

사회보장행정데이터의 구축과 사회정책의 근거 강화: 성과와 과제

2023. 9. 14.

이현주(한국보건사회연구원 선임연구위원)

목 차

1. 사회보장행정데이터의 개요
2. 사회보장행정데이터의 구축배경과 의의
3. 사회보장행정데이터의 정보구성과 구축과정
4. 정책 근거 강화와 데이터 활용사례
5. 쟁점과 향후 과제

1. 사회보장행정데이터의 개요

가. 행정자료의 정의

- 행정기관이 서비스를 포함한 행정 목적으로 개인이나 가구에 대한 정보를 수집하여 관리하는 데이터를 의미(Jones and Elias, 2006. 유종성 2020. p. 9에서 재인용).

나. 행정자료의 활용 목적

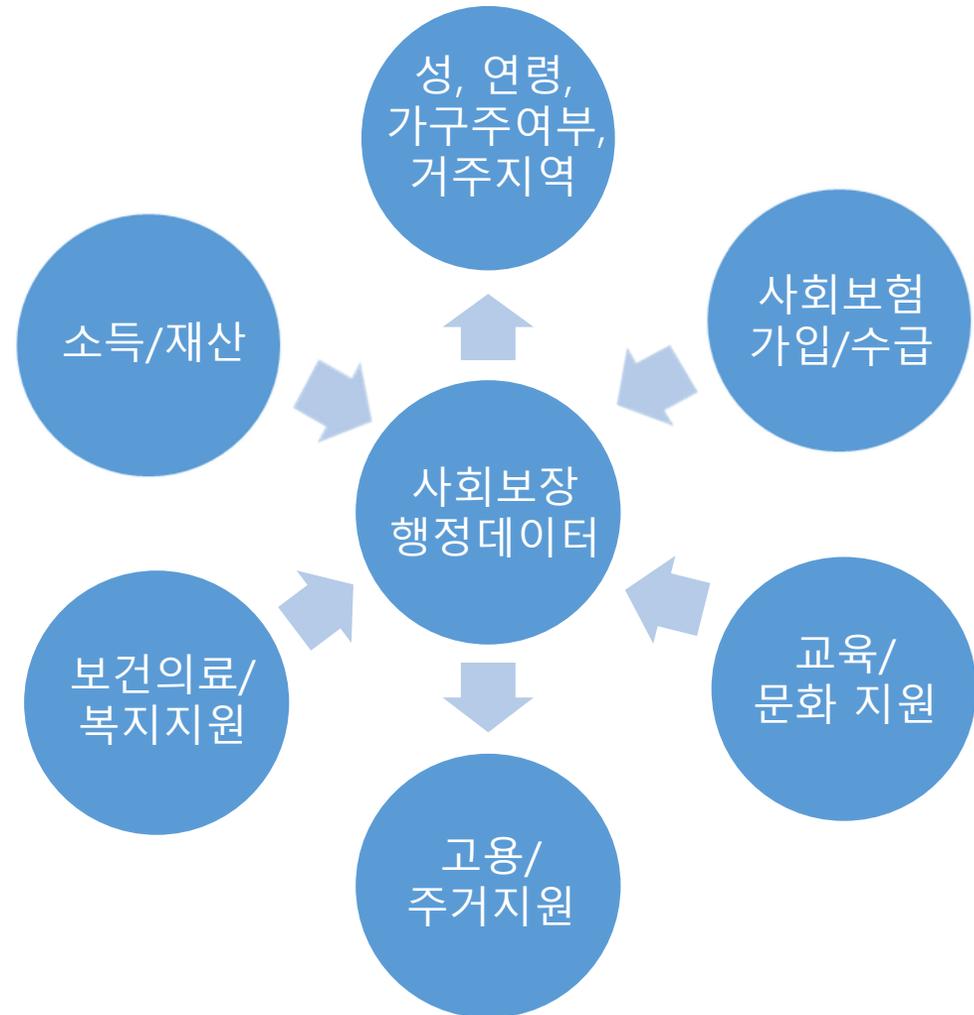
- 정확한 정책집행:
 - 사회정책집행 과정에서 수급 적격성 파악 등 행정자료 고유의 목적
- 정책의 근거강화:
 - 정책평가 및 국민의 생활여건 파악을 위한 연구, 분석
 - 행정데이터가 축적됨에 따라, 행정데이터를 활용, 증거(evidence)를 토대로 정책결정을 하려는 경향이 강화되는 중.

다. 논의 범위

- 정책의 근거 강화를 위한 행정데이터 구축과 활용으로 주 논의 한정
 - 또한 사회보장행정데이터 구축과 활용 사례를 중심으로 논의

1. 사회보장행정데이터의 개요

- 사회보장행정데이터는
 - 근거: 사회보장기본법 제42조 등
 - 정부, 지자체, 공공기관이 보유한 인구, 가구, 소득, 재산, 사회보장 수급 정보를 (가명처리하여) 개인 단위로 결합한 통합행정데이터
- 표본대표성
 - 전인구의 약 20%표본, 가중치 보정으로 대표성 확보
 - 2021년 기준 2023년 구축 중. 2022년 2020년 기준 데이터 구축
- 정보포괄성
 - 개인, 가구의 인구사회학적 기초 정보,
 - 경제여건, 건강 등 정보
 - 주거/의료/교육/소득지원 등 사회보장수급정보
- 정보 연계 통합
 - 생년월일, 성, 이름을 가명처리하여 결합키생성
 - 개인 결합키의 중복, 해당 가구원 가구의 가구원 제거 등 결합키 중복으로 개인 표본의 약 30%정도 손실



1-1. 사회보장행정데이터 구축 배경과 의의: 구축배경

- 사회정책 분석 기반의 취약성과 사회보장기본법의 개정
 - 정책근거강화에 대한 사회적 요구 대비 데이터기반의 취약성 잔존
 - 기존의 주된 정책분석 자료는 설문조사자료, 통합되지 않은 행정데이터(집계 정보 또는 특정 정책 정보)

<표 1> 설문조사자료와 행정자료의 장단점 비교

	설문조사자료	(통합)행정자료
장점	높은 접근성	정보의 포괄성 표본규모의 제약이 거의 없음 다양한 형태의 자료 구축 잠재력 낮은 구축비용
단점	제한된 정보 높은 비용(별도 조사 수행 시)	낮은 접근성(정보 정의 부재 등) 인식과 태도 파악 취약 지출 부담 등 욕구 정보 부족

- 통합되지 않은 행정데이터는 정보 포괄성의 제약으로 각 제도의 특성 이해에 한계 초래
 → 통합행정데이터의 수요 증대

1-1. 사회보장행정데이터 구축 배경과 의의: 주요 설문조사자료의 한계

- 사회정책의 빠른 확충 대비 설문조사 문항의 포괄성 취약
 - 사회서비스에 대한 정확한 정보 수집에서는 더 큰 한계 노정

<표 2> 대표적인 설문조사자료의 사회정책 관련 정보

가계금융복지조사자료	한국복지패널조사자료
공적연금 기초연금 양육수당(출산장려금포함) 장애인 연금, 수당 공공부조(현금성급여) 근로장려금, 자녀장려금 기타 정부보조금	공공부조수급(의료급여등수급액정보 부재)/ 장애수당 연금가입과수급, 수급액, 기초연금수급액 고용보험, 건강보험 가입 보육료지원, 아동수당, 양육수당, <u>한부모지원</u> 학비지원/ 주거보장수급 긴급복지/ <u>근로장려세제</u> <u>에너지감면과 보조금</u> / <u>통신비보조등 보조금</u> 기타 <u>바우처</u> 수급액 장애인·노인 <u>돌봄서비스이용 경험</u>

1-1. 사회보장행정데이터 구축 배경과 의의: 법적 근거

- 사회보장기본법의 개정(2021년 6월)과 사회보장행정데이터의 구축
 - 2019, 2020년 실험 구축을 위한 연구수행 등 준비 추진
 - 데이터 안정성을 위하여 사회보장기본법 개정, 2022년 동법에 근거한 1차년도 데이터 구축

<참고> 사회보장기본법 중 사회보장행정데이터의 목적과 개념

○ 사회보장기본법 제42조

① 위원회는 사회보장 정책의 심의·조정 및 연구를 위하여 관계 기관의 장에게 사회보장 행정데이터가 모집단의 대표성을 확보할 수 있는 범위에서 다음 각 호에 해당하는 사회보장 행정데이터의 제공을 요청할 수 있다. 이 경우 사회보장 행정데이터의 제공을 요청받은 관계 기관의 장은 특별한 사유가 없으면 이에 따라야 한다.

○ 사회보장기본법 제3조

⑥ “사회보장 행정데이터”란 국가, 지방자치단체, 공공기관 및 법인이 법령에 따라 생성 또는 취득하여 관리하고 있는 자료 또는 정보로서 사회보장 정책 수행에 필요한 자료 또는 정보를 말한다.

1-2. 사회보장행정데이터 구축 배경과 의의: 구축 목적

□ 사회보장행정데이터의 의의와 목적

○ 정부의 정책 지향과 사회보장행정데이터의 의의

- '사회보장행정데이터·통계 체계 구축 등 객관적 분석' 기반을 마련하는 것은 중요한 정책목표 중 하나
- '모든 데이터가 연결되는 세계 최고의 디지털플랫폼 정부 구현'은 현 정부의 국정과제 중 하나

○ 사회보장행정데이터 활용 목적

- 사회보장제도의 심의·조정을 위한 과학적인 정책 분석을 지원하는 사회보장행정데이터의 구체적인 활용 목표를 정리하면 아래와 같음.
 - 사회보장 관련 주요 기초 통계를 생산, 안정적이고 합리적인 정책 추진 환경 조성
 - 개별 정책의 경계를 넘는 지표를 개발, 정책목표 공유 및 정책 협력의 객관성 제고
 - 시계열데이터, 패널 데이터 기반, 사회보장제도 발전사 이해와 정책 전망 합리화

3. 사회보장행정데이터의 정보구성과 구축과정

□ 사회보장행정데이터의 정보구성 방식

관련 법 고려, 전문가의 의견수렴을 거쳐 정보목록을 구성, 보완

○ 정보 선정 기준 1

- 사회보장제도의 법적 범위 반영
- 사회보장기본법 제42조, 사회보장기본법 시행령 제20조, 사회보장기본법 시행령 제20조 2항 관련 별표에 사회보장행정데이터로 요청할 수 있는 정보의 목록 참조

○ 정보 선정 기준 2

- 정책분석에서 수요가 예상되는 정책정보인가?(제도 규모와 정보의 유의성 고려)
- 정보의 안정성, 정의의 명확성이 확보될 수 있는가?
- 정보의 관리조직, 부처의 경계를 넘어 분석하게 될 때 상대적으로 큰 의미를 지니는가?
- 향후 정책의 변화를 전망할 때 중요한 의미를 지니는 정보인가?
- 향후 사회보장위원회에서 정책기획, 평가에서 유념하여야 하는 정보인가?

3. 사회보장행정데이터의 정보구성과 구축과정

<표 1> 사회보장행정데이터의 정보 구성

사회경제적 특성	사회보장제도 수급				기타
	사회보험	수당 등 현금지원	사회서비스	공공부조	
성, 연령, 가구주여부, 가구규모, 지역(시군구) 소득(근로소득, 사업소득, 기타소득 등) 재산(부동산, 퇴직연금적립금, 전월세 보증금 일부) 세금과 사회보험기여금	건강보험, 고용보험(육아 휴직급여 등 포함) 연금보험(특수 직역연금 포함)산재보험 요양보험	아동수당, 기초연금, 장애인연금, 보훈급여, 장애수당, 장애수당, 장애아동수당, 아동수당, 양육수당, 농어촌양육수당, 장애아동양육수당	주거급여, 임대주택, 교육급여, 재난적 의료비지원, 입환자의료비지원, 장학금지원, 취업지원 등 고용서비스, 노인돌봄, 아이돌봄, 장애인활동지원, 장애아동가족지원, 발달장애인 지원 보육서비스, 농어촌 양육수당, 자활근로, 기저귀조제분유지원, 노인일자리 사업, 지역사회서비스 투자, 한부모가족지원, 문화이용권	생계급여 , 긴급복지 지원, 근로장려 금, ,	퇴직연 금 등

3. 사회보장행정데이터의 정보구성과 정보구성

<표 2> 사회보장행정데이터의 데이터 파일 예시

구분	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ		
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
2	가구식별	가구식별	일반사항	일반사항	일반사항	일반사항	지역구분	지역구분	지역구분	지역구분	지역구분	지역구분	가중치	가중치	가중치	가중치	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	소득/재산	
3	SAMPLE7	가구일련번호	성별	출생년도	만나이	가구주여부	가구원수	시군구코드	시도	수도권여부	시군구	구분	인구감소지	가구가중치	가구원가중치	가구가중치	가구원가중치	근로소득	사업소득	부동산소득	일용소득	기타소득	공적이전소득	공적이전소득	일차소득	일차소득	경상소득	경상소득	가치분소득	가치분소득	균등화시장	균등화시장	균등화시장	재산_개인	재산_가구	시장소득		
536																																						
537																																						
538																																						
539																																						
540																																						
541																																						
542																																						
543																																						
544																																						
545																																						
546																																						
547																																						
548																																						
549																																						
550																																						
551																																						
552																																						
553																																						
554																																						
555																																						
...																																						

3. 사회보장행정데이터의 정보구성과 구축과정

- 통계청의 가구 및 개인 정보 또는 행안부의 세대 정보를 참고하여 표본추출가능
 - 자료정제과정에서 정보의 가명처리 강화를 위한 가공·정제
 - 자료활용과정을 지원하기 위한 신청에 대한 심의, 지원절차 진행

[그림 1] 사회보장행정데이터 구축 과정



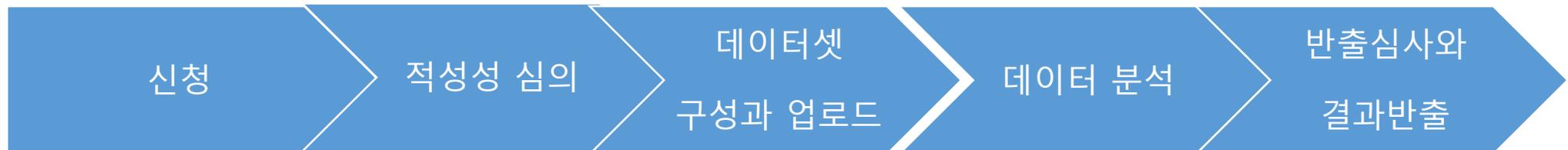
<참고> 사회보장행정데이터의 활용 예상 절차(안)

- 행정데이터의 국내외 활용 절차를 참고하면 아래와 같은 절차로 사회보장행정데이터 활용이 가능할 것

○ 활용 신청과 활용 절차 개요

- 데이터 활용 신청: 정보 목록을 참고하여 해당 연구에 필요한 정보를 요청
- 데이터 이용의 적정성 심의: 연구목적의 적정성, 요청한 자료의 연구목적과의 부합 여부 및 최소한의 정보 요청인지 여부 등 검토
 - 사회보장제도 조정을 위한 기초자료 마련 등 사회보장행정데이터 구축 목적에 부합하는 연구에 한정하여 승인될 것으로 예상
- 데이터 활용을 위한 데이터 셋 생산, 분석 센터에 데이터 업로드
- 데이터분석, 분석결과에의 검토와 반출

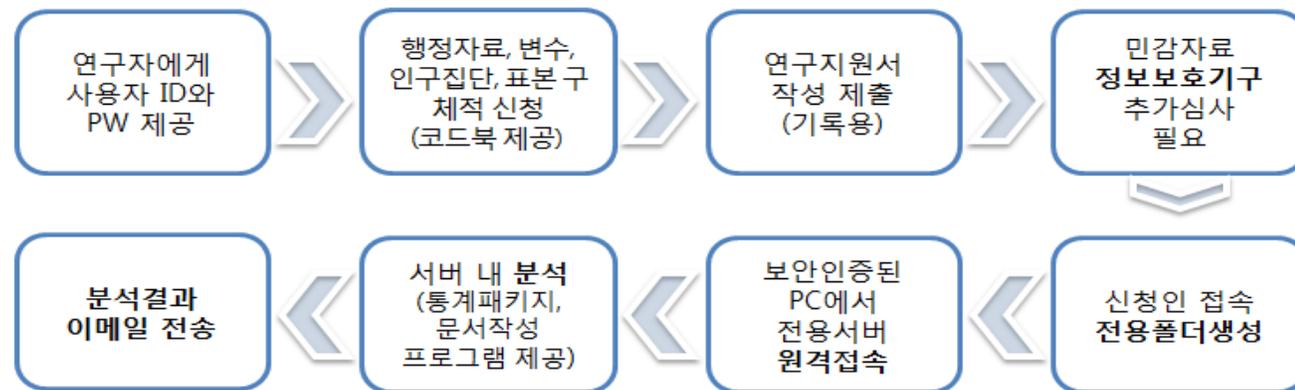
[그림 5] 사회보장행정데이터 활용 절차(안)



4. 정책근거 강화와 행정데이터 활용 사례: 해외 사례

○ 북유럽국가의 행정데이터 활용(이현주 외. 2020. pp. 376-7.)

- 자료신청을 하기 위해서 연구자는 사용자 ID와 비밀번호를 제공받아야 함. 뿐만 아니라 안내서와 자료의 설명서를 제공받음.
 - 덴마크 통계청은 승인된 연구기관의 연구자, 학생들이 통계청 행정등록 자료에 접근할 수 있도록 서비스를 제공.
 - 승인 기관 중 공공기관은 약 80-90%. 최근에서야 안전한 환경을 조성한 민간기관의 비율이 증가 중.
- 자료 결합이 필요하면 이 결합자료에 대한 승인 필요. 의료정보, 범죄정보와 같은 민감정보는 추가적 심사를 받아야 함.



4. 정책근거 강화와 행정데이터 활용 사례: 해외 사례

- 영국의 공공목적의 데이터 활용을 위한 거버넌스 체계 구축
 - 2010년 행정데이터 태스크포스를 구성
 - 2012년 태스크포스의 보고서는 행정자료 연구 네트워크(Administrative Data Research Network; ADRN)의 구성과 부처 간 갈등과 데이터 공유에 있어 발생하는 문제들을 정리할 법 제정을 제안.
 - 태스크포스(ADT)는 제3자(TTP: Trusted Third Party)를 통한 연계 방식을 제안. 모든 주체들이 부분 정보만을 다루게 함으로써 프라이버시 침해를 방지.
 - 행정자료 연구센터(ADRC)는 행정데이터를 요청한 연구 프로젝트의 적합성 여부를 따지는 심의기능도 수행.
 - 적합성 기준은 비상업적 연구, 공중 이익에 대한 기여 등임.
 - 자연 및 이공계의 연구보다 경제 및 사회과학 연구 지원에 중점을 둠.
 - ADRN의 후속으로 행정데이터 연구(ADR UK: Administrative Data Research UK) 구성
 - ADR UK는 영국통계청(Offices for National Statistics)와 협력, 연구자들에게 행정데이터 서비스를 제공, 국립통계자료윤리자문회의(NSDEAC)와 협력하여, 행정데이터 활용이 윤리적이고 공익에 부합하도록 하도록 함.
 - NSDEAC의 허가를 받은 연구자는 5년 간 행정데이터 접근을 보장받음.

4. 정책근거 강화와 행정데이터 활용 사례: 해외 사례

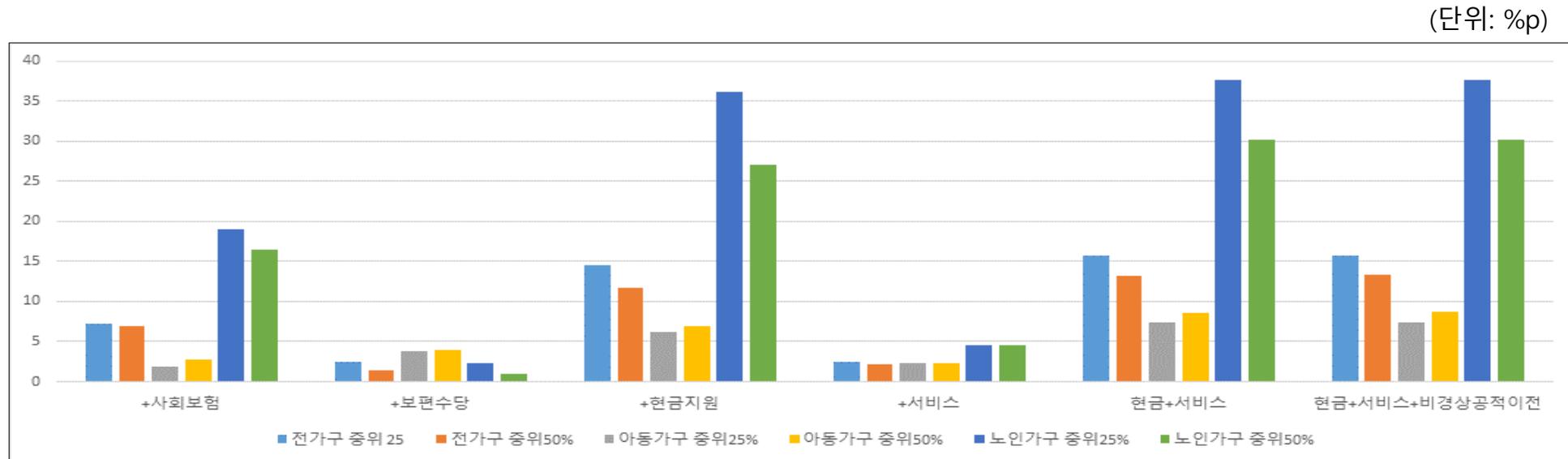
- 미국의 근거기반정책위원회(Commission on Evidence-Based Policymaking)의 구성과 활동
 - 행정자료 활용 가이드라인을 마련, 개별 기관이 권고를 준용하도록 함.
 - 자료 연결과 개인정보보호를 함께 추진되도록 하고, 소득과 임금정보에 대한 접근성을 높이는 등 정책 개선추진
- 2018년 수립된 미국의 증거기반 정책수립 기초법(Foundations for Evidence-Based Policymaking Act of 2018, 일명 Evidence Act)
 - 동 법 제3511조는 '데이터 목록 연방데이터 카탈로그'
 - 동법 제3520조 데이터담당위원회(Chief Data Officer council): 기관 간 데이터 공유에 대한 합의를 촉진하고 장려

4. 정책근거 강화와 행정데이터 활용 사례: 사회보장행정데이터 활용 사례

○ 사회보장급여의 빈곤율 완화효과

- 사회서비스를 비롯한 사회정책 영역, 유형별 효과분석 가능성 확인
- 사회보험의 빈곤율 완화효과는 매우 크고 특히 노인가구에서 그 효과가 크게 나타남.
- 보편수당의 효과를 보면 아동가구에서 빈곤율 완화효과가 큰데 중위 25%와 중위 50% 기준 적용에서 그 차이가 크지 않음.
- 반면 노인가구의 경우 중위 25% 기준일 때 빈곤율 완화효과가 중위 50%일 때보다 큰 것으로 나타남. 노인의 일차소득이 낮고 급여액수준도 높지 않은 때문으로 추정됨.

[그림 2] 사회보장수급의 빈곤율 완화 효과



4. 정책근거 강화와 행정데이터 활용 사례: 사회보장행정데이터 활용 사례

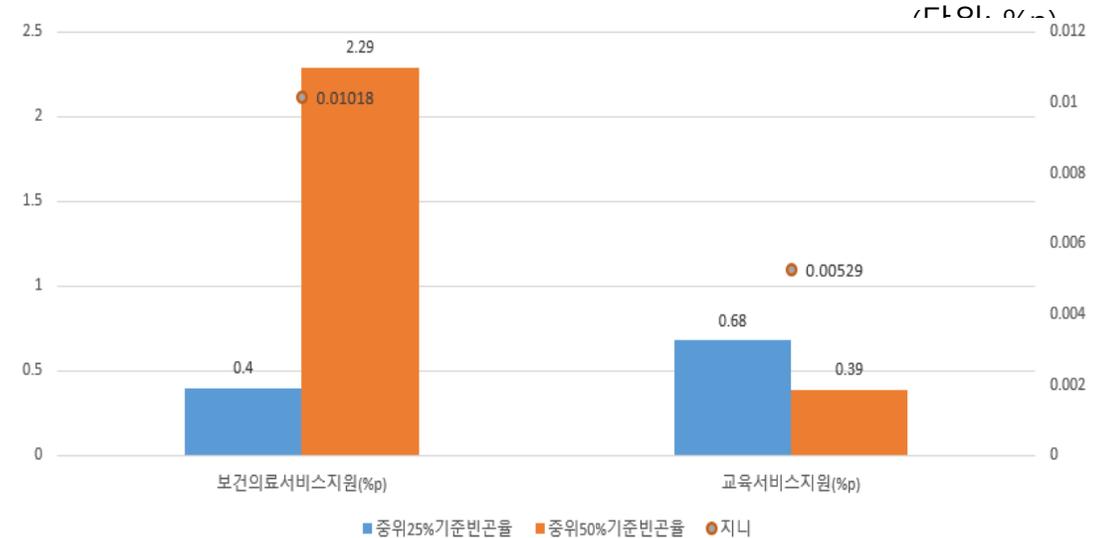
○ 조정가처분소득 불평등

- 서비스 수급 정보(수급여부와 수급액) 활용이 가능하여 관련 지표의 추정 오차를 줄일 수 있음
- 서비스 가산 조정가처분소득 = 가치분소득 + 현금화된 서비스 가치
- 중위소득 50% 기준, 보건의료서비스 반영 조정가처분소득 가구빈곤율은 가치분소득 가구빈곤율보다 약 2.3%p 낮고 교육서비스 반영 조정가처분소득 가구빈곤율은 0.39%p 낮음.

<표 3> 조정가처분소득의 소득재분배 효과

구분	가구						개인				
	가치분 소득 A	조정가처분소득		A-B	A-C	가치분 소득 D	조정가처분소득		D-E	D-F	
		보건의료 B	교육 C				보건의료 E	교육 F			
빈곤율	중위 25% 기준	16.75	16.35	16.07	0.4	0.68	13.63	13.33	13.26	0.3	0.37
	중위 50% 기준	30.05	27.76	29.66	2.29	0.39	25.77	24.73	25.87	1.04	-0.1
빈곤갭 (중위 50% 빈곤선)	17.03	16.08	16.34	0.95	0.69	13.88	13.23	13.47	0.65	0.41	
지니	0.47144	0.46126	0.46615	0.01018	0.00529	0.43430	0.42608	0.42849	0.00822	0.00581	

<그림 3> 조정가처분소득과 가치분소득의 불평등 차이



4. 정책근거 강화와 행정데이터 활용 사례: 사회보장행정데이터 활용 사례

- 경제활동인구의 경제활동 관련 사회보장제도의 소득효과
 - 특정 정책영역의 정책 전반의 효과를 파악할 수 있게 지원
 - 18-64세 가구주 가구의 일차소득을 기준 빈곤율은 28.1%. 경제활동 관련 사회보장제도 수급은 빈곤율을 2.2%p 완화시킴

구분	일차소득 (가)	(가)+근로 장려금	(가)+실업 급여	(가)+모성 보호	(가)+산재	(가)+두루 누리	(가)+청년 내일채움 공제	(가)+청년 구직활동 지원금	(가)+옆의 모두
전가구	0.372	0.366	0.360	0.371	0.368	0.371	0.371	0.371	0.350
18-64세 가구주 가구	0.281	0.276	0.268	0.280	0.279	0.280	0.281	0.281	0.259

4. 정책근거 강화와 행정데이터 활용 사례: 사회보장행정데이터 활용 사례

- 2017년 표본에 2015년 정보를 추가 연계한 개인 단위 불완전 균형패널 분석
 - 패널로 자료를 구축하여 사회이동성 등 분석이 가능함을 시사
 - 2017년 장학금 지원 수급자 중 2015년 교육급여 수급자인 학생은 0.67%, 비수급이었던 학생은 7.68%.
 - 저소득층 자녀의 낮은 대학진학률을 시사. 학력격차를 완화하기 위한 정책 지점을 찾고 저소득층의 취업지원에 대한 적극적 고려 필요

〈표〉 2017년 교육지원 수급자의 과거 수급이력

(단위: 명, 비율)

구분	만15-17	만18-25	계(만 12-25)	계(전인구)
17 장학-15교육	-	1.05	0.67	0.11
17장학-15장학	-	5.61	3.58	0.66
17장학-15비수급	0.01	12.04	7.68	1.44
17교육-15교육	2.07	0.34	0.92	0.27
17교육-15비수급	2.07	0.34	0.92	0.27
17비수급-15교육	0.89	1.47	1.22	0.23
17비수급-15비수급	91.24	67.93	76.68	95.23
기타	3.73	11.22	8.33	1.79

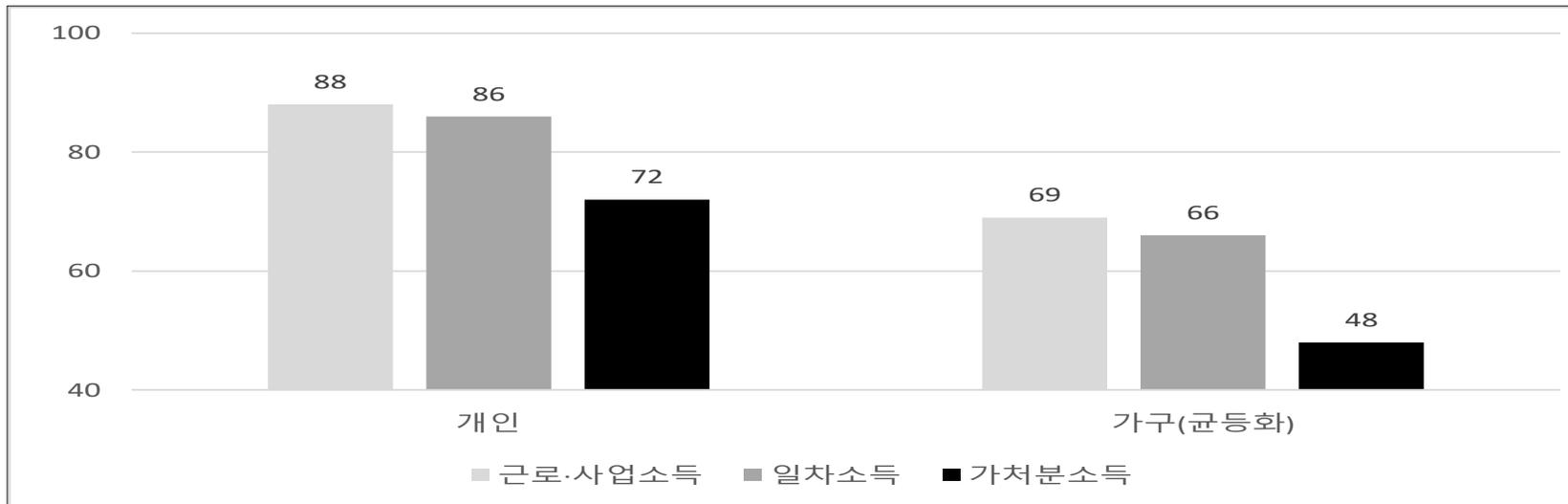
자료: 행정안전부 주민등록명부, 16개 기관 및 국세청 추출결합 원자료.

4. 정책근거 강화와 행정데이터 활용 사례: 사회보장행정데이터 활용 사례

○ 노인 개인과 가구의 사회보장제도수급 효과

- 설문조사 자료에서 근로소득은 가구원 개인 단위 근로소득, 1차 소득 등의 파악이 쉽지 않았음. 개인 일차소득을 별도 분석할 수 있는 사회보장행정데이터의 이점을 보여줌.
- 노인의 개인 일차소득 빈곤율은 86%, 개인 가처분소득 빈곤율은 72%.
- 노인 가구의 일차소득 빈곤율은 66%, 노인가구 가처분소득 빈곤율은 48%.
- 가구 내 소득 공유가 노인 빈곤을 크게 완화시킴.

[그림 4] 노인의 개인소득 및 가구소득(균등화) 빈곤율



5. 쟁점과 향후 과제: 사회보장행정데이터의 강점과 개선과제

<표 5> 사회보장행정데이터의 강점과 개선과제

	해당 내용	잠재력과 과제
강점	<표본의 크기> - 전 가구의 약 20% 표본	- 연령별, 지역별 분석 구체화 가능 - 작은 규모의 정책분석도 가능
	<높은 대표성> - 통계청 인구·가구 DB 기반 표본추출	- 통계 안정성 확보
	<정보 포괄성> - 성, 연령, 거주지, 소득, 재산, 사회정책수급 등 주요 정보의 연계(30 이상 기관의 570여개 정보)	- 정책이나 부처의 경계를 넘는 종합적인 정책분석, 조정 기반 마련
	<정보의 정확성> - 공적이전(현금과 현물이전), 세금, 사회보험료 - 상위 소득 및 재산	- 공적이전 수급효과에 대한 정확한 파악이 가능
개선 필요 사항	<표본 중 일부 누락> - 표본 중 집단시설가구거주자, 외국인의 누락	- 취약계층인 시설거주 국민에 대한 분석에서 한계 - 증가하는 외국인에 대한 고려 필요 - 시설거주자, 외국인을 표본으로 통합, 데이터 보완 필요
	<정보포괄성 부족> - 일부 소득재산정보의 누락 · 조세행정으로 농업소득, 비과세 소득 등 누락 · 금융재산과 금융소득 등 금융정보 누락 - 일부 공적이전정보의 누락 · 매입전세임대주택정보 누락 · 지방자치단체의 현금지원, 주거지원, 돌봄 등 사회서비스 정보의 누락	- 금융정보, 임대인의 부채정보와 재산소득, 임차인의 금융재산정보 보완 - 지방자치단체 정책 정보의 연계 추진, 구축데이터 공유
	<욕구, 연성(인식)정보의 부족> - 필수재 구입을 위한 지출 등 욕구정보의 부족 - 정책에 대한 태도, 만족도 등 정보 부족	- 인식, 태도 정보를 파악하는 설문조사자료와의 연계

5. 쟁점과 향후 과제: 사회보장행정데이터의 형태 다양화와 과제

- 현재 사회보장행정데이터는 횡단면 자료 형태로 매년 구축하는 방식
 - 데이터가 축적되면서 사회정책 효과에 대한 변화 분석 지원
- 향후 다양한 형태의 데이터를 구축하고 활용 활성화 강화를 위한 과제

패널데이터 구축

교육과 중장기 사회이동성 분석, 수급 역동성 분석

시설정보와 결합

지역사회기반 서비스 지원 효과 파악 및 정책기반 점검

지방자치단체 정보와 결합

정부 간 정책 상보성, 사회정책 최종 효과 분석

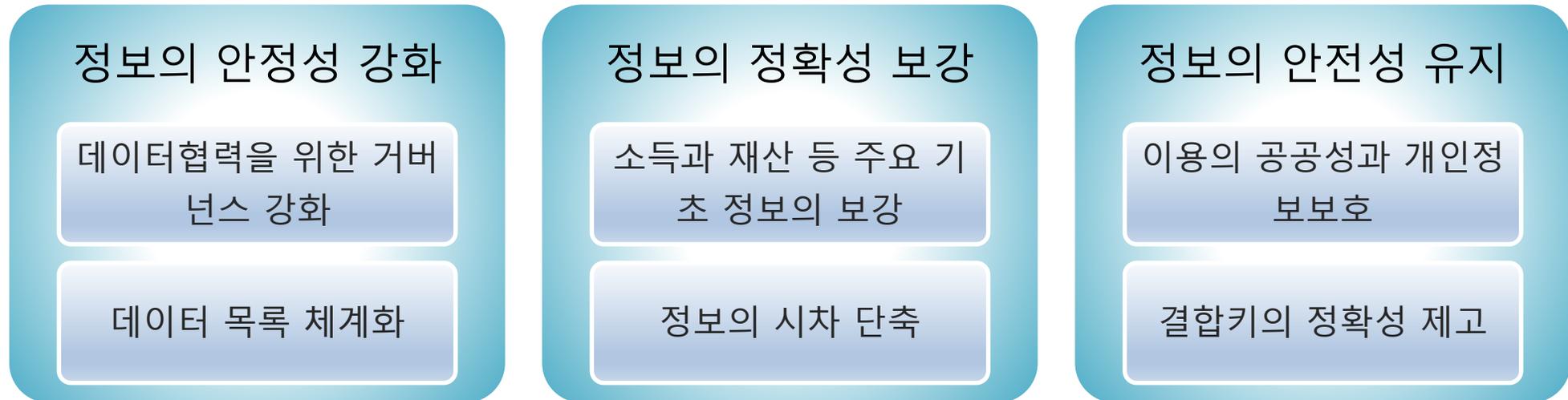
설문조사자료와 결합

정책에 대한 태도변화, 정책 만족도 등 인식 분석

5. 쟁점과 향후 과제: 사회보장행정데이터의 활용을 위한 과제

- 활용의 공공성 고려
 - 사회보장제도의 심의·조정, 이를 뒷받침하는 연구의 수행을 목적으로 제한
 - 공공 우선의 단계적 데이터 개방과 이용 절차의 객관성 확보
- 투입(input) 중심의 평가를 극복하고 정책의 효과, 영향 평가 강화
 - 공동의 정책목표를 가지는 정책군 평가 등 정책평가의 고도화 지원(사례: 영국 2000년대 중반 공공서비스협정(Public Service Agreement, PSA))

[그림 6] 사회보장행정데이터 활용을 위한 과제



5. 쟁점과 향후 과제: 유럽의 경험과 시사점

- 행정자료 활용에서 협력 문화의 영향
 - 유럽의 통계청들이 행정자료에 대한 성공적인 접근·활용에 기여하는 요인으로 꼽은 것들은 아래와 같음.
 - 자료보유자들 간의 협력이 가장 중요한 기여 요인
 - 데이터의 질, 데이터 교환을 위한 기술을 언급
 - 요는 데이터 추출에서 협력문화, 거버넌스가 주요

<표 > 성공적인 행정자료에 대한 접근과 활용에 기여하는 요인들

Success factor	Number of data sources
Close cooperation with data owner	43
Legal framework	27
Data quality	20
Data exchange technology	17
Existence of uniform / unique identifiers	10
Collaboration agreement	10
Availability of linkage information	5
Timeliness in data release	5
Right granted by the data protection authority	4
Experience of the NSI with the data	3
Political support	2
Data owner's understanding of NSI's needs	2
Availability of metadata	2
Adoption of common technical solutions	2
Data not subject to changes over time	1
Data protection technology	1
Payment agreement	1
Commitment of resources by the NSI	1
Recognition by the data owners of the importance of the data for the NSI	1

Note: Multiple answers were allowed.

5. 쟁점과 향후 과제: 유럽의 행정자료 구축과 활용 평가 및 시사점

- 유로스타트(eurostat)는 회원국의 법적, 제도적 환경에 대한 평가를 담은 보고서를 발간

○ 해당 평가의 제 측면은 방대

① 법적 기반,

② 제도적·조직적 측면: 전문성 등을 포괄하는 자원과 능력의 적절성, 행정자료 활용을 점검하고 지원하는 절차와 정책들, 협력의 기제, 개인정보보호를 보증하는 실천들, 자료의 질을 보증하는 노력 등,

③ 데이터 처리자들과의 협력과 광의의 제도적 환경: 행정자료 활용에 영향을 주는 담당자의 문화와 공공의 신뢰, 협력을 위한 협정 등

④ 데이터의 질과 기술적 측면: 방법론에 대한 설명, 자료의 질 점검과 자료의 질 보증을 위한 협력, 공통의 식별자, 자료 연계를 위한 충분한 정보 교환을 위한 기술 공유 등

- 유럽 통계청의 경험을 기반으로 보면,

★ 법적 개선만으로는 행정데이터 기반 정책근거 강화가 완성되기 어렵다는 점,

★ 협력 문화 조성, 데이터 질 개선을 위한 협력 등이 병행추진되어야 함을 확인가능

5. 쟁점과 향후 과제: 행정데이터 구축과 활용 관련 쟁점 정리

- 아래의 쟁점들에 대한 관계자의 입장은 서로 상이, 논의와 대안 모색이 필요
- 연구목적과 이용자 범위의 제한: 산업용 연구 포함(우선) vs 공공 목적으로 한정(우선)
 - 북유럽 사례는 선행사례이자 모범사례임에도 행정자료 활용의 이용자 범위를 상당히 제한적으로 설정하고 점차 확대하는 방식을 선택
- 활용 목적: 개인표적화 vs 철저한 개인정보 보호
 - 기능적 분리(functional separation)의 원칙(1977년 미국에서 개인정보보호연구위원회)
 - 집단차원의 분석을 넘어 개인이나 개별 가구를 표적화하는 행정자료 활용에 대한 우려 존재
- 표본의 포괄성: 전수 vs 표본
 - 정책분석을 위한 자료에서 전수 정보의 필요성이 뒷받침되지 않는다면 개인정보보호의 원칙, 즉 최소정보 수집에 위배될 수 있음.
 - 매우 소규모집단에 대한 분석에서는 전수 필요성 있을 수 있으나 근거확인 필요

5. 쟁점과 향후 과제: 행정데이터 구축과 활용 관련 쟁점 정리

- 연계키: 정확성(주민번호) vs 안전성(비식별조치된 성, 연령 등 대체정보)
 - 현재 활용 중인 결합키(생년월일, 성, 이름의 가명처리된 정보)의 한계
 - 2019년 관련 연구의 경우 표본 중 중복개인은 12.6%, 중복개인이 속한 가구는 25.7%, 중복가구에 속한 개인은 33.1%, 2022년 표본 30%의 가구 중 약 20% 잔존
 - 중복 결합키를 제거 후 가중치 조정으로 필요한 수준의 대표성 확보 가능
 - 활용 환경의 체계화: 플랫폼의 다변화 vs 데이터 신청 및 지원체계 일원화
 - 어느 정도의 다변화를 허용하는 것이 효과적이고 효율적인가에 대한 검토 필요
 - 정보의 관리: 여러 영역 정보의 통합 관리 vs 개별 영역의 정보관리와 연계강화
 - 산업, 사회보장 등 다양한 영역의 정보를 하나의 데이터 셋에 모으자는 주장도 존재.
 - 효율성과 정보보호의 위험을 모두 고려할 필요. 정보보호에서의 위험이 우선 고려되어야 할 것
 - 정보의 보관: 맞춤형 정보의 파기 vs 연구용 정보의 보관
 - 해외에서는 안정성이 확보된 자료의 경우 보관, 연구와 결과 확인 기반 마련
- ➡ 해법 모색을 위해 행정데이터 관리 및 활용을 체계화하기 위한 거버넌스 강화 필요

감사합니다