

표본선택편의를 고려한 오피스 매매가격 결정요인 분석 및 매매가격지수 산정*

Price Determinants and Transaction-Based Price Indices under Sample Selection Bias
in the Seoul Metropolitan Office Market

이 영 유 (Lee, Youngyoo)**

이 상 경 (Lee, Sangkyeong)***

< Abstract >

Price determinants and indices that based on the sample of transacted offices may not accurately represent a population due to sample selection bias. We analyse sample selection bias in estimating parameters of a hedonic price equation in the Seoul metropolitan office market. We use Heckman two-stage procedure(Heckman), maximum likelihood estimation(ML) and ordinary least square estimation(OLS) as analysis methods. We find the impact of sample selection bias by comparing results of Heckman and ML with OLS. Some Independent variables like building height, age, lot area etc are turned out statistically significant in Heckman and ML but some of them are not significant in OLS. We calculate office price indices based on Heckman, ML and OLS. Consequently, Heckman and ML price indices are a little different from OLS price index but not strongly.

주 제 어 : 표본선택편의, 오피스 매매가격 결정요인, 오피스 매매가격지수, 헤크만 2단계 모형

Keywords : Sample Selection Bias, Determinants of Office Price, Transaction-Based Office Price Indices, Heckman Two-Stage Procedure

* 이 논문은 이영유의 석사학위논문을 전면 수정·보완한 것임

** GKL(그랜드코리아레저), 경원대학교(현 가천대학교) 석사, daesungycn@naver.com

*** 가천대학교 도시계획학과 부교수, skylee@gachon.ac.kr(교신저자)

I. 서론

오피스 시장이 활성화되면서 시장 참여자들을 중심으로 매매가격의 움직임에 대한 관심이 점점 커지고 있다. 매매가격을 전망하기 위해서는 기초 연구에 해당하는 가격결정요인과 매매가격지수 연구가 선행되어야 한다. 그러나 이 같은 필요성에도 불구하고 이와 관련된 연구는 소수에 그치고 있는 실정이다. 국내 학술지에 발표된 연구로는 헤도닉 매매가지수를 추정한 이상경(2005)과 시변모수 매매가격지수를 다룬 이상경(2007), 반복매매가격지수를 다룬 이상경(2009)과 최성호 외(2010), 그리고 오피스 매매가격의 거래 특성 영향과 공간종속성을 다룬 박종기 외(2011) 정도가 있을 뿐이다. 이들 연구들은 모두 거래된 오피스만을 분석 대상으로 했다는 점에서 공통점을 가진다. 그러나 이 같이 거래되지 않은 오피스를 분석 자료에서 제외할 경우 표본선택편의(sample selection bias)를 야기할 수 있다는 점에서 주의가 필요하다.

표본선택은 모집단으로부터 자료들이 체계적 또는 비무작위(non-random)로 배제되는 경우로 표본선택편의는 이 같은 자료를 이용하여 회귀분석과 같은 통계분석을 수행할 경우 모수 추정치에서 발생하게 되는 편의를 말한다(남춘호, 1998; 예민규·이상경, 2011). 오피스 매매가격 연구에서 거래되지 않은 오피스가 제외되는 것은 거래된 오피스만이 가격 정보를 가지고 있기 때문이다. 그러나 이는 Heckman(1979)이 표본선택편의 발생 이유 중의 하나로 제시한 자료의 자기선택에

해당되며 오피스 매매가격 연구에서도 표본선택편의에 대한 고려가 필요함을 의미한다.

표본선택편의를 보정하기 위한 방법으로 여러 모형들이 개발되었는데 Heckman의 2단계 모형이 대표적인 것이라고 할 수 있다. 이 방법은 1970년대 후반 Heckman(1976, 1979)에 의해 제안된 것으로 프로빗 모형(probit model)과 보통최소자승 회귀분석(OLS regression)을 2단계로 결합한 것이다. Heckman 모형과 함께 최대우도법(maximum likelihood estimation)도 널리 이용되고 있는 데¹⁾, ML 모형에서는 Heckman 모형의 2단계 추정 대신 제 1단계인 프로빗 모형의 추정계수를 초기 값으로 하여 제 2단계 계수를 추정하게 된다(이성우 외, 2006). Heckman 모형과 ML 모형은 계량경제학과 사회학 등의 분야에서 임금 및 소득 추정과 관련하여 활발하게 이용되고 있다(김강호, 2009; 김우영, 2012; 남춘호, 1998; 박재민, 2011; 오욱찬, 2011; 원종욱, 1999).

오피스 시장을 대상으로 표본선택편의를 다룬 연구는 그다지 많지 않은 실정이다. 해외의 경우 Jud and Winkler(1999), Munneke and Slade(2000, 2001) 등이 있고 국내의 경우에는 예민규·이상경(2011) 정도가 있을 뿐이다. 그러나 예민규·이상경(2011)은 오피스 매매가격이 아닌 투자 행태를 분석한 연구로 가격결정 요인이나 가격지수를 다룬 연구는 아니다. 따라서 오피스 매매가격을 대상으로 표본선택편의를 분석한 국내 연구는 없다고 할 수 있다. 이 같은 인식 하에, 본 연구에서는 서울 오피스 시장을 대상으로 표본선택편의를 보정할 수 있는 Heckman 모형과 ML 모형을 이용하여 매매가격 결정요인을 분석하고 매매가격

1) 이들 방법 외에도 비모수적 방법인 커널 추정(kernel estimation)과 패널자료(panel data)를 이용한 추정 등이 있다(남춘호, 1998).

지수를 산정한다. 이 과정에서 전통적인 OLS 모형과의 비교를 통해 표본선택편의 보정이 갖는 통계적 의미를 분석하고자 한다.

II. 선행연구 고찰

국내 오피스 시장을 대상으로 매매가격 결정요인을 분석하거나 매매가격지수를 산정한 연구를 살펴보면 다음과 같다. 이상경(2005)은 시간더미변수 헤도닉모형을 이용하여 서울시 오피스 매매가격 결정요인과 매매가격지수를 분석하였다. 이 연구에서는 가격결정요인으로 입지권역, 지하철역 거리, 건축면적, 높이, 준공연수, 전용률, 인텔리전트빌딩 유무를 제시하였다. 이상경(2007)은 시변모수모형을 이용하여 서울시 오피스 매매가격 결정요인을 분석하고 매매가격지수를 산정하였다. 독립변수로 입지권역, 지하철역 거리, 건축면적, 층수, 준공연수, 매입용도 등을 사용하였으며, 매매가격지수로 라스파이레스지수, 파쉐지수, 피셔지수, 연쇄지수, 시간더미변수 헤도닉가격지수를 산정하고 그 결과를 비교하였다. 박종기 외(2011)는 거래경험과 투자특성 등의 가격영향 분석과 함께 공간계량경제모형을 통해 오피스 가격에 내재된 공간 종속성을 규명하였다. 한편 이상경(2009)과 최성호 외(2010)는 오피스매매가격 자료를 이용하여 반복매매지수를 산정하였는데 별도의 독립변수를 필요로 하지 않는 반복매매법의 특성상 가격결정요인을 규명하진 않았다. 이들 연구들은 거래된 오피스만을 분석대상으로 했다는 점에서, 즉 거래되지 않은 오피스를 제외했다는 점에서 공통점을 가진다. 이는 Heckman(1979)이 지적한 표본선택편의에 대한

고려가 없었다는 것을 의미한다.

부동산 가격과 관련하여 표본선택편의를 다룬 연구들은 주택분야에서 상대적으로 많이 수행되어 왔다. 헤도닉 가격지수를 다룬 Jud and Seaks(1994)와 반복매매가격지수를 다룬 Case, Pollakowski and Wachter(1991), Gatzlaff and Haurin(1997) 등이 대표적인 연구라고 할 수 있다. 주택 외 분야의 연구로는 도시지역 토지가격 결정요인을 분석한 Colwell and Munneke(1997)와 Yamamura(2003) 등이 있으며, 오피스 분야의 연구로는 Jud and Winkler(1999), Munneke and Slade(2000, 2001) 등이 있다. 오피스 분야 연구는 주택분야에 비해 상대적으로 적은 실정이다. 오피스 분야 연구를 살펴보면 다음과 같다. Jud and Winkler(1999)는 미국 샬럿 지역의 상업용 부동산과 오피스를 대상으로 가격지수를 산정하였다. 이 연구에서는 Heckman 모형을 이용하여 표본선택편의를 보정하였다. 분석 결과 상업시설의 경우에는 표본선택편의 보정효과가 크지 않았으나 오피스의 경우에는 보정효과가 상당한 것으로 나타났다. Munnke and Slade(2000)는 미국 피닉스 지역의 오피스 매매가격을 대상으로 Heckman 모형과 OLS 모형을 적용하고 시간더미 헤도닉 가격지수를 산정하였다. 분석 결과 OLS 가격지수와 Heckman 가격지수의 차이는 그다지 크지 않은 것으로 나타났다. Munnke and Slade(2001)는 미국 피닉스 지역의 오피스 매매가격을 대상으로 시변모수법을 이용하여 연쇄지수, 라스파이레스지수, 파쉐지수를 산정하였다. Heckman 모형을 이용하여 표본선택편의 조정 전후를 비교한 결과 연쇄지수와 라스파이레스지수는 표본선택편의 조정 전 지수의 95% 신뢰구간 내에 있지만 파쉐지수는 이 범위를 벗어나는 것으로 분석

되었다.

예민규·이상경(2011)은 오피스 매매가격을 분석한 연구는 아니지만 국내 오피스 시장을 대상으로 한 최초의 표본선택편의 연구라는 점에서 주목할 필요가 있다. 이 연구에서는 거래된 오피스만을 대상으로 하는 투자주체(투자방식) 결정 모형이 표본선택편의를 발생시킨다고 보고 대안으로 표본선택 이변량 프로빗 모형을 분석에 이용하였다. 여기서는 투자결정과 투자주체(투자방식) 결정을 일련의 연속된 상황으로 가정하고, 전체 오피스 대상 투자결정 프로빗 모형과 거래 오피스 대상 투자주체(투자방식) 결정 프로빗 모형의 결합을 시도하였다. 분석 결과 표본선택편의를 고려하는 이변량 프로빗 모형의 경우 통계적으로 유의한 것으로 확인되었다. 본 연구에서는 예민규·이상경(2011)에 사용된 투자결정 프로빗 모형을 Heckman 모형의 1단계에서 이용하게 되는 데, 이 점에서 본 연구는 예민규·이상경(2011)의 연구 성과를 계승하는 후속 연구의 성격을 지닌다고 할 수 있다.

Ⅲ. 분석자료 및 모형 설정

1. 분석 자료

연구에 사용된 자료는 전체 오피스 현황 자료와 거래 오피스 현황 자료로 구분된다. 이는 예민규·이상경(2011)의 연구 자료를 기본으로 분석에 필요한 변수들을 추가한 것이다. 전체 오피스 현황 자료는 2010년도에 작성된 것이다. 이 자료는 메이트 플러스 홈페이지에서 제공하는 오피스 정보를 기본으로 하여 건축물 대장과 서울시

GIS 포털 시스템을 이용하여 구축한 것이다. 거래 오피스 현황 자료는 2000년 1분기부터 2010년 3분기까지 거래된 오피스를 대상으로 하고 있으며, 이상경(2009)의 연구 자료에 신영에셋 등 오피스 관리회사들이 제공하는 정보들을 보강한 것이다. 여기에는 전체 오피스 현황 자료에는 없는 거래특성 변수들이 다수 포함되어 있다. 연구의 공간적 범위는 서울시 전역과 경기도 성남시 분당구이며, 전체 오피스 현황 자료는 825개, 거래 오피스 현황 자료는 320개로 구성되어 있다. 오피스 거래 선택을 분석하기 위해 프로빗 모형을 이용하는 Heckman 모형과 ML 모형에서는 이들 두 자료가 통합되어 이용되지만 매매가격만을 분석하게 되는 OLS 분석에서는 거래 오피스 자료만 이용된다.

2. 모형

오피스 시장을 구성하고 있는 전체 오피스 중 일부가 거래되었다면 거래되지 않은 오피스도 있게 된다. 거래된 오피스는 투자자들의 의사결정에 따른 선택 결과라고 할 수 있다. 따라서 거래 선택에 대한 고려가 필요한 데 이를 무시하고 거래된 오피스만을 대상으로 OLS 분석을 수행할 경우 모수 추정 결과가 모집단의 특성을 대표하지 못하는 표본선택편의 문제가 발생하게 된다. Heckman(1979)에 의하면, 이를 보정하기 위해서는 일종의 보정변인인 λ 를 OLS 모형에 투입해야 한다. 보정변인 λ 는 다음과 같은 논의 전개를 통해 도출되게 된다.

오피스 매매가격 모형에서 종속변수인 매매가격 P 를 로그 변환할 경우 (1)식과 같이 나타낼 수 있으며, 거래 선택을 분석하는 프로빗 모형은

(2)식과 같이 나타낼 수 있다. 참고로 이하의 식에서는 단순화 차원에서 개별 오피스를 의미하는 하첨자 i 를 생략하고 논의를 진행하고자 한다.

$$\ln P = X\beta + v, \quad v \sim N(0, \sigma_v^2) \quad (1)$$

$$Z^* = W\alpha + \mu, \quad \mu \sim N(0, \sigma_\mu^2) \quad (2)$$

$$\text{Cov}(v, \mu) = \rho\sigma_v\sigma_\mu \quad (3)$$

여기서 α, β 는 모수(parameter) 벡터이며, v, μ 는 오차항(error term), X, W 는 독립변수 벡터이다. 여기서 Z^* 는 관측되지 않고 오직 Z 만이 관측되게 되며 이는 (4)식과 같이 표현될 수 있다(남춘호, 1998).

$$Z = \begin{cases} 1 & \text{if } Z^* > 0 \\ 0 & \text{if } Z^* \leq 0 \end{cases} \quad (4)$$

P 는 오직 거래된 경우에만 관측되므로 $Z=1$ 인 경우를 대상으로 매매가격함수를 유도하면 (5)식이 도출된다. 여기에 v, μ 에 대한 이변량정규분포 가정을 적용할 경우 (6)식이 도출된다. 여기서 ϕ 는 표준정규확률밀도함수이며, Φ 는 표준정규누적밀도함수이다(이성우 외, 2006).

$$\begin{aligned} E[\ln P | X, Z = 1] &= E[\ln P | X, Z^* > 0] \\ &= X\beta + E[v | \mu > -W\alpha] \quad (5) \end{aligned}$$

$$= X\beta + \rho\sigma_v\sigma_\mu \frac{\phi(W\alpha)}{\Phi(W\alpha)} \quad (6)$$

σ_μ 는 관측할 수 없는 Z^* 의 표준편차로 추정이 불가하므로 1로 정규화 되었다고 가정한다.

이때 선택편의를 나타내는 $\frac{\phi(W\alpha)}{\Phi(W\alpha)}$ 는 역밀의

비(inverse Mills ratio)로 이를 보정변인 λ 로 둘 경우 (7)식이 도출되게 된다(Munneke and Slade, 2000; 이성우 외, 2006).

$$\begin{aligned} \ln P &= X\beta + \rho\sigma_v\lambda \\ &= X\beta + \theta\lambda \end{aligned} \quad (7)$$

다만, λ 의 계수인 $\rho\sigma_v$ 는 일치성을 갖지 않는 관계로 추정 후 이를 조정하는 과정이 필요한데 본 연구에서 사용하는 통계 프로그램 LIMDEP의 경우 일치성을 갖는 추정치 θ 를 한 번에 제공하는 관계로 추가적 작업은 필요가 없게 된다(이성우 외, 2006).

ML 모형의 경우는 Heckman 모형의 (2)식인 $Z^* = W\alpha + \mu$ 로부터 프라빗 모형의 $Z=0$ 일 때와 $Z=1$ 일 때의 우도함수를 각각 추정한다. 이에 근거하여 (8)식과 같은 최대우도함수를 도출할 수 있으며 이때 α, β, ρ 및 σ_v 의 추정치들이 구해진다(이성우 외, 2006).

$$\begin{aligned} \sum \ln L &= \sum_{Z=0} \ln \Phi(-W\alpha) - \sum_{Z=1} (\ln \sqrt{\pi} + \ln \sigma_v) \\ &\quad - \sum_{Z=1} \ln \Phi(-W\alpha) + \frac{(P - X\beta)^2}{2\sigma^2} \quad (8) \\ &\quad + \sum_{Z=1} \ln \Phi[W\alpha + \frac{\rho(P - X\beta)}{\sigma_v(1 - \rho^2)^{1/2}}] \end{aligned}$$

ML 모형을 통한 표본선택편의의 보정은 Heckman 모형처럼 1, 2단계로 구분하여 추정하지 않고 한 번에 통합하여 추정하게 된다. 따라서 보정변인 λ 를 발견할 수 없으며, 이점에서 Heckman 모형과 차이가 나게 된다.

3. 변수 설정

본 연구에서는 모형 간 비교를 위해 Heckman 모형과 ML 모형, OLS 모형의 독립변수들을 단계별로 동일하게 설정한다. Heckman 모형의 1단계에서는 종속변수로 사건 발생 확률을 사건이 발생하지 않을 확률로 나눈 값인 승산(odds)을 프로빗 변환(probit transformation)하여 투입하게

된다. 본 연구에서는 사건을 오피스 거래로 설정하고 더미변수 형태로 자료를 구축하였다. 거래된 오피스의 경우에는 1의 값을 부여하고 거래된 적이 없는 경우에는 0의 값을 부여하였다. 독립변수의 경우 예민규·이상경(2011)과의 결과 비교를 위해 동일한 변수를 도입하였다. 입지특성 변수로 입지권역, 이면도로, 접면도로 수, 교차로 인접 여부를 선정하고 이들을 더미변수 형태로

〈표 1〉 변수내역 및 기초통계량

변수		변수 내역		전체 오피스		거래오피스	
				평균	표준 편차	평균	표준 편차
종속변수	프로빗변환(거래 선택)	거래된 적이 있으면 1, 아니면 0					
	로그(매매가격)	ln(오피스 실제 매매가격, 천원)					
독립 변수	권역	도심권	도심권 1, 아니면 0	0.25	0.44	0.30	0.46
		여의도권	여의도권 1, 아니면 0	0.12	0.33	0.16	0.37
		분당권	분당권 1, 아니면 0	0.03	0.18	0.07	0.25
		기타지역	기타지역 1, 아니면 0	0.21	0.41	0.11	0.32
	입지 특성	이면도로	이면도로 입지하면 1, 아니면 0	0.22	0.41	0.29	0.45
		접면도로 수 더미	접면도로 3개 이상 1, 아니면 0	0.37	0.48		
		접면도로 수	접면도로 수(개)			2.16	0.83
		교차로 인접	교차로 인접 1, 아니면 0	0.15	0.36	0.11	0.31
	구조 특성	지하철역 거리	최근접 지하철역 도보거리(분)	6.14	4.24		
		용도지역	상업지역 1, 아니면 0			0.79	0.41
		대지면적	대지면적(m ²)	4349	11745		
		로그(대지면적)	ln(대지면적)			7.79	0.75
		층수	오피스 빌딩의 최고 층수	15.45	7.05		
		로그(층수)	ln(오피스 빌딩의 층수)			2.71	0.43
건축구조		철골철근콘크리트 1, 아니면 0	0.57	0.50	0.68	0.47	
노후도		준공연수(년)			14.57	8.64	
거래 특성		부분거래	부분거래 1, 아니면 0			0.17	0.37
		거래경험	거래 경험 2회 이상 1, 1회 0			0.40	0.49
	거래 시점	2001년	2001년 거래 1, 아니면 0			0.06	0.23
		2002년	2002년 거래 1, 아니면 0			0.10	0.30
		2003년	2003년 거래 1, 아니면 0			0.08	0.27
		2004년	2004년 거래 1, 아니면 0			0.10	0.30
		2005년	2005년 거래 1, 아니면 0			0.09	0.29
		2006년	2006년 거래 1, 아니면 0			0.05	0.22
		2007년	2007년 거래 1, 아니면 0			0.08	0.27
		2008년	2008년 거래 1, 아니면 0			0.12	0.32
2009년		2009년 거래 1, 아니면 0			0.17	0.37	
2010년	2010년 거래 1, 아니면 0			0.10	0.30		

도입하였다. 접면도로 수의 경우 3개 이상 도로에 접하고 있으면 1, 그렇지 않을 경우에는 0의 값을 부여하였다. 여기에 접근성을 대표하는 변수로 지하철역 거리를 추가하고 가장 가까운 지하철역까지의 도보 시간거리를 측정하여 투입하였다. 구조특성 변수로는 대지면적, 층수, 건축구조 변수를 선정하였는데, 대지면적과 층수는 별다른 변환 없이 투입하지만 건축구조의 경우 철골철근콘크리트구조의 경우 1을 부여하고 그렇지 않을 경우에는 0을 부여하는 더미변수로 투입하였다.

Heckman 모형의 2단계에서는 보정변인 λ 를 모형에 도입하여 OLS 추정을 하게 된다. 이를 위해 종속변수로 오피스 매매가격을 설정하고 이를 자연로그로 변환하여 투입하였다. 독립변수들의 경우 선행연구들을 참조하여 입지특성, 구조특성, 거래특성으로 구분하여 투입하였다. 먼저, 입지특성 변수로 입지권역, 지하철역 거리, 이면도로, 접면도로수, 교차로 인접 여부, 용도지역 변수를 설정하였다. 입지권역 변수는 이상경(2005, 2007)의 연구에서 이미 사용되었는데, 이는 서울 오피스 시장이 도심권, 강남권, 여의도권, 분당권 같은 하위시장으로 구분된다는 인식을 반영한 것이다. 이면도로, 접면도로 수, 교차로 인접여부 변수는 박종기 외(2011)의 연구에서 변수로 사용되었다. 이면도로 변수는 더미변수로 이면도로에 입지하고 있을 경우 1, 아니면 0의 값을 부여하였다. 본 연구에서는 이면도로를 폭 25m 미만인 도로로 정의하였다. 접면도로 수는 오피스가 접한 도로수를 산정한 것이다. 교차로 인접 변수는 더미변수로 25m 이상의 도로가 서로 교차하는 지점에 해당 오피스가 접하고 있으면 1, 그렇지 않으면 0의 값을 부여하였다. 용도

지역 변수의 경우 더미변수로 상업지역에 입지할 경우 1을, 그 밖의 지역에 입지할 경우 0을 부여하였다.

구조특성 변수로 대지면적, 층수, 건축구조, 노후도 변수를 도입하였다. 오피스 규모를 대변하는 변수로 대지면적과 층수 변수를 설정하였는데 대지면적이 넓거나 층수가 높을수록 매매가격이 높을 것으로 전망된다. 건축구조 변수는 1단계 프로빗 모형과 동일하게 철골철근콘크리트구조에 1, 그 밖의 구조에 0을 부여한 더미변수 형태로 투입하였다. 고가 구조물인 철골철근콘크리트구조를 가진 오피스의 가격이 더 높게 형성될 것으로 전망된다. 노후도 변수는 빌딩의 준공연수를 산정하여 투입하였다. 리모델링한 오피스의 경우 리모델링 시점을 준공시점으로 하여 노후도를 산정하였다. 노후도에 비례하여 가격은 떨어질 것으로 전망된다.

거래특성 변수로 부분거래, 거래경험, 거래시점 변수를 도입하였다. 부분거래 변수는 오피스의 일부만 거래되었을 경우에 1, 전체가 거래된 경우에는 0의 값을 부여하였다. 부분거래 변수는 이상경(2007)과 박종기 외(2011)에서 전체 거래보다 가격을 낮추는 요인으로 분석되었다. 거래경험 변수는 더미변수로 거래경험이 2회 이상인 경우에 1을, 거래경험이 1회인 경우에는 0을 부여하였다. 거래시점 변수는 매매가격지수 산정하기 위해 도입한 것으로 본 연구에서는 2000년을 기저로 2001년부터 2010년까지 각 연도를 더미변수 형태로 도입하였다.

IV. 분석 결과

1. 추정 결과

<표 2>는 Heckman 모형과 ML 모형, OLS 모형을 적용한 결과를 정리한 것이다. 먼저 표본선택편의 보정의 타당성과 관련하여 Heckman 모형의 보정변인 λ 의 계수와 ML 모형의 ρ , σ 를 보면 모두 1% 수준에서 유의한 것으로 나타났다. 이는 표본선택편의를 보정하기 위해 Heckman 모형과 ML 모형을 적용한 것이 타당하다는 것을 의미한다. 한편, Heckman 모형의 보정변인 λ 의 계수 θ 값은 -0.89이고 ML 모형의 보정변인 $\rho\sigma$ 값은 -0.35($\rho=-0.75$, $\sigma=0.47$)로 계산되었다. (7)식에 따르면 이들 두 보정변인 값이 일치해야 하지만 차이가 발생한 것은 본 연구에서 사용한 Limdep 프로그램으로 인한 것이다. Limdep에서는 최종적으로 일치추정된 θ 를 다시 계산하기 때문이다. 두 값의 차이는 있지만 부호가 일치하므로 결과에서의 불일치 문제는 없는 것으로 판단된다. Heckman 모형과 OLS 모형의 설명력을 비교한 결과 Heckman 모형의 Adjusted R^2 가 0.8173으로 OLS 모형의 0.8083보다 높게 나타났다. 두 모형간의 설명력 차이가 크지는 않지만 이 결과로부터 표본선택편의 보정변인에 의해 OLS 모형의 설명력이 증대될 수 있다는 것을 알 수 있다.

Heckman 모형의 1단계 프로빗 모형 적용 결과, 분당권, 기타지역, 지하철역 거리, 이면도로 인접 여부, 접면도로수, 교차로 인접 여부, 대지면적, 건축구조, 빌딩높이 변수가 유의하게 나타났다. 반면 도심권, 여의도권 변수는 10% 유의수준에서 유의하지 않은 것으로 나타났다. 이 같은 결과는 ML 추정에서도 같게 나타났다. Heckman과

ML 추정에서 유의하게 나타난 변수들의 계수 부호는 모두 같게 나타났다. 강남권과 비교하여 분당권의 거래 확률은 높지만 기타지역의 거래 확률은 낮은 것으로 나타났다. 지하철역 거리가 가까울수록, 이면도로에 위치할수록, 접면도로가 3개 이하일수록, 대지면적이 작을수록, 철골철근콘크리트 구조일수록, 빌딩 높이가 높을수록 거래 확률은 높아지는 것으로 나타났다. Heckman 모형의 1단계 프로빗 모형 적용 결과는 동일한 오피스 표본을 이용한 예민규·이상경(2011)의 일변량 프로빗 모형 적용 결과와 일치하는 것으로 나타났다. 이는 예민규·이상경(2011)에서 사용된 투자 의사 결정 일변량 프로빗 모형이 Heckman 모형의 1단계 프로빗 모형과 동일한 구조를 갖기 때문이다.²⁾

Heckman 모형 2단계에서는 거래 오피스를 대상으로 OLS를 적용하게 되는 데 이 과정에서 1단계 프로빗 모형에서 도출된 λ 를 보정변수로 활용하게 된다. 분석 결과 여의도권, 분당권, 이면도로, 접면도로수, 교차로 인접 여부, 용도지역, 대지면적, 층수, 노후도, 부분거래, 거래경험 변수가 통계적으로 유의한 것으로 나타났다. 반면 도심권, 기타지역, 건축구조 등은 유의하지 않은 것으로 나타났다. 이 결과를 표본선택편의를 보정하지 않은 OLS 분석결과와 비교하면, Heckman 모형에서 유의하였던 변수 중 이면도로, 접면도로 수가 유의하지 않게 나타났으며 반대로 유의하지 않았던 기타지역, 건축구조 변수, 그리고 2004년 시간더미 변수가 유의하게 나타났다.

ML 추정 결과는 Heckman 추정 결과와 약간

2) 예민규·이상경(2011)에서 제시된 표본선택 이변량 프로빗 모형의 적용 결과는 본 연구 결과와 일치하지 않은 것으로 나타났다. 이는 모형들이 독립적으로 추정되는 일변량 프로빗 모형과 달리 이변량 프로빗 모형의 경우 오차항들이 서로 결합되어 추정되기 때문이다.

〈표 2〉 추정 결과

구분		Heckman 모형		ML 모형		OLS 모형		
		거래 선택(프로빗 모형)		거래 선택				
		계수	t값	계수	t값			
절편		-0.51	-3.05***	-0.52	-3.25***			
입지	권역	도심권	0.19	1.55	0.18	1.51		
		여의도권	0.23	1.51	0.21	1.33		
		분당권	1.18	4.09***	1.16	3.03***		
		기타지역	-0.38	-2.82***	-0.36	-3.00***		
	이면도로	0.31	2.67***	0.32	2.71***			
	접면도로수	-0.30	-2.98***	-0.24	-2.62***			
	교차로 인접	-0.30	-2.09**	-0.30	-2.25**			
구조	지하철역 거리	-0.03	-2.23**	-0.03	-3.09***			
	대지면적	-0.00	-2.17**	-0.00	-2.63***			
	건축구조	0.02	1.77**	0.02	2.29**			
	층수	0.43	3.82***	0.40	3.64***			
N		825						
2LL		1101.85						
Chi-squared		118.77***						
구분		매매가격(편의보정 OLS)		매매가격(편의보정)		매매가격(무보정 OLS)		
		계수	t값	계수	t값	계수	t값	
절편		10.93	19.63***	10.30	32.63***	10.08	33.51***	
입지	권역	도심권	-0.07	-0.64	0.01	0.11	0.06	1.05
		여의도권	-0.36	-2.71**	-0.28	-2.73***	-0.23	-3.10***
		분당권	-1.05	-3.45***	-0.67	-4.07***	-0.40	-3.74***
		기타지역	-0.24	-1.58	-0.39	-4.81***	-0.48	-6.14***
	이면도로	-0.19	-1.70*	-0.10	-1.45	-0.02	-0.42	
	접면도로수	0.11	2.27**	0.06	2.00**	0.04	1.38	
	교차로인접	0.34	2.44**	0.24	2.93***	0.19	2.49**	
건물	용도지역	0.15	1.88*	0.15	2.12**	0.14	2.01**	
	대지면적	0.72	12.74***	0.70	18.56***	0.65	17.58***	
	층수	0.61	4.47***	0.70	8.87***	0.80	10.01***	
	건축구조	-0.03	-0.19	0.14	1.78*	0.25	4.29***	
거래	시점	노후도	-0.01	-1.81*	-0.01	-2.27**	-0.01	-2.00**
		부분거래	-0.55	-6.84***	-0.55	-8.17***	-0.58	-8.85***
	거래시점	거래경험	0.12	1.97**	0.12	2.12**	0.14	2.68***
		2001년	-0.04	-0.23	-0.05	-0.41	-0.04	-0.32
		2002년	0.07	0.49	0.06	0.55	0.08	-0.32
		2003년	0.10	0.64	0.09	0.69	0.12	0.11
		2004년	0.21	1.52	0.21	1.99**	0.21	1.86*
		2005년	0.34	2.33**	0.36	3.31***	0.37	3.14***
		2006년	0.18	1.10	0.16	1.24	0.18	1.32
		2007년	0.72	4.59***	0.71	6.36***	0.71	5.66***
2008년	0.81	5.71***	0.83	8.06***	0.84	7.29***		
2009년	0.58	4.19***	0.57	5.59***	0.58	5.15***		
2010년	0.63	4.32***	0.64	5.79***	0.62	5.22***		
λ		-0.89	-2.62***					
σ				0.47	9.46***			
ρ				-0.75	-5.74***			
N		320		825		320		
2LL		234.64		1267.46				
Adj. R ²		0.8173				0.8083		

주) *, **, ***는 각각 10%, 5%, 1%에서 유의함을 의미

다르게 나타나고 있는 데, Heckman 모형에서 유의하지 않았던 기타지역 변수와 건축구조 변수가 유의한 반면 유의했던 이면도로 입지 변수는 유의하지 않게 나타났다. ML 추정 결과를 OLS 추정 결과와 비교하면 ML 추정에서 유의했던 접면도로 수는 더 이상 유의하지 않게 나타났다.

정리하면, 유의수준 10%를 기준으로 Heckman 모형과 OLS 모형 추정 결과를 비교한 결과 이면도로를 비롯한 5개 변수에서 유의 여부가 다르게 나타났고, ML 모형과 OLS 모형의 비교에서는 접면도로 수 변수에서 유의 여부가 다르게 나타났다. 이 같은 결과는 독립변수의 통계적 유의 여부를 엄밀하게 판정하기 위해서도 표본선택편의에 대한 보정이 필요하다는 것을 의미한다.

세 모형에서 유의하게 나타난 변수들의 계수 부호를 통해 결과를 해석하면 다음과 같다. 입지 권역과 관련하여 여의도권과 분당권 변수는 세 모형 모두에서 유의하게 나타났는데, 이들 지역의 가격은 모두 강남권보다 낮은 것으로 나타났다. 기타지역 변수의 경우 ML 모형과 OLS 추정에서만 유의하게 나타났는데 마찬가지로 강남권에 비해 낮게 나타났다. 그리고 오피스가 노후화될수록, 이면도로에 입지할수록, 부분거래일수록 가격은 낮아지는 것으로 나타났다. 그러나 접면도로가 많을수록, 교차로에 인접할수록, 상업지역에 입지할수록, 대지면적이 클수록, 층수가 높을수록, 거래경험이 2회 이상일수록 매매가격은 높아지는 것으로 나타났다. 건축구조 변수의 경우 ML 모형과 OLS 모형에서만 유의하며 철골철근콘크리트구조의 오피스가 여타 구조에 비해 더 높은 가격을 형성하고 있는 것으로 나타났다. Heckman 모형과 ML 모형, OLS 모형의 추정 결과를 정리하면, 모형별로 유의한 변수들은 약간씩 다르지만 세 모형 모두에서 통계적으로 유의한 변수들의 부호가 다른 경우는 없는 것으로 나

타났다.

2. 지수산정 결과

Heckman 모형과 ML 모형, OLS 모형의 시간 더미변수의 계수를 이용하여 2000년부터 2010년까지의 매매가격지수를 산정하였다. <표 3>은 이를 정리한 것이고 <그림 1>과 <그림 2>는 이 결과를 도시한 것이다. 결과적으로 3개 모형의 지수들은 유사한 움직임을 보여주고 있다. 매매가격지수는 2001년부터 2005년까지 계속 상승하다가 2006년 들어 14~18% 정도 급락하였다. 그러나 곧바로 반전하여 2007년의 경우 69~73% 정도 급등하였다. 글로벌 금융위기가 발생한 직후인 2009년에 다시 20~23% 정도 급락하였으며 다음 해인 2010년 다시 4~7% 내외의 반등이 있었다. 2000년대 들어 서울 오피스시장은 급격한 팽창을 경험하였는데, 이 같은 결과는 가격의 변동성도 그만큼 커졌다는 것을 알려 준다. 정리하면, 가격지수의 변동률 평균은 8.7%~9.1%로 물가상승률이나 금리보다 높은 수준이며, 변동성을 의미하는 표준편차도 24.7~26.3%로 상당히 높게 나타나고 있다.

<그림 1>과 <그림 2>는 지수 간 비교를 위해 OLS 지수를 중심으로 95% 신뢰구간을 설정하고 Heckman 지수와 ML 지수를 각각 도시한 것이다. 그 결과 표본선택편의를 보정한 두 지수 모두 OLS 지수의 95% 신뢰구간 내에 있는 것으로 나타났다. 이 같은 결과는 선택편의에 대한 보정이 이루어지고 이를 이용한 지수가 산정되더라도 지수 값의 차이가 통계적으로 유의하지 않을 수 있다는 것을 의미한다³⁾. 구체적으로 OLS 지수와 Heckman 지수는 0.03~7.31p 정도 차이가 나며,

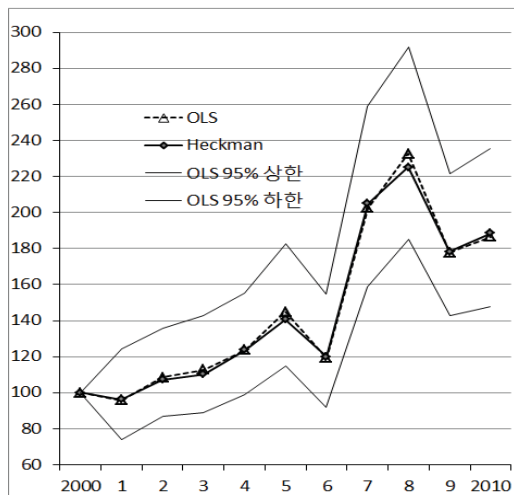
OLS 지수와 ML 지수는 0.44~3.27p 정도 차이가 난다. 지수 값들의 차이가 작은 것은 지수들 간의 상관관계 분석에서도 확인되고 있는 데,

OLS 지수와 Heckman 지수의 피어슨 상관계수 ρ 는 0.9982, OLS 지수와 ML 지수의 상관계수는 0.9991로 매우 높게 나타나고 있다. 이 같은 결

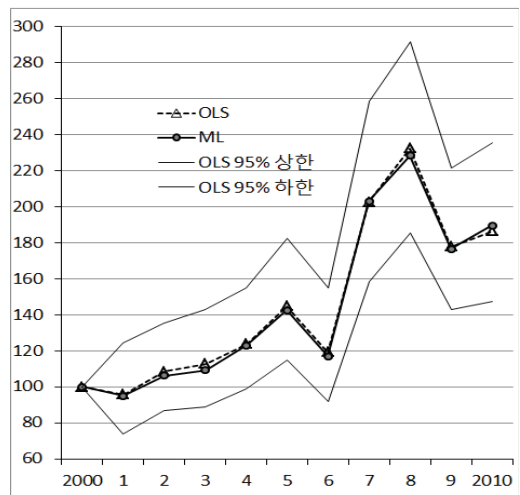
〈표 3〉 모형별 오피스 매매가격 지수 산정

년도	표본선택 무보정 OLS		Heckman 모형		ML 모형	
	가격지수	변동률	가격지수	변동률	가격지수	변동률
2000	100.00		100.00		100.00	
2001	95.87	-4.13%	96.35	-3.65%	94.96	-5.04%
2002	108.52	13.19%	107.25	11.31%	106.35	12.00%
2003	112.69	3.85%	110.18	2.73%	109.34	2.80%
2004	123.81	9.87%	123.84	12.40%	123.15	12.63%
2005	144.94	17.06%	140.73	13.64%	142.65	15.84%
2006	119.25	-17.72%	120.07	-14.68%	117.01	-17.97%
2007	202.59	69.88%	204.99	70.73%	203.19	73.65%
2008	232.59	14.80%	225.28	9.89%	228.58	12.50%
2009	177.87	-23.53%	178.27	-20.87%	176.62	-22.73%
2010	186.48	4.84%	188.56	5.77%	189.78	7.45%
연평균 변동률		8.81%		8.73%		9.11%
표준편차		25.41%		24.75%		26.32%
OLS지수와 ML 지수의 상관계수(ρ)				0.9982		0.9991

〈그림 1〉 OLS 지수와 Heckman 지수 비교



〈그림 2〉 OLS 지수와 ML 지수 비교



3) 이는 가격지수의 신뢰구간 폭과 밀접한 관계가 있다. 익명의 심사위원이 지적했듯이, 서울 오피스 시장의 경우 거래량이 적은 얇은 시장(thin market)에 해당되기 때문에 분석대상 자료의 크기가 작고 따라서 신뢰구간의 폭이 상대적으로 크게 될 수 있다.

과는 Heckman 모형을 이용하여 오피스 가격지수를 산정한 연구들인 Jud and Seaks(1994), Munnke and Slade(2000)에서도 나타난다는 점에서, 그리고 표본선택편의 보정이 상대적으로 미세 조정 성격을 갖는다는 점에서 본 연구에서만 나타나는 특이 현상은 아닌 것으로 판단된다.

V. 결론

본 연구에서는 서울 오피스 시장을 대상으로 표본선택편의를 보정할 수 있는 Heckman 모형과 ML 모형을 이용하여 매매가격 결정요인을 분석하고 매매가격지수를 산정하였다. 이 과정에서 표본선택편의를 보정하지 않은 전통적인 OLS 모형과의 비교 분석을 통해 표본선택편의 보정이 갖는 통계적인 의미를 분석하였다.

표본선택편의를 보정할 수 있는 Heckman 모형과 ML 모형을 적용한 결과, Heckman 모형의 보정변인 λ 의 계수와 ML 모형의 ρ 와 σ 가 모두 1% 수준에서 유의하게 나타났다. 또한 OLS 모형과 Heckman 모형의 Adjusted R^2 를 비교한 결과 Heckman 모형이 OLS 모형보다 설명력이 더 높은 것으로 나타났다. 이 같은 결과는 오피스 매매가격 모형 추정 시 표본선택편의를 보정하는 것이 통계적으로 유의하며 OLS 추정에 비해 더 높은 설명력을 가진 모형을 만들어낼 수 있다는 것을 의미한다.

분석 결과 표본선택편의에 대한 보정은 독립변수의 통계적 유의 여부에도 영향을 미치는 것으로 나타났다. 유의수준 10%를 기준으로 Heckman 모형과 OLS 모형을 비교한 결과 이면도료를 비롯한 5개 변수에서 통계적 유의 여부가

다르게 나타났고, ML 모형과 OLS 모형의 비교에서는 접면도로 수 변수에서 다르게 나타났다. 이는 독립변수의 유의 여부를 정확하게 판정하기 위해서도 표본선택편의에 대한 보정이 필요하다는 것을 의미한다.

Heckman 모형과 ML 모형, OLS 모형의 시간더미변수의 계수들을 이용하여 2000년부터 2010년까지의 매매가격지수를 산정한 결과 지수들의 변동률과 변동성은 상당히 크게 나타났다. 연평균 변동률이 8.7%~9.1%로 최근의 물가상승률이나 금리보다 높게 나타났으며, 변동성을 의미하는 표준편차도 24.7~26.3%로 높게 나타났다. OLS 가격지수와 Heckman, ML 가격지수를 비교한 결과, 두 가격지수 모두 OLS 가격지수의 95% 신뢰구간 내에 위치하고 있어 통계적으로 유의한 차이를 발견할 수는 없었다. 이는 서울 오피스 시장의 경우 거래건수가 적어 가격지수의 신뢰구간 폭을 줄이기 어렵다는 점과 표본선택편의 보정이 상대적으로 미세 조정 성격을 가진다는 점과 관계된다.

이 연구는 국내 오피스 매매가격 결정요인 연구에서 표본선택편의를 고려한 첫 시도라는 점에서 나름의 의미를 가진다. 한편 오피스 현황 자료 확보의 어려움으로 전수 자료를 구축하지 못한 점과 거래건수가 적은 얇은 시장(thin market)인 서울 오피스 시장의 특성을 충분히 고려하지 못한 점은 이 연구의 한계라고 하겠다.

논문접수일 : 2012년 10월 25일

논문심사일 : 2012년 11월 5일

게재확정일 : 2012년 12월 4일

참고문헌

1. 김강호, “학력과 직업훈련 참여가 임금에 미치는 효과”, 「농업교육과 인적자원개발」 제41권 제3호, 한국농산업교육학회, 2009, pp. 123-151
2. 김우영, “한국의 지역간 임금격차: 지역별 고용조사(RES)를 중심으로”, 「노동정책연구」 제12권 제1호, 한국노동연구원, 2012, pp. 1-28
3. 남춘호, “사회학적 연구에서 표본선택편의”, 「한국사회학」 제32집, 한국사회학회, 1998, pp. 99-136
4. 박재민, “대학 졸업생의 직업선택과 임금 수준”, 「기술혁신학회지」 제14권 1호, 기술혁신학회, 2011, pp. 22-39
5. 박종기·이상경·강승일, “오피스 가격 결정요인에 관한 연구-거래특성과 공간자기상관을 중심으로”, 「부동산연구」 제21집 제3호, 한국부동산연구원, 2011, pp. 91-108
6. 오욱찬, “장애인·비장애인의 취업확률 및 임금격차 분석”, 「사회보장연구」 제27권 제1호, 한국사회보장학회, 2011, pp. 1-25
7. 예민규·이상경, “표본선택 이변량프로빗모형을 이용한 오피스 투자 결정요인 분석”, 「국토계획」 제46권 제5호, 대한국토도시계획학회, 2011, pp. 331-340
8. 원종욱, “Heckman Selection Model에 의한 생업용자금수급자의 사업소득 결정요인 분석”, 「사회보장연구」 제15권 제2호, 한국사회보장학회, 1999, pp. 233-252
9. 이상경, “서울시 오피스 매매가격지수 개발에 관한 연구”, 「서울도시연구」 제6권 제4호, 서울연구원, 2005, pp. 121-134
10. 이상경, “시변모수법에 의한 오피스 매매가격지수 구축에 관한 연구”, 「국토계획」 제42권 제5호, 대한국토도시계획학회, 2007, pp. 233-245
11. 이상경, “오피스 투자 행태의 시공간적 특성에 관한 연구 - 투자 결정요인과 자본수익률을 중심으로”, 「서울도시연구」 제10권 제1호, 서울연구원, 2009, pp. 47-59
12. 이성우·윤성도·박지영·민성희, 「공간계량 모형 응용」, 박영사, 2006
13. 최성호·류강민·이건우·이창무, “반복매매 모형을 활용한 오피스 매매가격지수에 관한 연구”, 「국토계획」 제45권 제7호, 대한국토도시계획학회, 2010, pp. 119-131
14. Case, B., H. O. Pollakowski and S. M. Wachter, “On Choosing Among House Price Index Methodologies”, *Real Estate Economics*, Vol. 19, No. 3, 1991, pp. 286-307
15. Colwell, P. F. and H. J. Munneke, “The Structure of Urban Land Prices”, *Journal of Urban Economics*, Vol. 41, No. 3, 1997, pp. 321-336
16. Gatzlaff, D. H. and D. R. Haurin, “Sample Selection Bias and Repeated-Sales Index Estimates”, *Journal of Real Estate Finance and Economics*, Vol. 14, 1997, pp. 33-50
17. Heckman, J. J., “The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models”, *Annals of Economic and Social Measurement*, Vol. 5 No. 4, 1976, pp. 475-492
18. Heckman, J. J., “Sample Selection Bias as a

- Specification Error”, *Econometrica*, Vol. 47 No. 1, 1979, pp. 153-161
19. Jud, G. D. and T. G. Seaks, “Sample Selection Bias in Estimating Housing Sales Prices”, *The Journal of Real Estate Research*, Vol. 9, No. 3, 1994, pp. 289-298
20. Jud, G. D. and D. T. Winkler, “Price Indexes for Commercial and Office Properties: An Application of the Assessed Value Method”, *Journal of Real Estate Portfolio Management*, Vol. 5 No. 1, 1999, pp. 71-81
21. Munneke, H. J. and B. A. Slade, “An Empirical Study of Sample-Selection Bias in Indices of Commercial Real Estate”, *Journal of Real Estate Finance and Economics*, Vol. 21 No. 1, 2000, pp. 45-64
22. Munneke, H. J. and B. A. Slade, “A Metropolitan Transaction-Based Commercial Price Index: A Time-Varying Parameter Approach”, *Real Estate Economics*, Vol. 29 No. 1, 2001, pp. 55-84
23. Yamamura, Y., “Transaction-Based Land Price Indices for Commercial and Office Land During Bubble Economy in Japan”, *Kagawa University Economic Review*, Vol. 76 No.3, 2003, pp. 187-212
24. 메이트 플러스, <http://www.mateplus.net>
25. 서울시 GIS 포털시스템, <http://gis.seoul.go.kr>