

주거비용에 영향을 미치는 요소 분석: 시스템다이내믹스 계수추정을 위한 다층모형과 회귀모형의 비교

Determinants of Housing Cost: Hierarchical Linear Model for Estimating Coefficients of a Housing System Dynamics Model

강명구*

Kang, Myoung-Gu*

Abstract

To measure the effect of school zone on housing cost, Linear Regression Model is widely used, and school zone is known as a key determinant of housing cost in Korea. However, when the Hierarchical Linear Model (HLM) is applied with the same data, school effect on housing cost becomes statistically non-significant. It is because HLM effectively separates the effect of individual housing's attributes from the group effect. In sum, the housing cost of Kangnam, where good public schools are located, is apparently is higher than that of Kangbuk. However, the school effect on housing cost (Level 2) becomes non-significant when individual housing's attributes (Level 1) are controlled with HLM.

Keywords: 다층모형, 선형위계모형, 시스템다이내믹스, 주택, 주거비용
(Hierarchical Linear Model (HLM), System Dynamics, Housing)

* 서울시립대학교 도시공학과 교수 (mk@uos.ac.kr)

I. 서론

주택가격에 영향을 미치는 요소에 대한 논의나 계량적 분석은 많으나, 그에 비하여 주거비용¹⁾, 즉 임대료, 에 대한 논의는 매우 적다. 그 주된 이유는, 주택가격은 임대료의 자산화된 가격으로 여겨지면서 주택가격과 임대료가 같이 움직이는 것으로 생각하고 있기 때문이다. 즉, 임대료가 높아지면 주택가격이 오른다든지 또는 주택가격이 올라가면 임대료가 올라갈 것이라는 가정이다. 따라서, 주택가격에 미치는 영향과 임대료에 미치는 영향이 같을 것이라고 인식되고 있다.

하지만, 최근 몇년간의 한국의 경험을 보면 주택가격과 임대료가 별개로 움직이고 있다. 따라서, 임대료에 영향을 미치는 요소는 주택가격에 영향을 미치는 요소와 다를 수 있는지 검증할 필요가 있다.

이에 본 연구에서는, 주택가격과 임대료에 대한 개념적 정의를 명세화하고, 이를 바탕으로 임대료에 영향을 미치는 요인을 살펴보고자 한다. 주택가격은 크게 두가지 성분으로 구성되어 있다. 하나는 주택을 사용하면서 얻게되는 이용편익이다. 예를 들어, 주택의 평수, 지하철역까지의 거리, 또는 학군 등은 그 집에 거주하면서 누리는 편익에 해당된다. 또 하나의 성분은 자산으로서의 가치로서 미래기대수익이다. 이는 현재 사용하면서 누리는 이익은 아니지만, 미래에 상승할지 모르는 기회비용에 대한 기대치의 현재가치이다. 미래기대수익에 영향을 미치는 요소는 이용편익에 영향을 미치는 요소보다 더욱 다양할 뿐만아니라 미래라는 많은 불확실성과 모호성을 내포하고 있기 때문에 직접 계량화하기에는 많은 어려움이 있다.

본 연구에서 주거정책의 궁극적 대상은 주거비용(즉, 임대료) 이므로 주택 임대료에 초점을 맞추어, 그에 영향을 미치는 요소를 분석하고자 한다. 이는 한편으로 분석에 있어서의 불확실성을 많이 줄여주기도 한다.

이에 더불어, 본 연구에서는 연구방법론적인 측면에서 다층화된 구조를 파악하기 위한 다층모형분석을 적용하여 분석하는 것이 기존의 회귀분석모형을 사용하는 것보다 더 적절함도 함께 논의해 보고자 한다. 각각의 주택은 그 주택이 속해있는 커뮤니티에 속해 있다. 분석의 단위인 개별 주택의 임대료는 커뮤니티의 영향을 받게 되며, 그 영향은 커뮤니티별로 선형적으로 나타나지 않고 임의적으로 차이가 날 수 있다. 이런 경우, 개별 주택을 독립적이라고 말할 수 없다. 이는 회귀분석모형의 주요한 전제조건 중 하나인 분석단위의 독립성 가정을 만족하지 못함을 의미한다.

1) 개념적으로, 주거비용이란 주택과 토지에 대한 소비 지출로 정의된다. 주거비용에 대한 다양한 정의가 있으나, 본 연구의 목적상 주거비용과 임대료를 동일한 개념으로 사용한다.

이러한 위계적 특성을 갖고 있는 대상을 개별 주택별로만 분석 할 경우 집계화의 오류와 생태학적 오류를 범할 수 있다. 이렇게 다층자료에서 표집단위인 집단의 임의효과가 존재하는 경우에는 다층모형이 회귀분석모형보다 더 적절하다. (강상진, 1998)

II. 이론적 고찰

주거비용에 영향을 미치는 요인에 대한 직접적인 연구는 매우 드물다. 대신, 주택가격에 대한 연구는 매우 많다.

복잡하고 다양한 경로를 통하여 상호 인과관계를 지닌 주택가격은 시스템적 접근이 필요한 분야이며, Forrester (1969) 를 비롯한 여러 사람들에 의해서 시스템다이나믹스를 활용하여 설명하려 하였다. Forrester 이후 몇개의 도시동태모형이 더 논의되었지만 이들은 도시를 거시적으로 보고, 인구, 산업, 그리고 주택이라는 하위부문간의 연관관계를 살펴보고 있다.

이에, 권혁일과 이만형 (2005) 은 “동태적 주택가격모형과 응용”이라는 연구에서 시스템다이나믹스를 이용한 동태적 주택가격모형을 구축하면서, 현재 상태의 주거비용 결정요인들을 찾아내고 예측하는 횡단면분석을 추가로 보완할 필요가 있다고 지적하고 있다. 이를 통해서 자의적인 변수처리와 계수값의 임의적 결정과 같은 주관적 요소를 가능한 줄일 수 있다고 말하였다. 즉, 미시적관계를 보여주는 모델을 더 연구할 필요가 있다.

본 연구에서는 주택가격을 예측하고자 하는 노력의 연장선에서 시스템다이나믹스 모형 구축에도 활용될 수 있는 이용가치에 대한 보다 미시적이고 정밀한 분석을 시도하고자 한다.

주택가격은 다음과 같이 크게 두개의 부분으로 구성되어 있다. 이용가치는 다시 (1) 직접 이용편익, 즉 평수 또는 건축년수 등, 과 (2) 간접이용편익, 즉 공공서비스 또는 주변여건 등, 으로 구분할 수 있다. 이는 그 곳에 거주하면서 누리는 사용 가치라고 할 수 있다. 한편, 주택은 자산으로서의 가치도 내재하고 있어 향후 미래기대수익에 대한 기대가 담겨 있다²⁾.

2) 미래기대수익의 대표적인 예는 재건축아파트의 가격이 될 수 있다. 예를 들어, 서울 강남지역의 20평도 안되는 아파트 가격이 10억원을 넘는 사례들이 있다.

$$\begin{aligned} \text{주택가격} &= \text{이용가치} + \text{자산가치}^3) \\ &= (\text{직접 이용편익} + \text{간접 이용편익}) + \text{미래기대수익} \end{aligned}$$

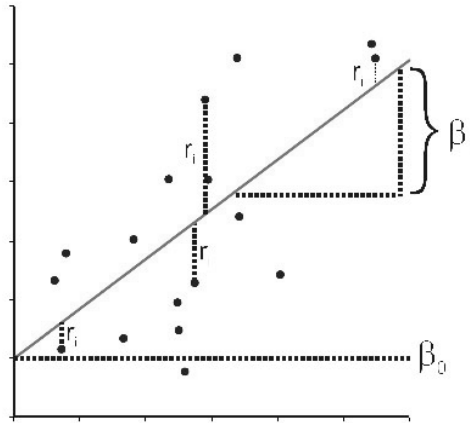
여기서, 주택정책의 궁극적인 목적은 주거복지 및 주거안정이라는 점이다. 즉, 좀 더 직접적으로 주거비용에 대한 연구가 필요하다고 할 수 있다. 더우기, 주택가격과 주거비용이 항상 같은 방향으로 움직이지 않거나, 그 두가지에 영향을 주는 요소 및 그 요소들의 영향 정도가 다르다면 주거비용에 대한 연구는 절실히 요구된다.

주택연구에서 주의할 점은, 여러학자가 지적하듯이 관련 변수들이 위계적 구조를 가지고 있다는 것이다. 위계적 구조를 가진 대상을 전통적인 통계모형인 일반선형회귀모형등을 사용하게되면, 분석의 단위, 회귀계수의 변산, 공상관요인, 층위간 상호작용, 그리고 신뢰도추정으로 인한, 연구의 타당성을 잃게될수 있다. 다시말하면, 일반선형모형에서 정상분포나 등분산의 가정과는 달리, 분석단위의 독립성의 가정이 지켜지지 않으면, 가설검정의 타당성은 심각하게 위협을 받는다 (이재열 외, 2005).

따라서, 본연구에서는 위계적 다층모형분석을 사용함으로써 연구방법의 타당도를 높이고자 한다.

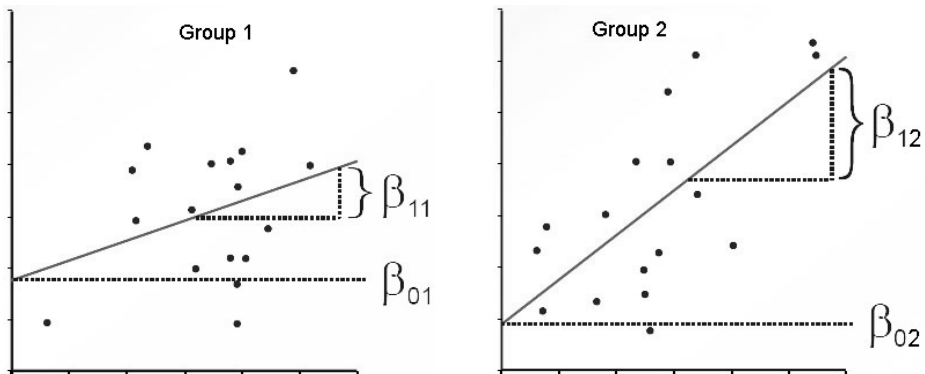
이를 그림으로 표현하면 다음과 같다. 개별 주택의 임대료에 영향을 주는 요인을 추정함에 있어 주택의 물리적 변수들 (예, 주택면적) 과 그 주택이 속한 커뮤니티 변수 (예, 학군) 를 일률적으로 회귀분석을 적용할 경우, <그림 1> 에서 보듯이 하나의 회귀분석 결과가 도출된다.

3) 주택가격에 관하여, 주택을 소유한다는 그 자체만으로도 연계되는 만족감이 있다는 의견이 있으나, 이는 좀 더 연구가 필요한 것으로 생각한다. 본 연구에서는 이용편익에 집중하므로 이 부분에 대한 논의는 다른 논문에서 다루겠다.



<그림 1> 집단화된 회귀분석모형

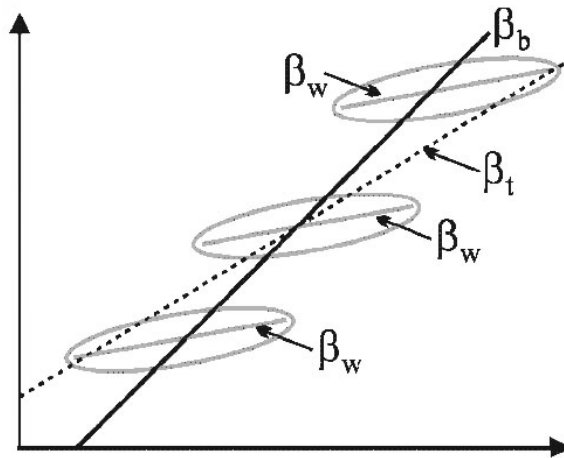
하지만, 이는 다음 <그림 2> 에서 볼 수 있는 것과 같이, 성격이 다른 두 집단의 특성이 무시된채 집계화의 오류를 범할 수 있게 된다.



<그림 2> 그룹별 변수간 연관성

<그림 2>에서 보듯이 개별주택의 변수 (예, 주택면적) 가 미치는 영향이 그룹별 (예, 학군) 로 다른 영향을 보일 수 있음을 시사한다. 이를 정리하면 다음 <그림3>과 같이 정리해 볼 수 있다. 즉, 회귀분석만을 사용할 경우 두 변수간의 관계가 β_t 로 보여질 수 있으나,

그룹별 통제를 할 경우 그 연관정도는 사실 β_w (Effect of Within-Group)이었을 것이다. 그리고 β_b (Effect of Between-Group)는 그룹에 따른 연관성 정도를 구별시켜준다.



〈그림 3〉 회귀분석과 다층모형분석

이에 따라, 본 연구에서는 회귀분석과 다층모형분석을 통해서 기존에 영향이 크다고 알려진 학군의 영향을 재고찰 해보고자 한다. 이를 위해서 주택의 크기와 경과년수 등을 제어변수로 사용할 것이다. 집단간의 차이로는 학군의 차이에 의한 영향을 고찰해 볼 것이다.

Ⅲ. 연구방법 및 결과

1. 자료와 통계 프로그램

본 연구에 사용된 자료는 2004년 7월에 서울시에서 실시한 “서울 및 수도권 주민 주거실태 및 정책수요 조사”이다. 본 연구의 목적상 다음과 같은 기준에 따라 6개구 63개 가구의 샘플을 추출하였다.

1단계: 지역적 대상은 서울로 한정하였다. 서울 주변지역을 포함시킬 경우 주택가격에 영향을 미치는 요소가 다양해 질 수 있어, 주변지역을 포함한 분석은 다른 연구에서 진행

하는 것으로 한다.

2단계: 조사에서는 여러가지 주거타입이 있으나, 주거 타입에 따른 변이를 줄이기 위하여 아파트만 추출하였다.

3단계: 아파트 경과년수가 20년 초과하는 경우는 제외하였다. 20년이 초과하는 경우 재건축 또는 리모델링이라는 가능성이 생기면서 가격에 불특정 왜곡이 발생하는 경우가 있어 본 분석에서는 20년 미만인 아파트만을 대상으로 하였다.

4단계: 본 연구에서 위계선형모형분석법을 사용함으로써 집단내와 집단간 비교를 할 필요가 있는데, 서울시 25개 구 중에서 추정할 계수들에 비하여 충분한 샘플갯수를 갖고 있지 않은 구들은 제외하였다.

최종적으로 6개 구 - 강남구, 서초구, 강서구, 마포구, 중랑구, 그리고 도봉구 - 63개 가구가 선택되었다.

또한 본 연구의 계량분석처리를 위해서는 SPSS 11.0.1. 버전과 Stephen Raudenbush, Tony Bryk, 그리고 Richard Congdon이 개발한 HLM 6.0를 사용했다

2. 연구문제

앞서 서술한 이론적 고찰과 자료특성에 근거하여 다음의 다섯 가설을 세우고, 계량분석을 통해서 각각의 연구문제에 대한 응답을 구하고자 한다. 그 각각의 가설을 요약하면 다음과 같다.

첫째, 서울지역에서 구별 아파트 전세가격은 차이가 있는가?

둘째, 만일, 서울지역에서 구별에 아파트 전세가격차이가 있다면, 학군이 어느정도 영향을 미친다고 할 수 있는가?

셋째, 학군을 고려하지 않고, 주택의 특성인 면적, 경과년수, 대중교통과의 접근성, 그리고 공원과의 접근성은 아파트 전세가격에 영향을 주는가?

넷째, 종합하여, 면적, 경과년수, 대중교통과의 접근성, 그리고 공원과의 접근성을 고려해서 얻은 예상 평균 주택가격은, 구별수준에서의 특성인 학군변인의해 그 차이를 보이는가?

다섯째, 같은 자료를 사용하여 회귀분석을 한 결과와 위의 다층모형의 결과가 상이한가?

3. 분석방법 및 결과

1) 구별간 아파트 전세가격은 차이가 있는가?

본 연구의 첫번째 연구과제는, 서울시 아파트 전세가격의 분산이 어느 정도 아파트들이 가지고 있는 속성에 의한 차이인지, 또한 어느 정도가 지역구별간의 차이에 의한 것인지를 밝히고자 하는데 그 목적이 있다. 따라서, 여기에서는 위계적 계층모형의 기본모형인 One-way ANOVA (일원분산분석) 모형을 사용하며, 아파트 전세가격의 분산을 구별내 (집단내) 분산과 구별간 (집단간) 분산으로 분할 추정하고자 한다.

다음은 위계적 계층모형 중, One-way ANOVA (일원분산분석통계) 모형을 수리적으로 정리한 것이다.

$$\text{Level-1 Model} \quad Y_{ij} = \beta_{oj} + r_{ij} \quad r_{ij} \sim N(0, \sigma^2)$$

$$\text{Level-2 Model} \quad \beta_{oj} = \gamma_{00} + \mu_{oj} \quad \mu_{oj} \sim N(0, \tau)$$

위의 모형에서 Y_{ij} 는 j 번째 구의 i 번째 아파트의 전세가격이며, β_{0j} 는 j 번째 구에 속한 아파트들의 평균 전세가격이며, r_{ij} 는 각각의 아파트들이 가지고 있는 개별효과이다. γ_{00} 는 전체 아파트들의 평균 전세가격이며, μ_{0j} 는 j 번째 구의 고유효과이다. 또한 σ^2 는 구별내 아파트들의 전세값차이에 의한 전세가격의 분산이며, τ 는 구별간의 차이에 의한 아파트 전세가격의 분산이다.

본 연구의 통계적 처리를 위해서 HLM 6.0 를 사용하였으며, 각각의 결과값은 APPENDIX에 첨부하였다. HLM 6.0에서 기본으로 사용하는 추정법은 최대우도추정법이다. 최대우도추정법에 의한 표본아파트들의 전체평균 전세값은 약 1억2천 만원이며, 표준오차는 1천3백만원이다.

부록의 Null Model 통계결과에서 보듯이, 각각의 아파트별 수준에서의 분산 (σ^2)은 8,638,991이고 각 구별 수준에서의 분산 (τ) 은 10,462,273 이다. 이와 같은 결과값은 아파트 전세가격의 분산의 차이는 구별간 차이와 아파트 특성에 기인한 차이가 반반 정도이다. 이에 대한 내용을 요약한 것이 집단내 상관계수 ICC (intra-class correlation)이며 그 값은 정리한 것은 다음과 같다.

$$ICC(\rho) = \tau / (\sigma^2 + \tau) = 10462273.15 / (10462273.15 + 8638991.02) = 0.548$$

따라서, 약 54.8%의 분산이 구별간의 차이에 의한 것임을 알 수 있고, 45.2%의 분산이 구별내에 속해있는 아파트 특성에 따른 차이임을 나타낸다.

구별 평균 전세가격대 한 구별간의 차이가 있는지에 대한 가설에 대해서는 [$X^2=63.93085$, (df=5), P-value= 0.000]이므로 통계적으로 유의미하게 구별 평균 전세가격은 구별간 차이가 있다고 본다.

2) 구별간 학군차이만으로, 아파트 전세가격의 차이를 설명할 수 있는가?

본 연구의 두번째 연구과제는, 아파트 전세가격의 차이가 구별간 유의미한 차이가 있다면, 구별간의 어떤 특성이 이러한 차이를 만드는가를 밝히는 분석의 다층모형분석의 두번째 단계이다. 이는 위계적 계층연구의 두번째 방법인 Means-as-outcomes model로 검증해 볼 수 있다. 이 모형은 전단계에서 활용한 One-way ANOVA model에, Level 2인 구별수준에 독립변인으로서 학군차이를 추가함으로써 학군차이에 의한 아파트 전세가격이 차이가 생기는지를 판별하고자 한다. 이때 학군변인은 dummy 변인 (8학군 = 1, 비8학군 = 0)으로 사용된다.

다음은 Means-as-outcomes Model을 수리적으로 정리한 것이다.

$$\text{Level-1 Model} \quad Y_{ij} = \beta_{0j} + \gamma_{ij}, \gamma_{ij} \sim N(0, \sigma^2)$$

$$\text{Level-2 Model} \quad \beta_{0j} = G_{00} + G_{01}(\bar{x}_j - \bar{z}) + \mu_{0j}, \mu_{0j} \sim N(0, \tau)$$

위의 모형에서, Y_{ij} 는 j 번째 구의 i 번째 아파트의 전세가격이며, β_{0j} 는 j 번째 구에 속한 아파트들의 평균 전세가격이 된다.

HLM 6.0. 결과값에 의하면, [t-value=0.603 with df=4, p-value= 0.579]이며, 따라서, 구별 전세가격이 학군에 의해서만 차이가 난다고는 말 할 수 없을 확인 할 수 있다. HLM6.0.에 의한 본 모형의 결과값은 APPENDIX에 정리해 놓았다.

3) 어떤 변인이 구별에 따른 아파트 전세가격에는 차이를 만드는가?

본 연구의 세번째 연구과제는 어떤 변인이 구별간의 아파트 전세가격의 차이를 형성하는지를 알아보는 것이다. 다시말하면, 아파트 전세가격이 다르게 형성되는데에는 구별간의 차이에 의한 원인도 있을 수 있지만, 각각의 개별 아파트들이 가지고 있는 속성때문에 아파트 전세가격이 다르게 형성될 수 있다는 가정이다.

이러한 연구과제는 위계계층모형분석중의 하나인 모형인 One-Way ANCOVA model를

사용하여 해결할 수 있다. 즉 이 모형은 주택수준(Level 1)에서 종속변인인 아파트 전세가격에 영향을 줄수 있는 독립변인들 (면적, 주택경과년수, 대중교통과의 접근성, 그리고 공원 등 녹지시설과의 접근성) 과의 관계를 밝히고자한다.

이때, Grand-mean Centering이라는 방법을 사용하게 되는데, 이 방법은 각각의 독립변인들을 통계적으로 처리함으로서, 이 속성들이 어떻게 아파트 전세가격형성에 영향을 주는지 여부를 밝힐 수 있는 장점을 가지고 있다.

다시말하면, Grand-mean Centering를 함으로써, 절편값 (β_{0j})값은 실제 평균 아파트 전세값에서 각각의 아파트들이 가지고 속성을 고려한 예상 평균 아파트 전세값으로 변한다. 따라서, 구별 평균 아파트 전세값은 구별내 속해있는 아파트들이 가지고 있는 속성에 따라서 조정된 값들이 된다.

다음은 위계적 계층 모형 중, One-way ANCOVA 모형을 수리적으로 정리한 것이다. 이때, 진한 이탤릭으로 표기된 변인은 Grand-mean Centering값을 표시한것이다.

Level-1 Model

$$Y_{ij} = \beta_{0j} + \beta_{1j}*(AREA) + \beta_{2j}*(YEAR) + \beta_{3j}*(TRANSP) + \beta_{4j}*(PARK) + \gamma_{ij}, \gamma_{ij} \sim N(0, \sigma^2)$$

Level-2 Model

$$\beta_{0j} = G_{00} + \mu_{0j} \quad , \quad u_{0j} \sim N(0, \tau)$$

$$\beta_{1j} = G_{10}$$

$$\beta_{2j} = G_{20}$$

$$\beta_{3j} = G_{30}$$

$$\beta_{4j} = G_{40}$$

HLM 6.0. 결과값에 의하면, 아파트별 면적(G10)는 [t - value=7.970 with df=58, p-value= 0.000] 으로 통계적으로 유의미하게 아파트별 평균 전세가격차이를 설명하는데 도움을 준다. 또한, 아파트별 경과년수(G20)는 [t - value=-2.338 with df=58, p-value= 0.023] 으로 통계적으로 유의미하게 아파트별 평균 전세가격차이를 형성하는데 영향을 준다. 그러나, 아파트별 대중교통과의 접근성(G30)과 공원 등 녹지등에 대한 접근성(G40)는 각각 [t - value=0.428 with df=58, p-value= 0.670], [t - value=-0.272 with df=58, p-value= 0.786] 으로 아파트별 전세가격차이에 영향을 준다고 말할 수 없다. HLM 6.0. 에 의한 본 모형의 결과값은 APPENDIX에 정리해 놓았다.

6) 면적, 경과년수, 대중교통과의 접근성, 그리고 공원과의 접근성을 고려해서 얻은 예상 평균 주택가격은, 구별수준에서의 특성인 학군변인의해 그 차이를 보이는가?

본 연구의 네번째 연구과제는, 면적, 경과년수, 대중교통과의 접근성, 그리고 공원과의 접근성을 고려해서 얻은 예상 평균 주택가격은, 구별수준에서의 특성인 학군변인의해 그 차이를 보이는지를 검증하는 것이다. 다시말하면, 비록 학군만으로 아파트 전세가격의 차이를 설명한다고는 말할 수 없으나, 만일 아파트 각각이 가지고 있는 속성을 고려한 상태에서 예상되는 아파트 전세가격은 학군에 의해 영향을 받을 지를 검증하고자 한다.

본 연구과제를 위해서는 위계계층모형분석측중의 하나의 모형인 Intercepts - and Slope - as - Outcomes Model를 사용하여 해결할 수 있고자 한다. 즉 이 모형은 주택수준(Level 1)에서 종속변인인 아파트 전세가격에 영향을 줄수 있는 독립변인들로서, 면적, 주택경과년수, 대중교통과의 접근성, 그리고 공원 등 녹지시설과의 접근성을 고려하고, 구별수준(Level2)에서는 구별속성변인인 학군변인을 독립변인으로 투입함으로써 아파트 전세가격변동에 대한 학군효과를 보고자 한다.

다음은 위계적 계층 모형 중, Intercepts-and-Slopes-as-Outcomes 모형을 수리적으로 정리한 것이다. 이때, 진한 이탤릭으로 표기된 변인은 Grand-mean Centering 값을 표시한것이다.

Level-1 Model

$$Y_{ij} = \beta_{0j} + \beta_{1j}*(AREA) + \beta_{2j}*(YEAR) + \beta_{3j}*(TRANSP) + \beta_{4j}*(PARK) + \nu_{ij}, \nu_{ij} \sim N(0, \sigma^2)$$

Level-2 Model

$$\begin{aligned} \beta_{0j} &= G_{00} + \mu_{0j}(\text{학군}) \quad , \quad u_{0j} \sim N(0, \tau) \\ \beta_{1j} &= G_{10} \\ \beta_{2j} &= G_{20} \\ \beta_{3j} &= G_{30} \\ \beta_{4j} &= G_{40} \end{aligned}$$

HLM 6.0. 결과값에 의하면, 앞서 모델에 근거한 결과값과 동일하게 아파트별 면적(G10)과 아파트별 경과년수(G20)는 각각 [t - value=8.038 with df=57, p-value= 0.000], 는 [t - value=-2.547 with df=57, p-value= 0.014]으로 아파트별 전세가격차이에 영향을 준다고 할 수 있다. 그러나, 아파트별 대중교통 접근성(G30)과 공원 등 녹지 접근성(G40)는 각각 [t - value=0.178 with df=57, p-value= 0.860], [t - value=-0.014 with df=57, p-value= 0.918] 으로 아파트별 전세가격차이에 영향을 준다는 통계적 유의성을 발견할 수 없다.

특히, 학군(G01)은 [t - value=1.349 with df=4, p-value= 0.249] 으로 아파트별 전세가격차이에 영향을 준다는 통계적 유의성은 없다. 다시말하면, 전세가격의 측면에서는 학군

의 영향이 과대평가되어있을 수도 있음을 시사한다. HLM 6.0.에 의한 본 모형의 결과값은 APPENDIX를 참조하길 바란다.

5) 같은 자료를 사용하여 회귀분석을 한 결과와 위의 다층모형의 결과가 상이한가?

SPSS 11.0.1. 버전을 사용하여 구한 회귀분석결과와 HLM 6.0. 버전을 사용하여 구한 다층모형분석결과를 비교한 결과, 두 분석방법 모두, 주택면적, 경과년수는 아파트 전세값차이형성에 유의미한 영향을 주며, 대중교통과의 접근성과 공원과의 접근성은 아파트 전세값 차이 형성에 유의미한 영향을 주지 않음을 보여준다.

그러나, 학군효과에 대한 분석결과는, 회귀분석의 경우, 각각 [t -value=3.338 with $df=59$, p -value= 0.001]로서 전세값과 학군간의 연관성이 통계적으로 유의미하다고 보여준다. 하지만, 위에서 보았듯이 다층모형분석의 경우는 [t -value=1.349 with $df=4$, p -value= 0.249] 으로 아파트별 전세가격과 학군의 연관성이 통계적으로 유의미하지는 않다고 보여준다.

IV. 결론 및 제언점

먼저, 전세값 형성에 미치는 학군효과에 대해서보면, 전통적으로 한국은 전세값 형성에 지대한 영향을 준다고 밝혀왔다 (송명규, 1992). 그러나, 다층모형분석결과와 회귀분석결과를 비교해 본 결과, 다층모형분석결과가 회귀분석 결과와 다르게, 다른 변수들을 통제한 뒤에는, 학군이 전세가격에 미치는 영향이 통계적으로 유의하지 않음을 보인다. 이는 표면적으로 강남구와 서초구의 전세가격이 다른 구에 비해서 비싼 것으로 비쳐질 수 있으나, 다층모형분석결과, 개별주택특성(Level 1)과 구별특성(Level 2)를 효과적으로 분리함으로써 구별특성 (Level 2)에 직접적인 영향을 주는 학군변인의 영향을 좀 더 정확히 변별해 내었다고 할 수 있다.

다만, 본 연구결과는, 제한적인 자료를 사용하여 검증한 결과이므로 위의 결과를 일반화시켜 단정짓기는 아직 무리가 있는 것은 사실이다. 좀 더 풍부한 자료와 다양한 변수를 활용한 검증이 추가로 이루어져야 할 것으로 보인다. 적어도, 다층모형분석을 통해서 집계화의 오류에 대한 추가적인 검증이 필요하다는 것을 시사하는 바는 크다고 할 수 있다.

본 연구결과는 주택가격을 추정하는 시스템다이내믹스 모델을 구축하는데 중요할 수 있다. 시스템다이내믹스 모델내의 계수들을 추정할 때, 전통적인 회귀분석을 주로 사용하는

데, 그 영향을 미치는 구조가 계층화 되었있는 경우에는 다층모형분석을 통한 계수 추정이 더 정확할 수 있음을 시사한다.

【참고문헌】

- 권혁일, 이만형. (2005). 동태적 주택가격모형과 응용. 국토계획, 40 (3). pp. 21- 36
- 강상진a. (1998). 교육 및 사회연구를 위한 연구방법으로서 다층모형과 전통적 선형모형과의 비교분석 연구. 교육평가연구, 11(1). pp. 207-258.
- 강상진b. (2003). 회귀분석의 이해. 교육과학사
- 송명규. (1992). 학군의 질 (고등학교 수준)과 명성이 주택가격에 미치는 효과에 관한 실증적 연구: 서울시의 경우. 지역사회개발연구, 17 (1). pp. 91-106.
- 이재열외. (2005). 사회과학의 고급계량분석: 원리와 실제. 서울대학교출판부.
- 이종성외. (2007). 사회과학 연구를 위한 통계방법. 박영사.
- 정수연. (2006). 교육요인이 서울아파트가격에 미치는 영향에 관한 연구. 국토계획, 41 (2). pp. 153 - 166
- 최열, 권연화. (2004). 위계선형모형을 이용한 교육환경이 주택가격에 미치는 영향 분석. 국토계획, 39 (6). pp. 71- 82
- Forrester, Jay W. (1969). Urban Dynamic. Cambridge: The MIT Press.
- Rumberber, R.W., & Palardy G.J. (2004). Multilevel Models for School Effectiveness Research.
- Raudenbush, S.W., & Bryk, A.S. (2002). Hierarchical linear models: Applications and data analysis methods, second edition. Thousand Oaks, CA: Sage
- Raudenbush, S.W., Bryk, A.S., Cheong, Y.F, & Congdon, R (2004). HLM6: Hierarchical linear and nonlinear modeling. Lincolnwood,IL: Scientific software international.

【부록】

I. Descriptive Statistics

LEVEL-1 DESCRIPTIVE STATISTICS

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
X19 (주택면적)	63	26.08	5.39	15.00	38.00
X32 (경과년수)	63	9.84	5.85	1.00	18.00
X38 (교통시설과의 접근성)	63	10.71	7.01	5.00	30.00
X40 (공원과의 접근성)	63	16.59	10.47	5.00	50.00
X48 (전세가격)	63	12141.27	4101.24	5000.00	21000.00

LEVEL-2 DESCRIPTIVE STATISTICS

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
MAXOFX52 (학군)	6	0.33	0.52	0.00	1.00

II. SPSS 11.0.1 version Regression Output

Regression

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.814 ^a	.663	.635	2479.720

a. Predictors: (Constant), X52, X19, X40, X32, X38

Coefficients ^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1103.463	1715.875		-.643	.523
	X19	556.948	58.993	.737	9.441	.000
	X32	-174.766	61.588	-.250	-2.838	.006
	X38	-41.144	54.510	-.070	-.755	.453
	X40	-5.465	31.765	-.014	-.172	.864
	X52	2653.248	794.861	.319	3.338	.001

a. Dependent Variable: X48

III. HLM 6.1. Multilevel Output

1. Null Model (One-way ANOVA Model)

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling
 Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
 Publisher: Scientific Software International, Inc. (c) 2000
 techsupport@ssicentral.com
 www.ssicentral.com

The outcome variable is X48

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + R$$

Level-2 Model

$$B0 = G00 + U0$$

Random level-1 coefficient Reliability estimate

INTRCPT1, B0	0.925
--------------	-------

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	Approx. T-ratio	df	P-value
For INTRCPT1, B0					
INTRCPT2, G00	11981.094234	1373.292037	8.724	5	0.000

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	3234.54373	10462273.14644	5	63.93085	0.000
level-1, R	2939.21606	8638991.01849			

Statistics for current covariance components model

Deviance = 1183.344363
 Number of estimated parameters = 2

2. Means-as-Outcomes Model

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling
 Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
 Publisher: Scientific Software International, Inc. (c) 2000
 techsupport@ssicentral.com
 www.ssicentral.com

The outcome variable is X48

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + R$$

Level-2 Model

$$B0 = G00 + G01*(MAXOFX52) + U0$$

Random level-1 coefficient Reliability estimate

INTRCPT1, B0	0.935
--------------	-------

The outcome variable is X48

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	Approx. T-ratio	df	P-value
For INTRCPT1, B0					
INTRCPT2, G00	11345.667668	1809.908478	6.269	4	0.000
MAXOFX52, G01	1881.008048	3119.710267	0.603	4	0.579

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
---------------	--------------------	--------------------	----	------------	---------

INTRCPT1, U0	3490.31916	12182327.84203	4	60.74975	0.000
level-1, R	2938.62187	8635498.48397			

Statistics for current covariance components model

Deviance = 1163.261784
Number of estimated parameters = 2

3. One-way ANCOVA Model

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

The outcome variable is X48
Summary of the model specified (in equation format)

Level-1 Model
 $Y = B0 + B1*(X19) + B2*(X32) + B3*(X38) + B4*(X40) + R$
Level-2 Model
B0 = G00 + U0
B1 = G10
B2 = G20
B3 = G30
B4 = G40

Random level-1 coefficient Reliability estimate

INTRCPT1, B0	0.897
--------------	-------

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	Approx. T-ratio	df	P-value
For INTRCPT1, B0					
INTRCPT2, G00	11952.023416	830.086769	14.399	5	0.000
For X19 slope, B1					
INTRCPT2, G10	502.082561	62.993960	7.970	58	0.000
For X32 slope, B2					
INTRCPT2, G20	-142.510383	60.959365	-2.338	58	0.023
For X38 slope, B3					
INTRCPT2, G30	22.161368	51.796856	0.428	58	0.670 (fail to reject)
For X40 slope, B4					
INTRCPT2, G40	-8.441425	31.002213	-0.272	58	0.786 (fail to reject)

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	1925.96786	3709352.20593	5	45.07594	0.000
level-1, R	2078.60656	4320605.24615			

Statistics for current covariance components model

Deviance = 1096.315187
 Number of estimated parameters = 2

4. Intercepts and Slopes - as - Outcomes Model

Program: HLM 6 Hierarchical Linear and Nonlinear Modeling
Authors: Stephen Raudenbush, Tony Bryk, & Richard Congdon
Publisher: Scientific Software International, Inc. (c) 2000
techsupport@ssicentral.com
www.ssicentral.com

The outcome variable is X48

Summary of the model specified (in equation format)

Level-1 Model

$$Y = B0 + B1*(X19) + B2*(X32) + B3*(X38) + B4*(X40) + R$$

Level-2 Model

$$B0 = G00 + G01*(MAXOFX52) + U0$$

$$B1 = G10$$

$$B2 = G20$$

$$B3 = G30$$

$$B4 = G40$$

Random level-1 coefficient Reliability estimate

INTRCPT1, B0	0.880
--------------	-------

Final estimation of fixed effects:

Standard	Approx.
----------	---------

Fixed Effect	Coefficient	Error	T-ratio	df.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	11174.189235	961.488357	11.622	4	0.000
MAXOFX52, G01	2295.718020	1701.751753	1.349	4	0.249 (fail to reject)
For X19 slope, B1					
INTRCPT2, G10	502.821870	62.555432	8.038	57	0.000
For X32 slope, B2					
INTRCPT2, G20	-156.928980	61.624804	-2.547	57	0.014
For X38 slope, B3					
INTRCPT2, G30	9.321052	52.511260	0.178	57	0.860 (fail to reject)
For X40 slope, B4					
INTRCPT2, G40	-3.246257	31.191572	-0.104	57	0.918 (fail to reject)

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	1761.25755	3102028.14095	4	31.35733	0.000
level-1, R	2078.41905	4319825.75752			

Statistics for current covariance components model

Deviance = 1076.002673
 Number of estimated parameters = 2