

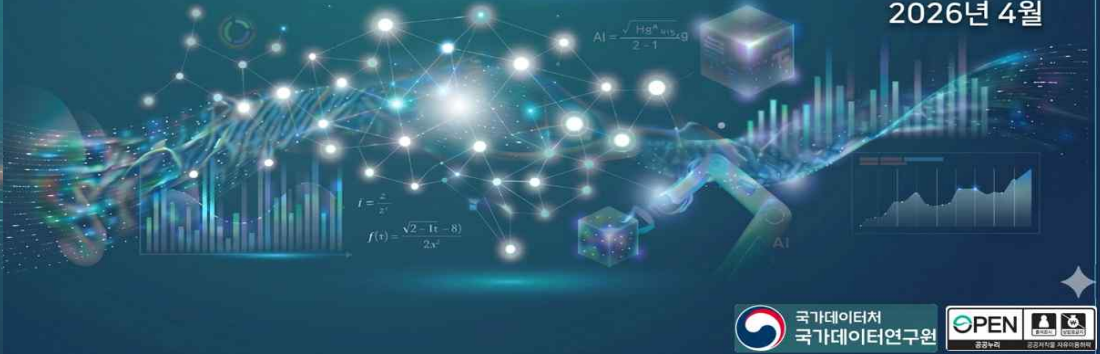
데이터 리서치 브리프

- ◆ 소득만으로는 보이지 않는 가계경제, '소비지표'로 더 정교하게 읽어낸다. 1
- ◆ '보는 통계'에서 '읽는 통계'로의 전환 6
- ◆ 소득이동통계로 본 한국의 이동성 구조: 이동성과 고착의 공존 ... 11



국가데이터처
국가데이터연구원





국가데이터서
국가데이터연구원



변준석

국가데이터연구원
정책통계연구실

송정현

국가데이터연구원
정책통계연구실

소득만으로는 보이지 않는 가계경제, '소비지표'로 더 정교하게 읽어낸다.

- 2011~2024 가계동향조사 기반의 계층별 소비 구조 및 생활비 부담 진단 -

- 소득 불평등 지수보다 **소비 불평등 지수(지니계수)**가 낮게 나타나며, 이는 가계 소득 변동에도 불구하고 일정한 생활 수준을 유지하려는 '소비의 하방 경직성'을 잘 보여준다.
- **팔마비율(소득 상위 10% 점유율/소득 하위 40% 점유율)**은 국제적으로 통용되는 지니계수와 상관관계가 매우 높을 뿐 아니라 분위별 점유율의 비율이라는 점에서 직관적으로 이해하기 용이한 지표이다.
- 외식과 배달 비중이 급증함에 따라, 단순히 '식료품비' 비중만 계산하는 전통적인 엔겔계수뿐 아니라 '식사비'까지 포함하는 **광의의 엔겔계수** 도입을 검토할 필요가 있다.
- **가계의 주거비 부담(슈바베계수)**을 산출할 때는 실제 지출하는 월세뿐 아니라 전세의 기회비용과 자가 가구의 의제주거비를 반영할 때 주거비 부담의 실질을 더 잘 나타낼 수 있다.
- **12대 소비지출 비목** 중 실질지출금액과 명목지출금액에서 차지하는 비중이 모두 하락하는 비목은 교육이고, 모두 상승하는 비목은 보건, 음식·숙박, 오락·문화이다. 교육과 보건 비목의 변화는 저출산·고령화라는 인구구조 변화의 결과이고, 음식·숙박, 오락·문화의 변화는 경제성장의 결과로 이해할 수 있다.
- 반려동물 관리, 성인 자기 계발, 운동시설 이용 지출이 새로운 필수 소비로 자리 잡으며 가계 소비 구조의 질적 전환이 일어나고 있음을 알 수 있다.

연구 배경 및 목적

- (연구 필요성) 소득은 가계의 자원 확보 능력을 보여주지만, '무엇을 소비하는가'는 삶의 질과 직결됨. 특히 고물가와 고금리가 지속되는 최근 경제 상황에서, 소득 지표만으로는 서민들이 체감하는 생활비 부담과 구조적 변화를 포착하는 데 한계가 있음

👉 (연구 질문) "가계동향조사 데이터를 활용해 소득 중심 지표의 사각지대를 해소하고, 국민의 삶의 질을 더 정교하게 측정할 수 있는 소비지표 체계는 무엇인가?"

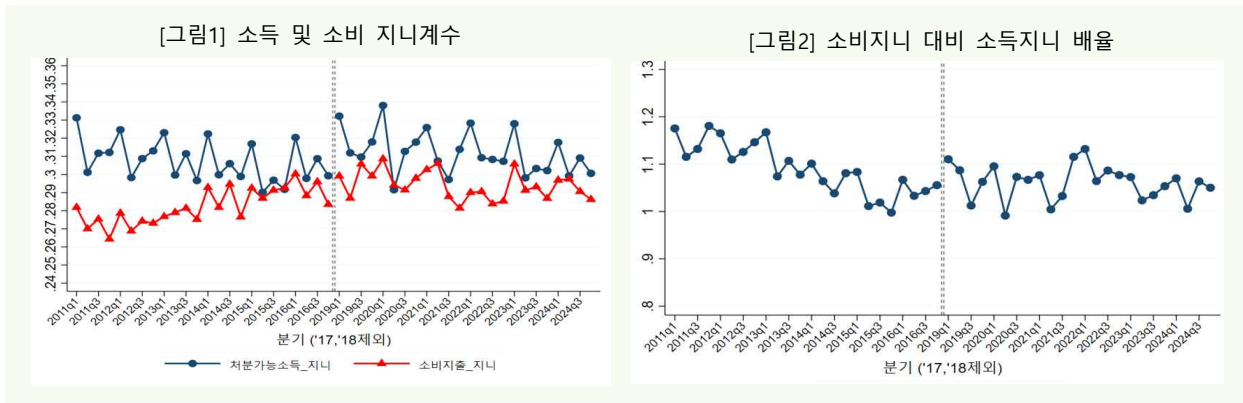
데이터 및 방법

- (활용 데이터) 가계동향조사 마이크로데이터 (2011년~2024년)
* 2017년, 2018년 제외
- (표본 및 주요 변수) 연간 약 7,200가구의 지출 항목(식료품비, 주류, 주거·수도·광열, 보건, 교육 등 12대 지출 비목)

주요 결과

① 소비로 본 분배: 소득보다 안정적인 삶의 흐름

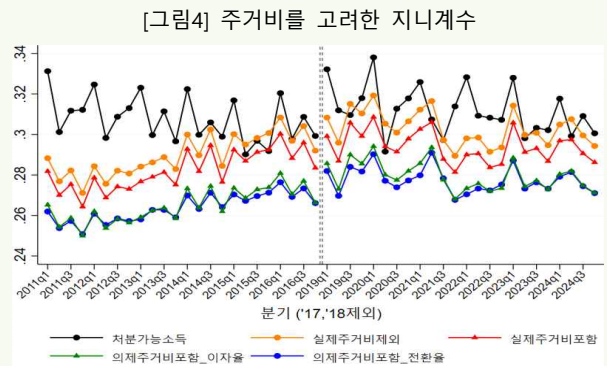
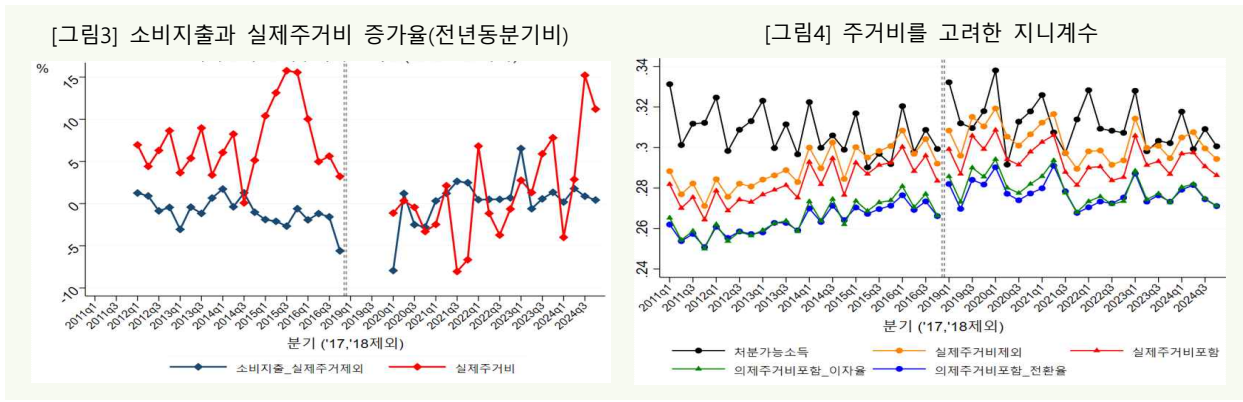
분석 결과, 소비 지니계수는 소득 지니계수보다 지속적으로 낮게 유지됨. 이는 가계가 소득이 일시적으로 줄어도 저축이나 대출을 통해 소비 수준을 일정 수준으로 유지하려고 소비 평활화가 나타나고 있음을 의미함[그림1]. 다만, 최근 소득과 소비지표가 동조화되는 경향을 나타내고 있음[그림2].



2010년 이후 다른 소비지출보다 실제 주거비(월세 등 실제 가구가 주거 서비스 비용으로 지출하는 비용)의 증가율이 더 심하게 증가함을 알 수 있음[그림3].

[그림4]를 보면 가구가 주거비를 포함하지 않았을 경우 소득(처분가능소득)과 소비(실제 주거비제외)의 지니계수는 2019년 이후 상당한 정도의 동조성을 보이지만, 주거비를 포함하면 소득(처분가능소득)지니에 비하여 소비지니가 낮아지는 경향을 확인할 수 있음. 특히 실제 주거비뿐 아니라 자가의 의제주거비(내 집에 살면서 주거 서비스를 스스로 소비하는 그것으로 보아 계산하는 가상의 임대료)까지 포함하면 소비지니(의제주거비 이자율 또는 전환율 모두)는 소득지니와 여전히 큰 격차를 보여주고 있음.

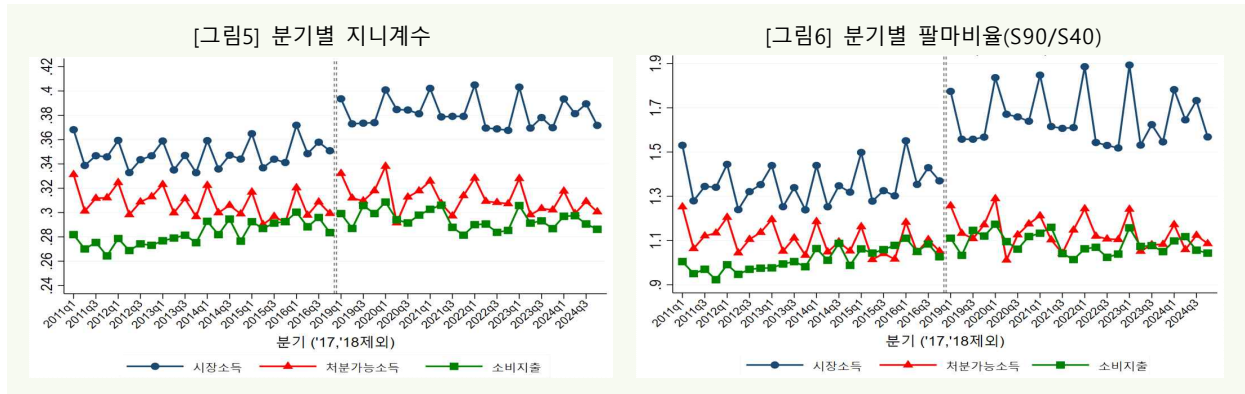
이 두 가지 사실을 고려할 때 2010년 이후 주거비가 꾸준히 상승하고 있음에도 자가와 전세의 경우 주거서비스 비용이 총소비지출에 포함되지 않음으로써 고소득층과 저소득층의 소비지출의 격차(소비지니계수)가 커지는데 이바지한 것으로 보임. 즉 고소득층에 비하여 저소득층에서 소비지출 중 주거비지출 비중이 더 높은 경향이 있기 때문에, 주거비를 제외할 경우 소비분배는 더 악화된 것으로 나타날 수 있음.



주요 결과
(계속)

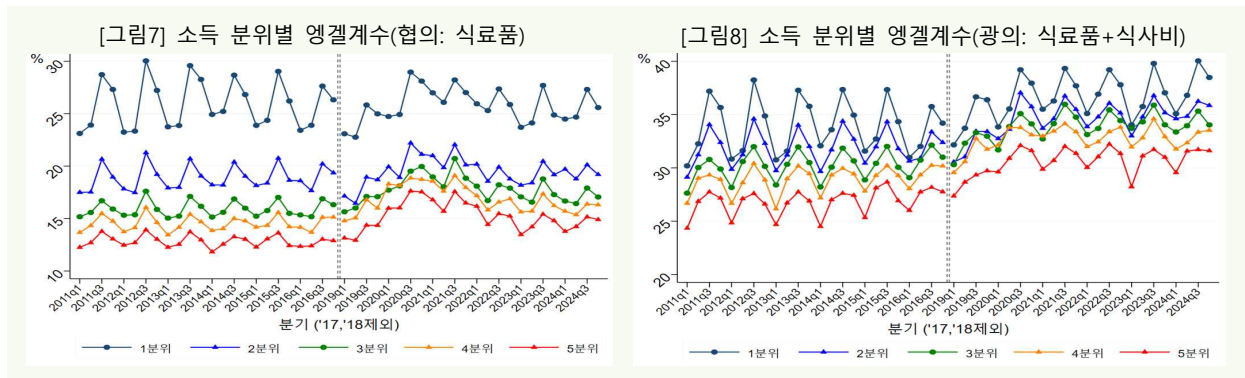
② 팔마비율의 재확인 : 대표성과 직관적 이해 두 가지를 모두 갖춘 지표

팔마비율 중간 50%는 전체 소득에서 차지하는 비중이 동질적이라는 전제하에 소득 상위 10%의 점유율이 소득 하위 40%의 몇 배인지를 나타내는 지표임. [그림5~6.]에서 볼 수 있는 바와 같이 시장소득, 처분가능소득, 소비지출 모두에서 지니계수와 팔마비율은 매우 유사한 흐름을 확인함. 지니계수와 상관계수는 순서대로 0.997, 0.995, 0.993, 부분결정계수도 팔마비율이 5분위 배율에 비하여 매우 높게 나타남.

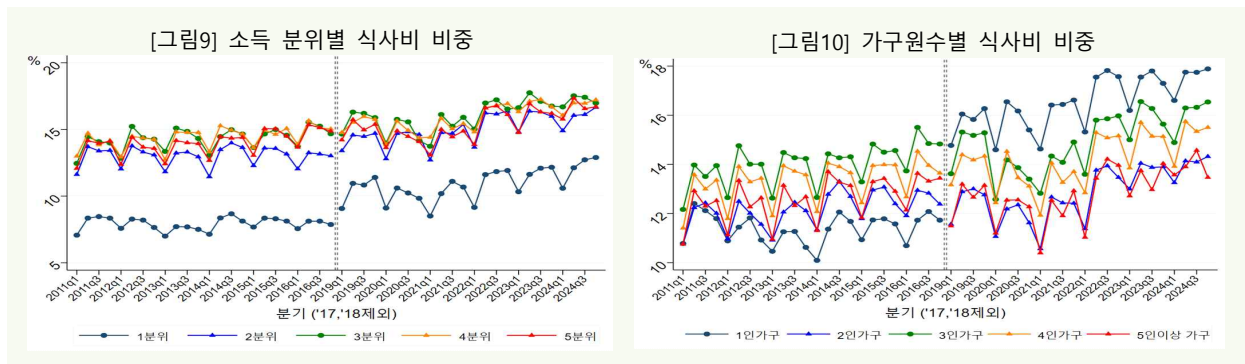


③ 앵겔계수의 재발견: '외식'이 필수재가 된 시대

전통적인 앵겔계수(식료품비/총지출)는 하락하는 추세지만, 식사비(외식(배달), 단체식, 주점 및 커피)를 포함한 '광의의 앵겔계수'는 전 계층에서 상승 추세에 있음.



외식 등을 포함하는 식사비는 모든 소득 분위와 모든 가구원수 가구에서 증가 추세에 있음. 특히 1인 가구의 경우 2019년 이후 외식비가 급격히 증가한 것을 확인할 수 있음. 이는 배달 플랫폼 활성화, 1인 가구 특히 경제력이 뒷받침되는 1인 가구의 증가에 기인하는 것으로 이해할 수 있음.

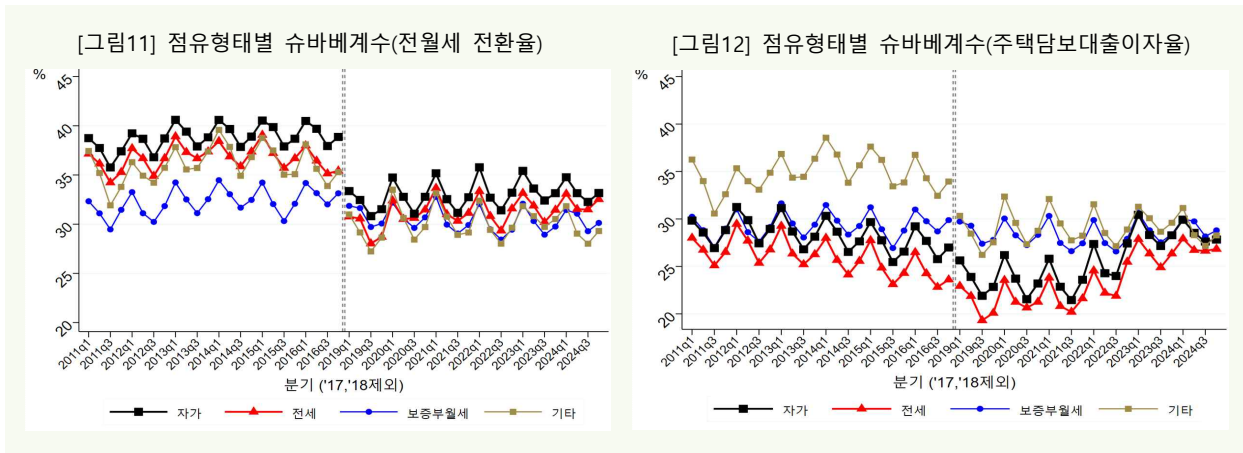


주요 결과

(계속)

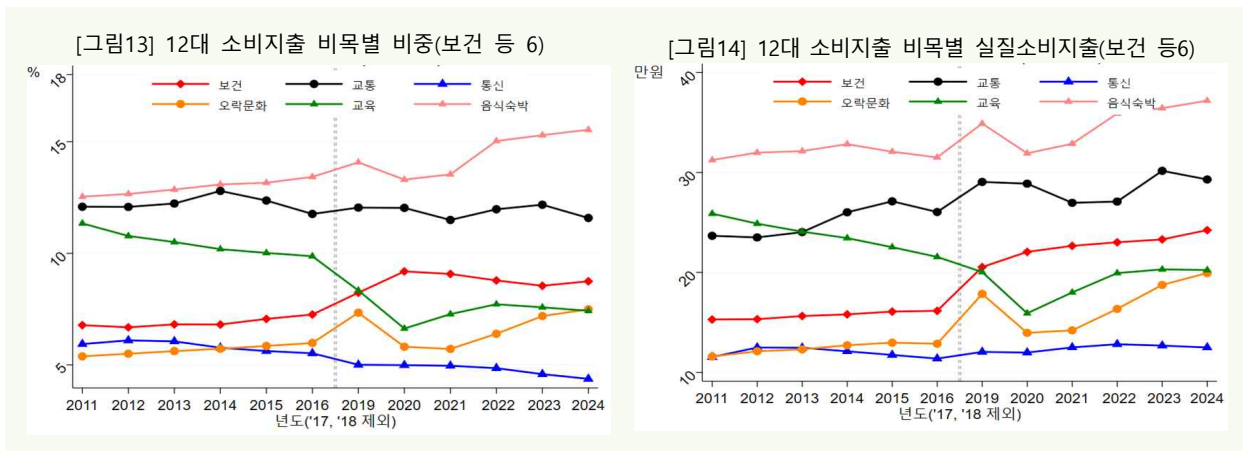
④ 주거비 부담: 숨겨진 비용 '의제주거비'의 영향

실제주거비는 월세가구에서만 측정되므로 주거비를 과소하게 측정할 가능성이 있음. 이에 전세 및 자가의 주거비를 의제주거비로 간주하여 추정함. 전월세전환율을 이용한 경우는 전세와 자가에 대하여 월세에 상당하는 비용을 지불하는 것으로 간주하는 것임(소위 주택시장을 통한 주거서비스비용 산정 방법). 이자율을 이용한 경우는 주택담보대출이나 전세자금대출 등 집주인이나 임차인의 자본 조달 비용을 측정(소위 주거 서비스를 가계의 신용에 따른 금융시장을 통한 부담)한 것임. 전반적으로 전월세 전환율을 적용할 경우의 주거비가 더 높음[그림11]. 또한 전월세 전환율을 사용할 경우 전세, 자가, 월세의 주거비 비중이 큰 차이가 없으나. 이자율을 사용한 경우 전세가 가장 부담이 적고 시기에 따른 편차가 상대적으로 큰 것이 특징임.



⑤ 소비트렌드의 구조적 변화 : 교육 감소와 보건, 음식·숙박, 오락·문화 증가

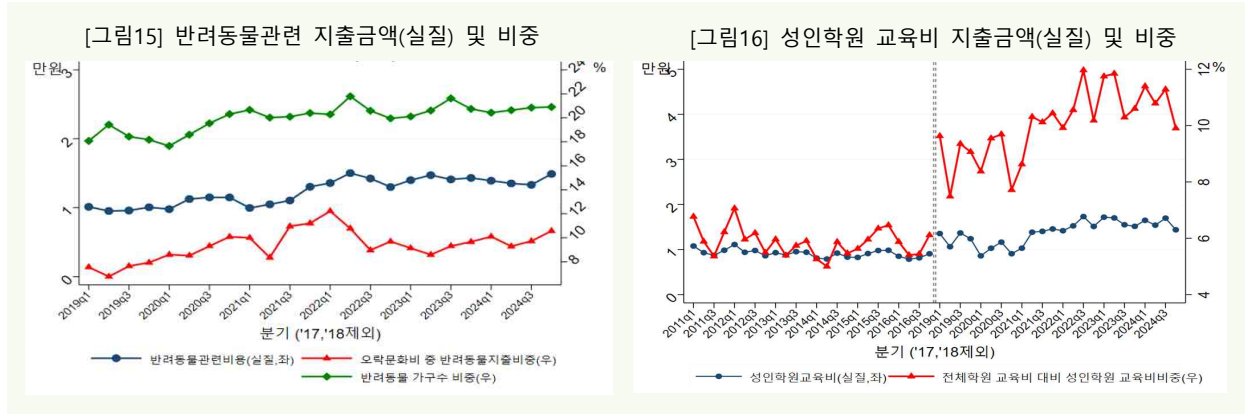
12대 비목 중 소비지출 금액에서 차지하는 비중과 실질 소비지출 금액 모두 하락하는 비목은 교육이고, 모두 상승하는 비목은 보건, 음식·숙박, 오락·문화임. 교육과 보건 항목의 변화는 저출산·고령화라는 인구 구조변화의 결과로 이해할 수 있음. 음식·숙박, 오락·문화의 증가는 경제성장의 결과로 이해할 수 있음.



주요 결과
(계속)

⑥ 새로운 소비 동향: '가치'와 '편의'에 지불하는 가계

오락문화비 중 반려동물 지출 비중은 10년 전 대비 2배 이상 상승함. 반려동물을 키우는 가구 수, 관련 지출 비용도 지속적으로 증가하고 있음. 자기계발 등 성인들의 직업훈련 및 학원비 지출 또한 2010년대 후반부터 급격히 증가하고 있음. 최근 건강, 운동에 관한 관심도 높아지고 있는 것은 운동시설 이용 금액과 오락·문화비 중 운동시설 비중이 지속적으로 증가하는 것으로 확인할 수 있음.



시사점

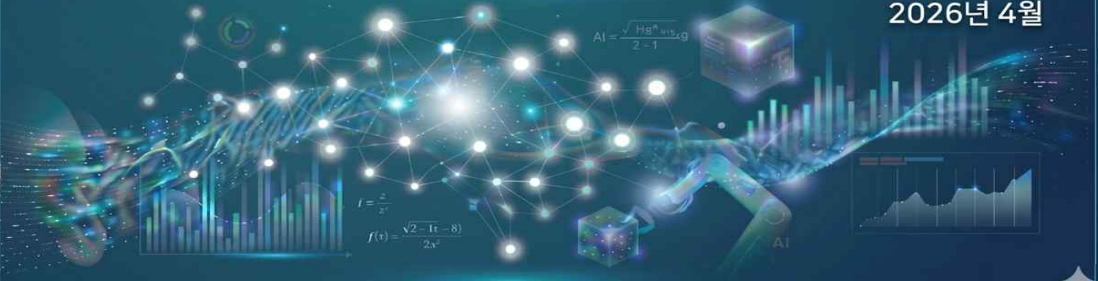
가계의 경제 상황 점검을 위하여 소득과 소비지표라는 이중 축(Dual-Pillar) 관리가 필요함. 특히 '소비 팔마비율' 지표는 소비의 분배 상태를 잘 나타내면서도 상하위 계층 간 소비 격차를 잘 나타내 대중적 수용성이 높을 수 있음. 엔겔계수와 슈바베 계수의 최근 방향을 반영한 개선 방안을 도출함. 소득 지표와 달리, 소비지표는 교육, 보건 등 소비 동향의 구조적 변화와 반려동물 등 새로운 소비 동향을 포착할 수 있다는 점에서 차별화됨.

한계 및 향후 연구

본 과제는 자료 활용 제약으로 인과관계보다는 전반적인 가계경제 현상만을 보여 주고 있다는 한계에 있음. 향후 지역별, 연령대별, 가구 특성별 연구를 통하여 지역별 소비 특성과 고령층 소비 특성 등을 연구할 필요가 있음

참고 문헌

- 김대일. (2015). "소득 및 소비의 불평등과 상호 연계". 노동경제 논집, 제38권(2), pp. 22~58
- 유경준. (2007). 소득불평등도와 양극화. 한국개발연구원
- 통계청. (2025). "가계동향조사 통계정보보고서". 통계청(현 국가데이터처)
- 한국재정학회. (2024). "가계동향조사 지표 개선방안 연구". 통계청(현 국가데이터처)



김정민

국가데이터연구원
데이터과학연구팀

'보는 통계'에서 '읽는 통계'로의 전환

- 나우캐스트 지표 생성형 AI 서비스 제공 방안 연구 -

- (연구 목적) 본 연구는 나우캐스트 지표의 가독성 향상을 위해 생성형 AI 기반의 통계 나레이션 모델을 시범 구현하고 실무 활용성을 평가하였다.
- (핵심 결과) LangGraph 기반의 자가 품질 점검 프로세스를 통계 나레이션 모델에 추가할 시, 신뢰할 수 있는 수준의 나레이션 도출이 가능함을 확인하였다.
- (향후 활용) 본 연구에서 제안한 LangGraph 기반 통계 나레이션 모델 구조는 나우캐스트 지표 외 다양한 통계서비스에 적용 시 참고하기에 용이하다.

연구배경 및 목적

- (연구 배경) 국가데이터처는 민간 자료를 활용한 속보성(주간, 월간) 지표인 나우캐스트 지표를 생산해 국가통계포털을 통해 서비스 중
- (문제 인식) 그간 자료 정제, 행정자료와의 연계 등 제공 지표의 정합성 확보에 집중해 온 관계로, 통계서비스 측면의 개선 검토는 상대적으로 부족
- (연구 필요성) 나우캐스트 시스템 고도화 및 이용자 편의성 향상을 위하여, 생성형 AI 기반의 현행 서비스 개선 방안을 모색해 볼 필요
- (연구 목표) 생성형 AI 기반의 통계 나레이션 모델을 시범 개발 후 나우캐스트 지표에 적용해 봄으로써 생성형 AI의 통계서비스 활용 가치 진단

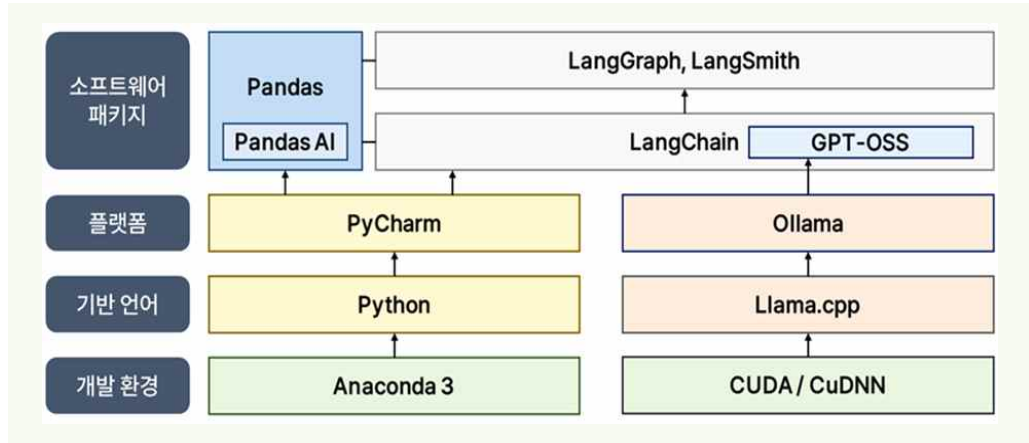
데이터 및 방법

- (활용 데이터) 나우캐스트 지표 9종 중, 신용카드 이용금액, 온라인채용 모집인원수, 모바일 인구이동 지표의 원천 자료 전수를 주요 실험 대상으로 선정
 - (구현 방법론) 통계 나레이션 시나리오를 6단계*로 구분하고, 단계별로 독립적인 시스템 프롬프트를 할당 후 답변 결과를 연계하는 방식을 채택
- * 지표 개요 → 최근 변화 추이 → 최대/최솟값 → 기초 통계량 → 종합 결론 → 유의사항

데이터 및
방법
(계속)

- (활용 패키지) LangChain, LangGraph, Pandas 등 다수

[그림1] 통계 나레이션 AI 모델 구현을 위한 주요 기술 스택



- (AI 모델 선정) 8종의 오픈소스 AI 모델을 정확성, 일관성, 지시 이행의 3가지 기준으로 정성 평가함으로써 GPT-OSS 20B*를 최적합 모델로 판단

* OpenAI가 '25년 8월 공개한 오픈소스 AI 모델로서, 파라미터 규모에 따라 20B와 120B 모델이 존재하며, MOE 기술을 채택해 적은 컴퓨팅 자원으로도 활용 가능

<표1> 오픈소스 생성형 AI 모델 8종의 성능 검토 결과

No.	개발사	모델명	규모(B)	성능 평가	
				정확성	지시 이행
1	네이버	HyperClova X SEED	1.5	정확성	○ · 자연스러운 문장 생성
				일관성	× · 자료 분석 오류 다수 발생
				지시 이행	× · 통계 설명 요청 이해도 부족
2	알리바바	Qwen3	32	정확성	○ · 자연스러운 문장 생성
				일관성	△ · 간헐적인 통계표에 대한 환각
				지시 이행	× · 간헐적인 중국어 혼용 답변
3	알리바바	QwQ	32	정확성	○ · 자연스러운 문장 생성
				일관성	○ · 일관성 있는 통계자료 해석
				지시 이행	× · 간헐적인 중국어 혼용 답변
4	LG AI 연구원	EXAONE 3.5	32	정확성	△ · 간혹 어색한 문장 생성
				일관성	△ · 간헐적인 통계표에 대한 환각
				지시 이행	△ · 작성 지침을 간헐적으로 간과
5	구글	Gemma3	27	정확성	△ · 간혹 어색한 문장 생성
				일관성	△ · 간헐적인 통계표에 대한 환각
				지시 이행	○ · 작성 지침에 대한 이해 탁월
6	메타	Llama 4 Scout	109	정확성	○ · 자연스러운 문장 생성
				일관성	△ · 간헐적인 통계표에 대한 환각
				지시 이행	△ · 작성 지침을 간헐적으로 간과
7	환광퀀트	DeepSeek-R1	32	정확성	△ · 간혹 어색한 문장 생성
				일관성	○ · 일관성 있는 통계자료 해석
				지시 이행	× · 간헐적인 중국어 혼용 답변
8	OpenAI	GPT-OSS	20	정확성	○ · 작성 지침에 따른 문장 생성
				일관성	△ · 간헐적인 통계표에 대한 환각
				지시 이행	○ · 작성 지침에 대한 이해 탁월

주요 결과

- (1차 구현) LangChain 기반 1차 구현 결과, 대체로 의도한 바와 일치하는 답변 생성을 확인하였으나, 기초 통계량 분석 과정에서 간헐적 오류 관측*

* 오류의 원인은 환각현상(hallucination)으로, 생성형 AI 모델의 고유 특성

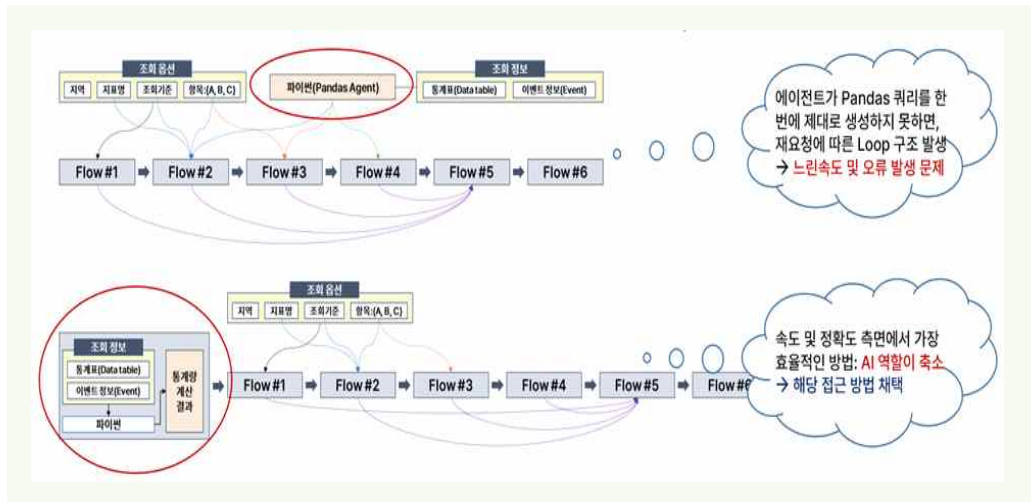
<표2> AI 나레이션 생성 결과의 통계 분석 오류 건수

(독립 시행 1,000회 기준)

모델명	정상 출력 건수	비정상 출력 건수	오류율(%)
GPT-OSS 20B	977	23	2.3
Gemma3 27B	951	49	3.9

- (오류 회피 전략) AI의 통계량 분석 오류를 회피하기 위하여 AI 에이전트를 활용하는 방법과 AI 동작 과정에서 일반 함수를 활용하는 방식을 검토

[그림2] AI 모델의 통계량 분석 오류 회피를 위한 방법 검토 결과



- (오류 회피 실증) 검토한 방법*을 실증해 본 결과, 일반 함수를 활용하는 방식이 추론 속도 및 오류 발생 건수가 가장 낮아 활용에 적합한 것으로 확인

<표3> 모델의 통계량 산출 방식별 모델 평균 실행 시간

(독립 시행 100회 기준)

모델 구분	AI가 직접 분석	AI 에이전트 활용	일반 함수로 산출
전체 실행	2분 50초	4분 47초	1분 21초
오류 발생 건수	5회	17회	2회

- (한계 인식) 1차 구현 방식은 지표의 종류에 따라 통계 나레이션 모델을 추가 개발해야 하는 비효율*과 답변 품질 고도화를 위한 대응책의 한계**가 뚜렷

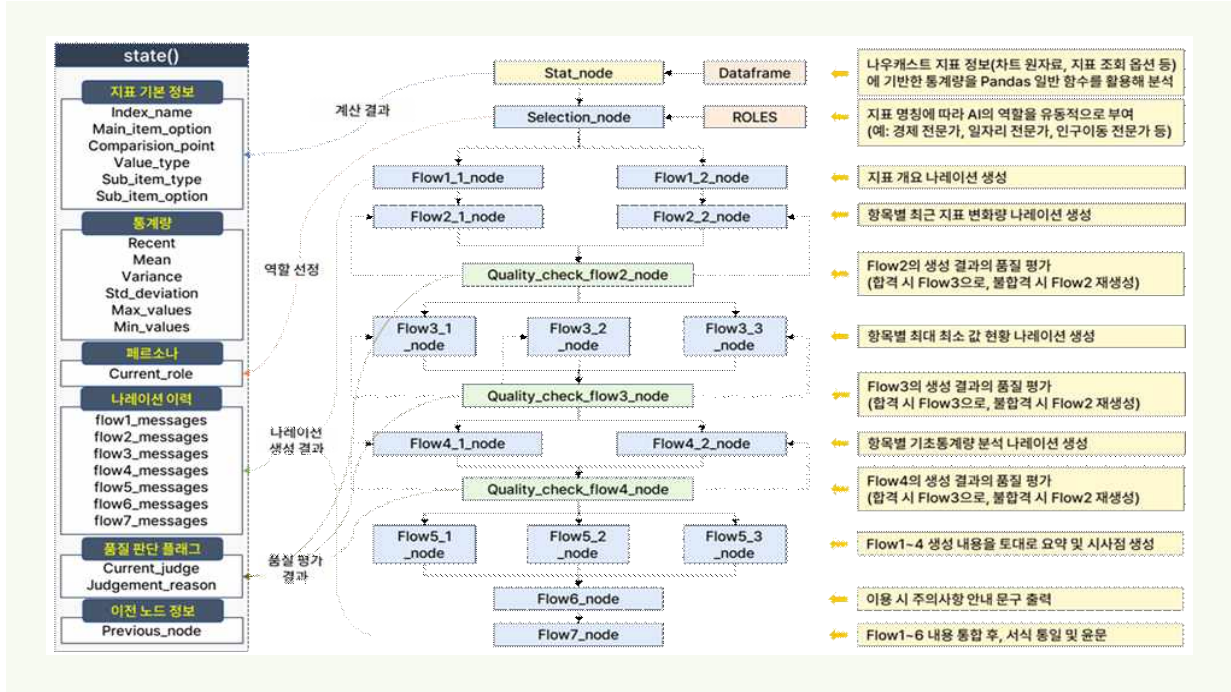
* LangChain은 다수의 AI 답변을 하나로 통합할 수 있는 이점이 있으나, 사전 정의된 일방향 동작만을 허용하므로 지표가 변경되면 모델도 새로 정의해야 하는 문제

** AI의 통계량 분석 오류를 해결하는 방식으로, AI가 오류가 없는 답변을 생성하도록 유도하는 방안 외 후처리(오류 발생 시 수정) 고려가 불가능

주요 결과
(계속)

- (2차 구현: 개선 모델) LangGraph 패키지를 도입하여, 조건부 분기와 순환 구조를 갖춘 그래프 형태의 통계 나레이션 모델로 재구현

[그림3] LangGraph 기반 통계 나레이션 모델 구조 설계 결과



- (2차 모델의 개선점) 나우캐스트 지표 종수를 확장하는데 있어 진입장벽이 낮고, 답변 품질에 대한 자가 점검 기능을 지원
 - ① 스테이트(state) 관리를 통해 모든 변수의 체계적 추적 및 제어가 편리
 - ② 페르소나 자동 선정 기능으로 지표 유형별 맞춤 나레이션 지원
 - ③ 품질 평가 노드를 통해 답변의 자가 검증 및 재생성 순환 구조 구현
 - ④ 윤문을 위한 AI 동작을 추가해 최종 나레이션의 서식 일관성 확보
- (자가 품질 평가*의 효과성) 통계량 오류 발생 시 답변 재생성을 지시한 결과, 최종 답변의 오류율을 0%에 수렴하도록 제어**할 수 있음을 확인

* LLM-as-a-Judge로 명명하는 기법으로서, 최근 다양한 논문에서 효과가 입증

** 단, 자가 품질 평가 노드를 모델에 추가하는 경우 품질 평가를 위한 AI 추론이 필요하므로, 최종 답변이 도출되는 데까지 걸리는 시간이 증가

시사점

- (서비스 활용성) LangGraph 기반의 품질 자가 점검 프로세스가 모델에 포함되어 있음을 전제로, 서비스에 적용할 수 있을 것으로 기대
- (실시간 서비스) 현재 기술 수준에서 실시간 나레이션 생성 서비스는 품질 보장을 위한 추론 속도 지연으로 인해 추천되지 않음

시사점 (계속)

- (AI 에이전트) 에이전트 기술은 나레이션 모델에 부적합하나, 향후 SPSS, SAS 등 통계 분석 패키지가 에이전트 기능을 지원할 시, 활용 여지 존재
- (연구 성과 활용) 본 연구의 모델 설계 구조는 나우캐스트 지표 외에도 다양한 통계서비스의 해설문 자동 생성에 확장 적용될 수 있음

한계 및 향후 연구

- (이벤트 정보) 당초 본 연구는 나우캐스트 지표상 제시된 이벤트 정보를 통계 나레이션에 연계하고자 하였으나, 등록된 이벤트 정보 규모* 문제로 한계
 - * 향후 언론 기사 빅데이터를 활용한 이벤트 정보 확충을 검토해 볼 필요
- (통계 벤치마크*) AI 모델의 데이터 테이블 해석 역량을 정량적으로 평가할 수 있는 벤치마크가 부재하므로, 관련 벤치마크 개발 검토 필요
 - * AI 모델의 성능을 진단하기 위해 개발된 도구로서, 측정하고자 하는 성능과 관련된 비공개 문항으로 구성
- (고도화 방안) 지표별 차별화된 시스템 프롬프트 정교화가 검토될 시, 연구를 위해 시범 구현된 결과 대비 높은 생성 품질을 기대할 수 있음
- (정부망 내 AI 도구 활용) 정부망의 보안 정책으로 인해 허깅페이스, VS-code 등 오픈소스 플랫폼의 접근이 제한되는 문제가 있어, 해소 방안 검토 필요

참고 문헌

- 김정민(2024), 국가통계 분야 인공지능 활용 체계 연구, 국가데이터연구원
- 김정민(2025), 통계 나레이션을 위한 생성형AI 개발: Nowcast 지표를 대상으로, 2025 국가통계연구 심포지엄
- Statistics Canada (2025), Project Examples of AI for Analytical Product Preparation and Validation`
- Sui, Y. et al. (2024), Table meets LLM: Can large language models understand structured table data?, WSDM 2024
- Zhang, X. et al. (2025), TableLLM: Enabling tabular data manipulation by LLMs in real office usage scenarios, ACL 2025
- Zhao, Y. et al. (2023), Investigating table-to-text generation capabilities of LLMs in real-world information seeking scenarios, EMNLP 2023

⇒ (연구보고서 원본) 국가데이터연구원홈페이지 > 연구간행물>연구보고서 > 나우캐스트 지표 생성형 AI 서비스 제공 방안 연구

이석민

국가데이터연구원
데이터과학연구팀

소득이동통계로 본 한국의 이동성 구조: 이동성과 고착의 공존

- 행정자료 기반 소득이동성과 청년층 지역이동 결합분석(2017~2023) -

- 본 연구는 소득이동통계 심층분석과 청년층 지역이동 분석을 결합하여 한국의 소득이동 구조를 실증적으로 정리하였다.
- 분석 결과, 한국의 소득이동성은 '계층 고착(하위/상위)'과 '선택적 이동(특히 청년층 지역이동에 의한 소득상승)'이 동시에 나타나는 구조를 보인다.
- 청년층의 수도권 이동은 개인 소득 상승(평균 +22.8%)의 주요 경로이나, 동시에 지역 간 인적자본 격차를 확대하는 메커니즘으로 작동한다.
- 소득이동통계는 단순 분배지표를 넘어 정책 수요 결정, 지역균형발전, 사회이동성 평가를 위한 핵심 지표로 활용 가능하다.

연구배경 및 목적

- 전통적 소득분배 지표는 특정 시점의 불평등 수준을 측정하지만, 개인의 경제적 위치 변화(이동성)는 포착하지 못함
- 한국은 수도권 집중, 청년 이동 증가, 노동시장 불안정성 확대 등 구조변화가 동시에 발생하며 '기회의 이동 경로'가 중요해지고 있음
- 특히 청년층은 지역 이동과 노동시장 진입이 결합된 집단으로, 소득이동 구조를 이해하는 핵심 대상임
- 본 연구는 소득이동통계의 구조를 정리하고, 청년층 지역이동과 결합하여 한국의 이동성 메커니즘과 정책적 함의를 도출함

👉 (연구질문) 한국의 소득이동은 어떤 구조를 가지며, 청년층의 지역이동은 이를 어떻게 재구성하는가?

데이터 및 방법

- **(활용데이터)** 국가데이터처 소득이동통계* 2017~2023년
 * 연간 약 1,100만 명 규모(전체 인구 약 20%) 종단 데이터(인구·가구 등록센서스 +국세청 자료 연계)
- **(주요 변수)** 근로·사업소득, 소득분위(5분위), 연령, 성별, 지역(수도권/비수도권, 5극3특)
- **(분석방법)** 소득분위 전이행렬, Fields-Ok 및 Shorrocks 이동성 지수, 소득 증감률 (절대 이동성), 분위 이동(상대 이동성), 지역 간/내 이동 비교 분석

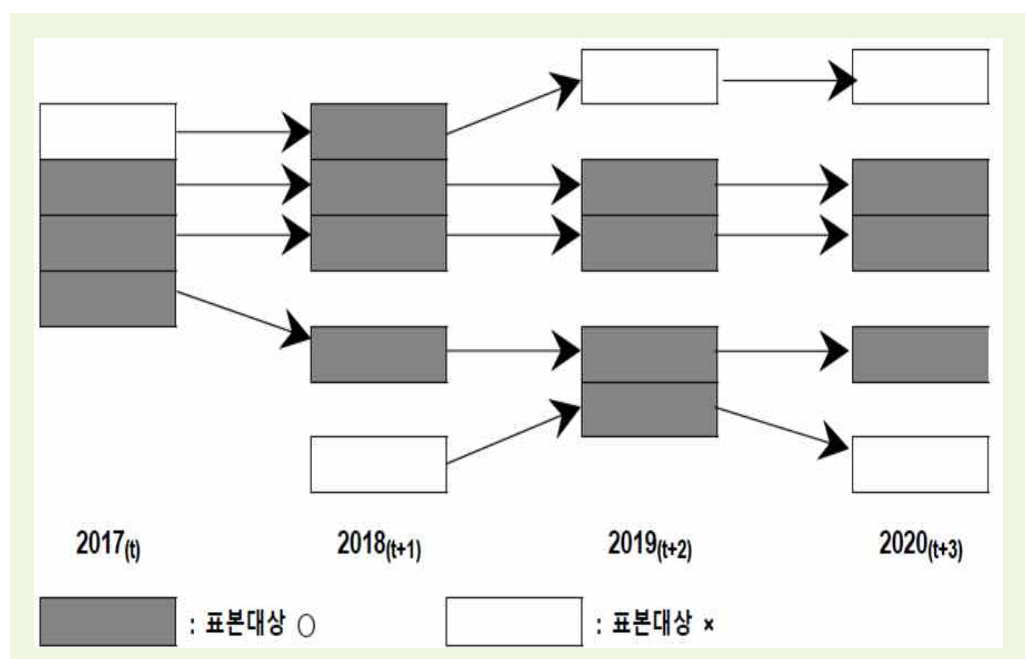
<표1> 소득이동통계 특징

① 다출처 데이터	<ul style="list-style-type: none"> · 통계청 등록센서스와 국세청 소득자료 결합 · 향후, 행정자료 추가 연계를 통한 데이터 확장성 및 효율성 증대
② 대규모 표본	<ul style="list-style-type: none"> · 등록센서스 가구단위 20%* 표본의 가구원(개인) 대상 * 2022년 표본 규모 : 1,162만 명(496만 가구)
③ 패널 데이터	<ul style="list-style-type: none"> · 데이터 총량 변화를 파악하는 횡단면 분석뿐만 아니라 우리 사회의 종단면 변화 및 이동상태를 파악할 수 있는 패널형태의 데이터 구축

패널 데이터(Panel data)

- 종단 데이터(Longitudinal data)라고도 불리며 시간의 흐름에 따른 다양한 상태의 변화를 포함하는 데이터
- 동일한 개인에 대해 매년 인구·가구·소득정보 등의 다양한 변화를 추적하여 데이터를 구축

[그림1] 소득이동통계 표본 가구원 유입 및 이탈 모형



주요 결과

○ 소득이동통계의 구조적 전환

- 기존 횡단면 통계의 한계를 보완하여 '동태적 분배 구조' 분석 가능
- 소득이동성은 경제적·사회적 성공을 위한 '기회의 평등'을 측정하는 지표로 '결과의 평등'을 특정하는 소득분배지표와 구별

○ 소득 계층 고착과 이동의 공존

- 하위 분위: 저소득층 잔류 비율 높음 (하위 고착)
- 상위 분위: 고소득 유지 확률 높음 (상위 고착)
- 중위 분위: 상·하향 이동이 동시에 발생하는 완충 구간

[그림2] 소득분위별 이동비율, 2022→2023년

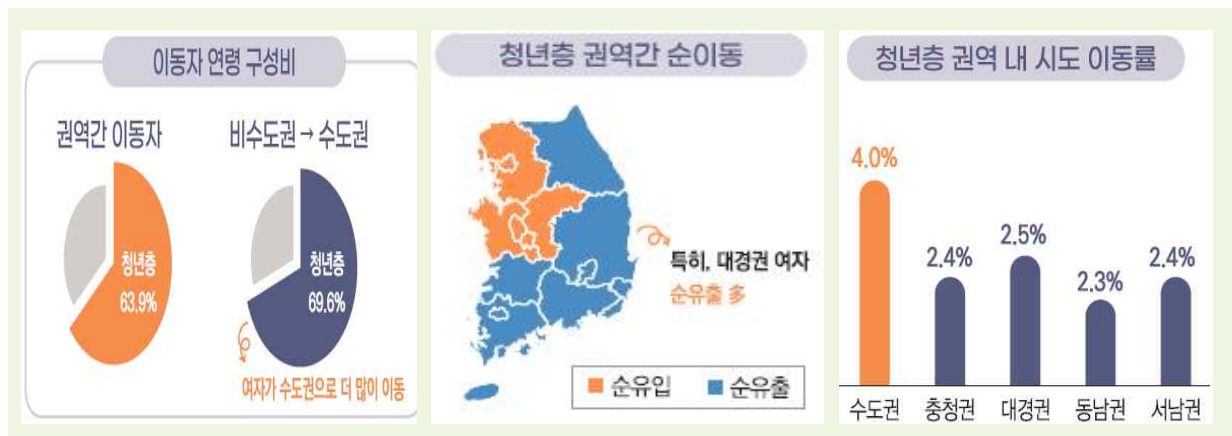


- 소득원천, 인구특성, 분석단위별로 소득이동성 격차 존재
- 소득이동성 단일 지표가 아닌 복합적 해석 필요

○ 청년층의 이동 집중 구조

- 전체 소득이 있는 인구의 권역 간 이동 중 청년층이 약 63.9% 차지
- 수도권 유입 인구 중 약 70%가 청년층, 청년층이 '이동성의 핵심 집단'

[그림3] 소득 있는 청년층(15~39세)의 이동특성, 2022→2023년



주요 결과 (계속)

○ 청년층 지역이동과 소득 상승 효과

- 수도권 이동이 소득 상승 경로로 작용
- 여성 청년의 수도권 이동 비중 및 소득 상승효과가 더 높음

<표2> 청년층의 수도권/비수도권 이동여부에 따른 평균소득 및 증감률, 2022→2023년

(단위 : 만 원, %)

성별	기준 연도	이동자		비이동자	
		비수도권→수도권	수도권→비수도권	수도권	비수도권
남녀 전체	2022년	2,439	2,693	3,187	2,884
	2023년	2,996	2,897	3,545	3,233
	증감률	22.8	7.6	11.2	12.1
남자	2022년	2,911	3,069	3,683	3,374
	2023년	3,531	3,423	4,100	3,796
	증감률	21.3	11.5	11.3	12.5
여자	2022년	1,918	2,222	2,670	2,287
	2023년	2,406	2,239	2,966	2,547
	증감률	25.5	0.8	11.1	11.4

- 저소득 청년일수록 이동 후 소득 증가폭 크게 나타남

[그림4] 수도권/비수도권 이동여부에 따른 소득분위 구성비 변화, 2022→2023년



- 권역별로는 대경권 및 동남권에서 타 권역으로 이동한 청년의 소득상승이 크게 나타남
- 특히, 청년 여자의 권역 간 이동과 비이동의 소득증감률 차이가 남자보다 크게 나타남. (특히, 대경권 청년 여자)

[그림5] 유출 권역별 이동에 따른 청년층 소득증감률, 2022→2023년



시사점

- **소득이동성의 정책적 재해석**
 - 소득이동통계는 '결과'가 아닌 '기회 구조'를 측정하는 지표
- **청년 이동의 이중 효과**
 - (개인) 소득 상승 기회 제공, (지역) 인적자본 유출로 격차 심화
 - 장기적으로 지역 불균형 고착 및 사회 이동성 제약 가능
- **정책 방향**
 - 이동 억제보다 지역 내 기회 창출 중심 정책 필요
 - 양질의 일자리, 교육·훈련 인프라, 정주 환경 개선
- **통계 인프라 활용**
 - 행정자료 기반 종단데이터는 정책 효과 평가 및 타겟팅 핵심 도구
 - 향후 사회이동성 통계로 확장 필요

한계 및 향후 연구

- **(소득원천 확장)** 및 공적이전 소득/지출 추가 반영 필요
- **(변수 추가)** 자산, 일자리, 교육 등의 사회 전반을 분석할 수 있는 변수 추가 필요
- **(향후 과제)** 일자리 관점의 소득이동성 분석, 지역 단위 이동성 지표 개발, 교육·건강 등 다차원 연계, 세대 간 이동성 분석 등

참고 문헌

- 국가데이터처(2025), 「2023년 소득이동통계 결과」 보도자료
- 국가데이터처(2025), 「청년 인구이동에 따른 소득변화 분석」 보도자료
- 국가데이터연구원(2025), "소득이동통계 심층분석 연구"
- OECD, (2018). A Broken Social Elevator? How to promote Social Mobility

⇒ (연구보고서 원본) 국가데이터연구원홈페이지 > 연구간행물 > 연구보고서 > 소득이동통계 심층분석 연구

출처를 밝히지 않고 국가데이터연구원 데이터 리서치 브리프의 내용을 무단전재 또는 복제하는 것을 금합니다.
본 리서치 브리프의 내용은 필자들의 개인적 의견이며, 국가데이터연구원의 공식적인 의견이 아님을 밝힙니다.

발 행 일 2026년 4월 30일

발 행 처 국가데이터처 국가데이터연구원

발 행 인 김 진

주 소 35220 대전광역시 서구 한밭대로 713 통계센터 6~8층

전화번호 042-366-7118

홈페이지 mods.go.kr/dsri