

## 국지가중회귀(Locally Weighted Regression) 모형을 활용한 아파트 실거래가격지수 산정방법에 관한 연구\*

A Study on the Construction of the Transaction-Based Real Estate Price Index  
Using Locally Weighted Regression (LWR) Model

박 헌 수 (Park, Heonsoo)\*\*

### < Abstract >

A real estate price index tracks the price of a standard unit of housing over time. The index typically is constructed from observations on the sale price and their characteristics. Due to the requirement of the stability of the index, I assumed four different types of houses. With centering on them, I estimated locally weighted regressions (LWR) of hedonic price models and constructed the transaction-based price indices from January 2006 to October 2010 for Gangnam Gu, Seoul. The appraisal-based KB indices have appraisal smoothing problems in which they have low volatility and time-lag but the proposed transaction-based price indices do not show those problems.

주 제 어 : 국지가중회귀, 실거래가격지수, 반복매매가격지수

Keywords : Locally Weighted Regression, Transaction-Based Real Estate Price Index, Repeated Sales Price Index

\* 이 논문은 2011학년도 중앙대학교 학술연구비 지원에 의한 것임.

\*\* 중앙대학교 도시계획·부동산학과 교수, heonsoo@cau.ac.kr

## I. 서론

2005년 7월에 「공인중개사의업무및부동산거래 신고에관한법률」이 제정되면서 부동산거래를 의무적으로 신고하도록 되어 있다. 국토해양부는 2006년 1월부터 정상거래(arm's-length transaction)라 판단된 실거래자료를 발표하고 있다. 실거래자료가 발표되기 이전까지는 시세자료를 바탕으로 주택가격지수를 발표하였다. 실거래자료를 이용하여 주택가격지수를 작성하는 것은 주택시장의 상황을 보다 정확하게 반영할 수 있기 때문에 실거래자료를 이용한 주택가격지수에 대한 관심이 높아지고 있다.

주택가격지수(housing price index)는 표준단위 주택에 대해 주택가격동향을 시간대별로 추적하여 나타낸 것으로 정의한다(박헌수 2001). 주택가격지수를 작성시 많이 사용되는 방법으로 표본에 의한 라스파이레스(Laspeyres) 방식을 들 수 있다. 이 방법으로 지수를 작성하기 위해서는 매 시점마다 표준단위 주택인 표본주택의 가격정보가 필요한데 표본주택이 매 시점 거래되는 것은 아니기 때문에 지수작성에 어려움이 있다. 실거래자료를 이용하여 주택가격지수를 작성할 경우, 기존의 표본에 의한 라스파이레스(Laspeyres)식 지수작성방법과는 다른 접근이 필요하다(한국감정원, 2007).

그동안 주택가격지수를 작성할 때 생기는 문제들을 해결하기 위해 많은 연구들이 이루어져 왔다. 이들 연구들을 크게 반복매매모형(repeated sales model)과 특성가격모형(hedonic price model)

으로 나눌 수 있다. 반복매매모형은 시간이 경과 하더라도 주택의 구조적 특성들이 변화하지 않는다는 가정 하에서 주택가격의 변화를 시간에 대한 더미변수에 대해 회귀시켜 지수를 산출하는 방법이다. 대표적인 연구로 Bailey, Muth and Nourse(1963)가 있으며, 국내에 적용된 대표적인 연구로 이창무·김병욱·이현(2002), 이창무·김진유·이상영(2005) 등의 연구가 있다.

특성가격모형은 주택가격을 주택특성들에 회귀시켜 특성가격을 추정한 후에 주택가격지수를 산출하는 방법으로 Rosen(1974)의 연구에서 그 이론적 틀이 마련되었다. 대표적인 선행연구 고찰은 Case, Pollakowski and Wachter(1991), Clapp and Giaccotto(1992)가 있으며, 국내 적용사례는 이용만(2007), 한국감정원(2007), 박헌수(2009) 등의 연구가 있다.

반복매매모형에 의해 주택가격지수를 작성하는 방법은 특성가격모형이 가지고 있는 모형규정(model specification)에 대한 오류를 줄일 수 있는 장점이 있는 반면 동일주택에 대한 자료쌍을 확보하는데 있어 다수의 실거래자료를 버리게 되는 문제가 발생한다. Case and Shiller(1989) 연구에서는 93%의 실거래자료를 분석에서 제외하였다.<sup>1)</sup> 실거래자료가 지니고 있는 많은 정보를 유실하기 때문에 주택가격지수를 작성하는 지역적인 범위가 커지거나 지역적인 범위가 줄어들 경우에는 지수작성 기간을 연간(annual) 또는 분기별(quarterly)로 집계하여(aggregate) 지수를 작성할 수밖에 없다.

특성가격모형을 이용하여 주택가격지수를 작성하는 방법은 실거래 자료가 가지고 있는 정보

1) Case and Shiller(1989) 연구에서는 반복매매쌍이라고 하더라도 주택특성이 내부인테리어 공사 등으로 주택의 질(quality)이 변한 경우는 제외하였다.

를 대부분 이용할 수 있는 장점이 있다. 주택가격지수를 작성하는 기간마다 거래된 주택들의 특성들이 모두 다르기 때문에 가격변화가 특성변화에 기인한 것인지 시장상황에 따른 것인지를 구별할 필요가 있다. 특성가격모형은 주택의 특성 차이에 따른 가격 차이를 통제할 수 있어 시장상황에 따른 주택가격의 변화를 파악할 수 있다. 하지만 주택특성들과 주택가격간의 함수형태에 대해 정확하게 모형을 규정하는 것이 쉽지 않으며, 중요한 주택특성변수들이 모형에서 고려되지 않을 경우에 오류가 발생하여 주택특성들의 특성가격이 편향(bias)되게 추정될 수 있어 결국 주택가격지수에도 영향을 미치는 약점이 있다.

실거래 자료를 이용하여 위의 두 가지 방법을 적용하는 경우에도 공통적으로 생기는 문제가 지수의 안정성(stability) 문제이다. 주택가격지수의 안정성이란 실거래 자료가 새로 추가되었을 때 기존에 작성된 주택가격지수가 변동되지 않는 것을 의미한다(이용만, 2007). 최근 미국에서는 주택가격지수와 연동되는 파생금융 상품이나 보험상품의 개발이 많아지면서 주택가격지수의 안정성이 주요 이슈가 되고 있다. 실거래자료가 추가될 때마다 주택가격지수가 변동한다면, 이에 연동된 파생금융상품이나 보험상품을 판매할 수가 없게 된다. 또한 정부의 주택정책의 결정이 주택가격지수에 크게 의존하여 이루어지고 있기 때문에 주택가격지수가 수시로 바뀐다면 정부정책은 신뢰를 얻을 수 없을 것이다.

본 논문의 목적은 주택가격지수의 안정성 문제를 해결하는 한편 다양한 주택유형에 대해 주택가격지수를 작성방법을 제시하고자 한다. 특히, 주택유형을 다양하게 구분한 다음 특정 유형에 대한 주택가격지수를 작성하는 방법을 제시하고

자 한다. 이를 위해 2006년 1월부터 2010년 12월 까지 서울특별시 강남구 지역에서 거래된 아파트 자료들을 이용하여 다양한 주택유형에 대해 비모수적(nonparametric) 방법인 국지가중회귀(Locally weighted regression, LWR) 모형을 사용하여 특성가격모형을 추정한 다음 이를 통해 주택가격지수를 작성하고 이를 국민은행 가격지수와 비교와 함께 연구의 한계와 추후 연구과제를 제시하고자 한다.

본 논문의 구성은 다음과 같다. II절에서는 국지가중회귀모형에 의한 가격지수 작성방법에 대한 모형을 소개하고, III절에서 지수작성에 사용된 자료에 대해 기술하고, IV절에서는 실거래가격지수를 작성과, 마지막 V절에서는 결론 및 향후 연구과제를 제시한다.

## II. 국지가중회귀모형에 의한 주택 가격지수 작성방법

주택가격지수는 표준화된 주택에 대해 주택가격의 변화를 지수화한 것이다. 주택가격의 변화는 주택이 갖고 있는 특성의 변화에 의해 발생될 수도 있고, 특성 변화가 없음에도 불구하고 시장상황의 변화에 의해 발생할 수도 있다. 예를 들어 주택의 내부시설을 최신 시설로 교체함으로써 주택가격이 상승할 수도 있으며, 근처에 지하철역이 개통되어 교통이 편리해짐에 따라 주택가격이 상승할 수도 있다. 이러한 가격변화는 주택특성의 변화에 따른 것으로, 전자를 주택의 내부특성 변화에 따른 가격변화라 한다면, 후자는 주택의 외부특성 변화에 따른 가격변화라 할 수 있다. 또한 주택의 내·외부 특성 변화가 없음에도

불구하고 주택수급상의 문제나 이자율의 변동 등과 같은 시장상황의 변화에 의해 주택가격이 변할 수도 있다.<sup>2)</sup>

이론적으로는 주택특성 변화에 따른 가격의 변화는 지수산정 시 가격변화에서 제외하는 것이 타당하다. 주택의 특성이 바뀌었다면, 이미 해당 주택은 품질면에서 이전의 주택과는 다른 주택이므로 지수 작성 시 신규 주택처럼 취급하는 것이 타당하다(이용만, 2007).

따라서, 주택가격지수를 작성하기 위해서는 주택의 특성변화에 따른 가격변화를 고려해야 한다. 반복매매모형은 동일한 주택특성에 대한 반복매매 쌍을 이용하여 주택가격지수를 작성하는 반면, 특성가격<sup>3)</sup>모형(hedonic price model)은 주택의 특성가격(hedonic price)을 추정한 후, 이를 이용하여 매 시점별로 가상적인 표준주택의 가격을 추정하여 지수화하는 방법이다. 본 논문의 목적이 지수의 안정성과 특정 주택유형에 대한 주택가격지수를 작성하는 점이라는 것을 감안하여 반복매매모형보다는 특성가격모형에 의한 지수작성 방법을 사용하기로 한다.

특성가격함수 모형에 의해 주택가격지수를 작성할 때 함수형태의 선택과 다양한 주택유형에 대한 지수를 산정하고자 사용되는 표본선택의 문제 등을 고려하여야 한다(Meese and Wallace, 1991).

국지가중회귀(locally weighted regression, LWR) 모형에 의한 주택의 특성가격함수를 추정하는 것은 함수형태에 대한 보다 유연한 비모수(nonparametric) 추정방법을 사용함으로써 특성가격함수의 형태들에 대해 사전적으로 함수형태를 규정하여야 하는

문제들을 완화시킬 수 있다. 또한 매 시점 거래되는 주택특성들을 이용하여 지수를 작성할 때 생기는 안정성(stability) 문제는 주택가격지수를 구하고자 하는 주택유형들에 대한 주택특성으로 고정한 후 이에 대한 특성가격의 국지적인 근사치 값들을 이용하여 가격지수를 작성한다.

Meese and Wallace(1991)에 따르면,  $t$  시점,  $j$  주택가격에 대해 자연대수 값을 취한 것을  $P_{jt}$  라 하고 다음 특성가격함수에 의해 설명된다고 가정하자.

$$P_{jt} = m_t + \beta_t' G(x_{jt}) + u_{jt}, \quad (1)$$

여기서  $m_t$ 는 주택특성으로 설명되지 못하는 잔차의 평균으로 시간에 따라 변화한다.  $\beta_t$ 는  $k \times 1$  모수벡터를 나타내며,  $x_{jt}$ 는  $t$ 시점,  $j$  번째 주택이 가지는  $k$ 개의 주택특성들로 이루어진 벡터이다.  $G$ 는 주택특성들의 함수를 말하며,  $u_{jt}$ 는 오차항을 나타낸다. 주택가격이 비정상적(nonstationary)인 특징을 보이는 것은 일정크기의 증가분(drift)에 기인한 것으로 보고  $m_t$ 는 다음과 같이 모형화한다.

$$m_t = \alpha_t D_t + e_t, \quad (2)$$

여기서  $D_t$ 는 더미변수로서  $t$  시점에 대해서는 1의 값을 가지며, 다른 시점에 대해서는 0의 값을 가진다.  $\alpha_t$ 는 회귀계수로서 주택특성 즉  $\beta_t' G(x_{jt})$ 에 의해 설명이 된 후에  $t$  시점에서의

2) 한 심사자는 Meese and Wallace (1991)의 예와 같이 주택수급상의 문제나 이자율의 변동 역시 주택특성으로 보아야 한다고 주장하였다.

3) 주택 특성가격은 헤도닉가격, 특성 잠재가격 등으로도 불리고 있다.



평균가격수준을 측정하는 계수이다.  $e_t$  는 시계열 오차항으로 백색잡음(white noise)라고 가정한다. 위 두 식을 결합하게 되면 다음과 같은 특성가격 함수를 얻을 수 있다.

$$P_{jt} = \beta_t' G(x_{jt}) + \alpha_t D_t + (e_t + u_{jt}). \quad (3)$$

식 (3)을 추정하기 위한 바람직한 방법은  $G(\cdot)$  함수에 대해 사전적으로 가장 적게 제약을 두는 것이다. Meese and Wallace(1991)는 모수적인 유연한 함수형태로서 초월대수와 이중로그 함수와 같은 형태들에 대해 검정을 해 본 결과 여러 도시들에 대해 하나의 함수형태를 규정할 수 없다고 주장하였다.

비모수적 방법을 사용하는 경우 함수형태에 대한 유연성이 가장 크다. 하나의 비모수적 방법으로 Cleveland and Devlin(1988)이 처음 소개한 국지가중회귀모형(locally weighted regression, LWR)이 있다. 이 방법은 회귀표면을 이동평균 방법으로 추정하는 방법으로 고정된 지점에서  $G(\cdot)$ 에 대한 국지적인 근사화를 제공해주는 장점이 있다.

Meese and Wallace(1991)은 (3)식을 다음과 같이 변형하여 사용하였다.

$$P_{jt} - P_{mt} = \beta_t' G(x_{jt}) + v_{jt}, \quad j = 1, 2, \dots, N_t, \quad t = 1, \dots, T \quad (4)$$

여기서  $P_{mt}$ 는  $t$ 시점의 주택가격 중위수 값의 로그값이며,  $v_{jt}$ 는 오차항이다. 비모수적방법에

의한 LWR 방법은 정상적(stationary) 종속변수와 독립변수들이 요구되기 때문에  $P_{jt}$ 에 존재하는 추세를 각 시점의 종속변수의 중위수값을 빼서 제거한다.<sup>4)</sup>

비모수적 방법은 다양한 주택에 대해 주택특성가격을 국지적(locally)으로 추정한다. 모든 주택유형에 대해서 추정이 가능하지만 일반적으로 중심지점에 대해 추정하는 것이 보편적이다. 중심 지점을 잡는 방법은 두 가지가 가능하다. 첫 번째 방법은 중심 지점을 고정하지 않고 LWR 추정량을 구하는 방법이다. 이는 각 시점  $t$ 에서 특성들의 중위수값으로 이루어진 벡터,  $X_{mt}$ 를  $G(\cdot)$ 를 적합시키는 국지적인 중심점으로 택하는 방법이다. LWR은  $X_{mt}$ 에 근접된 전체 관측치에 대한 일정 부분을 나타내는  $\delta$ ,  $0 < \delta < 1$ 를 결정하여야 한다. 여기서 근접성은 표본의 모든 지점과  $X_{mt}$ 간의 유클리드 거리를 사용하여 측정한다. 즉, 거리는 다음과 같이 정의할 수 있다. 식에서 합계는  $k$ 개의 주택특성들의 특성차이들의 제곱을 모두 합한 것이다.

$$D(X_{mt}, X_{jt}) = \sqrt{\sum (X_{mt} - X_{jt})^2} \quad (5)$$

두 번째 방법은 구하고자 하는 주택유형의 주택특성을 사전에 가정한 다음 이를 매 시기마다 고정시킨 후에 이를 중심으로 LWR 추정량을 구하는 방법으로 추정하는 방법은 첫 번째와 동일하다. 본 논문에서는 지수의 안정성과 다양한 주택유형에 대한 주택가격지수를 작성하는 것이 목

4) 본 논문에서는  $t$ 시점의 대표값으로 평균을 사용할 수 있으나 이상치(outlier)에 덜 영향을 받는 중위수(median) 값을 사용하였다. 평균값으로 추세를 제거하는 경우에는 표본의 표준편차로 나누어서 표준화한다. 이 경우는 일반적으로 독립변수에도 표준화방법을 적용한다.

적이기 때문에 후자의 방법을 사용한다.<sup>5)</sup>

특성가격곡면(hedonic surface)은  $X_{mt}$  에서 근접된 관측치들 가운데  $N_t$  의  $\delta$  부분에 해당하는 관측치를 사용하여 가중최소자승회귀분석에 의해 국지적으로 근사화시킨 것이다. 가중치는 Cleveland and Devlin(1988)에서 정의한 것을 사용한다. 즉,

$$W = V \left[ \frac{D(X_{mt}, X_{jt})}{D(X_{mt}, X_{[\delta N_t], t})} \right], \quad (6)$$

여기서  $D(X_{mt}, X_{[\delta N_t], t})$  는  $t$  시점에서  $X_{mt}$  와 그 지점에서 근접된 관측치 가운데  $[\delta N_t]$  번째 자료까지의 거리이다.<sup>6)</sup> 본 논문에서는  $V(\cdot)$  함수형태는 tricube 함수형태를 사용한다. 즉,

$$V(s) = \begin{cases} (1 - s^3)^3 & s < 1 \\ 0 & s \geq 1 \end{cases} \quad (7)$$

### III. 자료

본 논문에서는 실거래 자료가 신고되기 시작한 2006년 1월부터 2010년 12월까지 서울 강남구 지역에 대한 실거래 자료를 이용하였다. 현재 국토해양부에서 인터넷 홈페이지(<http://rt.mltm.go.kr>)를 통해 공개되고 있는 실거래 자료는 아파트 단

지명, 면적, 계약일자, 거래금액, 거래 층수에 대한 정보이다. 공개된 자료만을 가지고 주택특성가격함수를 추정하기에는 특성자료가 부족하다. 따라서 아파트에 대한 추가적인 특성자료는 부동산 114(<http://www.r114.co.kr>)를 통해 구하였다.

본 논문에서 고려한 주택특성 변수로는 전용면적, 방수, 욕실수, 출입문구조, 아파트 층수를, 단지특성 변수로는 입주년도를 이용한 건물경과연수, 세대수, 단지 내 동수, 단지 내 가장 높은 건물 층수, 난방방식, 사용연료, 재건축여부, 시공사 등을 고려하였다. 최종적으로 실거래주택가격지수를 작성하기 위해 사용된 설명변수는 면적, 건물연수, 방수, 단지 내 총세대수, 재건축유무, 계단식구조를 선정하였다.

일반적으로 방수와 욕실수는 많을수록 면적은 넓을수록 주택가격은 높다. 방수, 욕실수, 면적들 사이에는 높은 상관관계가 존재하기 때문에 실거래가격자료에서 공개되는 면적에 대한 자료만을 사용하였다. 아파트 단위면적당 주택가격에 대한 특성가격함수를 추정할 때 면적을 설명변수로 추가하는 것은 면적에 프리미엄이 존재하는지 여부를 검증하기 위해서이다. 추정계수가 통계적으로 유의하게 양(+)의 값을 가지는 경우 면적이 증가할수록 단위면적당 주택가격이 선형(linear)으로 증가하는 것이 아니라 면적이 증가할수록 주택가격에 프리미엄(premium)이 존재한다고 볼 수 있다.

건물연수는 아파트가 거래된 시점과 입주년도를 기준으로 계산하였다. 일반적으로는 건물연수가 오래될수록 주택가격은 하락하게 된다. 하지

5) Meese and Wallace(1991)은 전자의 방법을 사용하였다. 즉, 매 분기마다 평균값을 중심으로 모형을 추정하였기 때문에 지수를 작성할 때 매 분기 주택특성이 변하는 문제가 발생한다. 본 논문에서는 후자의 방법으로 주택유형을 미리 가정하여 고정하였기 때문에 매기마다 주택특성이 변하지 않는다.

6)  $[\delta N_t]$ 는  $j$  시점에 총 관측치가  $N_t$  라 하면  $\delta \times N_t$ 에서 정수부분을 말한다.

만 건물연수가 증가함에 따라 일정기간까지는 주택가격이 하락하지만 그 기간을 지나게 되면 재건축 가능성이 높아지면서 가격이 상승하게 된다.

재건축 여부를 모형에서 반영하는 방법은 두 가지 방법이 가능하다. 첫 번째 방법은 건물연수에 대한 이차항을 추가하여 추정하는 방법이며, 두 번째 방법은 더미변수를 사용하여 재건축효과를 추정하는 방법이다. 첫 번째 방법은 매기 특성가격함수모형을 추정시 모수추정의 통계적 유의성과 부호가 불규칙함에 따라 가격지수의 변동폭이 지나치게 크게 나타나는 경향이 있다. 두 번째 방법은 아파트 단지에 대해 재건축 대상지 여부를 파악하여야 한다. 본 논문에서는 실거래 가격자료가 방대하여 재건축여부를 파악하는 것이 쉽지 않아 첫 번째 방법을 사용하였다.

단지규모를 나타내는 세대수는 규모의 경제효과를 나타낸다. 단지규모가 클수록 주택가격에 양(+)의 영향을 미칠 것으로 기대된다. 기타 난방 방식, 사용연료, 가구당 주차대수 등은 아파트단지에 대한 자료가 없는 경우가 있으며, 또 특정 지역 특정시기에 실거래된 자료가 모두 동일한 특성을 가지는 경우가 많아 특성가격모형을 추정하는 변수에서 제외하였다.

#### IV. 실거래주택가격지수 작성

실거래가격자료에 대한 특성가격함수모형은 앞에서 결정한 주택특성변수들을 설명변수로 하고 주택가격을 피설명변수로 한 회귀모형을 의미한다. 피설명변수는 단위면적당( $m^2$ ) 가격(만원)으로 자연대수를 취한 값을 사용하였으며, 설명변수들은 더미변수인 계단식구조에 대한 변수만을 제외

하고 모든 변수들은 자연대수값을 취하였다.

모형추정은 피설명변수를 상수항을 포함한 여러 주택특성변수들에 회귀시켜 추정계수를 구한다. 이때 지수를 작성하기 위해 사용되는 방법은 세 가지 방법이 가능하다. 첫 번째 방법은 시간 더미변수를 이용한 방법으로 2006년 2월부터 2010년 12월까지 매월 거래된 시기에 대한 더미변수를 사용하여 추정된 모수를 통해 가격지수를 작성하는 방법이다. 두 번째 방법은 매월 거래된 자료들만 사용하여 특성가격함수모형을 추정 후 추정된 모수 값과 주택특성들을 이용하여 특성가격들의 합을 구한 후에 지수를 작성하는 방법이다. 세 번째 방법은 특정 주택유형의 아파트를 기준으로 실거래된 개별 아파트와의 특성들의 차이를 식 (5)로 구하고 이를 식 (6)으로 가중치를 정한 다음 식 (4)의 국지가중회귀(Locally weighted regression, LWR) 방법으로 모수를 추정한 다음 지수를 작성하는 방법이다.

첫 번째 방법은 새로운 실거래 자료가 추가될 경우 기존의 주택가격지수를 다시 산정해야하는 지수의 안정성 문제가 발생한다. 두 번째 방법은 매 기간마다 특성가격함수를 추정하기 때문에 새로운 실거래자료가 추가되더라도 기존의 지수를 수정할 필요는 없지만 매 기간마다 주택특성들이 변하기 때문에 주택특성의 변화가 주택가격지수에 반영되는 문제가 있다. 본 논문에서는 세 번째 방법을 이용하여 매기 주택특성가격모형을 추정하되 주택특성은 특정한 주택유형을 중심으로 고정시킴으로서 주택특성의 변화가 주택가격지수에 반영되지 않도록 하였다. 또한 함수형태에 대한 모형규정(model specification) 오류를 최소화하기 위해 모수추정을 비모수적(nonparametric) 방법인 국지가중회귀(LWR) 방법으로 추정하였다.

〈표 1〉 주택유형별 주택특성

주택특성	주택유형1 (HPIDX 1)	주택유형2 (HPIDX 2)	주택유형3 (HPIDX 3)	주택유형4 (HPIDX 4)
전용면적(m <sup>2</sup> )	85	100	85	100
주택연수(년)	10	10	25	25
방수(개)	3	4	3	4
총세대수(세대)	300	300	300	300
출입문구조	계단식	계단식	계단식	계단식

본 논문에서의 방법은 다양한 주택유형에 대해서 주택가격지수를 산정할 수 있다. 첫 번째와 두 번째 방법은 주택가격지수를 유형별로 구분하여 추정하고자 하는 경우 표본을 나누어야 하기 때문에 이 과정에서 특정 유형의 표본이 없을 경우가 생길 수 있다. 가령 가격지수를 소형 아파트와 중대형 아파트로 규모를 기준으로 구분하여 지수를 작성할 경우 유형별 아파트를 세분하는 과정에서 주택거래가 없는 유형이 발생하는 경우는 주택가격지수를 작성할 수 없게 된다.

주택거래가 있는 경우라 하더라도 두 번째 방법을 이용하여 지수를 작성할 때 비록 앞의 문제를 피하기 위해 표본을 구분하지 않고 주택특성들을 고려하여 특성가중합수를 추정하더라도 피설명변수에 대한 추정치의 오차는 설명변수들의 평균을 이용하여 추정하는 경우가 가장 작다. 다양한 주택유형에 대한 주택가격지수 작성이 비록 가능하다고 할지라도 매 기간 거래되는 실거래가격자료의 주택특성들의 평균값이 변하기 때문에 특정유형에 대한 주택가격 추정치에는 오차가 커지게 된다. 주택가격 추정치에 대한 오차가 클수록 이를 기초로 작성한 가격지수 역시 오차가 커지게 되는 문제가 있을 수 있다.

세 번째 방법은 원하는 주택특성들에 대한 다양한 유형의 실거래가격지수를 작성할 수 있다.

가령 주택특성들이 면적 85m<sup>2</sup>이면서, 건물연수가 10년이 경과되고, 계단식 복도구조를 가지고 있는 아파트에 대한 주택가격지수를 추정하고자 할 경우 해당 아파트를 기준으로 실거래된 아파트와의 주택특성들의 차이를 구한 후 이들 차이가 크면 클수록 가중치를 적게 부여하는 방식으로 국지가중회귀분석을 이용하여 모형을 추정하기 때문에 주택가격 추정치의 오차는 구하고자 하는 주택특성을 가지고 있는 지점에서 가장 낮게 되어 앞에서의 문제들을 해결할 수 있다.

이후에서는 서울특별시 강남구에 대한 2006년 1월부터 2010년 12월까지 국지가중회귀모형을 이용한 가격지수를 추정하였다. 먼저 구하고자 하는 주택유형, 즉 표준주택을 설정한다. 본 논문에서는 다음 <표 1>과 같이 4가지 유형의 주택으로 구분하였다.

LWR 방법으로 모형을 추정할 때 주택유형별 주택특성과 개별 아파트들의 주택특성과의 차이를 식 (5)를 이용하여 구한 다음 주택특성이 유사한 주택에 대해서는 가중치를 많이 부여하고 주택특성의 차이가 많이 나는 아파트에 대해서는 가중치를 적게 부여한다. 매 기간마다 실거래자료가 다르기 때문에 이들 가운데 몇 개의 표본을 모형추정에 사용할 것인지 비율  $\delta$  ( $0 < \delta < 1$ )은 AIC(Akaike Information Criteria)와 SC(Schwarz

Criteria) 통계량을 구한 다음 이들 통계량이 가장 낮은 값을 갖는  $\delta$  값을 구하였다.

특성가격모형은 2006년 1월부터 2010년 12월 까지 총 60개이다. 이들 추정결과를 모두 제시하는 것은 불가능하기 때문에 본 논문에서는 가격 지수만을 제시하기로 한다.<sup>7)</sup>

<그림 1>은 주택유형별 실거래가격지수를 보

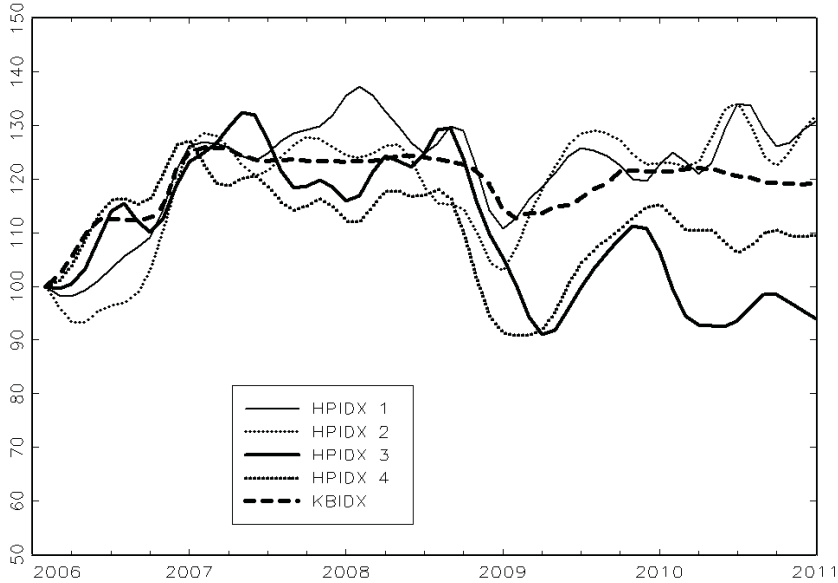
여주고 있다. 주택유형 1과 2는 2006년 상반기에는 가격이 하락하다 2분기부터 상승하는 반면 재건축 대상 가능성이 높은 주택유형 3과 4는 지속적으로 상승하고 있다. 2007년부터 2008년 상반기까지는 재건축 대상 아파트는 가격이 하락하다가 상승하는 모습을 보이는 반면 주택유형 1은 2007년 말까지 지속적으로 상승하고 있다. 주택

<표 2> 주택유형별, 월별 적정 q 값

구분	주택 유형1	주택 유형2	주택 유형3	주택 유형4	구분	주택 유형1	주택 유형2	주택 유형3	주택 유형4
200601	0.33	0.34	0.30	0.34	200807	0.30	0.36	0.34	0.34
200602	0.34	0.34	0.30	0.30	200808	0.32	0.33	0.36	0.36
200603	0.31	0.31	0.30	0.30	200809	0.52	0.52	0.30	0.30
200604	0.39	0.41	0.32	0.64	200810	0.30	0.33	0.39	0.30
200605	0.30	0.30	0.63	0.63	200811	0.98	0.98	0.30	0.30
200606	0.55	0.62	0.40	0.41	200812	0.82	0.82	0.33	0.33
200607	0.49	0.49	0.31	0.31	200901	0.84	0.41	0.55	0.55
200608	0.45	0.46	0.62	0.62	200902	0.44	0.38	0.35	0.30
200609	0.44	0.45	0.31	0.71	200903	0.99	0.99	0.37	0.37
200610	0.40	0.41	0.33	0.33	200904	0.52	0.45	0.48	0.49
200611	0.51	0.41	0.60	0.60	200905	0.43	0.40	0.39	0.40
200612	0.38	0.38	0.30	0.30	200906	0.47	0.42	0.50	0.51
200701	0.51	0.51	0.65	0.65	200907	0.40	0.50	0.45	0.47
200702	0.43	0.43	0.56	0.56	200908	0.51	0.51	0.48	0.51
200703	0.54	0.55	0.53	0.55	200909	0.58	0.52	0.43	0.43
200704	0.36	0.30	0.53	0.53	200910	0.31	0.30	0.30	0.43
200705	0.33	0.33	0.30	0.30	200911	0.42	0.52	0.43	0.45
200706	0.41	0.49	0.32	0.30	200912	0.55	0.55	0.31	0.33
200707	0.39	0.39	0.59	0.61	201001	0.51	0.51	0.39	0.47
200708	0.53	0.53	0.44	0.44	201002	0.34	0.33	0.33	0.34
200709	0.49	0.49	0.48	0.48	201003	0.31	0.35	0.32	0.35
200710	0.48	0.48	0.56	0.56	201004	0.34	0.41	0.44	0.44
200711	0.33	0.39	0.56	0.56	201005	0.34	0.30	0.38	0.32
200712	0.54	0.54	0.59	0.59	201006	0.32	0.36	0.37	0.32
200801	0.59	0.53	0.49	0.49	201007	0.30	0.30	0.35	0.35
200802	0.62	0.62	0.57	0.57	201008	0.30	0.38	0.35	0.30
200803	0.66	0.67	0.38	0.51	201009	0.31	0.31	0.39	0.39
200804	0.63	0.63	0.51	0.51	201010	0.59	0.59	0.39	0.39
200805	0.57	0.32	0.38	0.38	201011	0.55	0.55	0.45	0.46
200806	0.60	0.60	0.36	0.36	201012	0.57	0.57	0.32	0.40

7) 추정결과는 저자에게 요청하기 바람.

〈그림 1〉 강남구 주택유형별 아파트 실거래가격지수



유형 3과 4의 주택가격지수는 2008년 하반기부터 급속하게 하락하기 시작하면서 2009년 상반기까지 40% 정도 가격이 하락하였다. 이는 참여 정부가 지속적으로 추진하였던 서민주거복지증진과 주택시장 안정화를 위한 규제정책들이 발표되고, 특히, 개발이익환수제 도입과 재건축에 대한 규제강화에 기인한 것으로 보인다.

2009년에 들어와서는 재건축 대상 아파트들의 주택가격은 상승하고 있지만 2010년 들어와서 다시 하락과 상승을 반복하고 있다. 특히 소형 아파트에 대한 가격변동이 크게 나타나고 있다. 반면 주택유형 1과 2의 아파트는 2008년에는 하락세를 지속하고 있다. 소형보다 대형 아파트의 가격하락이 높게 나타나고 있다. 2009년을 지나서는 상승과 하락을 반복하고 있지만 전반적으로 가격이 증가하는 추세를 보이고 있다.

반면 국민은행 가격지수는 2006년 기간 동안

에는 상승하지만 2007년 이후부터는 2009년 1분기에 소폭 하락한 것 이외에는 가격변동이 거의 없는 모습을 보이고 있다. 이러한 경향은 실거래 가격지수는 시장에서 거래가 이루어진 가격을 기준으로 지수를 작성한 것인 반면 국민은행 가격지수는 판매자가 받고자 하는 호가(asking price)를 기준으로 작성된 것으로 차이가 있다. 전반적으로 국민은행 가격지수는 평활화(smoothing)된 지수를 보이는 반면에 국지가중회귀모형에 의한 실거래가격지수는 변동성이 크게 나타나고 있다.<sup>8)</sup>

## V. 결론

실거래가격에 기초한 새로운 주택가격지수는 시장상황에 대한 정확한 정보를 제공함으로써 올

8) 국민은행 가격지수의 평활화 문제에 대해서는 이용만·이상한(2008) 참조.



바른 주택정책을 이끌어낼 수 있는 기초가 될 뿐 아니라 다양한 파생적인 투자정보를 제공할 수 있다. 주택의 질적 차이에 따른 가격차이를 지수화할 수 있으며, 주택가격지수를 이용하여 투자 수익률 지표를 개발할 수 있다.

본 논문에서는 실거래가격에 기초한 주택가격 지수를 작성하는 방법들에 대해 살펴보았다. 주택가격지수는 최소한 주택의 특성차이에 따른 가격 차이를 걸러내야 하고, 지수의 안정성을 확보하고 있어야 한다. 실거래가격을 이용할 경우, 매기 표본주택이 실제 거래된다는 보장이 없으므로 표본 방식에 의한 라스파이레스 지수(Laspeyres Index)는 작성하는 것이 쉽지 않다. 실거래가격에 기초한 주택가격지수를 작성하기 위해서는 반복 매매가격지수모형 또는 특성가격지수모형을 사용하는 것이 바람직하다.

본 논문에서는 특성가격지수모형의 하나로서 비모수적 방법인 국지가중회귀모형을 이용한 실거래가격지수를 작성하는 방법을 제시하였다. 이 방법은 특정유형의 주택에 대해 개별 실거래 자료들의 특성들의 차이를 감안하여 가중치를 달리 부여함으로써 다양한 유형의 주택가격지수를 작성할 수 있다는 장점과 새로운 거래가 발생하더라도 가중치가 그리 크지 않은 경우, 기존 지수는 안정적인 범위내에서 변화하게 된다.

강남구를 예로 주택유형을 4가지로 구분한 다음 주택가격지수를 작성한 결과 주택유형에 따라 주택가격지수의 변화는 매우 상이하게 나타나는 것을 알 수 있다. 또한 현재 부동산시장의 가격지수에 대한 정보로 가장 많이 활용되고 있는 국민은행 가격지수와 비교하여 볼 때 국민은행 지수는 지수의 변동률이 매우 낮은 반면 실거래 가격지수는 변동률이 매우 크게 나타나고 있는

것을 알 수 있다. 이용만·이상한(2008)에서 주장하는 바와 같이 국민은행 지수는 호가를 기준으로 가격지수를 작성함에 있어 호가는 시장상황을 정확히 반영하지 못하고 시장가격을 평가하는 사람들의 현 상태를 유지하려는 관성때문에 지수가 시장에서의 부동산가격의 변화를 정확하게, 신속하게 반영되지 못하는 반면 실거래가격지수는 시장의 변화를 빠르게 반영하는 것을 알 수 있다.

본 논문에서 제시한 실거래가격지수 작성방법에는 여전히 해결해야 할 문제들이 남아 있다. 첫 번째로 주택특성함수의 모형설정에 대한 문제이다. 특성가격함수모형에서 사용하는 변수에 따라서 지수의 움직임은 다르게 나타난다. 따라서 어떤 변수들을 사용할 지에 대한 기준과 원칙이 세워져야 할 것이다.

두 번째의 문제는 실거래자료의 빈도가 낮은 경우 특성가격모형에 의한 지수작성은 매우 불안정한 모습을 보이고 있다. 향후 이에 대한 추가적인 연구가 필요하다.

세 번째의 문제는 특성가격모형을 추정할 때 사용한 회귀식에서 오차항은 모두 독립적이고 분포가 일정한 정규분포를 한다고 가정하였는데 많은 경우 분포가 일정하지 않는 이분산을 가지는 분포를 하는 경우가 많으며, 또한 실거래 자료는 부동산을 구입하고자 하는 사람이 제시하는 가격이 부동산을 판매하고자 하는 사람의 희망가격보다 높은 경우에만 거래가 이루어지기 때문에 모든 자료가 독립적인 자료라고 보기 어렵다. 이에 대한 해결방안도 필요하다.

논문접수일 : 2010년 11월 23일

논문심사일 : 2010년 12월 7일

게재확정일 : 2011년 3월 30일

## 참고문헌

1. 국토해양부 실거래가 조회 사이트  
<http://rt.mltm.go.kr>
2. 박헌수, “준모수방법에 의한 주택가격지수 추정에 관한 연구”, 「부동산학연구」 제7권 제1호, 한국부동산분석학회, 2001, pp. 1-16
3. 박헌수, “특성가격모형을 활용한 아파트 실거래가격지수 산정방법에 관한 연구”, 「부동산학연구」 제15권 제3호, 한국부동산분석학회, 2009, pp. 111-125
4. 부동산114, <http://www.r114.co.kr>
5. 이용만, “특성가격함수를 이용한 주택가격지수 개발에 관한 연구-시간변동계수모형에 의한 연쇄지수”, 「부동산학연구」 제13집 제1호, 한국부동산분석학회, 2007, pp. 103-125
6. 이용만 · 이상한, “국민은행 주택가격지수의 평활화 현상에 관한 연구”, 「주택연구」 제16권 제4호, 한국주택학회, 2008, pp. 27-47
7. 이창무 · 김병욱 · 이현, “반복매매모형을 활용한 아파트 매매가격지수”, 「부동산학연구」 제8집 제2호, 한국부동산분석학회, 2002, pp. 1-19
8. 이창무 · 김진유 · 이상영, “공동주택 실거래가 지수 산정에 관한 연구”, 「국토계획」 제40권 제4호, 대한국토도시계획학회, 2005, pp. 121-134
9. 한국감정원, 「부동산실거래가격에 기초한 주택가격지수 개발」, 2007
10. Bailey, M.J., R.F. Muth and H.O. Nourse, “A Regression Method for Real Estate Price Index Construction,” *Journal of the American Statistical Association*, Vol. 58, 1963, pp. 933-942
11. Case, K.E. and R.J. Shiller, “The Efficiency of the Market for Single Family Homes,” *American Economic Review*, 1989, pp. 125-37
12. Case, B., H.O. Pollakowski and S.M. Wachter, “On Choosing Among House Price Index Methodologies,” *Journal of the American Real Estate and Urban Economics Association*, Vol. 19, 1991, pp. 286-307
13. Clapp, J.M. and C. Giaccotto, “Estimating Price Indexes for Residential Property : A Comparison of Repeat Sales and Assessed Value Methods,” *Journal of the American Statistical Association*, Vol. 87, 1992, pp. 300-306
14. Cleveland, W.S. and S.J. Devlin, “Locally Weighted Regression : An Approach to Regression Analysis by Local Fitting,” *Journal of the American Statistical Association* Vol. 83, 1988, pp. 596-610
15. Meese, R., and N. Wallace, “Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Indices Price,” *AREUEA Journal*, Vol. 19, 1991, pp 308-332
16. Rosen. S., “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition,” *Journal of Political Economy*, Vol. 82, 1974, pp. 34-55