

연구보고서(수시) 2022-02

국민건강보험공단 행정자료의 사회정책적 활용을 위한 탐색적 연구



강신욱
이원진·이정민·강희정·하은솔·문석준
김재용·조재영·김은미·김지혜



사람을
생각하는
사람들



KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



한국보건사회연구원
KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



연구진

연구책임자	강신욱	한국보건사회연구원 선임연구위원
공동연구진	이원진	한국보건사회연구원 부연구위원
	이정민	서울대학교 경제학부 교수
	강희정	한국보건사회연구원 선임연구위원
	하은솔	한국보건사회연구원 전문연구위원
	문석준	한국보건사회연구원 연구원
	김재용	국민건강보험공단 빅데이터연구부 부장
	조재영	국민건강보험공단 빅데이터연구부 팀장
	김은미	국민건강보험공단 빅데이터연구부 주임연구위원
	김지혜	국민건강보험공단 빅데이터연구부 팀장

연구보고서(수시) 2022-02

국민건강보험공단 행정자료의 사회정책적 활용을 위한 탐색적 연구

발행일 2022년 12월
발행인 이태수
발행처 한국보건사회연구원
주소 [30147] 세종특별자치시 시청대로 370
세종국책연구단지 사회정책동(1~5층)
전화 대표전화: 044)287-8000
홈페이지 <http://www.kihasa.re.kr>
등록 1999년 4월 27일(제2015-000007호)
인쇄처 (주)정인애드

© 한국보건사회연구원 2022
ISBN 978-89-6827-925-6 [93330]
<https://doi.org/10.23060/kihasa.b.2022.02>

발|간|사

사회보장정책 연구분야에서 행정자료를 활용하는 사례가 점차 증가하고 있다. 정책 연구의 수요에 비해 활용할 데이터의 종류와 양이 충분하지 않은 사회정책분야의 특성을 고려할 때 바람직한 현상이라고 할 수 있다. 국민건강보험공단이 축적한 전 국민 대상의 방대한 행정자료에 대해 정책 연구자들의 관심이 집중되고 있었다. 하지만 개인단위 자료라는 한계로 인해 이 자료를 이용하여 소득계층별 사회경제적 실태를 파악하는데 어려움이 있었다. 이 연구는 이러한 어려움을 이론적으로, 그리고 자료 분석을 통해 해결하려는 의미있는 시도를 보여준다.

이 연구가 갖는 또 다른 중요한 의미는 두 기관 간 업무협약에 근거한 협동 연구라는 점이다. 2021년에 체결된 한국보건사회연구원과 국민건강보험공단의 업무협약을 바탕으로 이 연구가 기획되었고, 두 기관의 연구진들은 자료 처리 및 분석방법, 그리고 분석 결과의 해석에 이르기까지 깊이 있는 토론을 반복했다. 이 연구가 ‘탐색적’ 연구라는 유보적 표현을 제목에 포함하고 있기는 하지만 연구 과정에서 보여준 협력과 논의의 진지함은 결코 과소평가될 수 없을 것이다. 연구 책임을 맡은 한국보건사회연구원의 강신욱 선임연구원을 포함하여 연구에 참여한 두 기관의 모든 연구진, 그리고 서울대학교 이정민교수께 감사의 말씀을 드린다. 이 연구가 해당 주제의 연구 분야는 물론, 행정자료의 활용과 기관간 협업의 측면에서 더 큰 진전을 이루는데 기여할 것으로 기대한다.

2022년 12월

한국보건사회연구원 원장

이 태 수

목 차

KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



Abstract	1
요 약	7
제1장 서론	17
제1절 연구의 배경과 목적	19
제2절 연구의 방법과 구성	23
제2장 분석자료의 구축	27
제1절 자료원 소개	29
제2절 자료의 구축	32
제3절 소결	49
제3장 가구 단위 정보 생성 방법 및 가구 ID 생성방법	51
제1절 가구 개념에 대한 문헌 검토	53
제2절 가구 구성 방법	66
제3절 가구 단위 정보 생성 결과	92
제4절 소결	98
제4장 소득 분포 분석: 가계금융복지조사와의 비교	101
제1절 개요	103
제2절 맞춤형 소득-재산 DB와 가계금융복지조사의 차이	103
제3절 분석 결과	107
제4절 소결	142

제5장 소득과 재산의 결합 분포 분석	145
제1절 분석 자료 및 분석 방법	147
제2절 가구 구성안별 재산 분포	152
제3절 가구 구성안별 소득과 재산의 결합 분포	158
제4절 소결	163
제6장 소득-자산 결합 방법과 적용	165
제1절 소득-자산 결합 지표의 의미	167
제2절 연구의 내용 및 방법	169
제3절 결합 지수의 활용 및 특징 비교	176
제4절 소결	193
제7장 소득 수준별 사망률 차이에 대한 탐색적 분석	195
제1절 분석 배경과 방법	197
제2절 가구 구성안별 소득 100분위의 적합도 비교	206
제3절 소득 수준에 따른 사망률 차이	211
제4절 소결	220
제8장 결론 및 시사점	225
제1절 연구의 요약과 결론	227
제2절 자료의 활용도 제고를 위한 제언	231
참고문헌	235
부록	239

표 목차

KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



〈표 2-1〉 국민건강정보 DB의 영역별 주요 데이터 내용	30
〈표 2-2〉 국민건강정보 DB 자료 공개 및 제공 절차	31
〈표 2-3〉 자료원별 생산기관 및 자료확보·보험료 적용 시점	34
〈표 2-4〉 소득유형별 소득금액 출처	37
〈표 2-5〉 재산유형별 과세표준	38
〈표 2-6〉 연도별 구축과정별 인원(2009~2021)	41
〈표 2-7〉 연도별 내국인, 외국인 구성(2009~2021)	42
〈표 2-8〉 연도별 가입자 유형별 구성(2009~2021)	43
〈표 3-1〉 국외 조사 가구 기준 검토	57
〈표 3-2〉 가계금융복지조사 조사제외 가구	61
〈표 3-3〉 가계금융복지조사 조사대상 포함 및 제외 기준	62
〈표 3-4〉 인구주택총조사 조사대상 포함 및 제외 기준	63
〈표 3-5〉 인구주택총조사의 행정자료를 통한 가구보완	64
〈표 3-6〉 다섯 가지 가구 구성안의 명칭과 기본 개념	68
〈표 3-7〉 가구 ID 1안: A, B유형일 때 가구 ID 부여 예시	73
〈표 3-8〉 가구 ID 1안: C유형일 때 가구 ID 부여 예시	76
〈표 3-9〉 가구 ID 2안: A, B유형일 때 가구 ID 부여 예시	79
〈표 3-10〉 가구 ID 2안: C유형일 때 가구 ID 부여 예시	83
〈표 3-11〉 가구 ID 3안: A, B유형일 때 가구 ID 부여 예시	87
〈표 3-12〉 가구 ID 3안: C유형일 때 가구 ID 부여 예시	91
〈표 3-13〉 가구 구성안별 가구 수 및 가구원 수(2009, 2015, 2021)	93
〈표 4-1〉 가계금융복지조사 및 맞춤형 소득-재산 DB의 소득 개념·범위	104
〈표 4-2〉 2019년 개인 사례 수	108
〈표 4-3〉 2019년 개인 성·연령 분포	109
〈표 4-4〉 2019년 개인근로소득 분포, 전체 인구	111
〈표 4-5〉 2019년 개인근로소득 분포, 18~64세	112
〈표 4-6〉 2019년 개인사업소득 분포, 전체 인구	113

〈표 4-7〉 2019년 개인사업소득 분포, 18~64세	114
〈표 4-8〉 2019년 개인금융소득 분포, 전체 인구	115
〈표 4-9〉 2019년 개인공적연금소득 분포, 전체 인구	117
〈표 4-10〉 2019년 개인공적연금소득 분포, 65세 이상	118
〈표 4-11〉 2019년 개인총소득 분포, 전체 인구	119
〈표 4-12〉 2019년 개인총소득 분포, 18~64세	121
〈표 4-13〉 2019년 가구규모 분포, 가구 단위 분석	123
〈표 4-14〉 2019년 가구규모 분포, 개인 단위 분석	126
〈표 4-15〉 2019년 제곱근 균등화 가구근로소득 분포, 전체 인구	127
〈표 4-16〉 2019년 제곱근 균등화 가구사업소득 분포, 전체 인구	128
〈표 4-17〉 2019년 제곱근 균등화 가구금융소득 분포, 전체 인구	130
〈표 4-18〉 2019년 제곱근 균등화 가구공적연금소득 분포, 전체 인구	131
〈표 4-19〉 2019년 제곱근 균등화 가구총소득 분포, 전체 인구	132
〈표 4-20〉 2019년 제곱근 균등화 가구총소득 분포, 18~64세	135
〈표 4-21〉 2019년 제곱근 균등화 가구총소득 분포, 65세 이상	137
〈표 4-22〉 2019년 맞춤형 소득-재산 DB 소득분배지표	139
〈표 4-23〉 2019년 가계금융복지조사 소득분배지표	140
〈표 4-24〉 2019년 총소득 분배지표	141
〈표 5-1〉 〈가계금융복지조사〉에서의 가구 소득 및 자산 구성	151
〈표 5-2〉 가구 구성안별 재산 보유 실태	153
〈표 5-3〉 재산분위별 평균 재산	155
〈표 5-4〉 가구 구성안별 분위 평균 재산 비교(세대기준안=100)	155
〈표 5-5〉 가구 구성안별 재산 분배지표 비교(재산분위 이용)	156
〈표 5-6〉 가구 구성안별 소득-재산의 상관계수	158
〈표 5-7〉 소득분위별 평균 재산	159
〈표 5-8〉 소득 및 재산 5분위 배율(소득분위 기준)	160
〈표 5-9〉 재산분위별 평균 소득	162



〈표 5-10〉 소득 및 재산의 팔마 비율(재산분위 이용)	163
〈표 6-1〉 지역별 주거재산 인정 한도액 및 기본재산 공제액	175
〈표 6-2〉 기초생활보장제도의 가구원 규모별 가구 균등화 지수	176
〈표 6-3〉 가구의 정의와 소득-재산 결합방식에 따른 결합지수의 차이	177
〈표 6-4〉 결합지수가 0인 가구의 비율	178
〈표 6-5〉 0을 포함한 결합지수 중위값의 50% 미만 가구 특성: 세대보완 1안	179
〈표 6-6〉 0을 포함한 결합지수 중위값의 50% 미만 가구 특성: 세대보완 2안	180
〈표 6-7〉 0을 포함한 결합지수 중위값의 50% 미만 가구 특성: 세대보완 3안	180
〈표 6-8〉 0을 포함한 결합지수 중위값의 50% 미만 가구 특성: 세대기준안	181
〈표 6-9〉 0을 포함한 결합지수 중위값의 50% 미만 가구 특성: 건강보험증기준안	181
〈표 6-10〉 0을 포함하지 않은 결합지수 중위값의 50% 미만 가구 특성: 세대보완 1안 · 183	
〈표 6-11〉 0을 포함하지 않은 결합지수 중위값의 50% 미만 가구 특성: 세대보완 2안 · 184	
〈표 6-12〉 0을 포함하지 않은 결합지수 중위값의 50% 미만 가구 특성: 세대보완 3안 · 184	
〈표 6-13〉 0을 포함하지 않은 결합지수 중위값의 50% 미만 가구 특성: 세대기준안 · 185	
〈표 6-14〉 0을 포함하지 않은 결합지수 중위값의 50% 미만 가구 특성: 건강보험증기준안	185
〈표 6-15〉 상위 10% 가구: 세대보완 1안	186
〈표 6-16〉 상위 10% 가구: 세대보완 2안	187
〈표 6-17〉 상위 10% 가구: 세대보완 3안	187
〈표 6-18〉 상위 10% 가구: 세대기준안	188
〈표 6-19〉 상위 10% 가구: 건강보험증기준안	188
〈표 6-20〉 데이터 내 정합성: 0을 포함한 중위값의 50% 미만 가구	190
〈표 6-21〉 데이터 내 정합성: 0을 포함하지 않은 중위값의 50% 미만 가구	191
〈표 6-22〉 데이터 내 정합성: 상위 10% 이상 가구	192
〈표 7-1〉 사회경제적 수준과 사망률 간의 관계에 대한 실증적 연구(국외)	198
〈표 7-2〉 사회경제적 수준과 사망률 간의 관계에 대한 실증적 연구(국내)	200
〈표 7-3〉 가구 구성안별 분석 대상자(2020년, 30세 이상)	206

〈표 7-4〉 가구 구성안별 사망률에 대한 선형모형에서 가구 균등화 개인소득 100분위의 회귀계수와 설명력(2020년, 30세 이상)	207
〈표 7-5〉 가구 구성안별 소득분위별 대상자 분포	212
〈표 7-6〉 세대보완 1안(건강보험 피부양자 포함)에서 소득 10분위 연령보정 상대사망률비(RR)1	214
〈표 7-7〉 세대보완 2안(비동거 가족 포함)에서 소득 10분위 연령보정 상대사망률비(RR)1	216
〈표 7-8〉 세대보완 3안(분리세대 추정)에서 소득 10분위 연령보정 상대사망률비(RR)1 ..	217
〈표 7-9〉 세대기준안에서 소득 10분위 연령보정 상대사망률비(RR)1	218
〈표 7-10〉 건강보험증기준안에서 소득 10분위별 연령보정 상대사망률비(RR)1	219
〈부표 6-1〉 가구 정의와 결합방법에 따른 소득 중위값의 50%과 90분위수	239

그림 목차

KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



[그림 2-1] 구축 과정(2009~2021)	40
[그림 2-2] 연도별 성별 분포(2009~2021)	44
[그림 2-3] 연령대별 분포(2009~2021)	44
[그림 2-4] 성별 연령대별 분포(2009, 2015, 2021)	45
[그림 2-5] 성별, 연령별 총소득의 평균(mean), 중위값(median), 왜도지표(Skewness-index)	47
[그림 2-6] 성별, 연령별 총재산의 평균(mean), 중위값(median), 왜도지표(Skewness-index)	48
[그림 3-1] 1차 유형 분류 흐름도	70
[그림 3-2] 가구 ID 1안 생성 흐름: A유형, B유형	73
[그림 3-3] 가구 ID 1안 생성 흐름: C유형	75
[그림 3-4] 가구 ID 2안 생성 흐름: A유형, B유형	79
[그림 3-5] 가구 ID 2안 생성 흐름: C유형	83
[그림 3-6] 가구 ID 3안 생성 흐름: A유형, B유형	86
[그림 3-7] 가구 ID 3안 생성 흐름: C유형	90
[그림 3-8] 가구원 수별 가구 분포	93
[그림 3-9] 가구주 특성별 가구 분포: 성별	95
[그림 3-10] 가구주 특성별 가구 분포: 연령별	95
[그림 3-11] 가구주 특성별 가구 분포: 자격별	96
[그림 3-12] 노인 가구원이 있는 가구 비율: 65세 이상	97
[그림 3-13] 노인 가구원이 있는 가구 비율: 65세 이상 남자	97
[그림 3-14] 노인 가구원이 있는 가구 비율: 65세 이상 여자	98
[그림 4-1] 2019년 개인총소득 분위경계값, 전체 인구	120
[그림 4-2] 2019년 개인총소득 분위경계값, 18~64세	122
[그림 4-3] 2019년 개인 성·연령별 평균 가구규모, 10인 초과 가구 제외	125
[그림 4-4] 2019년 제곱근 균등화 가구총소득 분위경계값, 전체 인구	134
[그림 4-5] 2019년 제곱근 균등화 가구총소득 분위경계값, 18~64세	136

[그림 4-6] 2019년 제곱근 균등화 가구총소득 분위경계값, 65세 이상	138
[그림 7-1] 가구 구성안별 가구 균등화 개인소득 10분위에 따른 사망률의 변화	208
[그림 7-2] 가구 구성안과 연령구간별 소득 10분위에 따른 사망률의 변화	209
[그림 7-3] 세대보완 1안(건강보험 피부양자 포함)에서 소득 10분위별 연령보정 상대사망률비(RR)	214
[그림 7-4] 세대보완 2안(비동거 가족 포함)에서 소득 10분위 연령보정 상대사망률비(RR)	216
[그림 7-5] 세대보완 3안(분리세대 추정)에서 소득 10분위 연령보정 상대사망률비(RR)	217
[그림 7-6] 세대기준안에서 소득 10분위 연령보정 상대사망률비(RR)	218
[그림 7-7] 건강보험증기준안에서 소득 10분위 연령보정 상대사망률비(RR)	219
[그림 7-8] 가구 구성안별 소득 10분위 연령보정 상대사망률비(RR)(남자)	222
[그림 7-9] 가구 구성안별 소득 10분위 연령보정 상대사망률비(RR)(여자)	222



Abstract

Exploratory Researches on the Use of Administrative Data of the National Health Insurance Service for Social Policy

Project Head: Kang, Shin-Wook

This study aims to construct household information in various ways by using individual-level data from the administrative records of the Health Insurance Service, and to analyze economic and social disparities between income groups.

The data used in this study is per the customized income-property DB built based on the National Health Information DB of the National Health Insurance Service. This data contains individuallevel information of the entire population (57.21 million people) for 13 years(2009~2021).

In this study, households are organized in five packages per the information on individual resident registration, health insurance, family relationship, and income activities of the customized income-property DB. In the five proposals, households are made up of families (1) who have same Resident Registration Household ID + NHIS dependents(H1), (2) who have same Resident Registration Household ID + non-cohabiting family members(H2), (3) who have same Resident Registration

Co-Researchers: Lee, Wonjin·Lee, Jungmin·Kang, Hee-Chung·Ha, Eunsol·Moon, Seokjun·Kim, Jaiyong·Cho, Jae-young·Kim, Eunmi·Kim, Jihye

Household ID + disguised separated households(H3), (4) who have same Resident Registration Household ID(H4), and (5) who have same NHIS card ID(H5), respectively.

As a result of the construction of households, as of 2021, H1 had the smallest number of households with 21.21 million households, and H5 had the largest number of households with 27.62 million households. There was no significant difference between H2~H4 households with 23.45~23.95 million households. The proportion of elderly household owners over 65 years old was relatively low within H1 and H5.

Chapter 3 analyzes the income distribution of customized income-property databases and compares the result with the Household Financial Welfare Survey Data. First of all, comparing the distribution of personal income found that the income capture rate of the household financial welfare survey was generally higher than that of the customized income-property DB. The proportion of individuals with earned income was higher in the Household Financial Welfare Survey, and the average of individual's public pension was also higher in the Household Financial Welfare Survey. In terms of household income distribution, H2, H3 and H4 were quite similar. But significant differences were observed between the two groups, H2, H3, H4 and H1, H5. Compared with H2, H3, and H4, the household income level of the middle- and low-income groups was relatively high in H1, and the household income level of

the middle-and high-income groups was lower in H5. Regarding the equivalized gross household income, the Gini coefficient and poverty rate of the Household Financial Welfare Survey were 0.396 and 21.1%, respectively. The Gini coefficient and poverty rate of the customized income-property DB were 0.486~0.537 and 28.1~32.3% depending on the household composition methods. The poverty rate of the elderly in H1 and H5 were 32.1% and 32.9% respectively, which were much lower than the 56.6% in the Household Financial Welfare Survey.

When comparing the average household property, H1 was the highest and H5 was the lowest. The disparity between classes of property was also lowest in H1 and highest in H5. The difference between H2 and H3 was relatively small. The property gap by income class per the quintile ratio was the lowest in H1 and the highest in H4. The income disparity among property classes measured by the Palma ratio was relatively high in H2 and H4, and lowest in H5.

In chapter 6, we derived two income-property combination indices reflecting the economic solvency of households, by applying two methods: property-annuity method and the tariff income method. It was found that the average characteristics of households do not vary significantly by the household formation methods. When comparing the case of considering income only and the case of considering both income and property, there were differences in the age of the household

head and the presence of elderly household members. However, depending on how property was converted into and combined with income, the average characteristics did not change much.

On the other hand, the consistency between the indices was not high. We found that, depending on whether the property-annuity method or tariff income method is used, the identification of the lower or upper households is quite different. This means that it should be carefully considered in policy design how to define the solvency of households.

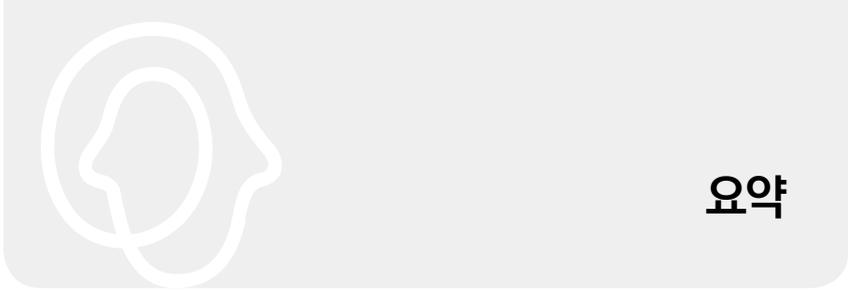
In Chapter 7, the explanatory power of the 100 percentiles of equilivalized income for the mortality rate was compared among the five household formation methods. And we also compared the five methods in terms of the relative mortality ratio of 10 income deciles. The explanatory power of the income percentile for mortality was the lowest in H5 and H1, while H2, H3 and H4 showed similar levels. Similarly, changes in the relative mortality rate ratio in the income decile showed the same pattern in H1 and H5, and the other three showed the same pattern. In H1 and H2, the mortality rate for both men and women decreased in line with the income increment. But in H3, H4, and H5, the mortality rate decreased sharply only in the top decile regardless of the gender.

In this study, we tried to propose a household composition methods by using the personal information in the customized

income-property DB. If this DB could be linked with the Population Census Data of KOSTAT, it would save a lot of effort and time for household formations and accuracy of the data. It is necessary to improve the system so that these two data can be easily linked.

If more detailed address information is available in the customized income-property DB, and, if information on other items not included in the income and property can be added or linked, richer analysis will be obtainable. As the specificity of various factors of households has varied according to the changes in the economic and social environment, it is necessary to continue theoretical researches on the concept of household in addition to the supplementing data.

Keyword : administrative data, National Health Insurance Service, household, definition of household, income distribution, wealth distribution, income-wealth combining index, mortality rate.



요약

제1장 서론

사회정책 연구에서 행정자료를 사용한 분석이 증가해 왔으나 자료상의 한계로 인해 가구 단위의 소득에 대한 분석은 충분히 이루어지지 못했고, 소득계층 간 경제·사회적 격차에 관한 연구도 제한될 수밖에 없었다. 본 연구는 건강보험공단 행정자료의 개인 단위 정보를 이용하여 다양한 방법으로 가구를 구성하고, 이를 바탕으로 소득계층 간 경제·사회적 격차를 시험적으로 분석하는 것을 목적으로 한다.

제2장 분석자료의 구축

본 연구를 위하여 국민건강보험공단의 국민건강정보 DB를 사용하였다. 국민건강정보 DB는 국민건강보험법 및 기타 법령에 따른 업무를 수행하는 과정에서 축적된 데이터를 비식별화 등 가명 처리한 뒤, 연구에 필요한 일부 자료를 추출하여 연구용 DB로 구축한 것이다.

이 DB에는 건보공단의 자격관리 및 보험료 부과 징수 업무 과정에서 수집된 소득 및 재산 자료가 포함되어 있다. 본 연구를 수행하기 위하여 경제적 수준을 측정하기 위한 자료로 맞춤형 소득-재산 DB를 구축하였다. 맞춤형 소득-재산 DB는 2009년부터 2021년까지 총 13년간의 자료이며, 매년 1월 1일 기준으로 국내에 실거주하는 전 국민을 대상으로 한다. 국내 거주 외국인과 재외국민도 포함한다.

DB에는 소득 변수로 연도별 근로소득, 사업소득, 금융소득(이자소득, 배당소득), 공적연금소득(공무원·사학·군인·국민(노령)·별정우체국연금), 기타소득이 포함되며, 재산 변수로 재산세 과세기준일 6월 1일

기준의 개인 소유인 주택, 건물, 토지, 선박·항공기, 전월세 보증금(지역 가입자)이 포함되었다. 주택, 건물, 토지, 선박·항공기의 재산 금액은 과세 표준 금액에서 공정시장가액 비율을 역산출하여 시가표준액을 제시하였다.

구축된 DB에는 2009~2021년 총 13년간의 국내 실거주자 정보 65,913만 건(실인원 5,721만 명)이 포함된다. 소득금액이 있는 인원은 2009년 39.8%에서 2021년 57.3%로, 재산금액이 있는 인원은 2009년 36.6%에서 45.6%로 연도별로 증가하는 경향을 보였다.

건보공단의 소득 및 재산 자료는 전 국민 당연가입인 사회보험 환경에서 구축된 것으로, 전 국민의 경제 수준에 대한 정보가 집약되어 있다는 점에서 의미가 있다. 다만 자료 수집의 목적이 건강보험의 자격관리 및 보험료 부과이기 때문에 관련 제도가 변화하면 자료 수집 대상에 변동이 발생하며, 조사 방식을 통해 수집한 자료와 수집항목에 차이가 있을 수 있다. 구축된 자료는 과세를 위한 소득·재산 금액으로, 연도별 물가 상승을 보정하지 않은 명목금액이며, 특히 재산의 경우 시장가격과 차이가 있을 수 있는 점을 고려해야 한다.

제3장 가구 단위 정보 생성방법 및 가구 ID 생성방법

맞춤형 소득-재산 DB에 포함된 개인별 세대 정보, 건강보험증 정보, 가족관계 정보, 소득활동 정보 등을 이용하여 다섯 가지 방법으로 가구를 구성하였다. 이 다섯 가지의 가구 구성안은 기존의 주민등록 세대를 보완한 세 가지 안(건강보험 피부양자 포함안(1안), 비동거 가족 포함안(2안), 분리세대 추정안(3안))과 세대기준안(4안), 건강보험증기준안(5안)이다.

기존의 주민등록 세대를 보완한 3개 안은 같은 곳에 거주하면 생계를 같이하는 것을 가구의 기본 개념으로 본다. 가구 구성안을 구축하기 위하여 1, 2차에 걸쳐 각 가구원의 유형을 구분하는 과정을 거쳤다.

1차 유형 구분 과정에서는 주민등록 세대별로 세대주 중심의 가구가 구성된다고 했을 때, 세대주와 각 세대원의 정보를 비교하여, 가구 구성원으로 인정 가능한 경우(세대주와 동일한 건강보험증(A), 세대주와 가족관계(B))와 그렇지 않은 경우(비가족 동거인(C))로 유형을 구분하였다.

2차 유형 구분 시에는 1차 유형에 기반하여 최종 가구 ID를 정의하게 된다. 1차 유형화가 각 세대원 본인과 같은 주소에 거주하고 있는 주민등록 세대주를 비교한 결과라면, 2차 유형 분류 시에는 주민등록 세대는 다르지만 건강보험증 대표의 가구로 묶어야 하는 상황이 고려된다. 건강보험증 대표의 가구원으로 귀속될 조건에 해당하는지를 확인하여, 이 조건에 해당한다면 ‘건강보험증 대표’의 ‘주민등록 세대 ID’ 및 ‘1차 유형(건강보험증 대표가 해당 주민등록 세대에서 가구원으로 인정되는지 여부)’에 따라 각 개인이 속할 최종 가구 ID가 정의된다. 건강보험증 대표의 가구원으로 귀속될 조건에 해당하지 않는다면 1차 유형 구분에서 잠정적으로 정의한 대로 가구가 정의된다.

1~3안의 차이는 건강보험증대표의 가구원으로 속할 조건에 있다. 1안은 각 개인이 건강보험증 대표와 다른 주민등록 세대를 구성하는 경우, 해당 건강보험증 대표의 가구로 귀속된다. 2안은 각 개인이 건강보험증 대표의 ‘배우자’이거나 ‘직계비속’인 경우, 건강보험증 대표의 가구로 귀속된다. 3안은 각 개인이 건강보험증 대표의 배우자 또는 비속이면서, 건강보험증 대표와 동일 주소일 때 건강보험증 대표의 가구로 귀속된다. 3안은 본래 같은 주소지에 거주하지만 세대가 분리된 가구를 한 가구로 식별하고자 하였으나, 자료의 한계로 주소지의 상세 정보를 이용할 수 없어 읍·면·동 지역이 같을 때 같은 가구로 정의하였다.

가구 구성 결과, 2021년 시점에서 비교해 보았을 때 1안이 2,121만 가구로 가구수가 가장 적었고, 5안이 2,762만 가구로 가구수가 가장

많았다. 1안은 주민등록 세대주 또는 건강보험증 대표를 공유하는 사람들을 한 가구로 보는 가장 포괄적인 가구 개념이다. 5안은 건강보험증 단위로, 같은 세대에 속함에도 불구하고 경제활동을 하는 가구원은 다른 가구로 분리되므로 가구수가 가장 많고 가구원 수는 적은 안이다. 2~4안에 따른 가구수는 2,345만~2,395만 가구로 큰 차이가 없었다.

가구 구성안별 특징을 보여주는 것은 65세 이상 노인에 대한 것이다. 65세 이상 노인 가구주의 경우에는 1안, 5안에서 그 비중이 상대적으로 낮고, 2~4안에서 더 높은 점이 두드러진다. 가구 구성 시 건강보험증의 영향이 큰 경우, 노인 세대가 더 젊은 연령대의 가구로 편입되는 특징을 보여준다. 65세 이상 노인이 있는 가구를 파악해본 결과, 특히 1안에서 노인이 있는 가구 비중이 가장 높았으며, 5안에서 노인이 있는 가구 비중이 가장 낮았다. 즉 1안에서 노인은 가구주가 아닌 가구원으로서 자녀 세대에 편입되는 특성을 보여준다.

본 연구는 선행연구를 참고하여, 같은 곳에 거주하면 생계를 같이한다고 보는 것을 가구 식별의 기본 원칙으로 하였다. 따라서 세대가 같을 때 가구원의 연령, 소득 활동 여부, 가족관계의 범위(촌수)에 대한 제한을 두지 않았으며, 동거인으로 신고한 경우에만 가구에서 제외하였다. 1~3안에서 세대가 다를 경우는 연령, 소득활동 여부, 가족관계 범위 등에 따라 가구원 여부를 결정할 필요가 있는데, 본 연구에서는 별도의 기준을 정하지 않고, '건강보험증'의 정보를 사용함으로써 건강보험 피부양자 인정기준을 준용하고자 하였다.

가구의 구성은 연도별 1월의 주민등록 세대 또는 건강보험증 정보에 기초하여 정의된 것이다. 따라서 실제 생계를 같이하지만 해당 시점에 다른 주소지에 거주하거나, 건강보험증이 분리되어 있는 경우 등 행정자료로 파악되지 않는 경우는 가구로 식별되지 않는다는 한계가 있다. 또한 시설

입소자로 구성되는 경우, 해당 가구의 경제적 수준을 가구원의 경제적 수준으로 보기 어려울 수 있으나, 본 자료의 정보로는 시설입소자에 대한 식별이 어려웠다. 따라서 시설입소자를 제외하고 분석해야 할 경우, 가구원 수 등을 이용한 조작적 정의를 적용할 수 있을 것이다.

제4장 소득 분포 분석: 가계금융복지조사와의 비교

제4장에서는 가계금융복지조사와 비교하여 맞춤형 소득-재산 DB의 소득 분포를 분석하였다. 우선, 가구 내에서 합산·균등화하지 않은 개인 소득 분포를 비교한 결과, 대체로 맞춤형 소득-재산 DB보다 가계금융복지조사의 소득 포착률이 높은 것으로 나타났다. 맞춤형 소득-재산 DB에 비해 서베이와 행정자료를 결합하여 구축한 가계금융복지조사에서 개인 근로소득이 존재하는 비율이 높았고, 특히 개인근로소득이 0원 초과~2,000만 원 이하인 비율이 상당히 높게 나타났다. 이는 가계금융복지조사가 조세 행정자료로 포착하지 못한 저임금 근로자의 근로소득을 상대적으로 정확하게 포착하였을 것이라는 예측을 지지한다. 또한 고소득 자영업자의 소득이 과소 보고되었을 가능성이 큰 맞춤형 소득-재산 DB에 비해 가계금융복지조사에서 개인사업소득 수준이 상대적으로 높게 나타났다. 개인 공적연금 역시 국민연금에 노령연금을 포함한 맞춤형 소득-재산 DB보다 국민연금에 노령연금, 장애연금, 유족연금 등을 포함한 가계금융복지조사에서 상대적으로 크게 포착되었다. 한편, 가계금융복지조사와 달리 비과세 원천분 금융소득 정보를 반영하지 못한 맞춤형 소득-재산 DB에서 개인금융소득이 존재하는 비율이 0.7%에 불과한 것으로 나타났다. 단, 제4장에서 맞춤형 소득-재산 DB와 가계금융복지조사의 개인소득 분포를 비교한 분석 결과는 탐색적인 수준에 그친다는 점에 주의해야 할 필요가 있다. 향후 맞춤형 소득-재산 DB가 입수한 소득 행정자료의 범위 및

정의를 꼼꼼하게 검토하여 가계금융복지조사와의 차이를 정교하게 이해하기 위한 후속 연구가 필요하다.

제4장에서는 또한 가구원의 개인소득을 합산·균등화한 가구소득 분포를 살펴보았다. 대체로 주민등록 세대를 부분적으로 보완한 2안(비동거 가족 포함안) 및 3안(분리세대 추정안)과 4안(세대기준안)의 가구소득 분포가 상당히 유사하였고, 2안, 3안, 4안의 가구소득 분포와 5안(건강보험증 기준안)의 가구소득 분포, 1안(건강보험 피부양자 포함안)의 가구소득 분포에 상당한 차이가 관찰되었다. 2안, 3안, 4안의 가구소득 분포와 비교할 때, 건강보험 피부양자인 노인을 건강보험증 대표인 자녀와 동일한 가구로 통합한 1안(건강보험 피부양자 포함안)의 가구소득 분포에서 중저소득계층의 가구소득 수준이 상대적으로 높게 나타났고, 건강보험 피부양자인 노인을 통합하지만 건강보험증을 달리하는 취업 배우자를 분리한 5안(건강보험증기준안)의 가구소득 분포에서 중고소득 계층의 가구소득 수준이 낮게 나타났다.

제4장의 분석에 따르면, 전반적으로 가계금융복지조사는 맞춤형 소득-재산 DB보다 균등화 가구총소득 수준이 높고 평등하였다. 이는 한편으로 맞춤형 소득-재산 DB에 비해 가계금융복지조사가 개인소득을 고르게 포착하여 개인총소득 수준이 높고 평등하게 나타났기 때문이기도 하고, 다른 한편으로 맞춤형 소득-재산 DB에 비해 1인 가구 비율이 훨씬 낮은 가계금융복지조사에서 가구 내 소득 공유가 강하게 반영되었기 때문이기도 하다. 근로소득, 사업소득, 금융소득, 공적연금소득을 합산하고 균등화한 가구총소득을 살펴보면, 가계금융복지조사의 지니계수 및 빈곤율은 각각 0.396, 21.1%였고, 맞춤형 소득-재산 DB의 지니계수 및 빈곤율은 가구 구성안에 따라 0.486~0.537, 28.1~32.3%로 나타났다. 단, 가계금융복지조사와 달리 1안(건강보험 피부양자 포함안) 및 5안(건강

보험증기준안)의 노인 빈곤율은 32.1~32.9%로 가계금융복지조사의 56.6%보다 훨씬 낮았다.

전반적인 분석 결과는 향후 맞춤형 소득-재산 DB의 소득 데이터를 더욱 발전시켜야 한다는 사실을 보여준다. 다양한 행정자료를 보완하여 소득 개념·범위를 확대하고 소득 측정의 정확도를 개선해야 한다. 또한 정확한 주소 및 가족관계 정보를 보완하고 실질적인 생계·주거 단위를 반영하여 가구 개념을 더 정교하게 다듬어야 할 것으로 판단된다.

제5장 가구 구성안별 재산 분포 및 소득-재산 결합 분포 분석

제5장에서는 개인의 주택, 건물, 토지 등 재산보유액 등 재산액을 이용하여 가구 단위로 합산한 후, 가구 간 재산 분포 및 소득-재산 결합 분포를 분석하였다. 가구당 평균 재산액이 가장 높은 가구 구성안은 1안(건강보험 피부양자 포함안)이고, 가장 낮은 것은 5안(건강보험증기준안)이었다. 재산의 계층 간 격차도 1안에서는 낮게, 5안에서는 높게 나타났다. 2안(비동거 가족 포함안)과 3안(분리세대 추정안)의 가구 평균 재산액 차이는 상대적으로 작았다.

가구 재산의 수준과 재산의 계층 간 분포를 중심으로 보았을 때, 1안 및 5안과 2안, 3안, 4안 사이에 차이가 존재한다는 점을 확인하였다. 1안은 가구의 범위를 확장하는 효과가 큰 것을, 5안은 반대로 가구의 범위를 축소하는 효과가 큰 것을 알 수 있다. 건강보험증 정보를 이용할 경우 소득은 적으나 재산이 있는 개인들(예, 피부양 노인)이나 반대로 소득은 있으나 재산이 적은 개인들(예, 청년층 직장가입자)의 분포가 가구 재산 분포나 소득-재산 결합 분포에 작지 않은 영향을 줄 수 있음을 염두에 둘 필요가 있다.

제6장 소득-자산 결합 방법과 적용

생애주기 가설의 관점에서 보자면 소득뿐만 아니라 재산 역시 가구의 경제적 지불능력에 영향을 미친다. 따라서 그동안 많은 학술 연구에서 소득과 재산을 모두 고려하여 경제적 지불능력을 평가하는 작업을 해왔다. 제6장에서는 맞춤형 소득-재산 DB를 이용하여 가구의 경제적 지불능력을 반영하는 소득-재산 결합지수를 도출했다. 특히 선행연구에서 제안된 대표적인 두 가지 방법인 연금소득화법과 소득인정액 환산법을 적용해 보고 결합방식에 대한 결과의 민감도와 지수 간 정합성을 살펴보았다.

분석 결과, 가구의 평균적인 특성은 결합방식에 따라 크게 변하지 않음을 발견하였다. 소득만을 고려할 경우와 재산까지 고려하는 경우, 가구주의 연령이나 고령 가구원의 유무가 달라지지만, 재산을 어떤 방식으로 환산하여 소득과 합산을 할 것인지에 따라서는 평균적인 특성이 크게 변하지 않았다. 이러한 결과는 하위 가구를 식별할 때나 상위 가구를 식별할 때 모두 동일하게 나타났다.

반면 지수 간 정합성은 높지 않은 것으로 나타났다. 재산을 고려하는 경우에도 연금화 방식을 사용하는가 혹은 기초생활보장제도에서 사용하는 소득인정액 환산방식을 사용하는가에 따라 하위 가구나 상위 가구의 식별이 상당히 다르게 나타난다는 사실을 발견했다. 이러한 사실은 정책 수립에 있어서 가구의 지불능력을 어떻게 정의할 것인가에 대해 신중하게 검토할 필요가 있다는 사실을 의미한다.

전 국민의 소득과 재산에 대한 행정데이터를 이용하여 가구의 지불능력의 분포를 추정하는 것은 경제적 취약계층을 파악하고 복지제도를 개선하는 데 도움이 될 수 있다. 경제적 취약계층은 설문조사 방식으로 파악하는 데 근본적으로 한계가 있을 수밖에 없기 때문이다. 소득과 재산에

대한 행정데이터를 연구자에게 개방하고 사용하기 쉽도록 접근성을 높이는 것은 근거 기반 정책 수립을 위한 기본적인 필요조건이라고 할 수 있다.

제7장 소득 수준별 사망률 차이에 대한 탐색적 분석

그간 개인의 사회경제적 특성과 사망률 간 관계를 실증적으로 분석하려는 수요가 있음에도 불구하고, 국내에서는 개인의 상병, 사망 정보와 사회경제적 정보의 개인수준 데이터 구축의 제약으로 인해 수요를 충족시키기 어려웠다.

제7장에서는 맞춤형 소득-재산 DB를 이용하여 5개 가구 구성안에 따라 가구 유형을 할당하고, 균등화 소득 100분위별로 사망률을 구하여 균등화 소득 100분위의 설명력을 가구 구성안별로 비교하였다. 그다음 100분위로 구분된 자료를 10분위로 재구성하여, 포아송회귀분석을 통한 상대사망률을 확인하였다.

사망률에 대한 소득분위의 설명력은 5안(건강보험증기준안)과 1안(건강보험 피부양자 포함안)에서 가장 낮았으며, 2안(비동거 가족 포함안), 3안(분리세대 추정안), 4안(세대기준안)은 유사한 수준을 보였다. 소득 10분위에서 상대사망률비의 변화도 이와 마찬가지로, 1안(건강보험 피부양자 포함안)과 5안(건강보험증기준안)이 동일한 패턴을 보였으며, 그 외에 세 개는 같은 양상을 보였는데, 앞의 두 안에서는 남녀 모두 소득이 증가할수록 사망률이 감소하였으나, 나머지 세 안의 경우 관련성이 없이 최상위 분위에서만 사망률이 급격히 감소하였다.

제8장 건보공단 행정자료의 활용도 제고를 위한 제언

본 연구에서는 건강보험공단의 맞춤형 소득-재산 DB에 포함된 개인 정보를 이용하여 가구 구성안을 제시하고자 하였다. 그런데 만일 이 DB를 통계청의 인구주택총조사와 연계할 수 있다면, 가구를 구성하기 위한 본 연구에서와 같은 많은 노력이 절감되고, 자료의 정확성도 높아질 수 있을 것이다. 이 두 자료의 연계가 용이하게 이루어질 수 있도록 제도 개선이 필요하다.

맞춤형 소득-재산 DB에서 주소 정보가 좀 더 구체적으로 주어진다면, 그리고 소득과 재산에 포함되지 않은 다른 항목들(공적 이전소득과 사적 이전소득, 금융재산과 직장가입자의 전월세 보증금 등)에 대한 정보가 추가되거나 연계될 수 있다면, 더 풍부한 분석이 가능할 것이다.

거주지를 달리하는 가족들 사이에 소득을 공유하는 방법이 다양해지고, 동일 거처에 거주하는 개인들의 구성 또한 다양해지고 있다. 경제·사회적 환경 변화에 따라 가구를 정의하는 다양한 요인들이 갖는 규정성이 달라지는 만큼, 가구 개념에 관한 이론적 연구도 지속될 필요가 있다.

주요 용어 : 행정자료, 국민건강보험, 가구, 가구 정의, 소득분배, 재산 분배, 소득-재산 결합지수, 사망률

사람을
생각하는
사람들



KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



제 1 장

서론

제1절 연구의 배경과 목적

제2절 연구의 방법과 구성

제 1 장 서론

제1절 연구의 배경과 목적

1. 사회정책 연구에서 행정자료 활용의 증가

행정자료는 행정기관이 행정 행위를 하는 과정에서 수집된 정보를 분석에 활용할 수 있는 형태의 자료(데이터)로 제공하는 것을 말한다. 자료를 수집하고 제공하는 과정에서 개인정보를 보호하기 위한 조치가 이루어지고 있으며, 특히 연구 목적으로 활용될 경우 일정한 기준에 따라 자료를 정제하고 표준화하는 등의 과정이 전제되기도 한다.

자료의 이용 측면에서 행정자료의 특징은 자료의 규모가 상대적으로 크고, 경우에 따라서는 분석 대상 집단 전체를 포괄하는 경우도 있다. 이 점이 표본 조사 자료와 가장 큰 차이점이다. 즉 행정자료를 사용할 경우 자료의 수집과 분석 과정에서 발생하는 통계적 오차를 줄일 수 있다는 장점이 있다. 또한 규모가 크지 않은 하위 집단에 대해서도 분석할 수 있다는 장점이 있다.

물론 행정자료의 한계도 있다. 행정 행위의 대상이 되지 않는 집단에 대한 정보가 포함되어 있지 않다거나, 행정 행위와 무관한 특성 항목이 빠져 있다는 한계가 있을 수 있다. 이런 점에서 행정자료를 정책분석에 사용하는 것은 다양한 장점과 한계를 동시에 내포한다.

최근 들어 사회정책 연구에서 행정자료를 이용하는 사례가 증가하고 있다. 다양한 행정자료 가운데에서도 특히 국민건강보험공단의 자료가 주목받고 있다. 국민건강보험공단의 자료는 대체로 전 국민을 포괄한다는

점에서 다른 자료와 구분되는 매우 중요한 특징을 지닌 자료이다. 이는 국민건강보험의 포괄성이 (국민기초생활보장제도의 의료급여까지를 고려할 경우) 거의 100%에 가깝다는 현실적 특징에서 비롯된 바 크다. 또한 건강보험료 부과를 위해 집적된 소득 및 재산 관련 정보를 제공한다는 점에서 국민건강보험공단의 자료는 다른 행정자료와 구분되는 특징이 있다. 전 국민을 포괄하는 주요 사회정책 관련 정보가 매년 집적되고 관리되고 있다는 점은 사회정책 연구에서 국민건강보험의 행정자료가 지니는 가치를 단적으로 설명해준다.

2. 가구 단위 정보 생성의 중요성과 행정자료의 한계

사회를 구성하는 하위 집단(sub-group) 간의 격차를 파악하고 바람직하지 않은 격차를 해소하는 것은 사회정책의 중요한 목표이자 관심사이다. 성, 연령, 지역, 인종, 국적, 문화적 배경 등 다양한 기준에 의해 구분되는 사회구성원들 간에 부당하거나 바람직하지 않은 격차를 줄이고 차별을 해소하여 사회 통합을 제고하는 것이 사회정책의 고유한 목표라고 할 수 있을 것이다.

하위 집단을 구분하는 기준은 매우 다양하다. 그 가운데 소득 수준에 따라 구별되는 집단, 즉 소득계층 간의 차이는 사회정책의 핵심적 관심사이다. 소득은 한 개인이 향유하는 행복의 크기를 정하는 데 결정적인 역할을 하며, 따라서 개인의 후생 수준에 대한 유력한 대리변수로 간주된다. 따라서 소득불평등은 그 자체로 사회적 분석의 대상이 되며, 소득계층 간에 존재하는 다른 자원 배분의 불평등, 예를 들어 건강, 주거, 교육기회, 문화적 자본 등의 불평등도 사회정책의 중요한 관심사이다.

그런데 사회과학적 분석에서 통상 소득계층은 가구를 단위로 규정된다.

가구를 구성하는 가구원들은 가구의 소득을 공동으로 사용하고(pooling) 그 소비에 따른 효과도 공동으로 누린다고 보기 때문에, 동일한 가구에 속하는 사람들은 동일한 소득계층에 속하는 것으로 간주된다. 이러한 이유로 소득분배에 대한 양적 분석에서는 가구원들의 소득을 합산한 후 가구원들에게 배분하는 방법을 적용한다.¹⁾ 전체 사회의 소득불평등도 역시 가구소득을 이용하여 계산되는 것이 일반적이다.

소득계층별 소득, 재산, 소비, 건강 등의 실태 차이를 비교하기 위해서는 소득계층을 구분할 수 있어야 하고, 이를 위해서는 가구 단위의 소득을 확정할 수 있어야 하며, 이를 위해서는 근본적으로 가구를 정의할 수 있어야 한다. 즉 어떤 개인들이 동일한 가구에 속하는지 여부를 구별할 수 있어야 한다.

그런데 국민건강보험공단 자료는 건강보험 가입과 관련된 개인 단위의 정보를 제공한다. 즉 가구 단위의 정보를 제공하지 않는 것이다. 이 자료에서는 특정 개인들이 동일한 가구에 속하는지 여부에 관한 정보도 제공되지 않으며, 가구의 수, 가구별 가구원의 수, 가구 단위의 소득 및 재산 등의 정보도 제공되지 않는다. 사실 이러한 특징은 대부분의 행정자료에서 공통적으로 나타나는 점이기도 하다. 가구 단위의 행정자료가 지극히 제한적인 이유는 행정 행위에서 가구를 대상으로 하는 경우가 드물기 때문이다. 국세 행정이나 사회보험의 경우로 한정하여 대부분 개인을 부과 징수의 대상으로 하고 있다. 건강보험의 경우, 피부양자 개념이나 세대원 개념이 있기는 하지만, 이는 전통적인 의미에서 가구 개념과 일치하지 않는다.²⁾

행정자료에 가구 단위의 소득이 제공되지 않는다면, 이를 이용하여

-
- 1) 물론 가구원들에게 합산된 소득을 균등하게 배분하지 않고 가구 균등화 지수를 적용하는 등의 조정 방법을 사용한다.
 - 2) 뒤에서 상세히 설명하겠지만, 동일 가구 안에서 다수의 건강보험증(번호)이 존재하는 경우가 적지 않다.

개인의 소득계층을 정의하는 것도 불가능하며, 따라서 소득계층 간 실태를 비교하는 것은 물론, 가구 소득불평등 자체에 대한 분석도 불가능하다. 국민건강보험공단의 행정자료가 사회정책 분야의 연구와 정책 수립에서 활용도가 높아지도록 하기 위해서는 개인 단위의 정보를 이용하여 가구 단위의 정보를 생성하는 작업이 필요하다. 가구를 구성하고 가구 단위의 소득을 계산하는 것이 각 소득계층별 생활실태를 파악하고 계층 간 격차를 파악하기 위한 전제 조건이기 때문이다.

3. 연구의 목적

이 연구는 개인 단위로 수집된 국민건강보험공단의 행정자료를 이용하여 가구 단위의 정보를 구성함으로써 소득계층 간 격차 분석에 이 자료를 활용할 수 있을지 탐색하는 것을 목적으로 한다. 본 연구가 '탐색적' 연구인 까닭은 개인 단위의 정보를 이용하여 가구 단위 정보를 형성하고자 시도했던 선행연구들이 거의 없었으며, 특히 국민 전수의 개인소득 자료를 이용해서 전 국민의 가구소득 분포를 추정하고자 시도했던 연구는 처음이라고 할 수 있기 때문이다. 예를 들어, 행정자료를 이용하여 가구 단위의 분석을 시도하였던 전병유 외(2018b), 구인회 외(2020), 이현주 외(2020)의 연구는 주민등록상의 세대를 가구로 간주하여 분석하였다. 전병유 외(2018a)와 강신욱 외(2021)의 연구는 가구를 구성하는 과정 없이 개인 단위의 분석을 수행하였다.

하지만 더 근본적으로는, 가구 개념 자체가 매우 복잡하고 각 정책 영역에서 상이한 의미로 사용된다는 점이 본 연구가 탐색적 성격을 갖는 중요한 이유이다. 통상 가구는 같은 공간에 거주하면서 생계를 공동으로 영위하는 단위로 이해된다. 하지만 경우에 따라 일시적으로 다른 공간에 거주하면

서도 공동으로 생활을 영유하고 있는 혈연관계의 사람들은 같은 가구를 이루고 있는 것으로 이해되기도 한다. 이 점이 같은 공간(주소) 거주를 중심으로 규정되는 주민등록상의 세대 개념과 다른 점이다. 통계청이 매년 인구총조사를 통해 발표하는 한국의 가구수가 행정안전부가 발표하는 세대 수와 다른 이유도 이 때문이다.

그런데 인구총조사가 전국 가구의 수와 인구학적 특성 분포 등에 대한 정보는 제공하지만, 전체 가구의 소득 분포와 관련된 정보는 조사하지도 제공하지도 않는다. 가구 단위의 소득 분포에 대한 모집단 정보가 존재하지 않는 상태에서 본 연구가 가구 단위의 소득 정보를 생성한다고 하더라도 그것이 현실과 일치하는지를 확인할 방법은 없다. 하지만 다양한 방식으로 가구 단위의 정보를 생성하고, 이를 이용하여 일부 주제에 대한 심층적 분석을 수행한 후 상호 비교할 수 있다면, 가구 단위의 특성에 대한 개략적 모습을 파악하는 데 도움이 될 수 있을 것으로 판단한다. 또한 각종 행정 자료를 이용하여 가구 단위 정보를 생성시키려는 이후의 연구와 분석에 본 연구의 시도가 중요한 참고가 될 수 있을 것으로 기대한다.

제2절 연구의 방법과 구성

1. 연구의 방법

개인 단위로 정보를 제공하는 행정자료를 이용하여 가구 단위의 정보를 생성하는 과정에서 우선적으로 검토해야 할 것은 가구 개념을 정의하고 이해하는 다양한 방식들이다. 이를 위해 본 연구에서는, 우선 국내의 주요 통계 및 행정자료에서 가구가 어떻게 이해되고 있는지를 비교했다. 나아가 국외의 사례를 분석해 가구에 대한 정의 방식이 어떻게 변해왔고 관점에

따라 어떻게 달라질 수 있는지를 살펴보았다. 특히 통계조사와 관련된 준거를 제공하려는 국제기구의 논의를 소개하고 참고하고자 하였다. 이러한 이론적 연구를 바탕으로 국민건강보험공단의 자료에 포함된 변수들과 정보를 이용하여 가구를 정의하는 복수의 방법을 도출하고자 하였다. 동일한 세대에 거주하는지 여부, 동일한 건강보험증을 사용하는지 여부, 소득활동 여부, 혈연(가족) 관계 여부 등의 정보를 이용하여 다양한 방식으로 가구 식별키를 생성하였는데, 이때 각각의 방식은 가구를 정의하는 핵심적인 규정성들과 맞닿아 있도록 하였다.

이렇게 가구 정보를 생성하고 난 후에는 실증적인 분석을 수행하였다. 다양한 방법으로 정의된 가구 구분방식을 이용하여 가구 소득을 계산하였고, 이를 이용하여 소득분배 실태, 소득계층별 재산보유 실태, 소득계층별 사망률 격차 등에 관해 실증적으로 분석하였다. 각각의 분석 주제와 관련된 구체적인 분석 방법론에서는 해당 장에서 좀 더 자세히 소개하도록 할 것이다.

2. 연구의 구성

본 연구는 다음과 같이 구성된다. 연구의 배경과 목적을 소개하는 제1장에 이어 제2장에서는 연구의 기본적인 분석 자료가 된 국민건강보험공단의 행정자료에 대해 상세하게 소개한다. 제2장에서는 본 연구에서 활용한 데이터셋에 포함된 각 정보의 자료원을 소개하고, 이들을 이용하여 데이터셋을 구축하는 과정과 결과를 소개할 것이다.

제3장은 제2장에서 소개된 데이터셋을 이용하여 가구 단위의 정보를 생성하는 과정을 설명할 것이다. 이를 위해 가구 개념과 관련된 국내외의 문헌들을 우선 검토할 것이다. 그런 후 개인 단위의 정보를 연계하여 각

개인들을 가구로 묶는 다섯 가지 방법들을 상세하게 소개할 것이다. 그리고 각 방법에 따라 주요 가구 특성들의 분포가 어떻게 다르게 나타나는지 비교할 것이다.

제4장부터 제7장까지는 이렇게 구축된 자료를 이용한 주제별 실증 분석을 수행한다. 먼저, 제4장에서는 개인과 가구소득의 분포를 분석하고 소득불평등 지표를 산출한다. 그리고 이 결과를 다른 자료를 이용하여 산출한 선행 연구들의 소득 불평등도와 비교하는 한편, 여러 해의 자료를 이용하여 소득분배 변화에 대한 시계열 분석을 시도할 것이다.

제5장에서는 소득과 재산의 결합 분포에 대한 분석을 시도한다. 제3장까지의 논의를 통해 가구 단위의 소득은 재산도 계산할 수 있는 만큼, 소득 계층별 재산보유 실태에 대한 다양한 분석을 시도하는 것이 제5장의 주요 내용이다.

제6장에서는 가구 단위의 소득과 자산을 하나의 정보로 결합하는 방법에 대해 검토한다. 소득과 자산이 각각 다른 성격을 지닌 변수이기는 하지만 모두 가구의 지불능력을 설명한다는 공통점이 있는 만큼, 이 두 변수를 결합하는 방법은 중요한 이론적, 정책적 관심사였다. 제6장에서는 연금 소득화법과 소득인정액 환산법을 국민건강보험공단의 자료에 적용하고, 그 결과에 대한 함의를 제공하도록 한다.

제7장에서는 소득계층에 대한 구분을 다른 영역의 정보와 결합하여 분석하는 작업을 시도하였다. 제3장까지의 분석을 통해 도출된 가구와 개인의 소득지위가 사망률의 차이에 대해 어느 정도의 설명을 제공할 수 있을지를 탐색하고자 한다.

마지막 제8장은 본 연구의 주요 결과를 요약하고, 이후의 연구를 위한 제안을 제시한다. 본 연구를 ‘탐색적’ 연구임을 전제로 수행한 만큼 데이터 셋 자체의 보완과 발전에 대해서도 많은 여지가 있으며, 이를 이용한 분석

26 국민건강보험공단 행정자료의 사회정책적 활용을 위한 탐색적 연구

에도 많은 한계점이 있는 것이 사실이다. 국민건강보험공단의 행정자료가 사회정책 관련 연구에 더 유용하게, 더 활발하게 이용될 수 있으려면 어떤 자료상의 보완과 연구·분석 측면의 발전이 필요한지 제안할 것이다.

사람을
생각하는
사람들



KOREA INSTITUTE FOR HEALTH AND SOCIAL AFFAIRS



제2장

분석자료의 구축

제1절 자료원 소개

제2절 자료의 구축

제3절 소결

제 2 장 분석자료의 구축

제1절 자료원 소개¹⁾

1. 국민건강정보 DB 개요

국민건강보험공단(이하 ‘건보공단’)의 국민건강정보 DB는 국민건강보험법 및 기타 법령에 따른 업무를 수행하는 과정에서 축적된 데이터를 비식별화 등 가명 처리하여 연구용 DB로 구축한 것이다. 국내 실거주 전 국민 자격 및 보험료, 진료내역, 건강검진 결과, 노인장기요양 정보 등 다양한 데이터가 수집되어 있으며, 가명 처리 후 정책 및 학술 연구에 필요한 일부 자료를 추출하여 연구용 DB로 구축하여 제공하고 있다.

건보공단은 2012년 6월, 전 국민 건강보험 자료를 비식별화 등 가명 처리하여 국민건강정보 DB로 구축한 후, 2008년부터 시행된 노인장기요양 자료를 추가하는 등 지속해서 업데이트를 추진하였다. 2014년 7월부터 국민건강정보 DB를 연구자에 제공하였으며, 2016년 4월 연구자 대상 DB 분석센터를 운영하였다. 2017년 6월부터 전 국민을 대상으로 표본 추출한 DB를 추가 구축·제공했고, 2019년 5월부터 맞춤형 DB 구축·분석 시스템을 운영하고 있다.

국민건강정보 DB는 2002년부터 현재까지 20년 이상의 전 국민 자료가 구축된, 국가를 대표하는 보건의료 분야의 자료원이다. 우리나라의 보건의료·복지 정책 지원을 위한 다양한 근거를 생성하여 국민건강 향상을

1) 국민건강보험공단(2012), 국민건강보험공단(2018)을 참고하여 작성함.

위한 지식 기반 구축에 기여하고 있다.

2. 국민건강정보 DB의 데이터 범위

국민건강정보 DB는 <표 2-1>과 같이 다양한 영역의 자료를 포함하고 있다. 의료이용 및 건강상태 관련 데이터뿐만 아니라, 출생, 출산(분만), 가계구조, 사망 등 인구학적 영역, 거주 지역 및 기간, 이동거리 등의 지리학 영역, 가구 및 세대, 직업, 국적 등 사회학적 영역, 보험료, 소득, 재산, 경제적 취약계층 등 경제학적 영역, 지역별 보건의료·복지 분야의 인력, 시설, 자원 분포 등과 관련된 사회자원 차원의 영역을 포괄하고 있다.

<표 2-1> 국민건강정보 DB의 영역별 주요 데이터 내용

구 분	내 용
인구학적 영역	출생, 출산(분만), 가계구조, 사망 등
지리학적 영역	거주 지역 및 기간, 이동거리 등
사회학적 영역	가구 및 세대, 업종, 국적 등
경제학적 영역	보험료, 소득, 재산, 경제적 취약계층 등
사회문제해결 및 취약계층	건강위험 노출집단, 경제적 취약계층, 장애인 등
사회자원	지역별 보건의료 및 사회복지 분야 인력, 시설, 자원 분포
개인 건강 영역	장애, 생활습관, 건강검진(신체계측, 일반건강상태) 등
개인 질병 및 의료이용 영역	주요 질환별 입원/외래 이용, 입원에피소드 등

자료: 국민건강보험공단. (2018). 국민건강정보 DB 2.0 매뉴얼. 내부자료.

3. 국민건강정보 DB의 공개 및 이용 절차

가. 연구 DB 제공 유형

현재 연구용으로 제공되고 있는 국민건강정보 DB의 경우 크게 ‘맞춤형

연구 DB', '표본 연구 DB'로 구분되어 있다.

'맞춤형 연구 DB'란 국민건강정보 DB를 연구자 개인의 정책 및 학술 연구를 수행하기 위한 목적으로, 이용에 적합하게 가공하여 제공하는 데이터를 말하며, 연구자 개인의 맞춤형 연구 DB는 건보공단 내의 지정된 장소인 '빅데이터 분석센터'에서만 열람 및 분석할 수 있다.

'표본 연구 DB'란 국민건강정보 DB를 주요 목적에 따라 전 국민, 노인, 건강검진 수검자 등 연구 대상에 적합한 기준으로 표본 추출하여 제공하는 데이터로, 전체 50만 명~100만 명의 연구 대상자의 자격, 진료, 건강검진 등의 주요 정보를 포함하고 있다.

나. 연구 DB 제공 및 공개를 위한 심의절차 운영

'맞춤형 연구 DB' 및 '표본 연구 DB' 등 연구용 DB를 이용하기 위해서는 데이터를 제공받기 위한 심의를 거쳐야 한다. 심의는 연구계획에 근거한 개별 신청에 의해 이루어진다. 통계 목적, 과학적 연구, 공익적 기록보존 등의 사유로 신청 가능하며, 연구 주제와 대상자, 연구 목적 및 내용 등을 검토하여 국민건강정보 자료 제공심의위원회에서 승인 여부를 의결한다. 심의·승인 절차에 따라 해당 연구계획상 필요하다고 인정되는 자료의 범위 내에서 공개·제공된다. 또한 연구별 연구참여자에게만 승인된 자료에 접근할 수 있도록 하고 있다. 세부 신청절차 및 필요 서류 등은 건보공단 건강보험자료 공유서비스 웹사이트를 참고하면 된다.

〈표 2-2〉 국민건강정보 DB 자료 공개 및 제공 절차

건강보험공단 자료공유사이트에서 신청 → 심의 → 자료구축 → 비용고지·납부 → 연구분석 수행 → 연구종료/성과물 등록
--

출처: 국민건강보험공단 공유서비스. <https://nhiss.nhis.or.kr/bd/ab/bdaba001cv.do>

제2절 자료의 구축

1. 활용 자료원

가. 개요

1) 인구집단 정의를 위한 자료

기본 자료는 국민건강정보 DB의 인구집단 DB(월별)을 사용한다. 이 DB는 국내에 실거주하는 전 국민을 포함한다. 자격 정보와 주민등록 정보를 바탕으로, 해당 월에 존재한 거주자를 정의한다. 국내 거주 외국인과 재외국민도 포함된다.

각 월에 존재한 가장 이른 시점의 건강보험 자격 정보(건강보험증 대표 ID, 건강보험증 대표와의 관계), 주민등록가구 정보(세대주 ID, 세대주와의 관계), 행정동 주소코드, 출생연도, 성별 등의 정보로 구성되어 있다. 모든 개인정보는 비식별화 등 가명 처리되어 있다.

2) 소득 및 재산 자료

건보공단은 국민건강보험법 제14조에 근거하여, 가입자 및 피부양자의 자격관리, 보험료 부과·징수 업무를 수행하고 있으며, 이를 위하여 건강보험가입자의 소득 및 재산 자료를 수집한다.

건강보험료 부과체계는 2017년 국회에서 통과된 건강보험료 부과체계 2단계 개편안에 따라, 2018년 1단계 개편과 2022년 2단계 개편이 시행되었다. 2단계 개편은 직장/지역 가입자 간 부과방식의 차이에 따른 형평성 문제, 고소득/고재산 피부양자의 무임승차 문제를 보완하고, 소득 중심

보험료 부과 원칙을 강화하는 데 초점을 맞추고 있다(보건복지부 2022).

개편의 목적을 요약하면, 공보험의 중요 원칙인 ‘능력에 따른 부담과 필요에 따른 이용’ 원칙을 실현하고자 하는 것이다. 여기에서 ‘경제적 능력 (=부담 능력)’과 부과된 보험료 사이의 관계를 모니터링하는 것은 부과체계 개편 정책의 단기 효과를 평가하는 작업과 연결된다. 다른 한편, 경제적 능력 → 합리적 부담 → 공보험의 유지 → 적정이용 → 건강수준 사이의 관계를 모니터링하는 것은 부과체계 개편의 중·장기 효과를 평가하는 것이다. 게다가 건강수준의 향상은 그 자체가 국민건강보험의 궁극적 목적인 가치의 측면과 경제적 능력의 전제조건이 된다는 수단의 측면을 함께 가지고 있다. 따라서 경제적 능력과 건강수준 사이의 관계를 평가하는 것은 국민건강보험의 자격·부과·급여 정책과 관련된 일련의 의사결정 과정에 중요한 과학적 근거를 제시할 수 있다. 또한, 2018년 1단계 부과체계 개편에서 지역가입자의 상당수에게 최저보험료가 부과됨으로써 보험료만으로는 정책으로 중요한 인구집단인 지역가입자 중·저소득 계층을 제대로 구분하기 어렵게 되었다. 보험료가 경제적 수준을 구분하는 변별력이 훼손됨에 따라 앞에서 기술한 연구 목적을 달성하기 위해서도 비식별화 등 가명 처리된 소득 및 재산 데이터의 활용이 필요했다.

그간 건보공단의 자료를 이용한 분석에서 경제적 수준을 설명하는 변수로 ‘월별 산정보험료’ 및 그 ‘분위’가 자주 사용되어 왔다. 그러나 직장 및 지역가입자는 보험료 산정 방식이 달라 직장가입자, 지역가입자 각각으로 구분하여 사용해야 했으며, 원천 자료에 따라 사용되는 소득과 재산의 귀속시점이 다른 한계가 존재하였다.²⁾ 이에 따라 동일 시점의 항목 간, 개인 간 비교가 가능하도록 귀속시점을 고려하였으며, 비식별화 등 가명

2) 예를 들어, 지역가입자의 2018년 2월 보험료는, 2016년 발생 소득, 2017년 소유 재산, 2018년 1월 소유 자동차로 산출한다.

34 국민건강보험공단 행정자료의 사회정책적 활용을 위한 탐색적 연구

처리한 후 DB를 구축하였다. 자료의 출처에 따라, 건보공단의 보험료 정산에 따른 직장가입자 보수, 국세청에 신고되는 종합소득 및 분리과세 소득, 공적 연금기관에서 지급하는 연금소득, 지자체 과세 대상 재산으로 구성된다.

〈표 2-3〉 자료원별 생산기관 및 자료확보·보험료 적용 시점

자료명		생산기관	원자료 귀속시점	자료 확보 (보험료 적용시점)	연구 DB 구축
소득	직장가입자 보수	연말정산	건보공단	전년도	3월 (4월~익년 3월)
		퇴직·중간정산		해당연도	수시 (매월)
	종합소득(이자·배당·사업·근로) 분리과세소득(이자·배당)		국세청	전년도	10월 (11월~익년 10월)
연금소득 (공무원, 사학, 군인, 국민, 별정우체국 연금)		5대 공적연금 기관	전년도	매월 (1월~12월) ¹⁾	
재산	재산세, 주택세, 종합토지세 (건물, 토지, 주택, 선박, 항공기 과세표준)		지방자치 단체	해당연도 (6월 1일 기준)	9월~10월 ²⁾ (11월~익년 10월)

주: 1) 귀속연도 2014년부터 전년도 연금소득이 1월 보험료부터 적용됨. 귀속연도 2013년 이전 자료는 매년 해당연도 11월 보험료부터 부과에 적용됨.
 2) 단 2005년 귀속자료는 익년 7월부터 보험료 부과에 적용됨.
 자료: 국민건강보험공단. (2021a, 2021b)를 참고하여 재구성.

각 자료는 수집 목적과 시기가 다르며, 건보공단에 자료가 확보되는 시점도 상이하다. 이에 따라 연구 목적의 소득-재산 DB를 구축할 때는 귀속시점을 기준으로 하여 기준연도를 재정의하였다.

지역가입자의 보험료 부과 시, 근로소득과 연금소득에는 평가율(20, 30%)이 적용되어 사용되나, 본 DB에는 평가율 적용 이전의 실제 소득 금액 그대로 구축되어 있다.

3) DB 공개 여부

본 연구는 건강보험 빅데이터를 활용하는 방향을 모색하고자, 건보공단 업무의 일환으로 수행된 것이다. 본 자료는 원칙적으로 「국세기본법」 등 수집의 근거 법령에 의해, 목적 외 사용이 제한된다. 이에 본 연구에서는 건보공단 소속이 아닌 연구자는 원자료에 접근하지 않았다. 건보공단 직원에 의해 가공된 맞춤형 소득-재산 DB 자료를 앞에서 기술한 연구 개발 목적을 달성할 수 있는 범위 내에서 사용하였다.

나. 소득 및 재산 변수 현황

1) 근로소득

직장가입자의 경우, ‘근로자 또는 법인사업장 대표’는 직장가입자 보수 자료 및 종합소득세 과세 대상 근로소득 자료로 구축하되, 두 원천자료에서 중복이 발생할 수 있으므로 중복분은 차감했다. ‘개인사업장 대표’는 종합소득세 과세 대상 근로소득을 이용했다. 지역가입자는 종합소득세 과세 대상 근로소득을 이용하여 구축하였다. 건강보험료 부과 대상인 근로소득은 소득 공제 전의 총급여액으로, 소득공제분을 차감하는 「소득세법」상 소득과 차이가 있다.

2) 사업소득

직장가입자의 사업소득은, ‘근로자·법인사업장 대표’의 경우 종합소득세 과세 대상 사업소득, 분리과세 주택임대소득이 포함된다. ‘개인사업장 대표’는 직장가입자 보수 자료, 종합소득세 과세 대상 사업소득, 분리과세

주택임대소득 자료로 구축하되, 중복이 있으면 중복분은 차감한다. 지역 가입자의 사업소득은 종합소득세 과세 대상 사업소득, 분리과세 주택 임대소득이 포함된다.

건강보험료 부과제도가 개정되어 귀속연도 2019년 소득분부터 분리 과세 주택임대소득³⁾이 보험료 부과 대상 소득에 포함되었다(2020년 11월 보험료부터 적용). 이에 따라 본 연구의 분석자료 중 기준연도 2020년(소득의 귀속연도 2019년)과 2021년의 자료에는 분리과세 주택임대 소득이 추가되었다.

3) 금융소득(이자소득, 배당소득)

금융소득은 이자소득과 배당소득으로 구분된다. 이자소득은 종합소득세 및 분리과세 대상 이자소득이 포함되며, 배당소득은 종합소득세 및 분리 과세 대상 배당소득이 포함된다.

건강보험료 부과제도 개정에 따라, 귀속연도 2019년 소득분부터 분리 과세 금융소득⁴⁾이 보험료 부과 대상 소득에 포함되었다(2020년 11월 보험료부터 적용). 이에 따라 본 연구의 분석자료 중 기준연도 2020년(소득의 귀속연도 2019년)과 2021년의 자료에는 분리과세 금융소득(이자소득, 배당소득)이 추가되었다.

4) 연금소득

연금소득은 공무원·사학·군인·국민·별정우체국연금의 5대 공적

3) 총수입금액의 합계액이 연 2천만 원 이하 시 분리과세

4) 이자 및 배당소득의 합계액이 연 1천만 원 초과~연 2천만 원 이하

연금기관에서 지급한 연간 총연금액이 포함된다. 다만, 국민연금의 경우 노령연금만 포함되며, 장애연금 및 유족연금 등 노령연금 외의 연금은 포함되지 않는다. 또한 개인연금, 퇴직연금 등의 사적연금도 포함되지 않는다.

건강보험료 부과 대상인 연금소득은 소득공제 전의 총공적연금액으로, 사적연금이 포함되고, 소득공제분을 차감하는 「소득세법」상 소득과 차이가 있다.

〈표 2-4〉 소득유형별 소득금액 출처

구 분	직장 가입자		지역 가입자
	① 근로자, 법인사업장 대표	② 개인사업장 대표	
근로소득	건강보험료 정산 (직장가입자 보수) 국세청 신고 소득 (종합소득세)	국세청 신고 소득 (종합소득세)	국세청 신고 소득 (종합소득세)
사업소득	국세청 신고 소득 자료(종합소득세, 분리과세 주택임대소득)	건강보험료 정산 (직장가입자 보수) 국세청 신고 소득 (종합소득세, 분리과세 주택임대소득)	국세청 신고 소득 (종합소득세, 분리과세 주택임대소득)
금융 소득	이자소득 배당소득	국세청 신고 소득 (종합소득세, 분리과세 금융소득)	
기타소득	국세청 신고 소득 (종합소득세)		
연금소득	공적 연금소득		

자료: 국민건강보험공단. (2021a, 2021b)를 참고하여 재구성.

5) 재산 변수 현황

주택, 건물, 토지, 선박, 항공기 등 「지방세법」에 따른 과세 대상 재산에 대해, 연도별 재산세 과세표준 금액을 시가표준액(공시가격)으로 변환하여 구축하였다. 개인 소유인 재산만 포함되며, 지방세법에 따라 매년 6월1일

시점(과세기준일)의 재산으로 구축되어 있다.

주택은 아파트, 연립, 다세대, 단독, 다가구 등이 해당된다. 원자료의 주택 과세표준 금액은 과세를 위하여 시가표준액에 공정시장가액 비율 60%가 적용된 금액이므로, 60% 적용 전 금액으로 변환하였다. 건축물은 주택 이외의 건물에 해당한다. 원자료의 건축물 과세표준 금액은 과세를 위하여 시가표준액에 공정시장가액 비율 70%가 적용된 금액이므로, 70% 적용 전 금액으로 변환하였다. 토지는 나대지, 사업용 토지, 농지, 공장용지 등이 해당된다. 원자료의 토지 과세표준 금액은 과세를 위하여 시가표준액에 면적과 공정시장가액 비율 70%가 적용된 금액이므로, 70% 적용 전 금액으로 변환하였다. 선박과 항공기는 시가표준액 그대로 과세표준 금액이 되므로, 금액을 변환하지 않았다.

전월세 보증금은 국민건강정보 소득재산 DB에 포함되지 않아, 보험료 부과 대상인 전월세 보증금 자료를 활용하였다. 이 자료는 부과를 위한 기초자료로, 주로 주택 또는 건물을 보유하지 않은 지역가입자(세대주)의 자료이다.

〈표 2-5〉 재산유형별 과세표준

구 분	과 세 표 준
주 택 ¹⁾	시가표준액 × 공정시장가액 비율 ³⁾ (60%)
건 축 물 ²⁾	시가표준액 × 공정시장가액 비율 ³⁾ (70%)
토 지 ¹⁾	시가표준액 × 면적 × 공정시장가액 비율 ³⁾ (70%)
선박, 항공기 ²⁾	시가표준액

주: 1) 「부동산 가격공시 및 감정평가에 관한 법률」에 따라 공시된 가액.
 2) 지방자치단체 결정가액.
 3) 공정시장가액 비율은 구축된 데이터에 해당하는 기간 중 동일(귀속연도 2008년부터).
 자료: 국민건강보험공단. (2021a). 지역보험료 부과관리 업무처리지침 재구성.

다. 자료의 장점 및 한계점

건보공단의 소득 및 재산 자료는, 전 국민 당연가입인 사회보험 환경에서 구축된 것으로, 전 국민의 경제 수준에 대한 정보가 집약되어 있다는 점에서 의미가 있다. 사회경제적 수준은 보건복지 분야뿐만 아니라, 여러 학문 분야에 있어 특정 현상을 설명할 수 있는 핵심적 요인이다.

다만, 이 자료는 건강보험료 부과를 위해 2차적으로 수집된 자료이므로, 보험료 부과제도의 변화에 영향을 받는다. 즉, 보험료 부과 제도가 변화하면 자료 수집 대상이 변동할 수 있다는 것이다. 예를 들어, 귀속연도 2019년부터 분리과세 금융소득 및 분리과세 주택임대소득이 지역가입자 보험료 부과 대상 소득으로 포함되었다. 또한 2022년 9월분 보험료부터 지역가입자 재산에서 부채를 경감하는 보험료 부과제도⁵⁾가 시행되는 등 부과제도는 계속해서 변화하고 있다(보건복지부 2022; 보건복지부 2020).

전월세 보증금과 관련해서는 전체 임대료를 대표할 수 없는 자료라는 한계가 있다. 현재 우리나라의 주택임대차 계약에 관한 사항은 의무 신고 대상이 아니며, 2023년 6월부터 국토교통부의 전월세신고제가 본격 시행될 예정이다(국토교통부 2022). 현재 보험료 부과는 가입자 본인이 건보공단에 신고한 임대차 계약 사항에 따르거나, 건보공단에 신고되지 않는 경우 주소에 기반한 추정금액에 따라 이루어진다. 본 자료는 후자의 경우가 우세하므로 실제 전월세 보증금액의 시세와 차이가 있을 것으로 판단된다.

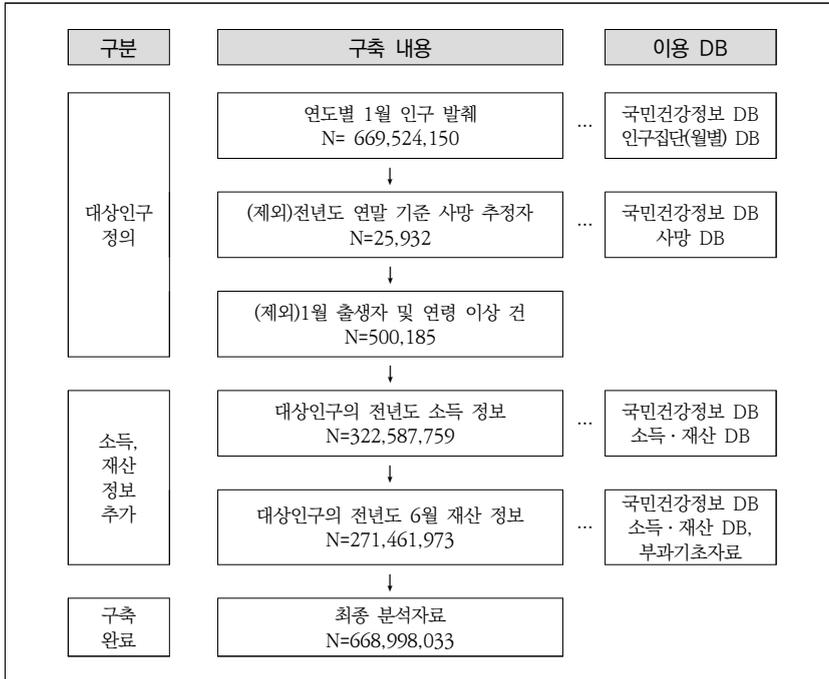
5) 지역가입자의 재산에 부과하는 보험료 산정 시, 1주택자이면서 공시가격 5억 원 이하 주택 등 일정 조건에 해당하는 경우,

2. 맞춤형 소득-재산 DB 구축과정

맞춤형 소득-재산 DB 구축을 위하여 먼저 인구집단 DB(월별)로부터 연도별 1월 대상인구를 발췌하였다. 이 중 전년도 12월 말까지의 사망 추정자를 제외하였으며, 해당 연도 1월 출생자와 연령 이상치를 제외하였다. 이 과정을 거쳐 매년 1월 1일 기준의 인구집단 자료를 구축하였다.

대상인구를 정의한 뒤, 소득 및 재산 정보를 추가하기 위해 국민건강정보 소득·재산 DB와 부과 기초자료를 연계하는 과정을 거쳤다. 소득은 전년도 연간 소득이며, 재산은 전년도 6월 1일 기준이다.

[그림 2-1] 구축 과정(2009~2021)



주: N은 13년간의 전체 건수가 포함됨.

2009년 4,993만 명부터 2021년 5,251명까지 총 13개년 기간의 66,900만 명의 연도별 자료가 구축되었다. 약 0.5%의 연중 사망자가 포함 되어 있다. 소득금액이 있는 인원은 2009년 39.8%에서 2021년 57.3%로, 재산금액이 있는 인원은 2009년 36.6%에서 45.6%로 연도별로 증가 하는 경향을 보였다.

〈표 2-6〉 연도별 구축과정별 인원(2009~2021)

(단위: 만 명, %)

구분	대상인구 정의				소득재산 정보 구축		
	1월 인구	제외 대상		최종 분석 대상	연중 사망자	소득 (전년도 연간)	재산 (전년도 6월)
		전년 연말 기준 사망 추정자	1월 출생자, 연령 이상 건				
2009	4,998	0.2	4.4	4,993 (100%)	25 (0.5%)	1,987 (39.8%)	1,825 (36.6%)
2010	5,025	0.2	4.3	5,021 (100%)	25 (0.5%)	2,089 (41.6%)	1,905 (37.9%)
2011	5,055	0.2	4.8	5,050 (100%)	26 (0.5%)	2,155 (42.7%)	1,941 (38.4%)
2012	5,085	0.1	4.5	5,080 (100%)	27 (0.5%)	2,205 (43.4%)	1,970 (38.8%)
2013	5,111	0.1	4.5	5,106 (100%)	27 (0.5%)	2,285 (44.7%)	1,998 (39.1%)
2014	5,137	0.1	4.2	5,133 (100%)	27 (0.5%)	2,362 (46.0%)	2,031 (39.6%)
2015	5,167	0.2	4.2	5,163 (100%)	28 (0.5%)	2,456 (47.6%)	2,032 (39.4%)
2016	5,192	0.2	4.0	5,188 (100%)	28 (0.5%)	2,549 (49.1%)	2,100 (40.5%)
2017	5,211	0.2	3.5	5,208 (100%)	29 (0.5%)	2,642 (50.7%)	2,178 (41.8%)
2018	5,225	0.1	3.3	5,222 (100%)	30 (0.6%)	2,739 (52.5%)	2,210 (42.3%)
2019	5,232	0.2	3.1	5,229 (100%)	30 (0.6%)	2,830 (54.1%)	2,250 (43.0%)
2020	5,259	0.4	2.7	5,256 (100%)	31 (0.6%)	2,949 (56.1%)	2,312 (44.0%)
2021	5,254	0.3	2.5	5,251 (100%)	32 (0.6%)	3,011 (57.3%)	2,396 (45.6%)
전체	66,952	2.6	50.0	66,900 (100%)	362 (0.5%)	32,259 (48.2%)	27,146 (40.6%)

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

3. 구축 결과

1) 분석 대상의 기본적 특성

맞춤형 소득-재산 DB에는 2009~2021년 총 13년간의 국내 실거주자 정보 65,913만 건(실인원 5,721만 명)이 포함된다.

2021년 기준, 내국인 5,136만 명(97.8%), 외국인 115만 명(2.2%)으로 구성되었는데, 2009년에 비해 외국인 비율이 다소 증가하였다(〈표 2-7〉).

〈표 2-7〉 연도별 내국인, 외국인 구성(2009~2021)

(단위: 만 명, %)

구분	내국인		외국인		전체	
	인원	비율	인원	비율	인원	비율
2009	4,953	99.2%	41	0.8%	4,993	100%
2010	4,976	99.1%	44	0.9%	5,021	100%
2011	4,999	99.0%	51	1.0%	5,050	100%
2012	5,024	98.9%	57	1.1%	5,080	100%
2013	5,046	98.8%	60	1.2%	5,106	100%
2014	5,067	98.7%	67	1.3%	5,133	100%
2015	5,086	98.5%	77	1.5%	5,163	100%
2016	5,105	98.4%	83	1.6%	5,188	100%
2017	5,120	98.3%	87	1.7%	5,208	100%
2018	5,130	98.2%	92	1.8%	5,222	100%
2019	5,134	98.2%	95	1.8%	5,229	100%
2020	5,136	97.7%	119	2.3%	5,256	100%
2021	5,136	97.8%	115	2.2%	5,251	100%
합계	65,913	98.5%	987	1.5%	66,900	100%

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

국민건강보험 가입자 유형별로는, 2021년 기준 직장가입자 3,678만 명(70.0%), 지역가입자 1,421만 명(27.1%), 의료급여 152만 명(2.9%)으로 구성된다(〈표 2-8〉).

성별로는 남녀 각각 약 50% 수준으로 연도별로 비슷한 수준이다. 연도별, 연령별 인원 분포는 최근일수록 20세 미만 연령대의 인구가 감소하며, 60대 이상 연령대의 인구가 증가하는 경향을 보인다(〈그림 2-2〉~〈그림 2-4〉).

〈표 2-8〉 연도별 가입자 유형별 구성(2009~2021)

(단위: 만 명, %)

구분	직장가입자		지역가입자		의료급여		전체	
	인원	분율	인원	분율	인원	분율	인원	분율
2009	3,031	60.7%	1,780	35.6%	183	3.7%	4,993	100%
2010	3,130	62.3%	1,725	34.3%	166	3.3%	5,021	100%
2011	3,228	63.9%	1,656	32.8%	166	3.3%	5,050	100%
2012	3,318	65.3%	1,602	31.5%	160	3.1%	5,080	100%
2013	3,404	66.7%	1,553	30.4%	150	2.9%	5,106	100%
2014	3,493	68.0%	1,495	29.1%	145	2.8%	5,133	100%
2015	3,559	68.9%	1,462	28.3%	142	2.8%	5,163	100%
2016	3,618	69.7%	1,418	27.3%	152	2.9%	5,188	100%
2017	3,654	70.2%	1,404	27.0%	150	2.9%	5,208	100%
2018	3,677	70.4%	1,392	26.7%	153	2.9%	5,222	100%
2019	3,678	70.3%	1,403	26.8%	148	2.8%	5,229	100%
2020	3,689	70.2%	1,418	27.0%	149	2.8%	5,256	100%
2021	3,678	70.0%	1,421	27.1%	152	2.9%	5,251	100%
합계	45,156	67.5%	19,730	29.5%	2,014	3.0%	66,900	100%

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

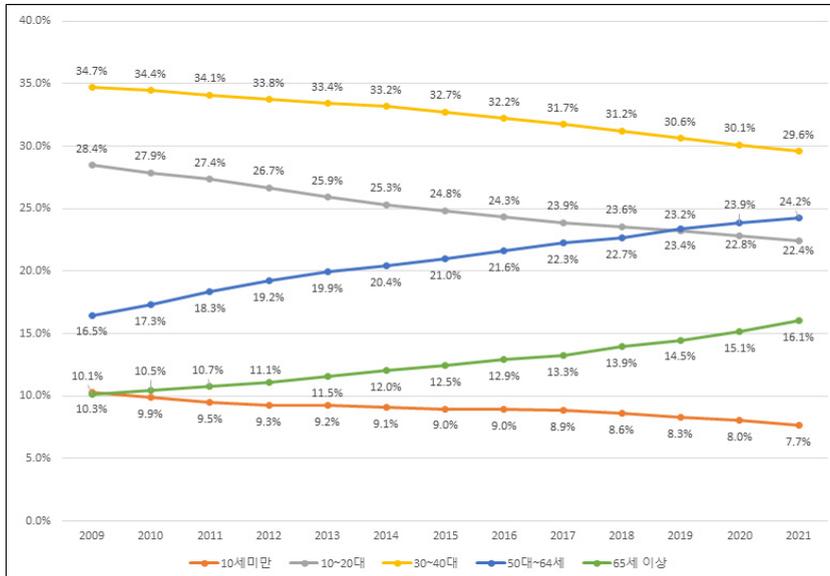
44 국민건강보험공단 행정자료의 사회정책적 활용을 위한 탐색적 연구

[그림 2-2] 연도별 성별 분포(2009~2021)



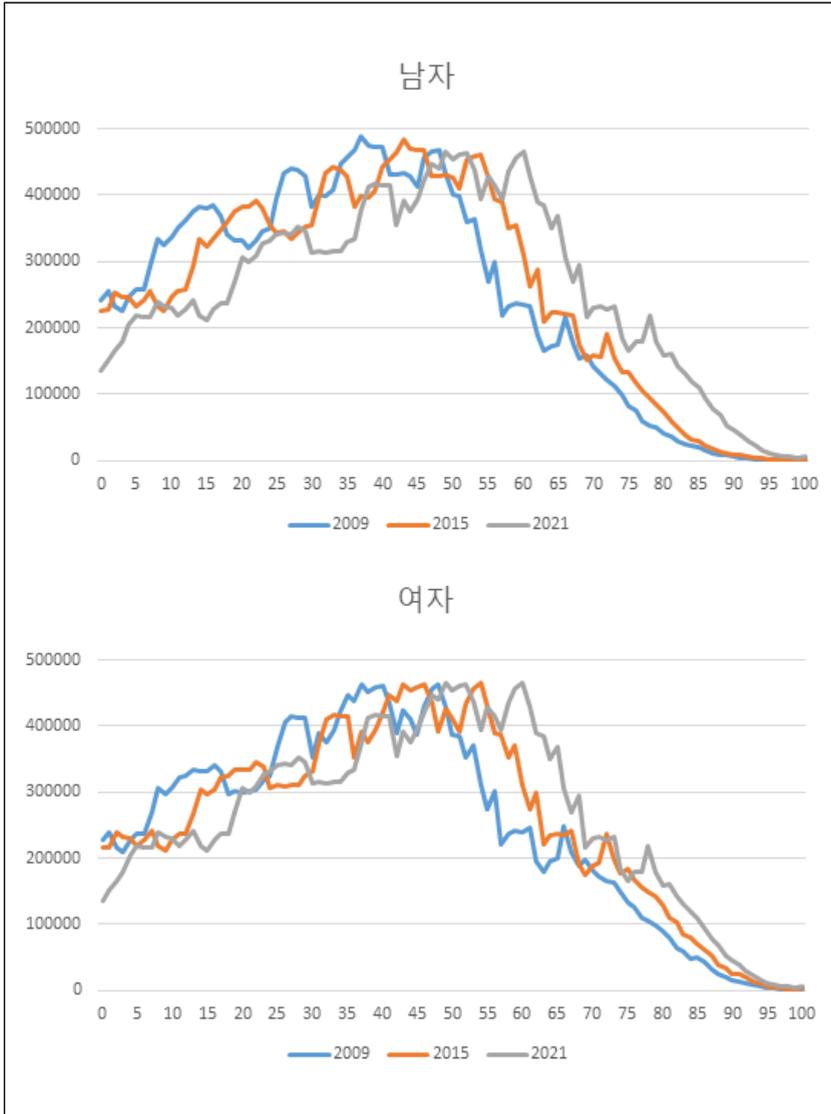
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

[그림 2-3] 연령대별 분포(2009~2021)



자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

[그림 2-4] 성별 연령대별 분포(2009, 2015, 2021)



자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

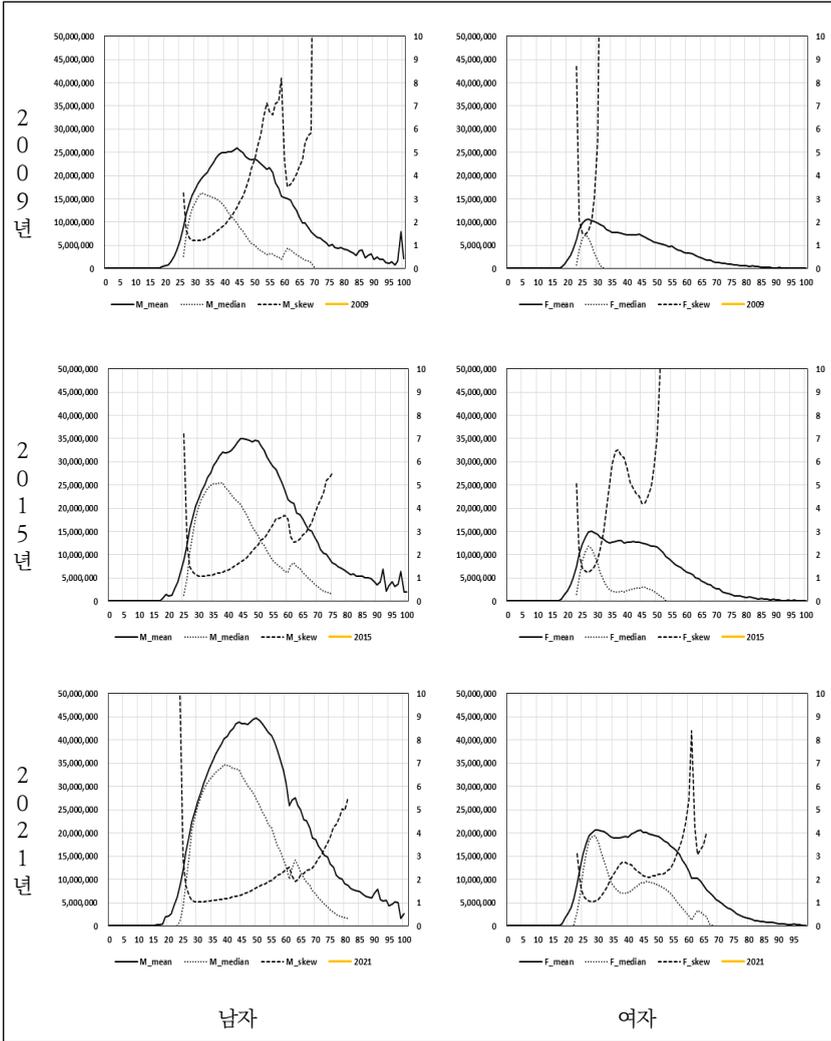
2) 소득 및 재산 관련 기초 현황

개인 단위 총소득금액을 기반으로, 성별, 연령별 소득의 기초 분석 결과는 [그림 2-5]와 같다. 평균 소득을 중심으로 보면, 남자는 40~45세 구간에서 가장 높으며, 2009년에 비해 2015년, 2021년에서 20~80세 구간 전반에 걸쳐 증가세를 보인다. 여자는 2009년에는 25~30세 구간에서 평균 소득이 가장 높지만, 2021년에는 30~50세로 넓은 범위의 연령대에서 평균 소득이 높은 결과를 보여준다. 가장 소득 수준이 높은 연령대를 기준으로 볼 때, 여자는 남자에 비해 소득이 50% 이하 수준이다.

개인 단위 총재산금액을 기반으로, 성별, 연령별 재산의 기초 분석 결과는 [그림 2-6]과 같다. 평균 재산을 중심으로 보면, 남자는 2009년, 2015년에 60~70세 구간에서 가장 높다, 2021년에는 75세 전후로 가장 높으며, 2009년에 비해 30세 이상 구간에서 전반적으로 상승하였다. 여자는 2009년, 2015년에 60세 전후로 평균 재산금액이 가장 높고, 2021년에는 55~70세 구간에서 높다. 가장 재산금액이 높은 연령대를 기준으로 볼 때, 여자는 남자에 비해 약 50% 또는 그 이하 수준이다.

다만 이 결과는 구축된 자료의 현황을 제시하기 위한 것으로, 연도별 물가 상승을 보정하지 않은 명목 금액인 점을 고려하여야 한다.

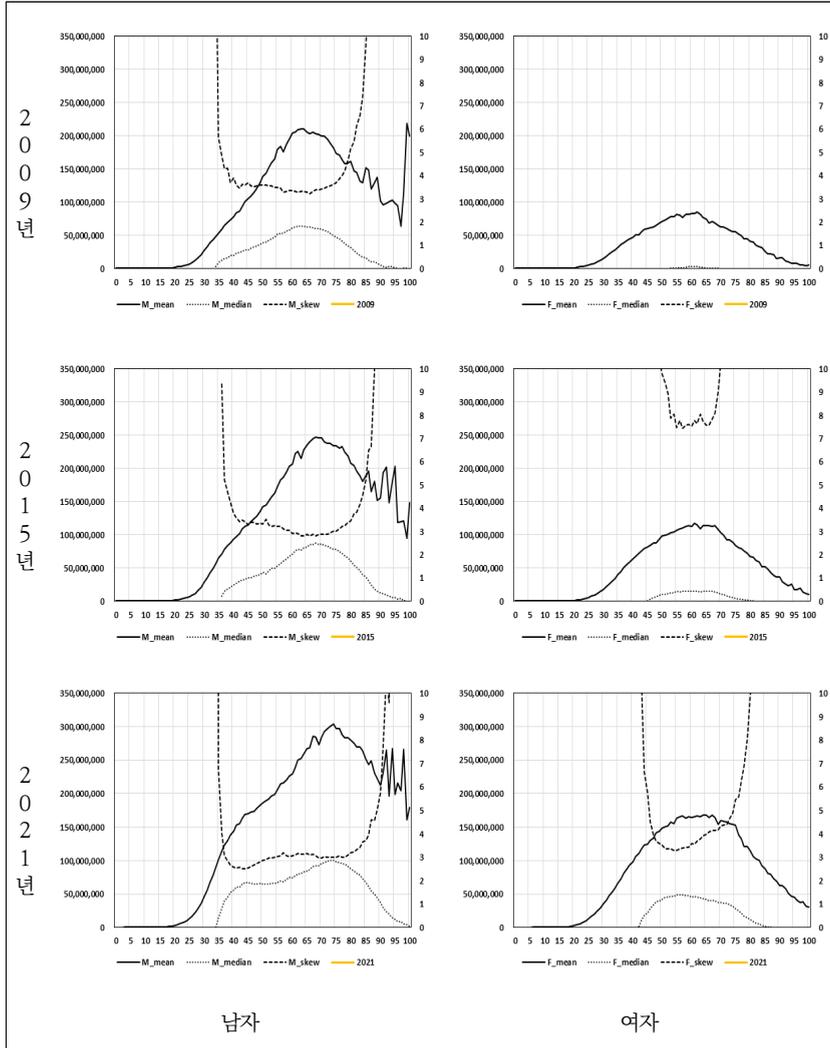
[그림 2-5] 성별, 연령별 총소득의 평균(mean), 중위값(median), 왜도지표(Skewness-index)



주: 'Skewness-index'는 각 '연령'별 평균을 중위값으로 나눈 값(mean/median)으로, 이 값이 클수록 평균과 중위값의 차이가 커, 소득이 특정 인원에 쏠려 있음을 의미함.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

48 국민건강보험공단 행정자료의 사회정책적 활용을 위한 탐색적 연구

[그림 2-6] 성별, 연령별 총재산의 평균(mean), 중위값(median), 왜도지표(Skewness-index)



주: 'Skewness-index'는 각 '연령'별 평균을 중위값으로 나눈 값(mean/median)으로, 이 값이 클 수록 평균과 중위값의 차이가 커, 재산이 특정 인원에 쏠려 있음을 의미함.

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

제3절 소결

연구를 위하여 비식별화 등 가명 처리된 국민건강보험공단의 국민건강정보 DB를 사용하였다. 이 DB에는 건보공단이 자격관리 및 보험료 부과 징수 업무 등을 수행하여 수집한 소득 및 재산 자료가 포함되어 있다. 본 연구를 위해서 건보공단 직원에 의해 맞춤형 소득-재산 DB가 구축되었다. 2009년부터 2021년까지 총 13년간의 자료이며, 국내에 실거주하는 전 국민을 대상으로 하고 있다.

이 자료는 전 국민이 가입한 국가사회보험 환경에서 구축된 것이라는 점에서 가치를 가진다. 다만 수집의 목적이 건강보험의 자격관리 및 보험료 부과이기 때문에, 관련 제도가 변화하면 자료 수집 대상에 변동이 발생하며, 조사 방식을 통해 수집한 자료와 수집항목에 차이가 있을 수 있다. 구축된 자료는 과세를 위한 소득·재산 금액으로, 연도별 물가 상승을 보정하지 않은 명목금액이며, 특히 재산의 경우 시장가격과 차이가 있을 수 있는 점을 고려해야 한다.



제3장

가구 단위 정보 생성 방법 및 가구 ID 생성방법

제1절 가구 개념에 대한 문헌 검토

제2절 가구 구성 방법

제3절 가구 단위 정보 생성 결과

제4절 소결

제 3 장

가구 단위 정보 생성 방법 및 가구 ID 생성방법

제1절 가구 개념에 대한 문헌 검토

가구는 통상 소득과 지출을 공유하면서 생계를 함께하는 사람들의 결합체로 이해된다. 공동으로 생계를 영위한다는 점에서 같은 거처에 거주할 것, 공동으로 음식이나 생활 필수품을 제공할 것, 친지(related person)일 것 등이 강조되기도 한다. 이 가운데 어떤 요인을 강조하는가에 따라 가구를 구분하는 기준에 차이가 있는 것이다.

가구에 대한 엄밀한 정의가 가장 강조되는 것이 통계조사, 그 가운데에서도 인구총조사(센서스)일 것이다. 조사를 하려면 어떤 기준에 따라 개인들을 동일한 가구의 가구원으로 볼 것인가를 사전에 정의해야 한다. 인구 조사와 마찬가지로 소득조사에서도 가구의 개념이 중요하다. 기본적으로 각 개인들은 가구의 소득을 공유하여 사용하는 것으로 간주되기 때문이다. 그리고 시기와 목적, 조사를 실시하는 사회적 상황에 따라 가구를 정의하는 방식이 다소 달라질 수도 있다(OECD, 2013). 이 절에서는 국제기구 및 인구조사 관련 국제적 논의에서 제안된 권고안들을 검토하고, 국내 주요 조사에서 활용하고 있는 가구 정의 방식을 정리하였다.

1. 가구 정의와 관련된 국외의 논의

가. UN World Populations Census

UN에서 1958년 처음 발행된 “Principles and recommendations

for population and housing censuses”는 세계 각국의 국가 통계 자료 생산을 지원하기 위한 인구 및 가구 센서스 조사의 국제적인 표준 권고안으로, 2017년 3차 개정판까지 발행되었다(UN, 2017). UN에서 권고하는 세계 인구 센서스에 따르면, 가구 개념은 1인 가구와 다인 가구로 구분되며, 생계를 같이하는 가구 개념(housekeeping concept)에 근거한 가구 정의를 채택하고 있다. 이때 1인 가구는 다른 가구원 없이 식료품과 기타 생활 필수품을 스스로 마련하는 가구이고, 다인 가구는 식료품과 기타 생활 필수품을 함께 마련하는 가구로 정의되며, 혈연이나 법적 관계가 아니어도 한 가구로 구성된다. 이러한 생계 공동체 기준에 따르면 가구원은 서로 다른 거처에 살 수 있지만, 하나의 가구로 정의될 수 있다. 가구원은 서로 떨어져 살아도 생계를 같이 하면 한 가구이기 때문에, 한 거처에 살고 있는 인원과 가구의 가구원 수가 반드시 일치할 필요가 없다.

나. UNECE/CES Population Census

유엔 유럽경제위원회(United Nations Economic Commission for Europe, UNECE)는 CES(Conference of European Statisticians) 회원국을 대상으로 실시하는 인구 및 가구 센서스를 위한 가이드라인을 제공한다. 2006년에 공개된 UNECE의 CES 권고안은 일부 국가의 경우 생계를 같이하는 가구 개념에 근거한 자료를 생산할 수 없다는 점을 고려하여 거주지 기반의 가구 개념(household-dwelling concept)을 제안하였다(UNECE, 2006, p.106). 여기서 말하는 거주지 기반의 가구 개념이란, 동일한 거처에 살고 있으면 하나의 가구로 정의하는 것을 말한다. 이에 따라 거처 단위에 살고 있는 사람의 수와 가구원 수는 일치하고, 거처 단위와 가구의 위치가 모두 동일하게 된다. 이러한 가구 개념은 실질적으로 한 가구이나 따로 살고 있는 가구의 경우 가구의 동거형태를 알 수

없다는 한계가 있다.

CES 권고안은 가구를 일반가구, 시설가구, 기타가구로 구분하고 있다. 먼저, 일반가구는 독립된 주택에 혼자 사는 사람 또는 주택 내 별도의 공간에서 다른 거주자와 다인 가구를 형성하지 않는 1인 가구 또는 주택 전체 또는 일부에서 거주하며 식료품 및 기타 생활 필수품을 함께 마련하는 2인 이상 다인 가구로 정의된다. 시설가구는 임시거처와 보급품이 시설에 의해 제공되는 가구로, 이때 시설은 장기 거주 목적 또는 서비스를 제공하기 위한 합법적 기관에 해당된다. 기타가구는 일반, 시설 가구에 포함되지 않는 가구로, 주로 노숙인이 여기에 해당한다.

다. ICLS

ICLS(International Conference of Labour Statisticians)는 ILO에서 개최하는 국제 컨퍼런스로서, 17차(2013) 컨퍼런스에서 가계 소득 지출 통계에 관한 가이드라인을 마련하였다. ICLS는 UN의 권고안을 기반으로, 동일하게 생계를 같이하는 가구 개념을 적용하고 있다. 가구는 일반가구와 집단가구로 구분되며, 모든 일반가구는 통계에 포함되나, 집단가구는 조사시 집단의 구성원들이 소비에 대한 의사결정을 하는 집단이라면 포함하고, 그 외는 제외할 것을 권고하고 있다. 시설인구는 일반적으로 제외하나 시설 내 식별 가능한 가구가 있다면 포함될 수 있다고 본다.

라. Canberra Group

캔버라그룹(Canberra Group)은 가구 소득 통계 국제 전문가 기구로 2011년 “Handbook on Household Income statistics”를 발간했다. Canberra Group은 UNECE/CES 기준에 근거한 가구 정의를 제안하고

있다. 이러한 정의는 가계소득조사(household income survey)의 기준으로 여겨지고 있으며, 혈연뿐만 아니라 비혈연만으로 구성된 가구도 가능하다는 것이 특징이다.

마. Euro area HFCS

Euro 지역의 가계금융소비조사(Household Finance Consumption Survey)에서 채택된 정의로, 가구는 혼자 사는 가구 또는 함께 살면서 지출을 공유하고 생활 필수품을 함께 마련하는 가구로 정의된다. 구체적인 가구원 기준을 제시하고 있는데 가구원이 지출을 공유하면서 재정적 독립성 조건을 추가하고 있는 점이 특징이다. 따로 살아도 지출을 공유하는 경우 다음과 같은 조건을 충족하면 가구원으로 보고 있다. 1) 휴가, 직장, 교육 등 이유로 일시적으로 부재하나 다른 곳에 개인 주소가 없고 실제 또는 예상 부재기간이 6개월 이내인 경우, 2) 실제 부재기간과 상관없이 배우자나 자녀가 가구와 밀접한 유대 관계를 유지하고, 정기적으로 거처에 돌아오는 경우, 3) 병원, 요양원, 기숙사, 기타 시설에 머무는 가구로, 가구에 재정적으로 의존하거나 실제 또는 예상 부재기간이 6개월 이내인 경우 가구원에 포함한다.

바. SNA

국민계정체계(System of National Accounts, SNA)에서 가구는 개인 또는 집단으로 구성되는 제도적 단위로 정의한다. 이때 다인 가구는 소득과 자산 일부 또는 전체를 함께 모으고 특정 재화(특히 주거와 식료품)와 서비스를 공동으로 소비하며 동일한 거처에서 함께 사는 개인들로 정의된다.

(표 3-1) 국외 조사 가구 기준 검토

구분	가구 정의			센서상 조사대상 가구 기준		
	가구 범위	가구 개념	가구 개념상 인구 범위	통계적 제약에 따른 제외	가구의 거처 기준	일상적 거처 기준
UN 세계 인구 센서스	-	<p>housekeeping concept (생계를 같이하는 가구 개념)</p> <ul style="list-style-type: none"> - 1인 가구: 다른 가구원 없이 식료품과 기타 생활 필수품을 스스로 마련하는 가구 - 다인 가구: 식료품과 기타 생활 필수품을 함께 마련하는 2인 이상 가구 • 혈연·법적 관계 아니어도 가구 구성 가능 	<p>시설인구 제외</p> <p>관습적 주택 거주자 및 주택 외 집단 거처 거주자(캠프, 하숙집, 호텔 등) 및 노숙인 포함</p>	-	<p>가구는 주택 전체 또는 일부, 여러 개 주택에서도 거주할 수 있음</p> <p>≠ household dwelling concept</p>	<p>12개월 기준 적용 권고:</p> <ul style="list-style-type: none"> - ① 최근 12개월의 대부분(일시적 부재 제외) 지속적으로 거주했거나, 최소 6개월 이상 거주할 곳 - ② 최근 12개월 이상 거주했거나(일시적 부재 제외), 최소 12개월 이상 거주할 곳
UNECE/CBS 인구 센서스	일반 시설 기타	<p>household-dwelling concept (거주지 기반의 가구 개념)</p> <ul style="list-style-type: none"> • 일부 국가의 경우 housekeeping 개념에 근거한 일반가구 자료를 얻을 수 없음으로 어려움 <p>- 일반가구: 독립된 주택에 혼자 사는 사람 또는 주택 내 별도의 공간에서 다른 거주자와 다인 가구를 형성하지 않는 1인 가구 또는 주택 전체 또는 일부에서 거주하며 식료품 및 기타 생활 필수품을 함께 마련하는 2인 이상 다인 가구</p> <ul style="list-style-type: none"> • 일반가구는 주택 외 집단 거처 거주자 일부도 포함될 수 있음. 	모든 인구	-	<p>일반 가구는 주택 전체 또는 일부에서만 거주</p>	<ul style="list-style-type: none"> - 사람들이 매일 밤 머무는 곳으로, 각 사람은 단 하나의 일상적 거처를 보유. <p>특수 상황에 대한 권고:</p> <ul style="list-style-type: none"> - 일 때문에 몇 주간 집을 비운 경우, 학기 중 집을 비운 경우(거처=가정), 대학을 다니는 동안 집을 비운 학생(거처=거주지, 단, 교육기관이 자국에 있는 경우, 거처=가정), 병원, 요양원, 교도소 등 시설에

구분	가구 정의			센서스상 조사대상 가구 기준		
	가구 범주	가구 개념	가구 개념상 인구 범위	통계적 제약에 따른 제외	가구의 거처 기준	일상적 거처 기준
		<ul style="list-style-type: none"> • 하숙은 가구원, 자취는 별도 가구로 간주. - 시설가구: 임시거처와 보금자리 시설에 의해 제공되는 사람들로, 이때 시설은 장기 거주 목적 또는 서비스 제공 위한 합법적 기관임 • 기숙사, 병원, 요양원, 생활 및 복지시설, 군대, 교정 및 형사시설, 종교시설, 직원숙소 - 기타가구: 일반, 시설거구에 포함되지 않는 가구(노숙인 등) 				<p>서 12개월 이상을 보냈거나 보낼 곳(거처=시설). 한 해 동안 두 곳 이상의 거주지에 머무름(거처=더 오래 머문 곳)</p>
ICIS	일반 집단	<ul style="list-style-type: none"> - 시설인구가 제외된다는 점에서 UN 세계 인구 센서스 기준과 거의 동일함 	UN 세계 인구 센서스 기준과 동일	<p>모든 일반가구 포함, 집단가구 중 주거 서비스 이용 가구 포함</p> <p>그 외 집단가구(하숙, 호텔 등)와 시설내거는 제외됨(단, 시설 내 식별 가능한 가구는 포함될 수 있음)</p>	<p>가구는 주택 전체 또는 일부, 여러 개 주택에서도 거주할 수 있음</p>	-

구분	가구 정의			센서상 조사대상 가구 기준		
	가구 범주	가구 개념	가구 개념상 인구 범위	통계적 제약에 따른 제외	가구의 거처 기준	
퀵버라 그들	일반 시설 기타	<ul style="list-style-type: none"> - 모든 인구 대상이라는 점에서 <u>UNECE/CES</u> 기준과 동일함 - 가계소득조사와 기준 - 혈연 및 비혈연 가구도 가능 	모든 인구	주택에 거주하는 일반가구만 포함 주택 외 거처 거주자 및 시설가구 제외	일반가구는 주택 전체 또는 일부에서만 거주	-
	Euro area HFCS	<ul style="list-style-type: none"> - 가구 자산조사 목적이라는 특수성, 지출 공유라는 점 이외에 재정적 독립성 조건 추가. - 혼자 사는 가구 또는 함께 살면서 지출을 공유하고 생활 필수품을 함께 마련하는 가구 - Euro지역의 가계금융소비조사에 적용 	시설인구를 제외한 모든 인구	국내 거주 일반가구만 포함. 집단가구(주택 외 거처 거주자), 시설인구 제외	일반가구는 주택 전체 또는 일부에서만 거주	지출을 공유하면서, 다음 경우에 해당하는 경우 가구원으로 분류: - 휴가, 직장, 교육 등 이유로 일시적으로 부재하나 다른 곳에 개인 주소가 없고 실제(예상) 부재기간이 6개월 이내, 실제 부재기간과 상관없이 배우자나 자녀가 가구와 밀접한 유대 관계를 유지하고, 정기적으로 거처에 돌아오는 경우, 병원, 요양원, 기숙사, 기타 시설에 머무는 가구로, 가구에 재정적으로 의존하거나 실제(예상) 부재기간이 6개월 이내

구분	가구 정의			센서상 조사대상 가구 기준	
	가구 범주	가구 개념	가구 개념상 인구 범위	통계적 제약에 따른 제외	가구의 거처 기준
SNA	가주 비가주	- 소득과 자산 일부 또는 전체를 함께 모으고 특정 제화와 서비스를 공동으로 사용(주거, 식료품 등)하며 함께 사는 가구 또는 혼자 사는 가구	국내 모든 인구	거주자가 구만포함(시설인구 포함)	하나 이상의 주거, 시설, 그 외 유형

주: 일상적 거처란 센서스 기간 당시 조사 대상자의 거처로, 열마칸 머물렀거나 머무를 것으로 예상되는 곳. 일상적 거처는 조사대상 가구의 기준이 됨.
 자료: OECD. (2013). OECD guidelines for micro statistics on household wealth. pp.47-49, pp.275-277를 참고하여 재구성함.

2. 국내 주요 조사에서의 가구 개념

가. 가계금융복지조사

통계청의 가계금융복지조사는 자산, 부채, 소득, 지출 등을 통해 미시적 재무건정성을 파악하여 사회 및 금융 관련 정책과 연구에 활용하기 위한 목적으로 실시하는 조사이다(통계청, 2020b, p.4). 가계금융복지조사에서 가구란 1인 이상이 모여 주거 또는 소득과 지출 등 생계를 같이 하는 경제적 생활 단위를 말한다. 가계금융복지조사는 가구의 재정 상태를 파악하기 위해 인구주택총조사의 가구 개념과는 다른 가족 단위에 가까운 개념을 채택하였으며, 학업 때문에 떨어져 살고 있는 미혼 자녀와 직장 때문에 외지에 살고 있는 배우자는 포함하고, 함께 살고 있는 가사사용인 등 동거인은 제외하고 있다(통계청, 2021, p. 26). 이런 개념에 따라 가계 금융복지조사는 조사제외(부적격) 가구를 여섯 가지 유형으로 명시하고 있다. <표 3-2>의 가구에 해당할 때 조사대상 가구로 보지 않는다.

<표 3-2> 가계금융복지조사 조사제외 가구

코드	조사제외(부적격) 가구
1	학업 때문에 본가에서 떨어져 살고 있는 미혼 자녀(학생, 학원생) 가구
2	직장 때문에 일시적으로 떨어져 살고 있는 배우자 가구
3	15세 미만 가구원으로 구성된 가구
4	전체 가구원이 사회시설(사회복지시설, 노인요양시설 등)에 거주하는 가구
5	비혈연 가구원만으로 구성된 가구
6	외국인으로만 구성된 가구

자료: 통계청. (2021). 2021년 가계금융복지조사 지침서. p.26.

가계금융복지조사에서 가구원은 가구주와 주거 또는 소득과 지출 등 생계를 같이하는 사람을 말한다. 이는 주민등록 등재 여부와 관계없이

사실상 가구를 구성하고 있는 구성원을 뜻하며, 취업이나 학업 등의 사유로 함께 생활하지 않더라도, 생계(소득과 지출)를 같이 하는 가구원은 조사 대상임을 의미한다(통계청, 2021, p. 27). 가계금융복지조사의 조사에 포함되는 가구원과 조사에서 제외되는 가구원의 구분은 다음과 같다.

〈표 3-3〉 가계금융복지조사 조사대상 포함 및 제외 기준

구분	기준
조사 포함 대상	<ul style="list-style-type: none"> - 혈연, 혼인, 입양 등으로 같은 거처에서 생계를 함께하는 가족 - 직장 때문에 떨어져 살고 있는 배우자 - 학업으로 인해 타 지역이나 외국에서 하숙, 자취하거나 기숙사에 있으면서 조사 대상 가구의 부담으로 공부를 하고 있는 직계 가족 - 시설(사회복지시설, 노인요양시설 등)에 단기(12개월 미만) 거주하는 가족 - 공익근무요원, 산업체근무요원으로 출·퇴근하는 가족 - 미결수나 구류처분을 받고 경찰서에 수감 중인 가족 - 선박 등의 승무원으로 기숙사가 아닌 선박 등에서 생활하고 있는 가족 - 가족과 함께 살고 있는 직업군인
조사 제외 대상	<ul style="list-style-type: none"> - 결혼·취업 등으로 생계를 같이하지 않는 분가한 형제나 자녀 - 같은 집에 살고 있으나 소득 및 지출을 달리하는 친인척 - 함께 살고 있는 가사사용인, 운전사, 하숙생 등 동거인 - 시설(사회복지시설, 노인요양시설 등)에 장기(12개월 이상) 거주하는 가족 - 군대, 의무경찰대 등으로 입대(의무 복무)한 가족 - 형이 확정되어 교도소, 소년원 등에 수용된 가족

자료: 통계청. (2021). 2021년 가계금융복지조사 지침서. pp.26-27.

나. 인구주택총조사

통계청이 매년 실시하는 인구주택총조사는 우리나라 모든 인구, 가구, 주택의 총수, 규모, 구조, 분포와 개별 특성까지 파악하여, 각종 경제·사회 발전계획 수립 및 평가는 물론이고 학술연구, 기업경영 등의 기초자료를 제공하기 위한 목적으로 실시되는 조사이다(통계청, 2020d, p.10). 인구주택총조사에서 가구란 1명이 단독으로 생계를 영위하거나 2명 이상이 공동으로 취사, 취침 등을 하며 생계를 영위하는 생활단위를 말한다

(「인구주택총조사 규칙」 제2조). 이때 가구의 가구원은 혈연이나 주민등록과 관계없이 이 가구에서 실제로 같이 살고 있는 사람에 해당한다(통계청, 2020). 조사를 위해 조사대상에 포함되는 사람과 조사에서 제외되는 사람은 다음과 같은 기준으로 나뉜다.

〈표 3-4〉 인구주택총조사 조사대상 포함 및 제외 기준

구분	기준
조사 포함 대상	<ul style="list-style-type: none"> - 출장, 친지 방문, 여행, 입원, 환자 간호 등으로 잠시 집을 떠나 있는 사람 - 예비군 훈련 또는 사회 복무(공익 근무) 요원으로 근무 중인 사람, 군부대 밖에서 거주하는 직업군인 - 선박, 항공기, 관광버스 등의 승무원 - 숙식을 함께하는 가사 도우미와 하숙인 - 한국인과 함께 사는 외국인
조사 제외 대상	<ul style="list-style-type: none"> - 군인, 의무경찰 등으로 입대한 가족 - 학업 때문에 따로 살고 있는 가족 - 직장 때문에 따로 살고 있는 가족 - 보육원, 노인요양시설(요양원, 요양병원), 부녀보호시설 등 사회복지시설에서 살고 있는 가족 - 외교관, 외국 정부 또는 국제기구 등에서 공무로 체류 중인 국내 거주 외국인 - 국내 주둔 외국 군인, 군무원 및 그 가족

자료: 통계청. (2020g). 2020 인구주택총조사 조사표. p.1.

인구주택총조사의 가구는 일반가구, 집단가구, 외국인가구로 구분된다. 집단가구, 집단시설가구를 분리하고, 특별조사구 대상자(의무소방대원, 의무경찰, 해외 주재 공간원 등)는 인구에는 포함하나 가구에서는 제외된다. 또한, 외국인가구는 외국인등록부를 통해 등록거주지 기반으로 통합 생성된다(통계청, 2020d, p.80). 인구주택총조사는 행정자료 보안을 통해 가구를 통합 및 분리 보완하여 조사대상 가구를 확정하고 있으며, 그 내용은 〈표 3-5〉와 같다.

64 국민건강보험공단 행정자료의 사회정책적 활용을 위한 탐색적 연구

〈표 3-5〉 인구주택총조사의 행정자료를 통한 가구보완

구분	보완 내용
집단가구, 집단시설가구 통합	집단시설 입소자는 집단시설가구 명부(대학기숙시설 입소자 명부, 사회시설입소자 명부)를 활용하여 하나의 집단시설가구인데 여러 가구로 분리된 경우 통합
동일 주소 친족가구 통합	주민등록상 동일 주소에 여러 세대가 있는 경우 가족관계등록 자료를 활용하여 친족인 경우 통합
동일 거처 가구 통합	일반단독주택, 공동주택(아파트, 연립, 다세대), 오피스텔은 1처처 1가구가 살도록 되어 있으므로, 전입전출로 인한 일시적으로 2인 가구 이상 이면 1년 전 거주지, 3개월 후 거주지가 빈집(빈방)이면 해당 거주지로 이동. 총조사에서 동일한 가구로 조사된 경우 한가구로 통합. 1인 가구 또는 남남가구인 경우 한 가구로 통합
부부 통합	실제 부부가 함께 살고 있을 확률이 높으나 주택청약, 진학 등의 이유로 주민등록상 주소가 다른 부부를 대상으로, 동일 특광역시, 동일 시군구, 인접 시군구(30km 미만) 내에 다른 주소로 등록된 부부 의 경우는 통합(부부 한쪽이 섬 지역인 경우 미적용)
부모와 자녀 통합	15세 이하의 자녀(기숙시설에 있지 않음)는 부모가구로 통합, 18세 미만 자녀(기숙시설에 있지 않음)의 부모가 동일 시군구 또는 동일 특광역시에 있는 경우 부모가구로 통합

자료: 통계청. (2020d). 인구총조사 통계정보보고서. pp.80-85.

다. 가계동향조사

통계청의 가계동향조사는 가구에 대한 가계수지 실태를 파악하여 국민의 소득과 소비 수준 변화의 측정 및 분석 등에 필요한 자료를 제공하기 위한 조사이다. 가계동향조사에서 가구란 1인 또는 2인 이상이 모여 취사, 취침 등 생계를 함께하는 생활 단위를 의미한다. 조사대상 가구는 전국의 농림 어가를 포함한 1인 이상의 일반가구 중 가계수지 파악이 가능한 가구이며, 조사대상 제외가구에는 집단가구, 음식숙박 겸용 주택가구, 장기출타가구, 비혈연가구, 외국인가구, 지출의존가구가 해당된다(통계청, 2022, p.6). 가구원은 주민등록 등재여부와 관계없이 생계를 같이하고 자산 또는

소득의 전부나 일부를 공동으로 사용하는 가구의 구성원을 말한다. 가구는 혈연관계인 가족으로 구성되기도 하고, 비혈연 관계의 가구원이 동거하는 경우도 있다. 또한, 가족이라 해도 취업, 유학 등의 사유로 함께 생활하지 않는 경우는 가구원에 포함하지 않는다(통계청, 2022, p. 34). 가계동향 조사는 취업, 학업 등의 이유로 떨어져 있는 가구원은 포함하지 않는 것이 가계금융복지조사와 주된 차이이다.

정리하면, 국내의 인구/소득통계 조사자료의 경우도 특성에 따라 상이한 가구 개념을 채택하고 있다. 통계청 <가계금융복지조사>에서는 같이 거주하지 않더라도 직장 때문에 따로 사는 배우자나 학업 때문에 따로 사는 자녀도 가구에 포함한다. 1인 또는 2인 이상이 모여 주거 또는 소득과 지출 등 생계를 같이 하는 생활 단위(가계금융복지조사의 개념)(통계청, 2021)와 같다. 통계청 <인구주택총조사>에서 가구는 ‘1인 또는 2인 이상이 모여서 취사, 취침 등 생계를 같이 하는 생활단위’로 정의되며, <가계동향조사>도 이 가구 개념을 따르고 있다. 1인 또는 2인 이상이 모여 취사, 취침 및 생계를 함께하는 생활단위로 일반가구, 집단가구, 외국인가구로 구분한다. 가계동향조사에서는 일반가구를 대상으로 조사하고, 집단가구나 외국인가구는 조사하지 않는다(통계청, 2022). 통계조사상의 가구 정의는 조사의 특수성을 반영한 것이며, 행정자료를 이용할 경우 가구 개념을 다양한 방식으로 정의하는 것이 가능할 것으로 생각된다.

제2절 가구 구성 방법

1. 다섯 가지 가구 구성안의 개요

가. 자료의 특징과 가구 ID 생성

제1절에서 소개한 가구 정의 방식은 모두 조사(survey)를 전제로 하고 있다. 인구총조사의 경우 한 국가의 모든 인구에 대한 직접 조사하는 과정이, 그리고 소득 조사의 경우 일부 표본에 대해 면접조사 하는 과정이 전제되어 있다. 실제로 조사를 수행하는 과정에서 각 개인들이 가구원에 해당하는지를 다양한 방식으로 확인하는 것이 가능하다. 그런데 본 연구에서는 이러한 과정이 없이, 주어진 행정자료를 이용하여 가구를 구성하여야 한다. 즉 개인 단위로 주어진 각 정보를 이용하여 어떤 개인들을 같은 가구로 묶을 수 있는지 검토해야 하는 것이다.

주어진 데이터를 이용하여 가구를 구성하는 작업은 달리 말해 각 개인들에게 가구식별번호(가구 ID)를 생성하여 부여하는 작업이다. 본 연구에서 사용하는 맞춤형 소득-재산 DB에는 비식별화된 개인 ID 이외에도 세대를 구분하는 비식별화된 번호(세대 ID), 국민건강보험증(이하 건강보험증)을 구분하는 비식별화된 번호(건강보험증 ID) 등 다수의 개인들이 공유할 수 있는 식별번호가 있다. 또한 주민등록상의 가구주 여부, 가구주와의 관계, 건강보험증상의 세대주, 세대주와의 관계 등 개인 간의 관계를 파악하는데 도움이 되는 항목들이 포함되어 있다. 거주지역(읍면동)코드도 개인들을 가구로 묶는 데 참고할 수 있다.

우선, 세대 ID가 같으면 주민등록상의 동일한 세대에 거주하는 것으로 볼 수 있다. 동일한 건강보험증 ID를 보유한, 다시 말해서 같은 건강보험증에 등재되어 있는 개인들은 생계를 같이하고 있을 가능성이 크다. 물론

생계를 따로 유지하면서도 건강보험증에만 같이 등재되어 있는 개인들도 있을 수 있다. 이런 점을 고려한다면, 어떤 두 사람이 동일한 세대 ID를 갖고 있다는 사실도, 또는 동일한 건강보험증 ID를 보유하고 있다는 것도 실제 생계를 같이하고 있다는 것을 곧바로 의미하지는 않는다.

결국 가구를 구성하고 가구 ID를 생성하는 과정은 이러한 여러 가지 정보를 조합하여 다양한 방식으로 개인들을 묶는 과정이다. 그리고 그 묶는 방법에 따라서 가구가 의미하는 바가 상이할 수 있다. 본 연구에서 시도하는 방법은 이론적으로 정의되는 가구 개념에 적합한 최선의 조합방식을 찾는 과정이라기보다 다양한 조합방식을 시도하고 그 결과를 상호비교하려는 것이다.

나. 다섯 가지 방안의 개요

맞춤형 소득-재산 DB에서 제공하는 정보를 이용하여 개인들을 묶는 몇 가지 방법을 생각해보자. 우선, 가장 단순한 방법은 별도의 과정 없이 현재 개인들을 묶고 있는 정보, 즉 세대 ID이나 건강보험증 ID를 이용하여 동일한 ID를 갖는 개인들을 하나의 가구로 간주하는 것이다. 이 가운데 특히 건강보험증 ID를 공유하는 개인을 하나의 가구로 간주하는 것은 여러 가지 한계가 있다. 실제 맞벌이 가구이거나 취업한(건강보험 직장가입자 자격을 지닌) 자녀가 있는 경우가 적지 않기 때문이다. 따라서 이 두 ID를 이용하여 가구를 구성하는 각각의 방법은 다른 방법을 구상하기 위한 일종의 중간 단계라고 이해해도 좋을 것이다.

이하에서는 동일한 세대 ID를 갖는(즉 같은 주민등록상 세대에 속하는) 개인들을 묶어 가구로 간주하는 방안을 ‘세대기준안’이라고 지칭할 것이다. 그리고 동일한 건강보험증에 기재되어 있는, 즉 동일한 건강보험증번호(ID)를 갖는 개인들을 묶어 가구로 간주하는 방안을 ‘건강보험증기준안’

이라고 지칭할 것이다.

다른 세 가지 구성안은 기본적으로 세대기준안의 변형이다. 어떤 내용으로 세대기준안을 변형할 것인지에 따라 차이가 있을 뿐이다. 첫 번째는 단순히 기존의 세대 구성원에다 건강보험증의 피부양자를 포함하여 하나의 가구로 구성하는 방법이다. 이를 ‘세대보완 - 건강보험 피부양자 포함안’이라고 지칭할 것이다. 그리고 간단히 지시하기 위하여 ‘1안’이라고 부를 것이다.

두 번째는 하나의 세대에 비동거 가족을 합하여 가구를 구성하는 안이다. 물론 모든 비동거 가족을 합하지는 않는데, 이에 대해서는 다음 절에서 상세하게 설명할 것이다. 이 방안을 ‘세대보완 - 비동거 가족 포함안’이라고 지칭하고, ‘2안’으로도 부를 것이다.

세 번째는 하나의 세대원 가운데 실질적으로 생계를 같이 하지 않는 세대원을 추정하여 이들을 제외한 후 가구를 정의하는 방법이다. 그 자세한 방법은 역시 다음 절에서 설명할 것이다. 이 방안을 ‘세대보완-분리세대 추정안(3안)’이라고 지칭할 것이다. 가구를 구성하는 다섯 가지 방안의 명칭과 특성을 간단히 정리하면 <표 3-6>와 같다.¹⁾

<표 3-6> 다섯 가지 가구 구성안의 명칭과 기본 개념

	명칭	기본 개념
1안	세대보완 1안(건강보험 피부양자 포함)	세대원 + 건강보험 피부양자
2안	세대보완 2안(비동거 가족 포함)	세대원 + 비동거(비동일 세대) 가족
3안	세대보완 3안(분리세대 추정)	세대원 - 생계분리 세대원
4안	세대기준안	세대원 = 가구원
5안	건강보험증기준안	동일 건강보험증 등재자 = 가구원

자료: 저자 작성

1) 각 안의 순서에는 특별한 의미가 없다.

가구를 구성하는 기본 개념은 <표 3-6>과 같지만 실제로 행정자료를 이용한 이 개념에 따라 가구원을 묶는 과정은 결코 단순하지 않다. 예를 들어 2안의 경우를 보자. 같은 세대원은 아니지만 실제로 생계를 같이 하고 있는 것으로 추정되는 가족(비동거 가족)을 함께 묶어 가구원으로 보는 것이다. 그렇다면 이 비동거 가족이 당초 속해 있던 세대의 구성원에 변동이 생긴다. 이와 같이 세대를 넘나드는 가구의 구성 방식은 연쇄적으로 발생하는 다양한 변화 등을 모두 고려해야 한다. 따라서 실제로 가구를 구성하고 가구 ID를 생성하는 방법은 매우 복잡할 수밖에 없다. 그 복잡함 내용은 다음 소절에서 설명하도록 한다.

2. 가구 구성안별 구축 과정(1~3안)

가. 1차 유형 구분: 1~3안 공통

□ 개인별 가구원 유형 구분: 주민등록 세대 기준

1~3안은 같은 곳에 거주하면 생계를 같이 한다고 보는 것이 기본 개념으로, 본 연구에서는 ‘주민등록상 동일 세대’일 때 같은 곳에 거주하는 것으로 본다. 이러한 관점을 바탕으로, 1차 유형 구분 단계는 ‘주민등록 세대’라는 동일 주소지를 갖는 사람들 중에서 가구구성원으로 보기 어려운 사람을 해당 가구에서 제외하는 과정으로 볼 수 있다. 세대주를 기준으로 해당 가구가 구성되는 것을 전제로, 세대주와 각 세대원의 정보를 비교하여 세대주의 가구원으로 귀속시킬지를 판단하는 것이다.

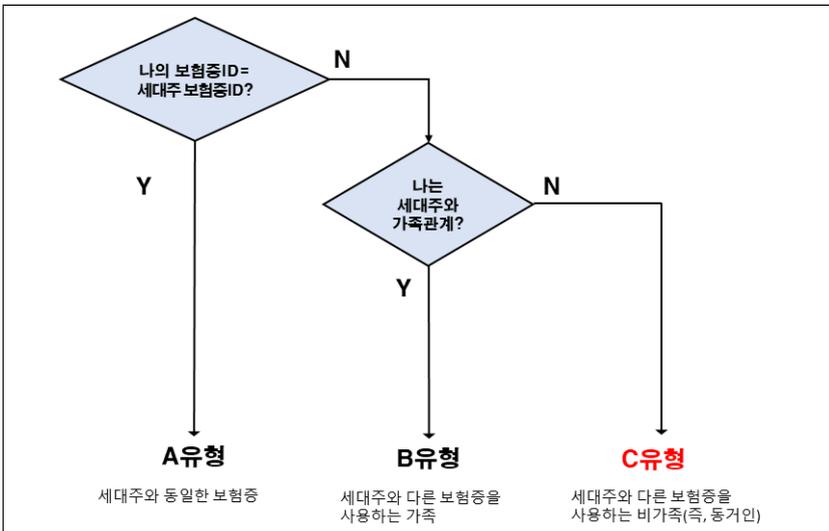
주민등록상 동일 세대인 상황에서, 각 세대원이 해당 주민등록 세대주의 가구원인지 판단하는 기준은 2가지이다. 첫 번째는 주민등록 세대주와 ‘동일한 건강보험증’인지 여부, 두 번째는 주민등록 세대주와 ‘가족관계’

인지 여부이다. 가족관계 여부는 주민등록상 세대주와의 관계 정보에서 ‘동거인’에 해당하지 않으면 가족관계로 판단하며, 촌수 등의 제한을 두지 않는다.

[그림 3-1]에 제시한 ‘A’ 유형은 주민등록 세대주와 동일한 건강보험증을 갖는 세대원으로, 가족관계 여부를 확인하지 않고 가구원으로 본다. ‘B’ 유형은 주민등록 세대주와 건강보험증이 다르지만, 세대주의 가족관계인 세대원으로, 세대주의 가구원으로 본다. ‘C’ 유형은 세대주와 건강보험증이 다르고, 가족관계가 아니므로 해당 세대주의 가구원에 속하지 않는다고 본다.

이러한 1차 유형 구분은 각 개인이 2차 유형 판단을 거쳐 최종적으로 어떤 가구에 속할지를 판단하는 기본적 정보가 된다. 2차 유형 구분 과정이 이루어지기 전까지의 잠정적인 상태로, 2차 유형 분류 과정에서 다른 가구의 가구원으로 정의될 수 있다.

[그림 3-1] 1차 유형 분류 흐름도



나. 2차 유형 구분: 1차 유형 구분에 기반한 개인별 최종 가구 정의

1차 유형은 각 세대원 본인과 같은 주소에 거주하고 있는 주민등록 세대주와 비교한 결과라면, 다음 단계로는 주민등록 세대를 달리하는 가구원을 하나의 가구로 묶는 상황이 고려된다.

이때 중요한 판단 기준은 각 세대원 본인의 건강보험증 대표의 정보이다. 건강보험증 대표의 가구원으로 귀속될 조건에 해당하는지를 확인하여, 해당이 된다면 '건강보험증 대표'의 주민등록 세대 정보 및 1차 유형에 따라 최종적으로 속하는 가구가 정의된다. 건강보험증 대표의 가구원으로 귀속될 조건에 해당하지 않는다면 1차 유형 구분에서 잠정적으로 정의한 대로 가구가 정의된다.

다만, 예외적으로 본인의 건강보험증 대표가 확인되지 않는 경우는 다음과 같이 처리한다. 본인의 1차 유형이 A 또는 B이면, 본인의 주민등록 세대주의 가구로 귀속되도록 세대주 ID를 부여한다. 본인의 1차 유형이 C이면, 주민등록 세대주 ID와 건강보험증 ID를 결합한 새로운 가구 ID를 부여하여, 다른 모든 유형의 가구와 분리되면서, 동일 세대이면서 동일 보험증의 개인들은 한 가구로 묶일 수 있도록 한다.

1~3안의 가구 구성안별 가구 정의 과정은 다음과 같다.

1) 세대보완 1안(건강보험 피부양자 포함)

□ 1차 유형이 A, B 유형일 때

1안에서는 각 개인이 건강보험증 대표와 다른 주민등록 세대를 구성하는 경우, 해당 건강보험증 대표의 가구로 귀속시킨다. 1차 유형이 A 또는 B인 경우는 2개의 판단기준에 따라 3개의 유형으로 구분된다.

○ 유형 1: 주민등록 세대주 중심의 가구 ID 생성

먼저, 본인의 주민등록 세대주와 건강보험증 대표의 세대주가 동일한지 여부를 확인한다. 동일하다면 본인의 건강보험증 대표와 같이 거주하므로 본인의 주민등록 세대주 중심의 가구 ID가 생성된다(본인의 세대주 ID 부여).

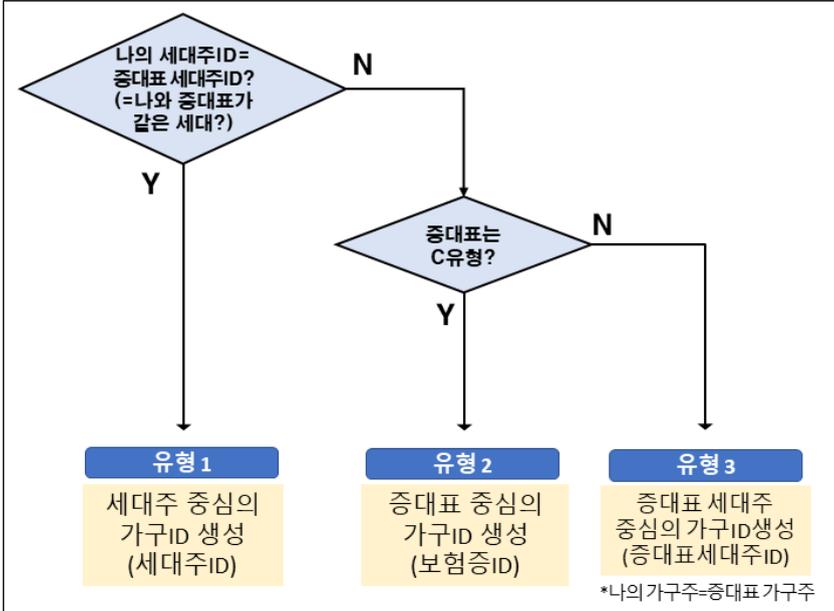
○ 유형 2: 건강보험증 대표 중심의 가구 ID 생성

첫 번째 단계에서 본인의 주민등록 세대주와 건강보험증 대표가 동일 주소에 거주하지 않는다는 것이 확인되면, 건강보험증 대표의 가구로 귀속된다. 이때 두 번째 단계에서 건강보험증 대표가 증대표의 주민등록 세대의 가구원인지 여부에 따라 가구 ID가 부여된다. 즉 건강보험증 대표의 1차 유형이 C라면 증대표의 주민등록 세대에서 가구원으로 속하지 않은 것으로 판단되므로, 증대표 중심의 가구 ID를 생성한다(건강보험증 ID 부여).

○ 유형 3: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

유형 3은 첫 번째 단계에서 본인의 주민등록 세대주와 건강보험증 대표의 세대주가 다른 것이 확인되고, 두 번째 단계에서 건강보험증 대표가 증대표의 주민등록 세대에서 A 또는 B로 분류된 경우이다. 본인의 건강보험증 대표는, 증대표의 주민등록 세대에서 가구원으로 귀속되므로, 증대표를 따라 증대표의 세대주 중심의 가구 ID를 생성한다(증대표 세대주 ID 부여).

[그림 3-2] 가구 ID 1안 생성 흐름: A유형, B유형



주: 1차 유형이 A 또는 B이면서 건강보험증 대표를 확인할 수 없는 예외적인 경우 세대주 ID를 부여함

<표 3-7> 가구 ID 1안: A, B유형일 때 가구 ID 부여 예시

구분	유형	예시	나의 가구 ID
A 또는 B	유형1	· 나는 세대주이며 증대표 · 나는 세대주이며 증대표는 같은 세대 내에 있는 나의 아들 · 나는 세대주의 아들이며, 내가 증대표. 세대주인 아버지는 나의 피부양자.	세대주 ID
	유형2	· 나는 딸(증대표)과 따로 사는 노인(할머니). 나의 세대주인 남편(역시 노인)도 딸의 보험증에 올라가 있음. 나의 딸은 직업적인 이유로 타 지역에서 하숙 중임.	보험증 ID
	유형3	· 나는 딸(증대표)과 따로 사는 노인(할머니). 세대주인 남편(역시 노인)도 딸의 보험증에 올라가 있음. 나의 딸은 사위가 세대주로 신고되어 있는 곳에 살고 있으며, 직장생활을 하고 있음.	보험증 대표의 세대주 ID

주: 1) 세대주: 주민등록상 세대주
2) 보험증: 건강보험증

□ 1차 유형이 C유형일 때

1차 유형이 C인 경우, 본인의 주민등록 세대에서는 가구원으로 속하지 않는 사람으로, 이 단계에서는 건강보험증 대표가 속하는 가구로 가구 ID를 부여하게 된다.

○ 유형 4: 새로운 가구 ID 생성(건강보험증 ID)

첫 번째 단계에서 본인의 주민등록 세대주와 건강보험증 대표의 세대가 동일하여 동일 세대에 거주하는 경우이다. 두 번째 단계에서 본인의 건강보험증 대표의 1차 유형이 C에 해당하면, 증대표는 증대표의 주민등록 세대주의 가구에 속하지 않으므로, 건강보험증 단위로 구분되는 새로운 가구 ID를 생성한다(건강보험증 ID 부여).

○ 유형 5: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

첫 번째 단계는 유형 4와 동일하나, 본인의 건강보험증 대표가 A 또는 B유형일 경우에 해당한다. 첫 번째 단계에서 본인은 건강보험증 대표와 같은 세대에 거주하는 것으로 확인되고, 건강보험증 대표는 증대표의 주민등록 세대에서 가구원으로 귀속되는 A, B의 유형이므로 증대표의 주민등록 세대주 중심의 가구 ID를 생성한다(증대표 세대주 ID 부여).

○ 유형 6: 새로운 가구 ID 생성(건강보험증 ID)

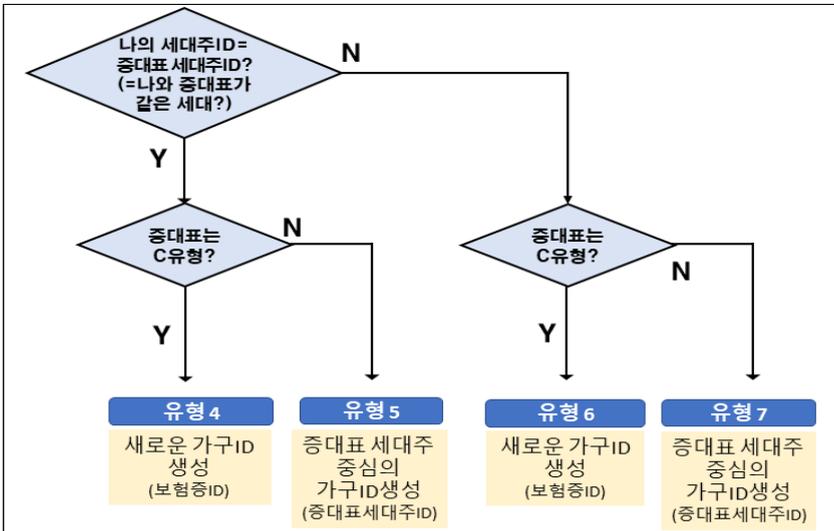
본인과 건강보험증 대표가 다른 세대에 거주하더라도(첫 번째 단계), 건강보험증이 동일한 것으로 확인되면 건강보험증 대표의 가구로 귀속된다.

두 번째 단계에서 건강보험증 대표가 C유형에 해당하면, 증대표는 증대표의 주민등록 세대에서 가구원으로 인정되지 않으므로, 본인 및 증대표가 한 가구로 묶일 수 있도록 건강보험증 단위의 새로운 가구 ID를 생성한다(건강보험증 ID 부여).

○ 유형 7: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

유형 6과 마찬가지로, 본인과 건강보험증 대표가 다른 세대에 거주하더라도, 건강보험증이 동일한 것으로 확인되면 건강보험증 대표의 가구로 귀속된다. 본인의 건강보험증 대표가 증대표의 주민등록 세대에서 A 또는 B로 분류되었다면, 건강보험증 대표의 세대주 중심으로 가구 ID를 생성하여, 증대표의 주민등록 세대에 귀속되도록 한다(증대표 세대주 ID 부여).

[그림 3-3] 가구 ID 1안 생성 흐름: C유형



주: 1차 유형이 C이면서 건강보험증 대표를 확인할 수 없는 예외적인 경우 '세대주 ID'와 '보험증 ID'를 결합한 새로운 가구 ID를 생성함.

〈표 3-8〉 가구 ID 1안: C유형일 때 가구 ID 부여 예시

구분	예시	나의 가구 ID
C	유형4 · 나와 형(나의 증대표)은 같은 곳에서 하숙 중임.	새로운 가구 ID (보험증 ID)
	유형5 · 나는 증대표의 미성년 자녀이고, 나의 증대표는 세대주와 재혼하여, 세대주와 모두 같은 집에 살고 있음. 주민등록 신고 시에는 세대주의 동거인으로 신고함.	보험증 대표의 세대주 ID
	유형6 · 나는 아들(증대표)과 따로 사는 노인(할머니). 나는 집주인과 같은 건물의 다가구 주택에 임대료 함께 살고 있으며, 주민등록상 집주인과 같은 세대로 묶여 있음. 나의 아들은 직업적인 이유로 타 지역에서 하숙 중임.	새로운 가구 ID (보험증 ID)
	유형7 · 나는 딸(증대표)과 따로 사는 노인(할머니). 나는 집주인과 같은 건물의 다가구 주택에 임대료 함께 살고 있으며, 주민등록상 집주인과 같은 세대로 묶여 있음. 나의 딸은 사위가 세대주인 곳에 살고 있으며, 직장생활을 하고 있음.	보험증 대표의 세대주 ID

주: 1) 세대주: 주민등록상 세대주
2) 보험증: 건강보험증

2) 세대보완 2안(비동거 가족 포함)

2안은 본인과 건강보험증 대표의 관계를 확인하여, 본인이 건강보험증 대표의 ‘배우자’이거나 ‘직계비속’인 경우, 건강보험증 대표의 가구로 귀속시킨다.

□ 1차 유형이 A, B 유형일 때

1차 유형이 A 또는 B이면, 본인의 주민등록 세대주와 동일 주소이면서, 건강보험증이 같거나 건강보험증이 분리되어 있더라도 가족관계이므로, 원칙적으로 본인의 주민등록상 세대주의 가구원으로 인정되는 사람으로 본다. 다만, 본인의 건강보험증 대표가 다른 세대를 구성하고 있고, 본인이

건강보험증 대표의 배우자 또는 비속이라면, 건강보험증 대표의 가구원으로 귀속되도록 한다.

3개의 판단기준에 따라 4개의 유형으로 구분될 수 있으며, 각 유형은 다음의 과정을 거쳐, 본인이 속할 가구의 ID가 부여된다.

○ 유형 1: 주민등록 세대주 중심의 가구 ID 생성

첫 번째 단계로, 본인의 주민등록 세대주와 건강보험증 대표의 세대주가 동일한지 여부를 확인한다. 전자와 후자가 동일하다면 본인의 건강보험증 대표도 동일한 주소에 거주하는 것이므로, 주민등록세대주의 가구원으로 귀속된다. 따라서 본인의 주민등록 세대주 중심의 가구 ID가 생성된다(본인의 세대주 ID 부여).

○ 유형 2: 건강보험증 대표 중심의 가구 ID 생성

첫 번째 단계에서 본인의 주민등록 세대주와 건강보험증 대표가 동일 주소에 거주하지 않는다는 것이 확인되면, 두 번째 단계에서 본인의 건강보험증 대표의 가구원으로 귀속되어야 하는지를 판단한다. 이는 본인의 건강보험증 대표와의 관계에서, 본인이 배우자 또는 비속에 해당하는지에 따라 판단된다. 배우자 또는 비속에 해당될 경우, 본인의 건강보험증 대표의 가구원으로 판단한다. 따라서 본인 건강보험증 대표의 1차 유형에 따라 증대표의 가구 ID와 동일한 가구에 속하게 한다.

세 번째 단계에서는 최종 가구 ID 부여를 위하여 건강보험증 대표의 1차 유형을 고려한다. 건강보험증 대표가 증대표의 주민등록 세대에서 C유형으로 분류되었다면, 증대표의 세대에서 가구원으로 속하지 않으므로, 증대표를 중심으로 하는 별도의 독립적인 가구 ID를 생성한다(본인의

건강보험증 ID 부여).

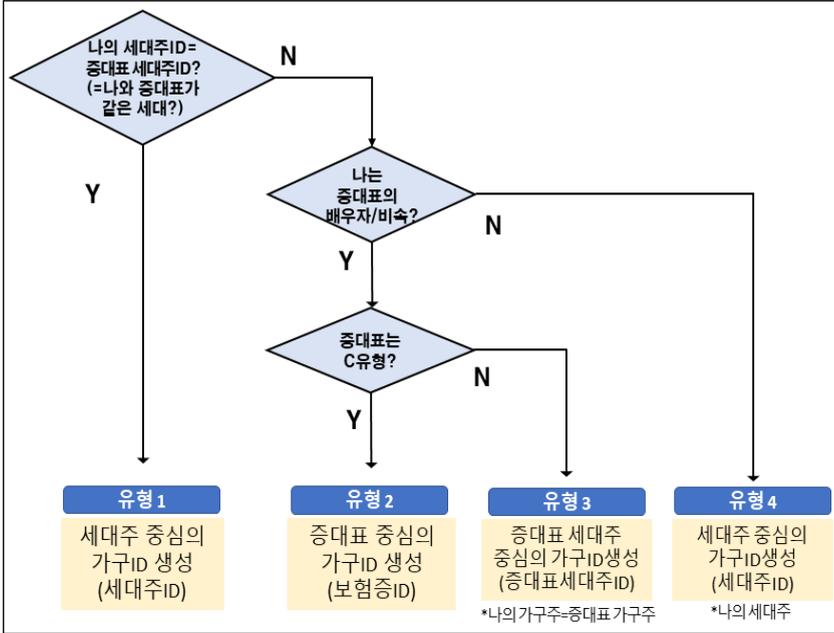
○ 유형 3: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

세 번째 단계에서 본인의 건강보험증 대표가 A 또는 B로 1차 유형이 부여되었다면, 증대표는 해당 주민등록 세대에서 가구원으로 인정되는 경우이다. 따라서 본인이 건강보험증 대표의 배우자 또는 비속일 때, 증대표의 1차 유형이 A 또는 B이면, 증대표의 주민등록 세대주의 가구로 귀속된다(증대표의 세대주 ID 부여).

○ 유형 4: 주민등록 세대주 중심의 가구 ID 생성

두 번째 단계에서 배우자 또는 비속이 아닌 것으로 확인되면, 건강보험증 대표의 가구와 관계없이, 본인의 주민등록 세대주의 가구로 가구 ID가 부여된다. 왜냐하면 본인은 1차 유형 분류 시, 이미 본인의 건강보험 세대주와 동일 세대이면서, 동일 건강보험증을 갖는 것으로 확인되어 A유형으로 판단되었기 때문이다.

[그림 3-4] 가구 ID 2안 생성 흐름: A유형, B유형



주: 1차 유형이 A, B이면서, 건강보험증 대표를 확인할 수 없는 예외적인 경우에 세대주 ID를 부여함.

<표 3-9> 가구 ID 2안: A, B유형일 때 가구 ID 부여 예시

구분	예시	나의 가구 ID
A 또는	유형1 · 나는 세대주이며 증대표 · 나는 세대주이며 증대표는 같은 세대 내에 있는 나의 아들 · 나는 세대주의 아들이며, 내가 증대표. 세대주인 아버지는 나의 피부양자.	세대주 ID
	유형2 · 나는 증대표(남편)와 따로 사는 아내이고, 세대주임. 증대표인 남편은 다른 세대의 동거인.	보험증 ID
B	유형3 · 나는 증대표(아버지)과 따로 사는 아들임. 나의 세대주는 증대표(아버지)와 따로 사는(나와는 함께 거주하는) 엄마임.	보험증 대표의 세대주 ID
	유형4 · 나는 아들(증대표)과 따로 사는 노인(할머니). 세대주인 남편(역시 노인)도 아들의 보험증에 올라가 있음.	세대주 ID

주: 1) 세대주: 주민등록상 세대주
2) 보험증: 건강보험증

□ 1차 유형이 C유형일 때

1차 유형이 C인 경우 본인의 주민등록 세대에서는 가구원으로 속하지 않는 사람이고, 이 단계에서는 건강보험증 대표의 가구에 속할 조건에 해당하는지를 중심으로 판단한다. 3개의 판단기준에 따라 유형 5~11까지로 구분할 수 있으며, 각 유형은 다음의 과정을 거쳐, 건강보험증 대표의 가구에 속하거나, 새로운 가구로 ID가 부여된다.

○ 유형 5: 새로운 가구 ID 생성(건강보험증 ID)

첫 번째 단계에서 본인의 주민등록 세대주와 건강보험증 대표의 세대주가 동일한 것으로 확인된 경우이다. 두 번째 단계에서 본인은 건강보험증 대표의 배우자 또는 비속으로 확인되어, 건강보험증 대표의 가구에 속하게 된다. 세 번째 단계에서 본인의 건강보험증 대표가 C에 해당한다면, 증대표는 증대표의 주민등록 세대주의 가구에 속하지 않으므로, 새로운 가구 ID가 생성된다(건강보험증 ID 부여).

○ 유형 6: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

첫 번째 및 두 번째 단계의 결과는 유형 5와 동일하나, 본인의 건강보험증 대표가 A, B유형인 경우이다. 본인은 건강보험증 대표의 가구에 속할 조건에 해당하며, 증대표는 증대표의 주민등록 세대에서 가구원에 속한다. 따라서 본인의 건강보험증 대표의 세대로 가구 ID가 생성된다(건강보험증 대표의 세대주 ID 부여).

C유형은 1차 유형 분류 시, 본인의 주민등록 세대주와의 조건으로는 가구원으로 속하지 못하였으나, 건강보험증 대표의 가구에 속하므로 증대

표의 세대주 ID를 부여한다. 사실상 건강보험증 대표의 세대주와 본인의 주민등록 세대주가 같으므로 본인의 세대주 ID와 동일한 결과이다.

○ 유형 7: 새로운 가구 ID 생성(건강보험증 ID)

첫 번째 단계에서 본인의 주민등록 세대주와 건강보험증 대표의 세대주가 동일한 것으로 확인되며, 두 번째 단계에서 본인은 건강보험증 대표의 배우자 또는 비속에 해당하지 않는다. 그러나 이 경우는 건강보험증 대표와 동일 세대이므로, 동일 세대-동일 보험증이면 동일 가구로 보는 원칙에 따라 건강보험증 대표의 가구로 연결한다. 건강보험증 대표가 C이면 증대표의 세대에서 가구원이 아니므로, 새로운 가구 ID를 생성하되, 본인과 증대표가 한 가구로 묶일 수 있도록 보험증 단위의 ID를 부여한다(건강보험증 ID 부여).

○ 유형 8: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

첫 번째, 두 번째 단계는 유형 7과 같으며, 세 번째 단계에서 본인의 건강보험증 대표가 A 또는 B유형에 해당하는 경우이다. 유형 7과 마찬가지로 건강보험증 대표와 동일 세대이므로, 동일 세대-동일 보험증이면 동일 가구로 보는 원칙에 따라 건강보험증대표의 가구로 연결한다. 다만 건강보험증대표가 증대표의 주민등록 세대에서 가구원에 속하므로 증대표의 가구 ID를 부여한다(건강보험증 대표의 세대주 ID 부여).

○ 유형 9: 새로운 가구 ID 생성(건강보험증 ID)

첫 번째 단계에서 본인과 건강보험증 대표가 다른 주민등록 세대임이

확인된 경우이다, 두 번째 단계에서 본인이 건강보험증 대표의 배우자 또는 비속에 해당하여 건강보험증 대표의 가구에 속하게 된다. 세 번째 단계에서 건강보험증 대표가 C이면, 증대표는 증대표의 세대에서 가구원이 아니므로, 새로운 가구 ID를 생성하되, 본인과 증대표가 한 가구로 묶일 수 있도록 보험증 단위의 ID를 부여한다(건강보험증 ID 부여).

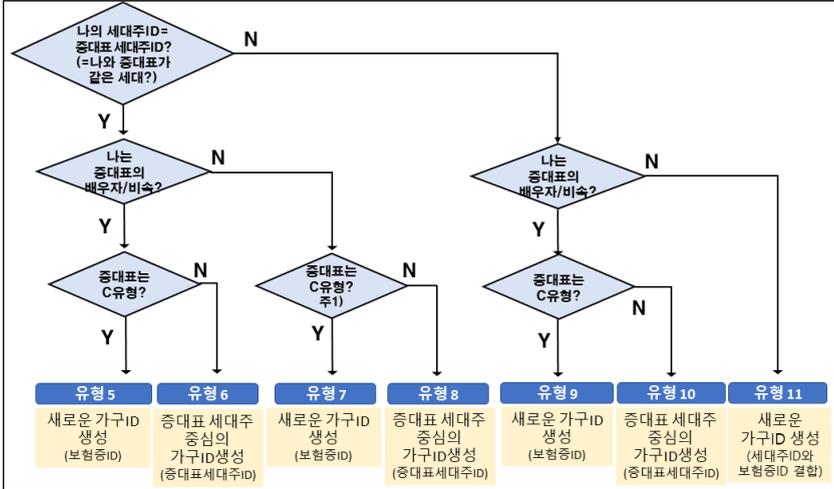
○ 유형 10: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

첫 번째, 두 번째 단계는 유형 9와 같으며, 건강보험증 대표의 1차 유형이 A 또는 B인 경우이다. 본인은 건강보험증 대표의 가구에 속할 조건에 해당하며, 증대표는 증대표의 주민등록 세대에서 가구원에 속한다. 따라서 본인의 건강보험증 대표의 세대로 가구 ID가 생성된다(건강보험증 대표의 세대주 ID 부여).

○ 유형 11: 새로운 가구 ID 생성(주민등록 세대주 ID와 건강보험증 ID를 결합한 새로운 ID 부여)

첫 번째 단계에서 본인과 건강보험증 대표가 다른 주민등록 세대임이 확인되며, 두 번째 단계에서 본인이 건강보험증 대표의 배우자 또는 비속에 해당하지 않는다. 따라서 본인의 주민등록 세대주와 건강보험증 대표와 다른 새로운 가구를 형성하게 된다. 이때 동일 세대-동일 보험증인 사람들이 있다면 동일 가구로 묶일 수 있도록 가구 ID를 부여한다(주민등록 세대주 ID와 건강보험증 ID를 결합한 새로운 ID 부여).

[그림 3-5] 가구 ID 2안 생성 흐름: C유형



주: 1) 건강보험증 대표의 배우자·비속에 속하지 않으나, 건강보험증 대표와 동일 세대이므로 동일 세대-동일 보험증이면 동일 가구로 보는 원칙에 따라 건강보험증 대표의 가구로 연결
 2) 1차 유형이 C이면서, 건강보험증 대표를 확인할 수 없는 예외적인 경우 '세대주 ID'와 '보험 증 ID'를 결합한 새로운 가구 ID를 생성함

<표 3-10> 가구 ID 2안: C유형일 때 가구 ID 부여 예시

구분	예시	나의 가구 ID
C	유형5 · 나와 남편(나의 증대표)은 같은 곳에서 하숙 중임. 이 경우 증대표를 중심으로 가구를 구성.	새로운 가구 ID (보험증 ID)
	유형6 · 나는 증대표의 미성년 자녀이고, 나의 증대표는 세대주와 재혼하여, 세대주와 모두 같은 집에 살고 있음. 주민등록 신고 시에는 세대주의 동거인으로 신고함.	보험증 대표의 세대주 ID
	유형7 · 나의 증대표는 아들임. 집주인과 같은 건물의 다가구 주택에 임대료 함께 살고 있으며, 주민등록상 집주인과 같은 세대로 묶여 있음.	새로운 가구 ID (보험증 ID)
	유형8 · 나는 누나의 부부와 함께 살고 있음. 나는 대학생으로 직장생활 중인 누나의 피부양자이며, 세대주는 매형임.	보험증 대표의 세대주 ID
	유형9 · 각각 하숙하는 부부. 증대표는 남편. 증대표를 중심으로 별도의 가구 구성	새로운 가구 ID (보험증 ID)

구분	예시	나의 가구 ID
유형10	· 나는 내가 거주하는 세대에서 비가족 동거인(하숙). 나의 증대표는 따로 사는 남편. 남편은 세대주임. 이 경우 나는 증대표를 따라 가구를 구성. 증대표의 가구주(세대주)가 나의 가구주가 됨.	보험증 대표의 세대주 ID
유형11	· 나는 하숙. 나의 증대표는 따로 사는 아들. 나는 아들과 독립적으로 생활하지만, 아들의 보험증에 올라가 있음. 이 경우 나의 세대주와도 나의 증대표와도 가구를 구성하지 못함. 즉 세대주 아이디나 증대표 아이디로만으로는 가구 구성 불가. · 나의 배우자도 나와 같은 곳에 하숙하며 아들의 보험증에 올라가 있다면 나와 배우자는 같은 가구에 속함.	새로운 가구 ID (보험증 ID+세대주 ID)

주: 1) 세대주: 주민등록상 세대주
 2) 보험증: 건강보험증
 3) 증대표: 건강보험증 대표

3) 세대보완 3안(분리세대 추정)

3안은 2안의 가구 ID 생성 방식에 조건을 한 가지 추가한 방식이다. 추가되는 조건은 건강보험증 대표의 배우자 또는 비속일 때, 건강보험증 대표와 동일 주소인지 여부이다. 본인의 건강보험증 대표와 주민등록세대가 다른 경우, 즉 나의 세대주와 건강보험증 대표의 세대주가 다를 때, 이 추가 조건을 적용한다. 개념적으로 3안은 생계를 같이 하나 주민등록상 세대 분리가 되어 있는 경우를 식별하여 동일한 가구로 묶는다는 것이지만, 주소자료의 특성상 세부 주소지 단위까지 동일 여부를 확인할 수 없는 한계가 있다. 따라서 활용 가능한 세부 주소 정보인 읍, 면, 동 단위 주소를 이용하였다.

□ 1차 유형이 A, B 유형일 때

○ 유형 1: 주민등록 세대주 중심의 가구 ID 생성

이 유형은 2안의 유형 1(1차 유형 A, B)과 동일하다. 건강보험증 대표와 동일한 주민등록 세대에 거주하고 있으므로, 본인의 세대주 중심의 가구 ID가 생성된다(세대주 ID 부여).

○ 유형 2: 건강보험증 대표 중심의 가구 ID 생성

이 유형은 건강보험증 대표가 다른 세대를 구성하고, 본인이 건강보험증 대표의 배우자 또는 비속이면서, 동일한 주소(읍면동)이므로 건강보험증 대표의 가구로 귀속되는 조건을 갖추는 경우이다. 이때 마지막 단계에서 건강보험증 대표의 1차 유형이 C유형에 해당하여, 증대표를 중심으로 하는 별도의 독립적인 가구 ID를 생성한다(본인의 건강보험증 ID 부여).

○ 유형 3: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

세 번째 단계까지 유형 2와 동일하므로 건강보험증 대표의 가구로 귀속 될 조건을 갖추었는지 확인되는 경우이다. 마지막 단계에서 건강보험증 대표의 1차 유형이 A 또는 B에 해당하므로, 건강보험증 대표의 세대주 중심의 가구 ID를 생성한다(증대표의 세대주 ID 부여).

○ 유형 4: 주민등록 세대주 중심의 가구 ID 생성

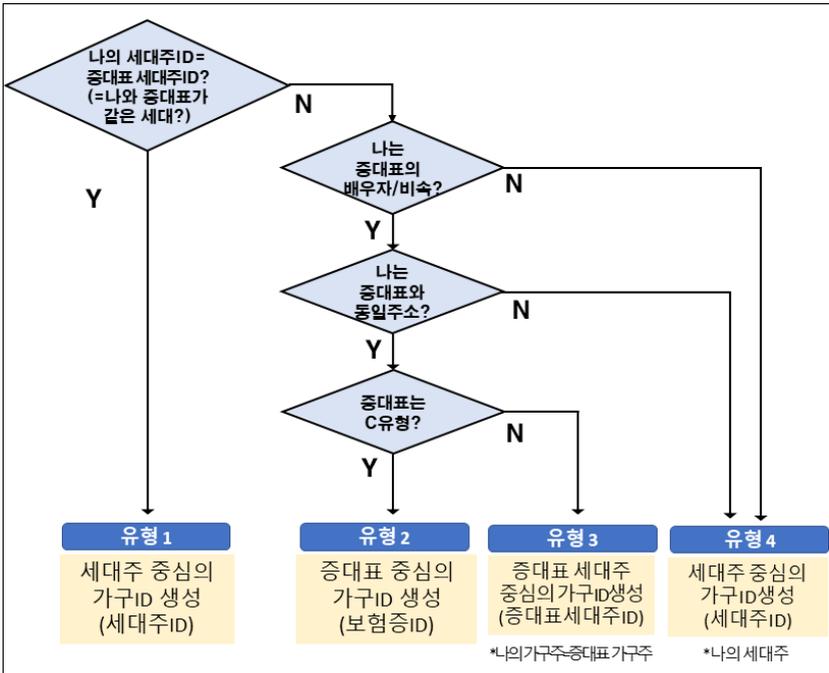
이 유형은 두 가지의 경우가 존재할 수 있다. 첫 번째는 본인의 건강보험

증 대표와 다른 주민등록 세대를 구성하고, 본인은 건강보험증 대표의 배우자 또는 비속에 해당하나, 동일 주소(읍면동)가 아니므로 건강보험증 대표의 가구로 귀속되지 않는 경우이다.

두 번째는 본인의 건강보험증 대표와 다른 주민등록 세대를 구성하고, 본인은 건강보험증 대표의 배우자 또는 비속에 해당하지 않는 경우이다.

따라서 본인의 1차 유형이 A, B이므로 본인의 주민등록 세대주에 귀속되는 가구 ID가 생성된다(세대주 ID 부여).

[그림 3-6] 가구 ID 3안 생성 흐름: A유형, B유형



주: 1차 유형이 A, B이면서, 건강보험증 대표를 확인할 수 없는 예외적인 경우 세대주 ID를 부여함.

〈표 3-11〉 가구 ID 3안: A, B유형일 때 가구 ID 부여 예시

구분	예시	나의 가구 ID
A 또는 B	유형1 · 나는 세대주이며 증대표 · 나는 세대주이며 증대표는 같은 세대 내에 있는 나의 아들임.	세대주 ID
	유형2 · 나는 세대주인 아내이며 증대표(남편)와 실제로는 같이 살지만, 주민등록상 다른 세대임. 증대표인 남편은 직업상의 이유로 옆집 세대에 방을 구하여 필요시마다 사용함. 이 방은 다가구주택으로 남편은 주민등록상 주인 세대의 동거인으로 되어 있음.	보험증 ID
	유형3 · 나는 증대표(아버지)와 같은 지역에 따로 사는 아들임. 증대표(아버지)는 할아버지(아버지의 세대주)와 살고 있음. 나의 세대주는 증대표(아버지)와 따로 사는(나와는 함께 거주하는) 엄마임.	보험증 대표의 세대주 ID
	유형4 · 나는 아들(증대표)과 따로 사는 노인(할머니). 세대주인 남편(역시 노인)도 아들의 보험증에 올라가 있음. · 나(아내)는 서울에 살고, 남편(증대표)은 직업상의 이유로 부산에 살고 있는 주말부부임. 각각은 각 주소에서 주민등록 세대주로 등록되어 있음.	세대주 ID

주: 1) 세대주: 주민등록상 세대주
 2) 보험증: 건강보험증

□ 1차 유형이 C유형일 때

1차 유형이 C인 경우 본인의 주민등록 세대에서는 가구원으로 속하지 않는 사람이고, 이 단계에서 본인이 건강보험증 대표와 다른 주민등록 세대일 때, 건강보험증 대표의 가구에 속할 조건에 해당하는지를 주소 조건을 추가하여 판단한다. 4개의 판단기준에 따라 유형 5~11로 구분될 수 있으며, 각 유형은 다음의 과정을 거쳐, 건강보험증 대표의 가구에 속하거나, 새로운 가구로 ID가 부여된다.

○ 유형 5: 새로운 가구 ID 생성(건강보험증 ID)

2안의 유형 5(1차 유형 C)와 동일한 판단과정을 거쳐 가구 ID가 생성된다. 본인이 건강보험증 대표와 동일한 주민등록 세대이며, 건강보험증 대표의 배우자 또는 비속의 관계이다. 따라서 동일 주소에 거주하는 것이므로 동일 주소 여부 확인은 생략하고, 건강보험증 대표의 1차 유형에 따라 가구 ID를 생성한다. 건강보험증 대표는 이 주민등록 세대에서 가구원으로 인정되지 않으므로, 새로운 가구 ID를 생성하되, 같은 보험증인 배우자와 비속이 가구로 묶일 수 있도록 한다(건강보험증 ID 부여).

○ 유형 6: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

위의 유형 5와 다른 점은 마지막 단계인 건강보험증 대표의 1차 유형이 A또는 B라는 점이며, 2안의 유형 6(1차 유형 C)과 동일하다. 따라서 건강보험증 대표의 주민등록 세대에 가구원으로 귀속된다(증대표 세대주 ID 부여).

○ 유형 7: 새로운 가구 ID 생성(건강보험증 ID)

2안의 유형 7(1차 유형 C)과 동일하게 가구 ID가 생성된다. 본인은 건강보험증 대표와 동일한 주민등록 세대이므로, 세대와 보험증이 같은 상황이다. 건강보험증 대표는 1차 유형이 C이므로 새로운 가구 ID를 생성하되, 본인과 증대표가 한 가구로 묶일 수 있도록 보험증 단위의 ID를 부여한다(건강보험증 ID 부여).

○ 유형 8: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

첫 번째, 두 번째 단계는 유형 7과 같으며, 세 번째 단계에서 본인의 건강보험증 대표가 A 또는 B유형에 해당하는 경우이다. 2안의 유형 8과 동일하다. 건강보험증 대표와 동일 세대이므로, 동일 세대-동일 보험증이면 동일 가구로 보는 원칙에 따라 건강보험증 대표의 가구로 연결한다. 다만 건강보험증 대표가 증대표의 주민등록 세대에서 가구원에 속하므로 증대표의 가구 ID를 부여한다(건강보험증 대표의 세대주 ID 부여).

○ 유형 9: 새로운 가구 ID 생성(건강보험증 ID)

본인과 건강보험증 대표가 다른 주민등록 세대이며, 본인은 건강보험증 대표의 배우자 또는 비속에 해당한다. 또한 본인은 건강보험증 대표와 동일한 주소(읍면동)이므로, 건강보험증 대표의 가구원에 해당하는 조건에 부합한다. 건강보험증 대표는 증대표의 주민등록 세대에서 가구원이 아닌 C로 분류되므로, 새로운 가구 ID를 생성하되, 증대표와 동일한 가구로 묶일 수 있도록 건강보험증 ID로 묶이는 가구 ID를 생성한다(건강보험증 ID 부여).

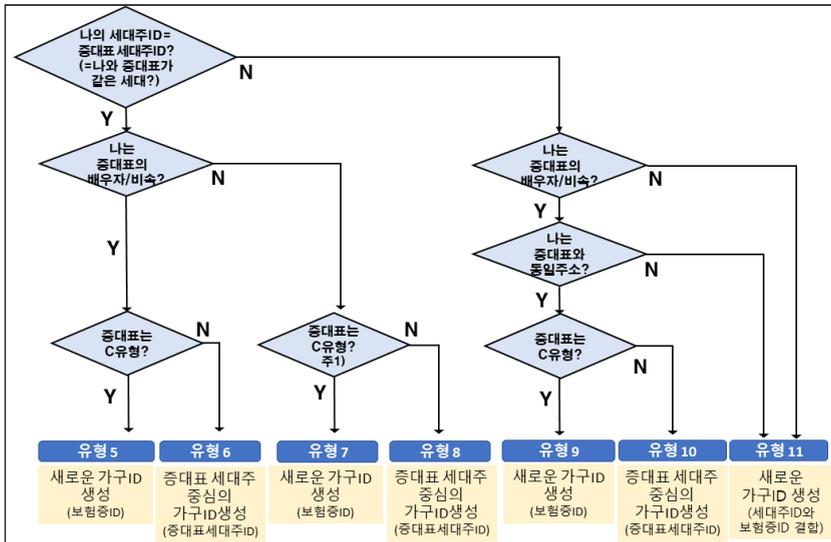
○ 유형 10: 건강보험증 대표의 주민등록 세대주 중심의 가구 ID 생성

마지막 단계를 제외하고는 유형 9와 같은 조건에 해당하여 건강보험증 대표의 가구원으로 판단되나, 건강보험증 대표의 1차 유형이 A, B라는 점에서 다르다. 건강보험증 대표는 증대표의 주민등록 세대에 가구원이므로 증대표 세대주 중심의 가구 ID를 생성한다(증대표 세대주 ID 부여).

○ 유형 11: 새로운 가구 ID 생성(주민등록 세대주 ID와 건강보험증 ID를 결합한 새로운 ID 부여)

이 유형은 본인과 건강보험증 대표가 다른 주민등록 세대이고, 건강보험증 대표의 가구에 속할 조건에 부합하지 않는 경우이다. 본인이 건강보험증 대표의 배우자 또는 비속이 아니거나, 건강보험증 대표와 동일 주소가 아닌 경우로, 건강보험증 대표의 가구와도 분리되고, 본인의 1차 유형은 C이므로 주민등록 세대주의 가구와도 분리된다. 다만 같은 경우에 해당하는 동일 세대-동일 보험증인 사람들이 있다면 동일 가구로 묶일 수 있도록 새로운 가구 ID를 생성한다(주민등록 세대주 ID와 건강보험증 ID를 결합한 새로운 ID 부여).

[그림 3-7] 가구 ID 3안 생성 흐름: C유형



주: 1) 증대표의 배우자·비속에 속하지 않으나, 증대표와 동일 세대이므로 동일 세대-동일 보험증 이면 동일 가구로 보는 원칙에 따라 증대표의 가구로 연결.
 2) 1차 유형이 C이면서, 건강보험증 대표를 확인할 수 없는 예외적인 경우 '세대주 ID'와 '보험 증 ID'를 결합한 새로운 가구 ID를 생성함.

〈표 3-12〉 가구 ID 3안: C유형일 때 가구 ID 부여 예시

구분	예시	나의 가구 ID
유형5	· 나와 남편(나의 증대표)은 같은 곳에서 하숙 중임. 이 경우 증대표를 중심으로 가구를 구성.	새로운 가구 ID (보험증 ID)
유형6	· 나는 증대표의 미성년 자녀이고, 나의 증대표는 세대주와 재혼하여, 세대주와 모두 같은 집에 살고 있음. 주민등록 신고 시에는 세대주의 동거인으로 신고함.	보험증 대표의 세대주 ID
유형7	· 나의 증대표는 아들임. 집주인과 같은 건물의 다가구 주택에 임대료 함께 살고 있으며, 주민등록상 집주인과 같은 세대로 묶여 있음.	새로운 가구 ID (보험증 ID)
유형8	· 나는 누나의 부부와 함께 살고 있음. 나는 대학생으로 직장생활 중인 누나의 피부양자이며, 세대주는 매형임.	보험증 대표의 세대주 ID
C 유형9	· 나는 주인 세대와 함께 살고 있는 비가족 동거인이며, 증대표(남편)와 실제로는 같이 살지만, 주민등록상 다른 세대임. 증대표인 남편은 직업상의 이유로 옆집 세대에 방을 구하여 필요시마다 사용함. 이 방은 다가구주택으로 남편은 주민등록상 주인 세대의 동거인으로 되어 있음.	새로운 가구 ID (보험증 ID)
유형10	· 나는 내가 거주하는 세대에서 비가족 동거인(하숙). 나의 증대표는 따로 사는 엄마이고, 엄마의 세대주는 아버지임. 이 경우 나는 증대표를 따라 가구를 구성. 증대표의 가구주(세대주)가 나의 가구주가 됨.	보험증 대표의 세대주 ID
유형11	· 나는 하숙. 나의 증대표는 다른 시 지역에 따로 사는 아들. 나는 아들과 독립적으로 생활하지만, 아들의 보험증에 올라가 있음. 이 경우 나의 세대주와도 나의 증대표와도 가구를 구성하지 못함. 즉 세대주 아이디어나 증대표 아이디어만으로는 가구 구성 불가. · 나의 배우자도 나와 같은 곳에 하숙하며 아들의 보험증에 올라가 있다면 나와 배우자는 같은 가구에 속함.	새로운 가구 ID (보험증 ID+ 세대주 ID)

주: 1) 세대주: 주민등록상 세대주

2) 보험증: 건강보험증

3) 증대표: 건강보험증 대표

제3절 가구 단위 정보 생성 결과

1. 가구 구성안별 가구수 및 가구원 수

2009~2021년의 가구 정보를 생성하였으나, 지면 관계상 2009, 2015, 2021년의 3개 연도를 발췌하여 제시하였다.

1안은 주민등록 세대주 또는 건강보험증 대표를 공유하는 사람들을 한 가구로 보는 포괄적인 가구의 개념으로, 다른 안에 비해 가구수가 가장 적고(2021년 2,121만 가구) 가구원 수는 가장 많다(2021년 평균 2.48명). 5안은 건강보험증 단위로, 같은 주민등록 세대라고 하더라도 경제활동을 하는 가구원은 다른 가구로 분리될 뿐만 아니라, 최근 건강보험 피부양자 기준의 강화 등의 영향을 받아, 가구수가 가장 많고(2021년 2,762만 가구), 가구원 수는 가장 적다(2021년 평균 1.9명). 2~4안은 1안, 5안에 비해 가구수와 가구원 수가 비슷한 수준이나, 2안과 4안에 비해 3안에서 상대적으로 가구수가 많고, 가구원 수는 적다(〈표 3-13〉).

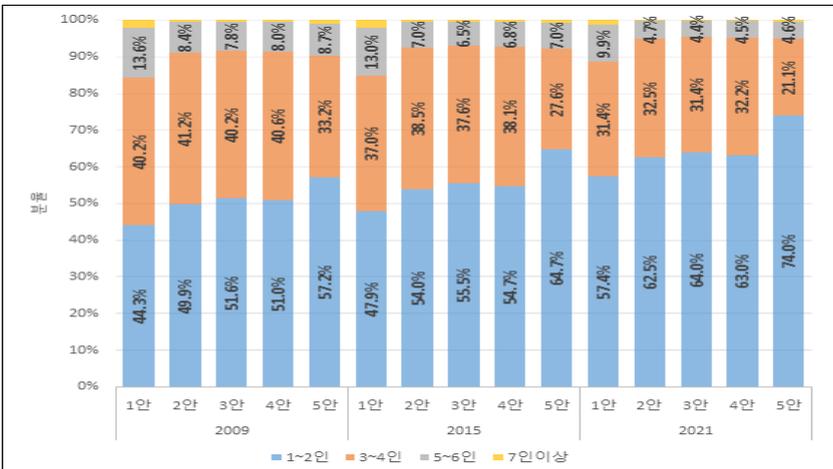
가구원 수별 가구 분포에서도 유사한데, 전체 가구 수 중 1~2인 가구의 비율은 1안이 가장 적고(2021년 57.4%), 5안이 가장 많다(2021년 74%). 연도별로는 2009년에 비해 2021년에 1~2인 가구의 비율이 높고, 3인 이상 가구의 비율은 낮다(〔그림 3-8〕).

〈표 3-13〉 가구 구성연별 가구 수 및 가구원 수(2009, 2015, 2021)

가구 구성안	기준 연도	가구 수 (단위: 만 가구)	가구원 수(단위: 명)			
			평균	SD	최소	최대
세대보완 1안 (건강보험 피부양자 포함)	2009	1,729	2.89	1.67	1	413
	2015	1,848	2.79	1.67	1	381
	2021	2,121	2.48	1.58	1	362
세대보완 2안 (비동거 가족 포함)	2009	1,906	2.62	1.47	1	413
	2015	2,066	2.50	1.42	1	381
	2021	2,345	2.24	1.32	1	362
세대보완 3안 (분리세대 추정)	2009	1,949	2.56	1.45	1	413
	2015	2,110	2.45	1.40	1	381
	2021	2,395	2.19	1.30	1	362
세대기준안	2009	1,930	2.59	1.70	1	1,861
	2015	2,082	2.48	1.57	1	1,163
	2021	2,352	2.23	1.40	1	709
건강보험증 기준안	2009	2,073	2.41	1.54	1	413
	2015	2,362	2.19	1.45	1	381
	2021	2,762	1.90	1.28	1	362

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〔그림 3-8〕 가구원 수별 가구 분포



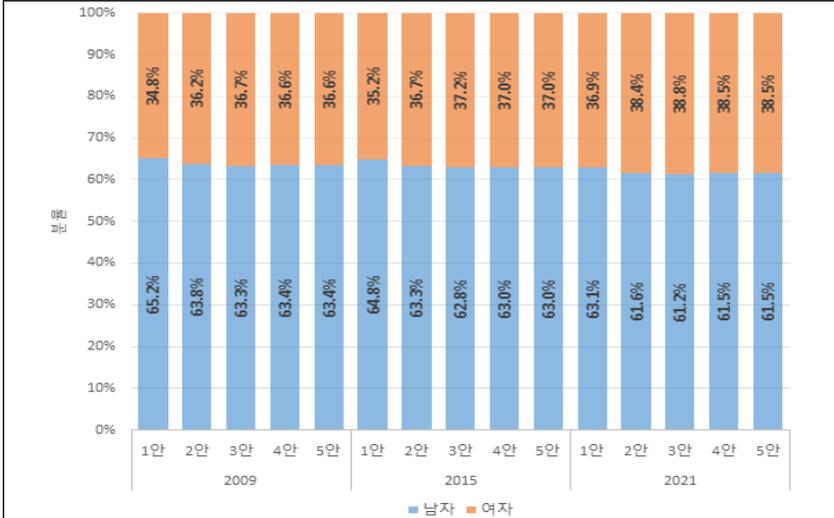
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

2. 가구 특성별 가구 분포

1) 가구주 특성별 가구 분포

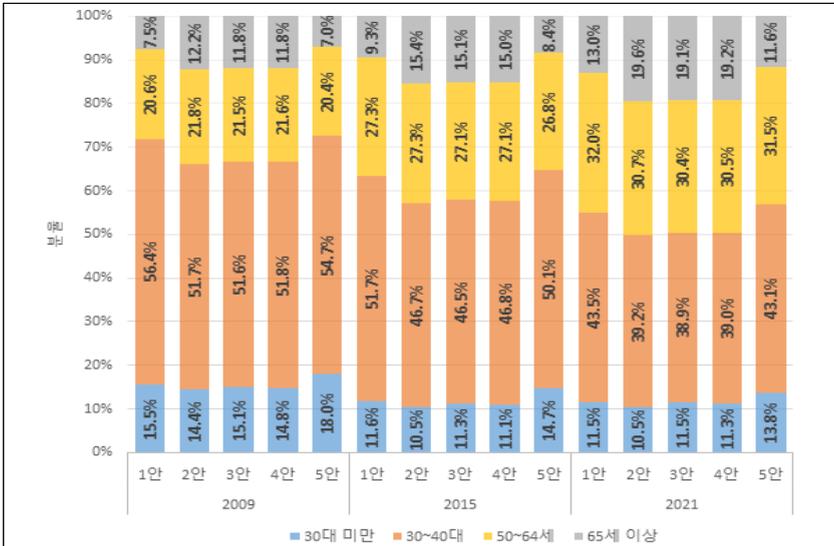
가구주 특성에 따른 가구 분포는 다음과 같다. 가구주 성별에 따라서는 가구 구성안별 연도별로 큰 차이가 없는 것으로 보인다([그림 3-9]). 가구주 연령별로 보면, 2~4안은 1안, 5안에 비해 65세 이상 가구주의 비율이 높고, 30~40대 가구주의 비율이 낮다. 2~4안은 주소가 다른 노인 부모 세대가 다른 가구로 분리되기 때문인 것으로 보인다. 5안은 1~4안에 비해 30대 미만 가구주의 비율이 높다. 주민등록 세대가 같더라도, 경제 활동을 하는 가구원이라면 별도의 가구로 분리되는 특징이 두드러진다. 연도별로는 2009년에 비해 2021년에 50세 이상 가구주의 비율이 현저히 증가하는 모습이다([그림 3-10]). 가구주 자격별로는 1안이 2~5안에 비해 직장가입자 가구주가 적고, 지역가입자 가구주가 많다. 연도별로는 모든 구성안에서 점차 직장가입자의 비율이 높아지고 지역가입자의 비율은 줄고 있다([그림 3-11]).

[그림 3-9] 가구주 특성별 가구 분포: 성별



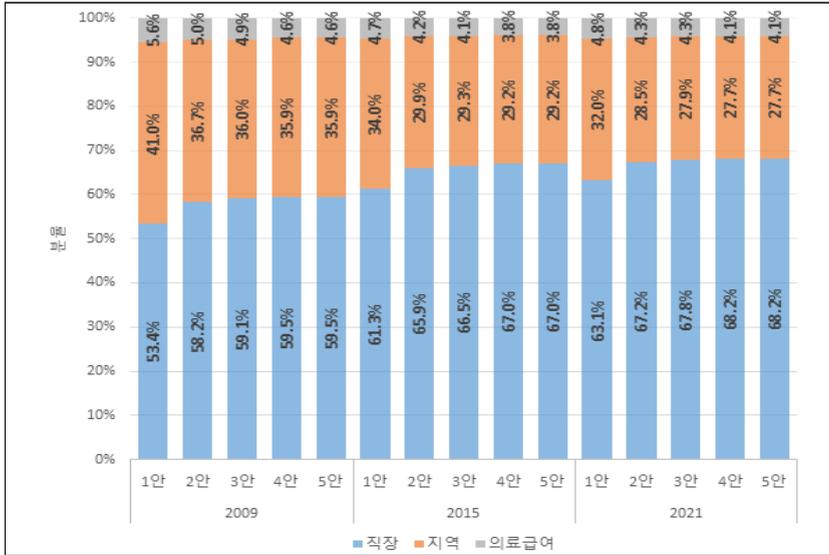
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

[그림 3-10] 가구주 특성별 가구 분포: 연령별



자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

[그림 3-11] 가구주 특성별 가구 분포: 자격별



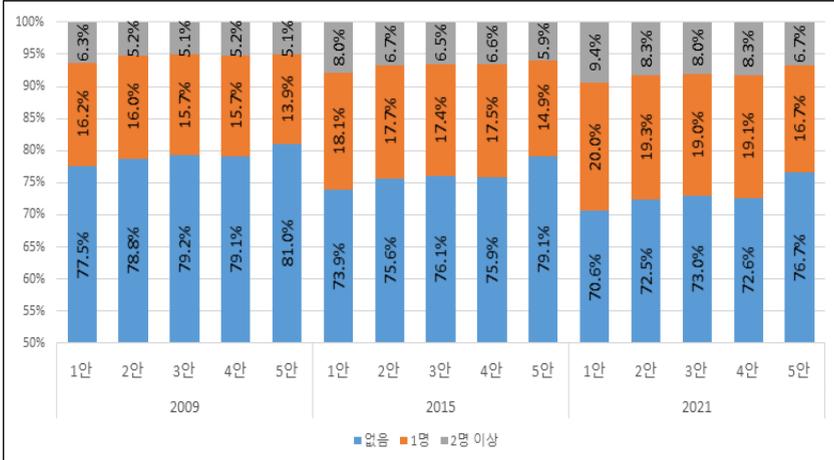
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

2) 65세 이상 노인 가구원 유무에 따른 가구 분포

65세 이상 노인 가구원 유무에 따른 가구 분율을 분석한 결과는 다음과 같다. 노인 가구원이 있는 가구의 비율이 가장 높은 가구 구성안은 1안이며, 가장 낮은 가구 구성안은 건강보험증기준안인의 5안이다. 2~4안은 유사한 수준이다([그림 3-12]).

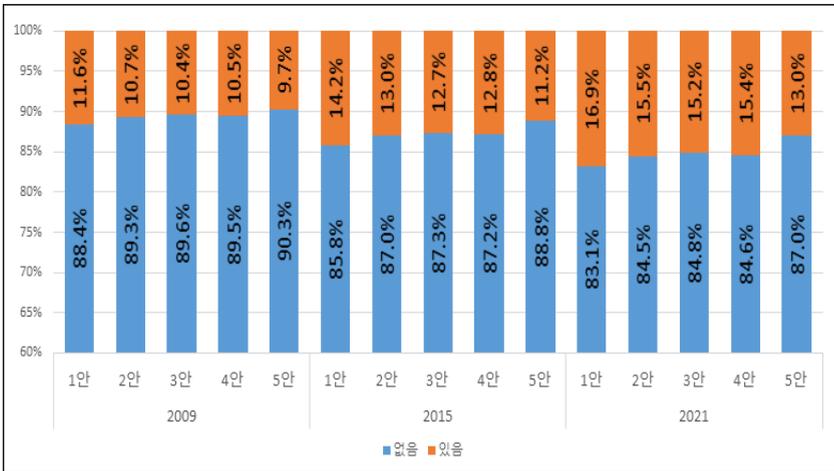
남자와 여자를 비교해보면, 여자 노인이 있는 가구의 분율이 전반적으로 더 높았다. 2021년 기준, 남자 노인이 있는 가구는 13.0~16.9% 수준이었고, 여자 노인이 있는 가구는 16.8~21.6% 수준이었다. 가구안별로는 노인 가구원 성별에 따라 패턴이 다르지 않은 것으로 보인다([그림 3-13], [그림 3-14]).

[그림 3-12] 노인 가구원이 있는 가구 비율: 65세 이상



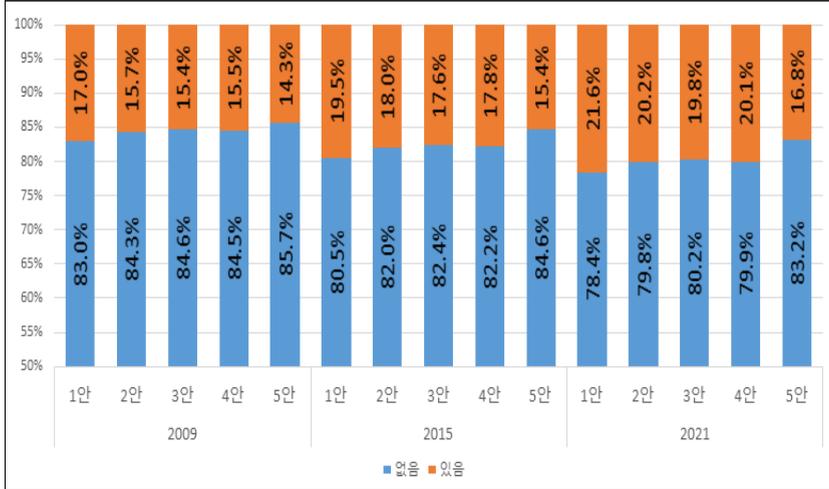
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

[그림 3-13] 노인 가구원이 있는 가구 비율: 65세 이상 남자



자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

[그림 3-14] 노인 가구원이 있는 가구 비율: 65세 이상 여자



자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

제4절 소결

가구 구성안은 총 5개로, 기존의 주민등록 세대를 보완한 3개 안(1안: 건강보험 피부양자 포함, 2안: 비동거 가족 포함, 3안: 분리세대 추정)과 세대기준(4안), 건강보험증기준(5안)으로 구성하였다.

가구 구성안 구축을 위해 1, 2차에 걸쳐 각 가구원의 유형을 구분하였다. 1차 유형 구분 시에는 각 주민등록 세대별로 세대주 중심의 가구가 구성된다고 했을 때, 세대주와 각 세대원의 정보를 비교하여, 가구 구성원으로 인정 가능한 경우와 그렇지 않은 경우로 유형을 구분한다. 2차 유형 구분 시에는 1차 유형에 기반하여 최종 가구 ID를 정의하게 된다. 주민등록 세대는 다르지만 건강보험증 대표의 가구원으로 귀속될 조건에 해당하는지를 확인하고, 건강보험증 대표의 가구로 정의한다. 이 조건에 부합하지

않으면 1차 유형 구분에서 잠정적으로 정의한 대로 가구가 정의된다. 1~3안의 차이는 건강보험증 대표와 다른 주민등록 세대일 때, 건강보험증 대표의 가구원으로 속하게 하는 조건에 있다. 1안은 조건 없이 해당 건강보험증 대표의 가구로 귀속된다. 2안은 건강보험증 대표의 '배우자' 또는 '직계비속'인 경우, 3안은 건강보험증 대표의 '배우자' 또는 '직계비속'이면서, 건강보험증 대표와 동일 주소일 때 건강보험증 대표의 가구로 귀속된다.

가구 구축 결과, 1안이 2,121만 가구로, 구축된 가구수가 가장 적었고, 5안이 2,762만 가구로, 가구수가 가장 많았다(2021년 기준). 1안은 주민등록 세대주 또는 건강보험증 대표를 공유하는 사람들을 한 가구로 보는 가장 포괄적인 가구 개념이며, 5안은 건강보험증 단위로, 같은 세대에 속함에도 불구하고 경제활동을 하는 가구원은 다른 가구로 분리되기 때문이다. 2~4안은 주민등록 세대 중심의 가구 구성으로, 구성인에 따라 다른 세대에 거주하는 배우자와 직계비속을 한 가구로 묶어주는 등의 차이가 있으나, 구축된 가구수는 2,345~2,395만 가구로 유사하였다.

65세 이상 노인 '가구주'의 경우에는 1안, 5안에서 그 비율이 상대적으로 낮고, 2~4안에서 더 높은 점이 두드러진다. 65세 이상 노인이 있는 가구를 파악해본 결과, 특히 1안에서 노인이 있는 가구 비율이 가장 높았으며, 5안에서 노인이 있는 가구 비율이 가장 낮았다. 즉 1안에서 노인은 가구주가 아닌 가구원 자격으로 자녀 세대에 편입되는 특성을 보여주고 있다.

본 연구는 선행연구를 참고하여 같은 곳에 거주하면 생계를 같이 한다고 보는 것을 가구 식별의 기본 원칙으로 하였다. 따라서 세대가 같을 때 가구원의 연령, 소득활동 여부, 가족관계의 범위(촌수)에 대한 제한을 하지 않고 동거인으로 신고한 경우에만 가구에서 제외하였다. 1~3안에서 세대가 다를 경우는 연령, 소득활동 여부, 가족관계 범위 등에 따라 가구원 여부를

결정할 필요가 있는데, 본 연구에서는 별도의 기준을 정하지 않고, ‘건강보험증’의 정보를 사용함으로써 건강보험 피부양자 인정기준을 준용하고자 하였다.

구축된 가구 ID는 연도별 1월의 주민등록 세대 또는 건강보험증 정보에 기초하여 정의된 것이다. 따라서 실제 생계를 같이하지만 해당 시점에 다른 주소지에 거주하거나, 건강보험증이 분리되어 있는 경우는 가구로 식별되지 않는다는 한계를 가진다. 또한 시설입소자로 구성되는 경우, 해당 가구의 경제적 수준을 가구원의 경제적 수준으로 보기 어려울 수 있으나, 본 자료의 정보로는 시설입소자에 대한 식별이 어려웠다. 따라서 시설입소자를 제외한 분석이 필요할 경우, 가구원 수 등을 이용한 조작적 정의를 적용할 수 있을 것이다.



제4장

소득 분포 분석: 가계금융복지조사와의 비교

제1절 서론

제2절 맞춤형 소득-재산 DB와 가계금융복지조사의 차이

제3절 분석 결과

제4절 소결

제 4 장

소득 분포 분석: 가계금융복지조사와의 비교

제1절 개요

이 장에서는 맞춤형 소득-재산 DB를 활용하여 소득 분포를 분석한다. 소득분배 분석을 위한 맞춤형 소득-재산 DB의 유용성을 판단하기 위해서는 공식 소득분배지표 작성에 활용되고 있는 가계금융복지조사와의 비교가 필요하다. 국세청, 보건복지부 등의 행정자료로 소득 및 비소비지출 조사 결과를 보완한 가계금융복지조사는 가장 신뢰할 만한 소득조사 마이크로데이터로 평가할 수 있다. 가계금융복지조사로 분석한 지니계수, 빈곤율 등의 소득분배지표는 각종 사회보장정책의 주요 근거 자료로 활용되고 있다. 이 장에서는 가계금융복지조사와 비교하여 맞춤형 소득-재산 DB의 소득분배 분석 데이터로서의 활용 가능성을 진단한다. 제2절에서는 맞춤형 소득-재산 DB와 가계금융복지조사의 소득 분포 차이에 영향을 미치는 요인을 논의하고, 제3절에서는 맞춤형 소득-재산 DB와 가계금융복지조사의 개인 및 가구 소득 분포를 실증적으로 비교·분석한다.

제2절 맞춤형 소득-재산 DB와 가계금융복지조사의 차이

이 장에서는 맞춤형 소득-재산 DB와 가계금융복지조사로 2019년 소득 분포를 분석한다. 두 데이터의 최신 소득 연도는 2020년이지만, 코로나 19의 소득 충격으로 인한 일반화의 한계를 고려하여 이 장에서는 2019년 소득 분포를 분석한다.¹⁾

〈표 4-1〉 가계금융복지조사 및 맞춤형 소득-재산 DB의 소득 개념·범위

가계금융복지조사		맞춤형 소득-재산 DB
근로소득		근로소득
사업소득		사업소득
재산소득	금융소득	이자소득
		배당소득
		분리과세금융소득
	임대소득	-
기타 재산소득	-	
공적이전소득	공적연금	국민연금소득
		지역연금소득
	기초연금	-
	아동 관련 급여	-
	장애 관련 급여	-
	기초보장급여	-
	근로·자녀장려금	-
기타 공적이전소득	-	
사적이전소득	가구 간 이전소득	-
	비영리단체 이전소득	-

주: 음영은 대체로 비교 가능한 소득 항목을 의미한다. 맞춤형 소득-재산 DB의 분리과세주택임대 소득과 기타소득은 가계금융복지조사 소득 항목과 비교하기 어렵다고 판단하여 제외하였다. 자료: 가계금융복지조사 조사표(통계청, 2020a)와 맞춤형 소득-재산 DB(국민건강보험공단, 2020)를 활용하여 필자가 작성하였다.

이론적으로, 맞춤형 소득-재산 DB와 가계금융복지조사의 가구소득분배 차이를 발생시키는 요인은 크게 다음과 같이 네 가지로 구분할 수 있다. 첫째, 두 데이터의 소득 개념·범위에 차이가 있다. 가계금융복지조사는 근로소득, 사업소득, 재산소득, 공적이전소득·지출, 사적이전소득·지출 등과 같이 가처분소득을 구성하는 데 필요한 소득 및 비소비지출 항목을 빠짐없이 조사한다. 하지만 이 연구가 구축한 맞춤형 소득-재산 DB는 근로

1) 2019년 소득 자료를 포함한 맞춤형 소득-재산 DB 기준 연도 및 가계금융복지조사 조사 연도는 2020년이다. 이 장에서는 소득 연도를 기준으로 서술한다.

소득, 사업소득, 이자소득, 배당소득, 분리과세금융소득, 분리과세주택 임대소득, 기타소득, 국민연금소득, 직역연금소득을 포함한다. 저소득층일수록 맞춤형 소득-재산 DB에 포함되지 않은 공적·사적이전소득을 많이 받으므로, 맞춤형 소득-재산 DB가 소득불평등 및 빈곤을 과대 추정할 가능성이 크다. 이 장에서는 <표 4-1>과 같이 가계금융복지조사 및 맞춤형 소득-재산 DB의 근로소득, 가계금융복지조사 및 맞춤형 소득-재산 DB의 사업소득, 가계금융복지조사의 금융소득과 맞춤형 소득-재산 DB의 <이자소득+배당소득+분리과세금융소득>, 가계금융복지조사의 공적연금과 맞춤형 소득-재산 DB의 <국민연금소득+직역연금소득>을 비교한다. 맞춤형 소득-재산 DB의 분리과세주택임대소득과 기타소득은 가계금융복지조사 소득 항목과 비교하기 어렵다고 판단하여 이 장의 분석에서 제외하였다. 서술의 편의를 위해, 이 장에서는 <근로소득+사업소득+금융소득+공적연금소득>을 총소득으로 부른다.

둘째, 맞춤형 소득-재산 DB와 가계금융복지조사의 모집단에 차이가 있고, 가계금융복지조사의 표본 대표성이 완전하지 않다. 가계금융복지조사의 목표 모집단은 전체 가구·인구이지만, 조사 모집단은 기숙시설·특수사회시설 거주자를 제외한다(통계청, 2020b, p.21). 또한 가계금융복지조사는 약 2만 가구에 대한 표본조사이므로, 무응답 및 조사 거부가 표본 대표성에 영향을 미쳤을 가능성이 있다. 가계금융복지조사는 지역 및 주택유형으로 층화하여 표본을 추출하지만(통계청, 2020b, p.24), 층화집단 내에서 소득계층별 무응답 확률이 상이할 가능성을 배제할 수 없다. 예를 들어, 고소득층이 상대적으로 조사를 강하게 거부하였다면 고소득층 누락으로 소득불평등이 과소추정되었을 것이다. 가계금융복지조사의 행정자료 보완 작업은 표본으로 추출된 사례의 소득 측정오차를 크게 감소시켰지만(이원진, 정해식, 전지현, 2019, pp.119-168), 표본

대표성의 한계를 완화하지는 못했다. 한편, 가계금융복지조사는 가구 단위 표본조사이므로 원칙적으로 전체 가구를 대표하도록 표본이 설계되었지만, 개인 단위로 사후층화하지 않으므로 표본가구에 속한 가구원의 집합과 모집단 개인의 특성에 약간의 차이가 존재할 수 있다. 이 장에서는 맞춤형 소득-재산 DB와 가계금융복지조사의 개인 단위 성·연령·소득 분포를 분석하여 모집단 차이 및 표본 대표성 문제를 탐색한다.

셋째, 맞춤형 소득-재산 DB와 가계금융복지조사의 소득 측정 정확도에 차이가 있다. 원칙적으로 맞춤형 소득-재산 DB는 행정자료를 수집하였으므로 소득 측정의 정확도가 상당히 높지만, 부분적으로 조세·건강보험 행정으로 포착되지 않는 소득이 존재한다. 통계청(2020c, p.56)은 가계금융복지조사를 행정자료로 보완할 때 미과세 금액 이하 근로소득은 조사자료와 행정자료 중 큰 값을 사용하고, 행정자료의 정확도가 낮은 사업소득은 기본적으로 조사자료를 사용하되 무응답을 행정자료로 대체한다. 따라서 서베이에서 누락되었을 가능성이 큰 초고소득층의 근로소득은 맞춤형 소득-재산 DB가 상대적으로 정확하게 포착하였을 것으로 예상되지만, 소득세를 납부하지 않는 저임금근로자의 근로소득은 가계금융복지조사가 상대적으로 정확하게 포착하였을 것으로 예상된다. 또한 사업소득은 대체로 맞춤형 소득-재산 DB보다 가계금융복지조사가 정확하게 측정하였을 것으로 예상된다. 한편, 통계청(2020c, p.56)은 가계금융복지조사를 행정자료로 보완할 때 금융소득과 공적연금만 조사자료를 행정자료로 전면적으로 대체하는데, 맞춤형 소득-재산 DB에 포함되지 않은 비과세 원천분 금융소득까지 포함한다(통계청, 2020c, p.44).

넷째, 맞춤형 소득-재산 DB와 가계금융복지조사의 가구 개념에 차이가 있다. 가계금융복지조사의 표본추출틀인 인구주택총조사는 주민등록 세대를 토대로 동일 주소 친족 가구 통합, 인접 지역 배우자·자녀 통합

등의 보완을 거쳐 가구를 정의한다(통계청, 2020d, pp.80-81). 또한 가계 금융복지조사는 인구주택총조사와 달리 학업·직장 때문에 따로 사는 미혼 자녀·배우자를 가구에 포함하고, 동거하지만 가족이 아닌 구성원을 가구에서 제외한다(통계청, 2021, p.26). 한편, 맞춤형 소득-재산 DB에서는 세대보완 1안(건강보험 피부양자 포함), 세대보완 2안(비동거 가족 포함), 세대보완 3안(분리세대 추정), 세대기준안, 건강보험증기준안의 5개 가구 구성안을 구축하였다. 맞춤형 소득-재산 DB에서 자녀·배우자·노부모를 통합하는 방향으로 가구를 구성할수록 가구소득분배가 개선될 것으로 예상된다.

지금까지 논의한 것처럼, 맞춤형 소득-재산 DB와 가계금융복지조사에서 가구소득분배의 차이는 크게 소득 개념·범위의 차이, 모집단 차이 및 가계금융복지조사 표본 대표성의 한계, 소득 측정 정확도의 차이, 가구 개념의 차이로 설명할 수 있다. 이와 같은 네 가지 요인의 영향을 정확하게 분해하기는 불가능하지만, 이 장에서는 두 데이터의 소득 분포를 다각적으로 검토하여 그 차이를 이해하기 위해 노력한다.

제3절 분석 결과

1. 개인소득 분포 분석

이하에서는 우선 맞춤형 소득-재산 DB와 가계금융복지조사의 개인소득 분포를 살펴본다. 두 데이터의 가구 내에서 합산·균등화하지 않은 개인 근로소득·사업소득·금융소득·공적연금소득 분포 차이는 주로 모집단 차이 및 가계금융복지조사 표본 대표성의 한계, 소득 측정 정확도의 차이로 설명할 수 있을 것이다.

〈표 4-2〉 2019년 개인 사례 수

(단위: 명)

구분		(A) 건보	(B) 가금복
남성	9세 이하	2,165,779	1,838
	10~19세	2,573,572	2,318
	20~29세	3,723,774	2,287
	30~39세	3,805,775	2,427
	40~49세	4,312,813	3,306
	50~59세	4,402,372	3,554
	60~69세	3,104,607	3,172
	70~79세	1,591,146	1,940
	80~89세	554,477	780
	90세 이상	48,306	55
여성	9세 이하	2,055,301	1,723
	10~19세	2,394,957	2,184
	20~29세	3,300,053	2,370
	30~39세	3,544,105	2,560
	40~49세	4,165,478	3,367
	50~59세	4,357,257	3,800
	60~69세	3,247,493	3,544
	70~79세	1,977,989	2,433
	80~89세	1,058,931	1,290
	90세 이상	174,305	147
계		52,558,490	45,095

주: 연령 기준 시점은 '건보' 2019년 12월 말, '가금복' 2020년 3월 말이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복: 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

우선 〈표 4-2〉에는 두 데이터의 사례 수를 보고하였다. 맞춤형 소득-재산 DB와 가계금융복지조사의 사례 수는 각각 5,256만 명, 45,095명이었다. 〈표 4-3〉에서 성·연령집단 비율을 분석한 결과, 두 데이터에 약간의 차이가 관찰되었다. 맞춤형 소득-재산 DB와 비교하면, 가계금융

복지조사는 20대 남성 비율 및 30대 남성 비율이 각각 1.9%p, 1.7%p 낮았고, 20대 여성 비율 및 30대 여성 비율이 각각 0.8%p, 0.7%p 낮았다. 이는 가구 단위 표본조사인 가계금융복지조사가 개인 성·연령 분포를 정확하게 포착하는 데 한계가 있다는 사실을 보여준다.

〈표 4-3〉 2019년 개인 성·연령 분포

(단위: %, %p)

구분		(A) 건보	(B) 가금복	(B-A)
남성	9세 이하	4.1	4.6	0.5
	10~19세	4.9	5.7	0.8
	20~29세	7.1	5.2	-1.9
	30~39세	7.2	5.5	-1.7
	40~49세	8.2	7.9	-0.3
	50~59세	8.4	8.0	-0.4
	60~69세	5.9	6.3	0.4
	70~79세	3.0	3.5	0.5
	80~89세	1.1	1.4	0.4
	90세 이상	0.1	0.1	0.0
여성	9세 이하	3.9	4.3	0.4
	10~19세	4.6	5.5	0.9
	20~29세	6.3	5.5	-0.8
	30~39세	6.7	6.0	-0.7
	40~49세	7.9	8.2	0.3
	50~59세	8.3	8.3	0.0
	60~69세	6.2	6.8	0.7
	70~79세	3.8	4.5	0.7
	80~89세	2.0	2.3	0.3
	90세 이상	0.3	0.3	-0.1
계		100.0	100.0	-

주: 연령 기준 시점은 '건보' 2019년 12월 말, '가금복' 2020년 3월 말이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복: 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

성·연령에 따라 소득 분포가 달라지므로, 맞춤형 소득-재산 DB와 가계금융복지조사의 성·연령 분포 차이는 두 데이터의 소득 분포 차이에 영향을 미칠 것이다. 이와 같은 영향을 확인하기 위해, 이하에서 가계금융복지조사를 분석할 때 실제 표본과 성·연령집단 비율을 맞춤형 소득-재산 DB와 동일하게 조정한 재가중 표본을 모두 살펴본다. 예를 들어, 가계금융복지조사 실제 표본에서 개인근로소득이 0원인 비율은 53.4%였지만, 고용률이 상대적으로 높은 20대·30대 비율을 맞춤형 소득-재산 DB와 동일하게 상향조정하면 개인근로소득이 0원인 비율이 50.5%로 감소하였다(〈표 4-4〉 참조). 가계금융복지조사 실제 표본과 맞춤형 소득-재산 DB의 개인근로소득 0원 비율의 차이는 10.1%p였지만, 가계금융복지조사 재가중 표본과 맞춤형 소득-재산 DB의 개인근로소득 0원 비율의 차이는 13.0%p였다(〈표 4-4〉 참조). 즉, 맞춤형 소득-재산 DB에 비해 가계금융복지조사에서 20대·30대가 과소포착되었다는 사실을 고려하면, 맞춤형 소득-재산 DB와 가계금융복지조사에서 개인근로소득 분포의 차이가 실제로 관찰된 것보다 좀 더 크다고 평가할 수 있다.

이와 같은 방식으로 〈표 4-4〉를 살펴보면, 맞춤형 소득-재산 DB보다 가계금융복지조사가 개인근로소득을 더 많이 포착한 것으로 나타났다. 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사 재가중 표본과 맞춤형 소득-재산 DB를 비교하면, 개인근로소득이 존재하지 않는 비율이 맞춤형 소득-재산 DB에서 13.0%p 높았고, 개인근로소득이 0원 초과 2,000만 원 이하인 저소득 비율이 가계금융복지조사에서 10.4%p 높았으며, 개인근로소득이 2,000만 원 초과 1억 원 이하인 중고소득 비율이 가계금융복지조사에서 2.6%p 높았다. 이러한 분석 결과는 맞춤형 소득-재산 DB에 비해 가계금융복지조사에서 행정자료로 포착되지 않는 근로소득을 상대적으로 많이 조사하였을 것이라는

예측을 지지한다. 한편, 개인근로소득이 1억 원을 초과하는 고소득 비율은 두 데이터에서 거의 유사한 수준으로 나타났다. 이는 가계금융복지조사에서 고소득층 누락 현상이 그다지 강하게 나타나지 않을 가능성을 시사한다.

〈표 4-4〉 2019년 개인근로소득 분포, 전체 인구

(단위: %, %p, 만 원/년)

구분	(A) 건보	(B) 가금복 실제	(C) 가금복 재가중	(B-A)	(C-A)	(C-B)
0원	63.5	53.4	50.5	-10.1	-13.0	-2.9
~연 100만 원 이하	0.4	2.0	2.1	1.7	1.7	0.1
~연 300만 원 이하	1.2	3.2	3.3	2.0	2.1	0.1
~연 500만 원 이하	1.3	2.4	2.5	1.1	1.2	0.1
~연 1000만 원 이하	2.8	4.8	5.1	2.0	2.2	0.2
~연 2000만 원 이하	6.0	8.7	9.1	2.7	3.1	0.4
~연 3000만 원 이하	8.3	8.5	9.0	0.2	0.7	0.5
~연 4000만 원 이하	4.9	4.8	5.3	-0.1	0.4	0.5
~연 6000만 원 이하	5.5	5.6	6.3	0.1	0.7	0.6
~연 8000만 원 이하	2.9	3.1	3.3	0.2	0.4	0.2
~연 10000만 원 이하	1.6	1.8	1.9	0.2	0.3	0.1
연 10000만 원 초과	1.6	1.6	1.7	0.0	0.1	0.1
계	100.0	100.0	100.0	-	-	-
평균	1361	1433	1539	72	178	106
표준편차	3328	2698	2757	-630	-571	59

주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정된 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

경제활동에 참여하지 않는 아동과 노인을 포함하여 분석한 〈표 4-4〉와 달리, 〈표 4-5〉에서는 18~64세 근로연령인구를 대상으로 동일한 분석을

실시하였다. 표를 살펴보면, 맞춤형 소득-재산 DB에 비해 가계금융복지조사 재가중 표본에서 개인근로소득이 0원인 비율이 15.9%p 낮았고, 0원 초과 2,000만 원 이하인 비율이 11.9%p 높았으며, 2,000만 원 초과 1억 원 이하인 비율이 3.9%p 높았고, 1억 원 초과인 비율이 2.3~2.4%로 유사하였다. 즉, 근로연령인구를 분석한 결과에서도 가계금융복지조사의 근로소득이 상대적으로 크게 포착된 것으로 나타났다.

〈표 4-5〉 2019년 개인근로소득 분포, 18~64세

(단위: %, %p, 만 원/년)

구분	(A) 건보	(B) 가금복 실제	(C) 가금복 재가중	(B-A)	(C-A)	(C-B)
0원	50.5	35.5	34.6	-15.0	-15.9	-0.9
~연 100만 원 이하	0.4	2.5	2.6	2.1	2.1	0.0
~연 300만 원 이하	1.5	3.3	3.3	1.7	1.8	0.1
~연 500만 원 이하	1.6	2.8	2.8	1.1	1.2	0.1
~연 1000만 원 이하	3.7	6.3	6.4	2.7	2.7	0.1
~연 2000만 원 이하	7.7	11.8	11.7	4.0	4.0	0.0
~연 3000만 원 이하	11.1	12.1	12.2	1.0	1.1	0.1
~연 4000만 원 이하	6.9	7.2	7.5	0.3	0.6	0.3
~연 6000만 원 이하	7.8	8.6	9.0	0.8	1.1	0.3
~연 8000만 원 이하	4.2	4.8	4.8	0.6	0.6	0.0
~연 10000만 원 이하	2.3	2.8	2.7	0.5	0.5	0.0
연 10000만 원 초과	2.3	2.4	2.4	0.2	0.1	0.0
계	100.0	100.0	100.0	-	-	-
평균	1886	2124	2146	238	260	22
표준편차	3685	3093	3082	-592	-603	-11

주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-6〉 2019년 개인사업소득 분포, 전체 인구

(단위: %, %p, 만 원/년)

구분	(A) 건보	(B) 가금복 실제	(C) 가금복 재가중	(B-A)	(C-A)	(C-B)
0원	79.8	80.9	80.4	1.1	0.6	-0.5
~연 100만 원 이하	5.7	4.4	4.6	-1.3	-1.1	0.2
~연 300만 원 이하	3.5	2.1	2.1	-1.4	-1.3	0.1
~연 500만 원 이하	2.2	1.2	1.2	-1.0	-1.0	0.0
~연 1000만 원 이하	3.1	1.8	1.8	-1.3	-1.3	0.0
~연 2000만 원 이하	2.7	2.7	2.7	0.0	0.0	0.0
~연 3000만 원 이하	1.1	2.2	2.2	1.1	1.1	0.0
~연 4000만 원 이하	0.6	1.5	1.6	0.9	1.0	0.1
~연 6000만 원 이하	0.6	1.8	1.9	1.3	1.3	0.1
~연 8000만 원 이하	0.3	0.6	0.7	0.4	0.4	0.0
~연 10000만 원 이하	0.1	0.4	0.4	0.2	0.3	0.0
연 10000만 원 초과	0.4	0.4	0.4	0.0	0.0	0.0
계	100.0	100.0	100.0	-	-	-
평균	278	435	449	157	171	13
표준편차	3531	2391	2429	-1140	-1102	38

주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정된 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-6〉에서 개인사업소득 분포를 살펴보면, 맞춤형 소득-재산 DB에 비해 가계금융복지조사에서 개인사업소득이 0원 초과 1,000만 원 이하인 비율이 낮았고 2,000만 원 초과 1억 원 이하인 비율이 높았다. 이처럼 가계금융복지조사가 개인사업소득 수준이 높은 집단을 좀 더 많이 포착한 결과 개인사업소득 평균이 훨씬 높게 나타났다. 이러한 분석 결과는 조세행정에서 고소득 자영업자의 소득이 과소 보고되고 있을 가능성을 시사한다.

단, 개인사업소득이 존재하는 비율은 가계금융복지조사보다 맞춤형 소득-재산 DB에서 조금 더 높게 나타났다.

〈표 4-7〉 2019년 개인사업소득 분포, 18~64세

(단위: %, %p, 만 원/년)

구분	(A) 건보	(B) 가금복 실제	(C) 가금복 재가중	(B-A)	(C-A)	(C-B)
0원	73.9	75.3	75.5	1.3	1.6	0.3
~연 100만 원 이하	7.4	6.0	6.0	-1.4	-1.4	0.1
~연 300만 원 이하	4.4	2.5	2.5	-1.9	-1.9	0.0
~연 500만 원 이하	2.9	1.4	1.4	-1.5	-1.5	0.0
~연 1000만 원 이하	4.0	1.9	1.9	-2.0	-2.1	0.0
~연 2000만 원 이하	3.5	3.2	3.1	-0.3	-0.4	-0.1
~연 3000만 원 이하	1.5	3.0	2.9	1.5	1.4	-0.1
~연 4000만 원 이하	0.7	2.1	2.1	1.4	1.4	0.0
~연 6000만 원 이하	0.7	2.6	2.6	1.9	1.8	-0.1
~연 8000만 원 이하	0.3	0.9	0.9	0.6	0.6	0.0
~연 10000만 원 이하	0.2	0.6	0.6	0.4	0.4	0.0
연 10000만 원 초과	0.5	0.6	0.5	0.1	0.1	0.0
계	100.0	100.0	100.0	-	-	-
평균	356	604	591	248	235	-13
표준편차	3632	2909	2867	-723	-765	-43

주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정된 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-7〉에서 18~64세 근로연령인구를 대상으로 개인사업소득 분포를 분석한 결과 역시 〈표 4-6〉과 대체로 유사하였다. 맞춤형 소득-재산 DB에 비해 가계금융복지조사에서 개인사업소득이 2,000만 원을 넘는 비율이

상당히 높았지만, 개인사업소득이 존재하는 비율은 다소 낮았다.

〈표 4-8〉 2019년 개인금융소득 분포, 전체 인구

(단위: %, %p, 만 원/년)

구분	(A) 건보	(B) 가금복 실제	(C) 가금복 재가중	(B-A)	(C-A)	(C-B)
0원	99.3	40.8	40.3	-58.5	-59.0	-0.5
~연 10만 원 이하	0.0	25.4	26.2	25.4	26.2	0.8
~연 50만 원 이하	0.0	14.4	14.5	14.4	14.5	0.1
~연 100만 원 이하	0.0	6.2	6.2	6.2	6.2	-0.1
~연 200만 원 이하	0.0	5.7	5.5	5.7	5.5	-0.2
~연 300만 원 이하	0.0	2.5	2.5	2.5	2.5	-0.1
~연 400만 원 이하	0.0	1.4	1.3	1.4	1.3	0.0
~연 600만 원 이하	0.0	1.4	1.4	1.4	1.4	0.0
~연 800만 원 이하	0.0	0.7	0.6	0.7	0.6	0.0
~연 1000만 원 이하	0.0	0.4	0.4	0.4	0.4	0.0
~연 2000만 원 이하	0.4	0.8	0.7	0.3	0.3	0.0
연 2000만 원 초과	0.3	0.4	0.4	0.1	0.1	0.0
계	100.0	100.0	100.0	-	-	-
평균	34	76	74	42	40	-2
표준편차	1474	614	618	-860	-856	4
평균, 연 1,000만 원 초과 집단	4726	2764	2784	-1963	-1942	21

주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-8〉에서는 개인금융소득 분포를 살펴보았다. 가계금융복지조사에서는 개인금융소득이 존재하는 비율이 약 60%였지만, 맞춤형 소득-재산 DB에서는 0.7%에 불과한 것으로 나타났다. 또한 가계금융복지조사의

금융소득 평균은 76만 원으로 맞춤형 소득-재산 DB의 34만 원보다 상당히 컸다. 이와 같이 두 데이터의 개인금융소득 분포에 매우 큰 차이가 나타난 주된 이유는 맞춤형 소득-재산 DB와 가계금융복지조사의 금융소득 행정자료 범위가 다르기 때문일 것으로 판단된다. 맞춤형 소득-재산 DB는 종합소득세 금융소득 및 분리과세금융소득 정보를 토대로 금융소득을 구축하였지만, 가계금융복지조사는 비과세 원천분 금융소득까지 포함하여 금융소득을 구축하였다. 이러한 금융소득 행정자료 범위 차이를 고려하여 1,000만 원 초과 집단을 대상으로 개인금융소득 평균을 살펴보면, 맞춤형 소득-재산 DB의 개인금융소득 평균이 4,726만 원으로, 가계금융복지조사보다 훨씬 크게 나타났다.

다음으로 <표 4-9>에서 개인공적연금소득 분포를 살펴보면, 가계금융복지조사의 개인공적연금소득 존재 비율이 11.9%로 맞춤형 소득-재산 DB보다 3.1%p 높았다. 하지만 가계금융복지조사의 노인 비율이 상대적으로 높다는 사실을 고려하여 성·연령 분포를 조정하면, 가계금융복지조사 재가중 표본의 개인공적연금소득 존재 비율이 10.6%로, 맞춤형 소득-재산 DB보다 1.8%p 높게 나타났다. 즉, 성·연령 분포의 차이를 고려하더라도 가계금융복지조사의 개인공적연금소득 존재 비율이 소폭 높았다. 또한 가계금융복지조사 재가중 표본의 개인공적연금소득 평균은 82만 원으로, 맞춤형 소득-재산 DB의 71만 원보다 크게 나타났다. 이와 같이 맞춤형 소득-재산 DB보다 가계금융복지조사의 개인공적연금소득이 더 크게 포착된 주된 이유는 행정자료 범위가 다르기 때문일 것으로 판단된다. 맞춤형 소득-재산 DB의 국민연금은 노령연금을 포함하는 반면, 가계금융복지조사의 국민연금은 노령연금, 장애연금, 유족연금 등 더 많은 급여를 포함한다.

〈표 4-9〉 2019년 개인공적연금소득 분포, 전체 인구

(단위: %, %p, 만 원/년)

구분	(A) 건보	(B) 가금복 실제	(C) 가금복 재가중	(B-A)	(C-A)	(C-B)
0원	91.3	88.1	89.4	-3.1	-1.8	1.3
~연 100만 원 이하	0.1	0.2	0.1	0.0	0.0	0.0
~연 200만 원 이하	1.3	1.8	1.5	0.5	0.3	-0.3
~연 300만 원 이하	1.5	2.3	2.0	0.8	0.5	-0.3
~연 400만 원 이하	1.3	1.9	1.7	0.6	0.4	-0.2
~연 600만 원 이하	1.6	2.1	1.9	0.6	0.4	-0.2
~연 800만 원 이하	0.7	0.8	0.7	0.1	0.0	-0.1
~연 1000만 원 이하	0.4	0.5	0.5	0.1	0.0	0.0
~연 1500만 원 이하	0.6	0.8	0.7	0.2	0.1	-0.1
~연 2000만 원 이하	0.3	0.3	0.3	0.1	0.0	0.0
~연 3000만 원 이하	0.3	0.5	0.4	0.1	0.1	-0.1
연 3000만 원 초과	0.6	0.8	0.7	0.2	0.1	-0.1
계	100.0	100.0	100.0	-	-	-
평균	71	92	82	21	11	-10
표준편차	378	417	395	39	17	-22

주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-9〉의 개인공적연금소득 분포 분석에서는 공적연금 수급률이 낮은 비노인을 포함하였다. 〈표 4-10〉에서 65세 이상 노인을 대상으로 개인공적연금소득 분포를 살펴보면, 가계금융복지조사의 개인공적연금소득 존재 비율이 약 49%로, 맞춤형 소득-재산 DB의 약 41%보다 훨씬 높게 나타났다. 또한 가계금융복지조사의 개인공적연금소득 평균은 356만 원으로, 맞춤형 소득-재산 DB의 299만 원보다 컸다.

〈표 4-10〉 2019년 개인공적연금소득 분포, 65세 이상

(단위: %, %p, 만 원/년)

구분	(A) 건보	(B) 가금복 실제	(C) 가금복 재가중	(B-A)	(C-A)	(C-B)
0원	59.3	51.1	51.2	-8.2	-8.2	0.1
~연 100만 원 이하	0.1	0.2	0.2	0.1	0.1	0.0
~연 200만 원 이하	7.3	8.6	8.4	1.3	1.1	-0.3
~연 300만 원 이하	7.7	9.9	9.8	2.2	2.1	-0.1
~연 400만 원 이하	6.7	8.2	8.3	1.5	1.6	0.1
~연 600만 원 이하	6.8	8.3	8.4	1.4	1.6	0.2
~연 800만 원 이하	3.0	3.2	3.2	0.1	0.2	0.0
~연 1000만 원 이하	1.8	1.9	1.9	0.0	0.1	0.1
~연 1500만 원 이하	2.4	2.7	2.8	0.3	0.4	0.1
~연 2000만 원 이하	1.0	1.2	1.2	0.2	0.2	0.0
~연 3000만 원 이하	1.4	1.9	1.9	0.5	0.4	-0.1
연 3000만 원 초과	2.3	2.8	2.8	0.5	0.4	0.0
계	100.0	100.0	100.0	-	-	-
평균	299	356	356	56	57	1
표준편차	714	754	752	40	39	-2

주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-11〉에서는 근로소득, 사업소득, 금융소득, 공적연금소득을 합산한 개인총소득 분포를 살펴보았다. 가계금융복지조사 재가중 표본의 개인총소득 0원 비율은 20.8%로, 맞춤형 소득-재산 DB의 44.3%보다 훨씬 낮았고, 개인총소득 평균은 2,143만 원으로, 맞춤형 소득-재산 DB의 1,744만 원보다 상당히 크게 나타났다. 가계금융복지조사 재가중 표본의 개인총소득 지니계수는 0.664로 맞춤형 소득-재산 DB의 0.752보다

작게 나타났는데, 초고소득이 지니계수를 왜곡할 가능성을 고려하여 상위 1%에서 탑코딩하더라도 두 데이터의 지니계수 차이는 그다지 크게 변화하지 않았다.

〈표 4-11〉 2019년 개인총소득 분포, 전체 인구

(단위: %, %p, 만 원/년, 0~1)

구분	(A) 건보	(B) 가금복 실제	(C) 가금복 재가중	(B-A)	(C-A)	(C-B)
0원	44.3	22.5	20.8	-21.8	-23.5	-1.7
~연 100만 원 이하	3.2	13.6	13.2	10.4	10.0	-0.4
~연 300만 원 이하	4.9	5.9	5.7	1.0	0.8	-0.2
~연 500만 원 이하	3.9	4.5	4.4	0.5	0.4	-0.1
~연 1000만 원 이하	5.9	6.8	6.8	0.9	0.9	0.0
~연 2000만 원 이하	8.5	11.4	11.7	2.9	3.2	0.3
~연 3000만 원 이하	9.4	10.7	11.1	1.3	1.8	0.5
~연 4000만 원 이하	6.0	7.3	7.8	1.3	1.8	0.5
~연 6000만 원 이하	6.6	8.5	9.1	1.9	2.5	0.6
~연 8000만 원 이하	3.3	4.2	4.5	0.9	1.2	0.3
~연 10000만 원 이하	1.8	2.5	2.6	0.6	0.8	0.1
연 10000만 원 초과	2.2	2.3	2.4	0.2	0.3	0.1
계	100.0	100.0	100.0	-	-	-
평균	1744	2036	2143	292	399	107
표준편차	5397	3580	3632	-1817	-1766	51
지니계수	0.752	0.680	0.664	-0.072	-0.087	-0.015
지니계수 (상위 1% 탑코딩)	0.732	0.668	0.652	-0.064	-0.080	-0.015

주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다. 탑코딩은 전체 인구를 대상으로 시행하였다.

(A) 건보: 맞춤형 소득-재산 DB

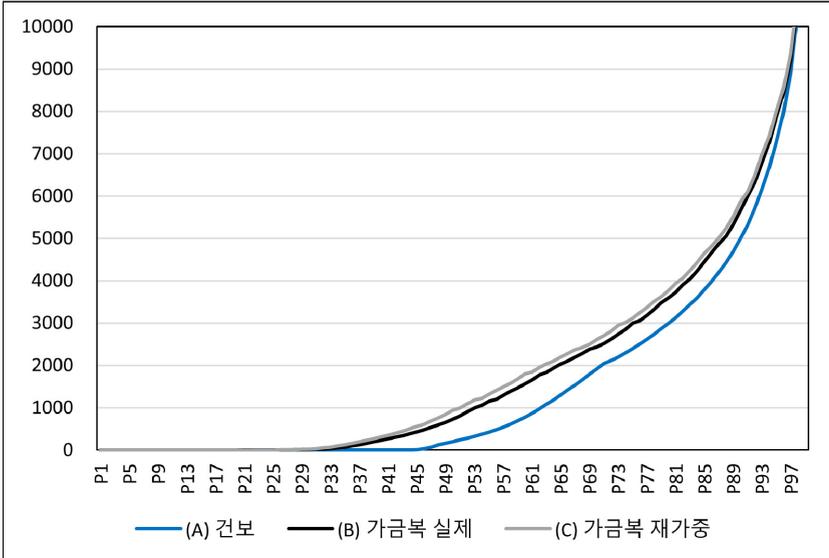
(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

[그림 4-1] 2019년 개인총소득 분위경계값, 전체 인구

(단위: 만 원/년)



주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정된 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

[그림 4-1]에는 개인총소득의 100분위 경계값을 보고하였다. 맞춤형 소득-재산 DB와 가계금융복지조사의 개인총소득 분포 차이는 대체로 중간 소득계층에서 가장 컸고, 초고소득계층으로 소득이 증가할수록 작아졌다. 맞춤형 소득-재산 DB와 가계금융복지조사의 성·연령 분포 차이가 개인 총소득 분포에 일정한 영향을 미쳤는데, 두 데이터의 성·연령 분포 차이를 통제하면 개인총소득 분포 차이가 더욱 확대되었다.

〈표 4-12〉 2019년 개인총소득 분포, 18~64세

(단위: %, %p, 만 원/년, 0~1)

구분	(A) 건보	(B) 가금복 실제	(C) 가금복 재가중	(B-A)	(C-A)	(C-B)
0원	31.1	9.2	9.2	-21.9	-21.9	0.0
~연 100만 원 이하	4.2	11.3	11.2	7.1	6.9	-0.1
~연 300만 원 이하	4.3	4.8	4.8	0.6	0.6	0.0
~연 500만 원 이하	3.8	3.7	3.8	-0.1	-0.1	0.0
~연 1000만 원 이하	6.8	6.9	6.9	0.1	0.1	0.0
~연 2000만 원 이하	10.6	14.2	14.1	3.7	3.5	-0.1
~연 3000만 원 이하	12.3	14.3	14.3	2.0	2.0	0.0
~연 4000만 원 이하	7.9	10.0	10.2	2.1	2.2	0.2
~연 6000만 원 이하	8.9	12.0	12.2	3.1	3.3	0.2
~연 8000만 원 이하	4.6	6.3	6.3	1.7	1.7	0.0
~연 10000만 원 이하	2.5	3.7	3.7	1.2	1.1	0.0
연 10000만 원 초과	3.0	3.5	3.4	0.5	0.4	-0.1
계	100.0	100.0	100.0	-	-	-
평균	2314	2854	2855	540	541	1
표준편차	5645	4091	4053	-1555	-1592	-37
지니계수	0.677	0.575	0.573	-0.101	-0.104	-0.003
지니계수 (상위 1% 탑코딩)	0.653	0.559	0.557	-0.094	-0.096	-0.002

주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다. 탑코딩은 전체 인구를 대상으로 시행하였다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정된 가계금융복지조사

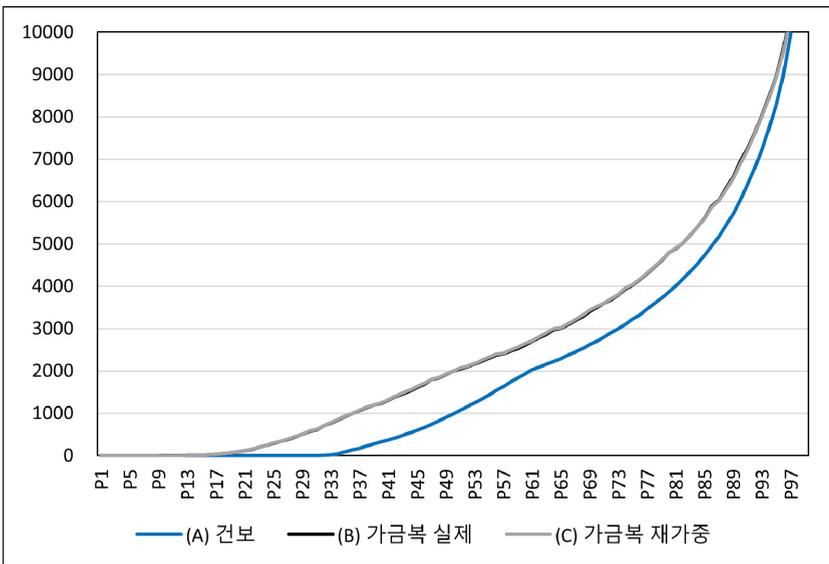
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-12〉와 [그림 4-2]에서는 18~64세 근로연령인구의 개인총소득 분포를 살펴보았다. 근로연령인구로 표본을 한정하면, 맞춤형 소득-재산 DB와 가계금융복지조사의 성·연령 분포 차이가 개인총소득 분포에 미치는 영향은 대체로 사라졌다. 가계금융복지조사의 개인총소득 0원 비율은

9.2%로 맞춤형 소득-재산 DB의 31.1%보다 21.9%p 낮았고, 가계금융복지조사에서 개인총소득이 0원 초과 100만 원 이하인 비율과 개인총소득이 1,000만 원 초과 1억 원 이하인 비율이 상대적으로 높게 나타났다. 가계금융복지조사의 개인총소득 평균은 2,854~2,855만 원으로 맞춤형 소득-재산 DB의 2,314만 원보다 컸고, 가계금융복지조사의 개인총소득 지니계수는 0.573~0.575로 맞춤형 소득-재산 DB의 0.677보다 작았다.

[그림 4-2] 2019년 개인총소득 분위경계값, 18~64세

(단위: 만 원/년)



주: 가구 내에서 합산·균등화하지 않은 개인소득 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정된 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

2. 가구소득 분포 분석

〈표 4-13〉 2019년 가구규모 분포, 가구 단위 분석

(단위: %, 가구)

구분	(A1) 세대보완 1인	(A2) 세대보완 2인	(A3) 세대보완 3인	(A4) 세대 기준안	(A5) 건보증 기준안	(B) 가금복 실제	
비율	1인	37.3	39.1	40.4	38.5	54.9	21.9
	2인	18.0	21.5	21.7	22.6	17.8	28.2
	3인	16.6	17.1	17.1	17.6	13.6	21.1
	4인	15.9	16.7	15.6	15.9	8.4	22.3
	5인	7.2	4.4	4.0	4.2	3.5	5.5
	6인 이상	5.0	1.3	1.2	1.2	1.8	1.0
	계	100.0	100.0	100.0	100.0	100.0	100.0
사례 수	1인	7,704,601	8,919,854	9,425,400	8,835,184	14,873,455	4,490
	2인	3,714,864	4,906,203	5,066,551	5,177,161	4,810,076	5,671
	3인	3,425,221	3,908,954	3,995,394	4,030,486	3,677,686	3,508
	4인	3,287,257	3,811,638	3,643,753	3,653,097	2,281,476	3,432
	5인	1,486,421	996,063	944,773	961,560	959,027	798
	6인 이상	1,023,750	293,396	269,235	283,365	491,641	165
	계	20,642,114	22,836,108	23,345,106	22,940,853	27,093,361	18,064

주: 가구 단위 분석이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복: 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

다음으로, 가구 내에서 합산하고 가구원 수 제공근으로 나누어 균등화한 가구소득을 개인 단위로 분석한 결과를 살펴본다. 두 데이터의 가구소득 분포 차이에는 앞에서 살펴본 개인소득 분포 차이와 이하에서 살펴볼 가구 개념 차이가 영향을 미쳤을 것이다. 이 연구의 맞춤형 소득-재산 DB에서는 세대보완 1인(건강보험 피부양자 포함), 세대보완 2인(비동거 가족 포함),

세대보완 3안(분리세대 추정), 세대기준안, 건강보험증기준안의 5개 가구 구성안을 구축하였다.

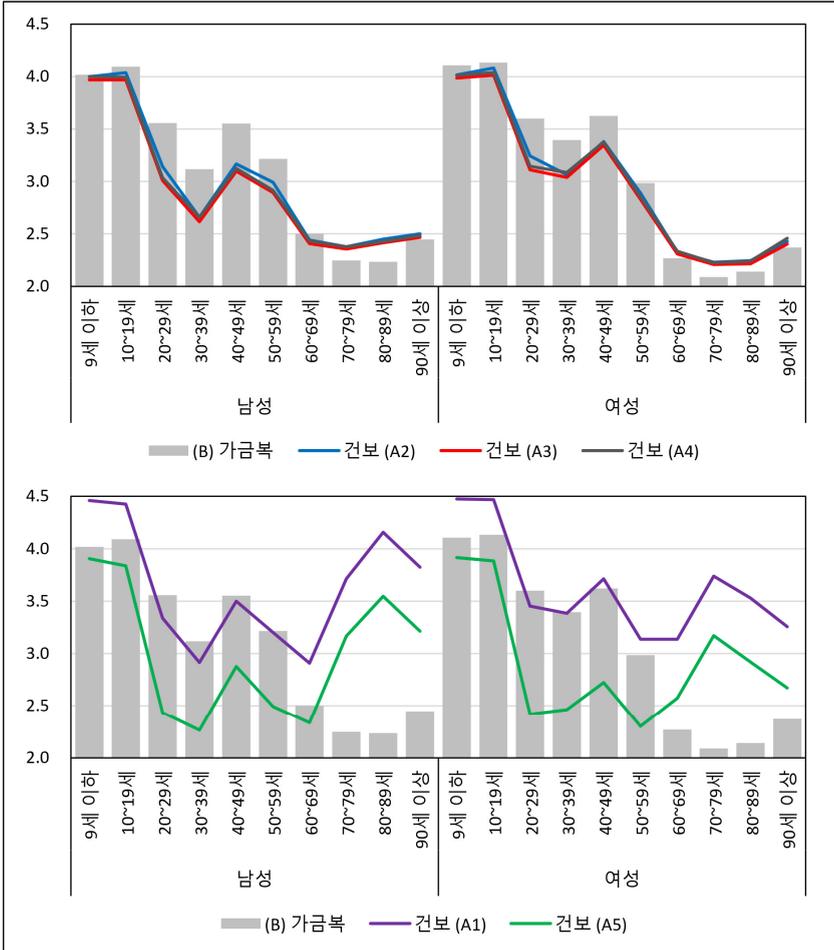
〈표 4-13〉에는 가구 단위로 가구규모 분포를 분석한 결과를 보고하였다. (A1)~(A5)의 맞춤형 소득-재산 DB 개인 수는 모두 5,256만 명이었지만, 가구 구성안이 다르기 때문에 가구수는 2,064만~2,709만 개로 상이하였다. 취업 배우자를 건강보험증 단위로 분리한 (A5)의 1인 가구 비율이 54.9%로 가장 높았고, 건강보험 피부양자를 가구로 통합한 (A1)의 1인 가구 비율이 37.3%로 가장 낮았다. (A2), (A3), (A4)의 가구규모 분포는 대체로 유사하였다. 한편, 가계금융복지조사의 1인 가구 비율은 21.9%로 전반적으로 맞춤형 소득-재산 DB보다 훨씬 낮았다.

[그림 4-3]에서는 성·연령별 평균 가구규모를 살펴보았다. 이때 가구 규모가 10인을 초과하는 사례는 일반적인 의미의 가구로 판단하기 어려워 제외하고 분석하였다. 그림은 맞춤형 소득-재산 DB (A2), (A3), (A4)의 성·연령별 평균 가구규모 패턴이 대체로 가계금융복지조사와 유사하다는 사실을 보여준다. 부모와 동거하는 아동기에서 부모로부터 독립하는 청년기로 이행할 때 가구규모가 감소하고, 자녀를 출산·양육하는 시기의 가구규모가 증가하며, 자녀가 독립한 후 노년기로 이행할 때 가구규모가 감소한다. 단, 맞춤형 소득-재산 DB (A2), (A3), (A4)와 비교할 때, 가계금융복지조사에서 20대·30대·40대의 가구규모가 다소 크고, 70대·80대의 가구규모가 다소 작았다는 사실은 언급해둘 필요가 있다.

한편, 맞춤형 소득-재산 DB (A1) 및 (A5)의 성·연령별 평균 가구규모 패턴은 가계금융복지조사와 상당히 다르게 나타났다. (A1) 및 (A5)는 건강보험 피부양자인 노인을 건강보험증 대표와 동일한 가구로 통합하기 때문에, 70대 이상 노인의 가구규모가 상당히 크게 나타났다. (A1)과 (A5)를 비교하면, 취업 배우자를 건강보험증 단위로 분리한 (A5)의 가구 규모가 상대적으로 작았다.

[그림 4-3] 2019년 개인 성·연령별 평균 가구규모, 10인 초과 가구 제외

(단위: 명)



주: 개인 단위 분석이다.

(A1) 건보: 맞춤형 소득-재산 DB 세대보완 1안(건강보험 피부양자 포함)

(A2) 건보: 맞춤형 소득-재산 DB 세대보완 2안(비동거 가족 포함)

(A3) 건보: 맞춤형 소득-재산 DB 세대보완 3안(분리세대 추정)

(A4) 건보: 맞춤형 소득-재산 DB 세대기준안

(A5) 건보: 맞춤형 소득-재산 DB 건강보험증기준안

(B) 가금부: 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-14〉 2019년 가구규모 분포, 개인 단위 분석

(단위: %, 명)

구분	(A1) 세대보완 1인	(A2) 세대보완 2인	(A3) 세대보완 3인	(A4) 세대 기준안	(A5) 건보중 기준안	(B) 가금복 실제	(C) 가금복 재가중
비율	1인	14.7	17.0	17.9	16.8	28.3	8.3
	2인	14.1	18.7	19.3	19.7	18.3	21.3
	3인	19.6	22.3	22.8	23.0	21.0	24.0
	4인	25.0	29.0	27.7	27.8	17.4	33.7
	5인	14.1	9.5	9.0	9.1	9.1	10.4
	6인 이상	12.5	3.6	3.3	3.5	5.9	2.4
	계	100.0	100.0	100.0	100.0	100.0	100.0
사례 수	1인	7,704,601	8,919,854	9,425,400	8,835,184	14,873,455	4,490
	2인	7,429,728	9,812,406	10,133,102	10,354,322	9,620,152	11,342
	3인	10,275,663	11,726,862	11,986,182	12,091,458	11,033,058	10,524
	4인	13,149,028	15,246,552	14,575,012	14,612,388	9,125,904	13,728
	5인	7,432,105	4,980,315	4,723,865	4,807,800	4,795,135	3,990
	6인 이상	6,567,365	1,872,501	1,714,929	1,857,338	3,110,786	1,021
	계	52,558,490	52,558,490	52,558,490	52,558,490	52,558,490	45,095

주: 개인 단위 분석이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정된 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

가구 단위 가구규모 분포를 살펴본 〈표 4-13〉과 달리, 〈표 4-14〉에서는 개인 단위 가구규모 분포를 살펴보았다. 예를 들어, 〈표 4-13〉에서는 가계금융복지조사 전체 가구 중에서 1인 가구가 차지하는 비율이 21.9%인 것으로 나타났는데, 〈표 4-14〉에서는 가계금융복지조사 전체 개인 중에서 1인 가구에 속한 개인이 차지하는 비율이 8.3%로 나타났다. 가계금융복지조사 표본을 재가중하여 가구규모가 작은 노인의 비율을 맞춤형 소득-

재산 DB와 동일한 수준으로 하향조정하면, 1~2인 가구에 속한 개인의 비율이 소폭 감소하였고, 3~4인 가구에 속한 개인의 비율이 소폭 증가하였다. 맞춤형 소득-재산 DB에서는 (A5)의 1인 가구 비율이 28.3%로 가장 높았고, (A1)의 1인 가구 비율이 14.7%로 가장 낮았다. (A2), (A3), (A4)의 1인 가구 비율은 16.8~17.9%로 대체로 유사하였다. 즉, (A2), (A3), (A4)의 가구 구성이 대체로 유사하였고, (A1), (A5)와 상이하게 나타났다.

〈표 4-15〉 2019년 제곱근 균등화 가구근로소득 분포, 전체 인구

(단위: %, 만 원/년)

구분	(A1) 세대보완 1안	(A2) 세대보완 2안	(A3) 세대보완 3안	(A4) 세대 기준안	(A5) 건보증 기준안	(B) 가금복 실제	(C) 가금복 재가중
0원	23.4	29.1	30.1	29.8	29.6	17.8	16.6
~연 100만 원 이하	0.6	0.7	0.7	0.7	0.7	1.4	1.3
~연 300만 원 이하	1.9	2.0	2.1	2.1	1.9	3.2	3.0
~연 500만 원 이하	1.9	2.0	2.1	2.1	2.0	2.5	2.4
~연 1000만 원 이하	6.0	5.9	5.9	6.0	5.7	7.3	7.3
~연 2000만 원 이하	16.3	15.2	15.0	15.1	16.5	16.8	16.9
~연 3000만 원 이하	16.3	13.9	13.6	13.6	16.9	15.5	15.8
~연 4000만 원 이하	12.8	10.9	10.6	10.6	11.5	12.6	12.9
~연 6000만 원 이하	13.3	12.2	11.9	11.9	10.3	13.9	14.3
~연 8000만 원 이하	4.6	4.7	4.7	4.7	3.0	5.8	6.0
~연 10000만 원 이하	1.6	1.7	1.8	1.8	1.0	2.1	2.2
연 10000만 원 초과	1.4	1.5	1.6	1.6	0.9	1.2	1.2
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
평균	2471	2323	2290	2306	2062	2572	2638
표준편차	3226	3322	3340	3349	3000	2597	2602

주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-16〉 2019년 제곱근 균등화 가구사업소득 분포, 전체 인구

(단위: %, 만 원/년)

구분	(A1) 세대보완 1안	(A2) 세대보완 2안	(A3) 세대보완 3안	(A4) 세대 기준안	(A5) 건보중 기준안	(B) 가금복 실제	(C) 가금복 재가중
0원	51.7	54.7	55.3	54.9	60.5	52.1	51.7
~연 100만 원 이하	14.7	13.2	13.0	13.1	12.6	11.5	11.8
~연 300만 원 이하	9.6	9.0	8.8	8.9	7.7	5.0	5.0
~연 500만 원 이하	5.5	5.3	5.2	5.3	4.3	2.9	2.8
~연 1000만 원 이하	7.5	7.3	7.2	7.2	5.9	5.2	5.2
~연 2000만 원 이하	5.6	5.5	5.4	5.5	4.6	9.4	9.5
~연 3000만 원 이하	2.1	2.0	2.0	2.0	1.8	6.7	6.8
~연 4000만 원 이하	1.0	1.0	1.0	1.0	0.9	2.9	3.0
~연 6000만 원 이하	0.9	0.9	0.9	0.9	0.8	2.8	2.8
~연 8000만 원 이하	0.4	0.4	0.4	0.4	0.4	0.8	0.8
~연 10000만 원 이하	0.2	0.2	0.2	0.2	0.2	0.2	0.2
연 10000만 원 초과	0.6	0.6	0.6	0.6	0.5	0.4	0.4
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
평균	489	475	467	470	408	773	777
표준편차	3595	3596	3597	3598	3531	2363	2310

주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융 복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융 복지조사 원자료.

〈표 4-15〉에서는 균등화 가구근로소득 분포를 살펴보았다. 가계금융 복지조사 재가중 표본의 가구근로소득 0원 비율은 16.6%로 맞춤형 소득-재산 DB보다 훨씬 낮았다. 가계금융복지조사에서 가구근로소득 존재 비율이 훨씬 높게 나타난 것은 〈표 4-4〉와 같이 개인근로소득 존재 비율이 훨씬 높게 포착되었기 때문이기도 하고, 〈표 4-14〉와 같이 1인 가구 비율이 낮아 가구 내 소득 공유가 강하게 반영되었기 때문이기도 하다. 맞춤형

소득-재산 DB에서는 취업 배우자와 건강보험 피부양자를 가구로 통합한 (A1)의 가구근로소득 존재 비율이 가장 높았고 가계금융복지조사에 가장 가까운 가구근로소득 분포를 나타내었다.

〈표 4-16〉에서는 균등화 가구사업소득 분포를 분석하였다. 표를 살펴 보면, 전반적으로 맞춤형 소득-재산 DB보다 가계금융복지조사에서 가구 사업소득이 1,000만 원 이하인 비율이 낮았고, 1,000만 원 초과 8,000만 원 이하인 비율이 높게 나타났다. 이처럼 가계금융복지조사에서 가구 사업소득 수준이 상대적으로 높게 포착된 것은 〈표 4-6〉에서 살펴본 것처럼, 가계금융복지조사에서 개인사업소득 수준이 상대적으로 높게 포착되었기 때문이기도 하고, 가계금융복지조사에 가구 내 소득 공유가 상대적으로 강하게 반영되었기 때문이기도 하다. 맞춤형 소득-재산 DB에서는 (A1)의 가구사업소득 수준이 가장 높게 나타났다.

〈표 4-17〉에서는 균등화 가구금융소득 분포를 살펴보았다. 앞에서 〈표 4-8〉에서 살펴본 것처럼, 맞춤형 소득-재산 DB에서 개인금융소득이 거의 포착되지 않았기 때문에, 가구 내에서 금융소득을 합산하더라도 가구금융소득이 존재하는 비율이 약 2%로 낮게 나타났다. 반면 가계금융복지조사에서는 개인금융소득이 존재하는 비율이 약 60%였고(〈표 4-8〉 참조), 가구 내에서 합산한 가구금융소득이 존재하는 비율이 약 89%로 높게 나타났다(〈표 4-17〉 참조). 앞에서 언급한 바와 같이, 맞춤형 소득-재산 DB와 가계금융복지조사의 금융소득 차이를 발생시킨 주된 원인은 금융소득 행정자료 범위의 차이인 것으로 판단된다.

〈표 4-18〉에서는 균등화 가구공적연금소득 분포를 살펴보았다. 가계금융복지조사 재가중 표본의 가구공적연금소득 존재 비율은 22.1%로, 맞춤형 소득-재산 DB (A2), (A3), (A4)의 18.3~18.5%와 (A5)의 21.5%보다 높았지만 (A1)의 25.5%보다는 낮았다. (A1)과 (A5)는 공적연금

수급률이 높은 노인을 자녀 가구에 강하게 통합하기 때문에 가구공적 연금소득 존재 비율이 상대적으로 높았고, 균등화 가구공적연금소득의 평균 역시 비교적 크게 나타났다. 특히 (A1)의 균등화 가구공적연금소득 평균은 124만 원으로 가계금융복지조사 재가중 표본의 130만 원에 근접한 수준이었다.

〈표 4-17〉 2019년 제급금 균등화 가구금융소득 분포, 전체 인구

(단위: %, 만 원/년)

구분	(A1) 세대보완 1안	(A2) 세대보완 2안	(A3) 세대보완 3안	(A4) 세대 기준안	(A5) 건보증 기준안	(B) 가금복 실제	(C) 가금복 재가중
0원	97.9	98.2	98.2	98.2	98.3	11.1	11.1
~연 10만 원 이하	0.0	0.0	0.0	0.0	0.0	31.7	32.1
~연 50만 원 이하	0.0	0.0	0.0	0.0	0.0	23.6	23.7
~연 100만 원 이하	0.0	0.0	0.0	0.0	0.0	10.7	10.6
~연 200만 원 이하	0.0	0.0	0.0	0.0	0.0	9.7	9.6
~연 300만 원 이하	0.0	0.0	0.0	0.0	0.0	4.6	4.6
~연 400만 원 이하	0.0	0.0	0.0	0.0	0.0	2.3	2.3
~연 600만 원 이하	0.4	0.2	0.2	0.2	0.3	2.4	2.3
~연 800만 원 이하	0.4	0.3	0.3	0.3	0.3	1.3	1.3
~연 1000만 원 이하	0.3	0.2	0.2	0.2	0.2	0.7	0.7
~연 2000만 원 이하	0.5	0.5	0.5	0.5	0.5	1.2	1.2
연 2000만 원 초과	0.5	0.5	0.4	0.5	0.4	0.6	0.6
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
평균	60	58	57	57	48	127	126
표준편차	1692	1691	1688	1689	1510	790	797

주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정된 가계금융 복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-18〉 2019년 제곱근 균등화 가구공적연금소득 분포, 전체 인구

(단위: %, 만 원/년)

구분	(A1) 세대보완 1안	(A2) 세대보완 2안	(A3) 세대보완 3안	(A4) 세대 기준안	(A5) 건보증 기준안	(B) 가금복 실제	(C) 가금복 재가중
0원	74.5	81.5	81.7	81.5	78.5	76.6	77.9
~연 100만 원 이하	3.8	1.4	1.4	1.4	2.8	1.5	1.4
~연 200만 원 이하	6.4	4.0	4.0	4.1	5.3	5.4	5.1
~연 300만 원 이하	4.3	3.3	3.3	3.4	3.7	4.5	4.2
~연 400만 원 이하	2.6	2.3	2.2	2.3	2.3	2.9	2.8
~연 600만 원 이하	2.8	2.5	2.5	2.5	2.4	3.0	2.8
~연 800만 원 이하	1.5	1.3	1.3	1.3	1.3	1.8	1.8
~연 1000만 원 이하	0.9	0.8	0.8	0.8	0.8	0.8	0.8
~연 1500만 원 이하	1.4	0.9	0.9	0.9	1.2	1.0	0.9
~연 2000만 원 이하	0.7	0.7	0.7	0.7	0.6	1.0	0.9
~연 3000만 원 이하	0.7	0.9	0.9	0.9	0.7	1.0	1.0
연 3000만 원 초과	0.3	0.3	0.3	0.3	0.4	0.5	0.4
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
평균	124	109	109	109	111	138	130
표준편차	391	393	393	393	390	443	428

주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-19〉에서는 균등화 가구총소득 분포를 살펴보았다. 앞의 〈표 4-11〉에서 살펴본 것처럼, 가계금융복지조사 재가중 표본에서 개인총소득이 존재하지 않는 비율은 20.8%였는데, 〈표 4-19〉에서 가구 내에서 합산한 가구총소득이 존재하지 않는 비율은 1.2%로 감소하였다. 이는 개인총소득이 존재하지 않는 아동, 노인, 비경활 배우자의 대다수가 개인총소득이 존재하는 가구원과 동거한다는 사실을 보여준다.

〈표 4-19〉 2019년 제공근 균등화 가구총소득 분포, 전체 인구

(단위: %, 만 원/년, 0~1)

구분	(A1) 세대보완 1안	(A2) 세대보완 2안	(A3) 세대보완 3안	(A4) 세대 기준안	(A5) 건보증 기준안	(B) 가금복 실제	(C) 가금복 재가중
0원	8.5	10.8	11.5	11.3	10.5	1.3	1.2
~연 100만 원 이하	2.0	2.2	2.3	2.3	2.3	1.8	1.7
~연 300만 원 이하	3.5	5.1	5.2	5.2	4.2	2.4	2.1
~연 500만 원 이하	3.0	3.8	3.9	3.9	3.7	2.3	2.1
~연 1000만 원 이하	7.0	7.7	7.8	7.8	8.5	5.7	5.4
~연 2000만 원 이하	16.4	16.2	16.2	16.2	18.9	15.3	15.0
~연 3000만 원 이하	17.9	16.2	15.8	15.8	19.1	19.5	19.6
~연 4000만 원 이하	14.8	12.8	12.5	12.5	13.4	17.7	18.0
~연 6000만 원 이하	16.2	14.5	14.1	14.2	12.5	20.0	20.5
~연 8000만 원 이하	5.9	5.8	5.7	5.8	3.9	8.3	8.6
~연 10000만 원 이하	2.2	2.3	2.3	2.3	1.3	3.4	3.5
연 10000만 원 초과	2.5	2.6	2.6	2.7	1.7	2.4	2.5
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
평균	3143	2965	2923	2942	2630	3610	3671
표준편차	5511	5571	5585	5593	5140	3348	3308
지니계수	0.486	0.528	0.537	0.535	0.506	0.396	0.388
지니계수 (상위 1% 탑코딩)	0.460	0.502	0.511	0.510	0.479	0.381	0.374

주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다. 탑코딩은 전체 인구를 대상으로 시행하였다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정된 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

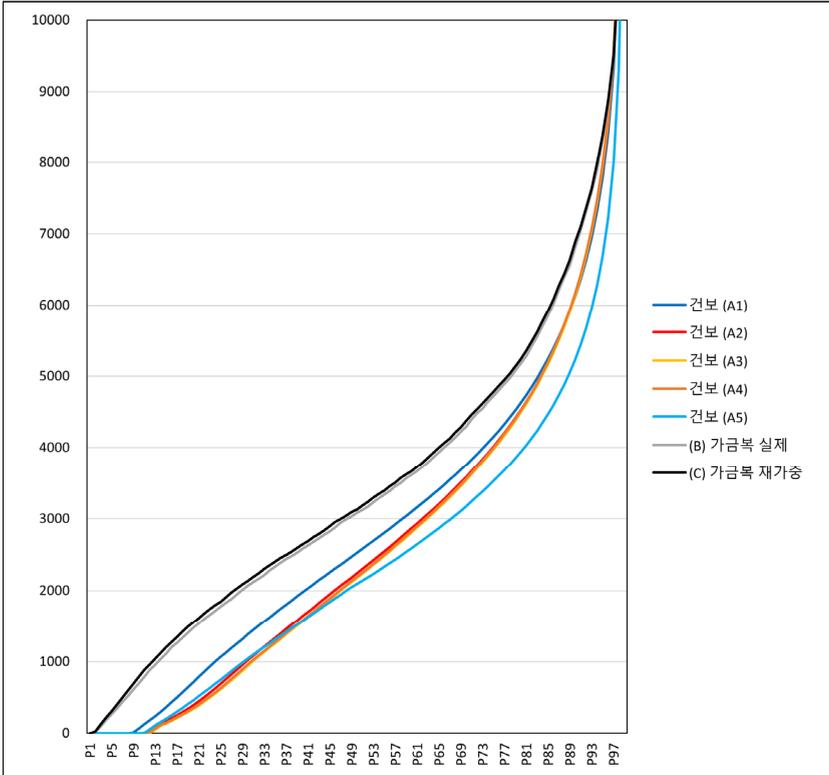
한편 맞춤형 소득-재산 DB에서는 개인총소득이 존재하지 않는 비율이 44.3%였는데(〈표 4-11〉 참조), 다양한 형태로 가구를 구성하면 가구총소득이 존재하지 않는 비율이 8.5~11.5%로 감소하였다. 특히 취업 배우자와 건강보험 피부양자를 가구로 통합한 (A1)에 가구 내 소득 공유가 가장

강하게 반영되었다. 전반적인 분석 결과는 맞춤형 소득-재산 DB에 비해 가계금융복지조사의 균등화 가구총소득 수준이 상당히 높고 평등하게 추정된다는 사실을 보여준다. 가계금융복지조사 재가중 표본의 균등화 가구총소득 평균은 3,671만 원으로, 맞춤형 소득-재산 DB의 2,630~3,143만 원보다 훨씬 컸다. 또한 가계금융복지조사 재가중 표본의 균등화 가구총소득 지니계수는 0.388로, 맞춤형 소득-재산 DB의 0.486~0.537보다 훨씬 작았다.

[그림 4-4]의 분위경계값 분석 결과는 가계금융복지조사의 가구총소득 수준이 상대적으로 높게 포착된다는 사실을 분명하게 보여준다. 맞춤형 소득-재산 DB에서 (A2), (A3), (A4)의 차이는 크지 않았고, (A1), (A2)·(A3)·(A4), (A5)의 차이가 두드러졌다. (A2)·(A3)·(A4)와 비교할 때 건강보험 피부양자인 노인을 자녀 가구로 통합한 (A1)에서 중저소득분위 경계값이 높게 추정되었는데, 이는 소득 수준이 낮은 노인을 자녀 가구에 통합하면 중저소득계층의 가구소득 수준이 향상된다는 사실을 보여준다. 이러한 분석 결과는 맞춤형 소득-재산 DB의 5개 가구 구성안 중에서 (A1)의 균등화 가구총소득 불평등 수준이 낮고, 빈곤, 특히 노인빈곤 수준이 양호하게 나타날 가능성을 시사한다. 또한 (A2)·(A3)·(A4)와 비교할 때 건강보험증 단위로 배우자를 분리한 (A5)에서는 고소득분위 경계값이 낮게 나타났는데, 이는 부부 단위 소득 공유가 고소득계층의 가구소득 수준을 향상시킨다는 사실을 보여준다. (A2)·(A3)·(A4)와 비교하면 (A5)의 고소득층 소득 수준이 낮은 편이므로, (A5)의 균등화 가구총소득 분배 수준이 상대적으로 양호하게 나타날 것으로 예상된다. 한편, 맞춤형 소득-재산 DB와 가계금융복지조사의 성·연령 분포 차이가 균등화 가구총소득 분포에 미친 영향은 그다지 크지 않았다.

[그림 4-4] 2019년 제급근 균등화 가구총소득 분위경계값, 전체 인구

(단위: 만 원/년)



주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다.

(A1) 건보: 맞춤형 소득-재산 DB 세대보완 1안(건강보험 피부양자 포함)

(A2) 건보: 맞춤형 소득-재산 DB 세대보완 2안(비동거 가족 포함)

(A3) 건보: 맞춤형 소득-재산 DB 세대보완 3안(분리세대 추정)

(A4) 건보: 맞춤형 소득-재산 DB 세대기준안

(A5) 건보: 맞춤형 소득-재산 DB 건강보험증기준안

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-20〉 2019년 제6급 근대화 가구총소득 분포, 18~64세

(단위: %, 만 원/년, 0~1)

구분	(A1) 세대보완 1안	(A2) 세대보완 2안	(A3) 세대보완 3안	(A4) 세대 기준안	(A5) 건보증 기준안	(B) 가금복 실제	(C) 가금복 재가중
0원	8.2	8.8	9.6	9.4	10.2	0.7	0.6
~연 100만 원 이하	2.1	2.2	2.4	2.4	2.5	0.7	0.7
~연 300만 원 이하	3.1	3.5	3.6	3.6	3.9	0.8	0.8
~연 500만 원 이하	2.9	3.3	3.4	3.3	3.8	1.0	0.9
~연 1000만 원 이하	7.0	7.4	7.5	7.4	8.8	3.7	3.6
~연 2000만 원 이하	16.7	16.5	16.3	16.4	19.3	13.9	13.7
~연 3000만 원 이하	18.1	16.9	16.5	16.5	19.3	19.9	19.9
~연 4000만 원 이하	14.7	13.6	13.2	13.2	12.8	19.3	19.4
~연 6000만 원 이하	16.4	16.0	15.5	15.6	12.2	23.0	23.1
~연 8000만 원 이하	6.1	6.5	6.4	6.5	4.0	10.1	10.1
~연 10000만 원 이하	2.2	2.6	2.6	2.6	1.4	4.1	4.1
연 10000만 원 초과	2.5	2.8	2.9	2.9	1.7	2.9	2.9
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
평균	3165	3195	3152	3173	2621	4021	4031
표준편차	5364	5514	5551	5558	4921	3218	3190
지니계수	0.482	0.501	0.511	0.510	0.507	0.352	0.350
지니계수 (상위 1% 탑코딩)	0.457	0.475	0.485	0.484	0.481	0.338	0.336

주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다. 탑코딩은 전체 인구를 대상으로 시행하였다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정된 가계금융복지조사

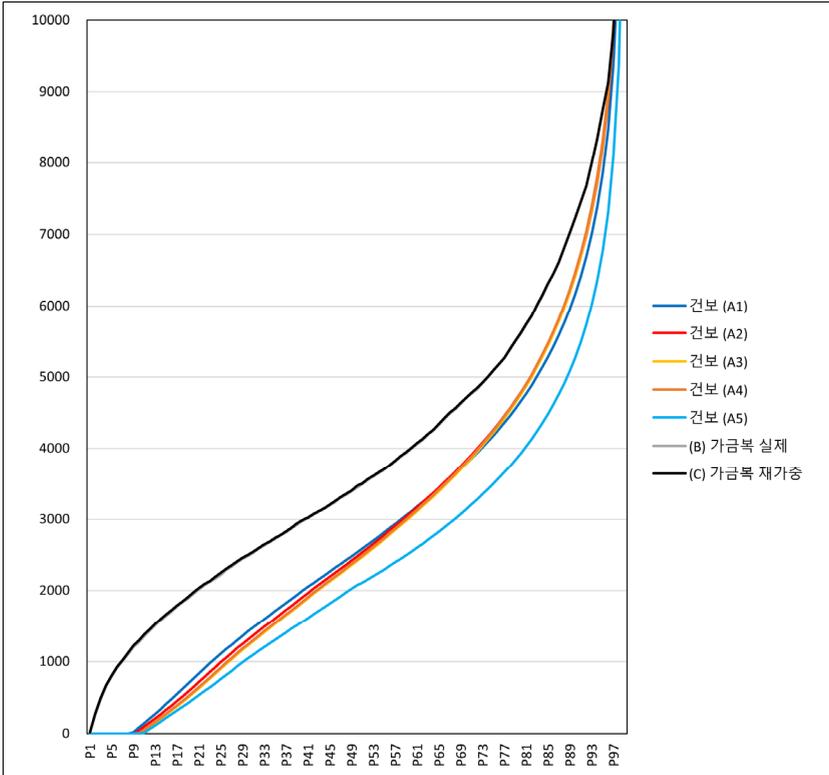
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-20〉과 [그림 4-5]에서 18~64세 근로연령인구의 근대화 가구총소득 분포를 살펴보면, 맞춤형 소득-재산 DB보다 가계금융복지조사에서 가구총소득 수준이 훨씬 높게 추정되었다. [그림 4-4]와 달리, 노인을 제외하고 분석하면 (A1)의 가구총소득 분포가 (A2)·(A3)·(A4)에 상당히

근접하였고, (A2)·(A3)·(A4)와 (A5)의 차이가 두드러졌다.

[그림 4-5] 2019년 제급근 균등화 가구총소득 분위경계값, 18~64세

(단위: 만 원/년)



주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다.

(A1) 건보: 맞춤형 소득-재산 DB 세대보완 1안(건강보험 피부양자 포함)

(A2) 건보: 맞춤형 소득-재산 DB 세대보완 2안(비동거 가족 포함)

(A3) 건보: 맞춤형 소득-재산 DB 세대보완 3안(분리세대 추정)

(A4) 건보: 맞춤형 소득-재산 DB 세대기준안

(A5) 건보: 맞춤형 소득-재산 DB 건강보험증기준안

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-21〉 2019년 제6급 균등화 가구총소득 분포, 65세 이상

(단위: %, 만 원/년, 0~1)

구분	(A1) 세대보완 1안	(A2) 세대보완 2안	(A3) 세대보완 3안	(A4) 세대 기준안	(A5) 건보증 기준안	(B) 가금복 실제	(C) 가금복 재가중
0원	12.4	24.6	24.6	24.2	13.7	4.3	4.3
~연 100만 원 이하	1.2	1.9	2.0	2.0	1.3	7.3	7.2
~연 300만 원 이하	5.4	14.3	14.3	14.3	6.0	9.6	9.3
~연 500만 원 이하	3.6	7.5	7.5	7.5	4.1	8.9	8.7
~연 1000만 원 이하	6.2	9.4	9.5	9.5	7.1	14.8	14.6
~연 2000만 원 이하	13.8	14.4	14.3	14.5	15.9	20.1	20.2
~연 3000만 원 이하	16.2	11.0	10.9	11.0	17.3	14.3	14.5
~연 4000만 원 이하	14.3	6.7	6.6	6.6	14.0	9.1	9.3
~연 6000만 원 이하	16.3	6.0	6.2	6.2	13.6	8.1	8.2
~연 8000만 원 이하	6.0	2.0	2.1	2.1	3.9	2.0	2.0
~연 10000만 원 이하	2.1	0.8	0.8	0.8	1.2	0.8	0.8
연 10000만 원 초과	2.6	1.3	1.3	1.3	1.8	0.8	0.8
계	100.0	100.0	100.0	100.0	100.0	100.0	100.0
평균	3089	1639	1649	1664	2665	1844	1862
표준편차	6796	6270	6294	6303	6853	2232	2240
지니계수	0.522	0.686	0.688	0.686	0.532	0.543	0.541
지니계수 (상위 1% 탑코딩)	0.491	0.660	0.662	0.660	0.498	0.536	0.534

주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다. 탑코딩은 전체 인구를 대상으로 시행하였다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

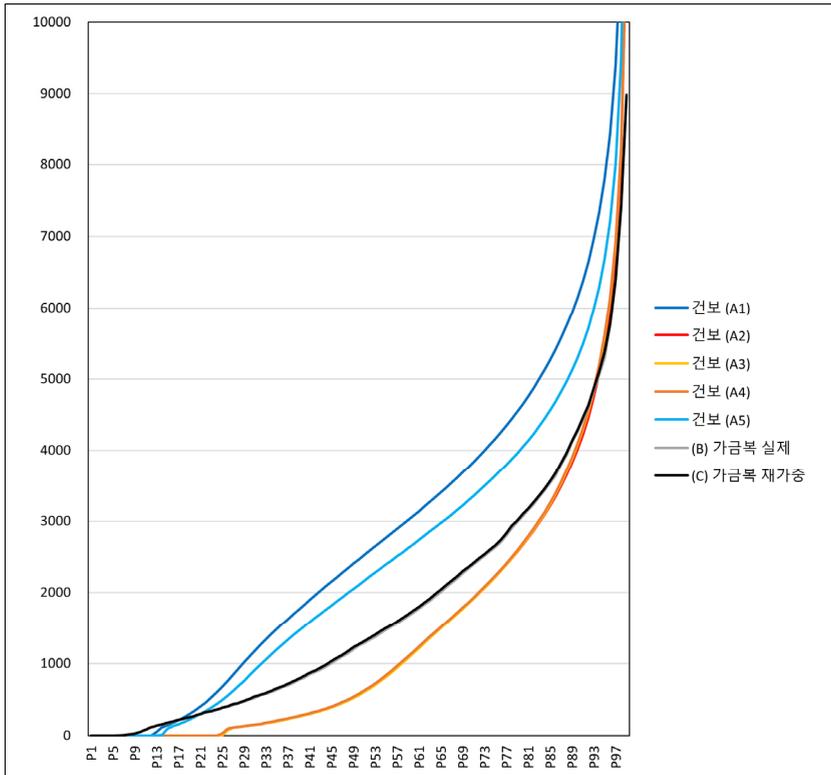
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-21〉과 [그림 4-6]에서 65세 이상 노인의 균등화 가구총소득 분포를 살펴보면, 대체로 건강보험 피부양자인 노인을 자녀 가구로 통합한 (A1)과 (A5)의 가구총소득이 가계금융복지조사보다 상당히 높게 추정되었고, 가구 내 소득 공유 수준이 낮은 (A2)·(A3)·(A4)의 가구총소득이

가계금융복지조사보다 상당히 낮게 추정되었다.

[그림 4-6] 2019년 제곱근 균등화 가구총소득 분위경계값, 65세 이상

(단위: 만 원/년)



주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다.

(A1) 건보: 맞춤형 소득-재산 DB 세대보완 1안(건강보험 피부양자 포함)

(A2) 건보: 맞춤형 소득-재산 DB 세대보완 2안(비동거 가족 포함)

(A3) 건보: 맞춤형 소득-재산 DB 세대보완 3안(분리세대 추정)

(A4) 건보: 맞춤형 소득-재산 DB 세대기준안

(A5) 건보: 맞춤형 소득-재산 DB 건강보험증기준안

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

3. 소득분배지표 분석

〈표 4-22〉 2019년 맞춤형 소득-재산 DB 소득분배지표

(단위: 만 원/년, 0~1, %)

구분		(A1) 세대보완 1안	(A2) 세대보완 2안	(A3) 세대보완 3안	(A4) 세대 기준안	(A5) 건보증 기준안
중위값	총소득1	2526	2244	2179	2193	2095
	총소득2	2535	2252	2187	2201	2100
지니계수	총소득1	0.486	0.528	0.537	0.535	0.506
	총소득2	0.486	0.527	0.536	0.535	0.506
지니계수 (상위 1% 탑코딩)	총소득1	0.460	0.502	0.511	0.510	0.479
	총소득2	0.460	0.502	0.511	0.510	0.478
전체 인구 빈곤율	총소득1	28.1	31.5	32.3	32.1	30.0
	총소득2	28.0	31.5	32.2	32.0	29.9
0~17세 빈곤율	총소득1	26.9	24.4	25.6	25.6	27.0
	총소득2	26.8	24.3	25.6	25.6	26.9
18~64세 빈곤율	총소득1	27.4	27.0	27.9	27.7	30.1
	총소득2	27.4	26.9	27.8	27.6	29.9
65세 이상 빈곤율	총소득1	32.1	59.5	59.2	58.9	32.9
	총소득2	32.0	59.5	59.1	58.9	32.8

주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다. 빈곤선은 균등화 총소득1/총소득2 개인 단위 중위값의 50%이다.

· 총소득1: 근로소득+사업소득+이자소득+배당소득+분리과세금융소득+국민연금소득+직역연금소득

· 총소득2: 근로소득+사업소득+이자소득+배당소득+분리과세금융소득+국민연금소득+직역연금소득+분리과세주택임대소득+기타소득

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

지금까지 분석에서는 가계금융복지조사와 비교하기 위해 근로소득, 사업소득, 금융소득, 공적연금소득을 합산하여 총소득을 정의하였지만, 〈표 4-22〉에서는 분리과세주택임대소득 및 기타소득을 추가로 합산한 총소득을 함께 분석하였다. 표를 살펴보면, 분리과세주택임대소득 및 기타소득의 합산은 중위값을 5~9만 원 증가시켰을 뿐, 총소득 분배지표에

그다지 큰 영향을 미치지 못한 것으로 나타났다. 가구 개념에 따른 분배 지표 차이를 살펴보면, 대체로 소득 수준이 낮은 노인을 자녀 가구와 강하게 통합한 (A1)과 (A5)의 불평등 및 빈곤 수준이 낮게 나타났다. 특히 65세 이상 노인 빈곤율은 (A1) 및 (A5)에서 32.0~32.9%로 상당히 낮았지만 (A2), (A3), (A4)에서 58.9~59.5%로 매우 높았다.

〈표 4-23〉 2019년 가계금융복지조사 소득분배지표

(단위: 만 원/년, 0~1, %)

구분		(B) 가금복 실제	(C) 가금복 재가중
중위값	근로·사업·금융·공적연금소득	3083	3132
	가처분소득	2875	2920
지니계수	근로·사업·금융·공적연금소득	0.396	0.388
	가처분소득	0.340	0.336
지니계수 (상위 1% 탑코딩)	근로·사업·금융·공적연금소득	0.381	0.374
	가처분소득	0.328	0.324
전체 인구 빈곤율	근로·사업·금융·공적연금소득	21.1	20.2
	가처분소득	16.3	15.8
0~17세 빈곤율	근로·사업·금융·공적연금소득	13.4	13.7
	가처분소득	10.6	10.8
18~64세 빈곤율	근로·사업·금융·공적연금소득	13.5	13.5
	가처분소득	11.0	11.1
65세 이상 빈곤율	근로·사업·금융·공적연금소득	56.6	56.5
	가처분소득	41.6	42.0

주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다. 빈곤선은 균등화 근로·사업·금융·공적연금소득/가처분소득 개인 단위 중위값의 50%이다.

· 가처분소득: 근로소득+사업소득+재산소득+공적이전소득+사적이전소득-공적이전지출-사적이전지출

자료: 통계청, (2020e). 2020년 가계금융복지조사 원자료.

한편, 지금까지는 가계금융복지조사 역시 맞춤형 소득-재산 DB와 비교하기 위해 근로소득, 사업소득, 금융소득, 공적연금소득을 합산하여 총

소득을 정의하였지만, <표 4-23>에서는 근로소득, 사업소득, 재산소득, 공적이전소득·지출, 사적이전소득·지출을 모두 반영한 가처분소득의 분배지표를 함께 살펴보았다. 예를 들어, 가계금융복지조사 근로·사업·금융·공적연금소득 지니계수는 0.396이었는데, 모든 소득 및 비소비지출 항목을 추가적으로 반영한 가처분소득의 지니계수는 0.340으로 감소하였다. 또한 근로·사업·금융·공적연금소득 빈곤율과 가처분소득 빈곤율은 각각 21.1%, 16.3%였다. 맞춤형 소득-재산 DB의 총소득 분배지표와 가계금융복지조사의 가처분소득 분배지표를 비교할 때는 소득 개념·범위 차이가 분배지표에 상당히 큰 영향을 미친다는 사실을 고려해야 한다.

<표 4-24> 2019년 총소득 분배지표

(단위: 만 원/년, 배, 0~1, %)

구분	(A1) 세대보완 1안	(A2) 세대보완 2안	(A3) 세대보완 3안	(A4) 세대 기준안	(A5) 건보증 기준안	(B) 가금복 실제	(C) 가금복 재가중	
중위값	2526	2244	2179	2193	2095	3083	3132	
P90/P10	90.69	-	-	-	-	9.74	8.75	
P90/P50	2.43	2.74	2.82	2.81	2.50	2.21	2.20	
P50/P10	37.40	-	-	-	-	4.41	3.98	
지니계수	0.486	0.528	0.537	0.535	0.506	0.396	0.388	
지니계수 (상위 1% 탑코딩)	0.460	0.502	0.511	0.510	0.479	0.381	0.374	
빈 곤 율	전체 인구	28.1	31.5	32.3	32.1	30.0	21.1	20.2
	0~17세	26.9	24.4	25.6	25.6	27.0	13.4	13.7
	18~64세	27.4	27.0	27.9	27.7	30.1	13.5	13.5
	65세 이상	32.1	59.5	59.2	58.9	32.9	56.6	56.5

주: 가구 내에서 합산·균등화한 가구소득의 개인 단위 분포이다. 총소득은 <근로소득+사업소득+금융소득+공적연금소득>을 의미한다. 빈곤선은 균등화 총소득 개인 단위 중위값의 50%이다.

(A) 건보: 맞춤형 소득-재산 DB

(B) 가금복 실제: 가계금융복지조사

(C) 가금복 재가중: 개인 성·연령 분포를 맞춤형 소득-재산 DB와 동일하게 조정한 가계금융복지조사

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료; 통계청. (2020e). 2020년 가계금융복지조사 원자료.

〈표 4-24〉에는 맞춤형 소득-재산 DB와 가계금융복지조사의 총소득 분배지표를 함께 보고하였다. 가구를 어떠한 형태로 구성하더라도, 맞춤형 소득-재산 DB의 지니계수는 0.486~0.537로 가계금융복지조사 재가중 표본의 0.388보다 훨씬 높았다. 이는 맞춤형 소득-재산 DB보다 가계금융복지조사의 개인총소득 분포가 평등하고(〈표 4-11〉 참조) 1인 가구 비율이 낮아, 가구 내 소득 공유가 강하게 반영되었기 때문일 것으로 판단된다(〈표 4-14〉 참조). 맞춤형 소득-재산 DB의 빈곤율은 28.1~32.3%로, 가계금융복지조사 재가중 표본의 20.2%보다 훨씬 높게 나타났다. 단, 건강보험 피부양자인 노인을 자녀 가구에 통합한 (A1)과 (A5)의 노인 빈곤율은 32.1~32.9%로, 가계금융복지조사 재가중 표본의 56.5%보다 훨씬 낮았다.

제4절 소결

이 장에서는 가계금융복지조사와 비교하여 맞춤형 소득-재산 DB의 소득 분포를 분석하였다. 우선, 가구 내에서 합산·균등화하지 않은 개인소득 분포를 비교한 결과, 대체로 맞춤형 소득-재산 DB보다 가계금융복지조사의 소득 포착률이 높은 것으로 나타났다. 맞춤형 소득-재산 DB에 비해 서베이와 행정자료를 결합하여 구축한 가계금융복지조사에서 개인근로 소득이 존재하는 비율이 높았고, 특히 개인근로소득이 0원 초과 2,000만원 이하인 비율이 상당히 높게 나타났다. 이는 가계금융복지조사가 조세 행정으로 포착하지 못한 저임금근로자의 근로소득을 상대적으로 정확하게 포착하였을 것이라는 예측을 지지한다. 한편, 가계금융복지조사와 맞춤형 소득-재산 DB에서 개인근로소득이 1억 원을 넘는 초고소득층의 비율이

1.6%로 거의 동일하게 나타났다는 분석 결과에 주목할 필요가 있다. 일반적으로 소득조사 서베이에서 고소득층의 무응답이 상대적으로 강하게 나타날 가능성이 있지만, 이를 실증적으로 평가하기는 어렵다. 이 연구의 분석 결과는 가계금융복지조사에서 고소득층 누락 현상이 그다지 강하게 나타나지 않을 가능성을 시사한다. 또한 고소득 자영업자의 소득이 과소 보고되었을 가능성이 큰 맞춤형 소득-재산 DB에 비해, 가계금융복지조사에서 개인사업소득 수준이 상대적으로 높게 나타났고, 개인공적연금 역시 국민연금에 노령연금을 포함한 맞춤형 소득-재산 DB보다 국민연금에 노령연금, 장애연금, 유족연금 등을 포함한 가계금융복지조사에서 상대적으로 크게 포착되었다. 한편, 가계금융복지조사와 달리 비과세 원천분 금융소득 정보를 반영하지 못한 맞춤형 소득-재산 DB에서 개인금융소득이 존재하는 비율이 0.7%에 불과한 것으로 나타났다. 단, 이 장에서 맞춤형 소득-재산 DB와 가계금융복지조사의 개인소득 분포를 비교한 분석 결과는 탐색적인 수준에 그친다는 점에 주의해야 할 필요가 있다. 향후 맞춤형 소득-재산 DB가 입수한 소득 행정자료의 범위 및 정의를 꼼꼼하게 검토하여 가계금융복지조사와의 차이를 정교하게 이해하기 위한 후속 연구가 필요하다.

다음으로는 세대보완 1안(건강보험 피부양자 포함), 세대보완 2안(비동거 가족 포함), 세대보완 3안(분리세대 추정), 세대기준안, 건강보험증 기준안의 5개 가구 구성안을 구축하여 가구 내에서 합산·균등화한 가구 소득 분포를 살펴보았다. 대체로 주민등록 세대를 부분적으로 보완한 세대보완 2안(비동거 가족 포함), 세대보완 3안(분리세대 추정)과 세대기준안의 가구소득 분포가 상당히 유사하였다. 세대보완 2안(비동거 가족 포함)·세대보완 3안(분리세대 추정)·세대기준안의 가구소득 분포와 건강보험증 기준안의 가구소득 분포, 세대보완 1안(건강보험 피부양자 포함)의 가구

소득 분포에 상당한 차이가 관찰되었다. 세대보완 2인(비동거 가족 포함)·세대보완 3인(분리세대 추정)·세대기준안의 가구소득 분포와 비교할 때, 건강보험 피부양자인 노인을 건강보험증 대표인 자녀와 동일한 가구로 통합한 세대보완 1인(건강보험 피부양자 포함)의 가구소득 분포에서 중저소득계층의 가구소득 수준이 상대적으로 높게 나타났다. 건강보험 피부양자인 노인을 통합하지만 건강보험을 달리하는 취업 배우자를 분리한 건강보험증기준안의 가구소득 분포에서 중고소득계층의 가구소득 수준이 낮게 나타났다.

이 장의 분석에 따르면, 전반적으로 가계금융복지조사는 맞춤형 소득-재산 DB보다 균등화 가구총소득 수준이 높고 평등하였다. 이는 한편으로 맞춤형 소득-재산 DB에 비해 가계금융복지조사가 개인소득을 고르게 포착하여 개인총소득 수준이 높고 평등하게 나타났기 때문이기도 하고, 다른 한편으로 맞춤형 소득-재산 DB에 비해 1인 가구 비율이 훨씬 낮은 가계금융복지조사에서 가구 내 소득 공유가 강하게 반영되었기 때문이기도 하다. 근로소득, 사업소득, 금융소득, 공적연금소득을 합산하고 균등화한 가구총소득을 살펴보면, 가계금융복지조사의 지니계수 및 빈곤율은 각각 0.396, 21.1%였고, 맞춤형 소득-재산 DB의 지니계수 및 빈곤율은 가구 구성안에 따라 0.486~0.537, 28.1~32.3%로 나타났다. 단, 가계금융복지조사와 달리 건강보험 피부양자인 노인을 자녀 가구에 통합한 세대보완 1인(건강보험 피부양자 포함) 및 건강보험증기준안의 노인 빈곤율은 32.1~32.9%로, 가계금융복지조사의 56.6%보다 훨씬 낮았다.

전반적인 분석 결과는 향후 맞춤형 소득-재산 DB의 소득 데이터를 더욱 발전시켜야 한다는 사실을 보여준다. 다양한 행정자료를 보완하여 소득 개념·범위를 확대하고 소득 측정의 정확도를 개선해야 한다. 또한 정확한 주소 및 가족관계 정보를 보완하고 실질적인 생계·주거 단위를 반영하여 가구 개념을 더 정교하게 다듬어야 할 것으로 판단된다.



제5장

소득과 재산의 결합 분포 분석

제1절 분석 자료 및 분석 방법

제2절 가구 구성안별 재산 분포

제3절 가구 구성안별 소득과 재산의 결합 분포

제4절 소결

제 5 장 소득과 재산의 결합 분포 분석

제1절 분석 자료 및 분석 방법

1. 소득-재산 결합 분포 분석의 의미

이 장에서는 2020년 맞춤형 소득-재산 DB를 이용하여 가구 단위의 소득-재산 교차분포를 분석할 것이다. 소득계층별로 재산의 보유실태는 어떠한지, 또는 역으로 재산 보유액에 따른 계층별로 소득의 보유실태는 어떠한지 살펴보는 것은 가구별 경제적 자원의 분배실태를 종합적으로 파악하는 의미가 있을 것이다. 또한 사회보장 영역의 많은 제도들이 대상자를 선정할 때 소득기준과 재산기준을 동시에 적용하고 있다는 점을 고려한다면 소득과 재산의 교차분포는 제도의 포괄성 확대를 논의할 때 반드시 고려해야 하는 사항이기도 하다.

2. 건보공단 맞춤형 소득-재산 DB에서의 소득과 재산

건보공단의 맞춤형 소득-재산 DB는 명칭 그대로 소득뿐만 아니라 개인의 재산보유액에 관한 정보를 제공하고 있다. 다만 모든 유형의 재산에 대해 보고하는 것은 아니다. DB에 포함된 재산 관련 항목은 주택시가표준액, 건물시가표준액, 토지시가표준액, 선박 및 항공기 시가표준액 등 부동산 자산과 전월세 보증금이다. 즉 개인 보유 자산의 중요한 구성항목인 금융자산 등 유동성 자산이 포함되어 있지 않다. 또한 전월세 보증금은 지역가입자에 대해서만 자료를 제공하고 있다. 이는 보험료를 부과할 때

직장가입자의 경우에는 재산을 고려하지 않기 때문에 관련 정보를 취합하고 있지 않기 때문이다.

이와 같이 맞춤형 소득-재산 DB가 제공하는 재산과 관련된 정보에는 한계가 있다. 그러나 이러한 점을 고려하더라도 전 국민의 소득과 재산 분포를 동시에 분석할 수 있는 자료로서 분석 가치는 충분하다고 볼 수 있다. 따라서 아래와 같은 전제하에 이 장에서는 가구 구성안별로 소득과 재산의 교차분포가 어떤 양상을 띠고 있는지 살펴보고 비교하고자 한다.

3. 주요 분석 방법

앞 장에서 살펴본 바와 같이 각 가구 구성안마다 도출되는 가구의 수는 상이하다. 그런데 어떤 방식에서든 가구원 수가 많은 가구의 경우, 시설일 가능성이 크다. 따라서 일정한 규모 이상의 가구를 분석에서 제외하는 것이 타당한데, 이 장에서는 10명을 그 기준으로 설정하였다. 즉 10명 이상의 가구는 분석에서 제외하고, 9명의 가구까지를 분석에 포함할 것이다.

가구의 소득 수준을 비교할 때에는 총소득을 이용할 것이다. 맞춤형 소득-재산 DB에서 총소득은 근로소득, 사업소득, 이자소득, 배당소득, 기타소득, 국민연금 및 직역연금소득의 합이다. 이 장의 분석 대상으로 삼는 2020년 DB는 2019년 발생 소득을 보고하고 있는데, 2019년부터는 분리과세금융소득과 분리과세주택임대소득도 보고되고 있어, 이 항목도 총 소득에 포함하였다. 가구 사이에 소득을 비교할 때에는 가구 단위의 총소득을 가구원 수로 균등화한 소득값(균등화된 총소득)을 이용하였다.

한편 소득분위를 구분할 때에는 동일 분위에 동일한 수의 개인이 포함되도록 하였다. 이를 위해 분위 구분 시에는 가구 단위의 (균등화된) 소득값을 이용하면서 가구원 수를 가중치로 사용하였다.

앞에서 언급한 바와 같이 맞춤형 소득-재산 DB에서 전월세 임대차 보증금은 건강보험 지역가입자에 대해서만 있다. 따라서 가구 자산에서 임대차보증금은 제외하기로 한다. 전월세 주택에 거주하는 자가의 비중이 작지 않다는 점을 고려한다면, 전월세 보증금을 분석에서 제외하는 것은 전체 가구의 재산 분포를 파악하는 데 큰 영향을 줄 수 있다. 그러나 직장 가입자의 경우 애당초 자료에 전월세 보증금 자료가 포함되어 있지 않은 점을 고려한다면, 전월세 보증금을 활용할 경우 집단 간 재산금액을 비교할 때 체계적인 오류를 발생시킬 수 있다. 물론 주택보유액 역시 주택을 보유하지 않은 가구나 개인에 대해서는 그 금액이 0이 된다. 그러나 주택자산이 없는 개인이 애당초 자료에서 제외된 것은 아니라는 점에서 전월세 임대보증금을 분석 대상 재산에 포함시키지 않는 것과는 차이가 있다. 이러한 이유로 이 장의 분석에서 활용하는 총재산은 가구가 보유한 주택, 건물, 토지, 선박 및 항공기 등의 시가 표준액을 합산한 값을 의미한다.

가구 재산의 규모를 분석할 때에는 소득과 달리 균등화지수를 적용하지 않았다. 재산 보유액에 따라 분위를 구분할 때에는 소득의 경우와 마찬가지로 가구원 수를 가중치로 이용하였고, 분위마다 같은 수의 개인이 포함되도록 구분하였다.

그런데 구분하는 분위의 수가 많아질 때 동일한 분위에 동일한 수의 개인이 포함되도록 분위가 정의되지 않는 경우가 있다. 0의 소득이나 재산을 보유한 개인의 수가 한 개 분위에 포함되어야 하는 개인의 수를 초과하는 경우가 그러한 경우에 해당한다. 이때 분위가 달라지면 소득(또는 재산)의 보유액이 달라져야 한다는 일반적인 분위 구분 원칙과 각 분위에 속하는 개인의 수는 같아야 한다는 원칙이 충돌하게 된다. 이 장의 연구에서는 불가피하게 0의 소득이나 0의 재산을 보유한 집단을 동일한 분위로 간주하는 방식을 적용하였다. 예를 들어 전체 인구를 10개 소득

분위로 구분하고자 하는데, 총소득이 0인 개인의 비율이 29%라면 3분위에 속하도록 하였다.¹⁾

4. <가계금융복지조사> 자료와의 비교

맞춤형 소득-재산 DB를 이용하여 가구 단위의 재산 분포를 분석하기에 앞서, 이 자료를 통계청의 <가계금융복지조사> 자료와 비교하는 것이 도움이 될 것이다. 이 장에서 분석할 가구의 소득이나 재산항목은 가구의 모든 소득과 재산을 아우르지 않는다. 따라서 이 장의 분석 대상 항목이 전체 가구의 소득과 재산에서 대략 어느 정도의 비중을 차지하는지를 파악하기 위해서는 다른 통계자료를 살펴볼 필요가 있다. 통계청의 <가계금융복지조사>는 가구 단위의 소득과 자산 및 부채에 대한 정보를 제공하는 대표적인 조사자료이다. 이하에서는 <가계금융복지조사> 2020년 공공용 마이크로데이터를 이용하여 가구의 소득 및 자산 구성에 대해 대략적으로 파악하고 자한다. 또한 다음절에 제시되는 분석 결과표에도 <가계금융복지조사>의 분석 결과를 참고로 제시하였다. 2020년의 가계금융복지조사는 2019년의 소득과 2020년 시점의 재산 상태를 보고하고 있다. 따라서 이 장에서 분석하게 될 자료의 소득 및 재산 파악 시점과 일치한다.

<표 5-1>은 가계금융복지조사 자료를 이용하여 분석한 가구 소득과 자산의 구성이다. 총소득(경상소득) 가운데 근로소득, 사업소득, 재산소득의 합이 차지하는 비율은 88.8%이다. 여기에 공적연금까지를 더하면 그 비율은 91.6%가 된다. 맞춤형 소득-재산 DB에 포함된 소득 범주는 <가계금융복지조사> 자료에서는 가구 소득의 약 93.3%에 해당하는 범주라고 볼 수 있다. 다음 절의 분석에서는 <가계금융복지조사>의 공공용 자료를

1) 즉 1, 2분위는 존재하지 않으며 3분위에는 인구의 30%가 포함되도록 하였다.

사용하는 만큼, 공적연금 부분을 제외한 근로+사업+재산소득을 총소득으로 간주하여 <가계금융복지조사>의 해당 수치를 제시할 것이다.

〈표 5-1〉 <가계금융복지조사>에서의 가구 소득 및 자산 구성

(단위: 만 원, %)

소득 항목	평균	경상소득 대비 비율	자산 항목	평균	자산총액 대비 비율
경상소득	3,544	100.0	자산	44,543	100.0
근로소득	2,203	62.2	금융자산	10,504	23.6
사업소득	679	19.2	저축액	7,632	17.1
재산소득	262	7.4	적립 및 예치식 저축	7,223	16.2
공적이전소득	318	9.0	기타 저축	408	0.9
공적연금	168	2.8	현 거주지 전월세 보증금	2,873	6.4
기초연금	63	1.1	실물자산	34,039	76.4
아동관련급여	13	0.2	부동산	31,962	71.8
장애관련급여	4	0.1	거주주택	18,945	42.5
기초보장급여	22	0.4	거주주택이외부동산	12,585	28.3
근로·자녀장려금	19	0.3	계약금 및 중도금 납입액	432	1.0
기타 공적이전소득	29	0.5	기타 실물자산	2,076	4.7
사적이전소득	82	2.3	자동차	1,111	2.5
			기타	965	2.2
			자동차이외기타실물자산	713	1.6
			권리금	252	0.6

주: 소득은 2019년, 재산은 2020년 기준임. 소득은 균등화지수를 적용한 값임.

색칠한 부분은 이 장에서 분석할 맞춤형 소득-재산 DB상의 소득 및 재산 범주에 해당하는 부분임.
자료: 통계청, (2020e), 2020년 가계금융복지조사 원자료.

마찬가지 방법으로 가구의 재산 구성을 살펴보자. 이 장에서 분석하는 맞춤형 소득-재산 DB상의 자산은 <가계금융복지조사> 자료의 거주주택, 거주주택 이외의 부동산, 자동차를 제외한 기타 자산이다. 이 세 항목이 전체 자산에서 차지하는 비율은 약 72.4%이다. 이하에서 제시하는 표에서 <가계금융복지조사>의 자산 금액은 이 세 항목을 더한 값을 이용하여 구한 값이다.

제2절 가구 구성안별 재산 분포

1. 가구 구성안별 재산 수준

다섯 가지 가구 구성안에 따른 가구원 수는 제3장에서 언급한 바와 같이 각각 다르다(〈표 3-1〉 참조). 이 장의 분석에서는 10인 이상의 가구를 분석에서 제외했기 때문에 각 구성안의 가구수는 달라진다. 〈표 5-2〉에서 보듯이 분석대상 가구수는 건강보험 피부양자까지를 포함한 가구 구성안의 경우 약 2,064만 가구이다. 다섯 가지 안 가운데 가장 가구수가 적다. 가구수가 가장 많은 안은 당연히 건강보험증을 기준으로 가구를 분리한 안으로, 가구수는 약 2천 7백만에 이른다.

전체 개인의 수가 동일한데 가구수가 적다는 것은 것은 가구당 많은 가구원들이 포함된다는 의미일 것이다. 따라서 가구수가 적은 경우 가구별 재산이 다른 안에 비해 많아질 가능성이 크다. 반대로 가구수가 많은 경우는 가구원 수가 적게 되고 가구별 평균 재산이 다른 안에 비해 낮아질 가능성이 크다. 이러한 상황은 〈표 5-2〉의 윗부분에서 확인할 수 있다. 건강보험 피부양자를 포함하는 세대보완 1안에서 가구당 평균 재산은 약 2억 6천만 원인 반면, 건강보험증 단위로 구분하였을 때에는 평균 재산이 약 1억 9천만 원이다. 세대를 단위로 가구를 구분하였을 때에나 비동거 가족을 포함하는 세대보완 2안에서는 평균재산이 약 2억 2천으로 비슷한 수준이다. 실제로 분리된 세대를 추정하여 별도의 가구로 구성하여 주는 세대보완 3안에서는 가구수가 늘어남에 따라 가구당 평균 재산도 세대 기준안에 비해 다소 적게 나타난다.

〈표 5-2〉 가구 구성안별 재산 보유 실태

(단위: 천 가구, 천 원, %)

구분		세대보완 1안	세대보완 2안	세대보완 3안	세대 기준안	건강보험증 기준안	가계금융 복지조사
가구수		20,636	22,833	23,343	22,937	27,092	20,291
평균		255,997	227,987	221,069	226,956	192,219	322,429
중위값		29,101	68,000	63,039	68,000	7,177	160,000
재산 보유 가구	비율	59.8	60.8	60.0	60.8	52.0	73.7
	평균	421,848	374,824	371,931	373,214	369,675	437,653
	중위값	214,393	191,500	190,000	191,000	192,000	260,000

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.
통계청. (2020f). 2020년 가계금융복지조사 원자료.

이러한 사정은 재산을 보유한 가구의 비율에도 반영된다. 〈표 5-2〉에서 건강보험증기준안의 재산보유 가구 비율이 가장 낮다. 세대기준안과 비동거 가족 포함안의 재산보유 가구 비율은 60.8%로 동일하다. 특이한 점은 건보 피부양자 포함안의 재산보유 비율이 세대기준안보다 낮다는 점이다. 건보 피부양자 가운데에는 재산이 있는 개인과 없는 개인이 혼재되어 있고, 재산보유액의 편차도 클 경우 이러한 현상이 나타날 수 있다고 추론할 수 있다. 예컨대 피부양자 재산이 없는 비동거(다른 세대) 자녀도 있을 수 있다. 이들을 가구로 묶을 경우 재산이 있는 가구의 비율과 평균 재산이 높아지는 방향으로 작용하게 된다. 하지만 재산이 없는 세대 간에 건강보험 부양-피부양 관계로 묶일 확률이 낮다면 전체 가구 중 재산이 없는 가구의 비율은 세대기준안에 비해 낮아질 수 있다.

재산이 있는 가구에 대해서만 재산의 평균을 비교했을 경우에도 각 구성안의 차이는 비슷한 양상으로 나타난다. 건보 피부양자를 포함한 세대보완 1안에서 약 4억 2천만 원으로 가장 높고 건강보험증을 기준으로 나누는 안에서 약 3억 7천으로 가장 낮다. 다만 구성안별 재산 평균의 상대

적 편차는 줄어들었다. 재산이 없는 가구까지 포함하여 비교했을 경우 세대보완 3안(분리세대 추정)과 건강보험증기준안의 평균 재산수준은 세대기준안 대비 약 0.97배와 0.85배였다. 재산이 있는 가구에 국한하여 비교하면 이 비율은 0.99배와 0.86배로 변한다. 재산이 없는 가구를 제외했을 때의 재산 중위값 순서는 전체 가구의 중위값 순서와 달라진다.

2. 가구 구성안별 재산 분포

재산 수준에 따라 전체 인구를 10분위로 구분하였을 때, <표 5-3>은 가구 구성안별로 분위별 평균 재산을 비교하여 보여준다. 앞서 언급하였듯이 재산액이 0인 가구는 모두 같은 분위에 속하도록 하였기 때문에 건강보험기준안을 제외하면 모든 가구 구성안에서는 3분위가 최하위 분위로 나타난다. 건강보험증기준안에는 4분위가 최하위 분위이다. 그만큼 재산이 없는 가구의 비중이 크다는 의미이다. 각 분위의 평균값을 비교하면, 세대보완 1안의 분위 평균이 다른 안들에 비해 가장 높고, 건강보험증기준안의 평균이 가장 낮다. 나머지 세 안의 분위별 평균재산 수준은 대체로 유사하지만, 세대기준안과 비교하면 세대보완 2안이 다소 높고 분리세대 추정안이 다소 낮다.

〈표 5-3〉 재산분위별 평균 재산

(단위: 천 원)

재산 분위	세대보완 1안	세대보완 2안	세대보완 3안	세대 기준안	건강보험증 기준안	가계금융복 지조사
3	764	35	1	11	-	4,916
4	62,154	39,568	33,507	36,508	1,433	97,373
5	125,395	101,353	96,909	99,223	70,969	170,768
6	188,756	154,979	150,575	152,321	133,868	241,262
7	273,902	221,037	215,350	216,926	205,425	329,427
8	400,884	320,951	313,489	315,485	311,648	458,316
9	630,282	509,458	498,263	500,882	501,008	684,199
10분위	1,804,652	1,567,975	1,539,261	1,545,081	1,457,516	1,682,442

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.
통계청. (2020f). 2020년 가계금융복지조사 원자료.

각 방안의 최하위 분위에는 재산이 0인 가구와 소액인 가구가 혼재되어 있는데, 그 비율의 차이에 따라 분위 평균 재산액의 차이가 크게 나타난다. 하위 분위에서는 구성안별로 편차가 크고 상위 분위로 갈수록 구성안별 편차는 점차 줄어든다. 이를 보여주는 것이 〈표 5-4〉이다.

〈표 5-4〉 가구 구성안별 분위 평균 재산 비교(세대기준안=100)

재산 분위	세대보완 1안	세대보완 2안	세대보완 3안	세대 기준안	건강보험증 기준안
3	7,193.6	332.3	6.9	100.0	-
4	170.2	108.4	91.8	100.0	3.9
5	126.4	102.1	97.7	100.0	71.5
6	123.9	101.7	98.9	100.0	87.9
7	126.3	101.9	99.3	100.0	94.7
8	127.1	101.7	99.4	100.0	98.8
9	125.8	101.7	99.5	100.0	100.0
10분위	116.8	101.5	99.6	100.0	94.3

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

소득분배를 나타내는 지표로 소득 5분위 배율 또는 10분위 배율이 사용된다. 이 지표를 재산의 분배상태에도 그대로 적용해 볼 수 있다. 그런데 <표 5-3>에서 보듯이, 이 장에서는 재산 하위 1, 2분위를 정의하지 않았다. 따라서 재산의 경우 분배실태를 비교하기 위해 다른 지표를 이용할 필요가 있다. 이 장에서는 두 가지 지표를 사용하였는데, 하나는 하위 4개 분위의 평균합 대비 상위 1개 분위의 평균이고, 다른 하나는 분위 구분 경계값의 비율 P90/P50이다.

첫 번째 지표는 소득분배지표로 사용되는 팔마 비율(Palma ratio)이다. 팔마 비율은 소득 하위 40% 인구의 점유율 대비 상위 10% 인구의 점유율을 비교하는데, 여기에서는 하위 4개 분위(즉 3~4분위)의 재산 점유율과 최상위 10분위의 재산 점유율 비교하였다. 두 번째 지표는 분위수 비율이다. 전체 인구를 재산 수준에 따라 10개 분위로 나눌 때, 각 분위를 구분하는 경계값(분위수)들이 모두 9개 존재하게 된다. 이 중 5분위와 6분위를 구분하는 경계값 P50²⁾ 대비 9분위와 10분위를 구분하는 경계값 P90의 비율(P90/P50)을 이용하면 중간층의 재산수준 대비 최상위층의 재산수준을 비교할 수 있다.

<표 5-5> 가구 구성안별 재산 분배지표 비교(재산분위 이용)

(단위: 배)

지표	세대보완 1안	세대보완 2안	세대보완 3안	세대 기준안	건강보험증 기준안	가계금융복 지조사
팔마 비율	19.0	27.9	32.3	30.0	131.7	11.6
P50/P10	5.3	5.3	5.3	5.3	6.4	4.4

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.
통계청. (2020f). 2020년 가계금융복지조사 원자료.

2) 즉 중위값에 해당한다.

재산의 팔마 비율을 비교해 보면, <표 5-5>에서 알 수 있듯이 세대보완 1안(건강보험 피부양자 포함)이 가장 낮고(19.0배) 건강보험증기준안이 가장 높다(131.7배). 건강보험증기준안이 이례적으로 높은 이유는, 이 안의 경우 재산이 없는 가구의 비율이 높아(68%) 3분위에 속한 개인들까지도 재산이 0이고, 4분위의 평균 재산도 낮은 수준이기 때문이다. 세대기준안의 경우 팔마 비율은 30.0배이고, 비동거 가족을 포함한 안의 경우는 이보다 다소 낮은 27.9배이다.

P90/P50에서는 각 가구 구성안별 차이가 축소되어 있다. 4개의 안에서는 그 비율이 5.3배로 동일한 반면 건강보험증기준안에서는 6.4배로 다른 안들보다 높은 수준이다. 그러나 하위 4개 분위 평균합 대비 상위 1개분위 평균 비율에서와 같이 이례적인 수치는 아니다.

이와 같이 가구 단위의 재산보유 수준과 계층별 재산의 격차는 가구 구성안별로 상이한 양상을 보인다. 더 많은 개인을 묶어 하나의 가구로 상정할 경우 당연히 가구의 평균 총재산이 높아질 가능성이 있다. 반대로 건강보험증기준안처럼 세대를 더 분할하여 가구원 수가 작아지는 경우에는 재산이 없는 가구의 비중이 커지고, 가구 평균 재산은 낮아지며, 계층 간 재산의 격차는 더 벌어지게 된다.

제3절 가구 구성안별 소득과 재산의 결합 분포

1. 가구 소득과 재산의 상관관계

일반적으로 재산의 형성은 소득을 저축하는 과정을 거쳐 이루어지게 된다. 특히 이 장의 분석에서 고려하는 재산이 주로 주택 등 부동산재산이므로 청년층보다는 중장년층에서 재산을 보유할 확률이 높는데, 이 시기는 생애주기 가운데 소득이 가장 높은 시기이기도 하다. 따라서 개인 단위로 보면 소득 수준과 재산 수준은 양의 상관관계를 보인다. 이러한 사정은 가구 단위로 상관관계를 살펴볼 때에도 마찬가지일 것으로 추론할 수 있다.

〈표 5-6〉은 다섯 가지 가구 구성안 별로 (균등화된) 가구 총소득과 총 재산의 상관관계를 비교하고 있다. 다섯 가지 방안 모두 소득과 재산 사이에 양(+의 상관관계를 보이고 있고 그 유의성도 매우 높다. 상관계수의 크기는 각 방안에 따라 차이가 있는데, 세대보완 1안(건보 피부양자 포함)에의 상관계수가 가장 높고 건강보험증기준안의 상관계수가 가장 낮다. 다른 세 안의 상관계수는 비슷한 수준이다.

〈표 5-6〉 가구 구성안별 소득-재산의 상관계수

	세대보완 1안	세대보완 2안	세대보완 3안	세대 기준안	건강보험증 기준안	가계금융복 지조사
상관 계수	0.3741 (〈.0001)	0.3579 (〈.0001)	0.3581 (〈.0001)	0.3574 (〈.0001)	0.3236 (〈.0001)	0.4142 (〈.0001)

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.
통계청. (2020). 2020년 가계금융복지조사 원자료.

건강보험증기준안의 상관계수가 낮은 이유는, 건강보험증 단위로 가구를 구분할 경우 가구원 중 재산은 없으나 소득은 있는 직장보험가입자가

별도의 가구로 간주되는 사례가 다수 있기 때문일 것으로 추론할 수 있다. 반대로 건보 피부양자를 포함할 경우 소득은 없으나 재산은 있는 노인 세대가 자녀 세대와 동일한 가구로 묶이면서 소득과 재산이 모두 있는 가구로 파악되기 때문일 것으로 유추할 수 있다.

2. 소득분위별 재산 격차

소득과 재산 사이의 상관관계가 어떤 모습을 보이는지는 소득분위별 평균 재산을 비교함으로써도 확인할 수 있다. 다만 가구 구성안별로 소득분위를 구분하는 지점과 재산분위를 구분하는 지점이 달라질 수 있음을 염두에 둘 필요가 있다. <표 5-7>은 소득분위별 평균 재산액을 가구 구성안별로 비교하여 보여준다.

<표 5-7> 소득분위별 평균 재산

(단위: 천 원)

소득 분위	세대보완 1안	세대보완 2안	세대보완 3안	세대 기준안	건강보험증 기준안	가계금융복 지조사
1	51,757	-	-	-	-	120,006
2	138,665	96,313	91,168	94,139	75,248	201,888
3	182,943	177,272	170,567	173,537	149,886	232,649
4	194,867	187,238	184,352	186,458	159,813	233,774
5	197,973	186,914	184,580	187,019	159,738	247,603
6	227,146	192,291	188,577	192,374	151,084	313,107
7	260,615	214,034	209,443	213,502	179,270	330,033
8	310,682	252,249	246,432	250,894	215,561	416,658
9	398,669	327,877	321,104	325,994	271,584	507,436
10분위	891,335	779,241	756,849	763,405	575,082	930,952

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

통계청. (2020). 2020년 가계금융복지조사 원자료.

소득분위가 높아질수록 재산의 크기도 증가한다. 소득분위가 높아질수록 재산이 증가하는 속도는 가구 구성안별로 차이가 있는데, 건보 피부양자 부양안이 소득분위별 평균 재산액이 타 구성안에 비해 높고, 소득분위가 높아질수록 평균액이 증가하는 정도도 크다. 분위별 평균재산액이 가장 낮은 것은 건강보험증 기준으로 가구를 구분했을 때이다. 다른 세 구성안의 분위별 재산을 비교할 때, 세대보완 2안(비동거 가족 포함)의 재산은 세대기준안에 비해 높고, 세대보완 3안(분리세대 추정)은 세대기준안에 비해 낮다.

소득분위별 평균 소득 격차와 평균 재산 격차를 비교하기 위해 5분위 배율을 분석하였다. <표 5-8>에는 소득을 기준으로 10분위를 구분하였을 때, 하위 2개 분위의 평균 재산 대비 상위 2개 분위의 평균 재산의 비율을 보여준다. 같은 방법으로 평균 소득의 비율도 계산하였다. 소득분위 구분에서 재산의 5분위 배율이 가장 높은 것은 세대보완 3안(분리세대 추정)이고, 가장 낮은 것은 세대보완 1안(건보 피부양자 포함)이다. 세대보완 1안의 배율이 가장 낮은 것은 재산을 기준으로 분위기를 구분했던 <표 5-4>와 같지만, 건강보험증기준안의 5분위 배율이 세대기준안과 큰 차이가 없는 것은 <표 5-4>와 다른 모습이다. 그 이유는 소득을 기준으로 분위기를 구분할 경우 소득이 0인 가구의 비율이 줄어들고, 따라서 분위 간 상하위 소득 분위 간 재산의 격차도 줄어들 것이기 때문으로 추정된다.

<표 5-8> 소득 및 재산 5분위 배율(소득분위 기준)

(단위: 배)

지표	세대보완 1안	세대보완 2안	세대보완 3안	세대 기준안	건강보험증 기준안	가계금융 복지조사
재산5분위 배율	6.77	11.50	11.82	11.57	11.25	4.47
소득5분위 배율	41.95	212.32	264.05	247.36	152.34	14.14

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.
통계청. (2020f). 2020년 가계금융복지조사 원자료.

한편 <표 5-8>을 보면, 소득분위 기준 소득 5분위 배율이 매우 높게 나타나는 것을 볼 수 있다. 이는 맞춤형 소득-재산 DB의 소득 항목이 지니는 자료상의 특징, 즉 다수의 공적이전소득 항목과 사적이전소득이 포함되지 않아 소득이 0인 가구의 비율이 많고, 반대로 사업소득과 임대 소득 등 상대적으로 고소득층에게서 많이 발생하는 소득이 포착되었기 때문인 것으로 추정된다.

소득분위별 소득 5분위 배율과 재산 5분위 배율을 비교해 보면, 재산의 5분위 배율이 절대적 수치 자체가 낮은 것을 쉽게 확인할 수 있다. 소득 분위에 따라 재산의 격차가 있지만 그 격차의 정도는 소득보다 재산이 작은 것인데, 이는 소득을 기준으로 분위기를 구분했고 소득과 재산의 상관 계수가 크지 않기 때문이라고 할 수 있다. <표 5-5>에서 볼 수 있듯이 재산을 기준으로 분위기를 구분했을 때, 상 하위 분위 간 재산 격차는 더 크게 나타난다.

3. 재산분위별 소득 격차

다음으로 이 소절에서는 재산을 기준으로 분위기를 구분하였을 때 나타나는 분위 간 소득 격차에 대해 살펴볼 것이다. 소득과 재산이 양의 상관 관계를 보이고 있기 때문에, 재산분위가 높아질수록 평균 소득이 증가할 것임은 쉽게 추측할 수 있다. <표 5-9>는 재산분위별 소득 격차의 구체적인 모습을 보여준다.

〈표 5-9〉 재산분위별 평균 소득

(단위: 천 원)

재산 분위	세대보완 1안	세대보완 2안	세대보완 3안	세대 기준안	건강보험증 기준안	가계금융복 지조사
3	16,740	17,473	17,170	17,237	-	21,632
4	20,904	16,558	16,150	16,475	18,048	21,189
5	25,375	21,262	21,059	21,273	21,241	26,614
6	29,015	25,281	25,117	25,305	24,703	31,804
7	32,711	28,709	28,690	28,808	28,278	35,348
8	36,641	32,522	32,523	32,688	32,121	40,099
9	42,026	37,510	37,544	37,703	36,927	45,235
10분위	67,789	62,805	63,155	63,384	58,553	63,836

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.
통계청. (2020f). 2020년 가계금융복지조사 원자료.

재산 4분위 이상에서 평균 소득이 가장 높은 것은 역시 건보 피부양자 포함안이다. 건강보험증기준안에서 소득 수준이 가장 낮다. 재산분위가 높아짐에 따라 평균 소득이 증가하는데, 이는 모든 가구 구성안에 공통된 현상이다. 그러나 3~4분위 구간에서는 재산분위가 높아질 때 평균 소득이 감소하는 특이한 현상이 나타난다. 4분위 구간에 재산은 있으나 소득이 없는 계층(예, 노인층)이 집중적으로 분포하면 이러한 현상이 나타날 가능성이 있다. 각각의 분위에서 평균 소득의 가구 구성안별 차이는 있지만, 재산분위가 높아질수록 평균 소득이 증가하는 속도는 6~9분위 구간에서는 다섯 가지 방안이 거의 동일하게 나타난다.

〈표 5-10〉 소득 및 재산의 팔마 비율(재산분위 이용)

(단위: 배)

구분	세대보완 1안	세대보완 2안	세대보완 3안	세대 기준안	건강보험증 기준안	가계금융복 지조사
소득	0.5	0.5	0.6	0.6	0.4	1.5
재산	19.0	27.9	32.3	30.0	131.7	11.6

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.
통계청. (2020f). 2020년 가계금융복지조사 원자료.

마지막으로, 재산을 기준으로 구분한 재산분위 사이의 소득 및 재산 격차를 살펴보자. 〈표 5-10〉에서 보듯이 재산분위별 소득 팔마 비율은 모든 가구 구성안에서 12보다 낮다. 재산분위별 재산 팔마 비율에 비해 크게 낮은 수치이다. 이는 낮은 재산분위와 높은 재산분위와의 소득 격차가 재산 격차만큼 심하지 않음을 의미한다. 소득 팔마 비율 자체는 세대기준안과 세대보완 3안(분리세대 추정)이 가장 높고 건강보험기준안이 가장 낮다. 〈표 5-10〉의 하단은 참고를 위해 〈표 5-5〉로부터 가져온 것이다. 재산을 기준으로 분위를 구분한 경우의 소득불평등은 재산불평등에 비해 훨씬 완화된 형태로 나타나는 것을 확인할 수 있다.

제4절 소결

이 장에서는 건보공단의 맞춤형 소득-재산 DB를 이용하여 3장에서 유형화한 다섯 가지 가구 구성안별로 가구 재산의 수준과 분포가 어떻게 다르게 나타나는지를 분석하였다. 상대적으로 더 많은 수의 세대와 개인을 묶는 세대보완 1안(건강보험 피부양자 포함)의 경우 총가구수는 적고 가구당 평균 재산액은 높다. 반대로 건강보험증기준안은 세대를 분할 하기도 하기 때문에 가구의 수가 많아지고 평균 재산은 감소한다. 직장

보험 보유자가 별도의 가구를 구성한다고 가정할 경우, 소득은 있으면서 재산이 없거나 적은 가구의 비중이 커질 가능성이 있기 때문이다. 이러한 이유로 재산의 계층 간 격차도 세대보완 1안(건강보험 피부양자 포함)에서는 작게, 건강보험증기준안에서는 크게 나타난다.

반면 세대를 기준으로 구성한 안과 세대 기준을 변형한 두 안(비동거 가족 포함안, 분리세대 추정안) 사이의 차이는 상대적으로 작다. 세대기준안에 비해 가구원 수가 늘어나는 세대보완 2안(비동거 가족 포함)은 평균 재산이 다소 많고, 반대로 세대보완 3안(분리세대 추정)은 세대원 가운데 일부를 별도 세대로 간주하게 되므로 평균 재산이 다소 적어지긴 한다. 그러나 세대기준안과의 차이는 두 안에서 크지 않다.

가구별 재산의 수준과 재산의 계층 간 분포를 중심으로 보았을 때, 건강보험증을 중심으로 가구를 구성한 두 종류의 안과 주민등록상 세대를 중심으로 가구를 구성한 세 종류의 안 사이에 차이가 존재한다는 점을 확인하였다. 세대보완 1안(건강보험 피부양자 포함)은 가구의 범위를 확장시키는 효과가 큰 것으로, 건강보험증기준안은 반대로 가구의 범위를 축소시키는 효과가 큰 것으로 추정된다. 특히 건강보험증 정보를 이용할 경우 소득은 적으나 재산이 있는 개인들(예, 피부양 노인)이나, 반대로 소득은 있으나 재산이 적은 개인들(예, 청년층 직장가입자)의 분포가 가구별 재산 분포나 소득-재산 결합 분포에 작지 않은 영향을 줄 수 있음을 염두에 두어야 할 것이다.



제6장

소득-자산 결합 방법과 적용

제1절 소득-자산 결합 지표의 의미

제2절 연구의 내용 및 방법

제3절 결합 지수의 활용 및 특징 비교

제4절 소결

제 6 장 소득-자산 결합 방법과 적용

제1절 소득-자산 결합 지표의 의미

본 장에서는 소득과 재산을 동시에 고려하여 가구의 실질적인 지불 능력을 반영하는 소득-재산 결합지수를 고안하고자 하는 선행연구를 소개하고, 이들 중 대표적인 두 가지 방법인 연금소득화법과 소득인정액 환산법을 건강보험공단 DB에 적용해 보고자 한다. 본 보고서의 앞에서 지적했듯이, 건강보험공단 DB는 몇 가지 한계에도 불구하고 전 국민의 소득과 재산에 대한 행정데이터를 제공한다. 이를 이용하여 가구의 지불 능력의 분포를 추정해 본다는 것은 큰 의미를 가지고 있다고 할 수 있다. 가구의 지불능력의 분포를 통해 경제적 취약계층을 파악하면, 기존의 복지 제도를 평가하고 개선하는 데 도움이 될 뿐만 아니라, 복지의 사각지대를 찾는 것에도 도움이 되며, 새로운 복지제도를 설계할 때도 중요한 통계를 제공할 수 있게 된다. 나아가 기술의 발전이나 경제환경의 변화가 사회 경제에 미치는 영향에 관한 연구에도 중요한 자료로 사용될 수 있다.

그동안 개인과 가구의 소득과 재산의 분포에 대해서는 ‘불평등’이라는 대주제하에 많은 연구가 이루어져 왔다. 그런데 선행연구들의 대부분은 소득을 가구의 경제적 상태의 척도(measure)로 사용하면서, 재산의 분포나 불평등에 대해서는 상대적으로 덜 주목해 온 것이 사실이다. 재산에 대한 선행연구가 상대적으로 부족한 직접적인 원인은 재산을 정확하게 측정하는 것이 어려워 신뢰성 있는 데이터가 많지 않기 때문이다. 예를 들어, 불평등의 정도에 관해 국가 간 비교를 하려고 한다든지, 한 국가의 불평등의 추세를 살펴보고 싶을 때, 국제적으로 혹은 시계열적으로 신뢰성

있고 일관된 재산의 분포 데이터를 구하기 어렵다. 따라서 데이터가 비교적 잘 정비되어 있다는 미국의 경우에서조차 아주 최근까지 재산의 분포에 대한 연구가 제대로 진행되지 못해 온 것이 사실이다(Piketty and Saez, 2003; Saez and Zucman, 2016).

하지만 이론적으로 보면, 개인과 가구의 경제적 상태는 단순히 현재의 소득에만 의존하는 것이 아니라 미래의 소득에 영향을 주는 현재의 재산에 의해서도 영향을 받는다(Ando and Modigliani, 1963의 생애주기가설; Friedman, 1957의 항상소득가설). 예를 들어, 현재 소득이 충분히 높다고 하더라도 재산이 없는 경우 미래의 경제적 충격에 대한 대비를 위해 현재 소비를 적게 할 수 있고, 반면 실직 같은 예상치 못한 경제적 충격에 의해 소득의 상당 부분을 상실했다고 하더라도 유동화할 수 있는 재산이 충분히 있다면 현재의 소비가 크게 영향받지 않을 수 있다. 따라서 재산에 대한 가용하고 신뢰성 있는 데이터가 점차 많아지면서 가구의 경제적 상태를 측정할 때 재산을 고려하는 시도가 많아지는 것은 이론적 입장에서 보면 당연한 것이라고 할 수 있다. 게다가 최근 기술의 발전에 따라 부의 집중에 대한 사회적 우려가 커지는 상황에서, 복지제도에서도 소득뿐만 아니라 재산까지 고려하여 정책의 지원 대상과 규모를 결정할 필요성도 제기되고 있다.

우리나라의 불평등 논의에서도 재산에 대한 관심이 점차 커지는 추세이다. 최근 우리나라에서도 주식투자를 하는 가구의 비중이 확대됨에 따라 가계의 금융재산 보유가 큰 폭으로 증가하고 있다. 또한 주택가격이 그동안 큰 폭으로 상승하여 주택 가치도 대폭 상승하였다(최요철·김은영, 2007).

본 장의 구성은 다음과 같다. 먼저 2절에서는 선행연구에서 소득과 재산을 결합하는 데 사용한 방법을 소개한다. 특히 여러 방법들 중 선행

연구에서 많이 논의된 연금소득화법과 우리나라의 여러 복지정책에서 사용하고 있는 소득인정액 환산법을 설명할 것이다. 그리고 이러한 방법을 어떻게 건강보험공단 DB에 적용할 수 있는지 논의한다. 3절에서는 두 가지 방법으로 계산된 경제적 지위 변수를 이용하여 인구집단별 혹은 시계열상 경제적 지위의 분포가 어떻게 변하는지를 살펴보고자 한다. 마지막으로, 4절에서는 본 장의 논의를 요약 정리하고 다음 단계의 연구 과제에 대하여 소개하고자 한다.

제2절 연구의 내용 및 방법

그동안 선행연구에서 이론적으로 검토되고 복지정책 등에서 실제로 시도해 온 소득과 재산을 결합하는 다양한 방법들은 아래와 같은 공식으로 일반화하여 나타낼 수 있다.

$$\begin{aligned} \text{소득 - 재산 결합지수} &= f(\text{소득}, \text{재산}, \text{환산율}) & (1) \\ &= \text{소득} + \text{환산율} \times \text{재산} \end{aligned}$$

그동안 제시되었던 소득과 재산을 결합하는 방법들은 위의 공식에서 볼 수 있듯이, 기본적으로 재산을 소득과 비교 가능한 수준으로 환산하여 이렇게 환산된 재산과 소득을 합치는 것이라고 할 수 있다. 그런데 위의 공식을 실제 데이터에 적용하여 소득과 재산을 결합하고자 할 때 결정해야 할 것이 있다. 첫째, 소득과 재산의 귀속 단위를 결정해야 한다. 선행연구 그리고 본 연구에서 귀속 단위는 개인보다는 가구를 사용한다. 왜냐하면 소득과 재산은 둘 다 가구 내에서 실질적으로는 가구원들끼리 공동으로 소유되고 사용되는 경우가 많기 때문이다. 예를 들어, 만약 개인으로

단위를 삼는 경우, 재산 중에 상당 부분을 차지하는 주택의 경우 가구원 중 누구에게 귀속시켜야 하는 문제가 발생하는데, 이에 대한 정답은 존재하지 않고, 임의로 나누게 될 경우 오히려 잘못된 결론에 이르게 될 가능성이 있다. 심지어 소득의 경우에도 가구원의 모든 소득을 합산하여 전체 예산에서 지출을 배분한다는 가정이 합리적일 수 있다.

둘째, 위의 공식을 보면 재산을 어떻게 소득과 비교 가능하도록 환산할 것인가의 문제가 중요하다는 것을 알 수 있다. 즉, 위의 공식을 실제로 데이터에 적용하기 위해서는 어떤 환산율을 사용할 것인지를 결정해야 한다. 환산율의 결정은 재산을 매각하여 사용하는 기간의 설정에 따라 달라질 수 있다. 예를 들어, 기초생활보장제도에서는 일반재산을 매각할 경우 2년 동안 균분사용하는 것을 기준으로 4.17%의 환산율을 적용하고 있다. 또 환산율은 재산의 유형에 따라 달리 적용될 수 있다. 예를 들어, 기초생활보장제도에서는 평가하는 재산을 주거재산, 일반재산, 금융재산, 자동차로 구분하고 각각 다른 환산율을 적용한다. 유동성이 높은 금융재산에는 일반재산보다 높은 환산율을 부과하고, 필수재산으로 간주되는 주거재산에는 일반재산의 1/4 수준인 낮은 환산율을 부과하여 필수재산을 가구의 지불능력으로 전환될 수 없다라는 관점을 반영하고 있다.

셋째, 위의 공식에 소득과 재산이라는 변수가 들어 있는데, 실제 데이터에서 어떤 종류의 소득과 재산을 포함시켜야 하는지 결정해야 한다. 이러한 문제는 소득의 경우에는 다소 쉽지만, 재산의 경우에는 어떤 재산을 포함시킬 것인지, 어떤 재산은 포함시키지 말 것인지를 결정하는 것이 쉽지는 않다. 위에서 지적했듯이 주택의 경우 필수재산이므로 아예 재산에 포함시키지 말아야 한다는 의견이 있을 수 있다. 다른 종류의 재산도 판단이 그리 쉽지는 않다. Saez & Zucman(2016)은 연금저축의 경우 재산에 포함을 시키지 않았다. 연금저축은 연금을 수령하게 되면 소득의 형태로

받게 될 미래의 소득이기 때문이다. 만약에 연금저축을 현재 재산에 포함시키게 되면 미래에 발생할 다른 소득(공적 이전소득 등)과 지출(세금 등)에 대해서도 포함시킬 것인지 아닌지를 결정해야 하는 어려운 문제에 봉착하게 된다. 본 연구는 건강보험공단 DB의 활용 가능성을 탐색하는 연구이므로 일단 재산에 포함되는 항목은 건강보험공단 DB에서 가용한 항목들로 제한할 예정이다.

마지막으로, 위의 공식을 적용하는 데 있어서 함수 $f(\cdot)$ 의 형태를 결정해야 한다. 가장 쉽게는 환산율을 하나의 독립적인 외생적 파라미터로 놓고 환산율이 적용된 재산을 소득과 선형적으로 합하는 함수를 생각해볼 수 있다. 그러나 환산율이 소득이나 재산의 크기에 의존하는 것도 이론적으로 충분히 지지가 될 수 있는 것이기 때문에 함수 $f(\cdot)$ 가 복잡한 비선형함수의 형태를 띠 수도 있다. 본 연구에서는 논의의 편의를 위하여 선행연구를 따라 환산율을 외생적 파라미터로 놓고 소득과 환산된 재산을 합하는 공식을 이용하고자 한다.

1. 연금소득화법

연금소득화법은 소득과 재산을 통합하여 개인의 경제적 상태를 평가하고자 Weisbrod & Hansen(1968)이 개발한 모형이다. 저자들은 소득의 분포와 재산의 분포가 매우 상이하며, 소득뿐만 아니라 재산을 함께 평가해야 개인의 진정한 구매력을 평가할 수 있다고 강조했다. 이 모형의 핵심은 재산을 매각하여 잔여기대수명 동안 연금처럼 사용할 수 있다는 가정이다.

$$\begin{aligned}
 AY_t &= Y_t + A_n NW_{t-1} \\
 &= Y_t + \frac{r}{1 - (1+r)^{-n}} NW_{t-1}
 \end{aligned}
 \tag{2}$$

Weisbrod & Hansen(1968)의 공식은 소득(Y_t)과 환산된 재산($A_n NW_{t-1}$)의 합으로 이루어져 있고, 여기서 A_n 이 일반화된 공식에서의 환산율에 해당한다. 환산율은 재산을 연금의 형태로 남은 생애기간 동안 사용한다는 가정하에 정의되었다. 여기서 순재산이 매각되어 연금으로 사용되는 기간(n)은 잔여기대수명으로 설정하고, 재산의 수익률(r)은 이자율로 설정한다. 연금을 잔여기대수명 동안 균분 사용한다는 가정은 생애주기가설(life-cycle hypothesis)에 기반한다. 생애주기가설은 개인이 평생소득(permanent income)을 고려해 현재의 소비와 저축을 결정한다는 이론이다. 이 가설은 은퇴, 실직 또는 예상치 못한 의료비 지출 등으로 소득이 감소할 때 이전에 축적한 재산을 처분하여 소비에 활용한다는 것을 가정한다(Ando & Modigliani, 1963). 특히, Weisbrod & Hansen(1968)은 잔여기대수명 계산을 위해 성별-연령별 평균수명을 이용했다.

연금화된 소득(AY_t)은 연소득(Y_t)과 연금화된 순재산의 연간 가치($A_n NW_{t-1}$, annual lifetime annuity value)의 합으로 구성된 선형함수이다. 연금화된 소득은 순재산의 가치가 클수록, 잔여기대수명이 짧을수록, 이자율이 높을수록 그 값이 커진다. 이 모형에서 연금화된 소득은 연령에 직접적으로 의존하며, 재산이 잔여기대수명 동안 모두 소진되므로 유산 등 사망 이후 잔여재산이 없다고 가정한다.

이처럼 Weisbrod & Hansen(1968)은 재산을 소득과 함께 평가할 수 있는 하나의 지수를 개발했고, 구체적으로 어떠한 재산을 포함하여 가구별로 어떻게 평가할 것인지에 대한 논의는 이후의 연구에서 이어져왔다. 그중 Brandolini, Magri & Smeeding(2010)은 Weisbrod & Hansen

(1968)의 연금화 모델을 Luxembourg Wealth Study의 10개국에 적용하여 국가별로 재산을 연금화한 소득을 산출했다. Brandolini et al. (2010)은 주거재산과 금융재산 등을 포함한 순재산에 대해 연금 가치를 계산했으며, 소득으로는 근로소득과 사업소득, 금융소득, 연금소득, 기타 이전소득 등을 포함했다. 또한, 가구원 규모를 반영하기 위해 가구 내 성인의 수의 제곱근을 가구 균등화 지수로 사용하였다. 따라서 연금화된 소득을 가구 균등화 지수로 나누어 가구원 규모별로 표준화된 연금화 소득을 구하였다.

본 연구는 Weisbrod & Hansen(1968)의 모형을 Brandolini et al. (2010)의 방식으로 건강보험공단 DB에 적용하고자 한다. 본 연구에서는 가구의 연소득을 Brandolini et al.(2010)의 방식과 마찬가지로 근로소득, 사업소득, 이자소득, 배당소득, 기타소득, 국민연금소득, 지역연금소득, 이자배당소득(분리과세금융소득), 사업소득(분리과세주택임대소득)의 합으로 구한다. 건강보험공단 DB에서는 금융재산의 정보가 없는 것이 한계이지만, 가용한 정보로부터 가구의 재산 능력을 평가하기 위해 재산의 가치는 전월세 보증금, 주택·건물·토지·선박항공기의 시가표준액의 합으로 구한다.

이자율(r)은 Brandolini et al.(2010)에서와 마찬가지로 2%와 10%로 설정하여 분석하고자 하며, 가구 균등화 지수는 연령 구분 없이 가구원 수의 제곱근을 사용하고자 한다. 잔여기대수명(n)은 Weisbrod & Hansen (1968)의 방식을 따라 통계청 『생명표』의 연도별 및 성별 기대수명과 가구주의 연령 사이의 차이로 계산할 수 있다.

2. 기초생활보장제도의 환산법

2003년부터 재산의 소득환산제를 실시한 기초생활보장제도는 소득과 재산 모두를 활용해 산정한 ‘소득인정액’을 기준으로 수급자 및 급여액을 결정하고 있다. 초기에는 평가재산의 유형을 일반재산, 금융재산, 자동차로 구분하였지만, 2013년부터 주거용 재산을 일반재산으로부터 분리했다. 주거용 재산은 필수재산으로서 이를 매각하여 다른 재화나 금전으로 변환하기 비교적 더 어렵기 때문이다(여유진 외, 2011).

재산가액을 소득으로 환산하는 환산율은 재산의 유형별로 상이하다. 일반재산은 2년 동안의 균분 사용을 전제로 4.17%의 환산율을 적용하고, 금융재산은 유동성이 높아 소득 전환이 더 용이하기 때문에 균등 사용기간 16개월을 전제로 일반재산의 1.5배인 6.26%의 환산율을 적용한다. 승용차는 사치재의 성격을 가진다는 국민 정서를 고려하여 월 100% 환산율을 적용했다(보건복지부, 2003). 주거재산에는 대출이자율과 전가임대료를 반영하여 일반재산보다 더 낮은 환산율 1.04%를 적용하였고, 그 대신 주거재산으로 인정할 수 있는 한도액을 적용하고 있다(여유진 등, 2011). 주거재산 인정 한도액을 초과하는 금액은 일반재산으로 포함하여 환원한다.

$$\text{소득인정액} = \text{소득평가액} + \text{재산의 소득환산액} \quad (3)$$

$$\begin{aligned} & \text{재산의 소득환산액} \\ &= (\text{재산} - \text{부채} - \text{기본재산공제액}) \times \text{환산율} \quad (4) \\ &= \{(\text{주거} \cdot \text{일반} \cdot \text{금융재산가액} - \text{부채} - \text{재산공제액}) + \text{승용차가액}\} \\ & \quad \times \text{종류별 환산율} \end{aligned}$$

기초생활보장제도에서 순재산 전체를 소득으로 환원하는 것은 아니다.

생활에 필수적인 재산을 보호하기 위해 기본재산을 공제하기 때문이다 (정은희 외, 2018). 기본재산 공제액은 지역을 대도시, 중소도시, 농어촌으로 구분하고, 지역별 최저 주거면적 전세가격을 기준으로 차등 적용한다. 2019년을 기준으로 지역별 구분과 주거재산 인정 한도액 및 기본재산 공제액은 <표 6-1>과 같다.

<표 6-1> 지역별 주거재산 인정 한도액 및 기본재산 공제액

분류	대도시	중소도시	농어촌
구성	서울, 광역시	도의 '시', 세종, 제주	도의 '군'
주거재산 인정 한도액	1억 원	6,800만 원	3,800만 원
기본재산 공제액	5,400만 원	3,400만 원	2,900만 원

자료: 보건복지부. (2019). 2019년 국민기초생활보장사업 안내.

예를 들어, 대도시에 거주하며 주택을 1억 4천만 원, 일반재산을 3천만 원 보유하고 있는 가구의 경우, 대도시 주거재산 한도액 1억 원까지 주거재산으로 인정한다. 1억 원에서 기본재산 공제액 5,400만 원을 제외한 4,600만 원에 대해 주거재산 환산율인 1.04%를 적용할 수 있고, 그 외 주거재산 한도액 1.2억을 초과하는 4천만 원과 일반재산 3천만 원에 대해서는 일반재산 환산율인 4.17%를 적용한다. 주거재산에서 기본재산 공제액 모두를 공제하였으므로, 일반재산에서는 공제금액이 없다.

기초생활보장제도에서 소득과 재산은 가구를 기준으로 산정되는데, 가구의 정의는 생계나 주거를 같이하는 가족(배우자, 직계혈족 및 형제자매 또는 직계혈족의 배우자 등)으로 한정된다.

본 연구에서는 기초생활보장제도의 재산 유형별 환산율을 건강보험공단 DB에 적용하여 재산을 소득으로 환원하고자 한다. 기초생활보장제도에서는 총재산에서 부채를 뺀 순재산을 소득으로 환산하지만, 건강보험공단 DB에서는 부채금액을 확인할 수 없다. 또한, 건강보험공단 DB에서는

주거재산과 일반재산만 확인 가능하다는 한계가 있다.

〈표 6-2〉 기초생활보장제도의 가구원 규모별 가구 균등화 지수

가구원 수	1	2	3	4	5	6	7	8
균등화 지수	1.00	1.70	2.20	2.70	3.20	3.70	4.20	4.70

자료: 노대명 외. (2013). 국민기초생활보장제도의 맞춤형 급여체계 개편방안 마련을 위한 연구. p.215.

본 연구에서는 전월세 보증금과 주택시가표준액을 주거재산으로, 건물, 토지, 선박항공기 시가표준액을 일반재산으로 포함한다. 기초생활보장제도와 마찬가지로 주거재산에는 1.04%, 일반재산에는 4.17%의 환산율을 적용하고자 한다. 연금소득화법에서는 연 단위의 소득을 환산했지만, 기초생활보장제도의 환산율은 월 단위의 소득인정액을 계산하기 위한 것이다. 따라서 연 단위의 소득인정액 산정을 위해서는 월 단위 소득인정액에 12를 곱해주어야 한다. 소득은 연금소득화법에서와 마찬가지로 근로소득, 사업소득, 이자소득, 배당소득, 기타소득, 국민연금소득, 직역연금소득, 이자배당소득(분리과세금융소득), 사업소득(분리과세주택임대소득)을 합산한 금액으로 계산한다. 마지막으로, 기초생활보장제도에서는 〈표 6-2〉에 제시되어 있는 옥스퍼드 방식의 균등화 지수를 사용하고 있지만, 본 연구에서는 연금소득화법의 결과와 비교하기 위해 동일한 방식으로 가구원 수의 제곱근을 사용하고자 한다.

제3절 결합 지수의 활용 및 특징 비교

본 절에서는 2절에서 소개한 두 가지 소득-재산 결합 방식인 연금소득화법과 기초생활보장제도의 환산법을 건강보험공단 데이터에 직접 적용

하여 두 가지 방식을 비교해 보고자 한다. 결합 방식은 두 가지이지만 연금 소득화법의 경우 이자율을 2%와 10% 두 가지를 적용해보려고 한다. 또 비교를 위해 재산을 무시하고 소득만으로 가구의 지불능력을 계산하는 경우도 함께 제시하고자 한다. 데이터는 2019년 한 해를 이용하였다.

〈표 6-3〉 가구의 정의와 소득-재산 결합방식에 따른 결합지수의 차이

(단위: 원)

가구정의	결합방법	평균	표준편차
세대보완 1안 (건강보험 피부양자 포함)	소득만 사용	25,729,824	58,429,323
	연금소득화법(2%)	33,003,597	80,655,358
	연금소득화법(10%)	41,674,451	101,827,462
	기초생활보장제도 환산법	92,022,341	277,857,221
세대보완 2안 (비동거 가족 포함)	소득만 사용	24,351,101	57,378,301
	연금소득화법(2%)	31,941,633	78,173,199
	연금소득화법(10%)	40,204,660	98,318,227
	기초생활보장제도 환산법	87,083,697	266,850,380
세대보완 3안 (분리세대 추정)	소득만 사용	24,212,073	58,029,751
	연금소득화법(2%)	31,715,342	78,570,584
	연금소득화법(10%)	39,918,479	98,872,722
	기초생활보장제도 환산법	86,401,562	268,921,975
세대기준안	소득만 사용	24,450,867	58,348,515
	연금소득화법(2%)	32,063,745	79,026,521
	연금소득화법(10%)	40,386,345	99,459,256
	기초생활보장제도 환산법	87,605,079	270,666,744
건강보험증 기준안	소득만 사용	23,784,467	63,026,079
	연금소득화법(2%)	30,346,565	83,646,061
	연금소득화법(10%)	38,111,654	104,118,844
	기초생활보장제도 환산법	82,023,411	278,139,336

주: “소득만 사용”은 다른 결합방법과 동일한 가구 균등화 지수를 이용하여 계산하였음.
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-3〉은 다섯 가지의 서로 다른 가구 정의와 네 가지의 소득-재산 결합 지수(소득만 사용, 연금소득화법 2%, 10%, 기초생활보장제도 소득 인정액 환산법)의 평균과 표준편차를 보여주고 있다. 두 가지 특징을 찾을 수 있는데, 첫째, 가구를 정의하는 방식은, 평균과 표준편차에 있어서는 큰 차이를 보이지 않는다는 점을 알 수 있다. 둘째, 하지만 하나의 데이터 내에서 결합방법에 따라서는 결합지수의 차이가 크게 나타난다. 당연히 소득만을 사용하는 경우에는 평균이 낮게 나오고, 연금소득화법으로 계산한 지수는 10% 이자율을 사용하는 경우가 2% 이자율 사용하는 경우보다 더 높게 나타난다. 기초생활보장제도의 소득인정액 환산법을 적용하면 결합 지수의 평균이 다른 방식들보다 훨씬 높게 나타나는 것을 알 수 있다.

〈표 6-4〉 결합지수가 0인 가구의 비율

가구정의	결합방법	0원인 가구 비율
세대보완 1안 (건강보험 피부양자 포함)	소득만 사용	15.7%
	연금소득화법(2%)	11.2%
	연금소득화법(10%)	11.2%
	기초생활보장제도 환산법	12.8%
세대보완 2안 (비동거 가족 포함)	소득만 사용	18.4%
	연금소득화법(2%)	10.9%
	연금소득화법(10%)	10.9%
	기초생활보장제도 환산법	13.0%
세대보완 3안 (분리세대 추정)	소득만 사용	19.2%
	연금소득화법(2%)	11.6%
	연금소득화법(10%)	11.6%
	기초생활보장제도 환산법	13.6%
세대기준안	소득만 사용	18.7%
	연금소득화법(2%)	11.1%
	연금소득화법(10%)	11.1%
	기초생활보장제도 환산법	13.1%

가구정의	결합방법	0원인 가구 비율
건강보험증 기준안	소득만 사용	15.3%
	연금소득화법(2%)	11.2%
	연금소득화법(10%)	11.2%
	기초생활보장제도 환산법	12.6%

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-4〉는 계산된 결합지수가 0으로 나오는 가구의 비율을 제시한 것이다. 가구의 정의와 결합 방법에 따라 다르게 나타나지만 전반적으로 크게 다르지는 않은 것을 알 수 있다. 다만 모든 경우에 결합지수가 0원으로 나오는 경우가 10% 이상으로 나오기 때문에, 이를 고려하여 아래의 분석에서는 0원의 가구를 포함하는 경우와 제외한 경우의 두 가지 경우를 따로 다루었다.

〈표 6-5〉 0을 포함한 결합지수 중위값의 50% 미만 가구 특성: 세대보완 1안

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.459	0.465	0.463	0.464
평균 가구원 수	1.900	1.857	1.816	1.814
1인 가구 비중	0.554	0.579	0.593	0.590
35세 이하 가구주 비중	0.166	0.190	0.200	0.217
36~54세 가구주 비중	0.414	0.431	0.423	0.417
55세 이상 가구주 비중	0.420	0.380	0.377	0.367
대도시	0.500	0.501	0.503	0.523
중소도시	0.380	0.381	0.380	0.372
농어촌(도의 군)	0.119	0.118	0.116	0.105
18세 미만 가구원 유무	0.179	0.185	0.179	0.173
18세 미만 가구원 수	0.283	0.292	0.282	0.269
65세 이상 가구원 유무	0.267	0.232	0.227	0.221
65세 이상 가구원 수	0.323	0.269	0.262	0.255

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-6〉 0을 포함한 결합지수 중위값의 50% 미만 가구 특성: 세대보완 2인

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.482	0.493	0.488	0.485
평균 가구원 수	1.783	1.784	1.762	1.772
1인 가구 비중	0.555	0.581	0.593	0.587
35세 이하 가구주 비중	0.125	0.158	0.169	0.186
36~54세 가구주 비중	0.314	0.359	0.360	0.360
55세 이상 가구주 비중	0.561	0.483	0.471	0.454
대도시	0.483	0.485	0.488	0.514
중소도시	0.378	0.380	0.380	0.371
농어촌(도의 군)	0.139	0.134	0.132	0.115
18세 미만 가구원 유무	0.133	0.154	0.154	0.153
18세 미만 가구원 수	0.209	0.244	0.243	0.239
65세 이상 가구원 유무	0.411	0.334	0.321	0.308
65세 이상 가구원 수	0.520	0.399	0.382	0.366

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-7〉 0을 포함한 결합지수 중위값의 50% 미만 가구 특성: 세대보완 3인

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.487	0.497	0.492	0.489
평균 가구원 수	1.759	1.757	1.735	1.743
1인 가구 비중	0.565	0.591	0.603	0.598
35세 이하 가구주 비중	0.151	0.187	0.200	0.213
36~54세 가구주 비중	0.310	0.352	0.352	0.352
55세 이상 가구주 비중	0.539	0.460	0.449	0.435
대도시	0.488	0.491	0.494	0.519
중소도시	0.375	0.377	0.377	0.368
농어촌(도의 군)	0.137	0.132	0.129	0.113
18세 미만 가구원 유무	0.136	0.157	0.156	0.155
18세 미만 가구원 수	0.213	0.246	0.245	0.240
65세 이상 가구원 유무	0.393	0.317	0.305	0.294
65세 이상 가구원 수	0.497	0.378	0.361	0.349

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-8〉 0을 포함한 결합지수 중위값의 50% 미만 가구 특성: 세대기준안

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.486	0.496	0.491	0.487
평균 가구원 수	1.794	1.799	1.779	1.789
1인 가구 비중	0.548	0.571	0.583	0.576
35세 이하 가구주 비중	0.146	0.182	0.194	0.208
36~54세 가구주 비중	0.312	0.355	0.355	0.355
55세 이상 가구주 비중	0.542	0.463	0.452	0.436
대도시	0.488	0.491	0.494	0.520
중소도시	0.375	0.377	0.377	0.368
농어촌(도의 군)	0.137	0.132	0.129	0.113
18세 미만 가구원 유무	0.138	0.159	0.158	0.157
18세 미만 가구원 수	0.218	0.253	0.252	0.248
65세 이상 가구원 유무	0.398	0.322	0.310	0.298
65세 이상 가구원 수	0.507	0.387	0.370	0.357

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-9〉 0을 포함한 결합지수 중위값의 50% 미만 가구 특성: 건강보험증기준안

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.156	0.161	0.156	0.149
평균 가구원 수	0.478	0.485	0.486	0.492
1인 가구 비중	1.689	1.636	1.592	1.568
35세 이하 가구주 비중	0.215	0.248	0.263	0.272
36~54세 가구주 비중	0.400	0.413	0.403	0.401
55세 이상 가구주 비중	0.384	0.339	0.333	0.328
대도시	0.509	0.508	0.511	0.528
중소도시	0.382	0.384	0.383	0.375
농어촌(도의 군)	0.109	0.108	0.106	0.096
18세 미만 가구원 유무	0.156	0.161	0.156	0.149
18세 미만 가구원 수	0.239	0.246	0.237	0.224
65세 이상 가구원 유무	0.225	0.191	0.185	0.180
65세 이상 가구원 수	0.269	0.218	0.210	0.204

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

우선, <표 6-5>에서 <표 6-9>까지는 결합지수가 0원인 가구를 모두 포함한 상태에서 결합지수의 중위값의 50% 미만인 가구의 인구학적 특성을 요약 제시하고 있다. 각 데이터와 결합방식에 따른 중위값의 50%는 부록에 표로 정리해 두었다.

가구의 정의를 가계금융복지조사의 정의에 가장 가깝게 한 경우(세대보완 2안)를 중심으로 중위 50% 미만의 가구의 특징을 살펴보면 다음과 같다(<표 6-6>). 먼저 재산을 무시하고 소득만으로 가구의 경제적 순위 지수를 계산한 경우, 가구주가 여성인 경우가 48.2%로 높게 나타났다. 평균 가구원 수는 1.8명이며, 1인 가구 비중이 55.5%이다. 가구주의 연령을 보면, 가구주가 55세 이상인 가구의 비중이 56.1%로 나타났다. 가구의 거주지를 보면, 서울을 비롯한 대도시에 거주하는 비율이 48.3%, 중소도시가 37.8%, 나머지 농어촌 지역이 13.9%이다. 18세 미만 가구원이 있는 가구는 13.3%이고, 65세 이상 가구원이 있는 경우는 41.1%로 나타났다.

결합지수 계산방식에 따라 중위 50% 미만 가구의 구성이 크게 바뀌지는 않는 것으로 나타났다. 특히 주목할 점은 재산을 고려하는지 여부에 따라 서도 평균적인 특성은 크게 바뀌지 않는 것으로 보인다. 예를 들어 여성 가구주의 비중은 결합지수의 종류에 상관없이 48~49%로, 상당히 유사하게 나타나는 것을 확인할 수 있다. 이러한 결과는 재산을 가지고 있는 가구의 비중이 그렇게 높지 않기 때문에 나타난 것으로 보인다. 특히 저소득 가구의 경우 재산이 적을 것이기 때문에 더욱 더 재산이 결합지수에 미치는 영향이 적게 나타난 것이다.

한 가지 주목할 것은 가구주의 연령과 (가구주를 포함한) 고령 가구원의 유무이다. 재산을 무시하고 소득만을 고려할 경우 중위의 50% 미만 가구 중 가구주가 55세 이상인 경우의 비중이 56.1%인 반면, 재산을 고려하는

나머지 세 가지 결합지수에서는 45~49%로 낮게 나타났다. 재산을 연금화 하여 소득에 더하는 방식에서는 47~49%로 나타났고, 기초생활보장제도의 방식에서 가장 낮게 45.4%로 나타났다. 아마도 가구주 때문인 것으로 보이지만 65세 이상 가구원 유무의 경우에도 비슷한 결과가 나타났다. 소득만을 고려하는 경우 41.1%의 가구에서 65세 이상 가구원이 존재하는 것으로 나타났지만, 재산을 함께 고려하면 그 비중이 31~33%로 낮아졌다. 이러한 결과는 단순히 가구주의 연령이 높을수록 재산을 가지고 있을 가능성이 크고 재산의 규모도 더 클 경향이 있기 때문에 나타난 것으로 보인다. 생애주기로 볼 때 재산은 과거 소득의 축적으로 형성되는 것이기 때문에 나타나는 당연한 결과로 볼 수 있다. 가구의 정의를 달리 한 경우 개별적인 가구의 구성은 바뀔 수 있지만 평균적인 특성은 큰 차이가 없는 것으로 나타났다.

〈표 6-10〉 0을 포함하지 않은 결합지수 중위값의 50% 미만 가구 특성: 세대보완 1인

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.459	0.465	0.463	0.459
평균 가구원 수	1.977	1.912	1.869	1.873
1인 가구 비중	0.528	0.559	0.573	0.569
35세 이하 가구주 비중	0.173	0.194	0.205	0.234
36~54세 가구주 비중	0.420	0.434	0.426	0.419
55세 이상 가구주 비중	0.407	0.372	0.369	0.347
대도시	0.502	0.502	0.504	0.524
중소도시	0.381	0.382	0.381	0.374
농어촌(도의 군)	0.116	0.116	0.115	0.102
18세 미만 가구원 유무	0.190	0.193	0.186	0.179
18세 미만 가구원 수	0.302	0.305	0.293	0.279
65세 이상 가구원 유무	0.264	0.230	0.225	0.213
65세 이상 가구원 수	0.321	0.269	0.261	0.247

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-11〉 0을 포함하지 않은 결합지수 중위값의 50% 미만 가구 특성: 세대보완 2인

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.478	0.491	0.486	0.478
평균 가구원 수	1.867	1.833	1.809	1.828
1인 가구 비중	0.525	0.561	0.572	0.566
35세 이하 가구주 비중	0.134	0.161	0.173	0.203
36~54세 가구주 비중	0.328	0.362	0.363	0.367
55세 이상 가구주 비중	0.537	0.477	0.464	0.431
대도시	0.486	0.486	0.488	0.514
중소도시	0.379	0.381	0.381	0.373
농어촌(도의 군)	0.134	0.133	0.130	0.112
18세 미만 가구원 유무	0.148	0.161	0.160	0.162
18세 미만 가구원 수	0.234	0.255	0.253	0.253
65세 이상 가구원 유무	0.393	0.332	0.319	0.293
65세 이상 가구원 수	0.500	0.400	0.382	0.352

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-12〉 0을 포함하지 않은 결합지수 중위값의 50% 미만 가구 특성: 세대보완 3인

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.481	0.494	0.489	0.479
평균 가구원 수	1.881	1.849	1.828	1.848
1인 가구 비중	0.517	0.551	0.561	0.553
35세 이하 가구주 비중	0.154	0.183	0.196	0.222
36~54세 가구주 비중	0.326	0.359	0.359	0.362
55세 이상 가구주 비중	0.520	0.458	0.445	0.416
대도시	0.491	0.491	0.494	0.519
중소도시	0.377	0.378	0.378	0.370
농어촌(도의 군)	0.132	0.131	0.128	0.110
18세 미만 가구원 유무	0.153	0.166	0.165	0.166
18세 미만 가구원 수	0.244	0.265	0.263	0.262
65세 이상 가구원 유무	0.382	0.321	0.308	0.285
65세 이상 가구원 수	0.489	0.389	0.372	0.344

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-13〉 0을 포함하지 않은 결합지수 중위값의 50% 미만 가구 특성: 세대기준안

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.482	0.495	0.490	0.482
평균 가구원 수	1.846	1.809	1.785	1.802
1인 가구 비중	0.534	0.569	0.581	0.575
35세 이하 가구주 비중	0.159	0.189	0.202	0.227
36~54세 가구주 비중	0.325	0.356	0.356	0.359
55세 이상 가구주 비중	0.517	0.455	0.443	0.414
대도시	0.491	0.492	0.494	0.519
중소도시	0.377	0.378	0.378	0.370
농어촌(도의 군)	0.132	0.130	0.128	0.111
18세 미만 가구원 유무	0.152	0.164	0.163	0.163
18세 미만 가구원 수	0.239	0.258	0.256	0.254
65세 이상 가구원 유무	0.376	0.316	0.303	0.281
65세 이상 가구원 수	0.479	0.380	0.363	0.337

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-14〉 0을 포함하지 않은 결합지수 중위값의 50% 미만 가구 특성: 건강보험증기준안

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.163	0.166	0.161	0.152
평균 가구원 수	0.480	0.488	0.488	0.495
1인 가구 비중	1.731	1.670	1.627	1.600
35세 이하 가구주 비중	0.219	0.249	0.265	0.280
36~54세 가구주 비중	0.407	0.418	0.408	0.404
55세 이상 가구주 비중	0.375	0.333	0.327	0.316
대도시	0.510	0.509	0.511	0.529
중소도시	0.383	0.385	0.384	0.377
농어촌(도의 군)	0.106	0.106	0.105	0.094
18세 미만 가구원 유무	0.163	0.166	0.161	0.152
18세 미만 가구원 수	0.251	0.256	0.246	0.230
65세 이상 가구원 유무	0.221	0.188	0.183	0.175
65세 이상 가구원 수	0.266	0.216	0.208	0.199

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-10〉에서 〈표 6-14〉까지는 결합지수가 0원인 가구를 제외한 후 결합지수의 중위값을 구하고 그것의 50% 미만인 가구를 식별하여 이들의 인구학적 특성을 정리한 것이다. 앞에서 보았듯이 모든 데이터에서 결합지수가 0원인 가구의 비중이 10% 이상으로 꽤 높게 나타난 사실을 생각해 보면, 이들 가구를 제외한 후의 특성은 상당히 다를 것으로 예상된다. 그러나 결과를 보면, 이러한 예상이 틀렸음을 알 수 있다. 예를 들어 〈표 6-5〉에서 〈표 6-9〉의 결과를 비교하면 통계값이 거의 같은 것을 알 수 있다.

〈표 6-15〉 상위 10% 가구: 세대보완 1안

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.227	0.227	0.249	0.285
평균 가구원 수	3.263	3.209	3.240	3.271
1인 가구 비중	0.201	0.185	0.173	0.163
35세 이하 가구주 비중	0.159	0.098	0.087	0.087
36~54세 가구주 비중	0.553	0.483	0.476	0.459
55세 이상 가구주 비중	0.288	0.419	0.436	0.453
대도시	0.603	0.614	0.619	0.605
중소도시	0.354	0.343	0.336	0.341
농어촌(도의 군)	0.043	0.044	0.044	0.053
18세 미만 가구원 유무	0.305	0.274	0.276	0.280
18세 미만 가구원 수	0.464	0.417	0.425	0.439
65세 이상 가구원 유무	0.334	0.423	0.451	0.482
65세 이상 가구원 수	0.488	0.630	0.678	0.730

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-16〉 상위 10% 가구: 세대보완 2안

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.225	0.229	0.254	0.303
평균 가구원 수	2.872	2.823	2.813	2.765
1인 가구 비중	0.226	0.207	0.198	0.197
35세 이하 가구주 비중	0.163	0.102	0.086	0.077
36~54세 가구주 비중	0.569	0.488	0.463	0.408
55세 이상 가구주 비중	0.268	0.410	0.451	0.515
대도시	0.599	0.608	0.611	0.595
중소도시	0.359	0.347	0.341	0.343
농어촌(도의 군)	0.042	0.045	0.047	0.062
18세 미만 가구원 유무	0.313	0.271	0.259	0.232
18세 미만 가구원 수	0.475	0.412	0.395	0.360
65세 이상 가구원 유무	0.122	0.237	0.278	0.347
65세 이상 가구원 수	0.159	0.339	0.398	0.494

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-17〉 상위 10% 가구: 세대보완 3안

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.217	0.222	0.247	0.300
평균 가구원 수	2.752	2.707	2.704	2.679
1인 가구 비중	0.255	0.235	0.225	0.216
35세 이하 가구주 비중	0.157	0.099	0.084	0.077
36~54세 가구주 비중	0.573	0.494	0.469	0.412
55세 이상 가구주 비중	0.270	0.408	0.448	0.511
대도시	0.591	0.601	0.605	0.590
중소도시	0.361	0.350	0.344	0.345
농어촌(도의 군)	0.048	0.049	0.051	0.064
18세 미만 가구원 유무	0.295	0.257	0.245	0.223
18세 미만 가구원 수	0.447	0.388	0.373	0.345
65세 이상 가구원 유무	0.123	0.234	0.275	0.343
65세 이상 가구원 수	0.160	0.333	0.391	0.487

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-18〉 상위 10% 가구: 세대기준안

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.215	0.221	0.246	0.299
평균 가구원 수	2.785	2.731	2.724	2.694
1인 가구 비중	0.236	0.222	0.215	0.210
35세 이하 가구주 비중	0.156	0.098	0.083	0.076
36~54세 가구주 비중	0.574	0.493	0.468	0.411
55세 이상 가구주 비중	0.270	0.408	0.448	0.512
대도시	0.591	0.601	0.606	0.591
중소도시	0.361	0.349	0.343	0.345
농어촌(도의 군)	0.048	0.049	0.051	0.064
18세 미만 가구원 유무	0.297	0.257	0.246	0.224
18세 미만 가구원 수	0.449	0.389	0.374	0.346
65세 이상 가구원 유무	0.124	0.235	0.276	0.345
65세 이상 가구원 수	0.162	0.336	0.394	0.491

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-19〉 상위 10% 가구: 건강보험증기준안

구분	소득 only	연금화(2)	연금화(10)	기초생활
여성 가구주 비중	0.202	0.181	0.188	0.204
평균 가구원 수	0.286	0.283	0.306	0.343
1인 가구 비중	2.132	2.175	2.255	2.406
35세 이하 가구주 비중	0.180	0.106	0.081	0.069
36~54세 가구주 비중	0.552	0.485	0.481	0.466
55세 이상 가구주 비중	0.268	0.410	0.438	0.466
대도시	0.604	0.614	0.619	0.603
중소도시	0.353	0.342	0.337	0.344
농어촌(도의 군)	0.044	0.044	0.044	0.053
18세 미만 가구원 유무	0.202	0.181	0.188	0.204
18세 미만 가구원 수	0.298	0.269	0.282	0.314
65세 이상 가구원 유무	0.241	0.337	0.367	0.405
65세 이상 가구원 수	0.327	0.471	0.519	0.583

주: 대도시= 서울, 광역시, 중소도시= 도의 시, 세종, 제주, 농어촌= 도의 군.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

재산을 고려하면 가장 크게 달라지는 부분이 저소득 가구에서와 마찬가지로 가구주 연령과 65세 이상 가구원 유무이다. 재산을 고려하면 소득으로 환산하는 방법에 크게 관계없이 55세 이상 가구주 비율이 41~51%로 높게 나타난다. 이는 소득만을 볼 경우 26~29%보다 상당히 높은 수준이다. 65세 이상 가구원이 존재할 가능성도 재산을 고려하면 가구주 비율이 소득만을 볼 경우보다 높게 나타난다. 고소득 가구의 경우 재산의 비중이 저소득 가구보다 크기 때문에 재산을 고려할지 여부에 따라 결합지수의 분포가 좀 더 민감하게 달라지는 것으로 보인다.

〈표 6-20〉에서 〈표 6-22〉에서는 데이터 내에서 결합지수 간 하위지수 가구나 상위지수 가구의 정의가 어느 정도로 정합한지를 살펴보고 있다. 〈표 6-20〉에서는 0을 포함한 경우, 중위값의 50% 미만인 하위지수 가구를 살펴보고 있다. 세대보완 3안(분리세대 추정) 데이터를 이용하여 표의 내용을 설명하면 다음과 같다. 우선, 소득만으로 식별된 중위값의 50% 미만 가구로 표본을 제한한다. 그리고 이들 가구 중 재산을 2% 이자율로 연금화한 지수에서도 중위값의 50% 미만 하위가구로 판별된 가구의 비율을 구하면 그 비율이 81.8%라는 것이다. 다음으로, 이자율을 10%로 하면 74.1%, 기초생활보장제도의 방식을 따르면 그 비율이 70.7%로 낮아진다. 비율이 낮다는 것은 두 개의 지수 사이에 정합성이 떨어진다는 것을 의미한다. 즉 〈표 6-20〉의 결과를 보면, 세대보완 3안(분리세대 추정) 데이터에서 소득만을 사용하는 경우와 재산을 고려한 경우의 지수 간 정합성은 그렇게 높지 않으며, 특히 기초생활보장제도의 방식이 가장 크게 다르다는 사실을 알 수 있다. 나아가 기초생활보장제도의 방식은 재산을 연금소득화하는 방식과도 정합성이 0.9 미만으로 낮게 나타나는 것을 알 수 있다. 이러한 특징은 가구 정의를 달리하는 경우에도 유사하게 나타난다. 데이터 내에서 지수 간에 정합성이 낮다는 사실은 가구의 평균적인 특성이

크게 다르지 않았던 사실과 대조되는 결과이다.

〈표 6-22〉에서는 상위 10% 가구를 판별하는 경우의 지수 간 적합성을 살펴보고 있는데, 하위 가구를 판별할 때보다 적합성이 더 낮게 나타나는 것을 알 수 있다. 인구총조사 가구 정의를 사용한 세대보완 3안(분리세대 추정) 데이터의 경우를 예로 보면, 소득만으로 판별된 상위 10%의 가구 중 오로지 39.1%의 가구만이 기초생활보장제도의 방식으로도 상위 10%에 속한다는 점을 볼 수 있다.

〈표 6-20〉 데이터 내 적합성: 0을 포함한 중위값의 50% 미만 가구

세대보완 1안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.877	0.803	0.762
연금화(2)	0.947	.	0.919	0.872
연금화(10)	0.893	0.946	.	0.946
기초생활	0.786	0.834	0.879	.
세대보완 2안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.814	0.736	0.700
연금화(2)	0.938	.	0.908	0.862
연금화(10)	0.887	0.949	.	0.943
기초생활	0.782	0.835	0.874	.
세대보완 3안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.818	0.741	0.707
연금화(2)	0.940	.	0.909	0.865
연금화(10)	0.889	0.950	.	0.945
기초생활	0.790	0.842	0.880	.
세대기준안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.816	0.737	0.702
연금화(2)	0.939	.	0.908	0.861
연금화(10)	0.887	0.949	.	0.943
기초생활	0.783	0.835	0.874	.
건강보험증기준안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.866	0.790	0.762
연금화(2)	0.946	.	0.916	0.883
연금화(10)	0.890	0.945	.	0.958
기초생활	0.812	0.861	0.906	.

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-21〉 데이터 내 정합성: 0을 포함하지 않은 중위값의 50% 미만 가구

세대보완 1안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.874	0.804	0.760
연금화(2)	0.963	.	0.922	0.871
연금화(10)	0.906	0.943	.	0.943
기초생활	0.748	0.778	0.823	.
세대보완 2안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.818	0.743	0.704
연금화(2)	0.971	.	0.910	0.858
연금화(10)	0.918	0.948	.	0.938
기초생활	0.770	0.790	0.829	.
세대보완 3안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.823	0.748	0.711
연금화(2)	0.973	.	0.911	0.860
연금화(10)	0.921	0.949	.	0.939
기초생활	0.779	0.797	0.836	.
세대기준안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.820	0.744	0.706
연금화(2)	0.972	.	0.910	0.857
연금화(10)	0.920	0.948	.	0.938
기초생활	0.773	0.792	0.831	.
건강보험증기준안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.866	0.792	0.759
연금화(2)	0.957	.	0.918	0.878
연금화(10)	0.898	0.942	.	0.952
기초생활	0.781	0.818	0.864	.

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 6-22〉 데이터 내 정합성: 상위 10% 이상 가구

세대보완 1안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.780	0.631	0.397
연금화(2)	0.780	.	0.846	0.608
연금화(10)	0.631	0.846	.	0.759
기초생활	0.398	0.608	0.759	.
세대보완 2안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.785	0.638	0.385
연금화(2)	0.785	.	0.849	0.591
연금화(10)	0.638	0.849	.	0.739
기초생활	0.385	0.591	0.739	.
세대보완 3안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.791	0.647	0.391
연금화(2)	0.791	.	0.852	0.591
연금화(10)	0.647	0.852	.	0.736
기초생활	0.391	0.591	0.736	.
세대기준안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.791	0.646	0.390
연금화(2)	0.791	.	0.852	0.590
연금화(10)	0.646	0.852	.	0.736
기초생활	0.390	0.591	0.736	.
건강보험증기준안	소득 only	연금화(2)	연금화(10)	기초생활
소득 only	.	0.772	0.614	0.371
연금화(2)	0.772	.	0.838	0.591
연금화(10)	0.614	0.838	.	0.749
기초생활	0.371	0.591	0.749	.

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

제4절 소결

경제 이론으로 보면, 가구의 실질적인 경제적 지불능력은 소득뿐만 아니라 재산에도 영향을 받는다. 그러나 소득은 유량(flow) 변수인 반면, 재산은 저장(stock) 변수인 관계로 두 변수를 어떻게 합산할 것인지에 대해서는 명확하게 정리되어 있지 않은바 그동안 선행연구에서도 몇 가지 방법들이 제시되어 왔다. 본 장에서는 그동안 선행연구에서 제시된, 대표적인 소득-재산 결합방식이라고 할 수 있는 연금소득화법과 소득인정액 환산법을 우리나라 건강보험공단 DB에 적용해 보았다. 이를 통해 복지 제도 수혜 대상인 하위 가구를 식별하거나 경제력의 집중을 학술적으로 분석하거나 누진세를 설계하고자 할 때 필요한 상위 가구의 식별에서 각 결합방식에 따라 식별 결과가 얼마나 달라지는지를 보기 위해 지수 간 비교와 정합성을 살펴보았다.

분석 결과, 가구의 평균적인 특성은 결합방식에 따라 크게 다르지 않음을 발견하였다. 소득만을 고려할 경우와 재산까지 고려하는 경우, 가구주의 연령이나 고령 가구원의 유무가 달라지지만, 재산을 어떤 방식으로 환산하여 소득과 합산을 할 것인지에 따라서는 평균적인 특성이 크게 변하지 않았다. 이러한 결과는 하위 가구를 식별할 때나 상위 가구를 식별할 때 모두 동일하게 나타났다.

그러나 지수 간 정합성은 그렇게 좋지는 않은 것으로 나타났다. 재산을 고려하는 경우에도 연금화 방식을 사용하는가 혹은 기초생활보장제도의 소득인정액 환산방식을 사용하는가에 따라 하위 가구나 상위 가구의 식별이 상당히 다르게 나타난다는 사실을 발견했다. 이는 학술적 연구나 정책 수립에 있어서 가구의 지불능력을 어떻게 정의할 것인가에 대해 신중하게 검토할 필요가 있다는 사실을 보여주는 단적인 예라고 할 수 있다.



제7장

소득 수준별 사망률 차이에 대한 탐색적 분석

제1절 분석 배경과 방법

제2절 가구 구성안별 소득 100분위의 적합도 비교

제3절 소득 수준에 따른 사망률 차이

제4절 소결

제 7 장

소득 수준별 사망률 차이에 대한 탐색적 분석

제1절 분석 배경과 방법

1. 분석의 배경

개인의 생명과 건강에 대한 ‘사회적 결정요인(Social determinants of health)’의 중요성이 강조되면서 이를 뒷받침하는 실증적인 연구의 중요성이 함께 주목받기 시작했다. 개인의 유병률 및 사망률을 설명함에 있어 초기에는 성별, 연령 같은 개인적 특성에 집중한 연구들이 주를 이루었으나, 시간이 지남에 따라 개인의 수요를 넘어 사회·경제적 요인, 환경적 요인, 제도적 요인 등 다차원적 요인들을 함께 고려한 사회생태학적 모형에 입각한 연구들이 주를 이루었다(송태민, 2013).

하지만 성인에서 개인의 사회경제적 특성과 사망률 간 관계를 실증적으로 분석한 연구 결과는 상대적으로 적었는데, 이는 상병 및 사망 정보와 사회경제적 수준을 개인 수준의 분석 데이터로 구축하여 분석하기가 어려웠기 때문이다.

이와 관련하여 해외에서 이루어진 일부 연구를 살펴보면, 먼저 Pekka et al.(2001)의 연구에서는 핀란드 국세청 및 사회보험국 자료와 인구센서스의 사망기록 자료를 매칭하여 30세 이상 성인 남녀를 대상으로 가구 소득 수준과 사망률의 선형관계를 파악하고자 시도하였는데, 전체적으로 모든 원인에 의한 사망률은 소득 1분위 대비 소득 10분위에서 남녀 모두 높게 나타났으며, 기타 질병이나 사고 및 상해로 인한 사망 역시 대체로 소득 수준과 밀접한 연관이 있음을 제시되었다. Sabanayagam &

Shankar(2012)의 연구에서는 미국 3기 국민건강영양조사와 국가 보건 통계센터의 사망자료를 활용하여 미국 20세 이상 성인 15,646명을 대상으로 빈곤소득비율(PIR)과 모든 원인에 의한 사망률 간의 연관성을 조사하였는데, 소득 수준이 가장 높은 4분위에서 사망률이 가장 낮고, 소득 분위가 낮아질수록 사망률이 통계적으로 유의하게 증가한다는 사실을 확인하였다. 비교적 최근 이루어진 Gleib et al.(2022)의 연구에서는 미국 코호트 자료를 활용하여 6,320명을 대상으로 18년간의 소득 수준별 사망률을 추적조사하였다. 이 연구에서는 65세 이하 연령 그룹의 경우, 40세 이상 구간부터 소득 수준이 높을수록 생존율이 높게 나타나는 뚜렷한 양상을 보였고, 65세 이상 연령 그룹에서는 소득 수준이 가장 높은 집단의 생존율이 가장 높고, 가장 낮은 집단의 생존율이 가장 낮게 나타나는, 소득 수준에 따른 사망률 양극화 현상을 보고한 바 있다.

〈표 7-1〉 사회경제적 수준과 사망률 간의 관계에 대한 실증적 연구(국외)

Martikainen et al.(2001)	
제목	Income differences in mortality: a register-based follow-up study of three million men and women
목적	핀란드 30세 이상 성인 남녀를 대상으로 가구소득 수준과 사망률의 선형관계를 파악
소득자료	핀란드 국세청 및 사회보험국(registers of the Finnish Tax Administration and the Social Insurance Institution)
소득정의	<ul style="list-style-type: none"> • 가구가처분소득(household disposable income)의 소득 10분위. • 임금, 자본소득, 과세소득 등 가구 구성원이 받는 모든 소득(세금 제외)
사망자료	인구센서스 사망기록(1991~1996)
주요결과	<ul style="list-style-type: none"> • 전체적으로 모든 원인에 의한 사망률(relative mortality ratio)은 소득 1분위 대비 10분위(소득 낮음)에서 남자는 약 2.37, 여자는 약 1.73으로 나타남. • 기타 다른 모든 질병(all disease)이나 사고나 상해(accidents and violence)로 인한 사망의 경우에도 대체로 소득 수준과 밀접한 연관이 있는 것으로 보임.

Sabanayagam & Shankar(2012)의 연구	
제목	Income Is a Stronger Predictor of Mortality than Education in a National Sample of US Adults
목적	미국 20세 이상 성인 15,646명을 대상으로 빈곤소득비율(PIR)과 모든 원인 사망률 간의 연관성을 조사
소득자료	미국 3기 국민건강영양조사(The Third National Health and Nutrition Examination Survey (NHANES III))
소득정의	빈곤소득비율의 4분위수
사망자료	국가 보건통계센터(National Center for Health Statistics (NCHS))
주요결과	<ul style="list-style-type: none"> 빈곤소득비율의 4분위(가장 높음)에 속한 대상자들의 사망률을 살펴보면, 소 4분위에서 6.6%, 3분위에서 8.1%, 2분위에서 14.9%, 1분위(가장 낮음)에서 15.0%로 나타남. 성연령을 보정한 사망률에 대한 오즈비를 살펴본 결과, 4분위(가장 높은 소득분위)를 기준으로 3분위는 1.14배, 2분위는 1.94배, 1분위는 2.86배 사망에 대한 가능성이 큰 것으로 통계적으로 유의하게 나타남.

Glei et al.(2022)의 연구	
제목	Assessment of Mortality Disparities by Wealth Relative to Other Measures of Socioeconomic Status Among US Adults
목적	소득과 기타 사회경제적 상태 지표에 따른 사망률 격차 비교
소득자료	US 코호트 자료(6,320명 대상)의 약 18년간 사망률 추적조사
사망자료	
주요결과	<ul style="list-style-type: none"> 65세 이하 연령을 비교한 결과 생존율(Probability of survival)은 40세 이상 구간에서부터 소득구간별로 차이가 나기 시작하며, 연령이 증가할수록 소득 수준이 높을수록 생존율이 높게 나타남. 65세 이상 연령을 분석한 결과도 마찬가지로 생존율은 소득수준과 비례하는 경향을 보이지만, 소득 수준이 가장 높은 집단(\$300,000 이상)의 생존율이 가장 높고, 가장 낮은 집단(\$0)의 생존율이 가장 낮게 나타나는 반면, 중간층 두 집단의 생존율의 차이는 미미하게 나타남.

국내에서 그간 이루어진 연구들을 살펴보면 다음과 같다. 강영호와 김혜린(2006)의 연구에서는 1998년 국민건강영양조사 자료와 통계청 사망확인 정보를 주민등록번호를 통해 매칭하여 각종 사회경제적 지표별 사망률의 불평등을 밝히고자 시도했다. 이들은 교육수준이 낮을수록, 육체

노동자가 비육체노동자에 비해, 임시·일용직 근로자가 상용(전일제, 시간제)에 비해, 가구 소득 수준이 낮을수록 사망위험이 높아짐을 실증적으로 분석하였다. 강민진와 손강주(2018)의 연구에서는 2006년 기준 전 국민 모집단 2%의 국민건강보험공단 표본 코호트 2.0 자료를 활용하여, 2007년 발병 당시 정보를 바탕으로 소득을 3분위로 구분하여 암, 뇌졸중, 심근경색 환자의 소득 수준에 따른 사망률을 확인하였는데, 일반적으로 저소득층에서 중·고소득층의 사망 비율이 높게 나타남과 동시에 위험비(Hazard ratio)가 낮게 나타는 것을 확인하였다. 비슷한 연구로는 국민건강보험공단 맞춤형 자료(2007~2018)를 후향적 코호트 형태로 전환하여 조기 대장암 환자의 소득 수준이 사망에 미치는 영향을 분석한 김동준(2020)의 연구가 있다. 이 연구에서는 연령별(50대, 60대, 70대) 조기 대장암 환자의 사망위험비는 저소득 그룹이 고소득 그룹에 비해 높게 나타남을 확인하였다. 비교적 최근 이루어진 강정희, 김철웅, 권순석(2022)의 연구에서는, 암센터등록 자료와 국민건강보험공단의 부과 자료를 바탕으로 소득 수준 20% 이하 그룹이 81% 이상 그룹에 비해 위암 5년 생존율, 대장암 5년 생존율의 사망위험비가 유의하게 증가함을 확인하였다.

〈표 7-2〉 사회경제적 수준과 사망률 간의 관계에 대한 실증적 연구(국내)

강영호 외(2006)의 연구	
제목	우리나라의 사회경제적 사망률 불평등: 1998년도 국민건강영양조사 자료의 사망추적 결과(2006)
목적	사회경제적 지위 지표별 사망률 불평등 실증분석
소득자료	국민건강영양조사(1998)
소득정의	<ul style="list-style-type: none"> • 가구 균등화 소득 산출 후 5분위로 구분 • 가구원 수 보정 월 가구 소득 = 월 가구 소득 / 가구원 수의 0.5승
사망자료	통계청 사망확인서비스
주요결과	<ul style="list-style-type: none"> • 교육수준이 낮을수록 사망위험은 높아짐. • 직업 표의 경우 비육체노동자에 비해 육체노동자가 2.73배 사망위험이 높음.

	<ul style="list-style-type: none"> • 종사상지위의 경우 상용근로자(전일제/시간제)에 비해 임시 및 일용근로자의 사망위험이 3.01배 높음. • 가구원 수 보정 가구소득 상위 20%를 기준으로 할 때, 20~40%와 40~60%에 해당하는 중간층 사망위험은 1.64배 높고, 60~80%와 80~100%(최하층) 사망비는 각각 2.3배, 2.29배 높게 나타남. • 최저 가구소득(월 50만 원 이하)의 경우 월 가구소득이 250만 원 이상인 사람들에 비해 사망위험이 2.37배 높게 나타남.
--	---

강민진 외(2018)의 연구

제목	암, 심뇌혈관 질환자의 소득 수준에 따른 사망 및 이차 질환 발생 위험
목적	암, 뇌졸중, 심근경색 환자의 소득 수준이 사망에 미치는 영향 확인
소득자료	국민건강보험공단 표본 코호트 2.0(2006년 기준 전 국민 모집단의 2%, 성, 연령, 가입자 구분, 보험료 분위, 지역별 층화 추출 100만 명, 2002년 1월 1일~2015년 12월 31일까지 14년간 사회경제적 현황, 의료현황 파악)
사망자료	
소득정의	2007 발병 당시 기준 3분위(low, mid, high) 구분 (의료급여대상자 low그룹)
주요결과	<ul style="list-style-type: none"> • 관찰기간 내 2,684(43.3%)명의 사망자가 발생했고, 소득 수준별로는 저소득층 1,672명 중 832명(49.3%), 중위층 1,943명 중 787명(40.5%), 고소득층 2,578명 중 1,065(41.3%)명의 사망자가 발생함. • 전체 대상으로 콕스 회귀분석을 통해 Hazard Ratio(HR)를 계산한 결과, 저소득층보다 중위층과 고소득층이 더 낮은 위험률을 보였음. • 연령을 보정한 결과, 65세 미만에서는 저소득층을 기준으로 중위층과 고소득층의 사망위험률이 통계적으로 더 유의하게 낮게 나타났으나, 65세 이상군의 경우 저소득층 대비 고소득층만이 유의하게 더 낮게 나타남.

김동준(2020)의 연구

제목	우리나라 조기 대장암 환자의 소득 수준이 사망에 미치는 영향
목적	조기 대장암 환자의 소득 수준이 사망에 미치는 영향 분석
소득자료	건강보험공단 맞춤형 자료(2007~2018)의 후향적 코호트 전환. 2007년부터 2018년까지 국민건강보험공단에 청구된 우리나라 전 국민 진료내역 중 '대장암'을 주진단으로 진단받은 청구자료로 분석 →건별로 되어 있는 건강보험자료의 한계점을 극복하여, 환자별 정보로 의료행태 전 주기적 파악 가능
사망자료	
소득정의	<ul style="list-style-type: none"> • 이 연구에서 독립변수는 소득 수준으로, 연구대상자가 처음 대장암 진단을 받았을 당시의 건강보험 소득분위를 바탕으로 기초생활수급자(소득분위 0과 결측)와 5분위수로 분류 • 5분위수는 소득분위가 1~4분위인 경우 20%ile 미만, 소득분위가 5~8인 경우 20%ile 이상~40%ile 미만, 소득분위가 9~12인 경우 40%ile 이상~60%ile 미만, 소득분위가 13~16인 경우 60%ile 이상~80%ile 미만, 소득분위가 17~20인 경우 80%ile 이상으로 분류

202 국민건강보험공단 행정자료의 사회정책적 활용을 위한 탐색적 연구

주요결과	<ul style="list-style-type: none"> • 소득 수준별 대장암 환자의 사망위험비는 80%ile 이상군에 비해 60%ile 이상~80%ile 미만군에서 1.09배(CI=1.04-1.14), 40%ile 이상~60%ile 미만군에서 1.19배(CI=1.13-1.25), 20%ile 이상~40%ile 미만군에서 1.25배(CI=1.19-1.31), 20%ile 미만군에서 1.25배(CI=1.19-1.31) 높음. 다만, 80%ile 이상군에 비해 기초생활수급자군에서 사망위험비가 1.1배 높았지만 통계적으로 유의하지 않음. • 소득 수준별 50세 미만인 조기 대장암 환자의 사망위험비는 80%ile 이상군에 비해 60%ile 이상~80%ile 미만군에서 1.53배(CI=1.28-1.83), 40%ile 이상~60%ile 미만군에서 1.85배(CI=1.54-2.23), 20%ile 이상~40%ile 미만군에서 2.06배(CI=1.70-2.49), 20%ile 미만 군에서 2.07배(CI=1.69-2.53) 높았음. • 50세 이상~60세 미만인 대장암 환자의 사망위험비는 80%ile 이상군에 비해 60%ile 이상~ 80%ile 미만군에서 1.25배(CI=1.11-1.41), 40%ile 이상~60%ile 미만군에서 1.37배(CI=1.21-1.55), 20%ile 이상~40%ile 미만군에서 1.42배(CI=1.26-1.61), 20%ile 미만군에서 1.59배(CI=1.40-1.81) 높았음. • 60세 이상~70세 미만인 대장암 환자의 사망위험비는 80%ile 이상군에 비해 60%ile 이상~80%ile 미만군에서 1.25배(CI=1.14-1.37), 40%ile 이상~60%ile 미만군에서 1.28배(CI=1.16-1.42), 20%ile 이상~40%ile 미만군에서 1.40배(CI=1.26-1.55), 20%ile 미만군에서 1.43배 (CI=1.30-1.50) 높았음. • 70세 이상인 대장암 환자의 사망위험비는 80%ile 이상군에 비해 40%ile 이상~60%ile 미만군에서 1.08배(CI=1.01-1.16), 20%ile 이상~40%ile 미만군에서 1.10배(CI=1.02-1.19), 20%ile 미만군에서 1.08배(CI=1.01-1.15) 높았음.
------	--

강정희 외(2022)의 연구	
제목	국가 암검진 사업의 주요 암종별 5년 생존율과 사회경제적 수준 및 요약병기의 관련성: 광주, 전남 지역 암등록본부 자료를 중심으로.
목적	위암, 대장암, 유방암, 자궁경부암 환자를 대상으로 요약병기를 보정한 상태에서 사회경제적 수준이 5년 생존율에 영향을 주는 요인을 확인
소득자료	국민건강보험공단 보험료 부과 자료
소득정의	<ul style="list-style-type: none"> • 국민건강보험공단 보험료 부과자료를 대리변수로 사용 • 환자들의 암 진단 당시 월 보험료를 기준으로 5개 그룹과 의료급여 대상자로 구분 • 월 보험료 0원(의료급여 대상자), 20% 이하, 21~40%, 41~60%, 61~80%, 81% 이상
사망자료	광주·전남 암등록센터 등록자료(2005~2007)
주요결과	<ul style="list-style-type: none"> • 위암 5년 생존율: 의료급여자는 81% 이상 소득자에 비해 사망위험비가 1.39배 유의하게 증가하였으며, 20% 이하 소득자는 81% 이상 소득자에 비해 사망비가 1.2배 유의하게 증가함. • 대장암 5년 생존율: 의료급여자는 81% 이상 소득자에 비해 사망위험비가 1.05배 유의하게 증가하였음. • 유방암과 자궁경부암의 경우 소득은 독립변수로서 유의한 효과 없음.

한국의 경우를 살펴보면, 개인 단위로 사회경제적 수준과 사망 자료의 연계가 어려웠던 초기에는 국민건강영양조사 데이터를 활용하였으며, 비교적 최근에는 국민건강보험공단 자료를 사용하여 분석하였지만, 일부 질환군에 한정하여 사망위험을 보고하는 양상을 보였다.

전 국민 건강보험 자료를 활용하여 개인 수준에서 소득과 사망의 관계를 탐색하고 의미를 해석하는 것은 구축된 자료의 질 제고와 활용 가능성을 제고하는 측면에서 의미가 있다. 특히, 빈곤과 질병의 악순환 트랩에서 소득 격차가 건강 격차로 이어지는 관련성에 대한 탐색은 건강 불평등 해소 정책의 실증적 근거 기반으로써 건강보험 자료 활용의 기회를 확대할 기회를 제공할 것이다.

2. 분석 방법

가. 분석 모형

이 연구를 목적으로 설정된 다섯 가지의 가구 구성안별로 구축된 전체 개인 단위 DB(맞춤형 소득-재산 DB) 중 2020년 DB를 활용하였다. 전체 가입자는 가구 구성안에 따라 5개의 가구 유형에 할당되고, 가구 단위로 합산된 소득을 가구원 수의 제곱근으로 나눈 균등화 소득을 갖게 된다. 모든 개인은 균등화 소득의 분포에서 100분위수(100percentiles)를 갖게 하였으며, 통계청의 사망일자를 연계하여 해당 연도에 사망했는지를 구분하였다.

분석을 위해 가구원 수가 10인 이상인 경우와 30세 미만은 제외하였다. 전자는 10인 이상의 가구는 시설로 간주하였기 때문이며, 소득과 사망의 관계 분석에서 부모 소득 등의 영향을 배제하는 등 분석의 신뢰도를 높이기

위하여 30세 이상 성인으로 대상을 좁혔다.

1) 소득의 설명력 비교

이 연구의 목적은 건강보험 행정자료를 활용하여 의미 있는 가구의 정의를 구성안별로 탐색하는 것에 있으므로, 첫 번째 분석은 일반적으로 확인된 소득과 사망의 관계를 활용하여 어떤 가구 구성안의 소득 분포가 사망률의 분포를 잘 설명하는지 비교하고자 하였다.

2020년 균등화 소득의 100분위별로 사망률을 구하고, 사망률과 균등화 소득 100분위의 선형관계에서 사망률에 대한 구성안별 균등화 소득 100분위의 설명력을 비교하였다.

2) 소득에 따른 사망률의 차이 탐색

2000년 단일연도에서 가구 구성안에 따른 5개 모형을 구분하고, 분석 대상은 30세 이상 성인, 10인 미만 가구의 개인으로 정의하였다. 이번 연구에서는 사망에 영향을 주는 사전적 요인, 즉 건강상태 등이 반영되지 못했으므로 향후 추가적인 분석을 위한 탐색적 목적으로 시행하였다.

소득은 총 100분위로 구분된 자료를 0이 있는 분위를 별도로 분리하여 최종 10분위로 구분하였으며, 소득 10분위에서 사망률의 차이를 확인하기 위해 포아송 회귀분석(Poisson regression)을 통해 상대사망비(Relative all-cause mortality ratio)를 구별하였다.

나. 소득과 사망의 정의

사망 불평등과 관련된 소득의 정의는 단위와 반영 범위에 따라 총 4개 (개인/가구 소득, 총소득/가처분소득)의 범주로 구분할 수 있다. 사망이 발생하는 것은 개인 단위이지만 사망에 영향을 주는 경제적 수준은 단순히 개인의 소득만을 포함하지 않기 때문에, 앞에서 검토한 연구들에서도 가구소득을 개인화하여 사용하고 있다. 소비 공동체로서 가구 단위 소득은 많은 연구에서 사용하고 있는데, 구성원의 모든 소득을 반영할 뿐 아니라, 구성원의 증가로 인한 규모의 경제를 반영하고자 하기 때문이다. 또한 소득도 총소득을 사용하기보다는 기본적인 생계 유지 비용을 제외한 가처분소득을 사용하여 분석하고 있다.

이 연구는 건강보험공단이 보유한 소득정보를 활용하여 전체 개인을 의미 있는 가구로 정의하고 이를 통해 개인화한 소득정보를 활용하여 사망률의 차이가 있는지 확인하고자 하는 것이다.

이 연구는 가구소득을 개인화(가구합산소득을 가구원 수의 제곱근으로 나눈 가구 균등화 소득)하여 개인 단위로 소득에 따른 사망률의 차이를 비교하였으며, 이를 통해 가구 구성안의 적절성을 검토하였다. 다만, 이 연구에서 사용한 소득은 국세청에 신고된 종합소득 정보의 총소득을 그대로 사용하였기 때문에 가처분소득을 사용하지 않은 것은 제한점이다.

이 연구에서 사망은 모든 사인의 사망이며, 건강보험공단이 통계청으로부터 제공 받는 사망자 정보를 이용하여 연초 생존자를 기준으로 작성된 연도별 DB에서 해당 연도 내 사망 발생 여부를 확인하였다.

제2절 가구 구성안별 소득 100분위의 적합도 비교

1. 분석대상

소득자료 연결의 완결성을 고려한 가장 최근의 자료로 2020년 자료만을 사용하였다. 가구 정의에 따라 소득에 따른 사망률 차이 분석에 포함된 30세 이상 개인은 다음 표와 같다. 구성안별로 가구를 구성하는 로직은 앞의 관련 내용에서 자세히 다루었으므로 생략한다.

〈표 7-3〉 가구 구성안별 분석 대상자(2020년, 30세 이상)

가구 구성안	남자(명)	여자(명)	계(명)
세대보완 1안(건강보험피 부양자 포함)	17,801,539	18,500,421	36,301,960
세대보완 2안(비동거 가족 포함)	17,811,737	18,511,989	36,323,726
세대보완 4안(분리세대 추정)	17,813,329	18,513,452	36,326,781
세대기준안	17,795,965	18,494,674	36,290,639
건강보험증기준안	17,816,197	18,516,696	36,332,893

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

2. 분석 결과

가구 균등화 개인소득의 100분위 별로 천 명당 사망률(사망자 수/대상자 수×1,000)을 구하고 사망률과 균등화 소득 100분위의 선형모형에서 소득 100분위의 설명력을 구성안별로 비교하였다. 분석에서는 100분위 중 소득이 실제 0보다 큰 수를 갖는 분위를 구분하였다. 예를 들면, **세대보완 1안(건강보험 피부양자 포함)**의 경우, 8분위까지 소득이 0이고 소득이 발생하는 경우는 9분위부터이므로, 0분위와 9분위부터 100분위까지가 반영되는 것이다. 같은 방식으로 **세대보완 2안(비동거 가족 포함)**은 0분위

와 11분위부터 100분위까지, **세대보완 3안(분리세대 추정)**은 0분위와 12분위부터 100분위까지, **세대기준안**은 0분위와 12분위부터 100분위까지, **건강보험증기준안**은 0분위와 11분위부터 100분위까지 포함되어 있다. 소득분위는 연속변수로 처리하였다. 구성안별로 분석모형은 소득분위만 포함한 경우, 소득분위에 성별을 추가한 경우, 소득분위에 성별과 연령군(그룹 1: 30~39, 그룹 2: 40~49, 그룹 3: 50~64, 그룹 4: 65~74, 그룹 5: 75~84, 그룹 6: 85+)을 추가한 경우로 구분하여 소득분위의 회귀계수와 설명력의 변화를 비교하였다.

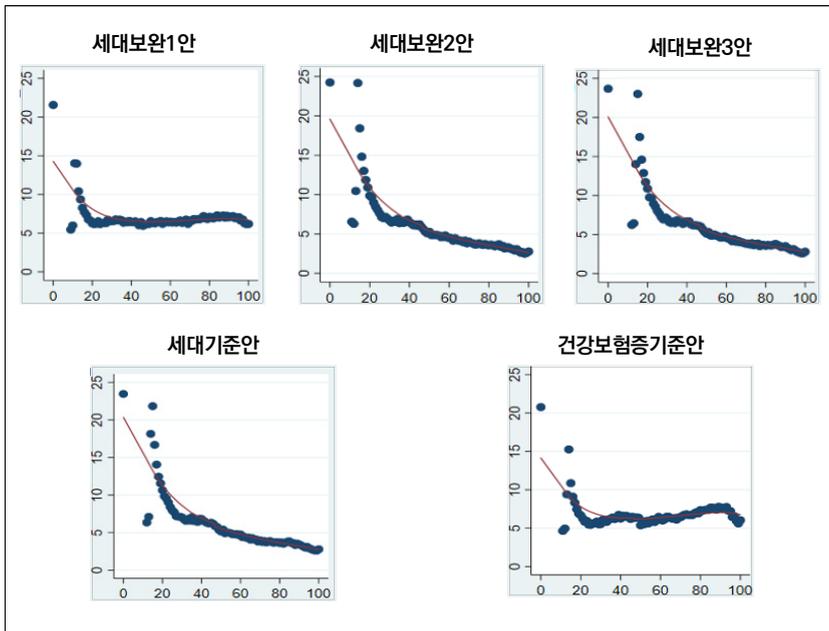
〈표 7-4〉 가구 구성안별 사망률에 대한 선형모형에서 가구 균등화 개인소득 100분위의 회귀계수와 설명력(2020년, 30세 이상)

가구 구성안별 모형		남자	
		회귀계수+	수정 R ²
세대보완 1안 (건강보험 피부양자 포함)	(a) 보정 안 함	-0.0206***	7.09%
	(a)+성별	-0.0213**	21.7%
	(a)+성별+연령군	-0.0650*	64.2%
세대보완 2안 (비동거 가족 포함)	(a) 보정 안 함	-0.1100***	56.6%
	(a)+성별	-0.1128***	50.0%
	(a)+성별+연령군	-0.0190	62.7%
세대보완 3안 (분리세대 추정)	(a) 보정 안 함	-0.11105***	58.5%
	(a)+성별	-0.11395***	51.4%
	(a)+성별+연령군	-0.0184	62.7%
세대기준안	(a) 보정 안 함	-0.1122***	59.9%
	(a)+성별	-0.1150***	52.3%
	(a)+성별+연령군	-0.0192	62.7%
건강보험증기준안	(a) 보정 안 함	-0.0116	1.4%
	(a)+성별	-0.0135*	17.8%
	(a)+성별+연령군	-0.0648*	64.1%

주: *: Pr <0.05, **: Pr <0.001, ***: Pr <0.0001
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

사망률에 대한 소득분위의 설명력은 건강보험증기준안에서 가장 낮았다(1.4%). 그다음은 세대보완 1안(건강보험 피부양자 포함)에서 7.09%로 낮았다. 세대보완 2안(비동거 가족 포함), 세대보완 3안(분리세대 추정), 세대기준안은 유사한 수준을 보였는데, 소득분위만 포함된 모형의 설명력은 세대기준안이 가장 높았다(59.9%). 세 개의 안은 연령군이 포함되면서 모두 통계적 유의성이 사라지는데, 이는 사망에 있어서 연령의 영향이 상당히 커서, 포함되기 전 소득과 사망 간의 관련성을 약화시키거나 사라지게 하는 기존의 연구와 동일한 결과이다(Pekka M et al., 2001).

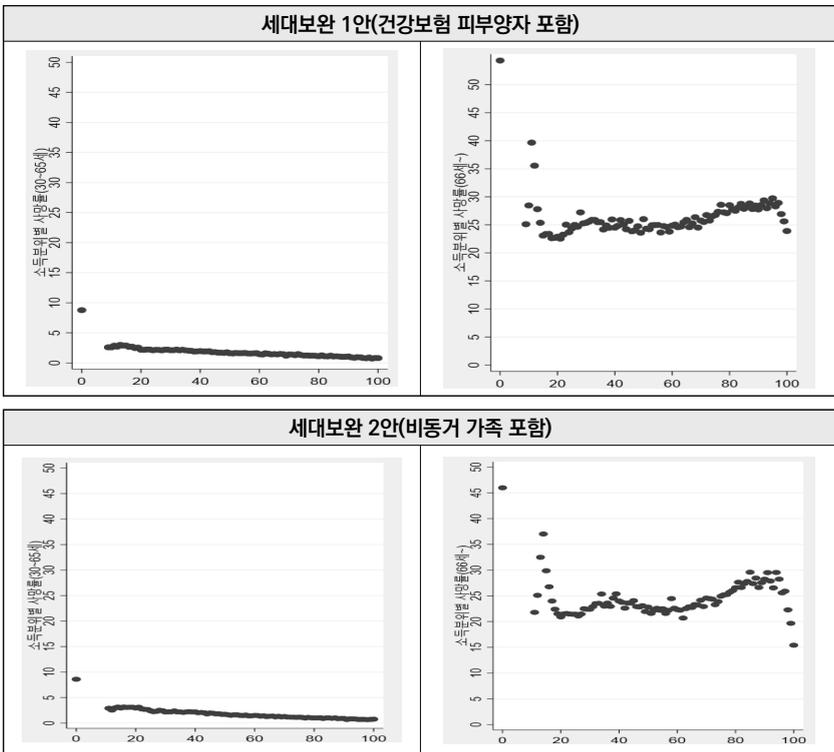
[그림 7-1] 가구 구성안별 가구 균등화 개인소득 100분위에 따른 사망률의 변화

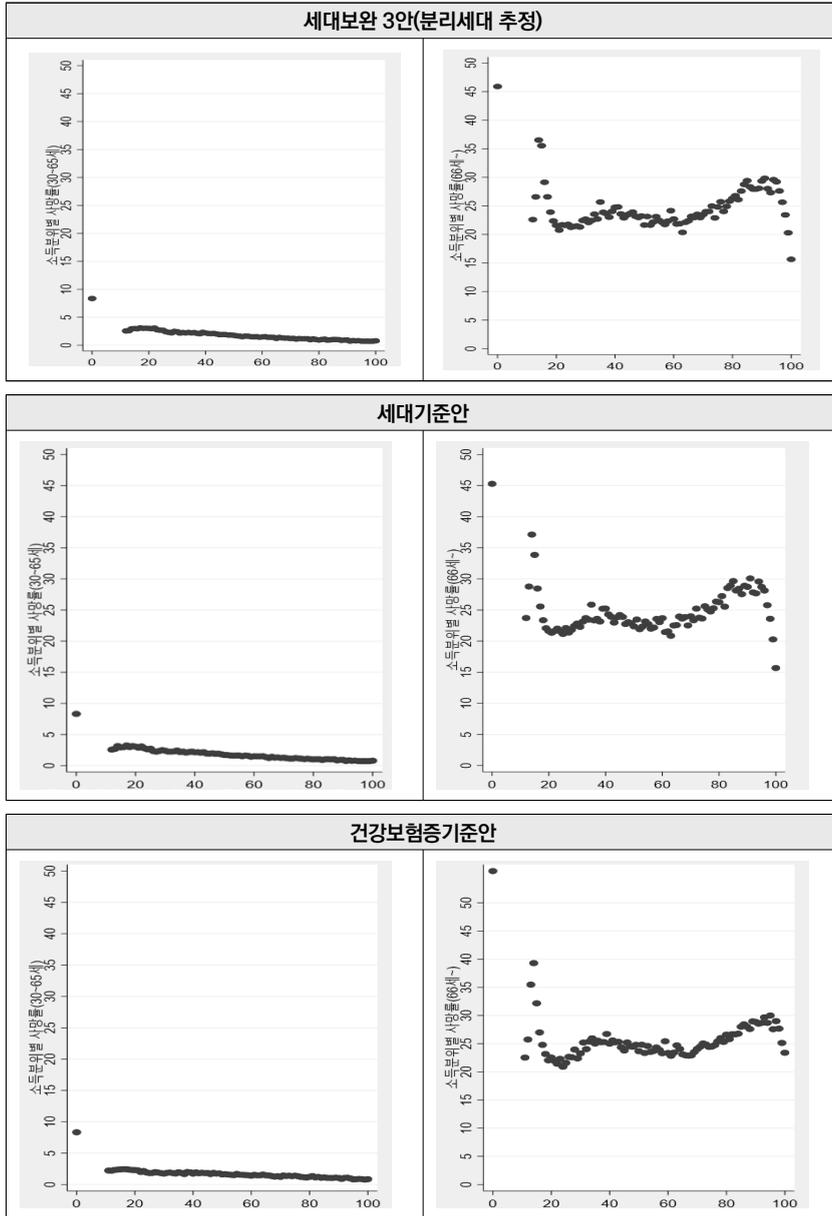


자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

노인 인구에서는 소득이 증가할수록 사망률이 감소하는 음의 관계가 약해지거나 사라질 수 있다. 이는 노쇠의 영향과 기존의 건강수준이 미치는 영향이 더 크게 작용하기 때문이다. 분석 자료에서도 65세 이하와 66세 이상으로 대상을 구분하여 비교하면, 65세 이하에서는 전반적으로 음의 관련성을 보이지만, 66세 이상에서는 음의 관련성이 사라진다. 다만, 최상위 소득분위에서는 사망률이 큰 폭으로 감소하는데 그 구간이 세대보완 2안(비동거 가족 포함), 세대보완 3안(분리세대 추정), 세대기준안에서 더 길게 떨어진다.

[그림 7-2] 가구 구성안과 연령구간별 소득 100분위에 따른 사망률의 변화





자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

66세 이상은 대부분 경제활동을 하지 않기 때문에 소득과 사망의 음의 관련성이 발생할 개연성이 사라지고 소득 외의 다른 요인들의 영향을 많이 받기 때문에 이들에 대해서는 향후 자산 등 보정된 효과를 확인할 필요가 있다.

제3절 소득 수준에 따른 사망률 차이

1. 분석대상

2000년 단일연도에서, 가구 구성안에 따른 5개의 DB에 대하여 반복적으로 소득에 따른 사망률 차이를 비교하는 것은 소득 수준과 사망률의 관련성에 대한 신뢰도를 확인하는 측면에서 의미가 있다. 향후, 적절한 가구 정의에 대한 검토가 완료된 이후, 세부적인 시계열적 분석을 진행할 필요가 있다.

소득 10분위에서 사망률의 차이가 있는지 검토하기 위하여 포아송 회귀 분석(poisson regression)을 통해 상대사망률비(relative all-cause mortality ratio)를 구하였다. 이를 위해 2020년 가구 균등화 소득 100분위 자료를 0이 있는 분위를 별도로 분리하여 모두 10개의 분위로 다음과 같이 재구성하였다. 소득 10분위 중 2분위의 범위가 가구 구성안별로 차이가 있다.

212 국민건강보험공단 행정자료의 사회정책적 활용을 위한 탐색적 연구

〈표 7-5〉 가구 구성안별 소득분위별 대상자 분포

가구 구성안	소득 10분위	소득 100분위	남자(명)	여자(명)	계(명)
세대보완 1안 (건강보험 피부양자 포함)	1: 소득 없음	0: 소득 없음	1728690	1754898	3483588
	2	9~20	2090012	2215423	4305435
	3	21~30	1683892	1793463	3477355
	4	31~40	1679684	1833756	3513440
	5	41~50	1721201	1825097	3546298
	6	51~60	1762441	1793042	3555483
	7	61~70	1787426	1780020	3567446
	8	71~80	1789840	1795989	3585829
	9	81~90	1779200	1833781	3612981
	10	91~100	1779153	1874952	3654105
세대보완 2안 (비동거 가족 포함)	1: 소득 없음	0: 소득 없음	2020790	2633635	4654425
	2	11~20	1780479	2045870	3826349
	3	21~30	1840917	1927625	3768542
	4	31~40	1698384	1817528	3515912
	5	41~50	1714513	1801170	3515683
	6	51~60	1719693	1720599	3440292
	7	61~70	1753630	1650410	3404040
	8	71~80	1760834	1615247	3376081
	9	81~90	1754949	1624613	3379562
	10	91~100	1767548	1675292	3442840
세대보완 3안 (분리세대 추정)	1: 소득 없음	0: 소득 없음	2061698	2711319	4773017
	2	12~20	1591706	1869521	3461227
	3	21~30	1833219	1946675	3779894
	4	31~40	1686751	1819379	3506130
	5	41~50	1709229	1814138	3523367
	6	51~60	1723657	1738816	3462473
	7	61~70	1765772	1665618	3431390
	8	71~80	1788149	1628277	3416426
	9	81~90	1801095	1638668	3439763
	10	91~100	1852053	1681041	3533094
세대기준안	1: 소득 없음	0: 소득 없음	2012755	2653314	4666069
	2	12~20	1640134	1921778	3561912
	3	21~30	1820399	1930005	3750404

가구 구성안	소득 10분위	소득 100분위	남자(명)	여자(명)	계(명)
	4	31~40	1682145	1815137	3497282
	5	41~50	1711820	1810183	3522003
	6	51~60	1728153	1736442	3464595
	7	61~70	1765972	1667710	3433682
	8	71~80	1783986	1632967	3416953
	9	81~90	1798088	1643479	3441567
	10	91~100	1852513	1683659	3536172
건강보험증기준안	1: 소득 없음	0: 소득 없음	2022638	2093616	4116254
	2	11~20	1699325	1932533	3631858
	3	21~30	1673971	1922776	3596747
	4	31~40	1549683	1907538	3457221
	5	41~50	1587268	1944242	3531510
	6	51~60	1639313	1905322	3544635
	7	61~70	1782751	1752836	3535587
	8	71~80	1883420	1663247	3546667
	9	81~90	1933902	1658727	3592629

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

2. 분석 결과

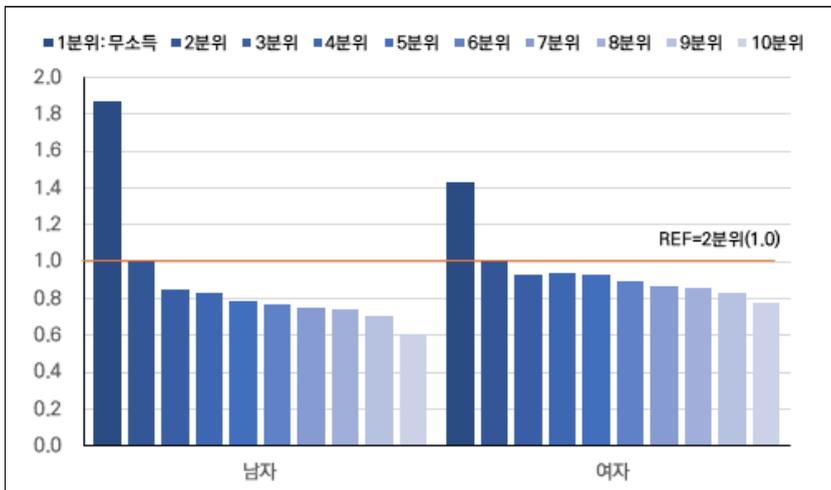
세대보완 1안(건강보험 피부양자 포함)에서 소득이 있는 최저 분위인 2분위를 기준으로 사망률을 비교하면, 소득 2분위 대비 최고 10분위의 사망률은 남자 0.61배, 여자 0.78배로 소득분위가 증가할수록 사망률의 상대비는 감소하는 선형의 관계를 보였다.

〈표 7-6〉 세대보완 1인(건강보험 피부양자 포함)에서 소득 10분위 연령보정 상대사망률비 (RR)¹

소득 10분위	남자		여자	
	RR ^a	95%CI	RR ^a	95%CI
1분위: 무소득	1.87	1.84-1.90	1.43	1.40-1.46
2분위	1.00		1.00	
3분위	0.85	0.83-0.87	0.93	0.91-0.96
4분위	0.83	0.81-0.84	0.94	0.92-0.97
5분위	0.79	0.77-0.81	0.93	0.91-0.96
6분위	0.77	0.76-0.79	0.89	0.87-0.92
7분위	0.75	0.74-0.77	0.87	0.84-0.89
8분위	0.74	0.72-0.76	0.86	0.84-0.88
9분위	0.71	0.69-0.72	0.83	0.81-0.85
10분위	0.61	0.60-0.63	0.78	0.76-0.79

주: Age adjusted relative all-cause mortality ratio. 6개 연령 그룹(40세 미만, 50세 미만, 65세 미만, 75세 미만, 85세 미만, 85세 이상)으로 보정함.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〔그림 7-3〕 세대보완 1인(건강보험 피부양자 포함)에서 소득 10분위별 연령보정 상대 사망률비(RR)



자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

세대보완 2안(비동거 가족 포함)에서 소득이 있는 최저 분위인 2분위를 기준으로 사망률을 비교하면, 소득 2분위 대비 최고 10분위의 사망률은 남자 0.37배, 여자 0.93배로 소득분위가 증가할수록 사망률의 상대비는 감소하는 선형의 관계를 보였다. 그러나 세대보완 1안(건강보험 피부양자 포함)에 비해 여성에서는 음의 관계가 사라지고 최상 소득분위에서만 감소를 보였다.

세대보완 3안(분리세대 추정)과 세대기준안에서는 모두 같은 경향을 보였다.

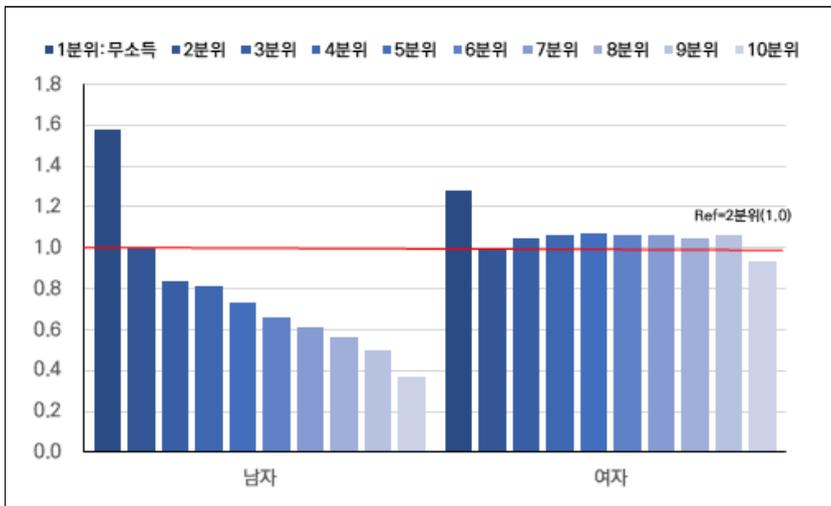
반면, 건강보험증기준안에서는 세대보완 1안(건강보험 피부양자 포함)보다 남자는 음의 관련성이 더 강해지고, 여자는 더 완화되며 동일한 경향을 보였다.

〈표 7-7〉 세대보완 2인(비동거 가족 포함)에서 소득 10분위 연령보정 상대사망률비(RR)1

소득 10분위	남자		여자	
	RR ^a	95%CI	RR ^a	95%CI
1분위	1.58	1.56-1.61	1.28	1.25-1.30
2분위	1.00		1.00	
3분위	0.84	0.83-0.86	1.05	1.03-1.08
4분위	0.81	0.79-0.83	1.06	1.04-1.09
5분위	0.73	0.72-0.75	1.07	1.04-1.10
6분위	0.66	0.64-0.68	1.06	1.03-1.09
7분위	0.61	0.60-0.63	1.06	1.03-1.09
8분위	0.56	0.54-0.57	1.05	1.02-1.08
9분위	0.50	0.49-0.51	1.06	1.03-1.09
10분위	0.37	0.36-0.38	0.93	0.90-0.95

주: Age adjusted relative all-cause mortality ratio. 6개 연령 그룹(40세 미만, 50세 미만, 65세 미만, 75세 미만, 85세 미만, 85세 이상)으로 보정함.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〔그림 7-4〕 세대보완 2인(비동거 가족 포함)에서 소득 10분위 연령보정 상대사망률비(RR)



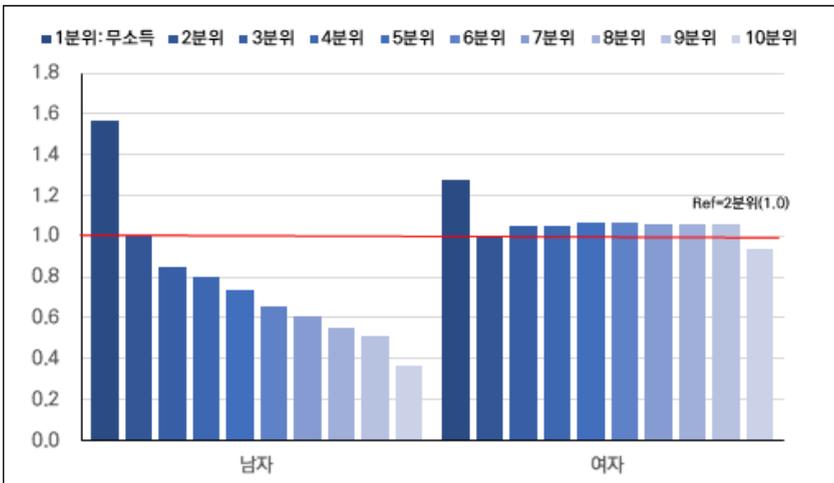
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 7-8〉 세대보완 3인(분리세대 추정)에서 소득 10분위 연령보정 상대사망률비(RR)¹

소득 10분위	남자		여자	
	RR ^a	95%CI	RR ^a	95%CI
1분위	1.57	1.55-1.59	1.28	1.26-1.30
2분위	1.00		1.00	
3분위	0.85	0.83-0.87	1.05	1.03-1.07
4분위	0.80	0.79-0.82	1.05	1.03-1.08
5분위	0.74	0.72-0.76	1.07	1.04-1.10
6분위	0.66	0.64-0.67	1.07	1.05-1.10
7분위	0.61	0.59-0.62	1.06	1.03-1.09
8분위	0.55	0.54-0.57	1.06	1.03-1.09
9분위	0.51	0.50-0.53	1.06	1.03-1.09
10분위	0.37	0.36-0.38	0.94	0.91-0.97

주: Age adjusted relative all-cause mortality ratio. 6개 연령 그룹(40세 미만, 50세 미만, 65세 미만, 75세 미만, 85세 미만, 85세 이상)으로 보정함.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〔그림 7-5〕 세대보완 3인(분리세대 추정)에서 소득 10분위 연령보정 상대사망률비(RR)



자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

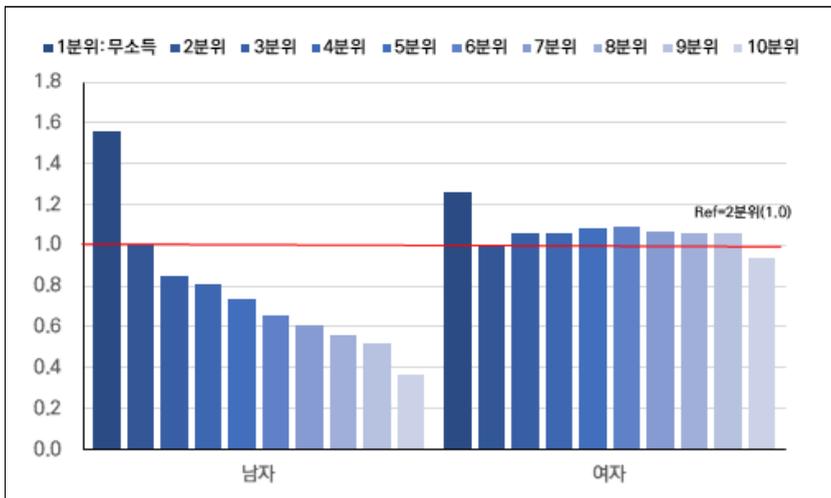
〈표 7-9〉 세대기준안에서 소득 10분위 연령보정 상대사망률비(RR)1

소득 10분위	남자		여자	
	RR ^a	95%CI	RR ^a	95%CI
1분위	1.56	1.54-1.58	1.26	1.23-1.28
2분위	1.00		1.00	
3분위	0.85	0.83-0.87	1.06	1.04-1.09
4분위	0.81	0.80-0.83	1.06	1.04-1.09
5분위	0.74	0.72-0.76	1.08	1.05-1.11
6분위	0.66	0.65-0.68	1.09	1.06-1.12
7분위	0.61	0.59-0.62	1.07	1.04-1.10
8분위	0.56	0.54-0.57	1.06	1.03-1.09
9분위	0.52	0.50-0.53	1.06	1.03-1.09
10분위	0.37	0.36-0.38	0.94	0.91-0.97

주: Age adjusted relative all-cause mortality ratio. 6개 연령 그룹(40세 미만, 50세 미만, 65세 미만, 75세 미만, 85세 미만, 85세 이상)으로 보정함.

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〔그림 7-6〕 세대기준안에서 소득 10분위 연령보정 상대사망률비(RR)



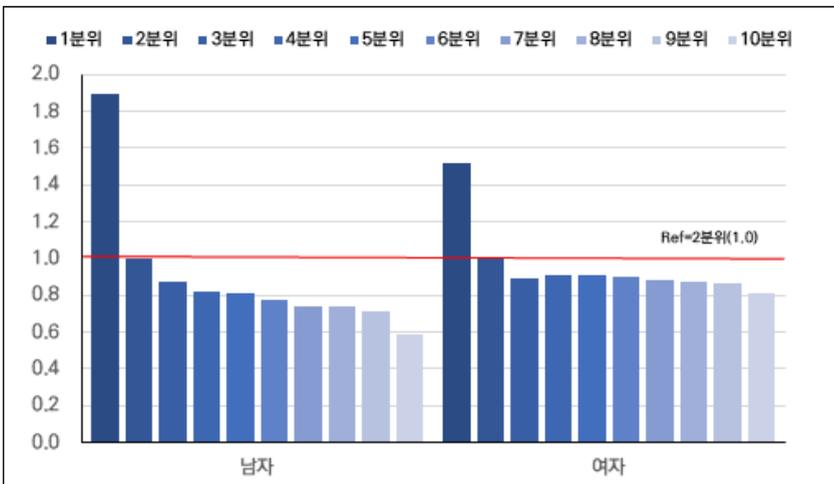
자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〈표 7-10〉 건강보험증기준안에서 소득 10분위별 연령보정 상대사망률비(RR)1

소득 10분위	남자		여자	
	RR ^a	95%CI	RR ^a	95%CI
1분위	1.89	1.86-1.92	1.52	1.49-1.55
2분위	1.00		1.00	
3분위	0.87	0.78-0.81	0.89	0.87-0.92
4분위	0.82	0.80-0.84	0.91	0.89-0.94
5분위	0.81	0.79-0.83	0.91	0.88-0.93
6분위	0.77	0.75-0.79	0.90	0.87-0.92
7분위	0.74	0.72-0.75	0.88	0.86-0.91
8분위	0.74	0.72-0.75	0.87	0.85-0.89
9분위	0.71	0.70-0.73	0.86	0.84-0.88
10분위	0.59	0.58-0.61	0.81	0.79-0.83

주: Age adjusted relative all-cause mortality ratio. 6개 연령 그룹(40세 미만, 50세 미만, 65세 미만, 75세 미만, 85세 미만, 85세 이상)으로 보정함.
 자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

〔그림 7-7〕 건강보험증기준안에서 소득 10분위 연령보정 상대사망률비(RR)



자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

제4절 소결

사회경제적 상태와 성인의 유병률 및 사망률 간 관계에 관한 관심은 높았지만, 상대적으로 개인 수준의 분석 데이터 구축의 한계로 실증분석 결과를 보고하는 연구는 많지 않았다. 한국의 경우는 대부분 국민건강영양조사 데이터를 활용하고 있으며, 일부 질환군에 한정하여 사망위험을 보고하는 연구가 대부분이다. 최근 양질의 행정데이터 구축을 활용하여 개인 수준에서 소득과 사망의 관계를 탐색하는 연구가 증가하고 있다 (Shi et al., 2021).

전 국민 건강보험 자료를 활용하여 개인 수준에서 소득과 사망의 관계를 탐색하고 의미를 해석하는 것은 구축된 자료의 질 제고와 활용 가능성을 제고하는 측면에서 의미가 있다. 특히, 빈곤과 질병의 악순환 트랩에서 소득 격차가 건강 격차로 이어지는 관련성에 대한 탐색은 건강 불평등 해소 정책의 실증적 근거 기반으로써 건강보험 자료를 활용하는 기회를 확대할 것이다.

이 연구는 향후, 소득 수준에 따른 건강 결과를 비교하기 위한 신뢰도 있는 데이터 구축을 위한 첫 단계로서, 건강보험 행정자료에서 가구를 정의하는 구성안별 모형을 검토하는 데 주요한 목적을 두고 있다. 개인 단위 건강 결과 비교에서 소득 수준은 중요한 영향요인이다. 개인 수준에서 건강에 대한 사회경제적 요인의 영향은 생활공동체의 경제적 수준에 영향을 받는다. 따라서 대부분 선행연구는 가구합산소득을 가구원 수의 제곱근으로 나눈 균등화 소득을 개인소득으로 반영하여 건강 결과에 대한 영향을 조정하거나 직접적인 관련성을 분석하고 있다.

이 연구의 목적을 위해, 이 장에서는 두 가지 측면에서 분석을 시행하였다. 첫째, 소득분위별 사망률에 대한 소득분위의 설명력을 비교하는 것이다. 어떤 가구 구성안의 데이터에서 분류 적합도가 가장 높게 나오는지

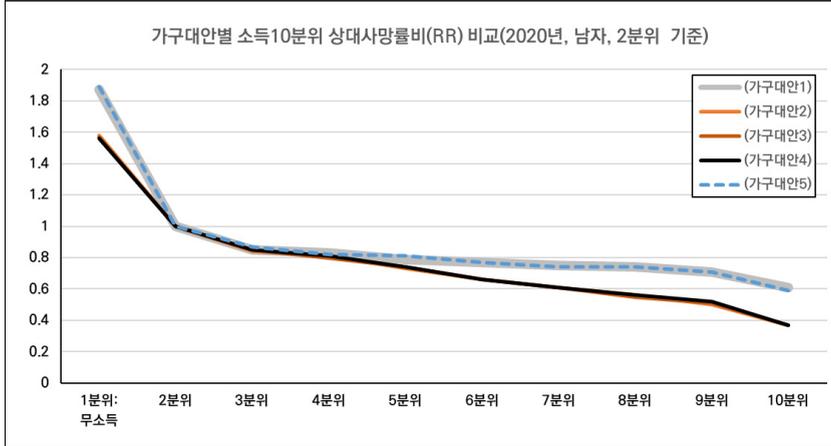
확인하였다. 소득 100분위별 사망률에 대한 선형모형에서 소득 100분위의 회귀계수와 수정설명계수를 활용하였다. 둘째, 가구 구성안 모형별로 소득 10분위에서 연령군 보정 상대사망률비(Relative all-cause mortality ratio)의 변화를 비교하는 것이다. 소득 100분위를 소득 10분위로 전환하여 성별, 연령별 대상자와 사망자 데이터에 대한 포아송 회귀분석을 시행하여 상대위험비를 추정하였다. 소득분위 구분에서 소득이 0인 대상자들은 모두 별도의 분위(소득 100분위에서는 0분위, 소득 10분위에서는 1분위)로 구분하였다. 이에, 두 번째 분석에서는 소득이 발생하는 최저 분위인 2분위를 기준으로 연령군 보정 상대사망률비를 추정하였다.

사망률에 대한 소득분위의 설명력은 건강보험증기준안에서 가장 낮았다(1.4%). 그다음은 세대보완 1안(건강보험 피부양자 포함)에서 7.09%로 낮았다. 세대보완 2안(비동거 가족 포함), 세대보완 3안(분리세대 추정), 세대기준안은 유사한 수준을 보였는데, 소득분위만 포함된 모형의 설명력은 세대기준안이 가장 높았다(59.9%).

소득 10분위에서 상대사망률비의 변화를 본 결과에서, 세대보완 1안(건강보험 피부양자 포함)과 건강보험증기준안이 동일한 패턴을 보였고, 이와 차이를 두고 기타 세 개의 구성안들이 동일한 패턴을 보였다.

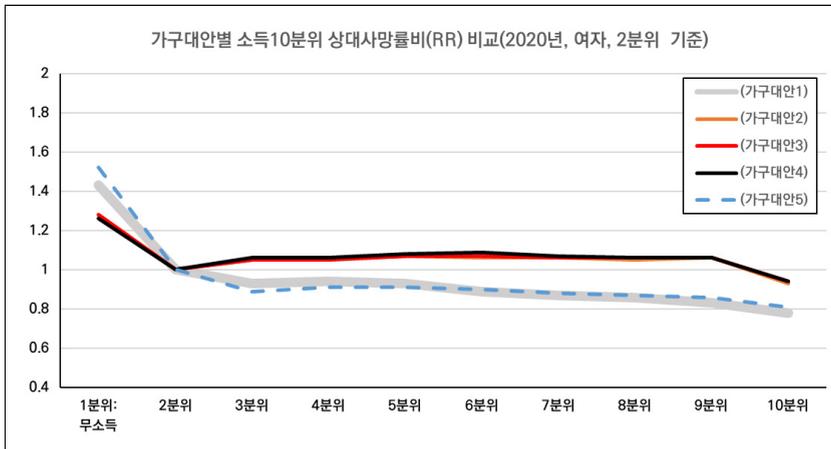
남자에서는 여자보다 소득 수준이 증가할수록 사망률이 감소하는 음의 관련성이 뚜렷했는데, 상대적으로 세대보완 1안(건강보험 피부양자 포함)과 건강보험증기준안보다 나머지 세 개의 구성안들이 더 가파른 기울기를 보였다. 여자에서 결과를 비교하면, 세대보완 1안(건강보험 피부양자 포함)과 건강보험증기준안이 기울기가 완만하지만 남자와 같이 소득이 증가할수록 사망률이 감소하는 동일한 경향을 보인 반면, 나머지 세 개 구성안들은 그러한 관련성이 사라지고 최상위 분위에서만 사망률이 급격히 감소하는 변화를 보였다.

[그림 7-8] 가구 구성안별 소득 10분위 연령보정 상대사망률비(RR)(남자)



자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

[그림 7-9] 가구 구성안별 소득 10분위 연령보정 상대사망률비(RR)(여자)



자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.

결론적으로, 가구를 정의하는 로직에서 가구 구성안 1과 5가 유사하고, 가구 구성안 3-5가 유사한 것이 결과에서도 체계적인 차이를 보여주고

있었다. 즉, 가구의 정의가 개인의 소득 수준 결정에도 영향을 미치며, 특히 여성에서 그러한 차이를 만들 수 있음을 보여준다.

향후, 연도를 확대하여 패턴의 변화를 비교하고 데이터 구조에서 체계적인 차이를 일으키는 요인에 대한 추가 검토가 필요하다.



제8장

결론 및 시사점

제1절 연구의 요약과 결론

제2절 자료의 활용도 제고를 위한 제언

제 8 장 결론 및 시사점

제1절 연구의 요약과 결론

최근 들어 사회정책 분야의 이론적, 정책적 연구에서 행정자료를 활용하는 사례가 증가하고 있다. 그런데 다수의 행정자료가 개인 단위의 정보를 제공하고 있고 가구 단위의 정보를 제공하지 못한다는 한계가 있다. 가구 소득을 계산하지 계산하지 못하면 소득계층을 구분하는 것도 어려워지고, 나아가 계층 간의 사회경제적 불평등을 분석하는 데에도 한계가 생긴다.

본 연구는 국민건강보험공단의 행정자료(맞춤형 소득-재산 DB)가 제공하는 개인 단위 정보를 이용하여 다양한 방법으로 가구 단위를 구성하고, 이를 사회정책 연구에 활용할 수 있는지를 탐색해 보는 것을 목적으로 수행하였다. 국민건강보험공단의 맞춤형 소득-재산 DB는 전 국민을 포괄하는 자료라는 점에서 본 연구를 통해 전 국민의 소득 및 재산 분포의 한 단면을 파악할 수 있을 것이다. 또한 이 DB가 다른 개인별 보건의료 특성과 연계할 수 있다는 점에서 보건의료 실태의 소득계층별 특성을 분석할 수 있을 것이다.

본 연구는 크게 두 부분으로 이루어져 있다. 먼저 제1~3장에서는 연구의 목적을 설명하고 본 연구에서 사용할 데이터를 구축하는 과정을 설명한다. 다음 제4~8장에서는 구축된 데이터를 이용하여 직접 분석을 수행하고 그 주요 결론과 함의를 설명한다.

제2장에서는 본 연구에 사용한 맞춤형 소득-재산 DB의 구축과정을 설명하였다. 이 자료는 국민건강보험공단의 국민건강정보 DB를 바탕으로 비식별화 등 가명 처리 과정과 연구에 필요한 정보만을 추출하는 과정을

거쳐 구축된 것이다. 이 자료는 2009년부터 2021년까지 총 13년의 자료로, 각 연도 1월 1일 기준의 국내에 실거주하는 전 국민(국내 거주 외국인과 재외국민 포함)을 포괄하고 있다. 건보공단의 소득 및 재산 자료는, 전 국민 당연가입인 사회보험 환경에서 구축된 것으로, 전 국민의 경제 수준에 대한 정보가 집약되어 있다는 점에서 의미가 있다. 다만 건보공단의 자료 수집 목적이 건강보험의 자격관리 및 보험료 부과이기 때문에, 관련 제도가 변화하면 자료 수집 대상에 변동이 발생하며, 조사 방식을 통해 수집한 자료와 수집항목에 차이가 있을 수 있음을 염두에 두어야 할 것이다.

제3장에서는 개인 단위의 세대 정보, 건강보험증 정보, 가족관계 등을 활용하여 가구 구성안을 제시하였다. 5개의 안은 기존의 주민등록 세대를 보완한 3개 안(건강보험 피부양자 포함안(1안), 비동거 가족 포함안(2안) 분리세대 추정안(3안))과 주민등록상 세대를 그대로 가구로 간주한 세대 기준안(4안), 동일 건강보험증 사용자를 가구로 간주한 건강보험증기준안(5안)으로 구성하였다.

기존의 주민등록 세대를 보완한 3개 안은, 같은 곳에 거주하면 생계를 같이 한다고 보는 것이 기본 개념이다. 가구 구성안을 구축하기 위하여 1, 2차에 걸쳐 각 가구원의 유형을 구분하는 과정을 거쳤다.

1차 유형 구분 과정에서는 주민등록 세대별로 세대주 중심의 가구가 구성된다고 했을 때, 세대주와 각 세대원의 정보를 비교하여, 가구 구성원으로 인정 가능한 경우와 그렇지 않은 경우를 별도의 유형으로 구분하였다. 2차 유형 구분 시에는 1차 유형에 기반하여 최종 가구 ID를 정의하였다. 1차 유형 구분이 각 세대원 본인과 같은 주소에 거주하고 있는 주민등록 세대주와 비교하는 과정이라면, 2차 유형 분류 시에는 주민등록 세대는 다르지만 건강보험증 대표의 가구로 묶어야 하는 상황을 고려하였다. 가구 구축 결과, 2021년 기준 1안이 2,121만 가구로 가구수가 가장 적었고,

5안이 2,762만 가구로 가구수가 가장 많았다.

제4장에서는 가계금융복지조사와 비교하여 맞춤형 소득-재산 DB의 소득 분포를 분석하였다. 우선, 가구 내에서 합산·균등화하지 않은 개인 소득 분포를 비교한 결과, 대체로 맞춤형 소득-재산 DB보다 가계금융복지조사의 소득 포착률이 높은 것으로 나타났다. 맞춤형 소득-재산 DB에 비해 서베이와 행정자료를 결합하여 구축한 가계금융복지조사에서 개인 근로소득, 개인사업소득, 개인금융소득, 개인공적연금 등이 모두 크게 포착되었다. 다음으로, 가구 내에서 합산·균등화한 가구소득 분포를 살펴본 결과, 가계금융복지조사는 맞춤형 소득-재산 DB보다 균등화 가구총소득 수준이 높고 평등하였다. 이는 한편으로 맞춤형 소득-재산 DB에 비해 가계금융복지조사가 개인소득을 고르게 포착하여 개인총소득 수준이 높고 평등하게 나타났기 때문이기도 하고, 다른 한편으로 맞춤형 소득-재산 DB에 비해 1인 가구 비율이 훨씬 낮은 가계금융복지조사에서 가구 내 소득 공유가 강하게 반영되었기 때문이기도 하다. 전반적인 분석 결과는 향후 맞춤형 소득-재산 DB의 소득 데이터를 더욱 발전시켜야 한다는 사실을 보여준다. 다양한 행정자료를 보완하여 소득 개념·범위를 확대하고 소득 측정의 정확도를 개선해야 한다. 또한 정확한 주소 및 가족관계 정보를 보완하고 실질적인 생계·주거 단위를 반영하여 가구 개념을 더 정교하게 다듬어야 할 것으로 판단된다.

제5장에서는 다섯 가지 가구 구성안별로 재산의 분포와 소득-재산의 결합 분포 특성을 비교 분석하였다. 상대적으로 더 많은 수의 세대와 개인을 묶는 세대보완 1안(건보 피부양자 포함)의 경우, 총가구수는 적고 가구당 평균 재산액은 높다. 반대로 건강보험증기준안은 세대를 분할 하기도 하기 때문에 가구의 수가 많아지고, 평균 재산은 감소한다. 재산의 계층 간 격차도 세대보완 1안(건강보험 피부양자 포함)에서는 작게, 건강

보험증기준안에서는 크게 나타난다. 세대를 기준으로 구성한 안과 세대 기준을 변형한 두 안(비동거 가족 포함안, 분리세대 추정안) 사이의 차이는 상대적으로 작다. 소득분위가 높아질수록 재산 보유액이 증가하는 것이 일반적이지만, 5분위 배율로 본 소득계층 간 재산 격차는 세대기준안에서 가장 크게 나타났다. 재산분위별 소득의 격차는 세대기준안과 세대보완 3안(분리세대 추정)이 가장 컸다.

제6장에서는 가구의 경제적 지불능력을 반영하는 소득-재산 결합지수를 도출했다. 특히 선행연구에서 제안된 대표적인 두 가지 방법인 연금소득 화법과 소득인정액 환산법을 적용해 보고, 결합방식에 대한 결과의 민감도와 지수 간 정합성을 살펴보았다. 분석 결과, 가구의 평균적인 특성은 결합방식에 따라 크게 변하지 않음을 발견하였다. 소득만을 고려할 경우와 재산까지 고려하는 경우, 가구주의 연령이나 고령 가구원의 유무가 달라지지만, 재산을 어떤 방식으로 환산하여 소득과 합산을 할 것인지에 따라서는 평균적인 특성이 크게 변하지 않았다. 이러한 결과는 하위 가구를 식별할 때나 상위 가구를 식별할 때 모두 동일하게 나타났다. 반면 지수 간 정합성은 높지 않은 것으로 나타났다. 재산을 고려하는 경우에도 연금화 방식을 사용하는가 혹은 기초생활보장제도에서 사용하는 소득인정액 환산방식을 사용하는가에 따라 하위 가구나 상위 가구의 식별이 상당히 다르게 나타난다는 사실을 발견했다. 이러한 사실은 정책 수립에 있어서 가구의 지불능력을 어떻게 정의할 것인가에 대해 신중하게 검토할 필요가 있다는 사실을 의미한다.

제7장에서는 소득계층별 사망률의 차이를 분석하였다. 사망률에 대한 소득 100분위의 설명력은 '건강보험증기준안'과 '세대보완 1안(건강보험 피부양자 포함)'에서 가장 낮았으며, '세대보완 2안(비동거 가족 포함)', '세대보완 3안(분리세대 추정)', '세대기준안'은 유사한 수준을 보였다.

소득 10분위에서 상대사망률비의 변화도 이와 마찬가지로, '세대보완 1안 (건강보험 피부양자 포함)'과 '건강보험증기준안'이 동일한 패턴을 보였으며, 그 외에 세 개는 같은 양상을 보였는데, 앞선 두 안에서는 남녀 모두 소득이 증가할수록 사망률이 감소하였으나, 나머지 세 안의 경우 관련성이 없이 최상위 분위에서만 사망률이 급격히 감소하였다.

제2절 자료의 활용도 제고를 위한 제언

1. 가구 구성안의 개선

본 연구의 일차적 과제는 건보공단의 행정자료, 그 가운데에서도 맞춤형 소득-재산 DB에 포함된 자료를 이용하여 가구 단위의 정보를 구성하는 것이었다. 이 과정에서 동일 세대 여부, 동일 건강보험증 보유 여부, 읍면동 단위의 주소 정보, 소득 활동 여부, 연령, 가족관계 정보 등 제한된 정보 항목을 최대한 이용하고자 하였다. 같은 건강보험증을 사용한다는 것은 생계를 같이하는지 여부를 판단하는 데 유용한 정보이다. 이와 같은 정보를 더 활용할 경우 일부 선행연구에서 주민등록세대와 가구를 동일시하는 방안에 비해 훨씬 다양한 분석을 할 수 있다.

그러나 이와 같이 복잡한 추론 과정을 거치지 않는 간단한 방법이 있는데, 이는 통계청의 인구주택총조사 자료와 연계하는 방법이다. 통계청의 인구주택총조사는 전국의 가구를 조사하는 유일한 자료이다. 전국 가구의 수와 분포에 대한 정보를 담고 있다. 만일 이 자료와 건보공단의 개인별 자료를 연계할 수 있다면, 다른 정보를 이용하여 가구를 '구성'하는 번거로운 시도를 피하고 정확성도 높일 수 있을 것이다.

물론 인구주택총조사와 맞춤형 소득-재산 DB의 개인별 정보를 연계하는 일이 쉬운 일은 아니다. 개인정보보호를 위해 각 데이터 보유 기관이 취하고 있는 여러 가지 법적, 제도적 제약이 있을 수 있으며, 이 제약을 극복하는 일은 한두 기관의 의지만으로 해결할 수 있는 일이 아니기 때문이다. 그러나 인구주택총조사 정보와 건강보험공단의 정보를 연계할 수 있다면, 가구 구성 과정에 소요되는 시간과 노력, 그리고 무엇보다 정보의 손실 등을 크게 줄일 수 있을 것이다.

물론 제3장에서 보았듯이, 인구주택총조사의 가구 개념은 여러 가지 가구 정의 방식 가운데 하나일 수 있다. 그러나 변화하는 현실에 맞추어 다른 가구 정의 방식을 찾는 이론적 노력과는 별개로, 현실적으로 존재하는 가구 정보를 활용하지 못하는 것은 무척 아쉬운 점이다. 대규모 행정자료, 특히 개인 및 가구 단위 자료의 연계가 가능하도록 제도적 개선이 이루어져야 할 것이다.

인구주택총조사의 가구 정보와 연계하는 것이 단기적으로 어렵다고 하더라도 기존 자료를 보완하여 가구 구성의 정확성을 높일 수 있는 방법도 있다. 본 연구의 가구 구성안 가운데 분리세대 추정안은 동일한 세대를 구성하고 있더라도 실제로 생계를 같이하고 있지 않다고 추정되는 세대를 분리하여 가구를 구성하는 안이었다. 그런데 이 작업은 서로 다른 세대 이더라도 실질적으로 같은 주소에 거주하면서 생계를 같이 하는 세대들을 하나의 가구로 묶는 작업과 동시에 진행되어야 한다. 그런데 후자의 작업을 위해서는 두 세대가 동일 주소에 거주하는지를 확인할 수 있어야 한다. 본 연구에서 활용한 맞춤형 소득-재산 DB에서는 개인별 주소 정보가 읍면동까지만 주어졌기 때문에 동일 주소 여부를 정확히 파악하는 것이 불가능하였다. 만일 상이한 세대 사이에 주소의 일치 여부를 파악할 수 있는 정보가 추가로 주어진다면, 실질적으로 생계를 같이 하는 두 개

이상의 세대를 하나의 가구로 간주하는 것이 가능해질 것이다.

2. 자료 항목의 보완

행정자료는 자료를 축적한 행정기관의 행정행위에 의해 자료의 내용과 범위가 제한되는 경우가 많다. 본 연구에서 활용한 맞춤형 소득-재산 DB의 경우도 예외가 아니다. 소득 항목의 경우, 가구 소득을 구성하는 소득 원천 가운데 다수의 공적 이전소득 항목과 사적 이전소득 전체가 자료에 포함되어 있지 않다.

공적이전소득을 구성하는 다수의 급여들이 건보공단의 자료에 포함되지 않은 것은 그것이 보험료 부과에 근거가 되지 않기 때문일 것이다. 그러나 소득분배 실태를 분석하거나 소득계층을 구분하는 데 있어서 공적이전소득을 정확히 반영하는 것은 매우 중요하다. 특히 공적이전소득의 경우 사회정책의 발전과 함께 저소득층 대상 공적이전소득의 규모와 종류가 대폭 증가한 만큼, 저소득층의 실태를 파악하기 위해서는 공적이전소득 자료가 대폭 보완될 필요가 있다. 사적이전소득을 행정자료로 수집하는 것은 원천적으로 불가능하다고 하더라도 좀 더 다양한 공적이전소득 관련 자료가 수집되고 연계될 필요가 있다.

재산 항목의 경우 자료의 한계는 더욱 분명하다. 개인과 가구 재산의 중요한 항목인 금융재산은 맞춤형 소득-재산 DB에 포함되어 있지 않고, 전월세 보증금의 경우 주로 지역가입자에 대해서만 자료 수집이 이루어져 있다. 이 역시 현행 건강보험료 부과체계의 특성을 반영한 것이다. 그러나 이러한 한계는 건보공단 자료의 사회정책적 적용 가능성이라는 관점에서 볼 때, 재산 자료의 유용성을 매우 제약하는 요인이다. 단기적으로 건보공단의 행정자료에 금융자산이나 직장가입자의 전월세 보증금 등의 항목이

추가되기 어렵다면, 다른 자료와의 연계 방안을 모색함으로써라도 재산 항목의 포괄성을 높이는 것이 필요하다.

3. 가구 개념에 대한 이론적 탐색과 행정자료

자료의 확충과 자료 간 연계 같은 노력과는 별개로, 가구를 어떻게 이해할 것인가에 대한 이론적 논의는 지속될 필요가 있다. 생계를 같이하는 것이 과거에는 같은 거주지에 사는 것과 거의 동일한 의미를 띠었다. 그러나 오늘날에는 거주지를 달리하면서도 실제로 소득을 같이 사용하고 지출의 효과를 공유할 수 있는 다양한 방법들이 생겨나고 있다. 부모의 신용카드를 자녀가 사용하는 방식으로 가족 간에 소득을 공유하는 것, 그리고 부모가 자녀에게 용돈을 주는 방식으로 소득을 공유하는 것은 공간적 제약의 범위가 다를 수밖에 없다. 동일한 물리적 장소에서 거주하는 것이 가구 정의의 기본 요건이 되어야 하는가는 사회 경제적 환경 변화에 영향을 받을 수밖에 없다.

다른 한편으로는 같은 거주 단위에 사는 인적 구성과 생활 양식도 다양해지고 있다. 혈연관계에 있는 가족이 분산되어 거주하는 경우가 늘어나는가 하면, 혈연관계가 아닌 개인들이 동일한 거주지에서 생활하는 경우도 늘어난다. 따라서 이러한 현실 변화를 반영한 유연한 가구 개념이 필요할 수 있다. 행정자료를 좀 더 폭넓게 활용할 수 있다면, 가구 개념을 확장하거나 다양화하기 위한 여러 가지 시험적 분석이 가능할 것이다.



- 강민진, 손강주. (2018). 암, 심뇌혈관 질환자의 소득수준에 따른 사망 및 이차 질환 발생 위험. 한국콘텐츠학회논문지, 18(10), 145-157.
- 강신욱, 이다미, 이원진, 한겨레, 문성웅, 이현옥, ..., 오하린. (2021). 『행정자료를 이용한 사회보험 사각지대 분석』, 소득주도성장특별위원회·한국보건사회연구원.
- 강영호, 김혜련. (2006). 우리 나라의 사회경제적 사망률 불평등: 1998 년도 국민 건강영양조사 자료의 사망추적 결과. 예방의학회지, 39(2), 115-122.
- 강정희, 김철웅, 권순석. (2022). 국가 암검진 사업의 주요 암종별 5 년 생존율과 사회경제적 수준 및 요약병기의 관련성: 광주·전남 지역암등록본부 자료를 중심으로. Journal of Korean Academy of Community Health Nursing, 33(2), 237-246.
- 구인회, 강상경, 김용득, 홍백의, 정찬미(2020). 『사회보장 행정데이터의 효과적 활용 및 발전방안 연구』, 보건복지부·서울대학교 산학협력단.
- 국가법령정보센터. 국민건강보험법. 법률 제19123호. <https://www.law.go.kr/>에서 2022. 12. 30. 인출.
- 국가법령정보센터. 인구주택총조사 규칙. 기획재정부령 제810호. <https://www.law.go.kr/>에서 2022. 12. 30. 인출.
- 국민건강보험공단. (2012). 국민건강정보 DB 사용자 매뉴얼: 자격·보험료.
- 국민건강보험공단. (2018). 국민건강정보 DB 2.0 매뉴얼. 내부자료.
- 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.
- 국민건강보험공단. (2021a). 지역보험료 부과관리 업무처리지침.
- 국민건강보험공단. (2021b). 직장보험료 부과관리 업무처리지침.
- 국민건강보험공단. (2022). 건강보험자료 공유서비스 웹사이트. <https://nhiss.nhis.or.kr> 2022. 9. 28. 인출.
- 국토교통부. (2022). “임대차 신고제’ 제도기간 내년 5월말까지 1년 연장”. 대한

민국정책브리핑. 2022.05.26.

김동준. (2020). 우리나라 조기 대장암 환자의 소득수준이 사망에 미치는 영향.
석사학위논문, 순천향대학교 대학원.

노대명, 이현주, 강신욱, 김문길, 신현웅, 신화연, ... 이주미. (2013). 국민기초생
활보장제도의 맞춤형 급여체계 개편방안 마련을 위한 연구. 보건복지부·
한국보건사회연구원.

보건복지부. (2003). 2003년 국민기초생활보장사업 안내.

보건복지부. (2019). 2019년 국민기초생활보장사업 안내.

보건복지부. (2020). “소득 중심의 건보료 부과체계에 한 발짝 다가서다!”. 보도자료.
2020.8.19.

보건복지부. (2022). “건강보험료 부과체계 2단계 개편 예정대로 9월부터 시행”. 보도
자료. 2022.8.30.

보건복지부. (2022). 2022년 국민기초생활보장사업 안내.

송태민. (2013). 앤터슨 행동모형을 이용한 노년기 외래의료서비스 이용에 대한
스트레스 취약요인의 매개효과 분석. 보건사회연구, 33(1), 547-576.

여유진, 김미곤, 김문길, 정재훈, 홍경준, 송치호. (2011). 국민기초생활보장제도
재산의 소득환산제 개선방안 연구. 보건복지부·한국보건사회연구원.

이원진, 정해식, 전지현. (2019). 소득조사 마이크로데이터 비교 분석: <가계동
향조사>와 <가계금융·복지조사>를 중심으로. 한국보건사회연구원.

이현주, 오욱찬, 이원진, 이윤경, 성재민, 이길제, ... 이병재. (2020) 『사회보장
정책 효과성 분석을 위한 행정데이터 연계·활용 연구』, 보건복지부·한국
보건사회연구원.

전병유, 계봉오, 최성수, 황선재, 최을, 이도훈, ... 고원태. (2018a). 제2장 행정
자료를 이용한 세대 간 분위소득 이동 분석, 정해식 외. (2018). 소득불평
등 심화의 원인과 정책적 대응효과 연구 2, 경제인문사회연구회.

전병유, 계봉오, 최성수, 황선재, 최을, 이도훈, ... 고원태. (2018b). 제3장 행정
자료를 이용한 한국사회 빈곤 추이. 정해식 외. (2018). 소득불평등 심화의
원인과 정책적 대응효과 연구 2, 경제인문사회연구회.

- 정은희, 강신욱, 김태완, 정해식, 임완섭, 오욱찬, 김근혜. (2018). 국민기초생활 보장제도 재산기준 개선방안 연구. 한국보건사회연구원.
- 최요철, 김은영. (2007). 가계소비의 자산효과 분석과 시사점. 한국은행.
- 통계청. (2020a). 2020년 가계금융복지조사 조사표. <https://www.narastat.kr> 에서 2021. 8. 18. 인출.
- 통계청. (2020b). 『가계금융복지조사』 통계정보보고서. 통계청.
- 통계청. (2020c). 가계금융복지조사에서의 조사자료와 행정자료의 통합방법 이해. 통계청.
- 통계청. (2020d). 『인구총조사』 통계정보보고서. 통계청.
- 통계청. (2020e). 2020년 가계금융복지조사 원자료. 통계청 MDIS 원격접속. DOI번호: 10.23333/R.930001.001.에서 2022. 12. 15. 최종접속.
- 통계청. (2020f). 2020년 가계금융복지조사 원자료. 마이크로데이터.
- 통계청. (2020g). 2020 인구주택총조사 조사표.
- 통계청. (2021). 2021년 가계금융복지조사 지침서. 통계청.
- 통계청. (2022). 2022 가계동향조사 이용자 가이드북. 통계청
- Ando, A., & Modigliani, F. (1963). The " life cycle" hypothesis of saving: Aggregate implications and tests. *The American economic review*, 53(1), 55-84.
- Brandolini, A., Magri, S., & Smeeding, T. M. (2010). Asset-based Measurement of Poverty. *Journal of Policy Analysis and Management*, 29(2), 267-284.
- Friedman, M. (1957). The permanent income hypothesis. In *A theory of the consumption function* (pp. 20-37). Princeton University Press.
- Glei, D. A., Lee, C., & Weinstein, M. (2022). Assessment of mortality disparities by wealth relative to other measures of socio-economic status among US adults. *JAMA Network Open*, 5(4), e226547-e226547.
- Martikainen, P., Mäkelä, P., Koskinen, S., & Valkonen, T. (2001).

- Income differences in mortality: a register-based follow-up study of three million men and women. *International Journal of Epidemiology*, 30(6), 1397-1405.
- OECD. (2013). *OECD guidelines for micro statistics on household wealth*. Paris: OECD Publishing.
- Piketty, T., & Saez, E. (2003). Income inequality in the United States, 1913-1998. *The Quarterly journal of economics*, 118(1), 1-41.
- Sabanayagam, C., & Shankar, A. (2012). Income is a stronger predictor of mortality than education in a national sample of US adults. *Journal of health, population, and nutrition*, 30(1), 82.
- Saez, E., & Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, 131(2), 519-578.
- Shi, J., Tarkiainen, L., Martikainen, P., & van Raalte, A. (2021). The impact of income definitions on mortality inequalities. *SSM-Population Health*, 15, 100915.
- UN. (2017). *Principles and recommendations for population and housing censuses(Revision3)*. https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Series_M67_rev3-E.pdf 에서 2023. 3. 6. 인출.
- UNECE. (2006). *Conference of European statisticians recommendations for the 2010 censuses of population and housing*. https://unece.org/fileadmin/DAM/stats/publications/CES_2010_Census_Recommendations_English.pdf 에서 2023. 3. 6. 인출.
- Weisbrod, B. A., & Hansen, W. L. (1968). An Income-net Worth Approach to Measuring Economic Welfare. *American Economic Review*, 58(5), 1315-1329.

부록



〈부표 6-1〉 가구 정의와 결합방법에 따른 소득 중위값의 50%과 90분위수

가구 정의	결합방법	0원 포함	기준값	
			중위값의 50%	90분위수
세대보완 1안 (건강보험 피부양자 포함)	소득만 사용	O	중위값의 50%	9,644,636
		X	중위값의 50%	11,906,426
		O	90분위수	54,400,428
	연금소득화법(2%)	O	중위값의 50%	11,775,000
		X	중위값의 50%	13,543,967
		O	90분위수	66,597,724
	연금소득화법(10%)	O	중위값의 50%	13,973,278
		X	중위값의 50%	16,085,224
		O	90분위수	84,301,627
	기초생활보장제도 환산법	O	중위값의 50%	17,862,800
		X	중위값의 50%	22,717,197
		O	90분위수	210,140,255
세대보완 2안 (비동거 가족 포함)	소득만 사용	O	중위값의 50%	8,323,224
		X	중위값의 50%	11,224,310
		O	90분위수	54,773,620
	연금소득화법(2%)	O	중위값의 50%	10,962,442
		X	중위값의 50%	12,723,543
		O	90분위수	66,370,216
	연금소득화법(10%)	O	중위값의 50%	13,275,000
		X	중위값의 50%	15,296,026
		O	90분위수	82,529,716
	기초생활보장제도 환산법	O	중위값의 50%	17,407,270
		X	중위값의 50%	22,019,881
		O	90분위수	197,052,618
세대보완 3안 (분리세대 추정)	소득만 사용	O	중위값의 50%	8,011,537
		X	중위값의 50%	11,118,723
		O	90분위수	55,134,378
	연금소득화법(2%)	O	중위값의 50%	10,664,539
		X	중위값의 50%	12,565,048
		O	90분위수	66,729,777

240 국민건강보험공단 행정자료의 사회정책적 활용을 위한 탐색적 연구

가구 정의	결합방법	0원 포함	기준값	
	연금소득화법(10%)	O	중위값의 50%	12,960,000
		X	중위값의 50%	15,135,767
		O	90분위수	82,841,250
	기초생활보장제도 환산법	O	중위값의 50%	16,954,293
		X	중위값의 50%	21,804,285
		O	90분위수	196,169,417
세대 기준안	소득만 사용	O	중위값의 50%	8,111,187
		X	중위값의 50%	11,150,000
		O	90분위수	55,580,040
	연금소득화법(2%)	O	중위값의 50%	10,797,567
		X	중위값의 50%	12,617,681
		O	90분위수	67,319,654
	연금소득화법(10%)	O	중위값의 50%	13,139,159
		X	중위값의 50%	15,229,266
		O	90분위수	83,600,578
	기초생활보장제도 환산법	O	중위값의 50%	17,360,522
		X	중위값의 50%	22,128,000
		O	90분위수	198,496,589
건강보험 증 기준안	소득만 사용	O	중위값의 50%	8,962,985
		X	중위값의 50%	10,776,900
		O	90분위수	49,620,280
	연금소득화법(2%)	O	중위값의 50%	10,714,281
		X	중위값의 50%	12,210,675
		O	90분위수	60,611,597
	연금소득화법(10%)	O	중위값의 50%	12,452,382
		X	중위값의 50%	14,272,922
		O	90분위수	76,691,211
	기초생활보장제도 환산법	O	중위값의 50%	14,597,640
		X	중위값의 50%	18,036,595
		O	90분위수	187,169,638

자료: 국민건강보험공단. (2020). 맞춤형 소득-재산 DB 원자료.