

한국국민의 가계 금융부채에 대한 체감도 분석

오만숙¹ · 현승미²

¹이화여자대학교 통계학과, ²이화여자대학교 통계학과

(2009년 1월 접수, 2009년 3월 채택)

요약

최근 금융위기의 요인이 되고 있는 가계부채에 대하여 가계구성원이 느끼는 부담감, 즉, 가계부채에 대한 체감도에 가계구성원의 속성들(주택집유형태, 가구주 학력, 가구주 연령, 월소득, 거주지역)이 미치는 영향을 2004년도 국민은행이 조사한 실데이터를 가지고 분석하였다. 체감도를 부채에 대한 부담감이 낮음과 높음의 이항자료로 구분하여 가계구성원의 속성들을 설명변수로 갖는 로지스틱 회귀분석을 수행하였다. 적합도에 대한 우도비 통계량을 이용한 후진제거법을 사용하여 간단하면서도 자료를 잘 적합시키는 모형을 선택한 결과 2개의 2차 교호작용을 갖는 모형이 선택되었다. 선택된 모형에 대한 계수 추정치를 통하여 각 속성이 부채 체감도에 대하여 미치는 영향을 분석하였다. 또한 가계부채의 유무에 대하여 가계구성원의 속성들이 미치는 영향을 로지스틱 회귀모형을 통하여 유사한 방법으로 분석하였다. 자가주택일수록, 월소득이 증가할수록, 가구주 학력이 낮을수록 그리고 가구주 연령이 낮아질수록 부채에 대한 체감도가 낮아짐을 알 수 있었다.

주요용어: 가계부채, 로지스틱 회귀모형, 후진제거법, 교호작용.

1. 서론

미국의 서브 프라임 모기지 부실로 촉발된 현재의 금융위기는 전 세계를 거의 공황상태로 빠뜨리고 있다. 부동산 구입을 위한 가계대출의 확대는 자산가격의 상승을 가져오고 자산가격의 상승은 다시 가계대출의 확대를 초래하는 이른바 대출과 자산가격이 상호 상승작용을 통해 자산가격의 버블을 유발하게 된다. 가계부채가 과도하게 누적된 상황 하에서 급격한 자산가격의 하락 혹은 금리의 급등은 가계소비의 급격한 위축을 초래하여 경기변동폭을 확대시키며, 금융기관의 부실채권을 증가시킴으로써 금융시스템의 안정성을 저해하는 부정적 요인으로 작용하게 된다.

한국의 가계 금융부채 상황을 보면 그림 1.1에서 보는 바와 같이 2000년 이후 급격한 가계대출의 증가가 나타나고 있다. 그림을 보면 1998년 외환위기와 2003년 카드대란 시 급감했던 가계 부채가 2008년 2/4분기 가계부채는 660조원으로 전분기 640조원에 비하여 20조원, 전년 동분기 596조원에 비하여 64조원으로 증가하여 증가폭이 확대되었다. 또한 이는 전년동분기 대비 10.7% 연간 증가율로 2007년도의 8.37%에 비하여 높은 수치로서 가계부채가 빠른 속도로 증가세가 확대되고 있음을 시사하고 있다.

가계부채와 소득간의 관계를 보면 신용카드 사용으로 불거진 “플라스틱 버블”이 붕괴된 이후 2004년 3/4분기까지 가계부채 증가율은 소득증가율을 하회하다가 2004년 4/4분기부터 부채 증가율이 소득 증가율을 상회하기 시작했다. 가계의 가처분 소득 대비 이자상환액 비율은 저금리의 기조로 인해 1999년 이후 지속적 하락세를 보이고 있어 가계 이자상환능력은 다소 개선된 것으로 보이는 반면 금융부채 대

이 연구는 한국과학재단의 기초연구지원 R01-2006-000-10563-0으로 수행되었음.

¹교신저자: (120-750) 서울시 서대문구 대현동 11-1, 이화여자대학교 통계학과, 교수. E-mail: msosh@ewha.ac.kr

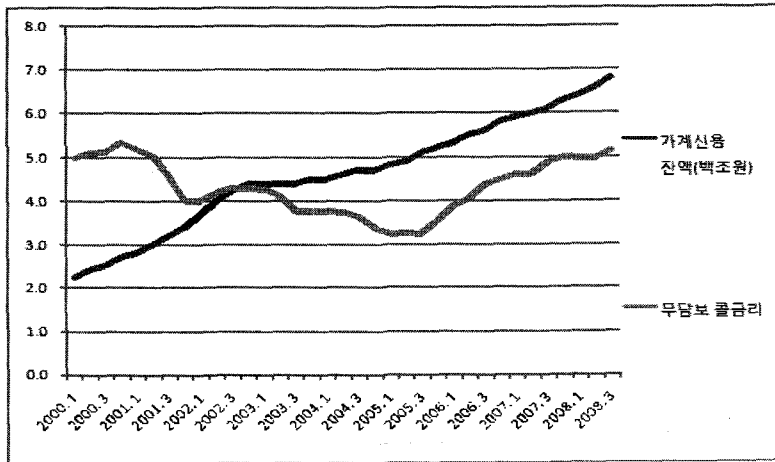


그림 1.1. 분기별 가계신용 잔액과 대출금리 변동 (한국은행 경제통계시스템 제공)

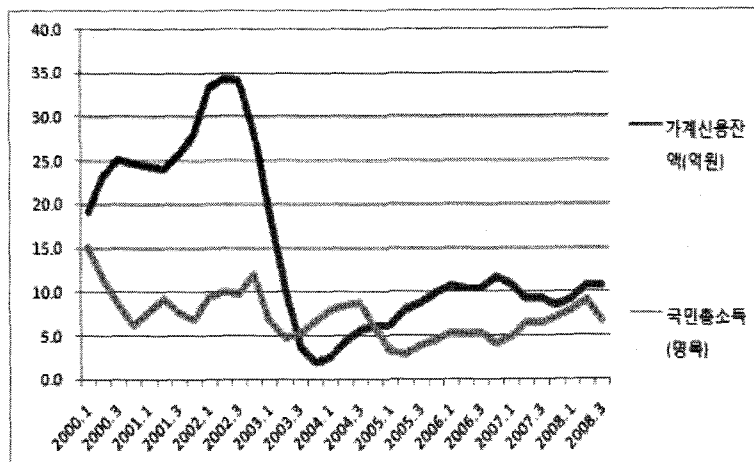


그림 1.2. 가계신용잔액과 국민총소득의 전년동기 대비 증가율

비 가치분소득 비율은 2001년 100%를 넘어 소득으로는 부채를 갚을 수 없는 상태이며, 가계의 금융자산 대비 금융부채 비율도 1999년 이후 점차 하락하고 있어 가계 부채상환능력이 약화되고 있음을 나타낸다 (김은혜, 2007). 구체적으로, 가계부채 증가율이 2004년 1/4분기 이후 상승 추세를 보이고 있는 반면 소득 증가율을 2004년 3/4분기 이후 둔화되고 있다. 가계부채 증가율을 전년 동기대비로 본다면 2004년 4/4분기가 6.1%, 2005년 4/4분기가 9.9%, 2006년 4/4분기가 11.6%, 2007년 4/4분기가 8.4%, 2007년 4/4분기대비 2008년도 2/4분기는 4.7%씩 증가하고 있는 반면 국민총소득은 전년동기대비 2004년 4/4분기 5.9%, 2005년 4/4분기 4.5%, 2006년 4/4분기 4.0%, 2007년 4/4분기 7.2%, 2007년 4/4분기 대비 2008년 2/4분기가 심지어는 -3.5%의 증가율을 보이고 있다 (윤소영, 2008). 그림 1.2와 같이 가계신용잔액의 전년동기대비 증가율과 명목 국민총소득의 전년동기대비 증가율을 비교하여 그래프로 그려보면 이러한 역전 현상을 뚜렷이 볼 수 있다.

가구당 빚이 상당히 높은 점과 가계의 소득대비 부채규모나 부채상환 부담 등을 고려할 때 가계부채 문제는 당분간 크게 개선되기 어려워 보인다. 가계부채의 증가는 가계부문의 부실화로 연결될 수 있고 가계부문의 부실화는 중산층의 붕괴 및 개인파산의 증가로 사회불안 요인으로 작용할 수 있을 뿐만 아니라 금융기관의 부실화 및 금융질서의 혼란을 초래하고 신용사회 정착의 걸림돌이 될 것이라는 부정적인 우려가 일부에서 제기되어 왔으며 현재는 이런 우려가 현실로 나타난 실정이다. 따라서 가계부채로 인한 경제문제를 막거나 줄이기 위해서는 가계부채에 대한 과학적인 분석과 가계부채의 증가에 미치는 영향들에 대한 연구가 요구되고 있다. 그러나 현재까지 수행된 분석들에서는 가계부채에 대한 원인으로 저금리, 정책부채, 환율 등 외적인 요인과의 관계를 분석하거나, 단순히 가계의 부채가 어느 정도이고 어떠한 그룹이 어느 정도의 부채를 갖고 있다는 기술적인 연구에 머무르고 있다. 또한 동시에 여러 요인들을 고려한 다변량 분석이 아닌 두 개의 변수에 대하여 이차원표를 정리하는 것에 머무르고 있다 (국민은행, 2004; 김은혜, 2007).

본 연구에서는 기존의 연구와 관점을 달리하여 가계부채에 대하여 가계구성원이 느끼는 체감도, 즉, 부담감의 정도에 주목하고자 한다. 가계부채는 소비의 감소, 즉, 내수의 감소로 이어져 디플레이션을 유발할 수 있는 점을 감안하면 단순히 가계부채의 액수만 고려할 것이 아니라 가계부채에 대하여 느끼는 가계구성원들의 체감도를 분석하는 것이 중요하다고 본다. 예를 들어, 같은 액수의 부채를 가지고 있다 할지라도 소득이 높은 사람과 낮은 사람이 느끼는 부담의 정도는 다를 것이며 또한 고정적인 수입이 있는 사람과 없는 사람이 느끼는 부담이 다를 것이고 이 체감도의 차이가 소비의 차이 또는 가계부채로 인한 경제적 활동에 차이를 가져올 것이다. 이러한 가계부채 체감도 분석은 DTI(Debt To Income ratio, 연소득 대비 대출금액 비율) 등과 같이 최근 주택담보 대출에서 정부가 소득에 따른 대출비율 등을 규제하는 것과도 연관이 있다고 할 수 있다.

가계부채에 대한 체감도에는 가계 구성원이 가지는 속성들, 예를 들면 주택이 자기소유인지 임차인지, 어느 지역에 거주하고 있는지, 어떤 직업을 가지고 있는지, 나이 등이 영향을 줄 수 있다. 따라서 본 연구에서는 이러한 가계구성원의 어떤 속성들이 체감도에 영향을 미치는지, 그 영향은 어떤 형태로 주어지는지 등에 관하여 분석할 것이다. 이를 통하여 각 속성의 수준별로 체감도가 얼마나 다른지 비교하고 각 수준별 부채체감 또는 부채 유무의 성향을 파악하여 향후 가계부채 산정에 있어 하나의 기준으로 사용하고자 한다. 이러한 기준이 확보되면 각 가계의 특성을 효과적으로 파악하여 가계대출 증가로 인한 금융기관 및 정부의 손실과 사전 대책 시스템을 구축할 수 있을 것이며, 또한 향후 특정 가계에 대한 부채 성향도 예측할 수 있을 것이다. 결국 본 논문의 결과물을 통하여 가계부채 부담의 완화를 위한 금융기관 또는 정부의 정책 시스템으로도 연계시킬 수 있을 것이다.

본 논문의 2절에서는 2004년 국민은행이 수집한 주택금융 수요실태조사 자료에 대해 설명한다. 3절에서는 위의 자료를 가지고 부채에 대한 체감도의 정도와 부채 유무에 영향을 미치는 가계 구성원의 속성들을 파악하는 분석을 수행한다. 4절은 요약과 결론에 할애한다.

2. 국민은행의 주택금융 수요실태조사 자료

국민은행(구 주택은행 포함) 부동산 Lab에서는 도시가구의 주택금융 이용실태와 이용계획 현황을 파악하여 주택금융제도 개선 및 주택정책 수립에 필요한 기초자료를 제공하기 위하여 1973년도부터 매년 주택관련 변수에 대한 조사를 실시하여 왔다 (국민은행, 2004). 조사지역은 서울(강남, 강북)과 6대 광역시(부산, 대구, 인천, 대전, 광주, 울산), 5대 신도시(성남(분당), 고양(일산), 안양(평촌), 군포(산본), 부천(중동)), 중소도시(용인, 의정부, 수원, 전주, 청주, 원주, 창원)이며 조사대상 지역내 거주하는 가구 중 가구주 연령이 만 20세 이상인 가구주 또는 가구주 배우자를 그 조사대상으로 하였다. 표본 할당

표 2.1. 분석에 사용된 설명변수들

변수명	수준
지역(area)	중소도시(city), 서울(seoul), 6대 광역시(wide)
주택점유형태(home)	자가(my), 임차(rent)
직업(job)	사무관리직(cont), 전문직(prof), 자영업(self), 기술기능직(tech), 기타(etc)
가구주학력(sch)	고등학교(12), 대학교(16)
월평균가구소득(won, 백만원)	100, 200, 300, 425, 700
가구주연령(age)	30, 45, 55, 70

표 2.2. 설명변수들 간의 공통 오즈비

변수	변수	오즈비	유의확률	변수	변수	오즈비	유의확률
area	home	6.4634	0.0395	home	area	6.4634	0.0395
	job	71.5242	<.0001		job	28.3396	<.0001
	sch	7.9401	0.0189		sch	3.7217	0.0537
	won	18.5610	0.0174		won	39.2004	<.0001
	age	29.7617	<.0001		age	189.0916	<.0001
job	area	71.5242	<.0001	sch	area	7.9401	0.0189
	home	28.3396	<.0001		home	3.7217	0.0537
	sch	432.0772	<.0001		job	432.0772	<.0001
	won	610.7917	<.0001		won	279.6581	<.0001
	age	621.4055	<.0001		age	117.4255	<.0001
won	area	18.5610	0.0174	age	area	29.7617	<.0001
	home	39.2004	<.0001		home	189.0916	<.0001
	job	610.7917	<.0001		job	621.4055	<.0001
	sch	279.6581	<.0001		sch	117.4255	<.0001
	age	265.0167	<.0001		won	265.0167	<.0001

방법은 지역별 인구비례에 의한 무작위 할당 추출법을, 자료수집 방법으로는 구조화된 질문지를 이용한 전화면접조사를 사용하였다. 본 연구에서는 위의 조사 중 가장 최근에 이루어진 2004년 10월 10일부터 10월 22일까지 13일간에 걸쳐 1차 랜덤조사에서 얻어진 2,000명의 자료를 분석하였다.

부채체감도의 네가지 수준(부채없음, 큰 걱정없음, 약간 걱정 있음, 걱정이 많음)을 반응변수로 보고 가계구성원의 속성에 대한 6개 변수(지역, 주택점유 형태, 직업, 가구주 학력, 월평균 가구소득, 가구주 연령)를 고려하였다. 그런데 지역을 너무 세분화할 경우 표본수가 매우 작은 셀이 나와 적합도 검정에 문제가 있을 수 있으므로 지역을 서울, 6대광역시, 중소도시(5대 신도시+중소도시)의 세 그룹으로 분류하였다. 또한 부채에 대한 무응답과 모름 응답은 상대적으로 전체 표본에서 차지하는 비율이 매우 적으므로 분석에서 제외하기로 한다. 이상의 설명 변수들에 대하여 이하 SAS 프로그램에서 사용된 변수명과 각 변수의 수준을 정리하면 표 2.1와 같다.

설명 변수들 중에서 지역, 주택점유형태, 직업은 질적 범주형 변수이고 가구주학력과 월평균가구소득, 가구주 연령은 양적 변수로 취급할 수 있다. 그런데 이상의 설명변수들을 모형에 모두 포함시켰을 때 중복성으로 인한 문제가 발생할 수 있으므로 설명변수들간의 연관성을 공통 오즈비를 통하여 아래와 같이 점검해 보았다. 표 2.2에 의하면 변수들간의 연관성이 대체적으로 매우 높는데 특히 ‘직업(job)’ 변수는 모든 다른 변수들과의 오즈비가 매우 높다. 변수의 개수가 많은 것이 아니므로 연관성이 높은 변수를 모두 제거하기는 어렵고 우선 ‘직업’ 변수를 제외하고 분석하기로 한다.

3. 로지스틱 회귀모형을 이용한 부채 체감도 분석

3.1. 부채 체감도 낮음과 높음에 대한 분석

부채에 대한 부담의 정도에 대한 문항을 살펴보면, 부채가 없음, 부채가 있으나 부채에 대하여 큰 걱정 없음, 약간 걱정 있음, 걱정이 많음의 네 범주가 있다. 큰 걱정 없음과 약간 걱정 있음의 문항 차이가 약간 모호하나 원자료에서 기술된 문맥과 범주의 차례를 고려하면 큰 걱정 없음, 약간 걱정 있음, 걱정이 많음의 차례대로 부담이 늘어남을 짐작할 수 있다. 부채가 없거나 부채가 있지만 부채에 대한 큰 걱정이 없는 경우에는 나머지 경우에 비하여 차후 소비활동에 영향을 덜 준다고 간주하여 이 두 범주를 체감도가 낮음으로 보고 나머지 약간 걱정 있음, 걱정이 많음의 두 범주를 체감도 높음으로 분류하여 반응변수를 이항변수로 (낮음 = 1, 높음 = 0)하는 로지스틱 회귀 모형을 사용하여 분석하기로 한다

본 논문의 데이터에서는 직업을 제외하더라도 최종 5차까지의 교호작용이 일어날 수 있다. 그러나 3차 이상의 교호작용에 대하여는 오즈비 해석이 어려울 뿐만 아니라 그 모형이 갖는 의미도 무의미해질 수 있으므로 우선 주효과 항과 5개의 변수에 의해 생길 수 있는 2차 교호작용을 모두 넣은 모형 M_0 를 고려하기로 한다. 즉, M_0 모형은 상수항과 주효과, 10개의 2차 교호작용을 포함한 다음과 같은 모형이다.

$$\begin{aligned} M_0 : \text{logit}(\pi) &= \log \frac{P(Y=1)}{P(Y=0)} \\ &= \alpha + \beta_1^{(1)} \text{city} + \beta_1^{(2)} \text{seoul} + \beta_2 \text{my} + \beta_3 \text{sch} + \beta_4 \text{won} + \beta_5 \text{age} \\ &\quad + \gamma_1^{(1)} \text{city} * \text{my} + \gamma_1^{(2)} \text{seoul} * \text{my} + \gamma_2^{(1)} \text{city} * \text{sch} + \gamma_2^{(2)} \text{seoul} * \text{sch} \\ &\quad + \gamma_3^{(1)} \text{city} * \text{won} + \gamma_3^{(2)} \text{seoul} * \text{won} + \gamma_4^{(1)} \text{city} * \text{age} + \gamma_4^{(2)} \text{seoul} * \text{age} + \gamma_5 \text{my} * \text{sch} \\ &\quad + \gamma_6 \text{my} * \text{won} + \gamma_7 \text{my} * \text{age} + \gamma_8 \text{sch} * \text{won} + \gamma_9 \text{sch} * \text{age} + \gamma_{10} \text{won} * \text{age}. \end{aligned}$$

위 모형에서 city와 seoul은 ‘지역(area)’이 각각 중소도시와 서울임을 나타내는 dummy 변수이고 my는 ‘주택형태(home)’이 자가인 경우 my = 1을 갖는 dummy 변수이다. ‘학력(school)’, ‘월소득(won)’, ‘연령(age)’은 양적 변수로 간주하고 로짓에 대한 조건부 선형관계를 가정하였다 (박태성과 이승연, 1998).

자료에 적합하면서도 단순한 모형을 찾기 위하여 우도비 적합 통계량의 차이를 통하여 복잡한 모형으로부터 시작하여 유의하지 않은 항들을 제거해나가는 후진제거법(Backward Elimination Procedure)을 적용하기로 한다. 표 3.1-3.2는 5개의 변수를 이용한 여러 로지스틱 회귀모형들을 적합하여 비교한 결과들이다. 각 모형에 대한 G^2 은 해당 모형의 적합도에 대한 우도비 통계량이고 두 모형의 ‘차이’는 비교대상인 두 모형의 우도비 통계량의 차이이다. 이 방법을 이용하여 각 단계에서 두 모형의 차이가 유의한지 검정해서 가장 큰 유의확률을 갖는 모형을 선택하였다. 예를 들면, 모형 1은 모형 M_0 에서 지역과 주택점유형태의 교호작용, 즉, area * home의 모든 수준 (city * my, seoul * my)을 제외한 모형이고 모형 2는 M_0 에서 교호작용 sch * area을 제외한 모형이다. 모형 1부터 모형 10은 M_0 에서 교호작용이 하나 제거된 모형이다. 후진제거법의 첫번째 단계로 모형 1-10 중에서 모형 7과 M_0 의 차이에 대한 유의확률이 가장 크고 또한 모형 7이 적합도에 대한 우도비 검정 통계량의 유의확률 (4번째 열)이 0.05보다 크므로 모형 7을 선택한다. 두번째 단계에서는 모형 7에서 교호작용을 하나씩 제거한 것이 모형 11부터 모형 19인데 이 중 모형 11의 차이에 대한 유의확률이 가장 크고 이 모형의 적합 유의확률이 0.05보다 크므로 모형 11을 선택한다. 세번째 단계에서는 모형 11에서 sch * area 교호작용이 제거된 모형 21이 선택된다. 이런 식으로 8번째 단계에서 모형 71이 선택된다. 9번째 단계에서 모형 71에서 교호작용을 하나씩 제거한 모형 81과 82가 있는데 이 두 모형 모두 차이에 대한 유의확률이 0.05보다 작으므로 후진제거를 종료하고 최종적으로 모형 71을 채택하게 된다.

최종적으로 선택된 모형은 주효과에 교호작용 $sch * won$ 과 $age * my$ 이 추가된

$$\begin{aligned} \text{logit}(\pi) = & \alpha + \beta_1^{(1)} \text{city} + \beta_1^{(2)} \text{seoul} + \beta_2 \text{my} + \beta_3 \text{sch} \\ & + \beta_4 \text{won} + \beta_5 \text{age} + \gamma_1 \text{sch} * \text{won} + \gamma_2 \text{age} * \text{home} \end{aligned}$$

이다. SAS의 PROC CATMOD 절차로부터 각 계수의 추정치를 구하고 이를 각 설명변수의 범주에 대하여 표시하면 다음과 같다 (조인호, 2004).

$$\begin{aligned} \text{logit}(\pi) = & -0.5601 + 0.1463 \text{city} + 0.0498 \text{seoul} + 0.7650 \text{my} + 0.0627 \text{sch} \\ & + 0.576 \text{won} - 0.0237 \text{age} - 0.0107 \text{age} * \text{my} - 0.034 \text{sch} * \text{won}. \end{aligned}$$

추정된 계수를 해석해 보자. 먼저 지역 변수를 보면 이차교호작용 없이 주효과만 존재한다. 계수의 크기를 보면 중소도시(city), 서울, 6대 광역시 순으로 부채에 대한 체감도가 낮음을 알 수 있다. 서울에 대한 계수는 0.0498인데 ASE가 0.718이므로 광역시와 유의하게 다르지 않음을 나타낸다. 따라서 지역 변수에서는 중소도시만 다른 지역에 비해서 체감도가 낮으며, 지역을 제외한 다른 변수들이 고정되어 있을 때 체감도 낮음과 높음의 오차가 중소도시의 경우 광역시 또는 서울의 $e^{0.1463} = 1.158$ 배, 즉, 약 15% 증가함을 의미한다. ASE를 고려하여 오즈비의 95% 신뢰구간을 구하면 (1.010, 1.327)로 1을 포함하지는 않지만 하한가가 1에 근접하므로 중소도시 또한 다른 지역과 유의하게 다르다고 결론 짓기에는 무리가 있다.

지역을 제외한 요인들인 주택점유형태, 가구주학력, 월평균가구소득, 가구주연령은 이차 교호작용 항을 가지고 있다. 다른 요인들이 고정되어 있을 때 고졸과 대졸의 학력에 따른 부채체감이 낮음의 로그 오즈비는

$$\log \frac{\pi_{sch=12}/(1 - \pi_{sch=12})}{\pi_{sch=16}/(1 - \pi_{sch=16})} = 0.0627(12 - 16) - 0.034(12 - 16) * \text{won} = -0.2508 + 0.1360 * \text{won}$$

이므로 소득이 늘어남에 따라 로그오즈비는 선형적으로 증가한다. 오즈비를 보면 월소득이 100만원 증가할 때 마다 오즈비는 $e^{0.136} = 1.146$ 배 증가한다. 월소득이 100만원일 경우 로그오즈비(오즈비)는 $-0.1148(0.8915)$, 200만원일 경우 $0.0212(1.0214)$, 300만원일 경우 $0.1572(1.1702)$, 425만원일 경우 $0.3272(1.3871)$, 700만원일 경우 $0.7012(2.0162)$ 로, 소득이 100만원인 경우는 가구주 학력이 대학교일 경우 고등학교인 경우에 비하여 부채에 대한 부담이 적지만 200만원 이상 부터는 고등학교인 경우가 대학교에 비하여 부담이 더 적음을 알 수 있다.

다른 요인들이 고정되어 있을 때 월소득에 따른 체감도의 변화를 살펴보자. 월소득과 학력이 교호작용이 있으므로

$$\log \frac{\pi_{won=200}/(1 - \pi_{won=200})}{\pi_{won=100}/(1 - \pi_{won=100})} = 0.576 - 0.034 * \text{sch}$$

이다. 가구주 학력이 고등학교일 때 로그오즈비는 $0.576 - 0.034 * 12 = 0.168$ 이다. 따라서 가구주 학력이 고등학교일 때 월소득이 100만원과 200만원을 비교하면 200만원의 경우 체감도가 더 낮다. 마찬가지로 가구주 학력이 고등학교일 때 월소득이 늘어남에 따라 부채에 대한 부담이 낮아지는 것을 알 수 있다. 가구주 학력이 대학교일 때 월소득 100만원과 200만원을 비교하면 로그오즈비는 $0.576 - 0.034 * 16 = 0.032$ 로 월소득이 늘어남에 따라 부채에 대한 부담이 낮아지기는 하지만 학력이 고등학교인 경우와 비교하면 로그오즈비의 증가분이 현저히 작아진다. 이는 대졸 학력의 경우 고졸에 비하여 월소득이 부채 체감도에 미치는 영향이 매우 낮음을 의미한다.

표 3.1. 로지스틱 모형에서의 후진제거법

소거항	자유도	G^2	p -value	모형비교	차이(자유도)	p -value
a. M_0	183	203.38	0.1439			
1. area * home	185	212.57	0.0804	1-a	0.38(2)	0.830
2. sch * area	185	212.63	0.0800	2-a	0.44(2)	0.800
3. sch * home	184	213.61	0.0666	3-a	1.42(1)	0.240
4. won * area	185	217.08	0.0532	4-a	4.89(2)	0.090
5. won * home	184	213.82	0.0653	5-a	1.63(1)	0.217
6. won * home	184	214.64	0.0605	6-a	2.45(1)	0.128
*7. age * area	185	212.50	0.0809	7-a	0.31(2)	0.859
8. age * home	184	220.22	0.0351	8-a	8.03(1)	< 0.005
9. sch * age	184	214.41	0.0618	9-a	2.22(1)	0.153
10. won * age	184	213.95	0.0645	10-a	1.76(1)	0.203
*11. area * home	187	212.74	0.0953	11-7	0.24(2)	0.896
12. sch * area	187	212.86	0.0944	12-7	0.36(2)	0.839
13. sch * home	186	213.93	0.0785	13-7	1.43(1)	0.238
14. won * age	187	217.25	0.0641	14-7	4.75(2)	0.095
15. won * home	186	214.03	0.0778	15-7	1.53(1)	0.227
16. sch * won	186	214.99	0.0714	16-7	2.49(1)	0.124
17. age * home	186	220.38	0.0430	17-7	7.88(1)	0.005
18. sch * age	186	214.66	0.0735	18-7	2.16(1)	0.159
19. won * age	186	214.16	0.0769	19-7	1.66(1)	0.213
*21. sch * area	189	213.11	0.1103	21-11	0.37(2)	0.834
22. sch * home	188	214.10	0.0930	22-11	1.36(1)	0.246
23. won * area	189	217.34	0.0772	23-11	4.6(2)	0.100
24. won * home	188	214.18	0.0924	24-11	1.44(1)	0.237
25. sch * won	188	215.23	0.0845	25-11	2.49(1)	0.124
26. age * home	188	220.56	0.0521	26-11	7.82(1)	0.005
27. sch * age	188	214.93	0.0866	27-11	2.19(1)	0.156
28. won * age	188	214.37	0.0910	28-11	1.63(1)	0.217
*31. sch * home	190	214.45	0.1078	31-21	1.34(1)	0.248
32. won * area	191	218.64	0.0830	32-21	5.53(2)	0.067
33. won * home	190	214.58	0.1067	33-21	1.47(1)	0.234
34. sch * won	190	215.65	0.0977	34-21	2.54(1)	0.118
35. age * home	190	221.05	0.0609	35-21	7.94(1)	< 0.005
36. sch * age	190	215.46	0.0993	36-21	2.35(1)	0.139
37. won * age	190	214.70	0.1057	37-21	1.59(1)	0.221

가구주 연령을 나타내는 변수 age는 주택의 점유형태를 나타내는 my(자가=1, 임차=0)와 연관되어 있다. 다른 요인들이 모두 고정되어 있을 때 가구주 연령 30과 45을 비교하면

$$\log \frac{\pi_{age=30}/(1-\pi_{age=30})}{\pi_{age=45}/(1-\pi_{age=45})} = -0.0237 * (30 - 45) - 0.0107 * (30 - 45) * my = 0.237 + 0.107 * my$$

이다. 위 식의 계수가 모두 양수이므로 연령이 낮아질수록 부채에 대한 부담감은 줄어드는데 자가 주택일 경우 연령에 따른 영향이 더 크다. 구체적으로, 30세 가구주는 45세 가구주에 비하여 체감도 낮음/높음의 오즈비가 임차인 경우 $e^{0.237} = 1.43$ 배, 자가인 경우 $e^{0.237+0.107} = 1.68$ 배 크다. 즉, 30세 가구주의 경우 부채에 대한 부담감이 45세 가구주에 비하여 매우 낮는데 특히 자가 주택인 경우 낮음의 오

표 3.2. 로지스틱 모형에서의 후진제거법 (표 3.1 연속)

소거항	자유도	G^2	p -value	모형비교	차이(자유도)	p -value
41. won * area	192	220.03	0.0808	41-31	5.58(2)	0.065
42. won * home	191	217.05	0.0951	42-31	2.6(1)	0.112
43. sch * won	191	217.36	0.0926	43-31	2.91(1)	0.091
44. age * home	191	221.25	0.0660	44-31	6.8(1)	0.009
*45. sch * age	191	215.90	0.1045	45-31	1.45(1)	0.236
46. won * age	191	216.31	0.1011	46-31	1.86(1)	0.192
51. won * area	193	221.45	0.0784	51-45	5.55(2)	0.066
*52. won * home	192	218.45	0.0924	52-45	2.55(1)	0.117
53. sch * won	192	218.67	0.0907	53-45	2.77(1)	0.097
54. age * home	192	222.33	0.0659	54-45	6.43(1)	0.012
55. won * age	192	219.06	0.0878	55-45	3.16(1)	0.080
61. won * area	194	223.11	0.0745	61-52	4.66(2)	0.098
62. sch * won	193	222.11	0.0740	62-52	3.66(1)	0.058
63. age * home	193	223.87	0.0632	63-52	5.42(1)	0.021
*64. won * age	193	220.20	0.0872	64-52	1.75(1)	0.204
*71. won * area	195	224.62	0.0717	71-64	4.42(2)	0.115
72. sch * won	194	225.09	0.0625	72-64	4.89(1)	0.028
73. age * home	194	225.16	0.0620	76-64	4.96(1)	0.026
81. sch * won	196	230.06	0.0482	81-71	5.44(1)	0.029
82. age * home	196	229.37	0.0514	82-71	4.75(1)	0.031

즈비가 68%나 증가함을 보여준다. 나이 차이가 커짐에 따라 차이는 더욱 뚜렷하여 30세와 70세를 비교하면 30세의 낮은 오즈비가 70세의 2.58배가 된다.

끝으로 주택의 점유형태를 나타내는 my 변수의 영향을 살펴보자. 다른 요인들이 모두 고정되어 있을 때 주택형태 my = 1(자가)와 my = 0(임차)를 비교하면

$$\log \frac{\pi_{my=1}/(1-\pi_{my=1})}{\pi_{my=0}/(1-\pi_{my=0})} = 0.7650 - 0.0107 * age$$

이다. 가구주 연령 30, 45, 55, 70을 위 식에 대입하면 로그오즈비가 0.444, 0.2835, 0.1765, 0.016로 주어진다. 모든 연령에서 자가 주택인 경우 부채에 대한 부담감이 적으나 연령이 늘어날수록 그 영향은 현저히 줄어들음을 알 수 있다. 특히 70세 가구주의 경우 로그 오즈비가 거의 0에 가까워 주택형태가 부채 체감도에 거의 영향을 미치지 못함을 보여주고 있다.

이상의 결과를 종합하여 부채 체감도에 대한 가계 구성원 속성들의 영향을 정리하면 다음과 같다. 먼저 거주지역이 미치는 영향을 보면 중소도시 거주자가 서울이나 광역시에 비하여 낮은 부채 부담을 가지고 있으나 그 차이에 대한 유의성이 약하다. 월평균소득 100만원 가구의 경우 대졸 가구주의 체감도가 더 낮으나 소득 200만원 이상에서는 고졸 가구주의 체감도가 대졸에 비하여 더 낮다. 월평균 소득이 늘어날수록 부담감은 낮아지는데 가구주 학력이 고등학교일 경우가 대학교일 경우 보다 월소득의 영향을 더 많이 받는다. 구체적으로 고졸의 경우 월소득 100만원 증가 당 높음 대비 낮은 오즈가 18% 증가하는데 반하여 대졸의 경우는 약 3% 증가에 불과하다. 가구주 연령과 주택 점유형태를 살펴보면, 가구주 연령이 낮을수록, 주택이 자가일수록 체감도가 현저하게 낮아진다. 주택형태의 영향은 연령이 어릴수록 더 뚜렷하고 연령이 높아질수록 작아져 연령이 70세일 경우는 주택형태의 영향이 거의 없다.

표 3.3. 부채 유무에 대한 로지스틱 모형에서의 후진제거법

소거항	자유도	G ²	p-value	모형비교	차이(자유도)	p-value
a. M ₀	183	203.38	0.1439			
1. area * home	185	203.51	0.1669	1-a	0.13(2)	0.938
11. sch * area	187	204.25	0.1841	11-1	0.74(2)	0.69984
25. age * area	189	204.89	0.2034	25-11	0.64(2)	0.727
36. sch * age	190	205.14	0.2145	36-25	0.25(1)	0.645
42. won * area	192	206.42	0.2260	42-36	1.28(2)	0.534
55. won * age	193	206.82	0.2355	55-42	0.40(1)	0.536
61. sch * home	194	207.30	0.2439	61-55	0.48(1)	0.492
71. won * home	195	209.02	0.2335	71-61	1.72(1)	0.193
82. age * home	196	211.25	0.2164	82-71	2.23(1)	0.145
91. sch * won	197	214.01	0.1930	91-82	2.09(1)	0.158
cc. sch	198	214.02	0.2068	cc-91	0.01(1)	0.924
aaa. Area	200	216.83	0.1970	aaa-cc	2.81(2)	0.246

3.2. 부채 유무에 대한 분석

이번에는 부채의 유무에 대한 설명변수의 영향을 분석하기 위하여 반응변수를 부채없음(1)과 부채있음(0)의 이항자료로 취급하여 로지스틱 회귀 모형을 사용하여 분석하기로 한다. 여기에서 부채있음은 부채에 대하여 큰 걱정 없음, 약간 걱정 있음, 걱정이 많음의 세 범주를 나타낸다.

3.1절에서와 같이 모든 이차교호작용을 포함한 M₀ 모형에서 출발하여 후진제거법을 적용한 결과 각 단계마다 채택된 모형이 표 3.3에 제시되어 있다. 최종적으로 선택된 모형은 주효과에서 ‘학력’과 ‘지역’이 소거된 aaa 모형으로

$$\log \frac{P(Y = 1)}{P(Y = 0)} = \alpha + \beta_1 my + \beta_2 won + \beta_3 age$$

이다. 이 모형을 SAS PROC CATMOD로 추정하면

$$\log \frac{P(Y = 1)}{P(Y = 0)} = -0.5565 + 0.2847my + 0.177won - 0.0302age$$

로 모든 계수는 유의하다.

이 모형은 교호작용이 없이 주효과만 존재하는 모형으로 해석이 매우 쉽다. 가구주 학력과 거주지역은 부채 유무에 유의한 영향을 미치지 않는다. 다른 변수들에 대해서는, 자가주택일수록, 월소득이 증가할수록 그리고 가구주 연령이 낮아질수록 부채가 없을 확률이 높아진다. 또한 두 변수가 고정되었을 때 그 고정된 값에 상관없이 남은 한 변수의 부채 유무에 대한 영향은 동일하다. 예를 들면, 소득과 연령이 고정되었을 때 소득과 연령 수준에 상관없이 자가 주택의 경우의 오즈가 임차인 경우의 오즈보다 $e^{0.2847} = 1.329$ 배 높다. 즉, 소득 100만원의 고등학교 학력인 가구주가 자가주택을 소유한 경우는 같은 소득과 학력에서 임차를 한 경우에 비하여 부채 없음/있음 오즈가 1.329배이고 마찬가지로 소득 700만원의 대학교 학력인 가구주가 자가주택을 소유한 경우는 같은 소득과 학력에서 임차를 한 경우에 비하여 부채 없음/있음 오즈가 1.329배라는 것을 의미한다. 마찬가지로 주택형태와 연령이 고정되었을 때 그 수준에 상관없이 소득이 100만원 증가하면 부채없음의 오즈는 $e^{0.177} = 1.194$ 배 만큼 증가한다. 또한 주택형태와 소득이 고정되었을 때 연령이 10세 감소하면 부채없음의 오즈는 $e^{0.302} = 1.353$ 배 만큼 증가한다.

4. 요약과 결론

최근 금융위기의 요인이 되고 있는 가계부채에 대한 체감도에 가계구성원의 속성들(주택점유형태, 가구주 학력, 가구주 연령, 월소득, 거주지역)이 미치는 영향에 대하여 2004년도 국민은행이 조사한 실제자료를 가지고 분석하였다. 체감도를 부채에 대한 부담감이 낮음과 높음의 이항자료로 구분하여 가계구성원의 속성들을 설명변수로 갖는 로지스틱 회귀분석을 수행하였다. 적합도에 대한 우도비 통계량을 이용한 후진제거법을 사용하여 간단하면서도 자료를 작 적합시키는 모형을 선택한 결과 2개의 2차 교호작용을 갖는 모형이 선택되었다. 선택된 모형에 대한 계수 추정치를 통하여 각 속성이 부채 체감도에 대하여 미치는 영향을 분석하였다.

분석 결과 월소득이 증가할수록, 그리고 월소득 100만원을 제외하고는 고졸 가구주가 대졸 가구주에 비하여, 부채에 대한 체감도가 낮음을 알 수 있었다. 특히 소득이 미치는 영향은 고졸 가구주에게 더욱 뚜렷하였고 대졸 가구주에게는 상대적으로 미미하였다. 또한 가구주 연령이 적을수록, 자가 주택이 임차에 비하여, 부채에 대한 체감도가 낮는데 특히 젊은 층에서 자가/임차의 영향이 더 컸다.

또한 가계부채의 유무에 대하여 가계구성원의 속성들이 미치는 영향을 로지스틱 회귀모형을 통하여 유사한 방법으로 분석하였다. 가계부채의 유무에 대한 모형에서는 ‘지역’과 ‘학력’ 변수가 제거되었고 나머지 속성들이 교호작용없이 주효과만 갖는 간단한 모형이 후진제거법에 의하여 선택되었다. 고려된 속성들 중 ‘지역’과 ‘학력’은 부채 유무와 관련이 없고, 모형의 계수 추정치를 볼 때 자가주택일수록, 월소득이 증가할수록 그리고 가구주 연령이 낮아질수록 부채가 없을 확률이 높아짐을 알 수 있었다.

국민은행에서 얻을 수 있는 자료로 2004년 자료가 가장 최근의 것이어서 아쉬운 점이 있다. 그러나 2004년 이후 2008년 말까지 급변하는 경제 상황이 없어 본 분석에서의 결과가 부채에 대한 가계 속성의 영향을 파악하는데 유용하게 사용될 수 있을 것으로 본다.

본 논문에서는 관심사를 부채 체감도의 낮음과 높음 그리고 부채의 유무에 두고 체감도 자료를 이항자료로 바꾸어 로지스틱 모형을 적용시켰다. 원 자료에서는 체감도의 정도가 네 단계로 주어지는데, 만약 체감도의 각 수준이 관심이라면 이를 순서형 범주형 자료로 간주하여 적합한 모형을 찾는 것도 시도할 만하다.

참고문헌

- 국민은행 (2004). 주택금융 수요실태조사, 부동산 Lab 보고서.
 김은혜 (2007). 가계부채 증가의 결정요인 분석, 건국대학교 경제학과 석사학위 청구논문.
 박태성, 이승연 공역 (1998). <Alan Agresti 범주형자료분석개론>, 자유아카데미.
 윤소영 (2008). <2008년 2/4분기중 가계신용 동향>, 한국은행 경제통계국 금융통계팀.
 조인호 (2004). <SAS 강좌와 통계컨설팅 >, 2nd Edition, 영진.COM.

Analysis of Stress Level of Korean Household Members due to Household Debt

Man-Suk Oh¹ · Seung Me Hyun²

¹Department of Statistics, Ewha Womans University;

²Department of Statistics, Ewha Womans University

(Received January 2009; accepted March 2009)

Abstract

Korean household debt is one of the main sources of the current financial crisis. This paper studies the impact of household members' attributes such as a type of housing(self-own or rent), education, age, average monthly income of the head of household, and the area of residence, on the stress level of the household members due to household debt. We analyze a real data set collected by KB Kookmin Bank in 2004. We consider low and high stress level as a binary response variable and use a logistic regression model with the attributes of household members as explanatory variables. A simple but well-fitting model is selected by backward elimination method based on the likelihood statistic for goodness-of-fit test, and the impact of the attributes on the stress level is studied from parameter estimates of the selected model. We also perform the similar analysis on a binary response variable which distinguishes households with no debt from the rest. From the analysis, the stress level tends to be low for households with self-own houses, high average monthly income, low education level, and young members.

Keywords: Household debt, logistic model, backward elimination, association.

This work was supported by Grant R01-2006-000-10563-0 from the Basic Research Program of Korea Science & Engineering Foundation.

¹Corresponding author: Professor, Department of Statistics, Ewha Womans University, 21 Daehyun-Dong, Seodaemoon-Gu, Seoul 120-750, Korea. E-mail: msoh@ewha.ac.kr