

준모수방법에 의한 주택가격지수 추정에 관한 연구

박 헌 수

안양대학교 도시정보공학과 조교수

A Study on the Semiparametric Estimation of the Housing Price Indexes

Park, Heon-Soo

Assistant Professor, Anyang University

Abstract: A semiparametric estimator is used to analyze the hedonic housing price indexes in a sample of 1,284 sales prices of apartments in the southern Seoul area. The hedonic price functions are expected to be nonlinear, but choosing the appropriate functional form is particularly difficult for location variables and time trends because it tends to be unique to the place and time. A semiparametric estimation offers important advantages for the hedonic housing price function estimation. It combines the benefits of the parametric and nonparametric estimation. The parametric portion of the model includes standard housing characteristics such as dwelling unit size, the size of the community, age of building, types of heating system and fuel material, school districts, among others. Distances to the CBD, several subcenters, and the nearest subway station are estimated nonparametrically. The sales year is modeled nonparametrically, as well.

The coefficients have the expected signs and are stable across parametric and semiparametric specifications. The semiparametric estimation has higher coefficient of determination than two parametric OLS estimations. Hausman (1978) tests for the specification of the hedonic price functional form prefer to the semiparametric estimation. At last, the housing price indexes for four subway station regions are estimated and compared.

I. 서론

주택가격지수(housing price index)는 표준 단위 주택에 대해 주택가격동향을 시간대별로 추적하여 나타낸 것으로 정의한다. 일반적으로 주택가격과 이에 영향을 미치는 주택 특성들에 대한 함수관계를 설정한 뒤에 자료를 사용하여 추정하게 된다. 즉, 주택가격에 영향을 미치는 요인들을 접근성(accessibility), 주변지역의 특성(neighborhood characteristics), 주택의 구조적 특성(structural characteristics) 등으로 구분한 다음 특성이격함수(hedonic price function)를 통해 주택특성들의 내재적 시장가격(implicit market price)을 파악하여 주택가격지수를 작성하게 된다.

특성이격함수를 이용하여 주택가격함수를 추정한 연구결과들을 보면 많은 경우 80%이상의 높은 결정계수(R^2)를 갖는 것으로 나타난다. 하지만 결정계수가 높다고 하더라도, 주택가격지수에 대한 정확한 시간경로를 제공한다고 볼 수는 없다. 특성이격함수에서 설명되지 못하는 변동의 많은 부분들이 시간효과(time effects)와 공간효과(spatial effects)들을 모형에서 적절하게 다루지 못함으로써 발생하게 된다.

특성이격함수모형에서 모수적(parametric) 방법으로 주택특성들의 효과를 추정하는 경우 시간효과들은 시간에 대한 여러 더미변수들을 사용하거나 시간변수를 고차 다항식(polynomials)으로 사용하는 경우가 대부분이다. 시간에 대한 여러 더미 변수들을 사용하는 경우, 주택가격은 사용한 더미변수들의 불연속적인 변화로 추정된다. 이와 같이 추정된 가격지수는 매우 비현실적일 뿐만 아니라 특정 시점에서 자료가 충분하지 않을 경우 부정확한 결과를 가져오게 된다. 반면, 시간변수를 고차 다항식으로 모형에서 사용하는 경우에는 주택가격이 시간변수와 복잡한 비선형(nonlinear) 관계를

가지고 있을 경우 다항식의 차수(order)를 아무리 증가하더라도 시간변수에 따른 주택가격 변동을 충분히 반영하기는 어렵다.

또한, 주택이 가지는 여러 특성들 가운데 '위치의 고정성(locational fixity)' 때문에 주택가격은 외부효과를 크게 받는다. 외부효과들은 지역마다 다르기 때문에 이들 효과들을 단순히 몇 개의 공간관련 모수들로 추정하는 데에는 어려움이 있다. 주택가격의 지역적인 변동을 고려하기 위하여 공간관련 변수들을 다항식으로 모형에 반영하여 추정할 수 있으나, 앞서 시간변수의 경우와 마찬가지로 주택가격이 공간관련 변수들과 복잡한 비선형 관계를 가지는 경우 모수적으로 추정하는 데에는 한계가 있다. 시간과 공간효과가 동시에 발생하는 경우에는 모수적 함수관계를 통해 추정하는 것은 더욱 어려운 문제가 된다.

이와 같이 모수적 방법에 의한 주택의 특성이격함수를 추정하는 문제들을 극복하기 위한 대안으로서 비모수적 방법(nonparametric approach)이 가능하다. 이 방법은 주택의 특성이격함수에서 시간과 공간효과들을 나타내는 변수들과 주택가격과의 비선형 관계를 자료를 통하여 추정한다. 함수형태에 대한 사전적인 제약이 없기 때문에 시간과 공간의 교차효과 등 다양한 관계들을 모형에서 추정이 가능하다. 따라서, 비모수적 방법은 전형적인 모수적 방법에 의한 주택시장의 특성이격함수 모형보다 장점을 가지고 있다.

하지만, 비모수적 추정방법의 단점을 살펴보면, 주요 정책변수들에 대한 가설검정이 쉽지가 않으며, 설명변수가 많아지면 정확한 추정을 위해서 많은 표본들이 필요하게 된다. 또한 모든 변수들을 비모수적으로 추정하는 경우 더미변수들을 모형에서 제외하게 되는데, 추후에 이들 더미변수의 효과를 파악하는 과정이 쉽지가 않다.¹⁾ 또한 비모수적 방법에서

1) 하지만 주택시장을 연구하는데 있어 더미

커널추정(kernel estimation)에서 점근적편향(asymtotic bias)문제에 따른 부정확한 신뢰구간(confidential interval)이 문제가 있다(Horowitz and Härdle, 1996).

준모수적 방법(semiparametric approach)은 모수적 방법과 비모수적 방법의 장점들을 합친 것이다. 예를 들어, 주택의 특성가격함수 모형에서 전용면적, 단지규모, 건물연수, 더미 변수들과 같이 주택가격에 미치는 영향이 비선형 관계가 낮은 경우에는 모수적으로 모수들을 추정하고, 반면 시간과 공간관련 변수들과 같이 주택가격에 미치는 영향이 비선형 관계가 큰 변수들에 대해서는 비모수적 방법으로 추정함으로써 추정량들의 효율성을 높일 수 있다.

본 논문에서는 주택의 특성들을 감안한 주택가격지수 추정을 시도하였다. 먼저, 서울 한강이남지역에서의 아파트 단지들에 대하여 먼저 아파트가 가지고 있는 특성들과 아파트 가격과의 특성가격함수를 추정하여 본 결과, 모수적 방법에 비해 준모수적 방법에 의한 주택가격함수의 추정이 모형의 적합도가 높은 것으로 나타났다. 특히, 특성가격함수모형에서 지하철역과 도심·부도심과의 거리, 그리고 매매연도와 같은 공간과 시간 변수들은 주택가격과 비선형 관계가 바람직한 것으로 나타났다. 추정된 특성가격함수를 이용하여 7호선 지하철 일부 전철역(대림역, 상도역, 이수역, 강남구청역) 반경 1km 이내 지역에 대해 월별 아파트가격지수를 추정하여 본 결과 시간과 공간효과가 지역별로 차이가 있는 것으로 나타났다으며, 특히 1992~1998년 기간동안 시간과 공간의 교차효과가 크게 나타났다.

본 논문의 구성은 다음과 같다. II 절에서는 주택가격지수에 대한 관련 문헌들을 고찰

하고, III 절에서는 준모수적 방법과 이의 추정과정을 소개하고 있다. IV 절에서는 사용된 자료를 기술하고 이의 기초통계들을 제시하고 있으며, V 절에서는 분석결과들을 제시하고 VI 절에서 결론을 내리고 있다.

II. 주택가격지수 관련 연구

주택가격지수는 표준단위 주택에 대하여 주택가격동향을 시간대별로 추적하여 나타낸 것이다. 주택 특성변수들은 시간이 지남에 따라서 변화한다. 예를 들어 투입자재의 가격, 소비자의 소득수준, 기호, 주택건설 기술 등은 시간이 지남에 따라 변화한다. 따라서 주택가격지수를 구하기 위해서는 이들 특성들이 시간이 지남에 따라서 주택 가격에 미치는 영향을 통제하면서 지수를 구하여야 한다.

주택가격 형성요인과 주택가격지수에 대한 특성가격함수(hedonic price function)에 의한 실증분석은 60년대부터 많은 연구가 이루어져 왔다²⁾. 특성가격함수 모형은 자료들을 임의로 선택한 다음 주택특성들의 차이들을 모수적 방법인 회귀분석방법으로 추정하는 것으로서, 이론적 골격은 Lancaster(1966)와 Rosen(1974)이 제시하였다. 이들은 추정된 특성가격함수 모형에서 회귀계수들의 의미를 주택특성들의 내재적 시장가격(implicit market price)으로 해석하였다. 또한, 1991년에 Journal of the American Real Estate and Urban Economics Association과 1996년의 Journal of Real Estate Finance and Economics에서 주택가격지수 추정을 특집으로 다루고 있다.

Haurin and Henderschoff (1991)과 Knight, Dombrow, and Sirmans (1995)는 특

변수들을 통상적으로 사용하게 되며, 종종 이들 변수들이 연구의 중점이 되기도 한다.

2) 특성가격함수와 관련한 기존 연구의 고찰은 서경천·이성호(2001)의 논문을 참조.

성가격함수모형을 사용하여 주택가격지수를 추정하는 과정에서 표본추출, 모형에서 주요 변수들의 누락, 함수형태의 선택, 그리고 주택 특성들의 내재된 가치들이 가지는 계열상관 등의 문제들을 지적하고 있다. 또한, 최근에는 부동산 자료들이 가지고 있는 공간자기상관 (spatial autocorrelation) 역시 중요한 문제로 부각하고 있다(Pace 1995).

Meese and Wallace(1991)와 Wallace (1996)는 16개 Bay Area 지역에 대해 모수적 방법으로 주택가격지수들을 추정한 다음 Cleveland and Devlin (1988)이 제시한 국지가중회귀(locally weighted regression) 방법에 의해 주택가격지수를 비모수적으로 추정하였다. 이들의 연구결과에 따르면, 비모수적 추정과 모수적 방법에 의한 지수 추정에는 큰 차이가 없는 것으로 나타났으나, 지역별로 가격 지수에 있어 상당한 차이가 있는 것으로 나타났다. 따라서, 여러 지역의 자료들을 통합(pooling)하여 도시 전체의 주택가격지수를 산정하는 방법에 신중을 기할 필요가 있다고 주장하고 있다.

Stock (1991)은 준모수적 방법을 사용하여 혐오시설 입지의 영향을 평가하고 있다. Thorsnes and McMillen (1998)은 토지가격과 필지규모와의 함수관계를 준모수 방법으로 추정을 시도하였다. 박헌수(2001)는 모수적 방법과 준모수적 방법을 이용하여 서울 한강이남 지역의 아파트 매매가격에 대한 특성가격함수를 추정을 하여 본 결과 모수적 방법보다 준모수적 방법에 의한 주택가격함수 추정이 바람직하다고 주장하였다.

III. 모형 및 추정방법

주택가격을 자연대수로 취한 값을 y 라고 하자. 주택가격에 대한 설명변수들은 두 개의

그룹으로 분류할 수 있다. 첫 번째 그룹을 x 라고 표시하고, 전용면적, 단지규모, 건물연수, 더미변수들과 같이 특성가격함수모형에서 모수적으로 추정하는 부분을 나타낸다. 모형에서 이들 변수들은 선형함수 $\beta'x$ 로 가정한다.

두 번째 그룹은 z 로 표시하고, 지하철역간 거리, 도심·부도심과의 거리, 그리고 연속변수인 시간변수들과 같이 주택가격에 미치는 영향이 비선형 관계를 가지는 변수들이다. 예를 들어, 지하철역에 가까울수록 주택가격은 올라가지만, 주택가격의 상승폭은 지역마다 다를 것이다. 이는 모형에서 고려하고 있지 못하는 여타 변수들의 영향이 간접적으로 반영되기 때문이다. 공원, 주요 간선도로와의 거리등은 주택 가격에 국지적으로 영향을 미치고 있다. 만일 이들 변수들과 지하철역과의 거리변수들이 상관관계를 가진다면 주택가격과 z 그룹의 변수들과의 관계를 추정하는데 매우 복잡하게 영향을 미치게 된다. 또 다른 예로서 주택가격은 시간변수와의 관계 역시 매우 복잡한 비선형 관계를 가진다.

i 번째 주택가격에 대한 특성가격함수는 다음 식과 같이 표현할 수 있다.

$$y_i = \beta'x_i + g(z_i) + u_i \quad (1)$$

여기서 $i=1, \dots, n$ 이고 함수 $g(\cdot)$ 는 2차 미분이 가능한 완만하고 연속적인 함수이라고 가정하자. 벡터 z 는 m 개의 시간과 공간관련 변수들로 구성되어있다고 가정한다. 오차항 u_i 는 서로 독립적이지만 이분산(異分散)을 갖는다고 가정한다.

모수적 방법의 추정과정을 간단히 설명하기 위해 먼저 $g(z_i)$ 가 선형함수, 즉 $\delta'z_i$ 인 경우 생각하여 보자. 모수벡터 δ 는 y 를 x 와 z 에 회귀분석을 통해 간단히 추정된다. 또한, 다음과 같이 세 단계 과정을 통해 동일한 추정치를 얻을 수도 있다. 즉, 첫 번째

단계에서 y 와 x 벡터내의 k 변수 각각을 z 변수에 회귀시켜서 \hat{y} 과 \hat{x}_k 을 구한다. 두 번째 단계에서는 1단계에서 구한 \hat{y} 와 \hat{x}_k 을 이용하여 $y - \hat{y}$ 을 $x_k^* \equiv x_k - \hat{x}_k$ 에 회귀시켜 β 의 추정치 $\hat{\beta}$ 을 구한다. 마지막 세 번째 단계에서는 $y - \hat{\beta}'x$ 를 z 에 회귀시켜서 δ 의 추정치 $\hat{\delta}$ 을 구한다. 첫 번째 단계에서는 y 변수에 z 의 효과들을 파악한다. 두 번째 단계에서는 x 의 y 에 대한 독립적인 효과를 추정한다. 마지막 단계에서는 두 번째 단계에서 구한 x 의 독립적인 효과부분을 제거한 다음 z 변수가 종속변수 y 에 미치는 효과를 추정한다.

준모수적 방법의 추정과정은 $g(\cdot)$ 의 함수형태가 비선형인 경우로서, 앞의 계산과정에서 z 와 관련하여 통상최소자승(OLS) 추정량 대신 비모수 추정량을 사용한다. 추정과정은 다음과 같다. 첫 번째 단계에서는 비모수적 추정과정으로서 x 벡터에서 k 설명변수들 각각에 대해서 $E(y_i | z_i)$ 와 $E(x_{ki} | z_i)$ 를 추정하고, 추정 값들을 \hat{y}_i 과 \hat{x}_{ki} 라고 한다. 두 번째 단계에서는

$y_i - \hat{y}_i$ 를 모든 $x_{ki}^* \equiv x_{ki} - \hat{x}_{ki}$ 에 회귀시켜 β 의 추정량 $\hat{\beta}$ 을 구한다. 세 번째 단계에서는 종속변수로 $E(y_i - \hat{\beta}'x_i | z_i)$ 사용하여 $g(z_i)$ 의 비모수 추정량 $\hat{g}(z_i)$ 를 구한다. 마지막으로 오차항의 분산

$$\delta^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\beta}'x_i - \hat{g}(z_i))^2 \text{ 을}$$

구한 다음 $\hat{\beta}$ 의 분산·공분산 행렬 $\hat{\delta}^2 (x^*{}' x^*)^{-1}$ 을 추정한다.

y 와 z 변수간의 선형 가정은 준모수적 방법에서 특수한 경우이다. 이때, $g(\cdot)$ 가 선형일 경우 준모수적 방법에서 추정량은 일치성(consistency)은 제공하지만 유효추정량(efficient estimator)은 아니다. 하지만 $g(\cdot)$ 가 비선형인 경우는 준모수적 방법은 추정량의 일치성을 갖지만 통상최소자승방법(OLS)의 추정량은 일치성을 제공하지 못한다(Härdle 1990).

i 번째 자료에 대해 비모수 추정치는 다음 식과 같다.

$$\hat{y}_i \equiv E(y_i | z_i) = \frac{\sum_{j=1}^n k_h(z_i - z_j) y_j}{\sum_{j=1}^n k_h(z_i - z_j)} \quad (2)$$

여기서

$$k_h(z_i - z_j) = k\left(\frac{z_i - z_j}{h}\right) \quad (3)$$

$k(\cdot)$ 는 커널함수이며 h 는 bandwidth이다. 커널함수로는 표준정규다변량커널(standard normal multivariate kernel)을 사용한다. 비모수 추정방법과 커널함수의 예에 대해서는 Härdle (1990), Härdle and Linton (1994) 및 Yatchew (1998) 문헌을 참조.

다변량정규커널 $k(\cdot)$ 는 다음 식과 같다.

$$k\left(\frac{z_i - z_j}{h}\right) = \prod_{l=1}^p \phi\left(\frac{v_{li} - v_{lj}}{h}\right), \quad (4)$$

여기서 ϕ 는 표준정규밀도함수이며, z 의 표본 분산·공분산행렬의 역행렬을 콜레스키분해(Choleski Decomposition)를 통해 구한 행렬을 S 라고 하면 새로운 벡터 $v = Sz$ 를 구한다.

i 관측지에 대한 추정치를 구할 때 커널 추정량은 z_i 에 가까운 z_j 값을 갖는 관측

지에 대해 가중치를 상대적으로 많이 부여한다. 커널추정량은 z 를 구성하는 m 변수들 상호작용들을 고려한다. 다변량정규밀도커널은 이들 변수들 사이의 상관관계를 고려하기 때문이다. 모수 h 는 인근 관측지에게 부여하는 가중치의 정도를 나타낸다. h 가 작을수록 인근 관측지에 보다 많은 가중치를 부여하게 된다.

h 값은 CV(cross-validation)방법을 사용하여 결정한다. 이 방법을 간단히 설명하면, 먼저 i 번째 관측지를 제외한 다음 전체 모형을 다시 추정하여 CV추정치를 구한다. 즉, i 관측지에 대한 잔차 $\hat{u}_i = y_i - \hat{\beta}'_i x_i - \hat{g}_i(z_i)$ 를 구한다. 여기서 $\hat{\beta}_i$ 와 \hat{g}_i 는 i 번째 관측지를 제외하여 얻는 추정치들을 나타낸다. 모수 h 선택은 잔차의 평균제곱합

$$n^{-1} \sum_{i=1}^n \hat{u}_i^2 \text{ 을 최소로 하는 } h \text{ 이다.}$$

그 다음 \hat{y}_i 를 구하기 위해 i 관측지를 포함하여 모형을 다시 추정한다.

준모수적 방법에서 모형의 적합도를 위하여 추정한 \hat{y} 을 실제 값 y 에 회귀시켜 R^2 을 구하면, 이는 모수적 방법에서의 R^2 와 동일하다.

$g(\cdot)$ 의 함수형태가 귀무가설 즉, 모수적 함수형태와 대립가설 즉, 준모수적 함수형태와의 추정량의 분포가 통계적으로 유의하게 다른지 유무를 파악하기 위하여 Hausman (1978) 검정방법을 사용한다. $\hat{\beta}_{SP}$ 와 $\hat{\beta}_{OLS}$ 를 각각 준모수적 방법과 모수적 방법에서의 β 의 추정치라고 하자. 통상최소사승법(OLS)에서는 상수항은 포함시키지 않는다. 준모수적 방법에서 추정치는 모형 규정이

옳다면 효율성이 보장되지 않더라도 일치추정량이기 때문에 $\hat{\beta}_{SP}$ 와 $\hat{\beta}_{OLS}$ 는 유사한 값을 가진다. $g(\cdot)$ 의 함수가 비선형이라면 준모수적 방법의 추정치는 일치성을 가지지만 모수적 방법의 추정치는 그렇지 않다. 따라서 $\hat{\beta}_{SP} \neq \hat{\beta}_{OLS}$ 이다. \hat{V}_{SP} 와 \hat{V}_{OLS} 를 각각 $\hat{\beta}_{SP}$ 와 $\hat{\beta}_{OLS}$ 의 분산·공분산 행렬이라고 한다면 검정통계량 H 는 다음과 같다.

$$H = (\hat{\beta}_{SP} - \hat{\beta}_{OLS})' (\hat{V}_{SP} - V_{OLS})^{-1} \times (\hat{\beta}_{SP} - \hat{\beta}_{OLS}) \quad (5)$$

이 검정통계량 H 는 귀무가설하에서 $(k-1)$ 자유도를 갖는 χ^2 분포를 한다.

마지막으로 주택가격지수를 추정하기 위하여 단위표준 주택 특성을 x_0 와 z_0 라 하자. 주택가격지수는 앞에서 정한 h 를 사용하여 종속변수 $y - \hat{\beta}' x_0$ 를 구한다음 z_0 변수에 대한 $E(y - \hat{\beta}' x_0 | z_0)$ 를 비모수적으로 추정하고 주택가격지수 $\hat{\beta}' x_0 + \hat{g}(z_0)$ 를 구한다. 주택가격지수의 표준오차는 Härdle(1990)이 제시한 부트스트랩(bootstrap) 방법을 사용한다.

IV. 자료

이 논문에서 사용된 자료는 부동산뱅크에서 격주로 발행하는 아파트 시세자료로서 1989년부터 2000년까지 12년간 아파트 단지별 평형별 매매 하한과 상한의 평균을 이용하였다. 시세는 매년 4월을 기준으로 하였다.

공간적 범위는 서울 한강 이남 11개 구(강남구, 강동구, 강서구, 관악구, 구로구, 금천구,

동작구, 서초구, 송파구, 양천구, 영등포구)이며, 2개 이상 건물을 가지고 있는 아파트 단지를 대상으로 하였다.

아파트 단지의 일반적인 정보는 부동산 114(www.r114.co.kr)에서 구하였다. 주요 단지 정보로서는 전용면적, 단지규모, 입주년도, 최고층수, 현관출입문의 구조, 난방방식, 사용연료, 건물연수 등의 속성자료를 이용하였다.

이러한 비공간적 속성자료와 더불어 아파트 단지의 공간상의 위치를 확인하기 위하여 서울·수도권 도로교통지도 (1:10,000)와 국가수치지형도 (1:5,000)를 활용하여 각 아파트 단지를 공간상의 (x, y) 좌표로 표시하였으며 이를 비공간 속성자료와 연결하여 분석에 이용하였다. 수치지형도에 누락된 아파트 단지는 수치지형도상의 도로경계선을 참조하여 입력하였다. 아파트 단지를 입력한 다음 GIS 프로그램을 이용하여 주요 위치와의 접근성에 대한 자료를 구하였다. 접근성에 대해 많은 연구에서 사용되는 변수들로서 CBD와의 거리, 부도심과의 거리, 지하철역과의 거리 등을 계산하였다³⁾.

최종적으로 확인된 1,044개 단지를 대상으로 13,848 개의 시세자료 가운데 계열상관(serial correlation)과 공간상관(spatial correlation) 문제를 보정하기 위해 난수표를 이용하여 1,284개의 표본을 임의로 추출하였다. 그림 1은 분석에 사용된 아파트 단지의 공간적 분포와 주변 여건들을 보여주고 있다.

아파트 매매 가격에 영향을 미치는 변수

3) 도심과 부도심을 결정하는 것과 관련한 연구들이 많으나, 본 논문에서는 과거 12년간의 아파트시세 자료를 활용하였기 때문에 매년도 도심과 부도심의 식별은 용이하지 않을 뿐만 아니라 식별한다고 하더라도 식별의 오류에 따른 문제가 발생하기 때문에 본 논문에서는 CBD는 태평로1가동, 부도심은 여의도동과 테헤란로 중심을 각각 가정하였다.

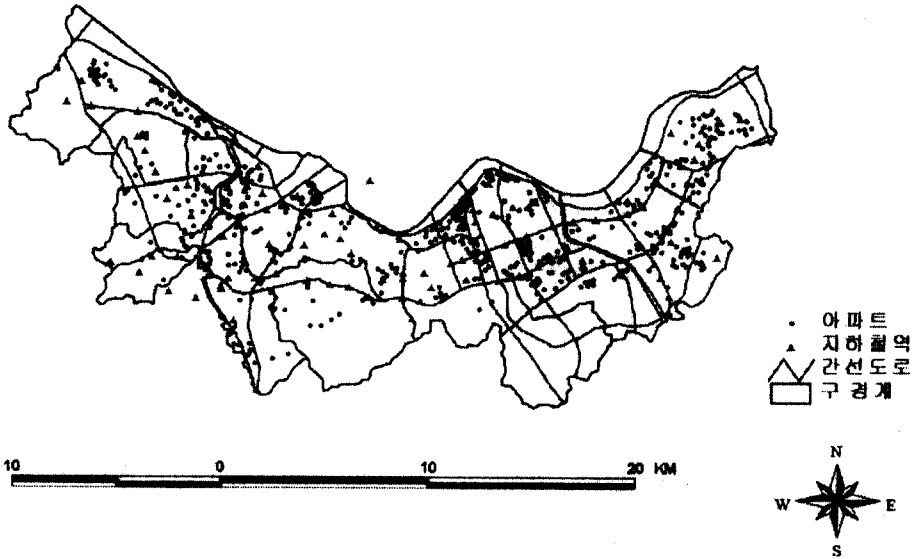
로서 전용면적, 방수, 건물연수를 사용하였으며, 더미변수로서 출입문 구조와 아파트 층수를 사용하였다. 아파트 층수는 10층 이상 아파트와 10층 이하 아파트로 더미변수로 구분하였다. 이는 일반적으로 10층 이하 아파트는 재건축 수요에 따른 아파트 가격의 프리미엄을 반영하기 위해서이다. 또 다른 더미변수는 난방방식과 연료방식을 사용하였으며, 학군 더미변수는 강서교육청을 기준으로 남부, 동작, 강남, 강동교육청으로 구분하였다. 마지막으로, 아파트 수요시장에 영향을 미치는 사회경제 변수로서 구별 인구규모를 사용하였다. 표 1은 분석에서 사용한 변수들에 대한 기초통계를 보여주고 있다.

사용된 소프트웨어로는 ArcInfo 8.0.1을 사용하여 아파트 단지를 점으로 입력하여 공간의 위치를 파악할 수 있도록 하였으며, 공간자료와 아파트 가격속성을 연결한 점 커버리지 형태로 저장할 수 있도록 하였다. 특성가격함수와 주택가격지수를 추정하기 위하여 GAUSS 프로그램을 사용하였다. 결과의 출력은 Arcview 3.1을 사용하였다.

V. 추정결과

준모수적 방법에서 중요한 모수로서 h 는 0.15에서 0.35까지 범위를 설정하여 0.01씩 h 를 증가하면서 최적의 h 를 탐색한 결과 0.25가 최적인 것으로 결정되었다. 이후의 모든 준모수적 추정은 h 를 0.25로 하여 추정을 하였다.

표 2는 두 개의 통상최소자승법(OLS)에 의한 모수적 추정치와 준모수적 방법에서 x 변수의 추정치를 나타내고 있다. 표에서 보는 바와 같이 추정계수들은 기대한 바와 같은 부



<그림 1> 분석 대상 아파트 분포 현황

<표 1> 기술통계 (1284 표본)

변수명	변수 설명	평균	표준편차	최소	최대
SPRICE	매매가격(백만원)	221.2	146.8	33.5	1,185.0
LPY	전용면적(평) (in log)	3.2	0.4	1.9	4.2
LNHHD	단지규모(100세대)	8.9	9.3	0.4	60.0
AGE	건물경과연수(년)	11.1	5.8	1.0	26.0
NROOM	방수	3.2	1.0	1.0	7.0
DENTR	출입문구조(복도식=1, 기타=0)	0.4	0.6	0.0	1.0
DHIGH	층수(고층=1, 저층=0)	0.8	0.4	0.0	1.0
DHEAT1	난방방식(중앙집중=1, 기타=0)	0.4	0.5	0.0	1.0
DHEAT2	난방방식(지역난방 =1, 기타=0)	0.4	0.5	0.0	1.0
DFUEL1	연료방식(기름 =1, 기타=0)	0.1	0.3	0.0	1.0
DFUEL2	연료방식(열병합 =1, 기타=0)	0.4	0.5	0.0	1.0
DSD1	남부교육청(영동포구, 구로구, 금천구=1, 기타=0)	0.2	0.4	0.0	1.0
DSD2	동작교육청(동작구, 관악구=1, 기타=0)	0.1	0.3	0.0	1.0
DSD3	강남교육청(강남구, 서초구=1, 기타=0)	0.4	0.5	0.0	1.0
DSD4	강동교육청(강동구, 송파구=1, 기타=0)	0.3	0.4	0.0	1.0
POP	구별 인구규모(만명)	50.3	9.4	27.1	68.9
DSTN	지하철역까지 거리(km)	0.9	1.0	0.04	8.4
DCBD	도심·부도심까지 거리(km)	5.0	2.9	0.3	12.1
YEAR	매매연도(1989=1)	7.7	3.4	1.0	12.0

호를 가지며 세 가지 방법 모두 비슷한 결과를 보이고 있다. 따라서, 이들 모수들의 추정치들은 모형 규정 방법에 크게 영향을 받지 않는 것으로 나타났다.

추정량들을 살펴보면, 먼저 전용면적은 0.982로서 1에 가까운 것을 알 수 있다. 이는 전용면적에 따른 주택가격 탄력성을 의미한다. 단지규모는 100 세대의 규모로 커질수록 0.6% 아파트 가격이 상승하는 것으로 나타났다. 아파트가 오래될 수록 아파트 가격은 연평균 0.4%~0.8% 떨어지는 것으로 나타났다. 방이 하나 추가됨에 따라 아파트가격은 8.4%~9.4% 증가한다. 현관출입문 구조를 보면 복도식이 계단식에 비해 8.6%~9.8% 가격이 낮다. 10층 미만의 아파트는 10층 이상 고층 아파트에 비해 5.0%~7.7% 아파트 가격이 높은 것을 알 수 있다. 이는 10층 미만의 아파트의 경우 재건축 가능성에 대한 프리미엄으로 판단된다.

난방방식은 개별난방에 비해 중앙난방과 지역난방이 각각 5.3%~6.3%, 6.9%~8.1%씩 높은 것으로 나타나 지역난방방식의 아파트가 가격이 높은 것으로 나타났다. 도시가스에 비해 연료 유형이 기름인 아파트는 가격이 낮은 반면 열병합 아파트는 높은 것으로 나타났지만 통계적으로 유의성은 없는 것으로 나타났다.

학군의 영향을 살펴보면 강서교육청을 기준으로 강남교육청은 14.6%~15.2% 정도 높으며, 강동교육청은 11.0%~13.5% 아파트 가격이 높은 것으로 나타났다. 반면, 남부교육청은 강서교육청에 비해 대략 10.0%~14.5% 낮은 것으로 나타났다. 동작교육청은 강서교육청과 차이가 없는 것으로 나타났다.

인구규모는 아파트 주변 지역의 도시화 규모를 나타내며 주택시장의 수요에 영향을 미친다. 추정결과 인구 만명이 늘어남에 따라 아파트 가격은 0.2% 높아지는 것을 알 수 있

다.

비모수적으로 추정하는 변수들의 영향을 알아보기 위하여 첫 번째 통상최소자승법(OLS 1)의 모수적 추정방법의 결과를 보면 먼저 지하철역까지 거리(DSTN)는 1km 멀어질수록 아파트 가격은 2.1% 하락하는 것을 알 수 있다. 또 다른 비모수적으로 추정하는 변수로서 도심·부도심과의 거리(DCBD)는 1km 거리가 멀어질수록 아파트 가격은 평균 3.2% 하락하는 것을 알 수 있다. 마지막으로 매매연도(YEAR)는 1989년을 기준으로 연평균 아파트가격이 2.8% 증가하여 온 것을 알 수 있다.

DSTN, DCBD 등 공간관련 변수와 매매연도(YEAR)의 시간변수들을 고차 다항식으로 모형에서 추정하여 본 결과 지하철역과의 거리는 3차 다항식까지만 유의성이 있으며, DCBD와 YEAR는 5차 다항식까지 통계적으로 유의성이 있는 것으로 나타났다. 이들 변수들간의 상호작용 효과는 1차식만 고려하였으며, 지하철역과 도심·부도심과의 거리에 대한 상호작용효과(DSTN×DCBD)만이 10% 유의수준에서 아파트 가격에 부(-)의 효과를 미치는 것으로 나타났다.

모형의 적합도 수준을 파악하기 위하여 결정계수를 보면, 준모수적 방법의 결정계수는 0.926으로서 두 모수적 방법의 결정계수 0.870, 0.916 보다 높은 것을 알 수 있다.

Hausman 검정통계량을 보면 모수적 방법의 모형규정을 귀무가설로 보고 준모수적 방법의 모형규정을 대립가설로 설정하였을 때 두 모수적 방법 모두 Hausman 검정 통계량이 귀무가설하에서 기각됨으로서 준모수적 방법의 모형규정이 바람직함을 알 수 있다. 따라서, 준모수적 방법과 모수적 방법간의 모수들의 추정량들의 차이는 크지 않지만 지하철역간의 거리(DSTN), 도심·부도심과의 거리(DCBD), 매매연도(YEAR) 변수들은 아파트

<표 2> 아파트가격함수 추정결과

변수	모수적 방법		준모수적 방법
	OLS (1)	OLS (2)	
상수항	1.564 (0.119)	-0.155 (0.151)	-
전용면적 (평, in log)	0.993 (0.035)	0.995 (0.030)	0.982(0.030)
단지규모 (100 세대)	0.006 (0.001)	0.006 (0.001)	0.006(0.001)
아파트년수 (년)	-0.004 (0.001)	-0.007 (0.001)	-0.008(0.001)
방수 (개)	0.084 (0.013)	0.087 (0.011)	0.094(0.011)
현관구조 (복도식=1, 계단식=0)	-0.098 (0.015)	-0.089 (0.013)	-0.086(0.013)
건물층수 (10층이상=1, 10층미만=0)	-0.050 (0.020)	-0.069 (0.017)	-0.077(0.018)
난방방식 (중앙난방=1, 기타=0)	0.053 (0.020)	0.072 (0.017)	0.063(0.017)
난방방식 (지역난방=1, 기타=0)	0.078 (0.025)	0.081 (0.023)	0.069(0.023)
연료유형 (기름=1, 기타=0)	-0.0180 (0.023)	-0.025 (0.020)	-0.040(0.021)
연료유형 (열병합=1, 기타=0)	0.004 (0.020)	0.020 (0.018)	0.014(0.018)
남부교육청 (더미)	-0.100 (0.032)	-0.145 (0.028)	-0.121(0.030)
동작교육청 (더미)	0.004 (0.031)	-0.020 (0.027)	-0.035(0.030)
강남교육청 (더미)	0.148 (0.027)	0.146 (0.024)	0.152(0.027)
강동교육청 (더미)	0.120 (0.026)	0.135 (0.023)	0.110(0.026)
인구규모 (만명)	0.002 (0.001)	-0.0003(0.001)	0.002(0.001)
지하철역과의 거리 (DSTN, 1km)	-0.021 (0.007)	-0.060 (0.084)	-
도심·부도심과의 거리 (DCBD, km)	-0.031 (0.004)	-0.277 (0.078)	-
매매연도 (YEAR, 1989=1)	0.028 (0.002)	2.204 (0.119)	-
$DSTN^2 \div 10$	-	0.361 (0.144)	-
$DSTN^3 \div 100$	-	-0.304 (0.122)	-
$DCBD^2 \div 10$	-	1.193 (0.370)	-
$DCBD^3 \div 100$	-	-2.308 (0.758)	-
$DCBD^4 \div 1000$	-	1.940 (0.687)	-
$DCBD^5 \div 10000$	-	-0.582 (0.227)	-
$YEAR^2 \div 10$	-	-7.862 (0.492)	-
$YEAR^3 \div 100$	-	12.755 (0.883)	-
$YEAR^4 \div 1000$	-	-9.536 (0.715)	-
$YEAR^5 \div 10000$	-	2.676 (0.214)	-
$DSTN \times DCBD$	-	-0.007 (0.004)	-
$DSTN \times YEAR$	-	-0.003 (0.003)	-
$DCBD \times YEAR$	-	-0.001 (0.001)	-
R^2	0.870	0.916	0.925
Hausman검정 (H0: OLS vs. H1: SP)	63.52 (0.000)	32.60 (0.005)	-

주 : ()은 표준오차임. 단 Hausman 검정에서의 ()은 귀무가설하에서 검정통계량의 유의확률임.

가격과 비선형 관계가 있음을 알 수 있다.

다음은 추정된 특성가격함수를 이용하여 7호선 지하철 일부 전철역 (대림역, 상도역, 이수역, 강남구청역) 반경 1km 이내 지역에 대해 월별 아파트가격지수를 추정하였다. 먼저, 4 개 지역에 대한 주택특성들은 표 3과 같다.

다음으로 표 3의 아파트 특성들을 사용하여 아파트가격지수를 추정하였다. 그림 2은 준모수적 방법에 의한 아파트가격지수(A)와 아파트가격지수증가율(B)을 나타낸 것이다. 아파트가격지수를 보면 1991년까지 모든 지역에서 아파트 가격은 상승하다가 1991년을 정점으로 하향곡면으로 바뀌고 있다. 1992년 중반까지 아파트 가격은 지속적으로 하락하고 있는데 이 기간동안은 수도권 신도시개발에 따른 주택시장의 공급측면의 충격효과가 주택 가격에 반영된 것이다. 1992년 중반부터는 아파트 가격은 모든 지역에서 상승하면서 1996년 중반까지 이어지다가 1997년 후반 IMF가 시작되면서 하락세로 돌아섰다. 1998년 중반 이후부터는 다시 아파트 가격이 상승하고 있다. (B)에서 아파트가격지수의 증가율 패턴을 보면 1992년부터 1998년 사이에 공간과 시간의 교차효과가 크게 나타나고 있다. 즉, 이 기간동안 지역별로 아파트가격지수의 증가율 패

턴은 복잡하게 나타나고 있다. 특히 IMF의 충격은 강남구청역 지역에서 크게 나타나고 있다. 하지만, 1998년 중반이후에는 모든 지역에서의 아파트가격지수 증가율은 매우 유사하게 나타나고 있다.

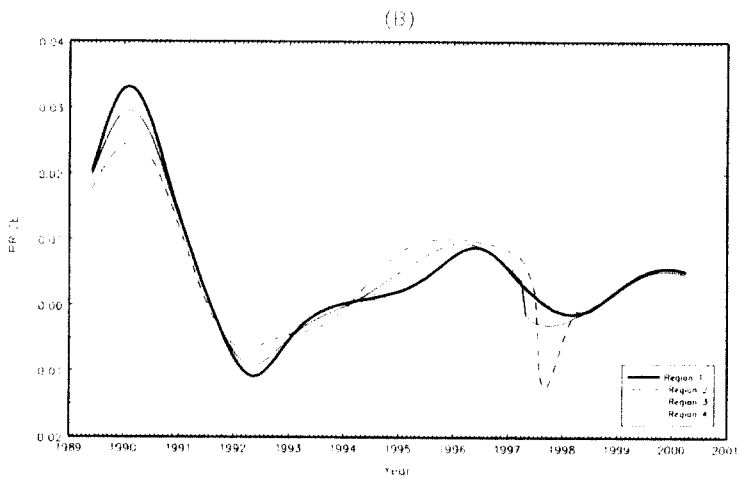
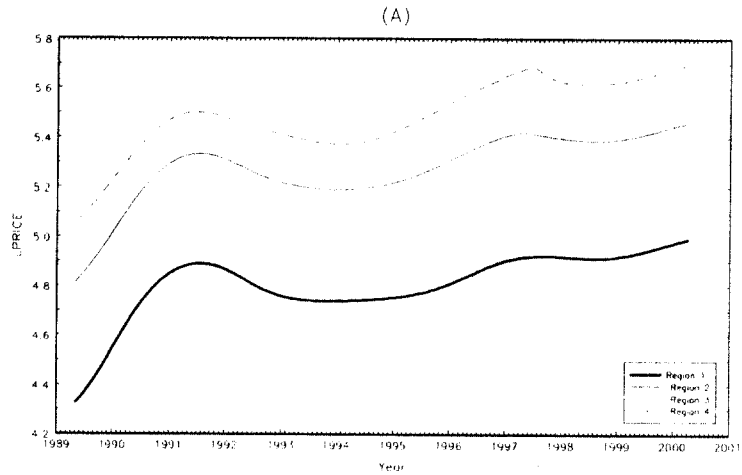
그림 3는 준모수적 방법과 모수적 방법에 의한 아파트가격지수를 비교하고 있다. 두 방법에 의한 지수는 큰 차이가 없으나, 초기에서는 모수적 방법이 준모수적 방법에 비하여 과소추정이 되고 있다. 이는 모수적 방법에서 모형에 가한 제약으로 인한 것이다. 즉, 매매 연도와 주택가격간의 비선형 관계를 추정하기 위하여 5차 다항식을 사용하였기 때문이다. 1997년이후에서도 모수적 방법이 준모수적 방법보다 과소추정을 보이고 있다.

그림 4는 준모수적 방법에 의한 아파트가격지수의 이의 표준오차를 제시하고 있다. 표준오차는 Härdle (1990)이 제시한 와일드 부트스트랩("wild" bootstrap)을 200회 반복하여 계산하였다.

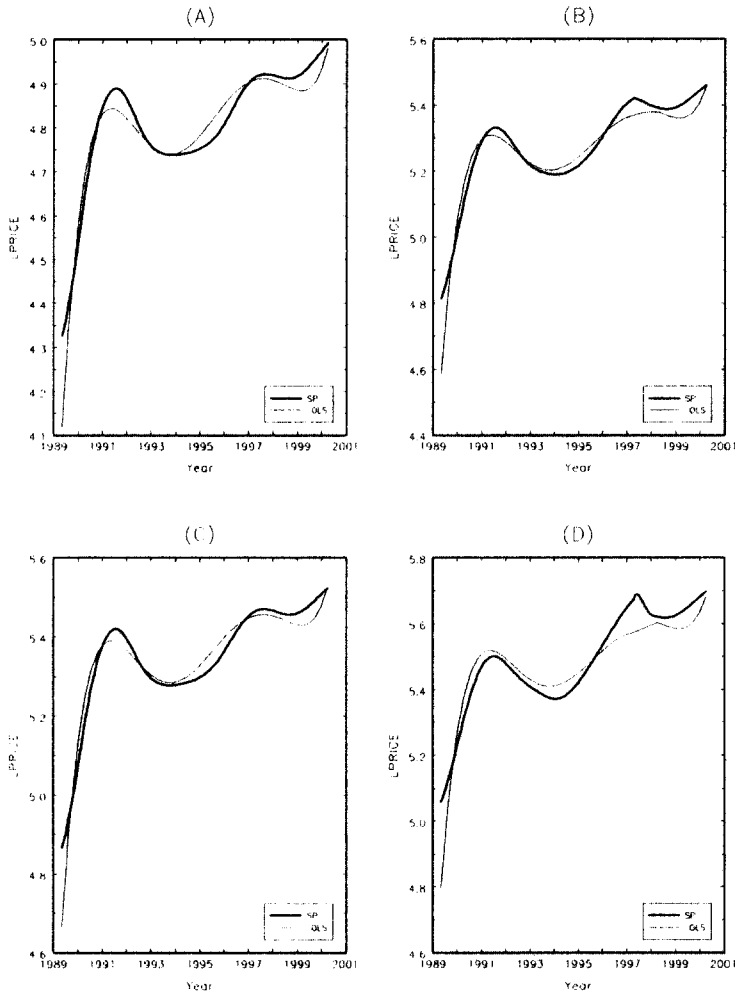
<표 3> 지역별 아파트 특성변수들의 평균치

변수명	대림역 지역 (Region 1)	상도역 지역 (Region 2)	이수역 지역 (Region 3)	강남구청역 지역 (Region 4)
log 전용면적 (평)	19.9	24.5	27.0	29.4
주택경과년수(년)	9.0	3.0	5.7	17.1
방수 (개)	2.5	3.2	3.1	3.4
단지규모 (100세대)	480.0	1036.6	1060.6	451.2
도심·부도심과의 거리(km)	4.7	2.8	6.0	1.4
지하철역과의 거리 (역 개통전, km)	0.46	0.90	0.66	1.27
지하철역과의 거리 (역 개통후, km)	0.45	0.55	0.60	0.44
구별 인구	40.5	42.9	43.4	54.8
학군	남부교육청	동작교육청	동작교육청	강남교육청

12 부동산학연구 제7집 제1호

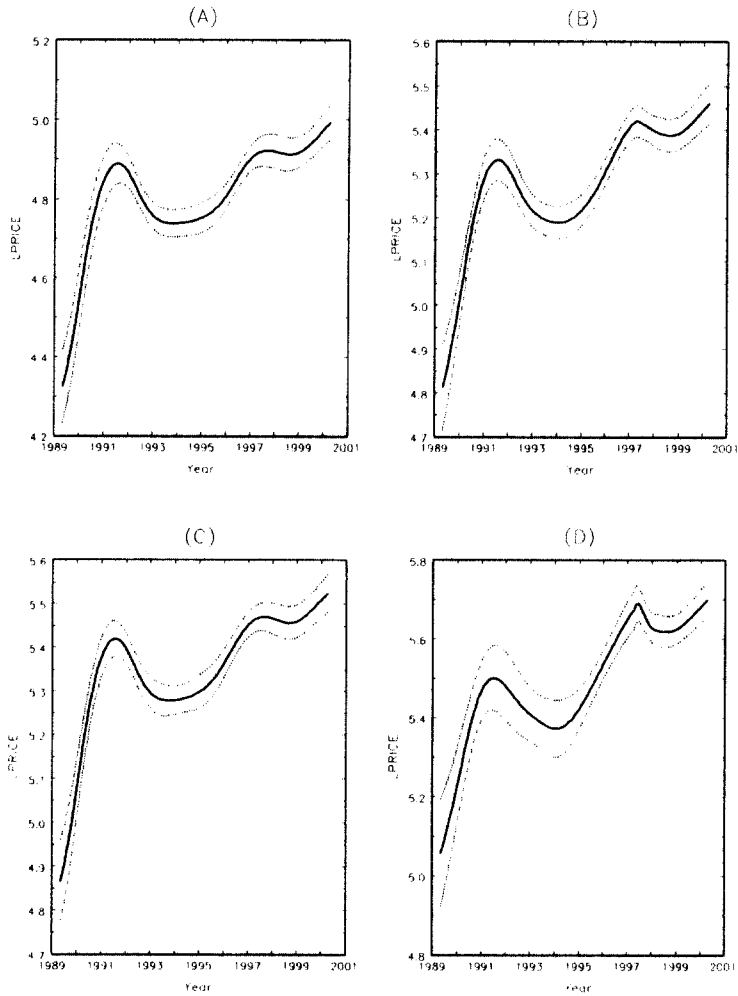


<그림 2> 준모수적 방법에 의한 아파트가격지수
 (A) 아파트가격지수 (B) 아파트가격증가율 지수



(A) 대림역 (B) 상도역 (C) 이수역 (D) 강남구청역

<그림 3> 모수적 방법과 준모수적 방법에 의한 아파트가격지수 비교



(A) 대림역 (B) 상도역 (C) 이수역 (D) 강남구청역

<그림 4> 준모수적 방법에 의한 아파트가격지수 추정 (점선은 95% 신뢰구간)

VI. 결론

주택 가격에 많은 영향을 미치는 공간 변수들과 시간 추세 등에 대한 특성가격함수들의 형태는 비선형으로 예상이 되지만, 비선형 함수 형태에 대해 이론적인 근거가 없어, 적절한 함수 형태를 선택하는 것이 쉽지 않다. 이는 적절한 함수 형태들이 장소와 시간에 따라 다르기 때문이다.

주택의 특성가격함수 추정에서 준모수적 방법은 모수적 추정과 비모수적 추정상의 장점들을 함께 제공함으로써, 주택특성변수들과 주택가격간의 비선형 문제를 쉽게 해결할 수 있음을 알 수 있다. 즉, 주택가격과 변수들의 효과가 선형 관계를 가지는 경우에는 모수적으로 추정을 하고, 주택가격과 변수들의 효과가 비선형 관계를 가지는 경우에 대해서는 비모수적으로 추정을 한다.

본 논문에서는 주택의 특성들을 감안한 주택가격함수를 모수적 방법과 준모수적 방법을 사용하여 추정하여 본 결과 모수적 방법보다 준모수적 방법에 의한 주택가격함수의 추정이 모형의 적합도가 높은 것으로 나타났다. 또한 지하철역과의 거리와 도심·부도심과의 거리, 그리고 매매연도와 같이 공간변수와 시간변수들에 대해 모수적 규정과 비모수적 규정에 대해서 Hausman(1978)의 모형규정에 대한 통계적 검정방법을 적용하여 본 결과 비모수적 함수 규정이 적합한 것으로 나타났다. 마지막으로 추정된 특성가격함수를 이용하여 7호선 지하철 일부 전철역(대림역, 상도역, 이수역, 강남구청역) 반경 1km 이내 지역에 대해 월별 아파트가격지수를 추정하였다. 준모수적 방법에 의한 아파트가격지수를 추정하여 본 결과 시간과 공간효과가 지역별로 차이가 있는 것으로 나타났으며, 특히 1992~1998년 사이가 시간과 공간의 교차효과가 크게 나타났다.

우리 나라에서의 주택가격지수는 한국주

택은행에서 발표하는 주택가격지수, 건설교통부의 지가지수, 서울상공회의소의 임대료 및 관리비지수들을 발표하고 있으나, 도시전체적으로 가격지수를 산정하기 때문에 국지적인 시·공간 효과의 차이를 나타내지 못하고 있다. 본 논문의 결과에서와 같이 주택가격에 시간효과와 공간효과가 영향을 미치는 것으로 보아 도시 전체적으로 가격지수를 산정하는 것에 신중을 기할 필요가 있다.

2001년 7월부터 시행되는 부동산투자회사(REITs)는 부동산투자에 대한 정확한 수익률을 계산하기 위해서는 부동산에 대한 보다 정확한 부동산가격함수체계와 부동산가격지수들을 필요로 하게 된다.

참고문헌

- 박헌수 (2001). 모수적 방법과 준모수적 방법에 의한 주택가격함수 추정에 관한 연구, 국토연구 34(4) 게재예정.
- 박헌수 (2000). 모수적 방법과 비모수적 방법에 의한 토지가격추정에 관한 연구. 한국지역개발 학회지 12(1): 101-109.
- 서경천·이성호 (2001). 지가의 공간적 변동에 따른 입지지의 분석에 관한 연구. 국토계획 36(1): 55-69.
- Cleveland, W.S. and S.J. Devlin.(1988). Locally Weighted Regression : An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association* 83 : 596-610.
- Härdle, W. (1990). *Applied Nonparametric Regression*. New York : Cambridge University Press.
- Härdle, W. and D. Linton (1994). *Applied Nonparametric Methods. Handbook*

- of *Economics*, Vol. 4, edited by R.F.Engle and D.L.McFadden. New York : Elsevier.
- Haurin, D.R. and P.H. Hendershott (1991). House Price Indexes: Issues and Results. *AREUEA Journal* 19: 259-269.
- Hausman, J.A. (1978). Specification Tests in Econometrics. *Econometrica* 46: 1251-1272.
- Horowitz, J.L., and W.Härdle (1996). Direct Semiparametric Estimation of Single-Index Models with Discrete Covariates. *Journal of the American Statistical Association* 91: 1632-1640.
- Knight, J.R., J. Dombrow, and C.F. Sirmans (1995). A Varing Parameters Approach to Constructing House Price Indexes. *Real Estate Economics* 23: 187-205.
- Lancaster, K.J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy* 74: 132-157.
- Meese, R., and N. Wallace (1991). Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housin Indices.g Price *AREUEA Journal* 19: 308-332.
- Pace, R.K. (1995). Parametric, Semiparametric, Nonparametric Estimation of Characteristic Values within Mass Assessment and Hedonic Pricing Models. *Journal of Real Estate Finance and Economics* 11: 195-217.
- Robinson, P.M. (1988). Root-N-Consistent Semiparametric Regression. *Econometrica* 56: 931-954.
- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 82: 34-55.
- Stock, J.H. (1991). Nonparametric Policy Analysis: An Application to Estimating Hazardous Waste Cleanup Benefits. *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, edited by W.A Barnett, J.Powell, and G.M. Tauchen. New York: Cambridge University Press.
- Thorsnes, P. and D.P.McMillen (1998). Land Value and Parcel Size : A Semiparametric Analysis. *Journal of Real Estate Finance and Economics*, 17(3) 233-244.
- Wallace, N. (1996). Hedonic-Based Price Indexes for Housing: Theory, Estimation, and Index Construction. Federal Reserve Bank of San Francisco Economic Review 1: 34-48.
- Yatchew, A. (1998). Nonparametric Regression Techniques in Economics. *Journal of Economic Literature* 36: 669-721.