

# 서울시 아파트 가격분위별 가격결정요인의 변동 추이에 관한 연구

Trends of Determinants of Housing Price in Seoul, Korea  
: A Quantile Regression Analysis

임재만 세종대학교 산업대학원 부동산학과 부교수  
Lim Jaeman Associate Professor, Dept. of Real Estate,  
Graduate School of Industry, Sejong Univ.  
(limjaeman@sejong.ac.kr)

## 목 차

- I. 서론
- II. 연구모형
  - 1. 특성가격모형
  - 2. 분위회귀접근법
- III. 자료와 실증결과
  - 1. 자료
  - 2. 2009년 회귀분석 결과
  - 3. 분위별 회귀계수의 연도별 변동 추이
- IV. 요약 및 결론

## I. 서론

주택시장은 한 나라의 경제 상태를 판단하는 데 매우 중요한 지표다. 주택 관련 논문 중에는 주택가격에 영향을 미치는 요인에 대해 설명하는 논문이 많다<sup>1)</sup>. 이들 연구는 주로 특성가격모형(hedonic price model)에 기초한 회귀분석으로 주택가격과 주택특성의 관계를 분석해 왔다. 특성가격모형은 주택가격이 주택을 구성하는 규모, 외관, 경과연수 등 내재적 특성과 학교나 쇼핑시설 접근성, 주택의 질, 근린 특성 등 외재적 특성에 기초하여 형성된다는 가정에 기초하고 있다. 그리고 소비자는 자신의 기대효용을 극대화하기 위해 주택특성의 묶음인 주택을 매수한다는 것이다.

Sirmans, Macpherson & Zietz(2005)는 특성가격모형을 적용한 125개의 실증연구를 분석한 결과 어떤 특성이 주택가격에 미치는 영향의 크기와 방향에 일관성이 없다고 보고했다. 예를 들면 방의 수에 대한 40개 실증논문 중 21개는 주택가격과 방수가 정의 관계에 있다고 했지만, 9개는 부의 관계, 그리고 나머지 10개는 유의적인 관계가 없다는 것이다. 문제의 핵심은 주택특성에 대한 모수(parameter)의 불확실성이다. 주택특성에 대한 모수의 불확실성은 대상지역, 모형의 설정, 자료의 차이에서 비롯한 것으로 설명할 수 있다.

Malpezzi(2003)는 주택특성이 주택가격의 모든 분포마다 동일하게 평가되지 않으며 소비자마다 주택특성을 다르게 평가할지 모른다고 주장했다. 특히 어떤 주택특성의 한계가치, 즉 탄력성은 주택가격의 범위에서 그 위치마다 다를 수 있다. 특성가격모형은 또한 주택특성의 공급함수가 특정 시점에 고정되어 있고 특성의 내재가격과 독립적

이라는 가정에 기초한다. 어떤 주택특성에 대한 우하향하는 수요함수와 수직의 공급함수가 교차하는 점에서 그 주택특성의 내재가격을 확인할 수 있는데, 이 내재가격은 특성가격모형으로 추정할 수 있다는 것이다. 주택시장에서 소비자가 모두 동일하다면 어떤 주택특성의 내재가격은 그 시장의 대표적인 소비자에게 내재하는 주택특성의 가격이 된다. 전형적인 최소자승법(Ordinary Least Square: OLS) 추정은 어떤 표본에서 주택가격의 평균을 최적으로 예측한 내재가격을 확인할 수 있으므로 대표적인 소비자 페러다임에 매우 잘 부합하는 모형이다.

그러나 예를 들어 주택시장에 소득과 금융에 제약이 있는 가난한 소비자와 소득과 금융에 제약이 없는 부유한 소비자, 두 소비자가 있다고 하자. 소득이 적어 축적된 부도 적고 금융기관에서 대출도 쉽지 않은 가난한 소비자는 고가 주택시장에 들어올 수 없을 것이다. 또한 부유한 소비자 역시 자신이 원하는 쾌적성이나 자신의 사회적 지위에 부정적인 영향을 주는 저가 주택을 매입하려고 하지 않을 것이다. 결국 두 시장은 분리되어 있다는 것이다.

주택시장의 분리(segmentation)는 가난한 소비자와 부유한 소비자가 가격이 다른 주택을 보유하고 있다는 것을 의미할 뿐만 아니라 어떤 주택특성에 대해 집단적인 호·불호를 발달시키고 있다는 것을 의미한다. 따라서 주택특성에 대한 수요함수가 두 집단에서 다르다면 주택특성의 내재가격도 다를 것이다. 동일한 논리를 공급곡선에 대해서도 적용할 수 있다. 주택건설업자가 이러한 소비자의 상황을 깨닫게 된다면 각 집단의 니즈에 부합하는 주택을 건축하려고 할 것이다. 따라서 두 집단에게 공급할 주택특성에 대한 공급곡선이 달라진

1) 유재우. 2007; 김경환·손재영. 2010. p166에서 재인용.

다.2)

OLS는 모형설정절차, 다중공선성, 독립변수 간 상호작용, 이분산성, 비선형성과 이상치(outlier) 등의 영향으로 특성가격모형의 성과에 심각한 영향을 미칠 수 있다(Limsombunchai, Gan & Lee, 2004). 최근 이러한 문제를 해결하기 위해 특성가격모형이 이상치나 오차항 분포에 민감하게 반응하지 않는 로버스트성(robustness)을 지니고 있는 분위회귀분석(Quantile regression; Qreg)으로 대체되고 있다(Gasparini & Escudero, 2004; Zietz, Zietz & Sirmans, 2008; Ebru & Eban, 2009).

Gasparini & Escudero(2004)는 소득불평등, 빈곤 또는 복지 수준의 측정에 주택소유주의 간주 임대료 소득을 고려하지 않아 발생하는 편의가 존재한다고 지적하고 브라질의 부에노스아이레스의 주택소유자의 간주임대료를 추정하기 위해 특성가격모형의 회귀계수를 Qreg으로 추정했다. 그 결과 소득분포의 하위계층에 속한 주택소유자가 상대적으로 비중이 높아 지니계수 등 불평등도 지표가 약간 개선되는 효과를 보였다.

Zietz, Zietz & Sirmans(2008)는 미국 유타주의 오렐/프로보의 공동중개서비스 목록에서 1999년 중반부터 2000년 중반에 걸쳐 1,366개 주택 매매사례에 대해 OLS와 Qreg 추정의 결과를 비교한 다음, 기존 OLS에서 그 크기와 방향이 엇갈렸던 규모나 경과연수의 차이는 주택가격의 차이로 설명할 수 있다고 주장했다. 토지면적, 화장실 수 등은 매매가격이 증가함에 따라 그 회귀계수도 함께 증

가했다. 경과연수와 산의 조망은 매매가격이 증가함에 따라 오히려 그 회귀계수는 감소했다. 차고, 외관, 스프링클러시스템, 주 경계나 도심까지 거리는 매매가격과 무관하게 거의 일정한 회귀계수를 보였다.

Ebru & Eban(2009)은 주택가격에 영향을 미치는 요인은 국가, 지역, 도시에 따라 다를 것이라고 보고 터키의 이스탄불에서 2007년 10월부터 12월 까지 부동산중개인 992명을 대상으로 조사한 자료에 OLS와 Qreg를 적용하고 그 결과를 비교했다. 차고에 대한 회귀계수는 30%분위까지 통계적으로 비유의적이었으나, 40%분위 이후 주택가격과 통계적으로 유의한 정의 관계로 나타났다. 또한 거리에 면한 주택은 30%분위 이후부터 주택가격과 부의 관계가 통계적으로 유의적인 것으로 나타났다.

Hung, Shang & Wang(2010)은 호텔 객실 이용료 결정요인을 분석하기 위해 Qreg 추정결과와 OLS 추정결과를 비교했다. OLS에서는 객실 수, 경과연수, 시장상황, 객실당 객실 담당 종업원 수가 주요 요인으로 나타났으나, Qreg 추정결과 저가격대 호텔에서는 객실 수와 객실당 객실 담당 종업원 수는 유의미한 변수가 아니었으며, 고가격대에서는 경과연수와 시장상황 그리고 외국인 고객 수가 중요한 변수로 나타났다.

본 논문은 OLS의 한계를 보완할 수 있는 Qreg을 적용한다.3) Qreg은 주택가격 분포의 여러 다른 위치에서 주택특성의 내재가격을 확인할 수 있는 방법이다. 이 방법은 어떤 주택특성의 내재가격이 고가 주택과 저가주택에서 서로 다른지 명시적으

2) 소형평형의무비율 규제에 대한 논란을 상기하면 쉽게 이해할 수 있음.

3) 박범조(2003)는 노동시장의 임금구조 분석, 실업지속기간에 미치는 영향 분석, 학교의 질이나 학급의 크기가 학생의 교육성과에 미치는 영향 분석, 기업 창업규모의 조건부 분포, R&D투자와 판매량의 관계, 정보기술수준과 기업성과, 포트폴리오 투자스타일 분석, 외환거래량과 원/달러 환율변동성 연구, 외환수익률의 비대칭 변동성 연구 등에 대한 국내외 선행연구를 소개하고 있음. Qreg의 활용 분야가 초기 노동경제분야에서 미시경제학 분야, 재무관리 및 국제금융분야에 대한 분석으로 확산되고 있음.

로 확인할 수 있게 해준다. Qreg은 종속변수의 조건부 분포에 대한 자세한 정보를 제공한다. 본 논문에서는 OLS와 Qreg의 결과를 비교한다. 또한 OLS 추정량에 이분산성이 있는지 여부와 정규성 가정에 부합하는지 여부도 검증할 것이고, 각 연도의 자료에 포함되어 있는 이상치의 존재 여부도 검증할 것이다. 이를 통해 Qreg의 적용이 타당함을 보일 것이다.

본 논문은 다음과 같이 전개한다. 제2장에서는 연구방법과 자료에 대해 설명하고, 제3장에서는 실증결과를 제시한다. 제4장에서는 연구의 성과를 요약한다.

## II. 연구모형

### 1. 특성가격모형

일반적으로 이질적인 주택의 가격은 주택을 구성하는 여러 특성에 대한 잠재가격의 합인 특성가격 모형으로 다음과 같이 표현할 수 있다.

$$p_i = X_i\beta_\tau + \epsilon_i$$

여기서  $p_i$ 는 주택가격<sup>4)</sup>으로 종속변수,  $X_i$ 는 설명변수,  $\beta$ 는 회귀계수,  $\epsilon_i$ 는 오차항을 뜻한다. 설명변수는 주택을 구성하는 주택특성이 된다. 주택가격을 설명변수에 대해 편미분하면 각 주택특성 1단위 추가에 대해 소비자가 지불하려고 하는 추가적인 한계효과를 알 수 있다.

### 2. 분위회귀접근법

Koenker & Basset(1978)은 OLS보다 더 유연하고 종속변수의 분포 전체에 대한 공변량의 영향을 분석할 수 있는 Qreg를 도입했다. Qreg은 주택가격 분포의 위치마다 주택가격에 대한 설명변수의 반응이 다를 수 있다는 점을 고려한다.

Qreg을 살펴보기 전에 분위의 정의를 살펴보자.<sup>5)</sup> 만일 어떤 도시에서 전체 주택 중 어떤 주택의 가격보다 가격이 낮은 주택의 비율이  $\tau$ 이고 가격이 높은 주택의 비율이  $(1 - \tau)$ 라면 이 주택의 가격은 도시 전체 주택의 가격 분포에서  $\tau^{\text{th}}$  분위 ( $\tau^{\text{th}}$  quantile)에 위치하게 된다.

Qreg은 조건부 분위 함수를 추정하기 위한 가중 절대편차(weighted absolute deviation)의 최소화에 기반하고 있다. 분위가 0.5인 중위수에 대해서는 대칭적 가중치가 사용되며, 다른 분위에 대해서는 비대칭적 가중치가 사용된다. 이외는 대조적으로 OLS는 조건부 평균 함수를 추정한다. 따라서 Qreg은 OLS와 달리 종속변수의 평균을 설명하는데 한정되지 않고 종속변수의 분포 중 어느 점에서든 종속변수의 결정요인을 설명할 수 있다. 특성가격모형에 Qreg을 적용하면 주택가격 분포의 위치에 따라 달라질 수 있는 주택특성에 대한 평가의 차이를 통계적으로 검토할 수 있다.<sup>6)</sup>

특성가격모형에서 Qreg은 다음과 같이 표현할 수 있다.

$$p_i = X_i\beta_\tau + \epsilon_{\tau,i},$$

$$Q_\tau(p_i | X_i) = X_i\beta_\tau, (i = 1, 2, \dots, n)$$

4) 실증분석에서는 아파트 매매가격에 자연로그를 취한 값을 사용했음.

5) 이 내용은 주로 박범조(2003)의 내용을 참조했음.

6) Newsome & Zietz(1992)은 종속변수를 몇 개의 집단으로 구분하고 OLS를 적용하면 Qreg과 유사한 결과를 도출할 수 있다고 하지만, Heckman(1979)에 따르면 이러한 종속변수의 단절은 모수추정에 편의가 발생할 수밖에 없으므로 적절한 대안이 될 수 없음.

여기서  $\beta_\tau$ 는  $\tau$ 분위 회귀계수를 뜻한다. 또한  $Q_\tau(p_i | X_i)$ 는  $X$ 가 주어진 상태에서  $p$ 의  $\tau$ 번째 조건부 분위함수다.<sup>7)</sup> 분위 회귀계수  $\beta_\tau$ 의 추정치는 주어진 분위  $\tau$ 에서 다음 최소화 문제의 해가 된다.

$$\text{Min } \frac{1}{n} \left( \sum_{p_i \leq X_i \beta} \tau |p_i - X_i \beta| + \sum_{p_i > X_i \beta} (1 - \tau) |p_i - X_i \beta| \right)$$

만약  $\tau$ 가 0.1이라면 우변 둘째 항의 비중이 크므로 관측치가 아래로 벗어날 때 큰 손실이 발생하나, 첫째 항의 비중이 작으므로 관측치가 위로 벗어날 때 손실은 작다. 이 경우 10%분위 회귀 추정식은 우변 둘째 항의 손실을 최소화하기 위해 산포도에 있어 아래 부분을 통과하게 될 것이다. 즉 산포의 하위 10%를 지나가는 주택가격을 추정한다. 주택가격 분포 하위 10%에 해당하는 주택가격을 결정하는 주택특성요인을 확인할 수 있다.

Qreg을 이용하여 추정한 계수에 대한 해석은 OLS와 동일하다. Qreg와 일반 선형회귀식의 구조가 동일하고, 다만 손실값에 붙는 비중( $\tau$ )만 다르기 때문이다(윤형호·임병인, 2009).

### III. 자료와 실증결과

#### 1. 자료

연구에 사용한 자료는 부동산114가 조사한 서울지역 아파트 호가 자료로 매매가격과 전세가격 조사가 모두 가능한 40세대 이상 단지에 대한 2000년부터 2009년까지 10년간 매년 말 기준 평형별 자료

다.<sup>8)</sup>

매매가격은 상한가와 하한가의 평균을 사용했다. 설명변수로는 특성가격모형을 적용한 선행연구에서 많이 사용한 물리적 특성과 입지 특성으로 한정했다. 투입된 설명변수는 경과연수, 면적, 현관구조, 난방연료, 주상복합 여부, 재건축 여부, 단지 세대수, 그리고 지하철접근성과 공원인접 여부이며, 강남지역의 입지적 특성을 반영하기 위해 강남지역 3개 구(강남구, 서초구, 송파구) 여부를 추가했다.

<표 1>은 2009년 자료의 기초통계량을 보여주고 있다. 종속변수든 설명변수든 모두 OLS의 정규성 가정에 위배된다. 자료에 포함된 서울시 아파트의 평균가격은 5억 9천만 원이며, 경과연수는 평균 약 12년이고, 평균면적은 112.6m<sup>2</sup>다. 공원인접 사례는 전체 자료의 59%, 현관구조가 계단식은 69%, 역세권에 해당하는 건수는 79%에 이른다. 반면 주상복합은 12%, 재건축 해당 건수는 4.5%에 지나지 않는다. 강남3구에 있는 표본은 전체 표본의 21%나 된다.

<표 2>는 2009년 자료에서 각 변수 간 상관관계를 보여주고 있다. 가격은 면적과 구조, 강남지역과 상관관계가 높으나, 경과연수, 난방연료, 공원인접 여부와 상관관계는 낮다.

여기서 종속변수인 가격은 전체 가격이므로 당연히 면적과 상관관계가 높게 나타난 것이며, 구조 역시 신축 주택일수록 현관구조가 복도식에서 계단식으로 변경되었기 때문에 구조와 가격의 상관관계가 높게 나타난 것으로 보인다.

경과연수는 재건축 여부, 구조와 상관관계가 높

7) Qreg도 일반적인 선형회귀모형의 기본가정인 모든  $i$ 에 대해서  $Q_i(\epsilon_{i,t} | X_i) = 0$ 이 성립함.

8) 2000년의 표본수는 5,873, 2001년은 6,730, 2002년은 7,669, 2003년은 8,527, 2004년은 9,373, 2005년은 10,118, 2006년은 10,813, 2007년은 11,299, 2008년은 11,979, 2009년은 12,568개임. 자료가 평형별 자료이니 각 세대나 단지에 대한 자료가 아님에 유의할 필요가 있음.

표 1\_ 변수의 기초통계량(2009년)

구분	n	mean	std	max	min	Jarque-Bera
가격(만 원)	12,568	59,281	49,338	750,000	5,750	918.68 (0.0000)
경과연수(년)		12.46	8.87	44	1	2330.39 (0.0000)
면적(3.3㎡)		34.11	12.56	124	4	13899.51 (0.0000)
구조 (복도식 = 1)		0.26	0.44	1	0	2841.33 (0.0000)
난방연료 (기름 외 = 1)		0.99	0.12	1	0	2670268 (0.0000)
주상복합 (주상복합 = 1)		0.12	0.33	1	0	16607.95 (0.0000)
공원 (공원인접 = 1)		0.59	0.49	1	0	2102.11 (0.0000)
재건축 (진행 중 = 1)		0.04	0.21	1	0	198576.1 (0.0000)
지하철 (도보10분 내 = 1)		0.79	0.41	1	0	4284.74 (0.0000)
세대수(세대)		518.09	723.23	6864	40	200610.6 (0.0000)
강남지역 여부 (강남지역 = 1)		0.21	0.41	1	0	3981.8 (0.0000)

표 2\_ 변수 간 상관계수(2009년)

구분	가격	경과연수	면적	구조	난방연료	주상복합	공원	재건축	지하철	세대수
경과연수	-0.0398	1.0000								
면적	0.7872	-0.1551	1.0000							
구조	-0.3368	0.3400	-0.3393	1.0000						
난방연료	0.0180	-0.0617	0.0182	-0.0110	1.0000					
주상복합	0.1796	-0.2939	0.2365	-0.0357	0.0269	1.0000				
공원	0.0217	0.0375	-0.0102	-0.0268	-0.0002	-0.0158	1.0000			
재건축	0.1021	0.4441	-0.0618	0.1012	-0.0248	-0.0811	-0.0190	1.0000		
지하철	0.1932	0.0718	0.0532	0.0344	0.0283	0.1081	0.0302	0.0745	1.0000	
세대수	0.1572	0.1473	-0.0261	0.0693	0.0471	-0.1246	0.1154	0.1233	0.1402	1.0000
강남지역	0.4371	0.0145	0.2072	-0.0051	-0.0007	0.1666	-0.0635	0.0889	0.1809	0.0430

다. 경과연수가 오래될수록 재건축 가능성이 높아 우가 많기 때문으로 보인다.<sup>9)</sup> 지며, 오래된 주택일수록 현관구조가 복도식인 경

9) 따라서 재건축 여부와 현관구조, 경과연수는 밀접하게 관련되어 있는 변수이나 아파트의 가격대에 따라 그 영향이 달라질 수

## 2. 2009년 회귀분석 결과<sup>10)</sup>

<표 3>은 OLS와 Qreg 추정결과를 보여주고 있다. 먼저 OLS 결과를 보면, 난방연료와 공원인접 여부만 제외하면 모든 설명변수가 통계적으로 유의하게 나타났다. 재건축 여부와 지하철접근성의 회귀계수는 이 변수들이 주택가격에 미치는 영향이 지대함을 보여준다. 각 설명변수의 회귀계수는 그 부호가 예상대로 도출되었으며, 결정계수도 76%로 모형은 비교적 적합하다고 할 수 있다.

그러나, OLS는 여러 통계적 오류 가능성이 있다. 횡단면 자료에 대한 분석에서 일반적으로 나타나는 이분산성 존재 여부에 대한 검증이 필요하다. White의 이분산성 검증<sup>11)</sup>에서 귀무가설은 공분산이 일정하다는 것이다. 유의수준 1%에서 귀무가설이 기각되고 있다. 또한 잔차의 정규성 여부는 Jarque-Bera의 검정방법을 적용한 결과 유의수준 1%에서 오차항이 정규분포라는 귀무가설을 기각하고 있다. 그리고 OLS 결과가 자료의 이상치의 영향을 받고 있는지 여부를 로버스트 회귀분석(Robust)을 통해 검토한 결과 <표 4>와 같이 2009년의 경우 1.82%가 이상치로 나타났다( $R^2 = 0.6264$ ). <그림 1>을 보면 경과연수, 공원, 재개발, 세대수에서 OLS와 Robust 추정 회귀계수의 차이가 나타나는 반면, 다른 변수에서는 큰 차이를 보이지 않고 있다.

Qreg 추정결과는 OLS 추정결과와 사뭇 다른 점이 눈에 띈다. 경과연수는 50%분위까지는 부의 영향, 60%분위부터는 정의 영향을 보이고 있다. 경과연수에 대한 OLS 추정결과는 70%분위의 Qreg 추정결과와 유사하다. 면적은 모든 분위에서 거의 일정하게 정의 영향을 보이고 있다. 이는 OLS 추정결과와 별다른 차이가 없다.

난방연료의 경우 10% 이하 분위에서 주택가격에 미치는 영향이 매우 크다. 10% 초과 분위에서는 OLS 결과와 마찬가지로 그 영향력이 크지 않을 뿐만 아니라 통계적 유의성도 없는 것으로 나타났다. 예외적으로 90%분위에서는 5% 유의수준에 통계적으로 유의한 것으로 나타났다. 10%분위에서는 정의 영향, 90%분위에서는 부의 영향으로 나타났는데, 이는 저가주택은 생활에 불편하더라도 저가의 난방연료를 사용할 수 있는 점이 정의 영향으로 나타난 것으로 보이며, 고가주택에서는 반대의 현상이 나타난 것으로 해석할 수 있다.

공원인접 여부는 모든 분위에서 통계적으로 유의하게 정의 영향, 현관구조는 부의 영향이 나타났다. 공원인접 여부는 주택의 쾌적성에 가장 큰 영향을 미치는 요소이기 때문에 보이며, 현관구조는 경과연수와 관련이 있으면서도 아파트의 사생활 보호와 같은 보안성과 관련이 높기 때문에 보인다.

주상복합 여부는 반대로 70%분위 이상에서만

이므로 모두 설명변수로 투입하기로 함. 실제 Qreg의 추정결과에서 이를 확인할 수 있음.

- 10) 본 논문에서는 아파트 전체 가격에 자연로그를 취한 값을 종속변수로 사용했으나, 단위면적당 가격에 자연로그를 취한 값을 종속변수로 할 수도 있음. 본 논문은 아파트 가격대에 따라 아파트시장이 분리되어 있다는 가정을 기존 연구에서 고려하지 못한 문제를 Qreg으로 분석할 목적을 지니고 있기 때문에 설명변수에 면적을 추가하여 면적의 영향을 통제했음. 단위면적당 가격에 자연로그를 취한 값을 종속변수로, 설명변수에서 면적을 제외한 모형을 추정한 결과  $\bar{R}^2=0.441$ 이었으며, 회귀계수의 통계적 유의성과 방향은 동일했으며, 그 상대적 크기도 큰 변화가 없었음. 다만, Qreg에서는 난방연료, 현관구조, 그리고 주상복합 여부의 회귀계수가 분위별로 다소 변동하는 모습을 보였음.
- 11) E-views에서 White의 이분산성 검증을 할 때, 변수 간 상호작용 항을 추가할 것인지 여부를 묻는 옵션이 있는데, 본문은 상호작용 항이 있는 경우의 결과를 보여주고 있음. 상호작용 항이 없는 경우에도 통계량(1484.11[0.0000])은 이분산성이 존재하고 있는 것으로 나타났음. 다른 검증모형에서도 결과는 모두 동일했음.

표 3\_ 회귀모형 추정결과(2009년)

구분	OLS	Q10	Q20	Q30	Q40	Q50	Q60	Q70	Q80	Q90
상수항	9.335*** (0.026)	9.094*** (0.038)	9.204*** (0.066)	9.306*** (0.052)	9.330*** (0.026)	9.360*** (0.032)	9.365*** (0.054)	9.377*** (0.047)	9.447*** (0.061)	9.562*** (0.047)
경과연수	0.003*** (0.000)	-0.007*** (0.001)	-0.007*** (0.001)	-0.005*** (0.001)	-0.003*** (0.001)	-0.002*** (0.001)	0.000 (0.001)	0.003*** (0.001)	0.007*** (0.001)	0.011*** (0.001)
면적	0.035*** (0.000)	0.033*** (0.000)	0.033*** (0.001)	0.034*** (0.000)	0.034*** (0.000)	0.035*** (0.000)	0.036*** (0.000)	0.037*** (0.001)	0.036*** (0.000)	0.037*** (0.001)
현관구조	-0.187*** (0.007)	-0.171*** (0.009)	-0.163*** (0.008)	-0.157*** (0.008)	-0.164*** (0.008)	-0.189*** (0.008)	-0.193*** (0.009)	-0.200*** (0.012)	-0.226*** (0.011)	-0.185*** (0.017)
난방연료	-0.001 (0.024)	0.113*** (0.040)	0.069 (0.067)	-0.005 (0.051)	-0.007 (0.022)	-0.005 (0.022)	-0.009 (0.052)	-0.013 (0.042)	-0.008 (0.060)	-0.069** (0.034)
주상복합	-0.013*** (0.009)	-0.200*** (0.020)	-0.108*** (0.017)	-0.066*** (0.013)	-0.040*** (0.015)	-0.022** (0.011)	-0.008*** (0.010)	0.012*** (0.014)	0.076*** (0.014)	0.099*** (0.018)
공원	0.030 (0.006)	0.031*** (0.007)	0.030*** (0.006)	0.017*** (0.006)	0.017*** (0.006)	0.014** (0.006)	0.018*** (0.006)	0.026*** (0.007)	0.041*** (0.008)	0.072*** (0.011)
재건축	0.283** (0.015)	0.260*** (0.035)	0.385*** (0.047)	0.456*** (0.020)	0.456*** (0.019)	0.416*** (0.017)	0.392*** (0.029)	0.380*** (0.027)	0.300*** (0.019)	0.225*** (0.032)
지하철	0.127*** (0.007)	0.129*** (0.009)	0.115*** (0.007)	0.111*** (0.006)	0.126*** (0.007)	0.122*** (0.007)	0.125*** (0.007)	0.132*** (0.009)	0.114*** (0.010)	0.092*** (0.012)
세대수	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
강남지역	0.397*** (0.007)	0.331*** (0.015)	0.377*** (0.013)	0.393*** (0.011)	0.422*** (0.011)	0.446*** (0.009)	0.439*** (0.010)	0.439*** (0.009)	0.422*** (0.011)	0.408*** (0.013)
White	2619.31 [0.000]	n.a.								
Jarque -Bera	336.63 [0.000]									
R2	0.757 [0.000]	0.470 [0.000]	0.483 [0.000]	0.495 [0.000]	0.508 [0.000]	0.521 [0.000]	0.530 [0.000]	0.539 [0.000]	0.550 [0.000]	0.561 [0.000]
회귀계수의 동일성 검증	$\chi^2$ value [Pr > $\chi^2$ ] = 1134[0.000]									
회귀계수의 대칭성 검증	$\chi^2$ value [Pr > $\chi^2$ ] = 492[0.000]									

주: 1) \*\*\*, \*\*, \*는 각각 1%, 5%, 10% 유의수준에서 통계적으로 유의함을 의미함.  
 2) R2: OLS는 Adjusted R2([ ]안 수치는 F-test p-value), Quantile regression은 pseudo R2([ ]안 수치는 quasi-LR통계량에 대한 p-value).  
 3) 괄호 안 수치는 표준오차, [ ]안 수치는 p-value.

표 4\_ 이상치 탐색 결과

구분	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
N	5873	6730	7669	8527	9373	10118	10813	11299	11979	12568
outlier/N	0.0133	0.0137	0.0110	0.0138	0.0135	0.0155	0.0100	0.0081	0.0171	0.0182

통계적으로 유의한 것으로 나타났다. 주상복합아파트가 도심지 고급주택으로서 의미를 지니는 것은 비교적 고가에 해당하고, 저가주택에서 주상복합은 경과연수가 오래된 상가겸용주택 수준의 주

택이 많기 때문으로 보인다. 재건축 여부는 40%분위까지 회귀계수의 크기가 계속 증가하다 50%분위 이후부터 점차 감소하는 모습을 보이고 있다. 재건축 대상 주택의 가격에 재건축 가능성이 미치는 영

향은 주택 가격대에 따라 다름을 알 수 있다.<sup>12)</sup>

강남지역 여부는 모든 분위에서 정의 영향을 미치고 있으며, 통계적으로도 유의하게 나타났다. 강남지역의 입지적 특성을 잘 보여주고 있다고 볼 수 있다. 이러한 점은 저가격대보다 고가격대에서 좀 더 높게 나타나고 있다. 즉 강남지역 소재 아파트가 비강남지역 소재 아파트에 비해 평균적으로 높은 가격대를 보이고 있음을 알 수 있는 대목이다.

OLS와 Qreg 추정결과에서 주택가격에 가장 영향을 미치는 주택특성은 강남지역 여부와 재건축 여부다. OLS에서는 강남지역 여부가 0.40, 재건축 여부가 0.28이다. Qreg에서는 10%분위에서 각각 0.33, 0.26이며, 20%분위에서 0.38, 0.39로 역전되어 분위가 상승함에 따라 점차 그 격차가 커져 40%분위에서 각각 0.42, 0.46로 최대가 되었다가 50%분위에서 0.45, 0.42로 다시 역전된다. 이후 그 차이는 더욱 벌어지는 양상을 보이고 있다. 즉, 저가격대에서는 강남지역 여부보다 재건축 여부가 매매가격에 더 중요한 영향을 미친다. 저가주택은 재건축을 통해 면적이 상향되고 그 결과 고가주택이 될 수 있기 때문에 재건축 여부가 중요한 요인으로 작용하며, 고가주택에서는 강남지역에 입지하고 있다는 특성이 더 중요한 요인으로 작용하고 있다고 볼 수 있다.

이상과 같이 Qreg의 결과는 OLS 결과가 큰 차이를 보인다. 이 두 변수의 회귀계수만 보아도 주택특성이 주택가격에 미치는 영향은 주택가격의 분위에 따라 다르며, 조건부 평균만 고려하는 전통적인 OLS로는 특성가격모형을 충분히 설명할 수 없음을 보여준다고 할 수 있다.

<그림 1>을 보면 OLS 추정 회귀계수와 비교할 때 Qreg 추정 회귀계수가 더 큰 분위와 작은 분위기를 확인할 수 있다. 경과연수, 면적, 공원과 지하철 접근성은 분위가 커질수록, 즉 주택가격이 고가일수록 상승하는 모습을 보이고 있다. 그러나 현관구조와 난방연료는 분위가 커질수록 하락하다가 일정 분위 이후부터 상승하는 모습을 보이고 있다. 난방연료와 공원인접 여부는 높은 분위보다 낮은 분위에서 더 주택가격에 영향을 미치는 요인으로 나타났다.

또한 주상복합 여부와 재건축 여부, 강남지역 여부는 저분위와 고분위에서보다 중간 분위에서 주택가격에 더 큰 영향을 미치는 것으로 나타났다. 총세대수는 10%분위에서 가장 영향이 낮으나 분위가 오르면서 점차 영향력이 상대적으로 커지다가 40%분위와 50%분위에서 최저의 영향을 보이고, 60%분위부터 다시 영향력이 커지고 있다.

한편, 각 분위별 회귀계수의 기울기가 통계적으로

표 5\_ 연도별 분위별 회귀계수의 동일성과 대칭성 검증 결과

구분	2000년	2001년	2002년	2003년	2004년	2005년	2006년	2007년	2008년
동일성	189 [0.000]	341 [0.000]	322 [0.000]	402 [0.000]	533 [0.000]	649 [0.000]	407 [0.000]	398 [0.000]	437 [0.000]
대칭성	127 [0.000]	305 [0.000]	756 [0.000]	220 [0.000]	160 [0.000]	202 [0.000]	350 [0.000]	234 [0.000]	181 [0.000]

주:  $\chi^2$  value [Pr >  $\chi^2$ ]

12) 재건축 여부는 경과연수와 높은 정의 상관관계(0.4441)를 보이고 있기 때문에 경과연수를 제외한 모형을 추정했으나, 설명력의 차이는 없었음. ( $R^2=0.756$ )

그림 1\_ OLS, Robust, Qreg 추정결과 비교

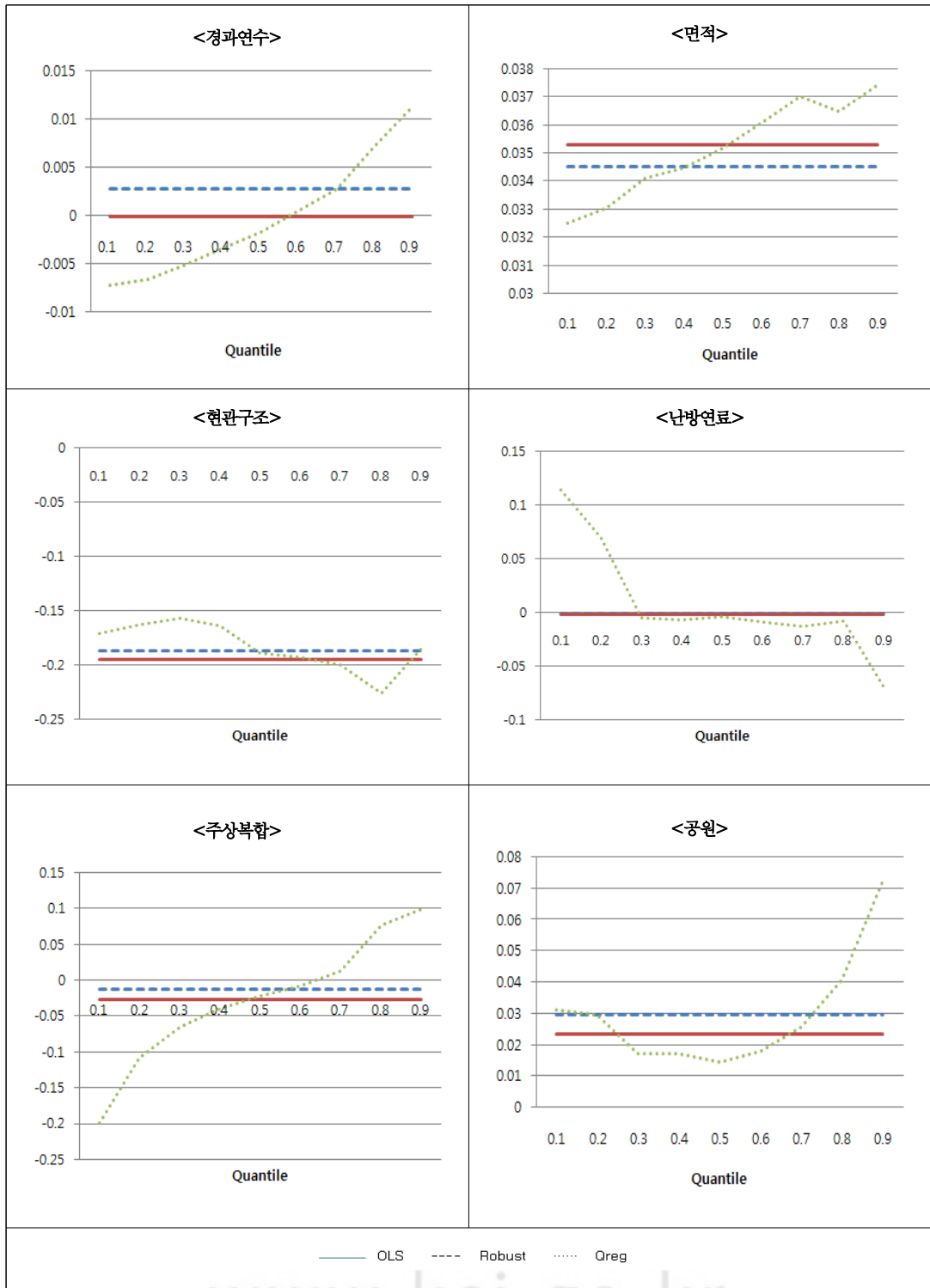
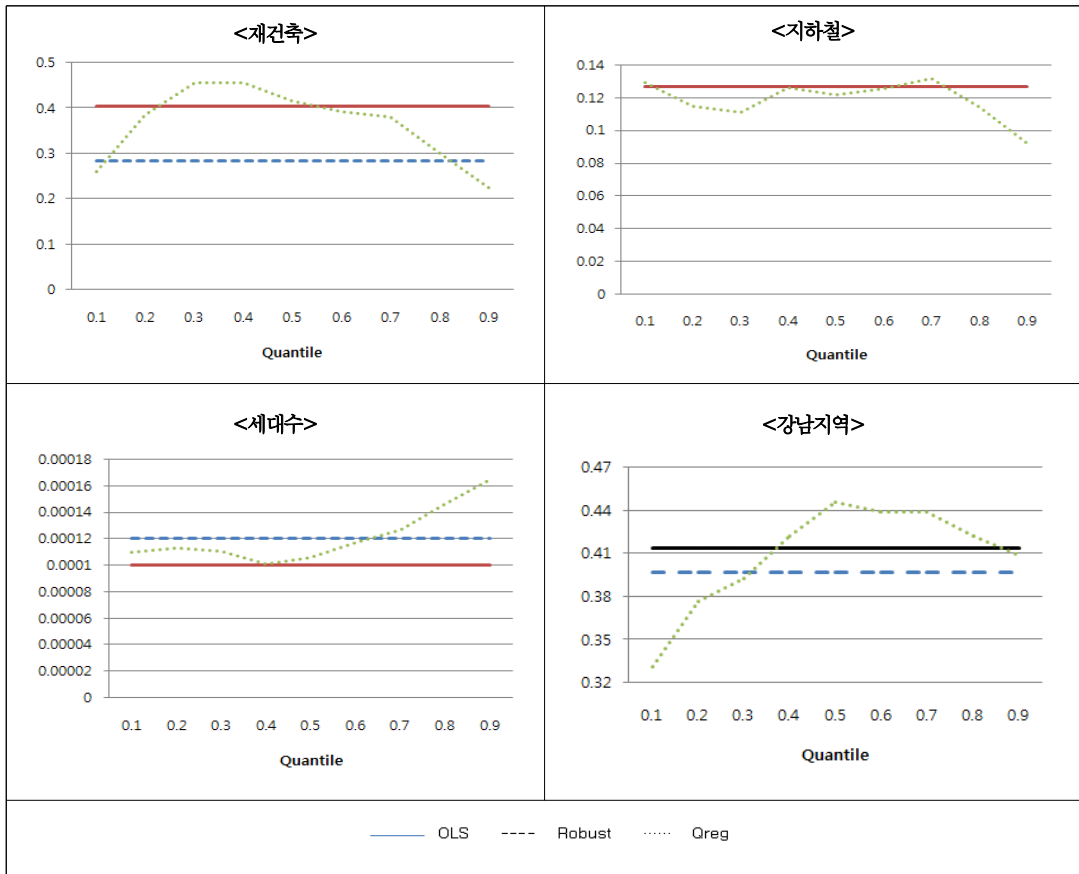


그림 1\_ OLS, Robust, Qreg 추정결과 비교(계속)



로 동일하지 여부를 검증한 회귀계수의 동일성 검증(Slope equality test) 결과  $\chi^2$ 값이 1,134로 유의수준 1%에서 통계적으로 유의하게 나타나, 분위 간 회귀계수는 같지 않으며, 조건부 분위는 동일하지 않음을 알 수 있다.<sup>13)</sup>

또한 분위의 수를 8개로 지정한 분위 간 회귀계수의 대칭성 검증(Symmetric quantiles test)<sup>14)</sup> 결과 분위 간 회귀계수는 중위수를 중심으로 비대칭적이다. 특히 25%분위와 75%분위 간 개별 회귀

계수에 대한 제약 검증에서는 유의수준 10%에서 난방연료, 공원인접 여부, 지하철역 인접 여부를 제외한 회귀계수에서 비대칭적으로 나타났다.

주택특성의 한계효과는 <표 6>과 같다. 한계효과는 해당 주택가격에 <표 3>의 회귀계수를 곱한 값이다. 90%분위의 매매가격은 10%분위의 매매가격의 약 5배 수준이다. 이에 반해 면적의 한계효과를 보면 90%분위는 10%분위에 비해 약 5.5배에 달하나, 재건축은 4.1배, 지하철접근성은 3.4배, 세

13) 귀무가설은  $\beta_1(\tau_1) = \beta_1(\tau_2) = \dots = \beta_1(\tau_k)$ 로 이분산성 검증 방법의 하나이며, 귀무가설이 채택되면 분위별 회귀계수가 같으므로 Qreg 추정의 실익이 없음을 의미함.

14) 귀무가설은  $\frac{\beta(\tau) + \beta(1-\tau)}{2} = \beta(1/2)$ 로, 귀무가설이 채택되면 50%분위를 중심으로 대칭적 위치에 있는 분위별 회귀계수의 평균값이 50%분위의 회귀계수와 동일함을 의미함.

표 6\_ 주택특성의 한계효과(2009년)

구분	OLS	Q10	Q20	Q30	Q40	Q50	Q60	Q70	Q80	Q90
주택가격	59,281	24,000	28,500	33,000	37,500	43,250	51,000	61,000	77,000	115,000
상수항	553,363	218,249	262,316	307,091	349,893	404,811	477,599	572,012	727,397	1,099,642
경과연수	159	-174	-187	-168	-126	-75	23	168	551	1,290
면적	2,046	780	942	1,125	1,293	1,521	1,840	2,257	2,810	4,304
현관구조	-11,085	-4,109	-4,659	-5,186	-6,153	-8,166	-9,835	-12,217	-17,379	-21,281
난방연료	-46	2,721	1,980	-155	-263	-197	-442	-812	-654	-7,922
주상복합	-761	-4,791	-3,091	-2,190	-1,514	-945	-421	759	5,847	11,335
공원	1,755	747	847	572	637	625	922	1,577	3,147	8,332
재건축	16,764	6,249	10,960	15,032	17,082	17,978	19,974	23,151	23,096	25,821
지하철	7,510	3,099	3,280	3,659	4,738	5,266	6,399	8,042	8,808	10,624
세대수	7	3	3	4	4	5	6	8	11	19
강남지역	23,505	7,946	10,737	12,969	15,813	19,276	22,401	26,750	32,500	46,944

주: 한계효과 = 추정 회귀계수 × 해당 주택가격  
 주택가격: OLS는 평균, 각 분위는 해당 분위수

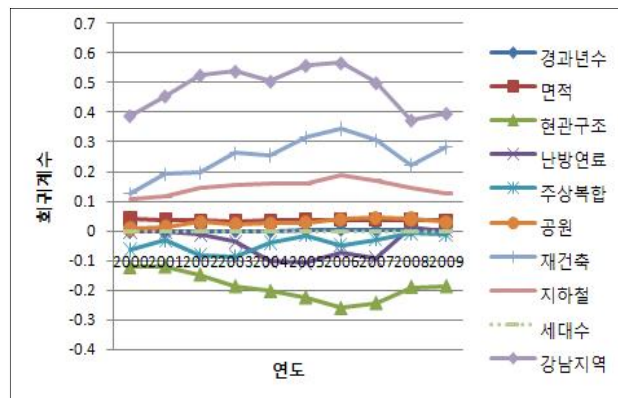
대수는 6.3배, 강남지역은 5.9배의 차이를 보이고 있다. OLS는 Qreg에 비해 저가주택의 매매가격에 대해 재건축과 지하철접근성은 과소평가하고 있으며, 면적과 세대수 그리고 강남지역 여부는 과대평가하고 있다는 것을 알 수 있다.

경과연수에 대한 한계효과를 보면 저가주택보다 고가주택에서 경과연수가 증가할수록 주택가격에 정의 영향을 미침을 알 수 있다. 이는 재건축 가능성에 대한 기대감의 표현이라고 할 수 있다. 주상복합도 경과연수와 유사한 패턴을 보인다.

3. 분위별 회귀계수의 연도별 변동 추이

먼저 연도별 자료에 대한 OLS 추정결과를 살펴보자. <그림 2>는 연도별 OLS 추정 각 회귀계수의 변동 추이를 보여주고 있다(상수항 제외). 연도마다 각 변수의 회귀계수는 고정적이지 않음을 알 수 있다. 이는 주택특성의 잠재가격이 일반경제상황

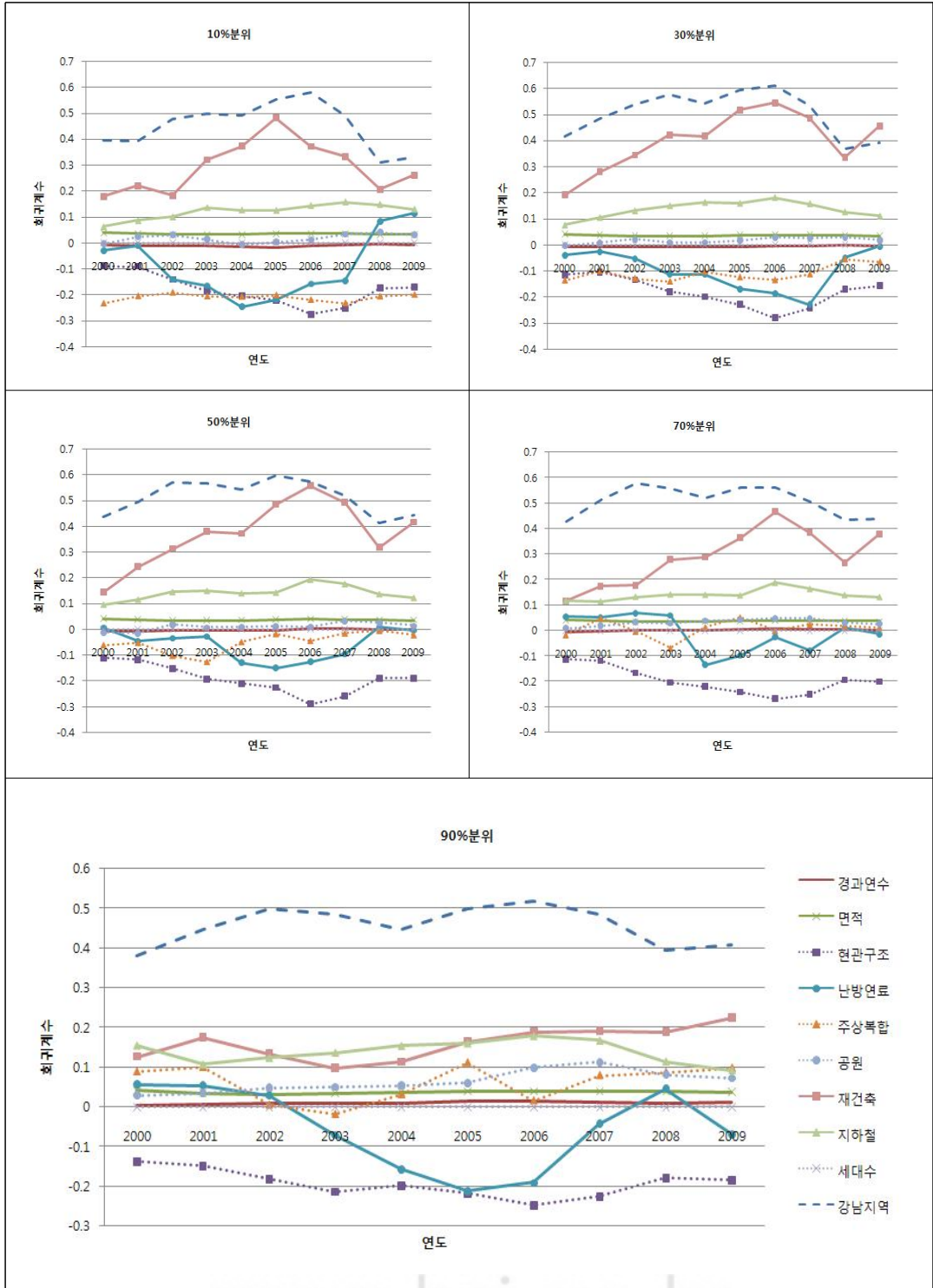
그림 2\_ OLS 추정 회귀계수의 변동 추이



에 따라 변동한다는 것을 시사한다. 대체로 2000년 이후 2006년까지는 재건축 여부와 지하철접근성, 주상복합 여부, 강남지역 여부는 상승하다 이후 하락하는 추세를 보이고 있으며, 난방연료와 현관구조는 반대로 하락하다 상승하는 추세를 보이고 있다. 그러나 면적, 경과연수, 총세대수는 연도별 변동이 거의 없는 것으로 나타났다.

Qreg으로 추정한 각 분위별 회귀계수의 변동추이는 다음 <그림 3>과 같다. 대표적으로 10%분위,

그림 3\_ Qreg 추정 회귀계수의 변동 추이



30%분위, 50%분위, 70%분위, 그리고 90%분위 결과만 보이고 있다.

면적과 경과연수는 연도별 변동이 거의 없다. 이는 OLS 추정결과와 같다. 그러나 재건축 여부는 90%분위에서 연도별 변동이 다른 하위 분위보다 상대적으로 안정적임을 알 수 있다. 고가주택일수록 재건축 여부가 주택가격에 미치는 영향은 연도별 변동이 적다는 것이다. 주상복합 여부는 2005년까지는 약간 상승하다 2006년 이후 다소 하락하는 추세를 보였다. Qreg에서는 각 분위마다 주상복합 여부의 회귀계수의 연도별 변동이 상대적으로 크게 나타났다. 하위 분위에서는 OLS 결과와 유사한 추세를 보였으나, 상위 분위에서는 상승과 하락을 반복하는 모습을 보여 대조적이다.

지하철접근성과 재건축 여부를 보면 70%분위까지는 지하철접근성보다 재건축 여부가 더 중요한 요인이나, 70%분위 이후부터는 지하철접근성이 더 중요한 요인이다. 또한 덜 중요한 요인의 연도별 변동도 크지 않은 것으로 나타났다. 즉 70%분위 이상에서 재건축 여부의 회귀계수는 지하철접근성의 회귀계수에 비해 연도별 변동추이가 상대적으로 안정적이다.

강남지역 여부는 모든 가격분위에서 유사한 모습을 보이고 있다. 2006년까지 중요성이 증대하다가 2006년 이후 그 크기가 다소 감소했다.

#### IV. 요약 및 결론

본 논문은 주택가격에 영향을 미치는 요인이 주택가격에 영향을 미치는 크기나 방향이 주택가격 분포의 위치에 따라 다를 수 있다, 즉 고가주택과 저가주택시장에서 수요자나 공급자가 판단하는 특정 주택특성의 잠재가격은 다를 것이며, 결국 두 시장은 서로 분리되어 있을 수 있다는 의문에서 출발했

다. 주택에 대한 특성가격모형을 적용한 연구는 주택시장을 가격대보다는 면적(평형)이나 지역에 따라 분류하고 있으나, 본 논문에서는 주택시장을 면적이나 지역에 따라 세분할 수도 있지만 가격대에 따라 세분할 수 있다는 점도 고려했다. 기존의 OLS 추정 특성가격모형은 주택가격 분포의 조건부 평균만 고려한 모형이기 때문에 전체 주택가격의 분포를 고려하지 못하는 단점이 있다. 따라서 본 논문에서는 특성가격모형을 OLS는 물론 주택가격의 조건부 분포의 특성을 알려줄 수 있는 Qreg으로 추정했다.

실증결과는 주택가격의 각 분위별 회귀식을 추정함으로써 주택가격에 영향을 미치는 주택특성에 대해 더 풍부한 이해가 가능했다. 예를 들어 경과연수는 50%분위까지는 주택가격에 부의 영향을 미치지만, 60%분위 이후부터는 정의 영향을 미치는 것으로 나타났다. 즉 경과연수의 회귀계수는 주택가격 수준에 따라 다른 영향을 미친다는 것이다. 고가주택수요자는 특정 주택특성에 대해 저가주택수요자와 다른 잠재가격을 지니고 있다고 할 수 있다.

2000년 이후 매 연말 기준으로 각 연도의 특성가격모형에서 각 주택특성의 회귀계수는 주택경기 변동에 따라 그 크기도 달라지고 있다. Qreg 추정결과는 대체로 OLS 추정결과와 유사한 변동 추이를 보였다. 그러나 재건축 여부는 고가주택일수록 그 변동이 상대적으로 안정적이며, 주상복합 여부는 고가주택에서 상승과 하락을 반복하는 추세를 보였으며, 강남지역 여부는 모든 주택가격 분위에서 유사한 형태를 보여 지역요인이 가격대마다 크게 다르지 않게 작용하고 있음을 알 수 있다.

본 논문은 서울시 아파트만을 대상으로 한 연구로 전국적인 현상으로 일반화하기에는 다소 무리가 있다는 한계가 있다. 그리고 실제 거래가격이 아닌 매도자의 호가 위주의 가격을 대상으로

한 점에서 실거래가격이 반영되지 않은 점은 큰 약점이라고 생각한다. 또한 주택특성이 물리적 특성과 접근성에 치우쳐 도심이나 학교, 편의시설에 서부터 거리, 조망, 교육시설 등 최근 연구주제가 되고 있는 특성이 포함되지 않아 변수누락 문제가 발생할 수 있으며, 공간자기상관을 고려하지 않은 계량경제학적 한계도 지니고 있다. 그리고 가격대에 따른 시장분리를 고려하다보니, 각 지역별 평균가격의 차이나 경과연수 등을 기준으로 본 지역별 아파트 구성비의 차이 등을 고려하지 못한 한계도 지니고 있다.

**참고문헌**

김경환·손재영. 2010. 부동산경제학. 서울 : 건국대학교출판부.  
 박범조. 2003. “분위수회귀접근법”. 계량경제학보 제14권 제4호. 서울 : 한국계량경제학회. pp93-122.  
 유재우. 2007. “아파트 외부공간 요소가 가격에 미치는 영향-강남-강북지역 아파트 사례를 중심으로”. 건국대학교 석사학위논문.  
 윤형호·임병인. 2009. “의료서비스특성을 감안한 고령자가구의 의료비지출 분석”. 재정학연구 제2권 제4호. 서울 : 한국재정학회. pp119-156.  
 Ebru, Çağlayan and Eban, Arıkan. 2009. “Determinants of House Prices in Istanbul: a Quantile Regression Approach”. *Quality and Quantity* vol.20, Dec 2009. OnlineFirst.  
 Gasparini, Leonardo and Escudero, Walter Sosa. 2004. “Implicit Rents from Own-Housing and Income Distribution: Econometric Estimates for Greater Buenos Aires”. *Journal of Income Distribution* vol.12, no.1-2. USA : Ad Libros Publication Inc. pp32-55.  
 Heckman, J. J. 1979. “Sample Selection Bias As a

Specification Error”. *Econometrica* 47. USA : John Wiley & Sons, Inc. pp153-161.  
 Hung, Wei-Ting, Jui-Kou Shang and Fei-Ching Wang. 2010. “Pricing Determinants in the Hotel industry: Quantile Regression Analysis”. *International Journal of Hospitality Management* vol.29, no.3. Netherlands : Elsevier B.V. pp378-384.  
 Koenker, R. and Basset, G. 1978. “Regression Quantiles”. *Econometrica* 46. USA : John Wiley & Sons, Inc. pp33-50.  
 Limsombunchai, V. Gan, C. and Lee, M. 2004. “House Price Prediction: Hedonic Price Model vs. Artificial Neural Network”. *American Journal Applied Science* vol.1, no.3. USA : Science Publications. pp193-201.  
 Malpezzi, S. 2003. “Hedonic Pricing Models: a Selective and Applied Review”. (in) O’Sullivan, T. Gibb, K. eds. *Housing Economics and Public Policy*. Malden, MA. : Blackwell Science. pp67-89.  
 Newsome, B. and Ziete, J. 1992. “Adjusting Comparables Using MRA-The Need for Segmentation”. *Appraisal Journal* 60. USA : Appraisal Institute. pp129-135.  
 Sirmans, G., Macpherson, D., and Zietz, E. N. 2005. “The Composition of Hedonic Pricing Model”. *Journal of Real Estate Literature* vol.13, no.1. Web Journal of American Real Estate Society. pp3-43.  
 Zietz, Joachim, Zietz, Emily Norman, and Sirmans, G. Stacy. 2008. “Determinants of House Prices: a Quantile Regression Approach”. *Journal of Real Estate Finance and Economics* vol.37. Netherlands : Springer Science & Business Media B.V. pp317-333.

- 논문 접수일: 2010. 9. 3
- 심사 시작일: 2010.10.18
- 심사 완료일: 2010.11.19

---

**ABSTRACT**


---

**Trends of Determinants of Housing Price in Seoul, Korea  
: A Quantile Regression Analysis**

Keywords: Quantile Regression, Hedonic Price Model, House Price

This article is written from idea that determinants of house prices are different across distribution of house prices. Ordinary least squares method, which estimates conditional mean function, and quantile regression method, estimated for each quantile of conditional distribution of house prices, are used to estimate hedonic price model.

The results show that buyers of higher-priced homes value certain housing characteristics differently from customers of lower-priced homes. For Example, the age of house is negative to house prices below 50% quantiles, however positive above 60% quantiles of house price distribution. Hedonic price models are estimated on the data from year 2000 to year 2009 of apartment sales prices in Seoul, Korea, the magnitude and importance of regression coefficients vary with the cycles of general and housing business. Trends of coefficients of two estimation methods show some similarities and differences, especially redevelopment and residential-commercial complex are differ across quantiles of house prices.

**서울시 아파트 가격분위별 가격결정요인의 변동 추이에 관한 연구**

주제어: 분위회귀접근법, 특성가격모형, 주택가격

본 논문은 주택가격에 영향을 미치는 요인이 주택가격에 영향을 미치는 크기나 방향이 주택가격 분포의 위치에 따라 특정 주택특성의 잠재가격은 다를 것이라는 의문에서 출발했다. 기존의 통상 최소자승법에 의한 추정은 주택가격 분포의 조건부 평균만 고려하기 때문에 주택가격의 조건부 분포의 특성을 알려줄 수 있는 분위회귀접근법으로 특성가격모형을 추정했다.

서울시 아파트 매매가격자료를 이용한 실증결과는 고가주택 수요자는 특정 주택특성에 대해 저가주택 수요자와 다른 잠재가격을 지니고 있음을 확인했다. 예를 들면 경과연수는 주택가격 분포의 50%분위까지는 주택가격에 부의 영향을 미쳤으나, 60%분위부터는 정의 영향을 미쳤다. 2000년부터 2009년까지 매 연도 말의 특성가격모형에서 각 주택특성의 회귀계수는 주택경기변동에 따라 그 크기와 중요도가 달라지고 있다. 분위회귀접근법 추정결과는 대체로 통상최소자승법 추정결과와 유사한 변동 추이를 보였으나, 재건축과 주상복합여부는 주택가격 분위에 따라 다른 변동 추세를 보였다.