

사회보장 급여의 시뮬레이션을 위한 시론적 연구

A Preliminary Study for Simulation of Social Security Benefit

추병주*, 김성훈**, 정상기***, 김경준****, 이인수*****, 김현정*****

명지대학교 행정학과*, (주)케이엘넷 빅데이터사업팀**, 한국조세재정연구원 장기재정전망센터***, 제주 인적자원개발위원회****, 사회보장정보원 사회보장정보연구소*****, 중앙대학교 사회복지학과*****

Byung-Joo Chu(bjchu@mju.ac.kr)*, Sung-Hoon Kim(ksh98@klnet.co.kr)**,
Sang-Gee Jung(skjung@kipf.re.kr)***, Kyung-Joon Kim(dr_june@nate.com)****,
In-Soo Lee(lis1018@ssis.or.kr)*****, Hyung-Jung Kim(guswjd0317@hanmail.net)*****

요약

국민의 복잡·다양한 복지수요에 대응하기 위하여 각종 사회보장 급여의 신설·변경(수정·보완)이 지속적 으로 이루어지고 있다. 이처럼 사회보장 급여를 신규로 만들고 수정하는데 앞서 행·재정적으로 발생가능한 문제점에 대비하기 위하여 수급변동을 정확하게 예측할 수 있는 시뮬레이션이 요구된다. 하지만 관련 선행 연구가 미흡하고 신규 대상자의 데이터 수집이 어렵기 때문에 수급 시뮬레이션을 수행하는 것은 불가능한 실정이다. 본 연구는 사회보장 급여의 수급 시뮬레이션을 수행하기 위한 시론적인 차원에서 '10년 인구 총 조사 마이크로 데이터를 기반으로 재정패널 데이터와 통계적 매칭을 통해 소득재산이 포함된 전국민 기초 데이터를 생성하고, '10년 기준으로 대상자 규모가 가장 큰 급여 가운데 하나인 기초노령연금에 대해 마이크로 시뮬레이션 방법론을 적용하여 수급 시뮬레이션을 수행하였다. 아울러, 시뮬레이션 기초데이터와 재정 패널 데이터, 기초노령연금 시범분석 결과와 기초노령연금 실제 수급자 데이터에 대한 비교를 통해 분석결 과의 정합성을 검증하였다.

■ 중심어 : | 사회보장 | 통계적 매칭 | 마이크로데이터 | 시뮬레이션 |

Abstract

This study performed statistical matching using population census microdata and financial panel data. It generated the national basic data for simulation including income and property. Using this data the basic old-age pension, which is one of the biggest benefits, simulation was performed by applying the micro simulation methodology. In addition, we verified the coherence of the analysis results by comparing simulation basic data and financial panel data, basic old-age pension pilot simulation analysis results and basic old&-age pension actual beneficiary data.

■ keyword : | Social Security | Statistics Matching | Microdata | Simulation |

I. 서론

각종 사회적 위협으로부터 국민을 보호하고 나날이 증가하는 국민의 복잡한 복지수요에 대응하기 위하여

정부는 복지예산 확충, 체계적인 계획수립 및 추진, 관 련 법률의 제·개정을 추진하는 등의 노력을 지속적으로 기울이고 있다. 이러한 노력은 실제로 기초노령연금 의 기초연금 확대 개편('14.7월)과 국민기초생활보장의

접수일자 : 2018년 03월 23일
수정일자 : 2018년 04월 11일

심사완료일 : 2018년 04월 11일
교신저자 : 추병주, e-mail : bjchu@mju.ac.kr

맞춤형 개별급여로의 전면개편('15.7월) 등과 같이 사회보장 급여의 확대·개편으로 이어져 국민의 복지수준을 제고하는 가시적인 성과로 나타나고 있기도 하다.

이처럼 국민의 복잡·다양한 복지수요에 대응하기 위하여 각종 사회보장 급여를 새롭게 만들거나 기존의 급여를 수정·보완을 하고자 한다면 이에 앞서 행·재정적으로 발생가능한 문제점 및 파급효과를 예상하고 대비책을 강구해야 한다. 이를 위해서는 급여의 신설·변경에 따라 수반되는 수급자 규모(소요재원 포함)의 시뮬레이션이 요구된다. 하지만 관련 선행연구가 미흡하고, 각종 개인정보 규제로 인해 데이터 수집이 불가능한 상황에서 각종 사회보장 급여의 수급 변동을 예측하는 시뮬레이션 모델을 구성하고 수행방안을 도출하는 것은 현실적으로 불가능한 실정이다.

이러한 시점에서 본 연구는 사회보장 급여의 시뮬레이션의 본격적인 수행에 앞서 시론적인 연구로써 시뮬레이션의 적용가능성을 검증하고자 하였다. 이를 위해 인구총조사와 재정패널 데이터의 통계적 매칭을 통해 시뮬레이션에 요구되는 데이터를 구성하고, 이를 활용하여 현행 사회보장 급여 가운데 대상자 규모가 가장 큰 급여에 속하는 기초노령연금의 대상자 선정기준을 적용하는 모델을 수립하여 마이크로 시뮬레이션을 수행하였다. 아울러 데이터 구성 및 시뮬레이션 결과를 검증하고, 이러한 시뮬레이션 모델의 활용방안을 제시하였다.

II. 선행연구

사회보장 급여의 시뮬레이션을 대상으로 하는 연구가 매우 미진한 상황으로 인해 여기서는 복지정책의 대상집단의 규모를 추정한 각종 연구를 중심으로 선행연구를 살펴보고 그 함의를 제시하였다.

먼저 차상위 계층 규모를 추정한 연구에 대해 살펴보면 이현주(2000)는 생활보호자를 제외한 최저생계비 이상 평균소득의 50% 이하 가구를 차상위 계층으로 정의하고, 도시가계연보의 원자료를 이용하여 그 규모를 14.2%로 추정하였다[1].¹ 이와 달리 박승후 외(2003)는

가구소득 최저생계비 120% 이하의 가구를 차상위 계층으로 정의하여 그 규모를 추정하였다. 가구소비실태조사의 원자료를 이용하여 차상위 계층에 대한 계층 연구를 수행하였는데, 분석 결과 2000년 차상위 계층 규모는 4.0% 수준이며, 가구소득에 재산기준까지 포함할 경우 그 규모는 3.1%로 축소될 것으로 제시하였다[2]. 또한 이현주 외(2008)는 차상위 계층에 대한 개념과 기준 설정 방식에 대한 종합적인 분석을 토대로 2006년 경상소득 기준 최저생계비 120% 이하인 차상위 계층의 규모는 우리나라 전체 가구의 5.04%인 82만 가구로 추정하였다[3].

한편 김교성(2009)은 한국복지패널조사의 1차년도와 2차년도 자료를 활용하여 우리나라 차상위 계층의 규모와 특성을 추정하였는데, 우리나라 차상위 계층 규모는 전체 가구의 8.9% (2005년)와 7.3%(2006년)에 이르는 것으로 나타나 앞선 연구들에 비해 그 규모가 크게 추정되었다. 더하여 이 연구는 여성, 노인, 한부모, 저학력, 실업이나 비정규직과 같은 빈곤계층의 특성이 차상위 계층에서도 동일하게 관찰되고, 차상위 계층의 지위 변화가 활발할 뿐 아니라 신규 진입하는 차상위 계층의 규모도 상당한 것으로 분석하였다[4].

그런가하면 복지대상자의 단편적인 규모가 아닌 우리나라 빈곤의 규모를 다각도에서 조망해 본 연구도 존재한다. 김교성 외(2008)는 복지패널의 2005년도 자료를 활용하여, 절대적 빈곤, 상대적 빈곤, 소비 빈곤, 주관적 빈곤, 상대적 발달 수준을 계측하고, 빈곤의 심도를 파악하기 위해 빈곤갭 비율을 측정하였는데, 분석 결과 절대적 빈곤율과 중위소득 50% 미만의 상대적 빈곤율은 14.2%와 20.3%로, 빈곤갭은 1.8%로 제시하였다[5].

끝으로 빈곤층이 아닌 대상자에 대한 연구 관심으로써 학교 밖 청소년의 대상자 규모를 추정한 연구가 있다. 오병돈·김기현(2013)은 청소년복지정책의 중요한 대상인 학교 밖 청소년을 「청소년 기본법」 으로부터 만 9~24세로 정의하고 여기에 학업중단과 미진학 대상자를 포함하여 그 규모를 추정하였다. 분석은 2011년 청소년 종합실태조사 자료를 이용하여 군집분석을 통

1 다만 이 연구는 국민기초생활보장제도가 도입되기 이전 상황의 분석으로 상대적 빈곤선을 사용하고 있어 전국을 대표할 수 없는 자료에 근거한 추정이라는 한계가 있다[4].

해 이루어졌는데, 분석 결과 전체 만9-24세 청소년 인구 중 학교 밖 청소년 규모는 19.73%에서 20.23% 사이인 약 200백만 명 내외로 나타났다[6].

이상의 연구를 종합해볼 때 그 간 연구들은 차상위 계층을 위시한 빈곤층 등 특정 정책집단의 규모를 추정하는 것을 중심으로 이루어졌으며, 추정방법 역시 소규모 조사자료를 활용하여 특정 기준선 아래 대상집단을 매크로한 접근을 통해 도출하였다. 이러한 모습을 볼 때 우리나라 사회보장 환경 및 제도의 변경에 따른 변화를 보다 정확하게 가늠하기 위해서는 국민 개개인의 속성을 반영하는 대규모의 데이터 구성과 시물레이션 분석이 요구된다.

III. 사회보장 급여 시물레이션

본 연구에서는 접근 가능한 통계를 이용하여 전국민을 모집단으로 하는 대표성 있는 인구집단을 설정하고, 이들의 인적·사회적 특성을 토대로 소득재산 자료를 생성하여 사회보장 급여의 수급 가능여부를 마이크로단위로 분석할 수 있는 시물레이션 모델을 개발하였다.

인구집단의 소득재산 데이터 생성과 기초노령연금을 대상으로 사회보장 급여 수급 시물레이션에 대한 내용을 세부적으로 살펴보면 다음과 같다.

3.1 소득재산 데이터 생성

본 연구의 인구집단은 통계청에서 5년마다 수행하는 인구총조사 결과를 활용하여 구성하였다. 가장 최근 연도인 '10년 조사결과 가운데 공개된 총인구 대비 2%를 인구집단으로 설정하였다[7].

인구집단의 소득재산 데이터 생성은 '10년 인구총조사 A형 샘플(2%) 데이터를 기준 파일(Recipient File)로 하고, 일반 국민의 소득재산을 가장 자세하게 조사하고 있는 재정패널 데이터를 제공파일(Donor File)로 하는 통계적 매칭을 통해 생성하였다[8][9].²

2 재정패널 데이터 가운데 '10년 인구총조사결과와 동일한 시점에 이루어진 재정패널 4차 데이터를 활용하였으며, 기준파일의 n이 제공파일보다 크기 때문에 제공파일의 레코드를 중복사용하는 비제한적 매칭을 사용하였다.

통계적 매칭은 세부적인 방법론에 따른 검증을 위해 다양한 기법 및 거리함수를 적용하였다. 랜덤핫덱(Random Hotdeck)과 최근접 이웃 핫덱(Nearest Neighbor Distance Hotdeck) 2가지 기법을 적용하였으며, 이를 위한 거리함수는 맨하탄(Manhattan), 마할라노비스(Mahalanobis), 유클리디안(Euclidean) 3가지를 활용하였다. 이를 통해 인구집단의 소득재산 데이터셋을 총 6개 생성하였다.³

통계적 매칭을 위한 공여계층(Donor Class)은 '성별'과 '거주지역'으로 설정하였고, 세부 매칭변수는 기준파일·제공파일 데이터 상관관계에 따른 설명력을 기준으로 최종학력, 연령, 혼인상태 등으로 선정하여 소득재산 데이터를 생성하였다.⁴ 이러한 일련의 과정을 거쳐 생성된 소득재산 데이터의 세부적인 사항을 재정패널 데이터를 기준으로 살펴보면 다음과 같다.

3.1.1 소득

소득은 가구원 개인 단위이며 '근로소득', '재산소득', '사업소득', '공적이전소득'의 합계이며, 세부사항은 [표 1]과 같다.

표 1. 소득 세부사항

중분류	소분류	재정패널 변수
근로소득	임금소득	근로소득 금액
사업소득	사업소득	사업소득 (순)소득 금액
	임대소득	부동산 임대 소득 총합 금액
재산소득	이자소득	이자 및 배당소득 금액
	연금소득	퇴직연금 소득 총금액
		연금신탁/펀드/저축 소득 총금액
		연금보험 소득 총금액
		보장보험 소득 총금액
		저축보험 소득 총금액
		자동차보험 소득 총금액
		주택연금 소득 총금액
공적이전소득	사회보험보조금 - 총금액01	
	사회보험보조금 - 총금액02	
	정부지원연금 소득 - 총금액 01	
	정부지원연금 소득 - 총금액 02	
	정부지원연금 소득 - 총금액 03	

출처: 재정패널의 소득을 바탕으로 재구성[9]

3 기준파일과 제공파일의 통계적 매칭은 R 패키지인 D'Orazio(2011)의 StatMatch를 활용하였다[10].

4 인구집단에 속하는 개개인의 사회보장 급여 수급여부를 분석하기 위하여 소득재산 수준을 조사하는 사회보장 급여의 일반적인 선정단위인 개별 가구단위로 데이터를 결합하였다.

3.1.2 재산

재산은 가구 단위이며, ‘금융재산’, ‘일반재산’ ‘자동차’, ‘부채’로 구분된다. 세부적인 사항은 [표 2]와 같다.

표 2. 재산 세부사항

분류	재정패널 변수
금융 재산	총 금액 - 은행 등 금융기관 예적금
	총 금액 - 펀드 가입 금액
	총 금액 - 채권 보유 금액
	총 금액 - 주식 보유 금액
	총 금액 - 저축성 보험, 연금성 보험
	총 금액 - 빌려준 돈
	총 금액 - 기타 금융재산
일반 재산	작년 주택마련 저축 및 펀드 불입금
	보증금액 - 가구원 거주주택 전월세 보증금
	보증금액 - 가구원 비거주주택 전월세 보증금
	보증금액 - 주택 이외 건물 전세 및 월세 보증금
	시가총액 - 현재 거주주택 외 보유주택
	시가총액 - 주택 이외의 보유 부동산 총액(토지, 건물)
	시가총액 - 보유 회원권(골프, 콘도 등)
시가총액 - 선박, 건설중장비	
시가총액 - 현재 거주주택 시가 총액	
자동차	보유차량 연식
	보유차량 금액
부채	총 대출금액 - 정부지원 주택자금 대출
	총 대출금액 - 학자금 대출
	총 대출금액 - 금융기관 주택담보 대출
	총 대출금액 - 금융기관 대출
	총 대출금액 - 세입자에게 돌려줄 전세금, 임대보증금
	총 대출금액 - 기타

출처: 재정패널의 재산을 바탕으로 재구성[9]

3.2 기초노령연금 수급 시뮬레이션 모델링

여기서는 사회보장 급여 가운데 하나인 기초노령연금을 적용하여 수급 시뮬레이션을 수행하는 과정을 소개하고, 그 결과를 검증하였다.

기초노령연금을 시뮬레이션 대상으로 선정할 이유는 현행 대상자 규모 및 소요예산이 가장 큰 급여 가운데 하나라는 점과 고령화 현상의 심화로 인해 향후 대상자 선정 기준 등의 제도변화가 일어날 가능성이 매우 높다는 점이 반영된 것이다.

기초노령연금 수급 시뮬레이션 과정은 크게 가구구성, 소득인정액 산출, 선정기준액 적용의 3가지로 구분할 수 있으며 도식화하면 [그림 1]과 같다. 세부적인 내용을 살펴보면 다음과 같다.

3.2.1 가구구성

지원대상자를 선정하는데 있어 경제수준이 반영되는 사회보장 급여는 대개 지원대상자 개인에 대하여 해당 가구의 경제수준을 토대로 이루어지는 경우가 많다.⁶ 기초노령연금 역시 지원대상자 본인에 대하여 배우자의 경제수준까지 포함하여 대상자를 선정하고 있다.

기초노령연금의 지원 가구유형은 지원대상자의 배우자 여부 및 배우자의 연령기준 충족 여부에 따라 단독노인가구, 부부1인수급, 부부2인수급으로 구분된다. 이러한 지원대상자 기준을 적용하면 조건식#1(65세 이상), 조건식#2(배우자여부), 조건식#3(배우자 연령기준)과 같다. 또한 인구집단에서 1개 가구가 기초노령연금 선정단위 ‘가구’는 2개 이상이 되는 경우를 구분하기 위해 조건식#4(가구주 포함 여부)를 구성하였다.⁷

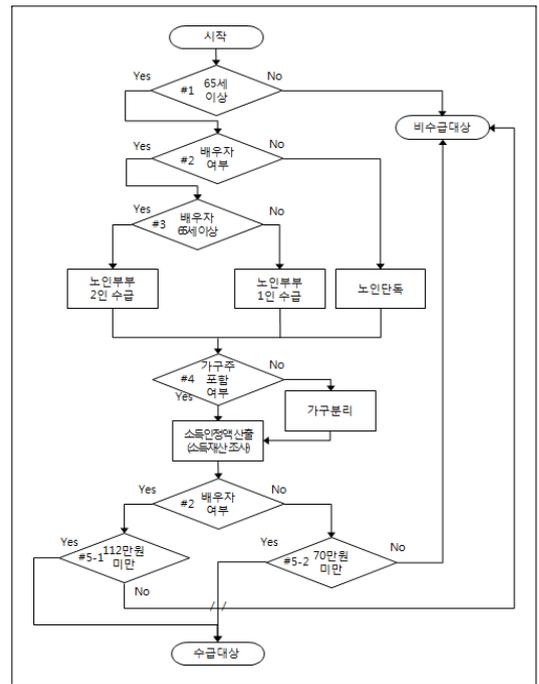


그림 1. 기초노령연금 수급 시뮬레이션 모델

6 사회보장 급여의 대상자 선정을 위해 구성되는 ‘가구’의 범위는 해당 사회보장 급여의 목적, 대상 등의 특성에 따라 다르게 구성된다.

7 예를 들어 90세 노모를 부양하고 사는 70세 아들과 그 배우자(68세)가 포함된 가구의 경우 일반적인 가구 기준으로는 90세 노모(혹은 70세 아들)가 가구주가 되는 1개의 가구이지만, 기초노령연금 제도의 지원대상 기준인 선정단위로 볼 때는 90세 노모 1가구, 70세 아들과 68세 배우자 1가부로 2가부로 구분해야 한다는 것을 들 수 있다.

5 자동차 가액은 재정패널의 자동차 구매가격, 구매년도를 토대로 과표액을 기준으로 재산출하였다.

3.2.2 소득인정액 산출

기초노령연금에서 대상자를 선정하기 위한 지원대상자 가구의 경제수준은 기초노령연금의 소득인정액 측정 및 산출 기준을 통해 반영된다. 소득인정액은 소득 평가액에 재산 소득환산액을 합한 값으로 세부적인 내용을 살펴보면 [그림 2]와 같다.

소득인정액 = 월 소득평가액 + 재산의 월 소득환산액* * 재산이 '음의 값' 인 경우 0으로 처리 월 소득 평가액 = 실제소득-가구특성별지출 비용-근로소득공제 금액 재산의 월 소득환산액 = {(일반재산-기본재산액)+(금융재산-2,000만원)-부채}×재산의 소득환산율÷12월
--

출처: 2010 기초노령연금 사업안내[11]

그림 2. 기초노령연금의 소득인정액[11]

기초노령연금의 재산의 월 소득환산액을 산출하는데 있어 반영되는 기본재산액은 대·중소도시, 농어촌 여부에 따라 10,800만원, 6,800만원, 5,800만원으로 구분되며 [표 3], 이는 조건식#5(기본재산액)로 구성하였다.

표 3. 기초노령연금의 기본재산액

구분	대도시	중소도시	농어촌
	특별시, 광역시의 '구'	도의 '시'	도의 '군'
기본 재산액	1억 800만원	6천 800만원	5천 800만원

출처: 2010 기초노령연금 사업안내[11]

3.2.3 선정기준액 적용

선정기준액은 기초노령연금 수급여부를 결정하는 기준금액이다. 이는 배우자 여부에 따라 2가지로 구분되며 [표 4],⁸ 조건식#5(#5-1, #5-2)(선정기준액)로 구성하였다.

표 4. '10년 기초노령연금의 선정기준액[11]

구분	노인단독	노인부부
기준금액	700,000	1,120,000

출처: 2010 기초노령연금 사업안내[11]

8 배우자여부는 앞서 조건식#2(배우자여부)와 동일하다.

3.3 데이터 생성 및 시뮬레이션 결과 검증

3.3.1 데이터 생성 결과의 검증

통계적 매칭을 통해 데이터 생성이 적절하게 이루어졌는지는 대표성의 문제로 귀결된다[12][13]. 여기서 대표성은 매칭결과가 원본파일의 성질을 그대로 유지하는가에 대한 문제이며, 이는 기준파일 변수의 평균, 표준편차, 분산과 얼마나 유사하게 나타났는가를 토대로 검증할 수 있다. 재정패널과 근사치로 나타난 생성 데이터를 기준으로 대표성을 검증한 결과를 제시하면 다음과 같다.

먼저 소득에 대해 가구를 기준으로 결합한 값을 기준으로 보면 재정패널의 평균이 2,520 (3,736)만원으로 최근접이웃하택이 각각의 거리함수 별로 맨하탄 2,479(3,876)만원, 마할라노비스 2,479(3,746)만원, 유클리디안은 2,484(3,869)만원으로 나타났다. 그 결과 개인 소득은 랜덤하택보다 유사하게 나타났으며, 가구소득은 랜덤하택이 보다 유사하게 나타났다. 로그를 적용하여 도출한 데이터별 분포는 [그림 3]과 같다.

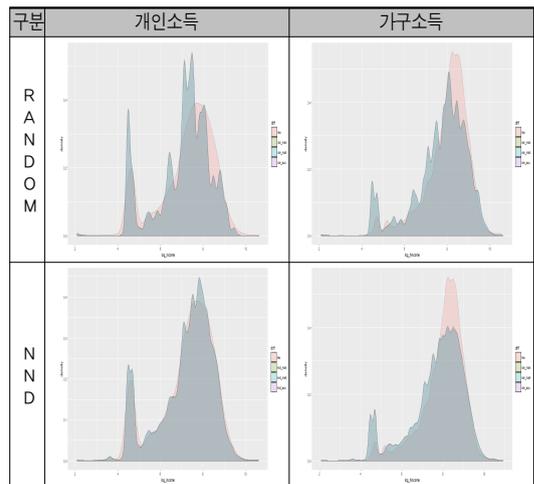


그림 3. 소득 생성 검증

일반재산과 자동차를 살펴보면, 재정패널의 일반재산 평균은 24,097만원이며, 최근접이웃하택이 각각 23,489만원, 23,581만원, 23,459만원으로 보다 유사하게 나타났다. 재정패널의 자동차 가액 평균은 850만원이

며, 랜덤햇택이 각각 856만원, 859만원, 855만원으로 최근접이웃햇택보다 유사하게 나타났다. 분포는 [그림 4]와 같다.

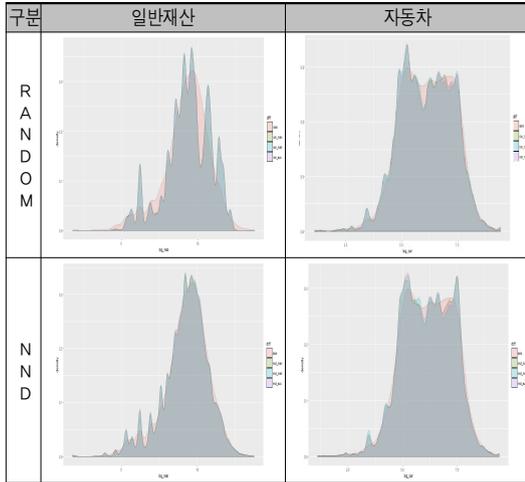


그림 4. 일반재산, 자동차 생성 검증

금융재산과 부채를 살펴보면 재정패널의 금융재산 평균은 4,070만원, 최근접이웃햇택이 각각 4,131만원, 4,157만원, 4,114만원이며, 재정패널 부채 평균은 8,632만원, 최근접이웃햇택이 각각 8,579만원, 8,641만원, 8,657만원으로 최근접이웃햇택이 보다 유사하게 나타났다.

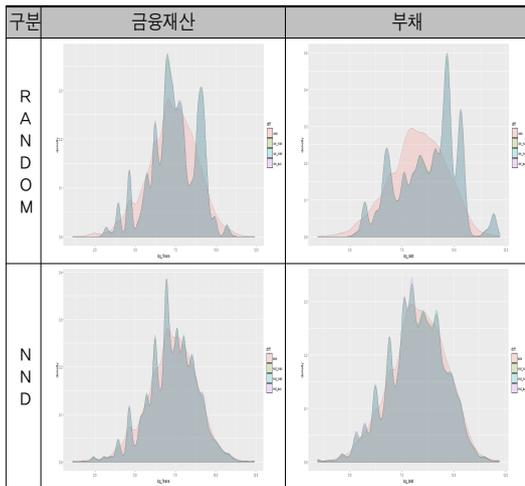


그림 5. 금융재산, 부채 생성 검증

3.3.2 시뮬레이션 결과의 검증

소득인정액이 선정기준액 이하인 가구에 수급판정을 내린 시뮬레이션 결과를 살펴보면 최근접이웃햇택의 수급율이 실제 수급율과 유사했으며, 마할라노비스 거리함수를 사용한 데이터가 67.78%로 가장 유사하게 나타났다[표 5].

표 5. 시뮬레이션 결과: 성별 대상자 규모 (단위: 명 / %)

구분	실제 수급	시뮬레이션 기초데이터					
		RAN man	RAN mah	RAN euc	NND man	NND mah	NND euc
전체 수급자	3,714,436 (67.46)	90,917 (69.50)	90,650 (69.30)	91,155 (69.68)	89,118 (68.12)	88,669 (67.78)	89,313 (68.27)
65세 이상 인구 수	5,506,352 (100)	130,817 (100)					

수급 시뮬레이션 결과를 좀더 검증하기 위하여 실제 수급 데이터와 성별, 거주지역별, 지원가구 유형별로 비교하여 살펴보면 다음과 같다[표 6].

먼저 성별의 경우 실제 수급대상자가 총 3,714,436명 중 남성이 1,304,632명(35.12%), 여성이 2,409,804명(64.88%)으로 나타나는 가운데 시뮬레이션의 경우 랜덤햇택 마할라노비스가 전체 90,650명 중 남성이 35,063명(38.86%), 여성이 55,587명(61.32%)으로 가장 유사하게 나타났다. 다음으로 지원가구 유형의 경우 노인단독 1,938,112명(52.18%), 노인부부 1인수급 351,246명(9.46%), 노인부부 2인수급 1,425,078명(38.37%)으로 나타나는 가운데 랜덤햇택 유클리디안이 단독가구 44,833명(49.18%), 노인부부 1인수급 10,114명(11.1%), 노인부부 2인수급 36,208명(39.72%)로 가장 유사하게 나타났다.

성별 및 지원가구 유형별 대상자 비율이 매칭기법 및 거리함수를 막론하고 전반적으로 유사하게 나타난 반면, 지역별 대상자 비율은 실제 수급 대상자 비율과 다소 차이가 나타나는 것으로 나타났다. 상대적으로 오차의 절대값이 가장 작게 나타난 랜덤햇택 유클리디안의 경우에도 대도시(5.59%), 중소도시(6.16%) 차이가 약 6% 나타나고, 농어촌(11.75%)은 약 12%의 수급비율 차이가 나타난 것이다.

표 6. 시뮬레이션 결과: 대상자 특성별 규모 (단위: 명 / %)

구분	실제 수급	시뮬레이션 기초데이터							
		RAN man	RAN mah	RAN euc	NND man	NND mah	NND euc		
성별	남성	1,304,632 (35.12)	35,288 (38.81)	35,063 (38.68)	35,282 (38.71)	34,557 (38.78)	34,395 (38.79)	34,681 (38.83)	
	여성	2,409,804 (64.88)	55,629 (61.19)	55,587 (61.32)	55,873 (61.29)	54,561 (61.22)	54,274 (61.21)	54,632 (61.17)	
지원가 구유형	노인단독	1,938,112 (52.18)	44,547 (49.00)	44,532 (49.13)	44,833 (49.18)	43,761 (49.1)	43,430 (48.98)	43,744 (48.98)	
	노인 가족 유형	1인	351,246 (9.46)	10,150 (11.16)	10,176 (11.23)	10,114 (11.1)	9,963 (11.18)	9,991 (11.27)	9,955 (11.15)
		2인	1,425,078 (38.37)	36,220 (39.84)	35,942 (39.65)	36,208 (39.72)	35,394 (39.72)	35,248 (39.75)	35,614 (39.88)
지역	대도시	1,378,100 (37.1)	28,557 (31.41)	28,397 (31.3)	28,720 (31.51)	28,425 (31.9)	28,353 (31.98)	28,500 (31.91)	
	중소도시	976,662 (26.29)	18,258 (20.08)	18,416 (20.3)	18,351 (20.13)	17,478 (19.61)	17,314 (19.53)	17,436 (19.52)	
	농어촌	1,359,674 (36.61)	44,102 (48.51)	43,837 (48.4)	44,084 (48.36)	43,215 (48.49)	43,002 (48.5)	43,377 (48.57)	
전체	3,714,436 (100)	90,917 (100)	90,650 (100)	91,155 (100)	89,118 (100)	88,669 (100)	89,313 (100)		

3.4 시뮬레이션 결과 요약

사회보장 급여의 수급 시뮬레이션 수행을 위해 인구 집단에 대해 통계적 매칭을 통해 생성한 기초데이터의 검증결과 핫택 및 거리함수에 상관없이 전반적으로 제공과일인 재정패널과 유사하게 나타났으며, 사회보장 급여 가운데 하나인 기초노령연금 수급 시뮬레이션을 구성하여 대상자 규모를 분석한 결과 실제 수급대상자 비율과 유사하게 나타났다.

본 연구에서 적용한 기초노령연금 수급 시뮬레이션 모델은 활용가능한 데이터를 고려하여 구성한 것으로 실제 기초노령연금 수급자를 판정하는데 적용되는 기준을 매우 단순화한 것이다. 그럼에도 불구하고 전체 대상자인 65세 이상 노인 가운데 수급자 비율에 있어서 시뮬레이션과 실체가 매우 유사하게 나타나는 것을 고려해 볼 때 시뮬레이션의 적용이 가능한 것으로 판단된다.

IV. 결론: 모델의 활용과 향후 연구방향

4.1 모델의 활용

사회보장 영역의 시뮬레이션은 대개 매크로한 수준

에서 장기적인 관점에서 이루어지고 있다. 대표적인 추계모형인 사회보장 중장기 재정추계의 경우 재정추계를 위한 사회보장지출 범주를 국제 비교가 가능한 OECD 사회복지지출(SOCX)을 토대로 9대 정책영역⁹로 구성하고 다양한 제도를 포함한다. 하지만 이러한 모형은 매크로한 수준의 변수만을 반영하기 때문에 사회변화에 따른 각종 사회보장 제도의 대상자 규모 및 소요재원을 추계할 수는 있지만, 세부 급여기준 변경에 의한 변화를 예측할 수는 없다.

하지만 본 연구의 사회보장 급여의 수급 시뮬레이션 모델은 사회보장 급여를 신설·변경하는 경우 지원대상을 선정하는 세세한 기준을 합리적으로 설정하기 위해 요구되는 정보를 제공할 수 있다. 현재도 이러한 기준 수급자 데이터를 활용하여 기준 변경에 따른 대상자 규모 변동을 수행하고 있지만 미신청자 데이터 미보유¹⁰ 등으로 인해 세부적인 변화를 살펴보는데 한계가 있다.

이러한 상황에서 전인구를 모집단으로 하는 대표성 있는 인구집단의 데이터를 토대로 각종 사회보장 급여의 대상자 선정기준을 적용하는 시뮬레이션을 수행한다면 불합리한 급여기준을 객관적 근거를 토대로 개선할 수 있을 것이며, 그 과정에서 다양한 조건식을 바탕으로 미세한 조정을 위한 심층분석을 수행하는 경우 기준변화에 따른 역효과를 줄이고, 긍정적인 효과를 제고할 수 있을 것이다. 아울러, 개별 사회보장 급여의 특성을 반영함으로써 표준화되지 못하고 복잡하게 구성된 각종 급여의 기준을 표준화하고 단순화하는데 있어서도 기여를 할 수 있을 것이다.

4.2 향후 연구방향

본 연구가 가지는 시론적 성격을 고려해볼 때 사회보장 영역에서 기존에 시도하지 않은 새로운 기법을 활용하여 분석데이터를 구성하고, 비록 기초노령연금이라

9 OECD SOCX 9대 정책영역: 노령, 유족, 근로무능력, 보건, 가족, 적극적 노동시장정책, 실업, 주거, 기타[14]

10 사회보장 급여의 업무는 대개 신청·조사·결정·지급·사후관리 단계로 이루어지며, 그 정보는 사회보장정보시스템에 축적된다. 따라서 급여를 신청한 대상자의 데이터만 보유하며 미신청자의 데이터는 보유하고 있지 않다.

는 단일 급여를 대상으로 하였지만 제도의 특성을 반영한 시뮬레이션 모델을 구성·적용하였다는 측면에서 의의를 찾아볼 수 있다.

하지만 다음과 같은 한계를 가지고 있다. 첫째, 인구 집단을 설정하는데 있어 현실에 존재하지 않지만 전국민에 대한 대표성을 가지는 인구집단을 새롭게 구성한 것이 아니라, 인구총조사라는 대규모 조사를 통해 구성되어진 집단을 사용하였다는 것이다. 둘째, 기초노령연금 수급 시뮬레이션 모델링 과정에서 제공파일 데이터 수집 및 활용의 한계로 인해 모델이 크게 단순화되었다. 마지막으로 과거 단일시점을 기준으로 수행한 정태적 시뮬레이션 연구의 특성으로 인해 시간의 흐름에 따른 변화를 추정하기 어렵다.

이러한 한계를 고려해볼 때 후속 연구는 현실에 존재하지 않지만 대표성을 가지는 대규모 인구집단을 구성할 수 있는 방법론을 모색해야 할 것이며, 실제 급여의 수급판정에 적용되는 기준을 적용할 수 있도록 해야 할 것이다. 아울러 대상자의 행태 변화 및 시간의 흐름을 반영하여 미래의 변화를 예측할 수 있는 다년도 시점의 동태적 시뮬레이션으로 수행되어야 할 것이다.

참 고 문 헌

- [1] 이현주, “빈곤과 차상위계층,” 보건복지포럼, 제41권, pp.26-38, 2000.
- [2] 박능후, 여유진, 김계연, 임완섭, 송연경, 박소연, *탈빈곤정책 현황과 발전방안 연구*, 한국보건사회연구원, 2003.
- [3] 이현주, 강신욱, 김안나, 김태완, 노대명, 박세경, 이태진, 정경희, 신현웅, 손창균, 최현수, 박능후, 이선우, 홍경준, 유진영, 전지현, 유정예, 윤필경, *2006년 차상위계층 실태분석 및 정책제안*, 한국보건사회연구원, 2008.
- [4] 김교성, 김성욱, 이정면, 노혜진, “차상위계층의 규모와 특성 및 지위 변화,” 한국사회복지조사연구, 제22권, pp.83-108, 2009.
- [5] 김교성, 김성욱, 이정면, 노혜진, “빈곤의 측정과

규모에 관한 연구,” 한국사회복지조사연구, 제19권, pp.297-320, 2008.

- [6] 오병돈, 김기현, “청소년복지정책 대상규모 추정: 학교밖 청소년을 중심으로,” 청소년복지연구, 제15권, 제4호, pp.29-52, 2013.
- [7] <http://www.census.go.kr>
- [8] <http://mdis.kostat.go.kr>
- [9] <http://panel.kipf.re.kr>
- [10] M. D’Orazio, “Statistical Matching and Imputation of Survey Data with the Package StatMatch for the R Environment,” R package vignette, 2011.
- [11] 보건복지부, *2010 기초노령연금 사업안내*, 2010.
- [12] X. Van Pelt, “The Fusion Factory: A Constrained Data Fusion Approach,” MSc Thesis, Leiden Institute of Advanced Computer Science, 2001.
- [13] 이영섭, 김선웅, 안홍엽, 임경은, 김희경, “통계조사자료와 행정자료 간의 통계적 매칭 기법에 관한 연구,” 통계연구, 제14권, 제1호, pp.82-98, 2009.
- [14] OECD SOCX(<http://www.oecd.org/social/expenditure.htm>)

저 자 소 개

추 병 주(Byung-Joo Chu)

정회원



- 2004년 2월 : 명지대학교 행정학과(행정학 학사)
- 2006년 2월 : 명지대학교 대학원 행정학과(행정학 석사)
- 2009년 8월 : 명지대학교 대학원 행정학과(행정학 박사)

▪ 2018년 3월 ~ 현재 : 명지대학교 행정학과 객원교수
 <관심분야> : 복지정책, 전자정부, 지방행정, 시뮬레이션

김 성 훈(Sung-Hoon Kim)

정회원



- 2002년 2월 : 건국대학교 철학과 (문학사)
- 2007년 2월 : 건국대학교 대학원 사회복지학과(사회복지학 석사)
- 2012년 12월 ~ 2016년 12월 : 사회보장 정보원 연구원

▪ 2017년 5월 ~ 현재 : ㈜케이넷 빅데이터분석팀 데이터분석가

<관심분야> : 데이터마ining, 시물레이션, 모델링

정 상 기(Sang-Gee Jung)

정회원



- 2012년 2월 : 광운대학교 경영학과(경영학 학사)
- 2014년 8월 : 한양대학교 대학원 경영학과(경영학 석사)
- 2015년 3월 ~ 2016년 8월 : 사회보장 정보원 위촉연구원

▪ 2017년 4월 ~ 현재 : 한국조세재정연구원 위촉연구원

<관심분야> : 복지정책, 시물레이션

김 경 준(Kyung-Joon Kim)

정회원



- 2005년 2월 : 인천대학교 경영학과(경영학 학사)
- 2007년 2월 : 인천대학교 대학원 경영학과(경영학 석사)
- 2017년 2월 : 인천대학교 대학원 경영학과(경영학 박사)

▪ 2018년 3월 ~ 현재 : 제주 인적자원개발위원회 책임연구원

<관심분야> : 정보화정책, 사회보장정보화

이 인 수(In-Soo Lee)

정회원



- 2007년 2월 : 명지대학교 행정학과(행정학 학사)
- 2009년 8월 : 명지대학교 대학원 행정학과(행정학 석사)
- 2015년 5월 ~ 현재 : 사회보장 정보원 연구원

<관심분야> : 복지정책, 전자정부

김 현 정(Hyung-Jung Kim)

정회원



- 2006년 2월 : 중앙대학교 사회복지학과(사회복지학 학사)
- 2012년 2월 : 중앙대학교 대학원 사회복지학과(사회복지학 석사)
- 2015년 3월 ~ 2016년 8월 : 사회보장정보원 위촉연구원

▪ 2016년 9월 ~ 현재 : 중앙대학교 사회복지학과 박사과정

<관심분야> : 복지정책, 빈곤정책