

집세조사 방법론

- 추정법 중심으로 고찰 -

2006년 12월

연구자 : 통계개발원 경제통계실 전 백근

Tel. (042)717-0227

e-mail : chunbg@nso.go.kr

主 要 內 容

- 본연구의 목적은 집세지수의 현실반영도를 제고하기 위하여 추정 방법을 제시하고자 한다.
- 집세지수에 관한 기본개념, 작성현황과 집세조사 표본현황과 문제점을 살펴보고, 문제를 해결하기 위하여 통계적 추정방법인 확률표본에서의 비추정법을 설명하고 가계조사자료를 이용하여 실증 분석하였다.
- 전·월세금액을 추정함에 있어 보조변수로 연면적(평)을 이용하여 평당 단가의 개념으로 집세지수를 추정하는 3가지 개선안과 현행 집세지수 집계 체계를 주택유형별로 분리하여 집계한 개선안과 함께 4개안을 제안하였다.
- 결론적으로 전·월세금액을 추정함에 있어 연면적, 사용방수, 가구원수 등 다양한 보조변수를 이용하는 것이 집세지수의 정확성과 신뢰성을 높이는 방안이라고 판단된다.

1. 서론

최근 부동산가격의 급등이 가구의 주거비를 증가시킴에 따라 가구에서는 물가 상승으로 느낀다. 부동산 가격의 상승이 소비자물가지수 항목인 집세지수에 영향을 주게 되고, 집세지수의 가중치가 가장 크기 때문에 집세지수의 큰 변동은 소비자물가지수에 많은 영향을 미친다. 따라서 소비자물가지수를 작성할 때 집세의 가격변동을 정확히 파악하는 것이 상당히 중요하다. 집세지수를 작성하기 위한 통계청 집세조사는 1955년부터 행하여져 왔으며, 현재 조사되고 있는 표본은 2000년 인구주택 총조사 결과를 기초로 2002년에 설계되어 집세가 조사되고 있다.

소비자물가지수(CPI)는 소비생활에 필요한 상품과 서비스를 구입하기 위해 지불하는 가격의 변동을 종합한 것으로 소비자가 구입하는 상품과 서비스의 가격변동이 가구의 소비생활에 얼마만큼 영향을 주는지를 알 수 있는 지표가 된다. 이에 따라 소비자물가지수는 한 나라의 경제동향분석이나 경제정책수립 등에 필요한 기초자료로 활용한다. 즉, 각 나라에서 소비측면에서 인플레이션을 측정하게 함으로써 거시경제 관리지표로서 사용되고, 화폐소득의 구매력 변화를 측정할 수 있어 임금과 연금, 보험료 등의 인상율과 연동되어 있으며, 경제성장률, 생산성, 소득, 임금 등 각종 실물 경제지표의 현재가치를 고정가치로 전환시키기 위한 디플레이터(deflator)로서 사용되기도 한다.

통계청에서는 소비자물가지수의 주거비부문에 있어 자가주거비는 보조지표로 작성¹⁾하여 발표해 오고 있다. 그동안 CPI에서는 자가주택을 자본재로 보아 지수에 반영하지 않았으나, 주거비(생계비)개념의 물가지수 작성의 필요성이 대두되면서 자가주거비 산출을 고려하여 '95년 기준 개편 당시 집세상당액 접근법을 채택하여 귀속가임을 산정하고 "자가주거비 포함 CPI"를 보조지표로만 작성하여 발표해 오고 있다.(통계청, 2005a)

1) 통계청 가계조사에서 자가주택에 대하여 자가 주거의 가격은 유사 주택을 임차하는 것과 같다고 보는 방법인 집세상당액 접근법으로 조사되어 공식적인 발표는 되지 않고 있지만, 매년 발간되는 소비자물가지조사 통계연보에 수록되고 있음

통계청의 집세조사¹ 이외에는 국민은행에서 주택시장의 전세가격지수²를 작성하고 있는데, 작성방법에서 실질적으로 거래되는 가격(contract price)으로 지수를 작성하기 보다는 시세 혹은 호가의 움직임을 지수에 반영하고 있다. 따라서 전세가격지수에 의해 소비자의 주거비에 대한 움직임을 파악하기에는 한계를 가지고 있다.

미국, 캐나다 등 외국에서는 자가주거비를 국별로 다양하게 추정하여 CPI에 포함시키고 있다. 자가주거비를 미국과 일본에서는 귀속임대로 개념으로 파악하는 임차료 상당액 접근법으로, 캐나다, 핀란드에서는 주택보유 관련 실제 비용 외에 기회비용까지 포괄하는 사용자비용 접근법으로, 호주 및 뉴질랜드에서는 신규주택 구입관련 제 비용을 반영하는 순취득가격 접근법으로, 아일랜드에서는 주택소유로 인해 현금지출된 비용을 반영하는 지급액 접근법으로 추정하고 있다.

본 연구에서는 집세조사의 현실반영도를 높이기 위하여 다음과 같이 수행하였다. 자가주거비 및 전·월세의 집세지수에 관한 기본개념과 소비자물가지수에 있어서 집세지수 기본개념 및 작성현황과 집세조사 표본현황을 살펴보고, 외국의 집세조사 사례를 살펴보았다. 집세조사의 표본설계개요, 모집단, 표본추출방법, 표본의 크기, 추정방법 등에 대해서 살펴보고, 표본오차 분석, 표본부문의 문제점, 가격반영 방법에 있어 무엇이 문제인지를 알아보고, 개선방법으로 새로운 표본규모와 집세가격 추정방법인 비추정, 회귀추정법에 대하여 설명하였다. 새로운 표본규모를 제시하고 집세조사 추정방법간의 실증분석을 통하여 어떤 방법을 사용하는 것이 좋은 것인지를 제안하고 마지막으로 결론을 맺고자 한다.

2) 국민은행에서는 1986년부터 주택매매가격지수와 주택전세가격 지수를 발표하고 있다. 조사 대상주택은 주택유형과 주택규모로 구분, 조사되며, 지수발표도 지역별, 주택유형별, 주택규모별로 집계, 발표되고 있다.

II. 집세조사의 작성 현황

1. 조사 개요

집세조사는 소비자물가조사에서 1955년 기준 지수 개편때 분류를 4개분류에서 5개분류로 나누면서 주거비가 포함되어 시작되었다. 당시에는 집세조사를 연 4회 조사(1, 4, 7, 10월)하다가 1960년 기준 개편때 부터 집세는 매월 조사가 실시되었다. 1965년 개편때에는 분기조사로 환원 되었으며, 주거비의 비중이 <표1>과 같이 높아지면서 1980년 지수 개편때에 월 조사로 전환하게 되어 현재에 이르고 있다.

<표 1> 주거비의 가중치 비중

년도	1955	1960	1965	1970	1975	1980	1985	1990	1995	2000
전국가중치	100	1000	1000	1000	1000	1000	1000	1000	1000	1000
주거비	13.05	84.9	75.2	97.4	110.1	149.0	129.4	141.7	148.3	156.4

자료 : 통계청, 2000년 기준 소비자물가지수 개편 보고서

집세의 조사품목과 규격은 주거비의 하위 품목으로 전세와 월세 2개 품목으로 구성되어 있고, 월세는 보증부 월세, 월세, 사글세로 구분한다. 집세 규격은 단독주택, 아파트, 연립주택, 기타주택으로 분류하였다.

집세조사는 소비자물가조사가 되고 있는 전국의 특·광역시를 포함한 36개 주요도시의 표본지역내 거주하고 있는 전·월세 가구를 대상으로 전·월세 금액을 매월 조사³⁾하고 있다. 그리고 매월 500명의 조사 직원이 조사하기 위해 약 10,000여 전·월세 가구(2002년 표본개편 당시 11,000가구)를 직접 방문한다.

집세(전세와 월세)는 매월 5일부터 7일까지 3일간 소비자물가조사와 별도로 월 1회 조사되고 있다.

3) 통계청 집세조사는 소비자물가조사 직원이 조사하는 것이 아니라, 가계조사를 담당하는 직원이 조사, 1,329개 표본조사구를 대상으로 전월세 대상가구를 직접 방문하여 조사하며, 자기집에 거주하는 자가(自家)가구 및 무상으로 거주하는 가구, 농가 및 어가, 외국인 가구(단, 외국인이지만 귀화한 경우는 적격), 음식점, 여관, 하숙업 등의 사업장과 주택을 같이 임차한 가구, 주택이외의 거처에 거주하는 가구, 임차가구가 다시 세를 놓은 임대가구는 조사에서 제외한다.

2. 주거비의 개념⁴⁾

집세지수는 소비자물가지수의 대분류에 주거비를 포함되어 있다. 주거비란 주택을 점유하고 있는 자가 주택이 주는 서비스의 대가로 지불하는 주거관련 비용의 총합이라고 할 수 있다.

(홍형욱·유병선, 2003)는 앵겔과 슈바베에 의한 전통적 주거비가 생존의 수단으로서의 주택에 필요한 비용으로 임대료와 광열비만을 의미하였다. (윤정숙·이경희·홍형욱, 1997)은 주거생활에 직접적으로 필요한 비용으로는 지대, 임대료, 설비비, 가구집기비, 수선비, 상하수도료, 오물수거비 등을 제시하고, 이외에 전기·가스·석유·연탄등의 연료를 위한 광열비와 통근이나 통학·외출 등을 위한 교통비, 가옥세, 토지세등도 주거비와 관련이 있다고 보고 있다. 한편 거주에 필요한 비용(dwelling expenses) - 모기지이자, 재산세, 자가소유자의 보험 및 기타 자가소유에 드는 비용에 추가적으로 광열·수도비, 가구 및 집기비용, 그리고 가구운영비용(household operation expenses)을 포함하여 정의하는 경우도 있다.(Dong-Hoon Oh, 1995).

우리나라의 가계조사에 따르면 주거비는 '주택 및 토지에 관련된 소비지출'로 정의되고 있으며 여기에서 재산증식을 위한 주택 및 토지등의 구입비용은 제외하였다. 통계청 분류에 따른 주거비에는 월세(실제로 지불한 주택 또는 주거용 토지의 임대료 및 권리금), 주택설비 및 수선비(수선재료비, 설비수리서비스비용) 및 기타주거비(공동주택관리비, 화재보험료, 승강기유지비)가 포함되어 있다. 이외에 주거생활과 관련된 비용이라 할 수 있는 전기·가스·난방비 등이 광열비, 상하수도료, 가구집기·가사용품비, 주택관련세금, 전세평가액, 보증부월세평가액, 자가평가액은 별도로 계상되고 있다(통계청, 2002). 이때 자가가구의 주거비용은 귀속임대료(imputed rents) 방식으로 물가지수에 포함한 지수인 자가주거비용포함 총지수를 별도로 작성하고 있으며, 임차형태가 월세인 경우에는 매월 지불하는 월세액을, 전세의 경우에는 전세보증금의 기회비용을 지수로 산출한다.

4) 강제춘(2006) 소비자물가통계의 집세지수에 대한 연구, 한남대학교 행정대학원 석사학위논문

3. 현행 주거비 체계

소비자물가지수 분류에서 현행 주거비는 대분류 기준으로 볼 때 10개 항목 중 하나이며, 소분류로 볼 때 516개 소비 품목 중 15개 품목에 해당한다. 주거비가 전체 소비지출에 차지하는 가중치는 현재 156.4이다. 그 중에서 전세 가중치는 93.5, 월세 가중치는 37.9이다.

<표 2> 2000년 기준 소비자물가지수 편제 내 주거비 비중

구분	품목수	가중치
	516	1000.0
총지수		
식료품	180	271.2
주거비	15	156.4
광열 · 수도	8	58.0
가구집기 · 가사용품	57	37.1
피복 및 신발	43	56.5
보건의료	42	43.9
교육	33	114.6
교양오락	58	53.6
교통 · 통신	44	159.3
기타잡비	36	49.4

자료 : 통계청(2005a), 물가연보

주거비 내에 세부 항목은 < 표3 >과 같이 집세, 주택설비수리비, 기타 주거의 세가지 항목으로 세분되어 있으며, 각각의 가중치는 1000포인트 내에서 131.4%, 16.4%, 8.6%를 차지하고 있다. 주거비 내에 전·월세에 해당되는 집세의 가중치는 물가지수를 구성하는 단일 소비 품목 중 가장 높다. 이러한 원인 때문에 여러국가에서는 ‘집세’를 소비자물가지수에 반영하는 문제에 대해 많은 논의가 있어 왔다.5)

집세는 주거서비스의 대가인데 주택점유형태별로 차이가 있다. 월세의 경

5) 우리나라에서는 자가주거비를 보조지표로만 활용하고 주거서비스에 대한 가치는 반영하고 있지 않다. 국내의 경우 주택점유형태(housing tenure type)는 크게 자가, 전세, 월세(보증부 월세 포함)의 세가지 유형으로 구분된다. 이 중 전세는 국외에는 존재하지 않는 우리나라만의 독특한 주택점유 형태이다. 월세가 보편화되어 있지 않는 국내 여건상 이러한 전세제도는 월세 전환, 주거서비스 가치 산출, 적정 전세가 비율, 경기 선 · 후행 논의 등에서 중요하게 다루어지고 있다.

우는 실질적으로 지급한 지출액을 기준으로 하면 되지만, 전세의 경우는 대부분 2년이라는 임대차 계약기간에 근거하여 목돈 형태의 주거서비스 가격을 선지출하는 개념이다. 엄밀하게는 전세 계약 시점의 지출액을 근거로 전세 집세를 측정하여야 하나, 사실상 이를 반영하기 어렵다. 따라서 2년 거주기간 동안 받는 서비스를 월별 가치로 환산할 필요가 있다. 이때 적용하는 개념이 '월세 평가액이다'이다. 전세의 월세평가액이란 전세 주택에 대하여 전세금 또는 보증금을 지불하는 가구가 실제 전세 주택과 같은 시설과 규모를 월세로 빌린다고 가정할 때 지불해야 하는 금액으로 간주하는, 이른바 수정된 지출개념을 적용하고 있다.

주거비 구성은 크게 집세, 주택설비수리비, 기타 주거의 세부분으로 되어 있으며, 2000년 기준으로 각각 131.4, 16.4, 8.6 의 가중치가 부여되어 있다. 전월세 집세의 경우 지출기준에 근거하여 주거서비스 가치를 주기적으로 관찰할 수 있으나, 자가에 대한 주거서비스 가치는 현실적으로 관찰하기 어렵다. 그것은 자가에 대한 주거서비스 가치가 전월세 집세와는 달리 투자 목적으로 지출된 비용도 포함하고 있기 때문이다. 이러한 이유로 현행 주거비 체계 내 '자가주택'에 대한 주거서비스에 대한 가치는 반영되어 있지 않다. 다만 보조지표로만 활용되고 있다.

<표 3> 주거비 품목

분 류	품목명(수)	가중치 (156.4)
1)집세		131.4
	전 세(1)	93.5
	월 세(1)	37.9
2)주택설비수리비		16.4
	설비수리재료(6)	9.9
	주택설비수리서비스(5)	6.5
3)기타 주거		8.6
	공동주택관리비(1) (일반관리비만포함)	8.0
	이삿짐운송료(1)	0.6

자료 : 통계청(2002), 소비자물가지수 개편 보고서

4. 표본선정

현행 조사는 2002년에 2000년 인구주택총조사 표본자료 24,998조사구를 대상으로 전국 36개 도시의 거주용 주택을 모집단으로 하였다.

경제활동인구 표본조사구내의 모든 집세 가구의 표본들은 경제활동인구 표본조사구 1,628 조사구의 32,580가구에서 표본규모를 도시지역 1,158조사구, 10,049가구(비조사구역 27조사구 포함)로 선정하였다. 그 외에 특수지역인 의정부, 안양 등 13개 중소도시의 경우 가구표본이 부족하여 경제활동인구 비표본 지역 포함하였고, 2005년 1월부터 고정표본방식에서 연동표본 방식으로 변경(매월 표본가구 중 1/36씩 교체)한다. 임차 기간(24개월)이 지나면 모든 전월세 가구가 계약을 갱신한다고 가정하면 지역별 최소 48개 표본 필요(전세 24가구, 월세 24가구)하다.

<표 4> 집세조사의 표본규모

구 분	7대광역시	29개 도시	총조사구수	집세대상가구수
조사구	781	350	1,131	10,947
가구수	15,620	7,000	22,620	

※ 보령시, 남원의 경우 표본가구수가 30~33가구에 불과
 자료 : 통계청, 2005년 12월 집세조사 결과

5. 집세지수 작성방법

집세지수는 기준시기를 고정한 가중산술평균법으로 라스파이레스 산식 (Laspeyres' Formula)으로 지수를 산출⁶⁾한다.

6) ① 도시별 가구당 평균 전(월)세액 계산

$$\text{가구당 평균 전(월)세액} = \frac{\text{연결가구 전(월)세액의 합}}{\text{연결가구수}}$$

② 도시별 전(월)세 지수 계산

$$\text{전(월)세지수} = \frac{\text{금월 가구당 평균 전(월)세액}}{\text{전월 가구당 평균 전(월)세액}} \times \text{전월 전(월)세지수}$$

③ 전도시 전(월)세 지수 계산

- 도시별 지수를 이용하여 전도시 전(월)세 지수 계산

$$\text{전도시 전(월)세지수} = \frac{\sum(\text{도시별 가중치} \times \text{도시별 지수})}{1,000}$$

계산과정은 도시별 가구당 평균 전세·월세액을 계산한 다음 도시별 전세·월세액을 계산하고, 마지막으로 전도시 전세·월세액에 도시별 가중치를 적용하여 전국평균 집세지수를 산출한다.

전세가격은 매월 5일(조사대상기간은 전월 6일 ~ 금월 5일)을 기준으로 계속가구 및 전출·전입이 동시에 이루어진 가구의 전세가격 변동을 반영⁷⁾(취득접근법)하고, 월세가격은 보증부월세의 경우 인근 주택의 월세에 근거하여 보증금을 월세액으로 환산한 가격을 반영⁸⁾(이자율 : 월 1% 적용, 지역에 따라 차이가 있음)하였다.

2000년 인구주택총조사의 결과자료를 기초로 하여 표본조사 항목중 점유형태별 도시가구의 가구수와 임차료자료를 이용하여 가중치 산출하였다. 집세가중치는 소비자물가지수 전체 가중치 1,000중 127.5이며 전세는 92.5, 월세는 35.0이다.

<표 5> 점유형태별 비중(전국기준) (2000년)

전 국		일반가구수	비중
계		14,311,807	100.0
자 가		7,753,100	54.2
전 세		4,039,514	28.2
월세	보증부 월세	1,524,279	10.7
	월세	279,141	2.0
	사글세	309,823	2.2
무 상		405,506	2.8
미 상		444	0.0

집세지수는 매월 가구변동이 다음과 같이 발생한다. 전출과 전입이 동시에 이루어진 경우와 집세종류가 변경된 경우가 있다. 즉 집세종류가 전월 전세에서 금월 월세로 변경되는 가구는 금월 월세를 민간 이자율을 적용하여 전세금액으로 환산하고, 전월 월세에서 금월 전세로 변경되는 가구는 금월 전세를 민간 이자율을 적용하여 월세금액으로 환산한다. 또한 6개월 이내에

7) 사용접근법이란 임차가구에 의해 소비된 주거서비스 비용의 변동을 측정하는 방법으로 사용자 비용(이자 등) 접근법, 집세접근법이 있다.

8) 취득접근법이란 다른 내구재와 같이 취득 시 발생한 화폐거래(전세의 경우 전세가격)를 근거로 주거서비스 비용의 변동을 측정하는 방법이다.

전입가구가 이사온 경우에는 전입 발생월에 전입과 전출을 동시에 처리하여 가격 반영하고, 전출 후 6개월이 지나도록 전입이 없는 경우 폐지 처리하고, 그 이후에 전입 온 경우 신규로 처리한다.

6. 외국의 집세조사 사례

가. 일본의 집세조사방법⁹⁾

1) 조사대상

우리나라와는 달리 전세제도가 없고 월세 제도가 정착된 국가로, 주택점 유형태는 자가와 월세 두 가지 유형을 조사한다. 월세의 경우는 공급 주체 (정부의 자금 지원 여부와 관련)에 따라 공영주택, 공단 주택 등으로 구분하고 있으며 임대료 가중치는 공급 주체별 가중치를 감안하여 산정하였다.

임대료는 전국 167개 시정촌에서 선정한 약 23,000세대로부터, 집세의 월액 및 연면적을 조사하고, 각 시정촌의 최근 국세조사조사구에서 일정수의 조사구를 무작위 추출해서 조사구내의 모든 세대를 조사한다.

2) 조사방법

민영임대주택은 주택의 면적 및 구조에 따라 목조소주택, 목조중주택, 비목조주택으로 구분하여 3.3㎡당 임대료를 조사하고, 공영임대주택은 구분없이 3.3㎡당 임대료를 조사한다.

3) 가중치 산정

농림어가 및 단독가구를 제외한 전국 167개 시정촌에서 매월 약 8000 가구를 대상으로 한 가계조사결과 소비지출액에서 임대료가 차지하는 비중을 1000분비로 산출하는 것이 우리나라와 같다.

9) 일본 총무성 홈페이지의 소매물가조사개요에서 발췌한 내용이며, 자세한 사항은 참고문헌을 참조 (일본 총무성, 2006)

<표 6> 일본의 집세조사 현황

구분	민영가임				공영가임	귀속가임		
	목 조 주택	목 조 중주택	비목조주택	계		목 조 소주택	목 조 중주택	비목조 주택
가중치 (1만분비)	30	81	189	316	48	1	1107	252

* 귀속가임 가중치 : 『1998년 전국소비실태』 결과 귀속가임에 민영가임 상승률과 가계조사소비지출과의 수준차 조정, 중복계산된 지대 등을 뺀 후 주택소유율을 곱해서 귀속가임 산출 후 귀속가임 가중치 산출

4) 조사주기

가임 조사지구는 1군 : 1, 4, 7, 10월 2군 : 2, 5, 8, 11월, 3군 : 3, 6, 9, 12월 3군으로 나누어서 각 군을 3개월 마다 조사한다.

나. 미국의 집세조사방법¹⁰⁾

1) 표본 및 조사방법

1차 주택 표본은 층화 집락 표본으로 최근의 10년 주기의 총조사(census) 이전에 지어진 주택단위를 대표하며, 총조사 이후에 지어진 주택단위는 신축조사(New Construction Survey)를 하여 설계되어 진다. 인구주택총조사(Census of Population and Housing)의 자료를 사용하여 CPI 지역들을 구획(segment, 지리적인 인접지역)으로 분리한 다음 구획들을 각 CPI 지역 안에서 층화 표본으로 추출한다.

집세는 변동성이 심하지 않기 때문에 CPI내의 다른 소비 품목을 조사하는 주기보다 가격조사 주기를 길게하여 조사한다. 주택표본은 약 36,000 임대 가구를 대상으로 임대료를 조사하며, 소구획(panel)이라 불리는 6개의 하위표본으로 나뉘어 지며, 각 소구획은 연속적인 순서로 가격을 조사한다. 각 소구획은 1년에 두 번 가격으로 조사한다. 예를 들면, 소구획 1은 1월과 7월에 가격을 조사하고, 소구획 2는 2월과 8월에 조사하면서 소구획 6까지 조사한다. 각 층(strata) 안에 있는 구획(segment)들은 소구획들에 할당되는데 각각

10) 미국 노동 통계국 홈페이지 미국의 집세 및 집세상당액에 대한 소비자물가지에서 발췌한 내용이며, 자세한 사항은 참고문헌을 참조(미국 노동통계국, 2006)

의 소구획은 CPI 지역에 대표적인 하위표본이 되도록 한다. 각각의 소구획이 전체 표본을 대표하고, 매월 조사가 되고 있기 때문에 한 소구획의 자료에 의해 작성되는 집세 및 집세상당액 지수는 매월 공표하기에 충분하다.

2) 가중치 및 지수반영의 특징

표본에 있는 각각의 타가 단위(renter unit)는 타가 모집단에 있는 타가 단위들을 대표한다. 구획안에 있는 주택의 타가 및 자가비용들이 그 구획을 위한 가격변동계산에 의해 사용되는 타가 및 자가가중치(renter weights)의 기초가 된다

집세지수는 임차인이 지불하고, 집주인이 수령한 집세의 변동을 측정하며, 노동통계국은 집세지수의 변동을 추정하기 위하여 집세와 타가가중치를 사용한다. 집세상당액지수는 자가에 의해 공급된 주거서비스의 가격 변동을 측정한다. 노동통계국은 집세상당액지수의 변동을 추정하기 위하여 집세 및 자가가중치(owner- weights)를 사용한다. 일반적으로 금월 및 6개월전월에 대하여 지역지수 내에 있는 주택단위들에 가중치를 부여하여 집세의 가격변동을 계산한다.

집세지수가 계산될 때에는 집세 및 타가 가중치가 사용된다. 즉, 각각의 CPI 지역 안에서, 금월 6개월 집세 변동 = 금월 타가가중된 집세 총액을 6개월 전 타가가중된 집세 총액으로 나누면 된다.

3) 지수 추정

금월 집세지수는 추정된 1개월 변동에 의해 곱해진 1개월 전 지수와 같다. 지수추정체계는 6개월 가격변동보다는 1개월 가격변동을 필요로 하므로, 6개월 변동의 6제곱근이 각각의 CPI 지역에서 사용된다.

$$\text{- 금월 집세지수} = \text{1개월전 집세지수} \times (\text{금월 6개월 집세변동})^{1/6}$$

Ⅲ. 개선방안

1. 집세지수의 개선점

집세조사 규격이 주택유형별(단독, 아파트, 연립, 기타)로 세분화 되어 있지 않아 유형별 집세의 동향을 파악하기 어렵다. 2000년기준 전세가중치는 92.5로 소비자물가지수에서 단일 품목으로는 가장 높으나, 조사규격으로 구분되어 있지 않고 1개 규격으로 통합되어 있다.

전세와 월세의 보증금에 대한 가격 반영 방법 차이가 있는데, 전세는 전세금액의 변동분을 적용(취득접근법)하는데 비해 보증금 월세의 경우 보증금을 월세로 환산하여 그 변동분을 가격에 반영(사용접근법)한다. 전세는 내구재의 구입이 아니라 주거서비스의 구입이므로 취득접근법보다는 사용접근법이 합리적이라고 생각한다.

집세 표본조사구가 집세조사 목적과 부합되어 있지 않다. 경제활동인구조사, 가계조사 등 다목적으로 설계된 표본으로 정확한 집세 가격변동 파악 어려움이 있는데, 2005년 1월 이후 표본선정 방법의 변경(고정표본 → 연동표본)은 집세 표본 오차를 더 크게 발생시킬 가능성이 있다. 그리고 현 집세 표본은 모집단의 집세변동을 제대로 반영하지 못하고 있다. 그 외 일부 지역 표본문제에 대해서는 (강재춘 2006)을 참고하면 된다.

집세조사의 대표성을 제고하고 정확한 집세 가격 변동을 측정하기 위하여 조사규격, 가격반영 방법, 표본선정 방법 등의 개선방안에 대해 (강재춘 2006)이 제안한 내용에 자료를 추가하여 검토해 보았다. 이에 따라 본 연구에서는 가격산출에 있어서 통계청에서는 가중단순평균방법을 사용하고 있어, 그 개선방안으로 확률표본 추정법에 대해 간략히 알아보고, 실질적으로 평균 전·월세금액을 추정하여 집세지수를 산출하고, 적절한 표본크기도 비추정법으로 추정하였다.

확률표본 중에서 비추정법으로 전·월세금액을 추정하기 위하여 표본을 주택유형별로 층화하였고, 자료는 2005년 인구주택총조사와 2005년1월부터 2006년9월까지의 가계조사를 이용하였다.

2. 통계적 추정방법 고찰

단순임의추출이나 층화임의추출 등 표본추출방법에 의해 얻어진 표본의 단위들로부터 관측된 관측값들인 y_1, y_2, \dots, y_n 을 기초로 하여 모집단 평균이나 모집단 합계를 추정하는 방법을 활용하였다. 때로는 반응변수인 y 값과 매우 밀접한 관계를 갖는 다른 변수들을 생각할 수 있는 경우가 있다. 이 경우 y 값 이외의 다른 변수들까지 관측하게 된다면 모집단 평균을 추정하는데 유용한 정보를 더 많이 얻게 된다. 즉 반응변수 y 의 모집단 평균을 추정하기 위해 보조변수들을 사용하는 방법이다. 보조변수 x 를 사용하는 추정방법으로 비추정(ratio estimation), 회귀추정법(regression estimation), 차이추정법(difference estimation)들이 있다. 3가지 방법 모두 각각의 표본단위들로부터 관심(반응)변수 y 와 보조변수 x 에 관한 측정을 요구하는 방법이다. 본 연구에서는 비추정법에 대해서만 다루어 보고자 한다.

가. 단순임의추출의 비추정법

비추정법은 두 변수의 비(比) R 을 추정할 뿐만 아니라, 모집단 합계 τ_y ¹¹⁾ 또는 모집단 평균 μ_y ¹²⁾를 추정한다. N 개의 모집단에서 n 개의 표본을 추출하고, 관심변수 y 와 보조변수 x 를 이용하여 비 R 은 다음과 같이 추정한다.

$$11) \text{ 모집단 총계 } \tau_y \text{의 비추정량 : } \hat{\tau}_y = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} (\tau_x) = r \tau_x ,$$

$$\hat{\tau}_y \text{의 추정분산 } \hat{V}(\hat{\tau}_y) = (\tau_x)^2 \hat{V}(r) = (\tau_x)^2 \left(\frac{N-n}{nN} \right) \left(\frac{1}{\mu_x^2} \right) s_r^2$$

$$12) \text{ 모집단 평균 } \hat{\mu}_y \text{의 비추정량 : } \hat{\mu}_y = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} (\mu_x) = r \mu_x ,$$

$$\hat{\mu}_y \text{의 추정분산 } \hat{V}(\hat{\mu}_y) = \mu_x^2 \hat{V}(r) = \left(\frac{N-n}{nN} \right) s_r^2$$

$$\text{모집단 비 } R \text{의 추정량} : r = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} = \frac{\bar{y}}{\bar{x}} \quad (1)$$

$$r \text{의 추정분산} : \hat{V}(r) = \hat{V} \left(\frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} \right) = \left(\frac{N-n}{nN} \right) \left(\frac{1}{\mu_x^2} \right) s_r^2 \quad (2)$$

$$\text{여기서 } s_R^2 = \frac{\sum_{i=1}^n (y_i - r x_i)^2}{n-1} \quad (3)$$

만약 보조변수 x 의 모집단 평균 μ_x 을 알지 못하는 경우에는 식(2)와 식(3)에서 μ_x^2 대신에 \bar{x}^2 을 사용하여 근사적으로 계산한다.

대표본에서의 정규분포에 기초한 신뢰구간의 추정은 비추정의 경우에도 적용할 수 있다. 예를 들어 비 R 에 대한 점근적인 90% 신뢰구간을 구해보면 다음과 같다.

$$R \pm 1.645 \sqrt{\hat{V}(R)}$$

R 의 분산과 추정량을 x 와 y 의 상관계수 ρ^{13} 를 포함한 R 의 추정분산은 다음과 같다.

$$\hat{V}(R) = \frac{1-f}{n} \left(\frac{1}{\bar{x}^2} \right) (s_y^2 + R^2 s_x^2 - 2R \hat{\rho} s_x s_y), \text{ 여기서 } f = n/N \text{이다.}$$

나. 총화임의추출에 의한 비추정법

모집단에서 모든 층에 대해 분산의 근사가 가능할 수 있도록 충분히 큰 표

$$13) \hat{\rho} = \frac{s_{xy}}{s_x s_y}$$

여기서,

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}), \quad s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

본을 취하여 x 와 y 값을 얻을 수 있다고 가정하자. 층화임의추출에 비추정량을 구하는 방법으로 두가지가 있다. 하나는 모든 층에서 각각 μ_y 와 μ_x 의 비를 추정하는 다음 각각의 추정값들을 가중평균하여 모집단의 비를 추정하는 방법인데 이것을 분리비추정량(separate ratio estimator)이라 한다. 또 다른 방법은 \bar{y}_{st} 와 \bar{x}_{st} 를 μ_y 와 μ_x 의 추정값으로 사용하는 방법으로 μ_y/μ_x 의 추정량은 $\bar{y}_{st}/\bar{x}_{st}$ 가 되는데 이것을 결합비추정량(combined ratio estimator)이라 한다. 확률변수들이 서로 독립일 경우에 그 확률변수들의 합의 분산은 각 변수의 분산의 합이다. 이에 따라 분리비추정량이든 결합비추정량이든 그 분산은 모집단 평균 비추정량의 분산식들의 합의 형식으로 표현할 수 있다.

예를 들어 2개층을 A와 B라 하자. A층의 비추정량 μ_{yA} 은 $(\bar{y}_A/\bar{x}_A)(\mu_{xA})$ 이며 분산추정량은 $\hat{V}(\hat{\mu}_{yA}) = \mu_{xA}^2 \hat{V}(r_A) = \left(\frac{N_A - n_A}{n_A N_A}\right) s_{rA}^2$ 이다 B층의 비추정량 μ_{yB} 도 $(\bar{y}_B/\bar{x}_B)(\mu_{xB})$ 이며 분산 추정량도 $\hat{V}(\hat{\mu}_{yB}) = \mu_{xB}^2 \hat{V}(r_B) = \left(\frac{N_B - n_B}{n_B N_B}\right) s_{rB}^2$ 와 같이 나타낼 수 있다. y 의 모집단 평균인 μ_y 의 추정량은 구하려면 층 크기에 비례하는 가중값을 구해 가중평균하면 된다.

즉 μ_y 의 비추정량 $\hat{\mu}_{yRS}$ 는 $\hat{\mu}_{yRS} = \left(\frac{N_A}{N}\right)\left(\frac{\bar{y}_A}{\bar{x}_A}\right)\mu_{xA} + \left(\frac{N_B}{N}\right)\left(\frac{\bar{y}_B}{\bar{x}_B}\right)\mu_{xB}$ 이다.

분산의 추정량은

$$\hat{V}(\hat{\mu}_{yRS}) = \left(\frac{N_A}{N}\right)^2 \left(\frac{N_A - n_A}{n_A N_A}\right) \frac{\sum_{i=1}^{n_A} (y_i - r_A x_i)^2}{n_A - 1} + \left(\frac{N_B}{N}\right)^2 \left(\frac{N_B - n_B}{n_B N_B}\right) \frac{\sum_{i=1}^{n_B} (y_i - r_B x_i)^2}{n_B - 1} \text{이다.}$$

또한 μ_y 의 추정량이 \bar{y}_{st} , μ_x 의 추정량이 \bar{x}_{st} 일 때 결합비추정량은 $\hat{\mu}_{yRC} = \frac{\bar{y}_{st}}{\bar{x}_{st}}(\mu_x)$ 이다. 만일 r_c 를 $(\bar{y}_{st}/\bar{x}_{st})$ 로 나타낸다면 $\hat{\mu}_{yRC}$ 의 분산추정량은

$$\hat{V}(\hat{\mu}_{yRC}) = \left(\frac{N_A}{N}\right)^2 \left(\frac{N_A - n_A}{n_A N_A}\right) s_{rA}^2 + \left(\frac{N_B}{N}\right)^2 \left(\frac{N_B - n_B}{n_B N_B}\right) s_{rB}^2 \text{이다. 여기서 } s_{rA}^2 \text{은 A층의 } (y_i - r_A x_i) \text{값들의 분산이고, } s_{rB}^2 \text{은 B층의 } (y_i - r_B x_i) \text{값들의 분산이다.}$$

일반적으로 결합비추정량의 추정분산이 분리비추정량의 것보다 크므로 분리비추정량을 사용하게 된다. 그러나 각 층별로 비추정량은 편향추정량이기 때문에 분리비추정량은 결합비추정량보다 편향을 크게 할 수 있다. 하지만 각 층별로 표본크기가 충분히 크다면¹⁴⁾(일반적으로 약 20개 이상)가 층별로 비추정량의 편향이 별 문제가 되지 않는 경우에는 분리비추정량을 사용하는 것이 효과적인 반면 각 층별 표본크기가 적거나 아니면 모든 층의 비가 근사적으로 같은 경우에는 결합비추정량을 사용하는 것이 더 낫다.

다. 표본크기 결정

자료의 변동과 표본크기 n 에 따라 표본에 포함된 정보의 양이 결정된다. 일단 표본추출법이 결정되고 나면 다음으로는 표본크기를 얼마로 할 것인가를 정하여야 한다. 단순임의추출법에서 비추정법을 사용하여 모수 R, μ_y, τ_y 에 대해 추정하여고 할 때 추정오차의 한계가 B 이내가 되게 하는 표본크기를 결정하는 방법이다. 모집단의 비 R 을 추정오차의 한계 B 이내로 추정하기 위한 표본크기 n 을 결정하려면 추정오차의 한계 B 를 비추정량 r 의 표준오차에 2배한 식과 같게 하여 n 에 관해 풀면 된다.

$$2 \sqrt{V(r)} = B \quad (6)$$

$$\text{분산의 추정량 : } \hat{V}(r) = \left(\frac{N-n}{nN} \right) \left(\frac{1}{\mu_x^2} \right) \frac{\sum_{i=1}^n (y_i - r x_i)^2}{n-1} \quad (7)$$

(7)식을 (8)과 같이 변형하고,

$$\hat{V}(r) = \left(\frac{N-n}{nN} \right) \left(\frac{1}{\mu_x^2} \right) s_r^2, \quad s_r^2 = \frac{\sum_{i=1}^n (y_i - r x_i)^2}{n-1} \quad (8)$$

(8)식에서 s_r^2 대신에 σ^2 을 대입하면 모분산 $V(r)$ 의 근사식으로 볼 수 있다. 따라서 오차의 한계 B 이내에서 R 을 추정하기 위한 표본크기 n 은 다음과

14) 김영원 외 3인 공역, 표본조사의 이해와 활용 참고

같이 구할 수 있다.

$$2\sqrt{\widehat{V}(r)} = 2\sqrt{\left(\frac{N-n}{nN}\right)\left(\frac{1}{\mu_x^2}\right)\sigma^2} = B \quad (9)$$

오차 한계가 B일 때, R을 추정하기 위한 표본크기는

$$n = \frac{N\sigma^2}{ND + \sigma^2}, \quad \text{여기서 } D = \frac{B^2\mu_x^2}{4} \quad (10)$$

결정한다.

실질적으로 σ^2 값을 알지 못하는 경우에 표본크기를 결정해야 하는 일이 있을 수 있다. 그런데 σ^2 의 추정량으로서 s_r^2 을 계산하여 사용하려고 해도 자료가 없는 경우에는 먼저 n' 개의 예비표본을 뽑아서 σ^2 의 추정값을 계산한다. 또한 μ_x 값을 모를 때에도 역시 예비표본의 자료에서 \bar{x} 값을 계산하여 대신 사용한다.

위와 비슷한 방법으로 추정오차의 한계 B 이내로 모집단 평균 μ_y 를 추정하기 위한 표본크기를 정하고자 할 때 표본크기 n은 다음 식을 n에 대해 풀어서 구할 수 있다.

$$2\sqrt{V(\hat{\mu}_y)} = B \quad (11)$$

식(11)을 다음과 같이 변형할 수 있다.

$$2\mu_x\sqrt{V(r)} = B$$

즉 오차의 한계가 B일 때 μ_y 를 추정하기 위한 표본의 크기는

$$n = \frac{N\sigma^2}{ND + \sigma^2}, \quad \text{여기서 } D = \frac{B^2}{4} \quad (12)$$

식(12)를 이용하여 표본크기 n을 결정할 때 μ_x 의 값을 몰라도 상관없지만 σ^2 에 대한 추정값은 알아야 한다. 활용 가능한 사전정보가 있는 경우에는 그 것을 사용하면 되고 사전정보가 없는 경우에는 예비조사 자료를 이용하여 추정하면 된다.

모집단 총계 τ_y 를 추정오차의 한계 B 이내로 추정하는데 필요한 표본크기

는 다음 식에서 n에 대해 풀면된다.

$$2 \sqrt{V(\hat{\tau}_y)} = B \quad (13)$$

식(11)을 다음과 같이 변형할 수 있다.

$$2 \tau_x \sqrt{V(r)} = B$$

즉 오차의 한계가 B일 때 τ_y 를 추정하기 위한 표본의 크기는

$$n = \frac{N\sigma^2}{ND + \sigma^2}, \quad \text{여기서 } D = \frac{B^2}{4N^2} \quad (14)$$

IV. 실증분석 및 결과

1. 모집단 변화

우리나라의 주택상황은 시간이 지남에 따라 변동이 큰 편이다. 2000년도에는 단독주택의 비율이 36.9%를 차지하였는데 매년 새로운 아파트 단지가 개발되어 2005년에는 31.6%로 축소된 반면 아파트는 2000년 47.4%에서 2005년에는 과반수가 넘는 51.6%를 차지하고 있다.

<표 7> 전국 / 거처의 종류별 주택수 및 구성비

전국 주택수	2000		2005	
	호수	구성비	호수	구성비
단독주택	4,269,180	36.92	4,263,541	31.58
아파트	5,479,828	47.39	6,962,689	51.58
연립주택	849,687	7.35	558,513	4.14
다세대주택	472,236	4.08	1,299,208	9.62
비거주용건물	401,470	3.47	208,690	1.55
주택이외거처	91,814	0.79	206,511	1.53

자료 : 2000년, 2005년 인구주택총조사

반면에 거주하는 가구수는 여전히 단독주택에 가장 높은 비중을 차지하고 있으나 점차 줄어들고 있다. 거처하는 가구수의 변동, 즉 모집단의 변동을 적절하게 반영하기 위해서는 표본설계를 정기적으로 보정하는 방법도 집

세조사의 신뢰성을 높일 수 있는 한 방법일 것이다.

그러나 표본설계 변경과 같은 개선 방안은 인력과 예산이 뒷받침 되어야 해결할 수 있는 것으로 사료된다. 따라서 현재 사용하고 있는 전·월세 가격 측정방법인 단순평균방법을 가격의 보조변수를 활용한 비추정법을 개선방안으로 제안하고자 한다. 실증분석을 위한 모집단은 2005년 인구주택총조사에서 80평이하에 거주하는 가구수를, 표본은 가구실태조사의 표본가구인 9000여 가구내에 전월세를 사는 가구를 대상으로 하였다.

집세지수를 산출함에 있어 분류를 점유형태별/주택유형별 체계로 다음 <표 8>과 같이 하였다. 즉 세분류별로 지수를 작성한 후 상위분류는 하위분류를 가중평균하는 방법으로 집세지수를 작성하였다.

<표 8> 집세지수 분류

대분류	중분류	소분류(점유형태)	세분류(주택유형)		
집세지수	전세지수	전세지수	단독주택		
			아파트		
			연립·다세대주택		
			기타		
	월세지수	보증금월세	보증금월세	단독주택	
				아파트	
				연립·다세대주택	
				기타	
		무보증금월세	무보증금월세	무보증금월세	단독주택
					아파트
					연립·다세대주택
					기타
사글세	사글세	사글세	단독주택		
			아파트		
			연립·다세대주택		
			기타		

세분류에서 소분류를 작성할 때 사용한 가중치는 모집단의 주택유형별 가구수의 비율을 사용하였다.

<표 9> 모집단의 주택유형별 거주 가구수의 비율

전세가구	단독	아파트	연립·다세대	기타
모집단 ^{주)}	38.79%	44.49%	13.38%	3.35%
표 본	47.14%	40.38%	11.12%	1.36%

주) 2005년 인구주택총조사 결과에서 80평이하 거주 가구수

<표 10> 각 월세지수에 사용된 가중치

		단독	아파트	연립·다세대	기타
보증부월세	모집단	46.00%	40.73%	6.36%	6.90%
	표본	58.47%	34.45%	4.43%	2.65%
무보증월세	모집단	80.32%	6.66%	4.61%	8.18%
	표본	90.55%	4.05%	2.92%	2.49%
사글세	모집단	85.21%	5.86%	4.26%	4.66%
	표본	89.94%	4.22%	1.17%	4.67%

주) 모집단은 2005년 인구주택총조사 결과에서 80평이하에 거주하는 가구수, 표본은 2005년도 연간 도시가계조사 결과

또한 월세지수는 보증금월세, 무보증월세, 사글세의 지수들의 가중평균으로 작성하는데, 이때 사용한 가중치는 2005년 도시가계조사 결과의 월세 점 유형의 주거비 비중으로 하였다.

<표 11> 2005년 도시가계조사 결과의 주거비 비중

2005년	보증금 월세지수	무보증금 월세지수	사글세 지수
주거비 비중(가중치)	(90.54%)	(5.12%)	(4.36%)

최상위 분류인 집세지수는 전세지수와 월세지수를 가중 평균하였다. 여기서 사용된 가중치는 2005년 도시가계조사의 결과에서 전세지출액과 월세지출액의 비율로 하였다.

<표 12> 2005년 도시가계조사 결과의 전·월세지출 비중

2005년	전세지수	월세지수
주거비 비중(가중치) ^{주)}	(69.55%)	(30.45%)
현행 집세지수 가중치의 비중	93.5 (71.2%)	37.9 (28.8%)

주) 전세금과 월세보증금을 월세평가액으로 환산한 금액 비중

2. 통계적 추정방법을 이용한 집세지수 산출

집세지수는 소비자물가지수의 항목으로 가격변동을 나타내는 것이다. 이에 따라 연면적당(평당) 전·월세 단가를 집세지수 산출에 적용하기 위하여 본 연구에서는 거주하는 연면적을 전·월세금액의 보조변수로 활용하였다.

앞에서 살펴본 비추정법을 적용하여 집세지수 추정한 3개의 개선안과 현행 집세조사 결과에서 집계분류와 가중치를 변경한 1개의 개선안을 제안하였다. 1안은 모수를 추정하지 않고 표본조사에 결과를 그대로 이용하여 주택유형별 모집단의 비중을 가중치로 활용하고, 2안은 표본조사 결과에서 모수를 추정하여 주택유형별 모집단의 비중을 이용한다. 3안은 현행 집세조사에서 주택유형별 전월비 증가율을 모집단의 주택유형별 비중을 가중치로 전·월세 증가율을 산출하였다. 4안은 전·월세 모두 월세평가액으로 단일화한 후 총화비추정법으로 지수를 산출하였다.

가. 개선 1안

집세지수는 소비자물가지수의 항목으로 가격변동을 나타내는 것이라 판단된다. 이에 따라 가구가 거주하는 연면적당(평당) 가구가 거주하는 지출비용을 산출하여, 즉 평당 전세·월세단가를 계산한 후 단가의 전월비를 이용하여 다음과 같은 계산과정으로 단가지수를 산출하였다.

첫째, 점유형태별, 주택유형별 평당 단가를 구한다. 둘째는 평당단가의 전월비를 구한다. 셋째는 2005년 인구주택총조사의 결과에서 점유형태별, 주택유형별 가구수의 비중을 곱한다. 넷째는 2005.1월을 100으로 한 단가지수를 구한다. 전세지수 산출과정을 살펴보자.

1) 전세지수 산출

주택유형을 단독주택, 아파트, 연립·다세대=(연립주택+다세대주택), 기타주택으로 구분하였다. 주택유형별로 가구당 주거 연면적의 평당 전세금 단가를 산출하였다.

$$\text{가구의 평당 전세금 단가} = \frac{\sum \text{표본가구의 전세금}}{\sum \text{표본가구의 주거연면적(평)}}$$

2005년 인구주택총조사에서 80평이하에 거주하는 가구수를 모집단으로 하여 주택유형별 거주 가구수의 비율을 가중치로 하여 평당 전세단가를 구한 후 전월비를 계산한다. 전세지수를 다음과 같이 산출한다.

$$\text{전세지수} = \text{평당 전세단가의 전월비} \times \text{전월 지수}$$

<표 13> 전세지수 산출과정

년월	단독주택	아파트	연립 다세대	기타	전세 평균	전월비	전세지수
2005.01	2125.8	2856.1	1872.0	1386.0	2392.0		100.0
2005.02	2101.0	2841.3	1845.6	1476.6	2375.3	0.9930	99.3
2005.03	2137.9	2836.9	1924.6	1454.8	2397.5	1.0093	100.2
2005.04	2100.1	2837.9	1939.3	1412.1	2383.8	0.9943	99.7
2005.05	2138.9	2859.9	1954.4	1394.1	2410.0	1.0110	100.8
2005.06	2139.5	2820.5	1945.4	1379.4	2391.0	0.9921	100.0
:	:	:	:	:	:	:	:
2006.07	2000.9	2978.1	2020.6	1509.7	2421.9	1.0067	101.3
2006.08	1959.3	2948.7	2072.1	1569.0	2401.5	0.9916	100.4
2006.09	1958.6	2955.2	2065.6	1651.3	2406.0	1.0019	100.6

주) 2005.1월을 100.0으로 지수를 산출

2) 월세지수 산출

월세 지수는 보증금 월세, 무보증금 월세, 사글세로 각각 전세지수와 같은 방법으로 산출한다.

<표 14> 보증금월세 지수 산출과정

년월	단독주택	아파트	연립 다세대	기타	보증금 월세 평균	전월비	보증금 월세지수
2005.01	27.2	23.9	25.0	24.5	25.6		100.0
2005.02	27.7	23.6	23.9	24.9	25.6	1.0025	100.3
2005.03	27.0	23.1	23.8	22.4	24.9	0.9709	97.3
2005.04	26.9	23.2	25.3	23.3	25.0	1.0070	98.0
2005.05	26.3	22.6	23.2	24.1	24.5	0.9770	95.8
2005.06	25.5	22.5	26.2	24.3	24.2	0.9910	94.9
:	:	:	:	:	:	:	:
2006.07	25.0	23.4	23.1	26.2	24.3	1.0060	95.2
2006.08	25.0	23.8	23.1	25.2	24.4	1.0038	95.6
2006.09	24.7	23.9	22.6	26.8	24.4	0.9986	95.5

주. 2005.1월을 100.0으로 지수를 산출

<표 15> 무보증 월세지수 산출과정

년월	단독주택	아파트	연립 다세대	기타	월세금액	변동률	월세지수
2005.01	17.1	11.7	15.5	10.7	16.2		100.0
2005.02	18.1	14.2	13.6	10.7	17.0	1.0544	105.4
2005.03	16.7	14.3	13.6	11.8	16.0	0.9403	99.1
2005.04	18.1	16.0	13.6	17.6	17.7	1.1062	109.7
2005.05	19.2	20.9	13.6	12.6	18.5	1.0455	114.7
2005.06	18.0	19.6	14.4	12.5	17.5	0.9464	108.5
:	:	:	:	:	:	:	:
2006.07	17.7	11.1	11.2	14.5	16.7	1.0123	103.4
2006.08	17.3	13.5	10.0	15.0	16.6	0.9912	102.5
2006.09	17.6	15.1	10.2	15.0	16.9	1.0194	104.5

주) 2005.1월을 100.0으로 지수를 산출

<표 16> 사글세 지수 산출과정

년월	단독주택	아파트	연립 다세대	기타	사글세 금액	변동률	사글세 지수
2005.01	15.2	20.4	11.5	19.0	15.5		100.0
2005.02	15.3	16.6	15.1	22.0	15.7	1.0102	101.0
2005.03	15.5	17.4	10.6	19.9	15.6	0.9971	100.7
2005.04	15.9	16.0	11.1	21.2	15.9	1.0197	102.7
2005.05	15.4	16.4	12.0	19.6	15.5	0.9746	100.1
2005.06	15.3	16.4	12.0	16.3	15.2	0.9805	98.2
:	:	:	:	:	:	:	:
2006.07	14.5	14.6	13.0	13.3	14.4	0.9745	92.9
2006.08	14.8	14.6	18.8	15.6	15.0	1.0372	96.4
2006.09	14.4	13.5	13.9	15.6	14.3	0.9592	92.4

주) 2005.1월을 100.0으로 지수를 산출

월세지수와 집세지수는 2005년 연간가계조사 결과를 이용하여 전·월세 지출금액의 비중을 구하고 그 비중을 가중치로 하여 산출하였다.

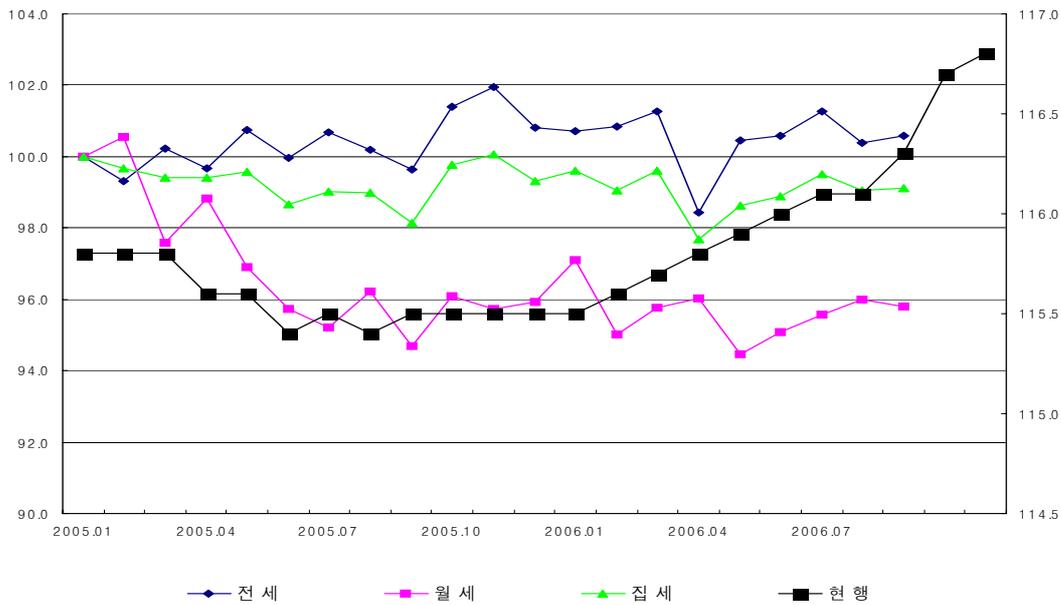
<표 17> 월세지수 산출과정

년도	월	보증금 월세지수	무보증 월세지수	사글세 지수	월세지수
2005년주거비 비중(가중치)		(90.54%)	(5.12%)	(4.36%)	
2005	01	100.0	100.0	100.0	100.0
2005	02	100.3	105.4	101.0	100.5
2005	03	97.3	99.1	100.7	97.6
2005	04	98.0	109.7	102.7	98.8
2005	05	95.8	114.7	100.1	96.9
2005	06	94.9	108.5	98.2	95.7
:	:	:	:	:	:
2006	07	95.2	103.4	92.9	95.6
2006	08	95.6	102.5	96.4	96.0
2006	09	95.5	104.5	92.4	95.8

< 표 18 > 최종 집세지수

년도	월	전세지수	월세지수	집세지수
2005년주거비 비중(가중치)		(69.55%)	(30,45%)	
2005	01	100.0	100.00	100.0
2005	02	99.3	100.55	99.7
2005	03	100.2	97.57	99.4
2005	04	99.7	98.81	99.4
2005	05	100.8	96.91	99.6
2005	06	100.0	95.73	98.7
2005	07	100.7	95.22	99.0
2005	08	100.2	96.22	99.0
2005	09	99.6	94.70	98.1
2005	10	101.4	96.10	99.8
2005	11	102.0	95.72	100.1
2005	12	100.8	95.93	99.3
2006	01	100.7	97.10	99.6
2006	02	100.8	95.00	99.1
2006	03	101.3	95.78	99.6
2006	04	98.4	96.03	97.7
2006	05	100.5	94.46	98.6
2006	06	100.6	95.08	98.9
2006	07	101.3	95.55	99.5
2006	08	100.4	95.99	99.1
2006	09	100.6	95.80	99.1

계산한 결과를 보면 개선1안 지수가 현행지수보다 더 불규칙한 움직임을 보이고 있는 원인은 개선1안의 표본가구수가 작은데도 불구하고 모수를 추정하지 않은 것과 일부 가구의 변동에 민감하게 반응하기 때문일 것으로 판단된다. 이에 따라 표본가구수를 늘리고, 모수를 추정한다면 해결될 수 있을 것으로 생각된다.



<그림 1> 현행지수와 개선1안 비교

나. 개선 2안

주거면적과 주거비(월세평가액)간의 상관관계수가 0.5117로 상관관계가 유의한 것으로 나타나고 있어, 비록 강한 상관은 아니지만 연관이 있는 것으로 판단하여 전월세 금액의 월 평균액을 추정함에 있어 보조변수로 주거면적을 사용하였다. 주택유형별로 총화하여 전·월세 평균금액을 추정하였다.

주택유형별 총연면적(τ_x)과 총가구수(N)은 2005년 인구주택총조사의 결과를 이용하였고, 매월 주거면적(X), 전·월세금액(Y), 표본가구수(n)는 가구실태조사를 이용하였다.

월별 전·월세 평균금액(μ_y)의 추정량은 평당 단가(r)에 가구당 평균면적(μ_x)을 곱한다. 여기서 가구당 평균면적(μ_x)은 총연면적(τ_x)을 총가구수(N)로

나누었다. 평당 단가(r)은 전·월세 금액총액($\sum Y$)을 주거총면적($\sum X$)을 나눈다. 예를 들어 2005년 1월의 아파트 전세금액을 비추정으로 추정해보자. 1월 가구실태조사에 의하면 아파트 전세에 살고 있는 표본가구수(n)는 684가구이고, 평균면적(\bar{X})은 19.08평이고, 평균전세금액(\bar{Y})은 5624.973만원으로 조사되었다. 2005년 인구주택총조사에서 아파트 전세에 살고 있는 총가구수(N)는 1,372,402가구이고, 가구당 평균면적(μ_x)은 20.74평이다. 모집단 평균 전세금액은 가구당 평균면적(μ_x)에 평당금액(r=294.848만원)을 곱하면 6114.97만원(표준오차는 140.828만원)이다. 이와 같은 방법으로 전·월세 금액의 총액을 <표 19>와 같이 추정하여 월별 변화률을 구하고, 변화률로 전·월 지수를 만들었다. 전세와 월세지수에 전세와 월세의 가중치를 이용하여 집세지수를 산출하였다.

<표 19> 전세지수 산출과정

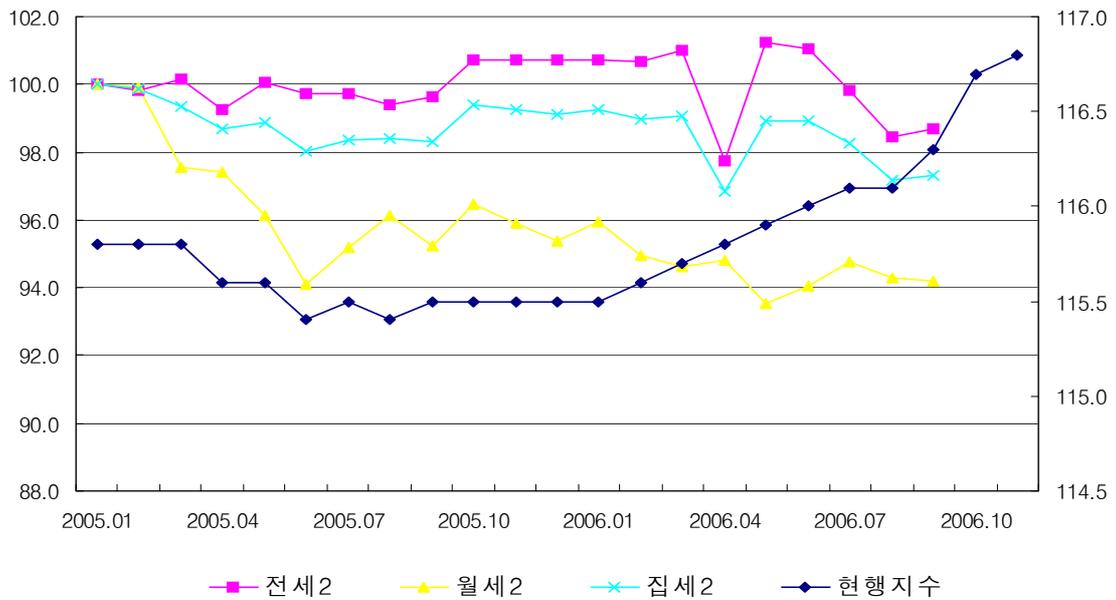
YEAR	단독주택	아파트	연립다세대	기타	평균금액	전월변동	전세지수
2005.01	88291.4	61149.7	31417.6	34256.7	66799.9		100.0
2005.02	87947.2	61124.7	31288.8	35998.7	66696.3	0.9984	99.8
2005.03	88546.9	60900.3	32270.0	34594.4	66913.4	1.0033	100.2
2005.04	87726.7	60388.4	32446.8	32349.8	66316.1	0.9911	99.3
2005.05	88666.6	60786.7	32607.1	31010.2	66834.5	1.0078	100.1
2005.06	89050.3	59932.8	32877.7	30612.1	66626.3	0.9969	99.7
:	:	:	:	:	:	:	:
2006.07	83492.5	64358.2	34515.1	31317.9	66682.0	0.9878	99.8
2006.08	81564.9	63622.5	35270.8	32993.0	65764.1	0.9862	98.4
2006.09	81482.0	63859.1	35220.4	35646.7	65919.3	1.0024	98.7

<표 20> 월세지수 산출과정

YEAR	단독주택	아파트	연립다세대	기타	합계(십억)	전월변동	월세지수
2005.01	925.3	360.2	334.8	443.4	667.1		100.0
2005.02	926.3	355.6	334.0	448.9	666.5	0.9990	99.9
2005.03	909.4	347.7	324.2	399.5	650.8	0.9765	97.5
2005.04	899.9	348.4	347.1	435.6	649.8	0.9985	97.4
2005.05	892.9	343.3	327.9	410.5	641.4	0.9871	96.1
2005.06	865.5	341.8	360.5	405.6	627.8	0.9788	94.1
:	:	:	:	:	:	:	:
2006.07	869.2	355.6	328.9	400.3	632.3	1.0078	94.8
2006.08	863.3	357.5	329.0	389.1	628.9	0.9947	94.3
2006.09	860.0	360.7	317.7	398.9	628.3	0.9989	94.2

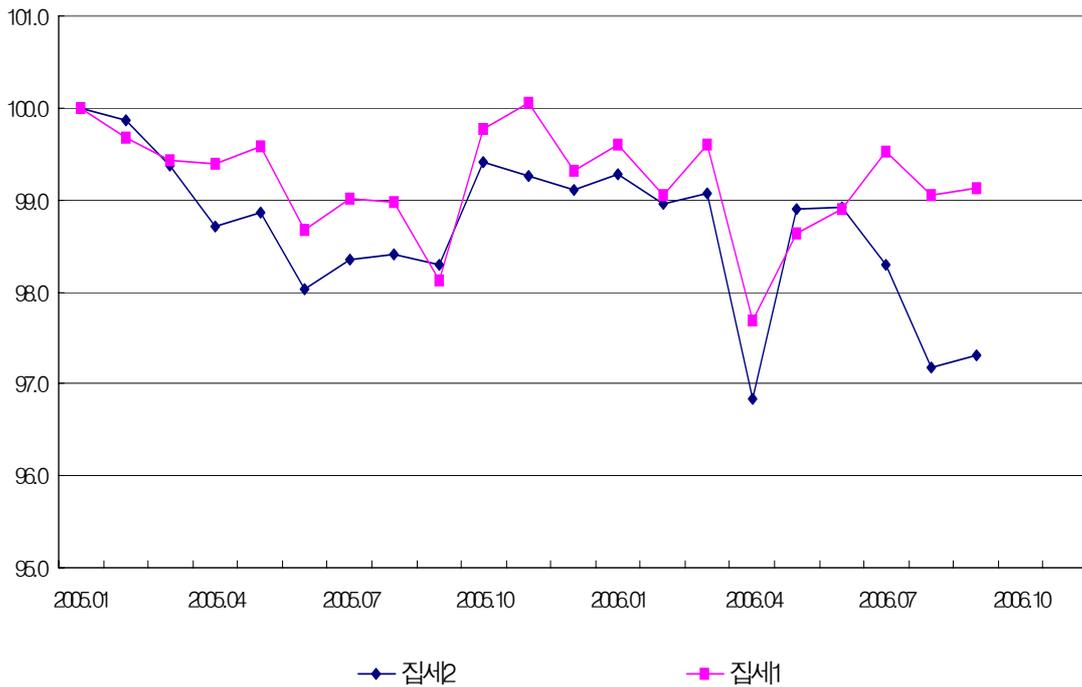
<표 21> 집세지수

	현행지수	전세지수	월세지수	집세지수
2005.01	115.8	100.0	100.0	100.0
2005.02	115.8	99.8	99.9	99.9
2005.03	115.8	100.2	97.5	99.4
2005.04	115.6	99.3	97.4	98.7
2005.05	115.6	100.1	96.1	98.9
2005.06	115.4	99.7	94.1	98.0
2005.07	115.5	99.7	95.2	98.3
2005.08	115.4	99.4	96.1	98.4
2005.09	115.5	99.6	95.2	98.3
2005.10	115.5	100.7	96.5	99.4
2005.11	115.5	100.7	95.9	99.3
2005.12	115.5	100.7	95.4	99.1
2006.01	115.5	100.7	95.9	99.3
2006.02	115.6	100.7	95.0	99.0
2006.03	115.7	101.0	94.6	99.1
2006.04	115.8	97.7	94.8	96.8
2006.05	115.9	101.3	93.5	98.9
2006.06	116	101.1	94.0	98.9
2006.07	116.1	99.8	94.8	98.3
2006.08	116.1	98.4	94.3	97.2
2006.09	116.3	98.7	94.2	97.3
2006.10	116.7			
2006.11	116.8			



<그림 2> 현행지수와 개선2안 비교

계산한 결과를 보면 개선1안 보다는 개선2안 지수가 다소 안정되었으나, 여전히 현행지수보다 불규칙한 움직임을 보이고 있다. 이와 같은 원인은 개선2안의 표본가구수가 작아서 일부 가구의 변동에 민감하게 반응하기 때문인 것으로 판단된다. 이에 따라 표본가구수를 늘리면 해결될 수 있을 것으로 생각된다.



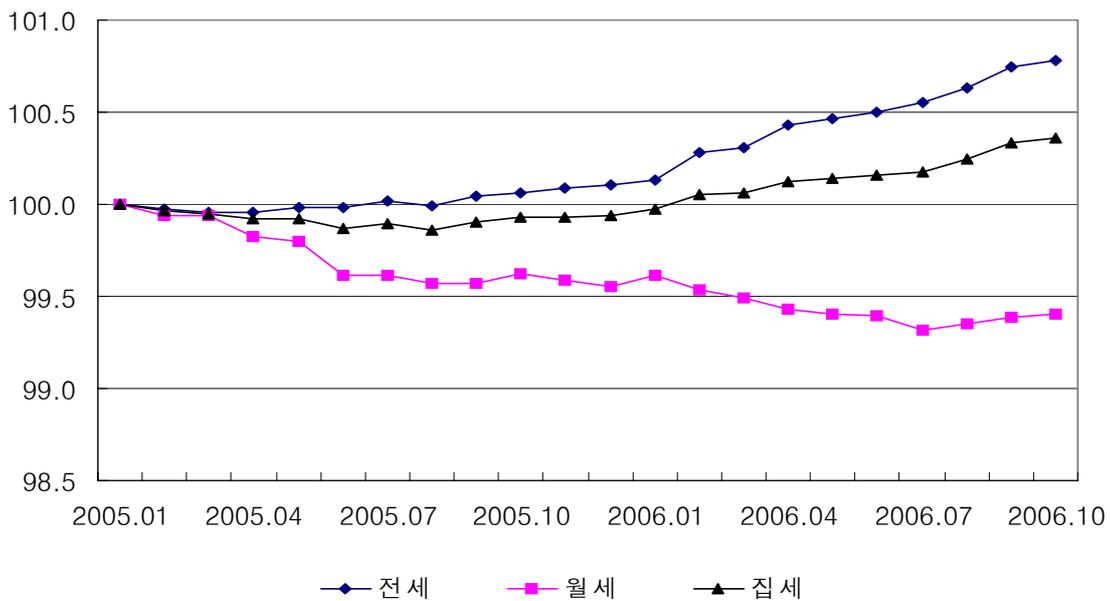
<그림 3> 개선 1안과 개선2안 비교

다. 개선 3안

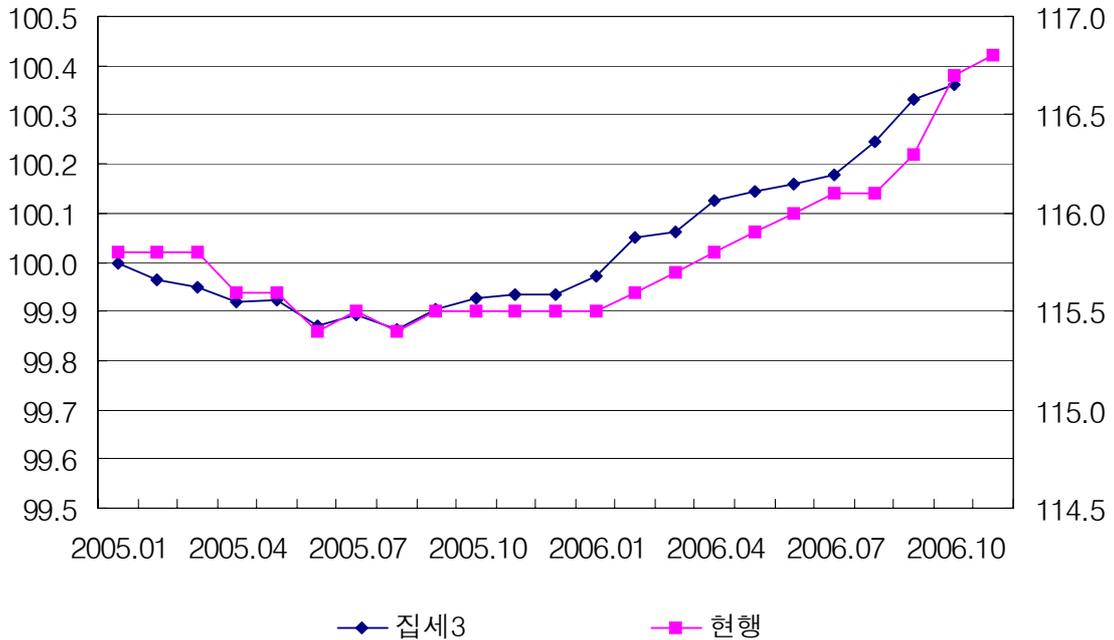
개선 3안의 집세지수는 현행 집세조사의 표본을 이용하여 주택유형별로 계산하였는데, 주택유형별 가중치는 인구주택총조사 결과 80평이하에 거주하는 가구수를 주택유형별로 비율을 사용하였다. 개선 3안의 점유형태는 현행 지수 산출방법과 같게하였다. 그 결과 개선3안 집세지수가 현행 집세지수보다 평활화되었다. 그 원인은 현행방법은 점유형태별로만 구분하였으나, 개선 3안은 주택유형별로 분류를 더 세분하여 전·월세의 변화를 반영하였기 때문인 것으로 생각된다.

<표 22> 집세지수 산출과정

년도	월	전세지수	월세지수	집세지수	현행지수
		가중치 69.55%	가중치 30.45%		
2005	01	100.0	100.0	100.0	115.8
2005	02	100.0	99.9	100.0	115.8
2005	03	100.0	99.9	99.9	115.8
2005	04	100.0	99.8	99.9	115.6
2005	05	100.0	99.8	99.9	115.6
2005	06	100.0	99.6	99.9	115.4
2005	07	100.0	99.6	99.9	115.5
2005	08	100.0	99.6	99.9	115.4
2005	09	100.0	99.6	99.9	115.5
2005	10	100.1	99.6	99.9	115.5
2005	11	100.1	99.6	99.9	115.5
2005	12	100.1	99.5	99.9	115.5
2006	01	100.1	99.6	100.0	115.5
2006	02	100.3	99.5	100.1	115.6
2006	03	100.3	99.5	100.1	115.7
2006	04	100.4	99.4	100.1	115.8
2006	05	100.5	99.4	100.1	115.9
2006	06	100.5	99.4	100.2	116.0
2006	07	100.6	99.3	100.2	116.1
2006	08	100.6	99.4	100.2	116.1
2006	09	100.7	99.4	100.3	116.3
2006	10	100.8	99.4	100.4	116.7



<그림 4> 개선 3안의 집세지수



< 그림 5 > 현행지수와 개선 3안의 집세지수 비교

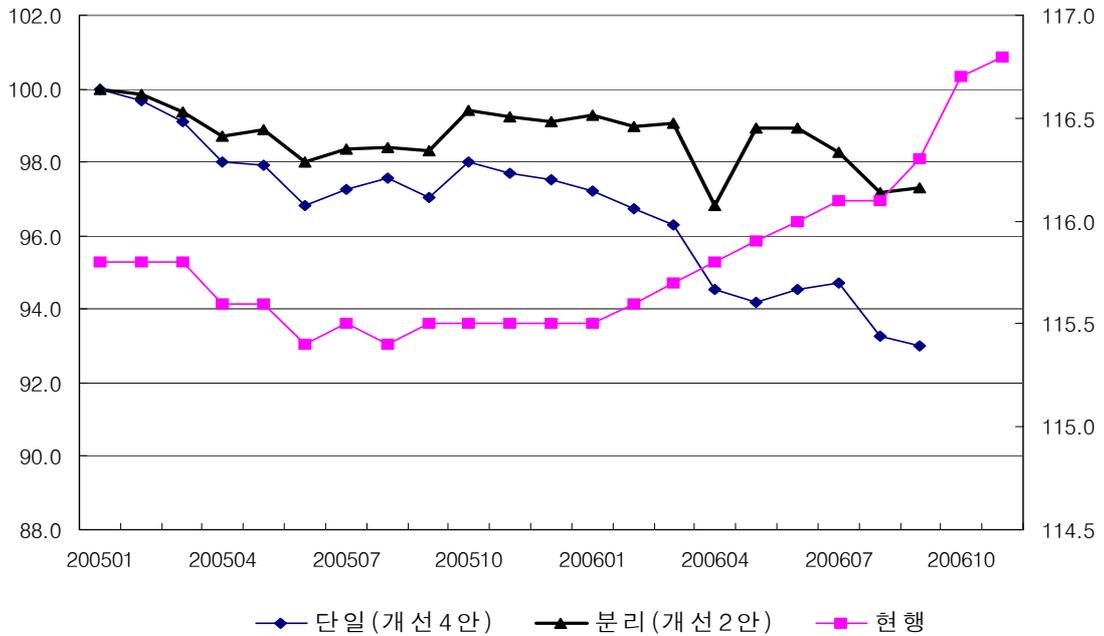
라. 개선 4안

개선 4안은 가계조사에서 전세 및 월세 보증금의 월세평가액에 월세금액과 사글세금액을 합하여 월세금액을 계산하였다. 집세의 점유형태는 단일화하였고 주택유형별 층화하여 월세금액을 층화 비추정법으로 추정하고, 월세금액 추정치의 전월비를 이용하여 집세지수를 산출하였다. 층화 비추정법으로 추정할 때 결합비추정법과 분리비추정법이 모두 추정하였으나 같은 결과가 추정되었다.

집세지수의 움직임이 평활화되어 보이지만 하락하고 있어 현행지수와는 다른 모습을 보였다. 이와 같은 원인은 앞서도 언급하였지만 이용한 자료의 차이에서 비롯된 것으로 판단할 수 있다. 하지만 같은 자료를 이용하여 집세를 추정한 개선 1과 2안 보다는 좋은 추정량을 갖는 것으로 생각되어 추정방법으로는 개선 4안을 제안하고자 한다. 단 전·월세 보증금을 월세평가액으로 환산할 때 사용할 이자율(주택대출금리 또는 정기예금금리)에 대한 연구가 선행되어야 할 것으로 판단된다.

<표 23> 집세지수 산출과정

년월	월세평가액		집세지수
	평균금액	증감률	
200501	717.4	-	100.0
200502	715.1	0.9968	99.7
200503	710.8	0.9940	99.1
200504	703.1	0.9891	98.0
200505	702.4	0.9990	97.9
200506	694.6	0.9889	96.8
200507	697.7	1.0045	97.3
200508	699.9	1.0031	97.6
200509	696.2	0.9947	97.0
200510	703.0	1.0099	98.0
200511	700.8	0.9968	97.7
200512	699.5	0.9982	97.5
200601	697.6	0.9972	97.2
200602	694.0	0.9948	96.7
200603	691.0	0.9957	96.3
200604	678.2	0.9816	94.5
200605	675.7	0.9962	94.2
200606	678.3	1.0039	94.5
200607	679.4	1.0016	94.7
200608	669.0	0.9848	93.3
200609	667.3	0.9975	93.0



< 그림 6 > 현행지수, 개선 2안과 개선4 안의 집세지수 비교

3. 표본크기 추정

자료의 변동과 표본크기에 따라 n에 따라 표본에 포함된 정보의 양이 결정된다. 일단 표본추출이 결정되고 나면 다음으로 표본크기를 얼마로 한 것인가를 정해야한다. 개선안으로 제시한 비추정법을 사용하여 R, μ_y, τ_y 에 대해 추정하려고 할 때 추정오차의 한계 B이내가 되게 하는 표본크기를 결정하는 방법으로 집세조사의 표본크기를 추정하였다. 추정오차 한계를 B 이내로 모집단 평균 μ_y 를 추정하기 위한 표본크기를 정하였다. 필요한 표본크기 n은 다음의 식을 n에 대하여 풀어서 구한다.

$$2\sqrt{V(\hat{\mu}_y)} = B \quad \text{또는} \quad \hat{V}(\hat{\mu}_y) = \mu_x^2 \hat{V}(r)$$

위 식을 n에 대해 풀면 $n = \frac{N\sigma^2}{ND + \sigma^2}$, $D = \frac{B^2}{4}$ 이다. 표본크기 n을 결정하는데 있어서 식을 이용하게 되면 μ_x 의 값을 몰라도 해결할 수 있지만 σ^2 에 대한 추정값은 알아야 한다. 활용 가능한 사전정보가 있는 경우에는 그것을 사용하면 되고 사전정보가 없는 경우에는 예비조사 자료를 이용하여야한다. 본 연구에서는 2005년 인구주택총조사 결과에서 사전 정보를 얻을 수 있었다. 예를 들면, 2005년 1월에 아파트 평당 전세금액을 추정하기 위하여 적절한 표본크기를 살펴보자. 아파트 분산이 1,357,220,926천원이고, 추정오차의 한계가 B 2366(=1.68×1408(= $\sqrt{V(\hat{\mu}_y)}$))이 되게 추정하면 n=968가구가 필요함을 알 수 있었다. 따라서 아파트가구는 285가구를 추가 조사해야 한다. 다른 주택유형별로 표본크기를 아파트와 같이 추정 했을 때 전체적으로 1291가구가 더 필요한 것으로 추정되었다.

<표 24> 주택유형별 전국집세지수 표본크기

2005년 1월		분산(σ^2)	추정오차 $\sqrt{V(\hat{\mu}_y)}$	추정오차한계 $1.68\sqrt{V(\hat{\mu}_y)}$	표본크기 추정	현 표본크기	차이
전 세	단독주택	319013666	617.9	1038	1183	835	348
	아파트	1357220926	1408.3	2366	969	684	285
	연립·다세대	333880154	1291.7	2170	283	200	83
	기타	154334875	2535.6	4260	34	24	10
월 세	단독주택	21423	4.9	8	1252	884	368
	아파트	25964	8.2	14	548	387	161
	연립·다세대	37878	27.5	46	71	50	21
	기타	21063	24.9	42	48	34	14
합계					4389	3098	1291

V. 결 론

본 연구의 목적은 전국 2005년 소비자물가조사 지수개편의 일환인 집세 조사를 위한 새로운 표본을 설계하는 방향을 제시하는 것보다는 최소한의 비용으로 집세지수의 정도를 제고하기 위하여 통계적 추정방법을 제안하고자 하는 것이다.

통계청 집세조사는 매월 조사가 계속되는 계속조사이기는 하지만, 경제활동인구 조사구에서 부수적으로 이루어지는 부가조사에 불과하다. 집세조사는 36개 지역별 통계를 생산할 뿐만 아니라 거의 10,000가구에 이르는 표본크기를 가지는 대규모 조사이다. 이런 대규모 계속조사에서 처음 표본 개편 시점 뿐 아니라, 시간이 경과한 이후에도 지속적으로 일정 수준 이상의 통계품질을 유지할 수 있도록 하는 것도 매우 중요할 것이다. 이를 위해서는 지속적인 표본설계와 추정방법 개선을 위한 예산과 인력, 그리고 시스템을 마련하는 것이 필수적이지만 실행하기는 현실적으로 어려운 실정이다.

본 연구에서는 이러한 한계를 통계적 추정방법으로 집세지수의 현실반영도를 높이하고자 4개의 개선안을 제안하고 집세조사의 결과를 직접 이용하여 지수산출을 해보지 못했으나 이와 유사한 가구조사내의 전월세금액을 활용하여 시산하여 보았다. 개선 3안을 제외한 개선안들이 현행지수와 다른 모습을 보이고 있으나 이것은 이용되는 자료의 차이 때문에 발생하는 원인이라 사료된다. 현재 집세지수는 모수를 추정하기 않고 표본의 움직임만으로 금액을 추정하여 사용하고 있다. 이에 따라 현행 집세지수를 개선 1,2,4안과 같이 가구의 평당금액을 도출하여 사용하였다면 정도 높은 개선안을 마련할 수 있었을 것으로 생각되고 특히 4안을 개선안으로 제안하고자 한다. 다양한 보조변수로 연면적뿐만아니라 사용방수, 가구원수 등도 활용하여 전월세 금액을 추정하는 것이 바람직한 것으로 판단된다.

그러나 본 연구에서 언급된 추정방법 이외에도 많은 부문이 고려되어야 할 것이다. 집세조사가 품질이 높은 조사가 되기 위해서는 실증 분석의 문제점을 토대로 중요하다고 생각되는 몇가지 사항을 제안하고자 한다.

첫째로 집세조사를 위한 독립표본이 필요하다. 개선2안에 의하여 표본크기를 추정한 결과, 전국단위 집세지수 작성을 위해서는 1291가구를 추가하는 되는 것으로 추정되었다. 그러나 지역별로 집세지수 단위를 확대할 경우에는 더 많은 표본가구수가 필요할 것이다. 현재 집세지수 표본은 1만여 가구 규모로 작은 표본이 아니므로, 추출방법을 비추정법으로 지역별/주택유형별/점유형태별로 표본설계를 하는 것을 고려해 볼 필요가 있다.

둘째로 전세가격을 월세평가액으로 환산하는 방법에서는 이자율을 적용하는 것을 신중히 고려해야 될 것이다. 집세지수가 이자율에 따라 움직일 수 있기 때문이다. 전세가격을 반영시에는 같은 규모의 주택에 월세를 적용하는 것이 자연적인 집세의 움직임을 반영하는 것이라 생각된다. 부족하면 국민은행에서 조사되고 있는 주택가격동향조사를 매월 참고하여 전세가격의 월세평가액을 산출했으면 한다.

셋째로 전세 가격반영 방법을 월세평가액으로 환산함으로써 취득접근법에서 월세지수 산출방법과 같은 사용자비용접근법으로 가격반영방법의 이원화에서 일원화를 이룰 수 있을 것이다. 국제노동기구(ILO)에서는 소비자 물가지수 편제시 자가계정 주거서비스의 평가액을 측정하기 위한 방법으로 집세상당액접근법, 사용자비용접근방법, 취득가액접근방법, 지불액접근법을 제시하고 있으나, 우리나라는 자가주거비를 산출하지 않고 있다. 전세금액에 이자율을 반영하는 지수 산출방법은 사용자비용접근법과 유사하다. 그리고 전세금과 보증금에 대하여 월세평가액으로 환산하는 것과 같이 자가주거비를 산출토록 권고하는 국제적 기준에 한 걸음 더 다가갈 수 있는 계기도 된다.

마지막으로 표본조사에 의한 통계들의 표본 규모를 늘리려고 하는 것보다는 다양한 추정방법으로 통계값을 추정하고 서로의 특징을 분석해보는 것도 통계의 신뢰성을 확보할 수 있는 한 방안이라고 생각된다.

본 연구에서 아쉬운 것은 표본설계에 있어서 지역별/주택유형별/점유형태별로 표본크기를 제안하고자 했으나 집세조사 자료를 수집하지 못해 실증분석까지는 하지 못했다. 추후에 집세관련 자료를 더 수집하여 심층적인 연구가 이루어졌으면 한다.

참고문헌

- [1] 강재춘(2006), “소비자물가통계에 있어 집세 지수에 대한 연구”, 석사학위논문, 한남대학교 행정정책대학원.
- [2] 김영원, 류제복, 박진우, 홍기학 (2000), 「표본조사의 이해와 활용」, 自由아카데미.
- [3] 박홍래(1996), 「표본조사론」, 英志文化社.
- [4] 통계청(2005a), 「2005 물가연보」
- [5] 통계청(2005b), 「국가통계 바로알기」
- [6] 통계청(2003a), 「소비자물가조사지침서 I」
- [7] 통계청(2003b), 「가구부문 표본개편 보고서」
- [8] 통계청(2002), 「소비자물가지수 개편 보고서」
- [9] 한국은행(2006), 「유로지역의 자가주거비 소비자물가지수 반영논의」 해외조사실 歐美經濟팀. 해외경제정보 제 2006-1호
- [10] Diewert, W.E.(2002). Harmonized indexes of consumer prices: their conceptual foundation. Paper presented at the Joint of the Centre for Economic Policy Research and the European Central Bank. Issues in the Measurement of Price Inflation, Frankfurt, November 16-17, 2001. This paper is available in the ECB Discussion Paper Series.
- [11] EU(2002). Housing Statistics in the European Union 2001.
- [12] Eurostat(2000), “The treatment of owner occupied housing in the HICP”, *Final Report by Task Force XII*, January, Luxembourg.

[13] ILO, IMF, OECD, Eurostat, UNECE and the World Bank(2004), *Consumer Price Index Manual* (Brussels/Luxembourg, Geneva, Washington, D.C.)

[14] ILO (2002), "*Some Special Cases*", *the CPI Manual* Chapter 10.

[15] Oh, Dong-Hoon. 1995. Housing Budget Share, Housing Expenditure, and Housing Affordability of U.S. Urban Households By Housing Tenure. Dissertation Committee for the Degree Doctor of Philosophy in the Graduate School of the Ohio State

[16] <http://www.kbstar.com>(국민은행, 2006년 9월 전국주택가격동향조사).

[17] U.S. Bureau of Labor Statistics. 2004. *BLS Handbook of Methods*. available at: <http://www.bls.gov>

[18] 日本, 總務省 ; 小賣物價調查概要 at : <http://www.stat.go.jp/index.htm>

< 부록 >

집세금액 추정에 사용한 SAS 프로그램

1. 비추정

```
-----;
* COMPUTES RATIO ESTIMATE, ESTIMATE OF POPULATION R, MEAN, OR ;
* FROM A SIMPLE RANDOM SAMPLE WITH AUXILIARY VARIABLE X. ;
-----;
%MACRO RATIO(SAMPLE=,SETUP=,NPOP=,N=,RESPONSE=,PARAM=,X=,TAU_X=,
MU_X=,FPC=);
%IF %LENGTH(&SAMPLE) = 0 %THEN %LET SAMPLE = %STR(SAMPLE);
%IF %LENGTH(&NPOP) = 0 %THEN %LET NPOP = %STR(NPOP);
%IF %LENGTH(&X) = 0 %THEN %LET X = %STR(X);

%IF %LENGTH(&FPC) = 0 %THEN %DO;
%LET FPC_ = %STR( (&NPOP - N_) / &NPOP );
%IF %LENGTH(&TAU_X) > 0 %THEN %LET MU_X = %STR(&TAU_X/&NPOP);
%END;
%ELSE %DO;
%LET FPC_ = 1;
%END;
%IF %LENGTH(&MU_X) = 0 %THEN %LET MU_X = %STR(XBAR_);

DATA MUX_; SET &SAMPLE(KEEP = &X &RESPONSE); CP_ = &X*&RESPONSE;
PROC MEANS DATA = MUX_ NOPRINT;
VAR &RESPONSE &X CP_;
OUTPUT OUT = EST_ MEAN = YBAR_ XBAR_
SUM = SUMY_ SUMX_ SUMXY_ N = N1_ N2_ N_ CSS = SSYY_ SSXX_;

%IF %LENGTH(&SETUP) > 0 %THEN %DO;
DATA EST_; MERGE EST_ &SETUP;
%END;
%ELSE %DO;
DATA EST_; SET EST_;
%END;
R_ = SUMY_/SUMX_;
SSXY_ = SUMXY_ - SUMX_*SUMY_/N_;
SR2_ = (SSYY_ + R_**2*SSXX_ - 2*R_*SSXY_)/(N_-1);
VAR_R_ = &FPC_*SR2_/N_*(&MU_X)**2);
STD_R_ = SQRT(VAR_R_);
BND_R_ = 2*STD_R_;
MU_HAT_ = &MU_X*R_;
STD_MU_ = &MU_X*STD_R_;
BND_MU_ = 2*STD_MU_;
%IF %LENGTH(&FPC) = 0 %THEN %DO;
TAU_HAT_ = &NPOP*MU_HAT_;
STD_TAU_ = &NPOP*STD_MU_;
BND_TAU_ = 2*STD_TAU_;
%END;
%ELSE %DO;
TAU_HAT_ = &TAU_X*R_;
STD_TAU_ = &TAU_X*STD_R_;
BND_TAU_ = &TAU_X*BND_R_;
%END;
%IF %INDEX(%UPCASE(&PARAM),R) > 0 %THEN %DO;
PROC PRINT DATA = EST_ NOOBS SPLIT='*';
TITLE1 'RATIO ESTIMATE OF POPULATION R';
TITLE2 'SIMPLE RANDOM SAMPLE DESIGN';
TITLE3 "RESPONSE VARIABLE = &RESPONSE";
TITLE4 "(AUXILIARY VARIABLE = &X)";
LABEL R_ = 'ESTIMATE';
LABEL STD_R_ = 'STANDARD*ERROR';
LABEL BND_R_ = 'BOUND';
LABEL SR2_ = 'S(R)^2';
LABEL N_ = 'SAMPLE*SIZE';
VAR R_ STD_R_ BND_R_ SR2_ N_;
```

```

%END;
%IF %INDEX(%UPCASE(&PARAM),MEAN) > 0 %THEN %DO;
PROC PRINT DATA = EST_NOOBS SPLIT='*';
  TITLE1 'RATIO ESTIMATE OF POPULATION MEAN';
  TITLE2 'SIMPLE RANDOM SAMPLE DESIGN';
  TITLE3 "RESPONSE VARIABLE = &RESPONSE";
  TITLE4 "(AUXILIARY VARIABLE = &X)";
  LABEL MU_HAT_ = 'ESTIMATE';
  LABEL STD_MU_ = 'STANDARD*ERROR';
  LABEL BND_MU_ = 'BOUND';
  LABEL SR2_ = 'S(R)^2';
  LABEL N_ = 'SAMPLE*SIZE';
  VAR MU_HAT_ STD_MU_ BND_MU_ SR2_ N_;
%END;
%IF %INDEX(%UPCASE(&PARAM),TOTAL) > 0 %THEN %DO;
PROC PRINT DATA = EST_NOOBS SPLIT='*';
  TITLE1 'RATIO ESTIMATE OF POPULATION TOTAL';
  TITLE2 'SIMPLE RANDOM SAMPLE DESIGN';
  TITLE3 "RESPONSE VARIABLE = &RESPONSE";
  TITLE4 "(AUXILIARY VARIABLE = &X)";
  LABEL TAU_HAT_ = 'ESTIMATE';
  LABEL STD_TAU_ = 'STANDARD*ERROR';
  LABEL BND_TAU_ = 'BOUND';
  LABEL SR2_ = 'S(R)^2';
  LABEL N_ = 'SAMPLE*SIZE';
  VAR TAU_HAT_ STD_TAU_ BND_TAU_ SR2_ N_;
%END;
RUN; TITLE;
%MEND RATIO;

```

2. 결합비추정

```

*-----*
* Combined ratio estimate of population mean and total, for ;
* a stratified random sample with auxiliary variable X. ;
*-----*
%macro c_ratio(sample=,response=,strata=,param=,setup=,
  npop=n,mu_x=,tau_x=,x=,fpc=,wts=);

%if %length(&sample) = 0 %then %let sample = %str(sample);
%if %length(&strata) = 0 %then %let strata = %str(strata);
%if %length(&setup) = 0 %then %let setup = %str(setup);
%if %length(&x) = 0 %then %let x = %str(x);

%if %length(&tau_x) > 0 %then %let mu_x = %str(&tau_x/&npop);
%else %put %str(*** Neither MU_X nor TAU_X was defined ***);

proc sort data = &setup; by &strata;
proc sort data = &sample; by &strata;

%if %length(&npop) > 0 %then %do;
proc means data = &setup noprint;
var &npop;
output out = ntot_ sum = ntotal_;
data ntot_; set ntot_;
call symput('ntot',trim(left(ntotal_)));
run;
%end;

data est_; set &sample( keep = &x &response &strata);
xy_ = &x*&response;

proc sort data = est_; by &strata;

proc means data = est_ noprint; by &strata;
var &x &response xy_;
output out = c_out_ mean = xbari_ ybari_
sum = sumxi_ sumyi_ sumxyi_ css = sxxi_ syyi_
n = n1_ n2_ ni_;

```

```

data c_out_; merge c_out_ &setup; by &strata;
    ssxyi_ = sumxyi_ - sumxi_*sumyi_/ni_;
%if %length(&npop) = 0 %then %do;
    wts_ = &wts;
%end;
%else %do;
    wts_ = &npop/&ntot;
%end;
    yy_ = wts_*ybari_;
    xx_ = wts_*xbari_;
    mu_temp_ = wts_*(&mu_x);

proc means data = c_out_ noprint;
    var xx_ yy_ mu_temp_;
    output out = combr_ sum = totx_ toty_ tu_x_;

data combr_; set combr_;
    mu_xpop_ = tu_x_;
    rc_ = toty_/totx_;
    mu_hat_ = rc_*mu_xpop_;

data final_; if _n_ = 1 then set combr_; set c_out_;
    sr2_ = (ssyyi_ + rc_**2*ssxxi_ - 2*rc_*ssxyi_)/(ni_-1);

%if %length(&npop) > 0 %then %do;
    fpc_ = (&npop-ni_)/&npop;
%end;
%else %do;
    fpc_ = 1;
%end;
%if %length(&fpc) > 0 %then %do;
    fpc_ = 1;
%end;
%if %length(&npop) = 0 %then %do;
    wts_ = &wts;
%end;
%else %do;
    wts_ = &npop/&ntot;
%end;

    vterm_ = (wts_)**2*fpc_/ni_*sr2_;
    rterm_ = vterm_/mu_xpop_**2;

proc means data = final_ noprint;
    var vterm_ rterm_;
    output out = est_cr_ sum = vnumer_ rnumer_;

data est_cr_; merge est_cr_ combr_;
    std_mu_ = sqrt(vnumer_);
    bnd_mu_ = 2*std_mu_;
    std_r_ = sqrt(rnumer_);
    bnd_r_ = 2*std_r_;
%if %index(%upcase(&param),TOTAL) > 0 %then %do;
    tau_hat_ = &ntot*mu_hat_;
    std_tau_ = &ntot*sqrt(vnumer_);
    bnd_tau_ = 2*std_tau_;
%end;
drop vnumer_;

proc print data = c_out_ noobs split='*';
    title1 'Actual Sample Sizes';
    title2 'Excludes Missing Data, If Any';
    label ni_ = 'n(i)';
    var &strata ni_;

%if %index(%upcase(&param),MEAN) > 0 %then %do;
proc print data = est_cr_ noobs split='*';
    title1 "Combined Ratio Estimate of Population Mean";
    title2 "Stratified Random Sampling Design";

```

```

        title3 "Response Variable = &response";
        title4 "(Auxiliary Variable = &x)";
        label mu_hat_ = 'Estimate';
        label std_mu_ = 'Standard*Error';
        label bnd_mu_ = 'Bound';
        label rc_ = 'r(c)';
        var mu_hat_ std_mu_ bnd_mu_ rc_;
    %end;
    %if %index(%upcase(&param),TOTAL) > 0 %then %do;
        proc print data = est_cr_ noobs split='*';
            title1 "Combined Ratio Estimate of Population Total";
            title2 "Stratified Random Sampling Design";
            title3 "Response Variable = &response";
            title4 "(Auxiliary Variable = &x)";
            label tau_hat_ = 'Estimate';
            label std_tau_ = 'Standard*Error';
            label bnd_tau_ = 'Bound';
            label rc_ = 'r(c)';
            var tau_hat_ std_tau_ bnd_tau_ rc_;
        %end;
    %if %index(%upcase(&param),R) > 0 %then %do;
        proc print data = est_cr_ noobs split='*';
            title1 "Combined Ratio Estimate";
            title2 "Stratified Random Sampling Design";
            title3 "Response Variable = &response";
            title4 "(Auxiliary Variable = &x)";
            label rc_ = 'r(combined)';
            label std_r_ = 'Standard*Error';
            label bnd_r_ = 'Bound';
            var rc_ std_r_ bnd_r_;
        %end;
    run; title;
    %mend c_ratio;

```

3. 분리비추정

```

*-----;
* Separate ratio estimate of population mean and total, and ;
* individual estimates of R, for a stratified random sample ;
* with auxiliary variable X ;
*-----;
%macro s_ratio(sample=,response=,strata=,param=,setup=,
    npop=,n=,mu_x=,tau_x=,x=,fpc=,wts=);
%if %length(&sample) = 0 %then %let sample = %str(sample);
%if %length(&strata) = 0 %then %let strata = %str(strata);
%if %length(&setup) = 0 %then %let setup = %str(setup);
%if %length(&x) = 0 %then %let x = %str(x);

%if %length(&tau_x) > 0 %then %let mu_x = %str(&tau_x/&npop);
%else %put %str(***) Neither MU_X nor TAU_X was defined (***) ;

    proc sort data = &setup; by &strata;

%if %length(&npop) > 0 %then %do;
    proc means data = &setup noprint;
        var &npop;
        output out = ntot_ sum = ntotal_;
        data ntot_; set ntot_;
        call symput('ntot',trim(left(ntotal_)));
    run;
%end;

    data est_; set &sample( keep = &x &response &strata);
        xy_ = &x*&response;
    proc sort data = est_; by &strata;

    proc means data = est_ noprint;
        by &strata; var &x &response xy_;
        output out = sep_out_ sum = sumxi_ sumyi_ sumxyi_
            css = sxxi_ syyi_ n = n1_ n2_ ni_ mean = xbar ybar;

```

```

data sep_out_; merge sep_out_ &setup; by &strata;
  ssxyi_ = sumxyi_ - sumxi_*sumyi_/ni_;
  ri_ = sumyi_/sumxi_;
  mu_hati_ = ri_*(&mu_x);
  sri_ = (ssyyi_ + ri_**2*ssxxi_ - 2*ri_*ssxyi_)/(ni_-1);

%if %length(&npop) > 0 %then %do;
  fpc_ = (&npop-ni_)/&npop;
%end;
%else %do;
  fpc_ = 1;
%end;
%if %length(&fpc) > 0 %then %do;
  fpc_ = 1;
%end;
%if %length(&npop) = 0 %then %do;
  wts_ = &wts;
%end;
%else %do;
  wts_ = &npop/&ntot;
%end;

  m_hati_ = wts_*mu_hati_;
  vterm_ = wts_**2*fpc_*sri_/ni_;
  drop sumyi_ sumxi_ sumxyi_ ssxxi_ ssyyi_ ssxyi_;

proc means data = sep_out_ noprint;
  var m_hati_ vterm_;
  output out = s_ratio_ sum = mu_hat_ vnumer_;

data est_sr_; set s_ratio_;
  var_mu_ = vnumer_;
  std_mu_ = sqrt(var_mu_);
  bnd_mu_ = 2*std_mu_;

%if %index(%upcase(&param),TOTAL) > 0 %then %do;
  tau_hat_ = (&ntot)*mu_hat_;
  std_tau_ = (&ntot)*std_mu_;
  bnd_tau_ = 2*std_tau_;
%end;

proc print data = sep_out_ noobs split='*';
  title1 "Ratio Estimates by Strata";
  title2 "Response Variable = &response";
  title3 "(Auxiliary Variable = &x)";
  label &strata = "Strata";
  label ni_ = "Sample*Size";
  label ri_ = "r(i)";
  var &strata ni_ ri_;

%if %index(%upcase(&param),MEAN) > 0 %then %do;
  proc print data = est_sr_ noobs split='*';
    title1 "Separate Ratio Estimate of Mean";
    title2 "Stratified Random Sampling Design";
    title3 "Response Variable = &response";
    title4 "(Auxiliary Variable = &x)";
    label mu_hat_ = 'Estimate';
    label std_mu_ = 'Standard*Error';
    label bnd_mu_ = 'Bound';
    var mu_hat_ std_mu_ bnd_mu_;
  %end;
%if %index(%upcase(&param),TOTAL) > 0 %then %do;
  proc print data = est_sr_ noobs split='*';
    title1 "Separate Ratio Estimate of Total";
    title2 "Stratified Random Sampling Design";
    title3 "Response Variable = &response";
    title4 "(Auxiliary Variable = &x)";
    label tau_hat_ = 'Estimate';

```

```

label std_tau_ = 'Standard*Error';
label bnd_tau_ = 'Bound';
var tau_hat_ std_tau_ bnd_tau_ ;
%end;
run; title;
%mend s_ratio;

DATA WORK.JTAX ;
%LET _EFIERR_ = 0; /* SET THE ERROR DETECTION MACRO VARIABLE */
INFILE 'I:\TAX\가계조사.TXT' DELIMITER='09'X MISSEVER DSD LRECL=32767 FIRSTOBS=2 ;
INPUT
  @1 HSYMM $6.
  @7 KEY $16.
  @24 PYON 8.
  @39 SIL_E $1.
  @40 SIL_2 8.
  @55 SIL_D_1 $1.
  @56 SIL_D_2 8.
  @64 SIL_D_3 8.
  @72 SIL_D_4 8.
  @84 RAT1 12.4
  @97 KUM 13.
  @115 PCODE $8. ;

YEAR=SUBSTR(HSYMM,1,4) ;
MON=SUBSTR(HSYMM,5,2) ;
SIDO=SUBSTR(KEY,1,2) ;

IF SIL_D_1 IN('4') THEN DO ;
  GECH= 'A' ;
  YKUM=SIL_D_2 ; END ;

ELSE IF SIL_D_1 IN('5','6','7') THEN DO ;
  GECH= 'B' ;
  YKUM=SUM(SIL_D_2,SIL_D_3) ; END ;
ELSE DELETE ;

IF PYON< 19 THEN PYY='S' ;
ELSE IF 19<=PYON< 39 THEN PYY='M' ;
ELSE IF 39<=PYON<=80 THEN PYY='L' ;
ELSE DELETE ;

IF _ERROR_ THEN CALL SYMPUTX(_EFIERR_,1); /* SET ERROR DETECTION MACRO VARIABLE */

RUN;

%MACRO SAMDATA(V1,V2) ;
DATA &V1&V2 ; SET JTAX ;
LENGTH SIL_A $2. ;
KEYY=HSYMM||GECH||SIL_E ;
SIL_A=GECH||SIL_E ;
IF HSYMM IN("&V2") THEN OUTPUT &V1&V2 ;
KEEP KEYY PYON YKUM SIL_A ;
RUN ;
%MEND ;
%SAMDATA(SM,200501) ; %SAMDATA(SM,200502) ; %SAMDATA(SM,200503) ;
%SAMDATA(SM,200504) ; %SAMDATA(SM,200505) ; %SAMDATA(SM,200506) ;
%SAMDATA(SM,200507) ; %SAMDATA(SM,200508) ; %SAMDATA(SM,200509) ;
%SAMDATA(SM,200510) ; %SAMDATA(SM,200511) ; %SAMDATA(SM,200512) ;
%SAMDATA(SM,200601) ; %SAMDATA(SM,200602) ; %SAMDATA(SM,200603) ;
%SAMDATA(SM,200604) ; %SAMDATA(SM,200605) ; %SAMDATA(SM,200606) ;
%SAMDATA(SM,200607) ; %SAMDATA(SM,200608) ; %SAMDATA(SM,200609) ;

DATA ONE ;
INPUT SIL_A $2. STR_SIZE TOTALX ;
CARDS ;
A1 1196563 53395643
A2 1372402 28462709
A3 412707 7465509
A4 103217 2498173
B1 1259621 53615586
B2 790536 11777420
B3 140254 2182768
B4 160502 3447350

DATA TWO ; SET ONE ;
DO YEAR='2005';

```

```

DO MON='01','02','03','04','05','06',
      '07','08','09','10','11','12' ;
KEYY=YEAR||MON||SIL_A ;
OUTPUT ;
END ;
END ;
DO YEAR='2006';
DO MON='01','02','03','04','05','06',
      '07','08','09' ;
KEYY=YEAR||MON||SIL_A ;
OUTPUT ;
END ;
END ;

%INCLUDE 'C:\C_RATIO' ;
%INCLUDE 'C:\S_RATIO' ;
%INCLUDE 'C:\RATIO1.SAS' ;

RUN ;

%MACRO RATT(C1,C2) ;

%S_RATIO(SAMPLE=&C1, SETUP=ONE, STRATA=SIL_A, X=PYON,
          RESPONSE=YKUM,TAU_X=TOTALX, PARAM=MEAN, NPOP=STR_SIZE) ;
DATA SR&C2 ; SET EST_SR_ ;
  Y_M="&C2" ;

DATA SR ; SET SR SR&C2 ;
IF Y_M='.' THEN DELETE ;

%C_RATIO(SAMPLE=&C1, SETUP=ONE, STRATA=SIL_A, X=PYON,
          RESPONSE=YKUM,TAU_X=TOTALX, PARAM=MEAN, NPOP=STR_SIZE) ;

DATA CR&C2 ; SET EST_CR_ ;
  Y_M='&C2' ;

DATA CR ; SET CR SR&C2 ;
IF Y_M='.' THEN DELETE ;

%MEND ;

DATA SR ;
DATA CR ;

%RATT(SM200501,200501) ; %RATT(SM200502,200502) ; %RATT(SM200503,200503) ;
%RATT(SM200504,200504) ; %RATT(SM200505,200505) ; %RATT(SM200506,200506) ;
%RATT(SM200507,200507) ; %RATT(SM200508,200508) ; %RATT(SM200509,200509) ;
%RATT(SM200510,200510) ; %RATT(SM200511,200511) ; %RATT(SM200512,200512) ;
%RATT(SM200601,200601) ; %RATT(SM200602,200602) ; %RATT(SM200603,200603) ;
%RATT(SM200604,200604) ; %RATT(SM200605,200605) ; %RATT(SM200606,200606) ;
%RATT(SM200607,200607) ; %RATT(SM200608,200608) ; %RATT(SM200609,200609) ;

PROC PRINT DATA=SR ;
PROC PRINT DATA=CR ;

RUN ;

```