

특성가격함수를 이용한 주택가격지수 개발에 관한 연구
- 시간변동계수모형에 의한 연쇄지수*

Estimation of Hedonic Price Models and Construction of New Housing Price Indexes
- Using Time Varying Parameter Model and Chain Index

이용만(Young Man Lee)**

<Abstract>

This paper is addressed a method for constructing house price indexes using transaction prices. Laspeyres index, which is made of appraisal prices of sample houses, can't be adapted as a method for transaction-based indexes. The method for transaction-based indexes must have abilities to control the difference of house attributes and the change of their intrinsic values over time. And indexes must not be revised when new transaction information is arrived. This paper adapts the time varying parameter model, one of hedonic price indexes. The model has advantages to get over the problems. It consider the regional effect on hedonic prices and the spatial auto-correlation in error term. The model is applied from Dec. 1999 to Jun 2006. The data is obtained from an information company so called 'r114'. And this paper produces a repeat sale price index to compare with the hedonic price index. The hedonic price index by the time varying parameter model appears to be more stable than the Laspeyres index which 'r114' has made. And the change of the hedonic price index is similar to that of the repeat sale price index.

주 제 어 : 특성가격, 주택가격지수, 시간변동계수모형, 연쇄지수, 실거래가격

Keyword : hedonic price, housing price index, time varying parameter model, chain index, transaction price

* 본 연구는 LG연합재단의 2005년도 교수해외파견제도에 의한 연구비 지원을 받아 수행되었다. 연구비 지원을 해준 LG연합재단에 감사드린다. 그리고 본 논문에 대해 유익한 논평을 해주신 익명의 세 심사위원께 감사드린다. 또한 자료를 제공해 준 부동산114와 본 논문의 초기 연구에 참여해 준 명지전문대학의 서후석 교수 및 R2Korea의 변재현 전 이사에게 감사드리며, 본 주제에 대해 토론에 응해 준 Univ. of Wisconsin, Milwaukee의 김선웅 교수께도 감사드린다. 본 논문에서 발견되는 오류는 전적으로 저자의 책임이다.

** 한성대학교 부동산학과 교수, ymlee@hansung.ac.kr

I. 머리말

우리나라는 현재 국민은행과 일부 민간 부동산정보업체에서 주택가격지수를 발표하고 있다. 이들 기관들은 표본주택을 선정한 후, 중개사들로부터 표본주택의 가격정보(대개 매각자가 희망하는 가격)를 월별로 받아 라스파이레스(Laspeyres) 방법으로 지수를 산출하고 있다.

해당 주택가격지수는 주택정책을 결정할 때 주요한 판단근거로 사용되고 있다. 예를 들어 투기과열지구나 주택거래신고지역을 선정할 때, 또는 선정된 지역을 해제할 때 해당 주택가격지수가 사용되고 있다.

그러나 이들 주택가격지수가 주택시장의 가격동향을 정확하게 보여주지 못하고 있다는 비판이 많이 제기되어 왔었다. 이들 주택가격지수에 대한 가장 큰 비판은 시세정보에 기초하여 지수가 작성되고 있다는 점이다. 시세정보가 현실을 왜곡하고 있기 때문에 이에 기초한 지수 역시 현실을 왜곡할 수 있다는 것이다. 특히 중개사들이 가격정보를 왜곡함으로써 시장 질서를 어지럽히고 있다는 비판이 많이 제기되었다.

이러한 비판의 사실 여부를 떠나¹⁾ 정부는 정확한 시장정보를 확보하기 위하여 2005년 7월에 「공인중개사의 업무 및 부동산 거래신고에 관한 법률」을 제정하였으며, 주택의 실거래가격에 대한 정보를 확보할 수 있는 법적 근거를 갖추었다. 동법에 따라 2006년 1월부터 모든 주택거래자는 실거래가격을 정부에 신고하도록 되어 있다.

하지만, 실거래가격이 수집되었다고 하여 주택시장 상황을 정확하게 보여주는 주택가격지수를 작

성할 수 있는 것은 아니다. 실거래가격을 기초로 하여 지수를 작성할 경우, 기존의 표본에 의한 라스파이레스(Laspeyres)식 지수작성방법을 사용할 수 없다. 표본에 의한 라스파이레스(Laspeyres)식 지수를 작성하기 위해서는 매 시점마다 표본주택의 가격정보가 필요하나, 표본주택이 매 시점 거래되는 것은 아니기 때문이다. 따라서 실거래가격을 기초로 하여 주택가격지수를 작성하고자 할 경우, 기존의 표본에 의한 라스파이레스(Laspeyres)식 지수작성방법과는 다른 새로운 지수작성방법을 고안해야 한다.

미국에서는 호가정보가 없기 때문에 실거래가격에 기초한 주택가격지수 작성 방법에 대한 논의가 오래전부터 이루어져 왔으며, 아직도 부동산학계나 주택경제학계, 도시 및 지역계획학계의 주요 이슈가 되고 있다. 논의의 초점은 상이한 주택의 질을 계량적으로 통제하여 표준적인 주택가격을 구하는데 있다. 매기 거래되는 주택의 질이 다양하기 때문에 주택의 질 차이에 따른 주택가격의 차이를 계량적으로 통제하지 않으면 현실을 왜곡할 수밖에 없다. 주택의 질을 계량적으로 통제하기 위해 가장 많이 사용하는 방법으로는 특성가격지수모형(hedonic price index model)이 있으며, 이 밖에 반복매매가격지수모형(repeat sale price index model)과 혼합가격지수모형(hybrid price index model) 등이 있다.

최근 미국에서는 주택가격지수에 연동되는 파생 금융상품이나 보험상품의 개발 가능성이 거론되면서, 주택가격지수의 안정성(stability)도 주요 이슈가 되고 있다²⁾. 새로운 실거래가격 정보가 추가될 때마다 주택가격지수가 변동한다면, 이에 연동되

1) 이러한 비판이 사실인지 여부는 실거래가격에 기초한 지수가 작성된 이후에나 가능해질 것이다. 일부 미공개 자료에 따르면, 실제 거래가격과 시세정보 간에 큰 차이가 없다는 보고도 있다.

는 파생금융상품이나 보험상품을 생각할 수가 없게 된다.

우리나라에서 실거래가격에 기초하여 주택가격지수를 작성하려고 할 경우 미국에서의 경험이 그대로 적용될 수밖에 없다. 우선 실거래가격을 기초로 하여 주택가격지수를 작성하기 위해서는 주택의 질 차이에 따른 가격 차이를 통제할 수 있는 지수작성방법을 사용하여야 한다. 그리고 해당 주택가격지수는 지수로서의 안정성을 확보해야 한다. 주택가격지수를 이용하여 정부정책에 대한 의사결정이 이루어지고 있기 때문에 주택가격지수가 수시로 바뀐다면 정부정책은 신뢰를 얻을 수 없을 것이다.

본 논문은 실거래가격을 이용한 주택가격지수 작성방법을 개발하는데 있다. 실거래가격에 기초한 주택가격지수의 작성은 올바른 주택정책의 수립을 위해, 그리고 금융기관과 가계 및 일반 투자자들에게 올바른 시장정보를 제공해 주기 위해 필요한 연구이다. 현실을 왜곡하지 않는 가격정보를 제공해야만 올바른 주택정책의 수립이 가능하며, 민간의 경제주체들도 올바른 가격정보를 갖고 있어야 합리적인 의사결정을 할 수 있다. 더 나아가 실거래가격에 기초한 주택가격지수는 새로운 금융상품의 개발을 위해서도 필요한 작업이다.

다행히 그동안 우리나라에서는 주택의 질 차이에 따른 가격 차이를 통제하는 방법에 대한 연구는 꾸준히 이루어져 왔었다. 그래서 본 논문에서는 주택가격지수의 안정성을 확보할 수 있는 지수작성방법에 대해 주로 논의하고, 이러한 논의에 근거하여 새로운 주택가격지수를 작성해 보고자 한다.

본 논문은 모두 4개의 장으로 구성되어 있다. 제

2장에서는 주택가격지수 작성에 관한 기존 연구를 검토할 것이다. 여기서는 특성가격지수모형(hedonic price index model)과 반복매매가격지수모형(repeat sales price index model)을 주로 살펴보고, 각 지수작성방법의 장단점과 이슈들에 대해 정리할 것이다. 제3장에서는 특성가격지수모형을 이용하여 주택가격지수를 작성할 것이다. 주택가격지수의 안정성을 위해 특성가격지수모형 중 시간변동계수모형을 이용할 것이며, 체인 지수작성방법(chain index method)으로 지수를 작성할 것이다. 마지막 제4장에서는 결론과 추후 과제에 대하여 설명할 것이다.

II. 주택가격지수 작성방법에 대한 기존 연구 검토

1. 주택가격지수 작성 방법에 대한 검토

주택가격지수(housing price index)는 주택가격의 변화를 지수화한 것이다. 주택가격의 변화는 주택이 갖고 있는 특성의 변화에 의해 발생할 수도 있으며, 특성 변화가 없음에도 불구하고 시장상황의 변화에 의해 발생할 수도 있다. 예를 들어 주택의 내부시설을 최신 시설로 교체함으로써 주택가격이 상승할 수도 있으며, 주택 근처로 지하철역이 생겨 교통이 편리해짐에 따라 주택가격이 상승할 수도 있다. 이러한 가격변화는 주택특성의 변화에 따른 것인데, 전자를 주택의 내부특성 변화에 따른 가격변화라고 한다면, 후자는 주택 외부의 환경적인 특성이 바뀐데 따른 가격변화라고 할 수 있다. 반면, 이런 특성 변화가 없음에도 불구하고 주

2) 여기서 주택가격지수의 안정성(stability)이란 새로운 실거래가격 정보에 의해 과거의 주택가격지수가 변동되지 않는 것을 의미한다.

택수급상의 문제나 이자율의 변동 등과 같은 시장 상황의 변화에 의해 주택가격이 변할 수도 있다.

문제는 무엇을 주택가격의 변화로 인식할 것이냐 하는 점이다. 이론적으로는 주택의 특성 변화에 따른 가격의 변화는 지수산정 시 가격변화에서 제외하는 것이 타당하다. 주택의 특성이 바뀌었다면, 이미 해당 주택은 품질 면에서 이전의 주택과는 다른 주택이므로 지수 산정 시 신규 주택처럼 취급하는 것이 타당하다.³⁾

따라서 주택가격지수를 작성하기 위해서는 명시적이든, 명시적이지 않던 간에 주택의 특성변화에 따른 가격변화를 고려해야만 한다. 주택가격지수 작성 방법은 주택의 특성변화에 따른 가격변화를 명시적으로 고려하였느냐, 아니면 고려하지 않았느냐에 따라 두 가지 범주로 구별할 수 있다.

먼저 주택의 특성변화에 따른 가격변화를 명시적으로 고려하지 않는 지수작성방법으로는 매매가격 중위수 방법(median sales price model)과 표본에 의한 라스파이레스(Laspeyres) 지수작성방법이 있다. 그리고 주택의 특성변화에 따른 가격변화를 명시적으로 고려하는 지수작성방법으로는 특성가격지수모형(hedonic price index model)과 반복매매가격지수모형(repeat sales price index model)이 있다. 이 네 가지 방법 중 매매가격 중위수 방법과 특성가격지수모형, 그리고 반복매매가격지수모형은 실거래가격에 기초하여 지수를 작성할 수 있는 방법이다.

1) 매매가격 중위수 방법

매매가격 중위수 방법은 실제 거래가 이루어진

주택매매가격을 각 시점별로 수집한 후, 각 시점별 매매가격의 중위수(median)를 연결하여 지수를 산정하는 방법이다.

$$M_t = \text{median}(V_{1t}, V_{2t}, \dots, V_{nt}) \\ t = 1, 2, \dots, T$$

위 식에서 V_{jt} 는 t 라는 시점에 거래된 j 번째 주택의 실거래가격을 의미한다. 결국 위 식이 의미하는 것은 t 라는 시점에 거래된 n 개의 실거래가격 중에서 중위수를 t 기의 대표치로 삼아서, 지수를 만드는 것이다.

미국의 NAR(National Association of Realtor : 전국부동산협회)은 단순하고 편리하다는 이유에서 이 방법으로 지수를 발표하고 있다⁴⁾.

2) 표본에 의한 라스파이레스(Laspeyres) 지수작성 방법

표본에 의한 라스파이레스 지수작성 방법은 표본 주택을 선정한 후, 비교시점의 가중평균 주택가격을 기준시점의 가중평균 주택가격으로 나누어 지수를 작성하는 방법이다. 구체적인 식은 다음과 같은데, 이 방식을 사용하면 지역별로나 주택종류별로 지수작성이 가능하며, 이를 모두 통합한 종합주택가격지수도 계산할 수 있다.

$$L_t = \frac{\sum q_1 V_t}{\sum q_1 V_1} \times 100$$

여기서 q_1 는 기준시점의 가중치, V_1 는 기준시점

3) 이런 문제는 주택가격지수를 작성할 때만 발생하는 것은 아니다. 물가지수를 작성할 때에도 동일한 문제에 직면하게 된다. 예를 들어 물가지수 산정 대상으로 20인치 TV가 포함되어 있을 때, 제조업체에서 TV의 품질을 올리면서 가격을 인상할 경우 품질 향상에 따른 가격 상승부분은 물가지수 산정 시 제외하는 것이 타당할 것이다.

4) NAR의 홈페이지(www.realtor.org)에서 발표하고 있다.

의 주택가격, V_t 는 비교시점의 주택가격을 의미한다.

표본에 의한 라스파이레스 지수작성 방법을 사용하기 위해서는 표본주택의 매기 가격정보가 필요하다. 표본주택이 매기 거래될 수가 없기 때문에 이 방법으로 주택가격지수를 작성하기 위해서는 호가나 감정가격을 조사할 수밖에 없다. 표본에 의한 라스파이레스 지수작성방법은 명시적으로 주택의 특성변화에 따른 가격변화를 고려하고 있지는 않지만, 표본주택을 선정하였기 때문에 암묵적으로는 주택의 특성에 변화가 없다는 전제가 깔려 있다.

우리나라에서는 현재 국민은행이나 일부 부동산 정보제공업체들이 이 방법으로 주택매매가격지수와 전세가격지수를 작성하고 있다.

는데, 특성가격함수의 형태를 이중로그함수로 가정할 경우 함수형태는 다음과 같다.

$$\ln V = \sum_{i=1}^k \beta_i \ln X_i + \sum_{t=1}^T r_t B_t + \varepsilon$$

여기서 V 는 주택가격이며, X_i 는 i 번째 주택특성변수, B_t 는 추세항으로서 t 시점에 대한 가변수(dummy variable)이다. 주택가격지수는 추세항의 계수인 r_t 을 통해 추정된다.

시간변동계수모형은 각 시점별로 다음과 같은 특성가격함수를 추정한 후(이중로그함수로 가정), 특성변수의 값을 일정한 값으로 대입하여 각 시점별로 주택의 잠재가격을 추정하여 이를 지수화하는 방법이다.

3) 특성가격지수모형(Hedonic Price Index Model)

특성가격지수 모형(Hedonic Price Index Model)은 주택의 특성가격(Hedonic Price)을 추정한 후, 이를 이용하여 매 시점별로 가상적인 표준주택의 가격을 추정하여 지수화하는 방법이다. 주택의 특성가격⁵⁾에 대한 이론적 틀은 Rosen(1974)에 의해 마련되었다. Rosen(1974)에 의하면 주택은 다양한 특성의 집합체이며, 주택이 입주자에게 제공하는 효용은 주택의 특성들이 입주자에게 제공하는 효용의 합과 같다는 것이다.

특성가격지수 모형을 이용한 지수작성 방법은 지수도출 방법에 따라 다시 추세항 방식과 시간변동계수모형(time varying parameter model) 등으로 분류할 수 있다. 추세항 방식은 주택의 특성가격 함수에 시간 추세항이 포함되어 있는 모형을 말하

$$\ln V_1 = \beta_{11} \ln X_{11} + \beta_{12} \ln X_{12} + \dots + \beta_{1k} \ln X_{1k} + \varepsilon_1$$

$$\ln V_2 = \beta_{21} \ln X_{21} + \beta_{22} \ln X_{22} + \dots + \beta_{2k} \ln X_{2k} + \varepsilon_2$$

...

$$\ln V_T = \beta_{T1} \ln X_{T1} + \beta_{T2} \ln X_{T2} + \dots + \beta_{Tk} \ln X_{Tk} + \varepsilon_T$$

이때 특성변수의 값을 무엇으로 하느냐에 따라 시간변동계수모형은 다시 기준년도 가중치 방법과 비교년도 가중치 방법, 연쇄지수방법(chain index) 등으로 나눌 수 있다. 예를 들어, 첫 번째 시점($t=1$)의 특성변수 평균값을 사용할 경우 기준년도 가중치 방법이 되고, 마지막 시점($t=T$)의 특성변수 평균값을 사용할 경우 비교년도 가중치 방법이 된다.

현재 미국에서는 상무부(Department of Commerce)의 통계국(Census Bureau)이 기준년도 가중치 방법으로 신규 단독주택 가격지수(Price Index

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

for New One-Family Houses Sold)를 작성하고 있다⁶⁾.

4) 반복매매가격지수 모형(Repeat Sale Price Index Model)

반복매매가격지수 모형은 지수작성 기간 중 두 번 이상 거래가 이루어진 주택의 매매가격을 이용하여 지수를 작성하는 방법이다. 반복매매가격 모형을 함수식으로 표현하면 다음과 같다.

$$\ln V_s - \ln V_f = \left(\sum_{i=1}^k \beta_{is} \ln X_i + r_s B_s + \varepsilon_s \right) - \left(\sum_{i=1}^k \beta_{if} \ln X_i + r_f B_f + \varepsilon_f \right)$$

여기서 V_s 는 두 번째 매매시점의 매매가격을, V_f 는 첫 번째 매매시점의 매매가격을 의미한다. 따라서 위 식의 좌변은 두 시점간의 주택가격상승률을 나타낸다. X_i 는 i 번째 주택특성변수이며, B_s 와 B_f 는 각각 첫 번째 매매시점과 두 번째 매매시점을 나타내는 가변수(dummy variables)이다. β 와 r_s 는 추정해야 할 계수이다. 위의 식에서 우변은 두 번째 매매시점의 특성가격추정치에서 첫 번째 매매시점의 특성가격 추정치를 뺀 것이다.

매매가 이루어진 두 기간 중에 주택의 특성이 변하지 않았고 특성의 잠재가격이 고정되어 있다고 가정할 경우, 위의 식은 다음과 같이 변하게 된다.

$$\ln \left(\frac{V_s}{V_f} \right) = r_s - r_f + \varepsilon$$

위 식을 이용하여 두 기간 중의 증가율을 계산하고⁷⁾, 이를 통해 지수를 작성하는 방법을 반복매매 가격지수 모형이라고 부른다.

반복매매가격지수 모형은 거래가 이루어진 두 기간 중에 해당 주택의 특성에 아무런 변화가 없다는 가정을 하고 있기 때문에 명목적으로는 주택의 특성변화에 따른 주택가격 변화를 고려하고 있다.

현재 미국에서는 OFHEO(Office of Federal Housing Enterprise Oversight)가 반복매매가격지수 모형으로 HPI (House Price Indexes)라는 주택가격지수를 작성하고 있으며⁸⁾, Freddie Mac에서도 CMHPI(Conventional Mortgage Home Price Index)라는 주택가격지수를 반복매매 가격지수 모형으로 작성하고 있다⁹⁾.

2. 주택가격지수 작성과 관련된 이슈들

각 지수작성방법에는 몇 가지 이슈들이 존재한다. 매매가격 중위수 방법의 경우, 지수작성의 편리성 때문에 NAR에서 이 방법을 사용하고 있기는 하지만 통계적 엄밀성을 갖고 있지 않다. 예를 들어 이 방법은 주택의 특성 차이에 따른 가격 차이를 전혀 감안하고 있지 않을 뿐만 아니라 지수산정에 사용된 자료가 편의(biasedness) 될 수 있다¹⁰⁾. 이런 통계적 엄밀성 문제는 매매가격 중위수 방법이 갖고 있는 본질적이면서도 극복하기 어려운 문제이기 때문에, 이 방법의 문제점들을 극복

6) 미 상무부 통계국의 홈페이지(www.census.gov) 참조
 7) 변수에 로그를 취한 후, 시간에 걸쳐 차분을 하게 되면 해당 기간 중의 증가율에 근사하게 된다.
 8) Calhoun(1996) 및 OFHEO 홈페이지(www.ofheo.gov) 참조
 9) Freddie Mac의 홈페이지(www.freddiemac.com) 참조
 10) 예를 들어 지난 달에는 고가주택만 거래되었고, 이번 달에는 저가주택만 거래되었다면 주택가격지수 상으로는 주택 가격이 하락하고 있는 것처럼 나타나게 된다.

하기 위한 논의는 거의 이루어지고 있지 않다.

표본에 의한 라스파이레스 지수작성방법에는 두 가지의 이슈가 존재한다. 첫째는 무엇을 가중치로 삼을 것이냐 하는 점이며, 둘째는 표본 변경기간을 어느 정도로 할 것이냐 하는 점이다. 먼저 가중치의 경우, 주택재고 비율을 기준으로 삼는 방법과 주택가치 비율을 기준으로 삼는 방법이 있다. 주택재고 비율을 가중치로 삼는 방법은 모든 개별주택에 동일한 가중치를 부여하는 반면, 주택가치 비율을 가중치로 삼는 방법은 고가주택에 더 큰 가중치를 부여한다¹¹⁾. 정책적인 관점에서 보자면, 주택재고 비율을 가중치로 삼는 것이 바람직하겠지만, 투자 성과를 측정하는 입장에서 보면 주택가치 비율을 가중치로 삼는 것이 바람직할 수 있다. 표본 변경기간의 경우, 그 기간이 짧을수록 좋지만, 그 대신 지수작성 비용이 많이 들어간다.

특성가격지수모형의 경우, 특성가격모형(hedonic price model)의 설정오류(mis-specification) 문제와 특성가격의 지수 전환 방법을 둘러싸고 여러 이슈가 존재한다.

이 중에서 특성가격모형의 설정오류 문제는 특성가격으로 지수를 작성할 때 부닥치는 가장 큰 어려움 중의 하나이다. 특성가격모형의 설정오류 문제는 다시 세 가지 문제로 분류된다. 첫째는 함수형태(functional form) 문제이고, 둘째는 변수누락(omitted variables) 문제이다. 그리고 셋째는 오차항의 이분산(hetero-scedasticity)이나 공간상의 자기상관(spatial autocorrelation) 문제이다.

먼저 함수형태 문제는 주택가격과 주택특성가격간의 관계를 나타내는 함수식을 어떻게 정의내릴 것인가의 문제이다. 일반적으로 선형모형이나 반

로그모형, 또는 이중 로그모형이 많이 사용되나, 사전적으로 함수형태를 알 수 없기 때문에 Box-Cox함수를 이용하기도 한다. 또 변수들이 정규분포를 한다는 가정 하에 계수들(coefficients)을 추정하는 것이 일반적인데, 변수들의 확률분포에 대한 사전적인 가정 없이 계수들을 추정하는 비모수(non parameter) 또는 반모수(semi parameter) 추정방법도 최근 많이 사용되고 있다.

Box-Cox함수를 이용한 모형의 함수형태 결정 방법은 Halvorsen and Pollakowski (1981), Cropper, Deck, McConnell(1988) 등에 의해 시도되었고, 이후 많은 연구에서 이 방법을 사용하고 있다. 국내에서는 허세림·곽승준(1994)이 이 방법으로 특성가격함수를 추정한 바 있다. 비모수 추정방법은 Meese and Wallace(1995)가 국지가중회귀모형(locally weighted regression)으로 특성가격함수를 추정한 바 있고, 반모수추정방법은 Stocker(1992) 등이 시도한 바 있다. 국내에서는 허세림·곽승준(1997)이 비모수추정방법으로, 그리고 박현수(2001)가 반모수추정방법으로 주택가격지수를 추정한 바 있다.

변수누락 문제는 특성가격지수모형이 근본적으로 해결하기 어려운 문제 중의 하나이다. 주택가격에 영향을 미치는 특성들을 모두 조사하는 것은 상당한 비용이 들어간다. 또 설령 많은 비용을 지불하여 모든 특성변수들의 값을 조사하더라도 또 다른 누락 변수가 있는지 알기가 어렵다. 변수누락 문제는 추정계수의 불편성(unbiasedness)을 보장하지 않는데, 이 문제는 종종 오차항의 이분산이나 공간상의 자기상관 문제로 나타나기도 한다.

오차항의 이분산이나 자기상관 문제는 추정계수

11) 국민은행의 주택가격지수는 주택재고 비율을 가중치로 삼고 있으며, 부동산114의 주택가격지수는 총주택가치 비율을 가중치로 삼고 있다. 서후석(1999)도 이 방법으로 주택가격지수를 작성한 바 있다.

의 효율성을 보장하지 않는다. 오차항의 이분산이나 자기상관 문제는 잘못된 함수형태나 변수누락에 의해 발생하기도 하고, 지역별로 특성가격이 다르기 때문에 발생하기도 한다. 또 주택가격이 주변 지역의 주택가격에 상호 영향을 주고받기 때문에 발생하기도 한다.

이 중에서 지역별 특성가격의 차이 문제는 위계적 선형모형(hierarchical linear model : HLM)에 의해 해결책이 강구되고 있다. 위계적 선형모형은 Can(1992), Can and Megbolugbe(1997), Orford(2000) 등에 의해 사용되었으며, 국내에서는 김주영·김주후(2002) 등에 의해 시도되고 있다. 주택가격이 주변지역의 주택가격에 의해 영향을 받는 문제는 공간상의 자기회귀(spatial auto regression)로 해결하고자 하는 시도가 있다. Pace, Barry, Clapp and Rodriguez(1998) 등이 공간상의 자기회귀 모형으로 이 문제를 해결하고자 하였으며, 국내에서는 박헌수(2003, 2004) 등이 공간상의 자기회귀모형을 도입한 바 있다¹²⁾.

특성가격지수모형은 특성가격함수의 추정 문제뿐만 아니라 특성가격을 이용한 지수작성 방법에서도 몇 가지 이슈가 존재한다. 먼저 거론될 수 있는 것은 특성가격이 시간에 따라 변화하는 문제이다. 주택이 갖고 있는 각 특성별 잠재가격은 고정되어 있는 것이 아니라 시간에 따라 변동한다고 보는 것이 타당하다. 그러나 앞에서 언급한 시간 추세항 방식은 특성가격이 변하지 않는다고 가정하고 있기 때문에 추정된 지수에 오류가 발생하게 된다. 시간변동모수모형(time varying parameter model)은 시간에 따라 특성가격이 달라질 수 있다는 점을 반영하기 위해 도입된 방법이다. 이 방법은 이미 앞에서 언급하였다시피 각 시점별로 주택 특

성가격 함수를 추정한 후, 특성변수의 값을 일정한 값으로 대입하여 지수를 작성하는 방법이다. 이에 대한 연구로 Knight, Dombrow and Sirmans(1995)가 있다.

시간변동모수모형은 특성변수의 값을 무엇으로 하느냐에 따라 기준년도 가중치 방법과 비교년도 가중치 방법이 있고, 기준년도 가중치와 비교년도 가중치의 기하평균인 피셔지수(Fisher Index)가 있다. 그리고 매기 기준년도를 바꾸면서 지수를 작성하는 연쇄지수(chain index) 방법도 있다.

반복매매가격 모형은 Bailey, Muth and Nourse(1963)에 의해 제안되었고, 이후 Case and Shiller(1987)에 의해 정형화되었다. 국내에서는 이창무 외(2002)와 이창무 외(2005)가 부동산114의 가격 자료를 이용하여 반복매매가격 모형으로 지수를 작성한 바 있다. 반복매매가격 모형은 변수누락 문제로부터 자유롭다는 장점을 갖고 있지만, 두 거래시점 간에 이루어진 주택의 특성변화 문제라든가 특성가격의 변화 문제, 그리고 실거래가격 정보를 충분히 사용하지 못하는 문제, 소규모 지역단위에서는 자료의 부족으로 지수작성이 어렵다는 문제 등을 안고 있다.

두 거래시점 간에 이루어진 주택의 특성변화 문제는 특성변화가 있는 실거래가격 자료를 분석대상에서 제외하는 방법으로 해결하고자 하는 시도가 있었다. Case, Pollakowski and Wachter(1991)이나 Clapp, Giaccotto and Tirtiroglu(1991)의 연구에 이에 해당되는데, 이 경우 실거래가격 정보를 충분히 이용하지 못하는 문제가 더욱 커지는 단점이 있다.

실거래가격 정보를 충분히 이용하지 못하는 문제는 반복 거래된 주택의 특성과 관련이 있다. Clapp

12) 이들 연구들은 공간상의 자기회귀뿐만 아니라 시간상의 자기회귀까지 모형에 포함하고 있다.

and Giaccotto(1992)에 따르면, 반복 거래된 주택은 질이 나쁘거나 소형주택인 경우가 많기 때문에 반복 거래된 주택가격에 기초한 반복매매가격지수는 편의(biasedness)가 클 수밖에 없다. 혼합가격지수모형(hybrid price index model)이나 감정가격 대비 매매가격 비율 방법(sale price appraisal ratio method : SPAR)은 이런 문제를 해결하기 위해 도입되었다.

Quigley(1995)에 의해 도입된 혼합가격지수모형은 반복 거래된 주택의 경우 반복매매가격지수모형을 적용하고, 한번만 거래된 주택의 경우 특성가격지수모형을 사용하는 방법이다. SPAR방법은 뉴질랜드에서 주택가격지수를 작성하기 위해 사용하고 있는 방법인데, 두 기간 중의 감정가격 대비 매매가격 비율을 이용하여 지수를 작성하는 방법이다. Bourassa, Hoesli and Sun(2006)에 따르면, SPAR방법은 특성가격함수에 의한 시간변동계수모형과 비슷한 결과를 보여주고 있어 반복매매가격지수모형이 가지고 있는 표본오차에 의한 편의(biasedness) 문제를 완화할 수 있다.

특성가격지수모형이나 반복매매가격지수모형은 모두 지수의 안정성이라는 이슈에서 자유롭지가 못하다. 특히 반복매매가격지수모형은 새로운 반복매매 사례가 추가될수록 과거의 지수가 달라지는 문제를 안고 있다. Clapham, Englund, Quigley and Redfearn(2006)는 반복매매 사례가 추가될수록 과거의 지수값이 하락하는 현상을 발견하였다. 특성가격지수모형의 경우에도 매기 특성가격함수를 별도로 추정하는 시간변동계수모형을 제외하고는 이런 문제에서 자유롭지가 못하다.

III. 지수작성모형의 설정과 지수개발

1. 지수작성 방법 및 자료

본 연구의 목적은 실거래가격에 기초한 주택가격지수를 작성하는데 있다. 실거래가격에 기초한 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 작성할 수가 없다. 특성 차이에 따른 가격 차이를 반영하지 못하는 중위수 방식도 사용할 수가 없다. 따라서 실거래가격에 기초한 주택가격지수를 작성하기 위해서는 반복매매가격지수모형 또는 특성가격지수모형을 사용할 수밖에 없다.

반복매매가격지수모형은 지수작성이 용이하다는 장점이 있지만, 주택특성의 변화나 특성가격의 변화를 반영하지 못하고 표본에 의한 편의(umbiasedness) 문제가 있다. 무엇보다도 반복매매가격지수모형은 지수의 안정성이 보장되지 않는 문제점을 안고 있다. 반면, 특성가격지수모형 중 시간변동계수모형은 특성가격함수 추정에 따른 모형설정오류 문제가 있지만 지수의 안정성이 보장된다.

이런 이유에서 본 논문에서는 시간변동계수모형, 그 중에서도 연쇄지수방식(chain index method)을 이용하여 주택가격지수를 작성하고자 한다. 연쇄지수방식의 경우, 매기마다 기준년도를 바꾸어나가기 때문에 가중치가 현실과 괴리되는 문제를 최소화할 수 있다¹³⁾. 다만, 여기서는 비교를 위해

13) 연쇄지수방식이 이런 장점이 있음에도 불구하고 각종 물가지수 작성에서 이 방식을 사용하지 않는 이유는 매기 표본을 새로 설정해야 하는데 따른 비용과 시간이 많이 들기 때문이다. 그러나 실거래가격에 기초한 주택가격지수의 경우, 매기 표본(거래가 일어난 주택)이 바뀌기 때문에 연쇄지수방식을 사용해도 별도의 비용이나 시간이 들어가지 않는

보조적으로 반복매매가격지수도 작성하도록 한다.

실거래가격을 이용하여 주택가격지수를 작성하는 것이 본 논문의 목적이지만, 실거래가격에 대한 정보가 공개되고 있지 않기 때문에 본 논문에서는 부동산114가 조사하는 아파트가격 관련 DB를 이용하였다. 본 연구는 실험적으로 새로운 주택가격지수를 작성하는 것이 목적이기 때문에 강남구만을 대상으로 삼았다.

부동산114가 갖고 있는 강남구의 아파트가격 관련 DB에는 주택특성에 관한 정보와 단지특성에 관한 정보, 그리고 지역특성에 관한 정보가 포함되어 있다. 주택특성에 관한 정보에는 분양평수와 전용평수, 건축년도와 시공사 정보가 있다. 단지특성에 관한 정보에는 세대수, 단지 내 동수, 단지 내 가장 높은 건물 층수, 단지 내 소형평형(20평 이하) 비율, 지역난방 여부, 가구당 주차가능 대수, 저밀도 단지 여부, 재건축대상 여부, 주상복합건물 여부 등에 대한 정보가 있다. 그리고 지역특성으로는 인근 버스정류장 버스 노선수, 도보 10분내 지하철역 존재 여부, 편의시설 수, 관공서시설 수에 대한 정보가 있다.

부동산 114가 갖고 있는 DB에는 교육환경에 대한 정보가 없다. 강남구의 경우 교육환경에 관한 변수가 주택가격을 결정하는데 중요한 역할을 할 것으로 보이기 때문에 여기서는 반경 1km 이내 우수 고등학교 수와 반경 500m 이내 학원밀집지역

수를 조사하여 변수에 포함하였다. 그리고 지역특성을 나타내는 변수에 각 아파트단지가 속해 있는 행정동의 중심지에서 서울 도심의 중심지까지의 거리, 그리고 강남구의 업무중심지까지의 거리를 추가하였다¹⁴⁾.

확보된 자료의 기간은 1999년 12월부터 2003년 6월까지이다¹⁵⁾. 2003년 6월을 기준으로 볼 때, 아파트가격 자료는 모두 586개이다¹⁶⁾.

각 특성별 변수들의 이름과 내용은 다음과 같다. 부동산114의 아파트가격DB에는 시공사 이름은 있으나 해당 시공사의 시공능력에 대한 정보는 없다. 이에 따라 건설협회에서 발표하는 2002년 말 기준의 시공능력 자료를 이용하여 시공사의 능력에 관한 자료를 만들었다.

다.

- 14) 행정동의 중심지는 해당 동사무소로 삼았다. 서울 도심의 중심지는 서울시청으로 삼았으며, 강남구 업무중심지는 테헤란로 포스코 사거리로 삼았다.
- 15) 최근 자료를 확보할 수 없어서 2003년 6월까지의 자료만을 사용하였다. 본 논문의 목적이 최근의 주택가격지수를 확인하고자 하는데 있는 것이 아니라 실거래가격을 이용한 주택가격지수 작성방법을 개발하는데 있기 때문에 최근 자료가 아니라고 하여 논문의 목적을 달성하지 못하는 것은 아니다.
- 16) 익명의 심사위원은 실거래가격 정보를 이용할 경우 매기 이용할 수 있는 자료의 수가 다르기 때문에 본 논문에서 사용한 방법을 그대로 사용하기 어려울 수 있다는 점을 지적하였다. 예를 들어 거래빈도가 매우 적을 때에는 추정방법을 달리해야 한다는 것이다. 이 심사위원의 지적은 정당한 지적이다.

<표 1> 주택 특성별 변수이름 및 내용

특성 구분	변수 이름	단위	변수 내용	
주택특성	주택가격	pr	만원/평	상한가와 하한가의 평균을 분양평수로 나눈 값
	평수관련 특성	py	평	분양평수
		realpy	평	전용평수
		dpy	평	분양평수와 전용평수 차이
	건물특성	dur	년	건축경과연수(2004년 기준)
		rdum1	dummy	시공사 도급순위 1-5위 = 1, 기타 = 0
rdum2		dummy	시공사 도급순위 6-20위 = 1, 기타 = 0	
단지특성	단지내부 특성	tdsu	세대	단지내 총세대수
		dsu	개	단지내 동수
		floor	층	단지내 가장 높은 건물의 층수
		small	%	단지내 소형평형(20평 이하) 구성비
		hdum	dummy	지역난방 = 1, 기타 = 0
	재건축관련 특성	park	대수	가구당 주차가능 대수
		low	dummy	저밀도 단지 = 1, 기타 = 0
	rest	dummy	재건축 대상 단지 = 1, 기타 = 0	
	건물종류 특성	multi	dummy	주상복합건물 단지 = 1, 기타 = 0
	지역특성	생활환경 특성	bus	개
stat			dummy	도보 10분내 지하철역 존재 = 1, 기타 = 0
envir			개	편의시설수
common			개	관공서 시설수
교육환경 특성		scdum	개	반경 1km 이내 우수학교 수
		stdum	개	반경 500m 이내 학원밀집지역 수
도심접근 특성		d_city	km	개별 행정동 중심지에서 서울시청까지 거리
		d_center	km	개별 행정동 중심지에서 강남업무중심지까지 거리

<표 2> 시공사 도급순위(2002년 말 기준)

도급순위 1위 ~ 5위	현대건설, 대우건설, 삼성물산, 엘지건설, 대림산업
도급순위 6위 ~ 20위	현대산업개발, 포스코개발, 롯데건설, 두산중공업, 한진중공업, 삼성중공업, SK건설, 태영, 쌍용건설, 벽산건설, 동부건설, 코오롱건설, 부영, 금호산업, 삼환기업

2. 주택특성가격 모형의 추정

1) 모형의 함수형태 및 변수 선정

주택 특성가격 모형은 주택특성변수들을 설명변수로 하고 주택가격을 피설명변수로 한 회귀모형을 의미한다. 설명변수들과 피설명변수간의 관계를 나타내는 함수 형태에 따라 주택 특성가격 모형은 선형모

형(linear model), 반로그모형(semi-log model), 이중로그모형(double log model), 트랜스로그 모형(translog model), 박스-콕스모형(Box-Cox model) 등으로 나누어질 수 있다.

이 중 어떤 모형을 사용할 것인가에 대한 일반적인 이론은 없으나, 여기서는 추정이나 해석상에서 몇 가지 이점이 있는 반로그모형(semi-log model)을 사용한다. 반로그모형(semi-log model)은 주택의 특성과 주택가격간의 관계를 비교적 잘 보여주고, 주택특성의 한 단위 변화에 따른 주택가격의 퍼센티지 변화를 계산하기 용이하다¹⁷⁾. 트랜스로그 모형(translog model)이나 박스-콕스모형(Box-Cox model)이 설명변수와 피설명변수 간의 관계를 좀 더 유연하게 보여주지만, 시간변동계수 모형으로 지수를 작성할 경우 추정된 모형 형태가 매기 바뀔 수가 있어 여기서는 사용하지 않았다.

반 로그모형은 다음의 식과 같이 설명변수가 피설명변수에 지수적(exponential)으로 영향을 미치는 모형으로, 피설명변수에 자연로그를 취한 비선형(non-linear)모형이다. 여기서 X_i ($i=1, \dots, k$)는 i 번째 주택특성변수를 의미하며, Y_j 는 j 번째 주택가격을 의미한다. α 와 β_i ($i=1, \dots, k$)는 각각 상수항과 i 번째 주택특성변수의 계수(coefficient)를 의미한다.

$$\log(Y_j) = \alpha + \sum_{i=1}^k \beta_i X_{ij} + \varepsilon_j$$

또는

$$Y_j = \exp(\alpha + \sum_{i=1}^k \beta_i X_{ij} + \varepsilon_j)$$

피설명변수는 평당 가격(pr)이며(단위는 만원),

반로그모형(semi-log model)에 따라 자연로그를 취한 값을 사용하였다.

설명변수로서 평수 관련 특성은 분양평수(py)와 실평수(realpy) 사이에 다중공선성이 존재하는 것으로 판단되어 py와 dpy(분양평수와 실평수 사이의 차이)를 변수로 사용하였다. 그리고 분양평수와 실평수 사이의 차이는 재건축대상 아파트에서 더욱 크게 주택가격에 영향을 미칠 것으로 판단되어 dpy*rest 변수를 추가하였다. 재건축대상 아파트 가변수(dummy variable)인 rest는 재건축대상 아파트에 대해 1을 부여하고, 그렇지 않은 아파트에 대해 0을 부여하고 있다.

분양평수(py)의 추정계수 부호를 사전적으로 예상하기는 어렵지만, 대형평형에 대한 선호와 마감재의 차이 등을 감안할 때 분양평수가 클수록 평당 가격이 높을 것으로 예상된다. 반면에 분양평수와 실평수의 차이가 클수록 평당 가격은 하락할 것으로 예상된다. 특히 재건축대상 아파트의 경우 분양평수와 실평수 차이가 클수록 부담금이 커지기 때문에 주택가격에는 마이너스 영향을 미칠 것으로 예상된다.

건물특성 관련 변수로 건물의 노후도(dur)는 건축년도를 2004로 뺀 값을 사용하였다. 건물의 노후도가 주택가격에 미치는 영향은 재건축대상 아파트이냐 아니냐에 따라 다를 수 있으며, 영향의 정도도 비선형일 가능성이 있다. 이에 따라 건물의 노후도는 $dur1=dur*rest$ 이라는 새로운 변수를 만들어 사용하였다. dur1은 재건축대상이 아닌 아파트에 대해서는 dur값을 부여하고, 재건축대상 아파트에 대해서는 0을 부여한 변수이다¹⁸⁾.

서울의 아파트들은 노후도가 증가함에 따라 가격이 하락하나, 일정 시일이 경과하고 나면 재건축

17) Malpezzi(2003) pp.79-81. Malpezzi(2003)에 따르면 반로그모형(semi-log model)은 5가지의 장점을 가지고 있다.

가능성이 높아지면서 다시 가격이 오르는 현상을 보인다. 이런 현상을 반영하기 위해 설명변수로 $dur1$ 외에 $(dur1)^2$ 을 추가하였다¹⁹⁾. 두 변수의 추정계수 부호를 사전적으로 알 수는 없지만, $dur1$ 은 마이너스 부호를, 그리고 $(dur1)^2$ 의 부호는 플러스 부호를 보일 것으로 기대된다. 이런 부호를 보인다면, 시간이 지날수록 노후화에 의해 주택가격이 하락하다가 일정 시점이 지나고 나면 재건축에 대한 기대 때문에 노후화가 진행되더라도 주택가격이 상승하게 된다.

시공사의 브랜드를 나타내는 가변수(dummy variable)인 $rdum1$ 와 $rdum2$ 는 플러스의 부호를 보일 것으로 예상된다.

단지의 특성을 나타내는 변수들 중에서 단지규모를 나타내는 세대수($tdsu$)와 동수(dsu)는 상호 중복적인 변수인 것으로 보여 유의성이 높고 추정계수가 안정적인 동수(dsu)를 사용하였다. 규모의 경제효과 때문에 단지규모가 클수록 주택가격에 플러스 영향을 미칠 것으로 기대된다. 단지 규모가 클수록 재건축의 이점이 클 것으로 보여 $dsu*rest$ 를 변수에 포함시켜 보았는데, 추정 결과 추정계수의 부호가 플러스를 보였으나 통계적으로 유의하지 않아 피설명 변수에서 제외하였다.

단지내 건물의 최고층수를 나타내는 변수인 $floor$ 도 단지특성을 나타내는 변수의 하나로 사용하였으나, 추정계수가 통계적으로 유의하지 않아

변수에서 제외하였다. 그러나 부호는 플러스를 보였다. $floor$ 는 해당 단지의 밀도를 보여 준다고 볼 수 있다. 즉, 고층일수록 밀도가 높고 저층일수록 밀도가 낮다고 볼 수 있다. 재건축 대상 아파트 단지의 경우 저밀도 일수록 재건축에 따른 이익이 크기 때문에 재건축대상 아파트에서는 $floor$ 가 주택가격에 유의미한 영향을 미칠 수 있다. 그래서 $floor*rest$ 라는 새로운 변수를 만들어 모형에 포함시켰다. 층수가 높을수록(고밀도일수록) 재건축 대상 아파트 가격이 하락할 것이므로 예상되는 부호는 마이너스이다.

단지 내 소형평형(20평 이하)의 비율을 나타내는 $small$ 의 경우, 사전적으로 부호의 방향을 알 수 없지만, 소형평형에 거주하는 가구들이 상대적으로 저소득층일 가능성이 높기 때문에 소형평형 비율이 높을수록 주택가격은 하락한다고 예상해 볼 수 있다.

지역난방 가변수(dummy variable)인 $hdum$ 은 부호가 일정하지 않아 변수에서 제외하였으며, 가구당 주차대수 변수인 $park$ 는 자료가 없는 부분이 많아 변수에서 제외하였다.

재건축관련 특성 중 저밀도 가변수(dummy variable)인 low 의 경우 추정계수의 부호가 플러스로 나타났으나 통계적으로 유의적이지 않아 변수에서 제외하였다²⁰⁾. 재건축 대상 아파트 단지에서 저밀도 효과를 보기 위해 $low*rest$ 라는 새로운 변수를

18) $dur1$ 에서 재건축대상 아파트란 재건축 가변수인 $rest$ 의 값이 1인 아파트 또는 건축연도가 25년 이상인 아파트를 의미하는 것으로 하였다. 재건축 가변수인 $rest$ 의 값이 0인 아파트 중에서도 건축연도가 25년 이상인 아파트들이 있는데, 이런 아파트들은 비록 재건축조합 등이 설립되지 않았더라도 재건축에 대한 기대가 형성되어 있을 것으로 판단된다.

19) 박현수(2004)는 건축 후 5년이 경과하기 이전에는 건물의 노후에 의한 가격하락 현상이 없고, 5년 이후부터 노후화에 따라 가격이 하락하는 것으로 보고 있다. 이러한 주장의 사실 여부를 확인하기 위해 3차 방정식을 고려하였으나 $(dur1)^3$ 의 추정계수가 유의적이지 않아 제외하였다.

20) 저밀도 단지는 거주자의 쾌적성이 높기 때문에 재건축대상 아파트가 아니라 하더라도 가격에 긍정적인 영향을 미치는 것으로 나타났으나 통계적으로 유의적이지는 않았다.

모형에 포함시켰는데, 부호는 플러스이지만 이 또한 통계적 유의성은 없는 것으로 나타났다. 이에 따라 low*rest 역시 모형에서 제외하였다.

한편 재건축 가변수(dummy variable)인 rest의 경우, 재건축 대상 아파트를 1로 하고, 나머지는 0으로 한 변수이다. 여기서 재건축대상 아파트란 부동산114에서 자체적으로 조사한 재건축 추진 아파트 외에 건축연령이 25년 이상인 아파트를 의미한다²¹⁾. 재건축 가변수의 추정계수는 플러스 부호를 보일 것으로 기대된다. 즉, 여타 조건이 동일하다면 재건축 대상 아파트는 그렇지 않은 아파트보다 주택가격이 높을 것으로 예상된다.

건물종류 특성인 주상복합 아파트 가변수(dummy variable) multi는 추정계수의 부호를 사전적으로 알 수 없다. 최근에 주상복합 아파트 붐이 일면서 주상복합 아파트의 평당 가격이 일반 아파트의 평당 가격보다 높을 것으로 예상해 볼 수는 있다.

지역변수들 중 생활환경의 특성을 나타내는 bus와 common은 추정계수의 부호가 일정하지 않고 유의적이지 않아 변수에서 제외하였으며, stat(지하철 역세권 여부를 나타내는 가변수)와 envir(편의시설수)는 추정계수가 유의적이어서 특성변수에 포함시켰다. stat와 envir의 경우, 추정계수 부호가 플러스를 보일 것으로 기대된다.

교육환경을 나타내는 변수 중에 우수학교 존재 여부 나타내는 scdum은 추정계수가 플러스의 부호를 보일 것으로 기대된다. 해당 아파트 단지 근처에 학원가가 어느 정도 존재하는가를 보여주는 변수인 stdum의 경우 추정계수의 부호가 플러스를 보였으나 통계적으로 유의미하지 않아 변수에서 제외하였다.

도심이나 업무중심지까지의 거리를 나타내는 변수인 d_city와 d_center의 경우, 추정계수가 마이너스를 보일 것으로 기대된다. 즉, 도심이나 업무중심지까지의 거리는 주택가격에 마이너스의 영향을 준다고 볼 수 있다. 일반적으로 도심이나 업무중심지까지의 거리는 다른 변수들에 비해 지수적으로 영향을 미치는 경우가 많으므로 여기서는 d_city^2 과 d_center^2 를 변수로 사용하였다.

그러나 이렇게 추정된 계수들은 지역에 따라 값이 달라질 수 있다. 위계적 선형모형(HLM)은 '주택의 호별 특성가격이 지역에 따라 달라질 수 있다'는 점에 착안하여, 주택의 호별 특성을 나타내는 변수의 추정계수가 지역변수에 의해 제약을 받도록 되어 있는, 일종의 모수제약 모형이다. 주택의 호별 특성이 지역에 따라 달라짐에도 불구하고 이를 모형에 고려하지 않을 경우 공간상의 자기상관(spatial auto correlation) 문제가 발생할 수 있다.

본 논문에서는 이런 문제를 해결하기 위해 위계적 선형모형(HLM)과 유사하게 주택변수에 지역변수를 결합하거나 또는 단지변수에 지역변수를 결합한 새로운 변수들을 모형에 추가하였다.

먼저 평형 변수의 추정계수가 지역특성에 따라 달라지는가를 보기 위해 py*stat, py*envir, py*scdum, py*d_city, py*d_center를 모형에 추가하였는데, py*stat를 제외하고는 모두 추정계수가 유의적이지 않았다. 이에 따라 py*stat만 모형에 추가하였다. 추정계수의 부호는 사전적으로 알 수 없다. 지하철 역세권이라면 평형이 클수록 주택가격에 플러스의 영향을 미칠 것으로 예상되지만, 대형 평형에 거주하는 가구들은 대부분 자가용으

21) 건축연령이 25년 이상이 된 아파트 중에는 재건축이 공식적으로 추진되지 않는 아파트도 있다. 이런 아파트는 부동산114의 자료에 재건축 대상 아파트가 아닌 것으로 분류되어 있는데, 건축연령이 25년 이상이 된 아파트들은 비록 재건축추진위원회 등이 설립되지 않더라도 재건축에 대한 기대가 형성되어 있을 것으로 판단된다.

로 출퇴근할 것이기 때문에 부호가 반대로 나올 수도 있다.

dpy1과 dur1, rdum1, rdum2, dsu, floor와 같은 주택의 호별 특성이나 단지특성을 나타내는 변수들과 지역변수들을 결합하여 모형에 추가하여 보았으나, 이들 결합변수들의 추정계수들은 통계적으로 유의하지 않았다. 이에 따라 이들 결합변수들은 모두 모형에서 제외하였다.

단지변수 중 small의 경우, small*stat, small*envir, small*scdum의 추정계수들이 모두 통계적으로 유의한 것으로 나타났다. 그리고 multi와 지역변수의 결합변수인 multi*stat, multi*envir, multi*scdum의 추정계수들 역시 통계적으로 유의한 것으로 타나, 모형에 추가하였다. 그러나 단지변수 중 rest는 지역에 따라 추정계수의 값이 달라지지 않는 것으로 나타나 rest와 지역변수들을 결합한 변수들은 모형에서 제외하였다.

일반적으로 공간 자료로 회귀분석을 할 경우, 공간상의 자기상관(spatial auto correlation) 문제가 나타난다. 그 원인이 어디에 있는지 명확하지는 않지만, 자기상관 문제가 존재함에도 불구하고 이를 무시하고 최소자승법(OLS)으로 추정할

경우 추정계수의 효율성(eficiency)이 보장되지 않는다. 본 논문에서는 아파트

의 호별 특성이나 단지 특성이 지역 특성에 의해 영향을 받는 것으로 하였음에도 불구하고 여전히 자기상관 문제가 존재하였다²²⁾.

자기상관 문제를 해결하기 위해 박헌수(2004)는

22) 오차항에 자기상관이 없다는 가정 하에 모형을 추정한 후, 잔차항에 자기상관이 존재하는가를 LM검정으로 검정해 본 결과, LM검정통계량이 134.59(확률 0.0000)로서 '잔차항에 자기상관이 존재하지 않는다'는 귀무가설이 1%의 유의 수준 하에서 기각되었다.

23) 동일 행정동 내에서의 자료 배열은 아파트 단지의 인접성을 고려하지 않고, 임의로 배열하였다.

24) 공간상의 자기상관은 시간상의 자기상관과는 달리 여러 방향으로부터 영향을 받는다(Anselin 1988, p. 9). 본 논문에서 사용한 자기회귀모형은 자료 배열에서 앞에 있는 주택가격이 뒤에 있는 주택가격에 영향을 주는 것으로 되어 있기 때문에 공간상의 자기상관 문제를 정확하게 반영하고 있다고 볼 수는 없다.

GIS자료를 이용하여 주변 지역에 있는 주택과의 거리를 자승한 후, 이의 역수를 가중치로 한 주변 주택가격의 가중평균치를 설명변수에 포함시킨 공간 자기회귀모형을 사용한 바 있다. 본 논문에서는 공간상의 자기상관 문제를 간단하게 해결하기 위한 방법의 하나로 오차항에 자기회귀(auto regression) 방정식을 포함시켰다. 그리고 공간상의 자기상관 현상이 명확하게 나타나도록 자료들을 행정동을 기준으로 그룹화시켰으며, 행정동 단위로 그룹화된 자료들은 다시 행정동 간의 인접성에 따라 재배열하였다²³⁾. 구체적으로 오차항에 자기회귀식이 포함된 특성가격모형은 다음과 같다.

$$\log(Y_j) = \alpha + \sum_{i=1}^k \beta_i X_{ij} + \varepsilon_j$$

$$\varepsilon_j = \gamma \varepsilon_{j-1} + u_j$$

위의 두 식을 통합하면 다음과 같은 단일 식으로 변형할 수 있는데, 이 식에서 보다시피 위의 식은 피설명변수인 주택가격이 인접한 주택의 가격에 의해 영향을 받는 구조로 되어 있다²⁴⁾.

2) 특성가격함수 추정 결과

위의 특성가격함수식에는 오차항의 자기회귀식이 포함되어 있기 때문에 고전적인 최소자승법

(OLS)으로 추정할 수가 없다. 여기서는 통계프로그램인 EViews를 이용하여 추정하였는데, EViews에서는 최우추정법(MLE)의 수치해석 방법(numerical method)인 마퀴드 알고리즘(Marquardt algorithm)을 이용하여 추정하고 있다.

본 논문의 목적은 실거래가를 이용한 주택가격지수 작성 방법을 검토해 보는 것이기 때문에 매월의 특성가격함수를 추정하지 않고, 1999년 12월, 2000년 12월, 2001년 12월, 2002년 12월, 2003년 6월의 특성가격함수만 추정하여 지수를 작성하였다. 구체적인 추정결과는 2003년 6월 및 2002년 12월 기준으로만 제시하였다.

2003년 6월 기준으로 특성가격함수를 추정한 결과, 모형의 설명력(R^2)는 81.3%이었으며, 수정된 R^2 는 80.5%으로 비교적 양호한 설명력을 보여주었다. 예측오차의 정도를 나타내는 평균평방근편센티지오차(RMSPE : root mean squared percentage error)는 1.7%로 예측오차가 크지 않은 것으로 나타났다.

Durbin-Watson 통계량은 1.98로 자기상관이 존재하지 않는 것으로 보였다. LM검정으로도 자기상관 여부를 검정하였는데, 자기상관이 없다는 귀무가설을 5%의 유의수준 하에서 기각할 수 없었다.

각 변수들 추정 결과를 살펴보면, 분양면적(py)의 경우 추정계수가 플러스를 보이고 있어 주택규모가 클수록 평당 가격이 높은 것으로 나타났다. 반면 분양면적과 전용면적간 차이(dpy)의 추정계수는 마이너스의 부호를 붙이고 있어 차이가 클수록 평당 가격이 하락하는 것으로 나타났다. 이런 효과가 재건축 대상 아파트에서 더욱 크게 나타난다.

건물연령이 주택가격에 미치는 영향은 비선형인 것으로 나타났다. $DUR1^2$ 과 $DUR1$ 의 추정계수로 계산해 보면, 건물연령이 11.32년이 될 때까지는 건물연령의 증가에 따른 주택가격의 하락폭이 커지다가, 11.32년 이후에는 점차 그 하락폭이 줄어들고, 22.65년 이후에는 오히려 건물연령의 증가에 따라 주택가격이 상승하는 것으로 나타났다²⁵⁾.

단지규모를 나타내는 동수(dsu)는 예상대로 플러스의 부호를 보였으며, 단지 내 소형평형 비율(small)은 마이너스의 부호를 보였다. 즉, 다른 조건이 동일하다는 가정 하에서 단지 내에 소형평형 비율이 높을수록 평당 주택가격은 하락하는 것이다. 또한 재건축 가변수(rest)는 예상대로 플러스의 부호를 보였으며, 추정계수값도 매우 크게 나타났다. 재건축 대상 아파트에서 밀도의 영향을 나타내는 floor*rest는 예상대로 마이너스의 부호를 보였다.

건물 형태를 나타내는 주상복합 가변수(multi)는 예상외로 마이너스의 부호를 보였다. 다른 조건이 동일하다면, 주상복합아파트가 일반 아파트보다 평당 가격이 낮다는 것이다.

지역변수인 지하철 가변수(stat), 생활환경 가변수(envir), 학교 변수(scddum)는 모두 예상대로 플러스의 부호를 보였다. 그러나 도심까지의 거리를 나타내는 d_city² 변수는 예상과는 달리 플러스의 부호를 보인 반면, 강남 업무중심지까지의 거리를 나타내는 d-center² 변수는 이론대로 마이너스의 부호를 보였다.

지역에 따른 호별 특성가격 또는 단지별 특성가격의 변화는 대부분 예상한 것과 같은 부호를 보였다. 즉, 지하철 역세권, 생활편의시설이 많은 지역,

25) $D=0.001573x^2-0.035624x=0$ 을 만족하는 x 값은 0과 22.647이다. 그리고 $dD/dx=0$ 을 만족하는 x 값은 11.32이다.

우수 고등학교 주변 지역에서는 그렇지 않은 지역보다 특성가격이 높게 나타나고 있는 것이다. 그러나 $py*stat$ 나 $multi*envir$ 에서는 예상과는 달리 부호가 마이너스로 나타났다.

오차항의 자기회귀 계수인 AR의 추정계수는 0.5689로서 모형의 수렴조건을 갖추었고, 추정계수의 유의성도 양호하였다.

2002년 12월 자료를 기준으로 할 때, 모형의 추정계수와 부호는 2003년 6월 자료로 추정된 모형과 큰 차이가 없었다. 다만, 추정계수의 유의성이 없는 변수들이 다소 발견되었다. 예를 들어 재건축 대상 아파트 가변수($rest$)와 지하철 역세권 변수($stat$), 그리고 지역별 추정계수의 변화를 보여주는 $py*stat$, $floor*rest$, $small*envir$ 등의 변수 계수가 통계적으로 유의하지 않은 것으로 추정되었다. 모형의 설명력(R^2)는 76.05%이었으며, 수정된 R^2 는 74.90%이었다. Durbin-Watson 통계량은 2.029이며, LM검정으로 자기상관 여부를 검정한 결과 자기상관이 없다는 귀무가설을 5%의 유의수준 하에서 기각할 수 없었다.

각 변수의 구체적인 추정계수와 t 통계량은 다음의 <표 3>과 같다.

<표 3> 주택가격의 특성가격함수 추정

변수	2003년 6월		2002년 12월	
	계수추정치	t 통계량	계수추정치	t 통계량
상수	7.00624	90.4206*	7.05443	79.7447*
py	0.00745	5.0727*	0.00505	3.0551*
dpy	-0.01127	-3.6689*	-0.01428	-4.2022*
dpy*rest	-0.01425	-2.8538*	-0.01152	-2.0918**
(dur1)^2	0.00157	5.1444*	0.00147	4.6411*
dur1	-0.03562	-4.4966*	-0.03495	-4.1058*
rdum1	0.06999	3.2276*	0.08360	3.5842*
rdum2	0.06159	3.1041*	0.07429	5.5254*
dsu	0.00381	4.8908*	0.00454	5.5254*
small	-0.54817	-5.8567*	-0.41709	-4.1387*
floor*rest	-0.00622	-2.1921**	-0.00044	-0.1341
rest	0.22183	3.5664*	0.09180	1.3210
multi	-0.51249	-3.6121*	-0.31788	-1.6557***
stat	0.15147	2.3244***	0.08413	1.1319
envir	0.02423	3.3902*	0.02415	3.1450*
scdum	0.02945	3.5608*	0.03316	3.7722*
d_city2	0.00111	4.2379*	0.00103	3.9688*
d_center2	-0.00451	-2.5909*	-0.00315	-1.8392***
py*stat	-0.00411	-2.8158*	-0.00185	-1.1099
small*stat	0.36738	4.3593*	0.48420	5.0945*
small*envir	0.04146	2.2076**	-0.01300	-0.6339
small*scdum	0.08330	4.1948*	0.03738	1.7310***
multi*stat	0.39893	4.5073*	0.32545	3.3406*
multi*envir	-0.11391	-3.1782*	-0.14759	-3.2255*
multi*scdum	0.12918	5.5511*	0.14047	5.1103*
AR	0.56892	15.7850*	0.51144	13.1650*

주 : (**, ***)는 유의수준 1%(5%, 10%) 하에서 귀무가설이 기각됨을 의미함.

시간변동계수모형은 추정계수의 값이 시간에 따라 변동한다고 보고 있다. 실제 2003년 6월과 2002년 12월 사이에 추정계수의 값이 변하였는지를 확인하기 위해서 Wald 검정을 하였다. 검정결과 두 시점의 추정계수값에 차이가 없다는 귀무가설이 1%의 유의수준 하에서 기각되었다.

<표 4> Wald 검정

귀무가설	Wald 통계량	확률
2003년 6월 기준 모형에 2002년 12월 추정계수값을 모수 제약	1.26E + 14	0.0000
2002년 12월 기준 모형에 2003년 6월 추정계수값을 모수 제약	48.0269	0.0054

3. 특성가격모형을 이용한 주택가격지수 작성

앞의 특성가격함수의 추정결과를 기초로 하여 주택가격지수를 작성하였다. 앞에서 언급하였다시피 본 연구에서는 시간변동계수모형을 사용하기로 하였고, 시간변동계수모형 중에서도 매기 기준년도를 바꾸어가는 연쇄지수방법을 사용하기로 하였다.

시간변동계수모형으로 지수를 작성하기 위해서는 매기의 특성가격함수를 추정한 후, 특성변수에 일정한 값을 부여하여 주택의 잠재가격을 추정하고, 각 기의 추정된 잠재가격을 이용하여 지수를 작성한다. 이때 특성변수의 값을 무엇으로 삼느냐에 따라 지수가 달라질 수 있다. 보통은 특정 기의 특성변수 평균값을 이용하여 지수를 작성하는데, 이때 특성변수의 값은 일종의 가중치 역할을 하게 된다. 이때 어떤 기의 특성변수 평균값을 사용하느냐에 따라 가중치가 달라지는데, 기준시점의 특성

변수 평균값을 사용하는 방법이 기준시점 가중치 방법이며,²⁶⁾ 비교시점의 특성변수 평균값을 사용하는 방법이 비교시점 가중치 방법이다²⁷⁾. 그리고 두 방법으로 작성된 지수의 증가율을 기하 평균한 피셔지수(Fisher Index)가 있다. 여기서는 세 가지 방법을 모두 사용해 보도록 한다.

그리고 상호 비교를 위하여 여기서는 반복매매가격모형으로도 지수를 작성해 보았다. 앞에서 살펴

보았다시피 반복매매가격 모형은 주택의 특성이 변하지 않는다는 가정 하에 두 기간간의 주택가격(또는 전세가격) 변화율을 이용하여 지수를 작성하는 방법이다. 즉, 반복매매가격모형에서 앞에서 보았다시피 두 기간 사이의 주택가격 변화율은 다음과 같이 표시된다.

$$\ln\left(\frac{V_s}{V_f}\right) = r_s - r_f + \epsilon$$

위의 식에서 $r_s - r_f = c$ 라고 하면, 결국 반복매매가격모형은 주택가격에 자연로그를 취한 후 차분한 값을 상수항에 대해 회귀하는 것이라고 할 수 있다. 이때 추정된 상수항이 곧 두 기간 중 주택가격의 증가율이 된다.

구체적인 지수작성방법은 먼저, 두 기간에 걸쳐 가격정보가 있는 자료들만 선택한다. 본 연구에서

26) 기준시점 가중치 방법을 흔히 라스파이레스(Laspeyres) 변동계수모형이라고 부르기도 하는데, 국민은행이나 부동산 114에서 작성하고 있는 라스파이레스 지수모형과 혼동이 될 수 있을 것 같아 여기서는 기준시점 가중치 방법이라고 부르도록 한다.

27) 파셰(Paasche) 변동계수모형이라고도 부른다.

사용하고 있는 부동산114자료는 표본을 추출하여 매 시점 가격을 조사하고 있기 때문에 1시점에 신규로 공급된 주택이나 또는 1시점에 소멸된 주택을 제외하고는 모두 추정대상에 포함된다.

그 다음, 두 기간의 가격자료에 자연로그를 취한 후 1기의 가격(자연로그를 취한 가격)에서 0기의 가격(자연로그를 취한 가격)을 빼준다. 이렇게 차분한 가격자료를 피설명변수로 하고, 상수항을 설명변수로 하는 회귀모형을 추정한다. 상수항의 추정값을 가지고 두 기간 중 증가율을 구한다. 0기에서 1기 사이의 증가율은 $g_1 = \exp(c) - 1$ 이다. 이러한 과정을 1기~2기, 2기~3기, ..., (n-1)기~n기까지 되풀이 시행하면, 매기의 증가율 g_2, g_3, \dots, g_n 을 구할 수 있다²⁸⁾.

마지막으로 0시점의 주택가격을 100으로 한 뒤, 위에서 구한 증가율을 이용하여 지수를 작성한다.

이상과 같은 방법으로 구한 주택가격의 증가율은 다음의 <표 4>와 같고, 이를 지수화한 주택가격지수들은 <표 5>와 같다. <표 5>에서 보다시피, 부동산114의 지수가 가장 빠르게 증가하는 것으로 나타나고 있으며, 시간변동계수모형 중 기준시점 지수가 그 다음으로 빠르게 증가하고 있다. 시간변동계수모형 중 피셔지수(Fisher Index)는 반복매매가격지수와 매우 유사하게 움직이고 있다.

반복매매가격지수의 경우, 반복매매된 주택의 특

성이 동일하고, 각 특성의 잠재가격이 동일하다고 하더라도 두 기간 중 건축연령은 달라지기 때문에 일반적으로 가격변화를 저평가하는 경향이 있다. 그럼에도 불구하고 반복매매가격지수가 시간변동계수모형 중 피셔지수와 유사하게 움직이는 것은 건축연령이 아파트가격에 비선형적으로 영향을 미치기 때문인 것으로 보인다. 즉, 강남구에 건축연령이 20년 내외인 아파트들이 많은 관계로 건축연령에 따른 아파트가격의 변화가 그리 심하지 않기 때문인 것으로 보인다.

논리적으로 볼 때, 부동산114의 지수와 반복매매가격지수는 거의 유사한 수준을 보여야 한다. 왜냐하면 매기 조사되는 표본이 동일하기 때문에 두 지수 사이에는 특성가격 차이에 따른 지수 차이 문제가 없다. 그럼에도 불구하고 부동산114의 지수가 가장 빠르게 상승하는 것은 가중치 때문인 것으로 보인다. 부동산114는 주택가치를 가중치로 삼기 때문에 고가 주택의 상승률이 높을 경우, 지수가 빠른 속도로 증가한다. 반면, 반복매매가격지수는 모든 주택에 대해 동일한 가중치를 부여하기 때문에 부동산114 지수처럼 고가주택 때문에 지수가 크게 좌우되지는 않는다.

28) 익명의 심사위원은 반복매매가격지수를 작성할 때, 분석기간 전체의 반복매매 자료를 pooling한 후 모수를 추정하는 것이 바람직하다는 지적을 하였다. 이 심사위원의 지적처럼 반복매매가격지수는 원래 분석기간 전체의 반복매매 자료를 pooling하여 모수를 추정한다. 그러나 본 논문에서는 전체 반복매매 자료를 pooling하지 않고, 쌍을 이루는 두 기간(예를 들어 2002년 12월과 2003년 6월, 또는 2001년 12월과 2002년 12월 등)의 반복매매 자료만을 가지고 모수를 추정하였다. 이런 방법으로 반복매매가격지수를 작성한 이유는 첫째, 추정계수의 안정성(stability)을 위한 것이다. 이미 앞에서 언급하였다시피 반복매매가격지수는 새로운 반복매매 자료가 생길 때마다 과거의 지수값이 달라지는 문제점이 있다. 이런 문제를 피하기 위해 쌍을 이루는 두 기간의 반복매매 자료만으로 지수를 추정하였다. 두 번째 이유는 부동산 114 자료의 특성 때문이다. 부동산 114 자료는 표본으로 선정된 아파트에 대해 매기 가격을 조사하기 때문에 매기 동일한 주택이 반복적으로 매매되는 것과 동일한 상황이다(물론 실거래가격에서는 이런 상황이 발생하지 않는다). 이런 상황에서는 분석기간 전체의 반복매매 자료를 pooling하여 모수를 추정하던, 아니면 본 논문과 같은 방법으로 모수를 추정하던 간에 결과는 차이가 없을 것으로 보인다.

〈표 5〉 지수작성방법별 지수 증가율

	반복매매	시간변동계수모형			부동산114
		기준시점	비교시점	기하평균	
2000년	8.67%	7.58%	7.56%	7.57%	7.36%
2001년	20.63%	21.37%	19.87%	20.62%	22.79%
2002년	37.32%	38.95%	37.73%	38.34%	40.83%
2003년6월	4.94%	5.28%	5.28%	5.28%	6.01%
평균	17.87%	18.29%	17.61%	17.95%	19.25%
표준편차	12.65%	13.42%	12.87%	13.14%	14.10%

주 : 전년말 대비 증가율임.

〈표 6〉 지수작성방법별 지수추이(1999년 12월 말 100 기준)

	반복매매	변동계수모형			부동산114
		기준시점	비교시점	두 시점 평균	
1999년	100	100	100	100	100
2000년	108.67	107.58	107.56	107.57	107.36
2001년	130.97	130.57	128.93	129.75	131.83
2002년	179.85	181.43	177.58	179.49	185.65
2003년6월	188.73	191.01	186.95	188.97	196.81

주 : 각 연도말 기준임. 단, 2003년은 6월말 기준임.

IV. 결론

‘공인중개사의 업무 및 부동산 거래신고에 관한 법률’에 의해 주택의 실거래가 신고가 의무화됨에 따라 실거래가격에 기초한 주택가격지수 작성이 가능하게 되었다. 실거래가에 기초한 주택가격지수는 시세정보에 기초한 주택가격지수보다 현실을

보다 정확하게 보여 줄 것으로 기대된다.

하지만, 실거래가격에 기초한 주택가격지수를 작성하기 위해서는 기존의 주택가격지수 작성방법과 다른 방법을 강구해야 한다. 기존의 표본추출을 통한 라스파이레스식 지수작성방법을 사용할 경우, 오히려 실거래가에 기초한 주택가격지수는 현실을 왜곡할 수가 있다.

현재 사용되고 있는 시세 기준의 라스파이레스식 지수작성방법은 매기 동일한 주택의 가격을 추적하기 때문에 주택의 특성차이에 따른 가격차이 가능성이 크지 않다. 또 시세정보의 경우 부동산 중개사들이 주택의 특성

차이에 따른 가격 차이를 걸러주는 역할을 수행해 왔기 때문에 주택의 특성차이에 따른 가격왜곡 현상이 상대적으로 작을 수 있다²⁹⁾.

그러나 매기 실제로 거래된 주택들은 특성들이 모두 다르기 때문에 가격변화가 특성변화에 기인한 것인지 아니면 시장상황에 따른 것인지를 구별해야 한다. 이를 감안하지 않고 주택가격지수를 작성할 경우 현실을 왜곡할 수 있는 것이다. 따라서

29) 부동산중개사들은 어떤 주택이 매물로 나왔을 때, 해당 주택의 내부시설 정도를 감안하여 호가정보를 제공한다. 예를 들어 어떤 지역의 30평형 아파트의 평균적인 호가가 1억원이라고 하자. 이때 내부시설을 잘 치장한 주택이 1억 2천만원에 나왔다고 하더라도 부동산중개사는 호가정보를 제공할 때 1억원으로 이야기한다.

실거래가격에 기초한 주택가격지수를 작성하기 위해서는 필연적으로 주택의 특성차이에 따른 가격 차이를 통제할 수 있는 지수작성방법이 필요하다.

이런 점에서 본 연구는 실거래가격에 기초한 주택가격지수 작성을 준비한다는데 의의를 갖고 있다.

본 연구에서는 위와 같은 인식 하에 주택특성가격모형을 이용하여, 실제 거래된 주택의 가격변화가 특성차이에 따른 가격변화인지 아니면 시장상황에 따른 가격변화인지를 구별하여 주택가격지수를 작성하였다.

주택특성가격모형을 추정하기 위해 본 논문에서는 반로그(semi-log)모형을 사용하였으며, 지역에 따라 호별 특성가격이나 단지별 특성가격이 달라질 수 있다는 점을 고려하였다. 그리고 공간상의 자기상관 문제를 해결하기 위해 오차항에 자기회귀식을 포함시켜 모형을 추정하였다.

주택가격지수는 매기 주택특성가격함수를 추정하는 시간변동계수모형을 이용하였으며, 시간변동계수모형 중에서도 기준년도를 매기 바꾸는 연쇄지수(chain index)방법을 사용하였다. 그리고 다른 모형에 의한 주택가격지수와 비교하기 위하여 반복매매가격지수도 병행하여 작성하였다.

주택가격지수를 작성해 본 결과, 표본에 의한 라스파이레스식 지수작성방법을 사용한 부동산114 지수는 본 연구에서 사용한 시간변동계수모형보다 증가율이 크고 변동성도 큰 것으로 나타났는데, 이는 주택가치를 가중치로 사용하였기 때문인 것으로 보인다. 또, 시간변동계수모형 중 피셔지수(Fisher Index)는 반복매매가격지수와 유사한 움직임을 보였다. 이는 본 논문에서 사용한 자료의 특성 때문일 수도 있으며, 건축연령이 주택가격에 미치는 영향이 비선형적이기 때문에 나타나는 현

상인 것으로 파악된다.

실거래가격에 기초한 새로운 주택가격지수는 시장상황에 대한 정확한 정보를 제공함으로써 올바른 주택정책을 이끌어낼 수 있는 기초가 된다. 더 나아가 실거래가격에 기초한 새로운 주택가격지수를 이용하여 다양한 파생적인 투자정보를 제공할 수 있다. 주택의 질적 차이에 따른 가격차이를 지수화할 수 있으며, 주택가격지수를 이용하여 투자 수익률 지표를 개발할 수 있다.

본 연구에서는 부동산114에서 조사하는 DB를 그대로 이용하였기 때문에 실제 주택가격에 영향을 미치는 주요 변수들 중 일부가 빠져 있을 것으로 보인다. 이러한 자료의 확보를 위해서는 주택의 실거래가격 정보를 신고할 때, 해당 주택의 특성에 관한 정보도 같이 신고 되도록 할 필요가 있다.

그리고 본 연구는 강남구에 한해 실험적으로 주택가격지수를 작성한 만큼 각 지역별 주택가격지수를 종합하는 작업이 필요하다. 각 지역별 주택가격지수를 종합할 때, 가중평균방법을 사용할 것인지 아니면 전체 표본을 대상으로 특성가격함수를 추정하여 지수를 작성할 것인지 여부를 결정해야 한다.

주택특성함수의 모형설정오류 문제나 지수 작성 방법상의 오류 문제를 해결하기 위한 노력은 지속적으로 이루어져야 한다. 이런 모든 작업들은 향후의 과제로 돌리도록 한다.

접 수 일 : 2007년 4월 16일

심사완료일 : 2007년 4월 28일

참고문헌

1. 김주영·김주후, "주택가격 평가를 위한 위계적 선형모델 적용", 「국토연구」 제33권, 2002, pp. 21-34
2. 박현수, "준모수방법에 의한 주택가격지수 추정에 관한 연구", 「부동산학연구」 제7권 제1호, 2001, pp. 1-16
3. 박현수, "시공간자기회귀모형을 이용한 서울 아파트가격지수 추정에 관한 연구", 「국토연구」 제42권, 2004, pp. 125-140
4. 박현수, "시공간자기회귀모형을 이용한 서울 아파트 가격의 추정", 「국토연구」, 제38권, 2003, pp. 95-106
5. 서후석, "아파트 투자지표와 부동산 포트폴리오 효과에 관한 연구", 중앙대학교 대학원 박사학위논문, 1999
6. 이창무·김병욱·이현, "반복매매모형을 활용한 아파트 매매가격지수", 「부동산학연구」 제8권 제2호, 2002, pp. 1-19
7. 이창무·김진유·이상영, "공동주택 실거래가지수 산정에 관한 연구", 「국토계획」 제40권 제4호, 2005, pp. 121-134
8. 허세림·곽승준, "한국주택시장에서의 주택가격지수 산출방법에 관한 연구", 「주택연구」 제5권 제1호, 1997, pp. 1-18
9. 허세림·곽승준, "헤도닉 가격기법을 이용한 주택 특성의 잠재가격추정", 「주택연구」 제2권 제2호, 1994, pp. 27-42
10. Anselin, Luc, *Spatial Econometrics : Methods and Models*, Kluwer Academic Publishers, 1988
11. Bailey, M. J., R. F. Muth, and H.O. Nourse, "A Regression Method for Real Estate Price Index Construction," *American Statistical Association Journal*, Vol. 58, 1963, pp. 933-942
12. Bourassa S. C., Martin Hoesli and Jian Sun, "A simple Alternative House Price Index Method," *Journal of Housing Economics*, Vol. 15, 2006, pp.80-97
13. Calhoun, Charles A., "OFHEO House Price Indexes : HPI Technical Description," Office of Federal Housing Enterprise Oversight, 1996
14. Can, A., "Specification and Estimation of Hedonic House Price Models," *Regional Science and Urban Economics*, Vol. 22, 1992, pp. 453-474
15. Can, A. and Megbolugbe I., "Spatial Dependence and House Price Index Construction," *Journal of Real Estate Finance and Economics*, Vol. 14, 1997, pp. 203-222
16. Case, B., H.O. Pollaskowski and S.M. Wachter, "On Choosing among House Price Index Methodologies," *Journal of AREUEA*, Vol. 19, 1991, pp.286-307
17. Case, Karl E. and Robert J. Schiller, "Prices of Single Family Homes since 1970 : New Indexes for Four Cities," *New England Economic Review*, 1987 September/October, 45-56
18. Clapp, John M. and Carmelo Giaccotto, "Estimating Price Indices for Residential Property: A Comparison of Repeat Sales and Assessed Value Methods," *Journal of the American Statistical Association*, Vol.87, 1992 pp. 300-306.
19. Clapp, J.M., C. Giaccotto and D. Tirtiroglu, "Housing Price Indices Based on All Transactions Compared to Repeat

- Subsamples," *Journal of AREUEA*, Vol. 19, 1991, pp. 270-285
20. Clapham, E., Peter Englund, J. M. Quigley and C. L. Redfeare, "Revisiting the Past and Settling the Score : Index Revision for House Price Derivatives," *Real Estate Economics*, Vol. 34, 2006, pp.275-302
21. Cropper, Maureen L., Leland B. Deck and Kenneth E. McConnell, "On the Choice of Functional Form Hedonic Price Functions," *Review of Economics and Statistics*, Vol. 70 No. 4, 1988, pp.668-675
22. Halvorsen R. and H. O. Pollakowski, "Choice of Functional Form for Hedonic Price Equations," *Journal of Urban Economics*, Vol. 10, 1981, pp. 37-49
23. Knight, J. R., Jonathan Dombrow, and C. F. Sirmans, "A Varying Parameters Approach to Constructing House Price Indexes," *Real Estate Economics*, Vol. 23 No. 2, 1995, pp. 187-205.
24. Malpezzi, S., "Hedonic Pricing Models : a Selective and Applied Review," in O'Sullivan, T. and Kenneth Gibb ed., *Housing Economics and Public Policy*, Blackwell Publishing, 2003
25. Meese, R. and N. Wallace. "Non-parametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Price Indices," *Journal of AREUEA*, Vol. 19, 1991, pp. 308-332
26. Orford, Scott, "Modelling Spatial Structures in Local Housing Market Dynamics : A Multi level Perspective," *Urban Studies*, Vol. 37, No. 9, 2000, pp. 1643-1671
27. Pace, R. K., R. Barry, J. M. Clapp and M. Rodriguez, "Spatio-temporal Auto regressive Models of Neighborhood Effects," *Journal of Real Estate Finance and Economics*, Vol. 17, 1998, pp.15-33
28. Quigley, J.M., "A Simple Hybrid Model for Estimating Real Estate Prices Indexes," *Journal of Housing Economics*, Vol.4, 1995, pp.1-12
29. Rosen, S., "Hedonic price and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, Vol. 82, 1974, 34-55
30. Stocker, Thomas M., "Lectures on Semi parametric Econometrics," in *Core Lectures Series*, Core Foundation, 1992
31. www.realtor.org
32. www.census.gov
33. www.freddiemac.com

