

제3장

집세조사의 표본규모 산정 및 관리방안

제3장



정미옥·박소현

제1절 서론

1. 연구배경

집세조사는 도시에 거주하는 가구가 주거생활을 위해 지출하는 전·월세금액을 파악하여 집세금액의 변동을 측정한다. 집세조사의 결과는 소비자물가지수의 9.75%를 차지하고 있는 집세지수를 작성하는데 활용된다. 전국 및 16개 시도¹⁾에 대해서는 전세지수와 월세지수를, 소비자물가지수 대상 38개 도시²⁾에 대해서는 전·월세 구분없이 집세지수를 공표하고 있다. 집세조사는 표본으로 선정된 임차가구로부터 직접 응답을 받는 가구대상 조사로 매월 조사를 수행한다. 집세조사의 표본은 기본적으로 경제활동인구조사의 표본을 공용하고 있으며, 표본수가 부족한 일부 도시에 대해서만 추가적으로 표본을 보완한다.

그러나 집세조사의 표본은 집세 측정이라는 목적에 맞춰 설계되지 않았기 때문에 표본의 대표성을 확보하기 위해 독립표본설계의 필요성이 제기되었다. 이에 따라 집세조사 목적에 부합하는 표본을 선정할 수 있도록 주택 시장 및 집세의 변동과 상관이 높은 요인들을 고려한 독립표본설계 방안에 대한 연구를 2009년도에 수행하였다. 이 연구에서는 집세 수준 및 변동을 설명할 수 있는 요인이 무엇인지 실증분석을 통해 확인하였다. 그리고 그 결과를 이용하여 임차가구의 거주 주택유형에 따라 아파트와 아파트 이외의 일반주택으로 모집단을 분리하여 독립적으로 표본설계하였다.

1) 1개 특별시, 6개 광역시, 9개 도

2) 서울, 부산, 대구, 인천, 광주, 대전, 울산, 수원, 성남, 의정부, 안양, 부천, 안산, 고양, 춘천, 원주, 강릉, 청주, 충주, 천안, 보령, 서산, 전주, 군산, 남원, 목포, 여수, 순천, 포항, 경주, 안동, 구미, 창원, 마산, 진주, 김해, 제주, 서귀포

이 연구에서 제안된 표본설계 관련내용을 간략히 요약해보면 다음과 같으며, 보다 자세한 내용은 정미옥 외(2009)를 참고하길 바란다. 먼저 일반주택과 아파트를 독립적으로 표본설계함에 따라 통계작성 범위를 좀 더 세분화시키고자 하였다. 집세지수의 작성 단위인 전국 및 16개 시도에 대해서 기존의 전세지수와 월세지수에 추가적으로 일반주택지수, 아파트지수 작성을 목표로 하였다. 일반주택에 대한 추출틀은 인구주택총조사 10% 표본조사구 정보를 이용하였고, 아파트에 대한 추출틀로는 부동산정보업체의 아파트 평형단지 정보를 사용하였다. 표본설계방법은 층화2단집락추출방법을 적용하였으며, 층내 자체가중을 만족할 수 있도록 설계하였다. 각 시도를 부차모집단으로 설정하고, 해당 시도를 여러 개의 층으로 구분하는데 있어서 가장 높은 효율을 보이는 지역적 층화방법을 제안하였다. 표본추출방법으로는 각 층내에서 1차추출단위인 집락을 확률비례계통추출하고, 각 집락 내에서 사전에 정해놓은 일정크기 만큼의 2차추출단위인 가구를 계통추출하였다. 이때, 1차추출단위는 일반주택의 경우 조사구로, 아파트는 평형단지로 정하였다. 그리고 표본의 크기는 기존 집세조사의 표본크기 및 변동률 평균 추정값에 대한 상대표준오차(CV: Coefficient of Variation)를 이용하여 산출하였다.

그러나 이 연구에서 집세조사의 표본을 독립표본으로 새롭게 설계하는 데 있어서 표본규모를 기존 집세조사의 결과만을 이용하여 적정 표본규모를 산정하는 데는 몇가지 문제가 있다고 지적하였다.

첫째, 기존 집세조사의 추정값에 대한 상대표준오차를 기준으로 표본크기를 산정하였을 때, 새로운 표본설계의 표본규모가 기존의 표본수와 조사결과에 매우 의존하게 된다는 것이다. 특히 최근 몇 년 동안 표본가구의 집세변동이 거의 일어나지 않아 변동률 평균의 상대표준오차가 매우 낮은 몇몇 도시들의 경우 표본의 크기가 너무 작게 계산되는 문제가 발생한다고 하였다. 따라서 이를 극복하기 위해서는 도시별로 기본적으로 유지되어야 하는 최소표본규모를 산정해야 할 것이라고 언급하였다.

둘째, 변동률의 경우 특정 시점이나 외부요인의 개입에 의해 상대표준오차의 크기가 극심하게 변동할 가능성이 있기 때문에 안정적인 표본규모를 계산하기 어려울 수 있다. 따라서 이와 같은 경우엔 집세 변동률과 같이 직접적으로 추정되어야 할 목표변수가 아니더라도 목표변수와 상관관계가 높은 집세 수준과 같은 기준변수를 고려한다면 안정적인 표본규모를 산정할 수 있을 것이라고 제안하였다.

셋째, 38개 도시별로 상대표준오차의 크기가 매우 차이가 나타나고 있어 도시별 목표정도를 다르게 책정해야 하는지, 목표정도는 어떠한 기준을 적용할 것인지에 대한 세밀한 검토가 필요하다고 언급하였다. 또한 이론상 계산된 표본의 크기뿐만 아니라 비용, 인력, 시간과 같은 현실적이고 경험적인 측면을 종합적으로 고려해야 한다고 하였다.

마지막으로 현행 집세조사에서는 표본조사구로 선정된 조사구에 거주하는 모든 임차



가구를 표본가구로 선정하고 있는데, 만약 특정 조사구의 임차가구비율이 높은 경우 그 조사구에서 집중적으로 표본이 추출되어 지역적으로 치우친 표본이 선정될 문제가 있다고 지적하였다. 이에 따라 조사구 내 표본 수를 조절하고 다양한 특성의 조사구를 추출함으로써 설계의 효율을 높이고자 일반주택은 조사구별 8가구, 아파트는 평형단지별 4가구를 조사하는 것을 원칙으로 하였다. 그러나 이러한 기준이 객관적이고 타당한지에 대해서는 추가적으로 검토되어야 한다고 밝혔다.

이와 같이 집세조사의 독립표본설계에 대한 선행연구에서 표본규모 산정과 관련하여 제기된 문제점들은 표본의 대표성과 신뢰성을 확보하고, 조사의 효율성과 정도를 제고하기 위해 실질적으로 해결해야 할 과제들이다. 모집단으로부터 관측할 표본의 수를 결정하는 문제는 반드시 비용을 수반한다. 만약 표본이 지나치게 크면 시간과 인력의 낭비를 가져오고, 반대로 표본의 크기가 너무 작으면 들인 시간과 노력에 비하여 신뢰할 수 없는 부적절한 정보를 얻게 되고 결과적으로 낭비를 초래하게 된다. 따라서 적절한 표본의 크기를 결정하는 것은 조사설계의 과정에서 매우 중요하다. 이에 본 연구는 선행연구에서 지적된 문제점들을 해결하는 후속과제의 성격을 가지며, 표본규모 산정방법에 대하여 보다 구체적이고 객관적인 검토기준을 마련하고자 한다. 또한 선행연구에서 미처 검토되지 못하였던 독립표본체계 하에서의 표본관리에 대한 방안을 제시하고자 한다.

2. 연구목적 및 내용

본 연구의 목적은 집세조사를 독립표본으로 설계하였을 경우에 필요한 표본규모의 적정 수준을 결정하고, 독립표본체계 하에서 검토되어야 할 표본 관리방안을 제시하는 것이다.

2절에서는 집세의 계약갱신과 관련한 특성들을 파악하여 집세조사에서 최소한으로 유지되어야 할 표본의 크기를 결정하고, 도시별 적절한 수준의 표본규모를 산정하였다. 3절에서는 신축 및 멸실 주택과 아파트 단지에 대한 표본선정 및 교체방법, 조사주기의 조정, 독립표본체계 하에서 거쳐야 할 표본관리방안을 제시한다. 마지막으로 4절에서는 이 연구를 통해서 확인한 시사점 및 결론을 제시하고자 한다. 본 연구는 신뢰성 있는 집세지수 작성에 필요한 적절한 표본의 크기를 결정하고, 제한된 조사여건 하에서 효율적으로 표본을 선정·관리하는 방안을 제시함으로써 결과적으로 집세지수의 정확도 향상에 기여할 것이라 기대한다.

제2절 표본규모 산정

1. 유효표본의 크기 결정

집세지수는 조사시점의 집세가 기준 월과 비교하여 얼마만큼 변동하였는지를 수치적으로 비교하고자 지수로 작성한 것이다. 임차가구는 임대차계약에 의해 정해진 기간 동안 집세를 지불하는데, 통상 전세는 2년 단위로, 월세는 1년 단위로 계약을 맺고 있다. 계약기간이 끝나게 되면, 재계약을 맺을 수도 있고 전출·입이 발생할 수도 있는데 이때 집세가 조정될 수 있다. 즉, 한번의 계약으로 맺어진 집세 수준은 정해진 기간동안 유지가 되며, 재계약이나 전출·입과 같이 거주주택의 임차가구에 대한 계약갱신이 발생하기 전까지는 변동이 발생하지 않는다.

따라서 집세조사에서는 전체적으로 약 1만여 가구의 표본가구를 조사하고 있지만, 조사대상 월에 계약갱신이 발생한 표본가구에 대해서만 집세의 변동 여부를 관측할 수 있다. 결국 집세표본으로 선정된 1만여 가구 중에서 해당 월에 계약갱신이 도래한 일부의 표본가구만이 집세지수에 영향을 줄 수 있는 유효한 표본인 셈이다. 만일 어느 지역에서 해당 월에 계약갱신이 일어난 가구가 한 가구도 없거나 매우 적다면, 그 지역에 표본의 크기가 크더라도 집세 변동을 충분히 관측할 수 없기 때문에 신뢰성 있는 지수를 작성할 수 없게 된다. 따라서 집세조사에서 전체 표본의 크기를 결정할 때 계약갱신 가구의 크기를 먼저 파악하는 것은 매우 중요하다. 안정적인 추정값을 제공하기 위해서는 이러한 유효표본은 매월 정해진 일정 수준 이상은 확보·유지되어야 하기 때문에, 최소한 유지되어야 하는 유효표본 크기를 결정해야 한다.

우선 현행 집세조사 시스템 하에서 매월 계약갱신이 이뤄지는 가구의 현황을 살펴봄으로써 집세의 변동을 파악할 수 있는 유효표본이 어느 정도인지를 가늠해보고자 한다. 2006년 1월부터 2008년 12월까지의 집세조사 자료를 이용하여 전체 표본 중 계약갱신이 발생한 표본의 비율을 분석해보았다. <표 3-1>에는 16개 시도별로 매월 평균적으로 조사되고 있는 표본 수와 이 표본들 중에서 계약갱신이 발생한 평균 표본수를 주택유형별로 구해보았다. 일반주택의 경우 전국적으로 매월 약 6,499개의 표본이 평균적으로 조사되고 있으며, 약 4.94%에 해당하는 320여 가구가 계약갱신이 발생한 유효표본임을 알 수 있다. 또한 이렇게 계약갱신이 있었던 가구 중에 금액변동이 있는 경우가 약 101건으로 전체 표본의 약 1.6%, 변동되지 않은 경우가 약 3.3%를 차지하고 있었다. 다시 말해, 472건의 유효표본 중 약 33%의 가구만이 계약을 갱신하면서 집세 수준이 변하고 있음을 확인할 수 있다. 아파트도 전체 표본 중 계약갱신 비율은 약 4.96%로 일반주택과 비슷한 수준이나, 금액변동 비율은 1.82%로 일반주택보다 다소 높다. 즉, 가구에 계약갱신

이 발생했을 때, 아파트가 일반주택보다 집세 수준이 변할 가능성이 더 높은 것으로 짐작해 볼 수 있다.

16개 시도별 계약갱신 비율은 다소 차이가 있는데, 일반주택의 계약갱신 비율은 특·광역시 중에서는 부산이 3.95%로 가장 낮고, 금액변동 가구의 비율도 1.0%로 가장 낮다. 그리고 9개 도 중에서는 전라북도가 계약갱신 비율이 3.85, 금액변동 비율은 0.86%로 가장 낮다. 아파트의 경우는 울산의 계약갱신 표본비율이 16개 시도 중 가장 낮고, 금액변동이 있는 비율이 0.99%로 전국 수준에 비해 매우 낮다. 정리해보면, 1만여 가구를 조사하더라도 매월 약 5%만이 계약갱신이 발생하여 집세변동을 관측 가능한 유효 표본의 수는 약 500가구가 됨을 알 수 있다. 또한 유효표본 500가구 중 집세의 금액이 변하는 경우가 33%로 평균적으로 매월 약 150여 가구가 금액크기에 상관없이 등락하고 있음을 확인할 수 있었다.

〈표 3-1〉 시도별 집세표본의 계약갱신 비율

(단위: 가구, %)

시도	일반주택				아파트			
	표본수	계약갱신 비율		표본수	계약갱신 비율		금액 미변동	
		금액 변동	금액 미변동		금액 변동	금액 미변동		
전국	6,499	4.94	1.55	3.39	3,040	4.96	1.82	3.13
서울	1,489	4.69	1.80	2.88	524	4.52	1.94	2.58
부산	567	3.95	1.00	2.94	258	4.84	1.63	3.22
대구	527	5.84	1.64	4.20	113	4.59	1.28	3.30
인천	376	4.50	1.49	3.01	205	4.57	1.75	2.82
광주	330	4.97	1.36	3.61	178	5.59	1.89	3.70
대전	337	5.32	1.45	3.88	222	4.85	1.41	3.44
울산	291	4.97	1.93	3.04	114	3.68	0.99	2.69
경기	658	4.86	1.86	2.99	292	4.07	1.84	2.23
강원	264	4.85	1.55	3.31	189	5.62	2.87	2.75
충북	175	4.95	1.58	3.37	158	4.83	1.49	3.35
충남	176	6.16	1.98	4.19	175	5.83	1.60	4.23
전북	176	3.80	0.86	2.94	92	4.21	0.86	3.35
전남	154	5.00	1.14	3.86	161	6.68	2.74	3.93
경북	460	5.77	1.03	4.73	215	5.89	2.09	3.80
경남	338	4.83	1.35	3.49	104	4.00	1.06	2.95
제주	181	6.62	2.35	4.28	42	7.75	3.95	3.80

주) 2006년 1월부터 2008년 12월까지 집세조사 자료의 36개월 평균 수준



이와 같이 전체 표본 중 유효표본의 비율도 중요하지만, 유효표본 가구들의 응답결과가 추정값에 바로 영향을 미치기 때문에 매월 계약갱신이 발생하고 있는 표본가구의 수도 중요하다. 예를 들어, 어느 도시에 유효표본이 한 가구밖에 없어서 그 한 가구가 도시의 변동 전체를 설명하게 된다면, 그 결과는 신뢰할 수 없을 것이다. 따라서 해당 도시에 매월 일정수준 이상의 가구들이 계약갱신이 일어날 수 있도록 유효표본의 수를 안정적으로 확보하고, 유지해 나갈 필요가 있다.

집세조사의 지난 3년간 발생한 시도별 계약갱신 수를 <표 3-2>와 같이 정리해 보았다. 표본의 크기가 다른 시도에 비해 월등히 많은 서울을 제외한 나머지 시도의 일반주택의 월별 평균 계약갱신 수는 16.8건이며, 아파트의 평균 계약갱신 수는 약 8.5건이다. 일반주택의 경우 충북, 전북, 전남, 제주도는 평균 계약갱신 수가 10건 미만으로 전체시도의 평균 수준보다 낮다. 특히 아파트의 경우 5건 이하로 계약갱신이 발생하는 시도가 많은데, 대구가 5건, 울산, 전북, 경남이 4건, 제주가 3건 정도의 계약갱신이 매월 평균 발생하고 있다.

<표 3-2> 시도별 집세표본의 계약갱신 수

(단위: 가구, %)

시도	합계		일반주택				아파트			
	표본수	계약 갱신수	표본수	계약 갱신수	전입 비중	재계약 비중	표본수	계약 갱신수	전입 비중	재계약 비중
서울	2,013	94	1,489	70	43.6	56.4	524	24	32.5	67.5
부산	825	34	567	22	28.7	71.3	258	12	30.0	70.0
대구	674	36	527	31	25.9	74.1	113	5	29.3	70.7
인천	581	26	376	17	43.0	57.0	205	9	38.2	61.8
광주	508	26	330	16	29.7	70.3	178	10	14.0	86.0
대전	559	29	337	18	43.6	56.4	222	11	27.6	72.4
울산	405	18	291	14	34.8	65.2	114	4	27.5	72.5
경기	950	44	658	32	39.9	60.1	292	12	38.1	61.9
강원	453	24	264	13	47.2	52.8	189	11	24.3	75.7
충북	333	17	175	9	40.6	59.4	158	8	30.4	69.6
충남	351	21	176	11	42.2	57.8	175	10	29.1	70.9
전북	268	11	176	7	23.3	76.7	92	4	30.9	69.1
전남	315	19	154	8	30.1	69.9	161	11	22.4	77.6
경북	675	39	460	27	23.3	76.7	215	13	14.4	85.6
경남	442	20	338	16	36.3	63.7	104	4	39.5	60.5
제주	223	15	181	12	28.5	71.5	42	3	20.4	79.6

주) 2006년 1월부터 2008년 12월까지 집세조사 자료의 36개월 평균 수준



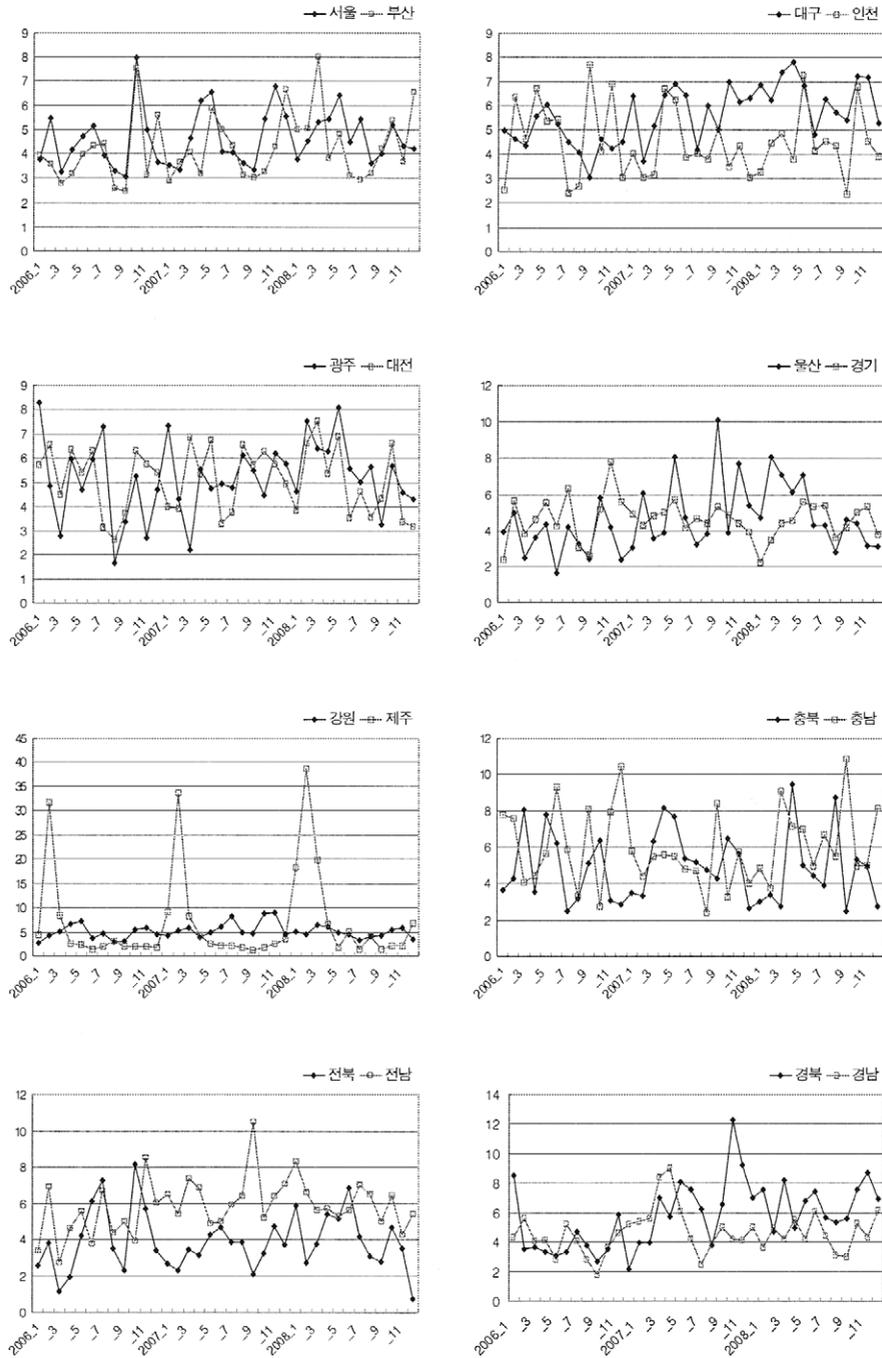
앞에서 설명한 바와 같이 계약갱신은 표본가구의 전출·입이 발생했거나, 표본가구의 임차기간이 만료되어 재계약일이 도래했을 경우에 이뤄진다. 계약갱신 수를 이와 같은 발생원인에 의해 구분해 볼 수 있다. 계약갱신 중 표본가구의 전출·입에 의한 비중이 낮은 지역은 일반주택의 경우 부산, 대구, 광주, 전북, 경북, 제주가 30% 미만이며, 아파트는 광주와 경북이 20%를 밑돌고 있다. 즉, 부산, 대구, 광주, 전북, 경북, 제주의 경우 타 시도에 비하여 전출·입이 빈번하지 않은 것으로 보인다.

2006년부터 2008년까지 3년 동안의 월별 평균을 구한 결과를 살펴보았는데, 매월마다 그 수준에 있어서는 차이가 존재한다. 특히 우리나라는 봄과 가을에 전출·입이 다른 계절에 비해 빈번하게 발생하여 이사철이라는 계절적 요인이 많은 영향을 미칠 것으로 예상된다. 월별 계약갱신 발생 추이를 [그림 3-1]과 같이 도시별로 살펴보았다. 표본가구의 수가 많은 서울에서 이사철의 효과가 뚜렷하게 보이는데, 3~5월, 9~10월 사이에 계약갱신이 많이 발생한다. 대부분의 시도에서 이사를 많이 하지 않는 7~8월과 12~1월 사이에도 계약갱신 비율이 약 2% 이상 유지되고 있는 것으로 나타났다. 특이한 점은 제주도의 경우 1~3월 사이에 계약갱신 비율이 매우 높으며, 특히 2월에는 전체 표본의 30%가 넘는 가구가 계약갱신을 맞고 있는 것으로 나타났다. 이러한 현상은 대한에서 입춘까지 신구간이라는 정해진 기간 동안에 이사를 하던 제주도만의 독특한 이사풍습 때문인 것으로 보인다. 이상으로 지역별로 계약갱신 발생과 관련된 특성들을 살펴보았다.

다음으로 계약갱신이 추정값에 미치는 영향을 살펴봄으로써 안정적인 추정값을 얻기 위해 최소한 요구되는 표본의 크기가 어느 정도가 될 것인지를 결정하고자 한다. 이를 위해 2006년 1월부터 2008년 12월까지의 36개월 동안의 16개 시도와 38개 도시별로 월별 변동률 평균에 대한 CV를 구해보았다. 집세조사의 경우 기존의 설계로부터 추출률에 대한 정보를 얻을 수 없는 상황이다. 이에 따라 집세 표본이 각 도시별로 단순임의추출(SRS)에 의해 추출되었다고 가정하고 집세의 평균변동률에 대한 CV를 구하였다. 추정 결과는 <부록 1>과 <부록 2>에 그래프로 첨부하였다. 지역 간 객관적인 비교를 위해 y 축의 최대값을 0.5로 고정시켰으며, 0.5를 초과할 경우 y 축의 범위 밖에 해당 월의 값이 존재하는 것으로 해석하면 되며, 변동률은 아래와 같이 구한다.

$$\text{변동률} = \frac{y_{ij,t}}{y_{ij,t-1}}$$

여기서 $y_{ij,t}$ 는 i 지역 j 가구의 t 시점(금월) 집세이고, $y_{ij,t-1}$ 는 i 지역 j 가구의 $t-1$ 시점(전월) 집세이다.



[그림 3-1] 시도별 표본가구 내 계약갱신 비율의 월별 추이



16개 시도별 일반주택과 아파트 부문에 대한 지수를 구하기 위해 필요한 표본의 수를 구한 다음에 각 도시별 지수를 작성하기 위해 유지되어야 하는 표본의 수를 결정하도록 한다. 각 지역에 대한 자세한 내용은 부록의 그림을 참고하고, 여기서는 결과만을 요약해본다.

우선 16개 시도의 주택유형별 평균변동률의 CV는 대체로 0.5% 이하임을 <부록 1>에서 확인해 볼 수 있다. 서울은 일반주택과 아파트 모두 0.2% 수준 이하이고, 월별로도 작게 움직이며 다른 시도에 비하여 매우 안정적이다. 부산, 대전, 경기, 경북도 두 주택유형별 CV가 큰 진폭없이 비교적 안정적이다. 그에 비하여 대구, 울산, 경남의 아파트의 CV와 충북, 충남, 전북, 전남, 제주의 일반주택과 아파트 CV가 월별로 큰 폭의 움직임을 보이며, 다소 불안정적인 모습을 보이고 있다. 이와 같이 안정적인 지역과 불안정한 지역은 평균 계약갱신 수에 있어서도 차이를 보이고 있다. 일반주택에 대해서 CV가 안정적인 부산, 대전, 경기, 경북 등의 지역은 평균 계약갱신 수가 최소 18건 이상인데 비해, 불안정한 충북, 충남, 전북, 전남 제주 지역의 월별 평균 계약갱신 수는 약 10건 정도이다. 또한 아파트에 대해서도 월별 CV가 비교적 안정적인 부산, 대구, 경기, 경북 지역은 월별 평균 10건 이상의 계약갱신 수가 발생하고 있었다. 이와는 대조적으로 불안정한 대구, 울산, 전북, 경남, 제주 지역은 5건 이하의 계약갱신이 일어나고 있음을 알 수 있었다.

마찬가지로 38개 도시별 평균 변동률의 CV와도 계약갱신 수와의 관계를 <부록 2>에서 도시별로 비교해볼 수 있다. 도시에 대한 지수 작성 범위는 주택유형을 구분하지 않은 단일지수로 작성하기 때문에 도시전체에 대해 평균변동률을 구하였다. 서울은 매월 약 60건 이상의 계약갱신이 발생하며, 집세의 평균변동률은 약 0.1% 내외의 수준이다. 또한 광역시의 평균변동률의 월별 CV도 0.1~0.2% 수준으로 매우 안정적인 형태를 보인다. 광역시들은 매월 약 15~30건 정도의 계약갱신이 발생하고 있다. 나머지 31개의 시는 월별 평균 계약갱신 수가 차이가 많이 난다. 청주, 천안, 구미에서 비교적 안정적인 월별 CV의 형태를 보이고 있는데, 이들 지역의 평균 계약갱신 수는 10건 이상이다. 다른 도시들은 대부분 월별 CV가 크게 진동하고 있는데, 5건 안팎의 계약갱신 수를 보이고 있다. 그리고 보령, 서산, 군산, 남원 등의 일부 지역에서는 계약갱신이 일어나지 않는 달이 빈번하여 변동률의 CV가 0인 개월 수가 많다.

지금까지 살펴본 내용을 종합해보면, 매월 신뢰성이 있는 지수를 안정적으로 제공하기 위해선 일정 수 이상의 계약갱신이 발생해야 한다는 것을 미루어 짐작해볼 수 있다. 이론상으로는 계약갱신이 많이 발생하면 금액변동이 발생할 가능성이 커지고, 금액변동이 발생을 많이 하면 결과적으로 변동률의 CV가 커지게 되는 게 보통의 경우이다. 그러나 CV가 월별로 매우 극심하게 차이가 나게 되는 데에는 표본의 규모가 최소한 확보되어야 하는 수준만큼 유지되지 못하고 있는 것이 원인이라고 판단된다. 따라서 월별 CV가

비교적 안정적인 지역에서 발생하고 있는 유효표본의 크기만큼을 다른 도시에서도 유효 표본으로 확보해야만 월별 CV가 안정적이 될 것이라고 기대할 수 있다. 이에 지역별로 유지해야 하는 유효표본의 크기에 대한 기준을 과거 집세조사 경험에 근거하여 실증적으로 찾았다. 시도 수준에서는 일반주택은 최소 18건 이상의 계약갱신이 발생하는 시도의 CV가 안정적이었기 때문에 유효표본을 최소 20건은 보유해야 할 것으로 판단되었다. 시도의 아파트는 계약갱신 수가 최소 10건 이상인 시도에서 CV가 안정적으로 보였으므로, 최소 유효표본 수를 15건으로 정하였다. 다음으로 도시별로는 최소 10건 이상의 계약갱신이 발생하여야 할 것으로 보인다.

이러한 유효표본을 확보하기 위해 요구되는 최소 표본수는 지역별 계약갱신 비율을 이용하여 구하였다. 목표 계약갱신 수를 달리 설정하였을 경우 필요한 최소 표본수를 산정해 보았다. <표 3-3>은 시도별로 구해보았는데, 일반주택의 목표 유효표본 수를 15개로 잡았을 때, 총 6,470가구의 표본이 필요하다. 또한 아파트의 목표 유효표본 수를 15건으로 설정하면, 총 4,890가구가 필요하다.

<표 3-3> 목표계약갱신 수에 의한 시도별 최소 유지 표본수

시도	일반주택					아파트				
	현행		최소요구표본수			현행		최소요구표본수		
	표본수	계약 갱신수	목표계약갱신수			표본수	계약 갱신수	목표계약갱신수		
			15	20	30			10	15	20
서울	1,489	69.7	320	427	641	524	23.8	220	331	441
부산	567	22.3	382	509	764	258	12.4	208	312	416
대구	527	30.8	257	343	514	113	5.1	221	332	443
인천	376	16.9	335	446	669	205	9.4	218	328	437
광주	330	16.3	304	405	607	178	10.1	176	264	353
대전	337	17.9	282	376	564	222	10.8	206	309	412
울산	291	14.4	304	406	608	114	4.3	269	403	537
경기	658	31.8	310	413	620	292	11.9	244	366	488
강원	264	12.8	308	411	616	189	10.6	177	266	355
충북	175	8.6	305	407	611	158	7.7	206	309	411
충남	176	10.9	243	325	487	175	10.2	171	257	342
전북	176	6.7	395	526	790	92	3.8	242	363	485
전남	154	7.7	300	401	601	161	10.8	149	223	298
경북	460	26.7	259	345	518	215	12.5	172	258	344
경남	338	16.2	313	417	626	104	4.2	246	369	493
제주	181	11.6	234	312	469	42	3.1	133	199	265
합계	6,499	-	4,852	6,470	9,705	3,040	-	3,260	4,890	6,519

또한 38개 도시에서 시도의 결과와 중복되는 특·광역시를 제외하고, 나머지 31개 도시에 대한 최소 표본수가 <표 3-4>에 정리되어 있다. 도시별 목표 계약갱신 수를 10건으로 설정할 경우 전체적으로 현행 약 4,000가구에서 약 2,200가구가 증가한 6,227가구를 확보·유지될 필요가 있다.

<표 3-4> 목표계약갱신 수에 의한 도시별 최소 유지 표본수

시도	현행			최소요구표본수		
	표본수	계약갱신		목표계약갱신수		
		건수	비율(%)	5	10	15
수 원	149	7	4.6	108	215	323
성 남	136	7	4.8	105	209	314
의정부	89	5	5.4	93	185	278
안 양	151	7	4.6	108	215	323
부 천	119	5	4.3	117	233	350
안 산	95	4	4.3	116	232	348
고 양	211	9	4.4	113	226	338
춘 천	145	7	5.1	98	195	293
원 주	173	8	4.7	105	211	316
강 릉	135	8	5.8	86	172	257
청 주	209	10	4.8	103	207	310
충 주	125	6	5.0	100	201	301
천 안	207	13	6.2	81	162	243
보 령	51	4	7.2	70	139	209
서 산	93	5	5.0	101	202	303
전 주	123	5	4.3	116	232	348
군 산	89	3	3.2	155	309	464
남 원	55	2	4.1	121	243	364
목 포	83	4	5.1	98	197	295
여 수	129	8	6.1	82	163	245
순 천	103	6	6.2	81	162	243
포 향	117	6	5.4	92	184	276
경 주	173	11	6.5	76	153	229
안 동	109	5	4.8	103	207	310
구 미	275	16	5.9	85	170	255
창 원	130	7	5.1	98	197	295
마 산	107	4	3.8	130	260	390
진 주	111	6	5.5	91	182	274
김 해	95	4	3.9	129	259	388
제 주	150	10	6.7	74	149	223
서귀포	73	5	6.4	78	156	234
합 계	4,007	208	—	3,113	6,227	9,340

2. 도시별 표본의 크기 결정

선행연구에서는 집세조사의 평균변동률에 대한 CV를 표본크기를 결정하기 위한 목표변수로 하였을 때 발생할 수 있는 여러 가지 문제를 지적하였다. 첫째, 현행 집세 최소유지표본을 확보하지 못하였음에도 불구하고 과거 표본크기에 의존하게 된다. 둘째, 평균변동률의 상대표준오차가 매우 낮은 도시의 표본크기가 신뢰할 수 없을 만큼 작게 계산된다. 셋째, 변동률이라는 것은 특정시점이나 외부효과의 개입에 의해 극심하게 변동할 수 있다. 또한 현재 활용 가능한 자료들을 통해서는 조사 모집단의 집세 변동률을 구할 수 없다. 대신 집세의 평균수준에 대해서는 인구주택총조사 10% 표본자료를 이용하면 조사 모집단의 분포 파악이 용이하고, 시점에 덜 민감하다. 조사설계의 효율성과 표본크기 산정방법의 안정성 측면을 고려하여 변동률 대신 집세 수준을 표본크기 결정을 위한 기준변수로 삼도록 하겠다.

표본 크기를 결정하기 위해서는 먼저 새로운 설계에서 달성하고자 하는 목표정도의 수준을 설정해야 한다. 즉, 집세 수준의 평균에 대한 상대표준오차의 수준을 정해야 하는데, 과거 집세조사의 경험을 참고해서 그 수준을 결정하고자 한다. 집세조사의 경우 기존의 설계로부터 추출률에 대한 정보를 얻을 수 없기 때문에, 집세 표본을 각 도시별로 단순임의추출(SRS)에 의해 추출되었다고 가정하고 집세의 평균에 대한 CV를 구하였다.

그러나 새로운 표본설계에서 필요한 표본의 크기를 결정하기 위해서는 다음과 같은 설계효과(design effect)가 먼저 고려되어야 한다. 설계효과는 두 가지 목적으로 사용된다. 하나는 표본의 크기를 결정하는 데 필요하고, 또 하나는 표본의 효율을 계산하는 데 필요하다. 즉, 목표수준 설정의 근거가 되는 집세 평균의 CV는 부차모집단별 단순임의추출을 가정하였다. 그러나 새로운 표본설계에서는 층화2단집락추출을 기반으로 하기 때문에, 층화와 집락효과가 고려되어야 한다. 그러므로 서로 다른 설계기반 하에서 목표정도의 비교와 표본크기 결정을 위해 설계효과를 고려해야 한다.

이러한 설계효과를 확인하기 위해 2005년 인구주택총조사 10% 표본자료에서 서울의 일반주택 자료를 이용하여 시뮬레이션 해보았다. 새로운 설계의 방법을 그대로 이용하여 층화2단집락추출을 수행한 후, 평균과 그의 분산을 추정하였다. 또한 동일한 표본의 단순임의추출을 가정한 경우에 대한 평균과 각각의 분산을 구해보았다. <표 3-5>와 같이 표본설계방법 하에서 얻게 되는 추정값의 분산을 비교함으로써 설계효과를 비교할 수 있다. 평균은 층화2단집락추출한 경우 4,473만원, 단순임의추출을 가정한 경우 4,469만원으로 추정되었다. 추정된 평균에 대한 층화2단집락추출의 분산이 11,644로 단순임의추출의 4,654보다 크게 추정된다. 표본틀에서 가구를 바로 선정하는 단순임의추출방법이 조사구를 먼저 추출하고, 추출된 조사구에서 다시 가구를 선정하게 되는 2단 집락

추출보다 더 효율적임을 알 수 있다. 단순임의표본의 분산에 대한 층화2단집락표본의 분산의 비로 설계효과를 구해보면, 2.5로, 층화2단집락추출에 필요한 표본크기가 n 이라면, 단순임의추출에 의하여 주어진 허용오차를 만족시키는 n_0 를 구하여 2.5배 하면 된다. 즉 $n = n_0(2.5)$ 가 된다.

〈표 3-5〉 표본설계방법에 따른 설계효과 추정

(단위 : 만원)

설계구분	Mean	Var	CV(%)	UpperCLMean	설계효과
층화2단집락추출	4,473	11,644	2.41	4,688	2.5
단순임의추출	4,469	4,654	1.53	4,603	1

부차모집단별로 추정값의 정도가 요구되면, 표본크기도 이에 따라서 부차모집단별로 계산하여 전체의 표본크기를 정해야 한다. 부차모집단별 단순임의추출을 가정하였을 때 평균 추정을 위한 표본크기(n_0)는 아래와 같은 식에 의해 구해진다.

$$n_0 = \frac{r^2}{CV^2 + \frac{r^2}{N}}, \quad r = \frac{\sigma}{\mu}$$

여기서 N 은 전체 모집단 크기, μ 는 모집단의 평균, σ^2 은 모집단의 분산, CV 는 상대허용오차이다. 부차모집단에 대한 표본크기가 결정되었으면, 각 층별로 층 내 임차가구 수에 비례하도록 배분하면 된다. 이와 같은 방식으로 n_0 를 구하고, 앞서 설명한 층화효과를 고려하여 n 을 구하였다.

새로운 설계에서 전국지수 뿐만 아니라, 16개 시도별 주택유형별 지수와 38개 도시별 지수를 제공하고자 하므로 이 모두를 만족시켜 주기 위해 공표범위별로 표본크기를 계산하여 그 중에서 큰 것을 택해야 한다. 목표정도를 정하기 위해 가장 최근에 조사된 2008년 1월부터 12월까지의 집세조사 결과를 사용하여 전국, 16개 시도, 38개 도시 수준에서 월별 집세 평균에 대한 CV 를 구해보았다.

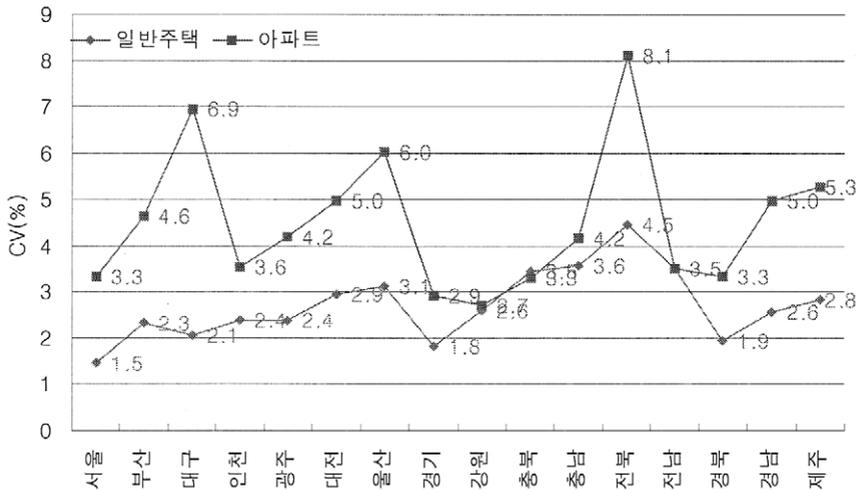
먼저 전 도시에 대해서 주택유형별 집세 평균에 대한 월별 CV 를 구하여 12개월의 평균 수준을 측정하면, 일반주택은 0.79%, 아파트는 1.57% 수준이었다. 이를 근거로 일반주택 0.8~1.5%, 아파트 1.6~3.0% 수준의 목표정도를 달성하기 위해 필요한 표본의 크기를 구하여, 이를 16개 시도에 임차가구의 수에 대해서 비례배분하였다. <표 3-6>이 그 결과표이며, 이때의 표본크기는 앞서 설명한 설계효과를 감안하여 계산하였다. 전 도시



일반주택에 대한 목표 CV를 0.8%로 설정할 경우, 12,987개의 표본이 필요하게 된다. 목표정도를 약 1.0%로 유지하게 된다면, 전국적으로 약 8,903개의 일반주택거주 임차가구의 표본 확보가 요구된다. 이러한 8,903개의 표본을 임차가구 수에 비례하게 시도별로 배분하면 서울에 전체의 약 39%를 차지하는 표본 수가 배분되고, 그 다음으로 경기도에 약 16% 정도의 표본이 배분된다. 또한 아파트 목표정도를 1.6%로 설정하면, 8,323개의 표본이 필요하고, 2.0%로 설정을 하면, 5,371개의 표본이 필요하다. 이와 같이 목표정도에서 요구되는 표본크기를 구하게 되면, 예산과 인력과 같은 현실적인 측면을 종합적으로 고려하여 주어진 여건 하에서 달성하고자 하는 목표정도를 조정함으로써 최종적으로 표본 크기를 정하게 된다.

〈표 3-6〉 전국 주택유형별 목표정도에 따른 표본 크기

전국	일반주택				아파트			
	목표정도 CV(%)				목표정도 CV(%)			
	0.8	1.0	1.2	1.5	1.6	2.0	2.4	3.0
합계	12,987	8,903	6,432	4,255	8,323	5,371	3,747	2,409
서울	5,023	3,443	2,488	1,646	2,284	1,473	1,028	660
부산	1,285	881	636	421	683	441	308	198
대구	841	576	416	276	497	321	224	144
인천	810	555	401	265	498	321	224	144
광주	319	219	158	105	476	307	214	138
대전	506	347	251	166	390	252	175	113
울산	312	214	155	102	179	115	81	52
경기	2,122	1,455	1,051	695	1,235	797	556	357
강원	225	154	111	74	281	182	127	81
충북	251	172	124	82	255	165	115	74
충남	149	102	74	49	235	152	106	68
전북	197	135	97	64	308	198	138	89
전남	102	70	51	34	327	211	147	95
경북	270	185	134	88	273	176	123	79
경남	513	352	254	168	351	227	158	102
제주	62	43	31	20	51	33	23	15



[그림 3-2] 시도별 주택유형별 집세 평균의 CV(%)

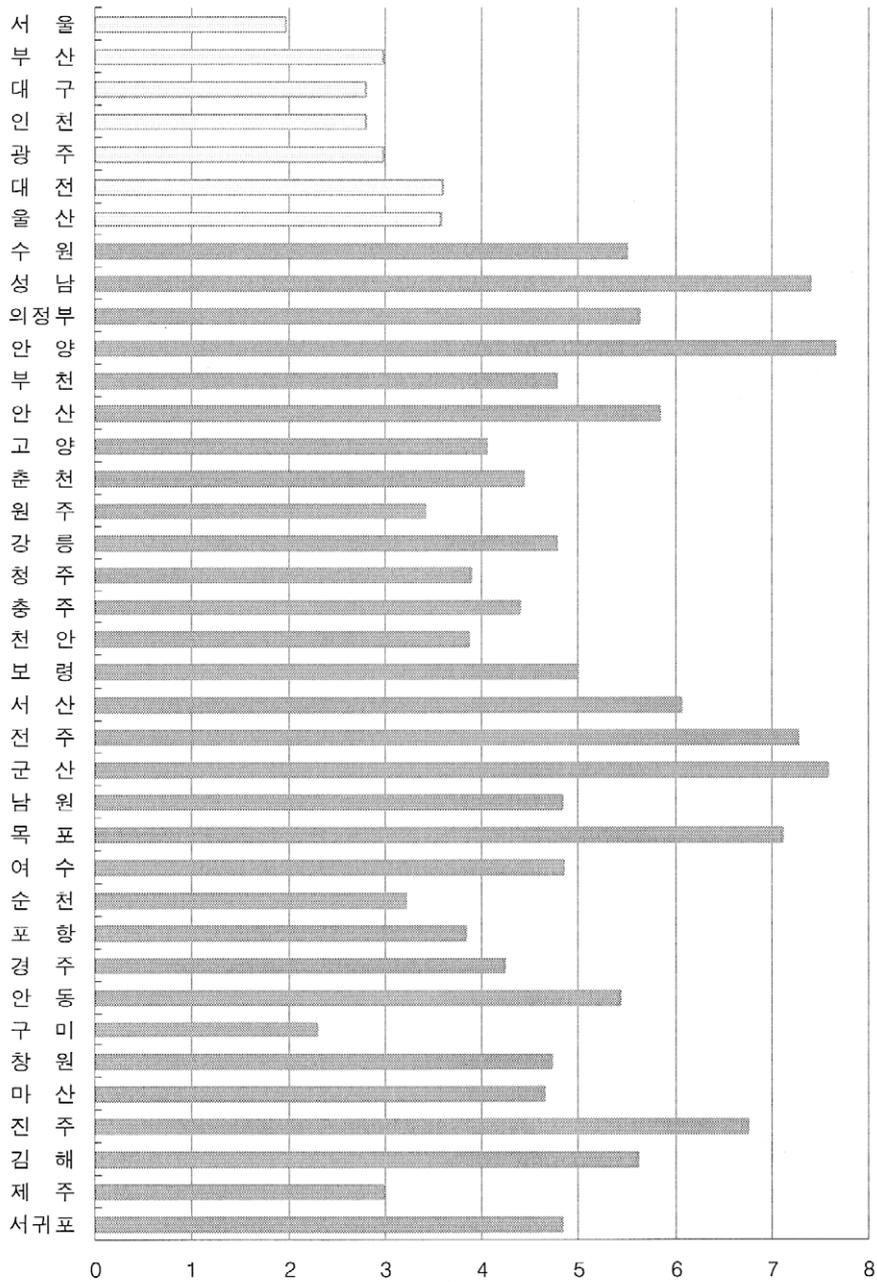
다음으로 16개 시도 수준에서의 표본크기를 살펴본다. [그림 3-2]는 16개 시도별 주택유형별 집세 평균에 대한 월별 CV를 구하여 12개월의 평균 수준을 측정하는 것이다. 먼저 일반주택 집세 평균에 대한 CV를 시도별로 비교해보면, 서울이 1.5%로 가장 낮고, 그 다음으로 경기와 경북이 각각 1.8, 1.9%로 낮다. 전북이 4.5%로 가장 높고, 울산, 충북, 충남, 전남이 3%를 상회한다. 그 외의 광역시와 도는 2% 내외 수준이었다. 다음으로 아파트 집세 평균의 CV를 보면, 강원 2.7, 경기 2.9%가 가장 낮은 수준이며, 대체로 3~5%의 수준을 보이고 있다. 전북이 8.1%로 가장 높고, 대구, 울산이 6%를 넘고 있다.

이와 같은 CV의 수준을 고려하여 시도의 일반주택의 경우 CV를 2.0~3.5%, 아파트의 경우 3~6% 수준의 목표정도를 달성하기 위해 필요한 표본의 크기를 <표 3-7>과 같이 구하였다. 여기서도 역시 앞서 설명한 설계효과를 감안하여 계산하였다. 일반주택의 평균 집세에 대한 시도 CV를 2.0%로 설정하면, 총 24,128가구의 표본이 필요하게 되고, 2.5%인 경우 15,991가구, 3.0%면 11,332가구가 필요하다. 아파트의 CV를 3.0%로 설정하면, 17,759개의 표본이 요구되며, 4.0% 이면 10,336가구, 5.0%를 유지하기 위해서는 6,725가구를 조사해야 한다. 16개 시도 중 9개의 도(道)에는 2~4개의 도시들이 포함되어 있다. 따라서 시도의 표본크기는 도시별 표본크기의 합과 연관되어 있으므로 도시별 표본의 크기를 구한 후 재조정하는 과정을 거치게 된다.

〈표 3-7〉 시도별 주택유형별 목표정도에 따른 표본 크기

시도	일반주택				아파트			
	목표정도 CV(%)				목표정도 CV(%)			
	2.0	2.5	3.0	3.5	3.0	4.0	5.0	6.0
서울	1,928	1,237	860	632	1,774	1,006	646	449
부산	1,838	1,188	829	611	1,538	885	572	400
대구	1,309	846	591	436	1,698	989	643	451
인천	1,198	774	541	398	996	571	369	258
광주	1,364	898	634	470	915	525	339	236
대전	1,489	972	682	505	1,325	771	501	351
울산	1,673	1,110	787	585	1,217	734	486	344
경기	1,895	1,220	850	626	1,109	629	404	281
강원	1,108	733	518	385	934	543	353	247
충북	1,508	1,005	714	532	1,158	682	446	313
충남	1,226	830	595	446	1,018	598	391	275
전북	1,920	1,314	948	714	991	576	374	262
전남	1,593	1,137	842	645	609	349	225	157
경북	1,643	1,096	779	581	981	572	373	261
경남	1,627	1,064	747	553	839	483	313	218
제주	809	567	415	315	657	423	290	210
합계	24,128	15,991	11,332	8,434	17,759	10,336	6,725	4,713

다음으로 38개 도시 수준에서의 표본크기를 살펴본다. 우선 38개 도시 수준에서의 평균 집세의 CV현황을 살펴보고자, 2008년 월별 평균 집세의 월별 CV를 구하여 12개월의 평균을 측정한 후 그래프로 [그림 3-3]과 같이 그려보았다. 서울은 약 2%, 부산, 대구, 인천, 광주는 3% 이하, 울산과 대전은 약 3.5% 수준이었다. 특·광역시 제외 나머지 31개 도시들은 3~6%의 수준을 보이고, 성남, 안양, 전주, 군산, 목포의 CV가 7%를 넘어 서고 있다. 7개 특·광역시는 개별 도시이긴 하지만 도시의 크기가 시도 수준이고, 앞의 시도의 표본크기를 결정하는 과정에서 먼저 고려되었기 때문에 도시 수준의 표본크기 계산에서는 제외하였다.



[그림 3-3] 도시별 집세 평균의 CV(%)



〈표 3-8〉 도시별 주택유형별 목표정도에 따른 표본 크기

도시	목표정도 CV(%)				
	4.0	5.0	6.0	7.0	8.0
수 원	743	479	334	246	189
성 남	1,336	866	606	447	343
의정부	451	293	205	151	116
안 양	874	569	399	295	227
부 천	758	491	343	253	194
안 산	609	393	274	202	155
고 양	678	438	306	226	173
춘 천	617	405	285	211	163
원 주	500	326	229	169	130
강 릉	438	287	202	150	115
청 주	691	448	313	231	177
충 주	527	350	248	184	142
천 안	643	418	293	216	166
보 령	511	388	300	236	190
서 산	380	256	183	137	106
전 주	625	406	284	210	161
군 산	722	480	341	253	196
남 원	480	335	244	185	145
목 포	543	358	252	187	144
여 수	430	282	198	147	113
순 천	335	218	153	113	87
포 향	685	450	317	235	181
경 주	642	431	307	229	177
안 동	384	257	184	137	106
구 미	425	276	194	143	110
창 원	722	470	330	244	187
마 산	644	421	296	219	168
진 주	675	444	313	232	179
김 해	479	311	218	161	124
제 주	423	279	197	146	113
서귀포	321	234	176	136	108
합 계	18,291	12,059	8,524	6,331	4,885



이와 같은 CV의 수준을 고려하여 각 목표정도에 의해 필요한 도시별 표본의 크기를 구해 보았다. 이때의 표본 크기는 각 도시별로 단순임의추출에 의해 가구를 추출한다고 가정하고 표본 수를 계산하였다. <표 3-8>을 보면, 도시별 평균 집세의 CV를 4.0%로 설정하면, 31개 도시에 대해 총 18,291가구의 표본이 필요하며, 6.0%를 설정하게 되면, 8,524가구를 조사해야 한다.

이와 같이 구해진 시도별, 도시별 표본의 크기는 앞서 언급된 최소유지 표본의 크기와 비교해야 한다. 만일 도시별로 목표정도에 의한 표본의 크기가 최소유지 표본의 크기보다 적다면, 최소유지 표본의 크기만큼 증가시켜야 한다. 그 다음 각각의 표본가구의 크기에 따라 소요될 비용을 예측한다. 표본설계에서 표본의 크기를 결정할 때 목표정도의 수준을 설정하는 데는 비용 예측이 필수적이다. 허용오차를 작게 설정하여 표본의 크기를 키우게 되면, 보다 정확한 추정값을 얻을 수 있지만 그만큼 비용이 많이 증가하게 된다. 전수조사를 하지 않고 표본조사를 선택하는 이유 중에 하나인 비용대비 조사의 효율성을 감안하여 목표정도를 적절하게 정하는 것이 바람직하다. 따라서 비용 예측 후, 적절한 수준에서 목표정도를 확정하여 최종적으로 표본의 크기를 결정하도록 하겠다.

3. 집락 당 표본의 크기 결정

집락추출은 단순임의추출과 비교할 때 보다 경제적이지만 효율은 떨어진다. 집락추출의 경제성을 유지하면서 효율을 높이는 한 가지 방법은 추출된 표본 집락 내의 모든 조사단위를 전부 조사하는 대신 각 표본 집락 내에서 다시 조사단위의 표본을 추출하는 방법을 생각할 수 있다. 이때, 집락 내의 조사단위가 동질적이고 집락이 서로 이질적이면, 표본 집락 내의 모든 단위를 조사하는 것은 유익하지 않기 때문에 집락 내의 표본 수를 작게 하고 그 대신 집락을 크게 하는 것이 유익하다. 반대로 집락 내의 조사단위가 이질적이고 집락이 서로 동질적인 경우는 표본크기가 동일하다면 집락 내에서 표본 수를 많이 추출하고 그 대신 집락 수를 줄이는 것이 효율적이다. 집세조사에서의 집락은 일반주택의 경우 조사구를 의미하고 아파트는 평형단지를 의미하며, 집락 내 조사단위는 임차가구를 뜻한다. 이러한 집락의 크기와 집락 내에서 추출되는 표본의 크기는 표본의 설계효과와 수집 비용에 많은 영향을 미친다. 설계효과를 고려한다면, 가능한 집락의 크기를 작게 만들고, 그 안에서 적은 수의 조사단위를 선택하는 것이 좋다. 가장 이상적인 방법은 하나의 조사단위만을 포함하도록 집락을 만드는 것인데, 이는 곧 임차가구를 임차가구 리스트에서 선택하는 것과 같다. 그러기 위해서는 모든 임차가구에 대한 명부(list)를 사전에 확보하고 있어야 하므로 불가능하다. 또한 집락 당 조사단위 수를 줄이면 더 많은 집락을 추출해야 하므로, 결과적으로 면접조사를 위한 이동거리와 시간이 늘어

나고 더 많은 집락에 대한 가구명부를 작성해야 한다. 이는 곧 비용의 증가를 의미한다. 따라서 적절한 집락의 크기와 집락 내에서 선택하는 최적의 조사단위 수를 결정하는 것은 필수적이다.

소비자물가지수의 집세지수 작성을 목적으로 조사하고 있는 외국의 집세관련 조사에서도 지역 추출틀(area frame)을 이용하여 다단집락추출을 하는 경우에는 집락의 크기와 집락 내 조사단위의 수를 효율적으로 관리하도록 조절하고 있다.

먼저 미국은 집세지수 작성을 위해 노동통계국(BLS : Bureau of Labor Statistics)에서 CPI Housing Survey를 수행하고 있으며, 가장 최근에 표본개편을 수행하여 2010년부터 적용될 예정이다. 이에 대한 주요 개편내용은 뒤에서 다시 설명하도록 하고, 여기서는 집락 당 표본크기만을 살펴보도록 하자. 먼저 추출단계는 38개 지수 작성지역으로부터, 87개의 1차 추출단위(PSU)를 추출하고, 1차 추출단위 내에서 10,000개의 세그먼트(segment)를 추출한 다음 각 세그먼트 내에서 5개의 가구를 선정한다. 여기서 1차 추출단위란 대도시와 중소도시를 의미한다. 세그먼트는 우리의 조사구 개념과 유사하며 도시 블록단위의 그룹을 말한다. 대도시의 경우 세그먼트 안에 최소 50개의 거처(housing unit) 이상을 포함할 수 있도록 그룹을 형성하며, 중소도시의 경우 최소 30개의 거처를 포함하도록 세그먼트를 구성한다. 즉, 미국의 경우 5만개의 표본을 추출하기 위해 세그먼트라는 집락을 1만 개 추출하고, 각 집락 내 조사단위 수를 5가구가 되도록 표본을 추출한다.

다음으로 일본에서는 House Rent Survey를 수행하고 있으며, 각 시정촌 내에 집세조사지구(Rent Survey District)를 구성하여 집락으로 활용하고 있다. 집세조사지구를 먼저 추출한 다음 선정된 집세조사지구 내에 거주하는 모든 임차가구를 표본으로 선택한다. 집세조사지구는 가장 최근 인구주택총조사의 조사구 중에서 매 5년마다 선정하는데, 조사구의 집세조사지구 선정을 위해 조사원의 업무 부담과 조사구 내 민영 임차가구 수를 고려하고 있다. 즉, 민영임차가구가 50가구 이상인 조사구는 제외시키며, 시단위의 지역에서는 15가구 미만인 조사구를 제외하고, 구단위 지역에서는 10가구 미만인 조사구를 제외시킨다. 2006년 12월 현재 1,212개의 집세조사지구에서 약 22,000 임차가구가 조사되고 있어 한 집세조사지구에서 평균적으로 약 18가구가 조사되고 있는 셈이다. 일본의 경우 집락 내 모든 조사단위를 선택하는 방법을 취하는 대신, 조사의 효율을 위해 집락의 크기를 제어하고 있다.

캐나다는 노동력조사(LFS)에 첨부된 Survey of Tenants를 통해 집세를 조사한다. 즉, LFS의 표본으로 선정된 거처에 대해 집세조사를 수행하고 있다. 따라서 LFS 표본체계를 살펴보면, 먼저 1차 추출단위(PSU)를 선택한 후, 2차로 거처를 선정한다. 집락에 해당하는 1차 추출단위는 여러 개의 센서스 블록을 묶어 생성하며, 보통 1차 추출단위는 200개 거처를 포함한다. 집락 당 선택되는 거처의 수는 시골(rural) 층에서는 집락 당 10가구를



선택하하며, 도시(urban) 층에서는 집락 당 8가구를 선정한다. 또한 대도시이면서 그 중요성이 큰 6개 도시에 대해서는 집락 당 6가구를 선정한다. 즉, 인구밀도가 낮은 지역은 수집비용이 매우 높은 반면, 추정값에 미치는 영향력은 상대적으로 낮은 특성이 있기 때문에 지역에 따라서 층을 구분함으로써 집락 당 조사단위의 수를 차등적으로 조절하고 있다.

우리가 집세조사에서 집락으로 설정한 조사구와 평형단지는 사전에 정의된 개념을 이용하고 있어 집락의 크기를 조절하는 것은 불가능하다. 따라서 집락의 크기가 일정한 상태에서 비용과 분산을 고려하여 임차가구의 수를 효율적으로 조절하는 방법을 이용해야 할 것이다.

집락 내 조사단위의 크기를 정하기에 앞서, 집락 내 조사단위 간에 유사한 특성을 보이는 지에 대한 탐색이 먼저 필요하다. 집락내의 조사단위 간의 동질성을 측정하는 측도(measure of homogeneity)로는 급내상관계수(ICC : Intraclass Correlation Coefficient)가 있다. 모집단에서 집락의 급내상관계수는 아래와 같이 구한다(Sharon L. Lohr, p.139).

$$ICC = 1 - \frac{M}{M-1} \frac{SSW}{SST}$$

여기서 M 은 집락 내 개체의 수이며, SSW 는 집락 내 제곱 합(Sum of Squares Within psu's)이며, SST 는 총 제곱 합(Total Sum of Squares)이다. 또한, $0 \leq SSW/SST \leq 1$ 이므로 급내상관계수는 $-\frac{1}{(M-1)} \leq ICC \leq 1$ 의 값을 갖는다. 만약, 집락 내의 단위들이 동질적인 경향을 보이면 급내상관은 1에 가까워진다. 반대로 집락 내 단위가 서로 이질적이고, 집락이 서로 동질적이면 급내상관은 0에 가깝거나 음의 값을 갖게 된다.

우리는 2005 인구주택총조사 10% 표본자료를 이용하여 일반주택과 아파트에 대해서 조사구 내 평균 집세에 대한 급내상관계수를 구해보았다. 일반주택은 조사구 내 집세의 유사성을 측정하면 된다. 그러나 아파트의 집락을 평형단지로 설정한 상황에서 평형단지 내 집세의 동질성을 측정할 만한 자료가 존재하지 않았다. 부동산 정보업체의 아파트 시세정보에서는 평형단지별 정보는 얻을 수 있지만, 평형단지 내 모든 가구의 가격이 동일하다. 반면, 인구주택총조사 자료로부터는 개별 가구의 집세정보는 얻을 수 있지만, 아파트의 평형단지 정보를 얻을 수 없다. 그러나 총조사의 아파트 조사구는 동일한 아파트 단지에서 한 개의 동 내지는 인접한 2개의 동이 한 조사구를 설정하고 있어 한 조사구에 한 개의 평형이 포함되어 있을 것으로 보았다. 집락 내 유사성을 측정하기 위해서는 집락 내 개별 가구의 정보가 필요하기 때문에, 우리가 정확히 얻고자 하는 평형단지의 정보는 아니지만 대체할 수 있는 정보로 판단되어 아파트 조사구에 대한 정보를 이용하였다.

주택유형에 따른 조사구 내 동질성을 확인하면서 7개 특·광역시와 9개 시도 간의 지리적 특성의 차이도 확인하기 위해 16개 시도별 급내상관계수를 구하였다. 인구주택 총조사의 조사구 크기는 적게는 40가구에서 많게는 80가구를 포함하고 있어 조사구별로 그 크기가 다르므로, 조사구 내 개체의 수(M) 대신에 해당 지역의 조사구별 개체 수의 평균(\bar{M})을 이용하여 계산하였다. <표 3-9>의 결과를 보면, 일반주택의 경우 경기도의 집락상관계수가 0.42로 16개 시도 중 조사구 내 임차가구 집세 수준이 가장 동질적으로 나타났다. 다음으로 울산 0.36, 서울이 0.32 순으로 조사구 내 단위들의 동질적인 집세경향이 어느 정도 있음을 확인할 수 있었다. 일반주택의 급내상관계수와 비교하여 아파트는 조사구 내 집세의 동질성이 매우 높게 나타나고 있는데, 대부분의 지역에서 0.85 이상의 값을 보이고 있다. 즉, 아파트 조사구 내의 임차가구의 집세 수준이 유사하므로 아파트 조사구 내의 표본 수는 줄이고, 다양한 아파트 조사구를 뽑는 것이 효율을 높일 수 있을 것이다. 일반주택은 조사구 내에서도 다양한 특성의 가구들이 존재하는 것으로 보이므로 아파트 조사구 내 추출 단위의 수보다는 더 많은 표본을 추출하는 것이 바람직하다고 보인다.

<표 3-9> 조사모집단의 주택유형별 급내상관계수(ICC)

	일반주택	아파트
전체평균	0.25	0.86
서울	0.32	0.88
부산	0.23	0.86
대구	0.22	0.87
인천	0.20	0.84
광주	0.19	0.87
대전	0.25	0.88
울산	0.36	0.84
경기	0.42	0.88
강원	0.20	0.88
충북	0.21	0.91
충남	0.25	0.84
전북	0.28	0.87
전남	0.16	0.88
경북	0.28	0.85
경남	0.21	0.85
제주	0.28	0.75

이상으로 한 조사구 내에서 집세 수준의 동질적 경향이 나타나고 있음을 확인하였다. 따라서 선정된 조사구 내에서 모든 가구를 조사하는 것보다는 조사구 내 표본 수를 일



정 수준으로 제어하면서 보다 다양한 특성의 조사구들을 추출하는 것이 효율적이라고 하겠다. 다음으로 일반주택과 아파트의 조사구 내에서 몇 개의 가구를 추출해야 할 것인지 비용과 효율성을 고려하여 구해보고자 한다. 비용함수의 추정은 매우 복잡하고 개별비용에 대한 세세한 정보를 필요로 하며, 추정된 비용함수가 객관성과 타당성을 확보하기 위해서는 많은 연구가 수행되어야 한다. 본 연구에서는 가장 간단한 비용함수를 이론상 접근해보도록 하겠다.

2단집락추출의 경우 비용함수는 두 개 부분으로 구성된다. 하나는 표본 1차 추출단위(PSU)의 수 n 에 의존하게 되고, 다른 하나는 표본 2차 추출단위(SSU)의 수 m 에 의존하게 된다.

$$C = c_0 + nc_1 + nmc_2$$

여기서 c_0 는 고정비용이 되고, c_1 은 표본 PSU당 비용이다. 즉, 표본 PSU의 현지확인 및 PSU간의 이동비용, 각 표본 PSU내 SSU를 추출하기 위한 리스트 작성, PSU의 추출확률 작업 등에 필요한 비용을 포함한다. c_2 는 표본 SSU당 비용으로서 SSU의 추출 및 확인에 소요되는 비용, 면접 또는 실측비용, 조사자료의 집계·분석 비용을 포함한다. 또한 표본의 최적값은 일정 비용 하에서 분산을 최소화 하는 n 과 m 을 정하는 문제가 된다. 일정 비용 하에서 구하고자 하는 m 의 최적값 m_0 는 다음과 같다.

$$m_0 = \sqrt{\frac{S_w^2 c_1}{S_b^2 c_2}}$$

여기서 $S_w^2 = MSW$, $S_b^2 = \frac{MSB - MSW}{M}$ 이다. 이에 대한 보다 자세한 내용은 박홍래(1999)를 참고하기 바란다.

조사구 당 최적의 추출단위 수를 구하기 위해서 2005 인구주택총조사 10% 표본자료를 이용하여 분산분석을 실시하였다. 조사구 당 비용(c_1)과 가구 당 비용(c_2)의 비율은 조사원의 방문기록과 업무량, 조사구 간 이동거리와 비용 등이 세세하게 측정이 되어야 가능하다. 조사구 당 비용 중 명부작성은 조사구 내의 전체 가구에 대한 정보를 얻어야 하므로, 현지 확인 작업과 전 가구 면접조사, 조사 홍보 등에 많은 시간과 비용이 소요되며, 조사구 간 이동비용도 매우 높은 편이다. 따라서 일반적으로 c_1 이 c_2 보다 더 크므로, c_1/c_2 의 비율은 1보다 큰 값을 갖게 된다. 그러나 현실적으로 c_1 과 c_2 를 파악할 수 없기 때문에 여기서는 비용의 비율에 대한 여러 가지 경우를 가정한 후, <표 3-10>과 같이 일정 비용 하에서 조사구 당 최적의 추출 가구의 수를 구해보았다.

〈표 3-10〉 일정 비용 하에서 조사구 당 최적의 추출단위 수

c_1/c_2	일반주택					아파트				
	5	10	12	15	20	5	10	12	15	20
전체평균	4.7	6.6	7.3	8.1	9.4	0.9	1.3	1.4	1.6	1.8
서울	3.0	4.2	4.6	5.2	6.0	0.7	1.1	1.2	1.3	1.5
부산	4.3	6.1	6.7	7.5	8.6	1.0	1.3	1.5	1.6	1.9
대구	4.5	6.4	7.0	7.8	9.0	0.9	1.3	1.4	1.5	1.8
인천	4.9	6.9	7.6	8.5	9.8	1.1	1.5	1.6	1.8	2.1
광주	5.0	7.1	7.7	8.7	10.0	0.8	1.1	1.3	1.4	1.6
대전	3.8	5.4	5.9	6.6	7.7	0.8	1.1	1.2	1.3	1.5
울산	3.1	4.4	4.8	5.3	6.2	1.1	1.6	1.7	1.9	2.2
경기	2.4	3.4	3.7	4.1	4.8	0.8	1.2	1.3	1.4	1.6
강원	5.1	7.3	8.0	8.9	10.3	0.8	1.1	1.2	1.3	1.5
충북	4.7	6.7	7.3	8.2	9.5	0.6	0.9	1.0	1.1	1.3
충남	5.0	7.1	7.8	8.7	10.1	0.9	1.3	1.5	1.6	1.9
전북	4.6	6.5	7.1	7.9	9.1	0.8	1.2	1.3	1.4	1.7
전남	7.7	10.8	11.9	13.3	15.3	0.7	1.0	1.1	1.3	1.5
경북	4.8	6.8	7.5	8.3	9.6	1.0	1.4	1.5	1.7	2.0
경남	4.6	6.5	7.1	8.0	9.2	1.0	1.4	1.5	1.7	2.0
제주	5.9	8.3	9.1	10.2	11.7	1.5	2.2	2.4	2.7	3.1

일반주택의 경우, c_1/c_2 의 비율을 5로 가정하였을 때 서울에서는 조사구 당 3.0가구, 전라남도에서는 7.7가구를 추출하는 것이 바람직하며, 16개 시도의 평균가구 수는 4.7가구이다. 또한 두 비용의 비율을 15로 가정하면, 평균 약 8가구를 조사구 내에서 추출하는 것이 적절한 수준으로 판단된다. 아파트의 경우, 조사구 내 집세의 동질성이 일반주택보다 높기 때문에 조사구 내 추출단위의 수가 훨씬 적게 계산된다. c_1/c_2 을 15로 가정했을 때 서울은 1.3가구, 제주도는 2.7가구를 조사구별로 선정하는 것이 최적이라 판단된다. 아파트의 조사구는 평형단지보다 작은 개념이므로 평형단지별 추출 표본의 크기는 이보다 다소 크게 정해야 한다. 따라서 조사구 당 비용이 조사구 내 비용보다 15배 정도 더 크다고 가정하면, 집락 당 추출단위의 수는 일반주택은 8가구, 아파트는 4가구가 만족스러운 수준으로 보인다.



4. 최종 표본 크기

최종 표본의 크기는 최소유지 표본 크기와 목표정도에 의한 표본 크기를 비교하여 월별 계약 갱신의 탐지가 가능하고, 추정의 정도를 만족시킬 수 있도록 두 방법 중 표본의 크기가 더 큰 쪽으로 결정하였다. 목표 CV는 전국 일반주택은 1.5%, 아파트는 3.0%로 삼았다. 또한 시도는 일반주택은 3.5%, 아파트는 6.0%로 목표수준을 설정하였으며, 31개 도시에 대해서는 목표 CV를 8.0%로 설정하였다. 하나의 집락 당 표본가구를 일반주택은 8가구, 아파트는 4가구를 포함하여야 하므로 8의 배수와 4의 배수가 되도록 조정해주었다. 최종 표본 크기는 <표 3-11>과 같다. 전국적으로 15,352가구의 표본이 필요하며, 이 중 일반주택이 8,696가구, 아파트가 6,656가구이다. 집락을 기준으로 보면, 일반주택 조사구는 1,087개, 아파트 평형단지는 1,664개가 필요하다. 이는 기존의 집세표본의 크기보다 약 6천여 가구가 증가한 것인데, 이는 시도별 주택유형별로 지수를 세분화시키고자 하였고, 도시별 최소유지 표본의 크기를 고려함으로써 더 많은 수의 표본이 필요하게 된 것이다.

이와 같이 결정된 표본의 크기에 대해서 어느 정도의 비용이 소요될 것인지 예측해보고자 한다. 하나의 조사에 소요되는 비용에는 여러 가지가 있는데, 조사기획 및 조사수행을 위한 자원에 대한 비용, 자료처리비용, 자료 수집을 위한 인건비용 등이 있다. 이 중에서 기존의 조사에서도 소요되고 있는 비용들은 제외하고, 조사비용 중 가장 큰 비중을 차지하며 표본 크기에 가장 민감하게 반응하는 인건비에 대해서만 살펴보고자 한다. 만일 집세조사가 독립표본으로 전환하게 된다면, 별도의 조사구에서 표본을 선정하게 될 것으로, 기존의 조사인력으로는 감당할 수 없게 된다. 따라서 인력의 보충이 필요할 것이며, 이를 집세조사만을 전담하는 상시조사원을 채용한다는 가정 하에서 인건비를 산출해보도록 하겠다.

상시조사원 인건비는 일당으로 계산하며, 2009년 기준단가가 일 42,660원이다. 현재 가구부문 경상조사를 담당하는 정규직원은 경제활동인구 조사구 2개와 가계동향 조사구 1개를 전담하고 있다. 이들은 경상조사 뿐 아니라 관리와 행정업무 등을 겸하고 있어서 이들의 업무량을 새로운 조사에 그대로 적용하는 것은 무리가 있다. 또한 상시조사원이 집세조사만을 전담한다고 하였을 때, 조사기간을 현행과 같이 일주일로 설정하는 것보다는 조사기간을 늘려주는 것이 필요하다. 여러 명의 상시조사원을 교육하고, 관리하는 것보다는 단련된 소수의 상시조사원이 집중적으로 조사를 수행하는 것이 관리측면에서 효율적이기 때문이다. 한 명의 상시조사원이 한 달 동안 4개의 일반주택 조사구(32가구)와 6개의 아파트 평형단지(24가구)를 전담한다고 가정하면, 연간 총 예상 인건비는 <표 3-12>와 같이 구해진다.

<표 3-11> 38개 도시별 최종 표본 크기

구분	기존 표본 가구 수			최종 표본 가구 수			
	합계	일반주택	아파트	합계	일반주택	아파트	
서울	2,013	1,489	524	2,300	1,640	660	
부산	825	567	258	1,008	608	400	
대구	640	527	113	880	432	448	
인천	581	376	205	764	440	324	
광주	508	330	178	728	464	264	
대전	559	337	222	852	504	348	
울산	504	291	114	984	584	400	
경기	수원	149			212	152	60
	성남	136			340	248	92
	의정부	89			180	104	76
	안양	151	658	292	224	128	96
	부천	119			224	160	64
	안산	95			228	184	44
	고양	211			216	104	112
강원	춘천	145			224	128	96
	원주	173	264	189	240	96	144
	강릉	135			192	96	96
충북	청주	209			544	312	232
	충주	125	175	758	288	128	160
충남	천안	207			284	120	164
	보령	51	176	175	192	96	96
	서산	93			224	96	128
전북	전주	123			412	216	196
	군산	89	176	92	388	112	276
	남원	55			264	80	184
전남	목포	83			304	104	200
	여수	129	154	161	268	64	204
	순천	103			276	64	212
경북	포항	117			180	112	68
	경주	173			172	88	84
	안동	109	460	215	200	72	128
	구미	275			164	88	76
경남	창원	130			188	120	68
	마산	107	338	104	256	192	64
	진주	111			180	144	36
	김해	95			256	112	144
제주	제주	150	181	42	332	200	132
	서귀포	73			184	104	80
합계	9,539	6,499	3,040	15,352	8,696	6,656	

〈표 3-12〉 예상 소요 총 인건비

조사원수	약 270명
인건비 ¹⁾	2,764,368천원

¹⁾인건비=42,660원*조사원수*20일*12개월



제3절 표본관리

표본설계에서는 가능한 최신의 모집단의 정보를 사용하여 모집단과 유사한 구조의 표본을 선정하는 것이 중요하다. 그러나 우리가 알고자 하는 모집단의 구조는 끊임없이 변화한다. 지속적으로 발생하는 모집단의 변동을 표본에 최대한 반영시켜주기 위해 경상조사에서는 표본 개편과 보완을 주기적으로 수행한다. 우리나라에서는 가구 모집단 구조를 파악할 수 있는 총조사가 5년에 한 번씩 수행되고 있어 가구표본의 개편도 5년 주기로 수행되고 있다. 따라서 표본개편 시 어떠한 설계방법을 적용하여 표본을 선정했는지의 문제만큼 선정된 표본을 다음 개편 때까지 어떻게 관리하고 유지해 나갈 것인지에 대한 문제도 그에 못지않게 중요하다. 표본을 관리하는 문제에 대해서 표본 교체방식과 운영방법 측면으로 구분해 고려해 볼 필요가 있다.

1. 표본교체방법

먼저 어떠한 관리방법을 적용하여 표본의 구조를 구성하고 교체할 것인가의 문제가 있다. 이는 표본을 설계하는 과정에서 고려해야 하며, 크게 두 가지로 구분해볼 수 있다. 하나는 고정표본관리방법으로 표본을 한번 선정하면 차기 표본개편 시까지 동일한 표본에서 계속 조사를 실시하다가 다음 개편 시 새로운 표본으로 일시에 교체해주는 방식이다. 이러한 고정표본관리체제 하에서는 최초의 표본이 시간이 지나면서 전출, 멸실, 불응, 상태변경 등의 여러 가지 사유로 조사 대상 표본으로부터 탈락함으로써 표본의 수가 점차 줄어드는 표본의 마모가 발생한다. 그리고 끊임없이 변화하는 모집단의 구조를 표본에 시의성 있게 반영하기가 어려워 그에 따른 대표성 상실의 우려도 있다. 따라서 고정표본관리방법을 이용하였을 때는 개편 사이에 보완이나 보정작업을 통해서 표본의 규모와 분포를 유지시켜줄 필요가 있다. 또한 고정표본제에서는 조사 대상 표본의 전부를 일시에 교체함에 따라 개편을 전후로 하여 시계열 단절의 문제도 발생하게 된다.

이와 같은 고정표본관리방법의 단점을 보완하기 위해 연동표본관리방법이 등장하였다.

이 방법은 표본의 전면개편 대신에 전체 표본의 일부를 일정기간마다 순차적으로 교체해주는 방법이다. 즉, 정해진 주기마다 표본을 부분적으로 교체해주기 때문에 모집단의 규모나 구조변화에 대해서 고정표본체에 비해 유동적으로 대처·관리할 수 있다. 기존 통계청의 집세표본은 고정표본관리방법에 의해 운영하였으나 2007년부터 연동표본관리방법으로 전환하여 관리되고 있다. 우리나라는 재고주택의 노후화, 지역 균형발전 및 수도권 분산화 정책 등에 의해 지난 수년간 재개발, 재건축, 신도시 개발 사업 등이 활발히 진행되었다. 이에 따라 단기간 특정지역에 신규주택이 집중적으로 공급되고 재고주택이 일시에 멸실되는 현상이 매우 빈번하게 발생하는 등 주택 임대시장의 구조가 빠르게 변하고 있다. 신축, 멸실, 점유형태 변경 등으로 변동이 발생한 모집단 정보를 주기적으로 수집하여 추출틀을 보완함으로써 표본 교체시기에 보다 효율적으로 대표성 있는 조사대상 가구를 선정할 수 있다.

해외의 집세관련 조사에서도 표본을 효율적으로 유지·관리하기 위해 노력하고 있다. 먼저 미국의 경우 집세표본을 표본 개편 시 전면 교체하는 고정표본관리방법을 이용해 왔다. 2009년 현재는 1999년에 개편한 표본을 이용 중이며 매년 줄어드는 표본을 조금씩 보완해주고 있다. 1999년 표본개편 당시에는 5만 가구를 목표로 표본을 설계하였으나, 최근 3~4년 동안은 예산부족 등의 이유로 줄어든 표본을 보완해주지 못해서 현재는 약 3만 6천의 가구만이 유지되고 있을 뿐이다. 이렇듯 개편 주기도 10년으로 우리에게 비해 매우 길고 표본 탈락 현상이 심각해지면서 2010년부터는 연동표본관리방법으로 변경할 예정이다. 매년 전체 표본의 1/6씩을 제거하고 새로운 표본을 투입하는 방법을 계획 중인데, 이 방법에 의해 6년이 지나면 최초의 표본이 새로운 표본으로 완전 교체된다. 또한 전체 표본을 6개의 패널로 구분하고, 동일한 가구에 대해서는 6개월에 한 번씩 조사를 수행하며 매월 하나의 패널을 조사한다. 즉, 6개의 패널 중 1번 패널은 1월과 7월에 조사를 하고, 2번 패널은 2월과 8월에 조사하는 방식으로 매월 하나의 패널을 조사하여 월별지수를 제공하고 있다.

다음으로 스위스의 집세조사에서는 매 분기마다 전체 표본의 1/8씩 교체된다. 따라서 모든 가구는 최대 8분기만 표본으로 유지된다. 새로 건축된 주택 뿐 아니라 임대방식이 변경된 새로운 임차가구도 표본으로 지속적으로 포함시키기 위해서 연동표본관리방법을 이용 중이다. 노동력조사의 표본을 이용하는 캐나다의 경우도 1/6 연동표본관리방법을 적용하고 있다. 집세의 변동은 계약갱신이 발생해야 가능하고, 일반적으로 월세의 계약기간은 1년, 전세의 계약기간은 2년이다. 표본가구가 계약을 해지하고 전출을 하지 않는 한, 계약기간 동안에는 집세는 변하지 않는다. 이렇듯 집세의 변동은 소비자물가조사의 다른 품목들 중에 빈번하지 않다는 특성으로 매월 조사 대신 비용과 조사의 효율을 높이기 위해 해외의 각국에서는 연동표본관리방법을 사용하고 있다.



우리나라의 집세조사가 새로운 표본설계에 의해 독립표본으로 전환하게 되면, 비용의 증가가 불가피하다. 앞의 <표 3-12>에서 보았다시피, 15,352가구의 표본을 조사하기 위해서는 약 28억 정도의 인건비가 소요되리라 예상된다. 하지만 연동표본관리방법을 적용하여 동일 표본가구에 대해서는 3개월에 한 번씩 조사를 하고, 매월 전체 표본의 1/3에 해당하는 표본가구만 조사를 수행하게 된다면, 인건비도 약 1/3이 줄어든 1억 3천여만원으로 절감할 수 있다. 이러한 연동표본관리 하에서 전체 표본을 3개의 패널로 구분한 후, 1번 패널은 1, 4, 7, 10월에 조사를 하고, 2번 패널은 2, 5, 8, 11월, 3번 패널은 3, 6, 9, 12월에 조사하게 된다. 그리고 매년 전체 표본가구의 1/3씩을 교체하게 된다면, 노후화된 표본을 신축 주택들과 자연스럽게 교체할 수 있다. 따라서 예산과 인력을 효율적으로 관리하고, 주택과 임차가구에 대한 모집단의 잦은 변동과 구조적인 변화를 표본에 반영할 수 있기 때문에 연동표본관리방법이 집세조사에 적합하다고 판단하고 이 방법을 제안하고자 한다.

2. 표본관리방법

다음으로 선정된 표본에 대해서 조사를 수행하는 기간에 걸쳐 실질적으로 운영·유지하는 방법에 대해서 고려해 보아야 한다. 표본설계 이후 실제 조사대상을 접촉하는 과정에서 최초 설계 시 선정해 놓은 표본조사구나 표본가구가 조사가 불가능한 다양한 상황에 직면하게 될 때가 있는데, 이를 대비하여 일관성 있는 규칙을 사전에 마련하여 표본가구의 관리에 혼란이 없도록 조치해야 한다.

가. 전입 및 전출가구에 대한 조치

표본설계를 통해 추출된 표본조사구에 대해서 조사를 수행하기 전에 표본조사구의 가구명부를 작성하여 조사대상 가구를 선정한다. 이 때 표본조사구 내에 실제 거주하는 가구들의 목록과 임차가구 현황을 조사하게 되는데, 해당 조사구에 철거, 재개발, 재건축 등으로 주변지역 전체의 주택속성이 바뀌거나 집단이주가 발생하여 조사가 불가능해진 경우에는 해당 조사구를 표본에서 제거하고 다른 조사구를 대체한다. 대체 표본조사구는 같은 동 내에서 특성이 유사한 인접 조사구를 선정하는 것을 원칙으로 한다. 또한 표본조사구 내에 조사대상 가구가 전출한 경우 동일 주택으로 전입해온 가구가 조사대상으로 적격하다면 전입가구에서 계속 조사를 진행한다. 만일 전출한 가구의 주택에 다른 가구가 전입해오지 않아 공실이 지속되거나, 자가 혹은 무상 거주 가구의 이전 등 부적격 가구가 전입을 하였다면 해당 주택 가구에 대한 조사를 중지하고 다른 가구로 대체한다. 대체 가구를 선정할 때는 아파트의 경우 같은 평형단지 동 내에서 대체 가구를

선정하며, 일반주택은 동일 조사구 내에서 주택특성이 유사한 인접 가구로 대체하는 것을 원칙으로 한다.

나. 무응답 처리

무응답은 크게 단위 무응답(unit nonresponse)와 항목 무응답(item nonresponse)로 나눌 수 있다. 단위 무응답은 표본가구의 모든 조사항목에 대해 응답을 얻을 수 없는 경우에 해당하고, 항목 무응답은 조사항목 중 일부 항목에 대해서만 응답을 얻지 못한 경우에 해당한다. 집세조사의 조사항목은 많지 않으며, 집세를 제외한 다른 항목들에 대해서는 응답자 이외의 경로를 통해 파악이 가능하다. 항목 무응답은 집세조사의 지침서에 근거하여 처리를 하되, 단위 무응답의 경우 표본의 크기 및 지수 작성에 큰 영향을 미치므로 단위 무응답은 표본관리 측면에서 좀 더 세밀한 주의가 필요하다.

집세조사에서 표본가구가 보수공사, 출타, 연락두절 등으로 응답을 하지 않아 일시적으로 단위 무응답이 발생한 경우에는 자료 대체 방법을 사용한다. 가능한 응답자를 교체하지 않는 것이 바람직하지만, 부득이하게 6개월 이상의 장기 불응이 지속될 시에는 해당 가구에 대한 조사를 중지하고 다른 표본가구로 교체한다. 또한 무응답에 관한 정보 및 관리상황은 해당 조사의 품질을 평가하고 관리하는 데 있어 중요한 요소이기 때문에 이에 대해 체계적으로 기록·관리해야 한다.

다. 표본들의 관리

표본설계가 수행된 이후에도 모집단의 상황은 계속 변하고 있기 때문에 변화된 모집단의 모습을 표본들에 반영해주지 않는다면 표본의 대표성이 떨어지게 될 것이다. 그러나 현실적으로 모집단의 모든 변화를 표본들에 반영해주는 것은 매우 어렵기 때문에, 정기적으로 혹은 필요에 따라서 표본들을 관리해줘야 한다. 본 연구에서 집세조사의 표본들은 일반주택은 인구주택총조사 결과를 이용하였으며, 아파트는 부동산정보업체의 아파트 평형단지 정보를 활용하였다. 이 중 인구주택총조사 자료는 5년에 한 번씩 수행되어 수시로 모집단의 변동을 확인하는 데는 제약이 있다. 그러나 재개발과 신도시 건설, 대규모 택지 조성, 신도시 건설 등 대규모로 철거와 건설이 대규모로 진행된 지역에 대해서는 모집단의 규모를 파악하여 표본에 반영해줘야 할 것이다.

일반주택에 비해 상대적으로 신축 및 입주계획 등에 대한 정보 수집이 용이한 아파트는 일 년에 한 번 정도는 주기적으로 표본들을 보정해줄 필요가 있다. 그러나 신축 아파트의 경우 해당 아파트의 시세가 형성되기까지는 일정기간이 필요하고, 매매가격 시세가 형성되고 나서 그에 준하여 전·월세 시세가 형성되는 경우가 일반적이다. 그리고

신규 입주를 전후하여 매매 및 전·월세 가격의 변동이 매우 심하다. 이러한 가격의 변동은 일반적인 물가상승을 반영하는 것이 아니기 때문에 신축 아파트 단지를 표본틀에 반영시키는 것은 지수의 불안정을 초래할 수 있다. 따라서 신축 아파트를 표본틀에 반영할 때에는 세심한 주의가 필요하다.

제4절 결론 및 요약

표본설계에서 당면하는 문제의 하나는 표본의 크기를 결정하는 것이라 할 수 있다. 표본의 크기가 필요 이상으로 크면 예산이 낭비될 뿐만 아니라 조사의 정확성도 감소될 수 있기 때문에 좋지 않고, 반대로 너무 작으면 표본의 정도가 낮기 때문에 좋지 않다. 본 연구에서는 표본규모 산정과 관련하여 객관적인 검토기준을 마련하고자 하였다. 먼저 집세의 계약갱신에 관한 특성을 고려함으로써 집세조사에서 최소한으로 확보·유지되어야 할 표본의 크기를 구하였다. 이에 지역별로 유지해야 하는 계약갱신의 크기에 대한 기준을 과거 집세조사 경험에 근거하여 실증적으로 찾았다. 시도 수준에서는 일반주택은 최소 20건의 유효표본을 보유해야 하며, 아파트의 최소 15건, 도시별로는 최소 10건 이상을 확보할 수 있도록 최소유지표본 수를 구하였다.

다음으로 1차추출단위인 집락 내 동질성을 측정하는 급내상관계수를 구하여 조사의 효율을 높일 수 있는 방법을 제시하고자 하였다. 일반주택 보다는 아파트가 조사구내 집세의 동질성이 높게 나타나고 있었다. 따라서 아파트 조사구 내의 임차가구의 집세수준이 유사하므로 아파트 조사구 내의 표본 수는 줄이고, 다양한 아파트 조사구를 뽑는 것이 효율을 높일 수 있다고 분석하였다. 집락 내 추출할 최적의 표본의 크기를 이론식에 의해 계산하였고, 선정된 집락 내에서 일반주택은 8가구, 아파트는 4가구를 추출하는 것이 효율적이라고 밝혔다.

그리고 지수 공표 범위별로 목표정도를 달성하기 위해 필요한 표본의 크기를 결정하였다. 집세 평균에 대한 목표 CV를 전국의 일반주택은 1.5%, 아파트는 3.0%, 시도의 일반주택은 3.5%, 아파트는 6.0%, 31개 도시는 8.0%로 목표정도를 설정하고, 표본크기를 구하였다. 또한 최종 표본의 크기는 최소유지표본 크기와 목표정도에 의한 표본 크기를 비교하여 월별 계약 갱신의 탐지가 가능하고, 추정의 정도를 만족시킬 수 있도록 두 방법 중 표본의 크기가 더 큰 쪽으로 결정하였다. 최종적으로 전국에서 15,352가구의 표본이 필요하며, 이 중 일반주택이 8,696가구, 아파트가 6,656가구가 필요하다고 제시하였다. 이는 기존의 집세표본의 크기보다 약 6천여 가구가 증가한 것인데, 이는 시도별 주



택유형별로 지수를 세분화시키고자 하였고, 도시별 최소유지표본의 크기를 고려함으로써 더 많은 수의 표본이 필요하게 된 것이라고 밝혔다. 이러한 표본을 독립적으로 조사하기 위해서는 매년 약 28억의 인건비가 소요될 것이라고 예상하였다.

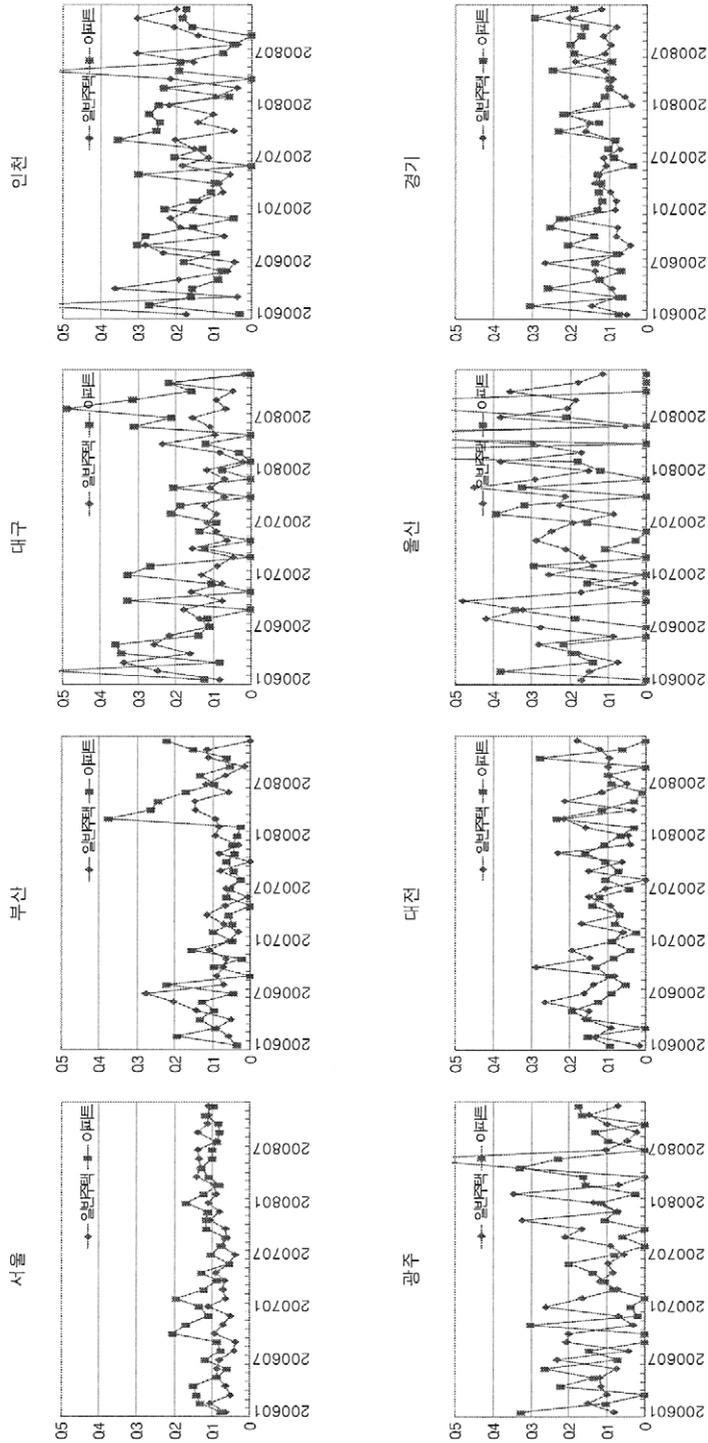
마지막으로 예산과 인력을 효율적으로 관리하고, 모집단의 변화도 주기적으로 반영시키기 위해서는 연동표본관리방법을 이용할 것을 제안하였다. 집세의 변동은 빈번하게 발생하지 않기 때문에 조사주기를 3개월로 조정해도 무리가 없다고 판단하였다. 즉, 동일 표본가구에 대해서는 3개월에 한 번씩 조사를 하고 매월 전체 표본의 1/3에 해당하는 표본가구에서 조사를 수행하면, 노후 표본을 자연스럽게 교체하고, 표본마모 현상을 방지할 수 있을 뿐 아니라 인건비도 약 1/3이 절감할 수 있을 것으로 보았다.

참고문헌

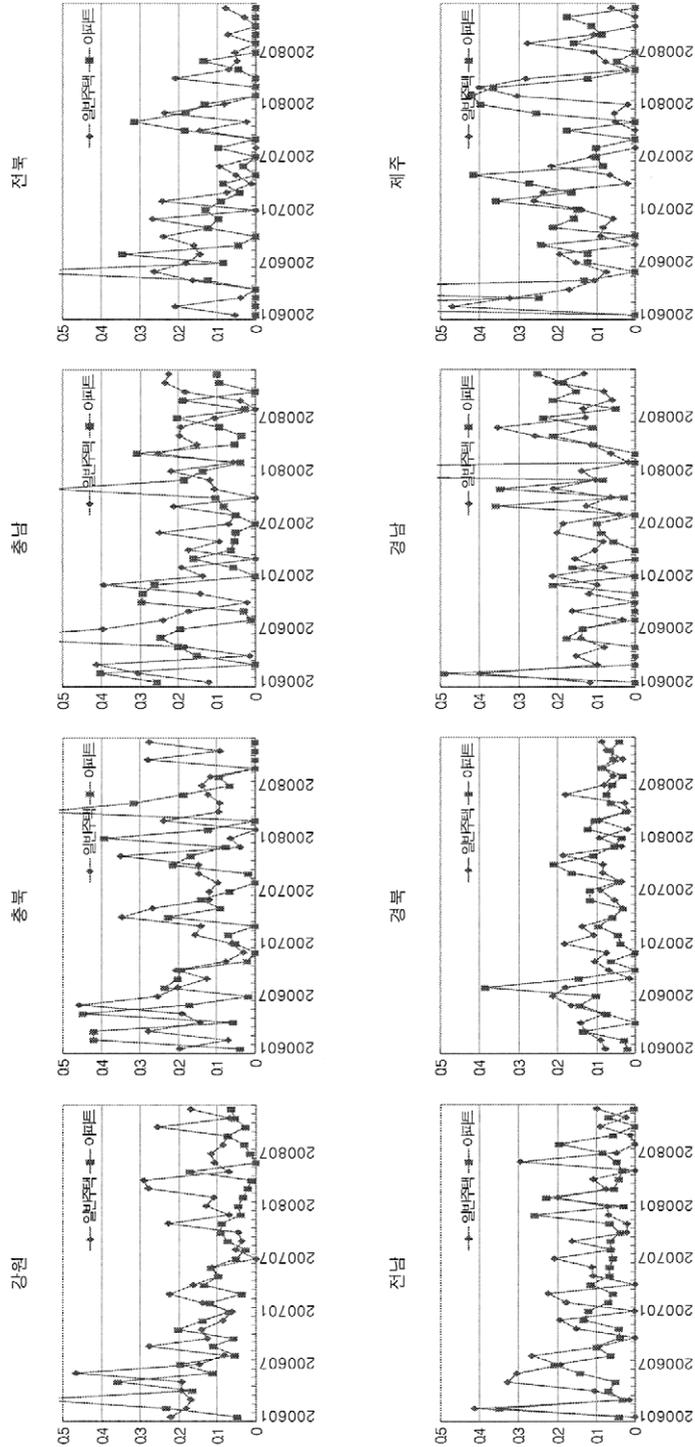
- 박홍래(1989), 「통계조사론」.
- 정미옥, 박소현(2009), “집세조사의 표본설계방안”, 「통계개발원 2009년 상반기 연구보고서 I」, pp.167-221, 통계개발원.
- 통계청(2009), 2008년도 표준업무량 측정 결과보고서, 조사기획과
- Sharon L. Lohr(1999), 「Sampling: Design and Analysis」.
- Statistics Canada(2008), “Methodology of the Canadian Labour Force Survey”.
- C. B. Vermeulen, H. M. Herren(2009), “Rents in Swizerland: sampling and qualilty adjustment”.



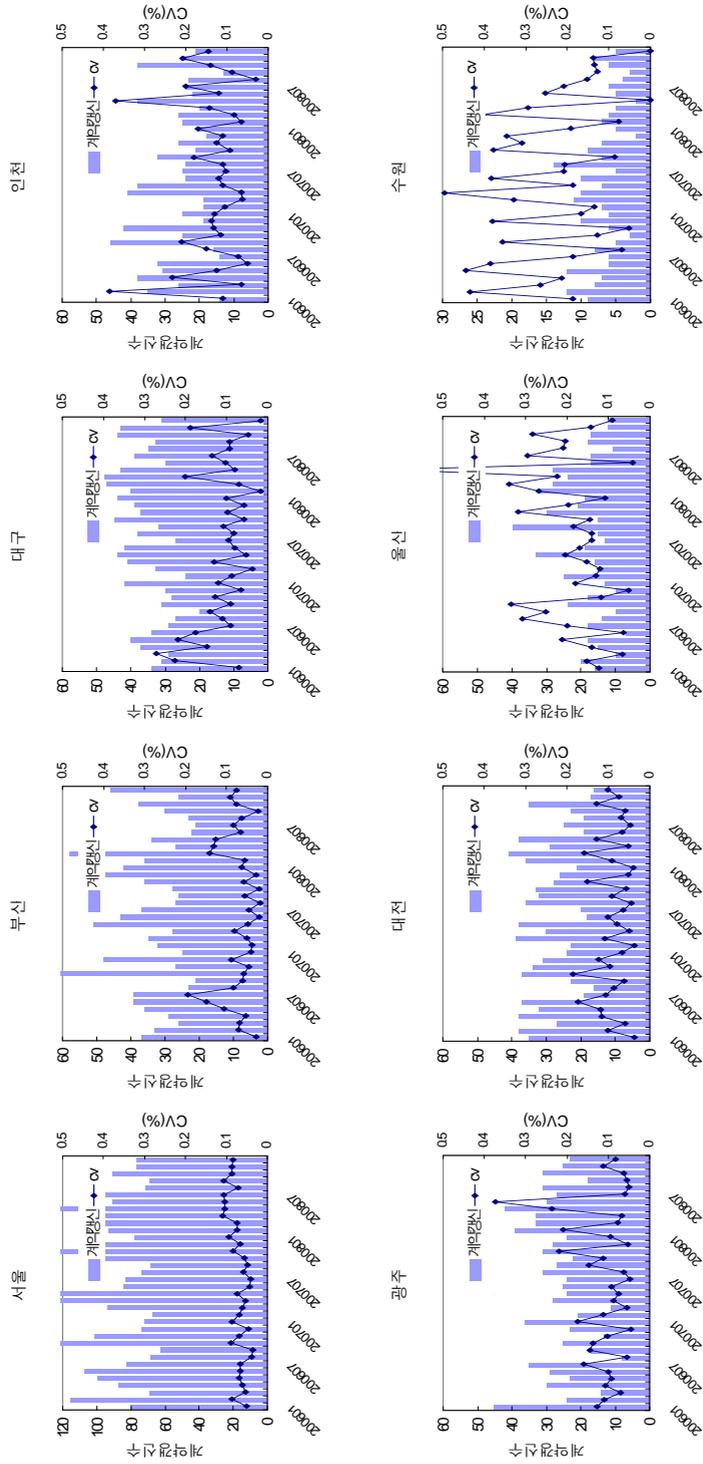
〈부록 1〉 시도별 주택유형별 평균 변동률의 CV(%)



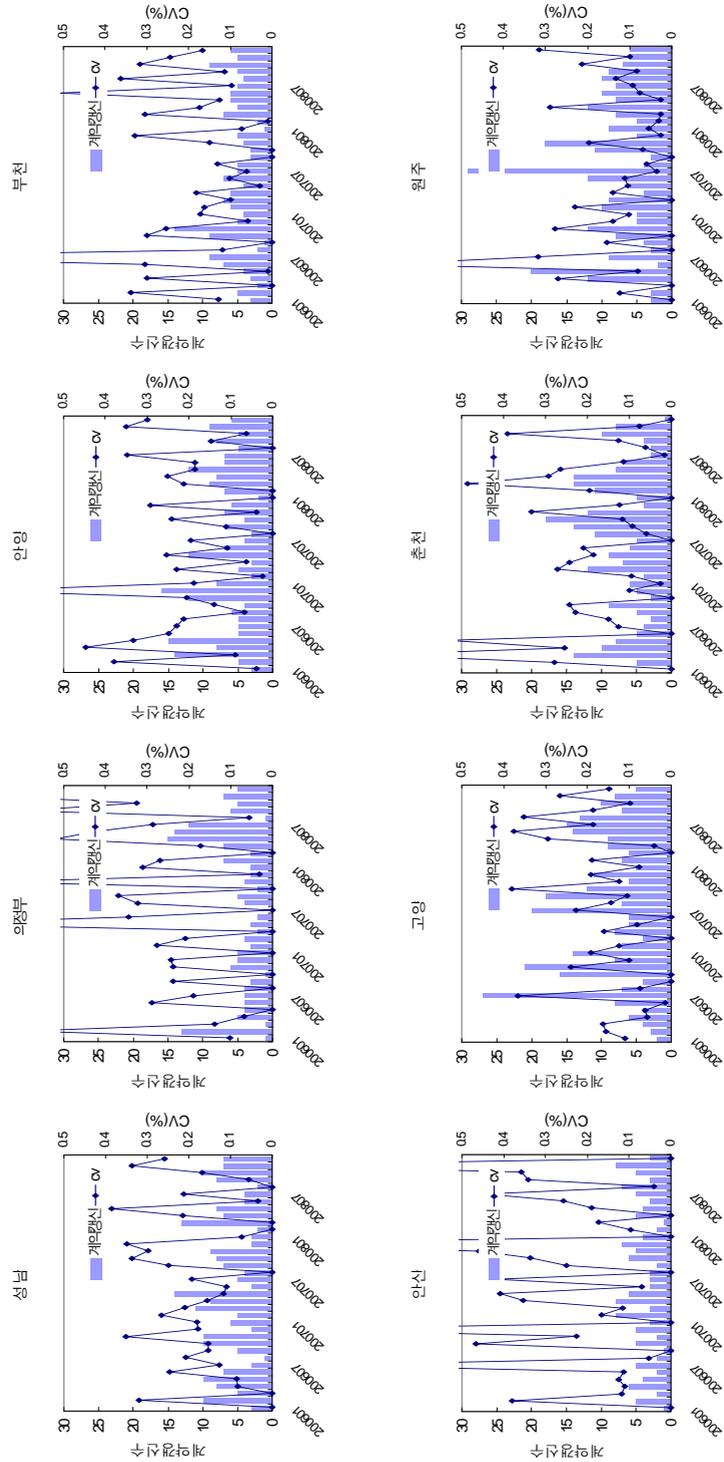
〈부록 1〉 시도별 주택유형별 평균 변동률의 CV(%) (계속)



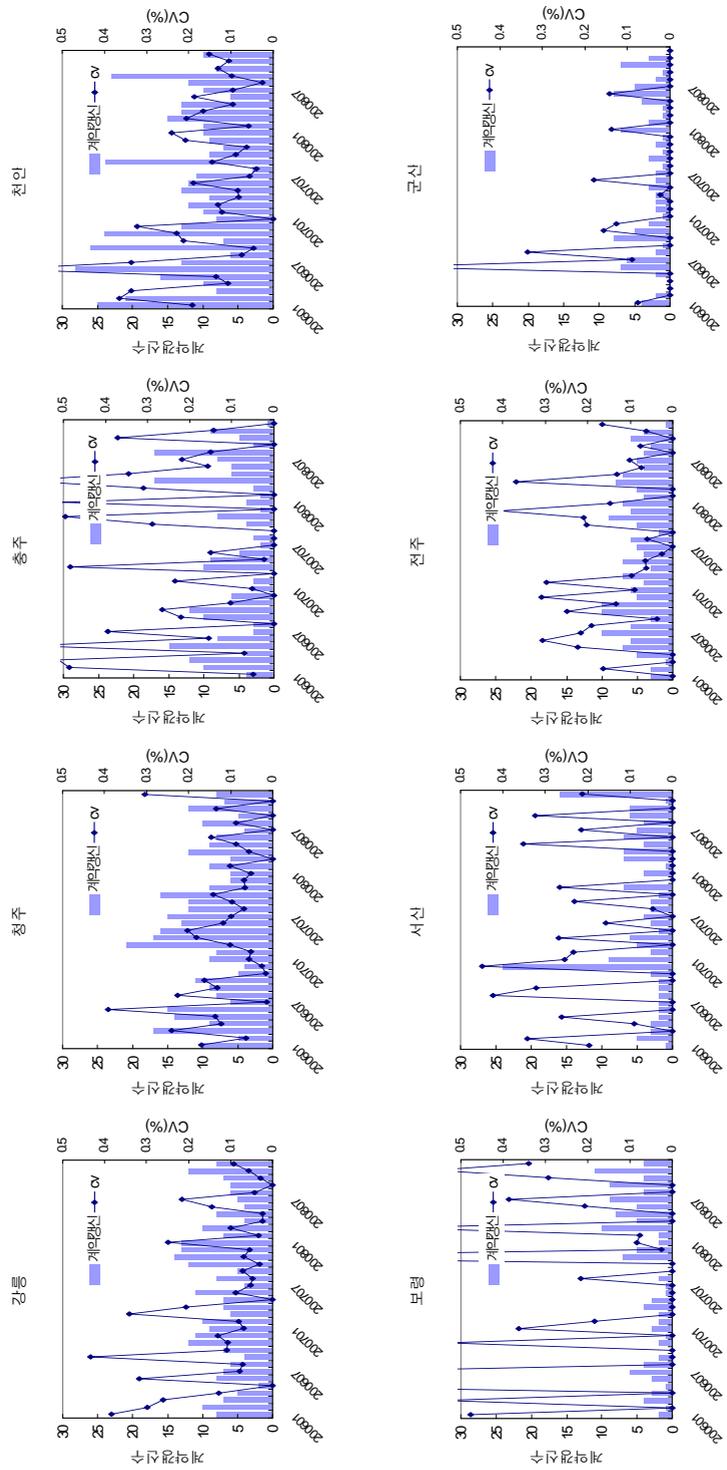
〈부록 2〉 도시별 평균 변동률의 CV(%)와 계약갱신수



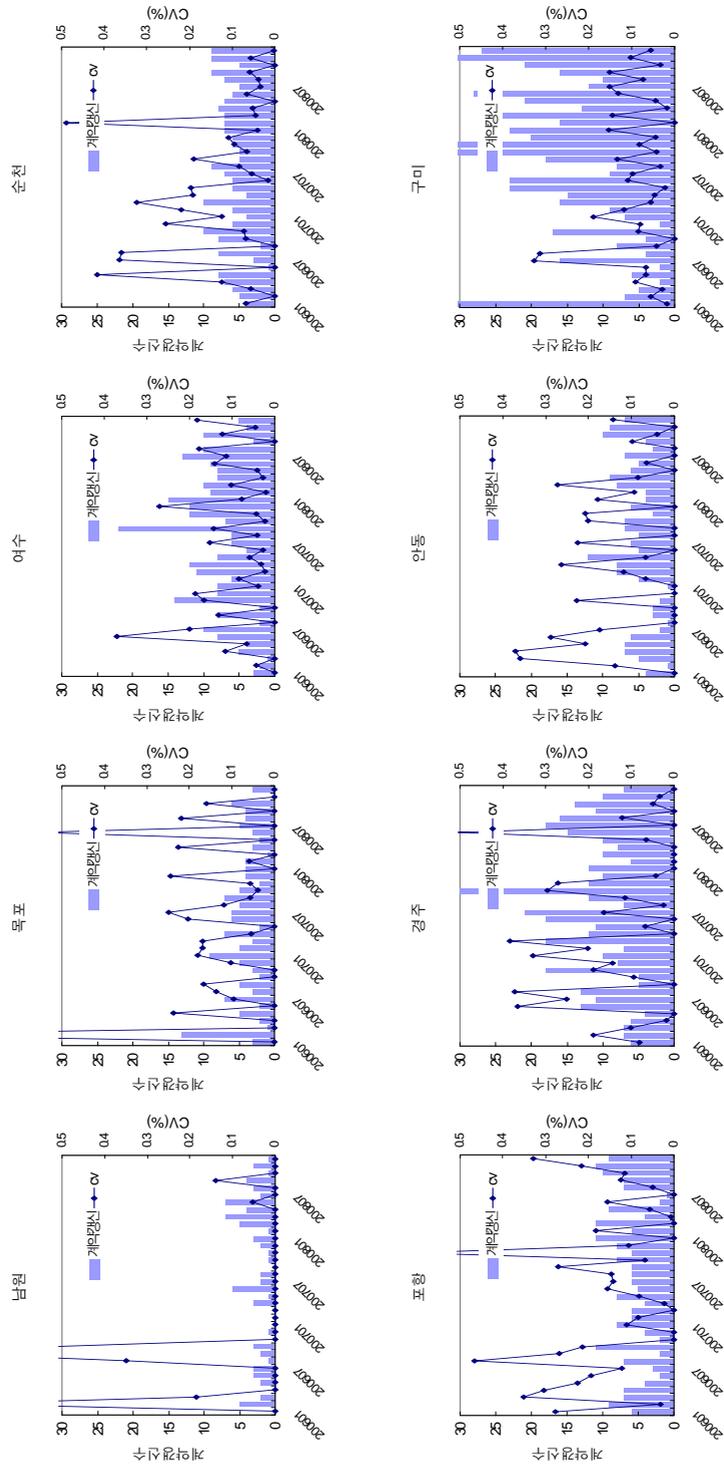
〈부록 2〉 도시별 평균 변동률의 CV(%)와 계약갱신수(계속)



〈부록 2〉 도시별 평균 변동률의 CV(%)와 계약갱신수(계속)



〈부록 2〉 도시별 평균 변동률의 CV(%)와 계약맹신수(계속)



〈부록 2〉 도시별 평균 변동률의 CV(%)와 계약갱신수(계속)

