수와 다중공선성 등의 문제로 인해 해결되지 못한 방법상의 문제 해결에 하나의 실마리를 제공해 주었다고 평가할 수 있다. 본 연구에서는 퍼지셋 방법론을 어떠한 경우에 활용해야 하는지에 대한 사용이유, fsQCA 2.0 소프트웨어를 넘어 Stata로의 확장, '검증'의 문제점을 극복하기 위한 Y-검증과 N-검증법을 소개함으로써 퍼지셋 방법론 활용의 어려움을 보완하고자 하였다. 이와 더불어 퍼지셋 방법론의 구체적인 절차 제시 및 그 절차를 활용한 실증 사례분석을 제시하였다. 이러한 이론적·방법론적 소개를 통해, 퍼지셋 질적비교분석은 양적 연구방법과 질적 연구방법의 대화에 주목하는 사회과학 연구자들을 위해 그 활용가치가 높음을 주장하였다.

[주제어] 퍼지셋 질적비교분석(FSQCA),
퍼지셋 다중 결합원인 분석, Y-검증과 N-검증

* 본 연구는 박사학위 논문을 수정한 것이다.

I. 서론

국가 간 비교연구는 오래된 방법론적 역사를 가지고 왔으며 진보해왔다(Merritt & Rokkan 1966; Tilly 1984). 본 연구에서 소개할 분석방법은 비교사회정책 연구에서 획기적 전환을 가져왔다고 평가되는 퍼지셋 질적비교분석(fuzzy-set qualitative comparative analysis) 방법이다. 퍼지셋 질적비교분석은 라긴(Ragin 1987, 2000, 2008a)에 의해서 태동, 구축, 발전되어 오고 있으며, 리쇼(Rihoux 2003, 2006), 크비스트(Kvist 1999, 2006), 최(Choi 2006)의 연구를 통해 이론적으로 보다 정교화 되어 왔다. 또한 fsQCA 소프트웨어의 방법론적 설명으로는 라긴(Ragin 1995, 2000, 2008b, 2009)을 필두로 켄트(Kent 2008), 슈나이더와 바게만(Schneider and Wagemann 2010), 바게만과 슈나이더(Wagemann and Schneider 2010)의 연구로 진화되고 있다. 그리고 롱이스트와 바이세이(Longest and Vaisey 2008)의 Stata Fuzzy 모듈이 「국제 STATA 저널」에서 검증·소개됨으로써 fsQCA 2.0이라는 통계 프로그램 이외에도 소프트웨어에서의 방법론적 확장을 시도하고 있다. 국내에서는 비교사회정책 연구방법론의 확장을 소개한 안상훈(2002)의 연구 이후, 사회과학에서 퍼지셋 활용의 장점을 안내한 최영준(2009)의 연구 및 최근 보다 방법론적으로 정교하게 퍼지셋 질적비교분석을 시행한 정해식(2012)의 연구로 발전되어 오고 있다.

퍼지셋 질적비교분석 방법의 출현을 보다 더 구체적으로 설명하면 다음과 같다. 일찍이 라긴(Ragin 1987, 34-52)은 사례중심적 접근(case-oriented approach)과 변수중심적 접근(variable-oriented approach)(ibid. 1987, 53-68)을 뛰어 넘어 조합전략(combined strategies)연구와 종합전략(synthetic strategies) 연구로의 발전을 제안한 바 있다(ibid. 69-84). 그것은 바로 불리언 대수(Boolean algebra)를 수학적 근거로 활용한 질적비교분석(QCA: Qualitative Comparative Analysis)(ibid. 85-102)이다. Ragin(2000, 120)은 QCA를 비교 사례연구가 정형화

되고 확장되어 발달한 것으로 본다. 특히 ‘집합 관계’를 통해 각각의 사례를 비교하고 그들의 유사점과 차이점의 패턴을 명확하게 보여주기 위한 분석적 도구임을 강조한다. 또한 QCA를 사용함으로써 사례들 간의 지형(configuration)을 볼 수 있으며 인과관계의 복잡한 패턴을 검증하고, 그 패턴에 기초하여 모집단을 재구성할 수 있다고 본다. 따라서 QCA는 기존의 교차사례분석의 대안인 동시에 다양성 지향 연구를 위한 이론적 기반이 될 수 있다(ibid.).

이후 Ragin(ibid)은 개념의 모호함 내지 흐릿함을 의미하는 fuzzy 개념을 추가한 질적비교분석인 ‘퍼지셋 질적비교분석(fuzzy-set QCA)’을 세상에 내놓았다. 그는 서문에서 퍼지셋 질적비교분석의 출현이 새로운 사회과학 연구방법의 ‘패러다임 전환’임을 3가지의 측면에서 강조한 바 있는데 그 주요내용은 다음과 같다. 첫째, 기존의 양적 방법론을 뒷받침하는 동질성 가정(homogenizing assumptions)은 생각과 증거 사이의 소통을 방해하므로 포기해야 하며, 둘째, 사회현상에 대한 배열적 접근(configurational approach)을 통해 사례의 이질성(heterogeneity)과 차이(difference)를 다룰 수 있는 다양성 지향(diversity-oriented) 연구가 필요하며, 셋째, 퍼지셋은 이론적 개념에 정확하게 들어맞을(tailored) 수 있으므로, 이론과 자료분석 간 연계가 향상될 수 있다고 한다(ibid. 5-6).

이렇듯 퍼지셋 방법론이 가지고 있는 다양한 장점은 그간 사회과학에서 양적 연구방법과 질적 연구방법에서 서로의 단점을 보완하고 장점을 취하고자 하는 노력에 큰 활력을 불어넣어 주었다. 그러나 여전히 퍼지셋 방법론의 활용에서 다음과 같은 어려움을 겪고 있다. 첫째, 퍼지셋 방법론을 ‘어떠한 경우에 사용하는지’에 대한 이해가 부족하다는 것이다. 둘째, fsQCA 2.0 소프트웨어는 2009년 1월 마지막 업데이트 이후에도 여전히 ‘점검 중(revisions in progress)’이므로 퍼지셋을 기술적으로 시행하기 어렵다는 인식이 있다. 셋째, 퍼지셋 다중 결합원인 분석(fuzzy-set multiple conjunctural causations analysis)에서 Y-검증과 N-검증에 대한 추가적 이해가 필요하다.

이러한 어려움에 도움을 제공하고자 본 연구는 퍼지셋 방법론의 사용이유를 보다 구체적으로 적시하고, Stata에서 제공하는 fuzzy 모듈을 소개하며, 퍼지셋 다중 결합원인 분석에서 필요한 Y-검증과 N-검증의 실증분석 결과를 제공하고자 한다. 특히 본 연구의 실증분석의 대상을 선진 서구국가 대신 탈사회주의 체제전환국으로 선정하고자 한다. 이를 통해 퍼지셋 방법론은 사회과학에서 익숙한 연구대상뿐만 아니라 특수한 환경에 놓인 새로운 연구대상에게도 적용 가능함을 강조하고자 한다. 이상의 문제제기와 연구 필요성을 통해 본 연구는 양적 연구방법과 질적 연구방법의 대화에 주목하는 연구자들에게 이론적 및 방법론적 도움을 주고자 한다.

II. 퍼지셋 방법론: 활용이유 및 절차

1. 퍼지셋 방법론의 활용이유

어떠한 경우에 퍼지셋 질적비교분석 방법을 활용하는가? 다양한 연구방법들 중에서 왜 퍼지셋 질적비교분석을 활용하는지는 다음과 같은 이유들로 열거될 수 있다. 첫째, ‘질적인, 사례중심의, small N의, 집중적인(intensive)’ 복잡성(complexity)을 추구하는 사례중심적 연구와 ‘양적인, 변수중심의, large N의, 광범위한(extensive)’ 일반화(generality)에 주목하는 변수중심적 연구 사이에서 다양성을 추구(*ibid.* 21)하기 때문이다. 요컨대 퍼지셋 연구의 핵심 철학은 변수중심 전략과 사례중심 전략을 넘어선 ‘다양성 지향’ 연구를 추구한다는 것이다. 다양성은 2가지 맥락에서 이해되어야 하는데 하나는 종류(kind)의 다양성이고, 다른 하나는 정도(degree)의 다양성이다. 먼저, 종류의 다양성은 질적 차이(qualitative distinctions)로 복합적 속성들을 배열화하는 것으로부터 나오는 종류의 차이(differences in kind)인 반면, 정도의 다양성은 양적 차이(quantitative distinctions)로 하나의 범주

에 속해있는 정도의 차이(variation by degree)라고 할 수 있다(*ibid.* 149). 퍼지셋 질적비교분석은 이러한 두 개의 다양성 문제를 해결하므로 강력한 분석도구인 것이다.

그런데 2개의 전통적 연구방법이 가지는 단점을 보완한다고 하지만, 퍼지셋 질적비교연구는 주로 변수중심적 접근을 비판하면서 사례중심적 접근의 장점을 취하고 있는 것이 사실이다. 예컨대 사례를 지형으로 이해하기, 모집단을 유연하고 조작가능한 구성체로서 이해하기, 구체적인 맥락 속에서의 질적 변화를 다룸으로써 결과중심적 연구를 강조하기, 인과관계를 결합적이고 이질적으로 이해하기라는 사례중심적 접근의 실용적 특징을 취하고 있다(*ibid.* 39-40). 이상의 설명을 바탕으로 퍼지셋 질적비교분석을 주요 연구방법으로 채택할 수 있는 이유는 무엇보다도 분석대상이 2-3개의 국가가 아니며, 30개 미만(large N)인 중범위 사례를 대상으로 하기 때문이다. 또한 파악하고자 하는 속성에 있어 분석국가들은 모두 동질적이지 않기 때문에, 이질적 속성을 파악함으로써 퍼지셋 질적비교분석의 장점인 유형과 정도에 있어서의 차이를 동시에 분석할 수 있기 때문이다. 그리고 추상적 개념을 조작적으로 정의한 기존의 ‘변수’에 해당하는 지수를 사용하기 어렵거나, 사용한다 할지라도 그 ‘해석’에 있어서 변수중심 해석에 머무르지 않고 변수지향적 및 사례지향적 연구 간의 다양성을 추구하면서 사례중심적 접근의 실용적 특징을 강조하고자 할 경우에 추천할 수 있다.

둘째, 사례를 맥락적으로 이해하고 지형으로 이해하는 핵심은 그 사례를 구성하는 주요한 속성들의 혼합(combination)으로 이해하는 것인데, 퍼지셋 질적비교분석 방법이 적합하기 때문이다(*ibid.* 66). 즉 퍼지셋 질적비교분석에서는 맥락 문제(context matter)가 바로 사례에 대한 지형적 접근의 핵심이며, 개별변수가 사례를 결정하는 것이 아닌 속성들의 혼합으로써 사례의 유형을 이해할 수 있다. 또한 지형적 사고의 주요 원칙은 중요한 요소가 변화하면, ‘사례 전체(whole case)’의 성격이 질적으로 바뀔 수도 있다는 것이다(*ibid.* 70). 예컨대, 북한출

신, 서울거주, 중산층, 남성 전문직의 속성들은 똑같은데, 핵심적 요소라고 여겨지는 정당지지 성향에서 보수정당과 진보정당으로 나뉜다면 사례 전체의 성격이 질적으로 다를 수 있다는 의미이다. 그리고 이러한 속성들은 5가지 변수(출신지, 거주지, 소득계층, 직업분류, 정당성향)로 보는 것이 아니라 사례들의 유형으로서 $32(2^5)$ 개의 지형으로 보는 것이다. 나아가 이러한 각각의 유형들은 특성 공간(property space), 즉 이상형(ideal type)을 구축한다. 퍼지셋을 활용하는 연구는 이러한 각 사례들의 배열이 만들어 낸 각각의 특성 공간을 통해 사례들을 유형화할 수 있다. 나아가 한 시점에 머무르지 않고 이러한 유형의 역사적 변화, 즉 어떠한 경로를 통해 변화 궤적을 보였는지에 대한 ‘질적 변화의 맥락’에 관심을 가질 때 퍼지셋 이상형 분석(fuzzy-set ideal type analysis)을 실시할 수 있다.

셋째, 사회현상의 원인을 단일 변수로 단순화하는 가정들을 지양해야 하며, 사회적 다양성을 평가하기 위해 인과적 복잡성(causal complexity)을 최대한 허용하는 연구전략이 필요(ibid. 93)할 때 퍼지셋 질적비교분석 방법이 장려될 수 있다. 요컨대, ‘단일’ 원인이 결과의 필요조건이나 충분조건이라고 추정하거나 인과관계의 본질을 ‘단순화’하는 가정은 지양되어야 한다. 즉 “특정한 사회현상을 낳는 원인조건들은 두 개 이상의 원인들이 결합되어 인과관계가 성립된다는 다중 결합적 인과관계를 의미한다”(Ragin 1987; Ragin, 2000, 104 재인용). 예를 들어, 한 나라의 노동시장정책은 노조조직률이라는 단일변수로 결정되는 것이 아니라 기업지배구조, 노동운동의 조직화, 노동시장의 유연화, 복지정권의 성격 등과 같은 복합적인 원인들의 결합으로 그 인과성을 검증할 수 있다. 또한 지형으로써 사례를 다룬다(working with cases as configurations)는 것은 퍼지셋의 백미인데, 각 사례에 대한 속성을 배열로 이해하고자 하기 때문에 퍼지셋 결합원인 분석을 실시하는 것이다. 즉 퍼지셋 질적비교분석에서의 각 사례는 다중 소속(multiple membership)으로 설명되는데, 다중 소속은 지형으로써 해

석할 수 있다. 특히 변수중심 연구에서 회귀계수와 같은 수치로만 해석되었던 한계를 넘어, 배열을 구성하는 해당 수치에 대해 언어적 의미를 투영할 수 있다는 점이 또 하나의 장점이다. 이처럼 본 연구는 하나의 변수가 독립적으로 종속변수를 설명하는 축소주의적 단순화 결론을 지양하고, 원인집합들의 결합적 산물로서의 해석을 지양하기 때문에 퍼지셋 다중 결합원인 분석을 지지한다. 결국 OLS 회귀분석에서는 독립변수 간 다중공선성의 문제로 인해 유사한 성격의 독립변수를 하나의 회귀모형에 투입할 수 없으나, 퍼지셋 질적비교분석에서는 유사한 성격의 원인조건이라 할지라도 연구자가 이론적으로 그 지수를 설명할 수만 있다면 모델에 포함할 수 있으며 오히려 그 원인조건들 간의 결합을 받기는 것이다.

2. 퍼지셋 방법론의 절차

이번에는 퍼지셋 질적비교분석 방법론의 실제 과정은 어떠한지에 대한 구체적인 절차를 살펴보고자 한다. 그 과정은 첫째, 원점수의 퍼지점수로의 변환의 과정인 ‘측정’, 둘째, 불리언대수를 이용한 퍼지 소속점수의 ‘연산’, 셋째, 필요조건과 충분조건을 집합관계의 ‘검증’, 넷째, 간명성 추구를 위한 ‘축약’이라고 할 수 있다(정해식 2012; Ragin 2000, 2008a).

1) 측정(calibration)

일부 사회과학자들이 퍼지셋을 비판하는 논거 중의 하나가 측정에 다. 그들은 퍼지점수로 측정한다는 것이 단지 이산변수를 연속변수로 변환한 것이라고 주장하지만 “퍼지셋은 보다 이론적이며 실제적인 지식을 포함한 값이므로 단순한 연속변수 이상의 의미를 갖는다”(Ragin 2000, 6). 또한 Ragin(2008b, 17)은 “변수의 측정값을 변환하는 것이 변환하지 않는 것보다 확실히 열등하다”고 지적하였고, “예를 들어 변

환되지 않는 민주주의의 측정은 한 국가가 다른 국가보다 더 민주적이라거나 평균보다 민주적이라는 것을 알 수는 있지만, 보다 민주적인지 보다 '전체주의적'인지는 알 수 없다"고 주장한다. 이러한 그의 주장은 요지는 바로 '언어적' 해석에 있다.

퍼지셋 질적비교분석이 도입되기 이전의 고전집합 질적비교분석(crisp-set QCA)에서의 멤버십은 0과 1이라는 '존재와 부존재'의 이분법으로 설명되지만, 퍼지셋에서는 0과 1사이의 다양한 값들을 이론 및 사례의 속성에 맞게 그 '언어적 의미'를 부여할 수 있다는 장점이 있다. 예를 들어, 임금협상에서 '산별노조의 강건성'을 5분위 퍼지셋으로 분류할 경우, 0(fully out: 완전 비소속), .25(more out than in: 소속보다 비소속), .5(neither out nor in: 소속도 비소속도 아님), .75(more in than out: 비소속보다 소속), 1(fully in: 완전 소속)로 멤버십 점수(membership score)를 부여할 수 있다. 그리고 그 해석을 산별노조의 강건성은 임금협상에 있어서 확실히 중요하지 않음(0), 아마도 중요하지 않음(.25), 중요하지도 중요하지 않지도 않음(.5), 아마도 중요함(.75), 확실히 중요함(1)으로 풀이될 수 있다.

따라서 이러한 퍼지셋 멤버십은 전통적인 비율척도보다 높은 측정 형태라고 평가된다. 왜냐하면 가장 낮은 수준의 측정 형태인 명목척도(전통적인 crisp set)부터, 서열척도(범주화, 위계적 순위), 등간척도(범주화, 위계적 순위, 등간격, 해석가능한 값의 표현), 비율척도(범주화, 위계적 순위, 등간격, 해석가능한 값의 표현, 의미있는 0값)까지로 분류할 때, 퍼지 멤버십은 비율척도의 특성을 모두 갖고 있으면서도 의미있는 '최댓값(1=full membership)'도 보유하고 있기 때문이다(Ragin 2000, 154-155). 이와 같은 퍼지점수의 측정은 5분위 퍼지셋 이외에도 3분위 퍼지셋과 7분위 퍼지셋 및 연속 퍼지셋으로도 가능한데 다음과 같다.

<표 1> 고전집합 대 퍼지셋(Crisp versus Fuzzy Sets)

Crisp set	3분위 퍼지셋	5분위 퍼지셋	7분위 퍼지셋	연속 퍼지셋
1 = 완전 소속	1 = 완전 소속	1 = 완전 소속	1 = 완전 소속	1 = 완전 소속
		.75 = 비소속보다 소속	.83 = 거의 소속이나 완전히 소속은 아님	멤버십 수준이 비소속보다 소속을 나타내는 수리점수 ($.5 < x_i < 1$)
			.67 = 다소 소속	
	.5 = 완전소속도 비소속도 아님	.5 = 분기점: 소속도 비소속도 아님	.5 = 분기점: 소속도 비소속도 아님	.5 = 분기점: 소속도 비소속도 아님
			.33 = 다소 비소속	
		.25 = 소속보다 비소속	.17 = 거의 비소속이나 완전히 비소속은 아님	멤버십 수준이 소속보다 비소속을 나타내는 수리점수 ($0 < x_i < .5$)
0 = 완전 비소속	0 = 완전 비소속	0 = 완전 비소속	0 = 완전 비소속	0 = 완전 비소속

출처: Ragin(2000, 156)

단지 소속과 비소속이라는 이분법적 구분만 존재했던 crisp-set 질적비교분석은 퍼지셋 질적비교분석의 도입을 통해 3분위, 5분위, 7분위 퍼지셋뿐만 아니라 연속 퍼지셋 개념까지 확장함으로써 해석의 다양성을 담보할 수 있게 된 것이다. 이후 퍼지셋 질적비교분석은 보다 수리적 접근에 친착함으로써 한 단계 발전하게 되었다. 그 발전은 바로 Ragin(2008a: 88)이 '언어적 표기의 수리적 변환(mathematical

translations of verbal labels)’을 통해 수준기표(benchmark proportion)의 언어적 해석 및 수학적 변환의 메커니즘을 다음과 같이 보다 선명하게 제시함으로써 가능하게 된 것이다.

<표 2> 언어적 표기의 수리적 변환

언어적 표기	소속의 정도 (Degree of membership)	소속의 공산 (Associated odds)	소속의 로그 오즈 (Log odds of full membership)
완전 소속(Full membership)	0.993	148.41	5.0
완전 소속의 한계점 (Threshold of full membership)	0.953	20.09	3.0
거의 소속(Mostly in)	0.881	7.39	2.0
비소속보다 소속(More in than out)	0.622	1.65	0.5
분기점(Crossover point)	0.500	1.00	0.0
소속보다 비소속(More out than in)	0.378	0.61	-0.5
거의 비소속(Mostly out)	0.119	0.14	-2.0
완전 비소속의 한계점 (Threshold of full nonmembership)	0.047	0.05	-3.0
완전 비소속(Full nonmembership)	0.007	0.01	-5.0

출처: Ragin(2008a, 88)

위 표에서 확인할 수 있는 것처럼 소속의 정도로 표현된 퍼지 소속점수는 0부터 1사이의 다양한 값을 갖는다. 소속의 공산에서 소속의 로그오즈로 전환한 후, 최종적으로 퍼지 소속점수로 변환하는 공식은 다음과 같다(Ragin 2008a, 85-88). 공식에서 나타난 바와 같이 먼저 사건이 발생하지 않을 확률과 발생할 확률의 비율인 공산(=odds, 승산)의 값에 자연로그를 취하여 로짓이라고 부르는 로그오즈(log odds)를 구한 후, 로지스틱 함수의 수식과 같은 ‘소속의 정도’가 바로 퍼지

소속점수가 되는 것이다.

$$\text{소속의 공산(associated odds)} = \frac{\text{소속의 정도}}{1 - \text{소속의 정도}}$$

$$\text{소속의 로그오즈(log odds of full membership)} = \ln \frac{\text{소속의 정도}}{1 - \text{소속의 정도}}$$

$$\text{소속의 정도(degree of membership)} = \frac{\exp(\log odds)}{1 + \exp(\log odds)}$$

이상의 설명이 퍼지점수로 변환되는 언어적 및 수학적 설명이었다면, 실제 실증적인 퍼지 소속점수로의 변환 단계는 다음과 같이 설명할 수 있다. Ragin(2000, 165-171)은 퍼지 소속점수의 측정을 총 6개의 단계로 구분하여 설명하는데, 1단계는 측정 범위(domain)의 구체화, 2단계는 개념으로부터 퍼지집합에 대한 정의, 3단계는 각각의 개념에 적합한 퍼지집합 유형에 대한 결정, 4단계는 퍼지 소속점수의 있을 법한 범위에 대한 결정, 5단계는 퍼지 소속점수를 매기기 위한 경험적 증거의 확인, 6단계는 경험적 증거를 점수로 변환하는 것이라고 한 바 있다. 이상과 같은 6단계를 거쳐 해당 지수의 원점수는 퍼지점수로 변환되게 된다. 특히 원점수를 연속 퍼지점수로 변환하기 위해서는 3가지의 중요한 고정축(anchors)이 선정되어야 하는데, 그것은 첫째, 완전한 소속의 값, 둘째, 완전한 비소속의 값, 셋째, 소속과 비소속을 가르는 분기점(crossover point)이다(Ragin 2008a, 89). 따라서 연구자는 하나의 퍼지집합을 구조화하는 세 개의 질적 기준점(benchmarks)에 대응하도록 본래의 변수값을 변환해야 한다. 원자료에 해당하는 등간 또는 비율 척도값을 퍼지 소속점수로 변환하기 위해서는 3개의 기준점이 필요한데, 완전한 소속(퍼지 점수 = 0.95), 완전한 비소속(퍼지 점수 = 0.05), 분기점(퍼지 점수 = 0.5)으로 구분되며, 그 측정을 위한 변환은 로그 오즈에 기초한다(Ragin 2008b, 18).

이후 본 논문의 퍼지셋 실증 연구에서는 연속 퍼지변수를 사용한다. 그 이유는 본 연구의 실증분석에서 사용되는 지수의 특성상 3분

위, 5분위, 7분위와 같은 구분은 오히려 정보의 손실을 야기하기 때문이다. 따라서 각 지수의 원 점수를 구한 후, 각 지수의 중위수를 분기점으로 하며 각 지수의 최댓값과 최솟값을 완전 소속과 완전 비소속으로 하는 3개의 고정축을 활용하여 퍼지 연속점수로 환산할 것이다. 따라서 완전소속의 퍼지값은 .95로, 완전한 비소속의 퍼지값은 .05, 분기점은 .50으로 고정하며, 각 지수의 최댓값, 최솟값, 중위값이 이에 해당한다. 평균을 분기점으로 하지 않는 이유는 중위수가 사례수의 50:50으로 소속과 비소속을 가르는 기준이 되며(Ragin 2008a, 87), 평균을 사용할 때 발생하는 불필요한 변량이 발생하지 않기 때문이다(정해식 2012). 무엇보다도 중위수를 사용하는 이유는 연구자의 ‘자의성’을 철저히 배제하기 위한 의도이다. 또한 환산된 각 퍼지점수의 왜도(skewness)를 계산한 결과, 정규분포를 크게 왜곡하지는 않아 중위수의 값을 사용해도 무방하다고 판단한다.

2) 연산(operation)

퍼지셋 연산은 불리언 대수를 활용한 방식을 차용하고 있으며 다음과 같이 요약된다(Ragin 2000, 71-176). 연산은 불리언 대수에서 활용하는 세 가지 집합 공식과 같은데 합집합(\cup), 교집합(\cap), 여집합(\sim)으로 구성된다. 먼저 합집합의 경우 퍼지셋 용어로는 logical or로 표현되며 기본법칙은 각 집합의 퍼지점수 중 최대 소속점수이며, 언어적 해석은 집합 A에 속하거나 집합 B에 속하거나 또는 둘 다에 속한다고 할 수 있다. 다음으로 교집합의 경우 퍼지셋 용어로는 logical and로 표현되며 기본법칙은 각 집합의 퍼지점수 중 최소 소속점수이며, 언어적 해석은 집합 A에 속하면서 집합 B에도 속한다고 할 수 있다. 마지막으로 여집합의 경우 퍼지셋 용어로는 negation으로 표현되며 기본법칙은 1에서 각 퍼지점수를 빼준 점수이며, 언어적 해석은 집합 A에 속하지 않는다고 할 수 있다. 이상의 내용을 다음의 표에 종합한다.

<표 3> 퍼지셋 연산(operations on fuzzy-sets)

	기호	표시	퍼지셋 용어	기본법칙	언어적 해석
합집합	$\cup(+)$	A \cup B	logical or	maximum	집합 A에 속하거나 집합 B에 속하거나 또는 둘 다에 속하는
교집합	$\cap(*)$	A \cap B	logical and	minimum	집합 A에 속하면서 집합 B에도 속하는
여집합(부정)	\sim	\sim A	negation	1-A	집합 A에 속하지 않은

출처: Ragin(2000, 71-176)

3) 집합관계의 검증(evaluating set relations)

(1) 일관성과 설명력(consistency and coverage)

Ragin(2006)은 집합관계를 검증하기 위해 일관성(consistency)과 설명력(coverage)이라는 두 개의 기준을 제시한다. “일관성이 필수적으로 고려되는 원인조건에 동의하는 결과집합의 정도를 평가하는 반면, 설명력은 필요조건과의 관련성을 평가하는데 필요조건이 결과의 사례들과 병행되는 정도를 평가한다”(Ragin 2008a, 44-45). 보다 구체적으로 설명한다면, 먼저 일관성은 유의성(significance)과 같은 경험적 연결이 연구자의 주목을 받을만한지 아닌지를 의미한다고 할 수 있는데, 가령 하나의 가설화된 집합관계가 일관적이지 않다면 연구자의 이론 또는 결론은 지지될 수 없다(ibid. 45). 따라서 일관성은 연구자의 ‘이론적 주장이 얼마나 지지될 수 있는지’를 검증한다. 다음으로 설명력은 강도(strength)와 같이 경험적 관련성 또는 집합이론적 연결의 중요성을 설명한다고 할 수 있다(ibid.). 따라서 설명력은 연구에 포함된 ‘사례들이 연구자의 이론에 의해 얼마나 설명될 수 있는지’를 검증한다. 예컨대 상관관계 분석에서 통계적으로 유의하지만 낮은 상관계수를 보이는 경우에는 일관성은 높으나 설명력은 낮다고 할 수 있다. 유사한 예로써, 회귀분석에서 F검증을 통한 모델검증 및 각 회귀계수가 통계적으로 유의하지만 r^2 값이 낮은 경우도 일관성이 높으

나 설명력은 낮다고 할 수 있다. 본 연구의 실증분석에서는 Ragin(2006, 2008a)의 권고대로 집합관계 확인을 위해 일관성과 설명력 모두를 검증하며 각각의 값을 제시하고 해석한다.

(2) Y-일관성 검증과 N-일관성 검증

먼저 일관성 검증을 위해서는 일관성 값의 검증 기준이라 할 수 있는 수준기표를 선정해야 한다. Ragin(2000, 109)은 “원인조합의 사례들이 .80보다 유의하게 크다면 그 원인조합은 ‘거의 항상(almost always)’ 충분하다고 주장할 수 있다”고 언급한 바 있다. 또한 “일반적으로(more often than not; .50), 대개(usually, .65), 거의 항상(almost always; .80)과 같은 언어적 검증을 이용한 원인조합의 준 충분성(quasi-sufficiency)을 검증하는 것도 가능하다”(ibid.)고 주장한다. 본 연구에서는 ‘체제전환국’에 대한 퍼지셋 질적비교분석이 처음의 시도라 추정되므로 보수적 기준을 적용하여 일관성 검증 기준값으로 수준기표 .80을 선정한다. 따라서 ‘거의 항상’ 충분하다고 주장될 수 있는 원인조건들을 해석한다.

일관성의 검증은 Y-일관성 검증과 N-일관성 검증으로 구분된다. Y-일관성 검증은 수준기표를 선정한 후, 이 수준기표와 비교하여 해당 모델이 얼마나 유의하게 차이가 나는지를 검증하는 방법이다. 다음으로 N-일관성 검증은 X집합이 Y집합의 하위 집합이 되는 일관성과 X집합이 Y 여집합의 하위 집합이 될 일관성이 얼마나 유의하게 차이가 나는지를 검증하는 방법이다(정해식 2012, 150). 예를 들어, 코포라티즘적 임금협상 집합의 소속점수가 강한 산별노조를 가지는 집합의 소속점수보다 크거나 같은 경향에 일관성이 있다면, 강한 산별노조는 코포라티즘적 임금협상을 실현하는데 충분하다고 평가할 수 있다. 반면 특정 사례들에서 강한 산별노조를 가지는 집합의 소속점수가 높음에도 불구하고 코포라티즘적 임금협상의 소속점수가 낮다면, 강한 산별노조가 코포라티즘적 임금협상에 충분하다는 주장은 완전히 타당하다고 할 수 없다. 바로 N-일관성 검증은 원인조건이

결과조건의 하위 집합이 되는 일관성과 원인조건이 결과조건의 여집합의 하위 집합이 될 일관성을 비교함으로써 일관성 정도를 검증하는 것이다. 기존의 퍼지셋 질적비교분석의 필요조건과 충분조건 검증을 통한 일부 연구들에서 Y-일관성 검증만으로 해석을 해 온 관습이 있는데, 이상의 설명과 같이 Y-일관성 검증과 N-일관성 검증을 동시에 검증하는 것이 논리적으로 보다 타당한 방법이라 할 수 있다. 따라서 본 연구의 충분조건 검증에서는 보수적 기준을 적용하여 Y-일관성 검증과 N-일관성 검증을 동시에 통과한 결합원인 배열만을 해석한다.

(3) 필요조건 검증

필요조건을 검증하기 위해서는 결과가 원인의 부분집합이라는 것을 밝혀야 한다. 즉 결과집합 Y가 원인집합 X의 하위 집합(subset)이 되는 경우를 결과집합에 대해 원인집합을 필요조건이라 부를 수 있다. 따라서 필요조건은 결과가 존재하면 원인이 항상 존재한다는 의미로써 결과집합의 소속점수가 원인조건의 소속점수보다 낮은 것이다. 필요조건의 일관성 값(Ragin 2008a, 61) 및 설명력 값(ibid. 53)은 다음과 같이 계산된다.

$$Consistency (Y_i \leq X_i) = \sum [\min (X_i, Y_i)] / \sum Y_i$$

$$Coverage (Y_i \leq X_i) = \sum [\min (X_i, Y_i)] / \sum X_i$$

(4) 충분조건 검증

충분조건을 검증하기 위해서는 원인이 결과의 부분집합이라는 것을 밝혀야 한다. 즉 원인집합 X가 결과집합 Y의 하위 집합(subset)이 되는 경우를 결과집합에 대해 원인집합을 충분조건이라 부를 수 있다. 따라서 충분조건은 원인이 존재하면 항상 결과가 존재한다는 의미로

써 결과집합의 소속점수가 원인조건의 소속점수보다 높은 것이다. 충분조건의 일관성 값(ibid. 52) 및 설명력 값(ibid. 57)은 다음과 같이 계산된다.

$$Consistency(X_i \leq Y_i) = \sum[\min(X_i, Y_i)] / \sum X_i$$

$$Coverage(X_i \leq Y_i) = \sum[\min(X_i, Y_i)] / \sum Y_i$$

4) 축약

이상의 과정을 통해 도출된 원인조건 조합들은 보다 축약된 형태로 제시될 수 있다. 퍼지셋 질적비교분석에서 원인조합의 배열을 최종적으로 결정할 때, 축약의 정도를 어느 수준까지 할 것인가는 매우 논쟁적인 문제이다. 왜냐하면 fuzzy set 2.0에서는 3가지의 해법을 동시에 제시하고 있기 때문이다. 첫째는 복잡한 해법(complex solution)이고, 둘째는 간명한 해법(parsimonious solution)이며, 셋째는 중간 해법(intermediate solution)이다. 먼저 복잡한 해법은 간명성 전략과 해석의 복잡성으로 인해 일반적으로 연구결과로 제시되지 않는다. 그리고 간명한 해법은 다중결합적 인과관계 전략을 강조하는 퍼지셋의 본성상 매우 간명한 1-2개의 원인조합으로 결과집합을 설명하는 단순화된 가정으로 인해 해석의 다양성에 제한을 두기 때문에 잘 활용되지 않는다. 결국 복잡성과 간명성 간의 제한된 다양성이라는 입장(Ragin & Sonnett 2004) 하에 대부분 중간 해법이 연구결과로 보고된다. 특히 Ragin(2008b, 79)은 퍼지셋 유저가이드에서 명시분석(Specify Analysis)보다 표준분석(Standard Analysis)을 추천하는데, 그 이유가 바로 표준분석에서만 중간해법을 제시해주기 때문이라고 안내한 바 있다. 다행히도 Stata 12.0 Fuzzy 모듈에서는 단 하나의 축소집합을 제시한다. 따라서 본 연구의 충분조건 분석에서는 Y-일관성 및 N-일관성 검증을 동시에 통과한 '최소배열 축소집합(Minimum

Configuration Reduction Set)'을 축약된 원인조건 배열로 제시한다.¹⁾

III. 실증 연구: 체제전환국의 사회지출에 대한 퍼지셋 결합원인 분석

1. 주요지수의 기술통계 및 퍼지점수의 측정

1) 주요지수의 기술통계

퍼지셋 질적비교분석을 위한 분석대상 국가는 22개의 탈사회주의 체제전환국이다. 22개 국가들은 그루지야(GEO), 라트비아(LVA), 러시아(RUS), 루마니아(ROM), 리투아니아(LTU), 몰도바(MDA), 벨라루스(BLR), 불가리아(BGR), 슬로바키아(SVK), 슬로베니아(SVN), 아르메니아(ARM), 아제르바이잔(AZE), 알바니아(ALB), 에스토니아(EST), 우크라이나(UKR), 체코(CZE), 카자흐스탄(KAZ), 크로아티아(HRV), 키르기스스탄(KGZ), 타지키스탄(TJK), 폴란드(POL), 헝가리(HUN)이다.

본 연구의 퍼지셋 결합원인 분석의 결과집합은 사회지출이며, 원인집합은 경제발전, 민주주의, 노인인구, 무역, 사회보장 기간, 사회주의 기간으로 구성된다. 먼저 결과집합인 GDP 대비 사회지출 수준을 보면, 우크라이나가 19.8%로 가장 높고, 슬로베니아, 헝가리, 폴란드, 에스토니아의 순이다. 대부분의 CIS 국가들은 낮은 사회지출을 보이며, 특히 중앙아시아에 속한 세 나라가 가장 낮은 사회지출을 보인다.

1) 본 연구는 퍼지셋 질적비교분석을 위해 Stata 12.0 Fuzzy 모듈을 이용한다. 그 이유는 첫째, fsQCA 2.0 소프트웨어는 2009년 1월 마지막 업데이트 이후에도 여전히 '점검 중'이기 때문이다(FSQCA, 2014). 둘째, 국제 STATA저널(The Stata Journal)에 Stata를 활용한 fsQCA분석이 소개됨으로써 훨씬 진보된 모듈을 제공하기 때문이다(Longest & Vaisey, 2008). 특히 Y 일관성(Y-consistency) 및 N 일관성(N-consistency) 동시 검증이 Stata Fuzzy 모듈에서는 가능하다.

반면 구 유고연방의 슬로베니아와 크로아티아, 비셰그라드 국가들 중 헝가리와 폴란드, 발틱 국가들 중 에스토니아와 리투아니아는 높은 사회지출을 보인다.

<표 4> 퍼지셋 결합원인 분석을 위한
주요지수의 원점수(N=22)

국가	원인집합						
	결과집합	원인집합					
	사회 지출	경제 발전	민주 주의	노인 인구	무역	사회 보장 기간	사회 주의 기간
	EXP	GDP	DEM	AGE	TRA	YIS	SOP
ALB	8.2030	8835.6037	7.4333	10.0929	43.0045	320	46
ARM	7.1976	5740.0786	5.4167	10.5746	32.8278	367	55
AZE	7.1746	9789.9735	3.9122	5.9273	39.1231	296	69
BLR	13.1772	13850.8713	3.9822	13.9617	66.4447	340	69
BGR	13.2711	14200.1147	8.7000	18.3509	58.8063	422	42
HRV	14.5372	19173.6573	8.5833	17.5606	40.2261	356	48
CZE	13.2206	25718.0666	9.6667	15.3936	64.2204	438	43
EST	15.1930	20519.7275	9.5667	17.4927	75.9274	370	51
GEO	6.8410	5082.9941	6.3500	14.2461	42.9220	168	70
HUN	18.2277	20812.4337	8.9833	16.7117	82.6315	441	40
KAZ	4.3844	12175.7604	4.1333	6.6802	37.7091	106	66
KGZ	4.2702	2303.2715	5.2600	4.4509	67.1320	424	69
LVA	13.0837	16991.5601	8.7833	18.3563	53.3022	301	51
LTU	14.4525	18873.7791	9.3333	15.5040	67.5633	363	47
MDA	14.2362	3108.2250	6.8500	11.1714	59.8496	141	47
POL	16.6799	19988.3418	9.0000	13.5570	41.1232	353	44
ROM	10.7171	15134.8112	8.5333	14.8497	37.5435	360	44
RUS	12.2462	20801.8009	5.3167	13.1214	25.1615	453	74
SVK	12.2145	23269.0040	9.1833	12.3137	80.2942	438	43
SVN	18.4481	28856.3658	9.7000	16.6563	64.5105	413	48
TJK	3.4845	1958.2042	3.6333	3.3126	35.5875	40	62
UKR	19.8270	6739.2453	6.8167	15.7646	52.1560	444	69

다음은 결과집합과 원인집합에 해당하는 주요지수의 기술통계이다. 평균, 중위수, 표준편차, 최댓값, 최솟값을 각각 제시한다.

<표 5> 퍼지셋 결합원인 분석을 위한
주요지수의 기술통계(N=22)

주요 기술통계	결과 집합	원인집합					
	사회 지출	경제 발전	민주 주의	노인 인구	무역	사회 보장 기간	사회 주의 기간
	EXP	GDP	DEM	AGE	TRA	YIS	SOP
평균	11.87	14178.36	7.23	13.00	53.09	334.27	54.41
중위수	13.13	14667.46	7.98	14.10	52.73	361.50	49.500
표준편차	4.76	7771.45	2.14	4.51	16.53	117.95	11.53
최솟값	3.49	1958.20	3.63	3.31	25.16	40.00	40.00
최댓값	19.83	26856.37	9.70	18.36	82.63	453.00	74.00

다음은 결과집합과 원인집합에 포함된 지수들 간 피어슨의 상관관계 분석표이다. 경제발전, 민주주의, 고령인구, 노인인구, 무역, 사회보장 기간은 결과조건인 사회지출과 정의 상관관계를 보인 반면, 사회주의 기간은 결과조건인 사회지출과 부의 상관관계를 맺고 있다. 상관관계의 정도로 설명하면, 결과조건인 사회지출은 노인인구와 가장 높은 상관관계($r=.799$)를 보이며 민주주의, 경제발전, 사회보장 기간, 무역, 사회주의 기간의 순이다.

<표 6> 퍼지셋 결합원인 분석을 위한
주요지수 간 상관관계(N=22)

	EXP	GDP	DEM	AGE	TRA	YIS	SOP
EXP	1						
GDP	.606**	1					
DEM	.683***	.694***	1				
AGE	.799***	.631**	.767***	1			
TRA	.456*	.364	.505*	.346	1		
YIS	.563**	.568**	.516*	.517*	.387	1	
SOP	-.447*	-.498*	-.818***	-.487*	-.409	-.261	1

주: 유의수준, * < .05, ** < .01, *** < .001

2) 퍼지점수의 측정

연구방법에서 설명한 바와 같이 각 지수의 최댓값, 최솟값, 중위값이라는 3개의 고정축을 이용하여 연속 퍼지변수로 변환한다.²⁾

<표 7> 퍼지셋 결합원인 분석을 위한
주요지수의 퍼지점수(N=22)

국가	원인집합						
	결과집합	경제	민주	노인	무역	사회보장	사회주의
	사회 지출 퍼지	발전 퍼지	주의 퍼지	인구 퍼지	퍼지	기간퍼지	기간퍼지
	EXPF	GDPF	DEMF	AGEF	TRAF	YISF	SOPF
ALB	0.1776	0.2016	0.4063	0.2469	0.2576	0.4044	0.2488
ARM	0.1364	0.1084	0.1455	0.2727	0.1029	0.5450	0.6623
AZE	0.1356	0.2402	0.0569	0.0934	0.1853	0.3518	0.9159
BLR	0.5052	0.4520	0.0596	0.4901	0.7984	0.4500	0.9159
BGR	0.5157	0.4724	0.7777	0.9524	0.6479	0.8791	0.0856
HRV	0.6525	0.7520	0.7405	0.9197	0.2041	0.4872	0.3837
CZE	0.5101	0.9382	0.9499	0.7130	0.7600	0.9247	0.1138
EST	0.7159	0.8085	0.9409	0.9161	0.9111	0.5692	0.5458
GEO	0.1239	0.0943	0.2448	0.5251	0.2559	0.1412	0.9249
HUN	0.9075	0.8194	0.8516	0.8629	0.9526	0.9313	0.0474
KAZ	0.0618	0.3571	0.0657	0.1127	0.1632	0.0844	0.8829
KGZ	0.0598	0.0512	0.1326	0.0640	0.8092	0.8859	0.9159
LVA	0.4964	0.6392	0.8019	0.9526	0.5144	0.3625	0.5458
LTU	0.6439	0.7379	0.9137	0.7286	0.8158	0.5123	0.3123
MDA	0.6245	0.0613	0.3140	0.3068	0.6714	0.1133	0.3123
POL	0.8306	0.7874	0.8553	0.4621	0.2205	0.4802	0.1497
ROM	0.3207	0.5287	0.7234	0.6286	0.1608	0.4965	0.1497
RUS	0.4317	0.8190	0.1372	0.4321	0.0474	0.9526	0.9526
SVK	0.4293	0.8926	0.8906	0.3781	0.9408	0.9247	0.1138
SVN	0.9155	0.9526	0.9526	0.8582	0.7653	0.8440	0.3837
TJK	0.0474	0.0474	0.0474	0.0474	0.1341	0.0474	0.8221
UKR	0.9526	0.1334	0.3090	0.7634	0.4844	0.9373	0.9159

2) 퍼지점수로 변환하는 기술적인 방법은 2가지이다. 기존의 방식대로 fsQCA 2.0 소프트웨어를 활용해도 된다. 점검 중이지만 변환값은 문제없이 산출된다. 다음으로 Stata 소프트웨어에 Fuzzy module을 검색하여 설치한 후 가능하다. 롱이스트와 바이세이(Longest and Vaisey 2008)의 논문을 통해 원하는 명령어를 확인할 수 있다. 다만 fsQCA 2.0이 소수점 둘째자리까지밖에 제공하지 못하여 유사한 값들 간 비교를 할 수 없지만, Stata는 그 이상의 수치제시가 가능하다.

2. 모델특정화

이론적 논의에 따라 체제전환국의 사회지출(EXP)을 결과집합으로 하고 원인조건으로는 1인당 GDP(PPP)(GDP), BTI 민주주의(DEM), 인구고령화(전체 인구 대비 65세 이상 노인인구 비율)(AGE), GDP 대비 무역(수입+수출) 비율(TRA), 사회보장제도(연금·장애·유족연금, 질병·출산수당, 실업, 산재, 가족수당) 기간의 총합(년)(YIS), 사회주의 기간(사회주의 건설년도부터 해체년도까지의 기간)(SOP)으로 구성한다. 이상의 설명을 토대로 한 분석모델은 다음과 같다. 분석모델은 모두 대문자로 표현하였고, 이후 표현되는 소문자는 해당 집합의 여집합이다.

$$EXPF = GDPF + DEMF + AGEF + TRAF + YISF + SOPF$$

3. 진리표

실제 분석을 시행하면 위의 모델에 해당하는 진리표(truth table)가 도출된다. 결과집합에 대한 원인집합의 진리표는 다음과 같다. 원인집합의 수가 6개이므로 실제 가능한 배열은 $64(2^6)$ 개다. 본 진리표에는 64개의 진리표를 모두 제시하지 않고 실제 사례수가 있는 16개의 배열만을 제시한다. 이 때 연구자는 어떤 수준에서 결과집합을 1로 할 것인지 0으로 할 것인지를 판단해야 한다. 본 연구에서는 다음과 같이 판단하였다. 체제전환국의 복지지출을 설명하는 원인을 검증하는 처음의 연구로 추정되므로 보다 보수적으로 판단하기 위해 '사례가 1개 이상 존재하고, 일관성 값이 .90 이상'인 배열만을 높은 사회지출 집합으로 판단하여 1로 표현하고 나머지 배열은 0으로 표현한다. 이 기준에 의거할 때 높은 사회지출 집합을 설명하는 배열은 총 8개이다.

8개의 배열 중 2개 이상의 사례가 포함된 2개의 배열만 설명하면 다음과 같다. 첫째, 높은 경제발전, 높은 민주주의, 높은 노인인구 비

을, 높은 무역 비율, 긴 사회보장 기간, 짧은 사회주의 기간의 배열 (GDATYs)은 사례수가 4개로 가장 많고, 일관성 값은 0.960으로 매우 높게 나타나며, 최적사례는 체코, 헝가리, 리투아니아, 슬로베니아이다. 둘째, 높은 경제발전, 높은 민주주의, 높은 노인인구 비율, 낮은 무역 비율, 짧은 사회보장 기간, 짧은 사회주의 기간의 배열(GDAtys)은 사례수가 2개이며, 일관성 값은 0.940으로 매우 높게 나타나며, 최적사례는 크로아티아와 루마니아이다.

<표 8> 결과집합에 대한 원인집합의 진리표

배열	원인집합						결과 집합	사 례 수	일 관 성 (Co ns.)	최적 사례 Bestfit
	경제 발전	민주 주의	노인 인구	무역	사회 보장 기간	사회 주의 기간	사회 지출			
	G	D	A	T	Y	S	E			
	GDPF	DEMF	AGEF	TRAF	YISF	SOFP	EXPF			
GDATYS	1	1	1	1	1	1	1	1	0.992	EST
GDATyS	1	1	1	1	0	1	1	1	0.985	LVA
gDATYs	0	1	1	1	1	0	1	1	0.973	BGR
GDatys	1	1	0	0	0	0	1	1	0.969	POL
GDATYs	1	1	1	1	1	0	1	4	0.960	CZE, HUN, LTU, SVN
GDAtys	1	1	1	0	0	0	1	2	0.940	HRV, ROM
GDaTYs	1	1	0	1	1	0	1	1	0.925	SVK
gdAtYS	0	0	1	0	1	1	1	1	0.920	UKR
GdatYS	1	0	0	0	1	1	0	1	0.890	RUS
gdaTyS	0	0	0	1	0	1	0	1	0.832	BLR
gdAtyS	0	0	1	0	0	1	0	1	0.790	GEO
gdaTYs	0	0	0	1	1	1	0	1	0.727	KGZ
gdatYS	0	0	0	0	1	1	0	1	0.726	ARM
gdatys	0	0	0	0	0	0	0	1	0.718	ALB
gdatyS	0	0	0	0	0	1	0	3	0.451	AZE, KAZ, TJK

4. 필요조건 검증

결과집합 Y가 원인집합 X의 하위집합(subset)이 되는 경우를 결과 집합에 대해 원인집합을 필요조건이라 부른다. 따라서 필요조건은 결과집합의 소속점수가 원인집합의 소속점수보다 낮은 것이다. 연구방법에서 밝힌 바와 같이 Y-일관성 및 N-일관성 검증을 모두 실시하며, 일관성 검증 기준(benchmark)은 .80으로 하고 유의수준을 .05로 설정하여 검증한다. 분석 결과, 사회지출에 대한 필요조건의 Y-일관성 및 N-일관성 검증을 ‘동시에’ 통과한 필요조건은 하나도 없다. 따라서 유의수준을 .10으로 낮추어 Y-일관성 및 N-일관성 검증을 ‘각각’ 실시하였다. 첫째, 사회지출에 대한 필요조건의 Y-일관성 검증 결과(benchmark \geq .80, $p < .10$), ‘높은 사회지출’ 집합에 소속되기 위한 필요조건은 없는 것으로 나타난다.

<표 9> 사회지출에 대한 필요조건의 Y-일관성 검증 결과

필요 조건	benchmark \geq .80, $p < .10$				설명력 (coverage)	검증결과	
	Y-일관성 (Y-Con)	검증기준값 (benchmark)	F	p		($p < .05$)	($p < .10$)
GDP	0.782	0.8	0.11	0.748	0.836		
DEM	0.739	0.8	1.37	0.255	0.821		
AGE	0.777	0.8	0.18	0.672	0.893		
TRA	0.735	0.8	0.90	0.352	0.779		
YIS	0.699	0.8	1.71	0.205	0.845		
SOP	0.503	0.8	9.08	0.007	0.558		

둘째, 사회지출에 대한 필요조건의 N-일관성 검증 결과($Y-con \geq N-con$, $p < .10$), ‘높은 사회지출’ 집합에 소속되기 위해서 경제발전, 민주주의, 노인인구가 대개(usually) 필요조건인 것으로 나타난다. Y-일관성 값이 아쉽게도 거의 항상(almost always, .80)의 수준에 조금 미치지 못하여 그 아래 수준인 대개(usually, .65)로 해석한다. 각각의 설명력은 경제발전 .836, 민주주의 .821, 노인인구 .893으로 높은 설명

력을 보이고 있다.

<표 10> 사회지출에 대한 필요조건의 N-일관성 검증 결과

필요 조건	검증기준 Y-con \geq N-con, p < .10				설명력 (coverage)	검증결과	
	Y-일관성 (Y-Con)	N-일관성 (N-Con)	F	p		(p<.05)	(p<.10)
GDP	0.782	0.563	3.51	0.075'	0.836		pass
DEM	0.739	0.505	4.03	0.058'	0.821		pass
AGE	0.777	0.551	3.85	0.063'	0.893		pass
TRA	0.735	0.581	1.33	0.261	0.779		
YIS	0.699	0.599	0.47	0.501	0.845		
SOP	0.503	0.786	2.94	0.101	0.558		

+ p < .10

이상과 같이 사회지출에 대한 필요조건의 Y-일관성 및 N-일관성 검증 결과, '높은 사회지출' 집합에 소속되기 위해서는 경제발전, 민주주의, 노인인구가 대개 필요조건으로 나타난다. 따라서 높은 사회지출을 보이기 위해서는 대개 높은 경제발전을 달성해야 하며, 높은 민주주의 수준을 구축해야 하며, 인구고령화 수준이 높아야 함이 전제된다.

5. 충분조건 검증

원인집합 X가 결과집합 Y의 하위집합(subset)이 되는 경우를 결과 집합에 대해 원인집합을 충분조건이라 부른다. 따라서 충분조건은 결과집합의 소속점수가 원인조건의 소속점수보다 높은 것이다. 연구방법에서 밝힌 바와 같이 Y-일관성 검증과 N-일관성 검증을 모두 실시하며, 일관성 검증 기준(benchmark)은 .80으로 하고 유의수준을 .05로 설정하여 검증한다.

먼저 Y-일관성 검증을 실시한 결과 총 49개의 배열이 주어진 검증기준(benchmark \geq .80, p < .05)을 통과하는 것으로 나타난다. 다음

으로 Y-일관성 검증을 통과한 배열 중 N-일관성 검증을 실시한 결과 총 49개의 배열 중 2개의 배열만이 주어진 검증기준(Y-con \geq N-con, p < .05)을 통과하는 것으로 나타난다.

<표 11> 사회지출에 대한 충분조건의 Y-일관성 및 N-일관성 검증 결과

Set	검증기준							최적 사례수 (Num BestFit)
	benchmark \geq .80, p < .05				Y-con \geq N-con, p < .05			
	Y-일관성 (Y-Con)	검증기준값 (benchmark)	F	p	N-일관성 (N-Con)	F	p	
GDATys	0.992	0.800	669.56	0.000	0.873	5.69	0.027	0
GDATYs	0.960	0.800	23.81	0.000	0.671	5.76	0.026	4

위와 같이 사회지출에 대한 충분조건의 Y-일관성 검증과 N-일관성 검증을 모두 통과한 원인조합의 충분조건 배열은 모두 2개임을 알 수 있다. 그런데 2개의 배열 중 GDATys 배열은 최적사례(bestfit)가 없다. 그러나 본 연구에서는 최적사례가 없는 원인조합의 배열도 포함하여 충분조건을 해석하고자 한다. 그 이유는 다음의 3가지이다. 첫째, Ragin(2000: 199)은 .50의 이상의 소속점수가 아닌 이유로 경험적 사례를 발견하지 못한 경우라 하더라도 모든 사례는 속성공간에 부분소속을 갖기 때문에 '이론적 설명이 가능'하거나 '원점수의 분석을 통해 정당'하다면 포함할 수 있다고 주장한다. 둘째, GDATys 배열은 최적 사례가 4개인 GDATYs의 사례와 이론적으로 함께 설명하여도 축약의 과정을 통해 보다 풍부한 해석을 가능하게 하므로 포함할 수 있다. 셋째, 원점수의 분석을 통해 퍼지점수의 배열인 GDATys 원인 조합을 찾아본 결과, '라트비아'는 해당 배열의 소속점수가 0.4856으로 분기점에 근접한 것으로 경험적으로 확인되기 때문에 포함할 수 있다.

이상의 논리에 따라 2개의 참인 원인조합의 배열은 GDATys와

GDATYs이며, 아래와 같은 축약의 과정을 통해 최소배열 축소집합 (Minimum Configuration Reduction Set)으로서 하나의 원인조합인 GDATs로 축약할 수 있다. 즉 GDATys * GDATYs = GDATs로 축약된다. 퍼지집합의 AND와 OR의 법칙을 통해 다음과 같이 부호화 및 해석할 수 있다.

$$\begin{aligned} \text{EXPF} &= \text{GDATys} + \text{GDATYs} \\ &= (\text{GDPF} * \text{DEMF} * \text{AGEF} * \text{TRAF} * \text{yisf} * \text{sopf}) + \\ &\quad (\text{GDPF} * \text{DEMF} * \text{AGEF} * \text{TRAF} * \text{YISf} * \text{sopf}) \\ &= \text{GDPF} * \text{DEMF} * \text{AGEF} * \text{TRAF} * \text{sopf} \end{aligned}$$

높은 사회지출 = 높은 경제발전 + 높은 민주주의
 + 높은 노인인구 + 높은 무역
 + 짧은 사회주의 기간

짧은 사회주의 기간의 역사를 가지면서(and) 높은 경제발전을 이루고(and), 높은 민주주의를 달성하며(and), 노인인구 비율이 높으며 (and), 무역비율이 높은 체제전환국은 높은 사회지출을 보인다고 할 수 있다. 충분조건 원인조합의 일관성은 .955이며, 설명력은 .565이다. 원인조합의 소속점수보다 결과집합의 소속점수가 높은 사례가 많으므로 일관성은 높지만 설명력은 상대적으로 낮게 나오게 된 것이다.

이러한 결과를 토대로 사회지출의 충분조건 원인조합이 설명하는 바는 다음과 같이 해석될 수 있다. 탈사회주의 체제전환을 시도한 지 25년을 목전에 둔 시점에서 서구 유럽의 발전된 복지국가의 수준까지 상승하지 못한 국내외의 여러 조건들이 있을 것이라 추정되어 왔다. 그러나 본 연구의 결과에 따르면, 경제발전, 민주주의 발전, 세계무역 활성화를 동시에 달성해야 한다는 것이다. 또한 동시에 인구고령화에 능동적으로 대처하며 사회주의의 부정적 유산을 청산하는 것이다. 결국 체제전환국의 사회지출은 단일한 요인(factor)에 설명되는 것이 아

닌 복합적 원인(causation)의 반영임을 이해할 수 있으므로, 퍼지셋 다중결합원인 분석의 적용이 적절하다고 할 수 있다.

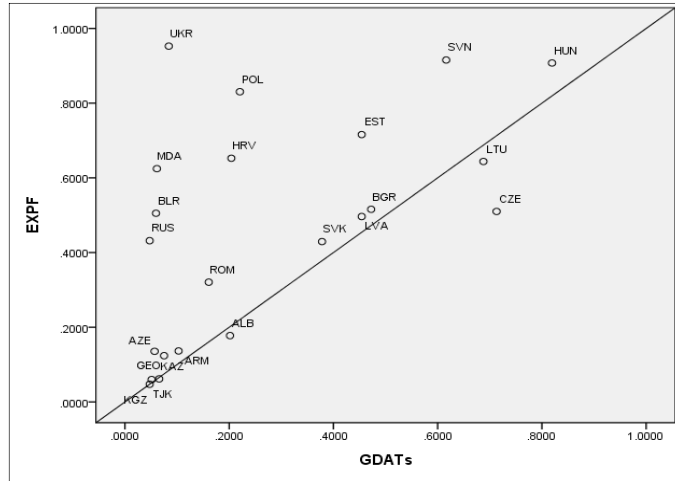
<표 12> 사회지출의 충분조건 원인조합

원인조합	개별설명력 (Raw Coverage)	구분설명력 (Unique Coverage)	일관성 (Consistence)
GDPF*DEMF*AGEF*TRAF*sopf	0.565	0.565	0.955

주: 원인조합이 1개이므로 총설명력(Total Coverage)과 결합식의 총일관성 값은 위의 표에 제시된 각각 개별설명력 및 일관성 값과 동일함

각각의 사례들에 대한 해석을 보다 구체적으로 설명하기 위해 결과집합인 사회지출 퍼지점수를 Y축에, 충분조건 원인조합의 퍼지점수를 X축에 놓은 산점도를 제시한다. 소수의 사례를 제외하고 거의 대부분의 사례들은 대각선의 위쪽, 즉 좌상단에 위치하여 높은 일관성을 보이고 있다. 퍼지셋 질적비교분석에서 주장하는 충분조건 산점도의 전형을 보여준다고 할 수 있다. 충분조건 원인조합을 가장 잘 설명하는 국가는 대각선의 오른쪽 모서리에 위치하고 있는 헝가리, 슬로베니아, 불가리아, 라트비아, 에스토니아이다. 모두 EU 국가들이며, 비셰그라드 그룹, 구유고연방, 남동부유럽, 발틱그룹이 고르게 분포되어 있다. CIS 국가들은 전반적으로 대각선의 위쪽에 위치하고 있지만, 대체로 대각선 위의 좌측에만 형성되어 있다. 즉 CIS 국가들이 충분조건 원인조합으로 설명되기는 하나 EU 국가들에 비해 상대적으로 낮은 설명수준을 보이는 것이다.

<그림 1> 사회지출과 충분조건 원인조합 간 산점도



대각선의 아래쪽, 즉 우하단에 위치한 소수의 사례들은 체코, 리투아니아, 알바니아, 카자흐스탄 이상 네 개의 나라이다. 그러나 네 개의 국가가 본 연구에서 주장하는 GDATs라는 충분조건 원인조합과 완전히 배치되는 증거라고 할 수 없다. 그 이유는 대각선으로부터 상당히 멀리 떨어져 있지 않기 때문이다. 즉 사회지출 퍼지점수에서 GDATs 조합의 퍼지점수를 빼주면 네 개의 국가만 음의 값을 갖는다. 그 크기는 체코가 -0.2029로 가장 크며, 리투아니아(-0.0438), 알바니아(-0.0240), 카자흐스탄(-0.0039)의 순이다. 이 크기는 그림 상에서도 보이는 것처럼, 본 연구의 이론적 주장을 뒤엎을 만큼 크지 않다. ‘사회지출 수준은 경제발전, 민주주의, 노인인구, 무역, 사회주의 기간의 다섯 개의 원인조합의 수준을 반영한다’는 이론적 진술보다 사회지출 수준이 낮게 나타나는 국가들일 뿐이며, 22개국 중 4개의 사례일 뿐이다.

대각선으로부터 어느 정도 떨어져 있는가를 구체적으로 파악하기 위하여 충분조건 원인조합 퍼지점수 및 예외사례에 대한 점수들을 제

시한다. GDATs는 충분조건 원인조합의 최소배열 축소집합의 소속점수, EXP-GDATs는 결과집합의 퍼지점수에서 GDATs의 퍼지점수를 뺀 값이다. 음의 값을 보이는 네 개의 국가들은 대각선의 우하단에 위치한 사례들로 알바니아, 체코, 카자흐스탄, 리투아니아이다.

<표 13> 충분조건 원인조합 퍼지점수 및 예외사례

국가	EXP	GDPF	DEMF	AGEF	TRAF	sopf	GDATs	EXP-GDATs
ALB	0.1776	0.2016	0.4063	0.2469	0.2576	0.7512	0.2016	-0.0240
ARM	0.1364	0.1084	0.1455	0.2727	0.1029	0.3377	0.1029	0.0336
AZE	0.1356	0.2402	0.0569	0.0934	0.1853	0.0841	0.0569	0.0787
BLR	0.5052	0.4520	0.0596	0.4901	0.7984	0.0841	0.0596	0.4457
BGR	0.5157	0.4724	0.7777	0.9524	0.6479	0.9144	0.4724	0.0433
HRV	0.6525	0.7520	0.7405	0.9197	0.2041	0.6163	0.2041	0.4484
CZE	0.5101	0.9382	0.9499	0.7130	0.7600	0.8862	0.7130	-0.2029
EST	0.7159	0.8085	0.9409	0.9161	0.9111	0.4542	0.4542	0.2617
GEO	0.1239	0.0943	0.2448	0.5251	0.2559	0.0751	0.0751	0.0487
HUN	0.9075	0.8194	0.8516	0.8629	0.9526	0.9526	0.8194	0.0881
KAZ	0.0618	0.3571	0.0657	0.1127	0.1632	0.1171	0.0657	-0.0039
KGZ	0.0598	0.0512	0.1326	0.0640	0.8092	0.0841	0.0512	0.0085
LVA	0.4964	0.6392	0.8019	0.9526	0.5144	0.4542	0.4542	0.0422
LTU	0.6439	0.7379	0.9137	0.7286	0.8158	0.6877	0.6877	-0.0438
MDA	0.6245	0.0613	0.3140	0.3068	0.6714	0.6877	0.0613	0.5632
POL	0.8306	0.7874	0.8553	0.4621	0.2205	0.8503	0.2205	0.6102
ROM	0.3207	0.5287	0.7234	0.6286	0.1608	0.8503	0.1608	0.1599
RUS	0.4317	0.8190	0.1372	0.4321	0.0474	0.0474	0.0474	0.3843
SVK	0.4293	0.8926	0.8906	0.3781	0.9408	0.8862	0.3781	0.0512
SVN	0.9155	0.9526	0.9526	0.8582	0.7653	0.6163	0.6163	0.2992
TJK	0.0474	0.0474	0.0474	0.0474	0.1341	0.1779	0.0474	0.0000
UKR	0.9526	0.1334	0.3090	0.7634	0.4844	0.0841	0.0841	0.8685

IV. 결론

본 연구는 사회과학에서 활용가치가 높은 퍼지셋 질적비교분석에 대한 이론, 방법론적 절차, 실증 분석결과를 제시함으로써 퍼지셋 방법론을 활용하고자 하는 연구자들에게 도움을 주고자 하였다. 특히 퍼지셋 방법론을 활용하는데 나서는 문제를 다음과 같이 3가지로 파

악하였다. 첫째, 어떠한 경우에 퍼지셋을 사용할 수 있는지 이해하기 어려움, 둘째, fsQCA 2.0 소프트웨어가 제대로 작동되지 않는다는 문제, 셋째, 퍼지셋 다중 결합원인 분석에서 모델검증의 어려움 이상 3가지를 지적하였다.

이러한 퍼지셋 방법론 활용의 어려움을 해결하고자 본 연구는 퍼지셋 방법론을 어떠한 경우에 활용가능한가라는 ‘사용이유’를 다음의 3가지로 요약하였다. 첫째, ‘질적인, 사례중심의, small N의, 집중적인(intensive)’ 복잡성(complexity)을 추구하는 사례중심적 연구와 ‘양적인, 변수중심적, large N의, 광범위한(extensive)’ 일반화(generality)에 주목하는 변수중심적 연구 사이에서 다양성을 추구(Ragin 2000, 21)하고자 할 때 사용할 수 있다는 점이다. 특히 소수 사례와 30개 이상의 사례 사이의 중범위 사례를 다룰 때 의미있다고 할 수 있다. 둘째, 사례를 구성하는 주요한 속성들의 혼합으로 이해할 필요가 있을 때, 본 방법론을 활용할 수 있다는 점이다. 특히 퍼지셋 이상형 분석에서 제기하는 특성공간을 통해 각 속성들의 혼합을 이해할 수 있으며, 이러한 이해는 한 시점이 아니라 시계열적 변화로 나타낼 수 있음을 강조하였다. 셋째, 사회현상을 단일 변수가 아닌 인과적 복잡성으로 이해하고자 할 때 퍼지셋이 의미있다고 제시하였다. ‘퍼지셋 다중결합적 인과관계 분석’은 독립변수의 상대적 영향력에 주목하기보다 결합관계를 용인하므로 다중공선성의 문제나 독립적 영향력의 문제를 해결할 수 있다는 점에서 의의가 있다. 예컨대 본 연구의 실증분석 중 충분조건 검증에서 나타난 <높은 사회지출 = 높은 경제발전 + 높은 민주주의 + 높은 노인인구 + 높은 무역 + 짧은 사회주의 기간>이라는 결과는 단일요인이 아닌 복합원인의 반영으로 해석할 수 있으므로, 다중회귀분석의 단점인 다중공선성과 독립적 영향력의 한계를 해결한다는 점에서 의의가 있다고 할 수 있다.

둘째, 더 이상 2009년부터 제대로 작동되고 있지 않는 fsQCA 2.0 소프트웨어에 의존할 필요가 없다는 점이다. 룽이스트와 바이세이(Longest and Vaisey, 2008)는 Stata Fuzzy 모듈을 소개한 바 있고

그의 주장대로 Stata를 통해 퍼지점수로 변환하는 것뿐만 아니라 진리표, Y-일관성 및 N-일관성 검증, ‘최소배열 축소집합(Minimum Configuration Reduction Set)’ 제시 등을 할 수 있다. 본 연구에서는 Stata를 활용한 실증분석 제시함으로써 퍼지셋 분석 도구(tool)로 인한 퍼지셋 방법론의 미활용에 대해 하나의 대안을 제시하였다.

셋째, 퍼지셋 다중 결합원인 분석시 활용될 수 있는 Y-검증과 N-검증의 실증분석 결과를 제시하였다. 그간 fsQCA 2.0 소프트웨어는 Y-검증만을 제공하여, 관행처럼 Y-검증만 해왔지만 논리적으로 N-검증도 동시에 시행해야 한다고 본 연구는 주장하였다. 그리고 실제 Y-검증과 N-검증을 동시에 통과한 충분조건 검증을 시행함으로써 방법론적 시현을 하였다.

그 동안 사회과학에서 양적 연구방법과 질적 연구방법은 서로의 단점을 보완하고 장점을 취하고자 노력하여 왔다. 그 중 라긴(Ragin)의 퍼지셋 질적비교분석 방법은 양적 및 질적 방법 사이에서 연구방법론상의 획기적인 기여를 하였다. 퍼지셋 이상형 분석과 퍼지셋 다중 결합원인 분석은 사례의 수와 다중공선성 등의 문제로 인해 해결되지 못한 방법상의 문제해결에 하나의 실마리를 제공해 주었다고 평가할 수 있다. 여전히 퍼지셋 이론이 발전하고 있으며 방법론적으로 진보하는 과정 중에 있으나 방법론적 어려움을 겪는 연구주체에 있어 하나의 대안적 연구방법론이라 할 수 있다. 본 연구를 통해 퍼지셋 방법론이 추구하는 양적 연구방법과 질적 연구방법의 대화에 주목하는 연구자들에게 방법론적 기여가 되기를 기대해 본다.

< 참고문헌 >

- 안상훈, 2002, “비교사회정책 연구방법론의 서설적 이해”, 『상황과 복지』, 13 pp. 47-71.
- 정해식, 2012, “사회의 질(SQ), 그 측정과 적용에 관한 비교 사회정책 연구: 복지국가와 국민행복을 중심으로”, 서울대학교 대학원 박사학위논문.
- 최영준, 2009, “사회과학에서 퍼지셋 활용의 모색: 퍼지 이상형 분석과 결합 요인 분석을 중심으로”, 『정부학연구』, 15(3) pp. 307-337.
- Choi, Y. J., 2006, “Pension reforms in East Asia: a comparative study”. PhD thesis, University of Bath.
- Cliff, T., 1974, *State capitalism in Russia*, London: Pluto.
- FSQCA, 2014, <http://www.u.arizona.edu/~cragin/fsQCA/>, 2014/07/10.
- Kent, R., 2008, “Using fsQCA: a brief guide and workshop for fuzzy-set qualitative comparative analysis”. Department of Marketing University of Stirling, in. <http://www.ccsr.ac.uk/publications/teaching/2008-10.pdf>, 2014/07/10.
- Kvist, J., 1999, “Welfare reform in the Nordic countries in the 1990s: using fuzzy set theory assess conformity to ideal types”. *Journal of European Social Policy*, 9(3) pp. 231-252.
- _____, 2006, “Diversity, ideal types and fuzzy sets in comparative welfare state research”, in. B. Rihoux & H. Grimm(eds.). *Innovative comparative methods for policy analysis: beyond the quantitative-qualitative divide*, New York: Springer-Verlag New York Inc.
- Longest, K. G. & Stephen Vaisey, 2008, “Fuzzy: a program for performing qualitative comparative analysis (QCA) in Stata”, *The Stata Journal*, 81(1) pp. 79-104.
- Merritt, R. L. & Stein Rokkan, 1966, *Comparing nations: the use of quantitative data in cross-national research*, New Haven, London: Yale University Press
- Ragin, C. C., 1987, *The comparative method: moving beyond qualitative and quantitative strategies*, Berkeley, Los Angeles, London: University of California Press.
- _____, 1995, “Using qualitative comparative analysis to study configurations”, in. U. Kelle(eds.), *In computer-aided qualitative data analysis*, Newbury Park: Sage Publications.
- _____, 2000, *Fuzzy-set social science*, Chicago, London: University of Chicago Press.
- _____, 2006, “Set relations in social research: evaluating their consistency and coverage”. *Political Analysis*, 14(3) pp. 291-310.
- _____, 2008a, *Redesigning social inquiry: fuzzy sets and beyond*, Chicago: University of Chicago Press.
- _____, 2008b, *User's guide to fuzzy-set / qualitative comparative analysis*, Tucson, Arizona: Department of Sociology, University of Arizona.
- _____, 2009, “Qualitative comparative analysis using fuzzy sets(fsQCA)”, in. B. Rihoux & C. Ragin(eds.), *Configurational comparative methods: qualitative*

- comparative analysis(QCA) and related techniques*, Thousand Oaks, CA and London: Sage Publications.
- Ragin, C. C. & John Sonnett, 2004, "Between complexity and parsimony: limited diversity, counterfactual cases, and comparative analysis. in. <http://www.sscnet.uci.edu/soc/soc237/papers/ragin.pdf>, 2014/07/10.
- Rihoux, B., 2003, "Bridging the gap between the qualitative and quantitative worlds? A retrospective and prospective view on qualitative comparative analysis", *Field Methods*, 15(4) pp. 351-365.
- _____, 2006, "Qualitative comparative analysis(QCA) and related systematic comparative methods", *International Sociology*, 21(5) pp. 679-706.
- Schneider, C. Q. & Claudius Wagemann, 2010, "Standards of good practice in qualitative comparative analysis (QCA) and fuzzy-sets", *Comparative Sociology*, 9(3) pp. 397-418.
- Tilly, C., 1984, *Big structures, large processes, huge comparisons*, New York: Russell Sage Foundation.
- Wagemann, C. & Carsten Q. Schneider, 2010, "Qualitative comparative analysis (QCA) and fuzzy-sets: agenda for a research approach and a data analysis technique", *Comparative Sociology*, 9 pp. 376-396.

Application of Fuzzy-Set Qualitative Comparative Analysis: : focusing on Y-Test and N-Test Using Stata

Min-Ki,chaе
(Institute of Social Welfare Seoul National University)

Abstract

Quantitative research and qualitative research methods in social science have been trying to complement the disadvantages of the other approach. The fsQCA (Fuzzy-Set Qualitative Comparison Analysis), a fuzzy-set ideal type analysis and fuzzy-set multiple conjunctural causations analysis developed by Charles Ragin as a convergence between quantitative and qualitative methods, has been represented as a methodological breakthrough. This study focused on the utilization of Fuzzy-Set methodology in any case (that is the reason for its use), the software tool's extension from fsQCA 2.0 to the STATA fuzzy module, and applying Y-tests and N-tests to overcome the problems of verification. In addition, specific procedures of fuzzy-set multiple conjunctural causations analysis are presented as demonstration studies. Through theoretical and methodological introduction to fsQCA, despite its limitations, Fuzzy-Set Qualitative Comparison Analysis has become highly utilized in the social sciences.

Keywords : fuzzy-set qualitative comparative analysis,
fuzzy-set multiple conjunctural causations analysis,
Y-test and N-test