

주택 매매가격 분포를 통한 지역별 · 가격대별 변동성 분석

민병철 연구위원



국민의 행복과 함께 성장하는 최고의 주택금융기관

주택시장/정책

주택 매매가격 분포를 통한 지역별·가격대별 변동성 분석

2021년 12월



■ 작성자 : 민병철 연구위원 (051-663-8177 / bcmin@hf.go.kr)

목 차

1. 연구배경 및 목적	1
2. 분석방법 및 자료	5
3. 분석결과	10
3.1 주요 특성별 가격 변화 분포	10
3.2 주택 가격 지수 변화율별 조건부 가격 변동 성 분석	12
3.3 주택 특성 클러스터링 결과	14
3.3.1 비조건부 분포 클러스터링 결과	15
3.3.2 조건부 분포 클러스터링 결과	19
4. 요약 및 결론	26
부록. 분석과정	28
참 고 문 헌	30

요 약

- 본 연구에서는 최초로 실거래 정보를 이용해 개별 아파트 가격에 대한 패널데이터를 구축하고, 이를 이용하여 가격 분포를 도출하여 기존 가격지수로 관찰하지 못하였던 주택시장의 특성을 분석하였음
- 일반적으로 동일 주택에 대한 거래는 빈번하지 않아 패널데이터 구축이 어려우나, 가격이 같다고 간주될 수 있는 아파트 세대가 다수 존재하는 한국 주택시장의 특성을 활용함
- 패널데이터 분석을 통해 가격지수만으로는 관찰할 수 없었던 주택시장 주요 특성들은 다음과 같음
 - 매매가격 변화율 분포의 꼬리는 오른쪽으로 긴 모습을 보여 중간값이 평균값보다 낮음
 - 주택가격 하락기에는 평균값인 지수가 측정하는 주택가격 하락폭보다 시장이 체감하는 하락폭은 더 클 수 있음
 - 주택 가격 하락기에는 개별 주택 가격이 대체로 같이 하락하며, 상승기에는 가격 상승률의 편차가 커지는 모습이 관찰됨
 - 가격지수가 상승할 때도 적지 않은 비중의 주택은 가격이 하락함
 - 매매가격지수가 0-10% 상승할 때는 18.5%, 10-20% 상승할 때는 5.2%가 가격 하락을 경험
 - 수도권 대형, 중·고가 주택은 주택가격 상승기에는 평균보다 많이 오르고 하락기에는 평균보다 더 하락함
- 본 연구는 경험적으로 체감하고 있었으나 기존 가격지수로 설명하기 어려웠던 주택시장의 현상들을 정량적으로 설명하였으며, 이러한 정량적 계산방식은 주택시장에 대한 이해뿐만 아니라 주택금융상품 리스크 측정 등에 유용하게 활용될 수 있을 것으로 기대됨

I. 연구 배경 및 목적

- 주택가격지수로 표현되는 주택시장과 시장 참여자들이 체감하는 주택시장 간에 괴리가 존재한다는 반응은 항상 존재해왔음
 - 매매가격에 변화가 없다고 여기는 사람과 크게 올랐다고 느끼는 사람이 동시에 존재하는 등 시장 참여자들 간에도 편차 존재
- 괴리 원인에 대한 기존 논의는 지수를 계산하는 방식 또는 가격 데이터 활용방식에 있다고 보는 기술적 분석이 주를 이룸
 - 재고주택 표본에 기인한 지수 계산 방식과 실거래 정보를 중심으로 계산하는 방식 간 차이, 평균 가격과 중위 가격 간 차이 등이 주요 원인으로 꼽힘
 - 이외 주택가격지수의 목적에 따라 달라지는 계산 방식에 대한 전반적인 논의는 이용만(2007) 참조
- 하지만 가격지수란 시장의 모든 정보를 하나의 숫자로 표현하는 것으로, 이 과정에서 많은 정보들이 소실되기 때문에 지수와 사람들이 체감하는 현실 간 괴리는 지수가 가지는 본질적인 한계일 수 있음
 - 설계된 지수가 통계적 오차 없이 계산되더라도, 사람들의 관심이 전체 시장이 아닌 특정 지역·가격대의 주택에 집중된다면 시장 전체의 움직임을 측정하는 주택가격 지수와 체감 가격의 변동성 간 괴리가 발생할 수 있음
 - 이는 지수가 시장 전체를 조망하도록 설계되었을 시 갖게 되는 한계임
 - 뿐만 아니라, 평균적인 움직임에서 벗어난 개별 거래 건들이 많을수록 가격지수는 시장 참여자의 인식과 동떨어진 값을 나타낼 가능성이 있음
 - 이는 지수가 평균적인 움직임을 나타내는 특성상 갖게 되는 태생적인 한계임
 - 예를 들어, 지수가 3년간 10% 상승하였더라도 거래건 중 50%는 30% 상승하고 50%는 10% 하락하였다면 지수의 움직임이 현실을 잘 반영한다고 보기 어려움

- 또한, 지수가 전체 시장을 적절히 대표하는지 알기 위해서는 주택 가격 변화율 분포에 대한 정보가 있어야 하나, 주택 가격의 분포에 대한 연구만 소수 있을 뿐 가격 변화율 분포에 대한 연구는 없음
 - 주택가격 분포를 다룬 최초의 연구들은 주택가격 분포와 소득 분포 간 관계를 다룸
 - Gyourko et al(2006), Van Nieuwerburgh and Weill (2010), Määttänen and Terviö(2010)
 - 주된 관심사는 소득의 불균형 심화가 주택 가격 차이의 심화로 이어지는지 여부임
 - Ohnishi, Mizuno, and Watanabe(2020)는 시기별로 도쿄의 주택가격 분포를 분석하여, 가격을 크기로 보정한 분포의 꼬리가 버블 시기에 두터워진다는 것을 밝힘
 - 분석기간 대부분에서 이론과 부합하게 로그 정규분포를 따르지만, 버블 시기에는 꼬리가 두터워짐
 - 위 연구들은 주택가격의 횡단면 분석을 수행하였으나, 패널데이터를 구축한 것이 아니므로 가격의 변화율 분포를 분석하지는 못함
 - 가격지수는 특정 시점의 값을 기준으로 정규화한 것이므로 지수의 수준 자체가 아니라 변화율이 중요함
 - 따라서 가격지수가 시장을 제대로 반영하는지를 알기 위해서는 가격 변화율 분포를 분석해야 하며, 이를 위해서는 패널데이터 구축이 필요함
- 뿐만 아니라, 최근 공급이 확대되고 있는 유한책임 대출과 보증금 반환보증의 위험은 시장의 평균적인 움직임을 나타내는 가격지수의 움직임으로 추정하기 어려움
 - 유한책임대출은 주택가격이 대출 잔액 이하로 떨어졌을 때 전략적 부도의 가능성이 생기므로 주택가격에 대한 풋 옵션을 판매한 것과 유사함
 - 임대인이 보증금을 반환하지 않아 대위변제가 발생했을 때, 대상 물건의

주택가격이 전세가격보다 낮다면 보증기관 입장에서는 손실이 발생할 수 있음

- 하지만 상기 금융상품의 손실은 기초자산인 주택가격의 하락폭에 비례하지 않고 특정 가격 이하부터 발생하는 비선형적인 특성을 가짐
 - 실제 손실은 가격이 평균보다 더 하락한 건들에서 발생하는 반면, 가격 지수는 평균만을 관찰하기 때문에 금융상품의 손실을 추정하는 작업에 한계가 있음
 - 예를 들어, 가격 지수는 3년간 10% 상승하여 가격 지수의 움직임만으로는 금융 상품의 손실이 예측되지 않더라도, 거래건 중 50%는 30% 상승하고 50%는 10% 하락하였다면 실제로는 손실이 발생할 가능성이 존재함
 - 가격지수가 하락하지 않았을 때도 실제로는 전세금 반환보증 사고가 발생하였으며, 이 때 가격지수가 아닌 가격 분포를 활용하여 사고율을 추정하였을 때는 실제와 유사한 움직임을 보임(민병철, 2021)
- 본 연구에서는 평균적 지수개념으로 주택시장을 이해하는 것이 아닌, 시장의 정보를 분포의 형태로 나타내어 정보의 소실을 최소화하는 방식이 시장 참여자들이 체감하는 시장을 이해하는데 보다 적합할 것이라는 관점에서 출발함
- 이를 위해 주택가격의 패널데이터를 구축하여 주택가격 변화를 분포를 살펴봄
- 일반적으로 동일 주택은 거래가 빈번하지 못해 패널데이터를 구축하기 어려우나, 한국 주택시장은 아파트 중심의 주택시장으로 같은 상품으로 간주할 수 있는 세대가 상당수 존재한다는 점에 착안하여 패널데이터 구축
- 또한, 아파트 매매시장을 주택의 크기, 가격대, 지역별로 나누되 유사한 특성을 가진 주택별로 묶어 분석함
- 크기: 생애주기에 따른 가구의 형성과정별로 소득 및 자산 수준, 가구의 규모가 달라 선호하는 주택의 크기가 다를 것이므로 주택

크기별로 수요층의 특성이 상이할 수 있음

- 가격대: 가구의 소득과 자산에 따라 살 수 있는 주택의 가격 범위가 한정되므로, 주택 가격수준에 따라 수요층의 규모 및 특성이 달라진다고 볼 수 있으며, 따라서 주택가격의 변동성의 특징 또한 달라질 가능성이 있음
- 가격대별 아파트 결정요인의 차이 분석에 관하여는 박운선·임병준(2012) 참조
- 지역: 부동산의 입지 및 주변 환경 발전 요소는 가격에 영향을 미치는 주요 요인임
- 크기별, 가격대별, 지역별로 다양한 세부 시장을 모두 관찰하고 분석하는 것은 비효율적이므로, 유사한 가격 움직임을 나타내는 세부 시장끼리 묶어 소수의 클러스터만을 분석함
- **본 연구는 주택가격의 패널데이터를 구축하여 주택가격 분포를 추정하고 이를 통해 주택시장 변동성을 분석한 최초의 연구로, 본 연구의 정량적 계산방식은 주택시장에 대한 이해뿐만 아니라 주택금융상품 리스크 측정 등에 유용하게 활용될 수 있을 것으로 기대됨**
- 유한책임대출과 보증금 반환보증은 잠재 손실액이 주택가격에 기반한 금융상품이므로, 주택가격에 대한 분석이 위험관리에 필수적 요소임
- 가격의 평균적인 움직임을 아닌 전체 움직임을 다루기 때문에 손실 발생 확률 및 크기를 분석 가능
- 실거래 데이터를 이용하여 주택가격 등락폭에 대한 분포를 구하고, 주요 특성별로 분포의 특징을 살펴봄

II. 분석 방법 및 자료

- 민병철(2021)에서 제안한 방식을 이용해 실거래 데이터 기반 주택가격의 패널데이터([표1])를 구축하고, 이를 통해 다음 두 가지를 계산함¹⁾

- 매매가격의 가격 변화 분포([표2])

- 주택 구입 시점 기준 이후 3년²⁾간 가격 변화율의 분포
- ‘06 - ’18년 사이 아파트를 구입한 사람들이 실제 경험한 3년 가격 상승률의 분포

[표1] 아파트 매매가격의 패널데이터 예시

아파트 구분	면적	층	...	2011.4	2011.5	2011.6	...
APT A	.	.	...	3,000	3,000	3,000	...
APT B	.	.		4,278	5,557	5,557	
APT C	.	.		6,500	7,583	6,833	
APT D	.	.		3,250	3,500	3,500	
APT E	.	.		6,316	6,316	6,272	

[표2] 비조건부 매매가격분포 예시

개별 매매건 거래 가격 변화율	표본수	확률
-50%이하	311	0.0%
-40%	1,765	0.2%
-30%	9,998	1.2%
-20%	46,815	5.5%
-10%	126,761	14.8%
0	225,016	26.3%
...		

- 매매가격지수 상승률에 대한 조건부 가격 변화율 분포([표3])

- 시군구 단위의 매매가격지수가 일정폭으로 변화할 때, 해당 지역에서 실제

1) 구체적인 계산 과정은 부록 참조

2) ‘12.6.29 이전에는 1세대 1주택에 대한 양도소득세 비과세 조건이 3년 이상 보유였고, 본 연구의 분석대상 기간이 ‘06년부터 ‘20년으로 절반가량의 기간이 이에 해당하므로, 평균적으로 주택 구입 후 3년은 보유한다고 가정하고 기준을 3년으로 잡음

거래된 아파트들의 가격 변화 분포

- 유한책임대출의 위험관리를 위해서는 향후 평균 주택가격 변화율에 대한 시나리오별로 손실률을 추정할 수 있어야 하는데, 이때 필요한 것은 과거 전체 기간에 대한 정보가 아닌 특정 가격 변화율에 대한 조건부 분포임
- 이후로 매매가격지수 상승률에 대한 조건부 가격 변화율 분포를 조건부 분포라 지칭하고, 이와 대비되는 첫 번째 매매가격의 가격 변화 분포는 비조건부 분포라 지칭함

[표3] 조건부 매매가격 분포 예시

개별 매매건 거래 가격 변화율	지역 매매가격 지수 변화율				
	−30%이하	−20%	...	30%	30%초과
−50%이하	3.9%	1.2%	...	0.0%	0.0%
−40%	13.4%	12.7%		0.0%	0.0%
−30%	23.5%	25.2%		0.0%	0.0%
...	...			...	
120%	0.0%	0.0%		1.2%	3.6%
130%	0.0%	0.0%		0.2%	2.2%
140%	0.0%	0.0%		0.2%	1.3%
150%	0.0%	0.0%		0.1%	0.7%

■ 위 방식으로 조건부 실거래 가격 분포를 생성하되 지역별, 크기별, 가격대별로 구분하여 도출

- 지역은 수도권·5대 광역시·기타 지역으로 나눔
- 크기: 85m² 이하, 85m²-135m², 135m² 초과
 - 일반적인 소형주택, 중형주택, 대형주택의 기준 사용
 - 1인 가구는 소형 주택을 선호할 것이고 자녀가 있는 가구는 중간-대형 주택을 선호하는 등 주택 크기별로 시장 참여자들의 특성이 다를 수 있음
- 가격은 지난 12개월간 해당 지역, 해당 크기 거래건 중 하위 30%(저가주택), 30~70%(중간 가격 주택), 상위 30%(고가주택)로 나누어 살펴봄([표4] 참고)
 - 서울에서는 상대적인 저가 주택일지라도 다른 지역에서는 고가주택으로

분류될 수 있으며, 반대로 다른 지역에서는 고가주택일지라도 서울에서는 중저가 주택일 수 있어 시도별로 저가·고가주택의 기준을 계산함

- 예시) 2017년 1월-2017년 12월 서울지역에서 거래된 85m²이하의 아파트 중 거래가 기준 하위 30% 가격을 2018년 1월 서울지역의 소형·저가 주택 기준으로 사용

[표4] 지역별·크기별 가격 수준 기준 예시
(소형, 하위 30%, 단위: 억 원)

연월	서울	인천	경기	부산	대구
2020.11	4.80	1.85	2.30	1.97	1.87
2020.12	4.85	1.85	2.30	1.90	1.88
2021.1	4.96	1.85	2.30	1.90	1.87
2021.2	5.00	1.84	2.27	1.88	1.86
2021.3	5.17	1.80	2.25	1.86	1.86

■ **계산된 지역별, 크기별, 가격대별 가격 변화율 분포를 이용해 비슷한 변동성을 보이는 특성끼리 묶는 클러스터링 분석을 수행함**

- 분포를 히스토그램의 형태로 만들고 히스토그램의 각 빈도를 변수로 사용하여 묶음
- 매매가격지수 상승률에 대한 조건부 가격변화율 분포([표5]), 비조건부 가격변화율 분포([표6])에 대하여 각각 클러스터링 분석 수행
- 조건부 가격변화율 분포에 대한 클러스터링은 지역별 매매가격지수 변화율 - 30% 이하, -20%, -10%, 0%, 10%, 20%, 30%, 30% 초과에 대한 분포를 활용함

[표5] 비조건부 분포를 이용한 클러스터링 예시

분포 특성			클러스터링 사용 변수 (개별 거래건 가격 변화율 히스토그램)				
크기	지역	가격대	...	-10%	0	10%	...
대형	수도권	고가		6.6%	14.0%	18.5%	
...				...			
중형	5대 광역시	저가		8.3%	16.1%	21.2%	
...				...			
소형	기타지역	중간가		5.9%	12.3%	16.0%	

[표6] 조건부 분포를 이용한 클러스터링 예시

분포 특성			클러스터링 사용 변수 (개별 거래건 가격 변화율 히스토그램)									
			지역 매매가격지수 변화율: -10%					지역 매매가격지수 변화율: +10%				
			..	-10%	0	10%	..	..	-10%	0	10%	..
크기	지역	가격대										
대형	수도권	고가		26.6%	9.7%	3.9%			12.2%	28.8%	24.7%	
	...			...								
중형	5대 광역시	저가		38.8%	21.1%	3.9%		..	2.8%	14.9%	33.5%	
	...			...								
소형	기타지역	중간 가		36.7%	31.4%	10.0 %			3.7%	15.3%	29.1%	

- 계층적 군집분석(Hierarchical Clustering)을 사용
 - 클러스터 개수 변화에 따른 클러스터 특성 변화를 해석하기가 상대적으로 용이
 - 두 군집이 합쳐졌을 때 군집 내 표본 간 오차제곱합(error sum of squares)을 최소화하는 와드 연결법(Ward linkage Method) 사용

■ (사용자료) 2006년 1월~2021년 3월 아파트 매매 실거래 정보

- 출처: 국토교통부 실거래가 공개시스템
- 전체 표본 수: 8,995,086
- 전체 표본 중 거래 후 3년 동안 가격이 150%를 초과하여 상승하였거나 50% 이상 하락하였다고 추정되는 건들은 일반적이지 않은 거래정보라 판단하여 제외함
- 3월까지의 자료만 사용한 2021년을 제외하면 매해 최소 45만 건의 표본이 존재함([표7])

[표7] 분석 자료의 기초통계량

연도	표본수	평균값	표준편차	최대값	최소값
2006	571,779	14,619	11,469	470,000	260
2007	465,007	13,093	10,025	310,000	300
2008	497,372	14,642	11,967	520,000	350
2009	600,249	20,205	20,036	565,000	155
2010	550,782	18,645	17,408	570,000	290
2011	614,333	20,068	16,593	438,000	320
2012	454,645	20,035	16,591	549,913	380
2013	584,208	22,096	16,870	520,000	275
2014	660,262	23,685	18,180	656,500	275
2015	730,992	26,088	19,995	770,000	350
2016	623,285	27,950	22,808	790,000	100
2017	569,389	31,387	27,203	780,000	700
2018	505,915	32,424	28,764	810,000	500
2019	534,069	34,756	33,771	840,000	550
2020	856,817	35,457	31,627	775,000	515
2021	175,982	33,791	35,080	1,150,000	400

단위: 표본 수는 건, 나머지는 만원

■ (사용자료) 한국부동산원 전국 주택가격동향조사 아파트 매매가격지수

- 기간: 2006년 1월-2021년 3월
- 출처: 한국부동산원 부동산통계정보시스템

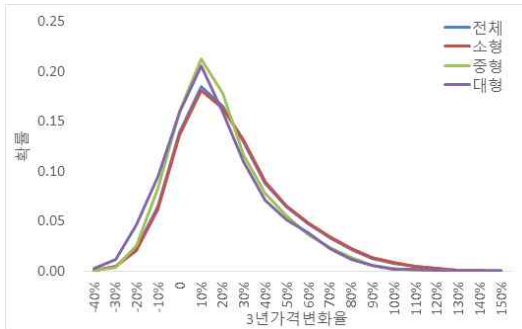
■ (한계) 본 연구의 결과는 재고주택의 분포에 대한 고려 없이 단순히 실제 거래된 건만을 바탕으로 산출하였으므로, 재고주택 표본을 기반으로 산출되는 가격지수의 적절성을 완벽하게 파악하는 데는 한계가 있음

- 하지만 주택매매 거래 발생과 대출 실행이 밀접하게 관련되어 있는 유한책임대출 등의 위험평가에는 적합함

III. 분석 결과

1. 주요 특성별 가격 변화 분석

[그림1] 주택 구입 이후 3년간 매매가격 변화를 분포(크기별)



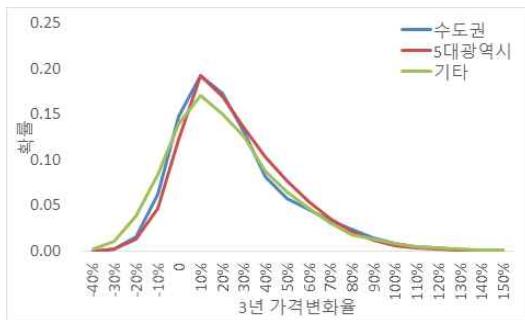
	전체	소형	중형	대형
평균	19.6%	20.4%	14.6%	12.8%
중간값	14.9%	15.6%	10.8%	8.6%
표준 편차	27.0%	27.4%	23.3%	24.8%
왜도	0.93	0.92	0.85	0.86
첨도	4.14	4.09	3.99	4.40

- (주의) 이 연구에서 도출한 매매가격 변화율 분포는 실제 거래된 매물의 3년 가격 변화율 분포로, 전체 재고주택에 대한 가격 변화율 분포와는 다름
- (전체) 분석기간 동안 아파트 가격은 구입 후 3년 동안 평균 19.6% 상승하였으며, 해당 기간 가격의 표준편차는 27.0% 수준으로 나타남 ([그림1])
- 꼬리가 오른쪽으로 길게 뻗어 있으므로 평균값이 중간값보다 커 평균을 통해 시장을 관찰하는 것은 현실을 왜곡할 우려 있음
 - 주택 크기와 무관하게 공통적으로 왜도가 0보다 커 꼬리가 오른쪽으로 길게 뻗어 있음
 - 전체 분포의 평균값은 19.6%, 중간값은 14.9%로 약 4.5%의 차이가 남
- 첨도 또한 정규분포의 첨도인 3보다 커 극단적인 가격 변화율이 발생할 확률이 다소 높음

■ (크기별) 소형 주택이 중대형 주택보다 평균 가격 상승률도 높고 가격 변동성도 큰 것으로 나타남

- 소형 주택은 구입 후 3년간 평균 20.4% 상승한 반면 대형 주택은 평균 12.8% 상승함

[그림2] 주택 구입 이후 3년간 매매가격 변화율 분포(지역별)



	수도권	5대광역시	기타지역
평균	19.8%	21.5%	18.0%
중간값	14.5%	17.2%	13.7%
표준편차	26.9%	25.3%	28.1%
왜도	1.05	0.82	0.86
첨도	4.33	3.86	4.04

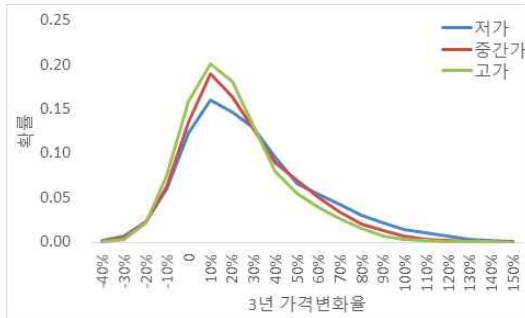
■ (지역별) 5대광역시의 평균 가격 상승률이 높고 표준편차가 작아 안정적으로 가격이 상승했던 것으로 나타남([그림2])

- 평균 상승률은 5대광역시>수도권>기타지역, 표준편차는 기타지역 > 수도권>5대광역시로 기타 지역은 평균 상승률은 낮으면서 변동성은 높았음

■ 수도권 지역의 가격 변화율 분포는 다른 지역과 비교해 왜도와 첨도가 높아, 평균과 중간값 간 차이가 클 뿐 아니라 평균에서 크게 벗어난 가격 변화율을 보이는 건들이 많음

- 주택시장의 가격 움직임을 지수로만 설명할 경우 특히 수도권에서 현실과의 괴리가 더 발생할 수 있음
 - 수도권의 평균값은 19.8%, 중간값은 14.5%로 약 5.3%의 차이가 있음

[그림3] 주택 구입 이후 3년간 매매가격 변화율 분포(가격대별)



	저가	중간가	고가
평균	24.3%	19.5%	15.8%
중간값	18.6%	15.0%	12.3%
표준 편차	31.0%	26.3%	23.4%
왜도	0.88	0.84	0.82
첨도	3.75	3.91	3.85

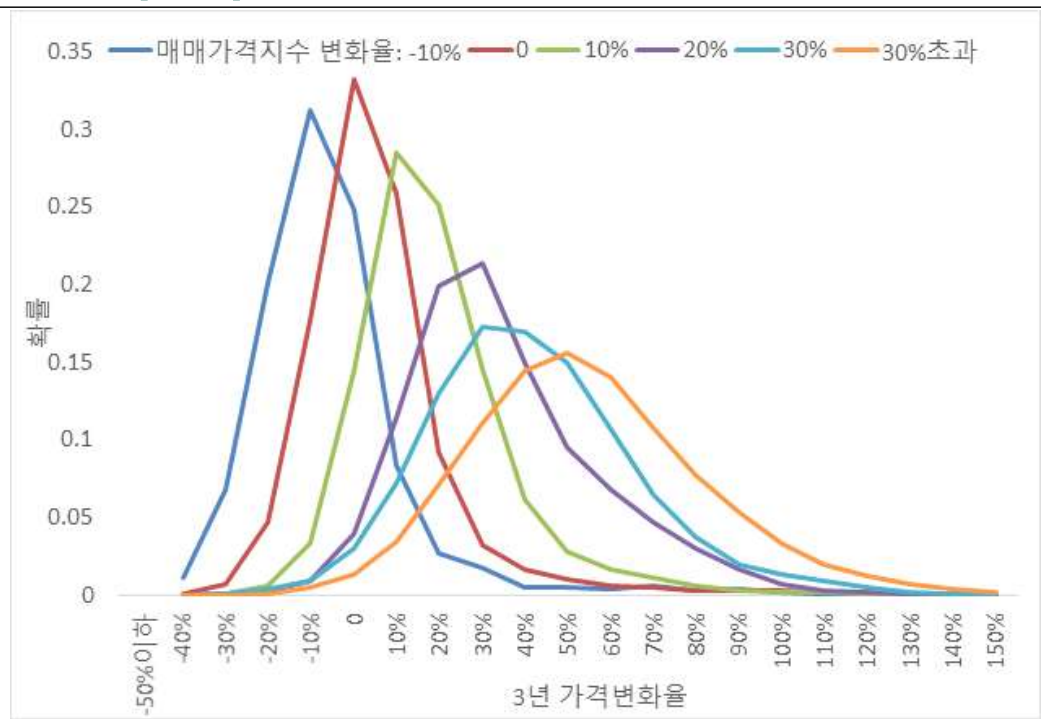
■ (가격대별) 가격 수준이 낮을수록 가격 상승률의 평균과 표준편차가 높은 것으로 나타남([그림3])

- 분석 기간 평균적으로 저가 아파트는 24.3%, 중간가 아파트는 19.5%, 고가 아파트는 15.8% 상승함
- 표준편차는 저가 아파트가 31.0%, 중간가 아파트가 26.3%, 고가 아파트가 23.4%로 가격 수준이 낮을수록 높음
- 가격 수준이 낮을수록 왜도 또한 높아, 평균값과 중간값의 차이가 큰 것으로 나타남

2. 주택 가격 지수 변화율별 조건부 가격 변동성 분석

- 주택 가격 상승기와 하락기를 구분하여 살펴보기 위해 지역별 주택 매매가격지수 변화율에 대한 조건부 매매가격 변화율 분포([그림4])와 분포의 모멘트를 도출함([표8])

[그림4] 매매지수에 대한 조건부 매매가격 변화율 분포



[표8] 조건부 분포의 모멘트

매매지수 변화율	-10%	0%	10%	20%	30%	30%초과
표본수	64,443	1,384,289	6,448,370	10,436,482	6,783,280	3,209,102
평균	-10.7%	0.7%	13.8%	29.9%	36.9%	49.8%
중간값	-13.0%	-1.9%	11.2%	26.3%	34.8%	47.6%
표준편차	18.4%	18.5%	18.2%	22.7%	24.6%	27.4%
왜도	2.21	2.38	1.50	0.82	0.60	0.46
첨도	11.87	13.27	7.89	4.03	3.86	3.35

■ 주택 가격 하락기에는 그래프 모양이 대체로 좁은 반면 상승기에는 넓게 퍼지는 모습을 볼 수 있음

- 주택가격 상승기에는 하락기에 비해 표준편차가 커, 하락기에는 대체로 같이 하락하는 반면 상승기에는 상승률 편차가 심함

- 주택가격 하락기에는 상승기에 비해 왜도와 첨도가 큰데, 이는 거래건의 대부분이 가격하락을 나타내지만 일부는 여전히 상당한 수준의 가격 증가율을 나타냈기 때문임
- **조건부 분포의 꼬리는 오른쪽으로 긴 모습을 보여 중간값이 평균값보다 낮은 모습을 보이는데, 이러한 특성은 주택가격 하락기에 좀 더 두드러짐**
- 조건부 분포의 왜도는 모두 0보다 큰데, 매매가격지수 변화율이 낮을 때 왜도가 큰 모습을 보임
- 매매가격지수가 30% 이상 상승할 때 중간값이 평균보다 2%-3% 낮는데, 지수가 10% 하락할 때도 중간값이 평균보다 2.3% 낮음
 - 중간값과 평균값의 차이가 주택가격 하락기에 더 두드러짐
 - 주택가격 하락기에는 평균값인 지수가 측정하는 주택가격 하락폭보다 시장이 체감하는 하락폭은 더 클 수 있음
- **주택매매가격지수가 상승할 때도 여전히 주택가격이 하락하는 비중이 상당히 있는 것을 관찰할 수 있음**
- 매매가격지수가 0-10% 상승할 때는 18.5%, 10-20% 상승할 때는 5.2%가 가격 하락을 경험함

3. 주택 특성 클러스터링 결과

- 이 절에는 주택 특성별로 27개 분류(주택 크기(3)*가격대(3)*지역(3)=27)로 나누어 가격 변화율 분포를 계산하고, 이를 토대로 클러스터링한 결과를 나타냄
- 27개 분류의 특성을 일일이 분석하는 것은 비효율적임
- 먼저 비조건부 분포를 이용하여 클러스터링을 한 후, 조건부 분포를 이용한 클러스터링도 수행함
- 비조건부 분포를 이용한 클러스터링은 과거에 발생한 주택가격 움직임이

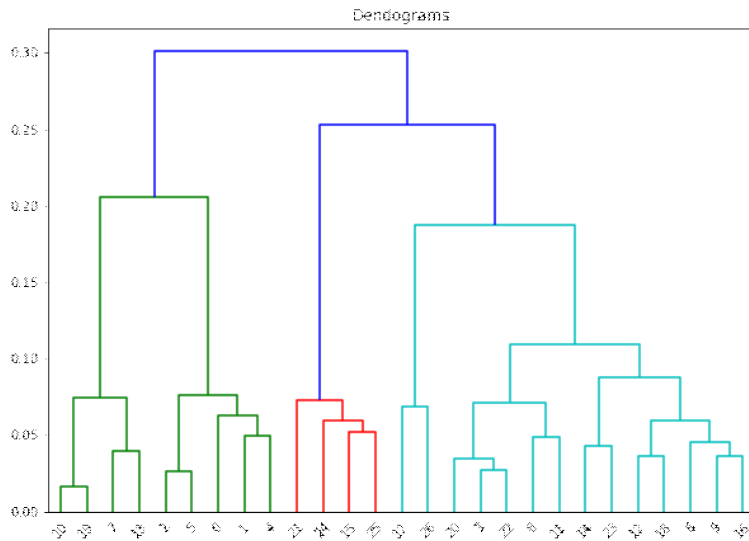
향후에 재현될 것이라고 가정하였을 때 활용 가능

- 조건부 분포를 이용한 클러스터링은 미래 주택가격 변화율에 대한 시나리오가 있을 때 예측 목적으로 활용 가능
- 지역별 가격 변동률 수준을 통제하였다고 볼 수도 있음

3.1 비조건부 분포를 이용한 클러스터링 결과

- [그림5]는 전체 비조건부 분포를 통해 클러스터링한 결과의 덴도그램을 나타내며, 이를 토대로 본 연구에서는 클러스터를 세 개로 결정하고 분석을 수행함([표9])
- 27개 분포를 유사도에 따라 하나씩 묶어가는 그림을 나타내며, 세로선의 길이가 길수록 인접 클러스터와의 차이가 크다는 것을 나타냄
- 클러스터를 자세하게 나눌수록 아파트 시장을 보다 복합적으로 살펴볼 수 있는 장점이 있지만 결과의 간결성이 떨어지게 됨
- (주의) 같은 클러스터로 묶인다는 것은 장기적으로 관측된 변동성의 특성이 비슷하다는 뜻이지, 동조화된 움직임을 보인다는 것은 아님

[그림5] 비조건부 분포의 클러스터링 과정을 나타낸 덴도그램

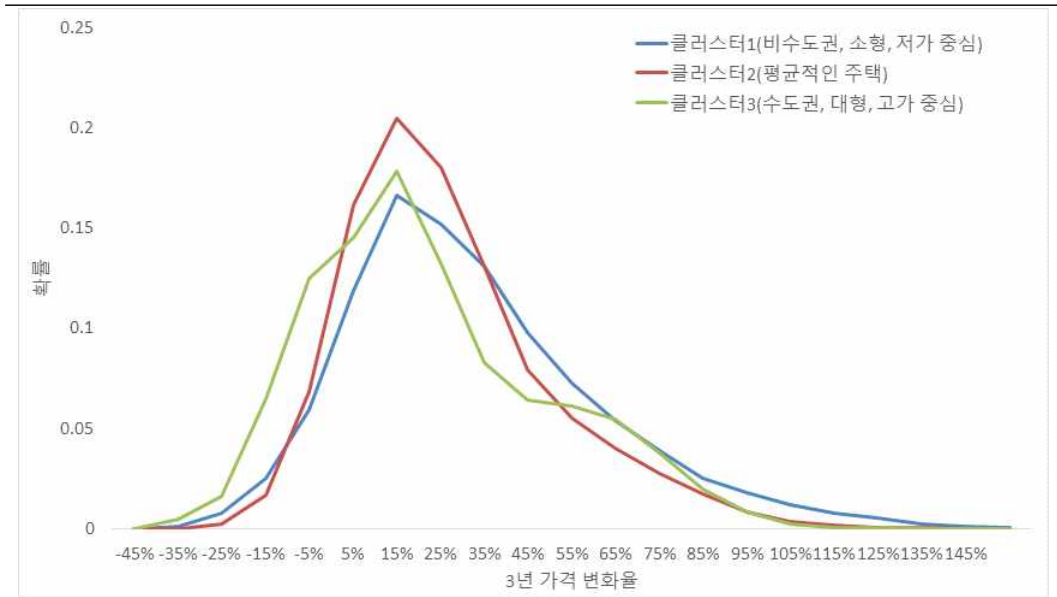


- 클러스터 분석 결과, 평균 가격 상승률 및 변동성이 큰 클러스터 1, 중간 수준의 변동성을 나타내는 클러스터 2, 평균 변동성은 낮으나 크게 하락할 확률이나 크게 상승할 확률이 높은 클러스터 3으로 나누어짐([그림6], [표10])

[표9] 비조건부 분포의 클러스터링 결과

클러스터	크기	가격	지역
클러스터1	소형	저가	수도권
	소형	저가	5대광역시
	소형	저가	기타
	소형	중가	5대광역시
	소형	중가	기타
	소형	고가	5대광역시
	중형	저가	5대광역시
	중형	중가	5대광역시
	대형	저가	5대광역시
클러스터2	소형	중가	수도권
	소형	고가	수도권
	소형	고가	기타
	중형	저가	수도권
	중형	저가	기타
	중형	중가	수도권
	중형	중가	기타
	중형	고가	5대광역시
	중형	고가	기타
	대형	저가	수도권
	대형	저가	기타
	대형	중가	5대광역시
	대형	중가	기타
	대형	고가	기타
클러스터3	중형	고가	수도권
	대형	중가	수도권
	대형	고가	수도권
	대형	고가	5대광역시

[그림6] 클러스터별 비조건부 분포



[표10] 클러스터별 비조건부 분포의 모멘트

	클러스터1 (비수도권, 소형, 저가 중심)	클러스터2 (평균적인 주택)	클러스터3 (수도권, 대형, 고가 중심)
평균	22.9%	16.6%	13.0%
표준편차	29.3%	23.7%	27.5%
왜도	0.85	0.93	0.61
첨도	3.87	4.08	2.88

■ (클러스터1) 비수도권, 작은 크기 및 저가 중심, 높은 가격 상승률과 높은 변동성

- 대체로 비수도권, 중·소형, 중·저가 주택들이 포함됨
- 평균 가격 상승률이 22.9%로 가장 높은 동시에 표준편차도 29.3%로 가장 높음
- 변동성은 높지만 크게 하락한 비중은 상대적으로 낮으며, 50% 이상 크게 상승한 비중은 높음

- (클러스터2) 전국, 소형 중 고가 · 중형 · 대형 중 저가 중심, 중간 수준의 가격 상승률과 변동성
 - 평균 가격 상승률은 16.6%로 전체 평균인 19.6%와 유사하며, 표준편차는 23.7%로 전체 평균인 27.0%보다 다소 낮음
 - 침도가 4.08로 가장 높아, 분포가 대체로 평균을 중심으로 좁게 뭉쳐 있음
- (클러스터3): 대도시의 대형 · 고가 주택
 - 평균 가격 상승률은 13.0%로 가장 낮지만, 표준편차는 27.5%로 높은 편임
 - 평균 상승률은 낮지만, 50%이상 큰 폭 상승한 비율은 클러스터 2보다 높음
 - 반면, 주택가격이 하락한 비중 또한 다른 클러스터에 비해 높음

3.2 조건부 분포를 이용한 클러스터링 결과

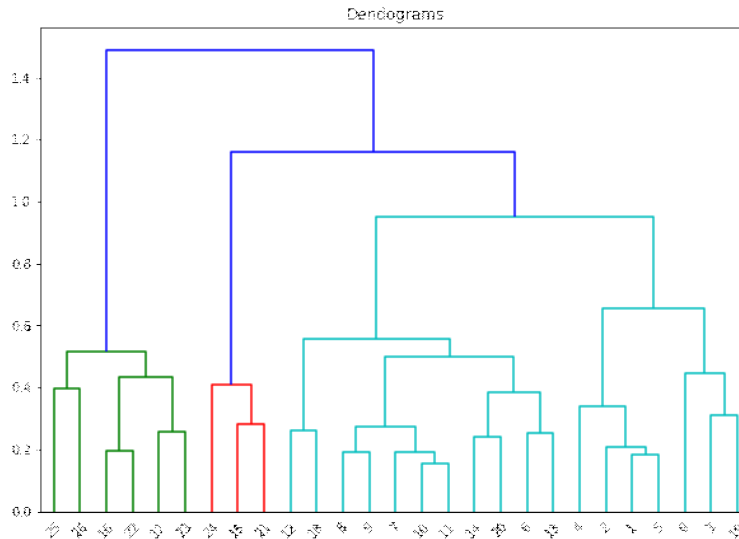
- 비조건부 분포를 이용한 클러스터링은 과거 지역별 주택가격 변화율을 전제로 한 것으로, 향후 지역별 주택가격 변화율 추이가 달라진다면 현실에 부합하지 않을 수 있음
- 조건부 분포를 이용한 클러스터링은 지역별 주택가격지수에 대하여 통제를 한 것과 같은 효과를 가짐
- [그림7]은 전체 조건부 분포를 통해 클러스터링한 결과의 덴도그램을 나타내며, 이를 토대로 본 연구에서는 클러스터를 네 개로 결정하고 분석을 수행함([표11])
 - 네 개로 나눈 후 다섯 번째 클러스터의 특성을 분석한 결과 변동성이 높거나 낮은 클러스터가 추가적으로 세분화되는 것이 아닌 평균적인 클러스터가 나뉘는 것이어서 큰 의미가 없다고 보고, 보고서에는 관련

내용을 포함시키지 않음

■ 비조건부 분포를 이용한 클러스터링 결과와는 달리, 대형 주택을 제외하면 대체로 지역보다는 주택 크기별로 클러스터링 됨

- 클러스터1: 소형 주택 중심
- 클러스터2: 중형 주택 중심
- 클러스터3: 수도권 대형 주택 중심
- 클러스터4: 비수도권 대형 주택 중심

[그림7] 조건부 분포의 클러스터링 과정을 나타낸 덴도그램



[표11] 조건부 분포의 클러스터링 결과

클러스터	크기	가격	지역
클러스터1	소형	저가	수도권
	소형	저가	5대광역시
	소형	저가	기타
	소형	중가	수도권
	소형	중가	5대광역시
	소형	중가	기타
	대형	저가	5대광역시
클러스터2	소형	고가	수도권
	소형	고가	5대광역시
	소형	고가	기타
	중형	저가	수도권
	중형	저가	5대광역시
	중형	저가	기타
	중형	중가	수도권
	중형	중가	5대광역시
	중형	중가	기타
	대형	저가	수도권
클러스터3	대형	저가	기타
	중형	고가	수도권
	대형	중가	수도권
클러스터4	대형	고가	수도권
	중형	고가	5대광역시
	중형	고가	기타
	대형	중가	5대광역시
	대형	중가	기타
	대형	고가	5대광역시
클러스터4	대형	고가	기타

- 클러스터별로 조건부 매매가격 변화율 분포를 살펴보면, 매매가격지수가 큰 폭으로 상승하거나 하락할 때는 클러스터별로 가격 변화율의 차이가 두드러지지만 매매가격지수의 변화폭이 작을 때는 클러스터별 차이가 미미함([그림8], [표12], [표13])
- 매매가격지수가 3년간 20%~30% 증가할 때는 클러스터별 평균 가격 증가율 차이가 최대 약 30%, 매매가격지수가 0%~10%변할 때는 최대

약 15%정도임

■ (클러스터1: 소형주택) 높은 평균 가격 상승률 및 편차

- 지역 매매가격지수가 10%, 20%, 30% 상승할 때 평균 15.6%, 31.2%, 42.9% 상승하여 지역 평균보다 높은 상승폭을 보임
- 지역 매매가격지수가 상승할 때 뿐 아니라 하락할 때도 높은 가격 편차를 보임
- 다른 클러스터들은 지역 매매가격지수가 상승할 때 가격 편차가 크고 하락할 때는 편차가 작음

■ (클러스터2: 평균적인 주택) 평균 수준의 가격 상승률 및 편차

- 지역 매매가격지수가 10%, 20%, 30% 상승할 때 평균 11.5%, 28.8%, 31.1% 상승하여 지역 평균과 유사한 수준의 상승폭을 보임
- 가격 상승률의 편차 또한 지역 매매가격지수가 상승할 때 높아지고 하락할 때 낮아지는 전형적인 모습

■ (클러스터3: 수도권 대형주택) 지역 가격지수 상승폭보다 높은 평균 가격 상승률, 하락폭보다 높은 평균 가격 하락률

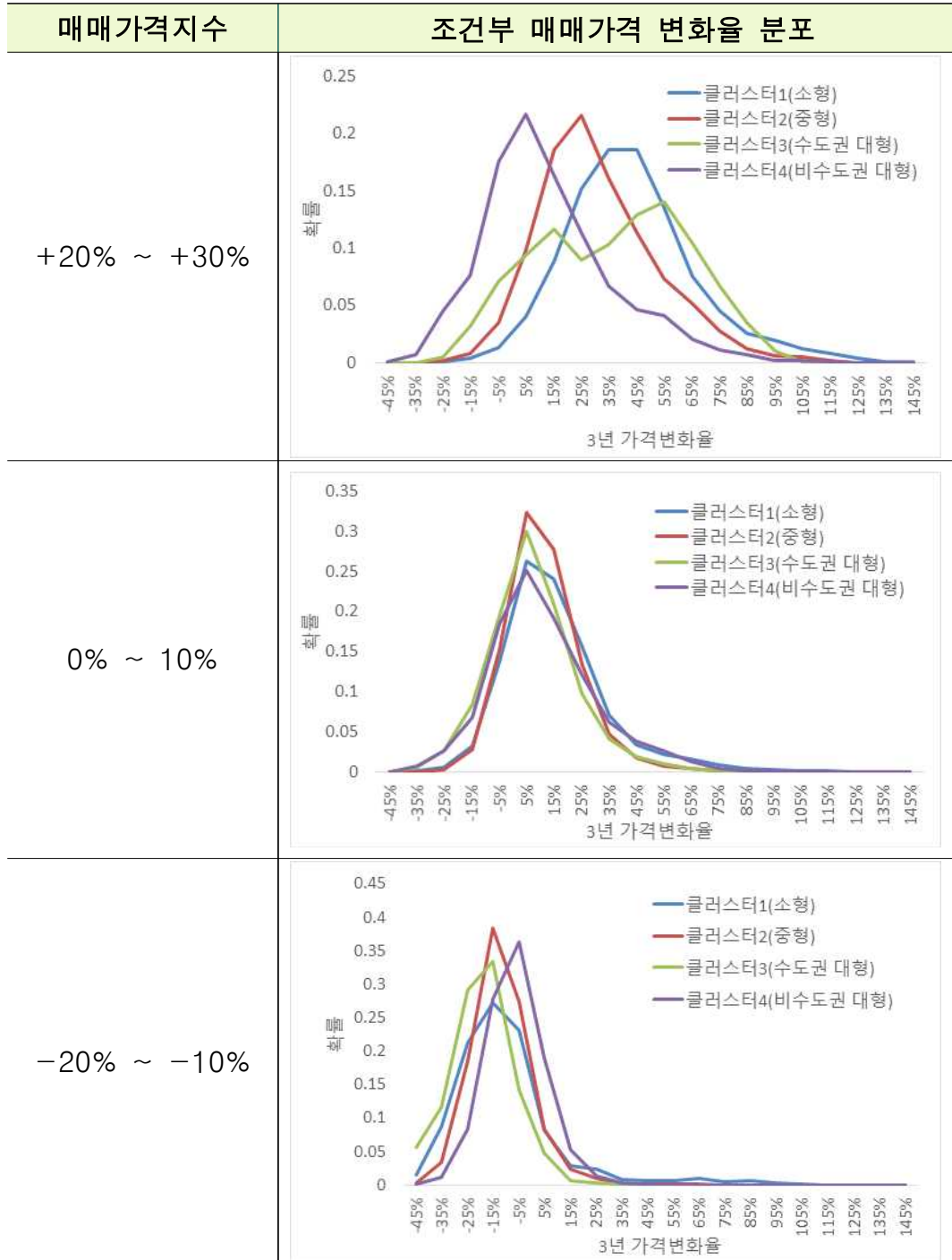
- 지역 매매가격지수가 30% 상승 할 때 평균 36.5% 상승하며 높은 상승률은 보이나, 분포의 모양은 두 개의 피크를 가지며 가격이 하락하는 건의 비중도 상당한, 양극화된 모습을 나타냄
- 지역 매매가격지수가 10% 하락할 때 평균 19.2% 하락하며 가장 낮은 하락폭을 보임
- 전체적으로 오를 때 많이 오르고 내릴 때 많이 내리는 등 변동성이 높은 특징을 지님

■ (클러스터4: 비수도권 대형주택) 지역 가격 지수 상승폭보다 낮은 평균 가격 상승률, 하락폭보다 낮은 평균 가격 하락폭

- 지역 매매가격지수가 10%, 20%, 30% 상승할 때 평균 11.3%, 17.3%, 13.0% 상승하여 지역 평균과 비슷하거나 낮은 수준의 상승폭을 보임

- 하지만 지역 매매가격지수가 10% 하락할 때 평균 6.3%만 하락하여 4개 클러스터 중 가장 낮은 하락폭을 나타냄
- 전체적으로 오를 때 적게 오르고 내릴 때 적게 내리는 등 변동성이 낮은 특징을 지님
- **클러스터별 평균 가격 상승률 수준은 앞서 살펴본 주택 크기별 가격 변동률 특성과 일치하나, 수도권 대형주택 중심인 클러스터3은 이와 별개로 높은 가격 변동성을 보임**
 - 평균 가격 상승률 수준은 클러스터1(소형주택 중심)>클러스터2(중형 중심)>클러스터4(비수도권 대형주택)로 대체로 주택 크기와 관련이 있음
 - 수도권 대형주택 중심인 클러스터3은 일반적인 크기에 따른 변동성 특성을 크게 벗어나, 비수도권 대형주택 중심인 클러스터 4와는 반대로 평균보다 높은 변동성 수준을 나타냄
 - 비수도권 대형주택(클러스터4)는 가격 상승기에 평균보다 적게 오르고 하락기에 평균보다 적게 떨어졌으나, 수도권 대형주택(클러스터3)은 가격 상승기에 평균보다 많이 오르고 하락기에 평균보다 많이 내렸음

[그림8] 클러스터별 조건부 매매가격 변화율 분포



[표12] 조건부 분포의 클러스터별 평균 가격 변화율 (단위: %)

매매가격지수 변화율	-10%	0%	10%	20%	30%
클러스터1(소형)	-9.9	3.1	15.6	31.2	42.9
클러스터2(중형)	-11.9	-2.3	11.5	28.8	31.1
클러스터3(수도권 대형)	-19.2	-12.2	7.4	31.4	36.5
클러스터4(비수도권 대형)	-6.3	1.6	11.3	17.3	13.0

[표13] 조건부 분포의 클러스터별 가격 표준편차 (단위: %)

매매가격지수 변화율	-10%	0%	10%	20%	30%
클러스터1(소형)	21.6	21.4	19.9	22.5	24.0
클러스터2(중형)	11.7	12.1	14.5	22.2	22.1
클러스터3(수도권 대형)	12.4	12.8	16.4	25.5	27.5
클러스터4(비수도권 대형)	11.6	14.2	20.2	22.9	24.0

IV. 요약 및 결론

- 본 연구에서는 실거래 데이터를 이용해 최초로 개별 주택에 대한 패널 데이터를 구축하고, 이를 통해 가격 변화율 분포를 도출하여 기존 가격 지수로는 관찰하지 못하였던 주택시장의 특성을 분석하였음
- 일반적으로 한 주택에 대하여 매매가 빈번하지 않으므로 개별 주택 가격의 시계열은 만들기 어려움
- 한국 주택시장은 아파트 중심의 주택시장으로 같은 특성을 지녀 가격이 같다고 간주할 수 있는 세대가 대량으로 존재하므로 이를 이용해 개별 주택가격의 패널데이터를 구축함
- 가격지수로는 관찰할 수 없었던 주택시장의 주요 특성들은 다음과 같음
 - 1. 매매가격 상승률은 꼬리가 오른쪽으로 긴 모습을 보여 평균값이 중간값보다 크므로, 특히 주택가격 하락기에 평균값을 이용하여 시장을 설명하면 체감 시장과의 괴리가 발생할 수 있음
 - 2. 지역 매매가격지수가 하락할 때는 개별 거래건의 가격 편차가 작는데 반해 지수의 상승폭이 커질수록 개별 거래건의 편차가 커져, 하락할 때는 같이 하락하고 오를 때는 차별적으로 오르는 특징이 관찰됨
 - 3. 주택매매가격지수가 상승할 때도 여전히 주택가격이 하락하는 비중이 상당히 있는 것을 관찰할 수 있어, 매매가격지수가 0-10% 상승할 때는 18.5%, 10-20% 상승할 때는 5.2%가 가격 하락을 경험함
- 클러스터링 기법을 이용하여 변동성의 특성이 유사한 주택 특성끼리 묶은 결과, 대체로 중간 크기의 주택보다는 소형 주택이나 대형 주택에서 높은 변동성이 나타난 빈도가 높았음
- 분석기간 비수도권의 중·소형 중·저가 주택의 가격 변동성이 컸으며

대도시의 대형·고가 주택은 평균 변동성은 낮지만 가격이 크게 오를 확률은 높았던 것으로 파악됨

- 비조건부 분포를 이용한 클러스터링

- 지역별 주택가격 변화율이 같다고 가정하면, 수도권의 중·대형, 중·고가 주택들은 지역 매매가격지수가 오를 때는 평균보다 더 오르고 지수가 내릴 때는 평균보다 더 내리는 높은 변동성을 가진 것으로 나타남

- 조건부 분포를 이용한 클러스터링

- 본 연구는 경험적으로 체감하고 있었으나 기존 가격지수로는 설명하기 어려웠던 주택시장의 현상들을 정량적으로 설명하였으며, 이러한 정량적 계산방식은 주택시장에 대한 이해뿐만 아니라 주택금융상품 리스크 측정 등에 유용하게 활용될 수 있을 것으로 기대됨

[부록] 분석 과정

- 이 연구에서는 민병철(2021)에서 제안한 방식을 통해 매매가격 변화율의 분포와 매매가격지수에 대한 조건부 실거래 가격 변화율 분포를 생성함

- 매매가격 변화율에 대한 경험적 누적분포함수(empirical cumulative distribution function) $F(x)$ 는 다음과 같음:

$$F(x) = \frac{\sum_i \sum_j \sum_t n_{i,j,t} I(r_{i,j,t} \leq x)}{\sum_i \sum_j \sum_t n_{i,j,t}}$$

- $r_{i,j,t}$: i 지역, j 아파트 단지, t 시점에 구매한 주택의 향후 3년 매매가격 변화율
- $r_{i,t}$: i 지역의 t 시점으로부터 3년간 매매가격지수 변화율
- $n_{i,j,t}$: i 지역, j 아파트 단지, t 시점의 거래량
- $r_{i,j,t}$ 을 t 시점에 구매한 주택의 향후 3년간 실현된 가격 변화율로 정의하면, $F(x)$ 는 분석 기간 임의의 시점에 주택을 샀을 때 기대할 수 있는 주택가격 변화율의 분포를 의미함
- k 번째 매매가격 구간에 대한 조건부 실거래 가격분포의 누적확률분포는 다음과 같음:

$$F_k(x) = \frac{\sum_i \sum_j \sum_t n_{i,j,t} I(l_{idx,k} < r_i \leq u_{idx,k}) I(r_{i,j,t} \leq x)}{\sum_i \sum_j \sum_t n_{i,j,t} I(l_{idx,k} < r_i \leq u_{idx,k})}$$

- 여기서, $I(\text{조건})$ 식은 조건을 만족할 경우는 1, 아니면 0임을 의미함
- 매매가격지수 분석 대상 구간의 상한과 하한:

$$[u_{idx,1}, u_{idx,2}, \dots, u_{idx,n_{idx}}], [l_{idx,1}, l_{idx,2}, \dots, l_{idx,n_{idx}}]$$

- n_{idx} : 매매가격지수 분석 대상 구간의 개수
- ex: 실제 사용되는 구간은 다음과 같음

$$l_{idx,1} < r_i \leq u_{idx,1}, l_{idx,2} < r_i \leq u_{idx,2}, \dots, l_{idx,n_{idx}} < r_i \leq u_{idx,n_{idx}}$$
- ex) 시군구 단위의 지역 매매가격지수가 3년간 5%~10% 하락했을 때, 같은 기간 동안 거래 이후 가격이 25% 이상 하락한 비중(해당 가격 하락폭이 네 번째 분석 대상 구간일 경우):

$$- F_4(-0.25) = \frac{\sum_i \sum_j \sum_t n_{i,j,t} I_{(-0.1 < r_i \leq -0.05)} I_{(r_{i,j,t} \leq -0.25)}}{\sum_i \sum_j \sum_t n_{i,j,t} I_{(-0.1 < r_i \leq -0.05)}}$$

■ 위 방식으로 조건부 실거래 가격 분포를 생성하되, 가격대별, 지역별, 크기별로 나눠서 도출

- 가격대: 해당 지역의 지난 12개월 거래가 상위 30%, 30-70%, 70%-100%
 - 고정된 금액을 기준으로 삼을 경우, 2006년에는 고가로 분류되었던 주택의 가격이 2021년에는 중고가로 분류되는 사례가 있을 수 있어 고가주택으로서의 기준이 일관되지 못함
- 지역은 수도권, 5대광역시, 기타 지역으로 나눔
- 크기: 85m² 이하, 85m²-135m², 135m² 초과
- 이러한 방식으로 계산된 분포는 가격지수가 일정폭 변할 때 해당 지역에서 실제 발생한 개별 거래건의 가격 변화폭을 나타냄

참 고 문 헌

- 민병철(2021), “전세가격 변동률 분포를 활용한 역전세 위험의 측정”, 부동산학연구 27(2), 2021.6, pp. 63-75
- 박운선, 임병준(2012), “지역 및 가격대별 아파트가격결정요인의 차이 분석”, 주거환경 10(1), pp. 91-113
- 이용만(2007), “주택가격지수의 목적과 방법을 둘러싼 쟁점 - 실거래가격에 기초한 지수를 중심으로-”, 부동산학연구 13(3), pp. 147-167
- Gyourko, J., C. Mayer and T. Sinai (2013) “Superstar Cities“, American Economic Journal: Economic Policy, 5(4), pp. 167-99.
- Ohnishi, T., Mizuno, T., & Watanabe, T. (2020). “House Price Dispersion in Boom-Bust Cycles: Evidence from Tokyo”, The Japanese Economic Review, 71(4), 511-539.
- Määttänen, N. and M. Terviö (2014) “Income Distribution and Housing Prices: An Assignment Model Approach“, Journal of Economic Theory, 151, pp. 381-410.
- Van Nieuwerburgh, S. and P. O. Weill (2010) “Why Has House Price Dispersion Gone Up?“, Review of Economic Studies, 77(4), pp. 1567-1606.
- 국토교통부 실거래가 공개시스템, rt.molit.go.kr
- 한국부동산원 부동산통계정보시스템, r-one.co.kr