

데이터 기반 사회정책 의제 도출

설동훈

한국사회학회 회장
전북대학교 사회학과 교수

개요

▶ 연구배경 및 필요성

- ▶ 교육부 연구용역 '데이터 기반 사회정책 실증연구 의제 도출'을 위해
설동훈(전북대)·계봉오(국민대)·최성수(연세대)·황선재(충남대)가 공동연구
- ▶ 행정 빅데이터의 특징과 필요성

▶ 사회정책 관련 연구주제 1: 생애과정의 이해와 정책적 대응

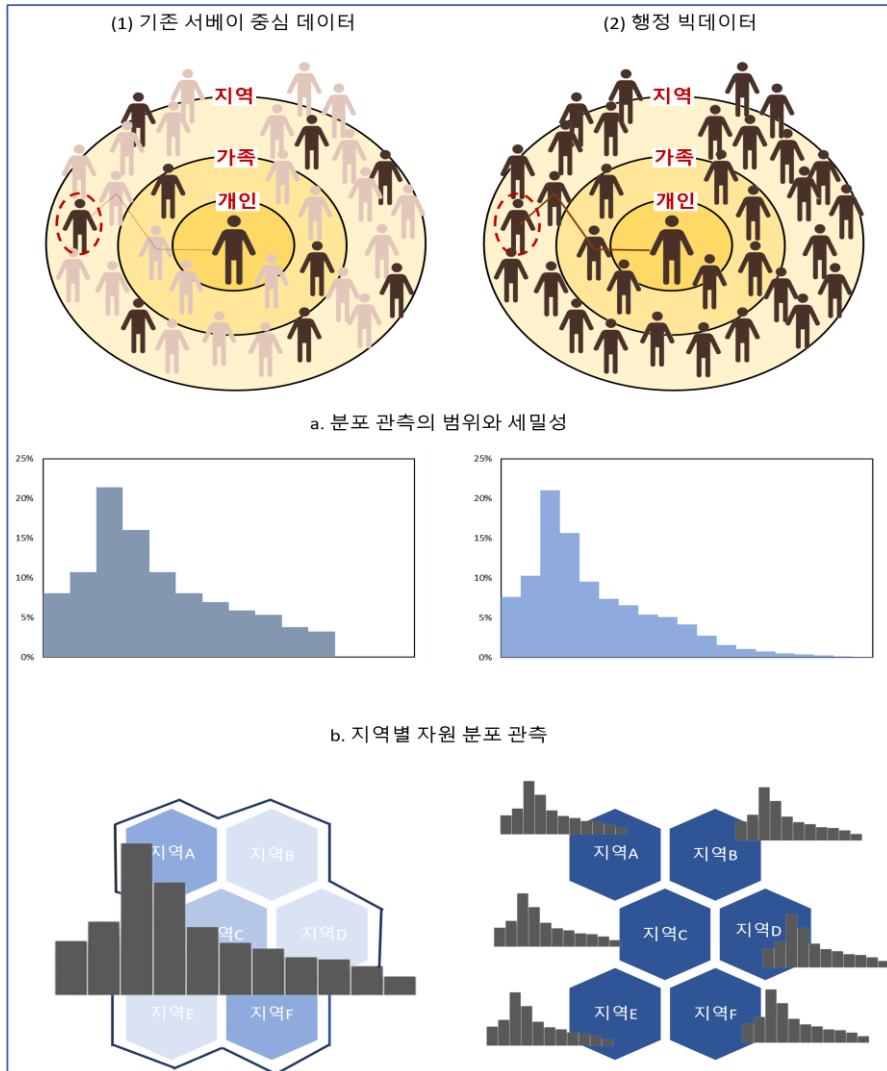
- ▶ 1. 초·중등학교 분리의 양상과 추세
- ▶ 2. 청년-성인기 이행
- ▶ 3. 가구소득 계층별 출산율 추이
- ▶ 4. 세대 간 소득계층 이동

▶ 사회정책 관련 연구주제 2: 지역사회 불평등과 정책적 대응

- ▶ 5. 소득수준별 거주지 분리 추이
- ▶ 6. 지역 소득불평등과 지역사회 삶의 질

▶ 요약 및 함의

연구배경: 행정 빅데이터 활용의 필요성



▶ 행정 빅데이터의 특성

- ▶ 전수자료
- ▶ 개인들의 연결망
- ▶ 관측의 범위와 세밀성
- ▶ 분포의 지역 간 비교 가능성

▶ 행정 빅데이터와 정책평가

- ▶ 행정자료
- ▶ 서베이 자료를 활용한 정책 평가의 어려움
- ▶ 정책 수혜자에 대한 직접적 관찰이 가능함

연구의 필요성: 행정 빅데이터, 생애과정, 지역

- ▶ 행정 빅데이터를 활용한 생애과정의 이해와 정책적 대응
 - ▶ 학교 교육, 성인기 이행(transition to adulthood), 가족형성과 출산 등 사회정책의 주요 의제는 생애과정의 단계별로 다양하게 나타남
 - ▶ 사회정책 의제 발굴은 생애과정에 대한 경험적 분석에 근거를 두어야 함
 - ▶ 행정 빅데이터를 활용한 경험적 분석과 정책적 대응 방안 모색
- ▶ 행정 빅데이터를 활용한 지역사회 불평등의 이해와 정책적 대응
 - ▶ 한국사회와 지역 사회 불평등: 지역 간 격차, 수도권 집중, 인구감소 위험 지역 증가
 - ▶ 지역 간 격차와 지역 내 격차가 공존
 - ▶ 행정 빅데이터를 활용한 지역 수준 분석과 정책적 대응 방안 모색

1. 초·중등 학교분리의 양상과 추세

▶ 학교분리(school segregation)의 정책적 중요성

- ▶ 학교분리: 지역마다 다양한 사회경제적 배경 출신(고소득 전문직 지위, 기초생활수급 지위 등) 학생들이 같은 학교에 함께 통합되어 다니거나 다른 학교들에 분리되어 다니는 정도를 의미
- ▶ 아이들이 자신과 다른 여건과 환경의 친구들과 어우러져 성장하는 경험이 공교육에서 어떻게 구현되고 있는지 보여줌
- ▶ 학교분리 양상 파악은 다양한 배경/가치의 이해와 포용을 바탕으로 사회통합을 이룰 성숙한 민주시민 양성이란 공교육의 본질적이면서도 절실한 시의성을 가진 목표를 이루기 위해 필수
- ▶ 학교분리가 심화된 환경일수록 교육성취 격차가 크고 개인중심적인 사고가 지배적이며 사회이동의 역동성도 약한 경향
- ▶ 지역별 격차 심화, 지역 재개발/재정비 사업, 부동산 가격 상승 등으로 거주지분리가 심해지고 이에 따른 학교분리 역시 심화되었을 개연성 높음

1. 초·중등 학교분리의 양상과 추세

▶ 학교분리 측정을 위한 데이터의 부재

- ▶ 현재 한국에서 학교분리를 측정할 수 있는 데이터는 부재하며 당연히 학교분리 양상에 대한 연구와 지식도 사실상 없음
- ▶ 학교분리는 지리적, 공간적 분포에 대한 것이므로 측정을 위해서는 다음과 같은 정보가 요구
 - ▶ 학교 내 학생들의 사회경제적 배경과 경험, 성취 정보 및 학교 행정자료 기반 정보가 담긴 마이크로데이터, 학교가 속해 있는 지역(읍/면/동, 시/군/구, 혹은 교육지원청 관할지역)의 식별정보
- ▶ 이들 정보는 아동, 청소년 서베이(교육종단연구, 교육고용패널, 아동청소년패널, 국가수준학업성취도평가 등) 및 EDSS의 행정자료 등을 통해 개별적으로 제공되고 있으나 연계되어 제공되고 있지 않음

▶ 대안과 제안

- ▶ 축적되어 있는 다년도의 조사자료와 행정자료를 연계함으로써 개인정보 유출의 위험 없이 학교와 지역 수준에서 거주지분리 및 학교분리의 다양한 지수를 측정 가능
- ▶ 이를 데이터베이스화하여 정책 및 연구를 위한 필수적 자료로 제공 가능

2. 청년 성인기 이행: 중등-고등교육과 노동시장 정착 및 가족형성

▶ 정책적 중요성

- ▶ 정치적 담론의 핵심 "청년": 정당, 정파를 막론하고 정치는 "청년"을 핵심 대상이자 주체로 호명해 왔으며, 수많은 (설익은) 정책과 담론들이 제안되고 소비되어 왔음
- ▶ 초저출산 현상을 비롯 한국사회의 심각한 인구학적 위기의 중심에 청년들이 존재하며 이들의 성인기 이행 과정에 대한 정확한 이해/진단이 필수
- ▶ 중등 및 고등교육, 노동 및 산업, 젠더와 가족의 변화와 정책이 중층적으로 교차하며 형성되기 때문에 전체를 함께 아우르는 총체적(holistic) 접근이 요구

▶ 데이터의 한계

- ▶ 풍성한 담론과 대조적으로 이런 총체적인 관찰과 분석, 진단에 필요한 데이터는 매우 빈약
- ▶ 고교 졸업 이후 30대 중반에 이르기까지 청년의 삶을 추적하는 코호트 기반 종단 데이터는 매우 제한적으로만 존재(교육고용패널). 시대적 변화를 파악할 수 있는 반복 횡단면 데이터 역시 한계가 뚜렷(대졸자직업이동경로조사)

2. 청년 성인기 이행: 중등-고등교육과 노동시장 정착 및 가족형성

▶ 대안과 제안

- ▶ 교육고용패널(직업능력연구원), 교육종단연구(교육개발원), 대졸자직업이동경로조사(한국고용정보원) 등 10대 후반-20대 초반 정보를 갖춘 서베이 자료를 행정자료(건강보험공단DB 및 통계청 인구·주택총조사)와 연계하여 30대 중·후반 소득, 직업, 거주지, 가족형성 정보를 파악 가능한 통합적 패널자료 구축
- ▶ 연계과정에서의 기술적 난점과 개인 식별 문제는 기계학습방법(machine learning) 등 통계적 연계 방법 활용해 해결
- ▶ 2019년 추적조사 자료를 수집/제공한 교육고용패널의 예를 따라, 대졸자직업이동경로조사와 같은 대규모 횡단자료도 졸업 10년 후 추적조사를 통해 데이터의 사회적, 정책적 가치를 높일 수 있음
- ▶ 이러한 자료의 구축을 통해서 성인기 이행 과정의 주요 사건들인 졸업, 취업, 결혼, 거주지 이동, 출산의 세밀한 분석이 가능해 질 것으로 기대

3. 가구소득 계층별 출산율 차이

▶ 연구의 배경 및 중요성

- ▶ 2022년 현재 합계출산율은 0.78로 유례를 찾기 어려운 초저출산이 지속되고 있음
- ▶ 불평등의 증가가 초저출산 지속의 원인으로 지적되고 있지만, 불평등과 출산율의 관계에 대한 경험적 연구는 일천한 상황임
- ▶ 불평등과 출산율의 관계를 이해하기 위한 일차적인 작업은 가구소득 계층별 출산율 차이에 대한 차이 분석임
- ▶ 그런데, 가구소득 수준에 따른 출산율 차이를 분석할 수 있는 대규모 서베이 자료는 부족함
- ▶ 이러한 자료의 부재 때문에 경험적으로 검증되지 않은 주장들에 기반을 둔 정책 실행으로 연결될 위험이 존재함

3. 가구소득 계층별 출산율 차이

▶ 대안과 제안

- ▶ 건강보험 빅데이터는 가구소득 정보와 출산 정보를 결합한 전수 자료를 제공해 줄 수 있음
- ▶ 건강보험 납부액을 소득 수준으로 활용해서 소득수준과 출산의 관계를 분석한 연구는 이미 존재하나(이철희 2022), 뚜렷한 한계를 안고 있음
 - ▶ 이 연구의 분석 단위는 건강보험 납부자와 이들의 피부양자임
 - ▶ 이는 일반적인 가구와 다른 경우가 많으며 해석이 어려움
- ▶ 따라서, 건강보험 빅데이터와 주민등록 자료를 연계해서 가구를 명확히 규정하고, 이를 분석 단위로 활용한 실증연구가 필요함
- ▶ 이를 통해 가구 소득 계층별 출산율 차이의 추이를 분석하고, 이를 기반으로 소득 불평등과 출산율 변화의 관계를 이해하는 데 기여할 수 있음

▶ 확장

- ▶ 인구주택총조사 마이크로 데이터, 다른 행정 빅데이터(고용보험 자료 등)의 연계를 통해서 개인의 사회경제적 특성(예: 교육수준)이나 생애과정의 사건(예: 취업, 실업)과 출산의 관계를 세밀하게 분석할 수 있음

4. 세대 간 소득계층 이동

▶ 연구의 배경 및 중요성

- ▶ 세대 간 사회이동은 공정성 문제와 직결되는 문제이며, 기회의 사다리가 좁아지고 있다는 일반적인 인식의 경험적 근거를 제공해 줌
- ▶ 직업지위를 기준으로 한 세대 간 사회이동 연구는 대체로 세대 간 사회이동이 줄어들지 않고 있음을 보여주고 있음
- ▶ 소득을 기준으로 한 연구는 자료의 한계로 인해서 정확한 추정치를 생산하지 못하고 있음
- ▶ 세대 간 사회이동이 갖는 중요성과 경험적 근거의 부재가 결합하여 건설적인 논의가 진전되고 있지 못하고 있음
- ▶ 따라서, 세대 간 소득계층 이동 연구는 매우 중요한 정책적 함의를 가짐

4. 세대 간 소득계층 이동

▶ 대안과 제안

- ▶ 건강보험 빅데이터의 가족DB를 활용은 세대 간 소득계층 이동 연구를 가능하게 함
- ▶ 한 선행연구(계봉오 외 2018)는 2002~2018년 건강보험 빅데이터를 활용해서 세대 간 소득계층 이동을 분석했는데, 2005년 이전 자료의 낮은 신뢰성 때문에 활용도가 낮았음
- ▶ 2005년에 청소년기를 보냈고 2023년에 30대에 이른 1990년 전후 출생 코호트의 청소년기 가구소득과 2023년 소득의 관계를 분석함으로써 한국사회의 세대 간 소득계층 이동의 정도를 분석할 수 있음
- ▶ 청소년기 거주지역 등에 따른 소득계층 이동의 차이 또한 분석할 수 있음

▶ 확장 및 함의

- ▶ 인구주택총조사 마이크로 자료나 각종 행정DB의 연계를 통해서 세대 간 소득계층 이동을 매개하는 요인(교육, 취업 등)을 분석할 수 있는 가능성을 열 수 있음
- ▶ 이를 통해 세대 간 소득계층 이동의 명확히 파악하고 공정한 기회 창출을 위한 정책적 과제를 도출할 수 있음

5. 소득수준별 거주지 분리 추이

▶ 연구의 중요성 및 의의

- ▶ 최근 한국 사회에서 소득 불평등 현실과 인식 간의 괴리가 증가하고 있음
 - ▶ 현실은 악화되었다고 보기 어려우나, 인식은 지속적으로 악화중임
- ▶ 이는 기회 및 과정에 대한 (불)공정성, 삶의 불안정성, 계층이동 가능성/기대와 밀접한 관련이 있음
- ▶ 또 하나의 중요한 측면은 사회계층별 거주지 분리(residential segregation) 증대에 따른 상호작용의 감소 및 상대적 박탈감 증가
- ▶ 특히 한국 사회에서 중요한 의미를 갖는 부동산, 거주지가 소득수준별로 분리되어 있는지, 그렇다면 그 추이는 어떤 지를 실증적으로 살펴보는 작업은 학술적인 측면 뿐 아니라 정책적으로도 매우 중요함

5. 소득수준별 거주지 분리 추이

▶ 기존 연구 및 데이터의 한계

- ▶ 소득수준별 거주지 분리 추이를 유의미한 차원에서 분석하기 위해서는 읍면동 수준에서 측정된 대량의 가구소득 및 주소지 정보가 필요함
- ▶ 일반적으로 활용되는 서베이 자료를 활용해서는 1) 읍면동 수준의 자료를 확보하기 어렵고, 2) 확보하였다 하더라도 표본수가 지나치게 적어 유의미한 실증분석 결과를 도출하기 어려움

▶ 대안 및 제안

- ▶ 가구소득과 주소지에 대한 전수에 가까운 자료를 읍면동 수준에서 확보할 수 있는 자료(예: 국민건강보험공단 연구DB)를 활용하여 원자료를 확보함
- ▶ 해당자료와 함께 거주지 분리를 측정하는 전통적인 분석도구(예: 상이지수, 고립지수, 델타지수 등)를 활용하여 소득수준별 거주지 분리 추이를 측정함

6. 지역 소득불평등과 지역사회 삶의 질

▶ 연구의 중요성 및 의의

- ▶ 최근 한국사회에서 인구감소, 인구절벽, 인구소멸 등과 같은 용어들이 회자되고 있음
- ▶ 해당 문제들은 기실 수도권만의 문제가 아니라 지방의 문제인만큼, 관련 현상들을 지방의 관점에서 집중적으로 (재)조명하는 작업이 요구됨
- ▶ 지방의 인구감소, 인구절벽, 인구소멸 현상의 핵심은 일자리 및 ‘삶자리’의 부재임
- ▶ 오늘날의 청년세대에게는 일자리 뿐 아니라 삶자리(또는 거주지의 삶의 질) 또한 중요한 문제이므로, 지역사회의 삶의 질을 측정할 수 있는 다양한 지표를 개발하고 분석하는 작업이 필요함
- ▶ 그 중에서 가장 중요한 지표는 지역의 소득수준 및 소득불평등도임

6. 지역 소득불평등과 지역사회 삶의 질

▶ 기존 연구 및 데이터의 한계

- ▶ 기존 연구는 주로 국가 전체를 측정 및 분석 단위로 삼고 있으나, 각 개인에 대한 불평등의 실질적인 영향력은 '거주지' 및 '생활권'이라는 물리적 공간과 인식의 범위에 의해 매개되는바, 불평등은 국가 단위에서 뿐 아니라 지역 수준에서도 측정되고 그 함의를 살펴보는 것이 바람직함
- ▶ 지역내 총생산 등으로 측정되는 지역의 사회경제적 '수준' 뿐 아니라 지역 지니계수와 같은 불평등 지표로 측정되는 사회경제적 '이질성'의 정도도 함께 분석해야 지역 단위의 삶의 질 및 사회의 질을 제대로 파악할 수 있음

▶ 대안 및 제안

- ▶ 가구소득과 주소지에 대한 전수에 가까운 자료를 읍면동 수준에서 확보할 수 있는 자료(예: 국민건강보험공단 연구DB)를 활용하여 원자료를 확보함
- ▶ 이 자료를 바탕으로 읍면동(혹은 시군구) 수준의 지역 소득 불평등도를 측정하고, 이와 관련된 다양한 삶의 질 지표들(예: 기대수명, 시민의식, 사회적 고립도 등)과의 관계를 실증적으로 분석함

요약 및 함의

- ▶ 행정 빅데이터 활용의 필요성
 - ▶ 전수자료: 개인들의 연결망, 관측의 범위와 세밀성, 지역 간 비교
 - ▶ 행정자료: 정책평가
- ▶ 사회정책 관련 연구주제
 - ▶ 생애과정의 이해와 정책적 대응: 초·중등학교 분리의 양상과 추세, 청년-성인기 이행, 가구소득 계층별 출산율 추이, 세대 간 소득계층 이동
 - ▶ 지역사회 불평등의 이해와 정책적 대응: 소득수준별 거주지 분리 추이, 지역 소득불평등과 지역사회 삶의 질
- ▶ 행정 빅데이터의 실증연구 활용을 통해 경험적 근거에 기반한 정책 실행 및 평가