

**해도닉 모델 추정시 GIS 공간분석기능에 의해 생성된 근린변수의 기여도에 대한 연구
- 토지이용도를 이용한 근린변수의 타당성을 중심으로 -**

손 철*

**A Study on the Contribution of GIS-Created Neighborhood Quality
Variables in Estimating Hedonic Price Models**

Chul Sohn*

요 약

본 연구는 근린유홍시설의 분포를 나타내는 해도닉 모델의 근린변수가 지리정보시스템의 공간분석기능을 충분히 이용하여 측정되었을 경우, 그렇지 않을 경우에 비해 해당변수를 포함하는 해도닉 모델의 통계적 질을 향상시킬 수 있는가를 평가하고 있다. 평가결과는 해당변수가 지리정보시스템의 공간분석기능을 충분히 이용하여 측정되었을 경우가 해도닉 모델의 설명력측면에서 우월함을 보이고 있다. 본 연구결과는 지리정보시스템이 단순한 직선거리를 추정하는 것 이상으로 해도닉 모델의 질을 향상시킬 수 있는 잠재력을 가지고 있다는 것을 보이고 있으며 주택시장의 행태를 설명하는 이론적으로 타당한 근린변수의 측정을 위해 보다 적극적으로 이용되어야 함을 실증적으로 보이고 있다.

주요어 : 지리정보시스템, 근린변수, 해도닉 모델

ABSTRACT : Variables representing neighborhood quality should be included in hedonic price models to control for the influences of negative or positive externalities from the quality of neighborhood on urban housing prices. This study proposes a GIS-based method to effectively measure the neighborhood quality variable when data on the neighborhood quality are aggregated by census sub area. This study also tests the superiority of the proposed neighborhood quality variable created by intensive use of GIS operations to a neighborhood

* 삼성SDS(Samsung SDS)

손 철

variable not based on GIS operations in explaining the housing price variations by using Seoul's apartment sales data. The results from this study show that the neighborhood quality variable based on GIS-based operations shows better performance in explaining the urban housing price variations in Seoul's housing market. The implication from the results is that the potentials of GIS-based spatial operations in creating neighborhood quality variables should be well acknowledged by the researchers in the area of urban housing market study and GIS-based spatial operations should be more actively applied to generate better neighborhood quality variables for hedonic price models.

Keyword : GIS, neighborhood variable, Hedonic Price Model

1. 서 론

부동산의 가격과 속성간의 통계적 관계를 나타내는 헤도닉 모델은 공공사업의 비용/편익분석, 기피시설 입지의 영향분석, 부동산가격지표의 작성 등에서 활발하게 이용되고 있다. 근래에 들어 지리정보시스템이 헤도닉 모델의 추정을 위해 활발하게 사용되고 있는데 이는 지리정보시스템의 사용을 통해 대용량의 부동산 데이터의 위치파악이 가능하고, 이를 통해 헤도닉 모델의 추정에 필수적인 다양한 형태의 근린특성(Neighborhood Quality)을 나타내는 근린변수들을 효과적으로 측정할 수 있기 때문이다.

근린변수의 측정과 관련하여 지리정보시스템 고유의 공현은, 공간분석기능 즉 베퍼링, 네트워크분석 등을 이용하면 그동안 이론적으로는 타당할지라도 종이지도와 자를 이용하여서는 측정하기 어려웠던 다양한 형태의 근린변수의 측정이 가능해졌다는 점이다. 그러나 이렇게 해서 측정된 변수들이 종이지도와 자를 이용하여 측정되었을 경우의

형태에 비해 헤도닉 모델의 통계적 질을 향상시키지 않는다면 헤도닉 모델의 추정을 위한 지리정보시스템의 기여는 단순히 모델링 과정에서의 시간을 절약하는 데 그치고 말 것이다.

최근 몇몇 연구들이 지리정보시스템의 공간분석기능을 통해 측정된 근린변수들이 실제로 헤도닉 모델의 통계적 질을 설명력의 향상과 설정의 오류 (Specification Error)의 감소라는 측면에서 높일 수 있음을 보이고 있다. 그러나 이들 연구들이 다루고 있는 근린변수의 유형이 제한되어 있어 근린변수 측정시 헤도닉 모델의 통계적 질을 향상시키기 위해 지리정보시스템의 공간분석기능을 적극적으로 이용하게 하는 데 근거자료로써 충분한 역할을 하지 못하고 있다.

이런 배경에서 본 연구에서는 선행연구에서 다루지 않았던 유형의 근린변수를 대상으로 과연 이 변수가 지리정보시스템의 공간분석기능을 충분히 이용하여 측정되었을 경우, 그렇지 않은 경우에 비해 해당변수를 포함하는 헤도닉 모델의 통계적 질을 향상시킬 수 있는 가를 평가함으

로써 해도닉 모델을 이용한 주택시장연구에서 지리정보시스템의 공간분석기능의 이용이 주는 가능성에 관한 추가적인 정보를 제공하고자 한다.

2. 지리정보시스템과 해도닉 모델의 추정

2.1 해도닉 모델에서의 근린변수의 역할

부동산이 가지는 독특한 두 가지 특성은 공간적 고정성 (Spatial Fixity)과 내구성 (Durability)이다. 이러한 특성은 부동산이 공간적으로 고정적이지 않거나 내구성이 없을 경우 중요시되지 않을 부동산이 위치한 근린의 특성을 부동산의 소유자와 이용자에게 중요시하게 하는 역할을 한다. 이러한 부동산 소유자와 이용자의 관심은 해당 부동산의 판매와 임대시 이들이 지불하고자하는 프리미엄에 반영된다.

부동산가격과 밀접한 관계를 가지는 근린의 특성은 접근도 (Accessibility), 물리적 환경(Physical Environment), 사회인구적 특성(Social & Demographic Context), 공공서비스 공급 (Public Service Provision) 등 몇 개의 유형으로 분류할 수 있다(Can, 1998). 이 가운데 접근도는 도시내 고용과 교통시설에 대한 접근도를 의미한다. 물리적 환경은 근린내의 기피시설의 존재유무, 근린내의 주택자체의 특성의 분포 등을 나타낸다. 사회인구적 특성은 근린내의 인종의 분포, 소득수준의 분포들을 의미한다. 공공서비스공급은 근린에 대한 공공지출과 밀접하게 관련된 특성으로 근린내 학교교육의

질, 치안의 수준 등을 나타낸다.

부동산의 가격과 속성간의 통계적 관계를 나타내는 해도닉 모델에서는 근린의 특성을 근린변수로 측정하여 적절하게 모델에 포함시키지 않으면 안된다. 이것은 적절하게 포함되어야 할 근린변수가 빠지게 되면 설정의 오류 (Specification Error)를 범하게 되며 만약 근린변수의 측정에 심각한 오차가 존재한다면 측정의 오류 (Measurement Error)에 의해 해도닉 모델에 포함된 모든 변수의 모수추정치가 편이되기 때문이다.

2.2 지리정보시스템을 이용한 근린변수의 측정

해도닉 모델의 추정을 위해 사용되는 근린변수를 작성하기 위해서는 미시데이터 (Micro Data), 집계데이터 (Aggregate Data), 공간데이터 (Spatial Data) 등 세 가지 유형의 데이터가 사용될 수 있다 (Sawiki and Flynn, 1996). 미시데이터는 개인에 대한 설문조사를 통해 얻어지는 데이터이다. 일반적으로 주민을 대상으로 한 주위환경에 대한 질문과 이에 대한 답변을 통해 얻게 된다. 집계데이터는 일반적으로 센서스 통계집계구역을 단위로 집계되어 보고되는 데이터로, 통계구역별 소득수준 등이 그 예이다. 마지막으로 공간데이터는 집계데이터가 센서스 구역별로 집계되기 전의 유형의 데이터로, 예를 들어 장소별 범죄발생과 같이 개별적인 데이터의 위치정보가 알려져 있는 상태의 데이터이다. 앞서의 데이터 가운데 집계데이터와 공간데이터만이 지리정보시스템을 이용한 근린변수의 측정에 이용될 수 있다.

손 철

이상의 데이터를 이용한 근린변수의 측정에서 지리정보시스템이 기여하는 점은 크게 두 가지 측면에서 정리할 수 있다.

첫번째 측면에서의 기여는 지리정보시스템의 활용을 통해 대용량 부동산 데이터의 손쉬운 위치파악이 가능하고, 이러한 위치파악을 통해 다양한 근린속성과 개개의 부동산간의 공간적 상호관계가 손쉽게 파악될 수 있다는 데 있다. 이러한 상호관계의 파악은 과거에는 종이지도와 자를 이용한 수작업에 의해 이루어졌다. 다루고자 하는 부동산 샘플의 수가 많을 경우 이러한 지리정보시스템의 자료처리기능을 이용한 상호관계의 파악은 분석의 시간을 절대적으로 줄여주는 역할을 한다.

두번째 측면에서의 기여는 다양한 형태의 근린변수측정이 지리정보시스템 특유의 공간분석기능을 통해 가능해졌다는 점이다. 예를 들어 종이지도와 자를 이용할 경우는 근린특성 가운데 접근도를 측정하기 위해서 직선거리 (Euclidean Distance) 만이 사용되었으나 지리정보시스템이 사용되면서 직선거리 대신 네트워크분석을 이용한 네트워크 거리가 접근도를 측정하는 데 사용되기 시작했다. 이상의 두번째 측면의 기여는 지리정보시스템 고유의 것으로 샘플의 수가 적더라도 종이지도와 자에 의해서 대신하기는 사실상 불가능한 것이다.

2.3 지리정보시스템의 공간분석기능의 기여에 대한 경험적 평가

현재까지 제한된 수의 연구가 지리정보

시스템의 공간분석기능을 통해 생성된 근린변수들이 과연 해도닉 모델의 통계적 질을 개선할 수 있는가 하는 문제를 다뤘다. 이러한 연구들은 크게 두 가지 부류로 나누어질 수 있다. 첫번째 유형의 연구는 해당변수들이 종이지도와 자를 이용하였을 경우 측정되었을 형태와 비교할 때 해도닉 모델의 통계적 질을 향상시킬 수 있는가 하는 문제를 다루고 있다. 두번째 유형의 연구는 종이와 자를 이용한 연구에서는 실제로 측정하기가 어려운 근린변수를 제안하고 과연 근린변수가 해도닉 모델에서 통계적으로 유의한가 하는 문제를 다루고 있다.

Rodriguez *et al.* (1995)은 첫번째 유형의 연구이다. 이 연구에서는 Baton Rouge, Louisiana의 주택시장을 대상으로 도심으로부터의 접근도가 도심으로부터의 직선거리와 도심으로부터의 네트워크 거리로 각각 측정되었을 경우 어느 것이 단독주택 판매가격의 분포를 설명하는 데 우월한지를 검토하였다. 연구결과는 지리정보시스템의 네트워크분석을 통해 얻어진 네트워크 거리가 단순한 직선거리보다 단독주택판매 가격의 분포를 설명하는 데 우월함을 보여주었다.

Geoghegan *et al.* (1997)의 연구는 두 번째 유형의 연구에 속한다. Geoghegan *et al.* (1997)은 Washington D.C. 주변 20마일 내의 단독주택 시장자료와 지리정보시스템의 베퍼링 기능을 이용, 각 주택 주변 지역 토지이용의 다양성(Diversity)¹⁾과 편린화(Fragmentation)²⁾를 측정하는 지표를 작성

1) Diversity Index: $H = - \sum(P_k) \ln(P_k)$, (P_k : k 유형의 토지이용)

2) Fragmentation Index: $R = \sum(P_i/A_i)$ (P : Perimeter length, A : Area of interior)

하였다. 그 후 이들 지표가 주택의 판매가에 미치는 영향을 분석하였다. 분석 결과 100미터 범위내의 토지이용 다양성과 편리화의 증가는 집값에 네가티브한 영향을 미치며 통계적으로 그 영향이 유의한 것으로 나타나고 있다.

이상의 경험적 연구들은 지리정보시스템의 공간분석기능을 근린변수의 측정에 이용함으로써 설명력의 향상과 통계적으로 유의한 변수의 포함을 통한 설정의 오류(Specification Error)의 최소화를 통해, 해도닉 모델의 통계적 질을 향상시킬 수 있음을 보여주고 있다. 그러나 이상의 연구에서 다루고 있는 근린특성의 유형은 상당히 제한된 것으로 보다 많은 유형의 근린특성에서 동일한 결과를 얻을 수 있다면 지리정보시스템의 공간분석기능이 해도닉 모델의 추정을 위해 보다 활발하게 이용되는 데 기여할 수 있을 것이다.

3. 분석의 틀

3.1 근린특성으로서 유통업소의 분포

본 연구의 목적은 지리정보시스템의 공간분석기능을 이용하여 측정된 근린변수에 대해 해당변수가 종이지도와 자료를 이용하여 측정되었을 경우의 형태를 제시하고 두 가지 형태 가운데 어떤 형태가 해도닉 모델의 통계적 질을 향상시킬 수 있는지를 분석하는 데 있다. 본 연구의 목적은 Rodriguez *et al.* (1995)의 연구목적과 동일하지만 다루고자 하는 근린변수와 사용되는 공간분석기능이 다르다.

본 연구에서 다루는 근린변수는 유통업소의 분포이다. 주택가 가까이 유통업소가 입지할 경우 거주자들은 유통업소로부터의 소음 등 부의 공간적 외부효과를 겪게 된다. 따라서 유통업소의 운영자와 주택소유자들의 분쟁이 상존하게 된다. 이를 효과적으로 해결하게 위해서는 관련법의 엄격한 집행과 함께 관련 당사자간의 활발한 협상이 선행되어야 한다. 해도닉 모델의 추정을 통한 주택가격과 유통업소 입지와의 관계분석은 이러한 협상에 커다란 도움을 줄 수 있다. 그 이유는 양자의 관계파악을 통해 주택의 소유자나 유통업소의 운영자가 현실을 바르게 파악하고 여러 가지 선택 가능한 대안에 대한 비용/편익분석이 가능하기 때문이다.

주택가격과 유통시설업소와의 관계를 파악하기 위해서는 통상적으로 해도닉 함수에 주택과 유통업소와의 거리를 포함하는 방법을 고려해 볼 수 있다. 이 경우 거리변수로부터 얻어지는 추정모수는 다른 모든 조건이 일정할 경우 주택으로부터 유통시설까지의 거리가 증가함에 따른 주택가격의 변동을 나타낸다. 앞서의 거리변수를 측정하기 위해서는 주택의 위치에 관한 자료와 유통업소의 위치에 관한 자료가 필요하다. 이러한 자료는 통상적으로 기존의 전산화된 데이터베이스나 데이터베이스가 존재하지 않을 경우 방문조사에 의해 얻을 수밖에 없다. 해도닉 함수의 추정이 국지적인 주택시장에 한정될 경우에는 방문조사가 가능하나, 광역적 지역을 대상으로 할 경우에는 방문조사가 사실상 불가능하며 존재하는 데이터베이스에 의존할 수밖에 없다.

우리나라의 경우 주택의 위치에 대해서는 부동산정보업체의 데이터베이스를 통해 비교적 상세한 파악이 가능한 반면, 유통시설의 위치에 대하여는 서울의 경우 제한된 연도에 한하여 행정동 단위의 연상면적에 대한 자료가 일반인에게 제공될 뿐이다(traffic.seoul.go.kr). 따라서 서울주택시장을 대상으로 주택의 유통업소에 대한 접근도를 측정하기 위해서는 주택이 속한 행정동의 유통업소 연상면적을 해당 주택에 부여하여 이를 유통업소에 대한 접근도로 이용할 수밖에 없다.

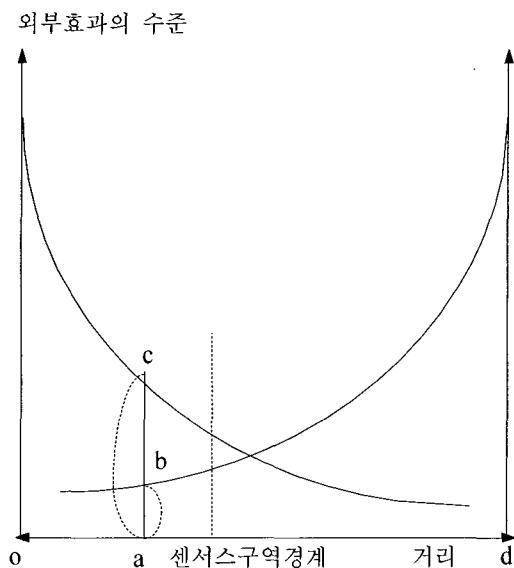
그러나 실제에 있어서 균린특성에 따른 외부효과는 행정동과 같은 센서스구역에 의해 제한되지 않고 센서스구역의 경계를 넘어 공간적으로 확산되는 특성이 있다. 즉 [그림 1]에서 볼 수 있는 것처럼 a지점의 주택은 o에 입지한 유통시설과 d에 입지한 유통시설로부터 ab와 ac에 해당하는

부의 외부효과를 감수해야 한다. 그러나 행정동 단위로 집계된 데이터를 이용하여 주택이 행정동의 어느 지점에 입지하는가에 관계없이 그 행정동이 가진 유통시설의 면적을 가졌다고 가정한다면 이 경우는 ac만큼에 해당하는 외부효과를 누리는 것으로 단순하게 가정하는 것이 된다.

즉 주택이 서로 다른 균린특성을 가지는 센서스 구역들의 경계선에 가까이 위치한다면 해당 주택은 자신이 속해 있는 센서스 구역뿐만 아니라 타 센서스 구역으로부터의 균린특성을 통해서도 영향을 받게 된다. 따라서 주택의 위치는 주소단위로 상세하게 파악되는 반면 관심의 대상인 균린특성이 센서스 구역단위로 보고되는 경우 균린변수를 측정하여 헤도닉 모델에 포함시키기 위해서는 기존의 센서스경계에 의존하지 않는 방법이 필요하다.

3.2 지리정보시스템의 공간분석기능을 이용한 유통시설분포의 측정

본 연구에서는 이러한 필요성에서 다음과 같은 새로운 측정방법을 제안한다. 첫째, 토지이용도(land-use map)상에서 특정한 균린특성이 존재하는 토지이용유형영역을 추출한다. 둘째, 특정근린특성이 그 토지이용영역 내에 고르게 분포한다고 가정하고 단위면적당 해당 균린특성의 밀도를 구한다. 셋째, 추출된 토지이용영역 내의 앞서 계산된 균린특성의 단위밀도를 속성으로 가지는 정밀한 그리드를 형성한다. 넷째, 주택을 중심으로 적절한 버퍼영역을 형성한 후 해당 버퍼영역 내의 특정 균린특성 밀도를 속성으로 가지는 그리드



[그림 1] 행정구역 경계와 균린환경의 외부효과

의 수를 집계한다. 마지막으로 집계된 그리드의 수에 단위밀도를 곱하여 해당버퍼 안에 존재하는 근린특성의 총량을 추정해낸다.

이 방법을 적용해서 유흥시설의 분포를 근린변수로 측정해 보면 [그림 2]와 같다. 즉 먼저 행정동 A와 B의 상업지역을 나타내는 폴리곤을 토지이용도로부터 추출해내고, 단위면적(1 제곱미터)당 유흥시설의 밀도(연상면적)를 값으로 가지는 1 제

곱미터 크기의 그리드를 상업지역 폴리곤을 이용하여 생성한다.

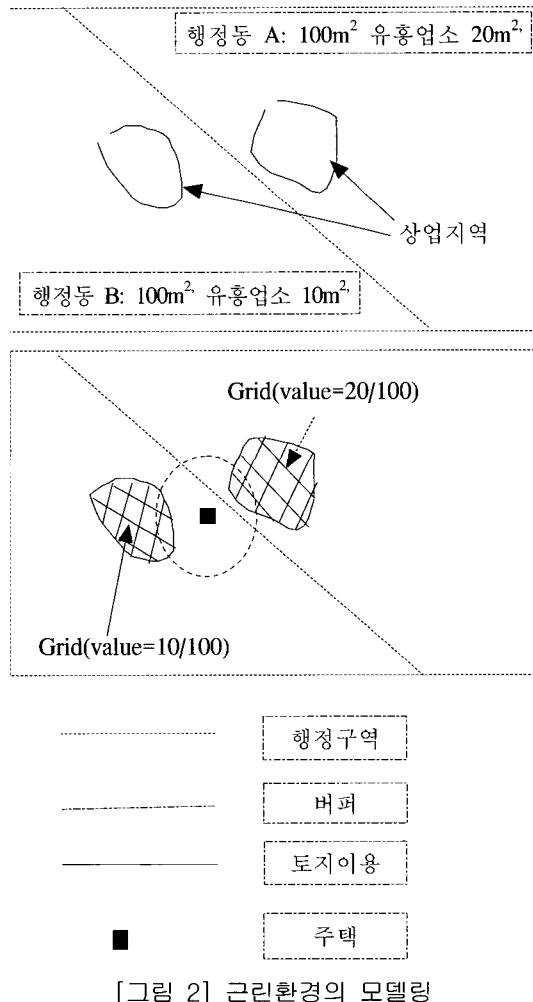
마지막으로 개별 주택을 중심으로 버퍼를 생성하고 버퍼내 그리드의 수를 집계한 후 집계된 그리드 수에 단위면적당 밀도를 곱하면 버퍼 내의 유흥시설의 연상면적이 얻어진다.

이 방법에서의 핵심은 토지이용도를 통해서 근린 특성의 위치를 센서스 경계에서 토지이용도상의 위치로 좁혀감으로써 센서스 통계보고구역이 아닌 개별주택의 버퍼지역으로 근린환경의 측정영역을 좁혀가는 데 있다. 이 방법은 토지이용도가 존재하고 센서스 통계보고구역 단위로 여러 가지 근린 속성에 대한 자료가 존재할 때 적용 가능하다. 예를 들면 특정 대기오염물질 기여도가 높은 산업의 분포가 센서스구역단위로 집계되고 공업지역에 대한 토지이용도가 존재하면 해당산업의 입지가 주택의 가격에 미치는 영향을 파악하기 위해서 위의 방법의 적용이 가능하다.

3.3 비계층적 (Non-Nested) 테스트

앞서 제안된 형태의 근린변수가 동일유형의 근린변수가 지리정보시스템의 공간 분석기능을 이용하지 않은 상황에서 측정되었을 형태에 비해 해도닉 모델의 설명력을 증가시키는가를 검정하기 위해서는 Davidson과 Mckinnon (1981)에 의해 제안된 J 테스트가 이용될 수 있다. J 테스트는 통상적으로 경쟁되는 두개의 모델 가운데 어떤 모델이 더 종속변수를 설명하는 데 우월한 모델인가를 검정할 때 이용된다.

J 테스트에 의해 (1) 과 (2)중 어떤 것이



더 선호되는 모델인지를 검정하기 위해서는 통산 이단계의 절차가 이용된다. 여기서 X , Z^0 , Z^1 은 설명변수의 벡터들이며 (1)과 (2)의 차이점은 (1)에서는 Z^0 가 사용된 데 반해 (2)에서는 Z^1 이 사용되었다는 점이다.

$$Y = X\beta + Z^0\gamma^0 + \varepsilon^0 \quad (1)$$

$$Y = X\beta + Z^1\gamma^1 + \varepsilon^1 \quad (2)$$

$$Y = (1 - \delta^0)(X\beta + Z^0\gamma^0) + \delta^0 \hat{Y}^1 + \varepsilon \quad (3)$$

$$Y = (1 - \delta^1)(X\beta + Z^1\gamma^1) + \delta^1 \hat{Y}^0 + \varepsilon \quad (4)$$

첫번째 단계에서는 (1)을 귀무가설로 그리고 (2)를 대체가설로 설정한 후 (2)로부터의 예측값 (\hat{Y}^1)을 이용하여 (3)을 추정한다. 만약 이때 귀무가설이 진실이라면 δ^0 의 참된 값은 0이어야 한다. 따라서 (3)의 추정결과, δ^0 값이 유의하지 않으면 귀무가설이 참인 것으로 받아들이고 δ^0 값이 유의하면 귀무가설을 기각하게 된다. 다음 단계에서는 (2)를 귀무가설로 그리고 (1)을 대체가설로 설정한 후 (1)에서의 예측값 (\hat{Y}^0)을 이용하여 (4)를 추정한다. (4)의 추정결과 δ^1 값이 통계적으로 유의하지 않으면 귀무가설이 참인 것으로 받아들이며 반대의 경우는 귀무가설이 기각된다.

경우에 따라서 δ^0 과 δ^1 의 통계적 유의성의 다양한 조합이 가능한데 Madalla (1993)는 각각의 조합에 대해 <표 1>과 같은 해석의 률을 제시하고 있다.

3.4 헤도닉 모델의 설정

지리정보시스템의 공간분석기능을 통해 생성된 근린변수와 그렇지 않은 근린변수의 상대적 기여도를 앞서 설명한 J 테스트를 통해 비교하기 위해 본 연구에서는 (5)와 (6)과 같이 헤도닉 모델을 설정하였다. (5)와 (6)의 차이점은 (5)에서는 유통시설의 밀도를 나타내는 근린변수가 앞서 설명한 지리정보시스템의 공간분석기능을 통해 측정되었으며 (6)에서는 단순하게 주택이 입지한 행정동이 가진 유통시설의 밀도를 사용했다는 점이다. 즉 (5)는 동일한 근린변수가 지리정보시스템의 공간분석기능을 이용하여 측정된 형태이며 (6)은 동일 근린변수가 공간분석기능을 이용하지 않고 측정된 형태이다. 본 연구에서는 (5)를 버퍼링 모델, (6)을 행정동 모델이라 부르기로 한다. (5)와 (6)에서 사용된 변수들에 대한 정의는 <표 2>에 설명되어 있다.

<표 1> J 테스트의 해석

	δ^0	δ^1	해석
1	유의함	유의하지 않음	(2)가 채택되어야함
2	유의함	유의함	(1)과 (2)가 모두 기각되어야함
3	유의하지 않음	유의함	(1)이 채택되어야함
4	유의하지 않음	유의하지 않음	(1)과 (2)가 모두 채택되어야 함

<표 2> 사용된 변수의 정의

변 수	정 의
PRICE	가격(만원)
AREA	평형(평)
TUNIT	단지내 전체 가구수(가구)
AGE	건축물의 연령(연)
DSUB	가장 가까운 지하철역으로부터의 거리(미터)
DGREEN	그린벨트 경계까지의 거리(미터)
RIVER	1km내의 하천 및 수변공간의 면적(제곱미터)
PBUFFER	반경 638 미터내의 유흥시설의 연상면적(제곱미터)
PDONG	행정동내의 유흥시설의 연상면적(제곱미터)

버퍼링 모델

$$\begin{aligned} PRICE = & \alpha + \beta^1 AREA \\ & + \beta^2 AGE + \beta^3 TUNIT + \beta^4 DSUB \\ & + \beta^5 DGREEN + \beta^6 RIVER \\ & + \beta^7 PBUFFER + e \end{aligned} \quad (5)$$

행정동 모델

$$\begin{aligned} PRICE = & \alpha + \beta^1 AREA \\ & + \beta^2 AGE + \beta^3 TUNIT + \beta^4 DSUB \\ & + \beta^5 DGREEN + \beta^6 RIVER \\ & + \beta^7 PDONG + e \end{aligned} \quad (6)$$

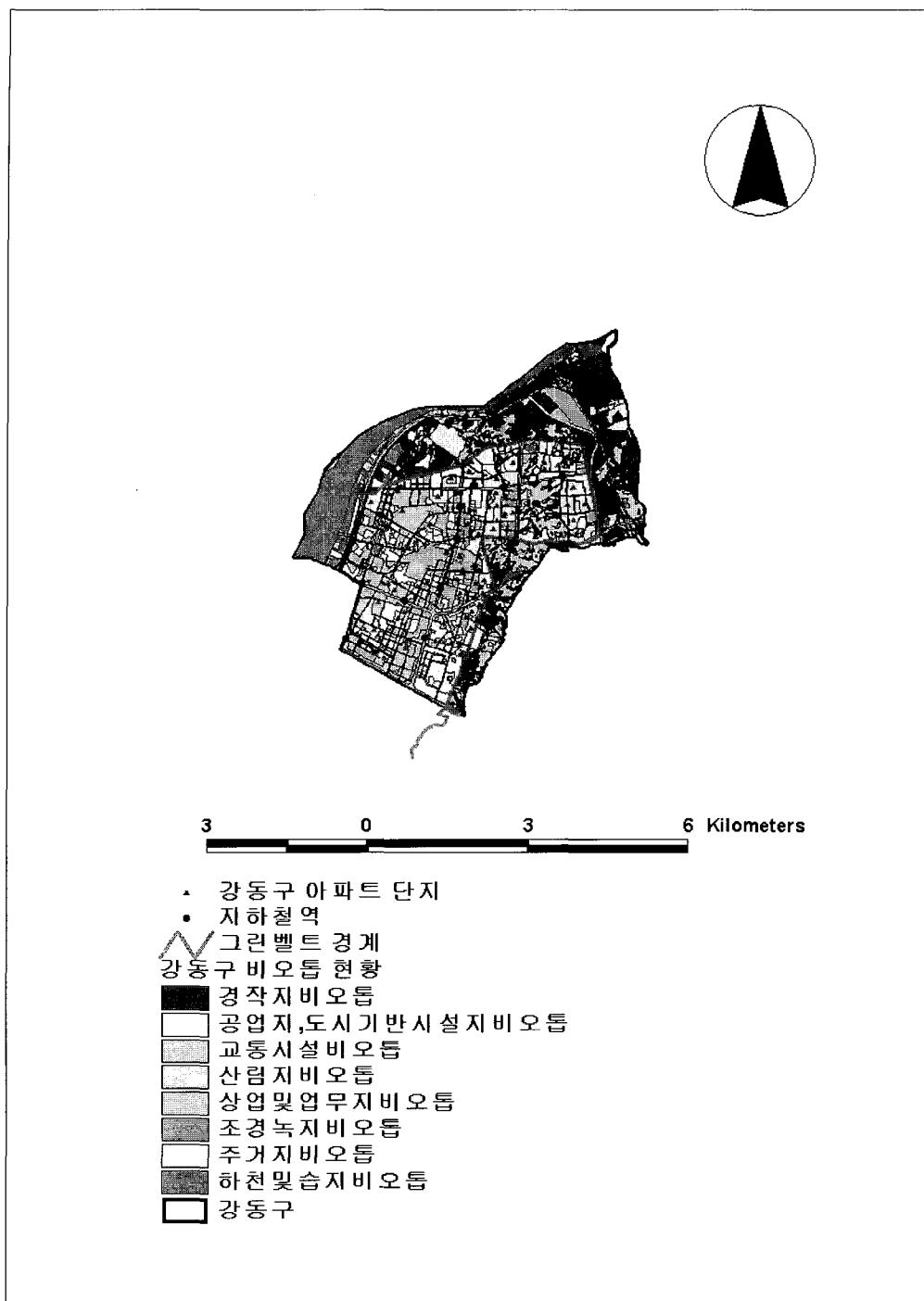
$$\begin{aligned} PRICE = & \alpha + \delta^0 \overbrace{PRICE}^{PDONG} + \beta^1 AREA \\ & + \beta^2 AGE + \beta^3 TUNIT + \beta^4 DSUB \\ & + \beta^5 DGREEN + \beta^6 RIVER \\ & + \beta^7 PBUFFER + e \end{aligned} \quad (7)$$

$$\begin{aligned} PRICE = & \alpha + \delta^1 \overbrace{PRICE}^{PBUFFER} + \beta^1 AREA \\ & + \beta^2 AGE + \beta^3 TUNIT + \beta^4 DSUB \\ & + \beta^5 DGREEN + \beta^6 RIVER \\ & + \beta^7 PDONG + e \end{aligned} \quad (8)$$

(5)와 (6)의 설정과 추정을 위해서는 강동구 아파트시장의 데이터가 사용되었다. 아파트의 가격 및 속성데이터는 부동산뱅크로부터 제공되었으며 아파트단지의 좌표는 타운넷(www.townnet.co.kr)으로부터 제공된 것이 사용되었다. 기타 토지이용도 (1/3000 지도) 및 필요한 지리정보시스템 데이터는 서울시정개발연구원으로부터 제공되었다. [그림 3]은 강동구 및 인근지역의 현황을 보여주고 있다.

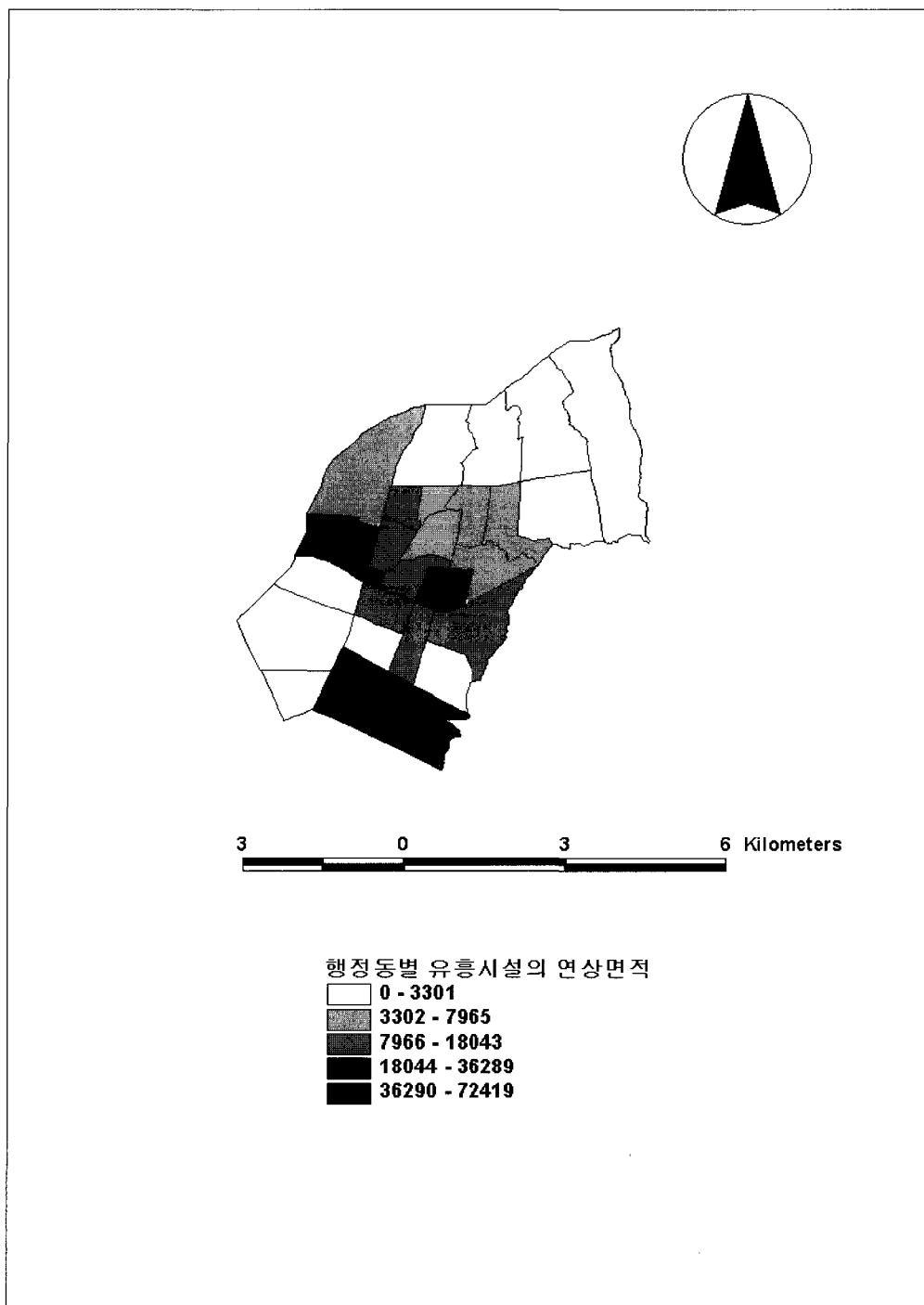
이상의 데이터를 이용하여 (5)와 (6)이 추정된 후 (5)로부터의 예측값, $\widehat{PRICE}^{PBUFFER}$ 와 (6)로부터의 예측값, \widehat{PRICE}^{PDONG} 을 이용하여 (7)과 (8)이 추정된 후 <표 1>에 기반하여 (5)가 강동구의 아파트가격의 변량을 설명하는 데 우월한지 아니면 (6)이 우월한지가 검토될 것이다. PBUFFER 변수를 측정하기 위해서는 [그림 4]에서 볼 수 있는 바와 같은 행정동별 유흥시설 연상면적의 분포와 [그림 5]에서 볼 수 있는 바와 같은 행정동별 상업/업무지역의 분포

손 철



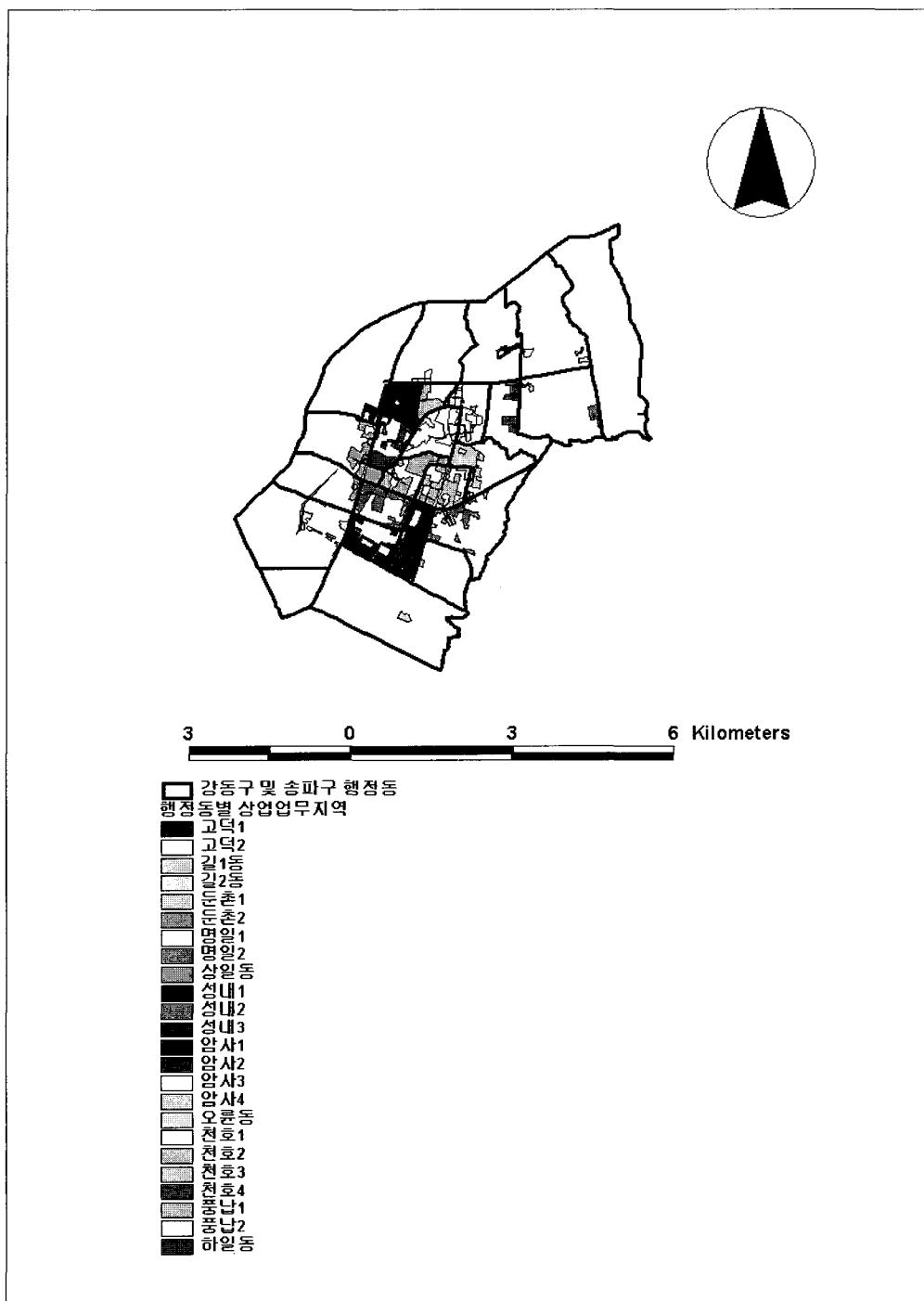
[그림 3] 연구대상지역 - 강동구

혜도닉 모델 추정시 GIS 공간분석기능에 의해 생성된 근린변수의 기여도에 대한 연구



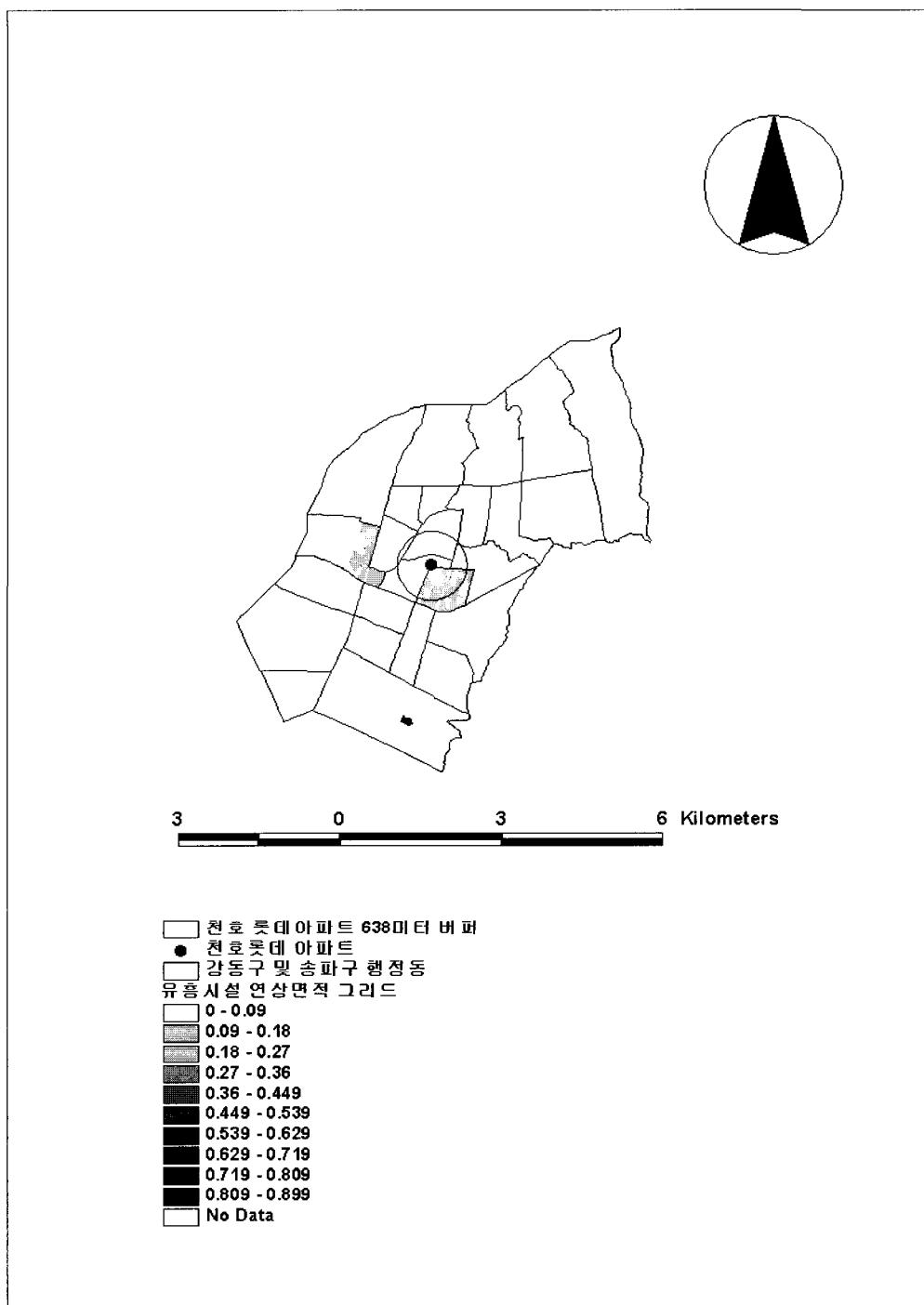
[그림 4] 행정동별 유통시설의 연상면적

손 철



[그림 5] 행정동별 상업업무지역

헤도닉 모델 추정시 GIS 공간분석기능에 의해 생성된 근린변수의 기여도에 대한 연구



[그림 6] 638 미터 버퍼

가 이용되었다. PBUFFER에서는 반경이 638 미터인 버퍼가 이용되었다. 이는 본 연구에서 이용된 행정동의 평균적 면적이 1,277,512 제곱미터임을 고려한 것이다.³⁾ [그림 6]은 행정동별 유흥시설연상면적의 분포와 상업/업무지역의 분포를 이용하여 [그림 2]에서 설명한 바와 같은 1 제곱미터 크기의 그리드가 형성되어 PBUFFER의 측정에 이용되었음을 보이고 있다.

4. 분석결과

모델 (5), (6), (7), (8)의 추정에 사용된 강동구 아파트시장의 데이터의 통계량을 <표 3>에 요약되어 있다. 그리고 앞서의 연구방법에서 설정한 모델들을 추정한 결과는 <표 4>와 <표 5>에 나타나 있다. <표 4>는 2000년 10월의 자료로부터 얻어진 결과이며 <표 5>는 1999년 1월의 자료로부터 얻어진 것이다.

각 기간별로 두개씩 추정된 모델들의 이분산 여부와 다중공선성의 심각성 여부를 판별하기 위해 White 테스트와 VIF (Variance Inflation Factor)가 이용되었다. 분석결과 결과 4개의 모델 모두에서 이분산의 존재가 발견되어 <표 4>와 <표 5>에서는 이분산에 로우버스트한 화이트 통계량이 추정되어 보고되었다 (White, 1980). 4개의 모델에서의 최고 VIF는 모두 2.1 미만으로 다중공선성은 그다지 심각하지 않은 것으로 드러났다.

추정결과를 살펴보면 4개의 모델은 각각 71%에서 79% 사이의 주택가격의 변량을 설명하는 것으로 나타났다. F 테스트 결과는 통계적으로 유의미하여 이 모델들이 모두 통계적으로 받아들여질 수 있음을 보여주고 있다.

2000년 10월을 대상으로 하는 <표 4>의 버퍼링 모델과 행정동 모델의 추정결과를 보면 건물면적(AREA), 건물나이(AGE), 단지의 전체가구수(TUNIT), 지하철역으로부

<표 3> 사용된 변수의 통계량

Variables	Obs	Mean	Std. Dev.	Min	Max
2000년 10월 가격	182.00	18223.63	7557.47	6500.00	45000.00
1999년 1월 가격	182.00	15447.25	6494.98	6000.00	36000.00
AREA	182.00	27.88	9.33	7.00	54.00
TUNIT	182.00	861.48	756.14	68.00	3000.00
AGE	182.00	13.04	5.77	2.00	21.00
DSUB	182.00	480.34	203.07	96.00	853.00
DGREEN	182.00	488.54	419.70	18.00	1872.00
RIVER	182.00	27696.80	93454.68	0.00	513270.00
PBUFFER	182.00	18229.80	16887.20	112.22	78098.37
PDONG	182.00	8156.48	10059.12	0.00	72419.00

3) Buffer Size = (평균면적/3.14)^{0.5}

해도닉 모델 추정시 GIS 공간분석기능에 의해 생성된 근린변수의 기여도에 대한 연구

<표 4> 회귀분석결과 – 2000년 10월

	버퍼링			행정동		
	Coef.	t		Coef.	t	
AREA	697.7606	17.3830	***	700.4722	16.6510	***
AGE	122.1714	1.7600	*	127.5265	1.8620	*
TUNIT	1.4366	2.4930	**	1.5970	2.8540	***
DSUB	-6.2884	-3.3830	***	-6.7209	-3.4570	***
DGREEN	-1.2560	-1.8900	*	-2.0688	-3.1000	***
RIVER	-0.0031	-1.0290		-0.0045	-1.4870	
PBUFFER	-0.0640	-2.9520	***			
PDONG				-0.0635	-2.1010	**
_cons	826.7754	0.4260		538.6970	0.2700	
	R-squared=0.71 F(7, 174)=55.40			R-squared=0.71 F(7, 174)=51.38		

***: 1% 수준에서 유의함, **: 5% 수준에서 유의함, *: 10% 수준에서 유의함

터의 거리(DSUB), 그린벨트로부터의 거리 (DGREEN), 주변지역의 유흥시설의 상면적(PBUFFER, PDONG) 등이 통계적으로 유의미하고 기대되었던 부호를 보여주고

있다. 반면 1킬로미터 내 수변공간의 면적을 나타내주는 RIVER는 주택가격과 부의 관계를 보이고 있으나 통계적으로 유의하지 않다. 통계적으로 유의한 변수 가

<표 5> 회귀분석 결과 – 1999년 1월

	버퍼링			행정동		
	Coef.	t		Coef.	t	
AREA	633.4077	21.4920	***	635.8795	20.5940	***
AGE	21.3862	0.4060		28.6049	0.5430	
TUNIT	0.9058	2.1800	**	1.0458	2.6200	***
DSUB	-4.6976	-3.5200	***	-5.0082	-3.6050	***
DGREEN	-0.7126	-1.4500		-1.3441	-2.7400	***
RIVER	-0.0027	-1.0350		-0.0037	-1.4640	
PBUFFER	-0.0496	-3.0580	***			
PDONG				-0.0437	-2.0600	**
_cons	333.8153	0.2430		-2.8914	-0.0020	
	R-squared=0.79 F(7, 174)=76.69			R-squared=0.78 F(7, 174)=70.39		

***: 1% 수준에서 유의함, **: 5% 수준에서 유의함, *: 10% 수준에서 유의함

손 철

운데 AREA, AGE, TUNIT 등이 주택가격과 정의 관계를 보여주고 있으며 DSUB, DGREEN, PBUFFER, PDONG 등이 부의 관계를 보여주고 있다. 이러한 결과는 대부분 서울지역을 대상으로 해도닉 모델을 추정한 선행연구와 일치하는 것이다. 추정된 파라미터의 절대적인 크기는 선행연구들과 다소 차이를 보이고 있다.

한 예를 들면 지하철역으로부터의 거리가 집값에 미치는 영향을 나타내는 DSUB 변수의 경우 모수추정치는 -6.3 ~ -6.7로 이는 주택의 위치가 가장 가까운 지하철 역으로부터 1 미터 멀어질 때 주택의 판매 가격이 6만3천원에서 6만7천원 정도 감소함을 나타낸다. 이는 비슷한 시기를 대상으로 한 시정개발연구원(2000)의 연구결과보다는 다소 큰 것이다. 그러나 시정개발연구원의 분석결과는 5호선이 통과하는 전체지역을 대상으로 한 것으로 강동구만을 대상으로 한 본 연구와의 절대적인 비교는 무의미하다. 본 연구에서의 주요관심대상인 근린의 유흥시설의 상면적은 1% 수준에서 아파트 가격에 통계적으로 유의한 부의 영향을 미치는 것으로 나타났다. 구체적인 파라미터의 크기를 살

펴보면 버퍼링 모델의 경우 -0.064이며 행정동 모델의 경우는 -0.0635로 두 모델간에 큰 차이는 없다. 이 결과는 근린의 유흥시설 상면적인 1제곱미터 증가할 때 아파트의 가격이 약 640원 정도 감소함을 나타낸다.

1999년 1월을 대상으로 한 <표 5>를 살펴보면 추정결과는 2000년 10월을 대상으로 하는 결과와 유사하게 나타나 있다. 다만 건물의 연령을 나타내는 AGE 변수가 통계적으로 유의하지 않는데 이는 재건축대상 아파트에 대한 투기적 수요가 경제적 침체기인 1999년 초반기에 악화되었기 때문으로 추측된다.

<표 6>은 2000년 10월과 1999년 1월을 대상으로 추정된 버퍼링 모델과 행정동 모델에 대해 J 검정을 실시한 결과를 보고하고 있다. <표 6>을 살펴보면 세번째 열에서 δ^0 와 δ^1 의 값을 보고하고 있다. 이 값들은 두 기간에서 모두 버퍼링 모델이 행정동 모델에 비해 우월한 모델 즉 아파트가격의 변량을 보다 잘 설명해주는 모델임을 보여주고 있다. 버퍼링 모델의 경우 분석의 대상인 유흥시설이 해당 토지이용영역 안에 고르게 분포되었다는 가

<표 6> J 테스트 결과

기간	귀무가설/대안가설	t(p-value)	해석	결론
2000년 10월	버퍼링/행정동	-0.383 (0.702)	유의성이 없음	버퍼링모델이 우월함
	행정동/버퍼링	2.217 (0.028)	유의성이 있음	
1999년 1월	버퍼링/행정동	-0.413 (0.680)	유의성이 없음	버퍼링모델이 우월함
	행정동/버퍼링	2.017 (0.045)	유의성이 있음	

정을 이용했음에도 불구하고 행정동 모델에 비해 우월한 모델임으로 보인 것은 유흥시설이 발생하는 부의 외부효과가 행정동의 경계를 넘어서는 공간적 파급효과가 있으며 베퍼링 모델에서 사용된 지리정보시스템의 공간분석기능에 의해 생성된 근린변수가 이를 효과적으로 잡아냈기 때문으로 판단된다.

5. 결 론

본 연구에서 다룬 질문은 과연 지리정보시스템이 해도닉 모델을 추정하는 데 얼마나 기여를 할 수 있는가이다. 앞서 언급한 바처럼 지리정보시스템은 해도닉 모델을 추정하는 데 두 가지 측면에서 기여할 수 있다. 그 가운데 하나는 대용량의 공간적 데이터를 효율적으로 처리하게 해준다는 점이다. 현재 이루어지고 있는 해도닉 모델을 이용한 경험적 연구를 보면 수만 개의 주택 데이터를 사용한 것을 흔히 보는데 이는 지리정보시스템이 가진 정보처리능력을 이용하지 않고는 불가능한 것이다. 그러나 이러한 장점도 단지 수백 개의 데이터를 이용하는 경험적 연구에서는 커다란 것이라 할 수 없다. 단지 수백 개의 데이터가 존재하는 경우 시간이 더 걸릴뿐이지 과거에 사용되었던 종이지도와 자를 이용한 수작업이나 비공간적 소프트웨어의 이용을 통해서도 해도닉 모델의 추정이 가능한 것이다. 이에 반해 근린변수의 측정을 위해 지리정보시스템이 가진 특유한 공간분석기능이 이용되는 것은 다른 소프트웨어나 수작업에

의해 대체될 수 없는 것이다. 그러나 이러한 기능이 실제 해도닉 모델이 추정하는 데 적극적으로 이용되지 않고 있는 것은 공간분석기능을 통해 측정된 변수들의 우월성이 경험적으로 적극적으로 검토되지 않고 있기 때문이다.

본 연구의 기여는 지리정보시스템의 공간분석기능에 의해 생성된 변수의 상대적 공헌도를 보다 폭넓게 경험적으로 평가하였다는데 있다. 본 연구는 지리정보시스템이 단순한 직선거리를 추정하는 것 이상으로 해도닉 모델의 질을 향상시킬 수 있는 잠재력을 가지고 있다는 것을 보이고 있으며 주택시장의 행태를 설명하는 이론적으로 보다 타당한 근린변수의 측정을 위해 보다 적극적으로 이용되어야 함을 실증적으로 보이고 있다. 이 결과는 해도닉 모델 추정시 단순한 직선거리를 측정하는데 있어서만 지리정보시스템의 공간분석기능을 이용하고 있는 우리나라의 연구자들에게도 유용한 정보라고 생각된다.

참고문헌

- 서울시정개발연구원, 2000, 지하철 5호선 건설 사후평가 연구.
Can, Ayse, 1998, "GIS and Spatial Analysis of Housing and Mortgage Markets.", Journal of Housing Research, 9, 1: pp. 61-86.
Davidson and Mackinnon, 1981, "Several Tests for Model Specification in the Presence of Alternative Hypotheses.", Econometrica, 49: pp. 781-793

손 철

- Geoghegan, Jacqueline, Lisa A. Wainger, and Nancy E. Bockstaal, 1997, "Spatial Landscape Indices in a Hedonic Framework: an Ecological Economics Analysis using GIS.", *Ecological Economics*, 23: pp. 251-264.
- Madalla, G. S., 1993, *Introduction to Econometrics*, Englewood, Cliff, Prentice Hall.
- Sawiki, David S. and Patrice Flynn, 1996, "Neighborhood Indicators: A Review of the Literature and Assessment of Conceptual and Methodological Issues.", *Journal of the American Planning Association*, 62, 2: pp. 165-183.
- Rodriguez, M., Sirmans, C. and Marks, A., 1995, "Using Geographic Information Systems to Improve Real Estate Analysis.", *Journal of Real Estate Research*, 10, 2: pp. 163-173.
- White, Hal, 1980, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.", *Econometrica*, 48, 4: pp. 817-838.