

주택시장 분석에 있어서 Lee-Carter 모형의 유용성 연구

마승렬*

<요약>

본 연구는 주택시장의 예측에 있어서 Lee-Carter모형 (LC모형)의 유용성을 평가하기 위해 아파트매매거래량을 대상으로 LC모형과 함께 VAR모형을 비교모형으로 설정하여 분석하였으며, 각각의 모형에 의한 예측치와 실제값을 비교하여 모형별 예측력을 평가하였다. 본 연구의 분석결과에 의하면 LC모형의 예측력이 VAR모형에 비해 열등하지는 않은 것으로 나타났다. 추가적 분석에서 AR모형 또는 ARCH모형 등 단변량 시계열예측모형을 이용하여 분석한 결과도 LC모형에 의한 예측치에 비해서는 더 나은 결과를 보여주지 못하였다. LC모형은 다른 모형에 비해 열등하지 않은 예측력을 보여 주면서도 모형의 단순성과 추정의 용이함으로 인해 주택시장의 실무에서 편리하게 사용할 수 있는 모형이 될 수 있을 것으로 기대된다. 한편, LC모형으로 추정된 16개 지역별 아파트매매거래량 예측치의 합계는 VAR모형의 예측치와는 달리 전국 아파트매매거래량의 단변량 예측치와 거의 유사한 값을 제공해 주었다. 이러한 관점에서도 LC모형은 아파트매매거래량 예측치를 실무에서 활용하는 경우 다른 예측모형에 비해 더 큰 장점을 가진다.

핵심주제어: 주택시장, 예측모형, 예측력, 아파트매매거래량, LC모형

I. 서론

주택시장에서 주택매매거래량은 주택매매가격 등과 함께 주택경기를 진단하는데 있어서 매우 중요한 지표중 하나이다. 예를 들어 주택시장 경기를 진단하는 모형 중 하나인 벌집순환모형(Honeycomb cycle model)에서는 주택매매거래량과 주택매매가격의 시계열자료만을 이용하여 주택매매거래량의 증가와 감소, 주택매매가격의 상승과 하락 등 양 시계열간의 조합에 근거하여 주택경기의 호황, 침체, 불황, 회복 등의 국면을 진단하고 예측한다(Janssen et al., 1994; Festa et al., 2012; 마승렬, 2016 등 참조). 전국의 주택시장 경기를 합리적으로 진단하고 전망하기 위해서는 전국의 주택매매거래량과 주택가격 등의 향후 추이 전망이 요구되고, 지역별로 세분화된 진단과 전망을 위해서는 지역별로 구분된 주택매매거래량과 주택가격 등의 향후 추이 전망이 요구된다.

주택가격과 거래량간의 장기적 관계를 규명하고자 시도된 기존의 실증 연구들을 살펴보면, Leung and Peng(2005)은 상업용빌딩(commercial building)의 경우 거래량이 가격을 단일 방향(one-way)으로 그랜저인과 한다는 분석결과를 보여주었으며 Clayton et al.(2010)은 주택가격과 거래량 간에는 양 방향(two-way)의 그랜저인과관계가 있다(Granger causing)는 분석결과를 보여주었다. Stein(1995)과 Genesove and Mayer(2001)의 연구에서는 주택가격이 하락하면 거래량이 감소한다는 분석결과를 보여주었다. 한편, 한국의 주택시장에서 임재만(2011)은 주택가격과 거래량간에는 양(+)의 상관관계가 있고 상호 양 방향으로 영향을 미친다는 분석결과를 보여주었고 임대봉(2014)은 아파트가격과 거래량간에는 선·후행 관계 또는 동행관계가 있는데 지역별 차이가 있음을 보여주었다.

본 연구는 주택경기를 진단하는데 사용되는 다양한 지표들 중 주택매매거래량 시계열을 대상으로 동 시계열의 합리적 예측을 위한 모형을 검토하며, 예측모형 중 LC모형(Lee-Carter model)의 유용성을 확인하고자 하는데 연구의 목적이 있다. 시계열의 미래값은 다양한 모형으로 추정해 볼 수 있다. 본 연구는 전국의 지역을 16개 시·도 지역으로 구분하여

분석하는데, 16개지역의 주택매매거래량을 동시에 예측할 수 있는 모형으로서 LC모형과 함께 VAR모형(vector autoregressive model)을 비교모형으로 설정하여 분석하며, 각각의 모형에 의한 주택매매거래량 예측치와 실제값을 비교하여 모형별 예측력을 비교한다.

LC모형은 횡단면시계열 자료인 연령별-사망률 프로파일을 이용하여 사망률의 예측에 주로 사용되는 모형으로서 본 연구에서와 같이 사망률 이외의 횡단면시계열 자료 예측에 LC모형을 사용하여 분석한 사례는 매우 미미하여 마승렬(2015), Suwanvijit et al.(2011) 등 몇몇 연구 외에는 분석결과를 찾아볼 수 없다. 마승렬(2015)은 공무원의 사망시점 연령(30세, 40세, 50세)에 따른 미래의 중단면 연령-소득프로파일을 예측하기 위해 LC모형을 이용하여 분석한 바 있으며, Suwanvijit et al.(2011)은 태국의 14개 지역별 탄산음료 판매수익의 장기예측을 위해 LC모형을 활용한 바 있다. 본 연구의 주택시장 분석에서 16개 지역별 아파트매매거래량을 개별적으로 예측하고자 하는 경우에도 Suwanvijit et al.(2011)의 연구에서 처럼 LC모형을 활용하면 16개 지역별 아파트매매거래량 예측치를 동시에 추정할 수 있을 것이다.

LC모형 이외에 16개 지역별 아파트매매거래량 예측치를 동시에 추정할 수 있는 예측모형으로 우리는 VAR모형을 고려해볼 수 있다. 지역별 아파트매매거래량은 서로 독립적으로 움직이는 것이 아니라 일정한 상관관계를 가지면서 움직이는 것으로 볼 수 있다. 즉 주택시장에 미치는 거시적 환경변화는 지역별로 각각 직접적인 영향을 미칠 뿐만 아니라 지역간 상관관계에 따라 상호 간에도 영향을 미치는 것으로 볼 수 있다. 이 경우 유용하게 활용할 수 있는 예측모형이 VAR(p)모형인데 단변량 모형과는 달리 자기 지역의 아파트매매거래량 시차변수뿐만 아니라 다른 지역의 시차변수들도 설명변수에 함께 포함하여 분석함으로써 예측의 합리성을 높일 수 있게 된다.

본 연구는 2006.01-2018.10까지의 지역별 월별 아파트매매거래량 시계열을 이용하여 분석하는데 모형의 추정에는 2006.01-2016.12까지의 자료를 이용하여 분석하며, 추정된 파라미터를 이용하여 2017.01-2018.10까지

의 시계열을 예측한다. LC모형과 VAR모형으로 예측한 16개 지역별 아파트매매거래량과 실제 아파트매매거래량을 비교함으로써 우리는 양 모형의 예측력을 비교해볼 수 있을 것이다.

한편, 전국 16개 지역의 개별적 아파트매매거래량을 합산하면 이는 전국 아파트매매거래량이 된다. 이때 지역별 예측치를 합산한 값이 단변량인 전국의 매매거래량을 대상으로 예측한 전국 매매거래량의 예측치와 비교하였을 때 그 크기에 있어서 분석에 사용한 예측모형에 따라 현저한 차이를 보여주는 경우가 많다. 따라서 단변량인 전국의 매매거래량을 대상으로 예측한 예측치와 전국 16개 지역의 개별적 아파트매매거래량 예측치를 합산한 값을 상호 비교해볼 필요가 있다.

실무 상 주택시장의 분석 및 진단은 전국 아파트매매거래량을 대상으로 수행되는 경우도 있고, 분석 목적에 따라서는 전국 16개 지역의 개별적 아파트매매거래량 예측치가 필요한 경우도 있다. 따라서 단변량인 전국의 매매거래량을 대상으로 예측한 전국 매매거래량의 예측치와 전국 16개 지역의 개별적 아파트매매거래량을 합산한 값이 현저한 차이를 보인다면 이들 예측치를 실무에서 활용할 때 전국 매매거래량의 예측치와 전국 16개 지역의 개별적 아파트매매거래량 예측치 모두 외적 타당성을 가지지 못할 수 있다.

본 연구는 단변량인 전국의 아파트매매거래량을 예측하기 위해 추가적으로 ARCH모형(autoregressive conditional heteroscedasticity model)에 의한 전국 예측치를 도출한 후 LC모형 또는 VAR모형으로 예측한 전국 16개 지역의 개별적 아파트매매거래량 예측치를 합산한 값과 상호 차이를 비교한다. 이때 양자간에 차이가 없거나 또는 거의 유사한 값을 보여주는 예측모형이 있다면 이러한 모형이 주택시장의 분석에 있어서는 상대적으로 더 유용한 모형이 될 수 있을 것이다.

본 연구의 구성은 다음과 같다. 먼저 제1장 서론에 이어 제2장에서는 본 연구의 분석에 사용하는 아파트매매거래량 자료와 아파트매매거래량의 예측에 사용할 예측모형인 LC모형, VAR모형, ARCH모형 등에 대하여 설명한다. 제3장에서는 예측모형별 예측치를 도출하고, 모형별 예측력

을 상호 비교하는 분석결과를 제시하고 마지막 제4장에서 본 연구의 결론을 도출한다.

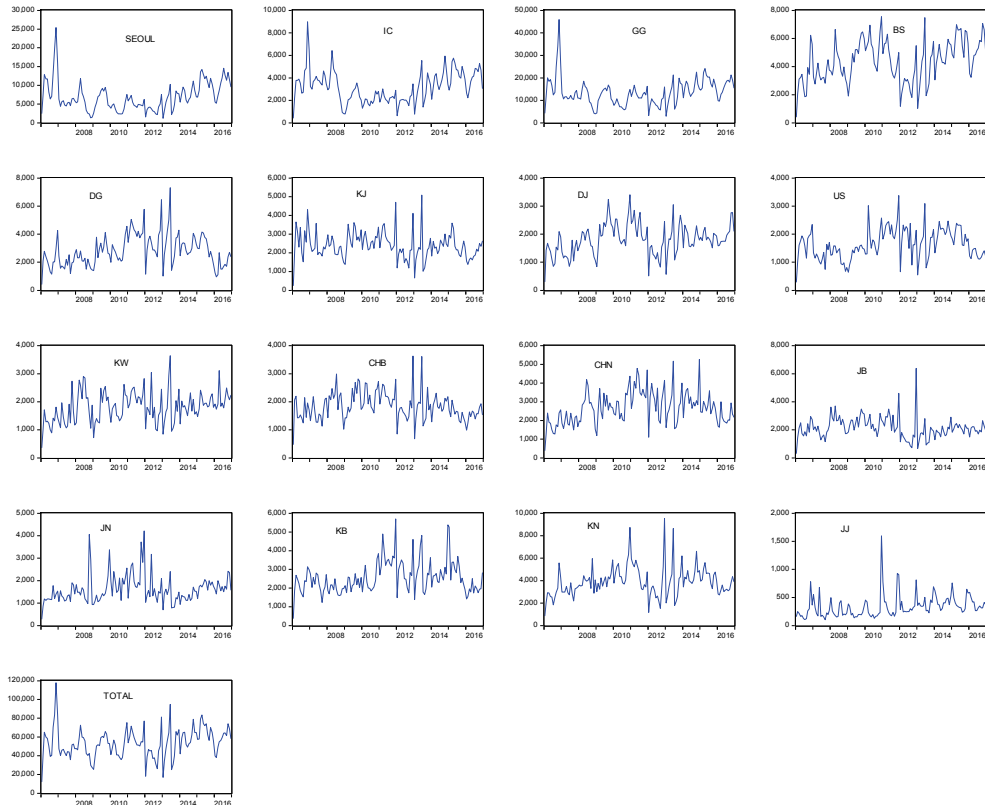
II. 자료 및 분석모형

1. 자료

본 연구에서 분석에 사용한 아파트매매거래량 자료는 국토교통부 DB의 2006.01-2018.10까지의 월별 자료이다. <그림 1>은 전국 및 16개 개별지역의 2006.01-2018.10간의 지역별 아파트매매거래량을 나타낸 것이다. 16개 지역 중 충남지역에는 세종시의 아파트매매거래량을 포함하였는데 이는 세종시의 시계열이 다른 16개 지역의 시계열에 비해서 이용가능한 시계열이 짧아서 전국의 지역을 17개 지역으로 나누어서 분석하는 것이 어려웠기 때문이다.

<그림 1>의 전국 및 16개 개별지역의 아파트매매거래량 추이는 그 형태에 있어서는 각각 상이하지만 장기적으로 평균값을 중심으로 평균복귀하는 형태의 시계열 추이를 가지고 있음을 확인할 수 있다. <그림 1>에서 최하단에 있는 그림은 전국(Total) 아파트매매거래량을 나타내는데 개별지역의 거래량과 마찬가지로 장기적으로 평균값을 중심으로 평균복귀하는 형태의 안정적 시계열 추이를 보여주고 있다.

<그림 1> 지역별 아파트매매거래량 추이 (2006.01-2018.10)



2. 예측모형

1) Lee-Carter모형

Lee-Carter모형(LC모형)은 미국의 통계학자들인 Lee and Carter(1992)에 의해 제안된 사망률의 확률 예측모형으로서 모형의 단순성과 추정용이함으로 인해 사망률 예측연구에 있어서 가장 기본적인 모형으로 사용되고 있다¹⁾. LC모형은 시간 순서에 따른 과거의 연령별 사망률 행렬

1) 본 연구는 Lee and Carter(1992)의 모형에 기초하여 분석하는데 1992년도 이후의 연구에서 지속적으로 개량되면서 발전되어 온 LC류의 사망률 예측모형들에 대한 다양한 논의에 대하여는 김세중(2012)의 연구를 참조할 수 있다.

(matrix) 즉, 행(row)은 시간, 열(column)은 연령으로 구성된 사망률 행렬 자료를 이용하여 미래의 시간 순서에 따른 연령별 사망률 행렬을 예측하는 모형이다.

본 연구는 횡단면시계열인 연령별-사망률 프로파일을 이용하여 미래의 사망률을 예측하는 LC모형을 응용하여, 횡단면시계열인 지역별-아파트 매매거래량 프로파일을 이용하여 미래의 지역별 아파트매매거래량을 예측한다. 즉, 사망률 예측 LC모형에서 사용되는 시간 순서에 따른 과거의 연령별 사망률 행렬을 행은 시간, 열은 지역으로 구성된 거래량 행렬로 대체하면 과거의 거래량 행렬을 이용하여 미래의 시간 순서에 따른 지역별 거래량 행렬을 예측할 수 있게 된다. 이때, 지역별-아파트매매거래량의 미래값을 예측하기 위한 LC모형은 다음과 같이 표현될 수 있다²⁾.

$$s_{x,t} = a_x + b_x k_t + \varepsilon_{x,t} \quad (x = 1, 2, \dots, n; t = 1, 2, \dots, T) \quad \langle \text{식 1} \rangle$$

$$\text{단, } \sum_{x=1}^n b_x = 1, \quad \sum_{t=1}^T k_t = 0$$

$s_{x,t}$: 지역 x 의 t 시점 아파트매매거래량

a_x : 지역별 아파트매매거래량의 평균형태를 반영하는 상수

b_x : 아파트매매거래량지수(k_t) 변화에 대한 지역 x 의 상대적 변화속도

k_t : t 시점의 아파트매매거래량지수

$\varepsilon_{x,t}$: 지역 x 의 t 시점 잔차항

식(1)에서 a_x 는 지역 x 의 분석대상 기간 동안의 아파트매매거래량 시계열의 평균값으로서 다음과 같이 생성한다.

2) 본 연구는 LC모형을 이용한 거래량 추정시 사망률 추정에서와는 달리 로그변환하지 않은 지역별 아파트매매거래량의 원시계열($s_{x,t}$)을 사용하여 분석하였다. 이는 본 연구에서 후술하는 바와 같이 원시계열을 사용하여 분석하면 전국아파트매매거래량 단변량을 이용하여 추정한 전국 예측치와 LC모형에 의해 예측한 각각의 지역별 예측치의 합계가 같아지는 합리적인 분석결과를 얻을 수 있기 때문이다. 단, 주택시장의 시계열적 변화 중 정책 변수가 미치는 영향 또한 큰 것으로 볼 수 있는데 LC모형에서는 이러한 정책변수의 통제 문제는 모형에 반영하지 못하는 한계점을 가진다.

$$a_x = T^{-1} \sum_{t=1}^T s_{x,t}; x = 1, 2, 3, \dots, n \quad \langle \text{식 2} \rangle$$

시점 t 에서의 아파트매매거래량지수 k_t 는 다음과 같이 생성한다.

$$k_t = \sum_{x=1}^n (s_{x,t} - a_x) \quad \langle \text{식 3} \rangle$$

k_t 의 변화에 대한 지역 x 의 상대적 변화속도를 나타내는 b_x 값은 식(2)와 식(3)을 이용하여 생성한 a_x 와 k_t 값을 이용하여 구할 수 있는데, 다음과 같이 최소자승법(LS)을 사용하여 상수항이 없는 회귀분석을 통해 도출한다.

$$b_x | LS(s_{x,t} - a_x) \text{ on } k_t \quad \langle \text{식 4} \rangle$$

2) VAR모형

LC모형의 예측력 평가를 위한 비교모형으로서 본 연구는 다변량시계열 예측모형인 VAR모형을 이용하였다. 차수 p 의 VAR(p)모형은 다음과 같은 수식으로 표현될 수 있다 (김명직·장국현(2003) 참조).

$$A(L)x_t = c + e_t \quad \langle \text{식 5} \rangle$$

$A(L) = (I_{(n)} - A_1L - \dots - A_pL^p)$: lag연산자 L 을 포함하는 행렬다항식,

$I_{(n)}$: n -차원 항등행렬,

x_t, c, e_t : 각각 n -차원 벡터 ($x_t = (x_{1t}, \dots, x_{nt})'$; $c = (c_1, c_2, \dots, c_n)'$;

$e_t = (e_{1t}, e_{2t}, \dots, e_{nt})'$)

3) ARCH모형

본 연구는 단변량 시계열 예측모형에 의한 예측치도 분석의 대상으로 한다. 이때, 전국 아파트매매거래량 등 단변량 시계열의 예측 시에 ARCH모형(autoregressive conditional heteroscedasticity model)을 이용하여 분석한다. 변동성 집중(volatility clustering) 또는 fat-tail의 특성을 가지는 단변량 시계열을 조건부 분산의 관점에서 모형화하기 위해 ARCH모형을 사용할 수 있다. 본 연구는 AR모형 추정식의 잔차제곱에 대한 ARCH효과를 확인하기 위해 LM검정(Lagrange multiplier test)을 시행한 후 ARCH효과가 확인되면 ARCH모형을 이용한 추정과 예측을 시행한다. ARCH모형의 형태를 설명하기 위해 AR(1)-ARCH(p)모형을 수식으로 표현하면 다음과 같다.

$$s_t = s_{t-1} + \varepsilon_t,$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 \quad \langle \text{식 6} \rangle$$

단, $\alpha_0 > 0$, $\alpha_1, \alpha_2, \dots, \alpha_p \geq 0$

III. 분석결과

1. 시계열의 안정성 검정

본 연구는 본격적인 분석에 앞서 분석대상 시계열들에 대한 안정성 여부를 확인하기 위해 ADF검정(Augmented Dickey-Fuller test) 및 PP검정(Phillips-Perron test)에 의한 단위근 검정(unit root test)을 시행하였다. <표 1>은 지역별로 구분하여 시행한 아파트매매거래량의 단위근 검정 결과를 나타낸 것이다.

<표 1> 단위근 검정결과 (2006.01-2016.12)

	ADF검정		PP검정	
	통계량	Prob.	통계량	Prob.
전국	-6.8099	0.0000	-6.7981	0.0000
서울	-3.9005	0.0027	-3.9755	0.0021
인천	-4.7722	0.0001	-4.9946	0.0000
경기	-4.7327	0.0001	-4.8116	0.0001
부산	-5.5092	0.0000	-5.3965	0.0000
대구	-6.6850	0.0000	-6.7422	0.0000
광주	-9.6940	0.0000	-9.7611	0.0000
대전	-6.2496	0.0000	-6.3037	0.0000
울산	-7.1290	0.0000	-7.5873	0.0000
강원	-8.1720	0.0000	-8.1928	0.0000
충북	-9.3325	0.0000	-9.7712	0.0000
충남	-7.5006	0.0000	-7.8163	0.0000
전북	-9.2986	0.0000	-9.6706	0.0000
전남	-8.1892	0.0000	-8.1475	0.0000
경북	-7.2602	0.0000	-7.4743	0.0000
경남	-7.6315	0.0000	-7.5980	0.0000
제주	-7.3042	0.0000	-7.3156	0.0000

주: 절편(intercept)만 있는 식에 대한 추정 결과임

<표 1>의 ADF 및 PP검정 결과에 의하면 본 연구의 분석대상 아파트 매매거래량 시계열들은 모두 1% 유의수준에서 해당시계열에 단위근이 존재한다는 귀무가설을 기각하여 모두 안정적 시계열로 확인되었다.

2. LC모형 추정결과

2006.01-2016.12 까지의 월별자료를 이용하여 LC모형을 추정한 결과 지역별 a_x 와 b_x 의 추정결과는 <표 2>와 같았다.

<표 2> 지역별 a_x 와 b_x 의 추정 (2006.01-2016.12)

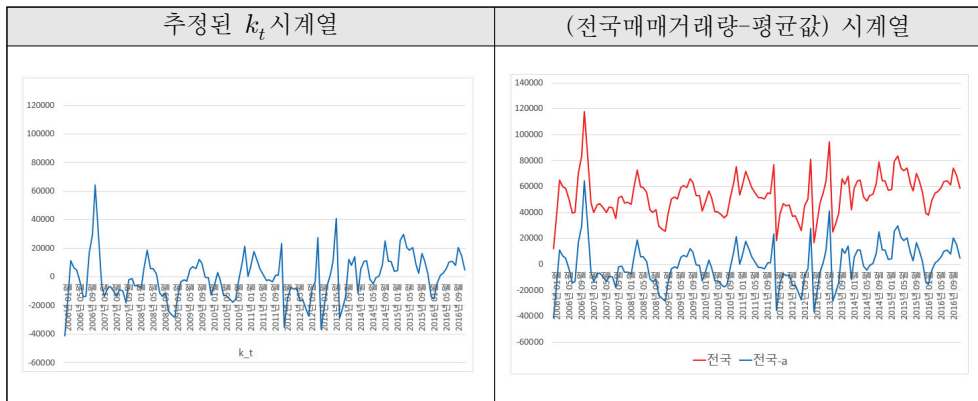
	서울	인천	경기	부산	대구	광주	대전	울산
a_x	6966	3173	13559	4362	2813	2392	1807	1645
b_x	0.2042	0.0702	0.3353	0.0740	0.0429	0.0328	0.0238	0.0221
	강원	충북	충남1	전북	전남	경북	경남	제주
a_x	1806	1841	2679	2139	1615	2560	3918	338
b_x	0.0205	0.0171	0.0269	0.0272	0.0130	0.0304	0.0552	0.0044

지역별 a_x 값을 살펴보면 경기지역의 평균 아파트매매거래량이 가장 많고 제주지역이 가장 적다는 사실을 확인할 수 있다. 아파트매매거래량지수 변화에 대한 지역 x 의 상대적 변화속도를 나타내는 b_x 의 값도 경기지역이 가장 크고 제주지역이 가장 작은 것으로 나타났다. 지역별 b_x 값은 $\sum_{x=1}^n b_x = 1$ 의 조건을 잘 만족시키고 있음을 확인할 수 있다. 월별 k_t 의

추정결과도 $\sum_{t=1}^T k_t = 0$ 의 조건을 만족시키고 있는데 <그림 2>와 같다.

<그림 2>에서 우리는 LC모형에서 추정된 k_t 시계열(좌측 그림)은 2006.01-2016.12간의 전국매매거래량에서 동일 기간의 전국매매거래량 평균값인 53613.01을 차감한 시계열(우측 그림)과 동일한 시계열임을 알 수 있다.

<그림 2> 월별 k_t 의 추정 (2006.01-2016.12)



2017년 01월 이후의 지역별 아파트매매거래량을 예측하기 위해서는 2017년 01월 이후의 k_t 에 대한 예측이 필요하다. <그림 2>에서와 같이 k_t 시계열은 평균 0을 중심으로 평균복귀하는 형태의 움직임을 보여주는 시계열이다. 본 연구는 일차적으로 AR(p)모형으로 추정된 k_t 시계열의 잔차제곱에 이분산성이 확인되어 최종적으로 변동성모형인 ARCH(p)모형을 이용하여 k_t 시계열을 예측하였다. ARCH(p)모형 추정결과는 <표 3>과 같았다.

<표 3> k_t 시계열 추정결과 (2006.01-2016.12)

AR(1)-ARCH(1)모형

$$k_t = 0.6212k_{t-1} + \varepsilon_t \quad (8.0580)$$

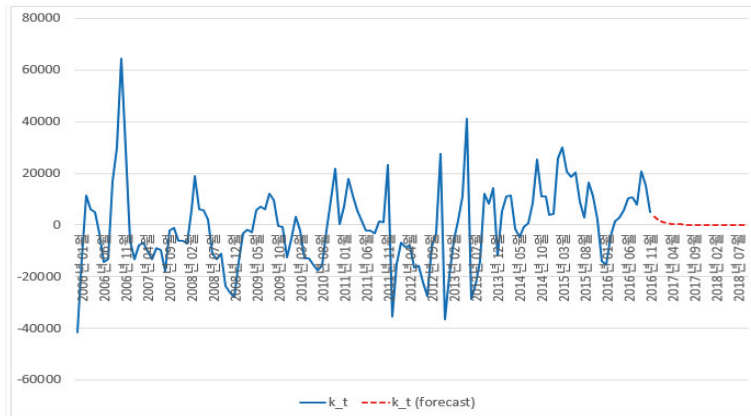
$$\sigma_t^2 = 82524267 + 0.6105\varepsilon_{t-1}^2 \quad (7.5620) \quad (4.0495)$$

- AIC : 21.5998, SC : 21.6656, Log likelihood : -1411.78
- 팔호속 수치는 z-Statistics임

주: AIC: Akaike information criterion, SC: Schwarz criterion

<표 3>의 ARCH모형 추정계수를 이용하여 예측한 2017.01-2018.10간의 k_t 시계열 예측치 추이는 <그림 3>과 같다.

<그림 3> k_t 시계열 예측치 추이 (2017.01-2018.10)



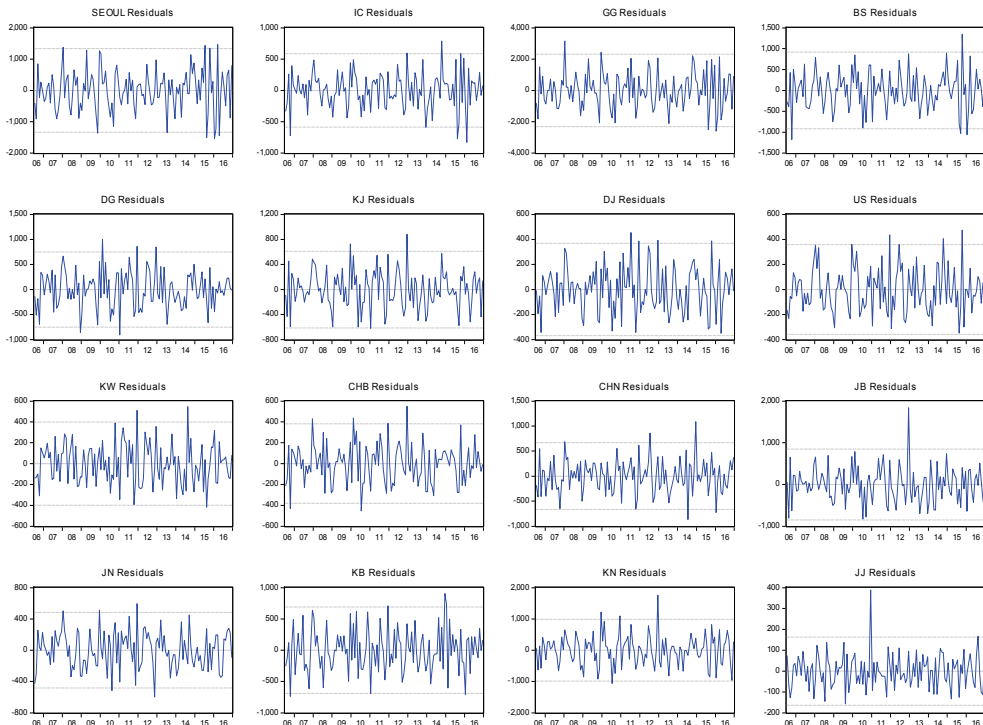
<그림 3>의 k_t 시계열 예측치를 지역별 a_x 와 b_x 값의 추정치와 결합하면 우리가 얻고자 하는 2017.01-2018.10간의 지역별 아파트매매거래량을 예측할 수 있게 될 것이다.

3. VAR모형 추정결과

본 연구는 16개 지역의 아파트매매거래량을 대상으로 VAR(p)모형을 추정하는데 있어서 VAR(p)모형의 적정 차수(p)선택을 위해 FPE (final prediction error), AIC 등의 차수선택기준 통계량을 이용하였는데, 차수 6 (lag 6)이 적정 차수인 것으로 확인되었다.

본 연구에서 추정한 VAR(6)모형의 추정계수와 통계량은 지면관계상 보고를 생략하였으며, 모형 추정의 안정성 여부 확인을 위해 추정된 VAR(6)모형의 지역별 잔차 시계열의 추이를 <그림 4>에 보고하였다. <그림 4>와 같이 VAR(6)모형을 추정한 후 잔차 시계열을 확인해본 바 분석대상 16개 지역 모두에서 백색잡음 (white noise) 과정 형태의 안정적 시계열 추이를 보여주었다.

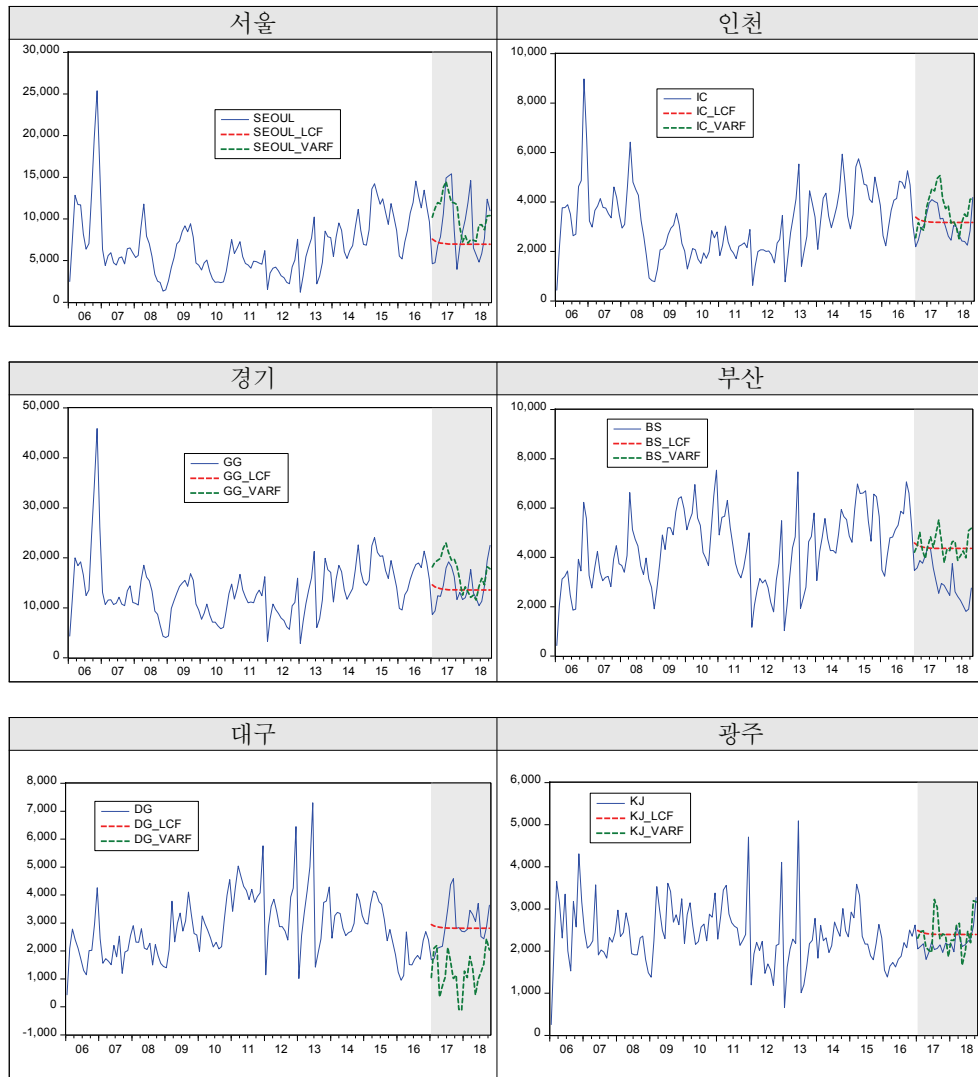
<그림 4> 추정된 VAR(6)모형의 16개 지역별 잔차시계열 추이 (2006.01-2016.12)



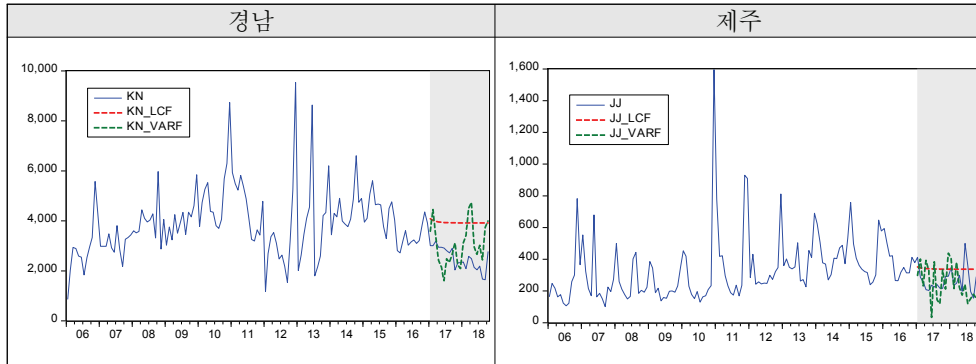
본 연구는 추정된 VAR(6)모형의 추정계수를 이용하여 지역별로 2017.01-2018.10간의 아파트매매거래량의 월별 추이를 예측하였다.

앞에서 추정한 LC모형의 추정계수를 이용하여 예측한 예측치와 VAR 모형의 추정계수를 이용하여 예측한 예측치 결과를 지역별로 동일 평면에서 비교해 보면 <그림 5>와 같다.

<그림 5> LC모형과 VAR모형의 지역별 예측치 추이 비교(예측: 2017.01-2018.10)



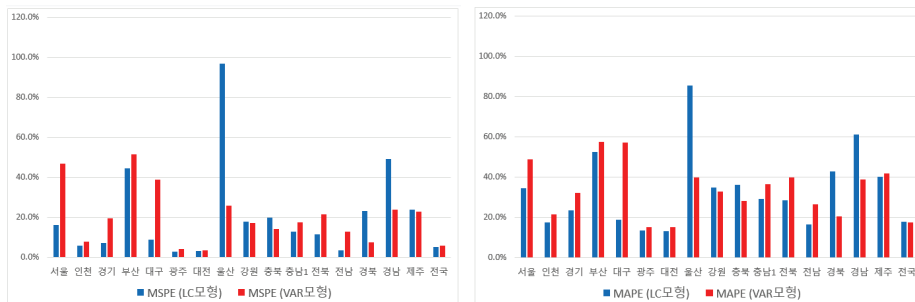




<그림 5>를 살펴보면 울산, 충북, 경북, 경남 등 지역은 VAR모형 예측치가 LC모형 예측치에 비해 실제값을 상대적으로 더 잘 반영하는 것으로 나타났고, 반면에 서울, 경기, 대구, 충남, 전북 등 지역은 LC모형 예측치가 VAR모형 예측치에 비해 실제값을 상대적으로 더 잘 반영하는 것으로 나타났다.

LC모형과 VAR모형의 지역별 예측력을 평가하기 위해 MSPE (mean squared percent error)와 MAPE (mean absolute percent error) 값을 지역별로 계산한 후 비교하였는데 그 결과는 <그림 6>과 같다.

<그림 6> 모형별 MSPE와 MAPE 비교



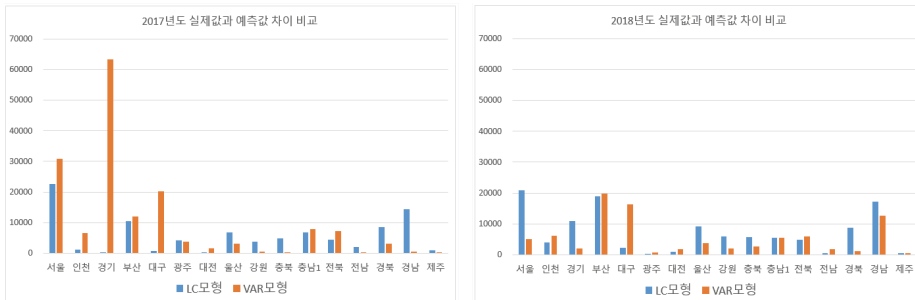
<그림 6>에 의하면 MSPE와 MAPE 기준으로 하였을 때 LC모형의 경우 부산, 울산, 경남지역에 있어서의 예측력이 상대적으로 가장 낮은 것으로 나타났으며, VAR모형은 서울, 부산, 대구지역이 가장 낮은 것으로 나타났다.

로 나타났다.

아파트매매거래량의 실제값과 예측값 간의 차이는 양(+)과 음(-)의 값이 교차할 수 있고, 이 경우 양(+)과 음(-)의 값의 상쇄효과로 인해 모형별 예측력은 월별 거래량을 기준으로 평가한 경우와 연간 거래량을 기준으로 평가한 경우 그 결과가 달라질 수 있다.

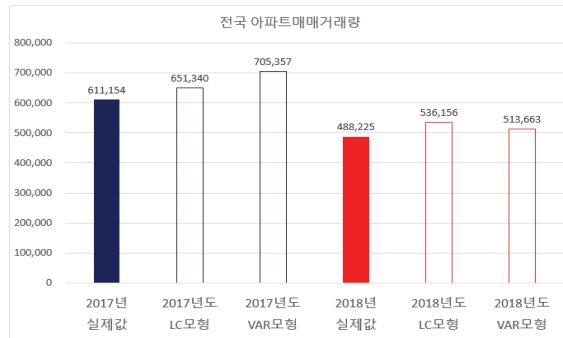
<그림 7>은 연간 거래량의 실제값과 예측값의 차이를 비교한 것인데 2017년도의 경우 LC모형은 서울과 부산, 경남지역에서 가장 큰 차이를 보여주었으며, VAR모형은 서울, 경기, 부산, 대구지역에서 가장 큰 차이를 보여주었다. 2018년도의 경우에는 1월~10월까지의 거래량을 나타낸 것인데 LC모형은 서울, 부산, 경남지역에서 VAR모형은 부산, 대구, 경남지역에서 가장 큰 차이를 보여주었다.

<그림 7> 연간 아파트매매거래량의 실제값과 예측값 비교 (지역별)



<그림 8>은 전국의 연간 아파트매매거래량의 실제값과 예측값의 차이를 비교한 것인데 2017년도의 경우 LC모형에 의한 예측치가 VAR모형에 의한 예측치에 비해 실제 거래량을 상대적으로 더 잘 반영 해주고 있으며, 2018년도에는 반대로 VAR모형에 의한 예측치가 LC모형에 의한 예측치에 비해 실제 거래량을 상대적으로 더 잘 반영 해주는 예측결과를 보여주고 있다.

<그림 8> 연간 아파트매매거래량의 실제값과 예측값 비교 (전국)



이상의 분석결과를 종합해 보면 LC모형과 VAR모형 모두 지역별로 예측력에 있어서 차이를 보여주고 있으며 몇몇 지역은 실제값을 잘 반영하지 못하는 예측치를 도출한 것으로 볼 수 있다. 그러나 LC모형과 VAR모형 양자간을 비교하였을 때 상대적으로 LC모형의 예측력이 VAR모형의 예측력에 비해 열등하지는 않는 것으로 나타났다.

본 연구는 VAR모형에 의한 월별 또는 연간 예측치 중 실제값과 큰 차이를 보여주고 있는 서울, 경기, 부산, 대구지역을 대상으로, 단변량 예측모형을 이용한 추가적인 예측치를 도출해 봄으로서 이들 지역에서 LC모형에 의해 도출된 예측치에 비해 상대적으로 더 높은 예측력을 가지는 예측치를 얻을 수 있는 방법이 있는지 여부를 추가적 분석을 통해 확인해 보았다.

4. 추가적 분석

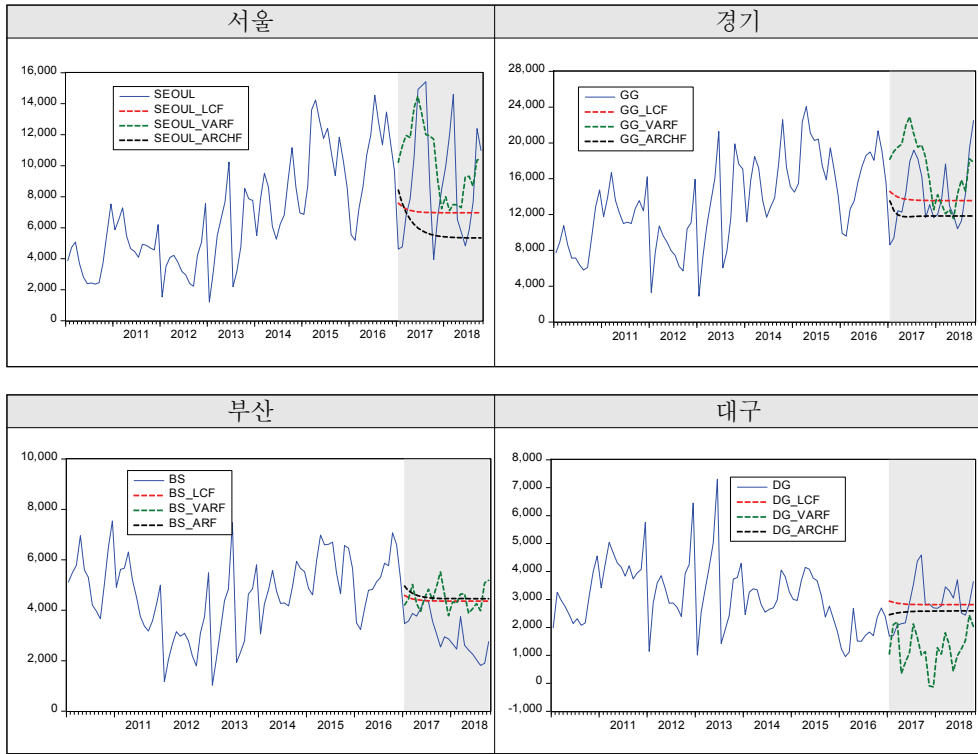
서울, 경기, 부산, 대구지역의 추가적 분석을 위해 본 연구는 AR모형 또는 ARCH류의 변동성 모형을 이용하여 예측치를 도출하였다. <표 4>는 이들 지역의 예측모형 추정결과를 나타낸 것이다.

<표 4> 서울, 경기, 부산, 대구지역의 예측모형 추정결과

서울 : AR(2)-ARCH(1)모형	
$S_t = 989.8 + 1.0380S_{t-1} - 0.2236S_{t-2} + \varepsilon_t$	
(2.3696) (11.012) (-2.854)	
$\sigma_t^2 = 2003611 + 0.6977\varepsilon_{t-1}^2$	
(4.9001) (2.9476)	
<ul style="list-style-type: none"> ◦ AIC : 18.0278, SC : 18.1381, Log likelihood : -1166.81 ◦ 괄호속 수치는 z-Statistics임 (이하 모두 동일함) 	
경기 : AR(2)-ARCH(1)모형	
$G_t = 3262 + 1.0188G_{t-1} - 0.2943G_{t-2} + \varepsilon_t$	
(4.0305) (12.009) (-4.087)	
$\sigma_t^2 = 5025801 + 0.8864\varepsilon_{t-1}^2$	
(3.8684) (3.6043)	
<ul style="list-style-type: none"> ◦ AIC : 19.0911, SC : 19.2014, Log likelihood : -1235.92 	
부산 : AR(1)모형	
$B_t = 1574.8 + 0.6468B_{t-1} + \varepsilon_t$	
(5.3397) (10.090)	
<ul style="list-style-type: none"> ◦ AIC : 16.8362, SC : 16.8801, Log likelihood : -1100.77 	
대구 : AR(1)-ARCH(1)모형	
$D_t = 729.7 + 0.7183D_{t-1} + \varepsilon_t$	
(4.8970) (12.862)	
$\sigma_t^2 = 243764 + 0.9322\varepsilon_{t-1}^2$	
(3.9663) (3.3841)	
<ul style="list-style-type: none"> ◦ AIC : 16.1700, SC : 16.2578, Log likelihood : -1055.13 	

<표 4>의 추정계수를 이용하여 2017.01-2018.10 (22개월간)의 4개지역 아파트매매거래량을 예측한 후 앞에서 살펴보았던 LC모형과 VAR모형 예측치와 비교해 보면 <그림 9>와 같다.

<그림 9> 4개지역 (서울, 경기, 부산, 대구)의 모형별 예측치 비교



<그림 9>의 예측치 결과에서 확인할 수 있는 바와 같이 분석대상 4개 지역 (서울, 경기, 부산, 대구)에서 AR모형 또는 ARCH모형을 이용하여 예측한 예측치 결과도 일부의 경우 VAR모형의 예측결과와는 현저히 개선시켜 주기도 하지만 (예: 대구지역), LC모형에 의한 예측치에 비해서는 더 나은 결과를 보여주지는 못함을 확인할 수 있다³⁾.

5. 전국 아파트매매거래량의 예측

주택시장을 분석할 때 분석 목적에 따라 전국의 거래량 예측치가 필요

3) 월별 거래량을 살펴보면 거래량 시계열이 계절성을 가지는 시계열임을 알 수 있는데 연간 거래량 비교에 있어서는 LC모형의 예측치가 VAR모형에 비해 우수한 결과를 보여준다 하더라도 월별 예측결과에서 시계열의 계절적 특성을 제대로 보여주지 못하는 한계점이 존재한다.

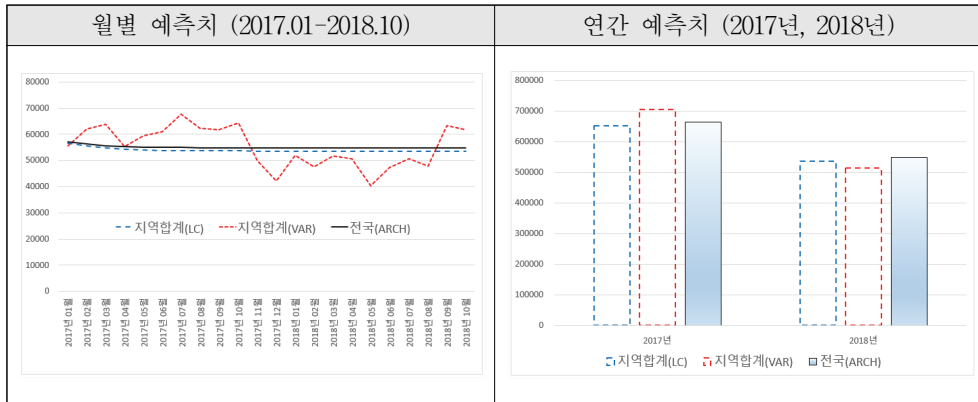
한 경우가 있고 때로는 지역별 거래량의 개별적 예측치가 필요한 경우가 있다. 본 연구에서 분석대상으로 한 16개 지역의 지역별 개별 아파트매매거래량 예측치를 모두 합산하면 이는 전국의 아파트매매거래량이 된다. 이때, 전국의 아파트매매거래량 시계열을 이용하여 예측한 전국의 아파트매매거래량 예측치와 지역별 아파트매매거래량을 개별적으로 예측한 후 이들 예측치를 모두 합산하여 구한 아파트매매거래량이 일치하면 주택시장의 전망 등 분석에서 이들 예측치를 유용하게 활용할 수 있다. 반면에 전국의 아파트매매거래량 시계열을 이용하여 예측한 전국의 아파트매매거래량 예측치와 지역별 아파트매매거래량을 개별적으로 예측한 후 이들 예측치를 모두 합산하여 구한 아파트매매거래량이 일치하지 않고 편차가 크다면 전국과 개별지역을 대상으로 추정된 예측치 모두 시장 진단에 활용하기에는 상당한 한계점을 가지게 될 것이다. 본 연구는 ARCH모형을 이용하여 전국 아파트매매거래량의 예측치를 도출하였는데, <표 5>는 이들 지역의 예측모형 추정결과를 나타낸 것이다.

<표 5> 전국 아파트매매거래량 예측모형 추정결과 (2006.01-2016.12)

전국 : AR(1)-ARCH(1)모형
$T_t = 21022 + 0.6166 T_{t-1} + \varepsilon_t$ <p style="text-align: center;">(5.1251) (8.0101)</p> $\sigma_t^2 = 81686147 + 0.6239 \varepsilon_{t-1}^2$ <p style="text-align: center;">(6.4519) (3.6316)</p>
<ul style="list-style-type: none"> ◦ AIC : 21.6135, SC : 21.7013, Log likelihood : -1411.69 ◦ 괄호속 수치는 z-Statistics임

<표5>의 ARCH모형 추정계수를 이용하여 예측한 2017.01-2018.10까지의 전국 아파트매매거래량 예측치(전국(ARCH))를 LC모형으로 예측한 지역 예측치의 합계(지역합계(LC))와 VAR모형으로 예측한 지역 예측치의 합계(지역합계(VAR))와 함께 동일 평면에 나타내면 <그림 10>과 같다.

<그림 10> 전국 아파트매매거래량 예측치 비교



시계열의 미래값은 다양한 모형으로 추정해볼 수 있지만, 개별지역의 매매거래량 예측치를 합산한 값이 단변량인 전국의 매매거래량을 대상으로 예측한 전국 매매거래량의 예측치와는 현저한 차이를 보여주는 경우가 많다. 그러나 본 연구에서 확인할 수 있는 바와 같이 LC모형에 의한 지역별 예측치는 다른 예측모형에 의한 예측치에 비해 결과가 열등하지도 않으면서, 한편으로는 LC모형에 의해 추정된 지역별 예측치의 합계액이 단일 변량인 전국 매매거래량만을 대상으로 예측한 단변량 예측치와 거의 유사한 값을 제공해 준다. 이러한 분석결과는 앞의 <그림 2>에서 살펴본 바와 같이 LC모형에서 추정된 k_t 시계열이 2006.01-2016.12 간의 전국매매거래량에서 동일 기간의 월별 전국매매거래량 평균값인 53613.01을 차감한 시계열과 동일한 시계열이기 때문에 가능한 결과이다. 따라서 LC모형은 실무에서 아파트매매거래량 예측치를 활용하는 경우에 있어서 다른 예측모형에 비해 상대적으로 더 큰 장점을 가지는 것으로 평가할 수 있다.

IV. 결론

본 연구는 주택경기를 진단하는데 사용되는 다양한 지표들 중 주택매

매거래량 시계열을 대상으로 동 시계열의 합리적 예측을 위한 모형을 검토하였다. 아파트매매거래량의 미래값 예측에 있어서 서로 다른 예측모형 간의 예측력 평가를 통해 아파트매매거래량 등 주택시장의 분석에 있어서 상대적으로 더 큰 유용성을 가지는 예측모형을 확인하였다. 본 연구는 전국의 지역을 16개 시·도 지역으로 구분하여 분석하였는데, 구체적으로 16개 지역의 아파트매매거래량을 동시에 예측할 수 있는 모형으로서 LC모형과 함께 VAR모형을 비교모형으로 설정하여 분석하였으며, 각각의 모형에 의한 주택매매거래량 예측치와 실제값을 비교하여 모형별 예측력을 비교하였다.

본 연구의 분석결과에 의하면 LC모형과 VAR모형 모두 지역별로 예측력에 있어서 다소간 차이를 보여주었으며, 16개 지역 중 몇몇 지역은 실제값을 잘 반영하지 못하는 예측치를 보여준 곳도 있었다. 그러나 LC모형과 VAR모형 양자간을 비교하였을 때 상대적으로 LC모형의 예측력이 VAR모형의 예측력에 비해 열등하지는 않는 것으로 나타났다.

본 연구는 VAR모형에 의해 추정된 월별 또는 연간 예측치 중 실제값과 큰 차이를 보여주고 있는 서울, 경기, 부산, 대구지역을 대상으로 LC모형에 의해 도출된 예측치에 비해 상대적으로 더 높은 예측력을 가지는 예측모형이 있는지 여부를 추가적 분석을 통해 확인해 보았다. 그러나 분석대상 4개 지역에서 AR모형 또는 ARCH모형을 이용하여 예측한 분석결과도 일부의 경우 VAR모형의 예측치에 비해서는 현저히 개선되었지만 (예: 대구지역), LC모형에 의한 예측치에 비해서는 더 나은 결과를 보여주지 못하였다.

한편, 개별지역의 아파트매매거래량 예측치를 합산한 값이 단변량인 전국의 매매거래량을 대상으로 예측한 전국 매매거래량의 예측치와는 현저한 차이를 보여줄 수도 있는데, LC모형에 의해 추정된 지역별 예측치의 합계액은 단일 변량인 전국 매매거래량만을 대상으로 예측한 단변량 예측치와 거의 유사한 값을 제공해 주었다. 이러한 분석결과는 LC모형에서 추정된 k_t 시계열이 분석대상 월별 전국매매거래량 시계열에서 동일기간의 평균값을 차감한 시계열과 동일한 시계열이기 때문에 가능한 결

과이다. 따라서 이러한 관점에서도 LC모형은 실무에서 아파트매매거래량 예측치를 활용하는 경우에 있어서 다른 예측모형에 비해 상대적으로 더 큰 장점을 가지는 모형으로 평가할 수 있다.

본 연구의 분석결과에 의하면 주택시장의 다양한 변수들 중 특히 전국 16개지역의 물량을 합제한 것이 전국의 물량과 동일하게 되는 변수들을 대상으로 지역별 각각의 개별 예측치를 동시에 얻고자 하는 경우 LC모형을 예측모형으로 유용하게 활용할 수 있을 것이라 판단된다. 이러한 주택시장 변수들 중 주택공급 부문의 대표적인 변수들로서 우리는 본 연구에서 분석한 아파트매매거래량 이외에 아파트 인허가물량, 아파트 분양보증물량, 아파트 미분양물량 등을 들 수 있다. 본 연구에서 제안한 LC모형은 지역별 아파트매매거래량의 평균 형태를 반영하는 상수(a_x)가 모형에서 차지하는 비중이 크므로 예측 결과 또한 과거의 평균치에서 크게 벗어나지 않는 특성을 가진다. 따라서 지역별 예측에서 주택시장 경기가 급격하게 침체기 또는 호황기를 맞이하는 특정지역의 예측치는 실제치와 많이 차이가 발생할 수도 있다(본 연구에서의 부산, 울산, 경북, 경남 등). 따라서 주택시장에서 향후 LC모형을 활용할 때에는 지역별 동시 예측의 유용성에도 불구하고 모형에 내재되어 있는 이러한 한계점을 고려하면서 예측결과를 해석하고 활용할 필요가 있을 것이다. 주택거래량 예측에 있어서 본 연구에서 제시한 LC모형이 보여주는 한계점은 본 연구에서 사용한 모형보다 더 개량된 LC모형들(김세중(2012) 참조)을 적용하여 분석하거나 또는 주택거래량과 가장 밀접한 관련성이 있다고 볼 수 있는 주택가격 등 다른 변수들의 영향을 함께 고려한 다변량 예측모형을 통해 보완할 수 있을 것으로 판단된다. 이 부분은 본 연구의 후속 과제로 남겨두기로 한다.

논문접수일	2019.06.18.
논문심사일	2019.06.19.
게재확정일	2019.06.24.

참고문헌

- 김명직 · 장국현, 2003, 금융 시계열 분석, 경문사.
- 김세중, 2012, “Lee-Carter모형을 이용한 사망률 예측에 관한 연구”, 계리학연구 4권 2호, pp.47-66.
- 마승렬, 2015, “공무원 일실회의 합리적 산정방법 연구”, 보험금융연구 26권 4호, pp.77-111.
- 마승렬, 2016, “별집순환모형을 이용한 지역별 아파트시장 분석 및 예측에 관한 연구”, 주택도시금융연구 1권 1호, pp.111-138.
- 임대봉, 2014, “주택거래량과 주택가격에 관한 연구”, 부동산학보 58집, pp.195-208.
- 임재만, 2011, “주택거래량은 주택가격 변동을 설명할 수 있는가?”, 국토연구 69권, pp.3-18
- Clayton, Jim, Norman Miller, and Liang Peng, 2010, “Price-volume Correlation in the Housing Market: Causality and Co-movements,” *The Journal of Real Estate Finance and Economics* 40, pp.14-40.
- Festa, Maurizio, Erika Ghirardo, and Alessandra Storniolo, 2012, “Analysis of Real Estate Market Cycles: An Application on Italian Data,” *Firenze University Press*.
- Genesove, David and Christopher Mayer, 2001, “Loss Aversion and Seller Behavior: Evidence from the Housing Market,” *The Quarterly Journal of Economics* 116(4), pp.1233-1260.
- Janssen, Jos, Bert Kruijt, and Barrie Needham, 1994, “The Honeycomb Cycle in Real Estate,” *The Journal of Real Estate Research* 9(2), pp.237-251.
- Lee, R. and L. Carter, 1992, “Modeling and Forecasting US mortality,” *Journal of the American Statistical Association* 87, pp.659 - 671.
- Leung, Charles Ka and Dandan Feng, 2005, “What Drives the

Property Price-Trading Volume Correlation? Evidence from a Commercial Real Estate Market,” *The Journal of Real Estate Finance and Economics* 31(2), pp.241-255.

- Stein, Jeremy C, 1995, “Prices and Trading Volume in the Housing Market a Model with Down-Payment Effects,” *The Quarterly Journal of Economics* 110, pp.379-406.
- Suwanvijit, Wassana, Thomas Lumley, Chamnein Choonpradub, and Nittaya McNeil, 2011, “Long-Term Sales Forecasting Using Lee-Carter and Holt-Winters Methods,” *The Journal of Applied Business Research* 27(1), pp.87-102.

Availability of Lee-Carter Model in Housing Market Analysis

Ma, Seungryul*

<Abstract>

In this analysis, we compared Lee-Carter (LC) model and VAR model to confirm the availability of LC model in the forecast of housing market using the data of apartment transaction volume, and then evaluated the forecast accuracy by comparing the actual values with the forecasted values from LC model and VAR model. According to the results of this analysis, it appeared that the LC model was not inferior to VAR model in the forecast accuracy. In the additional analysis, the univariate time series forecasting models such as AR model or ARCH model also could not show superior results to LC model. LC model is expected to be used usefully in the field of housing market because the structure of LC model is simple and the estimating process is easy as well as it has similar forecast accuracy compare to different models. Meanwhile, different from VAR model, the sum of forecasted values of all 16 area estimated by using LC model was almost the same as the forecasted total values estimated by the univariate model. In this viewpoint, the LC model also has its merits compared to different models when we use the forecasted apartment transaction volume in the business practice.

Key words: housing market, forecasting model, forecast accuracy, apartment transaction volume, Lee-Carter model

* Director, SonSa Research Institute, samhan12@hanmail.net