

부산 아파트 실거래가를 이용한 전통적 헤도닉모형과 공간계량모형간의 적합도에 관한 비교 연구

Comparative Study of the Fitness between Traditional OLS Models and Spatial
Econometrics Models Using the Real Transaction Housing Price in the Busan

김 성 우 (Kim, Sung Woo)*

정 건 섭 (Chung, Kyouun Sup)**

< Abstract >

The study is investigated the comparison in terms of $Adj R^2$ (i.e., goodness of fit) and several indicators such as LR(Log Likelihood), AIC(Akaike Information), SC(Schwarz Criterion) between traditional OLS(ordinary lest square) and spatial econometrics models. In order to do so, the traditional OLS models are used such as (i) the basic hedonic price regression model, (ii) basic model plus regional dummy model, and (iii) basic model plus regional dummy model plus quarterly time dummy model. On the other hand, the spatial econometrics models are used such as (i) SAR(spatial autoregressive model), (ii) SEM(spatial error model), and (iii) SAC(general spatial models). The major finding of the study is that the SAC is the best fit among several models considering the various indicators. And the SAC is parsimonious in terms of model specification, keeping explanatory power or better comparing other models. The study has a limitation, though. Maybe, the OLS is superior to the spatial econometrics models if the research area can be segmented into the correct sub-markets, which will be further studied.

주 제 어 : 공간계량모형, 헤도닉 모형, 공간적 자기상관

Keywords : spatial econometrics models, hedonic model, spatial autocorrelation

* 부경대학교 공공정책연구소 연구원, swkim@pknu.ac.kr (주저자)

** 부경대학교 행정학과 부교수, kschung39@pknu.ac.kr (교신저자)

I. 서론

주택은 자연적·인위적인 외부의 위협으로부터 인간을 보호하기 위한 단순한 은신처(shelter)의 개념뿐만 아니라 오늘날 경제적 재화, 사회적 신분의 상징 등의 한 부분을 차지하면서 주택을 둘러싼 주택시장이 갖는 중요성은 매우 크다. 주택시장 참여자간의 협상에 의해 결정되는 주택가격은 주택시장이 갖는 다양한 환경적 변화요인에 대한 정보를 내포하고 있어, 정부의 주택에 대한 정책 자료뿐만 아니라 금융, 건설 및 국내·외의 주요한 경제지표로서 활용되고 있다. 따라서 주택가격에 대한 예측력 높은 모델 개발이 갖는 의미는 크다.

일반적으로 주택가격 평가에 사용되는 방법은 주택을 구성하는 개별요소들의 잠재적 가격을 측정하여 이들 요소의 합으로 주택가격을 추정하는 헤도닉가격모형(hedonic price model)을 이용하고 있다(김주영·김주후, 2002:22). 헤도닉가격모형(hedonic price model)은 Rosen(1974)의 다중적 특성을 가진 단일재화시장에 대한 연구를 기초로 외국의 경우 1970-80년대부터 주택가격 추정연구에 이용되어 왔다. 최근 우리나라에서도 부동산의 가치를 산정할 때뿐만 아니라 부동산 주위의 환경적 가치를 평가하거나, 건설업체의 브랜드 가치를 평가할 때에도 헤도닉가격모형(hedonic price model)을 사용하고 있다. 나아가 주택가격 지수를 작성할 때에도 헤도닉가격모형이 사용되고 있다(이용만, 2008:81).

헤도닉가격모형(hedonic price model)에서 주택은 물리적 또는 지역적 속성 등 다양한 특성을 지닌 것으로 보고, 이들 개별특성들의 내재적 가치의 합으로써 가격이 설명된다는 것이다. 주택

등의 자산가치는 부지나 규모, 종류, 직장과의 통근거리, 공공시설 이용 편리성 이외에도 대기오염, 소음, 자연경관, 상하수도 등과 같은 생활환경 여건의 속성 등으로 인하여 지역마다 다르다. 따라서 헤도닉모형을 추정하는 일련의 과정을 통하여 자산가치에 영향을 미치는 물리적 속성, 주변속성, 교통 요인, 환경적 속성 등의 효과를 분석할 수 있다(양진우, 2000:234). 이러한 헤도닉가격모형 관련 연구들은 주택의 입지적 고정성을 인한 주택이 가지고 있는 단지 및 입지적 특성, 기타 환경 등 주택이 갖고 있는 개별특성 변수를 찾아 설명력(R^2)이 가장 높은 모델을 설정하는데 관심을 집중하였다.

많은 선행 연구들의 이론적 기반으로 삼고 있는 헤도닉가격모형(hedonic price model)은 주택가격함수 추정과 관련하여 다중공선성(multicollinearity)과 공간적 자기상관(spatial autocorrelation), 공간적 이분산성(spatial heterogeneity)에 대한 문제점을 항상 내포하고 있다.

기존 헤도닉 가격 모형에 따른 문제점을 보완하기 위해 최근 공간계량모형에 대한 관심이 높아지고 있다. 대표적인 연구로 Anselin(1988)은 공간효과를 적용하기 위하여 종속변수와 오차항에 공간가중행렬을 적용한 일반공간모형(SAC: general spatial models)을 기준으로 공간자기회귀모형(SAR: spatial autoregressive model), 공간오차모형(SEM: spatial error model)을 제시하였다. 특히 공간계량모형이 기존 헤도닉가격모형의 최소제곱법(OLS: ordinary least square)에 따른 오차항이 상호 독립적이라는 가정으로부터 자유롭다는 측면에서 모수 추정의 편의(bias)를 줄일 수 있다.

따라서 본 연구는 기존 헤도닉가격모형에서

다중공선성(multicollinearity)이 발생하지 않는 범위 내에서 독립변수 개수의 확장에 따른 설명력(R^2)의 변화와 Anselin(1988)이 주장한 공간계량모형과의 모형의 설명력을 비교 평가함으로써 주택가격추정에 보다 효율적인 추정모형을 제시하고자 한다. 즉, 주택의 특성 변수뿐만 아니라 지역변수와 시간변수를 포함한 32개의 독립변수를 이용한 전통적인 헤도닉 모형과 7개의 주요변수를 이용한 공간계량모형간의 적합도를 비교 분석하고자 한다. 이를 위해 2006년부터 2009년 2분기까지 국토해양부에 공개된 부산지역 아파트 실거래가 5,850자료를 대상으로 분석하였다.

본 연구의 구성은 II장에서 주택가격 추정에 있어 기존 헤도닉가격모형과 공간계량모형에 대한 선행연구를 살펴보고, III장에는 분석모형을 살펴보기로 한다. IV장에서는 분석에 이용된 자료와 변수에 대하여 설명하고, V장에서는 기존 헤도닉가격모형과 Anselin(1988)이 제안한 공간계량모형간의 적합도를 비교 하고자 한다. 마지막으로 VI장에서는 결과에 대한 요약과 향후 연구방향을 제시하고자 한다.

II. 선행연구

1. 전통적 헤도닉가격모형의 적용

주택가격 추정에 있어 Rosen(1974)에 의해 이론적 토대를 마련한 헤도닉가격모형(hedonic price model)은 주택구성요소들의 이질성을 실증적으로 분석할 수 있는 적절한 수단이라고 인식되면서 이후 많은 연구에서 일반적으로 사용되고 있다. Linneman(1980)은 미국 대도시권별로 주택가격

합수를 설정, 그 영향요인에 관하여 연구하였는데, 건축연수, 엘리베이터 유무, 상업활동 수준, 도로조건, 소음 등의 변수를 설정하여 부동산가격에 대한 헤도닉 가격합수를 추정하였다(양진우, 2000:234). Brookshire et al.(1982)은 LA지역의 주택매매자료를 이용하여 이산화질소와 미세먼지가 주택가격에 음(-)의 영향을 주고 있음을 밝혔다.

Dale-Johnson(1982)은 헤도닉가격합수를 이용하여 대기오염 등의 환경의 질에 관한변수, 주요직장 소재지로의 접근성, 학급규모, 방재수준, 1인당 재정지출 등 공공서비스수준에 관한 변수들을 사용하여 주택가격을 추정한 결과 접근성, 오염수준, 방수, 건축년도 등이 주택가격에 주요한 영향을 미치고 있다는 결과를 얻었다. Goodman and Thibodeau(1995)은 주택의 경과년수(age)은 주택가격에 미치는 효과가 비선형적 형태이며, 경과년수와 관련된 이분산이 존재 할 수 있음을 밝혔다. 한편 경과년수가 증가함에 따라 음(-)의 감가상각효과, 양(+)의 건설시기효과(vintage), 양(+)의 유지보수 및 개량효과(+)가 복합적으로 작용하고 있음을 규명하였다.

국내의 경우 김기호·이성우(1998)은 아파트가격에 대한 환경적 가치의 영향관계를 살펴보기 위해 부산지역 53개단지 369세대에 대한 조사 분석 결과 조망여부에 따라 주택가격이 차이를 보였으나, 바다와의 거리는 통계적으로 의미 있는 결과를 나타내지 못했다. 하지만 평형별로는 주택가격이 차이를 보이는 것을 밝혔다. 최종일·심성훈(2002)은 아파트가격과 대기질과의 관계를 분석한 결과 아파트 구조 및 접근성뿐만 아니라 아황산가스농도와 오존농도 간에 음(-)의 영향 관계가 나타나는 것을 밝혔다. 김태운 외(2007)은 주택에 내재된 경관가치를 추정하기 위하여 2006

년 1월부터 6월까지 실거래가 자료 중 분당의 경관자원을 조망 할 수 있는 40개 아파트단지 912가구를 대상으로 분석한 결과, 경관조망이 가능한 경우 주택가격에 양(+)의 영향을 미치고, 산, 하천, 공원의 개별 경관분석에서도 모든 경관적 요소들이 통계적으로 유의한 결과를 나타냈다. 송명규(1992)는 서울시 40개 행정 동에 거주하는 1,871개 가구 대상으로 교육환경이 주택가격 결정의 주요 요인임을 밝혔다. 오규석·이왕기(1997)는 환경의 질과 관계, 이상경·신우진(2001), 김석환·이현석(2005) 재건축 가능성과 주택가격간의 관계를 밝혔다.

2. 공간 계량모형의 적용

Anselin(1988)은 전통적인 선형모형에 대한 공간데이터 분석의 한계를 극복할 수 있는 다양한 공간계량모형이 제시하였다. 특히 주택가격에 미치는 공간효과를 공간적 종속성(spatial dependency)과 이분산성(heterogeneity)으로 구분하였다. 이모형을 시계열 분석의 자기상관을 공간차원에 적용한 것으로 종속변수나 오차항에 공간가중치를 적용하였다. Can(1992)은 ‘실제 부동산시장에서 이웃의 주택가격이 구매하고자 하는 주택가격에 하나의 변수로 작용 할 수 있다’는 가정 하에 공간 확장모형(spatial expansion model)에 대한 시도를 통해 전통적인 헤도닉모형에 공간적 자기상관이 존재한다는 것을 설명하였다.

국내 연구로 김종원(2000)은 서울시의 22개 자치구의 자가소유가구(609가구)와 세입자가구(512가구) 자료를 이용하여 전통적 OLS와 SAR, SEM모형을 비교 평가 하였다. 분석 결과 OLS모형은 일관되지 못한 추정결과를 초래할 가능성이

높으며, 자가와 전세모두 SAR모형이 OLS모형보다 더욱 적합한 것으로 나타났다. 서경천·이성호(2001)는 2000년도 부산지역 표준지 공시지가 자료 중 2,607개의 필지를 대상으로 일반적 헤도닉 접근방법(model 1)과 하위시장분할에 따른 구조적 특성변수(model 2), 주변지역 특성변수(model 3)의 모수추정방법에 대하여 OLS와 SAR이 결과를 비교 분석하였다. 분석결과 SAR분석 방법이 OLS보다 안정적이고 신뢰 할 수 있는 계수의 추정이 가능하다고 밝혔다. 박현수 외(2003)는 서울 광진구의 6개아파트 단지, 38개 동(棟)에 대한 설문조사결과를 바탕으로 공간계량모형(SAR, SEM)을 거리가중행렬을 이용하여 주택가격 형성요인을 분석하였다. 분석결과 공간자기회귀모형(SAR)이 가장 양호한 것으로 나타났다. 최열·이백호(2006)는 주거지역 주변의 용도지역과 시설에서의 접근성을 고려하여 주거지역내의 지가의 영향요인과 공간효과분석을 거리가중치를 이용하여 실시하였다. 분석결과 전통적인 OLS기법과 SAR, SEM중 SEM모형이 적합한 것으로 판명되었다. 허윤경(2007)은 서울 1,755개단지, 부산은 770개 단지의 아파트 시세자료를 활용하여 공간회귀모형인 SAR, SEM, SAC를 비교 분석하였다. 분석결과 서울, 부산 모두 OLS, SAR, SEM보다 SAC모형이 보다 우수한 것으로 나타났다.

III. 연구모형

1. 헤도닉가격함수

주택 분석에 있어 헤도닉 모형의 이론적 토대를 마련하는데 가장 큰 공헌을 한 Rosen(1974)의

모형은 주택구성요소들의 이질성을 실증적으로 분석할 수 있는 적절한 수단이라고 일반적으로 인식되고 있다. Rosen의 모형은 개별 주택구성요소에 대한 주택소비자들의 주관적 평가에 기초한 입찰가격(bid price)과 공급자의 공급가격(offer price)간의 상관관계에 대한 이론적 기반을 제공함으로써 주택특성가격을 분석하는 이론적 기초를 제공하고 있다(임재현, 1998:249). Rosen의 이론에 따르면 주택구성요소의 잠재가격은 입찰가격 함수와 공급가격함수간의 상호작용에 의해 결정되며, 재화의 가치는 해당재화에 내포되어 있는 특성에 의해 결정된다고 전제하고 있다.

헤도닉 모형에 따른 주택가격은 수많은 속성들이 내재된 가격에 의해 결정되는데, 내재가격이란 실제가치는 존재하지만 가격화 되지 않은 주택 특성들에 대한 추정치이다. 기본적으로 주택의 내재적 가치는 주택의 유형, 건축자재, 규모(면적), 층수, 경과년수 등 주택자체 특성에 의해 결정되지만, 주택을 둘러싸고 있는 물리적·자연적 환경요인들 역시 주택의 내재적 가치를 형성하는 중요한 요인들이 된다(김태경·박현수, 2008: 147).

주택에 대한 헤도닉모형은 주택가격을 주택상품 혹은 서비스를 구성하고 있는 개별 특성들의 함수로 표시한 것이다. 일반적인 주택가격 함수는 식(1)과 같이 나타낸다.

$$P = h(S, N, L) \quad (1)$$

여기에서 P는 주택의 가격이고, S(구조적 변수들: Structural variables), N(근린적 변수들: Neighborhood variables), L(장소적·지역적 변수들: Locational variables)은 주택의 개별특성을 의미한다(김성우

·정건섭, 2010:77). 이러한 헤도닉 함수는 일반적으로 선형함수 형태로 식(2)과 같은 기본모형을 갖는다.

$$y = \beta X + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I_n) \quad (2)$$

2. 공간계량모형

공간계량모형은 일반선형모형에서 공간가중치 행렬을 부가한 형태로 일반선형모형의 확장된 형태이다(허윤경 2007:8). 먼저 SAR은 기본 헤도닉 가격함수 모형에 설명변수 행렬을 첨가하여 준형태로 구성된다. 연구자에 따라 공간시차모형(spatial lagged model)로 명명되기도 하며, 시계열 분석과 유사한 형태를 띤다.

$$y = \rho Wy + \beta X + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I_n) \quad (3)$$

여기서 W는 공간가중행렬이다. ρ 는 공간자기회귀계수이고, X는 주택 속성별 변수의 벡터를 나타낸다.

다음 SEM 모형은 오차항에 공간자기상관을 포함하고 있는 모형으로 다음과 같이 정의된다.

$$y = \beta X + u$$

$$u = \lambda Wu + \epsilon \quad \epsilon \sim N(0, \sigma^2 I_n) \quad (4)$$

λ 는 모형의 설명변수가 아닌 관측 불가능한 생략된 변수의 오차항에 대한 충격에 속한다. 다시 말하면 어느 특정지점의 주택가격은 채택된 설명변수 뿐만 아니라 생략된 변수의 함수라는 것이다(국토연구원, 2004:438). W는 앞선 SAR

과 같은 공간가중행렬이다.

SAC 모형은 SAR과 SEM을 혼합한 모형으로 가장 일반적으로 공간중속성이 공간자기회귀형태와 오차항 모두에 포함되어 있다. 구체적인 식은 다음과 같다.

$$\begin{aligned} y &= \rho W_1 y + \beta X + u \\ u &= \lambda W_2 u + \epsilon \\ \epsilon &\sim N(0, \sigma^2 I_n) \end{aligned} \quad (5)$$

3. 모형의 적합도

모형 추정에 있어 어떤 추정모형이 보다 적합한가? 에 대한 비교가 필요하다. 즉, 추정모형이 얼마나 실제자료를 잘 요약하고 있는지에 대한 적절한 답을 할 수 있어야 한다. 이러한 문제를 해결하기 위한 척도가 모형의 적합도이다. 일반적으로 흔히 사용되는 적합도의 측정방식은 많은 연구자에 의해 다양하게 개발되어졌다. 본 연구에서는 전통적 헤도닉가격모형과 공간계량모형의 추정방법의 차이를 비교하기 위해 기존 연구에서 일반적으로 사용되는 결정계수(R^2)를 참고로 하여 log Likelihood, AIC(Akaike Information), SC(Schwarz Criterion)를 기준으로 각 모형별 적합도를 평가하였다.

결정계수(R^2)는 회귀식의 적합도를 평가함에 있어서 가장 많이 이용되는 것으로 독립변수가 종속변수를 설명할 수 있는 정도를 나타내는 지표이다. 이런 의미에서 결정계수를 설명력이라고 부르기도 한다(김호정 2005:473). 결정계수(R^2)는 그 값이 클수록 설명력이 증대하고, 회귀식의 활용가치가 높아지고, 분석결과는 신뢰성을 더욱

인정받게 된다. 결정계수 R^2 은 식(6)과 같다.

$$R^2 = \frac{\sum(\hat{Y} - \bar{Y})^2}{\sum(Y - \bar{Y})^2} \quad (6)$$

\bar{Y} 는 Y 의 평균, \hat{Y} 은 Y 의 예측치 이다. 결정계수(R^2)는 $0 \leq R^2 \leq 1$ 의 값을 가진다. R^2 값이 크다고 반드시 설명력이 높은 것인 아니다. 즉, 두변수가 직선관계가 아닐 경우 R^2 값이 예상보다 낮게 나타나는 경우가 발생할 수 있다. 한편 표본의 크기가 비교적 적고, 독립변수의 수가 많을 경우 R^2 값이 실제 모집단을 대상을 했을 때보다 크게 나타날 가능성이 높아 표본의 크기와 독립변수의 수를 고려하여 R^2 을 적절히 조정된 수정된 R^2 (adjusted R^2)를 이용한다.

Likelihood 값은 관측된 표본 값들을 얻을 수 있는 가능성을 나타는 계수를 말하며 보통 L 로 표시하고, 수식은 아래와 같다.

$$\begin{aligned} L &= \frac{1}{(2\pi^2)^{(2/n)}} |I_n - \rho W| \exp \\ &- \frac{1}{2\sigma^2} (y - \rho W y)' (y - \rho W y) \end{aligned} \quad (7)$$

일반적으로 L 자체보다는 L 에 log값을 취해 극대화 한다. 이는 log를 취해도 분석결과는 변하지 않을 뿐만 아니라 그 계산과정이 쉬워지기 때문이다(이종원, 1994:288). Log Likelihood 함수는 다음과 같다.

$$\ln L \propto -\frac{n}{2} \ln (y - \rho W y)' (y - \rho W y) + \ln |I_n - \rho W| \quad (8)$$

공간적으로 시차된 종속변수나 오차항이 포함된 경우 편의(bias)로 인해 OLS방법으로 추정되어 질수 없고, 대신에 최우추정량(maximum likelihood)이 사용된다. 따라서 OLS의 적합도를 나타내는 R^2 은 더 이상 공간모형과 비교될 수 없으며, Log Likelihood와 AIC(Akaike Information), SC(Schwarz Criterion)의 우도함수가 공간모형의 적합도를 측정하는데 사용되어 진다. 여기서 가장 높은 Log Likelihood나 가능 낮은 AIC와 SC를 가진 모형이 더 좋은 모형으로 고려되어진다(Anselin, 2005; 최열·이백호, 2006:49). AIC와 SC의 식은 다음과 같다(Anselin, 2005:175).

$$AIC = -2L + 2K \quad (9)$$

$$SC = -2L + K \times \ln(N) \quad (10)$$

여기에서 L 은 Log Likelihood값 이며 K 는 파라미터(parameter)의 수이다.

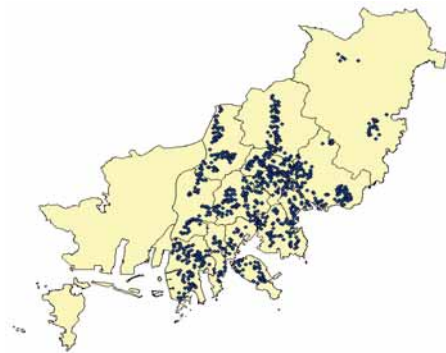
IV. 분석 자료

1. 자료 수집

연구대상 자료는 2006년 1분기부터 2009년 2분기까지 국토해양부실거래가 홈페이지(<http://rt.mltm.go.kr>)에 공개된 아파트 거래자료 중 5,850개를 분석에 이용하였다. 개별 아파트단지에 대한 기

본정보는 부동산114(<http://www.r114.co.kr>)의 자료를 활용하였다. 역세권 여부는 네이버지도(<http://map.naver.com>)를 통해 아파트와 지하철역까지 직선거리를 기준으로 도보 10분내 접근 가능한지에 따라 역세권과 비역세권으로 구분하였다. 단지의 재건축 추진 여부는 ‘부산시 정비사업정보센터(<http://renewal.busan.go.kr/>)’ 자료를 기준으로 선정하였다. 공간가중 행렬 구성을 위한 공간위치정보는 아파트주소를 기준으로 지오코딩(geo-coding)을 통하여 각 자료간의 거리를 측정하였다.

〈그림 1〉 아파트 위치도



2. 변수 설정

김기호·이성우(1998)은 주택특성과 단지(지역) 특성 변수를 기준으로 11개 변수를 독립변수로 분석에 이용하였다. 대지의 질과 주택가격의 관계를 분석한 최종일·심성훈(2002)은 아파트 구조변수, 접근도, 근린변수, 대기오염변수 등을 기준으로 13개 변수를 선정하였다. 김주영(2003)은 주택가격 추정요인으로 구조적 특징, 입지적 특징, 지역변수를 기준으로 10개 변수를 분석에 이용하였다. 최열 외(2008)는 구조적 특징, 단지 특징, 입지적 특징, 하위시장 특징으로 구분하여 28

개 변수를 설정하였다. 서울의 교육환경이 주택 가격에 미치는 효과를 분석한 진영남·손재영(2005)의 경우 아파트 특징, 환경요소, 교육환경 등의 요소로 15개 변수를 이용하여 아파트 매매 가격과 전세가격에 미치는 영향을 분석하였다. 기타 황형기 외(2008)은 35개, 김태운 외(2007)은 26개의 독립변수를 이용하여 아파트 가격에 미치는 영향을 분석하였다.

본 연구는 기존연구를 바탕으로 <표 1> 와 같이 7개 단지요인과 해운대구를 기준 변수로 12개의 지역요인, 2009년 2분기를 기준으로 13개의 시간 더미변수를 이용하여 총 32개 독립변수를 선정하였다.

변수 간 다중공선성 문제를 제거하기 위하여 단지요인에 대하여 요인분석과 단계적 방법(stepwise)방법에 따른 회귀분석을 반복 하여 모

형의 설명력이 높은 주택면적, 경과년수, 총세대수, 역세권, 재건축 변수를 선정하였다. 지역변수는 강서구의 경우 실거래된 아파트 단지가 적어 본 연구에서는 제외시켰다. 동구와 중구, 서구의 경우 실거래 자료가 적고, 구도심지역으로 노후도와 규모 등이 비슷한 점을 감안하여 동일 지역으로 간주하여 분석에 이용하였다. 기타 시간요인은 분기별 자료를 더미와 하여 실증 연구에 사용하였다.

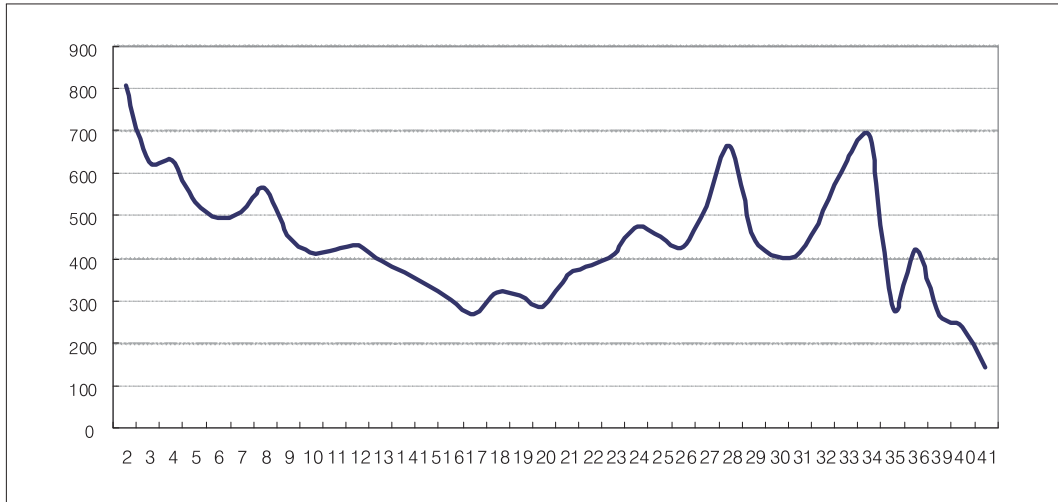
경과년수의 경우 제곱과 세제곱의 변수를 이용한 것은 <그림 2>에서와 같이 평균 평(3.3m²) 당 주택가격이 재건축의 영향으로 18~35년 사이에 다시 상승하는 형태를 보이고 있기 때문이다.

변수의 증가에 따른 모형의 적합성을 살펴보기 위해 OLS추정에 대한 전통적 헤도닉 모형은 단지요인(OLS-1), 단지요인+지역요인(OLS-2), 단

<표 1> 변수의 정의

변수		변수의 정의	비고
종속 변수	실거래가격	국토해양부 아파트 실거래가격(2006년~2009년)	만원
	주택면적	주택면적(전용면적+기타공용면적)	m ²
단지 요인	역세권	역세권여부 0 = 지하철 역과의 도보 10분 이상 1 = 지하철 역과의 도보 10분 이내	dummy
	재건축	재건축 추진여부 0 = 재건축 비추진, 1 = 재건축 추진 중	dummy
	총세대수	개별 아파트 총세대수	호
	경과년수	2010년 기준 아파트 경과년수 (2009년 입주=1, 2008년 입주=2, . . .)	년
	경과년수 ²	경과년수의 제곱	년
	경과년수 ³	경과년수의 세제곱	년
지역요인		각 구별 더미(금정구~동·중·서) 해운대구 : 기준 변수	dummy
시간요인		시간 더미(2006년 1분기~2009년 2분기) 2009년 2분기: 기준변수	dummy

〈그림 2〉 경과년수별 평균 평당 매매가격 곡선



지요인+지역요인+시간요인(OLS-3)으로 구분하여 공간계량모형인 SAR, SEM, SAC와 모형적합도를 비교 분석하였다. 모든 공간계량모형에 사용된 변수는 단지요인인 7개 변수를 사용하여 모형을 추정하였다.

3. 공간가중행렬

공간상에 분포해 있는 실체들의 관계를 정의하는 공간가중행렬(spatial weight matrix)은 주로 인접성척도(contiguity measure)와 거리척도(distance measure)를 기준으로 하고 있다. 인접성을 수량화하는 방법으로 Linear Contiguity, Rook Contiguity, Bishop Contiguity, Queen Contiguity등이 있다. 거리척도는 직접적으로 공간이 갖는 관계를 정의한다. 두 공간 실체의 거리를 공간가중행렬(spatial weight matrix)로 이용할 경우 거리가 멀어질수록 서로의 관계가 약화 될 수 있도록 식(11)과 같이 정의 한다(김성우·정건섭, 2010:81).

$$W_{ij} = 1/d_{ij}^{\alpha}, \quad (\alpha = 1, 2, 3, \dots) \quad (11)$$

인접성척도는 공간이 가지는 공간배열을 기준으로 간접적으로 정의하는 반면 각 지점의 위치를 나타내는 좌표나 지점간의 거리를 알고 있을 경우, 거리를 이용한 가중치 행렬이 인접성을 이용한 방법보다 바람직하다(Anselin, 1988).

이외에도 가장 근접한 이웃(nearest neighbors)의 개수와 공유경계선(shared boundary) 비율, 경제적인 차이 등을 이용하여 공간가중행렬을 구성할 수 있다.

본 연구에서는 인접성 보다는 공간가중행렬을 5,850개의 개별 아파트 단지간의 거리를 기준 ($W_{ij} = 1/d_{ij}$)의로 행표준화(row standardization)를 통해 5,850×5,850 형태의 공간가중행렬을 구성하였다. 공간계량분석은 Matlab을 이용하였다.

V. 분석 결과

1. 기초 통계량

분석에 이용된 5,850건의 아파트 실거래가 평균은 1억3천480만원으로 최저 1천600만원부터 15억원까지 분포하는 것으로 나타났다. 단지특성별로는 평균 면적은 99.45㎡(30.08평)이고, 전체 자료 중 재건축이 추진 중에 있는 것은 약 7%가량이다. 도보 10분이내의 역세권에 위치한 자료가 전체의 약 45%였다. 평균 총세대수는 667세대로 나타났는데, 이는 대규모 단지일수록 거래건수가 많기 때문이다. 경과년수는 1년부터 41년까지 자료가 존재하며 평균 13.84년이다.

지역별 평균 실거래가를 살펴보면 수영구가 1억8천633만원으로 가장 높았고, 다음 동래구 1억7천454만원, 연제구 1억5천987만원, 해운대 1억5천954만원으로 나타났다. 해운대구 지역이 높은 주택가격에 비해 평균 실거래가격이 낮은 것은 대부분 거래가 중소형을 중심으로 거래가 많기 때문이다.

시간별로는 2006년 1분기부터 2008년 4분기까지는 1억3천만원 전후의 평균 신거래 가격이

2009년 1Q 이후 증가세를 보였다. 이는 주택가격의 전반적인 상승이라기보다 글로벌 금융위기와 주택담보대출 금리인상 등에 따른 부담이 증가하면서 중소형 면적은 거래량이 줄어든 반면 고가아파트를 중심의 급매물 출시와 함께 거래가 늘었기 때문이다.

2. 모형별 적합성

모형간의 적합성을 평가하기 위해 결정계수(R^2), Log Likelihood, AIC(Akaike Information), SC(Schwarz Criterion)를 기준으로 이용하였다. 우선 전통적 헤도닉가격모형을 살펴보면 OLS-1(단지요인)은 $Adj R^2$ 값이 0.781로 7개의 독립변수로 78.1%를 설명하는 것으로 나타났다. 기타 log Likelihood -56,785.7, AIC 113,587, SC 113,641로 나타났다. 지역적 변수를 더한 OLS-2(단지요인+지역요인)은 $Adj R^2$ 값 0.807, Log Likelihood -56,425 AIC 112,868 SC 112,921으로 12개 변수의 추가로 설명력은 약 2.52%가량이 증가하는 것으로 나타났다. 지역변수와 시간 변수를 포함한 OLS-3(단지요인+지역+시간요인)의 경우 $Adj R^2$ 값이 0.810 Log Likelihood -56,376

〈표 2〉 단지특성 요인 기초 통계량

변수		N	평균	Std	최소값	최대값
실거래가격		5850	13,484.97	8519.38	1,600	150,000
단지 특징	주택면적	5850	99.45	30.89	33.00	306.83
	역세권	5850	0.45	.49	0	1
	재건축	5850	0.07	.25	0	1
	총세대수	5850	793.51	677.97	25	3382
	경과년수	5850	13.84	7.22	1	41
	경과년수 ²	5850	243.72	245.00	1	1681.00
	경과년수 ³	5850	5065.60	7853.03	1	68921.00

〈표 3〉 지역 및 시간별 실거래가

지역	N	평균	Std	시간	N	평균	Std
금정구	450	16774	10087	2006_1Q	424	13167	8151
기장군	450	9135	4037	2Q	481	12980	7930
남구	450	14644	8568	3Q	442	12599	7136
동래구	450	17454	10841	4Q	821	13548	7769
진구	450	12374	5909	2007_1Q	578	13090	7760
북구	450	11153	6005	2Q	400	13898	8624
사상구	450	9441	3740	3Q	347	13113	10621
사하구	450	10203	5143	4Q	365	12903	7538
수영구	450	18633	10504	2008_1Q	541	13089	7841
연제구	450	15987	7649	2Q	458	13817	8164
영도구	450	9548	5167	3Q	310	13908	9065
해운대구	450	15954	11963	4Q	202	13347	9419
동·중·서구	450	14000	450	2009_01Q	197	16974	13871
				2Q	284	14862	8723

〈표 4〉 전통적 헤도닉 모형 추정결과

	OLS-1	OLS-2	OLS-3		
			시간변수		
constant	4532.59***	7412.93***	7202.59***	2006_1Q	-249.85
주택면적	195.25***	188.057***	188.476***	2Q	-427.96
역세권	1841.51***	1706.89***	1724.86***	3Q	-586.16
재건축	3392.47***	2370.58***	2435.59***	4Q	-293.41**
총세대수	1.93***	2.08***	2.011***	2007_1Q	-354.49
경과년수	-2176.79***	-2114.02***	-2060.61***	2Q	-401.04
경과년수 ²	94.47***	91.63***	88.58***	3Q	-151.47
경과년수 ³	-1.17***	-1.14***	-1.09***	4Q	-452.01
금정구		-1603.34***	-1624.91***	2008_1Q	59.67
기장군		-3744.96***	-3780.76***	2Q	484.28
남구		-2572.49***	-2620.78***	3Q	722.07**
동래구		-1194.87***	-1281.60***	4Q	300.34
부산진구		-3074.59***	-3083.35***	2009_1Q	1737.01***
북구		-4468.61***	-4431.46***		
사상구		-3613.50***	-3619.30***		
사하구		-4448.12***	-4495.51***		
수영구		-145.25	-135.47		
연제구		-1661.08***	-1664.43***		
영도구		-2619.38***	-261196***		
동·중·서구		-2385.27***	-2470.12***		
R^2	0.7821	0.807	0.810		
$Adj R^2$	0.7818	0.807	0.8095		
Log Likelihood	-56785.7	-56425.8	-56376.3		
AIC	113,587	112,868	112,769		
SC	113,641	112,921	112,822		

주) *** p<0.01 ** p<0.05 * p<0.10

〈표 5〉 공간계량모형 추정결과

	SAR	SEM	SAC
constant	-361.03***	6777.85***	10045.74***
주택면적	190.3***	185.70***	174.73***
역세권	802.19***	947.51***	893.72***
재건축	2751.99***	2259.01***	2015.49***
총세대수	1.63***	1.79***	1.81***
경과년수	-2072.44***	-2356.54***	-2131.87***
경과년수 ²	92.69***	109.83***	99.80***
경과년수 ³	-1.19***	-1.49***	-1.36***
ρ	0.39***		-0.31***
λ		0.84***	1.14***
R^2	0.7888	0.8187	0.8255
$Adj R^2$	0.7885	0.8184	0.8253
Log Likelihood	-54415.93	-54283.66	-54217.26
AIC	108,850	108,585	108,455
SC	108,910	108,645	108,521

주) *** p<0.01 ** p<0.05 * p<0.10

AIC 112,769 SC 112,822로 OLS-2모형에 비해 13개의 시간 변수를 추가로 설명력은 0.25%가 증가한 것으로 나타났다. OLS-1과 OLS-3모형의 경우 25개의 변수 추가에도 불구하고 $Adj R^2$ 값의 차이가 2.77%로 나타났다.

공간계량모형의 경우 SAR모형의 $Adj R^2$ 값이 0.788으로 OLS-2, OLS-3모형에 비해 설명력은 낮지만 Log Likelihood -54,415, AIC 108,850 SC 108,910의 값은 OLS-2, OLS-3보다 우수해 SAR 모형이 OLS모형의 변수를 추가한 모형에 비해 모형적합도가 높은 것으로 나타났다. 공간계량모형 내에서도 SAC가 $Adj R^2$ 값이 0.8253, Log Likelihood -54,217 AIC 108,455, SC 108,521로 가장 모형이 우수한 것으로 나타났다.

OLS 중 모형이 가장 적합한 OLS-3과 공간계량모형중 모형 적합도가 높은 SAC를 기준으로 추정치를 살펴보면 우선 OLS-3모형의 경우 주택면적은 1㎡ 증가함에 따라 188만원(평당 623만원)이 증가하는 것으로 나타났고, 역세권의 경우 비역세권에 비해 1,724만원이 높은 것으로 나타났다. 재건축은 비재건축 단지에 비해 2,435만원이 높은 것으로 추정되었다. 평균 경과년수에 대한 기울기의 변화률²⁾은 529.72이다.

SAC 추정치와 비교할 경우 SAC의 주택면적에 대한 추정치는 1㎡당 174만원(평당 577만원)으로 OLS-3모형이 13만원 가량 높게 추정되었다. 역세권 역시 SAC는 893만원으로 OLS-3모형(1,724만원)이 831만원 가량 높게 추정됐다. 재건

2) $y = \beta_1 Age^3 + \beta_2 Age^2 + \beta_3 Age$ 의 평균 기울기의 변화률은 $\frac{dy}{dAge} = 3\beta_1 Age^2 + 2\beta_2 Age + \beta_3$ 이다.

축의 경우 SAC는 2,015만원으로 OLS-3(2,435)가 420만원 가량 높아 전반적으로 OLS-3에 의한 모수의 추정치 상향추정(over estimation)되고 있음을 보여준다.

따라서 전통적 헤도닉 모형에 따른 OLS에 따른 모수의 추정치보다 공간계량모형이 적은 변수의 숫자로도 안정적인 모형의 적합도를 유지할 수 있고, OLS모형이 갖는 몇 가지 기본 가정에 따른 문제점을 보완에 따라 모형의 예측력을 높일 수 있을 것으로 판단된다.

(residuals)는 독립적이며, 평균이 '0'이고, 분산이 일정하며, 공분산이 '0'이라는 가정에 적합하도록 하위시장으로 구분할 경우 OLS에 의한 추정결과가 보다 우수할 수 있다. 이러한 공간가중행렬의 구성과 하위시장 간의 모형 적합도에 대한 연구는 향후 과제로 남기고자 한다.

논문접수일 : 2010년 7월 5일

심사완료일 : 2010년 9월 27일

V. 결론

본 연구는 전통적 헤도닉가격모형에 따른 주택가격 추정에 있어 OLS방법의 추정치와 공간계량모형간의 모형의 적합도를 비교 평가하였다. 우선 OLS추정에 있어 모형의 설명력을 높이기 위한 변수를 증가시키는 것 보다는 공간계량모형에 따른 추정결과가 보다 우수한 것으로 나타났다.

한편 OLS추정의 경우 모수의 추정치 상향추정(over estimation) 될 수 있는 가능성을 보여줌으로써 기존 헤도닉가격모형의 모수추정에 있어 주의가 필요한 것으로 나타났다. 향후 주택가격 추정에 있어 공간계량모형의 적용으로 예측력을 높일 수 있을 것으로 기대된다.

하지만 이러한 공간계량모형의 우수성에도 불구하고 몇 가지 한계를 갖고 있다. 우선 공간가중행렬의 구성에 대한 문제이다. 아직 공간가중행렬구성에 대한 합의가 없고, 공간가중행렬을 어떻게 구성하느냐에 따라 추정결과가 달라질 수 있기 때문에 연구자의 주관에 개입될 가능성이 높다. 또한 OLS 추정에 대한 가정 즉, 잔차

참고문헌

1. 국토연구원, 「공간분석기법」, 한울아카데미, 2004
2. 김기호·이성우, “해안변 아파트 단지의 주택 가격 분포특성에 관한연구: 부산시 사례를 중심으로”, 「국토계획」, 제33권 2호, 1998, pp. 119-134
3. 김석환·이현석, “재건축 기대에 따른 아파트 가격변화 분석”, 「국토계획」, 제40권 3호, 2005, pp. 67-78
4. 김성우·정건섭, “견고한 공간계량경제모형의 추정: 층수를 고려한 3차원 공간가중행렬을 이용하여”, 「주택연구」, 제18권3호, 2010, pp. 73-92
5. 김종원, “주택시장에서 공간자가상관의 검증 및 회귀계수의 추정”, 「경제학연구」, 제48권 2호, 2000, pp. 155-173
6. 김주영, “위계적 선형모델을 이용한 주택가격 함수 추정”, 「국토계획」, 제38권 7호, 2003, pp. 223-234
7. 김주영·김주후, “주택가격평가를 위한 위계적 선형모델 적용”, 「국토연구」, 제33권, 2002, pp. 21-34
8. 김태경·박헌수, “주택가격을 결정하는 공간적 특성들의 시계열적 영향력 변화 분석”, 「국토계획」, 제43권 3호, 2008, pp. 145-166
9. 김태운 외, “경관특성 차이가 아파트가격에 미치는 영향: 주택실거래가를 사용하여”, 「부동산학연구」, 제13권 3호, 2007, pp. 169-186
10. 김호정, 「사회과학통계분석」, 삼영사, 2005
11. 박헌수·정수연·노태욱, “공간계량경제모형을 이용한 아파트가격과 공간효과분석” 「국토계획」, 2003, 제38권 5호, pp. 115-125
12. 서경천·이성호, “공간적 자기회귀모델과 도시시장분할에 의한 효율적 지가추정에 관한 연구”, 「국토계획」, 제36권 4호, 2001, pp. 77-94
13. 송명규, “학군의질(고등교육 수준)과 명성이 주택가격에 미치는 효과에 관한 실증연구: 서울시의 경우”, 「지역사회발전학회논문집」, 제17권1호, 1992, pp. 91-106
14. 양진우, “헤도닉기법을 이용한 환경소음개선 효과의 평가에 관한 연구”, 「국토계획」, 제35권 5호, 2000, pp. 233-244
15. 오규석·이왕기, “아파트 가격에 내재한 경관조망가치의 측정”, 「국토계획」, 제32권 3호, 1997, pp. 139-151
16. 이상경·신우진, “재건축 가능성이 아파트 가격에 미치는 영향”, 「국토계획」, 제36권 5호, 2001, pp. 101-110
17. 이용만, “헤도닉 가격모형에 대한 소고”, 「부동산학연구」, 제14권1호, 2008, pp. 81-87
18. 이종원, 「경제경영통계학」, 박영사, 1994
19. 임재현, “주택특성가격이론의 발전 모색”, 「한국행정학보」, 제32권1호, 1998, pp. 247-261
20. 진영남·손재영, “교육환경이 주택가격에 미치는 효과에 관한 실증분석: 서울시 아파트 시장을 중심으로”, 「주택연구」, 제13권 3호, 2005, pp. 125-148
21. 최열·김형수·박명제, “주택하부시장 특성을 고려한 신규분양가와 입주 후 가격변화에 관한 연구”, 「대한토지학회논문집」, 제28권 4호, 2008, pp. 523-531
22. 최열·이백호, “공간자가상관과 주변용도지역에서 접근성을 고려한 주거지 내 지가 추정에 관한 연구”, 「국토계획」, 제41권 5호, 2006, pp. 45-60

23. 최종일 · 심성훈, “서울시 아파트가격에 대한 대기질의 영향: 헤도닉기법을 이용하여”, 「자원 · 환경경제연구」, 제11권 2호, 2002, pp. 216-278
24. 허윤경, “도시별 주택가격의 공간적 영향력 검증: 서울과 부산의 아파트 가격을 중심으로”, 「주택연구」, 제15권 4호, 2007, pp. 5-23
25. 황형기 · 이창무 · 김미경, “한강조망이 주택 가격에 미치는 영향”, 「주택연구」, 제16권 2호, 2008, pp. 51-72
26. Anselin, L., *Spatial Econometrics: Method and Models*, Dordrecht: Kluwer Academic Publishers, 1988
27. Anselin, L., *Exploring Spatial Data with GeoDa : A Workebook*. Center for Spatially Integrated Social Science, 2005
28. Brookshire, D.S., M.A, Thayer, W.D, Schulze and R.C, D’Arge, “Valuing Public Goods : A Comparison of Survey and Hedonic Approaches”, *The American Economic Review*, vol.72(1), 1982, pp. 165-177
29. Can, A., “Specification and Estimation of Hedonic Housing Price Model”, *Regional Science and Urban Economics*, vol.22(3): 1992, pp. 453-473
30. Dale-jonson, “An Alternative Approach to Housing Market Segmentation Using Hedonic Price Data”. *Journal of Urban Economics*, vol.11(3), 1982, pp. 311-332
31. Goodman, A.C and T.G. Thibodeau, “Age-related heteroskedasticity in hedonic house price equations”, *Journal of Housing Research*, vol.6(1), 1995, pp. 25-42
32. Linneman, P., “Some empirical results on the nature of the hedonic price function for the urban housing market”, *Journal of Urban Economics*, vol.8(1), 1980, pp. 47-68
33. Rosen, S., “Hedonic Prices and Implicit Markets : Product Differentiation in Pure Competition”, *Journal of Political Economy*, vol.82(1), 1974, pp. 35-55
34. 국토해양부, <http://rt.mltm.go.kr>
35. 부동산 114, <http://www.r114.co.kr>
36. 네이버지도, <http://map.naver.com>
37. 부산시, <http://renewal.busan.go.kr>