

가계부채DB와 신용카드 이용액 데이터의 경제분석에서의 활용 및 연계가능성

2021.5.28

한국은행 경제통계국

통계기획팀 통계연구반 장근호



THE BANK OF KOREA

- ❖ 본 발표내용은 발표자 개인의견으로 한국은행의 공식견해와는 무관합니다.

목 차

- I 머리말
- II 가계부채DB의 내용 및 특징
- III 신용카드 데이터의 내용 및 특징
- IV 가계부채DB와 신용카드 데이터의 연계가능성
- V 맺음말



머리말



(검토 배경)

■ 금융 빅데이터로서 가계부채**DB** 및 신용카드 이용액 데이터

- 금융경제 데이터로서의 성격 + 빅데이터로서의 성격(**3V**)

※ **3V : Volume, Velocity, Variety**

→ ① 경제주체의 실질적인 특성을 반영한 미시적 분석

② 시의성 있는 신속한 분석

- 가계부채**DB** : 글로벌 금융위기 이후 가계부채 문제에 대한 심층적이며 다각적인 분석
- 신용카드 데이터 : 코로나19 이후 민간소비 동향의 신속한 파악

- <이하 내용> 가계부채**DB**와 신용카드 데이터의 주요 내용과 특징에 대한 소개, 경제분석에서의 활용사례, 양 데이터의 유용성을 제고할 수 있는 연계방안 등



가계부채DB의 내용 및 특징

1. 구조 및 특징

(개요)

■ 가계부채 미시통계에 대한 수요 확대에 부응하기 위해 2015년 3월 구축

- 미국 뉴욕연준이 2010년 구축한 소비자신용패널(CCP)을 벤치마크
- NICE 평가정보(주)가 보유한 개인신용정보 중 일부 표본을 입수
- 모집단은 만 18세 이상 내·외국인 거주자 중 신용이력 보유자이며 임의추출법을 통해 약 100만명의 표본 추출(표본비율 약 2.4%)

가계부채DB 모집단 대상

내·외국인 거주자(5,553만명)			
만18세미만 거주자 ¹⁾	만18세이상 거주자(5,358만명)		
	신용이력보유자 (모집단)	1회성 신용조회자	비신용활동인구
195만명(4%)	4,633만명(83%)	343만명(6%)	382만명(7%)

주: 1) 만 18세 미만 거주자 중 신용관련 금융거래 등이 존재하는 개인. ()내는 비중

자료: NICE 신용등급DB(2020.12월말 기준)

1. 구조 및 특징

(주요 내용)

- ① (고유 특성정보) 연령, 거주지, 신용등급 등
- ② (금융거래정보) 대출, 신용 · 체크카드 사용실적 등
- ③ (신용관련 정보) 연체, 세금체납, 채무불이행 등

가계부채DB의 주요 정보

구분	자료출처	주요 변수
식별정보	신용정보원 NICE	연령대, 성별, 거주지, 신용점수·등급, 소득·등급 등
가계대출	신용정보원	업권, 상품, 총건수, 총금액, 신규건수 및 금액 등
채무불이행	신용정보원	채무불이행(신용관리대상) 등록 총건수 및 총금액, 세금체납 관련 총건수 및 총금액, 파산/면책관련 확정 총건수 및 총금액, 개인회생 확정 총건수 및 총금액 등
기업여신공여	신용정보원	업종, 업권, 상품, 신용공여잔액, 만기도래액, 한도액 및 잔액, 담보잔액 등
대부업대출	신용정보원	대부업대출 기관수 및 총금액
계좌별 대출	NICE	업권, 상품, 상환방식, 대출용도, 만기일, 대출잔액, 거치기간, 연체대환대출 여부, 원리금상환액 등
카드	NICE	신용카드 한도, 이용금액, 체크카드 이용금액 등
연체 ¹⁾	NICE	업권, 상품, 미해제연체 총금액, 분기중 신규연체 등

주: 1) 채무불이행 정보에 비해 상세한 정보를 포함하고 있음

자료: 김성준 외(2018)

1. 구조 및 특징

(주요 특징)

- ① (실제 금융데이터) 서베이 통계인 가계금융복지조사, 노동패널과 차이
- ② (패널데이터) 매분기 약 **100**만명의 신용정보 수집
- ③ (속보성) 기존 서베이 데이터와 달리 해당 분기말 자료가 약 **2개월** 경과후 **DB** 수록

가계금융복지조사, 노동패널과의 비교

	가계금융복지조사	노동패널	가계부채DB
작성기관	통계청, 한국은행, 금융감독원	한국노동연구원	한국은행
작성주기 및 공표일자	연간, 익년도말 발표	연간, 익년도말 발표	분기, 분기말+1.5개월
시계열	2010~	1998년~	2012년~(일부 2010년~)
조사방법	조사원 면접 및 인터넷 조사	조사원 면접 조사	신용거래실적을 CB로부터 입수
표본	약 20,000가구	약 6,700가구(09통합표본) 약 12,000가구(18통합표본)	약 100만명
주요 변수	금융자산, 실물자산, 금융부채, 소득, 소비	고용형태, 근로계약, 임금 및 소득, 금융자산, 실물자산, 금융부채, 소 비 등	가계대출, 카드사용실적, 연체, 기업 여신잔액 등

2. 경제분석에서의 활용사례

(국내 연구)

관련 연구	주요 내용
한국은행 금융안정보고서 각 호	가계부채의 차주 특성별 비중 및 분포, 다중채무·저신용·저소득자 등 부채 취약차주 비중, 자영업자 보유 부채 현황 등
유경원·이상호(2020)	가계부채의 동태적 특성을 파악하기 위하여 가계부채DB의 가계부채의 변화를 유입과 유출로 구분하여 차주특성 등에 따라 분석
김성준 외(2018)	가계부채DB의 주요 내용과 특징을 정리하는 한편 기술분석을 실시
정호성(2017, 2018)	가계부채DB를 활용하여 주택담보대출의 연체 결정요인을 분석(2017), 자영업 차주의 부도요인과 비자영업 차주의 부도요인 간 어떤 차이가 있는지 분석(2018)
송상윤(2019)	가계부채DB를 이용하여 확장적 통화정책이 주택담보대출 차입자의 소비에 미친 영향을 분석

(해외 연구)

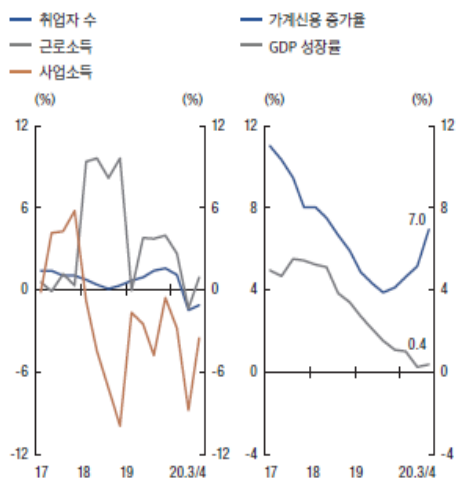
관련 연구	주요 내용
Bhutta(2015)	미국 CCP데이터를 활용하여 미국의 급격한 가계부채 감소에 대해 불황 시 채무불이행에 따른 부채유출보다 부채유입의 급감이 부채감소의 주된 요인임을 제시
Brown et al(2014)	학자금대출을 받은 경험이 향후 주택보유율, 여타부채상환, 신용접근성 등에 미치는 영향을 분석
Brown(2016)	미국 CCP데이터를 활용하여 분석한 결과, 2003년 이후 미국 가계대출 차주가 고령화되고 있다고 분석
Bleemer et al(2017)	미국 CCP와 교육부 자료를 연결하여 학자금대출을 받은 경험이 향후 주택보유율에 미치는 영향에 대해 분석
Lee and Tracy(2018)	미국 CCP자료를 활용하여 연방주택청(FHAI) 지원으로 최초로 주택을 구입한 경우 향후 주택보유 지속성에 미치는 영향에 대해 분석

2. 경제분석에서의 활용사례

(분석 사례 : 2020.12월 금융안정보고서*)

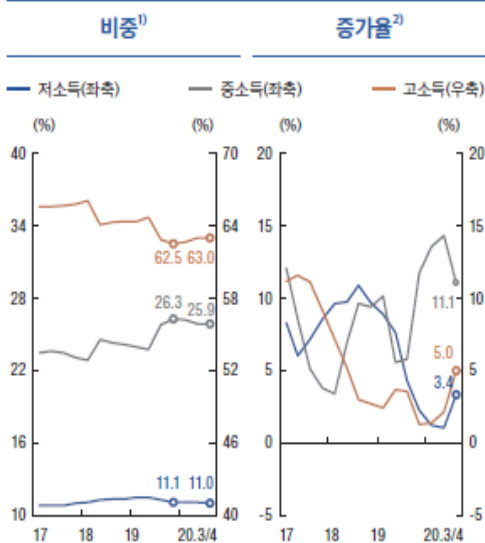
* 최근 가계대출 차주의 재무상환능력 변화

가계의 소득여건 및 부채 증가율¹⁾



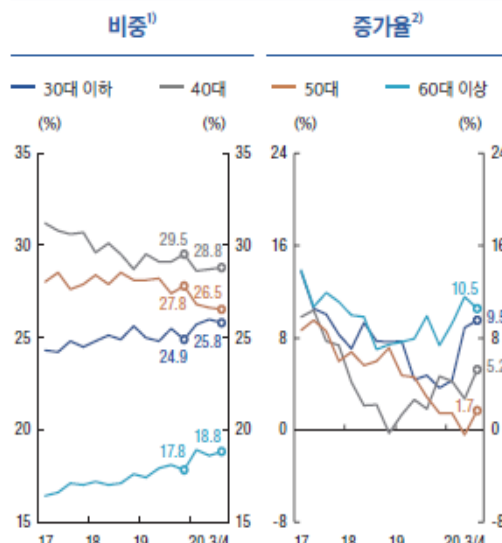
주: 1) 전년동기대비
2) 해당 분기 및 직전 3분기 명목GDP의 합
자료: 한국은행, 통계청

소득수준별 부채



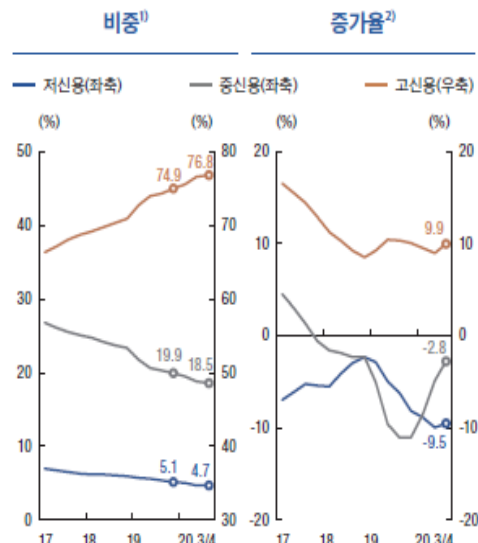
주: 1) 전체 부채액 대비 소득수준별 부채액 비중
2) 전년동기대비(2분기 이동평균)
자료: 한국은행(가계부채DB)

연령대별 부채



주: 1) 전체 부채액 대비 연령대별 부채액 비중
2) 전년동기대비(2분기 이동평균)
자료: 한국은행(가계부채DB)

신용등급별 부채



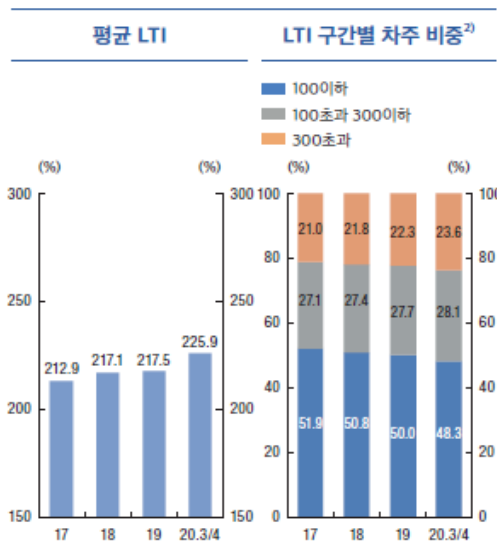
주: 1) 전체 부채액 대비 신용등급별 부채액 비중
2) 전년동기대비(2분기 이동평균)
자료: 한국은행(가계부채DB)

2. 경제분석에서의 활용사례

(분석 사례 : 2020.12월 금융안정보고서*)

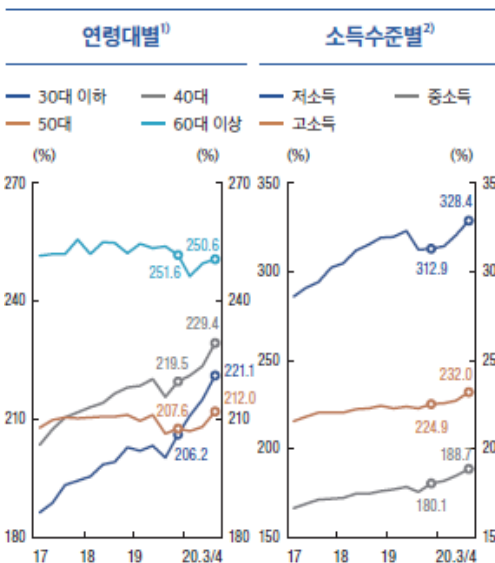
* 최근 가계대출 차주의 채무상환능력 변화

전체 차주 LTI¹⁾



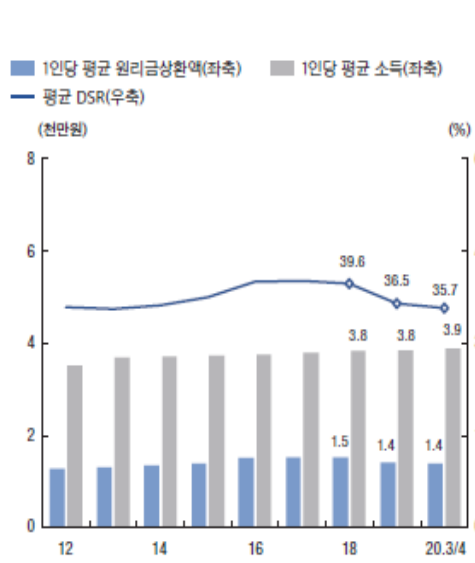
주: 1) 전체 차주 부채 합계 / 전체 차주 소득 합계, 기말 기준
2) 전체 차주 수 대비 비중, 기말 기준
자료: 한국은행(가계부채DB)

차주 특성별 평균 LTI



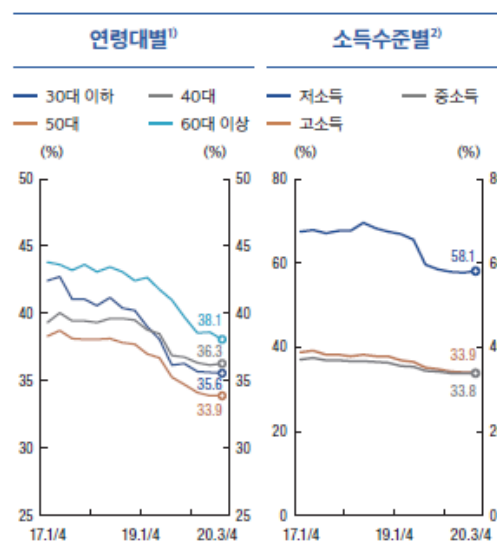
주: 1) 연령대별 차주 부채 합계 / 연령대별 차주 소득 합계
2) 소득수준별 차주 부채 합계 / 소득수준별 차주 소득 합계
자료: 한국은행(가계부채DB)

DSR¹⁾ 추이



주: 1) 전체 차주 원리금상환액 합계 / 전체 차주 소득 합계, 기말 기준
자료: 한국은행(가계부채DB)

차주 특성별 평균 DSR



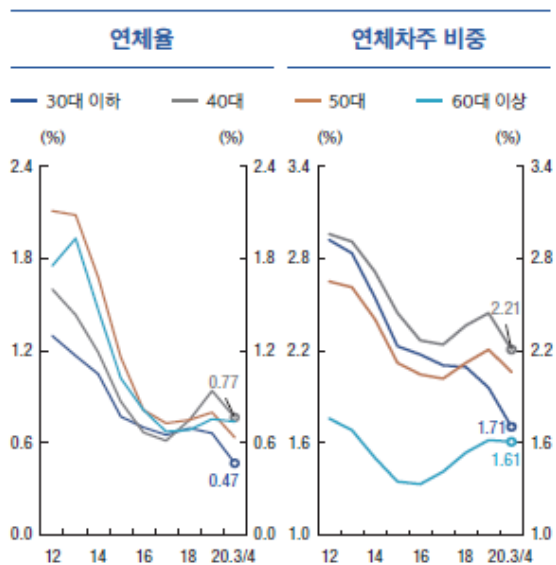
주: 1) 연령대별 원리금상환액 합계 / 연령대별 소득 합계
2) 소득수준별 원리금상환액 합계 / 소득수준별 소득 합계
자료: 한국은행(가계부채DB)

2. 경제분석에서의 활용사례

(분석 사례 : 2020.12월 금융안정보고서*)

* 최근 가계대출 차주의 채무상환능력 변화

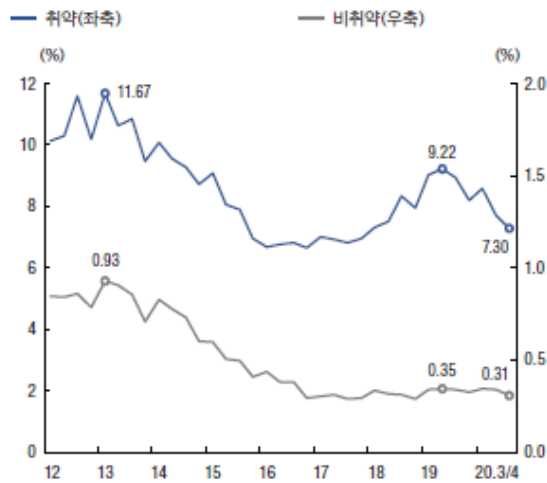
차주의 연령대별 연체율¹⁾



주: 1) 금융기관 업무보고서 상 집계된 연체액, 연체기간, 가계대출 포괄범위 등에서 차이가 있음에 유의(기말 기준)

자료: 한국은행(가계부채DB)

차주 특성별 평균 LTI



주: 1) 기말 기준

자료: 한국은행(가계부채DB)

3. 제약사항

- ① (소득·자산 정보 미흡) 가계의 부채상환 능력 파악에 일정부분 한계
- ② (내부적으로만 이용가능) 공식 승인통계가 아니어서 한국은행 내부적으로만 접근가능
- ③ (가구정보 미흡) 주소정보의 불완전성으로 경제공동체인 가구에 대한 정보 미흡,
가구원수 정보도 파악 어려움
- ④ (소비지출 파악관련) 개인별 카드 사용 총액 정보만 활용가능하고
세부 업종별 지출액이 제공 안됨



신용카드 데이터의 내용 및 특징

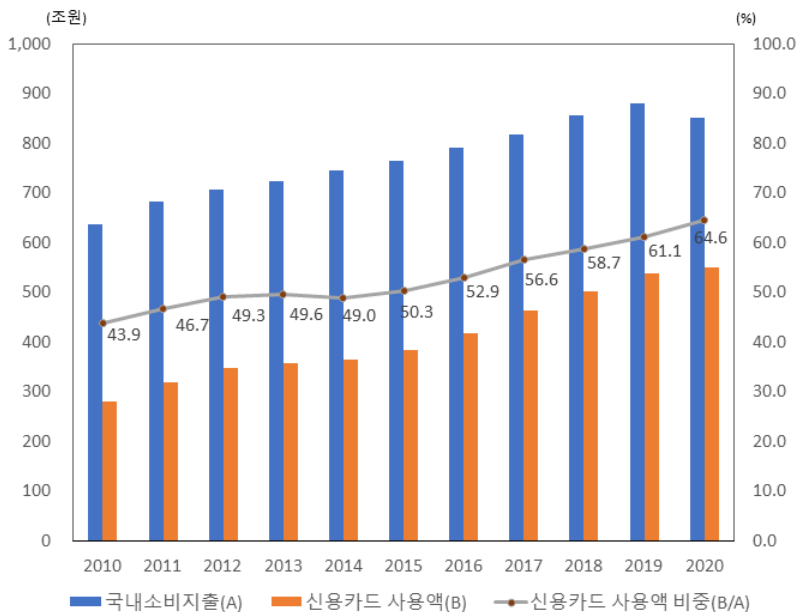
1. 구조 및 특징

(개요)

■ 신용카드사를 통해 실시간 집계되는 정형적(structured) 빅데이터

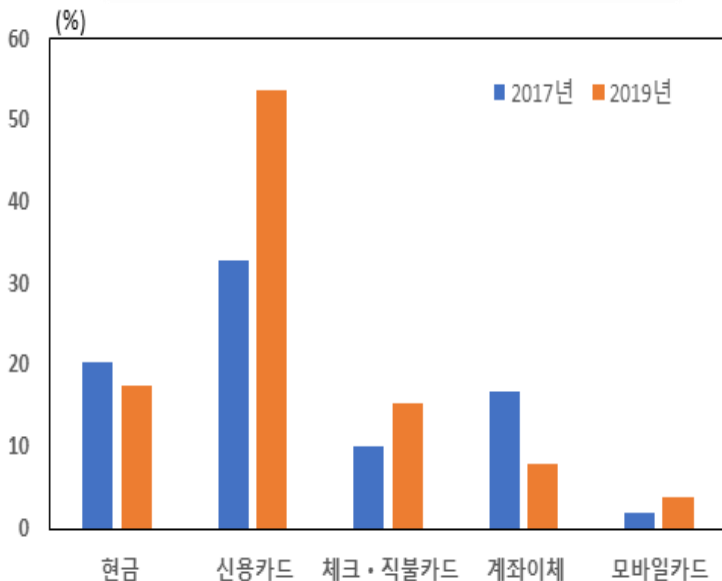
- 신용카드 이용액은 2020년 현재 전체 가계소비의 64.6% 차지, 민간소비 동향 파악에 유용
- <관련 데이터> ① 한국은행 지급결제 통계(월별) ② 여신금융협회 신용카드 승인실적(분기별) ③ 개별 신용카드사 입수 데이터

국내소비지출과 신용카드 사용액¹⁾ 추이



주: 1) 개인의 해외 신용카드 사용액 제외
자료: 한국은행 지급결제 통계, 국민소득 통계

지급수단별 이용¹⁾ 비중



주: 1) 금액 기준
자료: 한국은행 2019년 지급수단 및 모바일금융서비스 이용행태 조사결과

1. 구조 및 특징

(주요 내용)

- 신용카드 회사 보유 데이터는 소비자 개인정보와 가맹점별 결제금액으로 구성
- 한국은행 입수자료는 카드사 사용자의 연령대, 거주지별, 부채규모별 합산 신용카드 이용규모와 이용건수 금액으로 구성

신용카드 이용액 자료 형태

(신용카드사 보유 원자료)

	개인 정보					결제 금액			
	소득	주소	연령	성별	...	가맹점A	가맹점B	가맹점C	...
소비자1									
소비자2									
소비자3									
⋮									

	결제 금액 (합계)			
	업종1	업종2	업종3	...
20xx.xx월				
20xx.xx월				
20xx.xx월				
⋮				

(한국은행 입수자료 형태)

	이용 금액				이용 건수			
	업종A	업종B	업종C	...	업종A	업종B	업종C	...
전체 카드 사용자 합계								
연령대별								
20세 미만								
...								
60세 이상								
거주지(광역시도)별								
강원								
...								
충북								
부채규모별								
1분위								
...								
10분위								
부채미보유								

1. 구조 및 특징

(주요 특징)

- ① (속보성) 전자정보 형태로 실시간 집계, 기존 서베이 통계에 비해 기초자료 수집에 소요되는 시간 크게 단축
- ② (미시정보 포함) 소비지출 데이터를 소득, 연령, 성별, 지역 등 개인의 세부적인 특성과 업종, 매출규모, 지역 등 판매업체의 특성별 세분하여 분석가능
- ③ (실제 지급결제 데이터) 서베이 데이터와 달리 응답거부, 부정확한 응답 등 비표본오차 발생 우려 없음

2. 경제분석에서의 활용사례

(국내 연구)

관련 연구	주요 내용
한국은행 지역경제보고서 각 호	지역의 업종별 소비동향 파악을 위해 민간 카드사 데이터 활용
김경근.염명배(2015, 2017)	신용카드 사용데이터를 소비대리 변수로서 특히 지역단위에서 속보성 지표로 활용할 수 있음을 제시(2015), 신용카드 데이터를 활용하여 역외소비가 지역경제에 미치는 생산 및 고용 효과를 분석(2017)
이성호·최석환(2020)	수원시를 대상으로 신용카드 가맹점의 월별 매출액 자료를 통해 코로나19 발생이후 지역경제 및 상권별 영향을 분석한 결과 대면접촉 업종과 비필수 업종, 교통중심지 상권일수록 더 큰 매출감소를 보임
이유진.성영조(2020)	코로나19 발생 이후 경기도 신용카드 오프라인 매출을 분석한 결과 소득집단별로 재난기본소득 지급으로 인한 증가한 가처분소득의 주된 사용처가 상이하여 다양한 업종의 매출 증가로 이어지는 점을 발견
윤신희(2020)	고양시 전 지역에 대한 민간 카드사의 가맹점 개인 및 법인카드 사용내역을 이용하여 국내 코로나 확진자 최초 발생 이후 매주 업종별·지역별 소비동향을 분석하여 시 차원의 정책지원 방안을 모색

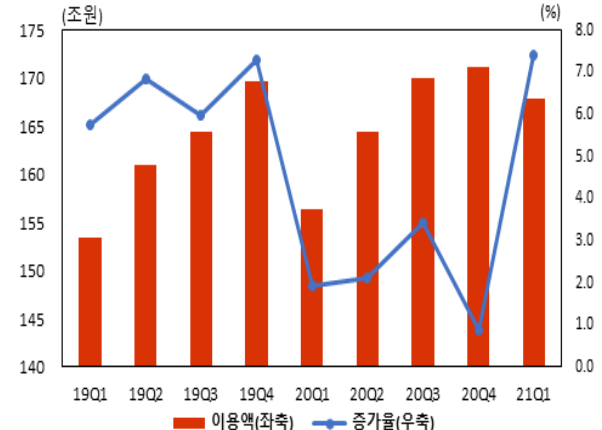
(해외 연구)

관련 연구	주요 내용
Dunn et al. (2020)	일간 신용카드 거래데이터를 이용하여 이벤트 스터디를 통해 코로나19 팬데믹이 지출항목별 민간소비에 미치는 영향을 평가
Aladangady et al. (2019)	신용카드 전표 매입사의 카드이용 승인신청액 정보를 활용하여 실시간으로 민간소비 동향을 파악할 수 있으며 허리케인과 같은 자연재해 등의 사건이 지역소비에 미치는 영향을 파악할 수 있음을 제시

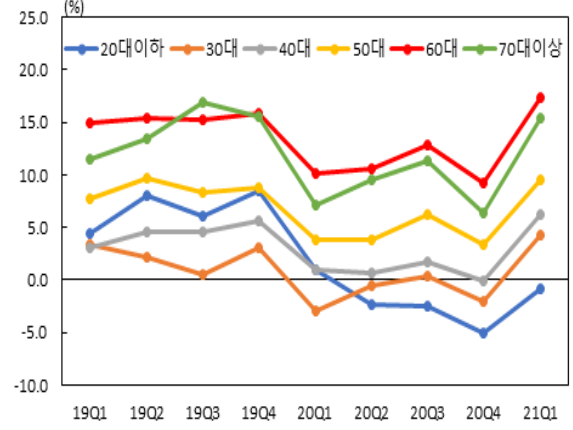
2. 경제분석에서의 활용사례

(분석 사례)

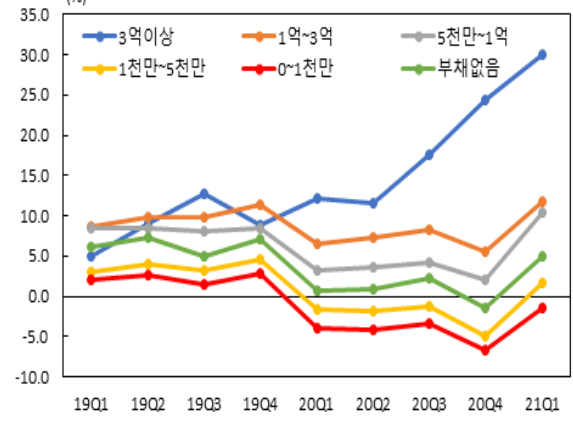
신용카드 이용액 및 증가율^{1),2)}



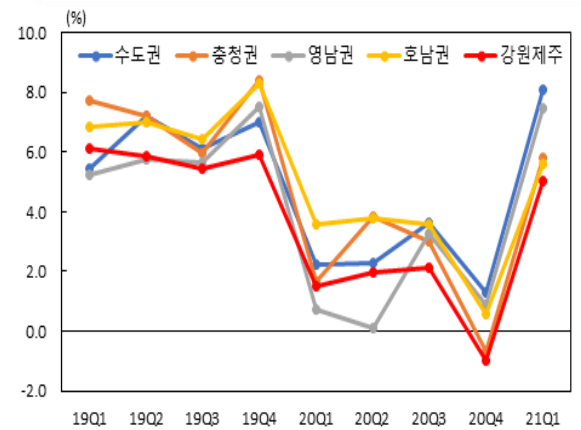
연령별 카드 이용액 증가율^{1),2)}



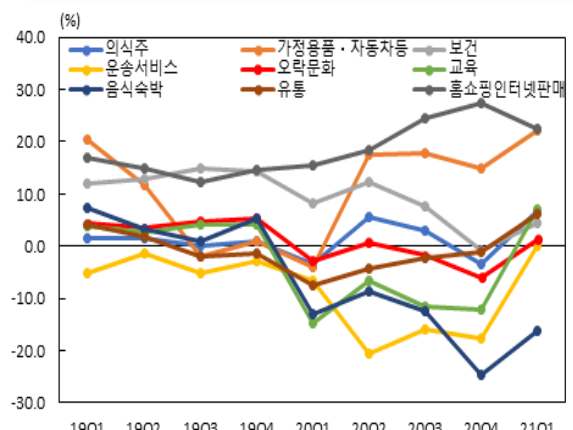
부채규모별 카드 이용액 증가율^{1),2)}



지역별 카드 이용액 증가율^{1),2)}



업종별 카드 이용액 증가율^{1),2)}



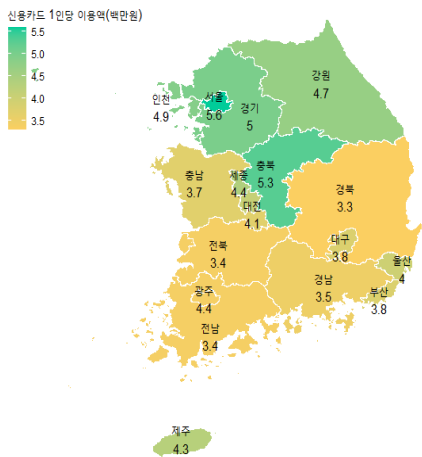
주: 1) 개인 이용액, 체크카드 이용액 포함, 해외이용액 제외, 2)전년동기대비 증가율
자료: ○○카드

2. 경제분석에서의 활용사례

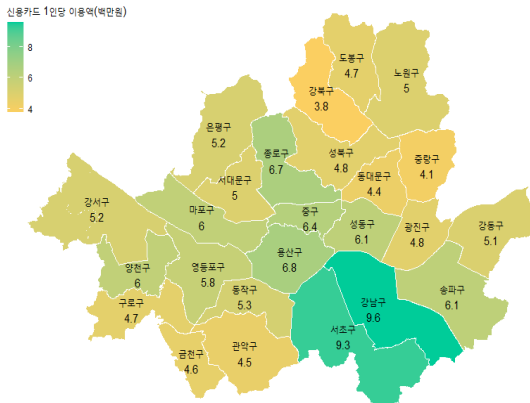
(분석 사례)

지역별 1인당 카드 이용액^{1),2),3)}

(전국 시도별)



(서울 자치구별)



주: 1) 개인 이용액, 체크카드 이용액 포함, 해외이용액 제외
2) 각 지역의 카드 이용액을 지역내 20세 이상 70세 미만 인구(통계청 인구동향조사의 2020년 주민등록 연앙인구 기준)로 나누어 산출
3) 2021년 1/4분기 기준

자료: ○○카드, 통계청

1인당 카드 이용액 분포^{1),2),3)}

(지역×연령)

(단위: 백만원)

	20대	30대	40대	50대	60대
서울	2.8	5.9	7.2	6.3	4.3
경기	2.5	5.3	6.4	5.4	3.7
인천	2.6	5.7	6.3	5.0	3.9
강원	2.7	5.9	5.8	5.0	3.1
충북	3.2	6.3	7.0	5.5	3.4
세종	1.6	4.3	6.0	4.9	3.1
충남	1.9	4.2	5.0	4.0	2.4
대전	1.8	4.5	5.5	4.8	2.9
경북	1.6	4.2	4.4	3.5	2.0
대구	1.6	4.4	5.1	4.0	2.7
울산	1.8	4.4	5.0	4.6	2.9
부산	1.6	4.2	5.4	4.0	2.6
경남	1.6	4.0	4.7	3.9	2.4
전북	1.5	3.9	4.5	4.0	2.2
전남	1.7	4.1	4.7	3.8	2.1
광주	1.8	5.0	5.8	5.2	3.1
제주	1.7	4.9	6.0	4.5	2.6

(업종×연령)

(단위: 만원)

	20대	30대	40대	50대	60대
식료품	3.3	8.4	10.5	11.4	10.2
의류및신발	3.3	5.4	5.1	4.2	2.9
주택및연료	0.4	2.4	4.3	4.4	3.1
가구·가사용품	3.6	14.5	15.8	14.9	9.6
보건	10.9	38.6	50.4	53.3	42.4
개인운송장비구매	2.1	8.8	10.8	11.3	7.2
개인운송장비운영	8.1	25.5	34.5	41.0	34.4
여객운송서비스	4.1	4.2	2.7	2.2	1.5
상품운송서비스	0.1	0.2	0.2	0.2	0.1
정보통신	14.3	28.3	29.4	22.1	10.7
오락용품및완동물	2.8	5.3	5.3	4.9	3.4
오락문화서비스·여행	5.0	6.4	8.1	6.7	3.2
교육서비스	1.8	14.7	23.6	7.6	0.8
일반음식점	33.4	43.0	39.5	33.8	18.5
주점 등	0.2	0.2	0.4	0.4	0.2
숙박서비스	1.2	2.2	1.8	1.4	0.8
개인미용·기타서비스	3.3	13.3	19.6	20.4	14.4
보험	5.4	10.8	11.4	10.3	6.5
백화점	7.3	16.4	12.3	8.8	4.9
할인점	11.5	32.3	43.9	40.5	28.5
편의점	11.2	14.6	10.8	6.7	2.9
홈쇼핑·인터넷판매	56.2	114.2	95.8	46.8	18.7
면세점·기타	6.2	12.9	14.0	10.3	6.4
상품권·세금공과금	5.2	29.4	73.7	65.3	47.5
기타	25.2	57.2	72.5	64.8	42.5

주: 1) 개인 이용액, 체크카드 이용액 포함, 해외이용액 제외
2) 각 지역(업종)의 연령대별 카드 이용액을 지역별(전국) 연령대별 인구(통계청 인구동향조사의 2020년 주민등록 연앙인구 기준)로 나누어 산출
3) 2021년 1/4분기 기준

자료: ○○카드, 통계청

3. 제약사항

- ① (신용카드 외 지급수단) 현금, 계좌이체 등 신용카드 이외의 다양한 지급수단이 활용되고 있으므로 신용카드 이용액이 전체 소비지출 내역을 적절히 반영하지 못할 우려
- ② (세부 소비항목 파악 어려움) 매출건별 총액만 기록되므로 종합소매업 등과 같은 다양한 상품을 판매하는 업체에서의 거래내역을 목적별 소비지출 통계에 편제하기 위해서는 추가적인 정보 필요
- ③ (신용카드 거래 비중 확대 추세) 신용카드 거래비중 확대 추세로 인해 신용카드 이용액 증가율로 가계소비 증가율 추정시 왜곡현상(상향 편의) 나타날 가능성

IV 가계부채DB와 신용카드 데이터의 연계가능성

1. 연계 필요성

(필요성)

- ① (소비활동과 부채·소득과의 연계분석) 가계의 소비에 대한 미시적 분석 영역 확대
- ② (소비지출의 세부정보 활용) 가계부채DB의 신용카드 사용액은 총액 기준으로 소비 규모의 측정은 가능하지만 세금 등 비소비성 지출이 포함되어 있으며 소비지출의 세부 내역은 파악하기 어려움

※ 연계데이터 분석 사례 : **Bleemer et al. (2017)**

(연계 방식)

- (제약사항) 신용카드 이용액 데이터는 개인정보 보호문제 때문에 개별 소비자별 데이터를 입수하지 못함

→ <코호트별 연계> 가계부채DB와 신용카드 데이터에 공통된 신용카드 이용액 데이터에 대해 연령, 지역, 부채규모별 정보를 기준으로 연계가능성 검토

2. 연계가능성 검증

(코호트 구성)

- ① (연령별) 6개 연령층 (20대 이하, 30대, 40대, 50대, 60대, 70대 이상)
- ② (지역별) 17개 광역시도(서울, 경기, 인천, 강원, 충북, 세종, 충남, 대전, 경북, 대구, 울산, 부산, 경남, 전북, 전남, 광주, 제주)
- ③ (부채규모별) 6개 구간 (3억원 이상, 1억~3억, 5천만~1억, 1천만~5천만, 0~1천만, 부채없음)

→ 총 612개 코호트

(양 데이터의 코호트별 구성 비교)

- 가계부채DB와 신용카드 이용액 데이터의 전체 이용금액의 규모와 대체적인 추이가 비슷

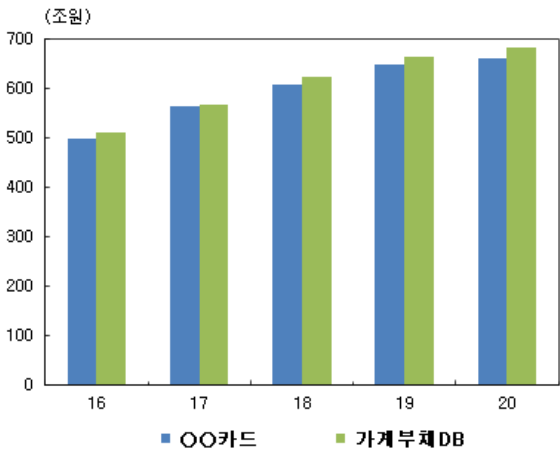
※ 해당 카드사의 시장점유율을 반영하여 전체 카드사 환산금액을 계산

- 연령, 지역, 부채규모별 신용카드 이용액의 구성비중도 양 데이터가 비슷한 모습

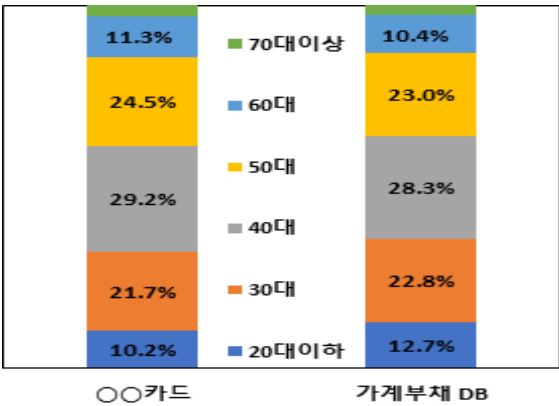
2. 연계가능성 검증

(양 데이터의 코호트별 구성 비교)

신용카드 이용액 총액 비교

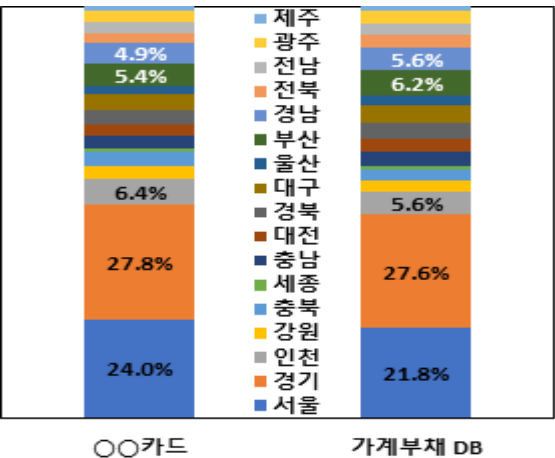


연령별 비중¹⁾ 비교

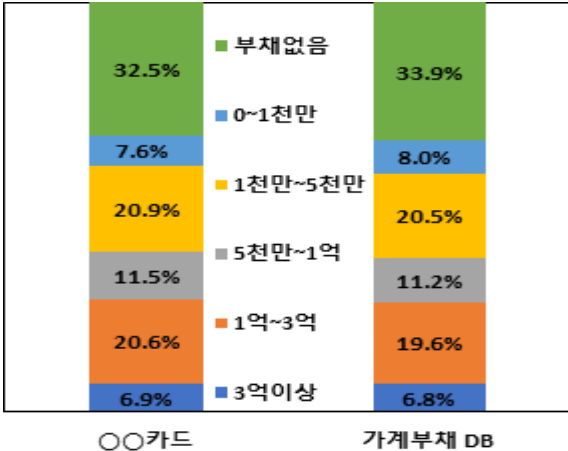


주: 1) 2020년 기준

지역별 비중¹⁾ 비교



부채규모별 비중¹⁾ 비교



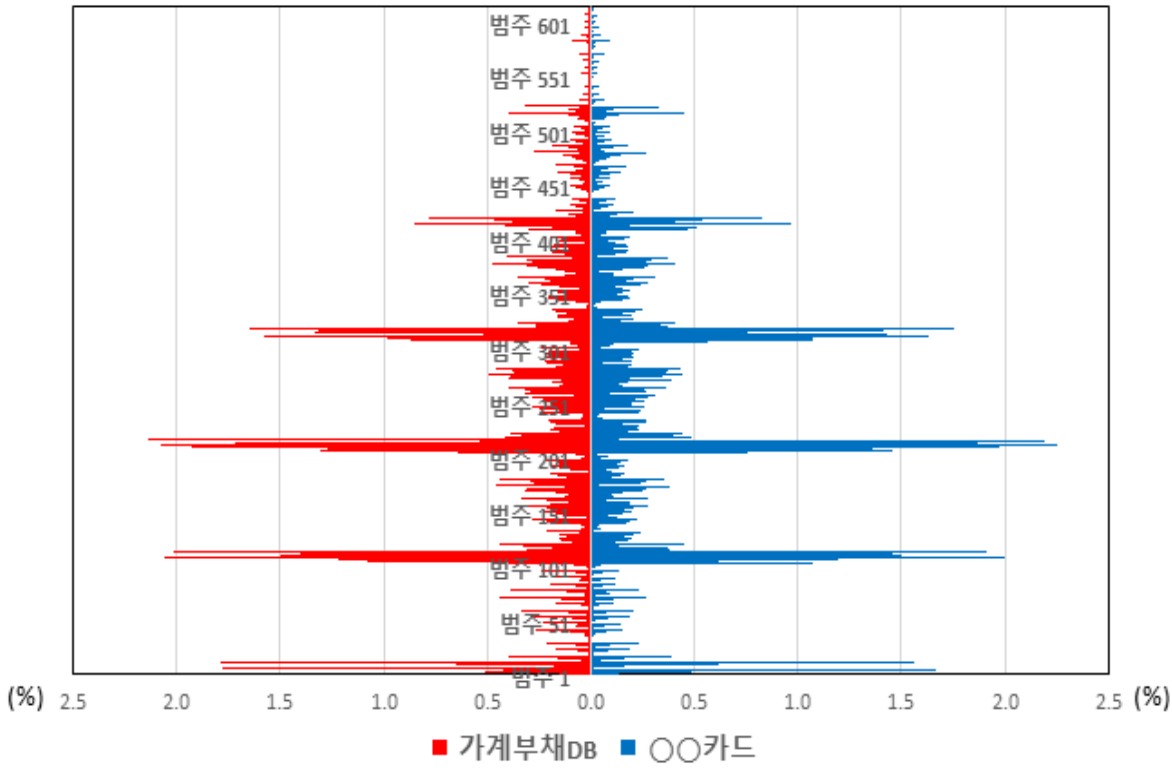
주: 1) 2020년 기준

주: 1) 2020년 기준

2. 연계가능성 검증

(양 데이터의 코호트별 구성 비교)

연령×지역×부채규모별 비교



주: 1) 2020년 기준

(연계데이터의 구조)

- 연령, 지역, 가계부채DB와 신용카드 데이터의 연령, 지역, 부채규모별 범주 및 조합의 카드이용액의 비중 분포가 유사한 것으로 나타남에 따라 612개 코호트에 대하여 양 데이터를 다음과 같이 연계

가계부채DB 및 신용카드사 데이터의 연계구조

[illegible]



(데이터간 연계로 유용성 제고)

- 가계부채**DB**의 다양한 대출 · 신용정보와 신용카드 데이터의 지출 내역을 연계 분석
- 조세 · 교육 · 행정 데이터 등 다양한 **DB**와의 연계를 통한 분석범위 확장
- 차주별 정보가 아니더라도 지역별 평균소득, 부동산 가격 상승률 등을 사용하여 분석범위 확대 가능

(향후 과제)

- 개인정보 보호와 빅데이터 활용 간 조화방안 모색
- 고빈도 데이터의 통계적 처리방안(계절조정 방법론 등) 연구
- 행정데이터, 스캐너 데이터 등 활용을 통한 가계부채**DB**와 신용카드 데이터의 유용성 제고

감사합니다.

