연구보고서 2023-29

공적연금의 개혁 효과 분석을 위한 동태적 미시 모의실험 모형 개발

류재린
이다미·이원진·남윤재·이병재·권혁진

사람을
생각하는
사람들

KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS

한국보건사회연구원
KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS

▌연구진

연구책임자	**류재린**	한국보건사회연구원 부연구위원
공동연구진	**이다미**	한국보건사회연구원 부연구위원
	이원진	한국보건사회연구원 부연구위원
	남윤재	한국보건사회연구원 연구원
	이병재	한국보건사회연구원 전문연구원
외부연구진	**권혁진**	경상국립대학교 교수

연구보고서 2023-29

공적연금의 개혁 효과 분석을 위한
동태적 미시 모의실험 모형 개발

발 행 일	2023년 12월
발 행 인	이 태 수
발 행 처	한국보건사회연구원
주 소	[30147]세종특별자치시 시청대로 370 세종국책연구단지 사회정책동(1~5층)
전 화	대표전화: 044)287-8000
홈페이지	http://www.kihasa.re.kr
등 록	1999년 4월 27일(제2015-000007호)
인 쇄 처	㈜에이치에이엔컴퍼니

ⓒ 한국보건사회연구원 2023

ISBN 979-11-7252-003-8 [93330]
https://doi.org/10.23060/kihasa.a.2023.29

발|간|사

최근 공적연금의 개혁 논의가 활발하지만, 큰 진전 없이 제자리걸음을 하고 있다. 이는 우리 제도가 직면한 상황이 매우 복잡하고, 그만큼 다양한 쟁점들이 존재하기 때문일 것이다. 이처럼 복잡한 사안에 대해 논의할 때에는 제도 변화의 파급 효과를 사전적으로 검토할 수 있는 분석 도구의 역할이 중요하다. 특히나 연금개혁의 효과는 장기간에 걸쳐 나타나며, 소득계층과 출생코호트에 따라 그 영향이 상이하므로, 연금개혁의 미시적 파급 효과에 대한 검토가 필요하다. 그러나 아직까지 연금개혁 논의에서 개혁 방안의 미시적 효과를 사전적으로 분석한 자료는 찾아보기 어렵다. 이는 연금개혁 논의가 진전되지 못하는 이유 중 하나일 것이다. 이러한 맥락에서 이 연구는 공적연금의 개혁 효과 분석을 위한 동태적 마이크로시뮬레이션 모형의 개발을 목적으로 하였다. 본 연구에서 구축한 모형이 향후 연금개혁 논의에서 중요한 도구로 쓰일 것을 기대한다.

이 연구는 류재린 부연구위원이 책임을 맡고, 경상국립대학교 권혁진 교수와 본 연구원의 이다미, 이원진 부연구위원과 남윤재 연구원, 이병재 전문연구원이 참여하였다. 고용정보원의 박진희 연구위원과 본 연구원의 송창길 부연구위원의 자문 또한 연구의 질을 높이는데 큰 도움이 되었다. 마지막으로 이 보고서의 내용은 한국보건사회연구원의 공식 견해가 아님을 밝힌다.

2023년 12월

한국보건사회연구원 원장

이 태 수

목 차

Abstract ····· 1

요약 ····· 3

제1장 서론 ····· 5

제1절 연구의 배경 및 목적 ····· 7

제2절 연구의 내용 및 방법 ····· 10

제2장 마이크로시뮬레이션 모형의 개념과 활용 ····· 13

제1절 개념과 발전 과정 ····· 15

제2절 기술적 특징과 분류 ····· 18

제3절 개발 및 활용 사례 ····· 27

제3장 KIHASA SIM의 구조와 활용 자료 ····· 35

제1절 모형의 개요 ····· 37

제2절 모형의 구조 ····· 40

제3절 활용 자료 ····· 46

제4장 기초 자료 구축 ····· 53

제1절 기초 자료 구축 ····· 55

제2절 인구 및 가구 보정 ····· 61

제3절 소득 분포 보정 ····· 66

제4절 국민연금 가입 및 수급 정보 보정 ····· 71

제5장 주요 모듈 구축 ····· 77

제1절 인구 모듈 ····· 79

제2절 가구 구성 모듈 ····· 84

제3절 교육·경제활동 모듈 ····· 87

제4절 소득추정 모듈 ····· 101

제5절 공적연금 모듈 ····· 113

제6절 고용 보험 모듈 ····· 127

제6장 시뮬레이션 결과 ····· 137

제1절 기본 시뮬레이션 결과 ····· 139

제2절 공적연금 개혁 효과 분석 ····· 144

제7장 결론 ····· 159

제1절 주요 내용 요약 ····· 161

제2절 한계와 향후 과제 ····· 163

참고문헌 ····· 169

부록 ····· 177

부록1. 변수일람표 ····· 177

표 목차

〈표 2-1〉 주요국의 MSM ····· 27
〈표 2-2〉 EU 회원국의 연금 분석·전망 모형 ····· 31
〈표 2-3〉 국가별 연금 모형에 대한 행정관리(administrations) 상의 분류 ····· 32
〈표 3-1〉 KIHASA SIM의 시뮬레이션 대상 생애 사건 ····· 41
〈표 3-2〉 KIHASA SIM의 폴더 배치와 기능 ····· 45
〈표 3-3〉 KIHASA SIM의 주요 생애사건과 활용 자료 ····· 46
〈표 3-4〉 DOSA와 KIHASA SIM의 활용 자료 비교 ····· 48
〈표 3-5〉 통계청(2021)과 송창길 외(2023)의 인구변동요인 결과 비교 (중위가정 적용, 2070년까지) ····· 50
〈표 3-6〉 통계청(2021)과 송창길 외(2023)의 인구변동요인 결과 비교 (LLG 적용: 중위, 2120년까지) ····· 50
〈표 3-7〉 경제변수 가정(기재부(2020)의 중위가정) ····· 51
〈표 4-1〉 주요 변수의 가공 및 처리 ····· 57
〈표 4-2〉 상용근로 확률 추정 결과 ····· 60
〈표 4-3〉 기초자료와 경제활동인구조사의 사회경제적 특성 ····· 65
〈표 4-4〉 기초자료의 소득 생성을 위한 소득방정식 추정 결과 ····· 66
〈표 4-5〉 기초자료의 성별 평균 근로(사업)소득 비교 ····· 68
〈표 4-6〉 기초자료와 노동패널의 종사상 지위별 평균 근로(사업)소득 비교 ····· 68
〈표 4-7〉 성-연령-가입종별 국민연금 가입자 규모 비교(2015년 기준) ····· 73
〈표 4-8〉 국민연금 가입확률 추정에 사용된 변수의 구성 및 특징 ····· 74
〈표 4-9〉 기초자료의 국민연금 가입확률 추정 결과 ····· 75
〈표 5-1〉 출산 모듈에 활용되는 기초 자료 ····· 79
〈표 5-2〉 주요 MSM의 출산 관련 변수 ····· 81
〈표 5-3〉 출산확률 추정에 사용된 변수의 구성과 특징 ····· 82
〈표 5-4〉 출산확률 추정결과 ····· 82
〈표 5-5〉 사망 모듈에 활용되는 기초 자료 ····· 83
〈표 5-6〉 가구 구성모듈에 활용되는 기초자료 ····· 84

〈표 5-7〉 혼인율 집계치 조정 자료(2022년) ····· 85
〈표 5-8〉 교육 모듈에 활용되는 기초자료 ····· 88
〈표 5-9〉 대학 진학 확률 추정에 사용된 변수의 구성 및 특징 ····· 89
〈표 5-10〉 대학 진학 확률 추정 결과: 다항로짓 분석 ····· 89
〈표 5-11〉 경제활동 상태 모듈에 활용되는 기초자료 ····· 91
〈표 5-12〉 주요 MSM의 경제활동 상태 설명 변수 ····· 93
〈표 5-13〉 경제활동상태 전이확률에 대한 다항로짓분석 결과(남성) ····· 94
〈표 5-14〉 경제활동상태 전이확률에 대한 다항로짓분석 결과(여성) ····· 95
〈표 5-15〉 경제활동상태별 이행 비율(40~44세) ····· 97
〈표 5-16〉 한국노동패널과 KIHASA SIM의 경제활동상태 전이확률 추정 결과 (2016~2021년) ····· 99
〈표 5-17〉 주요 MSM의 근로소득 추정 변수 ····· 103
〈표 5-18〉 소득방정식 패널분석 결과(AR(1) 확률효과 GLS 모형) ····· 104
〈표 5-19〉 성별 평균 근로(사업)소득 비교 ····· 106
〈표 5-20〉 종사상 지위별 평균 근로(사업)소득 비교(2021년) ····· 106
〈표 5-21〉 국민연금 모듈에 활용되는 기초 자료 ····· 113
〈표 5-22〉 제5차 국민연금 재정계산의 가입률 가정 ····· 115
〈표 5-23〉 지역가입자의 성별-연령별 징수율 지수 ····· 115
〈표 5-24〉 성-연령-가입종별 가입개월 수 분포(50-55세 남성) ····· 117
〈표 5-25〉 국민연금 가입확률 추정에 사용된 변수의 구성 및 특징 ····· 119
〈표 5-26〉 국민연금 가입확률 추정 결과 ····· 120
〈표 5-27〉 종사상 지위별 국민연금 가입률 ····· 122
〈표 5-28〉 고용 보험 모듈 ····· 128
〈표 5-29〉 고용보험 가입확률 추정(로짓 분석)에 사용된 변수의 설정 ····· 129
〈표 5-30〉 18~59세 임금근로자의 국민연금과 고용보험 가입률(2004~2022년) ····· 130
〈표 5-31〉 고용보험 가입확률 추정: 로짓 분석 결과 ····· 132
〈표 5-32〉 성별·연령별 연간 고용보험 가입 개월 수(2022년) ····· 134

〈표 6-1〉 제5차 재정계산보고서의 정책 조합(안) ····· 144
〈표 6-2〉 출생 코호트-생애 소득분위-시나리오별 노령연금 급여 변화 ····· 147
〈표 6-3〉 출생코호트-생애 소득분위-시나리오별 소득대체율 변화 ····· 150
〈표 6-4〉 소득대체율과 보험료율 조정에 따른 재정 전망 ····· 156
〈표 6-5〉 소득대체율과 보험료율 조정에 따른 재정 전망:기금운용 수익률 1%p 상향 ·· 157
〈부표 1-1〉 기초 자료의 변수일람표 ····· 177

그림 목차

[그림 2-1] MMESP의 구조 ····· 20
[그림 2-2] SESIM의 구조 ····· 21
[그림 2-3] 이(異)시점 모형과 공간 모형에서의 복잡성 증대 ····· 22
[그림 2-4] MSM의 기술적 분류 ····· 25
[그림 3-1] KIHASA SIM의 발전 과정 ····· 39
[그림 3-2] KIHASA SIM의 순환도 ····· 42
[그림 3-3] KIHASA SIM의 모듈 내 시뮬레이션 과정 ····· 43
[그림 3-4] KIHASA SIM의 폴더 구조 ····· 44
[그림 4-1] 인구 보정 전후의 성별 인구 수 비교 ····· 63
[그림 4-2] 2015년 전체 및 종사상지위별 근로소득 분포 비교 ····· 69
[그림 5-1] 경제활동 상태 시뮬레이션 과정 ····· 91
[그림 5-2] 2016~2021년의 경제활동 상태 배열 ····· 100
[그림 5-3] 2016년의 근로 및 사업소득 분포 비교 ····· 107
[그림 5-4] 2017년의 근로 및 사업소득 분포 비교 ····· 108
[그림 5-5] 2018년의 근로 및 사업소득 분포 비교 ····· 109
[그림 5-6] 2019년의 근로 및 사업소득 분포 비교 ····· 110
[그림 5-7] 2020년의 근로 및 사업소득 분포 비교 ····· 111
[그림 5-8] 2021년의 근로 및 사업소득 분포 비교 ····· 112
[그림 5-9] 가입종별 징수율 추이 ····· 117
[그림 6-1] 생애 소득분위별 가입기간 ····· 140
[그림 6-2] 생애 소득분위별 노령연금 수급률 ····· 140
[그림 6-3] 생애 소득분위별 노령연금 급여 수준 ····· 142
[그림 6-4] 생애 소득분위별 실질 소득대체율 ····· 142
[그림 6-5] 국민연금 재정수지 전망 ····· 143
[그림 6-6] 출생 코호트-시나리오별 신규수급자의 노령연금 급여 ····· 146
[그림 6-7] 생애 소득분위-시나리오-출생코호트별 노령연금 급여 변화 ····· 148
[그림 6-8] 출생코호트-시나리오별 소득대체율 변화 ····· 149

〔그림 6-9〕 생애 소득분위-시나리오-출생코호트별 소득대체율 변화 ····· 151
〔그림 6-10〕 소득대체율 인상에 따른 재정 전망 ····· 152
〔그림 6-11〕 보험료율 인상에 따른 재정 전망 ····· 154
〔그림 6-12〕 소득대체율과 보험료율 조정에 따른 재정 전망 ····· 158
〔그림 6-13〕 소득대체율과 보험료율 조정에 따른 재정 전망: 기금운용 수익률 1%p 상향 ·· 158

Abstract

Developing a Dynamic Microsimulation Model to Analyze the Effect of Public Pension Reforms

Project Head: Ryu, Jaerin

This study aims to develop KIHASA SIM, a dynamic microsimulation model, to assess the effects of public pension reforms. To enhance relevance and accuracy, KIHASA SIM is constructed by extensively updating and reorganizing DOSA, a microsimulation model developed by the Korea Institute of Health and Social Affairs in 2016. It uses a 2% sample (998,459 individuals) from the 2015 Population and Housing Census data as its base set. The model integrates a wide array of micro-data from Korea to simulate key life events, including childbirth, marriage, divorce, education, earnings, public pension, asset accumulation, and mortality.

We verify the simulation results by comparing them with the income distribution and economic activity transition probabilities observed in the Korean Labor Panel data, confirming their similarity. We also confirm that this model's long-term financial outlook is similar to the results of the National Pension Service's 5th actuarial forecast.

These results underscore the practical utility of KIHASA SIM, demonstrating its capability to provide insightful analyses of

Co-Researchers: Lee, Dah Mi · Lee, Wonjin · Nam, Yun-Jae · Lee, Byeongjae · Kwon, Hyuk-Jin

the nuanced effects of pension reforms. This validated model is essential for policymakers and researchers in developing informed, long-term management strategies for public pensions.

Key words : Pension Reforms, Dynamic Microsimulation Model, Policy Evaluation

요약

1. 연구의 배경 및 목적

본 연구는 연금개혁의 효과 분석을 위한 동태적 마이크로시뮬레이션 모형(microsimulation model, 이하 MSM) 개발을 목적으로 하였다. 이를 위해, 본 연구원에서 2016년에 구축했던 DOSA의 기초 자료를 업데이트하고, 그간 여러 연구자들이 개별적으로 개발하였던 코드를 정리하는 한편, 시뮬레이션 로직을 개선하여 KIHASA SIM으로 재구축하였다.

2. 주요 연구 결과

각 장의 주요 내용은 다음과 같다. 먼저, 2장과 3장에는 MSM에 대한 독자들의 이해를 도울 수 있는 내용들을 담았다. 2장에서는 MSM의 개념과 정의, 발전 과정, 기술적 특징을 소개하였고, 연금 모형을 중심으로 국내외 MSM을 검토하였다. 또한 3장에서는 본 연구에서 구축한 KIHASA SIM의 구조와 활용 자료에 대해 소개하였다.

다음으로, 4장에서는 기초 자료의 구축 과정을 소개하였다. KIHASA SIM은 2015년 인구주택총조사 자료의 2% 표본(998,459명)을 기초 자료로 이용한다. 이를 모형에 투입하기 위해서는 인구 및 가구 구성에 대한 보정, 소득 분포 추정, 국민연금 가입 및 수급 정보 보정 등 수많은 전처리 과정이 필요하다. 4장에는 이러한 과정을 가능한 상세하게 담았다.

5장에서는 주요 모듈의 구축 과정을 소개하였다. 본 연구에서 소개한 모듈은 인구(출산 및 사망), 가구 구성(결혼 및 이혼), 교육 및 경제활동, 소득추정, 공적연금, 고용보험 모듈이다. 각 모듈별로 핵심적인 시뮬레이션 내용과 방법, 주요 가정, 기초 자료 등을 자세히 소개하였다. 한편으로, 이 장에서는 경제활동 모듈과 소득추정 모듈의 시뮬레이션 결과에 대한 유효성

검토를 시행하였다. 이는 본 모형의 시뮬레이션 결과와 한국노동패널 자료를 비교하는 방식으로 이루어졌는데, 이를 통해 KIHASA SIM의 시뮬레이션 결과와 한국노동패널자료가 상당히 유사함을 확인할 수 있었다.

마지막으로 6장에서는 국민연금의 노후소득보장 기능과 관련된 주요 지표(국민연금 가입 기간, 노령연금 수급률, 급여 수준, 실질 소득대체율, 재정수지 등)들을 산출하여 시뮬레이션 결과가 안정적으로 나타나는지 살펴보았다. 또한 제5차 재정계산 보고서에서 제시한 연금개혁 방안(소득대체율-보험료율-기금운용 수익률)을 토대로 각 정책 조합이 급여 수준, 소득대체율, 재정에 미치는 영향을 검토하였다. 이를 통해 KIHASA SIM이 안정적이며 납득할만한 시뮬레이션 결과를 내놓고 있으며, 시나리오 분석 시에도 안정적인 결과가 도출됨을 확인할 수 있었다.

3. 시사점

KIHASA SIM의 시뮬레이션 결과가 안정적으로 나타나는 것은 이 모형이 연금 개혁의 미시적 파급 효과를 사전적으로 분석하여 공적연금의 장기 운용 전략을 수립하는데 기여할 수 있음을 의미한다. 연금개혁 논의에서는 재정적 지속가능성, 급여 적정성, 보편성 등의 여러 목표들을 동시에 고려한 균형 잡힌 접근이 필요한데, 향후 동 모형이 정책 입안자와 연구자들에게 유용한 도구가 되기를 기대한다.

주요 용어 : 연금 개혁, 동태적 마이크로시뮬레이션, 정책 평가

제 1 장

서론

제1절 연구의 배경 및 목적
제2절 연구의 내용 및 방법

제 1 장 서론

제1절 연구의 배경 및 목적

1. 연구 배경

최근 공적연금의 개혁 논의가 활발하다. 국회의 연금개혁특별위원회(이하, 연금개혁특위)를 중심으로 국민연금의 모수 조정 뿐 아니라 노후소득보장체계 전반을 재구조화하는 다양한 개혁 방안들이 논의되고 있다. 그런데 개혁 논의가 그리 쉽게 진전되는 것 같지는 않다. 연금개혁특위는 출범 당시 '23년 4월까지 활동하며 개혁안을 모색하였으나, 연금개혁특위 산하 민간자문위원회 위원들 간 이견으로 적절한 개혁안을 마련하지 못하였고, 결국 '23년 10월까지로 활동을 연장하였다.

이처럼 개혁 논의가 난항을 겪는 것은 우리 제도가 직면한 상황이 너무 복잡하기 때문일 것이다. 현행 공적연금제도는 재정안정성과 보장성을 모두 담보할 수 없고 내·외부적 환경 또한 계속 악화되고 있어 재구조화가 불가피하지만, 재정 안정성, 급여 적정성, 보편성, 세대 간 형평성, 타 노후소득보장제도와의 정합성 등 다양한 요소들을 고려해야 하는 복잡한 상황에 처해있다(류재린 외, 2022a, pp. 211-219). 사안이 복잡한 만큼 논의의 범위가 넓고, 쟁점들이 다양하며, 이를 달성하기 위한 정책 조합 역시 다양하다.

이렇게 복잡하고 다양한 개혁 방안들 중 최적의 방안을 찾기 위해서는 각 개혁 방안의 파급 효과에 대한 사전적 검토가 필수적이다. 사회정책의 영역에서는 가능한 대안을 포괄적으로 모색하는 것 못지않게, 대안 모색에

필요한 합당한 증거들을 검토하는 것이 중요하기 때문이다(오세영 외, 2017, p.3). 특히, 연금개혁은 생애기간 동안의 가계소득과 소비에 직접적인 영향을 미치므로, 연금개혁의 거시적 파급 효과 뿐 아니라 미시적 파급 효과에 대한 세밀한 검토가 필요하다.

그러나 현재의 연금개혁 논의는 미시적 파급 효과에 대한 제대로 된 검토 없이 이루어지고 있다. 그간 연금개혁 방안을 제시했던 많은 연구들은 연금개혁의 효과를 정성적으로 제시하거나, 성·연령별 구성 집단에 기반한 평균적인 추계 결과를 제시하는데 그쳤다. 그간 연금개혁 논의에서 주로 사용되었던 부과방식 비용률, 필요 보험료율, GDP 대비 급여 지출, 실질 소득대체율, 실질 급여액 등의 지표들은 모두 평균적인 접근을 통해 얻어낸 거시적 지표들이며, 이들을 통해서는 출생 코호트, 소득계층별로 상이한 연금개혁의 효과를 파악하기 어렵다. 본 연구원에서는 2016년에 DOSA를 개발한 뒤, 이를 공적연금의 개혁 효과 및 적정성 평가에 활용해 왔으나(고제이 외, 2016; 정해식 외, 2020; 주은선 외, 2017 등), 최근의 연금개혁 논의에는 이를 활용하지 못하고 있는 실정이다. 동 모형이 개발된 지 오래되어 기초자료와 기초율의 업데이트가 필요하기 때문이다.

한편으로, 연금개혁안에 대한 사전적 검토가 이루어지지 않는 것은 사회정책에 대한 모니터링과 평가(monitoring and evaluation) 체계가 부재함을 의미하기도 한다. 사회정책의 모니터링과 평가는 정책의 필요성과 작동 원리를 탐구하여 사회정책의 핵심 기능에 대한 증거(evidence)를 제시하는 과정을 의미한다(OECD, 2019, p.11). 또 이렇게 제시된 증거는 정책 결정의 기초자료가 되어, 정책 결정 과정을 투명하게 하고 정부 예산을 효율적으로 집행하는데 기여할 수 있다. 특히나 연금 개혁과 같이 정책의 파급 효과가 장기간에 걸쳐 나타나는 경우 모니터링과 평가 체계의 중요성이 더욱 커진다.

이에 선진국에서는 다양한 분석 도구들을 활용하여 사회정책에 대한 모니터링과 평가 체계를 구축하고 있는데, 그중 하나가 마이크로시뮬레이션 모형(Microsimulation Model, 이하 MSM)이다. MSM은 미시 자료에 기초하여 개별 개체에게 발생하는 주요 사건들을 세밀하게 모형화하므로, 정책 변화가 상호이질적인 미시적 개체에 미치는 영향을 분석하기에 적합하다(Zaidi et al., 2009). 이를 고려하여 다수의 국가들이 인구, 조세, 공적 이전, 연금 등 다양한 분야의 정책효과를 사전적으로 평가하는데 MSM을 이용하고 있다(Li et al., 2014, p.307).

반면, 우리나라는 사회정책에 대한 전문화된 모니터링과 평가 체계를 구축하기 시작한 지 얼마 지나지 않았으며(강신욱 외, 2015), 이를 위한 분석 도구인 MSM 개발 역시 아직까지 초기 단계라 할 수 있다. 현재 사회과학 분야에서 활용되고 있는 MSM은 손에 꼽을 만큼 드문 실정이다.

2. 연구 목적

이러한 맥락에서 본 연구는 공적연금의 개혁 효과 분석을 위한 Dynamic MSM인 KIHASA SIM의 개발을 목적으로 한다. 보다 구체적으로는, 생애기간 동안 개인이 경험하는 다양한 의사결정(출생·혼인·교육·근로·저축·사망 등)의 복잡한 상호 과정을 모형화하고, 이에 기초하여 공적연금 개혁안들의 장기적 분배 및 재정 효과를 평가할 수 있는 MSM 개발을 목적으로 한다. 이를 위해 기존 구축된 모형인 DOSA의 기초자료(base set), 모형 구조, 모의실험 알고리즘 등을 개선하고, 모형의 활용 가능성과 확장 가능성에 대해 검토하려 한다.[1)]

1) 이 모형에 대해서는 고제이 외(2016)을 참조하기 바란다.

제2절 연구의 내용 및 방법

1. 연구 내용

본 연구의 주요 내용은 다음과 같다. 첫째, MSM 방법론에 대한 최신 동향을 파악하고 해외 주요국의 MSM을 조사한다. 이는 본격적인 모형 개발에 앞서 MSM에 대한 이해도를 높이고, 최근의 기술 발전을 파악하기 위함이다. 특히, 해외 주요 모형들에 대한 검토를 통해 모형 개발에 대한 시사점을 도출하고 향후의 발전 방향을 설정하려 한다.

둘째, 기존 구축된 모형인 DOSA를 전반적으로 개선하여 새로운 모형을 구축한다. 최신 자료를 이용하여 기초자료와 집계치 조정값을 최신화하고, 각 모듈 내부의 행태방정식, 모의실험 알고리즘 등을 전반적으로 개선한다. 한편으로, 본 연구는 그간 여러 연구를 통해 개별적으로 개발된 모듈들을 통합한다는 점에서 의의를 지닌다. 예를 들어, 권혁진과 류재린(2018)이 개발한 국민연금의 모수 개혁 모듈과, 이다미와 권혁진(2019)이 개발한 국민연금의 실업크레딧 모듈, 류재린 외(2022b)가 개발한 보험료 지원 모듈 등을 결합하여 모형의 완성도를 높이는 것이 이 연구의 주요 과업 중 하나이다.

셋째, 모형의 기술적 안정성, 결과의 타당성을 검토하기 위한 유효성 검토(validation)를 시도한다. 특히, 그간 검토되지 않았던 근로 및 사업소득 추정, 노령연금 급여 등에 대한 유효성 검토를 시도한다. 이는 기본적으로 실적치가 존재하는 기간에 한해 시뮬레이션 결과와 실적치를 비교하는 방식으로 이루어질 것이다. 한편으로, 마르코프 체인 모형(Markov chain model)과 배열분석(Sequence Analysis, SA)을 이용해 노동시장 이행 과정의 모형화 정도를 점검한다. 이는 자산 모듈에 대한

유효성 검토만을 시행하였던 선행연구인 고제이 외(2016)와 차별되는 지점이기도 하다.

넷째, 최근 우리나라에서 논의되고 있는 연금개혁 방안들을 적용하여 시뮬레이션 결과가 안정적인지 살펴본다. 본 연구에서 적용하는 시나리오는 최근 발표된 제5차 국민연금 재정계산의 모수개혁 방안들(소득대체율 45%, 50% - 보험료율 12%, 15%, 18%)이다(국민연금 재정계산위원회, 2023). 동 방안들을 통해 출생코호트별, 생애 소득분위별로 급여 적정성과 재정적 지속가능성이 어떻게 달라지는지 살펴볼 것이다.

2. 연구 방법

이 연구는 이상의 연구 내용을 달성하기 위해 다음과 같은 방법들을 이용한다. 첫째, 문헌 연구를 통해 국내외 모형들을 검토한다. 이를 바탕으로 KIHASA SIM의 개발 및 개선 방향을 설정하고 모형의 확장가능성을 모색하려 한다. 최근까지 다양한 MSM들이 개발되고 있고, 모형의 구조와 시뮬레이션 알고리즘 등이 구체적으로 공개되는 추세이므로, 이러한 문헌 연구가 모형 구축에 큰 도움을 줄 것이다.

둘째, 전문가 자문 회의를 통해 모형의 방법론적 개선과 관련된 조언을 얻고, 각 생애사건별 행태방정식 추정 과정, 시뮬레이션 알고리즘 등이 논리적으로 설정되었는지 검토한다. 본 연구에서 구축하는 모형은 다양한 생애 사건들 반영하고 있으며, 각 생애 사건에 대한 가정의 합리성은 모형의 설명력과 직결된다. 연구진 외부의 전문가 자문은 모형에 적용된 여러 가정을 합리적으로 설정하는데 크게 기여할 것으로 보인다.

셋째, 모형의 구축 결과를 점검하는 차원에서 시나리오 분석을 시행한다. 소득계층별 가입기간, 노령연금 수급률, 소득대체율, 기금소진 시점 등의

지표들이 시나리오에 따라 어떻게 달라지는지 살펴봄으로써 모형이 안정적으로 구축되었는지 점검하려 한다. 이때, 이 연구에서 MSM 개발을 시도하는 궁극적 목적은 연금개혁의 미시적 파급 효과를 분석하는 것이지만, 모형 개발 초기 단계에서 이러한 목표를 달성하는 것은 매우 어려운 일이며, 따라서 이 연구의 시나리오 분석이 시뮬레이션 결과의 정합성과 안정성을 검토하기 위한 것임에 유의해야 한다.

제2장

마이크로시뮬레이션 모형의 개념과 활용

제1절 개념과 발전 과정
제2절 기술적 특징과 분류
제3절 개발 및 활용 사례

제 2 장 마이크로시뮬레이션 모형의 개념과 활용

제1절 개념과 발전 과정

1. 개념과 정의

오늘날 사회정책이 직면한 복잡한 문제들을 분석하고 해결하기 위해서는 시스템적 사고(systems thinking)가 필요하다. 사회 문제와 관련된 각각의 요소에 주목하기보다, 해당 문제가 발생한 전체적인 맥락을 파악하고 각 요소들이 어떻게 상호작용하는지 파악해야만 문제의 본질에 접근할 수 있기 때문이다.

시뮬레이션 모델링(Simulation modeling)은 이에 유용한 도구가 될 수 있다. 시뮬레이션은 시간의 흐름에 따른 실제 프로세스나 시스템의 작동을 모방하여, 현실 세계의 복잡한 문제들을 이해하는데 도움을 준다(Banks, 1998, p.3). 현실을 추상적으로 모사하고 분석하여, 문제의 본질에 접근하기 쉽게 도와주는 것이다. 특히, 다양한 시뮬레이션 모형들 중에서도, 상호이질적인 미시적 개체의 행태 변화와 상호작용에 주목하는 MSM은 복잡한 사회정책의 구조를 세밀하게 묘사하는데 적합한 분석 도구이다(Brown & Harding, 2002, p.4-6).

O'Donoghue(2014, p.1)에 따르면, MSM은 '개인, 가구 혹은 기업 등의 미시적 분석 단위 수준(micro-level)에서 작동하는 시뮬레이션 기반 도구'로 정의된다. 이는 고유한 식별자와 성별, 연령, 결혼 상태, 고용 상태 등의 정보를 지닌 미시적 개체들에 일정한 규칙(set of rules)을 적용하여 현실

세계를 모형화하는 것이 일반적이다(Brown & Harding, 2002, p.6). 이때, 적용되는 규칙들은 확정적(deterministic)일 수도 있고, 확률적(stochastic)일 수도 있다. 시뮬레이션 결과는 관련 규칙들을 적용하여 얻어낸 추정치이며, 전체적인 집계 차원의 변화와 분포 차원에서의 변화를 포함한다.

이처럼 미시적 분석 단위에 대한 시뮬레이션을 수행하는 MSM은 사회정책의 모니터링과 평가에 효과적으로 이용될 수 있다. 순차 과정(ageing process)[2]을 통해 각 개인이 경험하는 주요 생애사건들을 세밀하게 모형화하고, 정책 변화가 미시적 개체들의 장기적 이동 경로에 미치는 영향을 사전적으로 분석할 수 있기 때문이다(Figari et al., 2014, p.47). 즉, MSM은 어떠한 모형보다도 다양하고 구체적인 정책 시나리오 분석이 가능하다는 장점을 지니고 있다. 이에 Li et al.(2014, p.305)은 개인/가구로 구성된 가상적인 자료를 이용해 수많은 '만약'에 관한 질문('what-if' questions)들에 대답할 수 있는 분석 도구로 MSM을 정의하기도 한다.

한편으로, MSM은 사회정책의 복잡성과 이질성(heterogeneity)을 확인하기에 유용하다(Lay-Yee & Cotterell, 2015, p.306). MSM은 완전히 이질적인 개인들을 분석 대상으로 하며, 평균적인 경향에서 이탈한 개체들의 이동 경로를 추적·관찰할 수 있기 때문이다. 이에 Harding et al.(2010, p.46)은 '정책이 직접적으로 영향을 미치는 개인/가구를 분석 단위로 하여 정책 변화의 영향 혹은 개인/가구의 행태 변화에 따른 복잡한 현실을 모형화하는 모형'으로 MSM을 정의하기도 한다.

---

2) 순차과정(ageing process)은 미시적 개체들이 시간의 흐름에 따라 변화하는 과정을 의미한다. 이에 대해서는 2절에서 자세히 살펴본다.

2. 등장과 발전

MSM의 개념을 최초로 제안한 것은 Orcutt(1957)이다. 그는 당시 유행하던 거시 모형의 한계를 극복하기 위해 미시적 의사결정주체(개인, 가족, 그리고 기업) 간 상호작용을 중심으로 하는 새로운 분석 모형이 필요함을 제안하였다(Morgan, 2012, p.315). Orcutt(1957, p.117)은 입력(inputs)을 받아 출력(outputs)을 생성하고 다양한 종류의 개체들이 상호작용하는 모형을 통해 주어진 외부 조건들과 정부의 행태 변화를 분석할 수 있다고 주장하였다. 이후 MSM의 개념은 Orcutt(1960), Orcutt et al. (1961) 등을 통해 구체화되었고, 조세, 복지, 보건 등 여러 분야에서 활발하게 이용되고 있다. Lay-Yee and Cotterell(2015, p.306)은 MSM 개발이 사회과학 모델링의 새로운 청사진을 제시한 것이라 평가하고 있다.

그런데 MSM은 개념이 정립된 지 한참 뒤에서야 본격적으로 발전하기 시작하였다. MSM의 개발에는 방대한 미시 자료의 축적과 데이터 처리 능력이 필요하기 때문이다. 예를 들어, 프랑스에서 1970년대에 개발한 소득세 모형인 MIR(Modèle de l'impôt sur le Revenu)은 기초자료의 발전과 모형의 발전이 함께 이루어졌다. 동 모형은 개인 자료인 노동력 조사(Labour Force surveys)의 부가조사인 조세수입(Revenus fiscaux) 조사와 국세청의 조세 자료를 매칭·결합한 자료를 기초자료로 활용하였는데, MIR 개발 초기에는 조세수입 조사의 조사 주기가 5년이었기에 시간의 흐름에 따른 변화를 분석하는 동태모형을 구축하지 못하였다. 그러나 1996년부터 조세수입 조사의 조사 주기가 1년으로 단축되었고, 2005년부터는 동 조사가 조세 및 사회보장 소득(Revenus fiscaux et sociaux, RFS) 조사로 전환되면서 가구별 사회보장급여 정보를 결합할

수 있게 되었다. 이처럼 기초자료의 조사 주기가 단축되어 모형 개발에 필요한 자료가 충분히 구축된 이후에서야 MIR은 동태모형으로 발전할 수 있었다(Legendre, 2019; 권혁진, 2022, pp.14-15에서 재인용).

정보처리 기능의 발전이 MSM 개발에 미친 과정 역시 이와 유사하다. 1960~70년까지는 천공카드(punched card)로 데이터를 기록·저장하였고 대용량 자료는 마그네틱 테이프로 보관했었다. 따라서 재귀적(recursive) 반복 과정을 기본으로 하는 MSM의 개발이 거의 불가능하였다(권혁진, 2022, p.16). 그러나 1980년대 중·후반 이후부터 PC가 보급되고 정보처리 기술이 급속히 발전하면서 MSM의 방법론이 본격적으로 발전하기 시작하였다. 1980년대는 미국과 유럽에서만 30개가 넘는 MSM이 개발되었으며(Merz, 1991, pp.88-92), 이후에도 전 세계적으로 수많은 MSM이 개발되어 조세, 인구, 보건, 연금 등의 분야에 활용되고 있다(Anderson & Hicks, 2011, p.2).

제2절 기술적 특징과 분류

1. 기술적 특징

가. 기술적 특징

MSM은 미시적 개체를 대상으로 시뮬레이션을 수행하므로, 기본적으로 미시 자료에 기반한다. 이때, 시뮬레이션의 기초가 되는 자료를 기초 자료(Base set or baseline data)라 부르는데, 여기에는 미시적 개체들의 고유 아이디를 포함한 다양한 정보들과 가중치 등이 포함되는 것이 일반적이다.

특히, 사회정책 분석에 활용되는 MSM은 개인들의 다양한 사회·경제적 특성(소득, 고용, 교육 등)과 인구통계학적 특성(성, 연령, 가족구성 등)들로 이루어진 미시 자료에 기반한다. 최근에는 조사 자료와 행정 자료를 결합한 형태의 기초 자료가 많이 활용되고 있다.

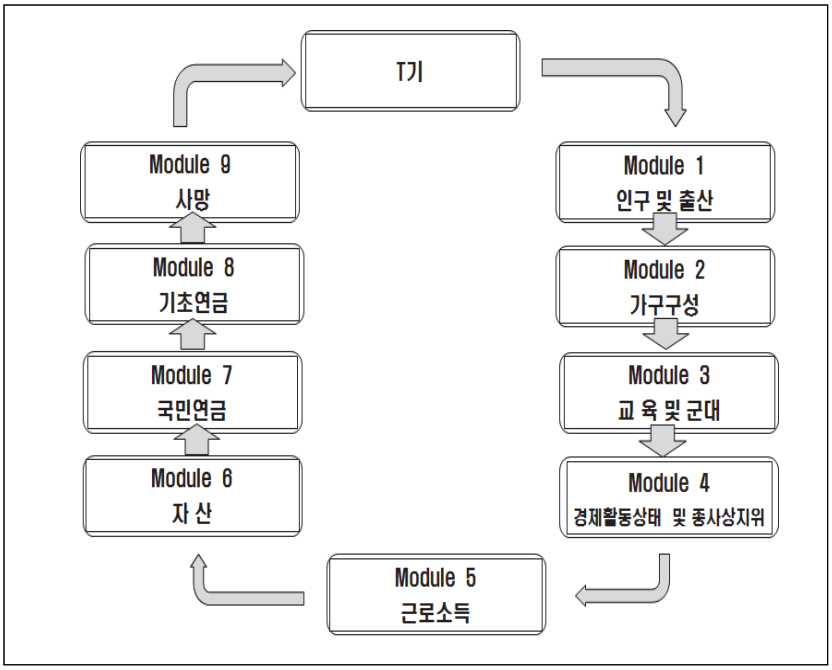
전술한 것처럼 MSM, 특히 동태적 MSM의 주된 관심사는 미시적 개체들이 경험하는 주요 사건들과 제도를 모형화하고, 그에 따른 장기적인 이동 경로를 사전적으로 분석하는 것이다. 이때 시뮬레이션 대상의 상태와 행태는 순차과정에 따라 변화하게 되는데, 순차과정은 경제학 모형에 근거한 행태방정식 추정 결과나 통계적 매칭(statistical matching)에 기반하는 것이 일반적이다. 이때 시뮬레이션 대상들은 직/간접적으로 상호작용하게 되는데, 예를 들어 개인의 생애 이력은 개인/가구가 속한 집단 구성원들의 특성과 그 사회가 직면한 사회·경제·인구적 환경의 영향을 크게 받는다.

그런데 생애주기 동안 경험하는 여러 사건들에 대해 각각의 개별 경제 주체가 보이는 다양한 행태를 모형화하는 것은 현실적으로 어려운 일이다(Klevmarken, 1997, p.9). 만일 행태 과정(behavioral process)이나 개인의 행태 간 상호작용에 대한 이론적 근거가 충분하지 않다면 설정 오류(misspecification error)의 문제가 발생할 수 있다(류재린, 2020, p. 20). 이러한 문제를 완화하는 방법들 중 하나는 시뮬레이션 결과를 거시적 차원의 외부 정보와 비교하여 수정하는 것이다(Li et al., 2014, p. 314). 모형 외부에 위치한 미래 전망치를 집계값으로 설정하고, 시뮬레이션 결과가 이와 일치 또는 유사하도록 조정하는 것이다. 이는 집계 차원의 목표를 맞추기 위해 시뮬레이션 결과를 제약하는 셈이지만(Li et al., 2014, p. 319), 비교적 쉽게 모형의 안정성을 확보할 수 있는 방법이다(고제이 외, 2016, p.61).

한편, MSM은 여러 개의 모듈(Module)로 구성되는 것이 일반적이다

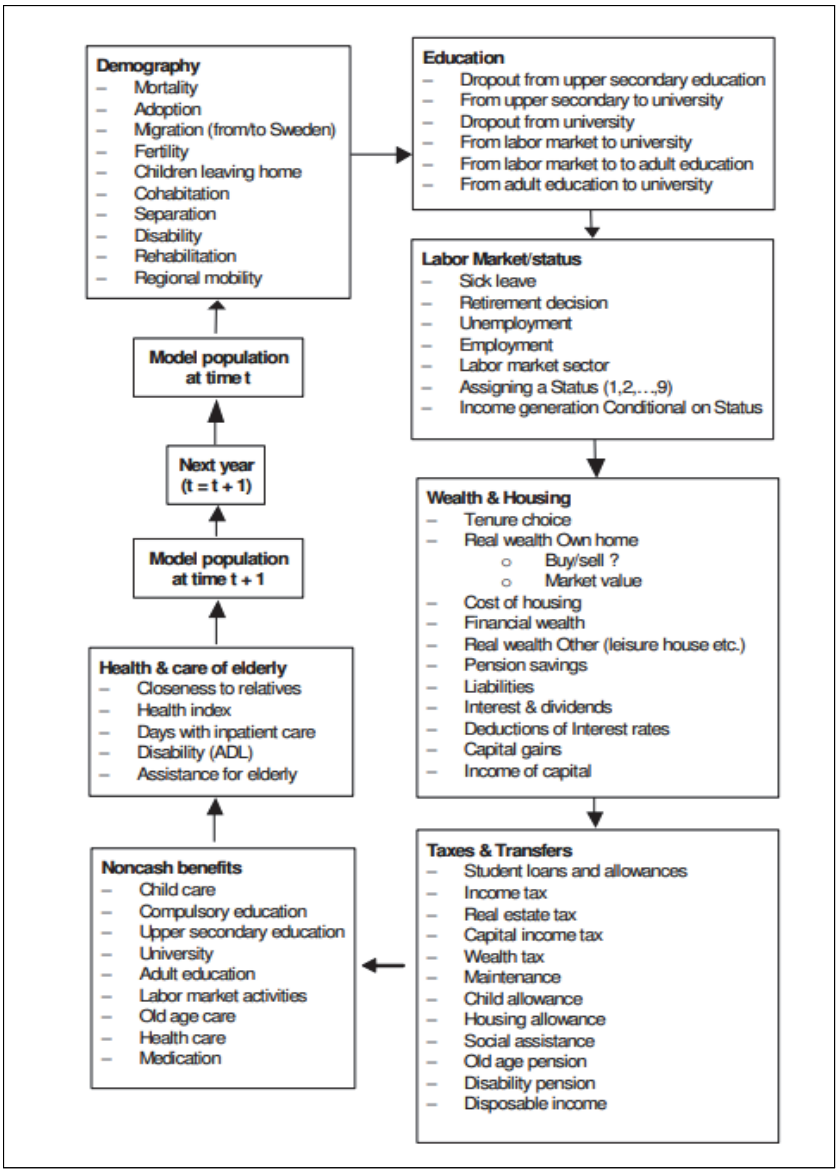
([그림 2-1], [그림 2-2] 참조). 이는 순차적으로 순환하는 모듈 형태로 모형을 구축하는 것이 생애사건 또는 제도들 사이의 이질성과 선후 관계를 고려하기에 용이하기 때문이다. 예를 들어, 기초연금의 지급을 시뮬레이션 하기 위해서는 근로 및 사업소득, 연금소득, 자산 등의 정보가 필요한데, 기초연금 모듈을 소득 추정 모듈과 자산 추정 모듈 뒤에 배치하여 필요한 정보들을 반영할 수 있다.

[그림 2-1] MMESP의 구조



자료: 권혁진, 류재린. (2015). 공적연금의 최저생계 보장 효과에 대한 장기 전망, 응용통계연구, 28(4), p. 746. Figure 2.1.

〔그림 2-2〕 SESIM의 구조

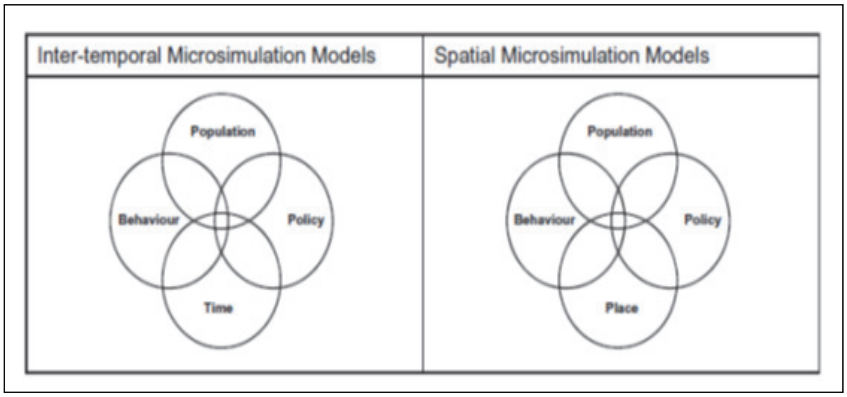


자료: Flood, L., Jansson, F., Pettersson, T., Pettersson, T., Sundberg, O., and Westerberg, A. (2012). The Handbook of SESIM-a Swedish dynamic micro simulation model. p. 6. Figure 2.1.

나. 장점과 단점

MSM은 미시적 개체를 대상으로 하므로, 개체별 이질성을 완전하게 반영할 수 있고, 이에 따라 평균적인 영향이 아닌 개별 단위에 미치는 영향, 즉 분포 변화를 살펴볼 수 있다. 이는 거시 모형이나 대표행위자 모형과 대비되는 MSM만의 특징이다. 또한 MSM은 개인 간 상호작용 뿐 아니라 제도 또는 정책 간 상호작용을 모형화함으로써 각 제도들이 미시적 개체에 미치는 복합적인 영향을 추정할 수 있다. 이처럼 평균 이면의 복잡한 역학관계(dynamics)를 모형화할 수 있다는 점도 MSM의 큰 장점이다. 이에 시간적(temporal), 공간적(spatial) 복잡성 등을 고려한 모형들이 개발되어 활용되고 있다(〔그림 2-3〕 참조). 한편으로 MSM은 제도별 효과를 분해하고 그 효과를 유발하는 요인을 사전적으로 파악할 수 있는데, 이 또한 MSM이 지닌 큰 장점 중 하나이다(O'Donoghue, 2014, p.4-6).

〔그림 2-3〕 이(異)시점 모형과 공간 모형에서의 복잡성 증대



자료: O'Donoghue. (2014). Introduction. in Handbook of Microsimulation Modelling. p.4, Figure 1.2.

그러나 MSM은 다음과 같은 한계들을 지니고 있다. 첫째, 전술한 것처럼 시뮬레이션 대상의 행태 과정이나 상호작용에 대한 정보가 충분하지 않다면, 모형의 설정 오류가 발생할 수 있다.[3] 이는 MSM의 현실 모방력(mimicking power)을 낮추어 시뮬레이션 결과와 현실 사이의 괴리를 낳는 주된 요인 중 하나이다.[4] 둘째, MSM은 거시적 단위에서의 상호작용을 고려하지 않으며, 경제의 일반균형 역시 고려하지 않는 것이 일반적이다. MSM의 시뮬레이션 결과로 얻어지는 거시적 집계치는 미시적 개체들의 단순 합일 뿐이다. 개별 주체들의 행태 변화에 주목하는 MSM의 특성을 고려하면, 이러한 한계가 이해되는 측면이 있으나, 모형의 이론적 기반이 약하다는 비판을 피하기는 어렵다. 이에 최근에는 CGE 등의 거시모형과 MSM을 top-down 방식으로 연결하거나, MSM에서도 균형을 찾아보려는 시도들이 계속되고 있다.[5] 셋째, 모형 개발에 많은 시간과 자원의 투입이 요구된다. 현실에서 나타나는 복잡한 사건들을 시뮬레이션하기 위해서는 방대한 자료와 복잡한 모형 뿐 아니라 고성능 컴퓨터와 테스트 과정이 필요하다(Spielauer, 2011, p.13). 모형의 구조를 투명하게 하고, 사용자의 접근성을 높이기 위한 적절한 인터페이스와 사용설명서를 제공할 수 있는 소프트웨어 패키지를 찾아야하는 어려움도 있다(Milne et al., 2014). 넷째, 시뮬레이션 결과에 대한 검증(validation)이 어렵다(류재린, 2020, p.21). 복잡한 현상을 다루는 MSM의 시뮬레이션

3) 이에 다수의 MSM이 집계값 조정 기법을 사용하고 있다. Li and O'Donoghue(2013) 따르면, 주요 동태적 MSM의 68%가 집계값 조정 기법을 사용하고 있다.

4) Morgan(2012, pp.320-337)은 MSM이 현미경처럼 현실을 자세히 들여다볼 수 있는 강력한 도구라고 평가하면서도, 모방력에 기반한 역추론을 경계할 것을 지적한다. 시뮬레이션을 통해 밝혀진 법칙들이 실물경제를 지배하는 행동 법칙들과 같은 영역에 속하지 않을 수 있기 때문이다.

5) 이는 경제학 이론과의 연결을 통해 MSM의 한계를 극복하기 위한 시도라 할 수 있다. 이에 관해서는 Zhang(2017), Zhang et al.(2022), Boulanger et al.(2022), Fredriksen et al.(2019) 등을 참조하기 바란다.

과정이 복잡해지는 경우, 어떠한 과정을 통해 시뮬레이션 결과가 산출된 것인지 정확히 파악하기 어려울 수 있다. 이에 다수의 MSM이 실적치가 존재하는 과거에 대한 시뮬레이션을 수행한 뒤 이를 실적치와 비교하거나, 집계값 조정 전후의 시뮬레이션 결과를 서로 비교하여 모형의 유효성을 검증하고 있다.

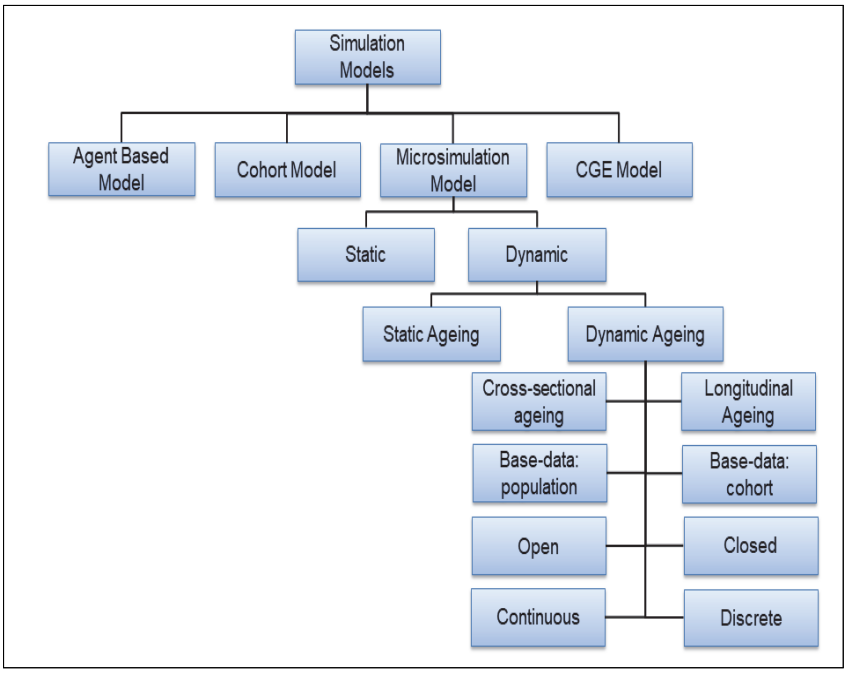
지금까지의 논의를 정리해보면, MSM은 정책결정자가 정책 효과를 이해하고 정책 개선의 필요성과 영향을 판단할 수 있는 증거들을 제공한다는 점에서 강력하고 효과적인 분석 도구가 될 수 있다(권혁진, 2022, p.7). 이는 MSM이 개인/가구의 다양한 사회·경제적 특성에 주로 의존하는 사회정책의 미래 수요를 예측하거나, 소요 재정, 소득재분배나 빈곤 등의 미시적 파급 효과를 사전적으로 분석할 수 있기 때문이다. 특히, 사회보장제도는 개인의 사회·경제·인구통계학적 특성에 따라 급여의 내용과 수준이 차별적으로 결정되며, 혼인, 출산, 경제활동, 퇴직, 사망 등 생애사건들에 따라 자격요건, 기여 및 급여 수준이 달라진다. 따라서 미시적 개체의 행태와 그들의 상호작용을 모형화하는 MSM의 강점이 두드러지는 분야라 할 수 있다.

2. 기술적 분류

Dekkers and Belloni(2009)에 따르면, MSM은 기술적 선택과 가정에 따라 세분화된다(〔그림 2-4〕 참조). 먼저, 순차과정의 유무에 따라 정태(static) 모형과 동태(dynamic) 모형으로 구분되며, 동태모형은 순차 방법, 기초자료의 종류, 시간에 대한 가정, 모형의 개방성 등의 기술적 선택과 가정에 따라 ①횡단면(cross-sectional) 모형과 종단면(longitudinal) 모형, ②전체집단(population) 모형과 코호트(cohort) 모형, ③개방(open)

모형과 폐쇄(closed) 모형, ④연속(continuous) 모형과 이산(discrete) 모형 등으로 세분화된다(Dekkers and Belloni, 2009, p.6)[6].

〔그림 2-4〕 MSM의 기술적 분류



자료: Dekkers, Gijs. and Belloni, Michele(2009), A classification and overview of micro simulation models, and the choices made in MIDAS, Chapter.2 in What are the consequences of the AWG-projections for the adequacy of social security pensions?, ENEPRI Research Report No.65, AIM WP4. p.10. Figure 1을 일부 수정.

전술한 기준들 중 가장 중요한 기준은 시간 흐름에 따라 개별 개체의 특성을 변화시키는 순차 과정, 즉 시간 차원(time dimension)의 유무라 할 수 있다(류재린, 2020, p.21). 이는 정태 모형과 동태 모형을 구분하는

6) 이에 대한 자세한 설명은 Dekkers and Belloni(2009)를 참고하기 바란다.

기준이기도 하다. 정태 모형은 특정한 제도 또는 정책 변화의 즉각적인 영향, 즉 정책의 투입(input)과 산출(output)의 직접적인 관계를 모형화한다. 반면, 동태 모형은 시간의 흐름에 따른 변화를 모형화한다. 이를 이용한 분석은 경제, 인구 성장과 같은 시기효과(time-effects)를 포함한다. 예를 들어, 전 세계에서 심화되고 있는 고령화는 한 사회 내 고령 인구의 증가를 의미하는데, 연간 단위로 순차하는 동태 모형을 이용하면 이를 쉽게 모형화할 수 있다. 또한 이를 토대로 고령화 대응 정책들의 중장기적 영향과 그 파급 효과(second-order effects)를 산출할 수 있다(권혁진, 2022, pp.6-7).

한편, 동태모형은 순차 과정을 어떻게 모형화하는지에 따라 정태적 순차모형과 동태적 순차모형으로 세분화된다. 먼저, 정태적 순차모형은 시뮬레이션 대상들의 특성을 그대로 유지한 상태로 순차과정을 시뮬레이션한다. 각 개체의 특성을 그대로 유지한 채 거시 집계치를 이용해 표본수나 가중치를 조정하는 방식이다(류재린, 2020, p.22). 반면, 동태적 순차모형은 시간의 흐름에 따른 각 개체들의 상태 변화를 반영하며, 이에 따른 다양한 행태 변화를 모형화할 수 있다. 국민연금 가입을 예로 들어보면, 정태적 순차모형에서는 개인의 특성 변화가 없으므로 개인의 사회·경제적 특성에 따른 국민연금 가입 확률이 변하지 않지만, 동태적 순차모형에서는 개인의 취업, 혼인 관계 등의 사회·경제적 상태 변화에 따라 국민연금 가입 확률이 변할 수 있다.

제3절 개발 및 활용 사례

MSM은 조세, 공적연금, 보건·의료, 교육 등 다양한 분야에서 활용되고 있다(〈표 2-1〉 참조). 특히나, 독일, 스웨덴, 영국, 미국, 캐나다 등 다수의 연금 선진국들이 소득재분배나 공적연금의 효과 분석에 MSM을 활용하고 있다. 한편으로 대부분의 모형들이 정부 또는 공공기관의 주도로 개발되었으나, SMILE, APPSIM 등 민간연구소에서 개발한 모형들도 있다.

〈표 2-1〉 주요국의 MSM

	지역/국가	기관	모형	유형	분야
정부주도	EU (15개국 공동)	EU	Euromod	static	조세
	벨기에 독일 이탈리아 (뉴질랜드)	정부부처	MIDAS	dynamic	공적연금
	노르웨이	Statistics Norway	MOSART 1/2/3	dynamic	인구, 교육, 연금, 노동공급
	스웨덴	Ministry of Finance	SESIM	dynamic	인구, 교육, 연금, 노동공급
	프랑스	French National Statistical Institute	Destinie	dynamic	공적연금
	이탈리아	Bank of Italy	DYNAMITE	dynamic	소득분배 변화
		Department of Employment and Social Policies/EC	CAPP_DYN	dynamic	장기 소득재분배
	영국	UK Department of Work and Pensions	PenSim 1/2	dynamic	공적연금
		UK Department of Health & PSSRU	CARESIM	dynamic	보건·의료
		Department of Health	DH aggregate model	static	보건·의료

	지역/국가	기관	모형	유형	분야
정부주도	미국	US Department of Health and Human Service	POLISIM	dynamic	사회보장제도
		US Department of Health and Human Service(with Urban Institute)	TRIM	dynamic	보건·의료
		Social Security Administration	MINT	dynamic	노후소득 재분배
	캐나다	Canadian Department of Human resources	DYNACAN	dynamic	공적연금
		Statistics Canada	LIFEPATHS	dynamic	보건·의료, 교육, 시간사용, 공적연금
	체코	Ministry of Labour and Social Affairs	Dynamic Microsimulation Model of the Czech Republic	dynamic	공적연금
	한국	국민연금연구원	KMAP	dynamic	공적연금
		한국보건사회연구원	DOSA	dynamic	노후소득 재분배, 공적연금
		한국고용정보원	KESIM	dynamic	노동공급
		한국조세재정연구원	KIPFSIM	static	조세
민간주도	미국	Urban Institute	DYNASIM 1/2/3	dynamic	인구, 경제
		RAND Corp.	FEM	dynamic	인구, 의료비
	아일랜드	-	SMILE	dynamic	인구
	영국	PSSRU/NCCSU	PSSRU/NCCSU-aged care models for UK	dynamic	사회보장제도
	호주	-	APPSIM	dynamic	인구, 장기정책효과
	일본	학계 전문가 협업	INAHSIM	dynamic	
	한국	-	MMESP	dynamic	노후소득 재분배, 공적연금

자료: 류재린. (2020). 국민연금제도의 사각지대와 지역가입자의 관리강화 및 지원 필요성. 연금포럼. 봄호 Vol.77. 국민연금연구원. p.24. 〈표 1〉

1. 해외 사례: EU 소속 국가들의 연금 모형[7)]

여기에서는 EU 소속 27개 국가에서 활용하고 있는 연금제도 관련 모형들을 살펴본다. 〈표 2-2〉에 따르면, 2009년 기준 EU 회원국 27개 국가에서 연금제도 모니터링에 활용하는 모형은 44개에 이른다. Gál et al.(2009)은 이들을 미시 자료의 가공·활용 여부, 개발 및 운영 주체, MSM의 역할 등을 기준으로 구분하였는데, 그 결과는 〈표 2-3〉과 같다.

이를 자세히 살펴보면, 코호트 모형을 활용하는 국가들은 미시 자료가 아니라 코호트별 평균, 가상적인 이력 혹은 거시 모형을 활용하여 중장기 연금 재정을 추계한다. 이때, 불가리아와 사이프러스 등 ILO나 World Bank 등의 국제기구에서 개발한 모형을 수입하여 활용하는 국가가 있는가 하면, 슬로바키아처럼 국제기구와 협력하여 개발한 독립적인 모형을 이용하는 국가도 있다. 한편으로, 체코, 스페인, 헝가리, 오스트리아 등과 같이 정부가 재정평가를 위해 코호트에 기반한 미시 모형의 개발을 시도하는 국가들도 존재한다.

MSM을 활용하는 국가들은 외부기관에 위탁하여 모형을 개발한 국가와, 정부 주도하에 MSM을 개발한 국가들로 나눌 수 있다. 이들 중 정부 주도하에 모형을 개발한 벨기에, 덴마크, 룩셈부르크, 네덜란드, 핀란드 등의 국가들은 자체적으로 MSM을 개발하여 재정평가 및 빈곤 분석에 활용하고 있다. 특히, 벨기에는 모형 개발과 유지를 담당하는 전문적인 기관을 두고 MSM을 계속해서 발전시켜 나가고 있으며,[8)] 덴마크는 재무부(Ministry

---

7) 이 부분의 내용은 Gál et al.(2009)과 권혁진(2022)를 주로 참고하였다.

8) Federal Planning Bureau(FPB)와 Federal Public Service Social Security(FPSSS) 등이 이에 해당한다. FPB는 코호트 모형인 MALTES, 행위자 기반 모형(ABM)인 MEP, 그리고 동태적 MSM인 MIDAS_BE를 개발·운용하고 있으며, FPSSS는 정태적 MSM인 MIMOSIS를 개발·운용한다. 이에 대한 자세한 내용은 Gál et al.(2009)을 참조하기 바란다.

of Finance)가 코호트 모형과 정태적 MSM을 동시에 개발하여 활용하고 있다.

한편, 재정 및 빈곤 평가에 MSM을 활용하는 국가들은 MSM이 주요 분석 도구인 국가들과 MSM을 보완적으로 활용하는 국가들로 나눌 수 있다.[9] 먼저, 독일, 이탈리아, 벨기에, 덴마크, 룩셈부르크, 네덜란드, 핀란드 등은 MSM을 보완적인 모형으로 활용한다. 이들은 공적연금의 재정 추계에는 코호트 모형을 사용하되, 제도개혁 시나리오 평가, 중·장기적인 소득대체율 계산, 소득분배, 불평등, 노인 빈곤 측정 등의 특별한 주제에 MSM을 활용한다.

반면, MSM을 주요한 분석 도구로 활용하는 국가들로는 프랑스, 스웨덴, 영국 등이 있다. 프랑스에서는 2000년 이후부터 연금자문위원회(Conseil d'Orientation des Reraites, 이하 COR)가 연금제도 전반에 대한 분석 작업들을 총괄하고 있다. COR이 제공하는 몇몇 가정을 공통적으로 사용하여 각각의 연금제도를 독립적으로 추계하는 방식이다. 여기에는 PRISME과 DESTINIE라는 두 개의 MSM이 이용되는데, PRISME은 연금제도 중 가장 중요한 전국노령보험금고(CNAV)의 추계에 이용되며, DESTINIE는 연금제도 개혁에 따른 잠재적 효과 분석에 이용된다. 스웨덴과 영국은 MSM 개발에 가장 적극적인 나라들이다. 스웨덴은 MSM을 활용해 공적연금의 재정을 추계하는 대표적인 나라이다. 현재 연금재정 추계에 활용되고 있는 모형은 MiMESIS인데, 그 전신은 1973년부터 개발되어 활용되었던 SESIM이다. SESIM은 학자금 대출의 효과를 분석하기 위해 개발되었으나, 이후 모형의 범위를 확장하여 생애주기별 소득과 관련된 다양한 분석에 활용되었다. 한편, 영국은 노동연금부(Department of Work and Pensions, DWP)가 2000~2004년 동안 개발한 PENSIM을 연금개혁의 효과 분석에 활용하고 있다.

---

9) 다만 이러한 경계가 항상 명확한 것은 아니다.

〈표 2-2〉 EU 회원국의 연금 분석·전망 모형

국가명	모형
Belgium	MALTESE; MEP; MIDAS_BE; MIMOSIS
Bulgaria	ILO PENS (BG)
Czech Republic	Czech Pension Model
Denmark	Danish Pension Model; LAW
Germany	AVID; German Pension Model
Estonia	Estonian Long Term Pension Budget Model
Ireland	Irish Pension Model
Greece	Greek Pension Model
Spain	Spanish Pension Model
France	DESTINIE; PRISME
Italy	CAPP_DYN; CeRP models; RGS
Cyprus	ILO PENS (CY)
Latvia	Latvian Pension Model
Lithuania	PRISM
Luxemburg	LuxMod; REDIS; SOBOLUX
Hungary	Hungarian Pension Models; NYIKA
Malta	PROST (MT)
Netherlands	GAMMA; MICROS; SADNAP
Austria	Austrian Applied Projection Models Austrian Microsimulation Model
Poland	FUS07
Portugal	ModPensPor
Romania	na
Slovenia	SIOLG 1.0
Slovakia	PROST (SK), MAJA
Finland	Finnish Centre for Pensions models
Sweden	MiMESIS, SESIM
United Kingdom	PENSIM2

주: 모형에 대한 공식적인 명칭이 없는 경우, "국가명 Pension Model"로 표기함

자료: Gál, R., Horváth, A., Orbán, G., and Dekkers, G. (2009). PENMICRO Monitoring pension developments through micro socioeconomic instruments based on individual data sources: feasibility study. TARKI Social Research Institute, Budapest. p.7. Table 1.

〈표 2-3〉 국가별 연금 모형에 대한 행정관리(administrations) 상의 분류

목적	내 용			
재정평가	코호트 모형만 활용			
		모형 수입		Bulgaria, Cyprus, Latvia, Lithuania, Malta, Portugal, Slovakia(MoF)
		자체 개발		
			미시 모형 개발 계획 없음	Estonia, Greece, Ireland, Poland, Slovakia (MoLSAF), Slovenia
			미시 모형 개발 진행 중	Czech Republic, Spain, Hungary, Austria
재정평가+분배/빈곤평가	(동학적) 미시 시뮬레이션 모형 활용			
		분배/빈곤 평가에 미시 모형 활용		
			외부 기관에 위탁 개발	Germany, Italy
			정부 부처 중심으로 개발	Belgium, Denmark, Luxemburg, Netherlands, Finland
		재정과 분배/빈곤 평가 모두 미시 모형 활용		France, Sweden, United Kingdom

자료: Gál, R., Horváth, A., Orbán, G., and Dekkers, G. (2009). PENMICRO Monitoring pension developments through micro socioeconomic instruments based on individual data sources: feasibility study. TARKI Social Research Institute, Budapest. p. 18. Table 4.

2. 국내 사례

우리나라의 경우 사회정책 분야에서 활용되는 MSM이 그리 많지 않다.[10] 패널자료를 활용해 간단한 시뮬레이션을 수행한 연구들은 상당수 존재하고 있으나, 동태적 MSM을 이용해 제도의 도입 또는 개선 효과를 사전적으로 비교·검토한 연구는 매우 드물다. 그간 MSM 개발을 시도한 연구가 존재하지 않는 것은 아니다. 권혁진과 한정림(2009), 김혜원 외(2010), 이상열 외(2019), 김대중과 이난희(2018) 등 다수의 연구들이 동태적 MSM 개발을 시도하였으나, 일회성 연구에 머물렀다. 이하에서는 MSM 개발에 성공한 국내 사례 몇 가지를 살펴본다.

먼저, 국내 최초로 개발된 정태적 MSM은 조세재정연구원에서 개발한 KIPFSIM이다(성명재 외, 2008; 성명재 외, 2010). 동 모형은 가계동향조사에 기반하여 소득세, 소비세, 재정지출(현금)의 흐름을 모형화한 조세-급여 모형(tax-benefit model)이다. 이는 정태적 모형이지만 제도 변화에 따른 경제 주체들의 행태 변화, 즉 소득세 한계세율 변화에 따른 노동 공급 효과, 소비세율 변화에 따른 소비지출 변화 효과 등 이차적인 파급 효과를 고려하였다.[11][12]

동태적 MSM으로는 한국고용정보원의 인력수급 전망 모형(KEISIM)이

---

10) 앞서 살펴본 것처럼 MSM이 미시 자료의 축적과 함께 발전된다는 점을 고려하면, 이는 다소 이례적인 상황이다. 우리나라는 경제활동인구조사, 가계동향조사(통계청), 한국노동패널(노동연구원), 한국재정패널(조세재정연구원), 한국복지패널(보건사회연구원) 등 다양한 미시 자료들이 조사·제공되고 있다. 매년 수많은 연구들이 미시 자료를 이용해 경제·사회정책들을 분석·평가하고 있으나, 대부분이 사후적(ex-post) 평가에 해당한다.

11) 다만, 성명재 외(2010)는 노동 공급 효과를 모형화하지 못하였으므로, 정책 변화에 따른 행태 변화를 명시적으로 모형화한 것은 아니라 할 수 있다.

12) 그러나 2010년 이후 KIPFSIM을 활용한 연구는 더 이상 수행되지 않고 있다. 이는 개발 과정에 필요한 자료의 부재와 인적 자원이 이탈하였기 때문으로 보인다. 다른 한편으로는, 재정패널 자료의 질적 개선에 따라 모형 개발의 유인이 저하된 것도 여러 원인 중 하나일 수 있다(권혁진, 2022).

있다. 동 모형은 고용정책기본법에 근거하여 중장기 인력수급 전망을 수행하기 위해 개발되었다. 현재 국내에서 동태적 MSM을 개발하여 지속적으로 활용되고 있는 사례는 본 연구원의 DOSA와 KEISIM이 유일하다. 동 모형은 2~3년의 개발 과정을 거친 뒤 2014년에 개발되었으며, 매년 향후 10년의 인력공급 전망을 수행하고 있다(이시균 외, 2014).

KEISIM은 「인구주택총조사」 2% 표본자료에 기초하며, '사망-출생-결혼-교육-노동시장(참가-고용-소득)' 모듈들로 구성되어 있다. 이는 정부 재정이 투입되는 일자리 사업이나 사회복지 정책에 따라 변화하는 인구구조나 노동시장 변화를 전망한다. 특히, 다양한 정책사업과 개인의 행태적 반응의 복잡한 상호작용을 직접 분석하고, 이를 반영한 인구구성 및 인력공급 전망치를 도출한다(박진희 외, 2020).[13)]

마지막으로 본 연구에서 개발하고 있는 KIHASA SIM의 전신인 DOSA는 공적연금 제도의 적정성에 대해 장기적인 전망을 목적으로 개발된 동태적 MSM이다. 이는 권혁진 외(2013)이 구축한 MMESP를 발전시킨 것으로, 2013년 개발 당시에는 '출산-교육-결혼-노동시장(고용-소득)-국민연금-사망' 모듈로 구성되어 있었으나, DOSA로 발전하면서 자산 모듈과 기초연금 모듈이 추가되었다. 동 모형은 연금개혁이 공적연금 소득과 빈곤에 미치는 영향(권혁진, 류재린, 2018; 주은선 외, 2017), 생애근로이력에 따른 공적연금 소득의 격차(신우진 외, 2016), 실업크레딧의 중장기 효과(이다미, 권혁진, 2019), 보충연금(GIS) 도입과 완전소득비례 연금 전환의 장기효과 분석(정해식 외, 2020), 지역가입자 보험료 지원 제도의 장기효과 분석(류재린 외, 2022b) 등에 활용된 바 있다.

13) 동 모형의 방법론에 대한 상세한 설명은 박진희 외(2020)를 참조하기 바란다.

제3장

KIHASA SIM의 구조와 활용 자료

제1절 모형의 개요
제2절 모형의 구조
제3절 활용 자료

제3장 KIHASA SIM의 구조와 활용 자료

제1절 모형의 개요

본 연구에서 구축하는 KIHASA SIM의 전신은 고제이 외(2016)이 개발한 동태적 MSM인 DOSA(Dynamic micro-simulation Outlook model for Social policy Analysis)이다. DOSA는 권혁진 외(2013)가 개발한 MMESP에 기반한다. 이는 국민연금의 노후소득보장 효과 분석을 위해 개발된 모형으로, 인구 모듈, 가구 구성 모듈, 교육 모듈, 경제활동 모듈, 소득 추정 모듈, 국민연금 모듈 등 DOSA의 뼈대가 되는 모형의 구조를 갖추고 있었다. DOSA는 이 MMESP에 자산 모듈과 기초연금 모듈을 추가한 뒤 각 모듈을 전반적으로 개선한 것이며, 고제이와 권혁진(2014)이 개발한 소득세 모형인 K.SIM과도 연동되는 구조로 개발되었다. DOSA가 개발된 이후에는 신우진 외(2016), 권혁진과 류재린(2018), 김준일과 류재린(2018), 주은선 외(2017), 정해식 외(2020) 등의 연구들이 DOSA를 활용한 바 있다.

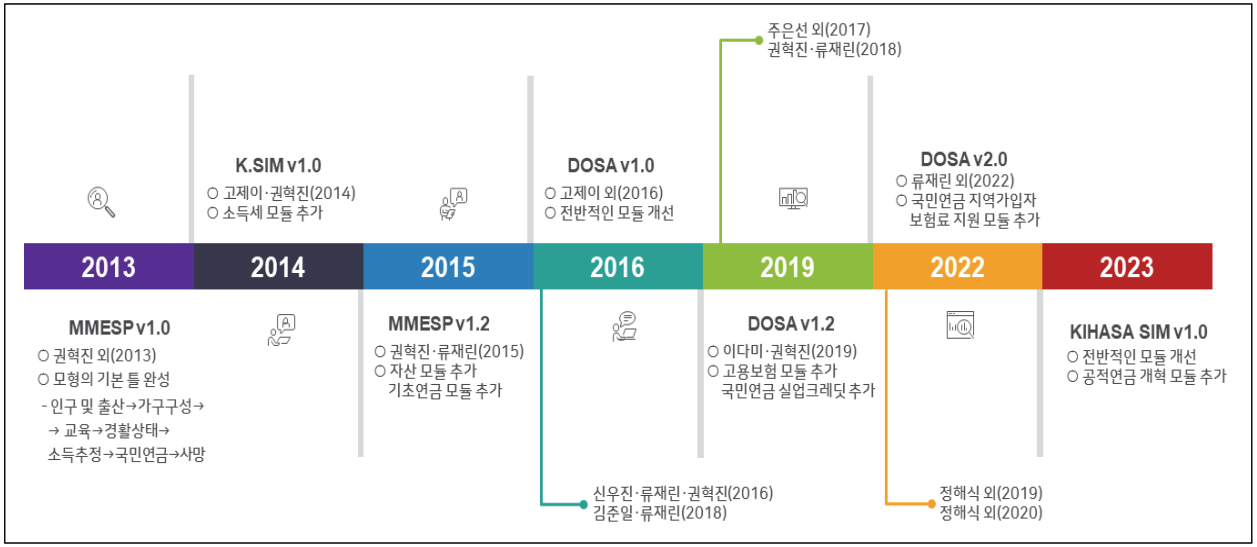
한편, DOSA가 개발된 이후 몇몇 연구들은 기존 모듈을 개선하거나 새로운 모듈을 추가하는 방식으로 이를 발전시켜 왔다(〔그림 3-1〕 참조). 예컨대, 이다미와 권혁진(2019)은 국민연금 실업크레딧의 효과를 분석하기 위해 고용보험 모듈을 추가하였고, 류재린 외(2022b)는 지역가입자에 대한 연금보험료 지원 제도의 장기 효과를 전망하기 위해 보험료 지원 모듈을 추가하였으며, 이용하 외(2022)는 국민-기초연금 간 연계감액 산식 변경 및 폐지 효과를 분석하기 위해 국민-기초연금의 연계 조정 모듈을

추가하였다.

그런데 주기적인 정리·통합 없이 여러 명의 연구자가 개별적으로 모형을 수정하면서, 동일한 버전 내 차이가 발생하고 모형이 복잡해지는 문제가 발생하였다. 모형의 버전이 동일함에도 불구하고, 모형에 사용되는 모듈의 종류, 실적치-집계치의 반영 기간, 시뮬레이션 방식 등의 차이가 발생하게 된 것이다. 예를 들어, 권혁진과 류재린(2018), 김준일과 류재린(2018)은 모두 DOSA v1.2를 사용하였다고 기술하였으나, 두 연구는 출산 시뮬레이션 방식, 실적치-집계치 반영 기간, 종사상지위 구분(고용주-자영자 구분) 등이 다르다. 이러한 버전별 차이는 DOSA가 발전하고 새로운 모듈이 추가되면서 더욱 커지게 되었다.

한편으로, 더 높은 버전의 모형에 이전 버전에서는 적용되었던 시뮬레이션 코드가 반영되지 않거나, 불필요한 로직이 시행되는 경우가 발생하기도 하였다. 물론, 시뮬레이션 목적에 따라 필요한 모듈을 개발하거나 자료를 업데이트하면서 발전되는 것은 일반적인 MSM의 발전 경로라 할 수 있다. 그러나 각 버전별 업데이트 결과를 주기적으로 통합하지 않고, 각 연구자가 필요에 따라 모형을 업데이트하면서, 모형의 복잡성이 커지고 효율성이 떨어지게 된 것은 분명히 개선이 필요한 지점이다. 기존 모형인 DOSA의 코드를 전반적으로 정리하는 동시에 각 모듈을 업데이트하여 새로운 모형으로 재구축할 필요가 있는 것이다. 이에 본 연구는 기존 모형인 DOSA의 코드를 전반적으로 정리 및 업데이트하여 새로운 모형으로 재구축하려 한다.

〔그림 3-1〕 KIHASA SIM의 발전 과정



자료 :필자 작성

제2절 모형의 구조

KIHASA SIM은 생애기간 동안 개인들이 경험하는 주요 생애사건들을 시뮬레이션한다. 이는 2015년 인구주택 총조사 자료의 2% 표본인 998,459명의 개인 자료를 기초자료로 활용한다. 모형 구축에 사용한 프로그램 언어는 STATA이다.[14] 동 모형은 기술적으로 횡단면적 순차과정을 포함한 동태적 전체집단 모형(population model with dynamic cross-sectional ageing)으로 분류된다(고제이 외, 2016, p.60).[15] 이는 모형 내 개인들의 사회·경제적 특성이 확률적으로 변하며 상호작용하는 동태적 모형이며, 전체 출생코호트를 대표하는 횡단면 자료에 기초하여 시뮬레이션을 수행하는 전체집단 모형이다. 이와 동시에 전체 시뮬레이션 대상자들이 연간 단위의 이산적인 시간 변화를 경험하는 횡단면적 순차과정을 지닌 모형이기도 하다.

이러한 유형의 모형은 기초자료에 여러 출생코호트를 포함하고 있어 각 코호트의 상대적 규모 변화를 고려할 수 있으며, 개인의 상호작용을 모형화하여 중요한 사회적 변화 또는 제도 변화를 세밀히 구현할 수 있다. 이에 여러 공적연금 모형들이 이러한 형태의 모형을 구축하여 노후소득 보장 효과, 생애 소비평탄화 효과, 공적 이전 효과 등을 분석하는데 활용하고 있다(Li et al., 2014, p.307).

한편, 횡단면적 순차모형의 특징 중 하나는 여러 개의 모듈로 구성된다는 점이다. MSM이 여러 생애사건들을 시뮬레이션하기 위해서는 서로 다른 연산 과정이 필요하며, 서로 다른 전이확률이나 행태방정식 추정 결과를 적용해야 하기 때문이다. KIHASA SIM 역시도 개인의 출생, 결혼, 이혼,

14) 즉, 이 모형은 STATA 프로그램으로 구현된다.
15) 이러한 분류는 Dekkers and Belloni(2009)의 기술적 구분을 따른 것이다.

재혼, 사별, 교육, 취업, 실업, 국민연금 가입, 연금 수급 등 다양한 생애 사건들을 시뮬레이션하는 모듈들이 연결된 형태로 구성되어 있다(〈표 3-1〉 참조).

각 모듈은 '출산- 결혼 및 이혼 - 교육 - 경제활동 - 소득 - 국민연금 가입 및 수급 - 자산 - 기초연금 - 고용보험 - 사망 및 유족연금' 순서로 배치되어 있다(〔그림 3-2〕 참조). 이는 개인이 경험하는 생애사건들의 순서를 고려하여 정한 것이다. 예를 들어, 교육시장 및 졸업을 시뮬레이션해야 졸업자의 노동시장 진입을 시뮬레이션할 수 있기에 경제활동 모듈을 교육 모듈 뒤에 배치하였다. 또한 기초연금의 소득인정액 산정 시 근로 및 사업소득, 공적연금 소득, 보유 자산에 대한 정보가 필요하기에 기초연금 모듈을 근로 및 사업소득, 국민연금 급여 수급, 자산 모듈 뒤에 배치하였다.

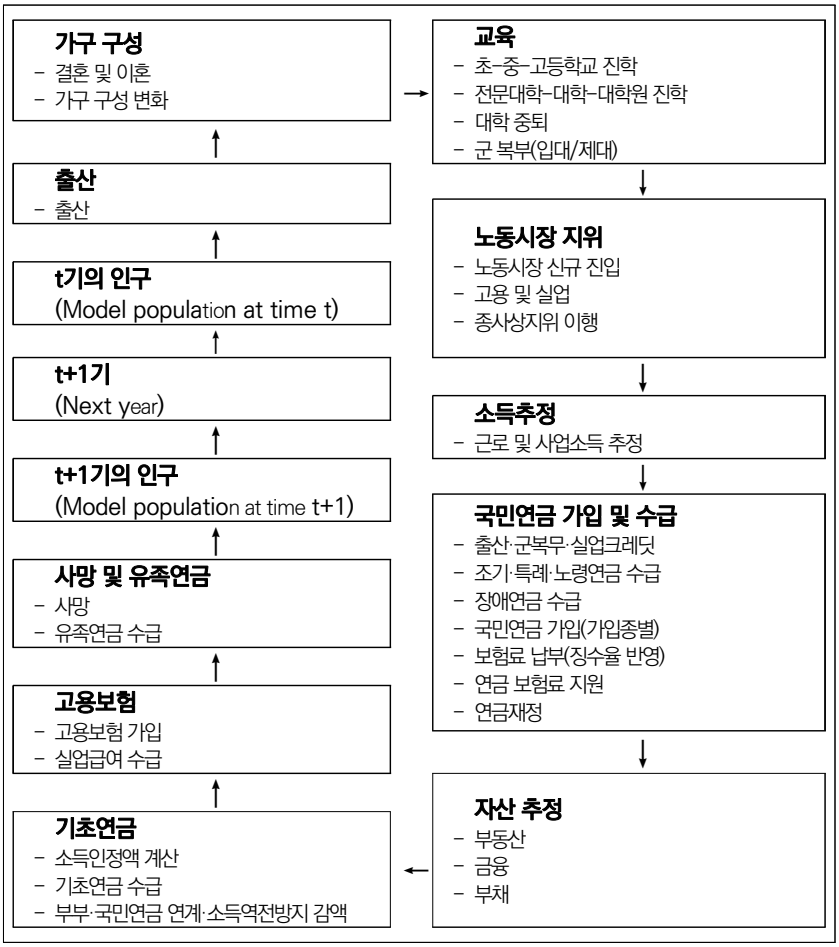
KIHASA SIM에 속한 개인들은 매년 각 모듈을 순차적으로 순환하여 여러 생애사건을 경험한다. t기에서 시작하여 모든 모듈을 순차적으로 경험한 이후에는 t기와 달라진 상태로 t+1기의 시뮬레이션을 수행한다. 이때, 한 개인은 t기에 동일한 생애 사건을 두 번 겪지 않는 것으로 가정한다. 예를 들어, t기에 대학 졸업도 하고 결혼도 하는 것은 가능하지만, 출산을 두 번 경험하거나 결혼을 두 번(즉, 결혼-이혼-결혼) 하지는 않는 것으로 가정한다.

〈표 3-1〉 KIHASA SIM의 시뮬레이션 대상 생애 사건

모듈	1	2	3	4	5
생애 사건	출생	결혼, 이혼, 재혼	교육 (진학 및 졸업)	경제활동 상태	근로(사업) 소득
모듈	6	7	8	9	10
생애 사건	국민연금 가입 및 수급 (노령, 장애)	자산 보유	기초연금 수급	고용보험 가입, 수급 및 실업크레딧	사망, 국민연금 (유족)

자료: 이다미, 권혁진. (2019). 국민연금 실업크레딧의 노후소득보장 효과 -미시 시뮬레이션(micro-simulation)을 활용한 장기 전망, p. 51. 〈표 5〉. 일부 수정.

〔그림 3-2〕 KIHASA SIM의 순환도

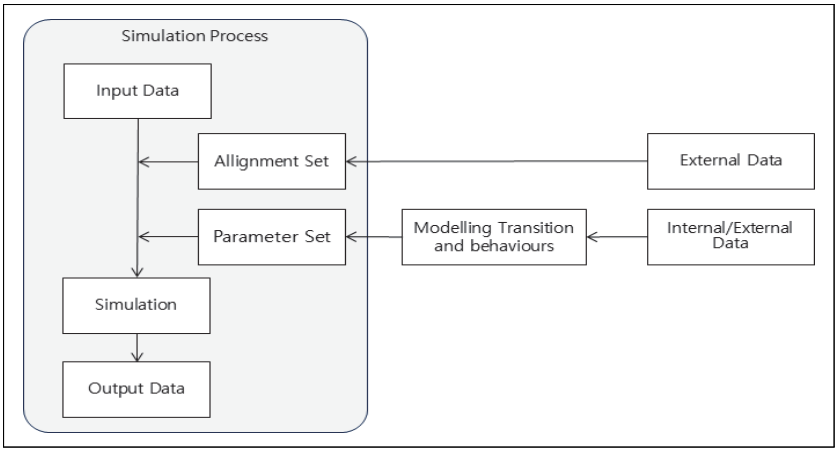


자료: 필자 작성

〔그림 3-3〕은 각 모듈에서 시행되는 기본적인 시뮬레이션 과정을 보여준다. 먼저, 모듈의 운용이 시작되면 앞선 모듈에서 생성된 자료(input data)를 불러온다. 여기에는 이전 기(t-1 기)의 시뮬레이션 결과가 반영되어 있으므로, 이번 기(t 기)의 시뮬레이션 수행을 위해 관련 변수를

초기화한다. 이후에는 시뮬레이션 수행에 필요한 거시적 집계 자료(alignment set) 또는 파라미터 자료(parameter set)를 결합한다. 여기서 거시적 집계 자료는 당해 년의 출생아 수, 결혼 및 이혼 건 수, 국민연금 가입률 등의 정보를 의미하며, 파라미터 자료는 미시적 개체들의 사회·경제적 상태 변화를 시뮬레이션하기 위해 모형 내/외부의 여러 자료를 이용하여 추정한 행태방정식 추정 결과이다. 예를 들어, 국민연금 가입확률 계산을 위한 계수 값들과 오차항이 이에 해당한다. 시뮬레이션에 필요한 자료들을 결합한 뒤에는 각 모듈별 고유의 시뮬레이션을 수행한 뒤, 다음 모듈 또는 다음 기(t+1)의 시뮬레이션을 위한 자료를 생성한다.

〔그림 3-3〕 KIHASA SIM의 모듈 내 시뮬레이션 과정

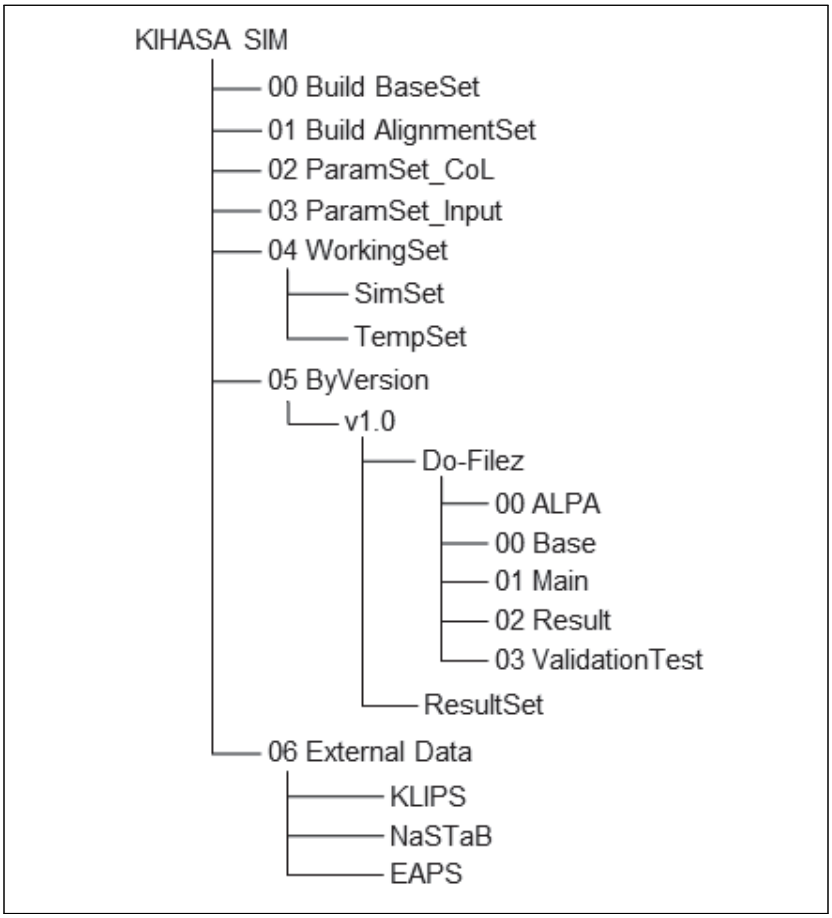


자료 : 고제이, 권혁진. (2014), 인구구조 변화에 따른 조세수입 전망과 사회복지 지출 부담 분석. 한국보건사회연구원 . p.55. [그림 14].를 토대로 필자 작성

KIHASA SIM은 STATA로 구축되어 있으므로, 각 모듈은 각각의 do 파일로 구성되어 있다. 시뮬레이션에 필요한 기초자료들과 do파일들은 여러 폴더에 정리되어 보관되며, 각 파일의 경로는 모두 프로그램화되어

시뮬레이션에 자동으로 반영된다. KIHASA SIM의 폴더 구조와 자료 배치에 관해서는 〔그림 3-4〕와 〈표 3-2〉를 참조하기 바란다.

〔그림 3-4〕 KIHASA SIM의 폴더 구조



자료: 필자 작성

〈표 3-2〉 KIHASA SIM의 폴더 배치와 기능

폴더명	기능
00 Build Base Set	BaseSet 구축에 필요한 기초자료 보관
01 Build Alignment Set	Alignment Set 구축에 필요한 기초자료 보관
02 ParamSet CoL	가공된 Alignment Set과 Parameter Set 보관
03 ParamSetInput	시뮬레이션에 과정에 직접 사용되는 Alignment Set과 Parameter Set 보관
04 Dofilez	버전 별 파일 보관
00 ALPA	Alignment Set과 Parameter Set의 가공을 위한 do 파일 보관
00 Base	Base Set의 가공을 위한 do 파일 보관
01 Main	메인 모듈의 do 파일 보관
02 Result	시뮬레이션 결과 추출을 위한 do 파일 보관
03 Validation Test	유효성 검토를 위한 do 파일 보관
05 Working Set	SimSet : 시뮬레이션 결과 저장 TempSet : 시뮬레이션 과정에서 발생하는 임시파일 저장
06 Result Set	시뮬레이션 결과로부터 추출된 여러 지표, 자료들 보관
10 External Data	기타 외부자료 보관

자료: 필자 작성

제3절 활용 자료

KIAHASA SIM은 생애주기 동안 발생하는 주요 생애사건들을 다루기 때문에 다양한 기초 자료를 이용한다. 〈표 3-3〉은 KIHASA SIM에서 시뮬레이션하는 주요 생애사건들과 기초자료를 보여준다. 각 자료들은 앞선 〔그림 3-3〕에서 살펴본 것처럼 시뮬레이션에 적합한 형태로 가공되어 투입된다. MSM의 시뮬레이션 결과가 높은 현실 설명력을 지니기 위해서는 기초자료들을 주기적으로 업데이트해야 하지만, 여기에는 많은 시간이 소요된다. 본 연구의 주요 목적 중 하나는 기존 모형에 사용되는 기초자료를 최신화하는 것이었으며, 〈표 3-4〉를 통해 대부분의 기초 자료들이 업데이트되었음을 확인할 수 있다. KIHASA SIM의 시뮬레이션에 사용하는 자료들에 대해서는 각 모듈의 시뮬레이션 과정과 함께 자세히 소개하며, 이하에서는 대부분의 모듈들에 공통적으로 적용되는 인구 및 거시경제 변수에 관해 자세히 살펴본다.

〈표 3-3〉 KIHASA SIM의 주요 생애사건과 활용 자료

구분	사건/변수	집계값 조정 셋	파라미터 셋
출산	출산	(출생아 수) 송창길 외(2023)	(출산확률) 한국노동패널조사
가구 구성	결혼	(초혼 및 재혼률) 인구동향조사	(결혼확률) 인구 · 주택총조사
	이혼	(이혼률) 인구동향조사	-
교육	상급학교 진학	(학교급별 진학률) 한국교육개발원 교육통계자료	(대학진학률) 한국노동패널조사
	대학 진학		
	군대 입대/제대	(군면제율) 병무청 자료	
노동 시장	경제활동상태 이행	(고용률) 기재부(2020)의 장기재정전망 (종사장 지위별 비율) 경제활동인구조사 (경제활동이행비율) 한국노동패널	(근로 및 사업소득) 한국노동패널조사
	근로소득		

구분	사건/변수	집계값 조정 셋	파라미터 셋
국민연금	국민연금 가입	(가입종별 가입률) 제5차 국민연금 재정계산, 경제활동인구조사 (조기수급률) 제4차 국민연금 재정계산 (과거 가입 및 이력) 노후보장패널 및 행정자료(2005~2009)	(국민연금 가입확률) 경제활동인구조사
자산	부동산, 금융, 부채	-	(자산) 재정패널조사
기초연금	신청/수급	-	-
고용보험	고용보험 가입	(연간 고용보험 가입기간, 가입률) 경제활동인구조사	(고용보험 가입확률) 경제활동인구조사
사망	사망	(사망률) 송창길 외(2023)	
거시경제변수	이자율, 임금상승률	2020년 기재부 장기재정전망	

자료: 필자 작성; 송창길, 나원희, 백주하, 류한별, 한신실, 고경표. (2023, 발간예정). 인구추계모형 구축과 활용 방안 . 한국보건사회연구원(발간예정); 한국노동연구원. (각년도). 한국노동패널조사. https://www.kli.re.kr/klips 에서 2023.5.10. 인출.; 통계청. (각년도). 인구동향조사. 원자료. https://mdis.kostat.go.kr/ofrData/selectOfrDataDetail.do?survId=38&itmDiv=1&nPage=3&itemId=2001&itemNm=%EC%9D%B8%EA%B5%AC 에서 2023. 5.10. 인출.; 통계청. (2015). 인구주택총조사. https://mdis.kostat.go.kr/ofrData/selectOfrDataDetail.do 에서 2023.3.10. 인출.; 한국교육개발원. (각년도). 교육통계분석자료집-고등교육통계편. 한국교육개발원. https://kess.kedi.re.kr/publ/list?itemCode=02&survSeq=&menuSeq=0&division=&word= 에서 2023. 5.10. 인출.; 한국교육개발원. (각년도). 교육통계연보. 한국교육개발원. https://kess.kedi.re.kr/publ/view?survSeq=2022&publSeq=2&menuSeq=0&itemCode=02&language=en 에서 2023. 5.10. 인출.; 병무청. (각년도). 병무통계. https://kosis.kr/statHtml/statHtml.do?orgId=144&tblId=TX_14401_A015&conn_path=I2 에서 2023.8.28. 인출.; 기획재정부. (2020. 9. 3.). 2020~2060년 장기재정전망. 기획재정부 보도자료 별첨자료. https://www.moef.go.kr/nw/nes/detailNesDtaView.do?searchBbsId1=MOSFBBS_000000000028&searchNttId1=MOSF_000000000045158&menuNo=4010100 에서 2023.1.10. 인출.; 통계청. (각년도). 경제활동인구조사 근로형태별 부가조사(8월). 원자료. https://mdis.kostat.go.kr/ofrData/selectOfrDataDetail.do?survId=47&itmDiv=1&nPage=3&itemId=2004&itemNm=%EB%85%B8%EB%8F%99 에서 2023. 5.10. 인출.; 국민연금재정계산위원회, 국민연금재정추계전문위원회, 국민연금기금운용발전전문위원회. (2023). 제5차 국민연금 재정계산 - 발표자료 -.; 국민연금재정추계위원회. (2018). 2018 국민연금 재정계산: 국민연금 장기재정추계.; 한국조세재정연구원. (각년도). 재정패널조사. 각년도. https://www.kipf.re.kr/panel/data/AllList/paData/view.do?cntNo=33 에서 2023. 5.10. 인출.

〈표 3-4〉 DOSA와 KIHASA SIM의 활용 자료 비교

모듈	사건	DOSA ver.1.1.	KIHASA SIM ver.1.0
출산	출산	제3차 재정계산	송창길 외(2023)
가구 구성	결혼	인구동향조사(2006-2015년) (통계청)	인구동향조사(~2022년) (통계청)
	이혼		
교육	상급학교 진학	한국교육개발원 교육통계자료(2006~2015년)	한국교육개발원 교육통계자료(2006~2021년)
	대학 진학		
	군대 입대/제대	병무청 자료 (2006~2013년)	병무청 자료 (2006~2020년)
노동 시장	경제활동상태 이행	경제활동인구조사 (2006~2015년) 2015년 장기재정전망 (기획재정부)	경제활동인구조사 (2006~2022년) 2020년 장기재정전망 (기획재정부)
	근로소득		
국민 연금	국민연금 가입	경제활동인구조사(~2015년) 제3차 재정계산	경제활동인구조사(~2022년) 제5차 재정계산
자산	부동산, 금융, 부채	한국재정패널(~2014년)	한국재정패널(~2021년)
기초 연금	신청/수급	-	-
고용 보험	고용보험 가입	-	경제활동인구조사(~2022년)
사망	사망	제3차 재정계산	송창길 외(2023)

자료: 1) DOSA ver.1.0. : 고제이 외. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 -노후 소득 보장 정책을 중심으로. p.70, 〈표 3-3〉.

2) KIHASA SIM ver.1.0 : 필자 작성; 송창길, 나원희, 백주하, 류한별, 한신실, 고경표. (2023, 발간예정). 인구추계모형 구축과 활용 방안 . 한국보건사회연구원(발간예정).; 통계청. (각년도). 인구동향조사. 원자료. https://mdis.kostat.go.kr/ofrData/selectOfrDataDetail.do?survId=38&itmDiv=1&nPage=3&itemId=2001&itemNm=%EC%9D%B8%EA%B5%AC 에서 2023. 5.10. 인출.; 한국교육개발원. (각년도). 교육통계분석자료집-고등교육통계편. 한국교육개발원. https://kess.kedi.re.kr/publ/list?itemCode=02&survSeq=&menuSeq=0&division=&word= 에서 2023. 5.10. 인출.; 한국교육개발원. (각년도). 교육통계연보. 한국교육개발원. https://kess.kedi.re.kr/publ/view?survSeq=2022&publSeq=2&menuSeq=0&itemCode=02&language=en 에서 2023. 5.10. 인출.; 병무청. (각년도). 병무통계. https://kosis.kr/statHtml/statHtml.do?orgId=144&tblId=TX_14401_A015&conn_path=I2 에서 2023.8.28. 인출.; 통계청. (각년도). 경제활동인구조사 근로형태별 부가조사(8월). 원자료. https://mdis.kostat.go.kr/ofrData/selectOfrDataDetail.do?survId=47&itmDiv=1&nPage=3&itemId=2004&itemNm=%EB%85%B8%EB%8F%99 에서 2023. 5.10. 인출.; 기획재정부. (2020. 9. 3.). 2020~2060년 장기재정전망. 기획재정부 보도자료 별첨자료. https://www.moef.go.kr/nw/nes/detailNesDtaView.do?searchBbsId1=MOSFBBS_000000000028&searchNttId1=MOSF_000000000045158&menuNo=4010100 에서 2023.1.10. 인출.; 국민연금재정계산위원회, 국민연금재정추계전문위원회, 국민연금기금운용발전전문위원회. (2023). 제5차 국민연금 재정계산 - 발표자료 -.; 한국조세재정연구원. (각년도). 재정패널조사. 각년도. https://www.kipf.re.kr/panel/data/AllList/paData/view.do?cntNo=33 에서 2023.5.10. 인출.

KIHASA SIM은 한국보건사회연구원이 장기재정추계를 위해 구축한 인구추계 모형인 송창길 외(2023)의 추계 결과(중위가정 적용)를 이용한다. 동 추계는 통계청의 인구추계 방법론을 구현한 뒤 추계기간을 2120년까지 확장한 것으로(송창길 외, 2023), 100세 이상자에 대한 추계 방법, 출산율 추정 방법 등에서 통계청의 방법론과 차이가 있으나, 전반적인 추계 결과는 통계청(2021)과 유사하다.[16]

다음의 〈표 3-5〉는 통계청(2021)과 송창길 외(2023)의 인구추계 결과를 보여준다. 양자 모두 통계청(2021)의 중위 가정을 적용한 결과이다. 이에 따르면, 천명 단위 비교에서는 양 추계 간 차이가 거의 없다. 단, 2070년의 사망자 수가 약 3천 명 정도의 차이를 보이는데, 이는 연앙인구 산출을 위한 생명표 적용 시 2개년의 평균치를 적용하기 때문으로 해석된다. 2070년과 2071년의 생명표를 통해 2070년의 사망자 수가 결정되는데, 통계청의 인구 가정이 2070년까지만 공개되어 있으므로 2071년의 사망률 가정이 제대로 적용되지 않아 이 같은 오차가 발생한 것이다(송창길 외, 2023, p.94). 다음으로 〈표 3-6〉은 LLG 모형을 적용하여 사망률을 추정한 뒤 추계기간을 2120년까지 연장한 송창길 외(2023, p.102)의 인구추계 결과와 통계청(2021)의 추계 결과를 비교하여 보여준다. 이에 따르면, 송창길 외(2023)와 통계청(2021)의 추계는 2120년 인구 기준 약 6천 명 정도 차이 난다. 이처럼 양자 간 차이가 거의 없으므로, 본 연구는 통계청의 중위가정을 적용한 송창길 외(2023)의 인구추계 결과를 이용한다.[17]

16) 본 연구는 통계청(2021)이 인구추계 결과를 제공하는 2070년 이후에 대해서도 시뮬레이션을 수행하므로, 송창길 외(2023)의 추계를 이용하는 것은 불가피한 선택이라 할 수 있다.

17) 이는 제5차 국민연금 재정계산의 인구 가정과 동일한 가정을 적용한 셈이다.

〈표 3-5〉 통계청(2021)과 송창길 외(2023)의 인구변동요인 결과 비교
(중위가정 적용, 2070년까지)

(단위: 천 명)

구분		통계청(2021) (A)			송창길 외(2023) (B)			차이(B-A)		
		2030	2050	2070	2030	2050	2070	2030	2050	2070
인구	계	51,199	47,359	37,656	51,199	47,358	37,656	0	-1	0
	남	25,534	23,495	18,788	25,534	23,494	18,787	0	-1	-1
	여	25,665	23,864	18,868	25,665	23,864	18,869	0	0	1
출생	계	305	236	196	305	236	196	0	0	0
	남	156	121	101	156	121	101	0	0	0
	여	149	115	96	149	115	96	0	0	0
사망	계	408	680	702	408	680	705	0	0	3
	남	215	345	345	215	345	347	0	0	2
	여	192	335	357	192	335	358	0	0	1
이동	계	46	43	40	46	43	40	0	0	0
	남	25	23	20	25	23	20	0	0	0
	여	21	20	20	21	20	20	0	0	0

자료 : 송창길 외, (2023, 발간예정). 인구추계모형 구축과 활용 방안. 한국보건사회연구원. p. 95 〈표 4-1〉.

〈표 3-6〉 통계청(2021)과 송창길 외(2023)의 인구변동요인 결과 비교
(LLG 적용: 중위, 2120년까지)

(단위: 천 명)

구분		통계청(2021) (A)			송창길 외(2023) (B)			차이(B-A)		
		2030	2070	2120	2030	2070	2120	2030	2070	2120
인구	계	51,199	37,656	20,953	51,196	37,650	20,946	-3	-6	-8
	남	25,534	18,788	10,571	25,532	18,784	10,562	-2	-4	-9
	여	25,665	18,868	10,383	25,664	18,866	10,384	-1	-2	2
출생	계	305	196	111	305	196	111	0	0	0
	남	156	101	57	156	101	57	0	0	0
	여	149	96	54	149	96	54	0	0	0
사망	계	408	702	300	407	702	300	-1	0	0
	남	215	345	151	215	345	151	0	0	0
	여	192	357	149	192	357	149	0	0	0
이동	계	46	40	36	46	40	36	0	0	0
	남	25	20	18	25	20	18	0	0	0
	여	21	20	19	21	20	19	0	0	0

자료 : 송창길 외, (2023, 발간예정). 인구추계모형 구축과 활용 방안. 한국보건사회연구원. p. 102. 〈표 4-5〉.

한편, 본 연구는 기재부(2020)의 거시경제전망치(중위 가정)를 적용한다(〈표 3-7〉 참조). 이는 실질 경제성장률이 2023~2030년 연평균 2.2%에서 2031~2040년 1.3%, 2041~2050년 0.8%로 둔화된 뒤, 이후의 기간에는 0.3~0.4%의 완만한 성장률을 유지할 것이라는 전망이다. 실질 임금 상승률은 2023년 2030년 기간 중 연평균 2.1%에서 점진적으로 하락하여 장기적으로 1.6% 수준에 이를 것으로 전망되며, 실질금리는 2031~2050년 연평균 1.4% 수준을 유지하다가 점진적으로 하락하여 장기적으로 0.9% 수준에 이를 것으로 전망된다. 경제활동 참가율은 2023~2030년 63.2%에서 점진적으로 하락하여 2071~2080년 50.6% 수준에 이른 뒤 52%까지 상승할 것으로 전망된다.

〈표 3-7〉 경제변수 가정(기재부(2020)의 중위가정)

(단위: %)

	2023~2030	2031~2040	2041~2050	2051~2060	2061~2070	2071~2080	2081~2090	2091~2100
실질경제성장률	2.2	1.3	0.8	0.5	0.3	0.3	0.4	0.4
실질임금상승률	2.1	2.1	2.1	2.0	1.9	1.7	1.7	1.6
실질금리	1.2	1.4	1.4	1.3	1.2	1.1	0.9	0.9
물가상승률	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
경제활동참가율	63.2	60.6	56.5	53.5	51.6	50.6	50.8	52.0

주: 제시된 경제변수 가정은 해당 기간의 단순평균임.
자료: 기획재정부. (2020. 9. 3.). 2020~2060년 장기재정전망. 기획재정부 보도자료 별첨자료.

제4장

기초 자료 구축

제1절 기초 자료 구축
제2절 인구 및 가구 보정
제3절 소득 분포 보정
제4절 국민연금 가입 및 수급 정보 보정

제4장 기초 자료 구축

제1절 기초 자료 구축

한 사회의 전체 인구집단을 대상으로 시뮬레이션을 수행하는 전체집단 모형은 사회 구성원 전체를 대표하는 동시에 다양한 사회·경제적 정보를 담고 있는 행정자료, 인구센서스, 설문조사 자료 등의 횡단면 자료를 기초자료로 활용하는 것이 일반적이다. 이때, 행정자료나 인구센서스의 경우 모집단의 특성을 온전히 반영한다는 점에서 대표성이 높지만, 미시적 개체들의 행태를 모형화하기 위한 세밀한 정보가 부족한 경우가 많다. 이에 다수의 MSM이 모형의 목적에 적합한 설문조사 자료를 기초자료로 활용하거나, 행정자료나 인구센서스에 설문조사 자료를 결합하여 기초자료로 이용하고 있다(Li et al., 2014, p.311).[18)]

본 연구에서 개발하는 KIHASA SIM은 998,459명의 개인이 포함된 2015년 인구주택총조사 자료의 2% 표본에 여러 설문조사 자료를 결합하여 기초 자료로 활용한다. 동 모형의 전신인 DOSA는 2005년 시행된 인구주택총조사 자료의 1% 표본을 사용하였으나, 그간 우리 사회가 빠르게 변화하였음을 고려하여 비교적 최신 자료로 기초자료를 교체한다. 2015년에 시행된 인구주택총조사가 2005년의 조사와 달리 새로운 조사 틀에 기초한다는 점, 2% 표본 자료를 공개하고 있어 기초 자료에 포함된 개체 수가

18) Li et al.(2014, p.311)에 따르면 Dynamic MSM의 43.8%가 설문조사 자료를 기초자료로 활용하고 있다. 그 외의 모형들은 인구센서스 자료(23.4%), 행정 자료(17.2%), 행정 자료에 센서스 자료를 결합한 자료(synthetic data)(15.6%) 등을 기초 자료로 활용하고 있다.

많아지게 된다는 점 역시 기초 자료의 교체를 결정한 이유이다. 이때, 가장 최근에 조사된 2020년 인구주택총조사 자료가 아니라 2015년의 자료를 활용한 것은 첫째, 2015년~2022년 사이의 시뮬레이션 결과를 해당 기간의 실적치와 비교하여 모형의 유효성을 검토하고, 둘째, 2020년 인구주택총조사의 조사 기간 동안 확산되었던 COVID19의 영향을 피하기 위함이다.

기초 자료의 구축 과정은 크게 둘로 나눌 수 있다. 첫 번째는 인구주택총조사 자료에서 제공하는 변수들을 시뮬레이션에 적합한 형태로 가공 및 보정하는 과정이다. 데이터의 누락, 입력 오류 등에 대한 보정 및 데이터 클리닝 과정이 여기에 해당한다. 이 과정은 개인 및 가구의 고유 ID, 성별, 만 나이, 출생연도, 가구원 수, 배우자 유무, 혼인 상태, 이혼·재혼·사별 여부, 가구주와의 관계, 경제활동 상태, 종사상지위, 재학 중인 학교, 최종학력, 졸업/재학 여부, 교육년수, 주택 점유 형태, 거처 종류, 기타 주택 소유 여부, 사망 등의 변수를 주로 다룬다. 둘째, 전술한 것처럼 인구주택총조사 자료에는 다양한 생애사건들을 시뮬레이션하기에 적합한 정보들이 부재하므로, 경제활동인구조사, 한국노동패널, 국민연금 가입자 이력 자료 등의 다양한 자료들을 이용하여 부족한 정보를 보완하는 과정이다. 근로·사업소득, 국민연금 가입 상태, 국민연금 가입 및 수급이력 등의 변수들이 이에 포함된다(〈표 4-1〉 참조).

구체적인 변수 가공 및 처리 내용은 다음과 같다. 첫째, 가구 및 개인의 고유 ID를 정리한다. 가구 ID의 경우 원자료를 그대로 사용하지만, 개인 ID는 시뮬레이션 과정에서 개인의 사망 후 상태를 식별할 수 있도록 새롭게 생성한다. 둘째, 성별과 만 나이의 경우 원자료의 정보를 그대로 사용하며, 만 나이를 이용해 출생연도를 생성한다.

〈표 4-1〉 주요 변수의 가공 및 처리

변수	자료	가공/처리내용
가구 관련		
가구ID	2015 인구주택총조사	- 원자료의 번호를 그대로 사용
개인ID		- 연도와 초기 데이터셋의 일련번호(7자리) - 이후 출생아는 출생년도+일련번호 20150000000+n
출생연도		- 만 나이를 고려하여 생성
배우자 유무		- 혼인은 '배우자 있음'으로, 미혼/사별/이혼은 '배우자 없음'으로 처리 - 14세 이하의 결측값은 '배우자 없음'으로 처리 배우자 있음 =1 / 배우자 없음 =0
혼인상태		- '미성년자/기혼/미혼/사별/이혼'으로 구분
이혼여부		- 원자료의 혼인상태 변수를 이용하여 생성 이혼경험 있음=1 / 이혼경험 없음=0
재혼여부		- 원자료에서는 재혼 여부를 판단할 수 없으나, 시뮬레이션 과정에서 재혼자가 발생 재혼=1 / 재혼 아님=0
사별여부		- 원자료 혼인상태변수를 이용해 생성 사별=1 / 사별하지 않음=0
교육 관련		
최종학력	2015 인구주택총조사	- 만 7세 이하 아동은 미취학으로 처리 미취학 아동=0 / 초등학교=1 / 중학교=2 / 고등학교=3 / 4년제 미만 대학=4 / 4년제 이상 대학=5 / 석사= 6 / 박사 = 7 / 무학=9
졸업/재학 여부		- 원자료 상의 졸업과 수료를 '졸업'으로, 재학/휴학/ 중퇴는 '재학으로 처리 미취학 아동 및 결측치=0 / 졸업=1 / 재학=2
학년		- 학교(최종학력)와 연령을 고려하여 설정
교육년수		- 최종학력에 따라 계산
경제활동 관련		
경제활동 상태 및 종사상 지위	2015 인구주택총조사	- 경제활동 상태 및 종사상지위와 관련된 이상치는 연령, 재학여부 등을 고려하여 실업 또는 비경활로 처리 - 가구 정리 과정에서 새롭게 생성된 개인의 경제활동상태는 경제활동인구조사의 수치를 이용하여 할당 상용근로자=11 / 임시·일용근로자=12 / 자영자= 13 / 고용주 =14 / 무급가족종사자= 15 / 실업자= 20 / 15세 이상 비경제활동인구=30 / 15세 미만 인구=50
소득 관련		
근로 및 사업소득	한국노동패널	- 근로 및 사업소득 추정
연금가입 관련		
국민연금 가입	경제활동 인구조사	- 경제활동인구조사 자료를 통해 개인의 국민연금 가입확률을 추정
국민연금 가입이력	국민노후 보장패널 및 국민연금DB	- 성, 연령, 종사상 지위별로 2016년 이전의 국민연금 가입 이력을 매칭함.

자료: 필자 작성

셋째, 원자료의 혼인상태 변수를 가공하여 배우자 정보를 정리한다. 혼인상태는 미혼, 기혼, 사별, 이혼, 미성년자로 구분하여 정리하는데, 미혼과 사별, 이혼은 '배우자 없음'으로, 기혼은 '배우자 있음'으로 처리한다. 이때, 14세 이하인 미성년자는 모두 결측값으로 처리한다. 이혼 여부는 원자료에서 혼인 상태 변수가 이혼인 경우 '이혼 경험 있음'으로 처리하며, 기타 상태는 '이혼 경험 없음'으로 처리한다. 사별 여부의 경우 원자료에서 혼인 상태 변수가 사별인 경우 '사별함'으로 처리하며, 다른 상태는 '사별하지 않음'으로 처리한다. 재혼 여부 변수는 2015년 원자료에서는 재혼 여부 상태를 판별하는 것이 불가능하지만, 이후 가구 구성을 시뮬레이션하는 과정에서 판별이 가능해진다. 이에 기초 자료에서는 해당 변수를 생성한 뒤, 모두 결측치로 처리한다.

넷째, 교육 관련 변수를 처리한다. 먼저, 최종 학력은 미취학 아동, 초등학교, 중학교, 고등학교, 4년제 미만 대학, 4년제 이상 대학, 석사과정, 박사과정, 받은 적 없음(無學)으로 처리한다. 이때, 만 7세 미만의 경우 학령인구에 도달하지 못하여 교육 경험이 없는 것으로 간주하여 미취학으로 분류한다. 졸업 및 재학 여부는 원자료 상에서의 졸업과 수료를 '졸업'으로 처리하며 재학, 휴학, 중퇴는 '재학'으로 처리한다. 학년의 경우 학교 변수와 연령 변수를 함께 고려하여 설정한다. 교육년수는 최종 학력과 출생연도를 고려하여 계산한다. 군대 복무의 경우 '군입대 전', '군입대 1년차', '군입대 2년차', '군 제대', '군 면제'로 구분한다. 입대 상태는 연령을 기준으로 만 19세가 되면 자동으로 입대하여 군 복무를 수행하는 방식이다.

다섯째, 경제활동 변수는 원자료 상의 경제활동상태와 종사상지위를 가공하여 '상용, 임시/일용, 고용주, 자영자, 무급가족종사자, 실업, 비경제활동, 15세 미만 인구로 구분한다. 보다 구체적으로는 경제활동상태가 '①주로 일하였음, ②가사, 학업(학교, 학원) 등 다른 활동을 하면서 틈틈이

일하였음, ③일자리를 갖고 있으나, 휴가 등으로 잠시 쉬고 있음.'이고, 종사상 지위 정보가 있는 경우 취업자로 간주한다. 또한 경제활동상태가 '일하지 않았음'이고, 종사상지위 정보가 없는 경우 실업 및 비경제활동으로 간주한다. 이때, 취업 상태라고 응답했으나 종사상 지위가 결측값인 경우는 최종적으로 실업 및 비경제활동으로 구분하며, 이상치는 연령과 재학 여부를 고려하여 실업 및 비경제활동으로 처리한다.

한편, 인구주택총조사에서는 상용-임시/일용과 실업-비경활을 구분하지 않고 있기에 경제활동인구조사 자료를 이용하여 각각을 분리한다. 먼저 상용-임시/일용은 경제활동인구조사 8월 부가조사에서 산출한 임금근로자 대비 상용 근로자 비율을 이용하여, 상용 확률이 높은 순으로 상용근로자를 결정한다. 이때, 상용근로 확률은 2011~2019년 한국노동패널 자료를 풀링(pooling)한 뒤 로짓모형을 이용해 추정한다. 추정 결과는 다음의 〈표 4-2〉와 같다. 실업과 비경제활동인구의 구분은 상용-임시/일용과 동일한 자료를 이용해 성-연령-학력수준별 비경제활동인구 및 실업 비율을 산출한 뒤 무작위 추출로 결정한다.

〈표 4-2〉 상용근로 확률 추정 결과

구분	남성		여성	
	모형1		모형2	
만나이	-0.102***	(0.009)	-0.019**	(0.010)
만나이^2	0.001***	(0.000)	0.000	(0.000)
배우자유무	0.362***	(0.044)	0.236***	(0.043)
교육년수	0.158***	(0.006)	0.021***	(0.007)
5세이하 자녀 유무	0.435***	(0.050)	0.264***	(0.055)
log(소득)	2.216***	(0.042)	2.497***	(0.043)
상수항	-10.087***	(0.253)	-11.007***	(0.272)
Log likelihood	-12,354		-10,909	
Pseudo R2	0.301		0.299	
표본개수	33,375		24,065	

주: 1) 종속 변수는 상용근로 여부임.
2) 소득은 월 평균 소득(만원)을 로그 불변연간소득으로 환산한 것임.
3) 괄호( ) 안은 표준 오차이며, ***, **, *는 각각 1%, 5%, 10%에서 통계적으로 유의함을 나타냄.
4) 횡단면가중치 부여
자료: 필자 작성

제2절 인구 및 가구 보정

KIHASA SIM은 인구주택총조사 자료의 '가구주와의 관계' 코드를 이용해 부부와 자녀를 중심으로 가구 구성을 정리하고, 새로운 가족 관계 코드를 생성하는 등 가구 구성을 시뮬레이션에 적합한 형태로 가공하여 사용한다. 현실의 복잡한 세대 및 가구 변화를 시뮬레이션하는 것이 어렵기에 시뮬레이션 결과가 왜곡되지 않는 범위 내에서 가구 구성을 단순화하는 것이다. 이는 모형 내 복잡성을 낮추고 시뮬레이션 과정에서 발생할 수 있는 오류를 최소화하여 시뮬레이션 속도와 안정성을 높일 수 있는 방법이기도 하다.[19]

주요 보정 내용은 다음과 같다. 첫째, 소득을 기준으로 가구주(가족 대표)를 변경한다. 인구주택총조사 원자료에는 가구주와의 관계가 경제활동과 관계없이 남성=본인, 여성=배우자로 설정되어 있으나, 본 연구는 양자의 소득을 비교하여 소득이 큰 자를 가구주(가족대표)로 설정한다. 둘째, 가구 내 복수의 가족이 존재하는 경우 고유의 가구 ID를 생성한 후 모두 분가하는 것으로 가정한다. 다만, 원가구와 분가 가구를 추적할 수 있도록 원자료의 가구 ID를 보존하고, 시조(始祖) 및 족보를 생성한다. 예컨대, 가구주와 형제, 자매들의 가족, 그리고 가구주의 부모가 한 가구에 존재하는 경우 가구주의 부모를 시초로 설정하고, 형제·자매들은 모두 다른 가구로 분가시킨다.

한편, 일반에게 공개되는 인구주택총조사의 표본자료는 20대 인구가 과소 추계되어 있고, 85세 이상 인구가 각 세별로 구분되지 않아 성·연령별

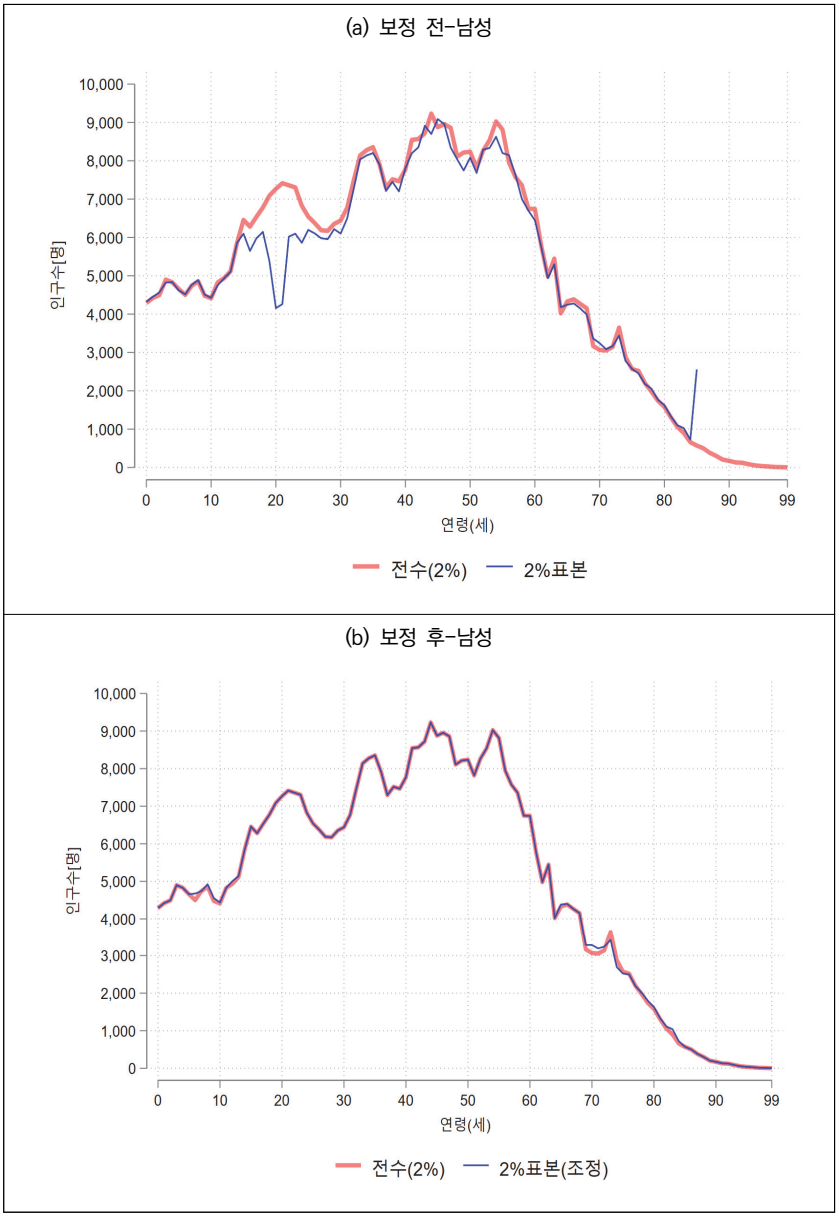
19) 물론, 이처럼 가구 구성을 단순화할 수 있는 것은 본 연구에서 구축하는 모형이 개인을 중심으로 하는 국민연금 제도와 부부를 중심으로 하는 기초연금 제도에 초점을 맞추고 있기 때문이다. 추후 기초보장제도 등 가구를 대상으로 하는 제도를 반영하기 위해서는 가구 구성의 시뮬레이션 과정을 정교화할 필요가 있다. 이는 추후의 연구과제이기도 하다.

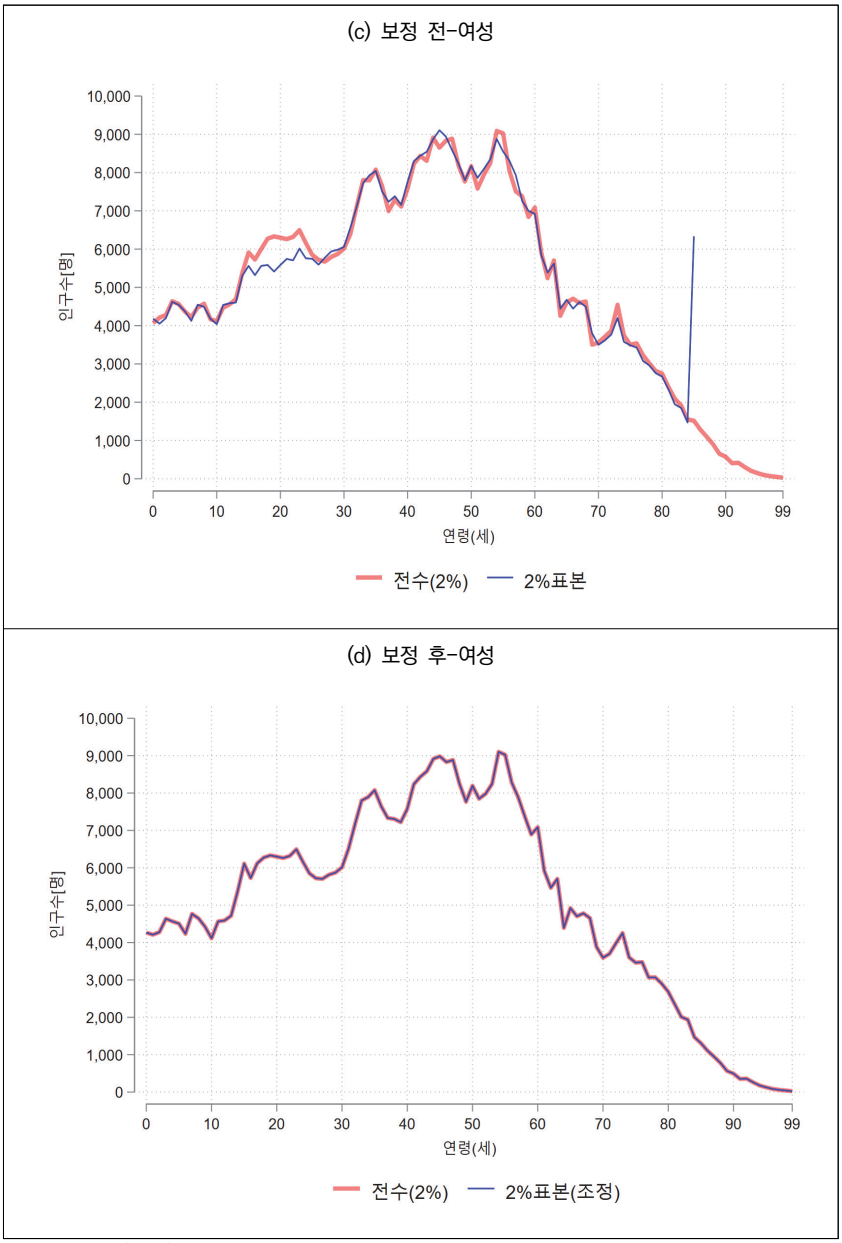
인구 수가 전수집계 결과와 다르다.[20] 이에 본 연구는 인구주택총조사의 성-각 세별 분포를 고려하여 가상 인구를 생성하여 두 자료 간 차이를 보정한다.

먼저, 20대의 경우 개인 ID를 생성한 뒤, 결혼 기간과 비동거 자녀 수 등의 정보를 이용해 생성된 개인을 기존 가구에 할당하는 방식으로 보정한다. 기본적으로 비동거 자녀가 있다고 응답한 가구를 추출하여 자녀를 할당하는데, 이때, 엄마와 자녀의 최대 연령 차이를 35세로 한정하며, 할당되는 자녀의 연령이 부모의 결혼 기간을 넘지 않는 것으로 한정한다. 85세 이상 인구는 인구주택총조사에 기초하여 인구 수와 조정 비율을 계산한 뒤, 각 세별로 새로운 인구를 생성한다. 전술한 과정을 통해 보정된 성-연령별 인구 분포는 〔그림 4-1〕과 같다.

20) 2005년 인구주택총조사 표본의 경우에도 유사한 문제가 존재하였다. 이에 대해서는 고제이 외(2016, p.68)을 참조하기 바란다.

〔그림 4-1〕 인구 보정 전후의 성별 인구 수 비교





자료: 통계청. (2015). 인구주택총조사. https://mdis.kostat.go.kr/ofrData/selectOfrDataDetail.do 에서 2023.3.10. 인출.

다음으로 생성된 인구의 사회·경제적 변수들을 설정하는 과정은 다음과 같다. 먼저, 교육 수준은 18세까지는 기존의 원칙(연령별 교육 수준)대로 지정한다. 19세부터는 통계청의 성-연령-교육 수준별 재학 인원을 이용하여 배정 인원을 산출한 뒤, 무작위 추출로 대상자를 선별한다. 다음으로 경제활동상태는 경제활동인구조사 자료를 이용하여 성-연령-교육 수준-종사상 지위별 인원을 산출한 뒤, 무작위 추출로 할당한다. 이때, 할당 순서는 '성-연령-학교-재학'→'성-연령-학교'→'성-연령' 순으로 반복적으로 이루어진다. 경제활동인구의 산업 및 직종 역시 이와 유사한 방법으로 '성-연령-교육 수준-종사상 지위-산업-직종'별 인원을 추출하여 할당한다. 한편, 19~21세 남자의 군복무는 학력, 연령, 경제활동상태를 고려하여 처리한다. 고졸자와 비 경제활동자는 19세에 입대하여 20세까지 복무하는 것으로 처리하며, 대학생은 20세 입대하여 21세까지 복무하는 것으로 처리한다. 가구 유형은 성별, 연령, 세대 유형, 가구주와의 관계, 혼인 관련 변수들을 이용하여 구분한다. 이상의 과정을 통해 생성된 기초자료의 사회경제적 특성은 〈표4-3〉과 같다.

〈표 4-3〉 기초자료와 경제활동인구조사의 사회경제적 특성

(단위: 명, (%), 세, %)

	KIHASA SIM			경제활동인구조사		
구분	인원 (비율)	평균연령	남성비율	인원 (비율)	평균연령	남성비율
상용	249,062 (29.4)	41.0	62.6	12,797,638 (29.6)	41.0	62.4
임시일용	134,088 (15.9)	44.8	45.8	6,676,663 (15.4)	45.1	45.3
고용주	31,806 (3.8)	48.1	75.6	1,623,516 (3.8)	48.0	75.8
자영자	104,164 (12.3)	53.3	56.8	5,270,837 (12.2)	53.2	56.8
실업	18,852 (2.2)	37.8	60.6	924,753 (2.1)	38.6	60.0
비경활	308,234 (36.4)	47.3	34.7	16,002,024 (37.0)	46.3	34.1
계	846,206 (100.0)	45.6	49.5	43,295,431 (100.0)	45.3	49.1

주; 1) 무급가족종사자는 자영자에 포함됨.
자료: 1) 통계청. (2015). 경제활동인구조사 근로형태별 부가조사(8월)
2) 필자 작성

제3절 소득 분포 보정

기초 자료의 소득추정 과정은 2015년의 한국노동패널조사 자료를 활용하여 인구주택총조사 자료에서 제공되지 않는 근로·사업소득을 보정하는 과정이다. 이는 한국노동패널조사에서 GLS 모형을 통해 소득을 추정한 뒤, 파라미터를 이용해 기초 자료의 소득 자료를 생성하는 것을 골자로 한다. 이때, 개인의 소득 수준은 임금/비임금 근로 여부에 따라 크게 달라지므로, 본 연구는 임금 근로와 비임금 근로를 구분하여 소득을 추정한다. 기초 자료의 소득추정 결과는 〈표 4-4〉와 같다.

〈표 4-4〉 기초자료의 소득 생성을 위한 소득방정식 추정 결과

구분	임금근로자	비임금근로자
	모형1	모형2
만나이	0.094***	0.11***
만나이^2	-0.001***	-0.001***
성별	0.281***	0.422***
가구주유무	0.131***	0.177***
배우자취업유무	-0.185***	
배우자유무	0.229***	0.295***
장애(육체적제약)여부	-0.368***	-0.325*
학력(고졸)	0.153***	0.262***
학력(전문대졸)	0.276***	0.288***
학력(대졸)	0.357***	0.393***
학력(석사)	0.408***	0.285**
학력(박사)	0.617***	0.475***
광공업	0.303***	0.637***
사회간접자본및서비스업	0.115***	0.262***
전문관리직		0.214***
사무직		0.602***
서비스판매직	-0.24***	0.317***
가구원 수		0.043**

구분	임금근로자	비임금근로자
	모형1	모형2
생산직	-0.26 ***	
상수항	14.473***	13.212***
잔차항 평균	0.041	-0.093
잔차항 표준편차	0.553	0.581
R^2	0.470	0.429
표본개수	5018	1731

주: 1) 종속 변수는 임금근로자 또는 비임금근로자의 월평균소득(만원)을 로그불변연간소득으로 환산한 것임.
2) 소득은 2015년 불변가격 기준임(2015년 기준 전국소비자물가지수로 환산)
3) ***, **, *는 각각 1%, 5%, 10%에서 통계적으로 유의함을 나타냄.
4) 횡단면가중치를 적용함
자료: 필자 작성

다음으로, 〈표 4-5〉와 〔그림 4-2〕는 기초 자료와 한국노동패널의 성별 평균 근로(사업)소득을 비교한 결과이다. 전체 및 성별에 따른 평균 근로(사업)소득 모두 노동패널 소득 대비 시뮬레이션의 추정치 비율이 105% 내외로, 기초 자료의 소득 수준이 다소 높지만, 대체로 유사한 수준이라 할 수 있다. 평균과 표준편차, 그리고 최대값 모두 시뮬레이션의 추정치가 약간 높은 수준이다.

〈표 4-6〉은 성-종사상 지위별 평균 근로(사업)소득 수준을 비교한 결과이다. 이에 따르면, 양자가 전반적으로 유사한 수준이다. 다만, 임금근로자의 근로(사업)소득은 한국노동패널자료에 비해 조금 높게 추정되었으며, 비임금근로자의 소득은 한국노동패널자료에 비해 낮게 추정되었다. 반면, 표준편차의 경우 임금근로자는 시뮬레이션의 추정치가 더 크고, 비임금근로자는 한국노동패널에서 산출한 값이 더 큰 것으로 나타났다.

〈표 4-5〉 기초자료의 성별 평균 근로(사업)소득 비교

(단위 : 명, 만원, %)

구분	연도	2015년		
	성별	남	여	전체
KIHASA SIM	인원	300,628	218,538	519,166
	평균(A)	3,647.4	2,213.9	3,043.9
	표준편차	2,725.4	1,635.6	2,434.8
	최소값	13.0	11.2	11.2
	최대값	63,105	32,926	63,105
노동패널	인원	4,719	3,251	7,970
	평균(B)	3,523.4	2,068.1	2,968.5
	표준편차	2,462.2	1,436.6	2,244.3
	최소값	0.0	0.0	0.0
	최대값	36,000	14,400	36,000
노동패널 대비 비율 (A/B)		103.5	107.0	102.5

자료: 1) 한국노동연구원. (2015). 한국노동패널조사. https://www.kli.re.kr/klips에서 2023.5. 10. 인출.
2) 필자 작성.

〈표 4-6〉 기초자료와 노동패널의 종사상 지위별 평균 근로(사업)소득 비교

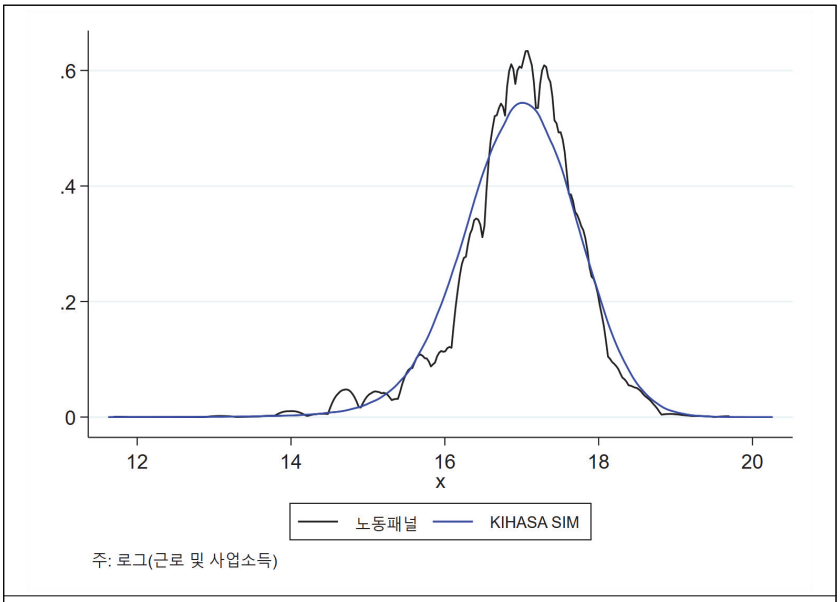
(단위 : 만원)

구분		임금			비임금		
		KIHASA SIM (A)	노동 패널 (B)	노동패널 대비 비율(A/B)	KIHASA SIM	노동 패널	노동패널 대비 비율(A/B)
전체	평균소득	3,044.4	2,867.7	106.2	3,042.5	3,263.6	93.2
	표준편차	2,394.0	1,826.9	·	2,546.3	3,146.4	·
남성	평균소득	3,670.0	3,460.4	106.1	3,588.3	3,677.9	97.6
	표준편차	2,690.8	1,940.2	·	2,812.9	3,416.9	·
여성	평균소득	2,224.6	2,015.7	110.4	2,180.1	2,280.3	95.6
	표준편차	1,602.4	1,222.5	·	1,735.5	2,079.0	·

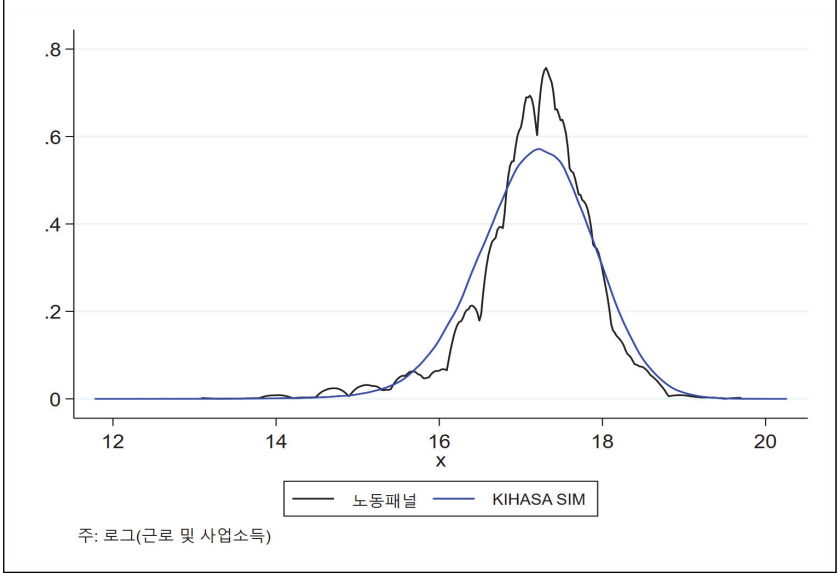
자료: 1) 한국노동연구원. (2015). 한국노동패널조사. https://www.kli.re.kr/klips 에서 2023.5. 10. 인출.
2) 필자 작성.

〔그림 4-2〕 2015년 전체 및 종사상지위별 근로소득 분포 비교

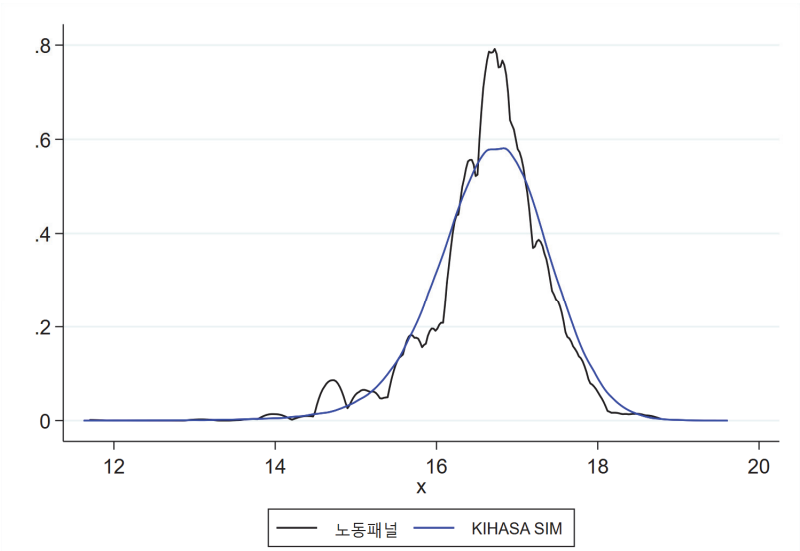
(a) 전체



(b) 남성

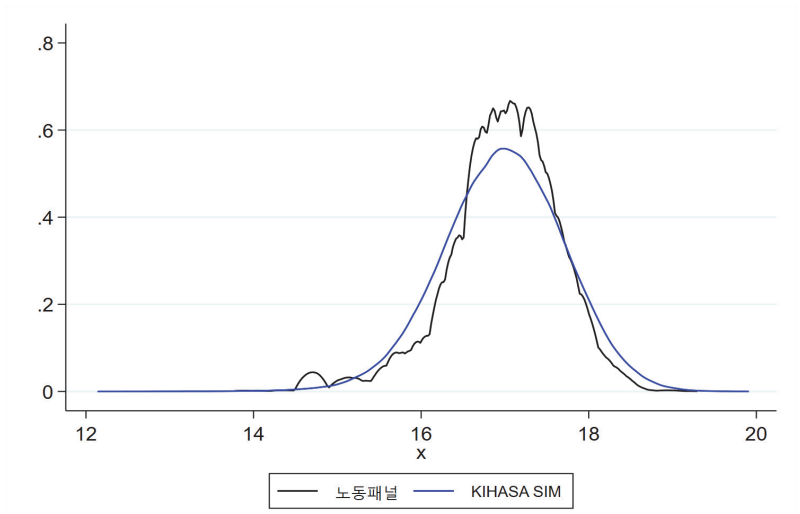


(c) 여성

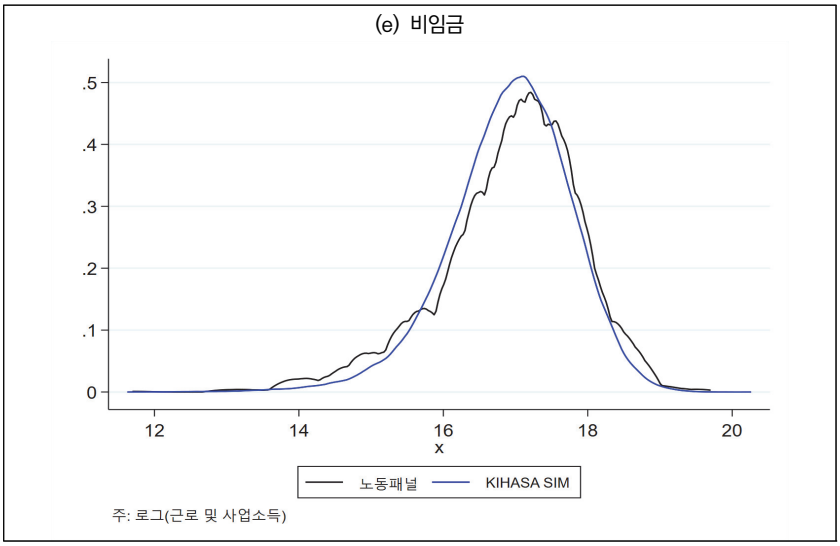


주: 로그(근로 및 사업소득)

(d) 임금



주: 로그(근로 및 사업소득)



자료: 필자 작성

제4절 국민연금 가입 및 수급 정보 보정

국민연금 가입에 대한 보정 과정은 성-연령-종사상 지위별 국민연금 가입자 규모를 산출한 뒤, 국민연금 가입 확률을 추정하고 이 확률이 높은 순서로 가입자를 선별하는 것을 골자로 한다.

먼저, 국민연금 가입자 규모는 경제활동인구조사 8월 근로형태별 부가조사 자료에서 산출한 임금/비임금 근로자 대비 국민연금 가입 비율을 이용한다.[21] 이때, 경제활동인구조사에서 사업장/지역 가입 여부만을 조사하고 있으므로, 본 연구에서도 임의가입과 임의계속가입을 고려하지 않는다.[22] 임의계속가입을 고려하지 않으므로, 국민연금 가입 대상 역시

21) 이처럼 국민연금 공표 통계가 아니라 경제활동인구조사를 이용해 가입자 규모를 파악하는 것은, 종사상 지위, 소득 수준 등의 경제활동상태와 국민연금 가입 상태 사이의 연결성을 반영하기 위함이다.

18~59세로 한정한다.

한편, 근로형태별 부가조사의 조사 시점이 8월이기에, 동 자료와 연말 기준의 국민연금 가입종별 현황 자료 사이에 시차가 존재한다. 또한 노동시장의 종사상 지위와 국민연금 가입종별이 완전히 일치하지 않으므로, 경제활동인구조사의 성-연령-종사상 지위별 가입자 비율과 통계청 자료를 이용해 산출한 가입자 비율이 다를 수 밖에 없다. 특히, 경제활동인구조사 자료 상에서는 임금근로자의 가입률이 높은 반면, 비임금근로자의 가입률은 지나치게 낮다.

이에 본 연구는 경제활동인구조사와 통계청의 국민연금 통계자료를 모두 이용하여 가입자 수를 산출한다. 먼저, 사업자 가입자 수는 성-연령(1세)별로 임금근로자 수 대비 가입자 수의 비율을 적용하여 산출한다. 반면, 지역가입자는 성-연령(1세)별 비임금근로자 수 대비 가입자 수의 비율을 적용하여 1차로 지역가입자를 산출한 뒤, 통계청의 성-연령대별 인구 수 대비 지역가입자 수 비율을 적용하여 가입자 규모를 재산출한다.

이상의 과정을 통해 산출된 성-연령-가입종별 국민연금 가입자 규모는 〈표 4-7〉과 같다. 이를 살펴보면, 남성의 경우 기초자료와 실적치 사이의 차이가 그리 크지 않은 것을 알 수 있다. 기초자료가 인구주택총조사의 2% 표본임을 고려하면, 남성의 경우 기초자료와 실적치 간 약 20,950명 정도의 차이가 발생하고 있는 셈이다. 반면, 여성의 경우 가입종별에 따라 기초자료와 실적치 간 상당한 차이가 발생하고 있다. 전체적으로는 약 7,800명 정도의 차이가 발생하고 있지만, 사업장가입자의 수는 실적치에

22) 국민연금 가입 유형은 크게 사업장가입, 지역가입(소득신고, 납부예외), 임의가입, 임의계속가입으로 구분된다. 2022년 12월 기준, 임의가입자는 36.5천명, 임의계속가입자는 50.1만 명 수준으로, 각각 국민연금 가입자 1,943만 명의 1.88%와 2.58%를 차지하고 있다(국민연금공단, 2022). 이처럼 두 가입종별이 전체 가입자에서 차지하는 비중이 그리 크지 않으므로 이들을 고려하지 않는 것의 부작용은 그리 크지 않으리라 판단된다. 다만, 이 경우 생애기간동안 누적되는 가입 기간이 다소 짧아질 수 있다는 점에 유의해야 한다.

비해 많고 지역가입자의 수는 실적치에 비해 낮게 산출되었다. 이러한 차이는 앞서 설명한 것처럼 경제활동인구조사와 통계청 자료의 연금 통계의 시점 차이, 노동시장의 종사상 지위와 가입종별 간 차이에 기인하는 것으로 판단된다.

〈표 4-7〉 성-연령-가입종별 국민연금 가입자 규모 비교(2015년 기준)

(단위, 명)

남성						
구분	사업장			지역		
	KIHASA SIM (A)	실적치 (B)	차이 (A-B)	KIHASA SIM (A)	실적치 (B)	차이 (A-B)
18~19세	873	1,163	-290	9	9	0
20-24세	5,530	5,529	1	148	185	-37
25-29세	15,961	15,961	0	713	773	-60
30-34세	23,912	23,911	1	2,289	2,292	-3
35-39세	24,557	24,557	0	3,598	3,598	0
40-44세	24,986	24,986	0	5,572	5,572	0
45-49세	22,789	22,787	2	7,319	7,318	2
50-54세	19,247	19,246	1	9,244	9,244	1
55-59세	14,829	14,829	0	10,720	10,755	-35
소계	152,684	152,970	-286	39,612	39,745	-133
여성						
연령	사업장			지역		
18~19세	828	1,031	-203	54	5	49
20-24세	9,012	9,012	0	180	180	0
25-29세	14,638	14,639	-1	734	734	0
30-34세	14,668	14,231	437	1,549	1,986	-437
35-39세	13,501	12,860	641	2,440	3,081	-641
40-44세	15,853	14,617	1,236	3,767	5,003	-1,236
45-49세	16,948	14,890	2,058	4,777	6,835	-2,058
50-54세	15,946	12,506	3,440	5,110	8,551	-3,441
55-59세	14,500	9,361	5,139	4,566	9,706	-5,140
소계	115,894	103,147	12,747	23,177	36,079	-12,902
전체						
구분	사업장	사업장	사업장	지역	지역	지역
계	268,578	256,117	12,461	62,789	75,825	-13,036

주: 국민연금 실적치는 국민연금 통계의 2%임.
자료: KOSIS, 국민연금통계, 2023.9.1. 인출

다음으로, 국민연금 가입 확률은 경제활동인구조사의 근로형태별 부가조사를 이용해 추정한다. 추정 모형은 아래의 (식 4-1)과 같으며, 가입 확률 추정에는 성별, 연령, 연령 제곱, 교육 수준, 배우자 유무, 임금 수준, 종사상 지위 등의 요인을 고려하였다. 가입확률 추정에 활용된 변수의 구성 및 특징은 〈표 4-8〉과 같으며, 추정 결과는 〈표 4-9〉와 같다. 국민연금의 가입 확률 추정 뒤에는 앞서 계산한 성-연령-가입종별 가입자 규모에 따라 가입 확률이 큰 순서로 국민연금 가입자를 선별한다.

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 w_i + \beta_4 marr_i + \sum_{j=1}^{k_1} \beta_{5j} D_{i,sj} \quad \text{(식 4-1)}$$

단, p_i : i의 국민연금 가입 확률, age_i : i의 연령(만나이), w_i : i의 평균임금.
$marr_i$: i의 혼인상태, $D_{i,sj}$: i의 교육더미(j=1,...,k_1),

〈표 4-8〉 국민연금 가입확률 추정에 사용된 변수의 구성 및 특징

변수 구분	변수	처리내용
종속 변수	국민연금 가입여부	미가입 =0, 가입=1
개인적 요인	성별	여성=0, 남성=1
	연령	만18-59세
	연령제곱	
	교육 수준	중졸 / 고졸 / 전문대·4년제 대졸 / 대학원 이상
	배우자 유무	미혼=1 유배우자=2, 이혼,사별=3
경제적 요인	log(평균 임금)	지난 3개월간 평균임금(만원)의 log 값
노동 요인	종사상지위	상용근로자=1, 임시일용근로자=2, 고용주=3, 자영자=4

자료: 필자 작성

〈표 4-9〉 기초자료의 국민연금 가입확률 추정 결과

		2015년	
성별		0.348***	(0.047)
연령		0.11***	(0.016)
연령제곱		-0.001***	(0.000)
배우자유무		0.064	(0.054)
교육수준 (기준: 고졸)	중졸미만 더미	-0.266***	(0.066)
	대졸 더미	0.245***	(0.052)
	대학원졸 더미	0.062	(0.160)
종사상지위 (기준: 자영자)	상용근로 더미	2.986***	(0.165)
	임시일용 더미	-1.518***	(0.090)
	고용주 더미	0.846***	(0.082)
평균 임금		0.005***	(0.000)
상수항		-2.333***	(0.319)
Pseudo R2		0.4230	
Obs.		28,091	

주: 1) 괄호 안은 표준오차임.
2) *** p〈0.01, ** p〈0.05, * p〈0.1
자료: 필자 작성

가입자 선별 이후에는 국민연금 가입자의 가입개월 수와 납부보험료를 설정해준다. 먼저 당해년의 가입개월 수와 납부보험료는 통계청 국민연금 통계를 이용해 계산한 성-연령-가입종별-기준소득월액 구간별 비율을 이용해 각 집단별 가입자 수를 산출한 뒤, 무작위 추출로 대상자를 선별한다.

또한 1988년부터 2015년까지의 연금 이력은 국민노후보장패널(KReIS)과 행정 자료를 결합한 자료(1988~2005년)와 DOSA의 시뮬레이션 자료(2006~2015년)를 이용하여 산출한다. 즉, 기존 모형에서 2015년까지 시뮬레이션한 결과를 이용하여 과거의 국민연금 가입 이력을 생성한다. 가입이력 생성 과정은 다음과 같다. 먼저, 1988~2005년 기간에

대해서는 월 단위로 되어 있는 행정 자료를 연간 단위로 변환한 뒤, 성-연령-가입종별로 연금 이력을 무작위 복원 추출하여 2005년을 기준으로 구축된 DOSA의 기초 자료와 매칭시킨다.[23] 이후 2015년까지 시뮬레이션을 수행하여 연장된 성-연령-가입종별 연금 이력을 무작위 복원 추출하여 KIHASA SIM의 기초 자료와 매칭한다.[24]

다음으로, 급여종별 수급자 규모는 통계 연보에 제시된 각 급여종별 수급자 수 비율을 경제활동인구조사의 성-연령별 인구에 적용하여 산출한다. 이때, 급여종별 수급자는 성-연령별로 무작위 추출하여 선정한다. 또한 이들의 수급액은 가입자의 연금이력 생성에 활용하였던 국민노후보장패널(KReIS)과 행정자료를 결합한 자료(1988~2005년)를 이용해 보정한다. 소비자물가지수를 이용해 2005년 기준의 수급액을 2015년 수준으로 조정한 뒤, 성-연령별로 무작위 추출하여 보정하는 방식이다.

23) 매칭된 이력은 별도의 파일로 보관한 후 향후 시뮬레이션 과정에서 수급권 판단 및 급여액 산정에 사용한다.

24) 이처럼 복잡한 과정을 통해 과거의 연금이력을 생성하는 것은 국민연금 가입기간 자료가 공개되어 있지 않기 때문이다. 자료 접근성의 한계는 본 연구의 연구진이 극복하기 어려운 문제로 보이며, 현재로써는 이러한 방법이 최선이라 판단된다. 추후 관련 자료가 공개된다면, 모형의 정확도를 제고할 수 있을 것으로 보인다.

제5장

주요 모듈 구축

제1절 인구 모듈
제2절 가구 구성 모듈
제3절 교육·경제활동 모듈
제4절 소득추정 모듈
제5절 공적연금 모듈
제6절 고용 보험 모듈

제 5 장 주요 모듈 구축

제1절 인구 모듈[25)]

1. 출산 모듈

출산 모듈에서는 출산 과정을 시뮬레이션한다. 핵심적인 시뮬레이션 과정은 출생아 수를 결정하고, 출산 확률에 따라 출산 대상자를 선별한 뒤 해당 가구에 출생아를 할당하는 것이다. 이에 활용되는 기초자료는 〈표 5-1〉과 같다.

〈표 5-1〉 출산 모듈에 활용되는 기초 자료

구분		DOSA	KIHASA SIM
출산	alignment set	• 2005년-2014년 실적치: 통계청 인구동향조사 • 2015년 이후는 3차 재정 추계의 연도별 출생아 수 및 성비	• 2016년-2022년: 통계청 인구동향조사 • 2013년~ : 송창길 외(2023)
	parameter set	• 출산확률셋: 노동패널(1-17차)	• 출산확률셋: 한국노동패널(3-22차)

자료 : 1) 고제이 외. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 -노후 소득 보장 정책을 중심으로. 한국보건사회연구원, p.79. 〈표 4-1〉
2) 필자 작성

25) KIHASA SIM은 고제이 외(2016)가 개발한 DOSA를 개선한 것이므로, 주요 모듈 내 시뮬레이션 과정에 대한 설명이 고제이 외(2016)과 유사할 수 밖에 없다. 즉, 이 장의 내용은 고제이 외(2016)를 주로 참고하였다.

구체적인 출산 시뮬레이션 과정은 다음과 같다. 먼저, 외부 자료로부터 성-연령별 출생아 수를 불러와 당해 연도의 출생아 수를 결정한다. 이때 실적치가 존재하는 2022년까지는 통계청의 인구동향조사 자료를 이용해 출생아 수를 산출하며, 2023년부터는 송창길 외(2023)의 인구전망치를 이용해 출생아 수를 산출한다.

출생아 수를 산출한 뒤에는 20-49세의 배우자가 있는 여성을 출산 대상으로 선별한 뒤 출산 확률 값을 계산한다. 이때, 출산 대상을 20세 이상자로 한정한 것은 KIHASA SIM에서는 20세 이상부터 결혼 시장에 진입하며, 혼외자를 고려하지 않기 때문이다. 출산 확률은 한국노동패널조사 자료를 통해 추정한 계수값을 이용해 계산하며, 이 출산확률이 높은 부부 순으로 각 가구에 출생아를 할당한다.[26]

한편, 출산확률은 3~22차(2000~2019년) 한국노동패널조사 자료를 풀링(pooling)한 뒤 Logit 모형을 이용해 추정한다. 분석 대상은 조사 기간 동안 아이를 출산한 부부이며, 추정식은 아래의 (식 5-1)과 같다. 출산 확률 추정에 사용되는 설명 변수들은 선행연구들이 고려한 주요 변수들 중 한국노동패널과 기초 자료에 존재하거나 계산가능한 변수들이다.[27] 〈표 5-2〉는 주요 MSM에서 사용한 출산 관련 변수들을 보여준다. 이에 따르면, 출산 확률 추정 시 연령, 고용 상태, 결혼 상태, 출산경력(자녀 수), 교육 참가 여부, 교육 수준, 출산 대기 기간 등의 변수들이 주로 사용되고 있다(고제이 외, 2016, p.83). 이에 본 연구는 연령, 연령 제곱, 취업 여부, 교육 연수, 기존 자녀 수, 출산 대기 기간, 출산 대기 기간의 제곱을 설명 변수로 활용한다(〈표 5-3〉 참조). 출산확률 추정 결과는 〈표 5-4〉와 같다.

---

26) 2019년까지의 자료를 사용한 것은 코로나19 확산의 영향을 피하기 위함이다.

27) 이처럼 행태방정식 추정 시 기초자료에 존재하는 변수들로만 설명 변수를 구성해야 하는 것은 MSM이 지닌 한계 중 하나이다. 다만, KIHASA SIM처럼 거시집계치를 조정하는 경우, 각 개인을 줄 세우는 것에만 행태방정식의 추정치가 활용된다는 점에서 추정 결과의 설명력이 다소 낮은 것은 큰 문제가 되지 않을 것으로 보인다.

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + X_i\beta + \epsilon_i \qquad \text{(식 5-1)}$$

단, p_i : i의 출산확률, X_i : i의 설명변수 벡터

〈표 5-2〉 주요 MSM의 출산 관련 변수

모형	사건	변수
DYNAMOD	첫번째 출산	교육참여 여부, 고용상태, 연령, 결혼 기간
	2번째 이후의 출산	출산경력, 교육참여 여부, 자녀연령, 최근 출산 이후 지속기간, 고용상태
FAMSIM	첫번째 출산	연령, 연령제곱, 동거여부, 결혼여부, 동거기간, 결혼기간, 교육참여여부, 교육지속기간, 노동여부, 노동기간 등
	두,세번째 출산	연령, 연령제곱, 동거여부, 결혼여부, 최근 출산이후 지속기간, 결혼기간, 교육참여여부, 교육지속기간, 노동여부, 노동기간 등
DYNACAM	출산	연령, 결혼 상태, 교육수준, 고용상태, 출산경력
LifePaths	출산	연령, 결혼상태, 자녀수.
DESTINIE	첫번째 출산	결혼 지속 년수, 교육수준
	두번째 출산	결혼상태, 출산경력
DYNAMITE	출산	연령, 결혼상태, 출산경력
MOSART	출산	연령, 출산경력, 가장 어린 자녀의 연령
SESIM	출산	연령, 결혼상태
LIFEMOD	출산	결혼상태, 연령, 출산 경력

자료: 고제이 외. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 -노후 소득 보장 정책을 중심으로. p.83. 〈표 4-3〉, 일부 수정

〈표 5-3〉 출산확률 추정에 사용된 변수의 구성과 특징

변수	처리내용
연령	만 20-49세
연령 제곱	
취업 여부	미취업 = 0, 취업 = 1
교육연수 (학력)	무학 = 0, 초등학교 중퇴 = 3, 초등학교 졸업,수료 = 6, 중학교 중퇴,재학,휴학 = 7.5, 중학교 졸업,수료 = 9, 고등학교 중퇴,재학,휴학 = 10.5, 고등학교 졸업,수료 = 12, 2년제 대학 중퇴,재학,휴학 = 13, 2년제대학 졸업,수료 = 14 4년제대학 중퇴,재학,휴학 = 14, 4년제대학 졸업, 수료 = 16, 대학원석사 중퇴,재학,휴학 = 17, 대학원석사 졸업,수료 = 18, 대학원박사 중퇴,재학,휴학 = 19, 대학원박사 = 20
기존 자녀 수	0-6 명
출산 대기기간	출산 대기 기간은 막내 자녀 출산 이후의 기간 또는 결혼 이후의 기간을 의미
출산 대기기간의 제곱	

주: 종속 변수는 출산 여부임.
자료: 필자 작성

〈표 5-4〉 출산확률 추정결과

	계수	표준오차
연령	0.373***	(0.075)
연령 제곱	-0.007***	(0.001)
취업 여부(취업=1)	-0.938***	(0.062)
교육 수준(학력)	0.022	(0.014)
자녀 수	-1.237***	(0.040)
출산 대기기간	0.418***	(0.040)
출산 대기기간의 제곱	-0.049***	(0.004)
상수항	-6.487***	(1.174)
Log likelihood	-5060.427	
LR chi2	3942.43	
Pseudo R2	0.2803	
관측치 수	48,495	

주: 1) 출산 대기 기간은 막내 자녀 출산 이후의 기간 또는 결혼 이후의 기간을 의미함.
2) *** p〈0.01, ** p〈0.05, * p〈0.1
자료 : 필자 작성

2. 사망 모듈

사망 모듈에서는 연도별 순차 과정의 마지막 단계인 사망을 시뮬레이션한다. 대부분의 MSM은 출생과 사망을 같은 모듈에서 시뮬레이션하지만(Li et al., 2014), KIHASA SIM은 사망 이후 발생하는 유족연금의 계산을 위해 별도의 사망 모듈을 구축한다. 한편으로, 이러한 모듈 배치는 연도별 생애 사건들을 모두 처리한 이후에 사망이 일어난다는 가정을 적용한 것이기도 하다.

KIHASA SIM의 사망자 수는 외부 자료를 통해 외생적으로 결정된다. 실적치가 존재하는 기간 동안에는 인구동향조사 자료의 성-연령별 사망 인구비(천분비)를 이용하여 사망자 수를 결정하며, 2023년부터는 송창길 외(2023)의 인구전망치에 따라 사망자 수를 결정한다.

성-연령별 사망자는 무작위 추출로 선별하며, 사망자 선별 이후에는 해당 가구의 가구 정보를 수정한다. 먼저, 가구 대표 사망 시 배우자가 있다면 배우자를 새로운 가구 대표로 설정하고, 배우자가 없거나 가구 대표와 함께 사망하여 자녀들만 생존하는 경우는 자녀들이 모두 개별 가구로 독립하는 것을 가정한다. 한편, 사망 모듈에서는 유족연금의 발생 및 소멸을 처리한다. 사망 모듈에서 이를 처리하는 것은 유족연금의 발생 요건이 국민연금 수급권자의 사망이기 때문이다.

〈표 5-5〉 사망 모듈에 활용되는 기초 자료

구분		DOSA	KIHASA SIM
사망	alignment set	• 3차 재정계산의 연도별 사망자수 및 성비	• 송창길 외(2023) 성-연령별 사망률

자료: 1) 고제이 외. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 -노후 소득 보장 정책을 중심으로. p.79. 〈표 4-1〉
2) 필자 작성

제2절 가구 구성 모듈

가구 구성 모듈에서는 결혼과 이혼, 그리고 그로 인해 발생하는 가구의 구성 변화를 시뮬레이션한다. 외부 집계치를 이용해 결혼 및 이혼 건수를 결정한 뒤(〈표 5-6〉 참조), 해당 사건의 대상자를 결정하고, 그에 따른 가구 변화를 정리하는 것이 핵심적인 시뮬레이션 과정이다.

구체적인 시뮬레이션 과정은 다음과 같다. 먼저, 결혼 건수는 실적치가 존재하는 2022년까지는 인구동향조사의 성-연령대별(20세-59세, 5세 단위) 결혼 건수의 천분비를 이용하여 산출한다. 2023년 이후의 혼인률은 2022년의 비율을 그대로 연장하여 사용한다. 2022년의 혼인율 집계치 조정 자료는 〈표 5-7〉과 같다.

〈표 5-6〉 가구 구성모듈에 활용되는 기초자료

구분		DOSA	KIHASA SIM
결혼 (초·재혼)	alignment set	• 통계청의 인구동향조사 자료의 인구 천명 당 결혼 건수 (2005-2015)	• 통계청의 인구동향조사 자료의 인구 천명 당 결혼 건수 (2016-2022)
	parameter set	• 결합확률 셋: 인구·주택총조사 내 결혼한 부부	• 결합확률 셋: 인구·주택총조사 내 결혼한 부부
이혼	alignment set	• 통계청의 인구동향조사 자료의 인구 천명 당 이혼 건수 (2005-2015)	• 통계청의 인구동향조사 자료의 인구 천명 당 이혼 건수 (2016-2022)

자료: 1) 고제이 외. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 -노후 소득 보장 정책을 중심으로. p.85. 〈표 4-5〉
2) 필자 작성

〈표 5-7〉 혼인율 집계치 조정 자료(2022년)

연도	성별	연령대	전체 혼인율	초혼 비율	재혼 비율	사별 후 재혼 비율	이혼 후 재혼 비율
2022	남성	15-19세	0.2	99.4	0.6	0.0	100.0
2022	남성	20-24세	2.6	97.6	2.4	5.6	94.4
2022	남성	25-29세	20.4	98.5	1.5	4.3	95.7
2022	남성	30-34세	40.3	97.1	2.9	3.2	96.8
2022	남성	35-39세	20.9	88.6	11.4	3.0	97.0
2022	남성	40-44세	8.1	68.9	31.1	2.9	97.1
2022	남성	45-49세	4.4	42.5	57.5	3.2	96.8
2022	남성	50-54세	3.2	18.8	81.2	4.9	95.1
2022	남성	55-59세	2.5	8.3	91.7	7.1	92.9
2022	남성	60-64세	1.8	4.4	95.6	10.6	89.4
2022	남성	65-69세	1.1	4.1	95.9	13.4	86.6
2022	남성	70-74세	0.8	4.2	95.8	15.7	84.3
2022	여성	15-19세	0.9	99.2	0.8	28.6	71.4
2022	여성	20-24세	7.0	96.1	3.9	4.2	95.8
2022	여성	25-29세	35.7	96.9	3.1	5.1	94.9
2022	여성	30-34세	41.3	92.6	7.4	3.9	96.1
2022	여성	35-39세	15.0	75.1	24.9	3.6	96.4
2022	여성	40-44세	5.6	40.2	59.8	4.1	95.9
2022	여성	45-49세	3.4	15.9	84.1	5.7	94.3
2022	여성	50-54세	2.7	6.8	93.2	7.3	92.7
2022	여성	55-59세	2.0	5.5	94.5	10.5	89.5
2022	여성	60-64세	1.2	5.5	94.5	13.8	86.2
2022	여성	65-69세	0.7	7.1	92.9	16.1	83.9
2022	여성	70-74세	0.5	7.6	92.4	15.3	84.7

자료: 2022년 인구동향조사 자료를 이용해 필자 작성; 통계청. (각년도). 인구동향조사. 원자료. https://mdis.kostat.go.kr/ofrData/selectOfrDataDetail.do?survId=38&itmDiv=1&nPage=3&itemId=2001&itemNm=%EC%9D%B8%EA%B5%AC 에서 2023. 5.10. 인출.

다음으로, 20-59세 인구 중 배우자가 없는 자들을 무작위 추출하여 결혼 시장 진입자를 선정한다. 이때, 초혼대상자와 재혼대상자의 결합

확률이 낮음을 고려하여, 결혼대상자를 초혼대상자과 재혼대상자로 구분한다. 재혼은 생애 1번으로 한정한다. 즉, 동 모형에서는 생애기간 동안 최대 2번까지 결혼할 수 있다. 재혼 이후에 이혼/사별을 경험한 자는 결혼대상자에서 제외되며, 군 복무 중인 자 역시 결혼 대상에서 제외된다. 한편으로, 가족 내에 여러 명의 결혼대상자가 존재하는 경우 당해 연도에는 연장자만 결혼 시장에 진입하도록 설정한다. 이는 가족끼리의 결혼을 방지하기 위함이다.

결혼 시장 진입자가 결정된 이후에는 Perses(2002)가 제안한 Mate Matching 알고리즘을 이용하여 결혼을 시뮬레이션한다. 이는 무작위로 남성 1명을 선정한 뒤, 이 남성과 여성들 간의 결합 지수(compatibility index)가 가장 높은 여성을 결혼시키고, 남은 남성 중 1명을 무작위로 선정하여 같은 방식으로 결혼을 시키고, 같은 과정을 계속해서 반복하는 방식이다. 전술한 매칭 과정은 혼인 건수만큼 반복된다. 이때, 결합 지수는 두 남녀가 배우자가 될 수 있는 확률, 즉 남녀의 결혼 확률을 의미한다(고제이 외, 2016, p.88). 본 연구는 인구주택총조사 자료에서 2012-2015년 사이에 부부가 된 17,133쌍을 대상으로 남녀의 결혼 확률을 추정하여 결합 지수로 사용한다.

마지막으로 이혼 건수는 결혼 건수와 마찬가지로 인구동향조사 자료에서 계산한 성-연령대(5세 단위)별 남성 기준 이혼 건수 천분비를 사용하여 산출한다. 결혼과 마찬가지로 실적치가 존재하는 2022년까지는 실적치를 사용하며, 그 이후에는 2022년의 비율이 그대로 지속한다고 가정한다. 이혼 부부는 무작위 추출로 선정한다. 이때, 이혼은 생애 1회로 한정한다.

제3절 교육·경제활동 모듈

1. 교육 모듈

교육 모듈에서는 상급학교로의 진학 과정과 군 입대를 시뮬레이션한다. 먼저, 진학 대상자는 만 7~32세로 한정한다. 이는 학제별 교육 연수를 고려한 것이다.[28] 한편으로, 이 모형에서는 학업과 노동을 동시에 수행하지 않으며, 모든 학업을 마치고 노동시장에 진입하는 것으로 가정한다. 또한 노동시장 진입 이후에는 다시 학교로 돌아가지 않는 것으로 가정한다.[29] 예컨대, 고등학교 진학 이전에 중퇴한 자는 새로운 학교로 진학하는 것이 아니라 노동시장에 진입하는 것으로 처리한다. 이 같은 가정들을 적용하는 것은 일-학습 병행자(returning-adults)의 행태를 구현하기 위한 기초 자료가 부족하기 때문이며, 교육-노동으로 이어지는 시뮬레이션 과정을 단순화하기 위함이기도 하다.

구체적인 교육의 시뮬레이션 과정은 다음과 같다. 먼저, 한국교육개발원의 교육통계자료에서 제공하는 성별 상급학교 진학률과 취업률을 이용해 학교급별로 상급학교 진학자와 미진학자 수, 그리고 미진학자 중 취업자 수를 산출한다. 이때, 실적치가 존재하는 2016~2021년 동안에는 실적치를 사용하며, 이후에는 2021년의 진학률과 취업률 수준이 지속된다고 가정한다(〈표 5-8〉 참조).

진학자와 취업자의 선별 과정은 다음과 같다. 먼저, 대학 진학을 제외한 상급학교 진학자는 균등 확률로 무작위 추출한다. 다음으로, 대학 진학자는 2년제와 4년제 대학에 대한 진학 확률을 이용하여 선별한다. 진학 확률

---

28) 학제별 교육 연수는 초·중·고등학교 12년, 학사 4년, 석사 2년, 박사 5년으로 총 23년이며, 이에 남성의 군 복무 기간 2년을 고려하면 총 25년이 된다.

29) 즉, 이 모형에서는 재수생을 고려하지 않는다.

계산에 사용되는 모수는 3-22차(2000~2019년) 한국노동패널 자료를 통합한 뒤 다항로직 분석으로 추정하였다. 이때, 다항로짓 분석에 사용되는 설명 변수는 가구 근로소득과 부모의 교육 수준이다(〈표 5-9〉 참조). 이처럼 단순한 모형을 설정한 것은 기초 자료 내에 진학 확률의 추정에 이용할 만한 정보가 많지 않았기 때문이다. 다항로짓 분석에 사용된 모형은 아래의 (식 5-2)와 같으며, 대학 진학 확률의 추정 결과는 〈표 5-10〉과 같다.

〈표 5-8〉 교육 모듈에 활용되는 기초자료

	구분	DOSA	KIHASA SIM
상급학교 진학 (대학 제외)	alignment set	• 한국교육개발원 교육통계:성별 상급학교 진학률과 취업률 (2005-2014)	• 한국교육개발원 교육통계: 성별 상급학교 진학률과 취업률 (2016-2021)
대학 진학	alignment set	• 한국교육개발원 교육통계:성별 상급학교 진학률과 취업률 (2005-2014)	• 한국교육개발원 교육통계:성별 상급학교 진학률과 취업률 (2016-2021)
	parameter set	• 대학진학 확률: 한국노동패널 (3~17차)	• 대학진학 확률: 한국노동패널 (3~22차)
군대 입대	alignment set	• 군 면제율: 병무청 병무통계 (2005-2015)	• 군 면제율: 병무청 병무통계 (2016-2022)

자료: 1) 고제이 외. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 -노후 소득 보장 정책을 중심으로. p.93. 〈표 5-1〉. 일부 수정
2) 필자 작성

〈표 5-9〉 대학 진학 확률 추정에 사용된 변수의 구성 및 특징

변수	처리 내용
대학 진학	미진학=0, 2년제 대학 진학=1, 4년제 대학 진학=2
로그 (근로소득)	연간 가구 근로소득(2019년 불변가격으로 현가화)의 로그값
부모 중 고학력자의 교육 연수 (학력)	무학 = 0, 초등학교 중퇴 = 3, 초등학교 졸업/수료 = 6, 중학교 중퇴/재학/휴학 = 7.5, 중학교 졸업/수료 = 9, 고등학교 중퇴/재학/휴학 = 10.5, 고등학교 졸업/수료 = 12, 2년제 대학 중퇴/재학/휴학 = 13, 2년제 대학 졸업/수료 = 14 4년제 대학 중퇴/재학/휴학 = 14, 4년제 대학 졸업/수료 = 16, 대학원 석사 중퇴/재학/휴학 = 17, 대학원 석사 졸업/수료 = 18, 대학원 박사 중퇴/재학/휴학 = 19, 대학원 박사 = 20

자료: 필자 작성

$$\log\left(\frac{p_{gi}}{p_{ri}}\right) = \beta_0 + \beta_1 \log(w_i) + \beta_2 pedu_i + \epsilon_i$$

(식 5-2)

단, p_i : i의 진학 확률 (r=미진학, g=2년제 진학, 4년제 진학)

$\log(w_i)$: i의 가구 근로소득

$pedu_i$: i의 부모의 교육 수준

〈표 5-10〉 대학 진학 확률 추정 결과: 다항로짓 분석

	2년제		4년제	
	계수	표준오차	계수	표준오차
로그(근로소득)	0.177***	(0.066)	0.321***	(0.063)
부모의 교육 수준	0.092***	(0.019)	0.237***	(0.018)
상수항	-2.765***	(0.763)	-5.801***	(0.733)
log likelihood	-3324.2005			
LR chi2(4)	348.88			
Pseudo R2	0.0499			
Obs.	3,481			

주: 1) 기준범주는 대학 미진학임.
2) *** p〈0.01, ** p〈0.05, * p〈0.1

자료: 필자 작성

군 복무에 대한 시뮬레이션 과정은 다음과 같다. 먼저, 병무통계의 면제율을 이용하여 입대자와 면제자 수를 산출한다. 이때, 실적치가 존재하는 2022년까지는 실적치를 이용하여, 이후의 기간에 대해서는 2022년의 면제율을 적용한다.

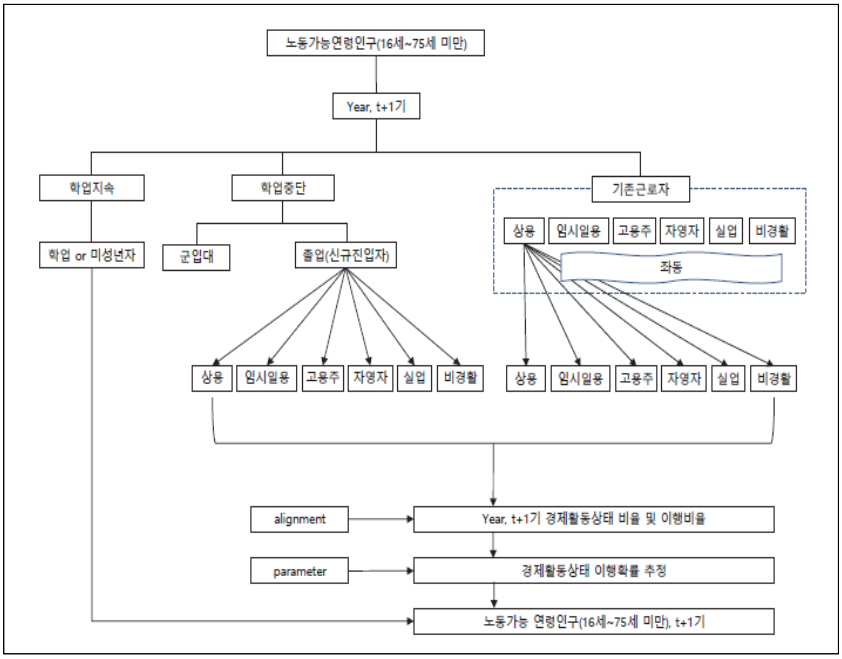
다음으로, 입영 대상은 균등 확률로 무작위 추출한다. 이때, 여성은 군 복무 대상에서 제외되며, 중학교 중퇴자도 면제로 간주한다. 입대 시기는 2년제 대학과 4년제 대학에 진학하지 못한 경우 고교 졸업 후 바로 입대하는 것으로 가정하며, 대학 진학자는 1학년을 마친 시점에 입대하는 것으로 가정한다. 복무기간은 2년이다. 군 복무가 완료된 이후에는 대학 재학자는 복학으로 처리하며, 취업자와 미취업자는 노동시장에 진입하는 것으로 처리한다.

2. 경제활동 모듈

경제활동 모듈에서는 매년의 경제활동 상태 변화를 시뮬레이션한다. 이때, 경제활동 상태는 '상용, 임시/일용, 고용주, 자영자, 실업, 비경활'의 6가지 상태로 구분한다(〔그림 5-1〕 참조). 경제활동 상태별 인구 수는 기재부(2020)의 거시경제 전망치 중 성-연령별 경제활동인구 비율에 경제활동인구조사에서 추출한 종사상 지위별 인구 비율을 적용해 산출하며, 개인의 경제활동 상태는 전이확률을 통해 결정한다.

경제활동 상태에 대한 자세한 시뮬레이션 과정은 아래와 같다. 시뮬레이션 자료에 성-연령-경제활동 상태별 비율을 적용하여 16~75세의 경제활동 상태별 규모를 산출한다. 이때, 실적치가 존재하는 2022년까지는 경제활동인구조사 자료에서 산출한 전체 경제활동자 대비 경제활동 상태별 비율을 적용하며, 이후의 기간에는 기획재정부(2020)의 전망치를 적용한다(〈표 5-11〉 참조).

〔그림 5-1〕 경제활동 상태 시뮬레이션 과정



자료: 고제이 외. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 -노후 소득 보장 정책을 중심으로. 한국보건사회연구원. p. 117. 〔그림 5-5〕.

〈표 5-11〉 경제활동 상태 모듈에 활용되는 기초자료

	구분	DOSA	KIHASA SIM
경제 활동 상태의 이행	alignment set	• 성·5세별 경제활동인구 비율: 경제활동인구조사(2001년~2015년) • 2015년 장기재정전망(기획재정부) • 경제활동상태별 이행비율: 한국노동패널조사(2005-2014년)	• 성·5세별 경제활동인구 비율: 경제활동인구조사(2015년~2022년) • 2020년 장기재정전망(기획재정부) • 경제활동상태별 이행비율: 한국노동패널조사(2015-2019년)
	parameter set	• 경제활동상태 전이확률 셋: 한국노동패널조사(2005-2014년)	•경제활동상태 전이확률 셋: 한국노동패널조사(2010-2019년)

자료: 고제이 외. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 -노후 소득 보장 정책을 중심으로. 한국보건사회연구원. p. 121. 〈표 5-11〉.

경제활동 상태별 인구 수를 산출한 이후에는 시뮬레이션 대상자를 기존 근로자와 노동시장 신규 진입자로 구분하여 경제활동 상태 변화를 시뮬레이션 한다.

먼저, 노동시장 신규 진입자의 경제활동 상태는 교육 수준을 고려하여 결정한다. 석사 이상 학위자는 상용 임금근로자로 구분하며, 학사 이하의 경제활동 상태는 무작위 추출로 결정한다.

다음으로, 이미 노동시장에 진입하여 경제활동 중인 근로자의 경제활동 상태는 전이확률을 통해 결정한다. 이에 활용되는 전이확률은 2012~2019년의 한국노동패널조사 자료에서 다항로짓 모형으로 추정한 것이다.[30] 전이확률의 추정식은 아래의 (식 5-3)과 같다. (식 5-3)에서 경제활동 상태의 선택확률 $\Pr(y_j = j)$는 0과 1 사이의 값을 가지며 모든 j에 대한 확률의 합은 1이 된다. 한편으로 전이확률 추정에 사용하는 변수는 해외의 주요 MSM을 참조하되, 한국노동패널과 기초자료에 존재하거나 계산가능한 변수로 한정하였다. 〈표 5-12〉를 보면, 주요 MSM은 연령, 성별, 교육 수준, 혼인 상태, 전년도의 경제활동참가 여부 등의 변수를 고려하였는데, 이 모형에서는 기초자료에 변수가 존재하는 연령, 연령 제곱, 가구주 여부, 교육 연수, 혼인 상태, 5세 미만 자녀 수, 전기의 경제활동 상태를 이용한다. 성별 경제활동 상태 전이확률의 추정 결과는 〈표 5-13〉과 〈표 5-14〉와 같다.

$$\Pr(y_j = j) = \frac{\exp(X_{ij}^{'}\beta)}{1+\exp(X_{i2}^{'}\beta)+\exp(X_{i3}^{'}\beta)+\ldots+\exp(X_{ik}^{'}\beta)},\ j = 2, \ldots, 6$$

$$\Pr(y_j = j) = \frac{1}{1+\exp(X_{i2}^{'}\beta)+\exp(X_{i3}^{'}\beta)+\ldots+\exp(X_{ik}^{'}\beta)},\ j = 1 \quad \text{(식 5-3)}$$

30) COVID19가 확산된 이후인 2020~2021년의 자료는 분석에서 제외하였다.

〈표 5-12〉 주요 MSM의 경제활동 상태 설명 변수

구분	설명 변수	경제활동 상태
DYNAMOD	연령, 성, 교육, 혼인상태, 자녀 유무 및 연령, 현재 학업, 계절	상용직, 임시직, 비경활, 실업
DYNASIMS	연령, 성, 교육, 인종, 지역, 장애, 혼인상태, 자녀, 학업, 배우자소득	연간 근로시간, 실업
DYNACAN	연령, 성, 교육, 장애, 가족구성, 노동 이력, 은퇴상태	연간 근로한 주 0, 1-47 또는 48+
Lifepaths	현재 직업 기간, 연령, 성, 지역, 교육, 자녀, 배우자, 배우자 고용	임금 근로, 자영, 미취업
SAGE	연령, 성, 가구 구성, 자녀, 고용 이력, 직업 유형, 산업, 건강 상태	상용직, 임시직, 실업, 비경활
MOSART	연령, 성, 자녀, 혼인상태, 교육, 연금 상태, 장애, 전년도 경제활동참가 여부	경제활동참가율
SESIM	연령, 성, 교육, 자녀, 배우자 고용, 여가 선호, 지역 규모, 출생국가, 전년도 고용/실업	연간 노동시간 0-3000, 실업, 질병
PENSIM	연령, 성, 교육, 전기 직업 특성	상용직, 임시직, 실업, 비경활

자료: 고제이 외. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 -노후 소득 보장 정책을 중심으로. 한국보건사회연구원. p. 114, 〈표 5-7〉. 일부 수정

〈표 5-13〉 경제활동상태 전이확률에 대한 다항로짓분석 결과(남성)

		상용	임시/일용	고용주	자영자	실업
가구주 (=1)		0.988***	0.904***	0.936***	1.105***	0.34*
		(0.152)	(0.149)	(0.232)	(0.223)	(0.178)
연령		0.194***	0.172***	0.223***	0.155***	0.224***
		(0.018)	(0.016)	(0.031)	(0.024)	(0.022)
연령 제곱		-0.003***	-0.002***	-0.003***	-0.002***	-0.003***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
혼인 상태 (기준: 미혼)	기혼 (=1)	0.120	-0.303**	0.672***	0.783***	-0.844***
		(0.150)	(0.152)	(0.244)	(0.245)	(0.193)
	이혼/사별 (=1)	-0.242	-0.345*	-0.031	0.605**	-0.556*
		(0.183)	(0.178)	(0.326)	(0.272)	(0.286)
교육 연수		0.135***	-0.003	0.085***	0.008	0.105***
		(0.012)	(0.012)	(0.017)	(0.015)	(0.019)
5세 미만 자녀 수(=1)		0.360	0.162	0.655**	0.562**	-0.068
		(0.230)	(0.239)	(0.272)	(0.267)	(0.301)
전기 경제활동 상태 (기준: 비경활)	전기(상용) (=1)	5.295***	1.466***	2.525***	2.159***	2.323***
		(0.106)	(0.143)	(0.283)	(0.191)	(0.178)
	전기(임시) (=1)	2.135***	4.751***	1.185***	2.144***	1.884***
		(0.136)	(0.109)	(0.439)	(0.202)	(0.200)
	전기(고용주) (=1)	1.880***	1.324***	7.682***	5.680***	2.270***
		(0.268)	(0.307)	(0.277)	(0.211)	(0.429)
	전기(자영자) (=1)	1.955***	1.745***	5.465***	6.824***	1.310***
		(0.188)	(0.187)	(0.256)	(0.151)	(0.363)
	전기(실업) (=1)	1.920***	1.988***	0.959*	1.550***	2.305***
		(0.180)	(0.173)	(0.497)	(0.321)	(0.209)
상수		-7.778***	-6.163***	-10.782***	-8.048***	-8.756***
		(0.367)	(0.380)	(0.724)	(0.568)	(0.452)
log pseudo likelihood		-60,636.372				
Wald chi2		20,412.73				
Pseudo R2		0.6599				
Obs.		38,249				

주: 1) ( )는 표준 오차임
2) ***, **, *는 각각 1%, 5%, 10%에서 통계적으로 유의함을 의미함
자료: 필자 작성

〈표 5-14〉 경제활동상태 전이확률에 대한 다항로짓분석 결과(여성)

		상용	임시/일용	고용주	자영자	실업
가구주 (=1)		0.344***	0.264**	0.413	0.564**	-0.047
		(0.112)	(0.110)	(0.257)	(0.219)	(0.209)
연령		0.194***	0.154***	0.322***	0.218***	0.197***
		(0.016)	(0.014)	(0.050)	(0.023)	(0.030)
연령 제곱		-0.002***	-0.002***	-0.003***	-0.002***	-0.002***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
혼인 상태 (기준: 미혼)	기혼 (=1)	-0.901***	-0.88***	-1.048***	-0.963***	-2.053***
		(0.135)	(0.146)	(0.364)	(0.258)	(0.238)
	이혼/사별 (=1)	-0.606***	-0.468***	-0.757*	-0.705**	-1.132***
		(0.171)	(0.168)	(0.444)	(0.287)	(0.307)
교육 연수		0.109***	0.032***	0.106***	0.052***	0.119***
		(0.010)	(0.010)	(0.026)	(0.017)	(0.026)
5세 미만 자녀 수(=1)		-0.355***	-0.504***	0.053	0.102	-0.737**
		(0.081)	(0.108)	(0.289)	(0.199)	(0.314)
전기 경제활동 상태 (기준: 비경활)	전기(상용) (=1)	5.217***	1.394***	2.04***	2.263***	2.406***
		(0.073)	(0.114)	(0.352)	(0.196)	(0.173)
	전기(임시) (=1)	2.424***	4.607***	1.682***	2.258***	2.227***
		(0.100)	(0.068)	(0.365)	(0.182)	(0.192)
	전기(고용주) (=1)	0.777*	0.718	7.943***	5.65***	-0.571
		(0.437)	(0.475)	(0.250)	(0.223)	(1.025)
	전기(자영자) (=1)	2.316***	1.993***	5.446***	7.11***	1.104**
		(0.186)	(0.172)	(0.230)	(0.136)	(0.438)
	전기(실업) (=1)	2.226***	1.979***	0.109	1.849***	2.208***
		(0.173)	(0.206)	(1.031)	(0.514)	(0.277)
상수		-7.415***	-6.122***	-13.89***	-9.962***	-8.558***
		(0.315)	(0.300)	(1.112)	(0.555)	(0.504)
log pseudo likelihood		-62,648.587				
Wald chi2		20925.16				
Pseudo R2		0.5856				
Obs.		43,105				

주: 1) ( )는 표준 오차임
2) ***, **, *는 각각 1%, 5%, 10%에서 통계적으로 유의함을 의미함
자료: 필자 작성

한편, 기존 근로자의 경제활동 상태 변화에 대한 시뮬레이션은 2단계로 이루어진다. 1단계에서는 전기(t-1)기의 경제활동 상태를 기준으로 이번 기(t)의 전이확률이 상대적으로 큰 개인부터 경제활동 상태를 부여한다. 부여되는 경제활동 상태의 순서는 '상용→임시/일용→고용주→자영자→실업→비경활' 순이다. 예를 들어, 전기(t-1)에 상용 근로자들였던 자들 중 이번 기(t)에도 상용근로자가 될 확률이 상대적으로 높은 자들을 선별한 뒤, 선별된 자들을 제외한 남은 자들(전기에 상용근로자였던 자들) 중에서 이번 기에는 임시/일용 근로를 할 확률이 상대적으로 높은 자들을 임시/일용 근로자로 선별한다. 이러한 과정을 이러한 과정을 '상용→임시/일용→고용주→자영자→실업→비경활' 순으로 반복하는 것이다. 이때, 경제활동 상태별 인구수가 성-연령별로 나누어져 있으므로, 전술한 과정을 성-연령-경제활동 상태별로 수행한다.

2단계에서는 1단계의 시뮬레이션 결과를 이번 기(t)기의 경제활동 상태별 인구 수와 비교하여 조정한다. 거시 집계치를 이용해 산출된 이번 기의 경제활동 상태별 인구 수에 비해 너무 많은 자들이 할당되는 경우, 전기의 경제활동 상태와 관계없이 이번 기의 종사상 지위별 전이확률 추정 결과를 이용하여 상대적으로 확률이 낮은 자들을 제외한다. 반대로 결정된 인구수에 비해 적은 인원이 할당된 경제활동 상태에 대해서는 전술한 경합 과정에서 탈락한 이들을 대상으로 인원을 조정한다.

마지막으로 앞선 과정들을 통해 결정된 경제활동 상태별 인구 수의 합, 즉 집계값을 성-연령대(5세)별 경제활동 이행 비율을 통해 조정해준다. 이때, 집계값 조정에 사용하는 경제활동 상태별 이행 비율은 2010-2019년의 한국노동패널자료를 이용해 추출한다. 추출된 이행 비율의 예시는 〈표 5-15〉와 같다.

〈표 5-15〉 경제활동상태별 이행 비율(40~44세)

성별	지난 경제활동 상태 [t-1기]	상용 [t기]	임시 일용 [t기]	고용주 [t기]	자영자 [t기]	실업 [t기]	비경활 [t기]
남성	상용[t-1기]	90.3	1.9	0.2	0.8	0.9	6.0
	임시일용[t-1기]	9.3	75.9	0.5	2.4	1.9	10.0
	고용주[t-1기]	3.9	2.1	69.8	19.2	0.4	4.6
	자영자[t-1]	4.2	0.9	9.9	78.8	0.4	5.7
	실업[t-1기]	34.7	10.2	0.0	2.0	10.2	42.9
	비경활[t-1기]	4.8	5.3	0.6	1.7	1.3	86.4
여성	상용[t-1기]	95.1	0.8	1.0	1.2	0.9	1.0
	임시일용[t-1기]	6.6	83.4	0.6	3.7	1.3	4.4
	고용주[t-1기]	4.1	1.1	75.6	17.1	0.9	1.3
	자영자[t-1]	4.3	1.8	12.6	79.9	0.6	0.9
	실업[t-1기]	32.9	28.8	2.7	6.8	6.8	21.9
	비경활[t-1기]	10.1	10.1	1.3	4.0	5.7	68.7

자료: 필자 작성

3. 경제활동에 대한 유효성 검토

여기에서는 경제활동에 대한 시뮬레이션 결과의 유효성을 검토한다. 이를 위해 2016~2022년의 시뮬레이션 결과를 같은 기간의 한국노동패널조사 자료에서 추출한 결과와 비교·분석한다.

먼저, 마르코프 체인 모형을 이용하여 2016~2021년의 경제활동 상태별 전이확률을 추정한다. 전이확률의 추정 방법은 다음과 같다. t기의 경제활동 상태(X_t)가 i인 개인이 $t+1$시점에 j로 전이될 확률 P_{ij}는 (식 5-4)와 같다. 이때, 과거의 경제활동 이력(X_0, X_1, ..., X_{t-1})과 현재의 경제활동 상태(X_t)가 결정된 상황에서 다음 기 경제활동 상태(X_{t+1})의 조건부 확률 분포는 현재의 경제활동 상태(X_t)에 전적으로 의존한다고 가정한다(식 5-5).

$$P_{ij} = P(X_{t+1} = j \mid X_t = i) \qquad \text{(식 5-4)}$$

$$P(X_{t+1} = j \mid X_t = i, X_{t-1} = i_{t-1}, \ldots, X_0 = i_0) = P(X_{t+1} = j \mid X_t = i) \quad \text{(식 5-5)}$$

이상의 과정을 통해 추정한 전이확률 추정 결과는 다음의 〈표 5-16〉과 같다. 이에 따르면, 한국노동패널조사를 이용해 추정한 '지난 기(t-1)와 이번 기(t)에 동일한 경제활동 상태를 유지할 확률'이 KIHASA SIM의 시뮬레이션 결과보다 다소 높지만, 그 외의 전이확률은 대체로 유사한 수준이다. 이때, 양자가 차이나는 것은 KIHASA SIM의 시뮬레이션 과정에서 경제활동인구조사의 경제활동 상태별 인구 비율을 적용하기 때문으로 보인다.

다음으로 〔그림 5-2〕는 한국노동패널과 KIHASA SIM의 경제활동 상태 배열을 보여준다.[31] 이에 따르면, 두 자료의 경제활동 상태 배열은 매우 유사하다. 다만, KIHASA SIM의 자영자 및 실업자 비율이 한국노동패널에 비해 높고, KIHASA SIM의 경제활동 상태 배열이 한국노동패널에 비해 조금 더 다양하다. 이는 동일 상태로의 전이확률이 한국노동패널에 비해 조금 낮게 추정된 〈표 5-16〉과 유사한 결과이다.

전술한 전이확률 추정 결과와 배열분석 결과는 KIHASA SIM의 개인의 경제활동 이행 과정에 대한 시뮬레이션이 안정적임을 보여준다. KIHASA SIM에서 개인의 경제활동 상태 변화가 전적으로 행태방정식 추정 결과에 의존한다는 점을 고려하면, 개인의 경제활동 상태별 전이확률과 배열이 한국노동패널조사 자료와 유사하게 나타나는 것은 고무적인 결과라 할 수 있다.

31) 이처럼 배열분석(Sequence Analysis, SA)을 이용해 MSM의 경제활동 상태의 이행 과정을 점검하는 것은 Salonen et al.(2020)이 제안한 것이다.

〈표 5-16〉 한국노동패널과 KIHASA SIM의 경제활동상태 전이확률 추정 결과 (2016~2021년)

(a) 노동패널

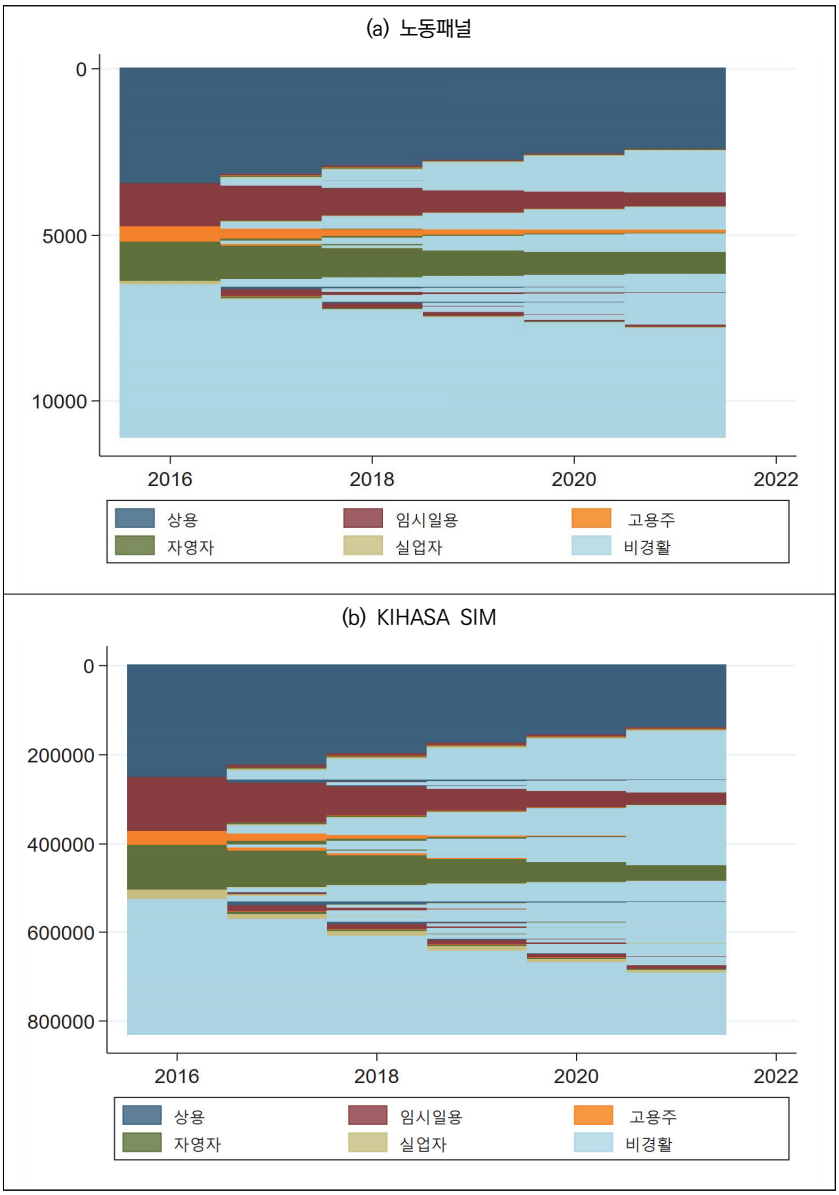
$t-1$ \ t	상용	임시일용	고용주	자영자	실업	비경활
상용	**0.859**	0.026	0.010	0.019	0.010	0.076
임시일용	0.148	**0.663**	0.012	0.026	0.017	0.133
고용주	0.073	0.019	**0.655**	0.153	0.003	0.098
자영자	0.104	0.036	0.082	**0.675**	0.005	0.099
실업	0.332	0.184	0.008	0.055	**0.115**	0.306
비경활	0.156	0.067	0.015	0.033	0.014	**0.715**

(b) KIHASA SIM

$t-1$ \ t	상용	임시일용	고용주	자영자	실업	비경활
상용	**0.811**	0.052	0.009	0.020	0.018	0.091
임시일용	0.163	**0.631**	0.012	0.039	0.024	0.132
고용주	0.101	0.043	**0.566**	0.208	0.012	0.070
자영자	0.088	0.043	0.096	**0.679**	0.011	0.083
실업	0.346	0.185	0.011	0.047	**0.105**	0.307
비경활	0.137	0.083	0.015	0.036	0.059	**0.671**

자료: 필자 작성

〔그림 5-2〕 2016~2021년의 경제활동 상태 배열



자료: 필자 작성

제4절 소득추정 모듈

1. 소득추정 모듈

근로(사업)소득은 15-22차(2012-2019년) 한국노동패널조사 자료와 1계 자기상관을 가정한 확률효과 패널 모형을 이용해 추정한다(Zaidi et al., 2009). 이때, 효율적인 추정량을 얻기 위해 Baltagi and Wu(1999)가 제시하는 GLS(generalized least squares)를 이용한다. 이때, 소득 추정에 활용하는 변수들은 선행연구를 참조하되, 한국노동패널자료와 기초 자료에서 활용가능한 변수들로 구성하였다. 〈표 5-17〉에 따르면, 해외의 주요 MSM들은 표준적인 임금방정식을 구성하는 성별, 연령, 학력, 고용 이력 등의 변수를 이용하고 있으며, 직업, 산업, 기업 규모 등의 임금에 영향을 미치는 제도·환경적 요인들을 포함하고 있다(고제이 외, 2016, p.115). 이에 본 연구에서도 전통적인 Mincer의 임금방정식에 근로소득에 대한 설명력이 높은 변수들을 추가한 형태로 모형을 구성한다((식 5-6) 참조).

$$\ln(w_{it}) = \beta_0 + X_{it}\gamma + u_i + \epsilon_{it} \qquad \text{(식 5-6)}$$

단, $\epsilon_{it} = \rho\epsilon_{i,t-1} + \nu_{it} \quad (j = 1,\ldots,N;\ t = 1,\ldots.t_i)$

w_{it}: i의 t년도 연간소득 (2020년 기준 불변가격)

X_{it} : 연령, 교육, 고용이력 등의 설명변수벡터

위의 (식 5-6)에서 u_i는 시간불변 개인고유 오차항를 의미하며, ϵ_{it}는 시간가변 오차항을 의미한다. 또한 $|\rho| < 1$, $v_{it} \sim i.i.d(0,\delta_v)$라고 가정한다. 한편, w_{it} 는 개인 i의 t년도 연간소득으로 월평균소득×12개월로 연간

화하여 산출한 뒤, 소비자물가지수를 이용해 2020년 기준 불변가격으로 환산한 값에 로그를 취해 사용한다. 고용이력 더미($D_{it,ej}$)는 이번 기(t) 취업자(e)를 대상으로 전기($t-1$)의 경제활동 상태가 실업(u), 비경활(n) 또는 취업자(e)인 경우를 구분하는 변수이다. 즉, 전기(t-1)와 이번 기(t)의 고용 이력을 '취업-취업', '실업-취업', '비경활-취업'으로 구분한다. 이외에도 연령, 연령 제곱, 학력 수준, 가구주 여부, 배우자 유무 및 배우자의 취업 여부, 종사상지위 등의 변수를 고려한다. 여성의 경우 8세 미만의 자녀 수의 변수를 추가한다. 소득 추정 결과는 〈표 5-18〉과 같다.

한편, 이상의 과정을 통해 추정된 소득은 각 연도별 경상가격으로 전환되어 시뮬레이션에 투입된다. 이때, 실적치가 존재하는 2016~2022년 동안에는 한국은행의 명목 임금상승률 자료를, 이후에는 기획재정부(2020)의 명목 임금상승률 전망치를 이용하여 경상가격으로 전환한다.

〈표 5-17〉 주요 MSM의 근로소득 추정 변수

Model	설명 변수	소득변수
Dynamod	나이, 성별, 혼인 상태, 자녀 수, 출생국가, 교육, 근무경력, 작년 개인의 평균 주급, 배우자 소득	주급
Dynasim	나이, 성별, 인종, 혼인 상태, 교육 수준	시급
SAGE	나이, 성별, 자격증, 직업, 산업, 자녀, 배우자의 고용 상태, 근무경력, 전일제 또는 시간제 고용 여부	월급
SESIM	나이, 나이 제곱, 교육, 자녀 수, 거주지역, 성별, 전기 고용 상태, 국적, 혼인 상태	시급
Dynacan	나이, 성별, 교육, 가구 구성, 은퇴상태, 실업률, 장애, 고용 상태, 전기의 소득 이력(여성의 경우 혼인 상태 및 자녀 유무)	주급
Lifepaths	나이, 성별, 지역, 교육, 학업 분야, 이민 상태, 학생 상태	주급
DESTINIE	성별, 교육, 근속연수, 근속연수 제곱	연봉
MIDAS	나이, 성별, 인종, 전기 소득분위	연봉
MOSART	전기 임금, 교육, 성별, 나이, 자녀, 연금 상태	연봉
PEMSIM	성별, 결혼 상태, 연령, 경제활동 참가에서 이탈한 연도, 노동시장 상태	연봉
CORSIM	연령, 전기소득, 교육, 교육×소득, 결혼 상태, 자녀 수, 현재 실업, 가장 어린 자녀의 연령, 인종, 성별	연봉

자료: 고제이 외. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 -노후 소득 보장 정책을 중심으로. 한국보건사회연구원. pp.115-116. 〈표 5-8〉. 일부 수정.

〈표 5-18〉 소득방정식 패널분석 결과(AR(1) 확률효과 GLS 모형)

구분		남성		여성	
		추정계수		추정계수	
연령		0.106***	(0.002)	0.079***	(0.002)
연령 제곱		-0.001***	(0.000)	-0.001***	(0.000)
종사상지위 (기준:임시일용)	상용근로(=1)	0.26***	(0.009)	0.366***	(0.009)
	고용주(=1)	0.374***	(0.013)	0.543***	(0.019)
	자영자(=1)	0.127***	(0.011)	0.185***	(0.013)
가구주(=1)		0.09***	(0.012)	0.027**	(0.014)
배우자(유=1)		0.183***	(0.011)	0.008	(0.015)
배우자 취업(=1)		0.13***	(0.011)	-0.044***	(0.014)
8세 미만 자녀 수					
학력 (기준: 전문대학)	중졸 이하(=1)	-0.23***	(0.015)	-0.099***	(0.019)
	고졸(=1)	0.092***	(0.014)	0.147***	(0.017)
	4년제 대졸(=1)	0.189***	(0.012)	0.265***	(0.015)
	대학원 졸(=1)	0.306***	(0.022)	0.388***	(0.032)
고용 이력 (기준: 취업-취업)	실업-취업 더미	-0.061***	(0.018)	-0.068***	(0.023)
	비경활-취업 더미	-0.039***	(0.005)	-0.052***	(0.006)
상수		14.694***	(0.048)	15.049***	(0.053)
rho_ar		0.320		0.346	
sigma_u		0.402		0.408	
sigma_e		0.281		0.298	
rho_fov		0.672		0.651	
within R-sq		0.0689		0.0689	
between R-sq		0.5417		0.5417	
overall R-sq		0.4729		0.4729	
Number of obs		37,761		24,131	
Number of groups		9,319		6,747	

주: 1) 종속 변수는 (월평균임금(원)×12개월)을 자연로그를 취한 값임
2) 연도더미를 포함하여 추정한 결과임
3) ***, **, *는 각각 1%, 5%, 10%에서 통계적으로 유의함을 의미함
자료: 필자 작성

2. 소득의 유효성 검토

여기에서는 소득의 시뮬레이션 결과에 대한 유효성을 검토한다. 이를 위해 한국노동패널 자료의 소득과 시뮬레이션으로 추정한 소득을 비교·검토하려 한다.

〈표 5-19〉는 2016~2021년 동안의 한국노동패널의 성별 평균 근로(사업)소득과 시뮬레이션 결과를 비교한 결과이다. 이에 따르면, 한국노동패널의 소득 대비 시뮬레이션의 추정치 비율은 105% 내외로, 시뮬레이션 결과가 다소 높은 수준이지만, 전반적으로 보면 유사한 수준이다. 〔그림 5-3〕~〔그림 5-8〕을 통해서도 이를 확인할 수 있다.

다음으로, 〈표 5-20〉은 성-종사상 지위별 평균 소득 수준을 보여준다. 이에 따르면, 한국노동패널과 시뮬레이션 결과는 전반적으로 유사한 수준이다. 다만, 종사상 지위별 차이가 있는데, 임금근로자의 소득은 시뮬레이션 결과가 한국노동패널에 비해 조금 더 높으며, 비임금근로자의 소득은 시뮬레이션 결과가 한국노동패널에 비해 조금 낮다. 표준편차 역시 평균과 유사한 패턴을 보인다.

〈표 5-19〉 성별 평균 근로(사업)소득 비교

(단위 : 명, 만원, %)

구분	연도	2016년			2021년		
	성별	남	여	전체	남	여	전체
KIHASA SIM	인원	297,031	218,949	515,980	305,017	230,146	535,163
	평균 근로소득	3,985	2,379	3,303	4,380	2,679	3,648
	표준편차	2,600	1,609	2,371	2,879	1,838	2,624
	최소값	39	39	39	65	55	55
	최대값	48,731	26,729	48,731	52,847	24,521	52,847
노동패널	인원	4,723	2,973	7,696	7,155	4,780	11,935
	평균 근로소득	3,769	2,255	3,184	3,875	2,340	3,260
	표준편차	2,609	1,570	2,382	4,286	1,640	3,558
	최소값	0	0	0	0	0	0
	최대값	37,585	18,792	37,585	234,146	28,098	234,146
노동패널 소득 대비 추정소득 비율		105.7	105.5	103.7	113.0	114.5	111.9

주: 2021년 불변가 기준
자료: 필자 작성

〈표 5-20〉 종사상 지위별 평균 근로(사업)소득 비교(2021년)

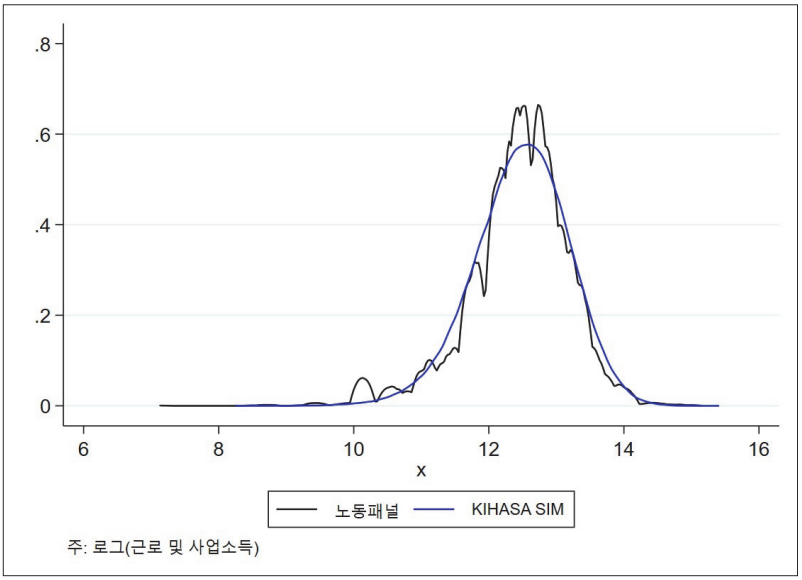
(단위 : 만원)

구분		상용직		임시일용직		고용주		자영자	
		KIHASA SIM	노동 패널	KIHASA SIM	노동 패널	KIHASA SIM	노동 패널	KIHASA SIM	노동 패널
전체	평균소득	4,075	3,652	2,303	1,781	4,646	6,008	2,856	2,675
	표준편차	2,616	1,929	1,655	1,248	2,954	7,594	2,097	2,607
남성	평균소득	4,708	4,223	2,933	2,255	5,100	6,558	3,415	2,985
	표준편차	2,830	2,045	1,915	1,420	3,083	8,495	2,314	2,814
여성	평균소득	3,089	2,745	1,808	1,388	3,364	4,308	2,091	1,955
	표준편차	1,853	1,284	1,206	916	2,074	3,009	1,440	1,860

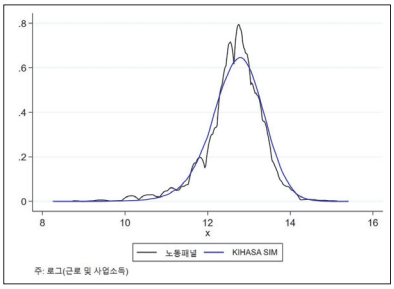
주: 2021년 불변가 기준
자료: 필자 작성

[그림 5-3] 2016년의 근로 및 사업소득 분포 비교

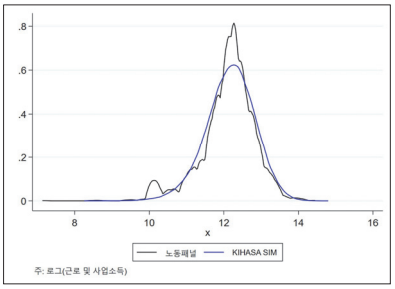
(a) 전체 근로자



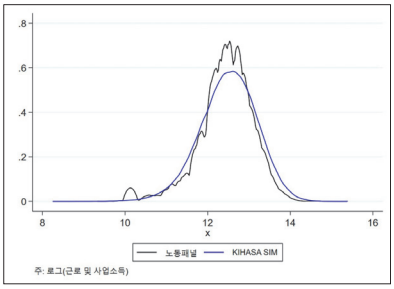
(b) 남성



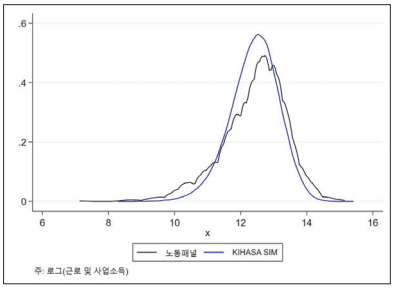
(c) 여성



(d) 임금근로자



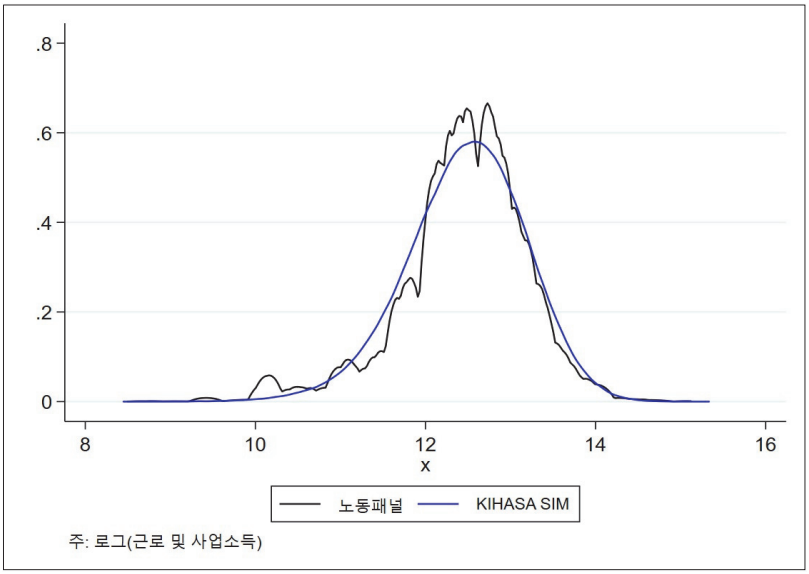
(e) 비임금근로자



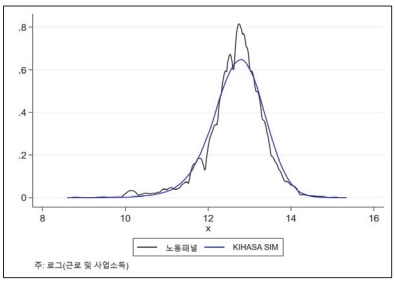
자료: 필자 작성

〔그림 5-4〕 2017년의 근로 및 사업소득 분포 비교

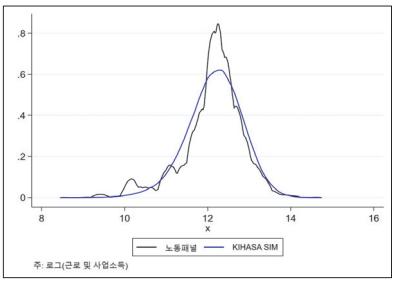
(a) 전체 근로자



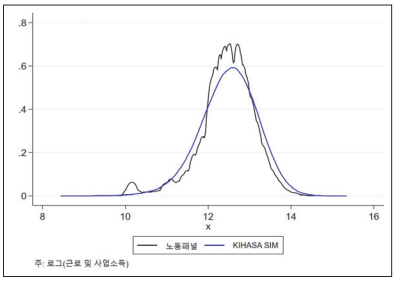
(b) 남성



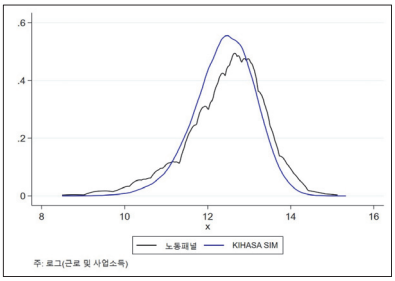
(c) 여성



(d) 임금근로자



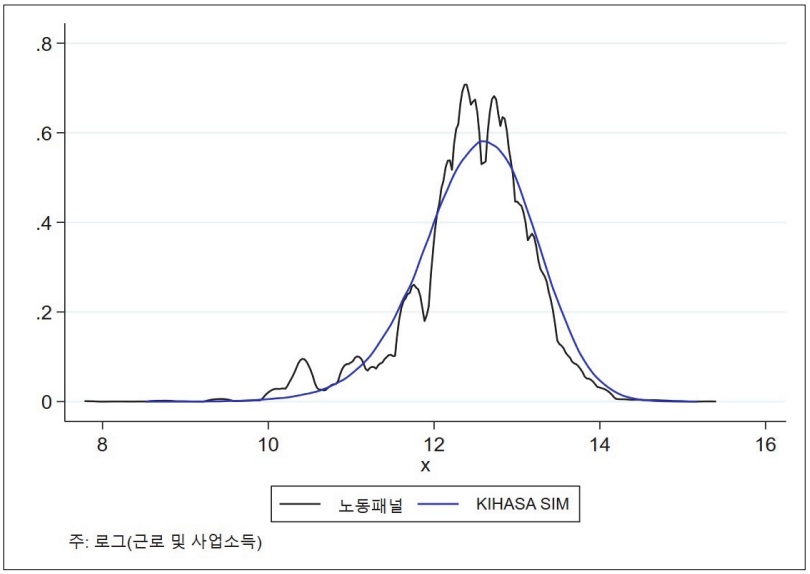
(e) 비임금근로자



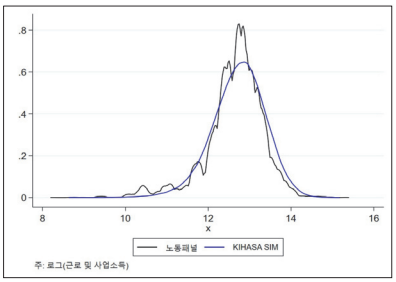
자료: 필자 작성

〔그림 5-5〕 2018년의 근로 및 사업소득 분포 비교

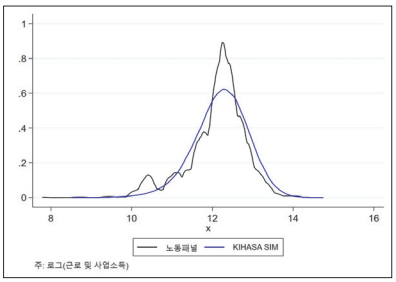
(a) 전체 근로자



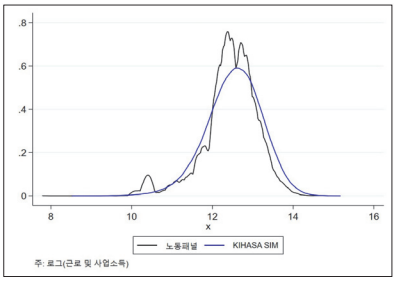
(b) 남성



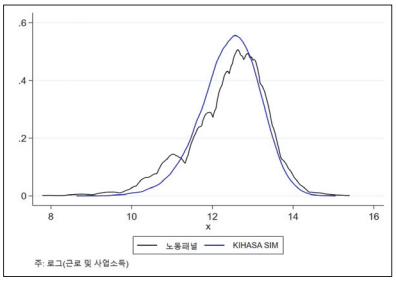
(c) 여성



(d) 임금근로자



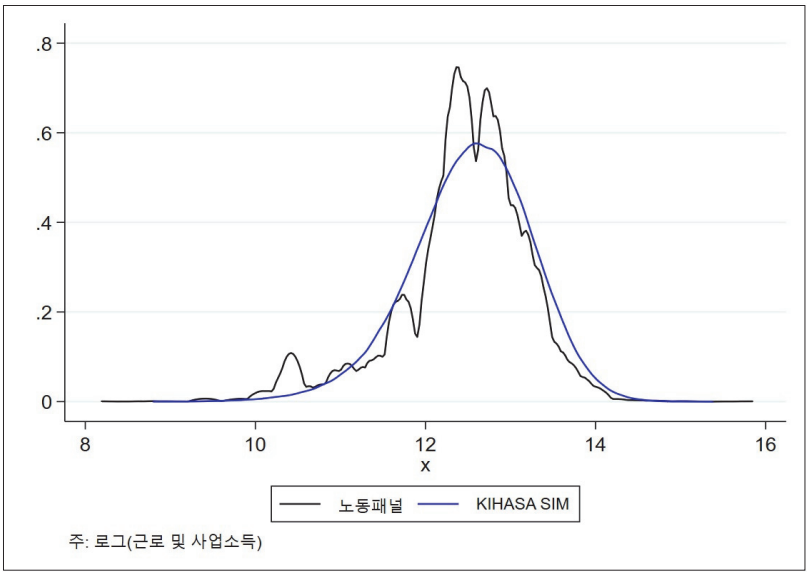
(e) 비임금근로자



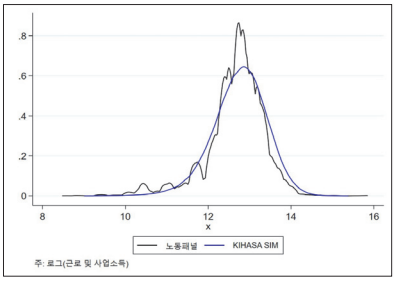
자료: 필자 작성

〔그림 5-6〕 2019년의 근로 및 사업소득 분포 비교

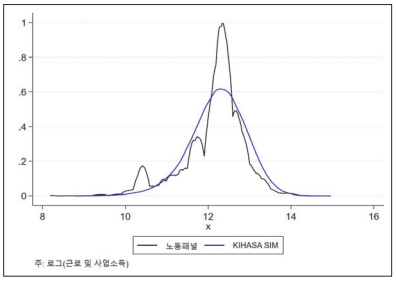
(a) 전체 근로자



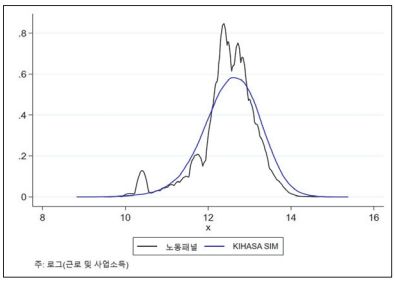
(b) 남성



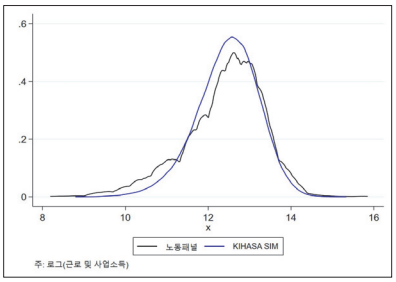
(c) 여성



(d) 임금근로자



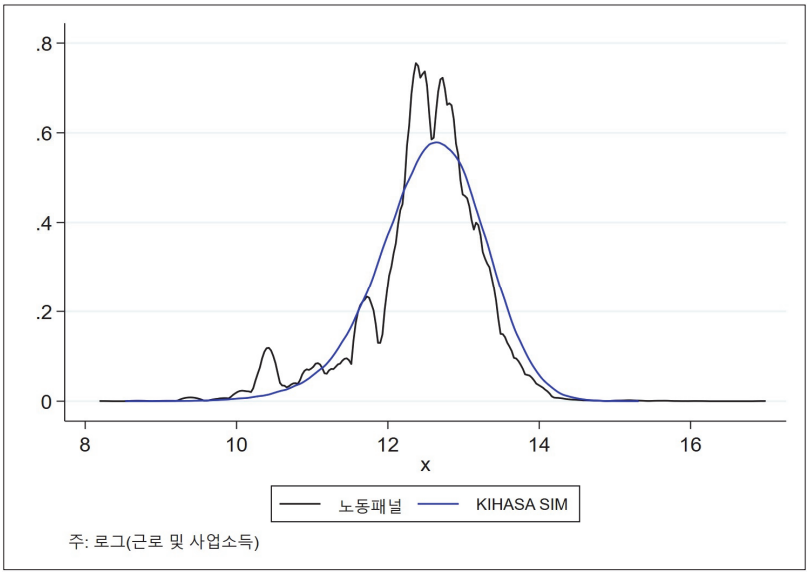
(e) 비임금근로자



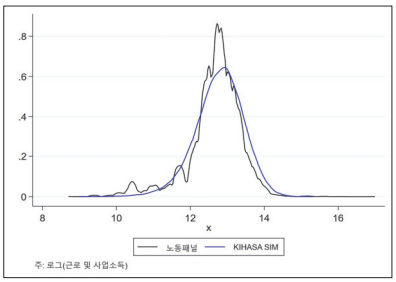
자료: 필자 작성

〔그림 5-7〕 2020년의 근로 및 사업소득 분포 비교

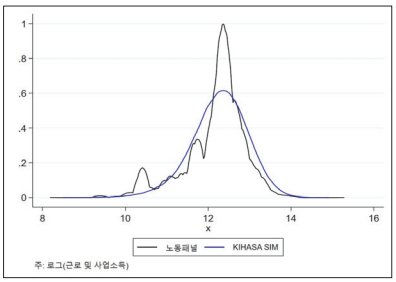
(a) 전체 근로자



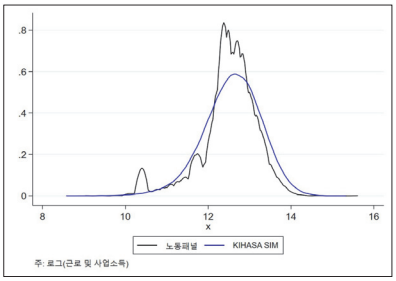
(b) 남성



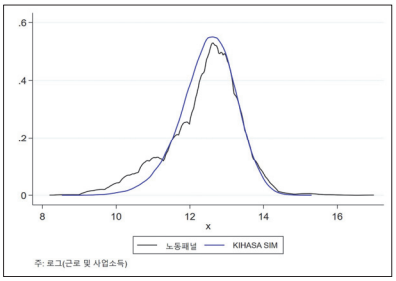
(c) 여성



(d) 임금근로자



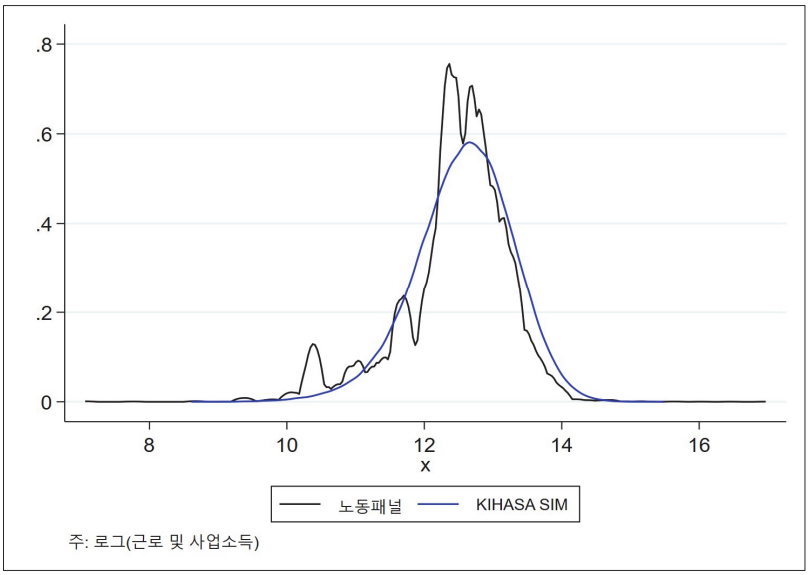
(e) 비임금근로자



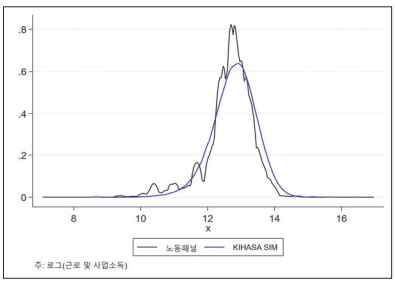
자료: 필자 작성

〔그림 5-8〕 2021년의 근로 및 사업소득 분포 비교

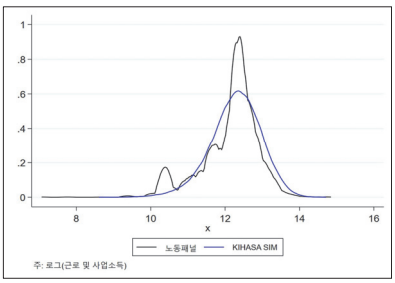
(a) 전체 근로자



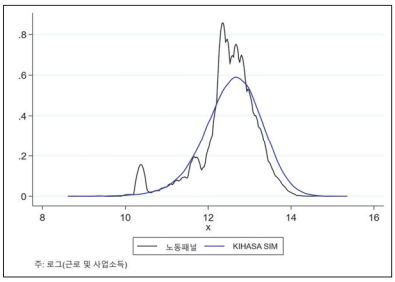
(b) 남성



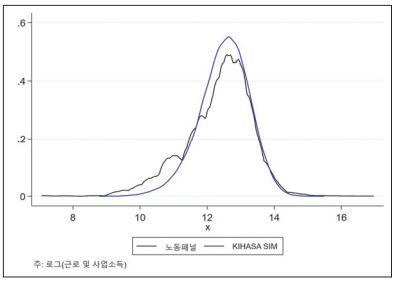
(c) 여성



(d) 임금근로자



(e) 비임금근로자



자료: 필자 작성

제5절 공적연금 모듈

1. 국민연금 모듈

가. 기초자료 및 기초율 가정

국민연금 모듈에서는 국민연금의 가입과 수급 과정을 시뮬레이션한다. 국민연금에 대한 시뮬레이션은 다양한 기초 자료와 가정들을 필요로 한다. 여기에서는 주요 자료와 가정들을 먼저 살펴본다.

먼저, 국민연금 제도와 관련된 기초율과 거시경제 변수의 경우, 실적치가 존재하는 2022년까지는 통계청, 한국은행 등을 통해 얻은 거시경제변수 실적치와 경제활동인구조사에서 산출한 국민연금 가입률 등을 활용하며, 이후의 기간에 대해서는 제4차 재정계산과 제5차 재정계산의 기초비율과 기획재정부(2020)의 주요 거시경제 변수 전망치(경제활동인구, 소비자물가 상승률, 명목임금 상승률, 이자율 등)를 활용한다(〈표 5-21〉 참조).

〈표 5-21〉 국민연금 모듈에 활용되는 기초 자료

구분		DOSA	KIHASA SIM
국민연금 가입과 수급	alignment set	• 제3차 재정계산의 거시·제도 기초변수	• 국민연금 관련 변수: 경제활동인구 조사(~2022년), 제4차 재정계산의 제도 기초변수(2023년 이후) • 거시경제 변수: 기획재정부(2020) 거시 기초 변수
	parameter set	• 국민연금 가입 확률 셋: 경제활동인구 8월 근로형태별 부가조사(2007~2011년)	• 국민연금 가입 확률 셋: 경제활동인구 8월 근로형태별 부가조사(2010~2022년)

자료: 필자 작성

보다 구체적으로 국민연금의 가입과 관련된 자료 및 가정을 살펴보면, 먼저 국민연금 가입자 규모는 2022년까지는 경제활동인구조사의 성-연령-종사상 지위별 가입 비율 적용해 산출한다.[32] 또한 이후의 기간에 대해서는 제5차 국민연금 재정계산에서 설정한 가입률 가정에 경제활동인구조사의 성-연령-종사상 지위별 가입 비율을 적용하여 가입자 규모를 산출한다. 이는 모형의 경제활동인구 대비 국민연금 가입률이 제5차 재정계산의 전망치와 동일하도록 조정하는 것을 의미한다.[33]

이때 한가지 유의할 점은 〈표 5-22〉의 국민연금 가입률이 자격 취득자 기준의 국민연금 가입률을 의미한다는 점이다. 즉, 2행의 국민연금 가입률은 실제로 보험료를 납부하지 않는 납부예외자, 미납자 등을 포함한 가입률이다. 여기에 납부예외자와 징수율을 적용하여 계산한 '실제 보험료를 납부하는 가입자'의 비율, 즉 '보험료 납부자'의 비율은 7행과 같다. KIHASA SIM은 이 '보험료 납부자'의 비율을 이용해 실질적인 가입자를 선별하며, 이들을 대상으로 보험료 산정 및 가입 기간과 관련된 시뮬레이션을 수행한다. 이때, 아직까지 제5차 재정계산에 적용된 세부적인 가정들이 공개되지 않았으므로, 여기에서는 제4차 재정계산에 적용된 성-연령별 징수율 가정을 이용한다(〈표 5-23〉 참조).

32) 이는 국민연금 공표통계에 성-연령별 경제활동인구 대비 국민연금 가입 비율에 대한 정보가 부재하기 때문이다.

33) 이처럼 거시 집계치를 조정하는 것은 성-연령-가입종별 가입 비율을 이용하여 가입자 수를 산출한 뒤 이를 상향식(bottom-up)으로 더한 결과가 전망치와 크게 달라질 수 있기 때문이다.

〈표 5-22〉 제5차 국민연금 재정계산의 가입률 가정

(단위: %)

구분	2023	2025년	2030년	2035년	2040년 이후
국민연금 가입률	92.6	93.5	93.8	93.9	94.1
지역가입자 비중	31.4	30.9	27.9	25.5	23.6
납부예외자 비율	40	39.2	38	36.3	34.3
지역가입자 징수율	72.1	73.9	75.9	76.8	77.7
보험료납부자 비율	75.2	76.7	79.0	80.7	82.2

주: 1) 국민연금 가입률은 경활인구 대비 납부예외 혹은 보험료 미징수자를 포함한 가입자 비율임.
2) 보험료 납부자 비율은 납부예외 혹은 보험료 미납자를 제외한 경활인구 대비 가입률임.
3) 보험료 납부자 비율 계산 시 사업장가입자의 징수율은 4차 재정계산의 가정인 98.6%를 적용함.

자료: 1행에서 4행은 국민연금재정계산위원회, 국민연금재정추계전문위원회, 국민연금기금운용발전전문위원회. (2023). 제5차 국민연금 재정계산 - 발표자료 -. p.7. 〈표4〉; 5행은 필자 작성.

〈표 5-23〉 지역가입자의 성별-연령별 징수율 지수

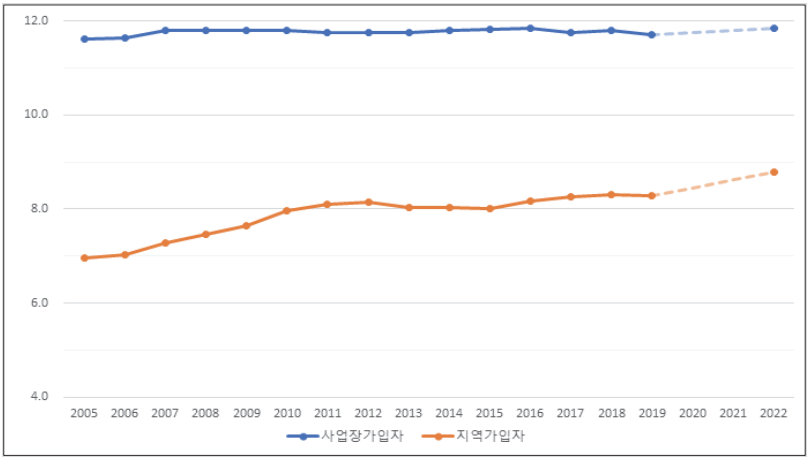
연령	남자	여자	연령	남자	여자
18~25	0.943	0.881	43	0.899	0.874
26	0.887	0.830	44	0.915	0.879
27	0.907	0.849	45	0.928	0.894
28	0.894	0.845	46	0.941	0.903
29	0.885	0.838	47	0.962	0.928
30	0.884	0.845	48	0.980	0.949
31	0.871	0.843	49	1.000	0.974
32	0.861	0.826	50	1.025	1.001
33	0.857	0.827	51	1.055	1.035
34	0.853	0.827	52	1.077	1.064
35	0.851	0.834	53	1.098	1.088
36	0.853	0.832	54	1.121	1.112
37	0.859	0.835	55	1.143	1.138
38	0.860	0.836	56	1.169	1.164
39	0.862	0.846	57	1.192	1.183
40	0.876	0.855	58	1.207	1.191
41	0.881	0.859	59	1.216	1.191
42	0.890	0.865	-	-	-

자료: 국민연금재정추계위원회. (2018). 2018 국민연금재정계산: 국민연금 장기재정추계. p.112. 〈부록 표 1-32〉.

다음으로, 가입자의 가입 기간 산정에 필요한 연간 가입 개월 수는 기존 모형인 DOSA에서 사용하였던 자료를 보정하여 사용한다. DOSA에서는 상용근로자에게 12개월의 가입 기간을 부여하였으며, 임시·일용근로자와 비임금근로자에게는 국민노후보장패널(KReIS)과 2005-2009년의 행정 자료를 결합하여 산출한 성-연령-가입종별 가입 기간 비율에 근거하여 가입 기간을 부여하였다. 또한 이와 관련된 정보가 공개되지 않았었기에 2010년 이후로는 2009년의 수치를 연장하여 적용했었다.

그런데 〔그림 5-9〕를 보면, 최근까지 지역가입자의 징수율이 꾸준히 증가하는 추세임을 알 수 있다. 이를 고려하면, 2009년의 연간 가입 개월 수 분포를 그대로 연장하여 사용하는 경우 장기적으로 가입 개월 수가 다소 짧게 추정될 수 있다. 실제로 제4차 재정계산에서는 2088년 노령연금 신규 수급자의 평균 가입기간이 26.8년으로 추계되었으나(보건복지부, 2018), DOSA에서는 약 23.4년 수준으로 추계되었다(류재린 외, 2022). 한편으로 2009년의 연간 가입 개월 수 분포가 향후에도 지속된다고 가정하는 것은 2035년까지 지역가입자 징수율이 꾸준히 증가하여 실질적인 보험료 납부자의 비율이 78.3%에 이를 것이라는 가정과도 맞지 않다. 이에 KIHASA SIM에서는 가입종별 평균 징수율 추이를 이용하여 성-연령-가입종별 평균 가입기간을 보정한 수치를 적용한다(〈표 3-24〉 참조). 이때, 2022년까지는 실적치를 이용해 연평균 가입 기간을 보정하며, 그 이후에는 제4차 재정계산의 징수율 가정을 이용한다.

[그림 5-9] 가입종별 징수율 추이



주: 징수업무가 건강보험공단으로 이관됨에 따라 2020년과 2021년의 징수율 정보는 공개되지 않았음.
자료: 국민연금 통계연보 각 연도.

〈표 5-24〉 성-연령-가입종별 가입개월 수 분포(50-55세 남성)

년도	성별	연령	가입종별	1개월	2개월	3개월	4개월	5개월	6개월	7개월	8개월	9개월	10개월	11개월	12개월
2009	남자	50	지역	0.011	0.020	0.030	0.040	0.051	0.061	0.074	0.087	0.101	0.124	0.145	1
2009	남자	51	지역	0.009	0.018	0.027	0.036	0.045	0.055	0.067	0.077	0.090	0.111	0.131	1
2009	남자	52	지역	0.009	0.017	0.025	0.034	0.043	0.053	0.064	0.075	0.087	0.108	0.128	1
2009	남자	53	지역	0.009	0.016	0.024	0.034	0.043	0.053	0.063	0.074	0.087	0.108	0.128	1
2009	남자	54	지역	0.008	0.015	0.022	0.031	0.039	0.048	0.059	0.069	0.081	0.101	0.124	1
2009	남자	55	지역	0.007	0.013	0.019	0.027	0.035	0.043	0.052	0.062	0.073	0.094	0.115	1
2020	남자	50	지역	0.033	0.047	0.059	0.071	0.084	0.098	0.112	0.125	0.143	0.163	0.199	1
2020	남자	51	지역	0.030	0.044	0.056	0.066	0.077	0.092	0.105	0.118	0.135	0.155	0.190	1
2020	남자	52	지역	0.029	0.042	0.053	0.062	0.074	0.087	0.100	0.112	0.128	0.147	0.181	1
2020	남자	53	지역	0.027	0.039	0.049	0.058	0.069	0.081	0.094	0.105	0.121	0.140	0.175	1
2020	남자	54	지역	0.025	0.036	0.045	0.054	0.065	0.077	0.089	0.100	0.115	0.134	0.170	1
2020	남자	55	지역	0.022	0.032	0.041	0.049	0.059	0.069	0.080	0.092	0.106	0.125	0.162	1

자료: 필자 작성

마지막으로, 국민연금의 급여 산정에 사용하는 자료는 다음과 같다. 조기노령연금 수급자의 수는 성-연령-가입자 유형(가입자, 대기자)별 조기노령연금 수급자 비율을 이용해 산출한다. 이때, 조기수급률 관련 실적치는 공개되고 있지 않기 때문에, 2017년부터는 제4차 재정계산의 수급률 전망치를 이용하여 조기노령연금 수급자를 선별하며, 2016년은 2017년의 비율을 적용하여 조기노령연금 수급자를 선별한다.[34] 장애연금 역시 관련 정보가 공개되지 않고 있으므로, 제4차 재정계산에서 공개한 성-연령(5세)-장애등급별 장애 발생률을 이용한다.[35]

나. 가입 단계

국민연금 가입에 대한 시뮬레이션 과정은 성-연령-가입종별 가입자 규모를 산출한 뒤, 국민연금 가입확률을 이용해 가입자를 선별하는 것을 골자로 한다.

먼저, 국민연금의 가입종별 가입자 규모는 앞서 설명한 자료들을 통해 결정된다. 이후의 가입자 선별 과정은 국민연금 가입확률에 따라 이루어진다. 이때, 국민연금 가입확률 계산에 사용되는 모수는 2016~2022년의 경제활동인구조사를 이용한 로짓 분석을 이용해 추정한다. 이에 사용되는 변수는 성별, 연령, 교육 수준, 배우자 유무, 평균 임금, 종사상 지위 등이다(〈표 5-25〉 참조). 이는 국민연금 가입 결정요인을 분석하였던 권혁진(2012), 김영옥 외(2011), 류재린과 오유진(2020), 고제이 외(2016) 등이 고려하였던 변수들 중 기초 자료에 존재하는 변수들을 추린 것이다. 국민

34) 2013년부터 2022년까지 실적치가 아닌 4차 재정계산의 수치를 이용하는 것은 최근 실적치에 대해 공개된 정보가 없기 때문이다.

35) 조기수급률에 대해서는 국민연금재정추계위원회(2018, pp.118-119)의 〈부록 표 1-11〉을, 장애발생률에 관해서는 국민연금재정추계위원회(2018, p.120) 의 〈부록 표 1-13〉과 〈부록 표 1-14〉를 참조하기 바란다.

연금 가입확률 추정 모형은 아래의 (식 5-7)과 같으며, 가입확률 추정 결과는 〈표 5-26〉과 같다.

〈표 5-25〉 국민연금 가입확률 추정에 사용된 변수의 구성 및 특징

변수 구분	변수	처리내용
종속 변수	국민연금 가입 여부	미가입 =0, 가입=1
개인적 요인	성별	여성=0, 남성=1
	연령	만18-60세
	연령 제곱	
	교육 수준	중졸/고졸/(초)대졸/석사 이상
	배우자 유무	미혼=1 유배우자=2, 이혼,사별=3
경제적 요인	평균 임금	지난 3개월간 평균임금 (단위: 만원)
노동 요인	종사상지위	상용근로자=1, 임시일용근로자=2, 자영(고용주, 자영자, 무급가족)=3

자료: 필자 작성

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 w_i + \beta_4 marr_i + \sum_{j=1}^{k_1} \beta_{5j} D_{i,sj} + \sum_{j=1}^{k_2} \beta_6 D_{i,ej} \quad \text{(식 5-7)}$$

단, p_i : i의 국민연금 가입 확률, w_i : i의 평균 임금
$D_{i,sj}$: i의 교육 더미(j=1,...,k_1), $D_{i,ej}$: i의 교육 더미(j=1,...,k_2)
age_i : i의 연령(만나이), $marr_i$: i의 혼인 상태

〈표 5-26〉 국민연금 가입확률 추정 결과

	2018년	2019년	2020년	2021년	2022년
성별	0.080	0.128*	-0.022	-0.021	-0.013
	(0.051)	(0.052)	(0.056)	(0.057)	(0.059)
연령	0.070***	0.058***	0.091***	0.068***	0.093***
	(0.017)	(0.018)	(0.018)	(0.019)	(0.020)
연령제곱	-0.001**	0.000	-0.001***	-0.001*	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
배우자유무	0.162**	0.102	0.081	0.148*	0.121
	(0.056)	(0.058)	(0.060)	(0.064)	(0.066)
중졸미만 더미	-0.346***	-0.517***	-0.452***	-0.595***	-0.345***
	(0.075)	(0.080)	(0.083)	(0.093)	(0.101)
(초)대졸 더미	0.339***	0.467***	0.331***	0.364***	0.339***
	(0.054)	(0.056)	(0.057)	(0.059)	(0.061)
대학원졸 더미	-0.160	-0.158	-0.263	-0.042	-0.091
	(0.175)	(0.165)	(0.180)	(0.183)	(0.181)
평균 임금	0.007***	0.006***	0.006***	0.007***	0.007***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
상용근로 더미	2.345***	2.505***	2.225***	2.223***	2.045***
	(0.164)	(0.163)	(0.175)	(0.178)	(0.185)
임시일용 더미	-1.939***	-1.793***	-2.021***	-1.911***	-1.859***
	(0.094)	(0.098)	(0.106)	(0.104)	(0.114)
고용주 더미	0.833***	0.897***	1.106***	1.153***	1.192
	(0.092)	(0.097)	(0.109)	(0.118)	(0.123)
상수항	-1.161***	-1.127***	-1.455***	-1.119**	-1.504***
	(0.327)	(0.341)	(0.354)	(0.367)	(0.373)
Obs.	27,044	26,533	25,807	24,871	24,568
Pseudo R2	0.4195	0.4229	0.4326	0.4332	0.4144

주: 1) 학력 더미는 고졸, 근로 더미는 자영이 기준임.
2) 괄호 안은 표준 오차임.
3) *** p〈0.01, ** p〈0.05, * p〈0.1
4) 2010년 이전의 가입확률 분석 결과는 권혁진 외6인(2013) 참조
자료: 필자 작성

한편, 선별된 가입자의 가입종별은 성-연령-종사상 지위별 가입자 비율에 지역가입 비율을 적용해 산출한다.[36] 이는 기존 모형인 DOSA와 크게 달라진 점인데, 기존 모형에서는 종사상 지위에 근거하여 국민연금 가입자의 가입종별을 결정하였다. 즉, 임금근로자인 가입자는 사업장 가입자로, 비임금근로자인 가입자는 지역가입자로 분류하였다. 그러나 비임금근로자 중 고용주는 사업장가입 대상이며, 실제로도 국민연금에 사업장가입자로 가입해 있는 고용주가 지역가입자로 가입해 있는 고용주보다 많다(〈표 5-27〉 참조). 또한 열악한 일자리에서 일하는 임시·일용직의 경우 사업장가입자로 가입할 수 없어 지역가입을 택하는 경우가 많다. 이에 KIHASA SIM에서는 성-연령-종사상 지위별 지역가입률을 적용하여 사업장가입자를 선별한다. 이때, 본 모형에서는 미래에 지역가입자가 감소할 것이라는 가정을 적용하고 있으므로, 가입종별 지역가입자 비율도 이에 맞추어 매년 보정하여 사용한다.

마지막으로 국민연금 가입자의 보험료는 근로(사업)소득에 기초하여 산정되며, 매년의 A값도 이에 따라 달라진다. 이에 KIHASA SIM은 A값의 실적치가 존재하는 2023년까지는 실적치를 사용하며, 이후에서는 모형에서 산출되는 A값을 이용한다. 한편, 기준소득월액 상·하한 금액 역시 2023년까지는 실적치를 이용하며, 이후에는 매년 A값 상승률을 반영하여 조정해준다. 연도별 가입 개월 수는 앞서 설명한 성-연령-가입종별 평균 가입 개월 수를 적용하여 계산한다.

36) 본 연구에서는 임의가입과 임의계속가입을 고려하지 않는다. 이는 두 가입종별은 경제활동과의 연계성이 떨어지며, 이들의 행태를 모형화할 수 있는 정보가 매우 부족하기 때문이다.

〈표 5-27〉 종사상 지위별 국민연금 가입률

연도	임금근로자						비임금근로자					
	상용			임시·일용			고용주			자영자		
	사업	지역	계	사업	지역	계	사업	지역	계	사업	지역	계
2011	98.5	0.5	98.9	21.0	15.6	36.6	42.3	38.8	81.1	11.0	52.2	63.2
2013	98.8	0.4	99.2	20.6	19.7	40.3	41.4	40.4	81.8	10.9	53.1	64.0
2015	98.7	0.4	99.1	19.9	21.6	41.5	40.5	42.8	83.3	10.8	55.8	66.6
2017	98.8	0.4	99.2	23.8	21.4	45.2	48.3	37.4	85.6	14.4	56.8	71.2
2019	98.6	0.6	99.2	24.1	22.2	46.4	52.6	36.7	89.3	12.3	61.0	73.3
2021	98.5	0.7	99.2	26.5	26.1	52.6	55.5	35.6	91.1	2.6	72.3	74.9

주: 1) 가입률은 18~59세 근로자 대비 국민연금 가입자 비율을 의미함.
2) 국민연금 가입자에는 특수직역연금 가입자가 포함되어 있음을 유의
자료: 1) 통계청,『경제활동인구조사 근로 형태별 부가조사』, 각 년도 8월.
2) 류재린, 문현경(2022). 코로나19 확산이 국민연금 가입에 미치는 영향 분석. 사회보장연구. p.36. 〈표 1〉 에서 재인용

다. 수급 단계

국민연금의 수급자의 선정과 급여액은 기초 자료와 결합된 과거의 국민연금 가입 이력과 시뮬레이션 기간 동안 누적된 가입 이력에 기초하여 결정된다. 노령연금 수급자는 법정 수급개시연령에 도달한 자들의 가입 이력과 급여종별 기준을 고려하여 선별한다. 조기연금 수급자는 앞서 설명한 방법대로 수급자 규모를 결정한 뒤 균등 확률에 따라 무작위 추출하며, 장애연금 수급자 역시 유사한 방식으로 선별한다. 유족연금은 사망 모듈에서 시뮬레이션하는데, 사망자가 국민연금 가입자이거나 수급자일 경우 유족연금 대상자로 설정한다. 한편, KIHASA SIM에서는 노령연금과 유족연금, 노령연금과 장애연금, 유족연금과 장애연금에 한해 중복급여의 조정을 시뮬레이션한다.[37][38] 노령연금과 유족연금의 경우 '노령연금 + 유

37) 이외의 중복급여를 고려하지 않는 것은 관련 기초자료가 부재하기 때문이다.

족연금의 20%'와 '유족연금' 중 큰 금액을 선택하는 것으로 가정하며, 노령연금과 장애연금은 더 큰 급여액을 선택한다고 가정한다. 유족연금과 장애연금의 중복수급권이 발생한 경우는 장애연금을 택하는 것으로 가정한다.

전술한 과정을 통해 수급자를 선별한 이후에는 급여액을 계산한다. 국민연금의 연금월액은 기본연금액((Basic Pension Amount, BPA)에 각 급여종별 수급 요건과 지급률 및 제한율을 곱한 후 부양가족 연금액을 더하여 산출된다((식 5-8) 참조). 이때, KIHASA SIM은 부양가족연금을 고려하지 않으므로, 연금월액은 기본연금액과 지급률에 따라 결정된다. 기본연금액은 "급여 계수×(A+재분배 상수×B)×가입 기간"으로 계산된다. 여기서 급여 계수는 소득대체율의 변화에 따라 매년 다르게 적용된다((식 5-9) 참조).[39]

연금월액 = 기본연금액(BPA) × 연금종별 지급률 + 부양가족 연금액
(식 5-8)

* 노령연금의 지급률 : 가입 기간 10년 50% (1년당 5% 증가)
* 장애연금의 지급률 : 장애1급 100%, 2급 80%, 3급 60%, 4급(일시금) 225%
* 유족연금의 지급률 : 가입기간 10년 미만 40%, 10년 이상 20년 미만 50%, 20년 이상 60%

38) 예를 들어 본인의 노령연금 수급권과 배우자의 사망에 따른 유족연금 수급권이 충돌할 수 있다.
39) 국민연금의 소득대체율은 1988년 국민연금제도가 도입될 당시 70% 수준이었으나, 1998년과 2007년의 연금개혁을 통해 2028년 40%까지 인하되는 중이다.

$$BPA = \underbrace{[2.4(A+0.75B)\times P_1/P}_{1988\sim1998년} + \underbrace{1.8(A+B)\times P_2/P}_{1999-2007년}$$

$$+ \underbrace{1.5(A+B)\times P_3/P}_{2008년} + \underbrace{\cdots\cdots\cdots\cdots}_{2009년-2027년} + \underbrace{1.2(A+B)\times P_{23}/P}_{2028년\ 이후}$$

$$+ \underbrace{X(A+A)\times C/P}_{출산크레딧} + \underbrace{X(A+A)\times 6/P]}_{군복무크레딧}\times(1+0.05n/12)$$

(식 5-9)

- A : 연금수급 전 3년간 전체 가입자의 평균소득월액의 평균액
- B : 가입자 개인의 가입기간 중 기준소득월액의 평균액(노령연금 산정 시에만 실업 크레딧 포함)
- P : 가입자의 전체 가입월 수 (노령연금액 산정 시에만 출산, 군복무 및 실업 크레딧을 포함한 전체 가입월 수)
- X : 1.5~1.2까지의 비례상수 중 노령연금 수급권 취득 시점의 상수
- $\cdots$: 2008년에서 2028년까지의 가입 기간에 대해서는 급여 계수가 1.5에서 매년 0.015씩 감소하여 최종적으로 2028년에 1.2를 적용.
- C : 추가가입기간 12, 30, 48, 50 (균분하는 경우에는 6, 15, 24, 25)
- n : 20년 초과월 수(노령연금액 산정시에만 출산, 군복무 및 실업 크레딧을 포함한 전체 가입월 수)

다음으로, A값은 국민연금 전체 가입자들의 기준소득월액의 3년 평균값(AW)을 아래의 (식 5-10)에 따라 소비자물가상승률(CPI)로 재평가한 금액을 의미한다. 앞서 설명한 것처럼 KIHASA SIM에서는 2023년까지는 A값 실적치를 사용하며, 2023년부터는 시뮬레이션을 통해 생성된 A값을 직접 계산하여 사용한다.

$$A_n = \frac{1}{3}\left(AW_{n-2}\times\prod_{k=1}^{2}(1+CPI_{n-k}) + AW_{n-1}(1+CPI_{n-1}) + AW_n \right) \quad (식\ 5\text{-}10)$$

마지막으로, 생애 평균소득(B값)은 국민연금 가입 기간 동안 적용된 기준소득월액(W_t)을 수급개시 전년도의 A값 기준으로 현재가치화한 금액(W_t')의 평균을 의미한다((식 5-11) 참조). 예를 들어, t기까지 국민연금에 가입하고 $t+1$기에 연금을 수급하는 가입자의 B값은 가입기간 동안의 기준소득월액(W_t)을 t시점의 A_t값을 이용해 재평가한 금액(W_t')의 평균으로 계산된다.

$$B = \frac{1}{n}(W_1' + W_2' + \cdots\cdots + W_n') = \frac{1}{n}\sum_{t=1}^{n} W_t' = \frac{1}{n}\sum_{t=1}^{n} W_t \times \frac{A_n}{A_t}$$

(식 5-11)

2. 기초연금 모듈

기초연금의 시뮬레이션은 별도의 집계값 조정 없이 이루어진다. 동 제도의 수급 대상은 만 65세 이상의 노인 중 소득인정액이 선정기준액 이하인 자이다. 이때, 특수직역연금의 수급권자 및 그 배우자는 수급 대상에서 제외된다.

기초연금액의 산정 절차는 다음과 같다. 먼저, 시뮬레이션 대상자의 소득인정액을 기준으로 65세 이상 노인 중 하위 70%를 선정한다. 이때, 소득인정액은 소득평가액과 재산환산액의 합으로 결정된다. 소득평가액은 아래의 (식 5-12)에 따라 계산되는데, KIHASA SIM에는 기타소득에 해당하는 소득이 연금소득만 존재하므로 근로소득과 연금소득만을 고려하여 소득평가액을 산정한다. 소득공제액은 2022년 기준인 103만원 적용 후, 2023년부터는 물가상승률에 연동하여 조정한다.

소득평가액 = 0.7 × (근로소득 - 소득 공제액) +기타 소득 (식 5-12)

자산의 소득환산액은 아래의 (식 5-13)에 따라 계산된다. 이때, KIHASA SIM이 가구 단위의 자산을 이용하며, 자동차와 회원권과 지역을 고려하지 않으므로, KIHASA SIM의 자산 소득환산액은 현실과 다소 달라질 수 있다. 이는 본 모형의 한계이기도 하지만, 전술한 내용은 특정 개인/가구가 아니라 모형 내 모든 개인/가구에게 적용되므로, 소득과 자산을 고려해 하위 70%를 선별하는 데에는 큰 영향이 없으리라 판단된다. 한편, 지역별 공제액은 대도시-중소도시-농어촌의 평균치인 9,750만원을 일괄 적용하며, 2023년 이후에는 물가상승률과 연동하여 조정한다. 금융자산 공제액은 2022년 기준 2,000만 원으로 가정하며, 마찬가지로 2023년부터는 물가상승률에 연동하여 조정한다. 마지막으로 부채액은 모형에서 산출한 선정기준액과 기준년도(2022년)의 실적치가 유사해지는 수준인 20%만을 적용한다.

자산의 소득 환산액= [{(일반 자산 -기본자산액)+(금융자산 -금융자산 공제액) -부채} ×자산의 소득 환산율÷12월] + P (식 5-13)

단, P: 고급 자동차 및 회원권

기본 자산액 공제: 대도시(1억 3,500만 원), 중소도시(8,500만 원), 농어촌(7,250만 원)

이상의 과정을 통해 기초연금 수급자를 선별한 뒤에는 감액(국민연금 연계감액, 부부감액, 소득역전방지 감액)을 적용하여 기초연금액을 산정한다. 먼저, 국민연금 수급자의 기초연금액은 아래의 (식 5-14)에 따라 연계 감액되어 산정된다. A급여액이 기준연금액의 150% 이하이면 기준연금액을 기초연금액으로 산정하며, A급여액이 150%를 초과하면, 감액 산식에 따라 부가연금액과 기준연금액 사이의 금액으로 기초연금액을 산정한다. 이후에는 부부감액과 소득역전방지 감액을 적용하여 최종

기초연금액을 산정한다.

(기준연금액-2/3 × A급여액) + 부가연금액 (식 5-14)

제6절 고용 보험 모듈

KIHASA SIM에서는 고용보험 모듈을 별도로 구축한다. 이는 국민연금의 가입대상 연령인 18~59세 임금근로자의 고용보험 가입확률을 추정하기 위함이다. 이는 국민연금 실업크레딧 적용에 따른 영향까지도 파악할 수 있다는 장점을 지닌다. 이때, '실업크레딧'은 고용보험 구직급여 수급자에 대하여 생애 최대 1년까지 국민연금 보험료를 일부 지원하고, 이 기간을 가입기간으로 인정하는 제도를 말한다(고용노동부 홈페이지, 2023).[40] 고용보험 모듈에서는 실업크레딧을 적용하기 위한 전 단계로 가입자의 고용보험 가입확률을 분석하게 되며, 이를 위해 통계청 경제활동인구조사 8월 근로형태별 부가조사(2004~2022년)를 활용한다(〈표 5-28〉 참조).

40) 고용노동부 홈페이지. 고용안전망 중 실업크레딧. https://www.moel.go.kr/policy/policyinfo/safety/list7.do (2023.11.13. 접속)

〈표 5-28〉 고용 보험 모듈

구분		DOSA[1)]	KIHASA SIM
고용보험 모듈	alignment set	• 고용보험 가입비율 - 통계청 경제활동인구조사 8월 근로형태별 부가조사(2008~2018년)	• 고용보험 가입비율 - 통계청 경제활동인구조사 8월 근로형태별 부가조사(2004~2022년)
	parameter set	• 고용보험 가입확률 추정: - 통계청 경제활동인구조사 8월 근로형태별 부가조사(2008~2018년)	• 고용보험 가입확률 추정: - 통계청 경제활동인구조사 8월 근로형태별 부가조사(2004~2022년)

주: 고용보험 모듈은 DOSA v1.0(고제이 외, 2016)이 아닌 DOSA v1.3(이다미, 권혁진, 2019)에서 구축되었음.
자료: 저자 작성

시뮬레이션 분석에 앞서 고용보험 가입자를 선별하기 위한 일련의 과정을 정리하면 다음과 같다. 전술한 것처럼 외부 집계치(alignment set)를 활용하여 고용보험 가입자 수를 정한 뒤에는 대상자들의 고용보험 가입확률을 산출하여 가입자를 결정하게 된다. 이때 고용보험 가입확률은 통계청 경제활동인구조사 8월 근로형태별 부가조사(2004~2022년)를 활용하여 로짓(logit) 분석을 실시하여 추정한다. 다만, 2023년부터는 집계자료가 부재하기 때문에 가장 최근인 2022년의 수치를 연장하여 가입확률을 추정한다.

고용보험 가입확률을 추정하기 위한 주요 변수는 다음의 〈표 5-29〉와 같다. 변수 설정은 본 연구에 앞서 실업크레딧 적용에 따른 영향을 분석한 이다미와 권혁진(2019)를 참조하였다. 개인적 요인으로 성별, 연령, 연령 제곱, 교육 수준, 배우자 유무를, 경제적 요인으로 로그 평균 임금을, 노동시장 요인으로 종사상 지위를 고려하였다.

〈표 5-29〉 고용보험 가입확률 추정(로짓 분석)에 사용된 변수의 설정

구분	변수	처리 내용
종속변수	고용보험 가입 여부	미가입=0, 가입=1
개인적 요인	성별	여성=0, 남성=1
	연령	만 18-59세
	연령 제곱	
	교육수준	중졸이하/고졸/(초)대졸/대학원 이상
	배우자 유무	미혼=1 유배우자=2, 이혼 또는 사별=3
경제적 요인	평균임금 (log)	지난 3개월간 평균임금(단위: 만원)
노동시장 요인	종사상 지위	상용근로자=1, 임시일용근로자=2, 자영(고용주, 자영자, 무급가족)=3

주: 성별, 종사상 지위는 교차항을 만들어서 적용함.
자료: 저자 작성

먼저, 동 자료를 활용하여 18~59세 임금근로자(정규직/비정규직)의 국민연금과 고용보험 가입률을 각각 산출하였다(〈표 5-30〉 참조). 정규직 여부와 무관하게 공통적으로 2004년 이후 최근까지의 가입률은 모두 매년 증가 추이를 나타내며, 최근으로 오면서 정규직과 비정규직 간 가입률 차이는 점차 감소하고 있다. 가장 최근인 2022년 기준으로 정규직에서는 국민연금과 고용보험 가입률이 각각 97.5%, 94.1%로 90%대 중후반을 상회하는 것에 반해, 비정규직은 72.8%, 65.5%로 정규직과 비교할 때 가입률이 저조한 것을 알 수 있다. 또한 비정규직과 정규직의 가입률 차이는 국민연금보다 고용보험에서 더 크게 나타난다.

〈표 5-30〉 18~59세 임금근로자의 국민연금과 고용보험 가입률(2004~2022년)

구분	정규직		비정규직	
	국민연금	고용보험	국민연금	고용보험
2004	73.8	62.6	40.5	39.2
2005	77.0	64.9	39.4	37.1
2006	77.4	65.9	41.8	39.6
2007	78.1	65.8	44.3	42.4
2008	83.0	67.2	53.7	42.1
2009	84.6	68.6	55.1	44.6
2010	84.6	76.9	56.1	43.6
2011	85.3	79.0	56.8	46.3
2012	87.0	80.7	59.4	48.1
2013	88.5	82.5	61.1	48.9
2014	89.1	84.0	61.1	48.7
2015	90.0	84.4	60.0	47.6
2016	90.8	86.2	61.0	48.0
2017	92.8	88.0	61.3	49.5
2018	93.7	88.8	63.1	49.2
2019	95.2	88.9	64.1	50.8
2020	96.0	90.9	67.9	53.9
2021	96.6	92.4	70.1	62.9
2022	97.5	94.1	72.8	65.5

주: 1) 가입률은 18~59세 임금근로자 대비 국민연금/고용보험 가입자 비율을 의미함.
2) 국민연금 가입자에는 특수직역연금 가입자가 포함되어 있음.
자료: 통계청. (각 년도). 경제활동인구조사 8월 근로형태별 부가조사 원자료를 이용하여 직접 산출

시뮬레이션 분석에 앞서 국민연금 가입확률을 추정하기 위한 모형은 다음의 (식 5-15)와 같다.

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + X_i\beta + \sum_{j=1}^{k_1}\beta_{5j}D_{i,sj} + \sum_{j=1}^{k_2}\beta_6 D_{i,ej} \qquad \text{(식 5-15)}$$

단, p_i : i의 고용보험 가입 확률, w_i : i의 로그 평균임금.
$D_{i,sj}$: i의 교육더미(j=1,...,k_1), $D_{i,ej}$: i의 교육더미(j=1,...,k_2), X_i: i의 기타 설명변수

위의 모형을 바탕으로 로짓분석을 실시한 결과는 다음의 〈표 3-4〉와 같다. 18~59세 임금근로자의 고용보험 가입에 있어서 가장 큰 영향을 주는 변수는 '소득'과 '종사상지위'인 것으로 나타난다. 이 두 변수의 추정계수는 다른 변수들과 비교해도 상당히 높게 나타난다. 종사상지위를 기준으로 여성 상용직 임금근로자의 고용보험 가입확률은 여성 일용직보다 높으며, 이는 남성 상용직과 비교해도 높은 편이다. 한편 연령 관련 변수들은 매 시기 유의하지 않은 것으로 나타난다. 교육수준의 경우, (초)대졸과 비교할 때 중졸 이하, 고졸, 대학원 졸업의 고용보험 가입 확률이 유의미하게 낮은 것을 알 수 있다. 특히 대학원 졸업의 고용보험 가입 가능성이 상당히 낮게 나타난다.

〈표 5-31〉 고용보험 가입확률 추정: 로짓 분석 결과

		2018년	2019년	2020년	2021년	2022년
log(소득)		0.862*** (0.061)	0.915*** (0.061)	0.939*** (0.061)	1.19*** (0.062)	1.312*** (0.065)
연령		-0.019 (0.021)	0.012 (0.020)	-0.035* (0.021)	0.001 (0.021)	0.016 (0.022)
연령제곱		0.0001 (0.0001)	-0.0003 (0.0001)	0.0003 (0.0001)	0.001 (0.8811)	-0.0002 (0.0001)
성별·종사상 지위 (기준: 여성 임시일용)	여성·상용	3.969*** (0.100)	3.624*** (0.089)	3.557*** (0.095)	2.936*** (0.095)	2.845*** (0.102)
	남성·임시일용	-0.332*** (0.073)	-0.45*** (0.076)	-0.471*** (0.078)	-0.623*** (0.078)	-0.615*** (0.082)
	남성·상용	3.549*** (0.093)	3.185*** (0.087)	3.278*** (0.095)	2.56*** (0.097)	2.374*** (0.103)
혼인상태 (기준: 미혼)	유배우자	-0.229*** (0.087)	-0.135 (0.084)	-0.26*** (0.088)	-0.118 (0.089)	-0.154 (0.096)
	사별·이혼	-0.377*** (0.128)	-0.221* (0.125)	-0.317** (0.125)	-0.03 (0.127)	-0.259* (0.135)
학력 (기준: 전문대·대졸)	중졸이하	-0.205* (0.109)	-0.035 (0.106)	-0.207* (0.114)	-0.473*** (0.115)	-0.498*** (0.123)
	고졸	-0.071 (0.068)	0.053 (0.064)	-0.01 (0.067)	-0.168** (0.066)	-0.266*** (0.071)
	대학원 졸	-1.156*** (0.172)	-0.99*** (0.154)	-0.868*** (0.169)	-0.934*** (0.175)	-1.077*** (0.188)
상수항		-4.256*** (0.412)	-5.23*** (0.422)	-4.345*** (0.424)	-5.781*** (0.416)	-6.484*** (0.456)
Pseudo R2		0.509	0.478	0.486	0.432	0.432
Log pseudolikelihood		-4,374,789	-4,793,618	-4,362,707	-4,434,482	-4,150,340
Wald chi2		4,547	5,021	4,560	3,521	3,075
Obs.		19,624	19,379	18,831	18,674	18,252

주: 1) 괄호 안은 표준오차임.
2) *** p〈0.01, ** p〈0.05, * p〈0.1
3) 2018년 이전의 고용보험 가입확률 분석 결과는 이다미, 권혁진(2019)을 참조할 것
자료: 저자 작성

다음의 〈표 5-32〉는 고용보험에 가입한 임금근로자의 고용보험 가입 개월 수를 성별, 연령별로 산출한 결과를 정리한 것이다.[41] 고용보험 가입 개월 수 분포는 조사 당시 고용보험에 가입한 임금근로자가 해당 일자리에 진입한 시기, 해당 일자리 지속 여부와 지속 기간에 응답한 결과를 활용한 것이다.[42]

성별을 막론하고 연령대가 낮을수록 연간 고용보험 가입 개월 수는 낮게 나타나며, 상용직에서 고용보험에 10~12개월 가입한 비율은 55~59세를 기준으로 남성 94.1%, 여성 91.8%로 나타난다. 종사상지위를 구분하면 임시일용직의 연간 고용보험 가입개월 수가 상용직과 비교할 때 크게 낮은 편이다. 임시일용직과 달리 상용직에서는 전 연령대에서 연중 3개월 미만 가입자가 부재한 것으로 나타난다.

한편 수급개시연령에 임박한 55~59세에서도 남녀 모두 임시일용직의 연 가입 개월 수가 현저히 낮다. 다만, 여성 임시일용직(75.9%)의 경우, 10개월 이상 가입한 비율이 남성 임시일용직(53.7%)과 비교할 때 크게 높은 것으로 나타난다. 남성 상용직은 19세 이하에서도 10개월 이상 가입 기간을 확보한 비율이 높은 반면, 여성 상용직, 남녀 임시일용직은 그 비율이 낮은 것을 확인할 수 있다. 이상의 결과들은 2014~2018년을 토대로 분석한 이다미와 권혁진(2019)과 비교할 때 수치 차이가 있긴 하나, 전반적인 비율 분포는 유사한 것을 알 수 있다.

---

41) 본 연구는 고용보험 가입과 관련하여 명확한 한계를 갖는다. 소득, 종사상 지위 외에 해당 일자리에서의 업종과 직종, 지역 등도 고용보험 가입에 영향을 줄 수 있다. 다만, 본 미시 시뮬레이션은 그 특성상 투입할 수 있는 변수들이 제한적인 한계를 갖고 있다(이다미, 권혁진, 2019, p.53).

42) 경활 부가조사는 조사 당시 고용보험 가입 유무만을 조사하기 때문에 본 연구에서 산출한 가입 개월 수 분포는 단순 추정치라는 것을 염두에 둘 필요가 있다. 그 예로, 임시직은 이미 일자리 시작 당시 고용보험에 미적용될 가능성이 존재하고, 현재 일자리의 지속성 역시 응답자의 주관적 판단에 따른 결과일 수 있음을 밝힌다(이다미, 권혁진, 2019, p.53).

〈표 5-32〉 성별·연령별 연간 고용보험 가입 개월 수(2022년)

구분		남성				여성			
		1~3 개월	4~6 개월	7~9 개월	10~12 개월	1~3 개월	4~6 개월	7~9 개월	10~12 개월
상용직	18~19세	0.0	0.0	26.7	73.3	0.0	29.9	23.5	46.5
	20~24세	0.0	13.3	16.9	69.8	0.0	12.1	17.9	70.0
	25~29세	0.0	9.8	12.0	78.3	0.0	7.5	10.4	82.1
	30~34세	0.0	5.6	5.1	89.3	0.0	6.0	6.3	87.7
	35~39세	0.0	3.6	3.7	92.8	0.0	3.7	7.4	88.9
	40~44세	0.0	2.9	3.5	93.6	0.0	3.8	7.8	88.4
	45~49세	0.0	2.7	2.6	94.7	0.0	5.4	7.7	86.9
	50~54세	0.0	2.6	2.4	95.1	0.0	4.4	6.4	89.2
	55~59세	0.0	2.6	3.3	94.1	0.0	3.0	5.3	91.8
임시일용직	18~19세	20.2	44.0	21.7	14.1	15.5	70.8	0.0	13.7
	20~24세	9.1	38.5	16.5	35.9	3.9	29.0	26.4	40.7
	25~29세	3.8	27.8	19.0	49.4	3.5	21.9	27.3	47.3
	30~34세	1.1	17.2	21.7	60.0	4.5	16.5	16.6	62.3
	35~39세	2.1	25.7	11.1	61.1	0.0	14.9	20.0	65.1
	40~44세	1.8	30.6	19.5	48.0	0.7	14.9	13.5	70.9
	45~49세	3.8	35.1	10.5	50.5	2.0	14.8	14.9	68.3
	50~54세	1.9	32.6	14.9	50.6	3.6	14.0	13.9	68.4
	55~59세	0.4	34.0	11.9	53.7	0.3	11.7	12.1	75.9

자료: 통계청. 경제활동인구조사 8월 근로형태별 부가조사 원자료.

이 같은 결과들에 기초하여 KIHASA SIM에서는 고용보험 가입에 관한 시뮬레이션을 실시한다. 고용보험 가입자 수는 임금근로자 대상으로 앞서 산출한 고용보험 가입 비율을 적용하여 추정한다. 이어서 확률적 요소(stochastic element)를 고려하여 임금근로자 개개인의 고용보험 가입 확률을 추정하고, 가입확률에 따라 가입자를 결정한 뒤 각각의 가입 개월 수를 부여하게 된다. 앞서 〈표 3-5〉에서는 연령대를 5세 단위로 제시하였으나, 실제 시뮬레이션 분석 시에는 각 세 별로 분석한다. 실측치가 존재하는

시점의 시뮬레이션은 해당 연도의 값을 적용하나, 2023년 이후 시뮬레이션은 2022년 수치를 연장하게 된다. 즉 현재 상황이 미래에도 동일하게 유지될 것이라고 가정하는 것이다(이다미, 권혁진, 2019, p.53).

고용보험의 구직급여 수급 및 국민연금의 실업크레딧에 관한 시뮬레이션을 실시할 때 별도의 외부 데이터를 활용하지 않는다. 시뮬레이션에 따라 만들어진 고용보험 가입 이력과 수급 요건에 따라 구직급여 수급과 그에 따른 실업크레딧을 적용하게 된다.

현재 고용보험법상 구직급여의 수급 요건은 이직일 이전 18개월간 피보험단위기간이 통산하여 180일 이상일 것으로 규정되어 있다.[43] 다만 시뮬레이션은 연간 단위로 이루어짐에 따라 본 연구에서는 별도로 '지난 2년간 6개월 이상' 가입 이력을 확보하는 경우에만 수급자로 정의하며, 이들의 누적 가입 기간(누적 피보험 기간)과 연령에 따른 소정급여일수에 대하여 구직급여 수급이 가능한 것으로 본다. 구직급여 수급액은 2019년까지는 전년도 평균임금의 50%, 2020년부터는 전년도 평균임금의 60%로, 이를 실업크레딧 신청 시 적용되는 기준소득월액으로 활용한다.[44] 이때 인정소득은 최대 70만원을 초과할 수 없다. 또한 실업크레딧의 지원 제한 요건[45]을 적용하여 전년도 소득이 상위 10%인 경우에는 신청 대상에서 제외한다.

---

43) 고용보험법 제40조(구직급여의 수급 요건)

44) 이직일이 2019년 10월 1일 이후인 경우부터는 '이직 전 평균임금의 60%*소정급여일수'를 적용하고 있다(고용보험 홈페이지, 구직급여액, https://www.ei.go.kr/ei/eih/eg/pb/pbPersonBnef/retrievePb0203Info.do, 2023.11.13. 접속)

45) 현행 실업크레딧은 재산세 과세표준의 합이 6억원 이하, 연간 종합소득(사업소득, 근로소득 제외)이 1,680만원 이하일 때 지원 대상이 될 수 있다.

제6장

시뮬레이션 결과

제1절 기본 시뮬레이션 결과
제2절 공적연금 개혁 효과 분석

제 6 장 시뮬레이션 결과

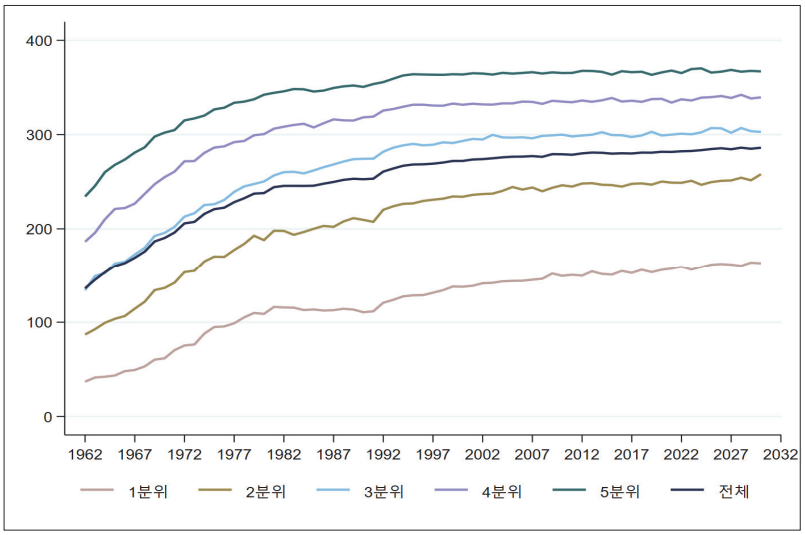
제1절 기본 시뮬레이션 결과

여기에서는 국민연금의 노후소득보장 기능과 관련된 주요 지표(가입기간, 수급률, 노령연금 급여 수준, 실질 소득대체율, 재정 수지 등)들을 중심으로 기본적인 시뮬레이션 결과를 살펴본다. 이때, 시뮬레이션 결과는 생애 소득분위를 기준으로 제시되는데, 소득분위는 생애기간(18~59세, 42년) 동안의 근로(사업)소득의 평균(즉, B값)을 기준으로 5분위로 구분한다.

먼저, 〔그림 6-1〕은 출생 코호트-생애 소득분위별 가입 기간 추이를 보여준다. 이에 따르면, 후세대로 갈수록 평균 가입 기간이 증가하여 장기적으로는 280개월 수준에 이를 것으로 전망된다. 이는 연금제도가 성숙됨에 따라 2040년 이후에는 실질적인 보험료 납부자가 경제활동인구의 82.2% 수준까지 증가할 것이라는 낙관적인 가정을 적용한 결과이다. 이처럼 가입기간이 점차 늘어남에 따라 노령연금 수급률도 장기적으로 90% 수준에 이를 것으로 전망된다(〔그림 6-2〕).

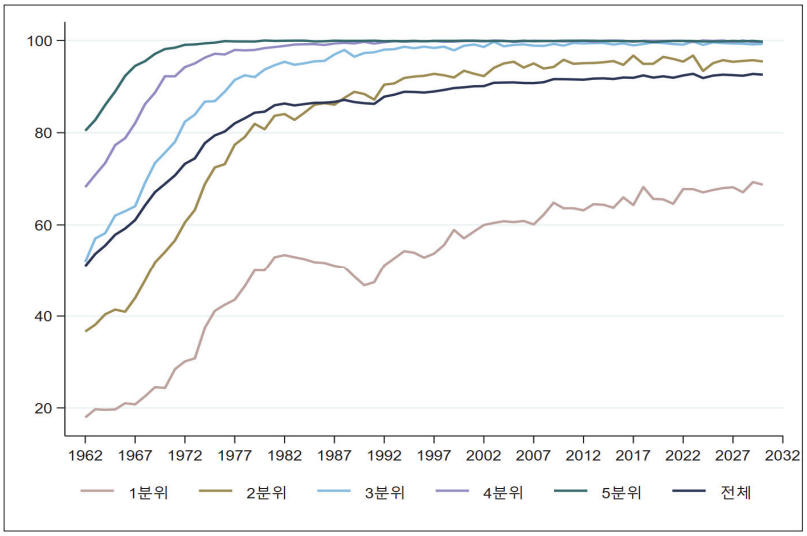
다만, 가입 기간과 노령연금 수급률 모두 생애 소득분위에 따른 격차가 존재할 것으로 전망된다. 〔그림 6-1〕을 보면, 5분위의 경우 가입 기간이 300개월을 훌쩍 넘길 것으로 전망된 반면, 1분위의 가입 기간은 그 절반 수준에도 미치지 못할 것으로 전망된다. 노령연금 수급률 역시 이와 유사하다. 상위분위의 노령연금 수급률은 100% 수준에 근접할 것으로 전망되나, 1분위는 70%에도 못 미칠 것으로 전망된다(〔그림 6-2〕). 이러한 결과는 소득 계층에 따른 가입 기간 격차가 존재한다는 최기홍(2015; 2018), 권혁진(2020), 류재린 외(2022a) 등의 주장을 재확인하는 것이다.

〔그림 6-1〕 생애 소득분위별 가입기간



자료: 필자 작성

〔그림 6-2〕 생애 소득분위별 노령연금 수급률

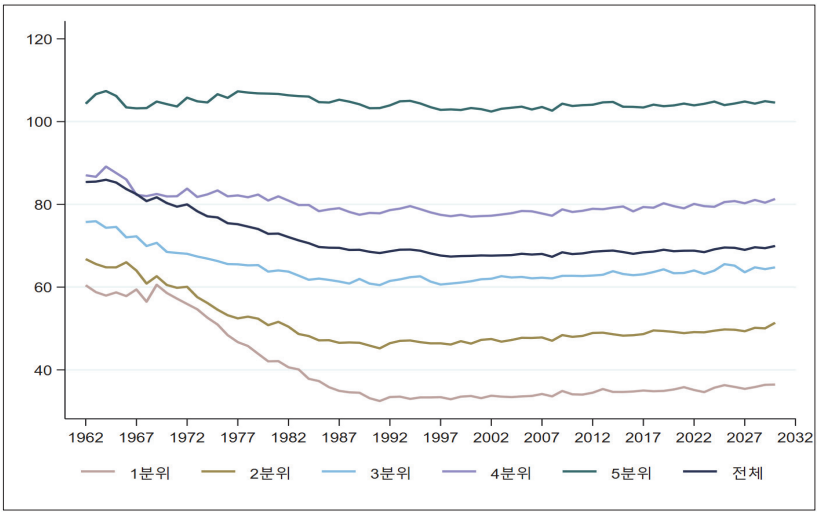


자료: 필자 작성

다음으로 신규 노령연금 수급자의 급여 수준은 후세대로 갈수록 감소하여 장기적으로 약 68만 원(2023년 불변가 기준)에 수렴할 것으로 전망된다(〔그림 6-3〕 참조). 이는 2028년까지 인하되는 소득대체율의 영향으로 풀이된다. 소득분위별로 살펴보면, 상위 분위에 비해 하위 분위의 급여 하락 폭이 더 클 것으로 전망된다. 앞서 살펴본 것처럼 하위 분위의 노령연금 수급률이 증가하여 저연금자가 많아지는 것도 하위 분위의 급여 수준을 낮추는 요인 중 하나이다.

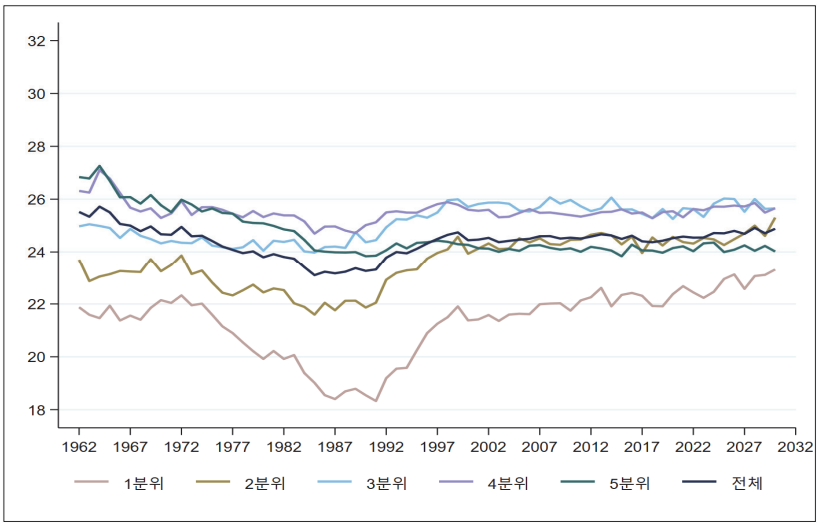
〔그림 6-4〕를 보면, 실질 소득대체율은 명목 소득대체율 인하의 영향으로 감소하지만, 후세대로 갈수록 가입 기간이 늘어나 장기적으로는 명목 소득대체율이 인하되기 이전 수준을 회복할 것으로 전망된다. 이에 따르면, 실질 소득대체율 역시 하위 분위에서 더 많이 하락하는데, 이는 하위 분위의 가입 기간이 소득대체율 인하 효과를 상쇄할 만큼 증가하지 않기 때문으로 풀이된다. 한편으로 1996년생 코호트부터는 5분위의 소득대체율이 수급자 전체의 평균 소득대체율에 비해 낮아질 것으로 전망되는데, 이는 2~5분위 사이의 가입 기간 격차가 완화되어 국민연금의 재분배 기능이 작동하기 때문으로 해석된다.

〔그림 6-3〕 생애 소득분위별 노령연금 급여 수준



자료: 필자 작성

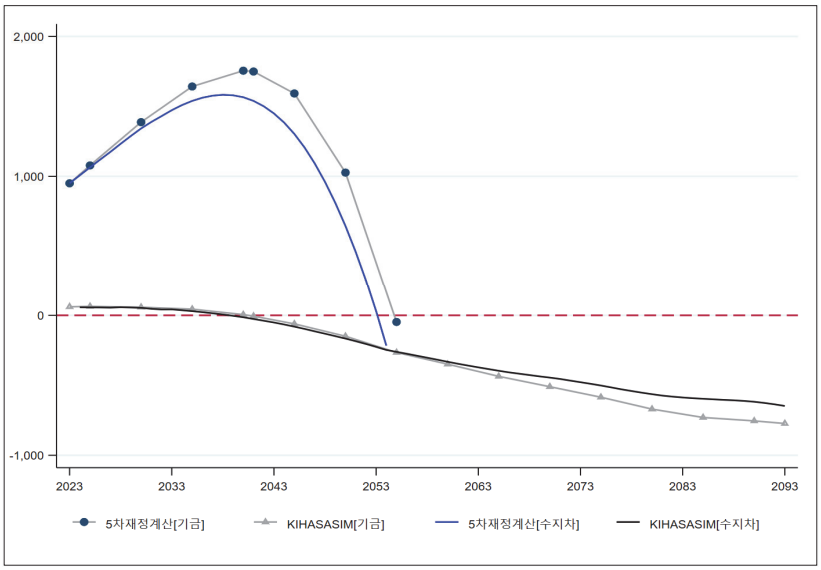
〔그림 6-4〕 생애 소득분위별 실질 소득대체율



자료: 필자 작성

마지막으로 국민연금의 재정수지 전망 결과는 〔그림 6-5〕와 같다. 이에 따르면, 국민연금의 적립 기금은 2038년 약 1,583조 원에 도달한 후, 수지 적자가 발생하는 2039년부터 급속히 감소하여 2054년에 소진될 것으로 전망된다. 기금이 소진되는 시점의 적자 규모는 약 217조 원으로 전망된다.[46]

〔그림 6-5〕 국민연금 재정수지 전망



자료: 1) 국민연금재정계산위원회, 국민연금재정추계전문위원회, 국민연금기금운용발전전문위원회. (2023). 제5차 국민연금 재정계산 - 발표자료 -. p.15. 〈표 5〉를 토대로 필자 작성
2) 필자 작성

46) 이는 제5차 재정계산의 결과와 비교할 때 다소 비관적인 추계이다. 다만, KIHASASIM과 국민연금 재정추계 모형은 모형의 특성과 구조, 주요 가정 등이 상이하므로, 양자의 결과를 직접적으로 비교하는 것은 무의미하다.

제2절 공적연금 개혁 효과 분석

1. 연금개혁 시나리오 설정

여기에서는 제5차 재정계산 보고서에서 제시한 보험료율 인상안(12%/15%/18%)과 소득대체율 인상안(45%/50%)의 조합이 연금 급여와 실질 소득대체율, 그리고 연금 재정에 어떠한 영향을 미치는지 살펴본다(국민연금 재정계산위원회. 2023). 두 모수의 조정 스케줄 역시 제 5차 재정계산 보고서의 설정을 차용한다. 먼저, 보험료율은 2025년부터 매년 0.6%p씩 인상하여 목표치에 도달하는 것으로 가정한다. 이에 따르면, 목표치인 12%에 도달하기까지는 5년, 15%에 도달하기까지는 10년, 18%에 도달하기까지는 15년이 소요된다. 소득대체율은 25년에 일시에 인상하는 것으로 가정한다.[47] 한편, 제5차 재정계산위원회는 기금운용 수익률 제고가 중요함을 크게 강조한 바 있다. 이에 본 연구도 기금운용 수익률 제고가 재정에 미치는 영향을 검토한다. 전술한 내용은 〈표 6-1〉에 정리되어 있다.

〈표 6-1〉 제5차 재정계산보고서의 정책 조합(안)

구분	소득대체율	보험료율	기금운용 수익률
S0	40%	9%, 12%, 15%, 18% (2025년부터 매년 0.6%p씩 점진적으로 인상)	제5차 재정계산의 수익률 가정/ 기금운용 수익률 1%p 상향
S1	2025년부터 45%		
S2	2025년부터 50%		

자료: 국민연금 재정계산위원회. (2023). 국민연금 재정계산 보고서. 2023.10. 국민연금 재정계산위원회를 바탕으로 필자 작성

47) 이에 대한 자세한 내용은 국민연금 재정계산위원회(2023, pp.24-39)를 참조하기 바란다.

2. 연금개혁 효과 분석

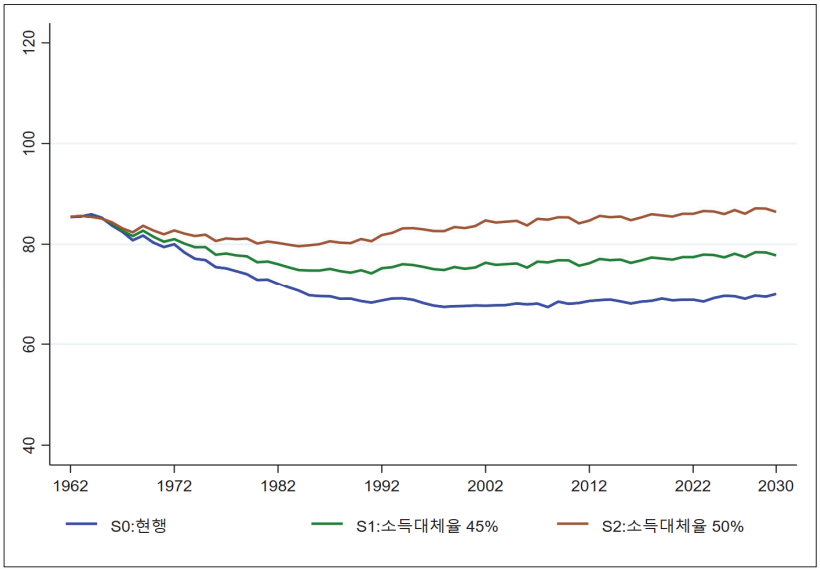
먼저, 〈표 6-2〉와 〔그림 6-6〕은 소득대체율 인상에 따른 급여액 변화를 보여준다. 이에 따르면, 현행 제도 하에서는 노령연금 급여가 서서히 감소한 뒤, 1990년대 후반 코호트부터 증가하여 장기적으로 약 70만 원 수준에 이를 것으로 전망된다. 앞서 살펴본 것처럼 이는 2028년까지 예정된 명목 소득대체율 인하의 효과로 풀이된다. 그러나 〈표 6-2〉를 보면, 명목 소득대체율을 높이는 연금 개혁을 통해 급여 수준이 크게 제고될 것으로 전망된다. 이에 따르면, 소득대체율을 45%로 인상하는 경우 장기적으로 급여 수준이 약 8.8만 원 정도 증가하며, 소득대체율을 50%로 인상하는 경우에는 급여 수준이 약 17.3만 원 증가할 전망이다.

다만, 〈표 6-2〉에 따르면, 소득 상위 분위의 급여액 증가 폭이 하위 분위에 비해 클 것으로 전망된다. 예를 들어, 2010년생의 경우 소득 5분위의 평균 노령연금액이 116.9만 원 내외로 현행제도에 비해 약 13.1만 원 정도 높아질 것으로 전망된다. 반면, 소득 1분위의 평균 노령연금액은 약 38.1만 원으로 현행에 비해 4만 원 정도 늘어날 것으로 전망된다. 명목 소득대체율을 50%로 인상하는 경우에는 소득분위에 따른 급여 격차가 더욱 심화되는데, 2010년생의 경우 소득 5분위는 노령연금액이 16.1만 원 늘어나지만, 하위 1분위는 8.3만 원 늘어나는데 그칠 것으로 전망된다. 이러한 격차는 〔그림 6-7〕을 통해서도 확인할 수 있다. 동 그래프에서 막대그래프는 현행 제도와의 급여 수준 차이를 보여준다. 이를 통해 하위 분위에 비해 상위 분위의 급여 증가 폭이 훨씬 크다는 점을 확인할 수 있다.

한편으로, 〈표 6-2〉와 〔그림 6-6〕를 통해 소득대체율 인상 효과가 나타나기까지 오랜 시간이 필요함을 알 수 있다. 이는 소득대체율을 조정하더라도, 현세대 노인의 빈곤 해결을 위한 별도의 대책이 필요함을 시사한다.

〔그림 6-6〕 출생 코호트-시나리오별 신규수급자의 노령연금 급여

(단위: 2023년 불변가, 만 원)



자료: 필자 작성

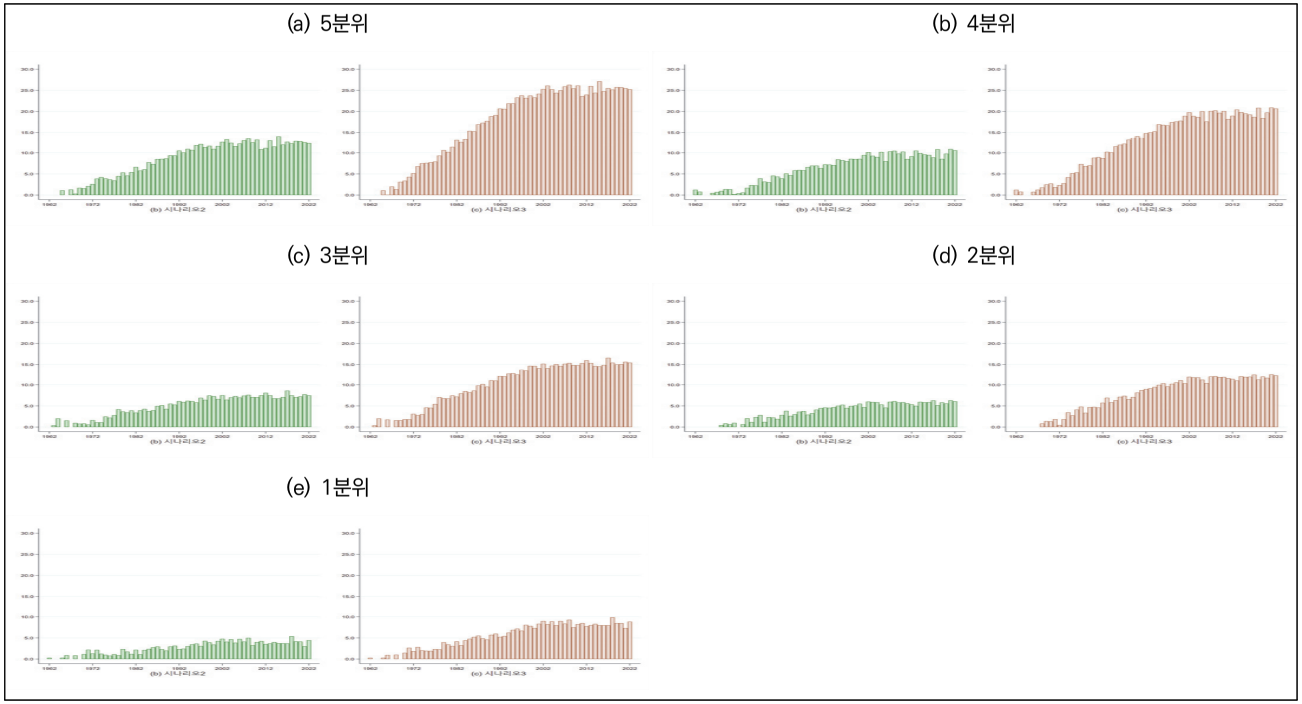
〈표 6-2〉 출생 코호트-생애 소득분위-시나리오별 노령연금 급여 변화

(단위: 2023년 불변가, 만 원)

출생 코호트	소득 분위	현행 (S0)	소득대체율 45% (S1)		소득대체율 50% (S2)	
				증감		증감
1970년생	1	58.6	59.7	1.2	60.1	1.5
	2	60.5	61.1	0.6	61.8	1.3
	3	68.5	69.3	0.8	70.3	1.8
	4	81.9	83.3	1.4	84.6	2.7
	5	104.2	105.8	1.6	107.6	3.4
	평균	**80.3**	**81.5**	**1.2**	**82.7**	**2.4**
1980년생	1	42.1	43.8	1.8	45.5	3.5
	2	50.8	53.0	2.2	55.6	4.8
	3	63.7	67.2	3.4	70.5	6.8
	4	80.9	85.5	4.6	89.8	8.8
	5	106.8	111.4	4.6	117.0	10.2
	평균	**72.9**	**76.4**	**3.5**	**80.1**	**7.3**
1990년생	1	33.2	36.1	3.0	38.9	5.7
	2	45.9	50.0	4.1	54.0	8.2
	3	60.9	66.4	5.5	71.9	11.0
	4	77.9	84.9	7.0	92.0	14.1
	5	103.2	112.6	9.4	122.1	18.8
	평균	**68.5**	**74.8**	**6.3**	**81.0**	**12.5**
2000년생	1	33.7	37.1	3.4	41.0	7.4
	2	46.3	51.9	5.5	57.4	11.1
	3	61.4	68.7	7.3	76.1	14.6
	4	77.0	85.6	8.6	94.8	17.8
	5	103.3	114.2	10.9	126.6	23.3
	평균	**67.5**	**75.1**	**7.6**	**83.2**	**15.7**
2010년생	1	34.1	38.1	4.0	42.4	8.3
	2	48.0	53.9	5.9	59.8	11.9
	3	62.7	69.8	7.1	77.5	14.8
	4	78.2	88.4	10.3	98.2	20.1
	5	103.8	116.9	13.1	129.9	26.1
	평균	**68.0**	**76.8**	**8.8**	**85.3**	**17.3**

주: 생애 소득분위는 18~69세까지 42년 간의 근로(사업)소득, 즉 B값을 기준으로 구분함.
자료: 필자 작성

[그림 6-7] 생애 소득분위-시나리오-출생코호트별 노령연금 급여 변화

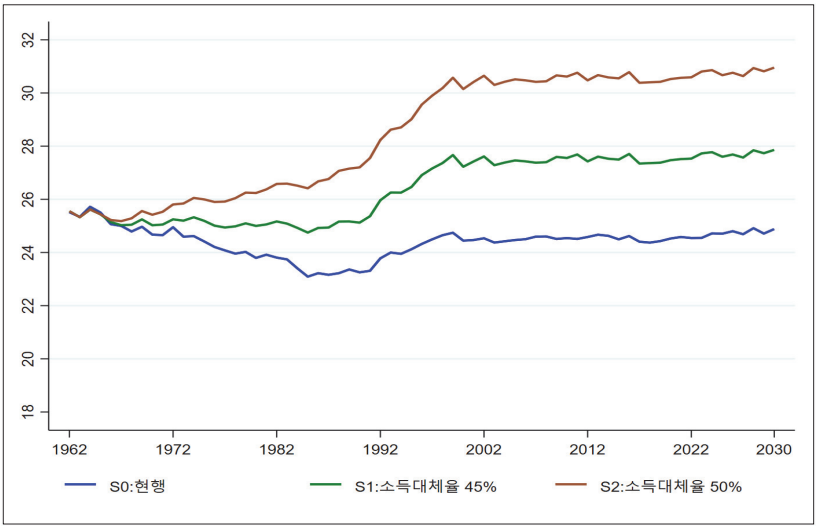


주: 생애 소득분위는 18~69세까지 42년 간의 근로(사업)소득, 즉 B값을 기준으로 구분함.
자료: 필자 작성

다음의 〔그림 6-8〕은 출생 코호트별 소득대체율 변화를 보여준다. 이에 따르면, 명목 소득대체율 인상 시 실질 소득대체율이 장기적으로 약 3~6%p 정도 높아질 것으로 전망된다. 이때, 소득대체율 인상 효과는 후세대로 갈수록 증가할 전망이다(〔그림 6-9〕 참조). 소득분위별로 살펴보면(〈표 6-3〉), 상위 분위의 소득대체율 인상 폭이 미세하게 더 크다. 이는 상위 분위의 가입 기간 증가 폭이 더 크기 때문으로 풀이된다. 다만, 후세대로 갈수록 상위 분위와 하위 분위의 소득대체율 인상 폭이 비슷해지는 것으로 전망된다. 이는 후세대로 갈수록 하위 분위의 가입 기간 증가 폭이 커지고 소득분위별 가입기간 격차가 감소하여 국민연금의 재분배 기능이 작동하기 때문으로 풀이된다.

〔그림 6-8〕 출생코호트-시나리오별 소득대체율 변화

(단위: %)



자료: 필자 작성

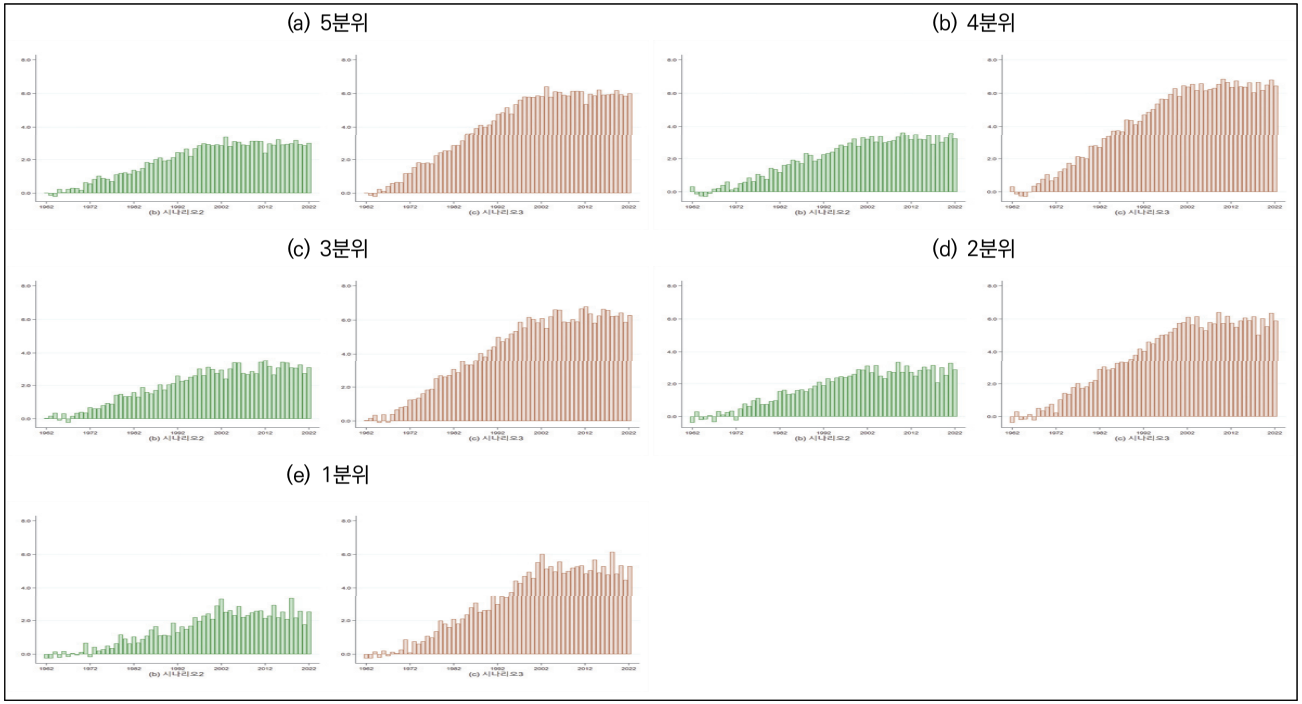
〈표 6-3〉 출생코호트-생애 소득분위-시나리오별 소득대체율 변화

(단위: 2023년 불변가, 만 원)

출생 코호트	소득 분위	현행 (S0)	소득대체율 45% (S1)		소득대체율 50% (S2)	
				증감		증감
1970년생	1	22.1	22.3	0.1	22.4	0.3
	2	23.2	23.5	0.3	23.8	0.6
	3	24.3	24.7	0.4	25.1	0.8
	4	25.3	25.9	0.6	26.3	1.1
	5	25.8	26.0	0.2	26.4	0.6
	평균	**24.7**	**25.0**	**0.4**	**25.4**	**0.7**
1980년생	1	19.9	20.8	0.9	21.7	1.8
	2	22.4	23.4	0.9	24.5	2.1
	3	24.1	25.4	1.3	26.7	2.6
	4	25.3	26.8	1.4	28.1	2.8
	5	25.1	26.3	1.2	27.6	2.5
	평균	**23.8**	**25.0**	**1.2**	**26.2**	**2.4**
1990년생	1	18.5	19.6	1.1	21.2	2.6
	2	21.9	23.7	1.9	25.7	3.8
	3	24.4	26.4	2.0	28.6	4.2
	4	25.0	26.9	1.9	29.1	4.1
	5	23.8	25.8	2.0	28.0	4.1
	평균	**23.3**	**25.1**	**1.9**	**27.2**	**3.9**
2000년생	1	21.4	23.5	2.1	26.0	4.6
	2	23.9	26.8	2.9	29.7	5.8
	3	25.7	28.7	3.0	31.8	6.1
	4	25.6	28.4	2.8	31.4	5.8
	5	24.3	27.1	2.9	30.1	5.8
	평균	**24.4**	**27.2**	**2.8**	**30.1**	**5.7**
2010년생	1	21.7	24.3	2.6	27.0	5.3
	2	24.5	27.2	2.7	30.2	5.7
	3	26.0	28.7	2.7	31.9	5.9
	4	25.4	29.0	3.6	32.2	6.8
	5	24.1	27.3	3.1	30.3	6.2
	평균	**24.5**	**27.6**	**3.0**	**30.6**	**6.1**

주: 생애 소득분위는 18~69세까지 42년 간의 근로(사업)소득, 즉 B값을 기준으로 구분함.
자료: 필자 작성

[그림 6-9] 생애 소득분위-시나리오-출생코호트별 소득대체율 변화

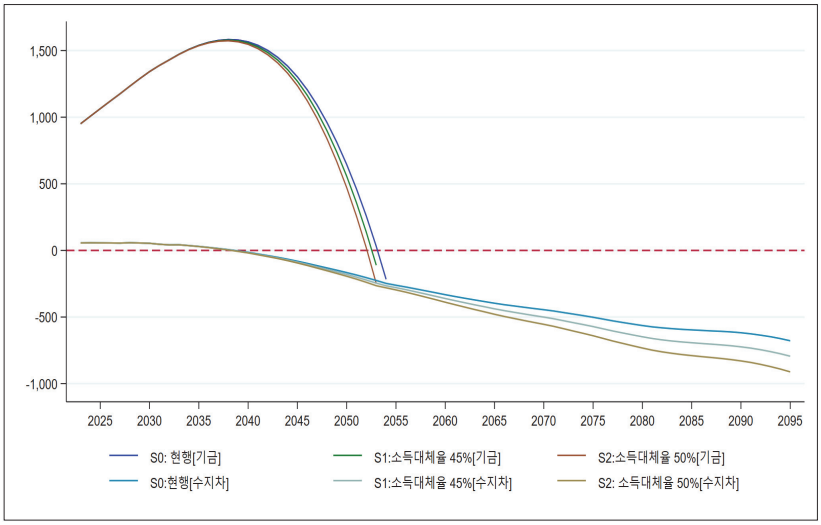


주: 생애 소득분위는 18~69세까지 42년 간의 근로(사업)소득, 즉 B값을 기준으로 구분함
자료: 필자 작성

마지막으로 〔그림 6-10〕은 시나리오별 국민연금 재정수지 전망을 보여준다. 이미 잘 알려져 있듯이, 소득대체율 인상은 재정적 지속가능성을 낮추는 요인이므로, 대체율이 높아질수록 연금 재정이 악화되는 것을 확인할 수 있다. 〈표 6-4〉에 따르면, 소득대체율이 인상되는 경우의 수지 적자 시점은 현행과 동일하지만, 기금소진 시점은 현행에 비해 각각 1년씩 앞당겨질 것으로 전망된다. 이처럼 소득대체율 인상이 기금에 미치는 영향이 미미한 것은 소득대체율 인상 효과가 장기에 걸쳐 서서히 나타나기 때문이다. 기금 소진 이후의 수지 적자 추이를 보면, 시간이 지날수록 소득대체율의 인상이 야기하는 재정부담이 점차 커지는 것을 알 수 있다.

〔그림 6-10〕 소득대체율 인상에 따른 재정 전망

(단위: 경상가, 조 원)



자료: 필자 작성

다음으로 〔그림 6-11〕은 보험료율 인상에 따른 재정 전망 결과를 보여준다.[48] 동 그림을 통해 보험료율 인상 시 기금소진 시기가 크게 늦추어지는 것을 확인할 수 있다. 〈표 6-4〉에 따르면, 보험료율을 12%로 인상하면 기금소진 시점이 4년 늦춰지며, 보험료율을 18%까지 인상하면 기금소진 시점이 21년 연장되는 것으로 전망된다. 다만, 보험료율을 18%로 인상하더라도 제5차 재정계산에서 설정한 재정 목표('추계기간 내 적립금 미소진)를 달성하지 못한다는 점에서, 보험료율 인상은 연금의 장기적 지속가능성을 확보하는 근본적인 대책이 아닐 수 있다.

한편, 〔그림 6-12〕와 〈표 6-4〉를 보면, 소득대체율과 보험료율을 동시에 인상하는 경우, 모든 정책조합이 보험료율 9% 유지하는 현행 제도에 비해 좋은 결과는 낼 것으로 전망된다. 디민. 보험료율만 높이는 방안보다는 재정안정성이 떨어지므로, 재정 목표 달성은 어려워 보인다. 이는 두 모수 간 상충관계가 작동하기 때문이다. 〈표 6-4〉에 따르면, 보험료율을 18%까지 인상하는 동시에 소득대체율을 50%로 인상하는 경우, 2071년에 기금이 소진되는 것으로 나타났다. 이는 보험료율만 18%로 인상하는 경우에 비해 4년 앞당겨진 것이다.

그런데 기금운용 수익률을 1%p 높이는 경우(〈표 6-5〉와 〔그림 6-13〕), 재정안정성이 크게 제고될 것으로 전망된다. 다만, 여전히 재정 목표 달성은 어려울 것으로 전망되는데, '보험료율 18% 인상-기금운용 수익률 1%p 상향' 조합이 재정 목표를 달성하는 유일한 정책 조합이다. 이외에 재정 목표 달성에 가장 근접한 방안은 '소득대체율 45% - 보험료율 18%'의 조합이다. 이러한 추계 결과는 모수 개혁이 근본적인 해결책이 될 수 없음을 시사한다.

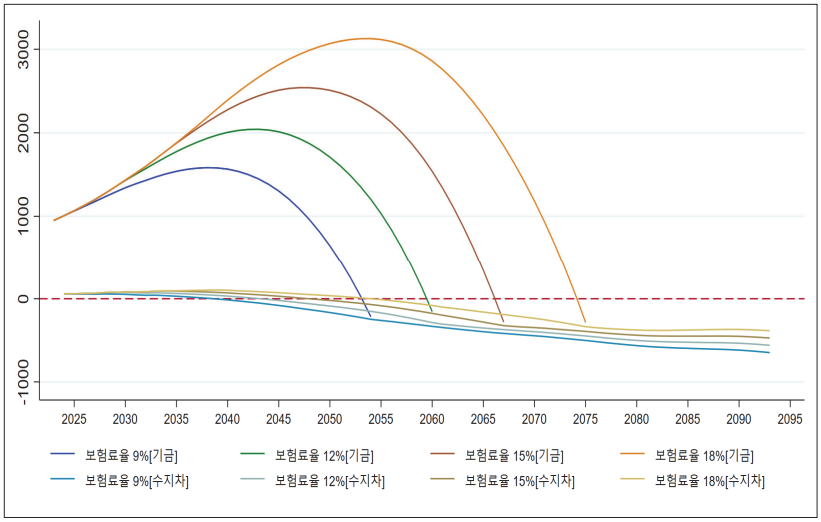
---

48) 본 연구의 추계 결과는 제5차 재정계산 결과에 비해 다소 비관적이다. 다만, KIHASA SIM과 국민연금 재정추계 모형은 모형 구조와 원리, 시뮬레이션 방법, 주요 가정 등이 상이하므로, 추계 결과에 대한 직접적인 비교가 무의미하다.

〔그림 6-11〕 보험료율 인상에 따른 재정 전망

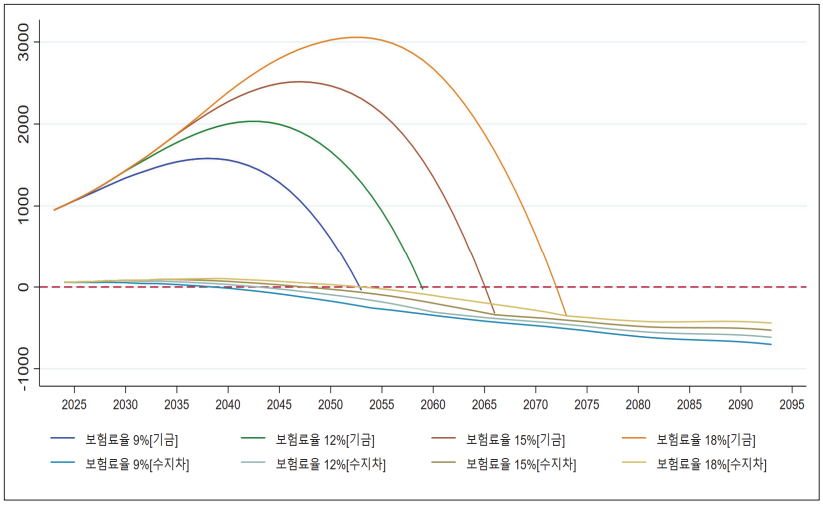
(a) 현행

(단위: 경상가, 조 원)



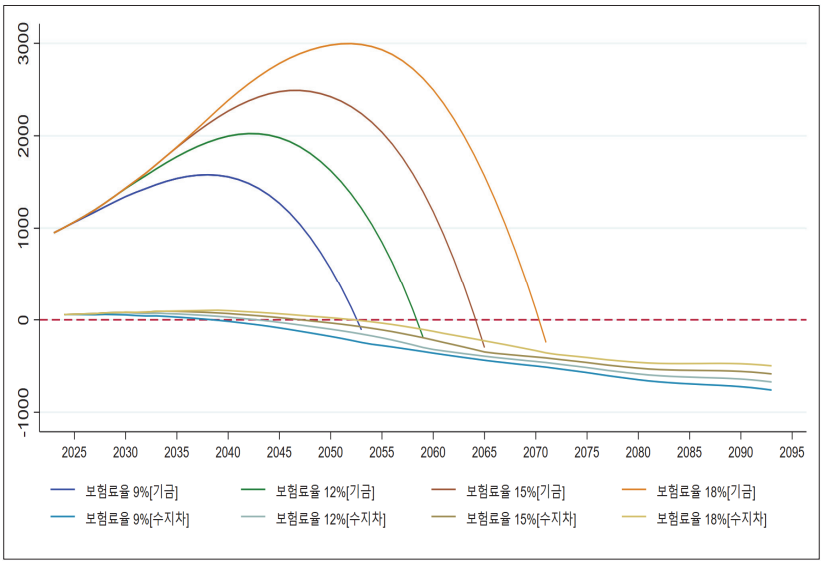
(b) 소득대체율 45%

(단위: 경상가, 조 원)



(c) 소득대체율 50%

(단위: 경상가, 조 원)



자료: 필자 작성

〈표 6-4〉 소득대체율과 보험료율 조정에 따른 재정 전망

구분	보험료율	보험료 수입 < 급여지출 시점	수지 적자 시점	기금소진 시점 (적자 규모)	최대적립 시점 (기금 규모) (GDP 대비 비율)
현행 (S0)	9%	2029년	2039년	2054년 (-217조원)	2038년 (1,583조원) (43.3%)
	12%	2033년	2044년	2060년 (-155조원)	2043년 (2,042조원) (49.1%)
	15%	2037년	2048년	2066년 (-282조원)	2047년 (2,540조원) (55.7%)
	18%	2040년	2055년	2075년 (-282조원)	2054년 (3,128조원) (58.5%)
소득 대체율 45% (S1)	9%	2029년	2039년	2053년 (-111조원)	2038년 (1,579조원) (43.2%)
	12%	2033년	2043년	2059년 (-31조원)	2042년 (2,032조원) (50.1%)
	15%	2037년	2048년	2066년 (-323조원)	2047년 (2,515조원) (55.1%)
	18%	2040년	2054년	2073년 (-352조원)	2053년 (3,056조원) (58.4%)
소득 대체율 50% (S2)	9%	2029년	2039년	2053년 (-248조원)	2038년 (1,574조원) (43.1%)
	12%	2033년	2043년	2059년 (-187조원)	2042년 (2,033조원) (49.9%)
	15%	2037년	2048년	2065년 (-300조원)	2047년 (2,491조원) (54.6%)
	18%	2040년	2053년	2071년 (-245조원)	2050년 (2,982조원) (58.6%)

주: 경상가임
자료: 필자 작성

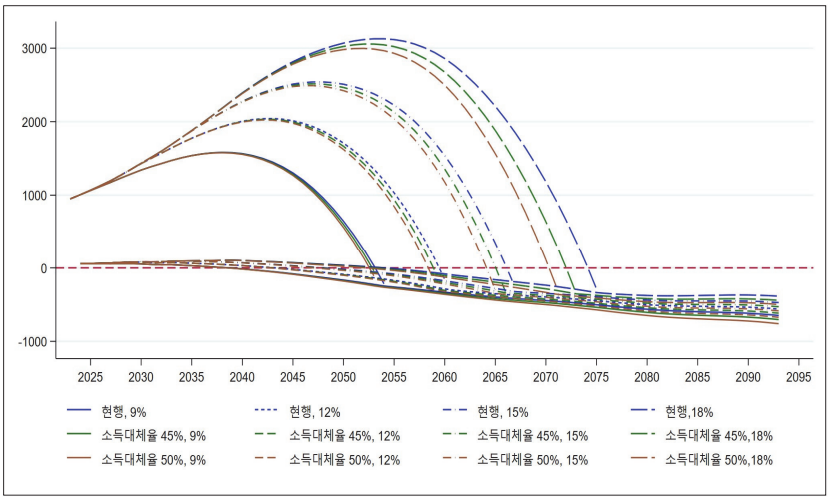
〈표 6-5〉 소득대체율과 보험료율 조정에 따른 재정 전망:기금운용 수익률 1%p 상향

구분	보험료율	보험료 수입 〈 급여지출 시점	수지 적자 시점	기금소진 시점 (적자 규모)	최대적립 시점 (기금 규모) (GDP 대비 비율)
현행 (S0)	9%	2029년	2042년	2058년 (-286조원)	2041년 (1,926조원) (48.7.3%)
	12%	2033년	2048년	2066년 (-180조원)	2047년 (2,654조원) (58.2%)
	15%	2037년	2057년	2077년 (-282조원)	2056년 (3651조원) (65.4%)
	18%	2040년	2067년	-	2054년 (5,156조원) (76.0%)
소득 대체율 45% (S1)	9%	2029년	2042년	2057년 (-118조원)	2041년 (1,920조원) (48.6%)
	12%	2033년	2048년	2065년 (-187조원)	2047년 (2,628조원) (66.4%)
	15%	2037년	2055년	2075년 (-418조원)	2054년 (3,551원) (66.4%)
	18%	2040년	2063년	2089년 (-206조원)	2062년 (4,833조원) (76.8%)
소득 대체율 50% (S2)	9%	2029년	2042년	2057년 (-247조원)	2041년 (1,914조원) (48.4%)
	12%	2033년	2047년	2064년 (-134조원)	2046년 (2,603조원) (58.3%)
	15%	2037년	2054년	2073년 (-379조원)	2054년 (3,466원) (64.8%)
	18%	2040년	2061년	2084년 (-334조원)	2062년 (4,608조원) (74.7%)

주: 경상가임
자료: 필자 작성

[그림 6-12] 소득대체율과 보험료율 조정에 따른 재정 전망

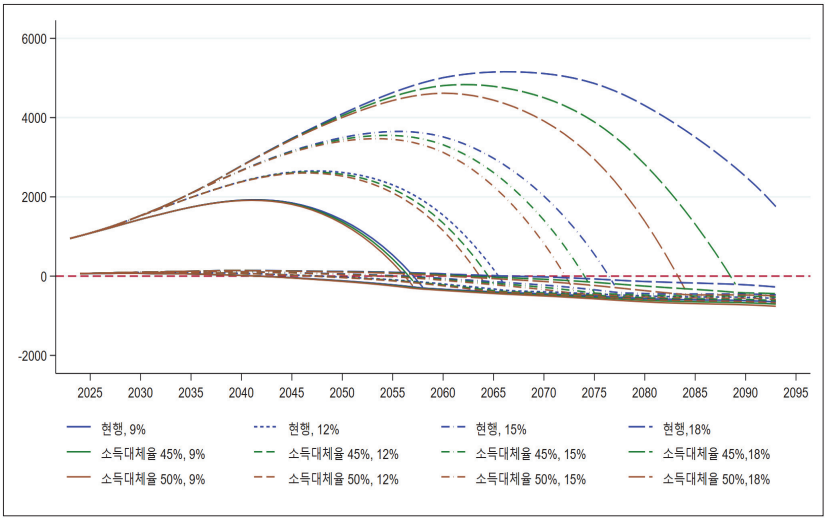
(단위: 경상가, 조 원)



자료: 필자 작성

[그림 6-13] 소득대체율과 보험료율 조정에 따른 재정 전망: 기금운용 수익률 1%p 상향

(단위: 경상가, 조 원)



자료: 필자 작성

제7장

결론

제1절 주요 내용 요약
제2절 한계와 향후 과제

제 7 장 결론

제1절 주요 내용 요약

본 연구는 연금개혁의 효과 분석을 위한 동태적 마이크로시뮬레이션 모형(Microsimulation model, MSM) 개발을 목적으로 하였다. MSM의 개발은 이질적 개인과 가구들이 경험하는 생애사건들과 여러 제도들을 모사하는 작업으로, 방대한 양의 자료와 시뮬레이션 기법은 물론 상당한 시간과 노력을 요구한다. 이를 고려하여 본 연구는 새로운 모형을 구축하기 보다 본 연구원에서 2016년에 구축했던 MSM인 DOSA를 전반적으로 업데이트하여 KIHASA SIM으로 재구축하는 방식으로 연구를 진행하였다. 주된 과업은 기초 자료를 업데이트하고, 그간 여러 연구자들이 개별적으로 개발하였던 코드를 정리하는 한편, 시뮬레이션 로직을 개선하는 것이었다. 이 같은 모형 개발 및 업데이트 과정은 보고서로 작성하기 어려운 측면이 있으나, 본 연구는 모형의 시뮬레이션 방법을 가능한 자세히 담아내고자 하였다.

각 장의 내용은 다음과 같다. 먼저, 2장과 3장에는 MSM에 대한 독자들의 이해를 돕기 위한 내용들을 담았다. 2장에서는 MSM의 개념과 정의, 발전 과정, 기술적 특징을 소개하였고, 연금 모형을 중심으로 국내외 MSM을 검토하였다. 또한 3장에서는 본 연구에서 구축한 KIHASA SIM의 구조와 활용 자료에 대해 자세히 소개하였다.

이후 4장에서는 기초 자료의 구축 과정을 소개하였다. KIHASA SIM은 2015년 인구주택총조사 자료의 2% 표본(998,459명)을 기초 자료로

이용한다. 이를 모형에 투입하기 위해서는 인구 및 가구 구성에 대한 보정, 소득 분포 추정, 국민연금 가입 및 수급 정보 보정 등 수많은 전처리 과정이 필요하다. 4장에는 이러한 과정을 가능한 상세하게 담았다.

5장에서는 주요 모듈의 구축 과정을 소개하였다. 본 연구에서 소개한 모듈은 인구(출산 및 사망), 가구 구성(결혼 및 이혼), 교육 및 경제활동, 소득추정, 공적연금, 고용보험 모듈이다. 각 모듈별로 핵심적인 시뮬레이션 내용과 방법, 주요 가정, 기초 자료 등을 자세히 소개하였다. 한편으로, 이 장에서는 경제활동 모듈과 소득추정 모듈의 시뮬레이션 결과에 대한 유효성 검토를 시행하였는데, 이는 본 모형의 시뮬레이션 결과와 한국노동패널 자료를 비교하는 방식으로 이루어졌다. 경제활동 상태의 경우 두 자료에서 추정한 경제활동 상태별 전이확률과 경제활동 상태의 배열을 비교하였으며, 소득 추정 모듈에서는 두 자료의 소득 분포를 비교하였다. 이를 통해 시뮬레이션 결과와 한국노동패널자료가 상당히 유사함을 확인할 수 있었다. KIHASA SIM이 거시집계치에 대해서만 집계값 조정 기법을 적용하고 있으며, 개인의 사회경제적 상태 변화는 전적으로 행태방정식에 기반하여 시뮬레이션하고 있음을 고려할 때 이는 고무적인 결과이다.

마지막으로 6장에서는 국민연금의 노후소득보장 기능과 관련된 주요 지표(국민연금 가입 기간, 노령연금 수급률, 급여 수준, 실질 소득대체율, 재정수지 등)들을 산출하여 시뮬레이션 결과가 안정적으로 나타나는지 살펴보았다. 또한 제5차 재정계산 보고서에서 제시한 연금개혁 방안(소득대체율-보험료율-기금운용 수익률)을 토대로 각 정책 조합이 급여 수준, 소득대체율, 재정에 미치는 영향을 검토하였다. 이를 통해 KIHASA SIM이 안정적이며 납득할만한 시뮬레이션 결과를 내놓고 있으며, 시나리오 분석 시에도 안정적인 결과가 도출됨을 확인할 수 있었다.

이처럼 KIHASA SIM이 안정적인 시뮬레이션 결과를 도출하는 것은

본 모형이 연금 개혁의 미시적 파급 효과를 사전적으로 분석하고, 공적연금의 장기 운용 전략을 수립하는데 활용될 수 있음을 시사한다. 연금개혁 논의에서는 재정적 지속가능성, 급여 적정성, 보편성 등의 여러 목표들을 동시에 고려한 균형 잡힌 접근이 필요한데, 향후 동 모형이 이를 위한 중요한 분석 도구로 활용되기를 기대한다.

제2절 한계와 향후 과제

본 연구에서 구축한 KIHASA SIM이 안정적인 값들을 산출하고 있다는 점에서 본 연구는 목표했던 바를 어느 정도 달성한 것으로 보인다. 그러나 아직까지 미진한 부분이 많이 존재한다.

첫째, 모형의 안정성과 효율성에 대한 제고가 필요하다. 본 연구가 기초 자료를 새롭게 구축하면서, 시뮬레이션에 이용되는 표본 수가 기존 모형의 2배 수준으로 늘어났다. 이는 시뮬레이션 수행에 소요되는 시간을 늘리는 주된 요인이다. 향후 이 모형의 활용성을 높이기 위해서는 표본 수를 유지하되 소요 시간을 단축시켜야 할 것이다. 한편으로, 기초 자료를 구축하는 과정에서 이상치, 결측치 등 자료 내 오류들을 세밀하게 수정하였으나, 표본 수가 크게 늘어난 만큼 오류들이 잔존하고 있을 가능성을 배제하기 어렵다. 이는 모형의 안정성을 저해하는 요인이므로, 향후 민감도 분석 등 추가적인 안정성 점검이 필요해 보인다.

둘째, 본 모형은 임의가입과 임의계속가입을 고려하지 못하였다. 앞선 5장에서 설명한 것처럼, 본 연구에서 임의가입과 임의계속가입을 모형화하지 않은 것은 두 제도가 노동과의 연계성이 떨어져 모형화하기 어렵고, 관련 기초자료가 부재하기 때문이다. 또한 임의가입과 임의계속가입이 추후

납부와도 직접적으로 연결되어 있어 그 추이가 불안정하다는 점도 이들을 반영하지 않은 이유이다. 그러나 두 가입종별이 국민연금 전체 가입자에서 차지하는 비중은 무시할 수 없는 수준이다. 2022년 12월 기준, 임의가입자는 36.5만 명, 임의계속가입자는 50.1만 명 수준으로, 각각 국민연금 가입자 1,943만 명의 1.88%와 2.58%를 차지하고 있다(국민연금공단, 2022). 즉, 두 가입종별을 배제하는 경우 그만큼 가입 기간이 과소추계될 수 있다. 이를 고려하면 향후 두 가입종별에 대한 모형화를 도모하여 국민연금 가입에 대한 시뮬레이션의 정확성을 높일 필요가 있어 보인다.

셋째, 건강 상태에 대한 고려가 필요하다. 본 모형은 개인의 건강 상태 변수를 고려하지 않았지만, 개인의 건강 상태는 경제활동 및 사망과 직결되므로, 향후 모형을 정교하게 개선하는 과정에서 건강 상태 변수를 추가할 필요가 있어 보인다. 특히, 소득계층별로 상이한 건강 상태 및 기대 여명의 차이가 공적연금의 급여 격차를 심화시킬 수 있으므로, 이에 대한 고려가 필요해 보인다. 최근 국내에서도 의료비 추계를 위한 MSM 개발을 시도하는 등 보건 영역에서의 MSM 개발이 계속되고 있으므로(박경선 외, 2023; 황인욱 외, 2023 등), 이들을 참조하여 모형 개발에 소요되는 시간을 단축할 수 있을 것으로 보인다.

한편으로 KIHASA SIM의 활용가능성을 높이기 위해서는 동 모형을 연금 중심의 노후소득보장 효과를 넘어 기초보장제도를 포함한 전반적인 노후소득보장체계 효과를 분석하는 도구로 확장할 필요가 있다.

이를 위해서는 첫째, 가구 구성 과정에 대한 보완이 필요하다. 현행 KIHASA SIM은 부부와 자녀 중심으로 가구 구성을 단순화하였고, 기초연금 모듈에서 부부 단위 소득 정보를 활용할 뿐, 배우자 외 가구원을 포함한 가구 단위 소득 정보를 구축하지 않는다. 하지만 기초보장제도는 배우자, 자녀, 부모 등의 동거 가구원을 포함한 가구의 최저생계를 보장

하는 가구 단위 공공부조이므로, KIHASA SIM에 기초보장제도를 반영하기 위해서는 반드시 주거·생계 기준 가구를 구성해야 한다.

현행 KIHASA SIM에서는 결혼·이혼·출산·사망에 따른 가구 변화를 반영하므로, 주거·생계 기준 가구를 구성하기 위해서는 우선적으로 이행기 청년의 부모로부터의 분가 및 중고령 부모의 성인 자녀와의 합가를 추가적으로 반영해야 한다. 이를 위해서는 외부자료로부터 10대 후반~40대 집단의 부모로부터 분가 확률, 50대 이상 집단의 성인 자녀와의 합가 확률을 추정하여 시뮬레이션에 반영하는 방법을 검토할 수 있다. 또한, 세부 가구유형별 추계가구 정보를 제공하는 통계청 장래가구추계의 집계치에 맞추어 부모-자녀 분가·합가를 할당하는 방법도 가능하다. 단, 결혼·이혼·출산·사망 및 추가적인 부모-자녀 분가·합가 시뮬레이션의 결과로 생성된 가구유형 분포가 1인 가구, 부부 가구, 부부-미혼 자녀 가구, 부부-부모 가구, 3세대 가구, 조손 가구, 한부모 가구, 비친족 가구 등 다양한 형태의 세부 가구유형 분포를 정확하게 반영하기는 쉽지 않을 것이다. 우선 노인-자녀 동거 가구 등과 같이 노인빈곤 또는 기초보장제도와의 관련성이 높은 주요 가구형태를 중심으로 모델을 개선하고, 장기적으로 가구 구성 모듈을 고도화하기 위해 노력해야 할 것으로 판단된다.

한편, 사회조사의 가구 개념이 단일하지 않다는 점도 고려해야 할 필요가 있다. 인구주택총조사는 주거 단위 가구 개념을 채택하기 때문에 가구에 직장, 학업 때문에 따로 사는 가족을 제외하고 비가족 동거인을 포함하지만, 가계금융복지조사는 생계 단위 가구 개념을 채택하기 때문에 가구에 직장, 학업 때문에 따로 사는 가족을 포함하고 비가족 동거인을 제외한다(통계청, 2021, p.26). 노후소득보장 분야 주요 정책지표인 통계청의 공식 노인빈곤율은 가계금융복지조사를 토대로 산출되므로 생계 단위 가구 개념을 따르지만, KIHASA SIM의 기초 자료인 인구주택총조사와 향후 가구

구성 모듈 개발에 잠정적으로 활용될 수 있는 장래가구추계는 주거 단위 가구 개념을 따른다. 이와 같은 가구 개념의 불일치에 대해서도 장기적인 검토가 필요하다.

둘째, KIHASA SIM의 소득 모듈을 궁극적으로 가구가처분소득 및 노인빈곤 시뮬레이션까지 가능하도록 소득 추정 과정을 보완해야 한다. 첫째, 앞서 논의한 것처럼 개인소득을 가구소득으로 전환하기 위한 가구 모듈 개선이 필요하고, 둘째, 근로·사업소득뿐만 아니라 재산소득, 사적이전소득, 공적이전소득 등으로 소득원천 포괄 범위를 넓혀야 한다. 가계금융복지조사에 따르면, 2022년 65세 이상 노인의 균등화 가처분소득 평균인 2,759만 원/년에서 근로소득, 사업소득, 재산소득, 공적이전소득, 사적이전소득, 공적이전지출, 사적이전지출이 각각 39%, 19%, 17%, 32%, 6%, -10%, -3%를 차지한다(통계청, 2023, 가계금융복지조사 소득분배지표 집계자료). 공적이전소득 중에서 가장 규모가 큰 국민연금과 기초연금, 고용보험은 이미 현행 KIHASA SIM에 반영되어 있으므로, 우선 기초보장제도, 근로장려금 등으로의 확장이 필요하다. 또한 재산소득의 규모도 상당히 크고, 사적 이전의 규모도 무시할 수 없는 수준이다. 단, 소득 모듈 확장의 우선순위를 따지면, 단기적으로 재산소득 및 기초보장제도 보완에 집중해야 할 것으로 판단된다.

이와 같이 소득 모듈을 확장하기 위해서는 가계금융복지조사 데이터를 활용하는 방안을 검토할 필요가 있다. 가계금융복지조사는 노인빈곤율 등의 공식 소득분배지표 작성에 활용되는 데이터로, 국세청, 보건복지부 등의 행정자료로 조사자료를 보완하였기 때문에 소득 데이터의 정확도가 높다. 특히, 재산소득과 근로장려금은 서베이에서 자기응답으로 잘 포착되지 않는 대표적인 소득원천으로(이원진 외, 2019, p.135), 국세청 행정 자료로 보완한 가계금융복지조사의 데이터가 유용하게 활용될 수 있다. 현행

KIHASA SIM에서 근로·사업소득은 외부 자료로부터 개인 및 가구 특성과 소득의 관계를 추정한 소득모형을 적용하여 시뮬레이션하고, 국민연금과 기초연금은 개인 및 가구 특성에 제도 규칙을 적용하여 시뮬레이션한다. 재산소득은 근로·사업소득과 마찬가지로 외부 자료로 추정한 소득모형을 적용하는 방식으로 시뮬레이션해야 한다. 기초보장제도 등의 공적이전소득은 외부자료로 추정한 소득모형을 적용하는 방식과 제도 규칙을 적용하는 방식을 모두 고려할 수 있다. 국민연금, 기초연금, 기초보장제도를 포함한 노후소득보장제도 조합 개편의 영향을 시뮬레이션하기 위해서는 제도 규칙을 적용하는 방식으로 기초보장제도 급여를 추정해야 하지만, 이와 같은 방식의 정확도를 담보하기 위해서는 복잡한 기초보장제도 규칙을 반영하는 데 필요한 가구, 소득, 재산 정보가 충분히 확보되어야 한다. 이상의 내용은 추후의 과제로 남긴다.

참고문헌

강신욱, 노대명, 이현주, 임완섭, 김현경, 권문일, ... ,박형존. (2015). 주요 소득 보장정책의 효과성 평가 연구. 한국보건사회연구원.

고용노동부 홈페이지, 고용안전망 중 실업급여. (2023). https://www.moel.go.kr/policy/policyinfo/safety/list7.do, 2023.11.13. 접속.

고용보험 홈페이지, 구직급여액. (2023). https://www.ei.go.kr/ei/eih/eg/pb/pbPersonBnef/retrievePb0203Info.do, 2023.11.13. 접속.

고제이, 권혁진, 신우진, 류재린, 하솔잎, 조남운. (2016). 미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발-노후 소득 보장 정책을 중심으로. 한국보건사회연구원.

고제이, 권혁진. (2014). 인구구조 변화에 따른 조세수입 전망과 사회복지 지출 부담 분석. 한국보건사회연구원.

국민연금 재정계산위원회. (2023). 국민연금 재정계산 보고서. 2023.10. 국민연금 재정계산위원회.

국민연금공단. (각년도). 국민연금 통계연보.

국민연금재정계산위원회, 국민연금재정추계전문위원회, 국민연금기금운용발전 전문위원회. (2023). 제5차 국민연금 재정계산 - 발표자료 -.

국민연금재정추계위원회. (2018). 2018 국민연금재정계산: 국민연금 장기재정 추계.

권혁진, 류재린. (2015). 공적연금의 최저생계 보장 효과에 대한 장기 전망. 응용통계연구, 28(4), pp.741-762.

권혁진, 류재린. (2018). 노후소득보장을 위한 공적연금의 적정성과 개인안정성에 대한 전망. 공공사회연구, 8(1), pp.38-82.

권혁진, 신우진, 한재명, 정동진, 류재린, 박충렬, 김균. (2013). 2007년 공적 연금 개혁의 적정성 전망과 평가-Dynamic Micro-Simulation을 통한 장기 분석. 사회보장연구, 29(4), pp.1-54.

권혁진, 한정림. (2009). 국민연금의 미시모의실험 모형 연구 및 개발, 국민연금연구원.

권혁진. (2012). 비정규 고용과 국민연금 사각지대. 응용경제, 14(2), pp.85-120.

권혁진. (2020). 노후소득 적정성에 대한 공적연금 개혁 방안들의 장기 효과 전망 및 평가 - 평균의 함정에 빠져버린 정부 정책의 한계. 연금연구, 10(1), pp.1-43.

권혁진. (2022). 동학적 미시 시뮬레이션 모형 구축 및 활용 방안. 이슈페이퍼. 저출산고령사회위원회(미발간 보고서).

기획재정부. (2020. 9. 3.). 2020~2060년 장기재정전망. 기획재정부 보도자료 별첨자료. https://www.moef.go.kr/nw/nes/detailNesDtaView.do?searchBbsId1=MOSFBBS_000000000028&searchNttId1=MOSF_000000000045158&menuNo=4010100에서 2023.1.10. 인출.

김대중, 이난희. (2018). 건강행태의 변화에 따른 질병 예측 및 질병부담 추계 연구. 한국보건사회연구원.

김영옥, 김종숙, 강성호, 이선행. (2011). 여성취업자의 국민연금 가입 및 수급의 실효성 강화 방안. 여성정책연구원.

김준일, 류재린. (2018). 국민연금기금의 사회책임투자와 인구구조 변화: 미시모의실험을 통한 연금재정분석. 경제발전연구, 24(4), pp.65-94.

김혜원, 김을식, 전승훈. (2010). 사회보장시스템의 생애소득 재분배 연구, 한국노동연구원.

류재린, 권혁진, 우선희. (2022b). 지역가입자 연금보험료 지원 제도의 효과 추정 - 마이크로시뮬레이션 모형을 중심으로. 한국보건사회연구원.

류재린, 문현경. (2022). 코로나 19 확산이 국민연금 가입에 미치는 영향 분석. 사회보장연구, 38(2), pp. 31-63.

류재린, 오유진. (2020). 국민연금의 임의가입 현황과 결정요인 분석. 사회보장연구, 36(2), pp.137-162.

류재린, 정해식, 이용하, 신화연, 이다미, 이지혜. (2022a). 공적연금 재구조화 방안 연구. 한국보건사회연구원.

류재린. (2020). 마이크로 시뮬레이션 모형의 소개 및 시사점. 연금포럼, 여름 (78). 국민연금연구원.

박경선, 이수연, 이상희, 송지은, 주시연. (2023). 미시실험모형 기반의 의료비 추계모형 개발 기초연구. 건강보험연구원(발간예정).

박진희, 이시균, 박세정, 강민정, 이혜연. (2020). 인력수급 전망 방법 개선 :KEISIM 개발과 개선을 중심으로. 한국고용정보원.

병무청. (각년도). 병무통계. https://kosis.kr/statHtml/statHtml.do?orgId=144&tblId=TX_14401_A015&conn_path=I2 에서 2023.8.28. 인출.

보건복지부. (2018). 제4차 국민연금 재정계산을 바탕으로 한 국민연금 종합 운영계획.

성명재, 송헌재, 전병목. (2010). 조세・재정 모의실험 모형: KIPFSIM10 모형의 구축. 한국조세재정연구원.

송창길, 나원희, 백주하, 류한별, 한신실, 고경표. (2023, 발간예정). 인구추계모형 구축과 활용 방안. 한국보건사회연구원.

신우진, 권혁진, 류재린. (2016). 불안정 노동이 국민연금 노후소득 보장기능에 미치는 영향. 사회보장연구, 32(1), pp.33-55.

오세영, 윤건, 오균. (2017). 증거기반정책을 위한 정부의 통계 구축 및 활용에 대한 현황 조사. 한국행정연구원 사회조사센터.

이다미. 권혁진. (2019). 국민연금 실업크레딧의 노후소득보장 효과 - 미시 시뮬레이션을 활용한 장기 전망. 사회보장연구, 35(3), pp.39-72.

이상열, 오진호, 이상조, 윤영규, 양건호. (2019). 장래인구추계 마이크로 시뮬레이션 방법론 연구. 한국통계학회, 통계청.

이시균, 강민정, 홍현균, 방 글, 공정승, 윤정향, … , 이진면. (2014). 중장기 인력수급 전망 2013~2023, I. 한국고용정보원.

이용하, 류재린, 이다미, 송창길, 최준영, 김원섭. (2022). 기초연금-국민연금과의 관계 및 향후 발전방안 연구. 한국보건사회연구원.

이원진, 정해식, 전지현. (2019). 소득조사 마이크로데이터 비교 분석: 〈가계동향조사〉와 〈가계금융·복지조사〉를 중심으로. 세종: 한국보건사회연구원.

정해식, 권혁진, 주은선, 최옥금. (2020). 노령층소득보장 제도 재구조화 방안. 소득보장체계 재구조화 방안 연구: 제도간 연계· 조정을 중심으로. 경제인문사회연구회. 한국보건사회연구원.

주은선, 권혁진, 유희원, 원종현, 이은주, 이재훈, 정해식, 김정목. (2017). 국민연금의 발전적 재구성. 공적연금강화 국민행동.

최기홍. (2015). 국민연금 급여산식의 구조와 소득재분배 기능에 대한 미시모의실험. 재정학연구, 8(1), pp.33-58.

최기홍. (2018). 급여산식의 누진성 비교연구: 국민연금과 미국 OASDI. 재정학연구, 11(3), pp. 85-115.

통계청 사회통계국 인구동향과. (2021). 장래인구추계: 2020~2070년. https://kostat.go.kr/board.es?mid=a10301020600&bid=207&act=view&list_no=415453 에서 2023.1.10. 인출.

통계청. (2015). 인구주택총조사. https://mdis.kostat.go.kr/ofrData/selectOfrDataDetail.do 에서 2023.3.10. 인출.

통계청. (2021). 2021년 가계금융복지조사 지침서. 대전: 통계청.

통계청. (2023). 가계금융복지조사 소득분배지표 집계자료. https://kosis.kr/ 에서 2023.12.30. 인출.

통계청. (각년도). 경제활동인구조사 근로형태별 부가조사(8월). 원자료. https://mdis.kostat.go.kr/ofrData/selectOfrDataDetail.do?survId=47&itmDiv=1&nPage=3&itemId=2004&itemNm=%EB%85%B8%EB%8F%99 에서 2023. 5.10. 인출.

통계청. (각년도). 인구동향조사. 원자료. https://mdis.kostat.go.kr/ofrData/selectOfrDataDetail.do?survId=38&itmDiv=1&nPage=3&itemId=2001&itemNm=%EC%9D%B8%EA%B5%AC 에서 2023.5.10. 인출.

한국교육개발원. (각년도). 교육통계분석자료집-고등교육통계편. 한국교육개발원. https://kess.kedi.re.kr/publ/list?itemCode=02&survSeq=&menuSeq=0&division=&word= 에서 2023.5.10. 인출.

한국교육개발원. (각년도). 교육통계연보. 한국교육개발원. https://kess.kedi.re.kr/publ/view?survSeq=2022&publSeq=2&menuSeq=0&itemCode=02&language=en에서 2023.5.10. 인출.

한국노동연구원. (각년도). 한국노동패널조사. https://www.kli.re.kr/klips 에서 2023.5.10. 인출.

한국조세재정연구원. (각년도). 재정패널조사. 각년도. https://www.kipf.re.kr/panel/data/AllList/paData/view.do?cntNo=33에서 2023.5.10. 인출.

황인욱, 이윤경, 임유나, 이태진. (2023). 미시모의실험모형을 이용한 장래 건강보험 진료비 추정. 보건경제와 정책연구, 29(2), pp.55-86.

Anderson, R. E., and Hicks, C. (2011). Highlights of contemporary microsimulation. Social Science Computer Review 29(1). pp.3-8.

Baltagi, B. H., and Wu, P. X. (1999). Unequally spaced panel data regressions with AR (1) disturbances. Econometric theory, 15(6), 814-823.

Banks, J. (1998). "Principles of simulation". in Handbook of simulation, 12, pp. 3-30.

Boulanger, P., Dudu, H., Ferrari, E., Mainar-Causapé, A. J. and Ramos, M. P. (2022). Effectiveness of fertilizer policy reforms to enhance food security in Kenya: a macro–micro simulation analysis. Applied Economics, 54(8), pp.841-861.

Brown, L. and Harding A. (2002). Social modeling and public policy: Application of microsimulation modeling in Australia. Journal of Artificial Societies and Social Simulation 5(4).

Dekkers, Gijs. and Belloni, Michele. (2009). A classification and overview of micro simulation models, and the choices made in MIDAS, Chapter.2 in What are the consequences of the AWG-projections for the adequacy of social security pensions?, ENEPRI Research Report No.65, AIM WP4.

Figari, F., Paulus, A., and Sutherland, H. (2014). Microsimulation and policy analysis. ISER Working Paper Series.

Flood, L., Jansson, F., Pettersson, T., Pettersson, T., Sundberg, O., and Westerberg, A. (2012). The Handbook of SESIM-a Swedish dynamic micro simulation model.

Fredriksen, D., Holmøy, E., Strøm, B., and Stølen, N. M. (2019). Fiscal effects of the Norwegian pension reform-A micro-macro assessment. Journal of Pension Economics & Finance, 18(1), 88-123.

Gál, R., Horváth, A., Orbán, G., and Dekkers, G. (2009). PENMICRO Monitoring pension developments through micro socioeconomic instruments based on individual data sources: feasibility study. TARKI Social Research Institute, Budapest.

Harding, A., Keegan, M., Kelly, S. (2010). Validating a dynamic population microsimulation model: Recent experience in Australia, International Journal of microsimulation 3(2) pp. 46-64.

Klevmarken, N.A. (1997). Behavioral modeling in micro simulation models: a survey, Working Paper No 1997:31. Department of Economics, Uppsala University.

Lay-Yee, R., and Cotterell, G. (2015). The role of microsimulation in the development of public policy. In Policy Practice and Digital Science(pp. 305-320). Springer.

Legendre, François. (2019). "The Emergence and Consolidation of Microsimulation Methods in France. In: Economie et Statistique", Economics and Statistics, n°510-512, Special Issue 50th Anniversary. pp. 201-217

Li, J., and O'Donoghue, C. (2013). A survey of dynamic microsimulation models: uses, model structure and methodology. International Journal of microsimulation, 6(2), pp.3-55.

Li, J., O'Donoghue, C., and Dekkers, G. (2014). "Dynamic models". In Handbook of microsimulation modelling(Vol. 293, pp. 305-343). Emerald Group Publishing Limited.

Merz, J. (1991). Microsimulation? a survey of principles, developments and applications. International Journal of Forecasting 7(1). pp.77-104.

Milne BJ, Lay-Yee R, McLay J, Tobias M, Tuohy P, Armstrong A, Lynn R, Davis P, Pearson J, Mannion O. (2014). A collaborative approach to bridging the research-policy gap through the development of policy advice software. Evidence Policy 10(1) pp.127-136.

Morgan, Mary S. (2012). The World in the model: How Economists Work and Think. Camgridge Press.

O'Donoghue. (2014). "Introduction" In Handbook of Microsimulation Modelling. Published online: 30 Oct 2014; 1-21.

OECD. (2019). Monitoring and evaluating social protection systems-Lessons from the EU-SPS Programme, OECD Development Centre

Orcutt, G. H., Greenberger, M., Korbel, J., and Rivlin, A. M. (1961). Microanalysis of socioeconomic systems: A simulation study.

Orcutt, Guy H. (1957). A New Type of Socio-Economic System. The Review of Economics and Statistics. Vol. 39, No. 2

Orcutt, Guy H. (1960). Simulation of Economic Systems. American Economic Review, 50:5. pp.893-907.

Salonen, J., Möttönen, J., Tikanmäki, H., and Nummi, T. (2020). Using sequence analysis to visualize and validate model transitions.

Spielauer, M. (2011). What is social science microsimulation?. Social Science Computer Review, 29(1), pp.9-20.

Zaidi A, Harding A, and Williamson P.(eds) (2009). New frontiers in microsimulation modeling. Ashgate.

Zhang, Q., Zhang, X., Cui, Q., Cao, W., He, L., Zhou, Y., ... and Fan, Y. (2022). The unequal effect of the COVID-19 pandemic on the labour market and income inequality in China: A multisectoral CGE model analysis coupled with a micro-simulation approach. International Journal of Environmental Research and Public Health, 19(3), 1320.

Zhang, Xiao-guang. (2017). Solving a Partial Equilibrium Model in a CGE Framework: The Case of a Behavioural Microsimulation Model. International Journal Of Microsimulation 10(3) pp.27-58.

부록1. 변수일람표

〈부표 1-1〉 기초 자료의 변수일람표

변수명	변수 내용	참고사항
sim_yy	연도	2015~
dfid	직계존속ID	정수(13자리)
fid	가족ID	정수(13자리)
pid	개인ID	정수(12자리)
frel	가족 내 관계	본인(1), 배우자(2), 자녀(3)
gend	성별	남성(1), 여성(2)
age	나이(만나이)	
birth_y	출생연도	정수(4자리)
fn	가족구성원 수	
children	가족 내 자녀 수	
sid	배우자 ID	
sid_d	사망배우자 ID	
mari_ok	배우자 유무	무배우자(0), 유배우자(1)
mari_sta	혼인 상태	미성년자(0), 기혼(1). 미혼(2), 사별(3)
mari_div	이혼 여부	이혼무경험(0), 이혼경험(1)
mari_re	재혼 여부	재혼무경험(0), 재혼경험(1)
mari_die	사별 여부	사별하지 않음(0), 사별함(1)
wed_age	최근 혼인연령	
fwed_age	초혼연령	
rwed_age	재혼연령	
sch_what	학교 수준	미취학아동(0), 초등학교(1), 중학교(2), 고등학교(3), 4년제 미만 대학(4), 4년제 이상 대학(5), 석사(6), 박사(7), 무학(9)
sch_goi	재학 상태	졸업/수료(1), 재학/휴학(2), 중퇴(3)
sch_grd	학년	
sch_yr	교육연수	

변수명	변수 내용	참고사항
gundae	군필 여부	군필(11), 여성(99), 입영대기(0), 복무년차(1,2), 면제
gundae_yr	입대연도	
eact	경제활동상태, 종사상지위	상용(11), 임시|일용(12), 자영자(13),고용주(14), 무급가족종사자(15), 실업(20), 비경활(30), 15세미만(50)
ind	산업	농업, 임업 및 어업(1), 광업(2), 제조업(3), 전기,가스,증기 및 수도사업(4), 하수 폐기물처리, 원료재생 및 환경복원업(5), 건설업(6), 도매 및 소매업(7), 운수업(8), 숙박 및 음식점업(9), 출판, 영상 방송통신 및 정보서비스업(10), 금융 및 보험업(11), 부동산업 및 임대업(12), 전문,과학 및 기술서비스업(13), 사업시설관리 및 사업지원서비스업(14), 공공행정, 국방 및 사회보장행정(15), 교육서비스업(16), 보건업 및 사회복지서비스업(17), 예술, 스포츠 및 여가 관련 서비스업(18), 협회 및 단체, 수리 및 기타 개인서비스업(19), 가구 내 고용활동 및 달리 분류되지 않은 자가생산활동(20), 국제 및 외국기관(21), 해당없음(99)
job	직업	관리자(1), 전문가 및 관련 종사자(2), 사무 종사자(3), 서비스 종사자(4), 판매 종사자(5), 농림어업 숙련 종사자(6),기능원 및 관련 기능 종사자(7), 장치기계 조작 및 조립 종사자(8), 단순노무 종사자(9), 기타(10), 해당없음(99)
inc_earn	연간근로(사업)소득	
ui	개인고유 오차항	
eit	시간가변 오차항	
dable_type	장애 유형	활동제약없음(0), 시청각장애(1), 치매와 중풍(2), 언어/자폐(3), 정신적 제약(4), 육체적 제약(5), 복합장애(6), 해당없음(9)
nps_part	국민연금 가입 상태	미가입(0), 가입(1)
nps_ptype	국민연금 가입종별	사업장(1), 지역소득신고(2), 미가입(9)
nps_pmon_yr	국민연금 가입월 수 [연간]	
nps_pwon	국민연금 납부보험료[연간,원]	
nps_pmon_cr	국민연금 가입월 수 [누적]	
nps_ben	국민연금 수급 상태	

변수명	변수 내용	참고사항
nps_btype	국민연금 급여 종류	
nps_bwon	국민연금 연금액(연간,원) [출산군 복무]	
nps_bstart	국민연금 최초 수급 연도	
nps_bdup	국민연금 노령(장애유족) 중 이미 선택한 자	
nps_bstop	국민연금 수급정지 사유	유족연금[3년지급 후 정지](1), 유족연금[소득있는 업무종사](2), 장애연금[소득있는 업무종사](11)
nps_credit_mat	국민연금 출산 크레딧	
nps_credit_unp	실업크레딧(개월수) (누적)	
nps_bwon_noc	연금액(연간,원) [크레딧제외]	
nps_bwon_mat	연금액(연간,원) [출산크레딧 반영]	
nps_bwon_mil	연금액(연간,원) [군복무크레딧 반영]	
nps_bwon_unp	연금액(연간,원) [실업크레딧 반영]	
nps_bwon_all	연금액(연간,원) [모든 크레딧 반영]	
nps_sub	보험료 지원(개월 수) (누적)	
nps_sub_yr	보험료 지원(개월 수) (당해)	
sub_mount_yr	보험료 지원(금액) (당해)	
sub_mount	보험료 지원(금액) (누적)	
nps_ptype2	가입종별 (납부예외 포함)	
Bpen_cal_noc	A급여액(연간,원) [크레딧 제외]	
Bpen_cal_mat	A급여액(연간,원) [출산크레딧 반영]	

변수명	변수 내용	참고사항
Bpen_cal_mil	A급여액(연간,원) [군대크레딧 반영]	
Bpen_cal_unp	[A급여액(연간,원) [실업크레딧:12개월]	
Bpen_cal_all	A급여액(연간,원) [모든 크레딧 반영]	
ei_part	가입여부	
ei_pmon_yr	가입개월수(연간)	
ei_pmon_yr_p	가입개월수(작년)	
ei_pmon_cr	가입개월수(누적)	
ei_pwon	고용보험 보험료(연간,원)	
ei_ben	고용보험 실업급여 수급 상태	
ei_bmonz	고용보험 실업급여 수급기간 (개월)	
ei_bwon	고용보험 실업급여액 (연간, 원)	
ei_pastw	고용보험 실업급여 기준임금(전년도임금) [연간]	
houstype	점유 형태	자가(1), 주택임차(2), 기타(3), 신생가구(99)
asset_hous	자가[만원]	
asset_rent	임대[만원]	
asset_real	일반자산[만원]	
asset_fina	금융자산[만원]	
asset_debt	부채[만원]	
gtype	세대유형	
hid_org	가구ID(원자료)	
hn_org	가구원수(원자료)	
rel_org	가구주와관계(원자료)	
fid_org	가족ID(원자료)	
is_die	사망여부	생존(0), 사망(1)

자료: 필자 작성