

동태적 모형 선택법을 이용한 주택가격 예측 변수 분석*

Forecasting Housing Prices Using Dynamic Model Selection

황 영 진 (Hwang, Youngjin)**

< Abstract >

This paper presents a time series forecasting model of the Korean housing prices using the dynamic model selection. This methodology is attractive in that not only model coefficients but also the entire forecasting model (i.e., predictors) are allowed to change over time. It is found that dynamic model selection performs better than other alternative popular forecasting models including conventional regressions, random walk, autoregressive models, and time-varying parameter models.

When looked at which sets of predictors are relevant for forecasting in each period, there are noticeable changes in predictors over forecast horizons. For short horizons, housing prices seem to be largely affected by housing market developments and business cycles; however, GDP and consumption seem to play a lesser role during housing market booms. For medium-run horizons, monetary aggregates and financial market variables appear to replace the aforementioned variables and have more predictive power, indicating that portfolio adjustment in asset markets may be a critical factor in housing price changes.

주 제 어 : 동태적 모형 선택법, 주택가격, 예측

key word : Dynamic Model Selection, Housing Prices, Forecast

* 이 논문은 2012년 한양대학교 교내연구비 지원으로 연구되었음(HY-2012-G)

** 한양대학교 경제학부 조교수, youngjinh@hanyang.ac.kr

I. 서론

최근 주요 선진국들에서 나타난 주택가격의 급등 및 그 이후의 조정 과정은 주택시장이 더 이상 경제 내의 한 독립된 소규모 영역이 아니라, 경제 전반과 밀접한 관련을 가지며 적지 않은 파급력을 지닐 수 있는 부문임을 시사하고 있다. 특히 최근의 미국 발 글로벌 금융위기 촉발의 주요 요인 중 하나로 주택시장의 과열 및 그와 관련된 주택 금융시장의 부실이 지목되고 있으며, 한국의 경우에도 2000년대의 가격 급등 이후 현재까지 주택시장은 계속 침체 양상을 보이며 여타 부문에 위협 요인으로 작용하는 등 우려가 지속되고 있다.

이와 같은 주택시장의 중요성을 감안할 때, 주택시장 변동 요인들에 대한 분석 및 그를 바탕으로 한 향후 전망 작업은 시장 불안정 요인들을 사전적으로 파악하여 적절한 정책 개입 등을 통해 주택시장의 안정 및 전반적인 거시경제의 건전성 도모를 위한 중요한 작업이라고 할 수 있다.

한편 이와 관련하여, 한국의 경우에 있어 특히 주목할 점은 외환위기 이후 전반적인 경제의 구조 변화와 더불어, 주택시장 또한 주택 금융시장의 발전 및 인구·가족 구조의 변화 등으로 인해 급격한 변화를 겪고 있다는 점이다(심성훈, 2006; 김경환, 2007; 송준혁, 2012; 이영수, 2012). 또한 외환위기 이후 주택시장 경기는 실물 부분의 경기와는 다소 분리되어 독자적 사이클을 가지는 것으로 보여, 한국 주택시장의 동학에 대한 분석에 있어서는 이러한 점들이 충분히 고려되어야 할 것으로 여겨진다.

이러한 주택시장 관련 분석 및 향후 전망 작업에 있어 핵심적인 작업 중 하나는 적절한 주택

시장 예측 모형의 구축 및 개발이라고 할 수 있다. 통상적인 예측에서는 예측 대상 변수와 이론적으로 밀접한 관련이 있거나 예측에 있어 유용한 변수들을 선정하여 예측 모형을 구성하고 추정된 뒤, 이를 이용한 예측 작업이 이루어지는 것이 일반적이다. 하지만 이러한 통상적인 예측 모형의 경우 암묵적으로 예측 변수들의 중요성 또는 역할이 표본 및 예측 기간 동안 일정한 것으로 전제하고 있으나, 앞서 언급한 주택시장의 구조 변화 가능성 등을 고려할 때, 적어도 한국의 경우에는 이러한 통상적인 예측 모형을 이용한 분석은 다분히 제한적이며 예측 대상의 특성을 적절히 반영하지 못할 가능성이 있다.

우리나라 주택시장 예측에 관한 기존 연구로는 손정식 외(2002), 최희갑·임병준(2009), 박천규·이영(2010), 정의철(2010) 등이 있으나, 이들 연구의 대부분은 고정된 단일 예측 모형이 매기 적용되는 통상적인 방법을 사용하여, 분석 기간 중 모형 또는 예측 변수가 변화할 가능성을 충분히 고려하지 않았다. 또한 예측 변수의 측면에서도 최근 연구의 상당수는 주택시장 관련 심리 지표들의 예측력 검토에 초점이 맞춰져 있는 등 주택시장 예측에 있어 거시 변수들의 역할에 초점을 맞춘 연구는 상대적으로 드문 실정이다.

이 논문은 동태적 모형 선택법(dynamic model selection, DMS)을 이용한 한국 주택가격 예측 모형을 제시하고 그 결과를 분석한다. 이 접근법은 주택가격 예측에 유용할 수 있는 변수들로 예측 후보 변수군을 구성한 뒤, 매 시점에서 향후 주택가격 변화를 가장 잘 예측할 수 있는 변수들을 선택하여 예측을 수행하게 된다. 따라서 이는 예측 변수와 그 영향력이 고정된 것으로 간주하는 통상적인 예측 모형과는 달리 매 시점마다 예

측의 유용성에 따라 예측 변수가 변할 수 있으며 그 변수의 영향력도 변할 수 있다는 점에서 매력적인 접근법이라 할 수 있다. 이러한 동태적 모형 선택법은 Raftery et al.(2010), Koop and Korobilis(2012) 등에 의해 제안되었으며, 인플레이션 등 주요 거시 변수 예측에 있어 주목할 만한 성과를 보이는 것으로 알려져 있다. 또한 이러한 접근법은 비록 명시적인 이론에 근거한 모형은 아니지만, 주택가격 변화에 선행하는 주요 변수들을 동태적으로 식별할 수 있다는 점에서 간접적으로 주택가격 변화 원인에 대한 분석에도 유용하게 사용될 수 있는 등 잠재적으로 결과에 대한 구조적 해석도 가능할 수 있다.

이 논문의 주요 결과를 간단히 정리하면 다음과 같다. 우선 동태적 모형 선택법은 자기회귀 모형 등의 간단한 일변량 모형, 고정된 예측 변수와 계수 추정치를 상정하는 통상적인 예측 모형, 시변모수(time-varying parameter) 모형 등 널리 쓰이는 예측 모형들과 비교하여 대부분의 예측 시계에서 우위를 보이며, 경쟁력 있는 예측 모형임을 확인할 수 있었다. 특히 시변모수 모형에 대한 예측 우위는, 한국 주택시장에 대한 예측의 경우, 잠재적인 구조 변화를 반영함에 있어, 모형의 계수 추정치의 변화 뿐 아니라 예측 변수 자체의 역할이나 중요성이 변화했을 가능성을 시사한다는 점에서 주목할 만하다고 할 수 있다.

이와 함께, 동태적 모형 선택법에 의한 예측 결과, 예측시계에 따라 주요 역할을 하는 예측 변수들이 상당정도 다른 양상을 보였다. 1~2 분기 후의 단기 예측의 경우, 실물경기 변수(GDP, 민간소비, 소비자물가지수, GDP 디플레이터 등) 및 주택가격 변수들(서울 주택가격, 전세가격)이 주된 역할을 한 것으로 보이며, 부분적으로 토지

거래량 및 주거용 건설투자 등의 주택시장 관련 변수 및 M1 등도 중요한 예측 변수 역할을 하는 것으로 나타났다. 하지만, 흥미롭게도 주택가격 급등기에는 실물경기 변수의 예측력은, 변수별로 정도는 다소 다르지만, 상당부분 감소하는 모습을 보였다.

하지만, 예측 시계가 증가하면서 이러한 변수들의 예측력은 전반적으로 감소하는 경향을 보였다. 예를 들어, 1년 이후의 중기 예측에서는, 주택가격 급등의 경우 주택가격 변수들의 예측력은 큰 폭으로 감소하거나 미미한 수준인 것으로 나타났으며, 토지거래량 및 주거용 건설투자의 경우에도 가격 급등기에는 소폭이긴 하나 오히려 예측력이 감소하는 모습을 보였다. 마찬가지로 소비 등 실물경기 변수들의 경우에도 주택가격과 관련한 뚜렷한 움직임은 찾기 힘들었다. 대신 중기 예측에 있어서는 M2, 회사채 수익률 및 주가 등 금융시장 관련 변수 및 생산자물가지수가 주된 예측 역할을 하는 것으로 나타났다.

이러한 결과는 주택시장이 단기적으로는 주택시장 및 실물경기와 밀접한 관련성을 보이면서 이들 변수에 영향을 받긴 했지만, 주택가격 상승기에는 실물경기 변수들의 영향력은 오히려 상대적으로 감소하는 모습을 보여, 주택시장 호황과 전반적인 거시경제 여건이 다소 괴리되는 양상을 확인할 수 있었다. 이와 함께, 일 년 이상의 중기에 있어서는 주택가격은 이러한 변수들 보다는 유동성 및 금융시장 관련 변수들에 의해 상대적으로 더 영향을 받아, 자산가격 변동에 따른 금융 및 부동산 자산시장에서의 포트폴리오 조정 과정이 주택가격 변동에 주요한 요인으로 작용했을 가능성을 시사한다.

또한 주택가격 예측에 있어, 구체적으로 어떤 변수/데이터를 사용하느냐에 따라 예측 결과에 있어 적지 않은 차이가 나타나는 것으로 나타났다. 예를 들어, 물가지수의 경우, 소비자물가지수, GDP 디플레이터, 그리고 생산자물가지수의 예측력이 시기별로 상당한 차이를 보였으며, 마찬가지로 유동성 및 이자율 관련 변수들의 경우에도 유사한 결과를 발견할 수 있었다. 이러한 결과는 주택시장에 대한 예측 모형을 설정함에 있어 예측 변수의 특성을 충분히 감안하는 등 구체적인 변수 선정에 있어 세심한 주의가 필요함을 시사한다.

이 논문의 구성은 다음과 같다. 2절에서는 동태적 모형 선택법을 이용한 예측의 기본적인 아이디어를 간단히 소개한다. 다음 3절에서는 사용된 데이터와 모형 설정 등에 관한 사항들을 간단히 논의하며, 이와 함께 통상적으로 널리 쓰이는 여타 예측 모형들과의 추정 및 예측 결과를 비교해 본다. 이어 4절에서는 이러한 동태적 모형 선택법을 이용한 추정 및 예측 결과를 바탕으로, 매 시기별로 주택가격 변화에 있어 어떤 변수들이 주된 예측 역할을 하였는지 살펴보고 그 함의를 분석해 본다. 마지막으로 5절에서는 논문을 요약하고 마무리한다.

II. 동태적 모형 선택법을 이용한 예측: 개요

다음과 같은 (주택가격) 예측 모형을 고려해 보자.

$$y_t = \beta_0 + \sum_{j=1}^p \beta_j y_{t-j} + \gamma X_{t-h} + \epsilon_t \quad (1)$$

여기서 y_t 는 주택가격을 나타내며, X_t 는 주택가격 예측에 잠재적으로 유용하게 쓰일 수 있는 거시·금융 및 주택시장 관련 변수들을 포함한 벡터이다. 또한, $\theta = (\beta_0, \beta_1, \dots, \beta_p, \gamma)$ 는 모형의 모수 벡터이며, ϵ_t 는 평균이 0인 예측 오차항을 나타낸다.

통상적인 예측 모형에서는 식 (1)의 우변에 포함된 예측 변수(predictor) 및 그에 대한 계수값이 표본 및 예측 기간 동안 일정한 것으로 간주된다. 하지만 앞서 언급한 경제의 구조 변화 및 주택시장 변수의 특성 등을 고려한다면, 적어도 한국의 경우, 각 예측 변수들의 역할 및 예측 기여도가 매 시기에 있어 동일하다고 간주하는 것은 그다지 현실적이지 않다.

이 논문에서는 각 시점에서 주택가격 예측에 있어 가장 예측력이 높은 변수들을 선택하여 예측 모형이 구성되는 동태적 모형 선택법(dynamic model selection, DMS)을 이용한 주택가격 예측 모형을 제시한다. 구체적으로, 이 방법에서는 매기에 벡터 X_t 에 포함된 개별 변수들의 포함 여부에 따라 모든 가능한 예측 모형들을 고려하고, 그 중 h 기 이후의 주택가격 변화에 대해 가장 예측력이 높은 변수들로 구성된 예측 모형을 선정하여 주택가격 변화를 예측하게 된다.¹⁾

이러한 동태적 모형 선택법에 의한 예측에서는 각 변수들의 예측력에 따라 매 시점마다 예측 변수들이 변할 수 있으며, 이들 변수들의 정성

1) 이와 유사한 방법인 동태적 모형 평균법(dynamic model average, DMA)에서는 이러한 각 모형의 확률을 가중치로 하여 각 모형별 예측치를 평균하여 그를 예측치로 사용한다. 보다 자세한 논의는 부록을 참고하기 바란다.

적·정량적 영향도 달라질 수 있다는 점에서 통상적인 예측 모형에 비해 매력적이다. 또한 이러한 예측 모형은 비록 명시적인 이론에 입각한 모형은 아니지만, 매기에 주택가격 변화와 밀접한 관련을 가지며 그에 선행하는 변수들을 식별할 수 있다는 점에서, 간접적으로나마 주택가격 변화의 장·단기적 원인을 살펴보는 작업도 가능할 수 있다.

이러한 동태적 모형 선택법을 이용한 예측에 있어 문제가 될 수 있는 부분은 잠재적으로 고려되는 모형의 수가 너무 많아져, 추정 및 계산 과정에서 상당한 비용이 수반되거나, 경우에 따라서는 실질적으로 추정이 불가능할 수도 있다는 점이다. 예를 들어, 위의 식 (1)에서 만일 m 개의 잠재적인 예측 변수가 있다고 하면, 개별 변수의 포함 여부에 따라 고려할 수 있는 가능한 모든 모형의 수는 매 시점에서 $K=2^m$ 이며, 표본의 τ 번째 시점에서 예측을 위해 추정해야 하는 모형의 수는 $2^{m\tau}$ 가 된다. 따라서 고려하는 예측 변수들의 수가 충분히 작지 않은 경우, 이러한 개별 모형들에 대한 추정 및 모형간의 이전 확률(transition probability)을 나타내는 ($K \times K$) 크기의 이전 행렬의 계산 등은 상당한 비용을 수반하게 된다.

이 논문에서는 Fagin(1964) 및 Jazwinsky(1970)의 연구를 바탕으로 Raftery et al.(2010) 및 Koop and Kobolis(2012) 등에 의해 제안된 방법을 따라, 일반적인 칼만 필터(Kalman filter)의 방법론 하에서 ‘망각 인자(forgetting factor)’를 도입한 적

절한 근사화 과정을 통해, 이러한 모형의 추정 및 계산 과정을 획기적으로 단축시킨 방법을 이용하였으며, <부록 1>에서는 이러한 근사화 및 추정 과정에 대한 기술적(technical) 논의를 간략히 정리하여 소개하고 있다.

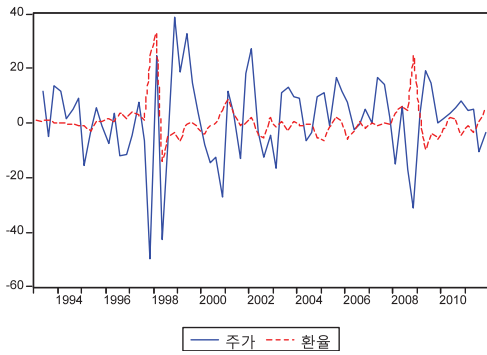
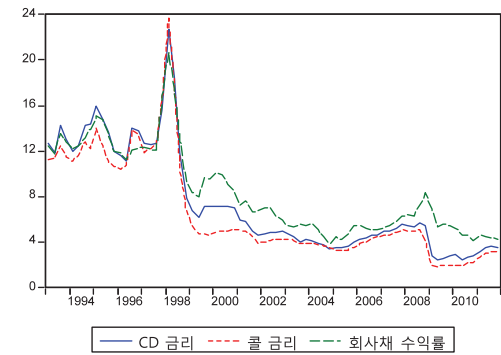
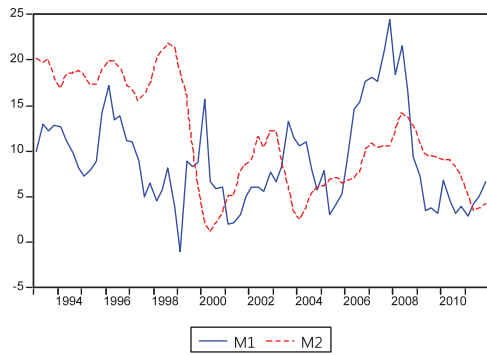
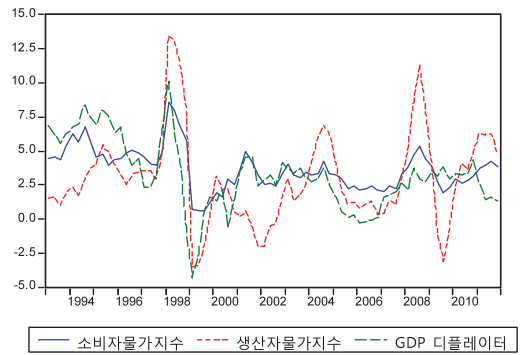
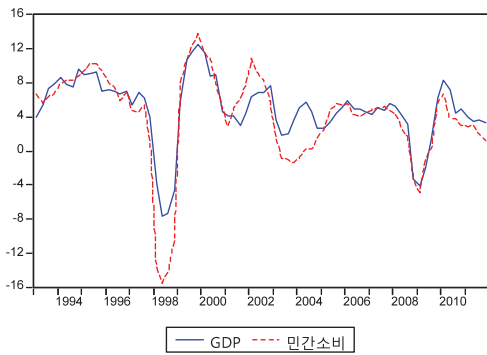
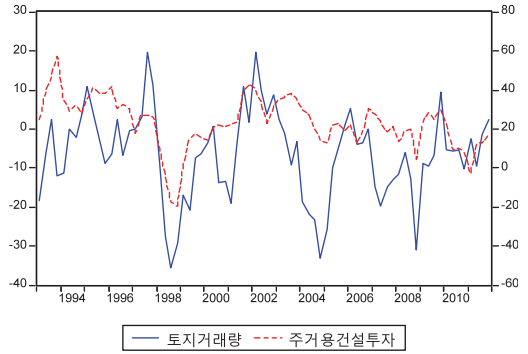
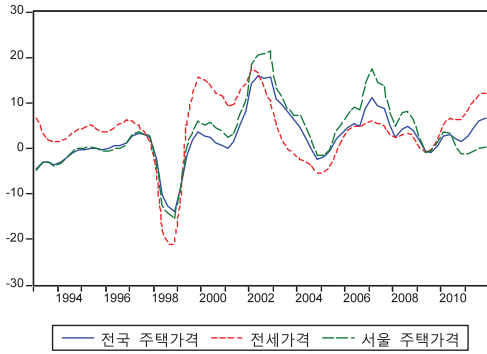
III. 데이터, 실증 분석 및 예측

1. 데이터 및 모형 설정

동태적 모형 선택법을 이용한 주택가격 예측 모형의 추정을 위해 고려된 예측 변수들은 손정식 외(2002), 최희갑·임병준(2009), 손종철(2010), 정의철(2010) 등을 참고하여 선정하였으며, 1992년 1분기부터 2011년 4분기까지의 분기별 데이터를 이용하였다. 논의의 편의를 위해 사용된 변수들을 몇 개의 그룹으로 나누면 다음과 같다. (1) 주택시장 관련 변수: 전국 주택가격, 서울 주택가격, 전세가격, 주거용 토지거래량, 주거용 건설투자; (2) 거시 및 물가 변수: GDP, 민간소비, 소비자물가지수, 생산자물가지수, GDP 디플레이터; (3) 통화 및 금융시장 관련 변수: M1, M2, CD 금리, 콜금리, 회사채 수익률, 주가지수, 원/달러 환율.²⁾ 이자율, 주가 및 환율을 제외한 다른 모든 변수들은 (전년동기대비) 증가율로 변환하여 이용하였으며, 자료의 출처 등 데이터에 대한 상세한 설명은 <부록 2>에 포함되어 있다.

2) 이들 변수 외에도 건축허가 (면적 및 동수), 지가 상승률 등의 변수들도 고려하였으나, 이러한 변수들의 경우 (외환위기 기간을 제외하고는) 표본 기간 내에 포함 확률이 0.5를 넘은 경우가 거의 없었으며, 또한 이들까지 다 고려하는 경우 기술적 제약으로 계산 과정에서 추정에 문제가 발생하는 경우가 있어 이들은 고려 대상에서 제외하였다. (추가로 이 변수들을 모두 포함한 예측 모형의 경우, 3.4GHz의 CPU 및 8G RAM의 PC 사양으로는 메모리의 제약으로 제대로 추정이 되지 않았다.)

〈그림 1〉 데이터



<그림 1>에는 이들 데이터의 전반적인 움직임이 나타나 있다. 앞서 언급한 바와 같이, 전국 주택가격은 외환위기시의 급격한 침체에 이어 2000년대에는 2001-2002년 및 2006-2007년의 두 번의 급등기를 보여주고 있으며, 이 과정에서 GDP 및 민간소비 등 실물경기 변수들의 움직임과는 밀접한 관련성을 보이지 않고, 대신 별도의 독자적인 사이클을 가지는 것으로 보이는 점 등은 주목할 만하다.

이러한 예측 변수들의 선정과 함께 모형 설정에 있어 결정되어야 할 사항은 식 (1)에 포함된 예측 대상 변수인 주택가격의 래그 차수 결정이다. 이에 대해서는 각 예측 시계별로 1기부터 4기까지의 다양한 래그 차수를 고려한 모형들을 이용하여 전체 표본에 대한 RMSE(root mean square error)와 로그 우도함수값(likelihood)을 계산하였으며, <표 1>에 보고된 결과를 바탕으로 1분기 및 2분기 후 예측에 대해서는 $p = 1$, 그리

고 4분기 후 예측 모형으로는 $p = 2$ 로 설정하였다.

2. 다른 모형과의 추정 결과 및 예측력 비교

동태적 모형 선택법을 이용한 주택가격 예측 결과를 살펴보기 전에, 우선 이 절에서는 이 모형의 유용성 및 적용 가능성을 점검해 보기 위해 통상적인 모형과의 추정 결과를 비교해 보기로 한다. <표 2>에서는 위에서 고려된 모든 예측 변수를 포함하며, 예측 변수와 계수 추정치가 매기 고정된 것으로 간주하는 통상적인 모형을 최소자승법(OLS)으로 추정한 결과와 함께, 동태적 모형 선택법을 이용한 예측 모형의 추정 결과가 비교되어 있다. 비교의 편의를 위해 동태적 모형 선택법의 경우, 모든 예측 변수를 고려한 1분기 후 예측에 있어, 각 변수 추정치의 최댓값 및 최솟값, 그리고 각 변수의 모형 내 포함 확률의 최댓

<표 1> DMS/DMA 모형의 설정

		RMSE (DMS)	RMSE (DMA)	log likelihood
$h = 1$	$p = 1$	1,474	2,035	-55.768
	$p = 2$	1,559	1,910	-58.435
	$p = 3$	1,605	2,113	-60.616
	$p = 4$	1,626	2,240	-61.954
$h = 2$	$p = 1$	2,594	4,083	-64.532
	$p = 2$	3,280	4,899	-67.356
	$p = 3$	2,919	4,689	-65.383
	$p = 4$	2,967	4,856	-65.967
$h = 4$	$p = 1$	4,583	7,145	-74.236
	$p = 2$	4,240	5,639	-71.463
	$p = 3$	4,480	5,707	-72.631
	$p = 4$	4,451	6,697	-73.764

주: 각 예측 시계별로 다양한 래그 차수를 고려한 경우, 전체 표본에 대한 RMSE와 로그 우도함수 값을 나타낸다.

〈표 2〉 추정 결과의 비교

	OLS		DMS			
			추정치		포함 확률	
	추정치	표준편차	최대	최소	최대	최소
상수항	0.196	0.607	9.707	-1.430		
주택가격(-1)	0.241	0.047	3.196	-2.968		
전세가격	0.080	0.030	0.480	-0.789	0.995	0.058
서울주택가격	0.518	0.039	3.285	-1.453	0.832	0.058
토지거래량	0.012	0.006	0.085	-0.146	0.753	0.002
주거용 건설투자	0.034	0.022	0.406	-0.152	0.619	0.005
GDP	0.017	0.061	1.374	-0.186	0.983	0.023
민간소비	-0.004	0.059	1.313	-0.168	0.994	0.011
소비자물가지수	-0.222	0.138	2.089	-0.234	0.947	0.077
생산자물가지수	0.088	0.052	0.483	-2.407	0.920	0.015
GDP 디플레이터	0.024	0.068	0.900	-1.192	0.980	0.001
M1	-0.049	0.019	0.067	-0.748	0.963	0.007
M2	0.102	0.026	0.187	-0.099	0.817	0.229
CD 금리	-0.444	0.239	0.667	-0.865	0.841	0.103
콜 금리	0.398	0.170	0.982	-0.861	0.461	0.091
회사채 수익률	-0.056	0.136	0.884	-1.754	0.977	0.070
종합주가지수	0.012	0.006	0.149	-0.183	0.842	0.002
원/달러 환율	-0.004	0.016	0.372	-0.420	0.945	0.019

주: OLS 모형은 표본 기간에 걸쳐 고정된 예측 변수와 계수를 상정한 모형이며, 동태적 모형 선택법에서는, 모든 예측 변수를 고려하여 $h = 1$ 및 $p = 1$ 로 설정하였을 때, 각 예측 변수 계수 추정치의 최댓값과 최솟값 및 각 변수들의 포함 확률의 최댓값과 최솟값을 나타낸다.

값 및 최솟값이 보고되어 있다.

우선 통상적인 예측 모형의 추정치는 동태적 모형 선택법의 결과와 비교하여 대략 평균적인 값을 보이는 것으로 나타났으나, 동태적 모형 선택법의 경우 개별 예측 변수들에 대한 추정치는 표본 기간 내에서 상당한 변동을 보이는 것으로 나타났다. 특히 동태적 모형 선택법의 경우에서 상당수의 예측 변수에 대한 계수 추정치는 일반적인 모형을 이용한 추정치에 대한 통상적인 유의수준 하에서의 신뢰 구간을 벗어나 변하는 모

습을 보여, 각 변수의 고정된 영향력을 가정하는 통상적인 예측 모형은 잠재적으로 적절하지 않을 가능성을 보여주고 있다. 또한 각 시점에서 개별 예측 변수들의 모형 내 포함 확률을 살펴보았을 때에도, 대부분의 변수가 최대 90% 이상의 높은 확률을 가지는 반면 동시에 다른 시점에서는 10% 미만의 확률을 가지는 등, 시간에 따라 모형 내 포함 여부에 상당한 변동이 있음을 알 수 있다. 이상의 결과는 주택가격 예측에 있어 주요 예측 변수들의 역할이 각 시기별로 바뀌었을 가

능성이 있으며, 통상적인 예측 모형을 이용할 경우, 이러한 점을 적절히 반영하지 못할 가능성이 있음을 시사한다.

〈표 3〉 모형별 예측력의 비교

모형	$h = 1$	$h = 2$	$h = 4$
DMS($p = 1$)	1.664	2.848	4.804
DMS($p = 2$)	1.607	3.620	4.650
DMA($p = 1$)	1.768	3.478	5.110
DMA($p = 2$)	1.769	3.392	5.078
임의보행	2.116	3.707	6.232
AR(1)	5.131	6.109	7.482
AR(2)	4.504	6.156	7.665
AR(3)	4.498	6.442	8.264
AR(4)	4.612	6.601	7.836
TVP($p = 1$)	2.117	3.589	5.886
TVP($p = 2$)	1.900	3.419	5.801
TVP($p = 3$)	1.912	3.338	6.428
TVP($p = 4$)	1.945	3.392	6.238
OLS($p = 1$)	3.600	4.746	5.997
OLS($p = 2$)	3.035	4.151	5.746
OLS($p = 3$)	3.019	4.076	5.735
OLS($p = 4$)	2.992	4.261	5.750

주: 모형별로 고려한 각 예측 시계에서 주택가격 증가율 예측에 대한 RMSE를 나타낸다. 각 모형에 대한 설명은 본문을 참고하기 바란다.

이제 동태적 모형 선택법과 일반적으로 널리 쓰이는 다른 예측 모형들과의 예측력을 비교해 보기로 한다. 비교 대상모형으로는 간단한 일변량 모형인 임의보행(random walk) 모형, 자기회귀(autoregression, AR) 모형, 그리고 시변모수(time varying parameter, TVP) 모형 및 통상적인

OLS 모형을 고려하였다. 우선 일변량 모형 중, 임의보행 모형은 시계열 데이터를 이용한 예측에서 일종의 기준 모형(benchmark model)으로 널리 사용되고 있으며, 자기회귀 모형은 간단한 시계열 모형임에도 불구하고 단기 예측에 있어 좋은 결과를 보이는 것으로 알려져 있다. 한편, 시변모수 모형은 위의 (1) 식과 같은 모형에서 예측 변수들이 고정되어 있지만, 각 예측 변수에 대한 계수 추정치가 시간에 따라 변하는 모형으로, 부분적으로 경제의 구조 변화를 반영할 수 있는 모형이다. 마지막으로, 통상적인 OLS 모형은 앞서와 마찬가지로 (1)과 같은 식을 예측 변수나 계수 추정치의 변동 등을 고려하지 않고 최소자승법으로 추정된 모형을 이용한 예측이다.

이상의 각 모형들에 대해, 우선 2000년 4분기까지의 표본을 이용하여 1분기, 2분기, 4분기 이후의 주택가격 예측치를 계산하였고, 이후 표본을 한 분기씩 늘려가며 2011년 4분기까지 축차적으로(recursively) 표본외예측(pseudo out of sample forecast)을 실시하였다. 이러한 예측 대상 기간은 2000년대 발생한 두 번의 주택가격 급등기를 포함하는 등 주택가격이 다양한 변동성을 보여 모형별 예측력 비교에 적절한 사례로 생각할 수 있다. 또한 다기간(multi-period) 예측의 경우에는 직접 예측(direct forecasting)을 이용하여 예측치를 계산하였으며,³⁾ 모형별 예측 결과의 비교를 위해서는 RMSE를 이용하였다.

〈표 3〉은 모형별 예측 결과를 요약해서 보여 주고 있다. 우선 동태적 모형 선택법과 동태적 모형 평균법의 비교에 있어서는, 대부분의 경우 동태적 모형 선택법이 우위에 있는 것으로 나타

3) 직접 예측은 반복 예측(iterated forecasting)에 비해 모형설정오류(misspecification) 등의 문제에 대해 상대적으로 결과가 강건한(robust) 것으로 알려져 있다(Marcellino et al. 2006).

났으며, 임의보행 및 자기회귀 모형과의 비교에서는 동태적 모형 선택법/평균법 모두가 모든 예측 시계에서 우위를 보였다. 다음으로 시변모수 모형의 경우, 통상적인 OLS 모형에 비해 대체적으로 예측력에 있어 우위에 있긴 했지만, 동태적 모형 선택법에 비해서는 역시 열등한 것으로 나타났다.

이상의 결과는 동태적 모형 선택법이 주택시장 예측에 있어, 통상적으로 널리 쓰이는 예측 모형들과 비교하여 상대적으로 우위가 있으며 충분히 경쟁력 있는 모형임을 보여주고 있다. 특히 시변모수 모형과의 비교에서 확인할 수 있듯이, 경제의 구조 변화를 반영한 예측에 있어, 모형의 계수 추정치 뿐 아니라, 예측 변수 자체의 예측력이 매 시기 변화할 가능성이 있으며, 주택시장 예측에 있어 이러한 점이 중요한 고려 요인임을 시사하고 있다.

IV. 예측 시계별 주요 주택가격 예측 변수분석

이 절에서는 동태적 모형 선택법을 이용한 앞 절의 결과를 바탕으로, 각 시점에서 주택가격 예측에 있어 주요한 변수들은 무엇이었는지 살펴보기로 한다. 구체적으로 매 시기 1분기, 2분기, 4분기 이후의 주택가격 예측에 있어 각 변수가 예측 모형에 포함될 확률의 변화를 살펴보면, 그 결과가 <그림 2~4>에 정리되어 있다. 이러한 작업은 외환위기시의 주택시장 침체 및 2000년대

들어 발생한 두 차례의 주택가격 급등과 관련하여, 이러한 주택가격 변화의 원인 또는 배경으로 작용한 변수들을 살펴보는 데에도 유용한 함의를 제공할 수 있을 것이다.⁴⁾

1. 예측 시계별 분석

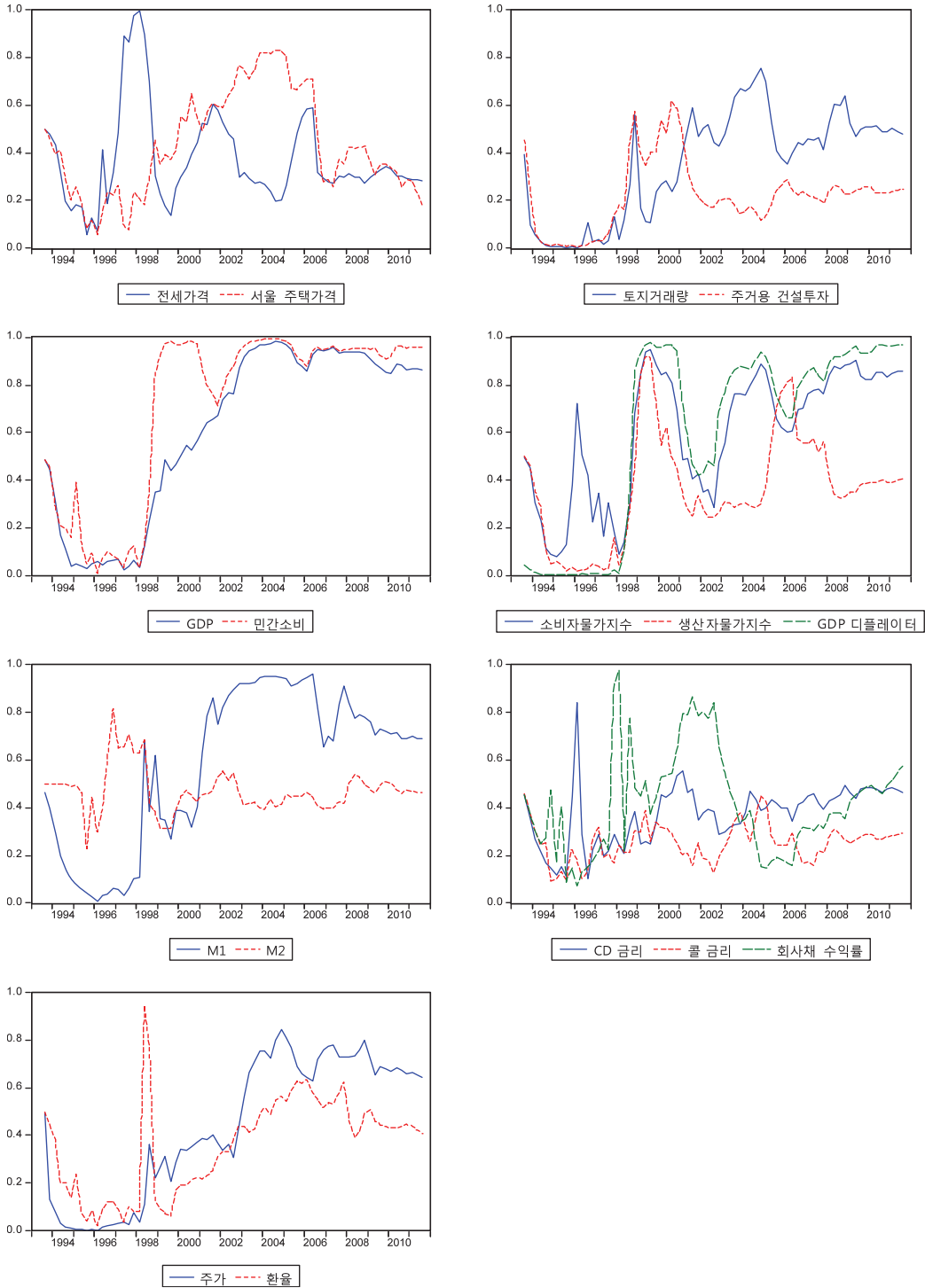
1) 1분기 후 예측

먼저 1분기 후 예측 결과를 살펴보면, 주택시장 관련 변수들의 경우, 전세가격은 외환위기, 2001-2002년 및 2006-2007년 등 3차례에 걸친 주택가격 급변기(특히 상승기)와 맞물려 단기 예측에 주요한 역할을 하는 것으로 나타났다. 또한 서울 주택가격도 2000년대 초반의 주택시장 호황기 무렵부터 2006년의 가격 급등기까지 줄곧 0.6~0.8 가량의 포함 확률을 나타내며 주요 예측 변수로 작용하였으나, 그 외 다른 기간 동안의 예측력은 전반적으로 미미한 것으로 나타났다. 토지거래량도 전반적으로 서울 주택가격과 유사한 결과를 보였으며, 반면 주거용 건설투자는 외환위기 무렵 0.5 안팎의 포함 확률을 보인 것 외에 다른 기간에는 대체적으로 미미한 예측력을 나타냈다.

한편 GDP 및 민간소비 등 실물경기 관련 변수들의 움직임도 흥미롭다. 두 변수 모두 (외환위기 기간 및) 2000년대 초반 이후 지속적으로 90% 안팎의 높은 포함 확률을 보여, 단기적으로는 실물경기가 주택시장 예측에 중요한 변수임을 시사하고 있다. 한편, 흥미롭게도 2000년대 두 차례의 가격 급등기에는 소폭이긴 하나 두 변수의 예측 중요도가 오히려 감소한 점은 주목할 만

4) 식 (1)을 변환하면 주택가격 증가율을 예측 변수들의 이동평균(moving average) 형태로 표현할 수 있고, 이는 주택가격 변화를 과거 각 예측 변수들(에 대한 충격)의 함수로 나타낼 수 있음을 의미한다.

〈그림 2〉 1분기 후 주택가격 예측에 대한 각 변수의 모형 내 포함확률



하다. 다음으로 물가지수의 경우 외환위기 시에는 세 변수 공히 중요한 예측력을 보였으나, 그 이후로는 변수별로 다소 다른 양상을 보였다. 소비자물가지수 및 GDP 디플레이터 경우에는 2000년대 이후 평균적으로 높은 예측력을 보였으나, 흥미롭게도 (앞서와 마찬가지로) 두 차례의 급격한 주택가격 상승기에 대해서는 다른 기간에 비해 포함 확률이 상당정도 감소하였다. 대신 생산자물가지수의 경우, 상대적으로 예측 중요도는 전체적으로 낮았으나 2006년경의 주택가격 상승기에는 주요 예측 변수로 작용하였다.

다음으로 금융관련 변수들을 살펴보면, 우선 M1 증가율은 2000년대 초의 주택시장 호황기 무렵부터 포함 확률이 급격히 증가한 이후 꾸준히 0.8 가량의 높은 확률을 나타낸 반면, M2 증가율은 외환위기 직전을 제외하고는 0.4~0.5 가량의 상대적으로 낮은 확률에 머물렀다.⁵⁾ 마지막으로 이자율 및 기타 금융자산 가격의 경우, 회사채 수익률은 외환위기 및 2001-2002년의 주택시장 호황기에 주요 예측 변수였던 반면, CD 금리 및 콜금리는 전반적으로 0.5 이하의 낮은 확률을 보였다. 주가는 2000년대 초반의 가격 급등기 이후 주요한 예측력을 보였으며, 환율은 외환위기시 일시적으로 확률이 급등하긴 했으나, 주가와 마찬가지로 전반적으로 주택시장 경기와는 관련된 뚜렷한 관련성은 보이지 않았다.

2) 2분기 후 예측

2분기 후 예측의 경우, 서울 주택가격은 앞서와 마찬가지로 외환위기 무렵 및 이후 가격 급등기

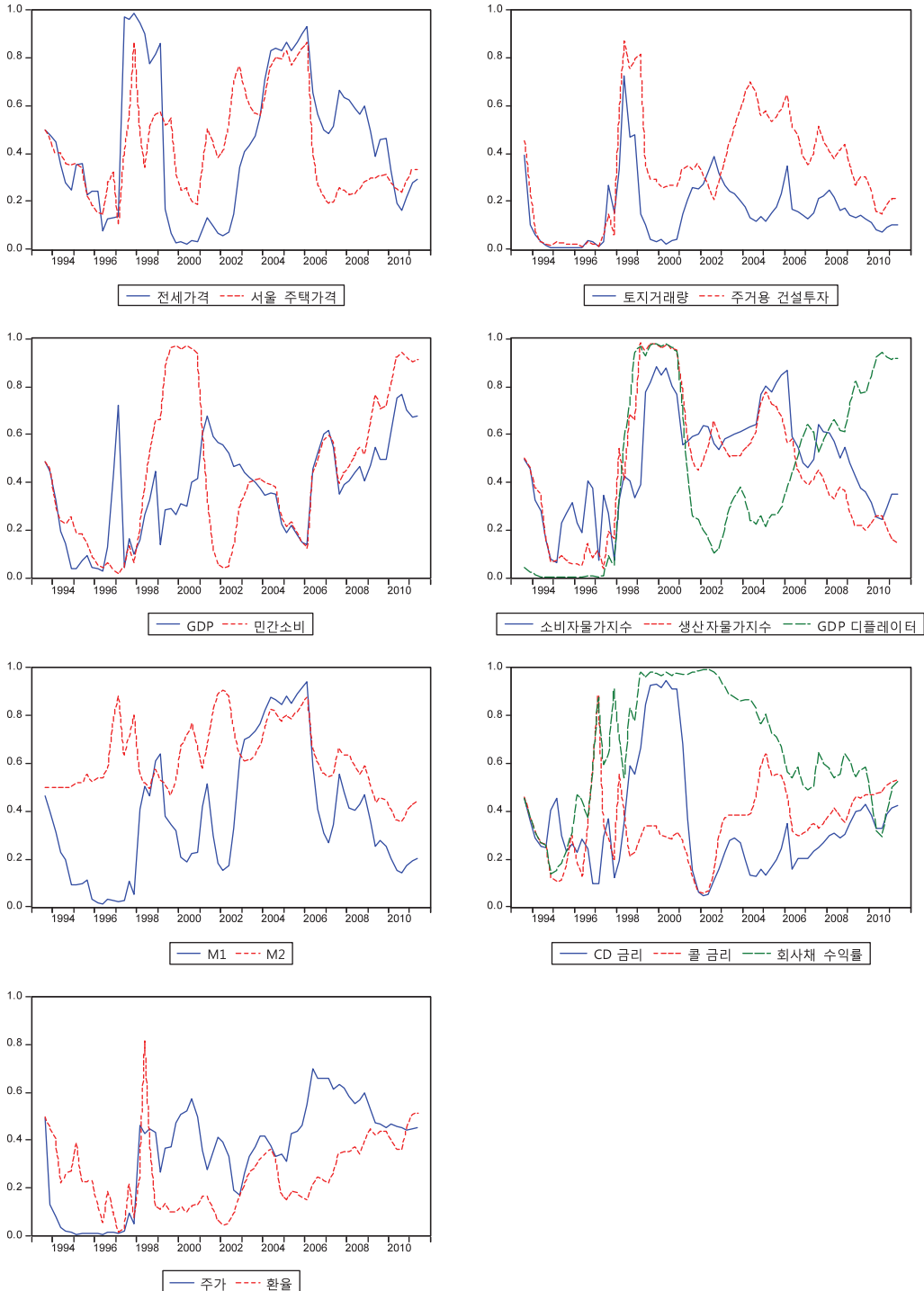
에 전세가격과 함께 중요한 예측 변수 역할을 한 것으로 나타났다. 또한 주거용 건설투자도 확률은 상대적으로 낮은 하지만 서울 주택가격과 유사한 움직임을 보이며 2000년대의 가격 상승기 사이에 중요한 역할을 한 것으로 보인다. 반면, 토지거래량은 2000년대의 주택가격 변동 사이클과 약한 연관성이 있긴 했으나, 외환위기 기간을 제외하고는 전반적으로 미미한 예측 확률을 나타냈다.

거시변수들의 경우, 우선 GDP는 주택시장 호황기 무렵에는 포함 확률이 0.6까지 상승하는 등 최근 두 차례의 주택가격 급등에 어느 정도 유용한 예측력을 보여주었다. 한편 민간소비는 두 차례의 위기에 주택시장 예측에 주요한 변수로 작용하였으나, 흥미롭게도 두 차례의 가격 급등기에는 포함 확률이 낮아지는 등 주택시장 호황기에는 오히려 예측력이 감소하는 양상을 보였다. 두 변수 모두 1분기 예측에 비해 예측력이 전반적으로 감소하는 양상을 보였으며, 대신 2008년 이후로는 꾸준히 예측력이 증가한 점은 주목할 만하다. 다음으로 세 물가변수는 모두 외환위기시 중요한 예측 변수 역할을 하였으나, 앞서와 마찬가지로 그 이후에는 서로 상이한 양상을 보였다. 2000년대 중반까지의 주택시장 호황기에는 소비자물가지수 및 생산자물가지수가 주요 예측 변수 역할을 하였으나 이후로는 예측력이 점차 감소하였고, 대신 GDP 디플레이터의 역할이 꾸준히 증가하는 모습을 보였다.

다음으로 대표적인 유동성 지표인 M1과 M2 모두 2000년대의 주택시장 호황기에 중요한 예측력을 보이는 것으로 나타났으나, 그 이후로는

5) M1의 경우 주택시장의 호황기와 맞물려 상대적으로 더 급격한 증가세를 보인 2006년경에는 예측력이 다소 감소하는 모습을 보였다.

<그림 3> 2분기 후 주택가격 예측에 대한 각 변수의 모형 내 포함확률



점차 예측력이 감소하는 모습을 보였다.⁶⁾ 한편, 이자율 변수들의 경우 외환위기 기간을 제외하고는 주택경기와의 뚜렷한 연관성은 찾기 쉽지 않았으며, 변수별로도 상당부분 상이한 예측력을 보였다. 회사채 수익률은 외환위기 이후 2000년대 중반까지 주요한 예측 변수 역할을 하였으며, 이후에는 콜금리가 단기적으로 하향 안정세를 보이던 2005-2006년 무렵 포함 확률이 크게 상승하는 모습을 보이긴 했으나, 두 차례의 가격 급등기에는 확률이 오히려 감소하는 양상을 보였다. 세 이자율 변수 모두 2006년 이후로는 상대적으로 부진한 예측력을 보였다. 주가는 전반적으로 0.5 안팎의 포함 확률을 보였으나, 주택가격 변동과의 뚜렷한 연관성은 찾기 힘들었으며, 환율은 위기 기간을 제외하고는 전반적으로 0.2 가량의 낮은 포함 확률을 나타냈다.

3) 4분기 후 예측

4분기 이후인 중기 예측에 있어서는, 앞의 단기 예측과는 다른 몇 가지 주목할 만한 차이점이 나타났다. 우선 전세가격 및 서울 주택가격 등 주택가격 변수는 2000년대의 주택시장 호황기에 있어서는 거의 예측력이 없는 것으로 나타났으며, 특히 서울 주택가격은 외환위기 이후 전 기간에 있어 거의 예측력이 없는 것으로 나타났다.

거시변수의 경우에도, 우선 GDP는 두 차례의 위기 기간(특히, 회복기)에 상당한 예측력을 보였으나 2000년대 초반의 주택시장 호황기에는 상

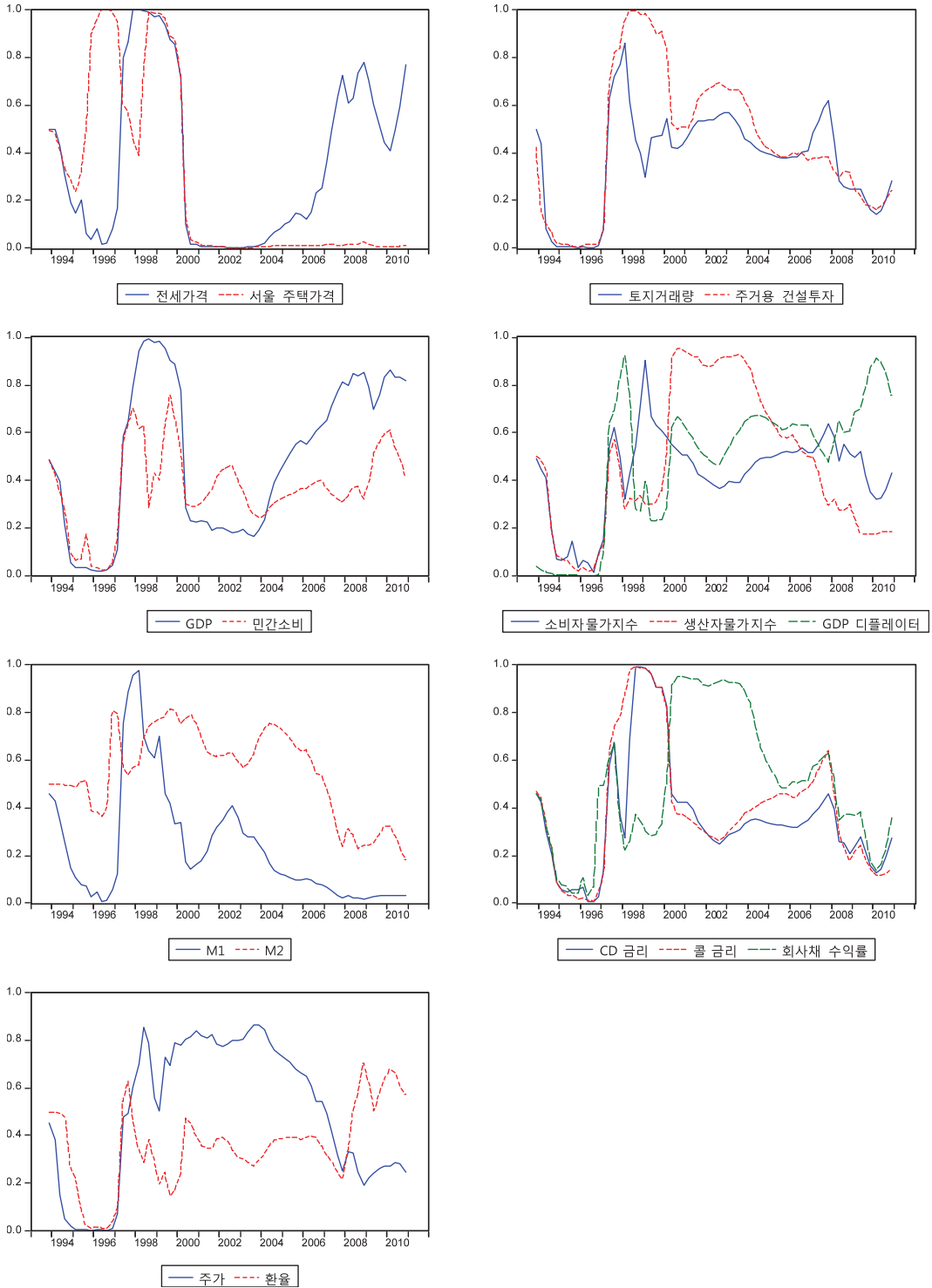
당정도 예측력이 약화되는 모습을 보였다.⁷⁾ 마찬가지로 민간소비의 경우에도, 주택 경기와 뚜렷한 연관성을 보이지 않으며 0.5 이하의 상대적으로 낮은 확률을 나타냈다. 한편 물가지수의 경우, 2000년대 초·중반의 주택시장 호황기에는 생산자물가지수가 주요 예측 변수 역할을 하였으나 이후로는 예측력이 점차 감소하였고, 대신 GDP 디플레이터의 역할이 꾸준히 증가하는 모습을 보였으나, 두 차례의 가격 급등기에는 상대적으로 확률이 소폭 감소한 점은 주목할 만하다. 소비자물가지수의 경우, 외환위기시 상대적으로 포함 확률이 증가하긴 하였으나, 여타기간에는 주택 경기와는 뚜렷한 관련성은 보이지 않으며 0.4~0.5 가량의 확률을 나타냈다.

다음으로 금융시장 관련 변수들을 살펴보면, M1은 외환위기시를 정점으로 그 이후로는 대체적으로 예측력이 감소한 반면, M2는 두 차례의 주택시장 호황기(특히 상승기)와 맞물려 0.6~0.8의 높은 확률을 나타내며, 주요한 예측력을 보인 것으로 나타났다. 이자율의 경우, 앞의 단기 예측의 경우와 유사하게 변수별로 다소 상이한 결과가 나타나, 외환위기 시의 주택시장 침체와 관련해서는 CD 금리와 콜 금리가 결정적인 예측력을 보인 반면, 2000년대 초반의 주택시장 호황의 경우 회사채 수익률이 주된 예측 변수 역할을 하였다. 주가는 주택시장 호황기를 포함한 1998~2006년 동안 0.8 가량의 높은 포함 확률을 보였으나, 그 이후로는 지속적으로 하락하는 모습을

6) 흥미롭게도 1990년대 후반부터 2000년대 초반까지 두 변수의 포함 확률이 상당정도 대칭적인 모습을 보이고 있는데, 이는 해당 기간 동안 대체적으로 경기순응적 움직임을 보인 M2 증가율과는 달리 M1 증가율은 경우에 따라 경기역행적인 움직임을 보인 점에 부분적으로 기인한 것으로 보인다.

7) 주택가격 급등에 별다른 예측력을 보이지 못하던 전세가격과 함께, GDP의 예측 중요도가 2000년대 중반 이후 꾸준히 상승하고 있는 점은 주목할 만하다.

〈그림 4〉 4분기 후 주택가격 예측에 대한 각 변수의 모형 내 포함확률



보였으며, 환율은 두 차례의 위기 기간에 상대적으로 다소 높은 확률을 나타냈지만 전반적으로 0.3 안팎의 낮은 포함 확률을 나타내는 등 주택시장 사이클과 관련된 뚜렷한 예측력은 발견하기 힘들었다.

2. 요약 및 논의

이상의 예측 결과를 2000년대의 주택가격 호황기를 중심으로 요약하자면, 다음과 같이 정리할 수 있다. 우선 1~2 분기 후의 단기 예측의 경우, GDP 및 민간소비 등의 실물경기 변수 및 서울 주택가격 및 전세가격 등 주택가격 변수들이 주된 역할을 한 것으로 보인다. 이와 함께, 부분적으로 토지 거래량 및 건설투자 등 주택시장 관련 변수 및 M1 등이 중요한 예측 변수 역할을 하는 것으로 나타났다. 하지만, 주택가격 급등기에는 민간소비, 소비자물가지수 및 GDP 디플레이터 등의 실물경기 변수의 예측력은, 변수별로 정도는 다소 차이가 있지만, 오히려 상대적으로 감소한 점은 주목할 만하다.⁸⁾

하지만 예측 시계가 증가하면서, 주된 예측 역할을 하는 변수들의 성격 및 그 구성에는 적지 않은 변화가 있었다. 예를 들어 1년 이후의 중기 예측에서는, 주택시장 호황기의 경우 전세가격 및 서울 주택가격 등 주택가격 변수들의 예측력은 큰 폭으로 감소하거나 미미한 수준으로 나타났다. 또한 민간소비 등 실물경기 관련 변수들도 주택가격과 관련된 뚜렷한 움직임은 찾기 힘들었으며, 토지거래량 및 주거용 건설투자의 경우에도, 가격 급등기 사이에는 상대적으로 높은 확률

을 보였으나, 실제 가격 급등시에는 소폭이긴 하나 오히려 예측력이 감소하는 모습을 보였다. 대신 M2, 주가 등 금융관련 변수 및 생산자물가지수 등의 예측력이 증가하며 주된 역할을 하는 모습을 보였다.

이러한 결과는 주택시장이 단기적으로는 주택시장 및 실물경기과 밀접한 관련성을 보이면서 이들 변수가 주요 역할을 했지만, 오히려 가격 상승기의 경우에는 주택가격은 실물경기 또는 경제 전반의 기본 여건(fundamentals)과는 다소 괴리되는 측면이 있었으며, 대신 금융 관련 변수 및 기대 등에 영향을 받았을 가능성을 시사한다. 또한, 일 년 이상의 중기에 있어서는 이러한 경향이 더 뚜렷하게 나타나, 주택가격은 주택시장 및 실물경기 변수들보다는 유동성 및 회사채 수익률, 주가 등 금융시장 관련 변수들에 의해 상대적으로 더 영향을 받았으며, 이러 결과는 자산 가격 변동에 따른 (금융 및 부동산을 포함한) 자산시장에서의 포트폴리오 조정 과정이 주택가격 변동에 주요한 요인으로 작용했을 가능성을 시사한다.

이 논문에서 사용된 예측은 축약형(reduced from) 모형을 기반으로 이루어졌다는 점에서 예측력의 우위를 보인 개별 변수들을 곧바로 주택가격 변동의 직접적인 원인으로 간주하기에는 다소 조심스러운 측면이 있다. 하지만, 이상의 결과는 기존의 연구와도 상당부분 부합되는 측면이 있다. 예를 들어, 주택가격 변동 원인에 대한 기존 연구들에 의하면, 2000년대 이후 주택시장에 대한 콜금리 변경 등을 통한 통화정책의 효과는 제한적이었으며, 대신 가계대출 등 유동성 증가

8) 물가변수와 관련하여, 이러한 결과는 한국의 경우, 주택가격이 인플레이션 헤징 역할을 제대로 하지 못하고 있다는 기존의 연구(최희갑·임병준, 2009; 이영수, 2012)와도 궤를 같이 하는 것으로 보인다.

가 주택가격 급등의 주요 요인 중 하나였으며(손중칠, 2010; Hwang 2012), 이와 함께 GDP, 민간 소비 등 실물경기 변수들도 중요한 역할을 한 것으로 여겨지고 있으며(심성훈, 2006; 손중칠, 2010), 이는 위에서 논의된 결과와 대략적으로 궤를 같이 한다고 할 수 있다. 따라서 이러한 점을 고려한다면 각 시점에서 주된 역할을 한 예측 변수들을 주택가격 변화의 원인으로 간주하는 것이 아주 무리한 해석은 아니라고 여겨지며, 더 나아가 이 논문에서는 이러한 변수들을 예측 시계별로 분석함으로써, 장·단기 영향력을 구분하여 살펴보았다는 점에서 의의가 있다고 할 수 있다.

한편, 주택가격 관련 분석에 대한 기존의 연구에서는 큰 주목을 받지 않았던 회사채 수익률 및 생산자물가지수 등이 주요 예측 변수로 나타난 점은 흥미로운 결과로 여겨진다. 특히, 통상적으로 CD 금리나 소비자물가지수 등의 변수가 주택시장과 더 밀접하게 관련되어 있다고 여겨지는 점을 고려한다면, 이러한 변수들에 비해 회사채 수익률이나 생산자물가지수의 전반적 예측력이 상대적으로 더 높게 나타난 점은 다소 의외의 결과라고 할 수도 있다.

이에 대해서는, 우선 이자율 변수들의 경우, 비록 장기적으로는 고려한 세 변수의 움직임이 유사한 추세를 보이긴 하나, 다른 두 이자율 변수들은 주택가격 상승기에는 뚜렷한 움직임을 보이지 않았던 데 반해, 상대적으로 장기 이자율인 회사채 수익률은 비교적 뚜렷한 하락세를 보이는 것에 부분적으로 기인하는 것으로 보인다. 또한 CD 금리는 주택 금융시장에서의 자금 수요와 직접적으로 관련된 변수로 볼 수 있으나, 앞 절에서의 분석과 함께 최근의 주택가격 상승은 이자율 하락과는 별개로 금융환경의 변화와 맞물려

발생한 가계대출 붐을 통해 촉발·심화되었다는 주장(서근우, 2011; Hwang, 2012) 등을 고려해 본다면, 이러한 이자율 변화를 매개로 한 주택 금융시장에서의 수요 경로는 주택가격 변동에 있어 그다지 중요하지 않았음을 시사하는 것으로 보인다. 대신 앞서 언급한 바와 같이, 유동성 증대 및 이와 맞물린 회사채 수익률의 하향 안정세는 자산 시장에서 포트폴리오 조정으로 이어지고, 이 과정에서 발생한 자금의 움직임이 주택가격 상승의 배경으로 작용했을 가능성을 생각해 볼 수 있다.

한편 생산자물가지수의 경우, 2000년대 초·중반의 주택가격 상승기에 하락·안정세를 보이며 주된 예측 변수 역할을 하는 것으로 나타났는데, 이는 2000년대에 있었던 두 차례의 주택가격 급등의 경우, 안정 추세를 보이던 소비자물가지수 및 GDP 디플레이터와는 달리, 생산자물가지수는 주택가격 급등기에 앞서 다소의 큰 폭의 하락세를 보이는 등, 다른 두 물가지수와는 독립적인 사이클을 보인 점 등에 부분적으로 기인한 것으로 여겨진다. 이러한 생산자물가지수의 변화는 비용절감 요인 등으로 작용하여, 주택 건설의 증가를 통해 주택시장의 공급 측면에서 주택가격에 영향을 미쳤을 가능성을 고려해 볼 수 있다.

V. 맺음말

최근 일부 선진국에서 발생한 급격한 주택가격의 변화 및 그 이후의 조정 과정은 주택시장의 중요성을 다시금 환기시켜 주는 계기로 작용하였으며, 한국의 경우에도 2000년대의 가격 급등을

경험한 이후 최근까지 침체 양상을 보이는 등 주택시장에 대한 우려가 커지고 있다. 이러한 점을 감안할 때, 주택시장에 변동 요인에 대한 실증 분석 및 예측 모형의 개발은 중요한 작업이며, 특히 한국 주택시장은 외환위기 이후 적지 않은 구조적 변화를 경험하고 있는 것으로 여겨져, 이를 적절히 고려하여 반영하는 것이 주택시장 관련 모형 개발에 있어 중요한 고려사항이라 할 수 있다.

이 논문은 위와 같은 점들을 감안하여 통상적인 예측 모형과는 달리, 예측 변수의 중요도와 그 역할이 매기 변할 수 있는 동태적 모형 선택법을 이용하여 한국 주택가격 변화에 대한 예측을 시도하였다. 이러한 접근법은 널리 쓰이는 여타 예측 모형들에 비해 우위를 보이며 경쟁력 있는 모형임을 알 수 있었고, 또한 매 시기별 주택가격 변동과 관련한 주요 변수들의 식별을 통해, 간접적으로 주택가격 변동 요인을 살펴볼 수 있는 등 주택시장 분석에 있어 여러 유용한 함의를 얻을 수 있었다.

하지만, 다음과 같은 점들은 향후에 추가 연구가 필요하다고 할 수 있다. 우선, 근본적으로 이 논문에서 이용된 모형은 축약형 모형으로써 이론적 근거가 다소 취약하여, 경제 이론과 결부시켜 결과를 적극적으로 해석하기에는 다소 조심스러운 측면이 있다. 따라서 동태적 모형 선택의 기본적인 접근법을 유지하면서, 적절한 경제 이론을 바탕으로 한 제약 등을 접목한 구조적 모형의 개발은 차후 유용한 연구가 될 수 있을 것이다. 또한 시계열 자료상의 제약으로 최근 여러 연구에서 주목받고 있는 다양한 주택시장 관련 심리지표들을 충분히 고려하지 못해 이에 대한 분석이 제대로 이루어지지 못한 것도 미흡한 점으로

여겨진다. 이후 충분한 시계열 자료가 축적될 경우, 이들 변수를 포함한 분석은 또 다른 유용한 연구가 될 것으로 기대된다.

부록 1: 동태적 모형 선택법

식 (1)을 동태적 모형 선택법을 반영할 수 있도록, 다음과 같은 상태-공간(state-space) 모형으로 나타내 보자.⁹⁾

$$y_t = Z_t^{(k)}\theta_t^{(k)} + \epsilon_t^{(k)} \quad (\text{A.1})$$

$$\theta_{t+1}^{(k)} = \theta_t^{(k)} + \eta_t^{(k)} \quad (\text{A.2})$$

여기서 $Z_t = [1, x_{t-1}, y_{t-1}, y_{t-2}, \dots, y_{t-p}]$, $\epsilon_t^{(k)} \sim N(0, H_t^{(k)})$ 및 $\eta_t^{(k)} \sim N(0, Q_t^{(k)})$ 이며, 상첨자 k 는 개별 모형 k 에 적용되는 식임을 나타낸다.

t 기까지의 데이터를 $y^t = (y_1, y_2, \dots, y_t)$ 로, t 기에 적용되는 모형을 $L_t \in \{M_1, M_2, \dots, M_K\}$ 로 나타내자. 이제 임의의 시점 $t-1$ 에서, H_t 와 Q_t 값 하에서 모형의 계수 추정치 θ 에 대한 조건부 분포는 다음과 같다고 하자.

$$\theta_{t-1}|y^{t-1} \sim N(\hat{\theta}_{t-1}, \Sigma_{t-1|t-1}) \quad (\text{A.3})$$

이로부터 (칼만 필터링에 의한) 예측(prediction)은 다음과 같이 나타나며,

$$\theta_{t|t-1} \sim N(\hat{\theta}_{t|t-1}, \Sigma_{t|t-1}) \quad (\text{A.4})$$

여기서 $\Sigma_{t|t-1} = \Sigma_{t-1|t-1} + Q_t$ 이다.

이 과정에서 $Q_t = (1 - \lambda^{-1})\Sigma_{t-1|t-1}$ (즉, $\Sigma_{t|t-1} = (1/\lambda)\Sigma_{t-1|t-1}$)로 대체함으로써 위의 계산이 훨씬 용이해지며(Raftery et al.2010), 이러한 근사화 과정을 이용하면 더 이상 Q_t 는 추정할 필요가 없게 된다. 여기서 λ 는 0과 1사이의 값을 가지는 ‘망각 인자(forgetting factor)’로서, j 분기 이전의 데이터는 λ^j 의 가중치를 가지게 된다.¹⁰⁾

마지막으로 칼만 필터링을 통해, 다음과 같은 업데이트(update) 식을 얻을 수 있으며,

$$\theta_t|y^t \sim N(\hat{\theta}_t, \Sigma_{t|t}) \quad (\text{A.5})$$

여기서 조건부 평균과 분산은 다음과 같다.

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \Sigma_{t|t-1}Z_t \times (H_t + Z_t\Sigma_{t|t-1}Z_t')^{-1}(y_t - Z_t\hat{\theta}_{t-1}) \quad (\text{A.6})$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - \Sigma_{t|t-1}Z_t \times (H_t + Z_t\Sigma_{t|t-1}Z_t')^{-1}Z_t\Sigma_{t|t-1} \quad (\text{A.7})$$

이상의 결과를 이용하면, 변수 y_t 에 대한 예측은 다음과 같은 조건부 분포로 나타난다.

$$y_t|y^{t-1} \sim N(Z_t\hat{\theta}_{t|t-1}, H_t + Z_t\Sigma_{t|t-1}Z_t') \quad (\text{A.8})$$

매기마다 위의 식 (A.3)-(A.8)의 과정이 반복되며, 이를 통해 매기 y_t 에 대한 예측치를 계산할 수 있다.

9) 여기서는 1기후 예측의 경우를 고려하였으나, 다른 예측 시계에 대해서도 유사하게 나타낼 수 있다.

10) 이를 달리 해석하자면, 실효 표본(effective window)의 크기는 실제 표본의 $1/(1-\lambda)$ 가 됨을 의미하며, 실제 λ 값으로는 0.95과 0.99사이의 값이 널리 쓰인다.

이제 s 기까지의 데이터를 이용한 t 기 주택가격 예측에 대한 모형 l 의 확률을 $\pi_{t|s,l} = P(L_t = l|y^s)$ 라 하고, 모형 k, l 사이의 이전 확률을 $P_{k,l}$ 로 나타내자. 이 경우 모형 예측식은

$$\pi_{t|t-1,k} = \sum_{l=1}^K \pi_{t-1|t-1,l} P_{k,l} \quad (\text{A.9})$$

로 나타낼 수 있으나, Raftery et al.(2010)에 의하면 이는 다음 식으로 근사화하여 대체할 수 있다.

$$\pi_{t|t-1,k} = \frac{\pi_{t-1|t-1,k}^\alpha}{\sum_{l=1}^K \pi_{t-1|t-1,l}^\alpha} \quad (\text{A.10})$$

여기서 α 는 0과 1사이의 값을 가지는 상수로서, 앞의 λ 와 유사하게 해석할 수 있으며, 이 경우에는 이전 확률을 계산할 필요가 없어 계산 과정이 훨씬 용이해 진다.

이제 앞서와 마찬가지로, 모형 업데이트 식은 다음과 같이 쓸 수 있으며,

$$\pi_{t|t,k} = \frac{\pi_{t|t-1,k} P_k(y_t|y^{t-1})}{\sum_{l=1}^K \pi_{t|t-1,l} P_l(y_t|y^{t-1})} \quad (\text{A.11})$$

여기서 $P_l(y_t|y^{t-1})$ 은 식 (A.8)을 통해 계산할 수 있는 개별 모형의 예측밀도(predictive density)이다.

한편 이상의 모든 계산 과정은 식 (A.2)에 나타난 H_t 의 값을 조건부로 얻은 것이며, 이에 대해서는 Koop and Koroblis (2012)를 따라, 다음

과 같이 계산한 값을 이용하였다.

$$\hat{H}_t = \frac{1}{t^*} \sum_{j=t-t^*+1}^t [(y_t - Z_t^{(i)} \hat{\theta}_{t-1}^{(i)})^2 - Z_t^{(i)} \Sigma_{t|t-1}^{(i)} Z_t^{(i)'}]$$

위 식에서 $t^* = 20$ 으로 두어 해당 시점에서 과거 5년간의 데이터를 이용하여 계산함으로써, 모형의 확률적 변동성(stochastic volatility)이 충분히 반영되도록 하였다.

이상의 과정에서, 동태적 모형 선택법에 의한 예측은 매 시점에서 $\pi_{t|t-1,k}$ 이 가장 큰 모형을 이용하여 예측하는 것이며, 동태적 모형 평균법에 의한 예측은 이러한 (조건부) 확률을 가중치로 개별 모형들의 예측치를 평균하여 예측하는 것이다. 보다 자세한 내용은 Raftery et al.(2010) 및 Koop and Koroblis(2012)를 참고할 수 있다.

본문에 보고된 추정 결과는 별도의 특별한 언급이 없는 한 $\alpha = \lambda = 0.99$ 를 이용했으며, 최초 시점에서 추정치의 사전 분포는 $\theta_0 \sim N(0,100)$, 각 변수의 예측 확률은 동일한 것으로 설정하였다.

부록 2: 데이터

모형의 추정에 사용된 변수 중 주택가격은 국민은행에서 매월 발표하는 ‘주택가격동향’ 자료 중 주택매매가격 및 주택전세가격 종합지수를 사용하였고, 주거용 토지거래량은 국토교통부가 발표하는 ‘지가동향’ 자료를 이용하였다. 또한, GDP, 민간소비, 주거용 건설투자 등의 기타 거시 및 금융 변수들은 한국은행 경제통계시스템(ECOS)으로부터 구하였다. 통화지표인 M1과 M2는 평잔을 이용하였으며, 이자율의 경우 각각 91일 CD 유통수익률, AA-등급 3년만기 회사채 장외거래수익률, 은행간직거래 1일물 무담보 콜금리를 이용하였다.

주택가격 등 계절조정 자료가 발표되지 않는 일부 자료들의 경우에는 X12-ARIMA 방법을 통해 계절조정을 거쳤으며, 자료가 월별로 존재하는 경우는 평균하여 분기별 자료로 변환하여 사용하였다. 마지막으로 이자율, 주가 환율을 제외한 다른 모든 변수들은 (전년동기대비) 증가율로 변환하여 이용하였다. 경기 순환과 관련한 각 변수들의 동태적 정보를 보다 적절히 반영하기 위해서는 자료의 전기대비 증가율을 이용하는 것이 더 바람직할 것이다. 하지만, 한국의 경우 경기동향 분석 등에 있어 아직 전년동기대비 증가율의 사용이 널리 통용되고 있으며, 본 논문에서 사용된 자료 중 주택가격, 물가지수 등 상당수가 계절 조정을 거치지 않은 원자료의 형태로 공표되는 점을 고려한다면, 명절 등 우리나라의 특수성을 고려하지 않은 외국의 계절조정 방법을 적용할 경우 조정된 데이터에도 여전히 계절성이 남아 있을 우려가 있다. 따라서, 이러한 계절성 등의 문제가 상대적으로 적은 전년동기대비 증가율

로 변환한 데이터를 사용하였고, 결과의 강건성(robustness)을 확인해 보기 위해 전기대비 증가율을 사용해서도 추정해 보았지만, 정성적인 결과에는 큰 차이가 없는 것으로 나타났다.

논문접수일 : 2013년 10월 4일

논문심사일 : 2013년 11월 13일

게재확정일 : 2013년 11월 27일

참고문헌

1. 김경환, “외환위기 전후 주택시장 구조변화와 주택정책,” 「경제학연구」 제55권 제4호, 한국경제학회, 2007, pp.369-399
2. 박천규·이영, “주택시장 체감지표의 주택시장 지표 예측력 분석,” 「부동산학연구」 제16집 제1호, 한국부동산분석학회, 2010, pp.131-146
3. 서근우, “은행의 가계대출 급증: 행태론적 원인 분석,” 「금융연구」 제25호 1권, 한국금융학회, 2011, pp.99-127
4. 손정식·김관영·김용순, “부동산가격 예측모형에 관한 연구,” 「주택연구」 제11권 제1호, 한국주택학회, 2002, pp.49-75
5. 손종철, “통화정책 및 실물·금융변수와 주택가격간 동태적 상관관계 분석,” 「경제학연구」 제58권 제2호, 한국경제학회, 2010, pp.179-219
6. 송준혁, “구조변화를 감안한 우리나라 주택시장 분석,” 「응용경제」 제14권 제1호, 한국응용경제학회, 2012, pp.151-186
7. 심성훈, “주택가격과 거시경제변수의 순환변동에 대한 연구-외환위기 전후기간의 비교분석,” 「부동산학연구」, 제12집 제1호, 한국부동산분석학회, 2006, pp.147-163
8. 이영수, “주택가격과 물가의 동학: 한국의 경험,” 「부동산학연구」 제18집 제4호, 한국부동산분석학회, 2012, pp.55-72
9. 정의철, “소비자 심리가 주택시장에 미치는 영향 분석: 주택매매가격을 중심으로,” 「부동산학연구」 제16집 제3호, 한국부동산분석학회, 2010, pp.5-20
10. 최희갑·임병준, “주택의 인플레이션 헤징효과,” 「부동산학연구」 제15집 제3호, 한국부동산분석학회, 2009, pp.51-69
11. Fagin, S., “Recursive Linear Regression Theory, Optimal Filter Theory, and Error Analyses of Optimal Systems,” *IEEE International Convention Record*, Part 1: 1964, pp.216-240
12. Hwang, Y., “Sources of Housing Price Fluctuations: An Analysis Using Sign-restricted VAR,” *Korea and the World Economy*, Vol. 13 No.2, 2012, pp.249-283
13. Jazwinsky, A., *Stochastic Processes and Filtering Theory*, New York, Academic Press, 1970
14. Koop, G. and D. Korobilis, “Forecasting Inflation Using Dynamic Model Averaging,” *International Economic Review*, Vol. 53 No.3, 2012, pp.867-886
15. Marcellino, M., J. Stock, and M. Watson, “A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series,” *Journal of Econometrics*, Vol. 135 No.1-2, 2006, pp.499-526
16. Raftery, A., M. Karny, and P. Ettler, “Online Prediction Under Model Uncertainty via Dynamic Model Averaging: Application to a Cold Rolling Mill,” *Technometrics*, Vol. 52 No.1, 2010, pp.52-66