

住宅市場에서의 空間自己相關의 檢證 및 回歸係數의 推定

金 鍾 元*

논문초록 본 연구는 주택가격과 같이 자료가 공간적 특성을 포함하고 있을 경우의 불편 추정통계량을 구하기 위한 바람직한 모형설정방법을 모색하고자 하였다. 최소자승법을 이용하여 공간자기상관에 대한 회귀진단을 통하여 공간시차 또는 공간오차 모형이 적합한지를 분석하였다. 이 두 모형을 최우추정법 및 공간 2단계 최소자승추정법을 통하여 추정한 결과, 여전히 오차항의 이분산성이 존재하여 추정치의 효율성에 영향을 미치고 있었다. 이 문제를 다루기 위하여 강건 2단계 추정(Rubust 2SLS) 방법을 사용하였다. 분석결과를 자가 및 전세 주택시장 모두 공간시차모형이 적합한 것으로 나타났다. 따라서 기존의 전통적 특성가격에 기초한 주택가격지수의 산정은 편기되고 일관(consistent)되지 못한 추정 결과를 초래한다는 것을 보여주고 있다. 공간시차모형의 또 하나의 장점은 다른 모형으로서는 구할 수 없는 간접효과를 공간승수(spatial multiplier)를 통하여 포착할 수 있다는 장점을 가지고 있다.

핵심주제어: 공간자기상관, 특성가격기법, 공간계량경제학
경제학문헌목록 주제분류: R1

* 국토연구원 국토환경연구실 책임연구원

I. 序 論

전통적인 특성가격분석기법은 함수형태, 식별의 문제, 편익추정상의 문제 등에 초점을 맞추어 왔지만, 주택가격의 공간적인 특성을 고려하지는 않았다. 단지, Dubin(1988, 1992)이 오차항이 공간적으로 상관관계에 있다는 점에 주안하여 공간 오차모형을 사용한 연구와, Can(1990, 1992)이 실제 부동산시장에서는 이웃의 주택가격이 구매하고자 하는 주택가격에 하나의 변수로 작용할 수 있다는 가정하에, 공간확장 모형(spatial expansion model)에 대한 시도가 있었다.¹⁾

최근 들어 지리정보시스템(GIS)의 발달과 공간계량경제학에 대한 이론적 발전이 급속히 전개되면서 공간적으로 상관되어 있는 자료의 공간계량분석방법이 새로운 연구분야로 제시되고 있다. Anselin(1988a, 1988b, 1998, 1999a, 1999b), Anselin and Bera(1998), Anselin and Kelejian(1997), Kelejian and Prucha(1998) 등의 연구가 공간자기상관이 있을 경우에 대한 이론 및 실증적 분석을 다루고 있다.

이러한 추세에 맞추어 본 연구는 우리나라 주택시장에 공간계량경제모형을 접합시켜 보고자 하였다. 특히 우리나라의 경우 주택가격의 결정은 주택 자체의 특성(면적, 방수, 욕실수 등)과 주택의 입지특성(주변 편익시설과의 거리, 위치 등)에 의해서만 좌우되는 것이 아니라 이웃의 주택가격이 또한 중요한 역할을 하고 있다. 왜냐하면, 일반적인 주택구입형태는 구입하고자 하는 주택의 위치를 선정하고 나면 그 지역의 부동산중개소 등을 통하여 주변 시세에 대한 정보를 획득하고 최종적으로 구입가격을 결정하는 것이 대부분의 주택가격 결정형태이다. 이러한 점에 비추어 보면 경험적으로 우리나라의 주택시장은 공간적으로 밀접하게 상관관계가 있을 것으로 예상할 수 있다. 이런 관점에서 공간자기상관을 모형에 접합시킬 경우의 자가 및 세입자의 경우에 어떠한 패턴을 보이는지를 함께 검토하고자 한다.

본 연구의 구성은 다음과 같다. 제 II 절에서는 실증연구에 필요한 자료와 이 자료의 공간적인 특성을 모형에 반영시킬 수 있는 공간 가중치 행렬의 작성을 다루고 있다. 제 III 절에서는 공간계량경제모형의 기본적 방법론인 공간시차 및 공간오차

1) Dubin의 연구는 주택가격의 공간시차에 대한 검증 없이 오차모형만을 대상으로 수행하였고, Can의 경우는 공간시차만을 대상으로 공간확산방법(spatial expansion method)을 이용하여 분석하였다. 본 연구에서는 주택시장에서의 공간시차 및 오차모형을 회귀진단을 통하여 검증한 뒤 가장 바람직한 모형을 선정하여 회귀계수를 추정하고 있다.

모형에 대한 소개를 다루고 있으며, 제 IV 절은 회귀진단(regression diagnostics) 절차를 통하여 모형의 적합성을 검증함과 동시에 분석결과를 제시한다. 제 V 절은 요약 및 결론으로 구성되어 있다.

II. 資料 및 空間加重值 行列의 작성

1. 資料

본 연구의 실증분석에 이용된 자료는 국민주거상태 및 주거의식조사(1993) 중에서 서울의 자가소유가구 및 세입자 가구자료²⁾를 이용하였다. 여기에 이용된 자가 가구는 609가구이며, 전세가구는 512가구이다. 이들 가구는 서울시의 22개 區, 한 區당 3~4개 洞에서 추출하여 총 78개 洞에 걸쳐서 조사되었다. 본 연구에서 의도하는 주택간의 상호작용의 크기를 분석하기 위해서는 주택의 위치가 필요하다. 그러나 현 주택자료는 주택의 위치가 동별로만 파악되고 있어 정확한 위치를 파악할 수 없었다. 여기서는 주택의 위치에 대한 가정으로 같은 洞에 있는 가구들은 洞의 중심(center)부터 균등하게 분포되어 있는 것으로 가정하고 洞間 거리는 각 洞의 중심에서의 거리로 나타내었다. 이 가정은 모든 가구원(1, 121가구)의 개인정보를 최대한 활용하기 위해서이다. 만약 각 洞에 있는 가구원을 종합할(aggregate) 경우, 각 동에서 얻은 개인 정보의 평균값에 의한 추정량이 되어 개별 표본의 정보를 상실함과 동시에 각 洞에서 추출된 표본수가 동일하지 않음으로 인해서 심각한 편기(bias)와 이분산(heterogeneity) 문제 등이 야기된다.

주택가격의 추정을 위하여 사용된 변수로는 크게 주택특성변수, 주변여건변수, 환경변수로 나눌 수 있다. 주택특성변수로는 주택의 연면적, 방의 수, 목욕탕의 수, 건축 연도, 그 외에 더미 변수로서 난방연로 형태, 주택유형을 사용하였다. 주변여건변수로는 이웃의 소득수준, 종합병원, 중·고등학교, 지하철역, 근린공원의 접근성을 이용하였다. 그리고 환경변수로는 점오염원의 대표적인 오염물질인 아황

2) 본 자료는 국토연구원에서 실시한 설문조사를 이용하였으며, 총 1,560가구(전세자 포함)가 설문조사의 대상이었지만 부실응답자를 제외한 자가소유 609가구 및 전세가구 512가구가 본 분석의 대상이다.

〈표 1〉 변수의 정의

변수명	변수 설명	자가가구(평균)	전세가구(평균)
PRVAL	주택가격 (백만 원)	194.46	-
RENT	전세가격 (백만 원)	-	37.77
TFLSP	주택의 연면적 (평)	37.82	20.71
NMRMS	방의 수 (개)	3.81	2.35
NMBATH	목욕탕의 수 (개)	1.37	1.054
HSAGE	건축 연수	11.96	-
DFUEL	사용연료 형태에 따른 더미변수 (석유, 가스=1, 연탄=0)	0.857	0.805
DHOUS	주택형태에 따른 더미변수 (단독주택=1, 아파트=0)	0.53	0.447
DINCOM	이웃의 소득수준에 따른 더미변수 (고소득층, 고·중소득층=1, 중간 이하=0)	0.23	0.123
ACSHPT	종합병원까지의 거리 (분)	11.33	11.66
ACSSCH	학교까지의 거리 (분)	13.31	15.02
ACSSUB	지하철역까지의 거리 (분)	18.42	18.22
ACSPRK	근린공원 및 체육시설까지의 거리 (분)	21.46	22.87
MSO2T	SO ₂ 수준 (unit: ppb)	24.24	24.54
MNO2T	NO ₂ 수준 (unit: ppb)	31.66	31.40

산가스와 자동차 등에 의하여 주로 배출되는 비점오염원의 대표적 오염물질인 이산화질소를 사용하였다.

한편 개별 주택자료 이외에 각 주택에서의 대기오염 수준을 파악하기 위하여 서울시정개발연구원(1994)에서 분석한 20개 조사지점(monitoring station)에 대한 연평균 SO₂ 및 NO₂에 대한 대기오염자료를 주 자료원으로 이용하였다.

20개의 대기오염 측정소에서 관측된 자료를 이용하여 어떻게 78개洞의 주택표본추출지점의 오염도를 측정하느냐의 문제가 발생한다. 기존의 연구들을 보면 가장 가까운 측정지점의 오염수준을 그대로 적용하거나 가장 가까운 두 측정지점의 오염수준의 평균값을 사용하고 있다. 여기서는 최근에 지리정보시스템을 이용한 공간적 배분방법인 추세곡면다항식 방법(polynomial trend surface), 크리깅(kriging), 스플라인 보간법(spline interpolator) 중에서 스플라인 보간법을 이용하였다. 이 기법은 대기오염이나 토양오염과 같이 측정지점의 오염도를 이용하여 공간적으로 汚染度

를 補間하는 데 유용한 방법으로 알려져 있다.³⁾

2. 空間加重值 行列의 作成

공간가중치 행렬은 회귀분석에서 공간효과(spatial effect)를 다루는 주요한 도구의 하나이다. 공간가중치 행렬은 두 공간단위간의 인접관계를 나타낸다. 일반적으로 두 공간단위 사이의 거리, 지리적 경계, 도로에 의한 인접도, 또는 이들의 복합적 사용 등에 의하여 공간가중치 행렬을 구할 수 있다.⁴⁾ 여기서는 주택들이 78개 洞으로만 구분되어 있으므로 현재의 상태에서 가능한 방법은 이들 洞의 중심간의 거리를 바탕으로 한 가중치행렬을 구하는 것이 가장 합리적인 대안이다. 거리를 바탕으로 가중치행렬을 구하는 절차는 다음과 같다. 우선 78개 洞간의 거리를 측정하여 78×78 행렬을 만든다. 다음으로 동일한 洞 내의 주택은 洞의 중심으로부터 균등하게 분포된 것으로 가정하였으므로 표본이 속한 洞의 반경을 주택간의 거리로 환산하였다.⁵⁾ 이는 같은 洞 내의 주택간의 상관관계가 이웃 洞의 주택과의 상관관계보다 크게 작용한다는 가정에 이루어졌다. 이를 바탕으로 주택 표본수와 동일한 609×609 거리행렬을 만들었다.⁶⁾ 이를 바탕으로 주택간의 상호작용 크기는 거리와 역으로 비례하는 것으로 가정하고, 이 개념을 바탕으로 다음과 같은 9개의 空間加重值行列을 구하였다.

$$(1) w_{ij} = \frac{1}{d_{ij}^{\alpha}}, \quad \alpha = 1, 2.$$

$$(2) w_{ij} = 1 \text{ if } d_{ij} < D_c, \quad D_c = 2, 3, 4, 5, 7, 9, 10, \dots \text{ (단위: km)}$$

여기서 w_{ij} 는 행렬(W)의 i 번째 행과 j 번째 열의 요소를 의미하고, d_{ij} 는 두 공간단위 i 와 j 의 거리를 나타낸다. $\alpha=1$ 인 경우는 가중치가 두 지점간의 거리에 반

3) 자세한 설명은 Kim(1997) 또는 Chou(1997)을 참조바람.

4) 공간가중치 행렬 설정에 대한 문제는 아직 학문적으로 논쟁의 대상이 되고 있으나, 최선의 방법은 자료의 성격, 가능한 방법, 경제적 이론에 부합하는 것이 바람직하다. 자세한 논의는 Cliff and Ord(1981)나 Anselin(1988a) 참조.

5) 각 洞의 중심의 TM 좌표를 바탕으로 GIS를 이용하여 거리를 구하였고, 각 洞의 평균 반경 또한 GIS를 이용하여 구하였다.

6) 전세가구를 위한 512×512 행렬도 동일한 방법으로 구하였다.

비례하는 것을 의미하고 $\alpha=2$ 인 경우는 두 지점간의 거리의 제곱에 반비례하는 것을 의미한다.⁷⁾ D_c 는 두 공간단위간 거리의 임계치를 나타낸다. (2)식의 의미는 두 주택간의 거리가 D_c 보다 짧을 때는 두 주택간의 상호작용이 있는 것으로 간주하였고, 두 주택간의 거리가 임계치인 D_c 보다 클 경우에는 두 주택간에는 직접적인 연관성이 없는 것으로 가정하여 가중치 행렬을 구성하였다.⁸⁾ 이와 같이 생성한 공간행렬을 최종적으로 행단표준화행렬⁹⁾ (row-standardized matrix)을 만듦으로써 본 연구에서 사용하고자 하는 최종 공간가중치 행렬은 다음과 같다.

$$W = \begin{bmatrix} 0 & w_{11} & w_{12} & \cdots & w_{12} & \cdots & w_{1,J} & \cdots & w_{1,J} \\ w_{11} & 0 & w_{12} & \cdots & w_{12} & \cdots & w_{1,J} & \cdots & w_{1,J} \\ w_{21} & \cdots & w_{21} & 0 & w_{22} & \cdots & w_{2,J} & \cdots & w_{2,J} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{21} & \cdots & w_{21} & w_{22} & 0 & \cdots & w_{2,J} & \cdots & w_{2,J} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{J,1} & \cdots & w_{J,1} & w_{J,2} & \cdots & w_{J,2} & 0 & \cdots & w_{J,J} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ w_{J,1} & \cdots & w_{J,1} & w_{J,2} & \cdots & w_{J,2} & w_{J,J} & 0 & 0 \end{bmatrix}$$

그 의미는 자가주택의 경우 609×609 행렬 중에서 첫번째 블록의 차수는 첫번째 동의 자가소유 주택수와 일치하며, 행렬의 대각요소의 값은 “0”으로 주어지는데 이

7) 실제 주택가격의 결정에 있어서 두 주택가격의 거리가 멀면 멀수록 상호작용이 약해진다는 중력모델에서 쓰이는 방법을 원용하였다.

8) 여기서 두 지점간의 거리를 2km부터 시작하여 10km까지로 구분하여 각각의 행렬을 만들고 거리에 반비례하는 경우와 거리의 제곱에 반비례하는 경우를 합쳐 총 9개의 가중치 행렬을 구성한 뒤 제 3장에서 언급한 모형 속에서 가장 적합한 가중치 행렬을 추출하여 추정결과를 제시하였다. 한 가지 부연 설명하면, 1km 이내의 것을 제외한 이유는 하나의 동 이내에 있는 주택들의 실제거리가 1km 이상인 경우가 있는 반면, 동간의 거리가 1km 이내인 경우도 있어서 임계치를 1km로 할 경우에 동일한 구역 이내에 있는 주택간에는 상호작용이 없고 행정 구역이 바뀌는 다른 행정구역의 주택가격은 밀접한 관계가 있는 것으로 간주되는 모순을 없애기 위하여 최소 단위를 2km로 하였다. 실제 1km를 임계치로 하여 검정한 결과도 바람직하지 않게 추산되었다.

9) 행단표준화라 함은 모든 행의 합이 1이 되도록 함을 의미한다. 즉, $\sum_{j=1}^n w_{ij} = 1$ 을 의미한다. 이는 산업연관분석에서 기술계수의 의미와 유사하다. 보다 자세한 설명은 Anselin (1988a) 참조.

는 $i = j$ 인 경우에 $d_{ij} = 0$ 이 되기 때문이고, 동일 블록 내에서의 대각원소 이외의 값은 모두 같은 값을 취하게 된다. 이는 동일 洞 내에서의 주택은 균등한 거리에 분포되어 있는 것으로 가정한 데에 기인한다.

자료의 위와 같은 공간적 처리는 블록내 및 블록간의 상관관계라는 복합효과를 낳게 된다. 또한 각 블록내의 자료의 수가 상이한 데 따른 이분산성(heterogeneity) 문제도 야기된다. 이는 통계적으로 주택가격과 오차항이 독립적이지 못하게 되고(즉, $E[(WP)_i \cdot \varepsilon_i \neq 0]$), 결과적으로 오차항 $E[\varepsilon_i \varepsilon_j] \neq 0$ 이 된다. 이 경우 오차항의 공분산 행렬은 다음과 같은 구조를 띠게 된다.

$$E[\varepsilon\varepsilon'] = \begin{bmatrix} \sigma_{11} \otimes [i_{n1} i_{n1}'] & \sigma_{12} \otimes [i_{n1} i_{n2}'] & \dots\dots\dots & \sigma_{1n} \otimes [i_{n1} i_{nk}'] \\ \sigma_{21} \otimes [i_{n2} i_{n1}'] & \sigma_{22} \otimes [i_{n2} i_{n2}'] & \dots\dots\dots & \sigma_{2n} \otimes [i_{n2} i_{nk}'] \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} \otimes [i_{nk} i_{n1}'] & \sigma_{n2} \otimes [i_{nk} i_{n2}'] & \dots\dots\dots & \sigma_{nn} \otimes [i_{nk} i_{nk}'] \end{bmatrix}$$

여기서 n_i 는 i 번째 블록의 표본수, \otimes 는 크로네커 곱, " i "는 "1"의 열 벡터를 의미한다. 공간자기상관 및 이분산성 문제는 제 IV 절 회귀방정식의 추정시 구체적으로 논의되고 있다.

III. 理論的 模型의 設定

회귀모형에 공간효과를 통합하는 두 가지 기본적인 방법은 공간시차(spatial lag) 및 공간오차(spatial error) 모형이 있다(Anselin, 1988a). 이 두 모델 설정은 계량경제적인 측면에서는 서로 유사하지만, 경제적인 의미의 해석에는 각기 다른 의미를 지니고 있다. 공간시차모형은 주택가격이 전통적인 설명변수(주택의 특성 및 주변여건 변수)에 의하여 직접적으로 설명되는 부분 외에, 암묵적으로 이웃 주택가격의 공간 가중치 평균이 주택의 가격결정에 영향을 미치고(간접효과) 있다는 가정이다. 한편, 공간오차모형은 간접효과를 포함하지 않는 대신 회귀방정식에서 설명변수가 하나 이상 생략되어 있고 이 생략된 변수가 공간적으로 서로 다른 값을 가진다는 가

정에서 출발한다. 따라서 이 생략된 변수의 공간적 패턴에 기인하여, 오차항이 공간적으로 자기상관관계를 띠고 있다는 것이다.

1. 공간시차모형

일반적으로 공간특성가격 시차모형(spatial hedonic model)은 시계열분석(auto-regressive time series)과 유사하다. 일반적인 형태의 공간시차 주택특성가격분석모델은 다음과 같다.

$$P = \rho WP + x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + \varepsilon \quad (1)$$

$$\varepsilon \sim N(0, \Omega)$$

여기서 P 는 $(n \times 1)$ 주택가격벡터를 나타내고, ρ 는 공간자기회귀계수, W 는 $(n \times n)$ 공간가중치 행렬(spatial weight matrix), x_1 은 주택특성가격변수의 벡터, x_2 는 이웃특성가격변수의 벡터, x_3 는 환경변수의 벡터, ε 은 오차항을 나타낸다고 하자.

공간시차모형은 이웃의 溢出效果(spill-over effects)를 포착하고자 할 때 유용하게 쓰일 수 있는 도구이다. 즉, 공간시차모형은 주택 1의 가격은 이웃의 주택들의 가중합($p_1 = \rho(w_{11}p_1 + w_{12}p_2 + \dots + w_{1n}p_n)$)이 하나의 설명변수 역할을 한다.

위의 공간시차모형에서 보듯이 모형의 오른쪽에 종속변수의 존재로 인하여 오차항과 종속변수(WP)가 서로 연관(correlated)되어 있다. 즉, $E[(WP_i)\varepsilon_i] = E[(W(1-\rho\omega)^{-1}\varepsilon)\varepsilon_i] \neq 0$, 그리고 $E[ee'] = (1-\rho\omega)^{-1}\Omega(1-\rho\omega)^{-1}$, 여기서 $e = (1-\rho\omega)^{-1}\varepsilon$ 이다. 그러므로 OLS를 이용한 계수의 추정은 편기(biased)될 뿐만 아니라 효율적(efficiency)이지 못함을 의미한다. 이 경우 일치추정치(consistent estimator)를 얻기 위해서는 최우법(Anselin, 1988a)이나 대변수(IV) (Kelejian and Robinson, 1993; Kelejian and Prucha, 1998) 추정방법이 필요하다.

2. 공간오차모형

공간종속성이 오차항에 나타날 경우, 모형의 설정시 종속성을 포함하기 위한 공간오차모형은 다음과 같다.

$$P = X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \varepsilon$$

$$\varepsilon = \lambda W\varepsilon + \mu \quad (2)$$

여기서 λ 는 공간회귀계수(spatial autoregressive coefficient), W 는 공간가중치 행렬, μ 는 오차항으로서 i.i.d.이다. 이 모형은 non-spherical 오차공분산행렬을 가진 회귀방정식의 설정을 위한 특별한 경우이다. 여기서 공간회귀계수(λ)는 모형의 설명변수가 아닌 관측 불가능한 생략된 변수(unobserved variable)의 오차항에 대한 충격에 속한다. 다시 말하면 어느 특정 지점의 주택가격은 채택된 설명변수뿐만 아니라 생략된 변수의 함수라는 것이다.

공간오차모형 식(2)의 양변에 $[I - \lambda W]$ 을 곱한 뒤 재정리하면 공간오차모형은 다음과 같다.

$$P = \lambda WP + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 - \lambda WX_1\beta_1 - \lambda WX_2\beta_2 - \lambda WX_3\beta_3 + \mu \quad (2)'$$

여기에 계수에 대한 공통인자(common factor) 제약조건($H_0: \gamma = -\lambda\beta$)을 부과하면,

$$P = \lambda WP + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + WX_1\gamma_1 + WX_2\gamma_2 + WX_3\gamma_3 + \mu \quad (2)''$$

이 형태는 소위 "Spatial Dubin Model"로서 상대적으로 복잡한 공간시차모형을 보다 다루기 쉬운 공간오차모형으로 바꿀 수 있다. Anselin(1988a)은 공통인자 가설에 대한 검정방법으로 왈드(Wald) 테스트를 제안하였다. 여기서 공통인자 가설을 기각하면 공간시차모형이 적합하다는 것을 의미한다.

IV. 模型의 適合性 檢定 및 推定

1. 공간상관관계의 분석

주택가격의 공간적 상관관계를 측정하기 위하여 Moran's I와 Geary's C 통계량을 이용하였다.¹⁰⁾ Moran's I 및 Geary's C 통계량은 다음과 같다.

$$I = \left(\frac{N}{S} \right) \cdot \left(\frac{\sum_i \sum_j w_{ij} (x_i - u)(x_j - u)}{\sum (x_i - u)^2} \right) \equiv \left(\frac{N}{S} \right) \cdot \frac{e' W e}{e' e}$$

$$C = \left(\frac{N-1}{2S} \right) \cdot \left(\frac{\sum_i \sum_j w_{ij} (x_i - x_j)^2}{\sum (x_i - u)^2} \right)$$

여기서 N 은 표본수, S 는 공간가중치 행렬의 모든 요소의 합, e 는 최소자승 추정치의 잔차, w_{ij} 는 공간가중치 행렬의 i 행 j 열의 값, x 는 주택가격, u 는 주택가격의 평균을 나타낸다. 통계량 추정치를 보면 두 통계량 모두 1% 수준에서 유의함을 보여주고 있다. 여기서 Moran's I 의 값이 평균보다 크거나 Geary's C 통계량이 "1"보다 적다는 것은 양의 자기상관관계를 의미한다.¹¹⁾ 이는 주택가격이 공간적으로 서로 상관관계가 있다는 것을 보여주고 있다. 이를 바탕으로 다음 항에서는 주택가격의 공간적 상관관계를 모형 속으로 접목시키는 방법인 공간시차 및 공간오차 모형을 다루고자 한다.

〈표 2〉 자기상관계수

구 분	자가가구		전세가구	
	계 수	평 균	계 수	평 균
Moran's I	0.11***	-.0002	0.133***	-0.002
Geary's C	0.88***	1.00	0.855***	1.00

*** 통계 유의수준 1%를 의미함.

2. 공간시차모형

(1) 자가가구 거주자

우선 출발점으로서 최소자승 추정법에 의한 특성가격모형을 기초로 한 계수의 추정과 회귀진단을 하였다. 본 연구의 초점이 공간계량적 측면이므로 추정 함수형태에 대한 자세한 내용과 설명은 생략하기로 한다. 잘 알려진 바와 같이 경제이론이

10) 시계열분석에서 쓰이는 Durbin-Watson 검정과 Durbin's h 통계량과 유사하지만, 공간자료는 시계열과 달리 양방향으로 영향을 미치므로 이들 통계량을 이용할 수 없다.

11) 보다 자세한 내용은 Cliff and Ord (1973, 1981) 참조.

특성가격함수형태에 대한 제약이 없음으로 인하여 특별한 목적의 함수형태 및 적합도 기준 등이 최상의 모형을 선택하는 가이드라인이 되고 있다. 박스-콕스(Box-Cox) 변환과 같은 보다 유연한(flexible) 함수형태에 비중을 두고 있지만, 현재로서는 공간중속성(spatial dependence)과 함께 추정하는 프로그램이 없는 상태이다. 그러므로 여기서는 로그-선형모형의 추정 결과만을 고려하고자 한다.¹²⁾ 또한 앞에서 설명한 것처럼 자가 및 전세 가구를 대상으로 9개의 공간가중치 행렬을 모형에 이용하였지만 여기서는 대표적인 공간가중치 행렬을 접목시킨 결과만을 제시하기로 한다. 한 가지 흥미있는 결과는 9개의 공간가중치 행렬 중에서 자가가구는 4 km, 전세가구는 9 km 범위 이내에서 상관관계가 높고 모형의 적합도가 가장 좋은 것으로 나타났다.¹³⁾ 전세가격의 공간적 상관관계가 주택가격보다 공간적 범위가 넓게 나타난 것은 전세자가 주거입지를 결정할 때 보다 광범위하게 그리고 전세가격에 맞추어 주거지를 선택한 것으로 해석할 수 있다.¹⁴⁾

특성가격에 의하여 추정된 계수들은 대부분 기대부호와 일치하게 추정되었다. 주택가격은 주택특성변수인 면적, 방의 수, 욕실 수, 난방연료 형태와는 양의 상관관계를 나타내고 있고, 단지 주택건축 연수와는 음의 상관관계를 가지는 것으로 나타나고 있다. 이웃 특성변수인 이웃의 소득수준과는 양의 관계를, 그리고 접근성(거리와 반비례) 변수인 병원, 학교, 지하철역, 근린공원과는 기대부호와 일치하는 음의 상관관계를 보이고 있다. 한편 환경의 질을 나타내는 변수로서 이산화탄소 및 이산화질소 농도변수를 사용하였다. 경험적으로는 오염수준과 주택가격은 반비례하는 것이 정상적이다. 여기서의 추정결과를 보면, 아황산가스는 기대치와 같이 음의 관계를 보이고 또한 유의수준이 1%로서 통계적으로도 매우 유의함을 나타내는데 반하여, 이산화질소는 주택가격과 양의 상관관계를 가지지만 통계적으로 유의하

12) 다양한 함수형태에 대한 추정과 최적함수 형태에 대한 자세한 설명은 Kim(1997) 참조. 또한 Box-Cox 변환을 통한 공간계량모형의 추정은 개략적으로 grid search를 통하여 할 수는 있지만 추정절차가 번거롭고 본 연구의 초점에서 벗어나므로 다루지 않았다.

13) 거리를 기준으로 작성된 9개의 공간가중치 행렬 중에서 자가가구는 $d_{ij} \leq 4$ km, 전세가구는 $d_{ij} \leq 9$ km를 기준으로 작성한 횡단표준화 행렬을 이용한 결과만 제시하였다. 여기서 생략된 가중치 행렬을 이용한 경우도 유사한 결과를 보여주고 있으나, 적합도 기준으로 가장 설명력이 높은 자가 및 전세가구를 선택하였다. 모든 가중치행렬을 사용한 결과는 Kim(1997) 참조.

14) 지금까지의 연구에서 주택시장의 상관관계를 공간적 범위로 접근한 연구가 없음으로 인해 이 부분에 대해서는 향후 보다 심도있는 연구가 진행된 후에 다시 한번 재해석 또는 검토하는 것이 타당할 것으로 판단된다.

〈표 3〉 자가가구의 공간시차모형의 추정방법별 회귀계수 추정 결과

변 수	OLS	MLE ^{a)}	S-2SLS ^{b)}	S-2SLS Robust ^{c)}
ρ		0.470*** (0.069)	0.596*** (0.095)	0.554*** (0.069)
CONSTANT	18.35*** (0.107)	9.356*** (1.319)	6.959*** (1.83)	7.772*** (1.555)
DHOUS	0.118*** (0.041)	0.129*** (0.039)	0.131*** (0.039)	0.116*** (0.042)
DINCOM	0.217*** (0.041)	0.156*** (0.040)	0.140*** (0.042)	0.1500*** (0.040)
DFUEL	0.185*** (0.053)	0.187*** (0.050)	0.188*** (0.051)	0.175*** (0.055)
TFLSP	0.011*** (0.0008)	0.011*** (0.0008)	0.011*** (0.0008)	0.011*** (0.002)
NMRMS	0.081*** (0.014)	0.078*** (0.013)	0.078*** (0.013)	0.078*** (0.018)
NMBATH	0.078*** (0.031)	0.082*** (0.030)	0.083*** (0.030)	0.070*** (0.033)
HSAGE	-0.0046** (0.002)	-0.0057*** (0.002)	-0.0059*** (0.002)	-0.0067*** (0.002)
ACSHPT	-0.0035 (0.002)	-0.0041* (0.0021)	-0.0043** (0.0021)	-0.0040** (0.0019)
ACSSCH	-0.0072*** (0.002)	-0.0058*** (0.002)	-0.0054*** (0.002)	-0.0059*** (0.002)
ACSSUB	-0.0024 (0.001)	-0.0026 (0.0016)	-0.0026 (0.0016)	-0.0018 (0.0015)
ACSPRK	-0.003* (0.0015)	-0.0030** (0.0015)	-0.0029** (0.0015)	-0.0035** (0.0013)
MSO2T	-0.0149*** (0.003)	-0.0077** (0.003)	-0.0058** (0.003)	-0.0065** (0.003)
MNO2T	0.002 (0.0023)	0.001 (0.002)	0.0007 (0.002)	0.00107 (0.002)
R^2	0.629	0.639	0.650	0.655
BP/KB ¹⁾	206.17***	380.68***	380.68***	-
LM-Error ²⁾	54.66***	0.168	0.065	-
LM-Lag ³⁾	62.09***	-	-	-

주: a) 최우추정법, b) 공간 2단계 추정법, c) 공간 2단계 강건추정법을 의미함.

1) PB/KB는 이분성을 검증하는 Breusch-Pagan 및 Koenker-Bassett 검증,

2) 공간오차(λ)에 대한 라그랑제승수 검증,

3) 공간시차(ρ)에 대한 라그랑제승수 검증을 의미한다.

*** 1%에서 유의함, ** 5%에서 유의함, * 10%에서 유의함. (): S. D.

지 않는 것으로 나타나고 있다. 이 결과는 매우 흥미있는 것으로, 그 이유는 오염 물질의 특성에 기인하는 것으로 보인다. 아황산가스는 눈에 보이는(visible pollutant) 오염물질인 반면, 이산화질소는 보이지 않는 오염적 특성에 기인하는 것으로 해석할 수 있다.¹⁵⁾

이분산성(heteroscedasticity)과 공간종속성(spatial dependence)에 대한 회귀진단 결과는 오차항의 비정규성(Jarque-Bera 통계량: 82.2 with $\alpha < 0.000$)과 더불어 심각한 OLS 모형의 결점이 존재함을 나타내 보이고 있다. Koenker-Bassett 검증결과($\alpha < 0.000$)는 이분산성이 존재함을 알 수 있고, 공간오차 및 공간시차에 대한 LM 검증결과는 강력한 공간종속성이 존재함을 보여주고 있다.

이 문제를 해결하기 위하여 공간시차모형을 이용한 최우추정법(MLE)과 공간 2단계추정법(S-2SLS)을 사용하였다. 결과를 보면, 공간시차모형을 이용할 경우 공간자기계수(ρ)가 통계적으로 매우 유의하게 나타남을 알 수 있다. 공간시차모형을 사용할 경우에 오차의 공간종속성에 대한 LM 검증치가 귀무가설($H_0: \lambda = 0$)을 기각하지 못하는 것으로 나타나고 있어 모형의 적합성을 보여주고 있다.¹⁶⁾ 그러나, 비록 최우추정법이나 공간 2단계 추정법을 사용하여도 여전히 이분산성(Koenker-Bassett 값이 5% 통계수준에서 유의함)이 존재하고 있어 추정치의 효율성에 영향을 미치고 있음을 알 수 있다. 이분산성 문제를 해결하기 위하여 강건 불편추정 방법인 공간 강건 2단계 추정법(Robust 2SLS)을 사용하여 추정하였다(〈표 3〉 참조).¹⁷⁾

(2) 전세가구 거주자

전세가구에 대한 추정 결과도 앞에서 설명한 자가가구와 동일한 방법으로 추정하였다. 전세가격에 대한 설명변수도 자가가구와 동일하지만 주택의 건축연수에 대한 항목은 자료의 신뢰성 측면에서 생략하였다.¹⁸⁾ 전세가구자에 대한 계수의 추정 결

15) 대기오염물질의 특성에 대해서는 Seinfeld (1975) 등에서 물리적 화학적 특성을 잘 다루고 있다.

16) 자세한 설명은 Anselin (1998b) 참조.

17) 보다 자세한 공간 회귀진단 절차 및 상세한 설명에 대하여는 Anselin (1988a, 1999a), Anselin and Bera (1998) 참조.

18) 실제 설문조사 결과에서 건축연도에 대한 질문에 응답하지 않는 가구가 상당히 많았는데, 이는 전세자의 경우 정확히 건축연도(특히, 단독주택)를 알 수 없는 상황으로 판단하여 제외하였다.

〈표 4〉 전세가구의 공간시차모형의 추정방법별 회귀계수 추정 결과

변 수	OLS	MLE	S-2SLS	Robust S-2SLS
ρ	-	0.322*** (0.12)	0.397*** (0.13)	0.422*** (0.11)
CONSTANT	16.58*** (0.10)	10.94*** (2.11)	9.649*** (2.31)	9.197*** (2.01)
DHOUS	0.103*** (0.029)	0.095*** (0.029)	0.094*** (0.029)	0.103*** (0.027)
DINCOM	-0.027 (0.044)	-0.044 (0.043)	-0.0486 (0.044)	-0.0574 (0.044)
DFUEL	0.100*** (0.036)	0.100*** (0.036)	0.100*** (0.036)	0.086** (0.040)
TFLSP	0.038*** (0.002)	0.037*** (0.002)	0.037*** (0.002)	0.037*** (0.002)
NMRMS	0.116*** (0.030)	0.116*** (0.029)	0.116*** (0.030)	0.126*** (0.028)
NMBATH	0.0075 (0.07)	-0.032 (0.07)	-0.041 (0.07)	-0.072 (0.066)
ACSHPT	-0.0068*** (0.002)	-0.0065*** (0.002)	-0.0064*** (0.002)	-0.0057*** (0.002)
ACSSCH	-0.005*** (0.0018)	-0.0046*** (0.0018)	-0.0045** (0.0018)	-0.0049*** (0.0018)
ACSSUB	-0.0034** (0.001)	-0.0038** (0.001)	-0.0039*** (0.001)	-0.0038*** (0.001)
ACSPRK	-0.0015 (0.0014)	-0.0012 (0.0013)	-0.0012 (0.0013)	-0.0007 (0.0013)
MSO2T	-0.0089*** (0.002)	-0.0056*** (0.002)	-0.0048*** (0.002)	-0.0038*** (0.002)
MNO2T	-0.0007 (0.0019)	0.0001 (0.0019)	0.0003 (0.0019)	0.0005 (0.0018)
R^2	0.73	0.74	0.74	0.74
BP/KB	24.4***	24.9**	24.9**	-
LM-Error	0.19	2.07	2.07	-
LM-Lag	7.43***	-	-	-

주: *** 1%에서 유의함, ** 5%에서 유의함, * 10%에서 유의함.

() : S. D.

과도 자가가구와 유사한 결과를 보이고 있다. 한 가지 흥미있는 결과는 지하철역 접근도(ACSSUB)가 자가가구에서는 유의하지 않지만 전세가격의 결정에는 매우 유의한 변수로 나타난 반면, 근린공원 접근도(ACSPRK)가 자가가구에서는 주택가격에 유의한 변수인 데 반하여, 전세가격의 결정에는 유의하지 않음을 보이고 있다. 이는 전세자의 경우 근린공원이라는 쾌적성보다는 교통이 편리한 지하철 역세권 주택을 선호하고 있음을 입증하고 있다. 일반적으로 전세자의 경우 자가가구 거주자보다 소득수준이 낮은 데 기인하여 쾌적성보다는 편리성을 더 선호한 결과로도 해석할 수 있다.¹⁹⁾

회귀진단 결과를 보면, OLS를 이용하여 추정할 경우 이분산성과 공간시차가 존재함을 알 수 있다. 자가가구와 달리 공간오차에 대한 회귀진단 결과는 공간오차가 존재하지 않음을 보여주고 있다. 공간시차를 고려한 최우추정법과 2단계 최소자승 추정법의 결과를 보면, 공간시차계수(ρ)가 통계적으로 매우 유의한 수준을 보이고 있지만, 이분산성 문제가 여전히 심각함을 보여주고 있다. 이분산성 문제를 해결하기 위하여 자가가구와 동일한 방법인 강건 2단계 추정량을 통하여 최종적으로 추정치를 구하였다(〈표 4〉 참조).

3. 공간오차모형

OLS 회귀진단의 결과는 전세가가구는 공간시차만이 존재하는 반면, 자가가구의 경우 공간시차 및 공간오차 모두가 존재함을 보여주고 있다. 따라서 공간오차모형을 자가가구에 대해서만 적용·추정하였다. 앞에서 살펴본 바와 같이 공간시차모형을 적용할 경우의 추정 결과는 공간오차가 더이상 존재하지 않음을 알 수 있었다. 여기서는 또 하나의 모형정립 방법인 공간오차모형을 적용하여 추정하였다. 추정방법으로는 최우추정법과 GM(generalized moment method)을 이용하였다.²⁰⁾

MLE 추정의 회귀진단을 보면, 공간오차모형을 적용할 경우, λ 값은 통계적으로 유의하나(1%), 여전히 이분산성 문제가 심각함을 보여주고 있다. 또한 공통인자가

19) 현재와 같이 지하철 노선이 보다 많이 있을 경우에는 자가주택도 지하철 접근성이 유의한 결과로 나타날 수도 있을 것으로 예상된다.

20) GMM 추정방법에 대하여는 Anselin and Bera(1998)을 참고하기 바라며, 여기서 추정에 사용된 소프트웨어는 SpaceStat (Version 1.90, 1997)이다.

〈표 5〉 자가가구의 공간오차모형의 회귀계수 추정 결과

변 수	MLE	GM-Two Step
CONSTANT	18.166*** (0.157)	18.239*** (0.141)
DHOUS	0.144*** (0.040)	0.140*** (0.040)
DINCOM	0.139*** (0.041)	0.151*** (0.041)
DFUEL	0.195*** (0.0007)	0.195*** (0.051)
TFLSP	0.011*** (0.0008)	0.011*** (0.0008)
NMRMS	0.085*** (0.013)	0.084*** (0.013)
NMBATH	0.074*** (0.030)	0.074** (0.030)
HSAGE	-0.0044*** (0.002)	-0.0045** (0.002)
ACSHPT	-0.0033 (0.0021)	-0.0033 (0.0021)
ACSSCH	-0.0064*** (0.002)	-0.0066*** (0.002)
ACSSUB	-0.0034* (0.0017)	-0.0033* (0.0017)
ACSPRK	-0.0025 (0.0015)	-0.0025* (0.0015)
MSO2T	-0.0105* (0.005)	-0.0127*** (0.0047)
MNO2T	0.0049 (0.003)	0.0042 (0.003)
λ	0.628*** (0.083)	0.515*** (0.000)
R^2	0.570	0.580
BP/KB	364.24***	-
Common Factor (LR)	39.39***	-
LM-Lag	4.184**	-

주: *** 1%에서 유의함, ** 5%에서 유의함, * 10%에서 유의함.

(): S. D.

설을 기각함으로써 공간시차모형이 더욱 적합한 것으로 나타나고 있다. 공간오차모형을 적용하여도 여전히 공간시차가 존재(LM-Lag: 통계적으로 5%에서 유의)하는 것으로 나타나고 있다. 이는 계량경제학적인 측면에서 보면 관련변수의 생략(omit of relevant variable)으로서, 이렇게 추정된 계수는 편기되거나 일관되지(biased and inconsistent) 못한 추정결과를 초래한다.

이상의 분석 결과, 공간중속성을 계량화하는 두 가지 기본적인 방법 중 공간시차모형이 더 적합한 것으로 판명되었다. 기존의 특성가격을 이용한 특정변수의 한계 변화에 따른 가격변화는 추정된 계수(β)이지만, 공간시차모형의 경우는 $\beta k \cdot (1 - \rho W)^{-1}$ 가 된다. 의미를 해석하여 보면 지역 1의 주택가격은 지역 1의 하나의 특성가격의 변화에 의해 영향을 받을 뿐만 아니라 공간가중치 행렬을 통하여 다른 지역의 특성가격의 변화도 지역 1의 주택가격에 영향을 미치고 있음을 나타낸다. 여기서 우리는 횡단표준화된 공간가중치 역행렬($[I - \rho W]^{-1}$)은 무한기하과정의 합($1 + \rho W + (\rho W)^2 + (\rho W)^3 + \dots$)이므로, 주택가격에 대한 한 특성변수에 대한 총 효과(amplified effect)를 '空間乘數'(spatial multiplier)라고 말할 수 있다. 그러므로 공간시차모형은 주택가격결정의 직접효과뿐만 아니라 간접효과도 함께 파악할 수 있는 장점을 가지고 있다.

V. 要約 및 結論

서울의 주택시장에 대한 특성가격모형은 자가 및 전세 모두 공간오차모형보다는 공간시차모형이 더 적합한 것으로 판단된다. 본 연구의 결과는 공간시차를 고려하지 않은 기존의 특성가격모형은 설명변수 중에서 관련변수의 누락으로 추정된 통계량이 편기될 뿐만 아니라 일관된 추정량을 구할 수 없는 것이다. 모형에 의해서 추정된 계수들을 살펴보면, 대부분의 추정계수 부호가 기대부호와 일치하는 것으로 나타나고 있다. 추정계수 중에서 흥미있는 결과는 환경변수 중에서 눈에 보이는 아황산가스는 주택가격 결정에 통계적으로 유의하게 영향을 미치고 있지만 눈에 보이지 않는 이산화질소는 통계적으로 유의하지 않은 변수로 나타나고 있다는 점이다. 또 다른 하나는 쾌적성을 나타내는 근린공원 접근도는 자가주택가격 결정에는 의미 있는 변수이지만 전세가격 결정에는 유의하지 않은 변수로 나타나고 있다. 전세가

격 결정변수로는 쾌적성보다는 접근성을 나타내는 지하철역의 근접도가 유의한 변수로 나타나고 있다는 점이다.

공간시차모형을 적용하여 최우추정법이나 공간 2단계 최소자승법을 이용하여도 여전히 남아 있는 이분산성 문제는 강건 2단계 추정방법을 사용하는 것이 바람직한 추정방법의 하나이다. 또 하나 공간시차모형은 기존의 모형으로는 구할 수 없는 환경오염과 같은 외생변수의 변화에 따른 직접효과 $\left(\frac{\partial P}{\partial x_{1k}}\right)$ 뿐만 아니라 간접효과 $\left(\sum_{i=2}^n \frac{\partial P_i}{\partial x_{1k}}\right)$ 도 동시에 포착할 수 있다는 점이다.

최근 들어 발달되고 있는 지리정보시스템(GIS)으로 우리나라도 공간자료가 상당 부분 축적되고 있고, 공간계량분석 방법론에 대한 이론적 연구와 관련 소프트웨어 또한 급속히 발달되고 있음을 감안할 때 이 분야에 대한 지속적인 연구가 필요할 것으로 판단된다.

■ 參考文獻

1. 국토개발연구원, 『국민 주거상태 및 주거의식 조사』, 1993.
2. 서울시정개발연구원, 『서울市 大氣汚染 特性 研究』, 1994.
3. Anselin, L., *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer Academic Publishers, 1988a.
4. ———, "Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity," *Geographical Analysis*, Vol. 20, 1988b, pp. 1~17.
5. ———, "GIS Research Infrastructure for Spatial Analysis of Real Estate Markets," *Journal of Housing Research*, Vol. 9, 1998, pp. 113~133.
6. ———, "Spatial Econometrics," in B. Baltagi(ed.), *Companion in Theoretical Econometrics*, Oxford: Basil Blackwell (forthcoming), 1999a.
7. ———, "Raos Score Test in Spatial Econometrics," *Journal of Statistical Planning and Inference* (forthcoming), 1999b.
8. ——— and A. Bera, 1998, "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics," in A. Ullah and D. Giles(eds.), *Handbook of Applied Economic Statistics*, New York: Marcel Decker, 1998, pp. 237~289.
9. ——— and H. Kelejian, 1997, "Testing for Spatial Error Autocorrelation in the Presence of Endogenous Regressors," *International Regional Science Review*, Vol. 20, 1997, pp. 153~182.
10. Can, A., "The Measurement of Neighborhood Dynamics in Urban House Prices," *Economic Geography*, Vol. 66, 1990, pp. 254~272.

11. ———, "Specification and Estimation of Hedonic Housing Price Models," *Regional Science and Urban Economics*, Vol. 22, 1992, pp. 453~474.
12. Cliff, A. and J. Ord., *Spatial Processes, Models and Applications*, London: Pion, 1981.
13. Chou, Yue-Hong, *Exploring Spatial Analysis in Geographic Information Systems*, Santa Fe: Onword Press, 1997.
14. Dubin, R. A., "Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms," *The Review of Economics and Statistics*, Vol. 70, 1988, pp. 466~474.
15. ———, "Spatial Autocorrelation and Neighborhood Quality," *Regional Science and Urban Economics*, Vol. 22, 1992, pp. 432~452.
16. Kelejian, H. and I. Prucha, "A Generalized Spatial Two Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances," *Journal of Real Estate Finance and Economics*, Vol. 17, 1998, pp. 99~121.
17. Kelejian, H. H. and D. P. Robinson, "A Suggested Method of Estimation for Spatial Interdependent Models with Autocorrelated Errors, and an Application to a County Expenditure Model," *Papers in Regional Science*, Vol. 72, 1993, pp. 297~312.
18. Kim, Chongwon, "Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach," Unpublished Ph. D. Dissertation, West Virginia University, 1997.