

우리나라 스태그플레이션 발생확률 추정

박근형* 강규호**

2022년 5월

장단기 스태그플레이션 발생확률 예측은 한국은행의 경기 및 물가안정을 위한 선제적 통화정책 의사결정에 지대한 영향을 미친다. 본 연구는 향후 2년간 우리나라 스태그플레이션 발생확률을 추정한다. 이를 위해 일변수 및 이변수 자기회귀시차(Autoregressive Distributed Lag, ADL) 모델을 이용하여 물가상승률과 실질 경제성장률의 결합예측분포를 생성한다. 이 과정에서 표본외 예측을 통해 베이지안 변수선택과 정밀한 튜닝과정을 거쳐 최적 예측모형을 선택한다. 최근 5년을 대상으로 한 장단기 표본외 예측결과, 대부분의 예측시계에서 일변수 ADL 모형으로 각 변수를 독립적으로 예측했을 때 결합분포 예측력이 극대화되었다. 최적 모형을 이용한 실제 예측결과, 스태그플레이션이 발생할 확률은 2021년 4분기를 기점으로 점차 높아져 2023년 1분기에는 38.4%에 이르는 것으로 나타났다. 이후 다소 하락하기는 하지만 2023년 3분기까지 32.3%로 높은 수준을 유지하다가, 그 이후에는 그 확률이 20%대로 하락하며, 대신 경기호황 확률이 높아진다. 따라서 향후 1년에서 1년 반 동안은 일시적으로 스태그플레이션이 발생할 확률이 정상기에 비해 높은 것으로 추정된다.

JEL Classification Number: G12, C11, E43

핵심주제어: 표본외 결합예측, 통화정책, 경기변동, 베이지안 변수선택

본 연구내용은 집필자들의 개인의견이며 한국은행의 공식견해와는 무관합니다.

* 한국은행 인사경영국 소속 과장 (이메일: keunhyeong.park@bok.or.kr, 전화: 02-759-4114)

** 교신저자, 고려대학교 경제학과 교수 (이메일: kyuho@korea.ac.kr, 전화: 02-3290-5132)

I. 서론

2012년 이후부터 10여 년간 지속된 전 세계적인 저인플레이션 기조가 2021년 들어 다른 양상을 보이기 시작하였다. 미국의 경우 2021년 3월 개인소비지출 물가지수(PCEPI)의 상승률이 2%를 상회한 이후 지속적으로 높아져 2022년 3월 현재 6.6%(CPI 기준 8.5%)에 달하고 있으며, 우리나라도 2021년 3월 소비자물가 상승률이 물가안정목표 2%를 상회한 이후 2022년 3월 현재에는 4.1%를 기록하고 있다. 이와 같은 고인플레이션이 우리나라와 미국뿐 아니라 전 세계 주요국에서 공통으로 나타나는 현상¹⁾임에 따라 향후 물가경로에 대한 불확실성이 크게 확대되었으며 이와 같은 오름세가 구조적인 현상인지, 일시적인 현상인지에 대한 논쟁이 이어지고 있다²⁾. 예를 들어 Summers (2021)는 2009년 글로벌 금융위기 당시 재정지출 규모가 GDP 갭의 절반 정도 수준이었으나 금번 부양책의 경우 GDP갭의 3배를 초과하는 규모라고 평가하면서 인플레이션 위험이 높다고 지적하였다. Goodhart·Pradhan (2021)은 팬데믹 초기 대규모 유동성 지원 정책으로 인해 통화량이 크게 확대된³⁾ 가운데 경기회복 과정에서 pent-up 수요 확대 등과 맞물릴 경우 인플레이션 압력이 증폭될 가능성을 제기하였다. Blanchard (2021)은 경기부양책의 규모가 GDP갭 대비 과도한 수준임에 따라 재정정책에 따른 인플레이션이 수요 압력 뿐만 아니라 기대인플레이션 상승을 통해서도 유발될 가능성을 제기하였다.

반면 인플레이션 상승압력에 대한 반론도 존재한다. Krugman, IMF (2021), Hazell et al. (2020) 등은 대규모 경기부양책이 기대인플레이션을 자극하더라도 필립스곡선의 기울기가 평탄화되어 지속적인 인플레이션 압력으로 작용하기에는 한계가 있다고 평가하였다. 또한 Daly·Chankova (2021)는 과거 감염병 위기 사례를 보면 감염병 위기시에는 물적자본 파괴를 수반하지 않기 때문에 인플레이션 압력이 대규모 전쟁 이후에 비해 제한적인 경향을 보이는 것으로 분석하였다. 이처럼 최근의 고인플레이션 현상에 대한 논쟁이 이어지고 있는 가운데 최근에는 미 연준이 3월 FOMC에서 향후 물가의 상방압력 억제하기 위해 정책금리를 인상하는 등 긴축기조를 본격화하기 시작하였다. 그리고 코로나19의 경제에 대한 부정적 영향이 상존하는 가운데 러시아의 우크라이나 침공 사태까지 발생함에 따라 스태그플레이션에 대한 우려가 고조되고 있다. 이러한 상황에서 경제전망을 통해 선제적 통화정책을 운용하는 중앙은행에 있어 물가상승률과 경제성장률을 정확히 예측하는 것은 무엇보다도 중요한 과제이다.

이에 본 연구에서는 Univariate ADL(이후 Uni-ADL) 모형과 Bivariate ADL(이후 Bi-ADL) 두 종류의 모형을 이용하여 물가상승률과 경제성장률을 예측해 보고자 한다. 우선 Uni-ADL 모형의 경우, 강규호·김정성·신세림 (2021)에서 고안된 Spike-and-Slab prior 기반 베이지안 변수선택기법과 정밀한 튜닝과정을 물가상승률과 경제성장률 예측에 각각 적용한다. 이와 같은

1) 2022년 3월 현재 주요국 물가상승률

G20	영국	유로지역 (19개국)	(독일)	(프랑스)	(이탈리아)	캐나다
7.9%	7.0%	7.4%	7.6%	5.1%	6.8%	6.7%

주: 1) 유로지역(독일, 프랑스, 이탈리아)은 HICP 기준 자료: OECD Stat

2) 이 같은 고인플레이션 현상의 주요 원인으로는 코로나19로 인한 기저효과, 수요회복에 따른 공급병목 등이 지적되고 있다.(박경훈 외, 2021)

3) 글로벌 금융위기시에는 양적완화 등으로 늘어난 본원통화가 금융기관 대출태도 강화, 안전자산 선호 등에 따라 금융기관 지준으로 존재하면서 유동성 증가가 제약된 반면에, 최근에는 상당한 규모의 유동성이 민간부문에 직접 공급되어 이를 적절히 환수하지 못할 경우 물가 상승압력으로 작용할 것이라는 우려를 나타내었다.

예측기법을 사용하는 이유는 다음과 같다. 기본적으로 물가상승률과 경제성장률은 자기회귀 및 기타 외생변수의 시차항이 설명변수로 포함되는데, 물가상승률이나 경제성장률은 이론적으로나 경험적으로 관련이 있을 것으로 예상되는 변수가 상당히 많이 존재한다. 그렇지만 이러한 변수들을 모두 설명변수로 고려하게 되면 종속변수의 움직임을 설명하는 데 있어 통계적으로 유의하지 않은 변수임에도 불구하고 설명변수로 포함되는 과도적합 문제(over-fitting problem)가 가능성이 높아지고 이로 인해 모형의 예측력이 크게 저하될 수 있다. 또한 연구자가 임의로 소수의 예측변수를 사용할 경우에도 누락변수 문제로 인해 예측력이 낮아진다. 이와 같은 누락변수와 과도적합 문제를 큰 계산상의 비용 없이 해결해주는 방법 중 하나가 최근 거시변수 예측에 널리 사용되고 있는 Spike-and-Slab prior 기반 베이지안 변수선택기법이다. 동 기법의 기본적인 아이디어는 k 번째 설명변수 X_k 의 중요성 여부를 나타내는 지시(indicator)함수 γ_k 를 도입함으로써 회귀계수 β_k 가 중요한 설명변수의 계수라면 평균은 0이지만 분산이 큰 정규분포(Slab prior)를, 중요하지 않은 설명변수의 계수라면 β_k 가 0에 가까운 값을 갖도록 분산이 아주 작은 정규분포(Spike prior)를 따르게 하는 것이다. 이를 통해 사후분포 샘플링 과정에서 중요한 설명변수의 계수는 자료의 정보에 의해서 추정되도록 하는 반면, 중요하지 않은 설명변수의 계수는 임의로 0에 더욱 가깝게 한다. 따라서 불필요한 설명변수의 역할이 보다 축소되면서 중요한 설명변수의 예측력이 강화되는 효과가 발생한다⁴⁾는 점에서 장점이 있다. 한편 예측시계 별로 선택되는 변수와 예측의 정확도는 표본내 기간, 최대시차, 설명변수의 단위조정 방법, 사전분포를 구성하는 하이퍼-파라미터 등에 따라 예민하게 달라질 수 있다. 따라서 예측력을 극대화하기 위해서는 수십가지의 사전설정 조합 중에 예측력 평가지표인 로그 사후예측우도(posterior predictive likelihood, PPL) 혹은 RMSE(root mean squared error)가 가장 큰 경우로 최적의 사전설정을 서치(search)하는 튜닝과정을 진행한다.

다음으로 Uni-ADL 모형과 더불어 Bi-ADL 모형을 고려하는 이유는 이론적으로 물가상승률과 경제성장률은 상호 간 영향을 주며 상관관계를 가질 가능성이 있기 때문이다. 본 연구의 목적은 스테그플레이션 발생확률을 추정하는 것이고, 그러기 위해서는 두 변수간 결합예측분포를 생성해야 한다. 이들은 수요충격에 대해서는 양(+)의 관계를, 공급충격에 대해서는 음(-)의 관계를 가지게 되는데, 현실에서는 수요충격과 공급충격 중 어느 충격이 경기변동에 큰 역할을 하는 지에 따라서 부호가 결정된다. 또한 두 충격이 비슷한 비중으로 혼재되어 발생하면 두 변수 간 상관관계가 통계적으로 유의하지 않을 가능성도 있다. 따라서 결합예측력이 극대화되는 모형을 모색하기 위해서는 Uni-ADL과 Bi-ADL 모형을 비교해야 한다. 기본적으로는 Bi-ADL 모형이 상대적으로 일반적인 모형설정이지만 충격의 상관관계가 0일 경우에는 Uni-ADL 모형이 우월할 수 있다. 튜닝과 모형선택은 표본외 예측력을 기준으로 이루어지며 본 고에서 사용된 표본외 예측력 평가기준은 분포예측력 기준인 로그 PPL과 RMSE이다.

본 연구는 2003년 1분기부터 2021년 4분기까지의 분기별 거시자료를 사용하였으며, 최근 5년을 표본외 예측기간으로 설정하였다. 주요 추정결과는 다음과 같다. 첫째, 두 모형을 통한 표본외 예측 및 실제예측 결과, 물가상승률과 경제성장률간 통계적 상관관계가 강하지 않아 두 변수간 상관관계를 고려한 Bi-ADL 모형보다는 두 변수를 독립적으로 예측하는 Uni-ADL 모형이 보다 유용한 것으로 나타났다. 표본외 예측시 로그 PPL과 RMSE 모두 예측시계가 2인 경

4) 보다 자세한 내용은 '장규호 (2016), 『베이지안 계량경제학』, 박영사.'를 참조하기 바란다.

우를 제외하고는 모든 예측시계에서 Uni-ADL 모형의 예측력이 보다 우월한 결과를 나타내었다. 둘째, 표본외 예측에서의 변수선택 결과, 상수항과 시차가 1인 자기회귀항이 모든 예측시계에 걸쳐 물가상승률과 경제성장률 예측에 결정적인 역할을 하는 것으로 나타났다. 아울러 외생 변수 중에서는 주택가격, 주택담보대출 금리 등 주택시장 요인 변수와 환율, 유가 및 미국의 물가지수, GDP, 금리 등 해외요인 관련 변수가 일부 예측시계에서 중요한 변수인 것으로 분석되었다. 이는 그간 우리나라의 주택담보대출 급증과 주택가격 상승세, 높은 해외 의존도 등을 반영하는 결과라 판단된다.

마지막은 실제예측 결과인데, 물가상승률의 추세적 하락세가 이어져 2023년 4분기에는 물가목표 수준에 근접하는 것으로 나타났다. 그리고 경제성장률의 경우 2023년 1분기까지는 하락세가 이어지다가 이후 상승하여 3%대를 기록하는 것으로 나타났다. 이러한 결과로 볼 때 최근의 높은 물가상승률이 당분간은 지속되겠으나 구조적인 현상은 아닌 것으로 추정되며 경제성장률도 점차 회복될 것으로 예상된다. 다만 물가상승률의 신용구간 폭은 2~3%p, 경제성장률은 4~6%p 수준으로 모두 예측에 적지 않은 불확실성이 존재하는 것으로 나타났다. 이러한 결과는 중기적으로 물가안정목표 달성이 어려워 질 수 있음은 물론이고, 자칫하면 스태그플레이션이 발생할 수도 있음을 의미한다. 실제로 물가상승률과 경제성장률에 대한 결합 사후예측분포를 이용하여 등고선(Contour) 그래프를 그려본 결과, 향후 1년~1년반 동안은 스태그플레이션이 나타날 확률이 상당한 것으로 추정되었다. 먼저 Uni-ADL 모형의 경우 2021년 4분기를 기점으로 스태그플레이션 확률이 점차 높아져 5분기 이후 시점인 2023년 1분기에는 38.4%에 이른다. 이후 다소 하락하기는 하지만 7분기 이후 시점인 2023년 3분기까지 32.3%로 높은 수준을 유지하다가, 그 이후에는 20%대로 하락하여 대신 경기호황 확률이 높아지는 모습을 나타냈다. Bi-ADL 모형의 경우에도 Uni-ADL 모형과 유사한 모습을 나타냈다. 먼저 2021년 4분기를 기점으로 스태그플레이션 확률이 점차 높아지다가 3분기 이후 시점인 2022년 3분기에 발생확률 43.9%에 이르고, 이후 시점부터 하락하기 시작하여 8분기 이후 시점인 2023년 4분기에는 17.4%로 스태그플레이션 발생확률이 크게 낮아지는 것으로 추정되었다.

한편 우리나라의 물가상승률 혹은 경제성장률을 예측한 연구로는 성병희(2001), 최인·황승진(2012), 김현학(2015), 강규호·김정성·신세림(2021) 등이 있다. 먼저 성병희(2001)는 원재료가격, 총통화, 쿨금리, 소비자물가, GDP, 실업률 등 6개 변수를 이용한 베이지안 VAR 모형으로 물가상승률과 경제성장률을 예측하였다. 이를 통해 베이지안 VAR 모형을 이용한 예측력이 일반적인 VAR모형이나 대규모 거시계량모형에 뒤떨어지지 않을 뿐만 아니라 예측의 불확실성을 제시할 수 있다는 점에서 장점이 있음을 보여주었다. 최인·황승진(2012)은 Stock and Watson(1999)의 주성분 분석을 이용한 차원축소 방법론을 활용하여 AR 모형과 ADL 모형을 통해 2000년 11월부터 2009년 9월까지의 154개 월별 거시금융 자료로 2009년 10월부터 2011년 6월까지 물가상승률을 예측하였다. 표본외 예측결과, 실업률 또는 이자율 차이와 주성분이 예측변수인 ADL 모형이 가장 정확한 점예측치를 산출하였다.

한편 김현학(2015)도 월별자료를 사용하여 물가상승률을 예측하였는데, 표본기간은 1970년 1월 이후이며 표본외 예측기간은 2000년 1월부터 2013년 12월까지이다. AR 모형과 더불어 주성분 분석과 LASSO, Ridge 회귀분석 등 다양한 차원축소 방법을 채택하여 표본외 예측을 실시한 결과, 개별 모형 중에서는 AR 모형이 가장 좋은 점 예측성적을 보여주었다. 또한 각 모형별

로그 스코어(score)를 기반으로 예측조합을 실시하면 예측력이 다소 향상된다는 것을 보고하였으며, 비모수적 부트스트랩을 이용해서 분포 예측을 실시하였다. 강규호·김정성·신세림 (2021)은 네 개의 물가상승률 예측모형(ADL 모형, AR 모형, 소규모 및 대규모 VAR 모형)을 설정하고 표본외 예측을 통한 모형선택결과를 바탕으로 실제 우리나라 물가상승률 예측치와 불확실성을 수치화하였다. 모형선택결과, 베이지안 변수선택기법이 적용된 ADL 모형과 AR 모형으로 직접 예측하는 것이 VAR 모형으로 반복예측하는 것보다 우월함을 보였으며, 특히 베이지안 변수선택기법 적용 시 광범위하고 엄밀한 튜닝과정을 거치는 것이 예측력 향상에 결정적임을 발견하였다. 그리고 향후 물가상승률은 2022년 하반기까지 2% 내외 수준을 유지하다가 2023년 상반기부터는 1%대로 하락할 것으로 예상하였다.

상기 연구들에 비교할 때 본 연구의 학술적 기여는 베이지안 변수선택과 튜닝과정을 적용함으로써 물가상승률과 경제성장률에 대한 표본외 결합예측력 극대화를 통해 결합 사후예측분포를 제시한 첫 국내연구라는 점이다. 특히 본 연구는 두 모형의 표본외 예측력을 비교함으로써 예측시계가 길어질수록 Bi-ADL 모형보다는 Uni-ADL 모형이 예측에 우월함을 보였으며, 최적 모형으로부터 산출된 결합 사후예측분포를 통해 경기호황, 불황, 스태그플레이션, 안정기의 예측시계 별 발생확률을 제시하였다. 최근과 같이 불확실성이 크게 확대된 여건 하에서 향후 스태그플레이션 발생확률을 구체적인 수치로 제공된다는 것은 정부와 중앙은행 등 정책당국과 민간 경제주체의 효율적이고 선제적인 의사결정에 매우 유용할 것으로 판단된다.

본고의 구성은 다음과 같다. 2장에서는 각 예측모형의 설정과 추정, 표본외 예측 및 튜닝과정을 구체적으로 설명한다. 3장에서는 물가상승률과 경제성장률에 대한 표본외 예측결과 및 실제예측 결과와 실제예측 결과로부터 도출된 결합 사후예측분포를 제시한다. 끝으로 4장에서는 분석결과에 대한 요약 및 정책적 시사점, 본 연구의 한계와 추가 연구방향을 논의한다.

II. 계량분석

앞서 언급한 바와 같이 본 연구는 Uni-ADL 모형과 Bi-ADL 모형 두 종류의 모형을 이용하여 물가상승률과 경제성장률을 예측한다. 주어진 예측변수 하에서 Uni-ADL 모형은 물가상승률과 경제성장률이 각각 독립적으로 예측된다. 이 과정에서 물가상승률의 시차변수는 경제성장률의 예측변수로 사용되며, 반대의 경우도 마찬가지이다. 하지만 총수요나 총공급 충격과 같이 이론적으로 물가상승률에 오는 충격과 경제성장률에 오는 충격이 상관관계를 가질 수 있는데, Uni-ADL 모형은 이를 반영할 수 없다는 한계가 있다. 이런 이유로 충격의 상관관계까지 반영하여 결합 예측분포를 추정할 수 있는 Bi-ADL 모형도 함께 고려된다. 한편 현실경제에서는 수요충격과 공급충격이 혼재되어 발생하기 때문에 두 변수간 상관관계가 통계적으로 유의하게 나타나지 않을 가능성도 있다. 따라서 본고는 Uni-ADL과 Bi-ADL 모형의 표본외 예측력을 비교하여 두 모형 중 최적 예측모형을 선택한다.

좀 더 구체적으로 설명하자면, 우선 Uni-ADL 모형으로는 변수선택 및 사전설정 튜닝과정을 적용하여 물가상승률과 경제성장률을 종속변수로 한 2개의 모형을 개별적으로 추정하고, 표본외 예측과 실제예측을 수행한다. 이를 단계적으로 살펴보면 표본외 예측에서는 변수선택 및 사

전선택 튜닝의 결과 각 모형의 로그 PPL이 가장 큰 경우를 최적의 사전선택이 적용된 최적 모형으로 선정한다. 그리고 실제예측시에는 표본외 예측에서 도출된 예측시계 별 최적 사전선택을 적용하고, 동 사전선택 하에서 다시 한번 변수선택을 실시하여 향후 8분기 중 물가상승률과 경제성장률을 예측한다.

Bi-ADL 모형에서는 변수선택 및 사전선택 튜닝과정이 직접 이루어지지 않는다. 대신 Bi-ADL 모형 추정에 앞서 Univariate CPI 및 GDP ADL 모형을 동시에 추정하고, 이를 이용한 표본외 예측 결과 로그 PPL 합계가 극대화⁵⁾되는 최적 사전선택을 구한다. 그리고 동 최적 사전선택과 함께 이를 실제예측시 적용하여 선택된 변수조합을 그대로 Bi-ADL 모형에 적용하여 추정과 예측을 수행한다.⁶⁾ 이하부터는 각 모형별 설정, 추정방법 및 예측분포 생성과정과 함께 표본외 예측과 튜닝과정에 대해 보다 구체적으로 설명하고자 한다.

II-1. 예측모형 설정

가. h -기 예측 Uni-ADL 모형

본 연구에서는 물가상승률과 경제성장률 예측을 위하여 우선 식 (1)과 같은 Uni-ADL 모형을 고려한다.

$$y_t = \beta_0 + \sum_{j=1}^J \rho_j y_{t-j-h+1} + \sum_{k=1}^K \sum_{j=1}^J \phi_{k,j} z_{k,t-j-h+1} + e_t \quad (1)$$

$$\text{where } e_t | \sigma^2 \sim N(0, \sigma^2), \sigma^2 \sim IG(\alpha_0/2, \delta_0/2)$$

y_t 는 t 시점의 전년동기대비 소비자물가 상승률 또는 실질GDP 증가율이고 y_t 의 조건부 평균이 y_t 의 자기시차와 외생변수($z_{k,t}$)의 시차변수에 의해서 설명되는 것으로 가정한다. 그리고 오차항은 베이지안 추정방법상 편의를 위해 $e_t | \sigma^2 \sim N(0, \sigma^2)$ 으로 조건부 정규분포를 따르며, 오차항의 분산은 $\sigma^2 \sim IG(\alpha_0/2, \delta_0/2)$ 로 켈레분포인 역감마분포를 따르는 것으로 설정한다. $z_{k,t}$ 는 t 시점의 $k(=1, 2, \dots, K)$ 번째 외생변수(예측변수)를, J 는 고려되는 최대 시차의 크기를 나타내는데, 이 때 설명변수는 상수항을 포함하여 최대 ' $J+K \times J+1 = (K+1) \times J+1$ ' 개까지 고려될 수 있다. 예를 들어 $K=4, J=4$ 인 경우 설명변수의 수는 최대 21개가 된다. 이렇게 고려되는 설명변수의 수가 많을 경우 종속변수의 움직임을 설명하는 데 있어서 유의하지 않은 변수임에도 불구하고 설명변수로 포함될 가능성이 높아지고 이로 인해 모형의 예측력이 크게 저하될 수 있다. 이와 같은 과도적합 문제(over-fitting problem)를 완화하기 위해 본 연구에서는

5) 두 Uni-ADL 모형의 로그 PPL 합계 대신 각각의 Uni-ADL 모형의 로그 PPL이 가장 큰 경우로 최적 사전선택을 결정할 수도 있겠으나, 이렇게 할 경우 두 Univariate 모형의 최적 사전선택이 상이해져 Bi-ADL 모형에 그대로 적용하기가 어려워진다.

6) Bi-ADL 모형에 직접 변수선택과 사전선택 튜닝을 적용할 경우 계산부담이 지나치게 커지는 문제가 있다. 뿐만 아니라 Bi-ADL 모형은 SUR 모형 형태이기 때문에 VAR 모형과 달리 두 식의 설명변수나 사전선택을 동일하게 할 필요는 없다. 그러나 Bi-ADL 모형에 변수선택과 사전선택을 적용하여 추정할 경우 CPI와 GDP식의 변수선택과 사전선택을 통일해야 하는 제약이 발생하여 각 식의 예측력을 극대화할 수 있는 최적의 변수선택과 사전선택이 적용되기 어려워진다.

Spike-and-Slab 베이지안 변수선택기법을 도입하고, 예측력을 높이하고자 한다.

이 예측모형은 기본적으로 강규호·김정성·신세림 (2021)에서 소비자물가 상승률 예측을 위해 사용된 모형과 동일한 형태를 따르며, 예측방식도 강규호·김정성·신세림 (2021)과 마찬가지로 반복예측(Iterated forecasting)보다는 장기예측에 보다 유리한 직접예측(Direct forecasting) 방식을 이용한다. 반복예측을 위해서는 모든 변수를 내생변수로 취급하여 예측해야 하기 때문에 축약형 벡터자기회귀모형이 사용된다. 강규호·김정성·신세림 (2021)의 연구에 의하면 반복예측의 표본외 정확도가 직접예측에 비해 현저히 낮게 나타났으며, 이런 이유도 본 연구에서도 반복예측을 고려하지 않는다.

Spike-and-Slab 베이지안 변수선택기법의 기본적인 아이디어는 모형에 포함될만한 후보변수들을 모두 모형에 포함하는 대신 데이터를 통해 해당 변수의 중요도를 판단하여 중요한 변수를 확률적으로 선택하는 것이다. 이를 위해 ‘포함여부를 결정하는 지시변수(inclusion indicator) 파라미터’를 도입하고, 동 변수를 매개로 하여 중요하지 않은 변수들의 계수 값은 0에 가깝게, 중요한 변수의 계수 값은 0과 상당한 차이가 있게 샘플링 되도록 한다. 이와 같은 변수선택 방식은 추정이 쉬운데다, 변수의 수가 많은 경우에도 비교적 계산부담을 줄일 수 있다⁷⁾는 장점이 있다. 이를 위해 우선 식 (1)을 다음과 같이 종속변수의 시차항과 외생변수의 시차항을 포괄한 형태로 변형하여 나타낸다.

$$y_t = \beta_0 + \sum_{k=1}^{K+1} \sum_{j=1}^J \beta_{k,j} x_{k,t-j-h+1} + e_t \Rightarrow Y = X\beta + e \quad (2)$$

Y 는 소비자물가 상승률 혹은 실질GDP 증가율의 시계열벡터, X 는 Y 와 외생변수의 시차항 및 상수항으로 이루어진 설명변수⁸⁾ 행렬이고, $\beta = (\beta_0, \beta_{1,1}, \beta_{1,2}, \dots, \beta_{K+1,J-1}, \beta_{K+1,J})'$ 는 각 예측변수의 회귀계수 벡터이다. 이 때 $\beta_{k,j}$ 에 대한 Spike-and-Slab 조건부 사전분포는 식 (3)과 같이 사전평균이 0이고, 사전분산은 (0 보다) 충분히 큰 값인 b_1 과 0에 매우 가까운 (그러나 0은 아닌) b_0 로 구성된 혼합 정규분포를 따르는 것으로 가정한다. 이러한 사전분포를 설정함으로써 실제 계수가 중요하지 않은 경우 임의로 계수 값을 0에 극도로 가깝게 하고, 중요한 경우(즉 실제 계수가 0과 크게 다른 경우), 데이터에 의해 추정되도록 할 수 있다.

$$\begin{aligned} \beta_{k,j} | \gamma_{k,j}, b_0, b_1 &\sim N(0, b_{k,j}) \text{ for } k = 1, 2, \dots, K+1, j = 1, 2, \dots, J \\ b_{k,j} &= \gamma_{k,j} \times b_1 + (1 - \gamma_{k,j}) \times b_0 \end{aligned} \quad (3)$$

여기서 $\gamma_{k,j}$ 는 (k, j) 번째 예측변수가 예측에 있어 중요한 변수라면 1의 값을 가지고, 그렇지 않은 경우 0의 값을 가지는 지시변수를 나타낸다. 만약 어떤 예측변수가 중요하다면($\gamma_{k,j} = 1$), $\beta_{k,j}$ 의 사전분산은 (0 보다) 충분히 큰 값인 b_1 이 되고, 0과 상당히 다른 값으로 추정될 확률이 높아

7) 예를 들어, 설명변수가 h 개인 단변량 회귀분석에서 주변우도를 통해 모형을 선택하는 경우 비교대상이 되는 모형의 수는 2^h 개가 된다. 이에 비해 Spike-and-Slab 베이지안 변수선택기법 베이지안 MCMC(Markov Chain Monte Carlo) 확률적 탐색 알고리즘을 통해 계산비용을 크게 경감시킬 수 있다.

8) 이 때 두 Uni-ADL 모형에서 최초의 외생변수 집합은 CPI와 GDP를 제외한 모든 외생변수들을 공통적으로 포함하는 가운데, CPI와 GDP는 각각 GDP ADL식과 CPI ADL 식에 교차하여 포함된다.

지게 된다. 반면 어떤 예측변수가 중요하지 않다면($\gamma_{k,j} = 0$), $\beta_{k,j}$ 의 사전분산은 0에 매우 가까운 값인 b_0 가 되고, 0에 가까운(즉, 유의하지 않은) 값으로 추정될 확률이 높다. 이러한 $\gamma_{k,j}$ 의 사전분포는 $\gamma_{k,j} = 1$ 일 확률이 p (inclusion probability)인 베르누이 분포를 따르며, 동 지시변수의 사후분포에서 도출되는 사후평균이 각 예측변수의 중요도를 나타낸다. 각 예측변수의 중요도는 0과 1사이의 연속적인 값으로, 1에 가까울수록 깃스 샘플링 반복시행 중에서 그 예측변수가 사후예측 분포 생성에 사용되는 빈도가 많으며 예측력이 높다는 것을 의미한다.

한편 변수선택과정에서 하이퍼-파라미터에 대한 임의성을 최대한 배제하고 데이터로부터 많은 정보를 모형에 반영하기 위해 p 와 b_0 , b_1 에 대해서도 사전분포를 설정하였다. p 는 0과 1사이의 확률을 나타내는 연속된 값이므로 식 (4)와 같이 베타분포를, b_0 와 b_1 은 역감마분포를 가정하였다.

$$\begin{aligned} p &\sim B(a_0, c_0) \\ b_0 &\sim IG(\alpha_{00}/2, \delta_{00}/2), b_1 \sim IG(\alpha_{01}/2, \delta_{01}/2) \end{aligned} \quad (4)$$

단, 여기서 b_1 은 충분히 큰 값이고 b_0 는 0과 매우 가까운 값이어야 하므로 역감마분포의 사전평균이 $0.5\delta_{00}/(0.5\alpha_{00} - 1)$ 임을 감안하여 하이퍼-파라미터 δ_{00}/α_{00} 를 0에 가깝게, δ_{01}/α_{01} 은 충분히 크게 설정하도록 한다. 이렇게 설정하는 경우 그 분포의 형태로 인해 b_0 를 Spike 분산, b_1 을 Slab 분산이라고 지칭한다.

나. h -기 예측 Bi-ADL 모형

앞서 설명하였듯이 Uni-ADL 모형을 이용하여 물가상승률과 경제성장률을 독립적으로 예측하는 방식은 잠재적으로 존재할지도 모르는 두 변수의 (오차항간) 상관관계가 고려되지 않을 수 있다. 이에 두 변수간 통계적 상관관계가 강하게 존재할 경우 이 점을 모형에 반영할 수 있도록 다음과 같은 Bi-ADL 모형을 이용하여 물가상승률과 경제성장률을 예측(직접예측)한다.

$$\begin{aligned} y_{1t} &= x_{1t}'\beta_1 + u_{1t} \\ y_{2t} &= x_{2t}'\beta_2 + u_{2t} \\ \text{where } u_t|\Omega &= \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \sim \text{Normal}(0, \Omega) \end{aligned} \quad (5)$$

y_{1t} 와 y_{2t} 는 각각 소비자물가 상승률과 실질GDP 증가율이다. 두 종속변수에 대한 선형회귀식은 Uni-ADL 모형과 동일한 형태로 y_{1t} , y_{2t} 각각의 조건부 평균이 상수항, 자기시차 및 외생변수의 시차변수에 의해서 설명된다. 그리고 오차항도 Uni-ADL 모형과 마찬가지로 조건부 정규분포를 따른다. 다만 Uni-ADL 모형과 달리 오차항의 분산-공분산 행렬 Ω 가 두 변수의 오차항간 상관관계를 고려한(즉, 공분산이 0이 아닐 수 있음) 2×2 행렬이기 때문에 $\Omega \sim IW(v_0, R_0)$ (단, v_0 는 자유도, R_0 는 스케일 행렬)로 역위샷분포를 따르는 것으로 설정한다.

아울러 동 Bi-ADL 모형에서는 변수선택 및 사전설정 튜닝과정이 직접 이루어지지 않는다. 대신 Bi-ADL 모형 추정에 앞서 Univariate CPI 및 GDP ADL 모형을 동시에 추정하고, 이를 이용한 표본외 예측 결과 로그 PPL 합계가 극대화되는 최적 사전설정을 구한다. 그리고 동 최적 사전설정과 함께 이를 실제예측시 적용하여 선택된 변수조합을 그대로 Bi-ADL 모형에 적용하여 물가상승률과 경제성장률을 예측을 수행한다. 예를 들어 Uni-ADL 모형으로 실제예측시 최대시차 4를 기준으로 Univariate CPI ADL 모형에서는 [상수항, $y_{1,t-3}$, $y_{2,t-1}$, $z_{2,t-2}$]의 설명변수 조합이 선택되고, Univariate GDP ADL 모형에서는 [$y_{1,t-2}$, $y_{2,t-2}$, $z_{3,t-2}$]의 설명변수 조합이 선택되었다면 h -기 예측 Bi-ADL 모형은 식 (6)과 같이 구성될 것이다.

$$\begin{aligned}
 y_{1,t+h} &= \beta_{1,0} + \beta_{1,13}y_{1,t-3+1} + \beta_{1,21}y_{2,t-1+1} + \beta_{1,22}z_{2,t-2+1} + u_{1,t+h} \\
 y_{2,t+h} &= \beta_{2,12}y_{1,t-2+1} + \beta_{2,22}y_{2,t-2+1} + \beta_{2,32}z_{3,t-2+1} + u_{2,t+h} \\
 \Rightarrow \begin{bmatrix} y_{1,t+h} \\ y_{2,t+h} \end{bmatrix} &= \begin{bmatrix} x_{1,t-j+1}' & 0_{(1 \times K_2)} \\ 0_{(1 \times K_1)} & x_{2,t-j+1}' \end{bmatrix} \begin{bmatrix} \beta_{1,kj} \\ \beta_{2,kj} \end{bmatrix} + \begin{bmatrix} u_{1,t+h} \\ u_{2,t+h} \end{bmatrix} \Rightarrow y_{t+h} = x_{t-j+1}' \beta_{kj} + u_{t+h} \\
 (2 \times 1) \quad (2 \times 7) \quad (7 \times 1) \quad (2 \times 1)
 \end{aligned}
 \tag{6}$$

단, $x_{1,t-j+1} = (1, y_{1,t-3+1}, y_{2,t-1+1}, z_{2,t-2+1})'$, $x_{2,t-j+1} = (y_{1,t-2+1}, y_{2,t-2+1}, z_{3,t-2+1})'$,
 $\beta_1 = (\beta_{1,0}, \beta_{1,13}, \beta_{1,21}, \beta_{1,22})'$, $\beta_2 = (\beta_{2,12}, \beta_{2,22}, \beta_{2,32})'$ 이다.

이 때 회귀계수 $\beta = [\beta_1 \ \beta_2]'$ 에 대한 사전분포는 다음과 같이 β_1 과 β_2 가 독립인 정규분포를 따르는 것으로 설정한다. 그리고 K_1 과 K_2 는 각각 첫번째와 두번째 회귀식의 (상수항 포함) 설명변수의 수를 나타내는 것으로 동 예시에서는 각각 4, 3의 값을 가진다.

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \sim Normal \left(\beta_0 = \begin{bmatrix} \beta_{1,0} \\ \beta_{2,0} \end{bmatrix}, B_0 = \begin{bmatrix} B_{1,0} & 0_{(K_1 \times K_2)} \\ 0_{(K_1 \times K_2)} & B_{2,0} \end{bmatrix} \right) \tag{7}$$

II-2. 추정방법

우선 Uni-ADL 모형의 사후분포 추정과정은 강규호·김정성·신세림 (2021)에서 이미 설명된 바가 있으므로 <부록 A>를 참조하길 바란다. 그리고 Bi-ADL 모형의 사후분포 추정과정은 SUR(Seemingly Unrelated Regression) 모형과 동일하므로 이를 중점적으로 설명하고자 한다.

가. SUR 모형의 사후분포 추정

SUR 모형의 사후 샘플링 알고리즘은 일반적인 다중선형회귀의 사후분포 추정과정과 기본적으로 동일한 가운데, 대상변수가 스칼라(scalar)에서 벡터(vector)로 바뀐다는 점에서만 차이가 있다. 따라서 SUR 모형의 파라미터인 β 와 Ω 의 사후분포는 이들의 완전조건부분포를 구한 후 깃스 샘플링 알고리즘을 통해 도출할 수 있다.

(β , Ω 의 완전조건부분포)

식 (5)에서와 같이 SUR 모형이 표현될 경우, β 의 완전조건부분포는 다음과 같은 정규분포를 따른다.

$$\begin{aligned} \beta | Y, \Omega &\sim N(\beta_1, B_1) \\ \text{단, } B_1 &= \left(B_0^{-1} + \sum_{t=J+1}^T x_t \Omega^{-1} x_t' \right)^{-1} \\ \beta_1 &= B_1 \left(B_0^{-1} \beta_0 + \sum_{t=J+1}^T x_t \Omega^{-1} y_t \right) \end{aligned} \quad (8)$$

그리고 Ω 의 완전조건부분포는 다음과 같은 역위샷분포를 따른다.

$$\Omega | Y, \beta \sim IW \left(v_0 + (T - J), R_0 + \sum_{t=J+1}^T (y_t - x_t' \beta)(y_t - x_t' \beta)' \right) \quad (9)$$

(깃스 샘플링 알고리즘)

β 와 Ω 의 완전조건부분포를 이용한 깃스 샘플링 알고리즘은 다음과 같다.

Step 0) 초기값 $\Omega^{(0)}$ 를 사전평균($=R_0/(V_0 - K - 1)$)으로 설정하고, 샘플링 차수 $i=1$ 로 둔다.

Step 1) β 의 완전조건부분포 $\beta | Y, \Omega^{(i-1)} \sim N(\beta_1, B_1)$ 으로부터 $\beta^{(i)}$ 를 샘플링한다.

$$\text{where } B_1 = \left(B_0^{-1} + \sum_{t=J+1}^T x_t \Omega^{(i-1)^{-1}} x_t' \right)^{-1}, \quad \beta_1 = B_1 \left(B_0^{-1} \beta_0 + \sum_{t=J+1}^T x_t \Omega^{(i-1)^{-1}} y_t \right)$$

Step 2) Ω 의 완전조건부분포 $\Omega | Y, \beta^{(i)} \sim IW \left(v_0 + (T - J), R_0 + \sum_{t=J+1}^T (y_t - x_t' \beta^{(i)})(y_t - x_t' \beta^{(i)})' \right)$ 로부터 $\Omega^{(i)}$ 를 샘플링한다.

Step 3) $i=i+1$ 로 설정하고, $i \leq n$ (반복시행 총 횟수)이면 1단계로 돌아간다.

(사후예측분포 도출)

파라미터의 사후분포가 해석적으로 도출되지 않으므로⁹⁾ 사후예측분포도 파라미터의 사후분포와 마찬가지로 수치적 적분법을 통해 근사적으로 도출할 수 있다. 예를 들어, Uni-ADL 모형과 Bi-ADL 모형의 h 기 이후 물가상승률(또는 경제성장률)은 각 깃스 샘플링 반복시행마다 파라미터의 사후 샘플이 주어지면 다음과 같은 조건부 정규분포에서 생성된다.

$$\text{Uni-ADL 모형: } y_{t+h} \sim N\left(\beta_0 + \sum_{k=1}^{K+1} \sum_{j=1}^J \beta_{k,j} x_{k,t-j+1}, \sigma^2\right) \quad (10)$$

$$\text{Bi-ADL 모형: } y_{t+h} \sim N(x_{t-j+1}' \beta_{kj}, \Omega)$$

II-3. 표본외 예측

본 연구에서는 2003년 1분기부터 2021년 4분기 자료를 이용하여 향후 2년(8분기)에 대한 실제예측을 수행하고자 한다. 하지만 그에 앞서 표본외 예측(out-of-sample forecasting)을 통해 모형의 예측력을 평가할 필요가 있다. 본 연구에서는 최근 5년 기간(2017년 1분기~2021년 4분기)을 대상으로 각 모형의 표본외 예측력을 평가하는데, 한 가지 유의할 점은 예측시계(h)에 따라 사용되는 표본내 기간(in-sample period)이 달라진다는 것이다. 예를 들어 표본내 기간(in-sample period) 구성이 확장 방식(recursive)인 경우, 예측시계가 3분기 이후라면 2003년 1분기부터 2016년 2분기까지의 자료를 이용하여 2017년 1분기 값을, 2016년 3분기까지의 자료를 이용하여 2017년 2분기 값을 예측한다. 이러한 방식으로 2021년 4분기까지 총 20번의 표본외 예측이 수행된다. 그리고 표본내 기간 구성이 롤링 방식(rolling)인 경우에는 2017년 1분기는 확장 방식과 동일하게 2003년 1분기~2016년 2분기이지만, 2017년 2분기는 2003년 2분기~2016년 3분기가 표본내 기간이 된다.

표본외 예측시 모형의 예측력 판단을 위한 지표로는 로그 PPL을 이용한다.¹⁰⁾ 로그 PPL은 각 모형으로부터 도출된 사후예측분포로부터 실제 값이 발생할 확률 즉, 실제 값의 로그 사후 예측밀도(로그 Posterior Predictive Density, 이하 로그 PPD)를 계산하고, 이를 표본외 예측 전 기간에 걸쳐 합계하여 산출된다. 다만 여기서 한 가지 유의할 점은 Uni-ADL 모형은 단변량 정규분포로부터 로그 PPD가 계산되지만, Bi-ADL 모형의 경우 분산-공분산 행렬에서의 correlation을 고려하여 다변량 정규분포로부터 결합(Joint) 로그 PPD가 산출된다는 것이다. 각 모형에 대한 로그 PPD 산출산식은 다음과 같으며, 표본외 예측시 Univariate CPI ADL 및 GDP ADL 모형의 로그 PPL을 합계한 값과 Bi-ADL 모형의 결합(Joint) 로그 PPL 값을 비교함으로써 예측력을 평가할 수 있다.

9) $\beta \sim N(\beta_0, B_0)$ 와 같이 β 의 사전분포 분산이 오차항의 분산-공분산 행렬(Ω)에 비례하는 형태가 아닌 경우 사전분포를 보다 용이하게 설정할 수 있다는 장점이 있으나, β 와 Ω 의 결합사후분포 및 각각의 주변사후분포를 해석적으로(analytically) 도출할 수 없게 된다. 대신 이런 경우 β 와 Ω 각각의 완전 조건부 사후분포를 도출하고 깃스 샘플링을 통해 수치적으로(numerically) β 와 Ω 의 (주변)사후분포를 도출해 낼 수 있다.

10) 로그 PPL 이외에 RMSE(root mean squared errors)도 모형의 예측력(주로 점예측) 평가를 위해 자주 이용되는 지표이나 본 연구의 주목적이 구간예측에 있으므로 로그 PPL을 예측력 평가지표로 이용한다.

(Uni-ADL 모형의 로그 PPD)

$$\text{Ln}\left(\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(y_{t+h}-\mu)^2\right)\right) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}(y_{t+h}-\mu)^2 \quad (11)$$

$$\text{단, } \mu = \beta_0 + \sum_{k=1}^{K+1} \sum_{j=1}^J \beta_{k,j} x_{k,t-j+1}$$

(Bi-ADL 모형의 로그 PPD)

$$\text{Ln}\left(\frac{1}{(2\pi)^{k/2}|\Omega|^{1/2}}\exp\left(-\frac{1}{2}(y_{t+h}-\mu)'\Omega^{-1}(y_{t+h}-\mu)\right)\right) \quad (12)$$

$$= -\frac{1}{2}\ln((2\pi)^{k/2}|\Omega|^{1/2}) - \frac{1}{2}(y_{t+h}-\mu)'\Omega^{-1}(y_{t+h}-\mu)$$

$$\text{단, } \mu = x_{t-j+1}'\beta_{kj}$$

본 연구에서는 표본외 예측결과, (i) 물가상승률과 경제성장률 Uni-ADL 모형 각각의 로그 PPL이 가장 큰 경우를 최적 사전설정이 적용된 Best 모형으로 선정한다. 실제예측시에는 표본외 예측에서 도출된 예측시계 별 최적 사전설정을 적용하고, 동 사전설정 하에서 다시 한번 변수선택을 실시하여 향후 8분기 중 물가상승률과 경제성장률을 예측한다. 그리고 (ii) Univariate CPI 및 GDP ADL 모형을 동시에 추정하여 표본외 예측한 결과 로그 PPL 합계가 극대화되는 최적 사전설정을 구한다. 그리고 동 최적설정과 함께 이를 실제예측시 적용하여 선택된 변수조합을 Bi-ADL 모형에 그대로 적용하여 물가상승률과 경제성장률을 예측한다.

II-4. 튜닝

예측시계 별로 선택되는 변수와 예측 정확도는 표본내 기간, 사전분포를 구성하는 하이퍼-파라미터 등에 따라 예민하게 달라질 수 있다. 이에 Uni-ADL 모형에 강규호·김정성·신세림 (2021)에 따른 다섯 가지 사전설정에 대한 튜닝과정을 적용함으로써 표본외 예측력을 극대화할 수 있다. 튜닝의 대상이 되는 사전설정은 (i) Spike 분산(b_0)의 사전분포(δ_{00}/α_{00}), (ii) 예측변수의 시차, (iii) 표본내 기간, (iv) 설명변수의 단위조정(rescaling) 방법, (v) 변수집합 필터링이다.

(Spike 분산의 평균)

베이시안 변수선택 모형에서 Spike 분산 b_0 의 값이 너무 작은 경우 중요하지 않은 변수가 모형에서 제외되지 않을 수 있고, 반대로 b_0 가 너무 큰 경우에는 중요한 변수가 누락될 수 있다. 따라서 b_0 의 사전평균($0.5\delta_{00}/(0.5\alpha_{00}-1)$)을 어떤 값으로 하는지에 따라 선택되는 변수와 예측력이 크게 달

라진다. 이에 본 연구에서는 α_{00} 를 1,000으로 고정하는 한편, b_0 의 사전평균을 반영하는 δ_{00}/α_{00} 값을 세 가지 경우(10^{-6} , 10^{-8} , 10^{-10})로 비교하여 예측력을 극대화하는 최적의 사전평균을 탐색한다.

(시차)

마찬가지로 시차를 어느정도까지 고려하느냐에 따라 예측력이 크게 달라진다. 이에 본 연구는 예측변수의 시차를 1분기부터 최대 4분기까지 고려한다.

(표본내 기간)

예측 시 표본내 기간을 구성하는 방법에는 표본내 기간의 첫 시점이 고정된 채로 마지막 시점이 점차 확대되는 확장 방식(recursive)과 표본내 기간의 크기가 고정된 채로 첫 시점과 마지막 시점이 동일하게 변동하는 롤링 방식(rolling) 방식이 있다. 분석대상 기간중 구조변화가 존재할 경우 두 방식에 따른 추정 및 예측결과는 크게 달라질 수 있다. 특히 우리나라의 소비자물가는 2012년을 전후하여 한국은행의 물가안정목표치인 2%를 하회하는 등 구조변화가 존재할 가능성이 있다. 따라서 확장 회귀분석과 더불어 롤링 회귀분석을 롤링 윈도우의 크기 2가지 경우(20, 40분기)에 대하여 적용하고 예측력을 비교한다.

(설명변수의 단위조정)

설명변수의 단위 크기에 따라 계수 값이 달라지고 이로 인해 변수선택 및 예측력에도 차이가 발생할 수 있다. 따라서 각 예측변수의 예측력 기여도를 정확히 판단하기 위해서는 예측변수의 크기를 사전에 조정할 필요가 있다. 변수의 단위조정 방식은 Z-score 방법과 미네소타 사전분포 형태(Minnesota prior type)가 주로 이용된다. Z-score 방법은 각 변수의 표본평균은 0, 편차는 1이 되도록 표준화하는 방법이다. 미네소타 형태는 베이지안 VAR 모형 추정시 차원축소를 위해 자주 이용되는 미네소타 사전분포¹¹⁾에서와 같이 각 변수에 (ω_y/ω_k) 를 곱하여 상대적으로 잡음이 많은 설명변수에는 작은 가중치를 부여하는 방식이다. 여기서 ω_y 와 ω_k 는 Giannone and Primiceri (2015)가 제안한 방식에 따라 각각 종속변수(물가상승률 또는 경제성장률)와 설명변수를 AR(4)로 추정했을 때 오차항 표준오차의 OLS 추정치로 산출된다.

(변수집합 필터링)

어떤 변수집합에 변수선택을 적용하는지에 따라서도 예측결과가 크게 달라질 수 있다. 변수집합이 너무 적으면 중요한 변수를 누락시킬 수 있으며, 반대로 변수집합이 너무 크면 불필요한 변수가 다수 포함되어 중요한 변수의 예측력이 비효율적으로 추정되고 예측력이 저하될 수 있다. 따라서 강규호·김정성·신세림 (2021)에서 고안된 다음과 같은 4단계 변수집합 필터링을 적용하여 효율적인 변수집합을 구성할 필요가 있다.

11) 미네소타 사전분포가 부여된 VAR 모형에 대해서는 Dieppe, Legrand, and Roye (2016)를 참조하기 바란다.

Step 1) 전체 변수집합을 대상으로 변수선택 알고리즘을 적용하고 사후 중요도(=γ의 사후평균)가 0.1 이하인 변수들을 변수 집합에서 제외

Step 2) 1단계에서 축소된 변수집합에 다시 변수선택을 실시하고 중요도가 0.2 이하인 변수들을 제외

Step 3) 2단계에서 축소된 변수집합에 다시 변수선택을 실시하고 중요도가 0.3 이하인 변수들을 제외

Step 4) 3단계에서 축소된 변수집합에 다시 변수선택을 실시하고 중요도가 0.4 이하인 변수들을 제외

이제 각 예측시계 별로 $3_{(Spike분산)} \times 4_{(시차)} \times 3_{(표본내 크기)} \times 2_{(단위조정)} \times 4_{(변수집합)} = 288$ 번 추정 후, $3_{(Spike분산)} \times 4_{(시차)} \times 3_{(표본내 크기)} \times 2_{(단위조정)} \times 1_{(변수집합)} = 72$ 가지 사전설정 조합¹²⁾에 대해서 표본외 예측력을 평가한다. 이 때 평가기준은 로그 PPL이며, 2개의 Uni-ADL 모형 각각의 로그 PPL 이 극대화되는 최적의 사전설정 조합을 구한다. Bi-ADL 모형의 경우 2개의 Uni-ADL 모형을 동시에 추정하여 로그 PPL 합계가 가장 큰 경우를 최적 사전설정 조합으로 선정하여 그대로 적용한다. 그리고 Uni-ADL과 Bi-ADL 모형의 하이퍼-파라미터 및 사전설정을 종합하면 각각 <표 1>과 <표 2>와 같다.

<표 1> Uni-ADL 모형의 하이퍼-파라미터 및 사전설정

항목	하이퍼-파라미터 및 사전설정
Spike 분산(b_0)의 사전평균(δ_{00}/α_{00})	$10^{-6}, 10^{-8}, 10^{-10}$ (단, $\alpha_{00}=1,000$)
Slab 분산(b_1)의 사전분포(δ_{01}/α_{01})	0.16 (단, $\alpha_{01}=20$)
오차항 σ^2 의 사전분포	$\alpha_0 = 5, \delta_0 = 0.1$
inclusion probability p 의 사전분포	$a_0 = c_0 = 5$
예측변수의 시차	1, 2, 3, 4분기
표본내 기간	확장, 롤링(윈도우의 크기: 20, 40분기)
설명변수의 단위조정	Z-score, 미네소타 형태
변수집합 필터링	4단계(중요도: 0.1 → 0.2 → 0.3 → 0.4)

<표 2> Bi-ADL 모형의 하이퍼-파라미터¹⁾

항목	하이퍼-파라미터
β 의 사전분포	$\beta_0 = [0 \ 0]', B_0 = I_2$
Ω 의 사전분포	$v_0 = 10, R_0 = (1 \times I_2)^{-1}/v_0$

주: 1) $R_0 = (1 \times I_2)^{-1}/v_0$ 에서 $(1 \times I_2)$ 가 사전평균이다.

12) 필터링이 총 4단계이기 때문에 네 번의 모형 추정이 필요하지만, 최종적으로 사용하게 되는 변수집합은 4단계까지의 필터링을 모두 거친 결과이므로 비교대상이 되는 케이스는 한 가지라는 점에 유의하기 바란다.

III. 추정 결과

III-1. 통계자료

추정에 사용할 자료의 표본기간은 2003년 1/4분기에서 2021년 4/4분기이며 월 단위로 자료가 제공될 경우에는 월별자료를 분기 중 평균을 취하여 변환하였다. 예측의 속보성 등을 고려하면 분기자료 보다는 월별자료를 사용하는 것이 보다 바람직하겠으나 GDP를 포함한 많은 외생변수가 월 단위로는 통계가 제공되지 않기 때문에 분기자료를 이용하여 분석을 수행하였다.

본 연구에서 사용된 자료의 목록은 다음 <표 3>과 같다. 우선 종속변수는 총지수 기준 소비자물가지수와 실질GDP 계절조정계열을 이용하였다. 두 변수 모두 전년동기대비(%)로 변환하여 사용하기 때문에 계절조정계열 대신 원계열을 사용해도 무방하지만 계절조정계열로 추정시 원계열에 비해 결과가 보다 안정적으로 나오는 경향이 있어 계절조정계열이 제공되는 실질GDP는 이를 활용하였다. 외생변수의 경우 변수집합 필터링과 변수선택을 통해 결국은 중요한 변수들이 선택될 것이므로 이론적으로나 경험적으로 물가상승률과 경제성장률을 예측하는 데 도움이 될만한 변수들을 충분히 많이 포함하였다.

<표 3> 변수일람¹⁾

항목	세부기준	변수변환	출처
소비자물가지수(CPI)	총지수	전년동기대비	통계청
국내총생산(GDP)	실질, 계절조정	전년동기대비	한국은행
기대인플레이션(IE)	일반인 기대, 향후 1년	수준	한국은행
유가(Oil)	두바이유 현물, 배럴당 달러	수준	IMF
장단기금리스프레드(Sp)	국고채10년물-1년물	수준	한국은행
미국 물가지수(PCEPI)	총지수	전년동기대비	US BEA
통화량(M2)	M2 말잔, 계절조정	전년동기대비	한국은행
산업생산지수(IP)	전산업, 계절조정	전년동기대비	통계청
주택매매가격지수(HP)	KB부동산, 총지수	전년동기대비	KB부동산
원달러 환율(EX)	종가, 말일기준	전년동기대비	한국은행
주택담보대출 금리(MR)	예금은행 가중평균 신규취급액	수준	한국은행
미국 금리(USR)	T-note 10년물	수준	Fed
경기선행지수(LI)	순환변동치	수준	통계청
미국 GDP(USGDP)	실질, 계절조정	전년동기대비	US BEA
실업률(U)	—	수준	통계청
콜금리(Call)	—	수준	한국은행
고용률(E)	—	수준	통계청
실질임금(W)	제조업 명목임금지수/CPI 공업제품 계절조정	전년동기대비	고용노동부

주: 1) ()내는 각 변수의 기호로서 변수선택 결과 보고서 이용한다.

외생변수 중 국내지표로는 물가상승률과 경제성장률간 상관관계를 고려하여 GDP와 CPI를 상호 간에 교차하여 포함하였으며 기대인플레이션, 유가, 장단기금리스프레드, M2, 전산업생산지수, 주택매매가격, 원달러 환율, 주택담보대출 금리, 경기선행지수, 실업률, 콜금리, 고용률, 실질임금을 공통적으로 포함하였다. 기대인플레이션과 유가, 실질임금은 CPI에, 전산업생산지수와 경기선행지수는 GDP에 직·간접적으로 영향을 줄 수 있는 대표적인 변수이다. 주택매매가격과 주택담보대출 금리는 조병수·이종웅·권인하 (2021)에서 우리나라 주택가격과 가계부채가 물가상승률에 미치는 영향이 유의하게 나타난 점을 고려하여 포함하였다. 이와 관련하여 금융위기 이후 크게 늘어난 유동성이 가계부채 등에 미친 영향을 고려하기 위해 M2를 포함하였고, 장단기 금리스프레드는 채권시장의 물가상승률과 경기에 대한 기대를 반영하므로 포함하였다.

실업률과 고용률의 경우 고용상황이 물가상승률과 경제성장률 모두에 영향을 주고 받을 수 있다는 점을 고려하였으며 두 변수의 변동 방향이 고용상황에 대해 상반되는 신호를 나타내는 경우도 있어 두 변수를 모두 포함하였다. 끝으로 콜금리는 통화정책이 물가와 실물경기에 미치는 영향을 고려하였다. 이 외에 우리나라는 해외시장 의존도가 높고 글로벌 경기의 영향을 많이 받는 점을 고려하여 미국의 물가지수(PCEPI), GDP와 장기금리, 원달러 환율 등 해외요인들도 외생변수로 포함하였다. 동 변수들은 그 성격에 따라 전년동기대비로 변환하거나 수준변수 그대로 사용하였다.

III-2. 표본외 예측 결과

표본외 예측결과를 설명하기에 앞서 분석과정이 다소 복잡한 만큼 이를 한번 더 요약하여 설명하고자 한다. 우선 첫번째 단계에서는 Univariate CPI 및 GDP ADL 모형을 개별적으로 추정하여 물가상승률과 경제성장률을 예측한다. 이 때 실제예측에 앞서 표본외 예측을 실시하는데, 이 과정에서 Spike 분산의 평균(10^{-6} , 10^{-8} , 10^{-10}), 시차(1~4분기), 표본내 기간(확장회귀, 롤링윈도우가 20분기 또는 40분기인 롤링회귀), 설명변수의 단위조정(Z-score 방법, 미네소타 사전분포 형태)이라는 네 가지 범주¹³⁾에 대한 최적 사전설정이 로그 PPL을 극대화하는 경우로 결정(튜닝)되고, 이와 함께 변수선택이 이루어진다. 그리고 표본외 예측에서 결정된 최적 사전설정 세팅이 실제예측시 그대로 적용되고 변수선택의 경우 표본외 예측 결과와 상관없이 실제 예측과정에서 다시 한번 이루어진다.

두번째 단계에서는 Bi-ADL 모형의 추정시 적용될 사전설정과 변수조합을 결정하기 위한 선행과정으로 Univariate CPI 및 GDP ADL 모형을 동시에 추정한다. 이 때 첫번째 단계와 마찬가지로 실제예측에 앞서 표본외 예측이 수행되는데, 표본외 예측 결과 로그 PPL 합계가 극대화되는 경우로 최적 사전설정 세팅이 결정된다. 그리고 이 사전설정 세팅이 실제예측에 그대로 적용되어 실제예측에서 변수선택이 이루어진다. 세번째 단계에서는 두번째 단계에서 결정된 최적 사전설정 세팅(표본외 예측에서 결정)과 변수선택 결과(실제예측에서 결정)를 Bi-ADL 모형에 그대로 적용하여 추정 및 예측을 실시한다. 즉, Bi-ADL 모형에서는 튜닝과 변수선택은 수

13) 앞 절에서는 이해가 용이하도록 변수집합 필터링을 포함하여 다섯 가지 사전설정으로 설명하였으나, 사실 변수집합 필터링의 경우 나머지 네 가지 사전설정과 달리 여러 대안이 존재하는 것이 아니므로 최적 사전설정은 네 가지 범주에서 이루어진다는 것이 보다 정확한 표현이다.

행하지 않는다. 이 과정에 따르면 본 연구의 분석대상 모형은 ① CPI와 GDP간 통계적 상관관계가 고려되지 않은 Univariate CPI/GDP ADL 모형과 ② CPI와 GDP간 통계적 상관관계를 고려한 Bi-ADL 모형이다.

다음 <표 4>는 각 모형별 표본외 예측시 결정된 예측시계 별 최적의 사전설정 세팅과 예측력 평가지표인 로그 PPL 및 RMSE이다. H는 예측시계, P는 예측변수의 시차(1, 2, 3, 4분기), b_0 는 Spike 분산의 사전분포(δ_{00}/α_{00}), WS(Window Size)는 표본내 기간 설정방식(확장회귀, 20분기 또는 40분기 롤링회귀), Unit은 설명변수의 단위조정 방식(Z-score, 미네소타 방식)을 나타낸다. 패널(a)와 (b)는 각각 Univariate CPI ADL 및 GDP ADL 모형에서의 결과로 두 모형을 개별적으로 추정하여 각각에 대한 로그 PPL값이 가장 큰 경우의 결과이다. 패널(c)는 Univariate CPI ADL 및 GDP ADL 모형을 동시에 추정하여 각 모형에서 산출된 로그 PPL 합계가 가장 경우의 최적 사전설정과, 실제예측시 변수선택 결과가 Bi-ADL 모형에 적용된 결과이다. 그리고 패널(c)의 로그 PPL과 RMSE는 Bi-ADL 모형의 추정 및 예측에 따른 결과이다.

먼저 로그 PPL과 RMSE 결과를 보면 모든 모형에서 예측시계가 길어질수록 로그 PPL은 작아지고 RMSE는 커져 예측력이 악화되는 경향이 있는 것으로 나타났다. 이는 뒤에서 소개될 변수선택 결과를 보면 알 수 있듯이 시차가 1인 자기회귀항의 중요도가 여타 변수들에 비해 큰 데 주로 기인한 것으로 판단된다. 아울러 Uni-ADL 모형이 Bi-ADL 모형에 비해 로그 PPL¹⁴⁾은 크고, RMSE는 작아 구간예측과 점예측 모두에서 보다 우수한 것으로 나타났다. 이는 물가상승률과 경제성장률이 통계적인 상관관계가 약할 가능성을 나타내는데 이에 대해서는 뒤에서 살펴볼 등고선(Contour) 그래프를 통해 다시 설명하고자 한다. 그리고 GDP에 대한 예측력이 CPI에 비해 다소 떨어지는 것으로 나타나는데 이는 코로나의 영향으로 GDP 상승률이 급격히 하락하였다가 상승하는 구간이 존재하는 데 주로 기인한 것으로 판단된다. 다만 이러한 점들에도 불구하고 <그림 1~4>와 <부록 B>의 표본외 예측결과 그래프를 살펴보면 전반적으로 예측치가 실현치를 비교적 잘 fitting하는 것으로 보인다.

<부록 B>를 보면 표본외 예측기간 20분기 중 10~20%를 제외하고는 대체로 90% 신용구간 안에 실현치가 포함되어 있으며, <그림 1~4>를 보면 CPI의 경우 대체로 2020년 이후의 상승세를 잘 잡아내고 있다. GDP의 경우에도 코로나 발생 이후 급락시기를 제외하고 보면 예측치가 실현치의 흐름을 비교적 잘 따라가고 있는 것으로 보인다.

다음으로 최적 사전설정 결과를 살펴보면 시차(P)는 2~3이, b_0 는 10^{-10} 이, 표본내 기간(WS)은 주로 확장회귀(Recursive)가 선택되었으며, 단위조정은 Z-score와 미네소타 방식이 골고루 선택되었다. 표본내 기간으로 주로 확장회귀가 선택된 것은 코로나 발생 이후 물가상승률의 최근 레짐(regime)이 저인플레이션 시기가 도래한 2012년 이후보다는 2012년 이전의 레짐에 보다 가깝기 때문인 것으로 추정된다.

14) Uni-ADL 모형에서는 CPI ADL과 GDP ADL에서 산출된 로그 PPL을 합계하고 Bi-ADL 모형에서는 다변량 정규 분포로부터 산출된 Joint 로그 PPL을 사용하여 두 값을 비교함으로써 두 모형의 예측력 우위를 판단할 수 있다.

<표 4> 모형별 표본의 예측력 및 최적 사전설정

(a) Best Univariate CPI ADL 모형

H	lnPPL	RMSE	P	b ₀	WS	Unit
1	-19.00	0.61	2	10 ⁻¹⁰	3	1
2	-24.63	0.83	2	10 ⁻¹⁰	3	2
3	-28.44	0.99	3	10 ⁻¹⁰	3	2
4	-29.41	0.90	2	10 ⁻¹⁰	1	2
5	-25.62	0.69	2	10 ⁻⁶	2	2
6	-24.65	0.80	1	10 ⁻⁸	1	2
7	-30.07	1.09	3	10 ⁻¹⁰	3	2
8	-31.11	1.09	2	10 ⁻¹⁰	3	2
평균	-26.62	0.88	2.13	—	—	—

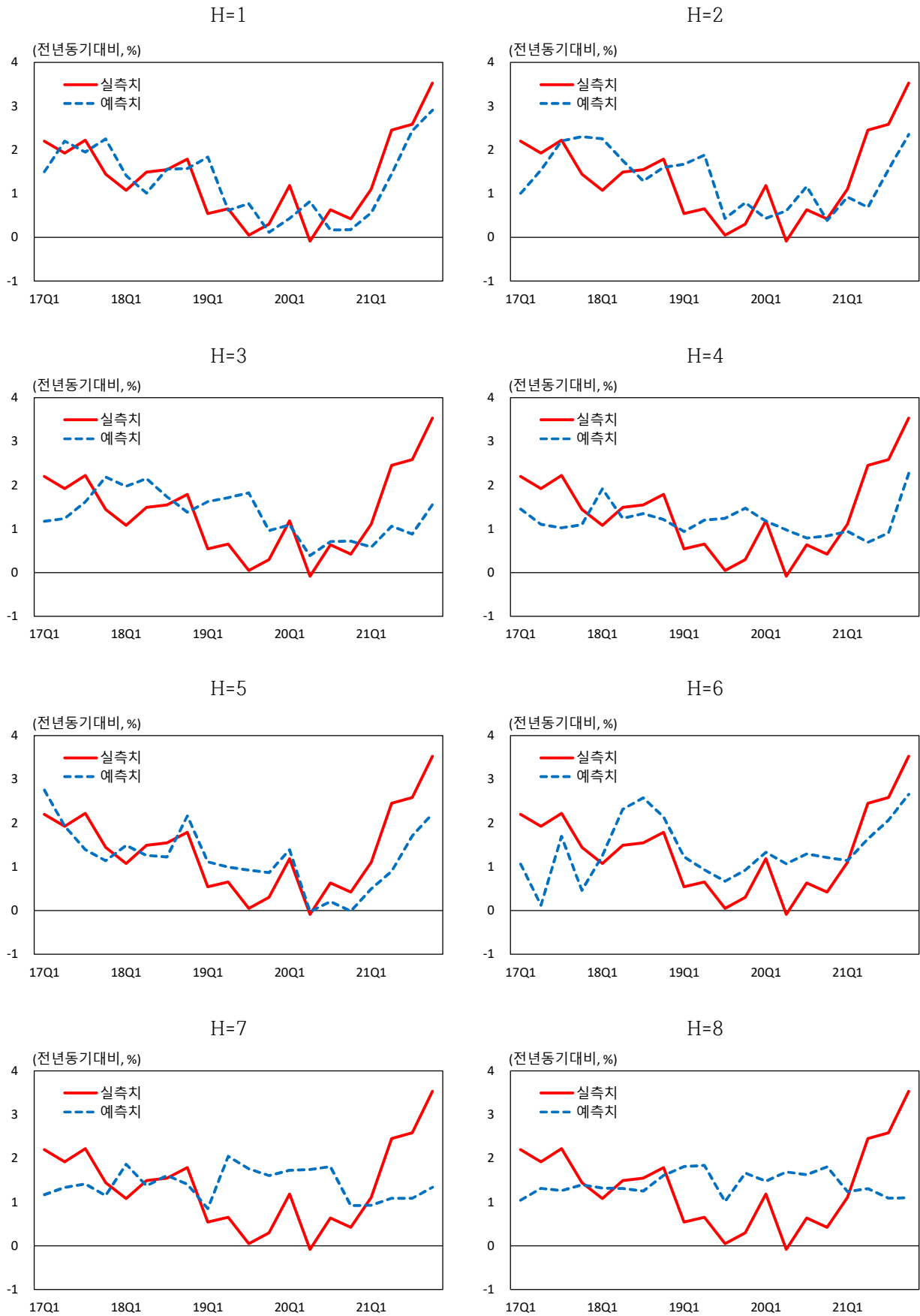
(b) Best Univariate GDP ADL 모형

H	lnPPL	RMSE	P	b ₀	WS	Unit
1	-34.61	1.35	3	10 ⁻¹⁰	3	1
2	-41.22	1.88	1	10 ⁻¹⁰	3	1
3	-41.37	1.79	3	10 ⁻¹⁰	3	2
4	-42.78	1.93	3	10 ⁻¹⁰	3	2
5	-42.99	1.95	4	10 ⁻¹⁰	3	1
6	-42.04	1.87	1	10 ⁻¹⁰	3	2
7	-41.78	1.80	2	10 ⁻¹⁰	3	1
8	-42.69	1.94	1	10 ⁻¹⁰	3	1
평균	-41.19	1.81	2.25	—	—	—

(c) Bi-ADL 모형

H	Joint lnPPL	(a), (b)의 lnPPL 합계	RMSE		P	b ₀	WS	Unit
			CPI ADL	GDP ADL				
1	-57.08	-53.61	0.70	1.56	4	10 ⁻¹⁰	3	2
2	-63.03	-65.84	0.80	1.85	3	10 ⁻¹⁰	3	2
3	-78.22	-69.82	1.04	2.45	2	10 ⁻¹⁰	3	1
4	-76.39	-72.20	1.18	2.35	1	10 ⁻¹⁰	3	2
5	-77.13	-68.61	1.08	2.55	3	10 ⁻¹⁰	3	1
6	-77.58	-66.70	1.15	2.41	4	10 ⁻¹⁰	3	1
7	-77.18	-71.85	1.16	2.32	3	10 ⁻¹⁰	3	2
8	-77.85	-73.80	1.20	2.28	4	10 ⁻¹⁰	3	1
평균	-73.06	-67.80	1.04	2.22	3	—	—	—

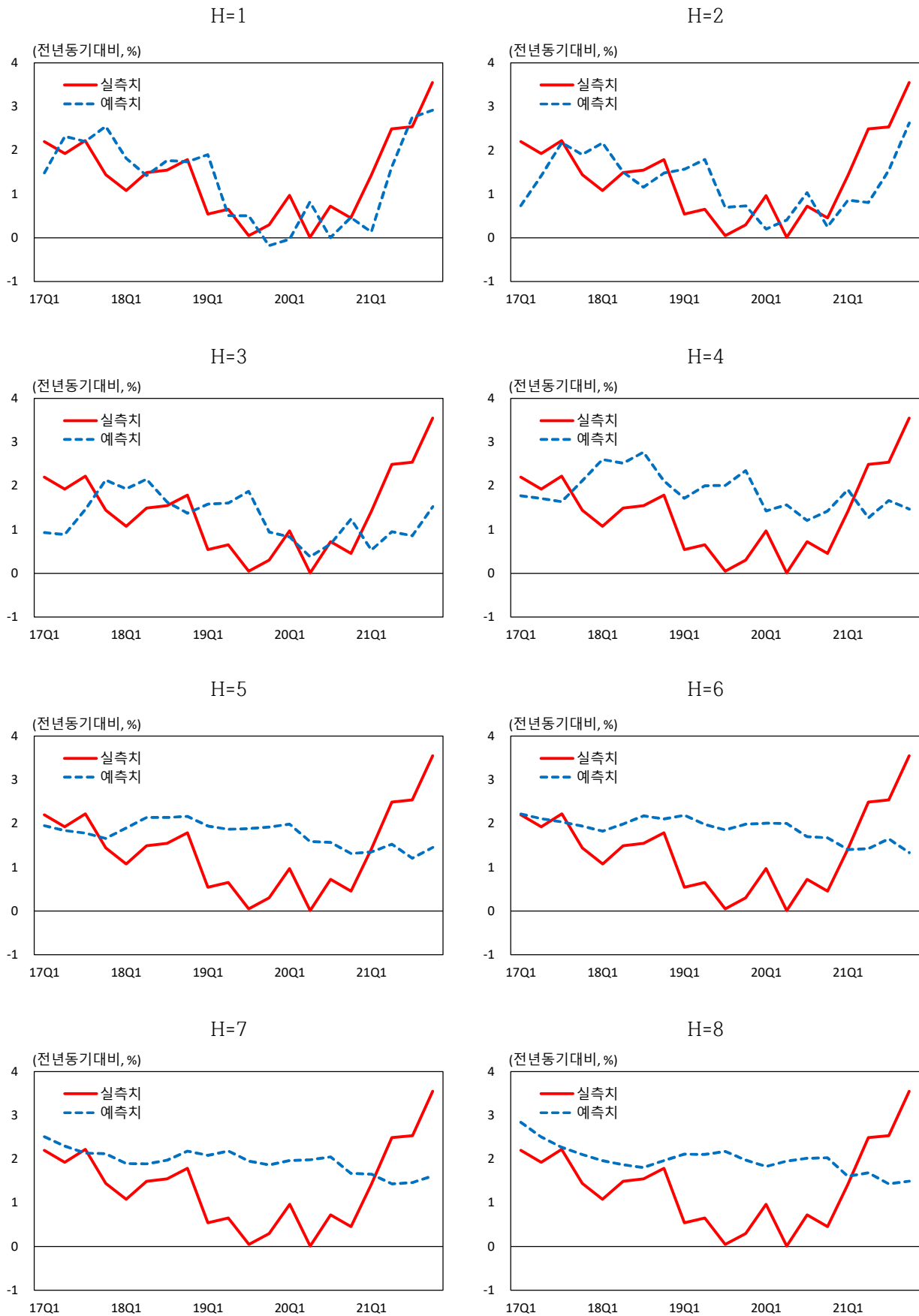
<그림 1> 표본의 예측결과: Uni-ADL(CPI)



<그림 2> 표본외 예측결과: Uni-ADL(GDP)

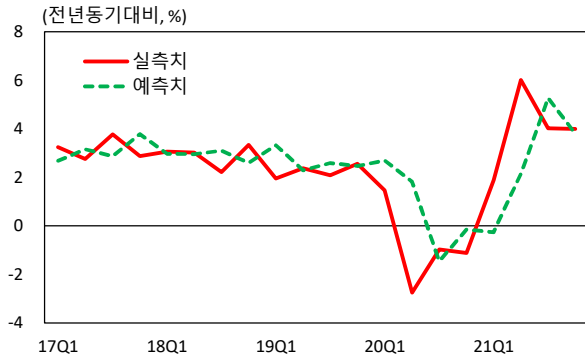


<그림 3> 표본외 예측결과: Bi-ADL(CPI)

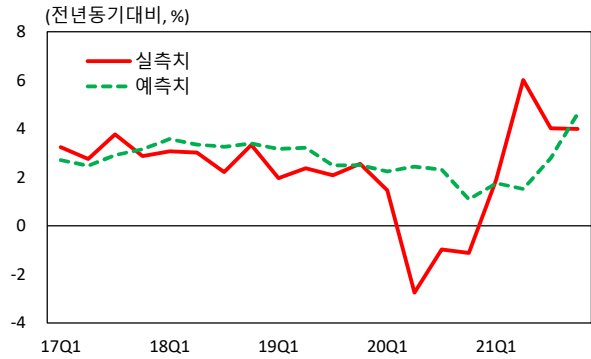


<그림 4> 표본외 예측결과: Bi-ADL(GDP)

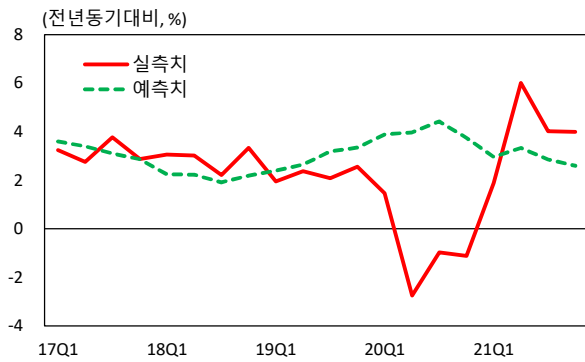
H=1



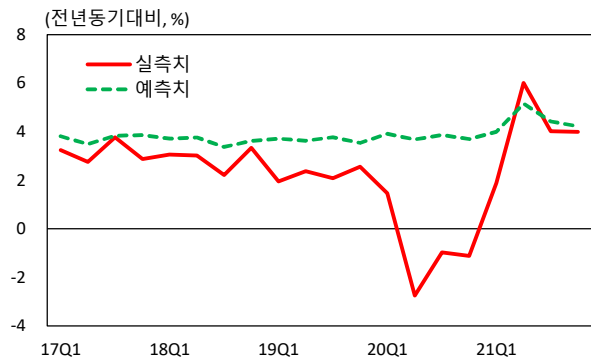
H=2



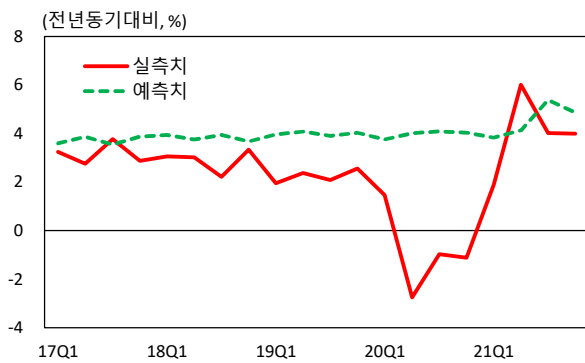
H=3



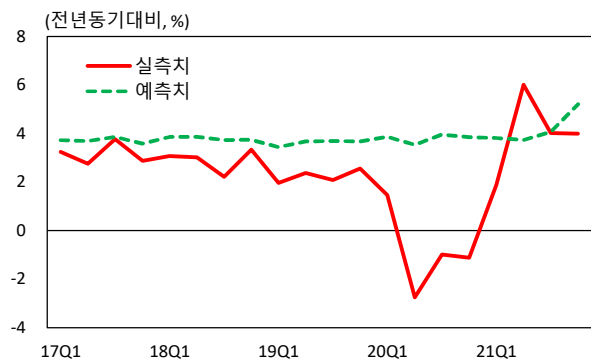
H=4



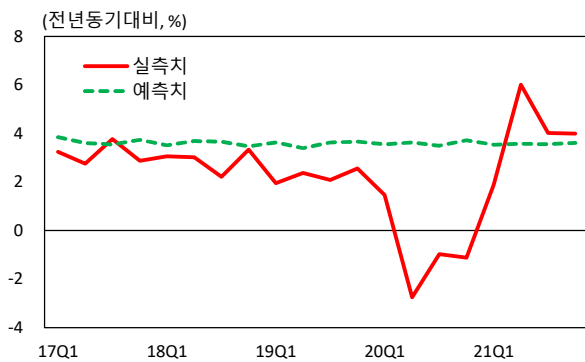
H=5



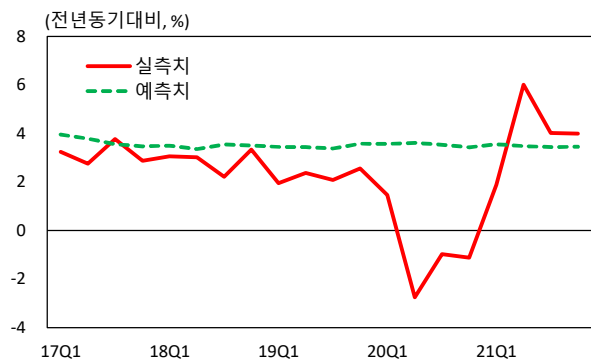
H=6



H=7



H=8



다음으로 표본외 예측시 변수선택 결과는 다음 <표 5>와 같다.¹⁵⁾ 이를 살펴보면 크게 2가지 정도의 특징적인 점이 드러난다. 우선 첫번째는 Univariate CPI/GDP ADL 모형에서 공통적으로 상수항과 첫번째 자기회귀 변수(CPI(-1), GDP(-1))가 모든 예측시계에서 중요변수로 포함되어 있다는 것이다. 이로 인해 직접예측의 특성상 예측시계가 길어질수록 예측치의 실현치 fitting이 약해지는 경향이 있는 것으로 판단된다. 다음으로 두번째 특징은 CPI ADL 모형과 GDP ADL 모형 모두 주택시장 요인과 해외요인 관련 변수가 다수 포함되어 있다는 점이다. 주택시장 요인과 관련한 변수로는 주택가격(HP)과 주택담보대출 금리(MR)가 있으며 해외요인 관련 변수로는 환율(EX)과 미국시장 지표(PCEPI, USGDP, USR) 및 유가(Oil)가 포함되어 있다. 이러한 결과는 동 연구의 예측기간(2017~2021년) 중 국내 주택시장과 해외경제 여건 등에 따라 우리나라 물가상승률과 경제성장률의 움직임이 크게 영향을 받았으며 예측에 있어서도 두 시장의 여건이 중요함을 시사한다. 이 외에도 고용관련 지표(E, U, W), 유동성(M2), 산업생산지수(IP) 등이 CPI 및 GDP ADL 모형에 공통적으로 중요한 예측변수인 것으로 나타났다.

<표 5> 표본외 예측 변수선택 결과¹⁾

H	Uni-ADL	
	CPI	GDP
1	상수항, CPI(-1), EX(-1), E(-1)	상수항, GDP(-1, -3), CPI(-3), E(-1)
2	상수항, CPI(-1)	상수항, GDP(-1), CPI(-1), MR(-1), USGDP(-1), U(-1)
3	상수항, CPI(-1)	상수항, GDP(-1), M2(-2), MR(-1), LI(-2), E(-3)
4	상수항, CPI(-1), USGDP(-2)	상수항, GDP(-1), Oil(-1, -2), M2(-3), LI(-1)
5	상수항, CPI(-1 to -4), GDP(-1, -2), IE(-2), Oil(-1, -2), Sp(-1, -2), PCEPI(-1, -2), M2(-1, -2), IP(-1, -2), HP(-1, -2), EX(-1, -2), USR(-1, -2), LI(-1, -2), U(-1, -2), Call(-1, -2), E(-1)	상수항, GDP(-1, -3), IE(-1), IP(-1, -2), HP(-4), MR(-3)
6	상수항, CPI(-1), GDP(-1), IE(-1), Oil(-1), Sp(-1), PCEPI(-1), M2(-1), IP(-1), HP(-1), EX(-1), MR(-1), LI(-1), U(-1)	상수항, GDP(-1), PCEPI(-1), IP(-1)
7	상수항, CPI(-1), EX(-2), MR(-1)	상수항, GDP(-1, -3), PCEPI(-2), HP(-2), MR(-1, -2), USR(-1), Call(-1), W(-1)
8	상수항, CPI(-1)	상수항, GDP(-1 to -3), Sp(-1), PCEPI(-1), M2(-1), IP(-1), EX(-1), MR(-1), U(-1)

주: 1) CPI 모형, GDP 모형 각각에 대하여 표본외 예측기간 중 평균 중요도(=γ의 사후평균을 표본외 예측기간 중 단순평균)가 0.1 이상인 변수들을 기재하였다.

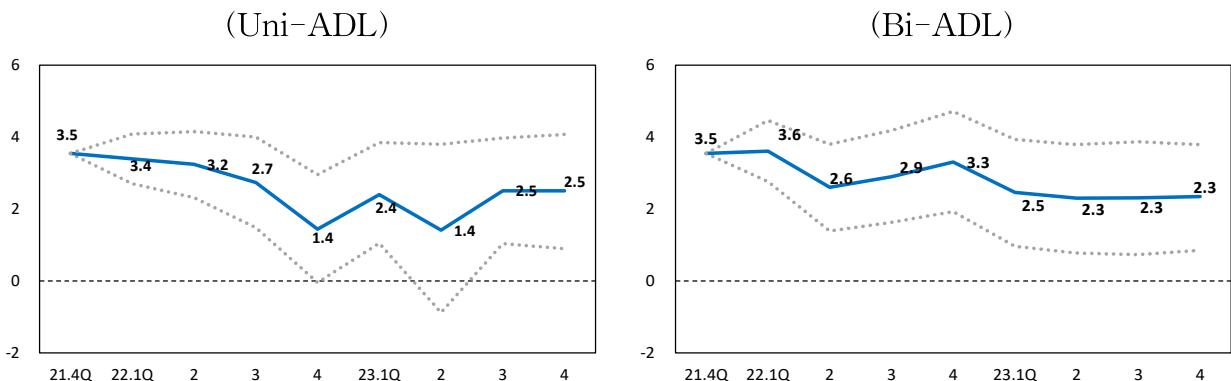
15) Bi-ADL 모형에서는 변수선택이 수행되지 않았고, Univariate CPI와 GDP ADL 모형을 동시에 추정할 경우는 Bi-ADL 모형에 적용하기 위해 로그 PPL 합계로 최적 기준만 다르게 하여 변수선택이 이루어진 것이다. 따라서 Univariate CPI와 GDP ADL 모형을 개별적으로 추정한 것과 본질적으로 다르지 않으므로 변수선택 결과는 기재하지 않는다.

III-3. 실제예측 결과

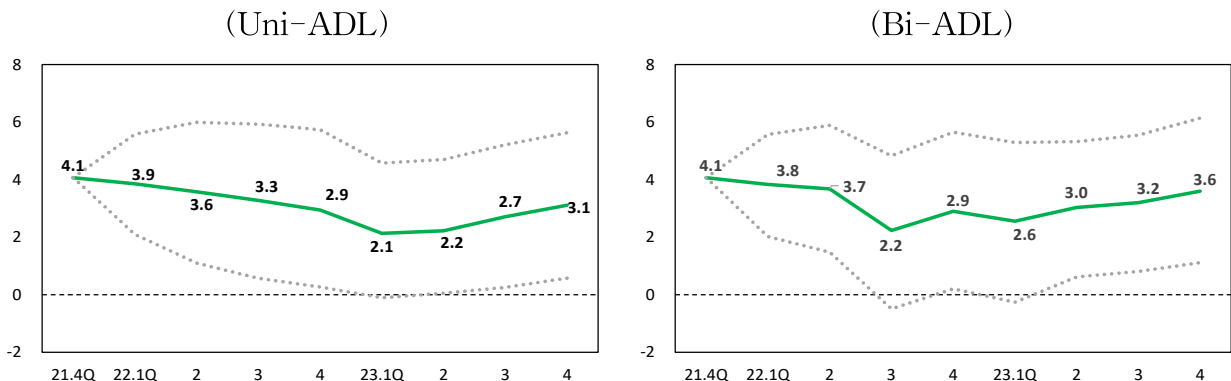
<그림 5~6>은 2021년 4분기 시점에서 물가상승률과 경제성장률에 대한 향후 2년(8분기) 동안의 실제예측 결과이다. 실선은 물가상승률과 경제성장률의 사후예측분포로부터 계산된 사후 평균이며 점선은 90% 신용구간이다. 우선 물가상승률은 일부 상승구간이 존재하기는 하지만 Uni-ADL과 Bi-ADL 모형 모두 추세적으로 하락세가 이어져 2023년 4분기에는 물가목표 수준에 근접하는 것으로 나타났다. 그리고 경제성장률의 경우 Uni-ADL 모형과 Bi-ADL 모형 모두 2023년 1분기까지는 하락세가 이어지다가 이후 상승하여 3%대를 기록하는 것으로 나타났다. 이러한 결과로 볼 때 최근의 높은 물가상승률이 당분간은 지속되겠으나 구조적인 현상은 아닌 것으로 추정되며 경제성장률도 점차 회복될 것으로 예상된다.

다만 신용구간을 살펴보면 물가상승률과 경제성장률 모두 예측에 적지 않은 불확실성이 존재하는 것으로 나타났다. 먼저 물가상승률의 신용구간 폭은 2~3%p 수준으로 예측분포가 0%대에서 5%에 근접한 수준까지 퍼져있어 중앙은행의 중기적인 물가목표 달성이 어려워질 수 있음을 보이고 있다. 경제성장률의 경우에는 신용구간의 폭이 4~6%p 수준으로 물가상승률에 비해 불확실성이 훨씬 큰 것으로 나타났다. 이에 따라 2022년 3분기~2023년 1분기 중에는 경제성장률이 아주 낮은 확률로 마이너스를 기록할 가능성도 있는 것으로 예측되었다.

<그림 5> CPI 실제예측 결과



<그림 6> GDP 실제예측 결과



다음으로 실제예측시 변수선택 결과를 보면 표본외 예측과 마찬가지로 모든 모형과 예측시계에서 시차가 1인 자기회귀 변수가 포함되어 있는 것으로 나타났다. 외생변수의 경우 Univariate CPI ADL 모형은 콜금리(Call)가 3회, 유가(Oil)와 미국 물가(PCEPI)가 2회씩 포함되었고 이외에도 원달러 환율(EX), 미국 금리(USR), 유동성(M2), 주택가격(HP), 고용률(E), 실업률(U), 실질임금(W), 장단기금리스프레드(Sp), 경기선행지수(LI), 산업생산지수(IP)가 1회씩 선택되었다. Univariate GDP ADL 모형은 장단기금리스프레드(Sp)가 2회 선택된 가운데 원달러 환율(EX), 유동성(M2), 고용률(E), 실업률(U), 콜금리(Call), 산업생산지수(IP)가 1회씩 예측시계에서 선택되었다. 그리고 Bi-ADL 모형은 CPI 식의 경우 미국 GDP(USGDP)가 2회, 그 외 경기선행지수(LI), 실업률(U), 기대인플레이션(IE)이 1회씩 선택되었고, GDP 식에서는 미국 GDP(USGDP), 유동성(M2), 경기선행지수(LI), 장단기금리스프레드(Sp)가 1회씩 선택되었다. 전반적으로 볼 때 외생변수 중에서는 해외요인 변수(환율, 미국관련 지표 등)와 콜금리, M2, 장단기금리스프레드 등이 물가상승률과 경제성장률 예측에 중요한 변수인 것으로 판단된다.

<표 6> 실제예측 시 선택된 변수집합

H	Uni-ADL 모형	
	CPI	GDP
1	상수항, CPI(-1)	상수항, GDP(-1)
2	상수항, CPI(-1, -3)	상수항, GDP(-1)
3	상수항, CPI(-1), Call(-2)	상수항, GDP(-1), E(-3)
4	상수항, CPI(-1)	상수항, GDP(-1), M2(-1)
5	CPI(-1, -3), Oil(-1), PCEPI(-1 to -2), M2(-1), EX(-1), USR(-1 to -2), E(-1), W(-1), Sp(-2), LI(-2), U(-2), Call(-2)	상수항, GDP(-1, -4), EX(-3), Call(-3)
6	CPI(-1), Oil(-1), PCEPI(-1), IP(-1), HP(-1)	상수항, GDP(-1), Sp(-1), IP(-1)
7	상수항, CPI(-1)	상수항, GDP(-1, -4)
8	상수항, CPI(-1), Call(-1)	상수항, GDP(-1), Sp(-1), U(-1)

H	Bi-ADL 모형	
	CPI	GDP
1	상수항, CPI(-1), LI(-2), USGDP(-3)	상수항, GDP(-1)
2	상수항, CPI(-1), USGDP(-3)	상수항, GDP(-1), Sp(-2)
3	상수항, CPI(-1), U(-2)	상수항, GDP(-1), LI(-2)
4	상수항, CPI(-1), IE(-1)	상수항, GDP(-1)
5	상수항, CPI(-1)	상수항, GDP(-1 to -2), M2(-2)
6	상수항, CPI(-1)	상수항, GDP(-1)
7	상수항, CPI(-1)	상수항, GDP(-1)
8	상수항, CPI(-1)	상수항, GDP(-1), USGDP(-1)

다음으로 Uni-ADL 및 Bi-ADL 모형에서 도출된 물가상승률과 경제성장률에 대한 결합 사후예측분포를 <그림 7>, <그림 8>과 같은 등고선(Contour) 그래프로 그려봄으로써 두 변수 간 상관관계와 함께 향후 경제상황을 스태그플레이션(stagflation), 경기호황(boom), 불황(recession) 및 안정기(moderation)로 구분하여 각 상태별 발생확률을 추정해 볼 수 있다. 동 등고선 그래프는 실제예측 결과 산출된 물가상승률과 경제성장률의 예측분포로부터 각각의 사후평균과 분산-공분산 행렬(Ω)의 사후평균을 구하고, 이를 이용하여 다변량 정규분포의 확률 밀도를 계산하여 그릴 수 있다. 이 때 물가상승률의 경우 물가안정목표 수준인 2%를, 경제성장률은 지난 10년(2012년 1분기~2021년 4분기)간 실질GDP 성장률의 평균인 2.6%¹⁶⁾를 기준선으로 삼는다. 이 두 기준선을 통해 구분된 사분면은 1사분면부터 4사분면까지 차례대로 경기호황(boom), 스태그플레이션(stagflation), 불황(recession) 및 안정기(moderation)를 나타낸다.

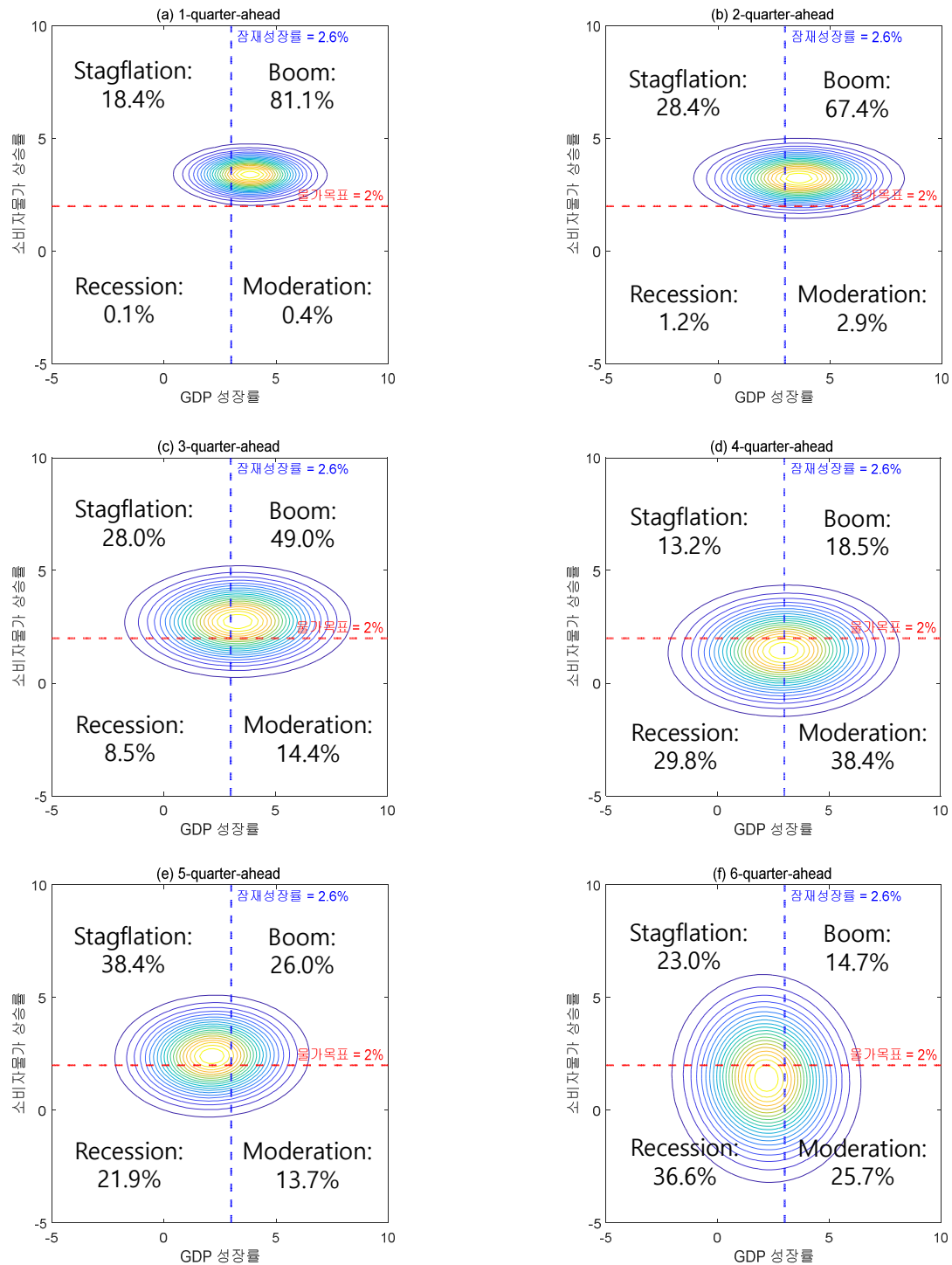
예측시계 별로 각 상태별 확률을 살펴보면 Uni-ADL 모형의 경우 2021년 4분기를 기점으로 스태그플레이션 확률이 점차 높아져 5분기 이후 시점인 2023년 1분기에는 38.4%에 이른다. 이후 다소 하락하기는 하지만 7분기 이후 시점인 2023년 3분기까지 32.3%로 높은 수준을 유지하다가, 그 이후에는 20%대로 하락하여 대신 경기호황 확률이 높아지는 모습을 나타냈다. Bi-ADL 모형의 경우에도 Uni-ADL 모형과 유사한 모습을 나타냈다. 먼저 2021년 4분기를 기점으로 스태그플레이션 확률이 점차 높아지다가 3분기 이후 시점인 2022년 3분기에 발생확률 43.9%에 이르고, 이후 시점부터 하락하기 시작하여 8분기 이후 시점인 2023년 4분기에는 17.4%로 스태그플레이션 발생확률이 크게 낮아지는 것으로 추정되었다. 이와 같은 결과로 봤을 때 향후 1년에서 1년반 동안은 일시적으로 스태그플레이션이 발생할 확률이 상당한 것으로 추정된다.

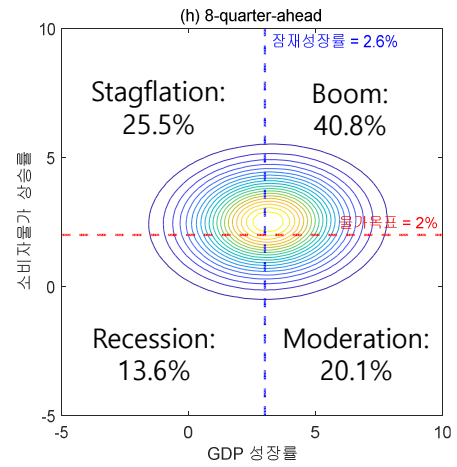
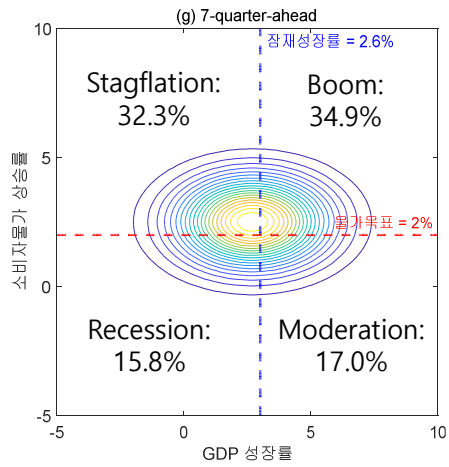
다음으로 등고선의 형태를 통해 두 변수 간 상관관계 여부를 가늠해볼 수 있다. 먼저 Uni-ADL 모형의 경우 모든 예측시계에서 두 변수 간 상관관계가 매우 약한 것으로 나타났다. 반면 Bi-ADL 모형은 예측시계 1~6분기 이후까지는 완만한 양(+)의 기울기를 보이다가 7~8분기 이후에는 기울기가 0에 가까워진다. 실제로 Bi-ADL 모형의 물가상승률과 경제성장률 예측분포 간 상관계수를 살펴보면 등고선에서 나타난 바와 같이 예측시계 1~6분기 이후까지는 상관계수가 0.3~0.4 내외 수준이다가 7분기 이후로는 0에 가까워지는 것으로 나타났다.

한편 Uni-ADL 모형과 Bi-ADL 모형 간 표본외 예측력 평가에서 Uni-ADL 모형이 우월했다는 사실로부터 물가상승률과 실질성장률 간 단기적 상관관계는 충격의 상관관계가 아니라 예측변수 간 상관관계로부터 비롯된 것으로 봐야 한다. 또한 Bi-ADL 모형의 경우 예측시계가 길어질수록 물가상승률과 경제성장률 간 관계가 약해지므로 Uni-ADL 모형이 Bi-ADL 모형에 비해 장기예측에 보다 적합한 모형일 수 있음을 시사한다. 이러한 결과는 코로나 위기를 전후로 하여 공급병목 현상 등으로 인해 총수요 충격과 총공급 충격 간 상관관계에 구조적 변화가 발생하였음에도 불구하고, Bi-ADL 모형이 이러한 구조변화를 반영하지 못하고 있어 예측력이 저하된 것에 주로 기인한 것으로 판단된다.

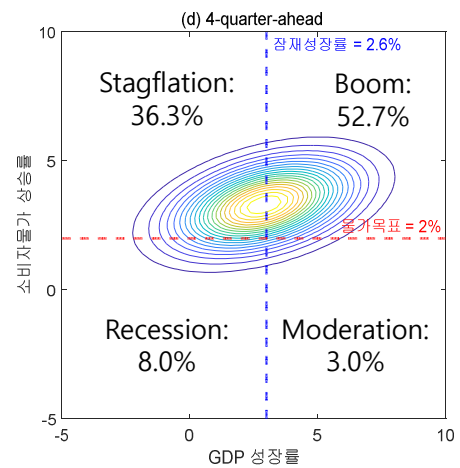
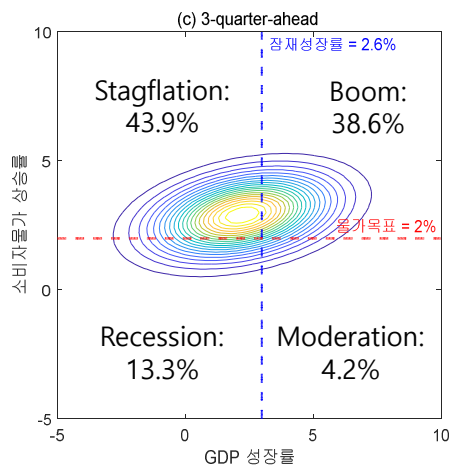
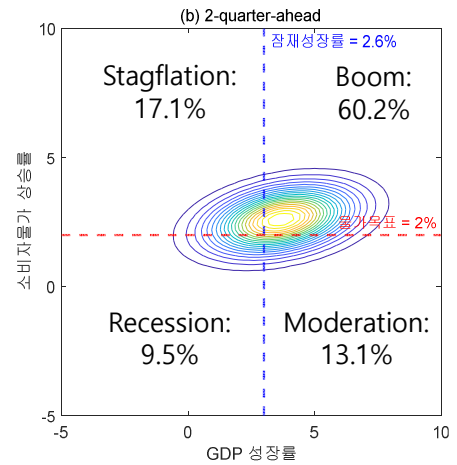
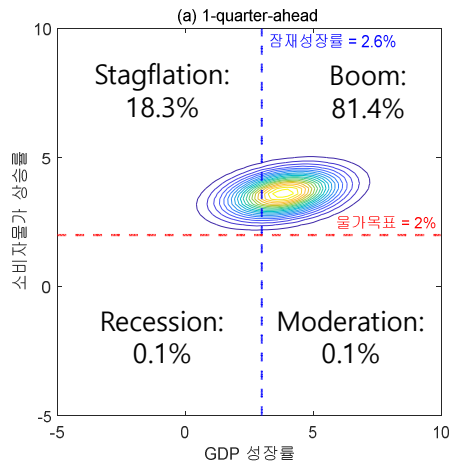
16) 참고로 우리나라 실질GDP의 코로나 발생 이전 최근 5년간(2015.1/4분기~2019.4/4분기) 평균 성장률은 2.8%이며, 권지호·김도완·지정구·김건·노경서 (2019)의 우리나라 잠재성장률 추정치는 2.5~2.9%이다.

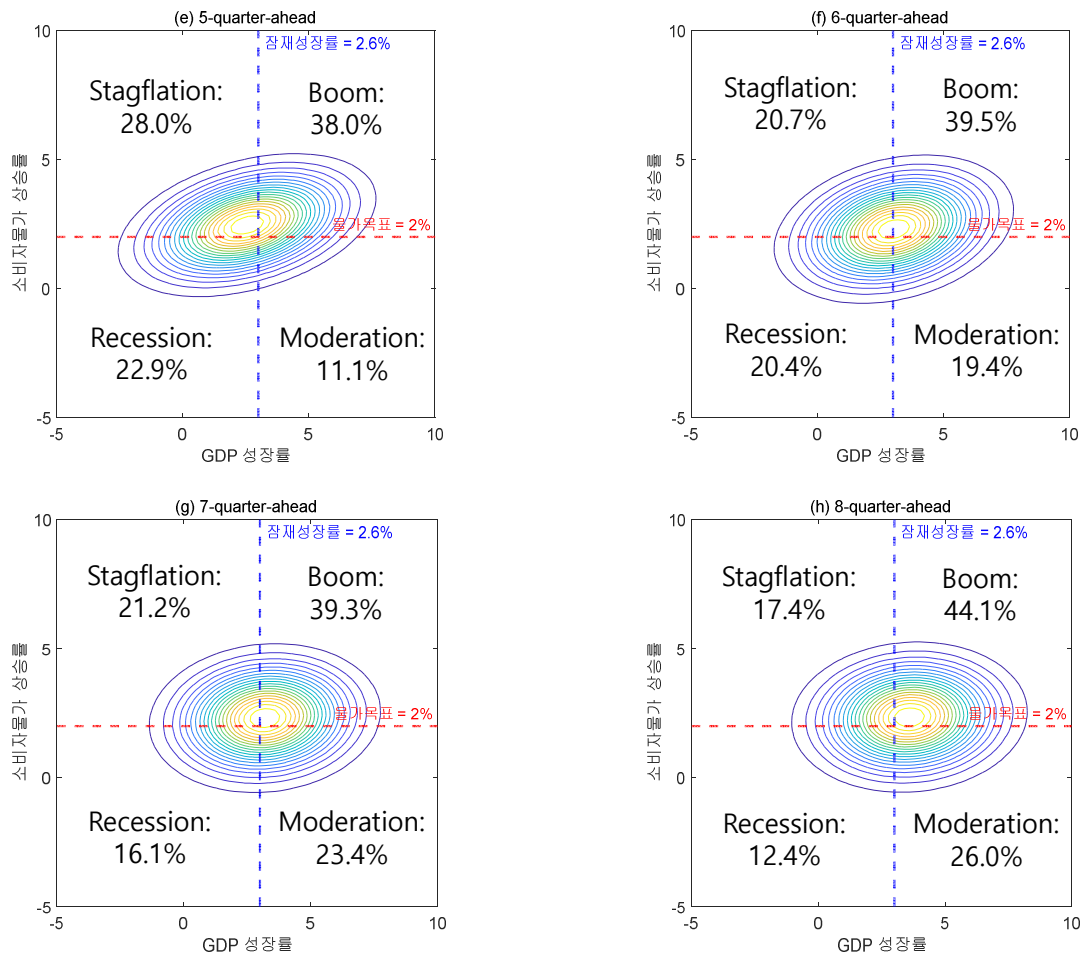
<그림 7> Uni-ADL 모형의 CPI·GDP 결합 사후예측분포





<그림 8> Bi-ADL 모형의 CPI·GDP 결합 사후예측분포





<표 7> 예측시계 별 CPI·GDP 예측분포간 상관계수

	H=1	2	3	4	5	6	7	8
Uni-ADL	-0.01	0.001	0.01	0.03	0.03	-0.04	0.002	0.03
Bi-ADL	0.32	0.33	0.33	0.41	0.37	0.26	0.09	0.07

IV. 결론 및 정책적 시사점

IV-1. 요약

본 연구는 Uni-ADL 모형과 Bi-ADL 모형에 변수선택 및 엄밀한 튜닝과정을 적용하여 우리나라 물가상승률과 경제성장률을 결합예측하였다. 두 모형을 통한 표본외 예측과 실제예측 결과, 물가상승률과 경제성장률 간 통계적 상관관계가 강하지 않아 두 변수 간 조건부 상관관계를 고려한 Bi-ADL 모형보다는 두 변수를 독립적으로 예측하는 Uni-ADL 모형이 보다 유용한 것으로 나타났다. Bi-ADL 모형에서 두 변수의 사후예측분포를 통해 파악한 상관계수는 예측시계 6분기 이후까지는 0.3~0.4 내외로 나타났으나 7분기 이후로는 0에 근접하였다. 이는 단기

적 예측시계에서 예측변수 간 상관관계로 인해 물가상승률과 경제성장률이 동행성을 띠는 것과 예측시계가 길어질수록 Uni-ADL 모형을 이용한 예측이 보다 우월해진다는 것을 의미한다. 실제 분석결과도 표본외 예측시 로그 PPL과 RMSE 모두 예측시계가 2인 경우를 제외하고는 모든 예측시계에서 Uni-ADL 모형의 예측력이 보다 우월한 결과를 나타내었다. 다만 표본외 예측 결과를 보았을 때 Uni-ADL 모형과 Bi-ADL 모형 모두 실현치가 10~20%를 제외하고는 90% 신용구간 내에 포함되어 있어 예측력이 우수한 것으로 나타났다. 이는 베이지안 변수선택기법과 함께 사전설정에 대하여 엄밀한 튜닝과정을 적용하는 것이 예측력을 높이는 데 매우 중요함을 시사하는 결과라 판단된다.

한편 표본외 예측에서의 변수선택 결과, 상수항과 시차가 1인 자기회귀항이 모든 예측시계에 걸쳐 물가상승률과 경제성장률 예측에 결정적인 역할을 하는 것으로 나타났다. 아울러 외생변수 중에서는 주택가격, 주택담보대출 금리 등 주택시장 요인 변수와 환율, 미국의 물가지수, GDP, 금리 등 해외요인 관련 변수가 일부 예측시계에서 중요한 변수인 것으로 분석되었다. 이는 그간 우리나라의 주택담보대출 급증과 주택가격 상승세, 높은 해외 의존도 등을 반영하는 결과라 판단된다.

표본외 예측에서 결정된 최적의 사전설정을 적용하여 변수선택 기법이 도입된 실제예측 결과, 물가상승률은 추세적으로 하락세가 이어져 2023년 4분기에는 물가목표 수준에 근접하는 것으로 나타났다. 그리고 경제성장률의 경우 2023년 1분기까지는 하락세가 이어지다가 이후 상승하여 3%대를 기록하는 것으로 나타났다. 이러한 결과로 볼 때 최근의 높은 물가상승률이 당분간은 지속되겠으나 구조적인 현상은 아닌 것으로 추정되며 경제성장률도 점차 회복될 것으로 예상된다. 다만 신용구간을 살펴보면 물가상승률의 신용구간 폭은 2~3%p, 경제성장률은 4~6%p 수준으로 모두 예측에 적지 않은 불확실성이 존재하는 것으로 나타났다. 이러한 결과는 중기적으로 물가안정목표 달성이 어려워 질 수 있음은 물론이고, 자칫하면 스태그플레이션이 발생할 수도 있음을 의미한다.

실제로 Uni-ADL 모형과 Bi-ADL 모형에서 도출된 물가상승률과 경제성장률에 대한 결합사후예측분포를 이용하여 등고선(Contour) 그래프를 그려본 결과, 향후 1년에서 1년반 동안은 스태그플레이션이 나타날 확률이 상당한 것으로 추정되었다. 먼저 Uni-ADL 모형의 경우 2021년 4분기를 기점으로 스태그플레이션 확률이 점차 높아져 5분기 이후 시점인 2023년 1분기에는 38.4%에 이른다. 이후 다소 하락하기는 하지만 7분기 이후 시점인 2023년 3분기까지 32.3%로 높은 수준을 유지하다가, 그 이후에는 20%대로 하락하여 대신 경기호황 확률이 높아지는 모습을 나타냈다. Bi-ADL 모형의 경우에도 Uni-ADL 모형과 유사한 모습을 나타냈다. 먼저 2021년 4분기를 기점으로 스태그플레이션 확률이 점차 높아지다가 3분기 이후 시점인 2022년 3분기에 발생확률 43.9%에 이르고, 이후 시점부터 하락하기 시작하여 8분기 이후 시점인 2023년 4분기에는 17.4%로 스태그플레이션 발생확률이 크게 낮아지는 것으로 추정되었다.

IV-2. 정책적 시사점

본 연구의 주안점 중 하나는 2021년 이후 지속되고 있는 글로벌 인플레이션이 일시적 현상인지 구조적 현상인지 판단해보고, 향후 스태그플레이션으로 발전할 가능성이 있는지 등을 예

측해 보는 데 있다. 분석결과 현재의 높은 물가상승률이 중장기적으로 지속되지는 않을 것이고 경제성장률도 잠재성장률 수준으로 점차 회복될 것으로 예측되었으나, 예측의 불확실성이 매우 크다는 점에 유의할 필요가 있다. 이는 향후 물가와 경제성장 경로가 우리가 예측한 바와 상당히 다르게 움직일 가능성이 낮지 않음을 의미하기 때문이다. 이처럼 불확실한 경제상황 하에서 경제주체들은 일반적으로 소비와 투자를 지연시킨다.

또한 예상치 못한 충격이 발생했을 경우 거시경제에 여러 부정적인 영향이 발생할 수 있다. 예를 들어 갑작스럽게 물가상승률이 높아지면 한국은행에서는 금리 인상 등을 통해 물가상승을 억제하려 할 것이다. 그러나 현재와 같이 가계부채가 상당히 누적되어 있는 가운데 글로벌 금융위기 이후 장기간 지속된 저금리 여건에 익숙해져 있는 경제주체 입장에서 큰 폭의 금리 인상은 이자 부담 등 큰 충격으로 다가오게 될 것이다. 이는 현금흐름 경로(Cash flow channel)를 통해 소비를 크게 위축시킬 수 있으며 주택수요 급감으로 주택가격이 크게 하락할 경우 음(-)의 부의 효과 경로(wealth effect channel)를 통해 소비 위축이 더욱 확대될 수 있다. 그리고 이로 인해 경기부진이 심화될 경우 스태그플레이션으로 발전하게 될 가능성도 있다. 따라서 한국은행 등 정책당국은 민간 경제주체가 가지고 있는 불확실성을 최소화하기 위해 노력할 필요가 있다.

예상치 못한 충격으로 인해 발생하는 불확실성 자체는 막을 수 없겠으나 대신 민간 경제주체에 대한 시의적절한 정보제공을 통해 불확실성과 충격의 영향을 최소화할 필요가 있다. 이를 위해서는 예측시계 별로 물가상승률과 경제성장률을 예측하는 데 있어 중요한 변수가 상이하다는 점에 유의하여 전망을 수행할 필요가 있다. 아울러 본 연구의 분석결과 국내 주택시장과 해외요인 관련 변수가 물가상승률과 경제성장률 예측시 특히 중요한 변수인 것으로 나타난 만큼 이들에 대한 각별한 모니터링이 필요할 것으로 판단된다.

IV-3. 한계 및 추가 연구방향

본 연구에서 물가상승률과 경제성장률 예측을 위해 이용한 모형은 몇 가지 한계점을 지니고 있다. 첫째, 익히 알려져 있듯이 축약형 모형은 특정 변수가 다른 변수에 미치는 영향의 경로를 파악할 수 없다는 단점이 있다. 따라서 본 연구에서 이용한 축약형 모형은 예측 그 자체에 있어서는 유용할 수 있겠으나, 물가상승률 및 경제성장률과 외생변수간 동태적 상관관계를 구조적으로 파악하기 어려우며 어떠한 경로로 영향을 미치는지를 설명할 수 없다. 둘째, 본 연구에서 이용된 베이지안 변수선택 기법의 경우 변수선택이 전적으로 통계적 기준에 의해 결정되기 때문에 경제학적인 근거를 파악해 내기 어렵다. 즉, 특정 예측시계에 특정 변수가 중요변수로 선택되었다 하더라도 그 근거를 경제학적인 논리에 따라 구체적으로 설명하는 것이 불가능하다. 셋째, 추정 과정에서 구조 충격을 식별해내는 과정이 존재하지 않기 때문에 시기별로 어떠한 충격이 물가상승률과 경제성장률에 영향을 주었는지 파악할 수가 없다. 이는 어떤 설명변수의 변동이 특정 구조 충격을 그대로 대변하는 것이 아니라 그 변수에 내재되어 있는 여러 충격의 효과가 복합적으로 작용하여 나타난 결과이기 때문이다. 예를 들어 주택가격이 상승하였을 때 이는 부동산 정책, 유동성, 통화정책, 재정정책, 주택가격 상승기대, 국내경기 호조 등 수 많은 충격요인에 기인한 것일 수 있다. 따라서 이와 같은 축약형 모형의 한계점을 보완하기

위해 DSGE 모형 등 구조 모형을 보조적으로 활용하여 물가상승률과 경제성장률을 예측하고, 예측치의 변동 원인과 경로를 함께 살펴볼 필요가 있다.

아울러 본 연구에서는 표본내 기간 구성시 확장 방식과 롤링 방식을 모두 고려하였다는 점에서 모수의 구조변화를 간접적으로 반영하기는 하였지만, 마코프-스위칭(Markov-Switching) 모형 혹은 시변계수(Time-varying parameter) 모형 등을 통해 구조변화를 명시적으로 고려하지 않았다는 한계점도 존재한다. 본 연구의 분석대상 기간인 2003년 1/4분기에서 2021년 4/4분기 중에는 글로벌 금융위기와 함께 2012년 이후 저물가 기조, 2019년 코로나 위기와 최근의 물가상승세 등 구조변화 가능성이 잠재되어 있는 시기가 여럿 포함되어 있다. 이와 같은 구조변화 가능성은 마코프-스위칭(Markov-Switching) 모형 혹은 시변계수(Time-varying parameter) 모형 등을 통해 분석할 수 있다. 그러나 본 연구에서는 그렇게 하지 않았는데, 이는 강규호·김정성·신세림 (2021)에 따르면 일반적으로 구조변화 여부나 구조변화 발생 시점을 예측함으로써 발생하는 비효율성이 구조변화를 고려함으로써 제거되는 잠재적인 편의보다 크기 때문이다. 그럼에도 불구하고 향후 보다 많은 시계열 자료가 축적되면 본 연구에서 사용한 모형에 구조변화 요인을 도입하여 물가상승률과 경제성장률을 예측해 보는 것도 좋은 연구 사례가 될 것으로 사료된다.

<참고문헌>

- 강규호 (2016), 『베이지안 계량경제학』, 박영사
- 강규호·김정성·신세림 (2021), "우리나라 소비자물가 상승률 예측," 『경제분석』, 제27권 제4호, 한국은행
- 권지호·김도완·지정구·김건·노경서 (2019), "우리나라의 잠재성장률 추정," 『조사통계월보』, 2019년 8월호, 한국은행
- 김현학 (2015), "예측조합 및 밀도함수에 의한 소비자물가 상승률 전망," 『경제분석』, 제21권 제3호, 한국은행, pp. 103-136.
- 박경훈·김찬우·이종웅·지정구·권인하·김철주·장준호 (2021), "최근 인플레이션 논쟁의 이론적 배경과 우리경제 내 현실화 가능성 점검," 『BOK 이슈노트』, 제2021-16호
- 성병희 (2001), "Bayesian VAR模型을 이용한 經濟展望," 『경제분석』, 제7권 제2호, 한국은행
- 조병수·이종웅·권인하 (2021), "주택가격 변동이 실물·물가에 미치는 영향의 비대칭성 분석," 『BOK 이슈노트』, 제2021-17호, 한국은행.
- 최인·황승진 (2012), "한국의 인플레이션 예측," 『응용경제』, 제14권 제3호, 한국응용경제학회, pp. 133-168.
- Blanchard, O. (2020), "Is there Deflation or Inflation in our Future?," *VoxEU.org*, 24 April.
- Blanchard, O. (2021), "In Defence of Concerns over the \$1.9 Trillion Relief Plan," *Peterson Institute for International Economics*, 18 February.
- Daly, K. and R. D. Chankova (2021), "Inflation in the Aftermath of Wars and Pandemics," *VoxEU.org*, 15 April.

- Dieppe, Alistair, Romain Legrand, and Bjorn van Roye (2016), “The Bear Toolbox,” *European Central Bank Working Paper*, No. 1934.
- Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri (2015), “Prior selection for vector autoregressions,” *Review of Economics and Statistics*, Vol. 97(2), pp. 436–451.
- Goodhart, C. and M. Pradhan (2020), “The Great Demographic Reversal: Ageing Societies, Waning Inequality, and an Inflation Revival,” *Cham, Switzerland, Palgrave Macmillan*.
- Goodhart, C. and M. Pradhan (2021), “Friedman vs Phillips: A Historic Divide,” *VoxEU.org*, 26 February.
- Hazell, J., J. Herreño, E. Nakamura, and J. Steinsson (2020), “The Slope of the Phillips Curve: Evidence from US States,” (No. W28005), *National Bureau of Economic Research*.
- IMF (2021), “US Inflation: Set for Take-off?,” *VoxEU.org*, 7 May.
- Krugman, Paul. (2021) “A Conversation with Paul Krugman and Lawrence H. Summers,” Princeton university Bendheim Center for Finance, *MARKUS’ ACADEMY*, February 12.
- Stock, James H., and Mark W. Watson (1999), “Forecasting inflation,” *Journal of Monetary Economics*, Vol. 44(2), pp. 293–335.
- Summers, Lawrence H. (2021), “My column on the stimulus sparked a lot of questions. Here are my answers,” *Washington Post*, February 7.

부록 A. 변수선택 적용 ADL 모형의 사후 샘플링 알고리즘

Spike-and-Slab 사전분포가 적용된 ADL 모형의 파라미터인 $\theta = (\beta, \sigma^2, b_0, b_1, p)$ 및 γ_k 들의 행렬인 $\Gamma = (\gamma_0, \gamma_{1,1}, \gamma_{1,2}, \dots, \gamma_{K+1,J-1}, \gamma_{K+1,J})$ 의 사후분포는 이들의 완전조건부분포를 구한 후 깃스 샘플링 알고리즘을 통해 도출할 수 있다.

A1. (σ^2, β) 샘플링

$Y = X\beta + e$ 의 식을 기준으로 β 의 완전조건부분포는 다음과 같은 정규분포를 따른다.

$$\begin{aligned}\beta | Y, \theta_{-\beta}, \Gamma &\sim N(\beta_1, B_1) \\ B_1 &= (\sigma^{-2} X'X + B_0^{-1})^{-1} \\ \beta_1 &= \sigma^{-2} B_1 X'Y\end{aligned}$$

$\theta_{-\beta}$ 는 θ 에서 β 를 제외한 여타 파라미터의 집합이며 β 의 사전 분산-공분산 행렬 B_0 는 Γ, b_0, b_1 이 주어졌을 때 식 (3)에 따라 다음과 같이 나타낼 수 있다.

$$B_0 = \text{diag}[(1 - \Gamma) \times b_0 + \Gamma \times b_1]$$

B_0 는 대각원소의 갯수가 $(K+1) \times J+1$ 개인 대각행렬이다. Γ 의 초기값을 모두 1로 설정하는 경우 첫번째 반복시행시 대각원소의 값이 모두 b_1 이지만, 이후 반복시행이 진행됨에 따라 Γ 가 샘플링되어 새로운 값이 저장되면서 데이터의 정보가 반영된 사전 분산-공분산 행렬이 생성된다. 그리고 오차항의 분산 σ^2 의 완전조건부분포는 다음과 같은 역감마분포를 따른다.

$$\sigma^2 | Y, \beta \sim IG((\alpha_0 + T)/2, (\delta_0 + (Y - X\beta)'(Y - X\beta))/2)$$

A2. Γ 샘플링

γ_k 의 완전조건부확률은 베이즈룰(Bayes' rule)에 따라 다음과 같이 도출된다.

$$\begin{aligned}\pi[\gamma_k | Y, \beta_k, p, b_0, b_1] &\propto f(Y | \beta, \sigma^2) \pi(\beta | b_0, b_1, \Gamma) \pi(b_0) \pi(b_1) \pi(\Gamma | p) \pi(p) \\ &\propto \pi(\beta_k | b_0, b_1, \gamma_k) \pi(\gamma_k | p)\end{aligned}$$

여기서 두번째 줄 식의 경우 β_k, p, b_0, b_1 이 주어지면 γ_k 의 완전조건부확률이 Y 에 의존하지 않고, β_k 와 γ_k 를 제외한 다른 계수와 지시함수에도 사전적으로 독립이라는 점에 기인한다. 이에 따라 γ_k 가 1의 값을 가질 완전조건부확률을 구하면 다음과 같다.

$$\begin{aligned}\Pr(\gamma_{k,j} = 1 | \beta_{k,j}, b_0, b_1) &= \frac{\pi(\beta_{k,j} | b_0, b_1, \gamma_{k,j} = 1) \Pr(\gamma_{k,j} = 1 | p)}{\pi(\beta_{k,j} | b_0, b_1, \gamma_{k,j} = 1) \Pr(\gamma_{k,j} = 1 | p) + \pi(\beta_{k,j} | b_0, b_1, \gamma_{k,j} = 0) \Pr(\gamma_{k,j} = 0 | p)} \\ &= \frac{p \times N(\beta_{k,j} | 0, b_1)}{p \times N(\beta_{k,j} | 0, b_1) + (1-p) \times N(\beta_{k,j} | 0, b_0)}\end{aligned}$$

$$\Pr(\gamma_{k,j} = 0 | \beta_{k,j}, b_0, b_1) = 1 - \Pr(\gamma_{k,j} = 1 | \beta_{k,j}, b_0, b_1)$$

A3. (b_0, b_1) 샘플링

$\gamma_{k,j}$ 가 1에 해당되는 모든 계수들의 집합을 $\beta_{\Gamma=1}$, 그리고 $\beta_{\Gamma=1}$ 의 원소의 갯수를 K_1 이라고 하자. 반대로 $\beta_{\Gamma=0}$ 은 $\gamma_{k,j}$ 가 0에 해당되는 모든 계수들의 집합이고, K_0 는 $\beta_{\Gamma=0}$ 의 원소의 갯수이다. $\beta_{\Gamma=1}$ 이 주어졌을 때, $\beta_{\Gamma=1} | b_1 \sim N(0, b_1)$ 을 따르고, b_1 의 사전분포가 $b_1 \sim IG(\alpha_{01}/2, \delta_{01}/2)$ 을 따른다. 이 때, 두 분포를 보면 $\beta_{\Gamma=1}$ 이 종속변수이고 설명변수는 존재하지 않으며 b_1 이 오차항의 분산에 해당하는 선형회귀식과 같다는 것을 알 수 있다. 따라서 일반적인 선형회귀식의 완전조건부분포를 적용하여 b_1 과 b_0 의 완전조건부분포를 다음과 같이 나타낼 수 있다.

$$\begin{aligned}b_1 | \beta_{\Gamma=1} &\sim IG((\alpha_{01} + K_1)/2, (\delta_{01} + \beta_{\Gamma=1}' \beta_{\Gamma=1})/2) \\ b_0 | \beta_{\Gamma=0} &\sim IG((\alpha_{00} + K_0)/2, (\delta_{00} + \beta_{\Gamma=0}' \beta_{\Gamma=0})/2)\end{aligned}$$

A4. p 샘플링

p 의 완전조건부밀도는 베이즈룰(Bayes' rule)에 따라 우도와 사전밀도의 곱으로 표현한 후 p 와 독립인 항들을 제외하면 다음과 같이 도출된다.

$$\begin{aligned}\pi(p | Y, \beta, \Gamma, b_0, b_1, \sigma^2) &\propto \pi(p) \pi(\Gamma | p) \\ &= \pi(\gamma_1, \gamma_2, \dots, \gamma_K | p) \pi(p) \\ &= p^{K_1} (1-p)^{K_0} \times p^{a_0-1} (1-p)^{c_0-1}\end{aligned}$$

위의 식의 결과는 베타분포의 밀도와 같은 형태이기 때문에 결국 p 의 완전조건부분포는 다음과 같은 베타분포를 따른다.

$$p | \Gamma \sim B(a_0 + K_1, c_0 + K_0)$$

그리고 각 파라미터들의 완전조건부분포를 이용하여 다음과 같은 알고리즘에 따라 사후분포가 샘플링될 수 있다.

Step 0) 초기값 $\theta_{-\beta}^{(0)}$, $\Gamma^{(0)}$ 를 사전평균으로 설정하고, 샘플링 차수 $i=1$ 로 둔다.

Step 1) β 의 완전조건부분포 $\beta | Y, \theta_{-\beta}^{(i-1)}, \Gamma^{(i-1)} \sim N(\beta_1, B_1)$ 으로부터 $\beta^{(i)}$ 를 샘플링한다.

$$\text{단, } B_1 = (\sigma^{-2} X'X + B_0^{-1})^{-1}, \beta_1 = \sigma^{-2} B_1 X'Y$$

Step 2) σ^2 의 완전조건부분포 $\sigma^2 | Y, \beta^{(i)} \sim IG((\alpha_0 + T)/2, (\delta_0 + (Y - X\beta^{(i)})'(Y - X\beta^{(i)}))/2)$ 로부터 $\sigma^{2(i)}$ 를 샘플링한다.

Step 3) 다음과 같은 b_0 과 b_1 의 완전조건부분포로부터 $b_1^{(i)}$ 와 $b_0^{(i)}$ 를 샘플링한다.

$$\begin{aligned} b_1 | \beta^{(i)}, \Gamma^{(i-1)} &\sim IG((\alpha_{01} + K_1)/2, (\delta_{01} + \beta_{\Gamma^{(i-1)}=1}^{(i)} \beta_{\Gamma^{(i-1)}=1}^{(i)})/2) \\ b_0 | \beta^{(i)}, \Gamma^{(i-1)} &\sim IG((\alpha_{00} + K_0)/2, (\delta_{00} + \beta_{\Gamma^{(i-1)}=0}^{(i)} \beta_{\Gamma^{(i-1)}=0}^{(i)})/2) \end{aligned}$$

Step 4) p 의 완전조건부분포 $p | \Gamma^{(i-1)} \sim B(a_0 + K_1, c_0 + K_0)$ 로부터 $p^{(i)}$ 를 샘플링한다.

Step 5) 각 k 에 대하여 만약 균등분포 $u \sim Unif(0, 1)$ 에서 임의로 추출한 값이

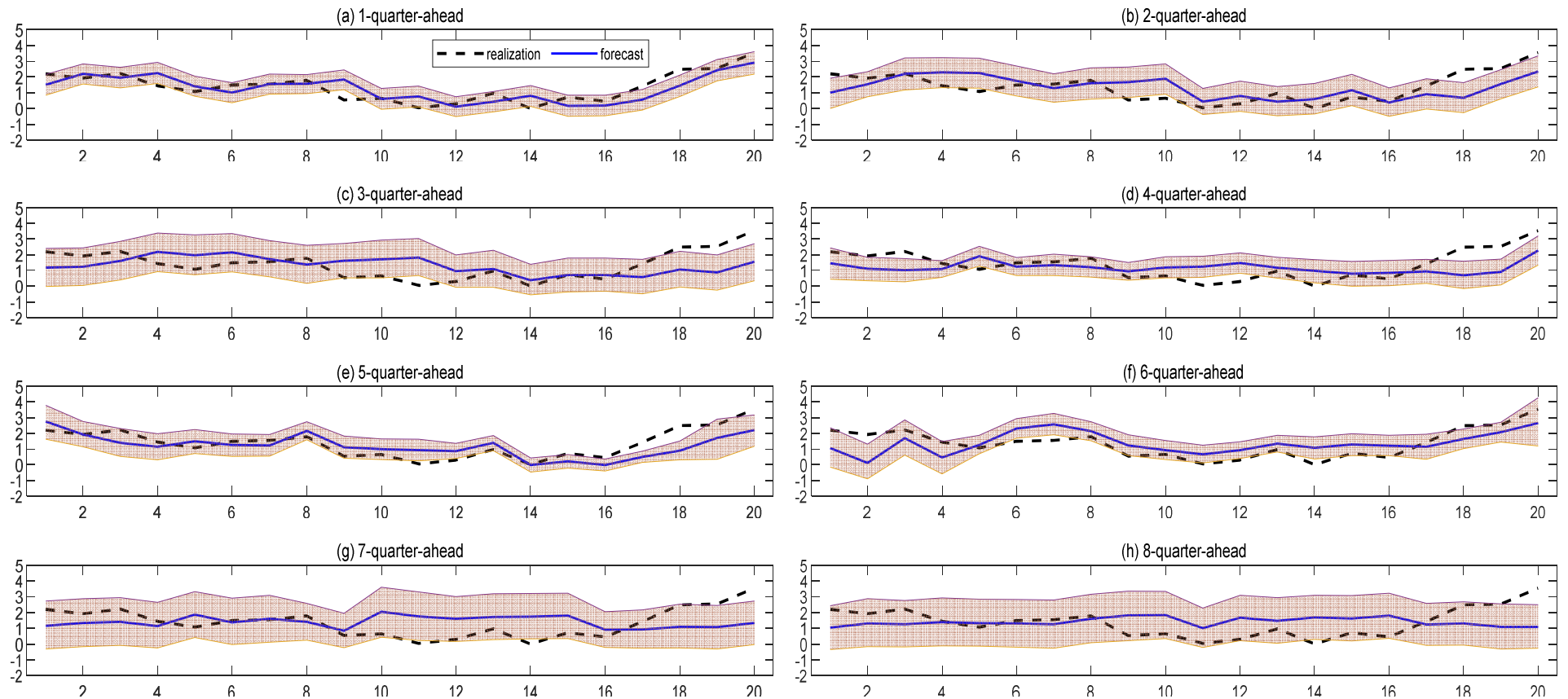
$$\Pr(\gamma_{k,j} = 1 | \beta_{k,j}^{(i)}, b_0^{(i)}, b_1^{(i)}) = \frac{p^{(i)} \times N(\beta_{k,j}^{(i)} | 0, b_1^{(i)})}{p^{(i)} \times N(\beta_{k,j}^{(i)} | 0, b_1^{(i)}) + (1 - p^{(i)}) \times N(\beta_{k,j}^{(i)} | 0, b_0^{(i)})}$$

보다 작으면 $\gamma_{k,j}$ 가 1, 아니면 0의 값을 가지는 것으로 지정하여 샘플링한다.

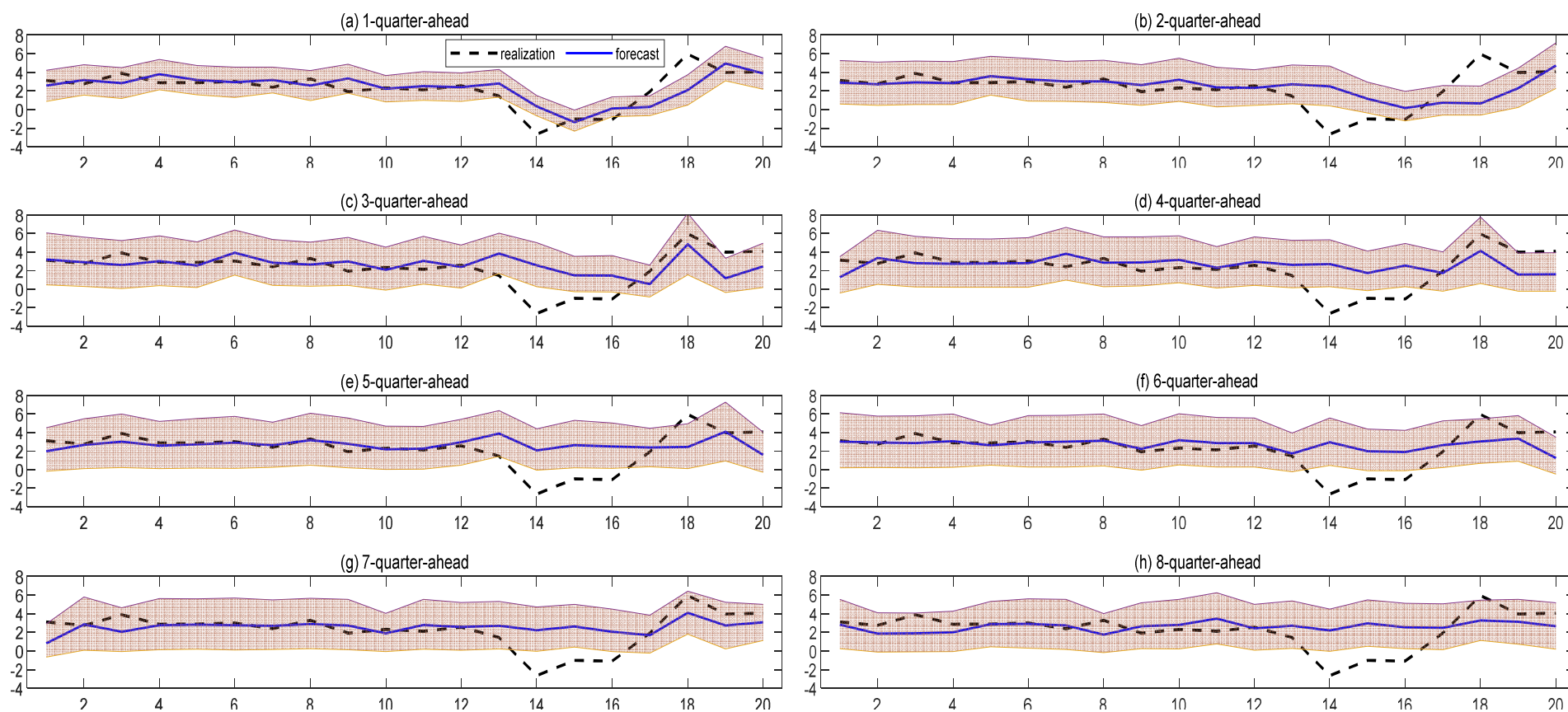
Step 6) $i = i + 1$ 로 설정하고, $i \leq n$ (반복시행 총 횟수)이면 1단계로 돌아간다.

부록 B. 표본외 예측결과 - 90% 신용구간

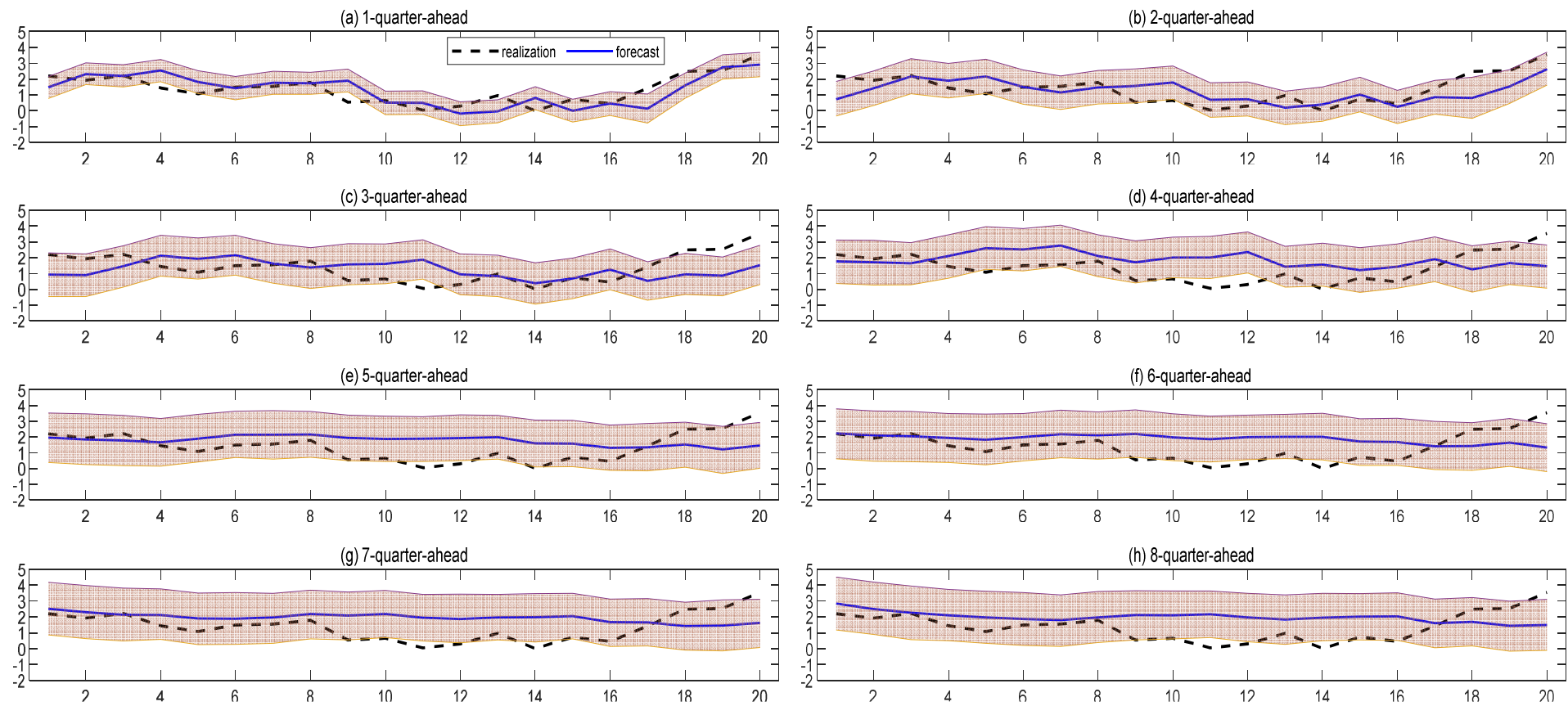
1. Uni-ADL(CPI)



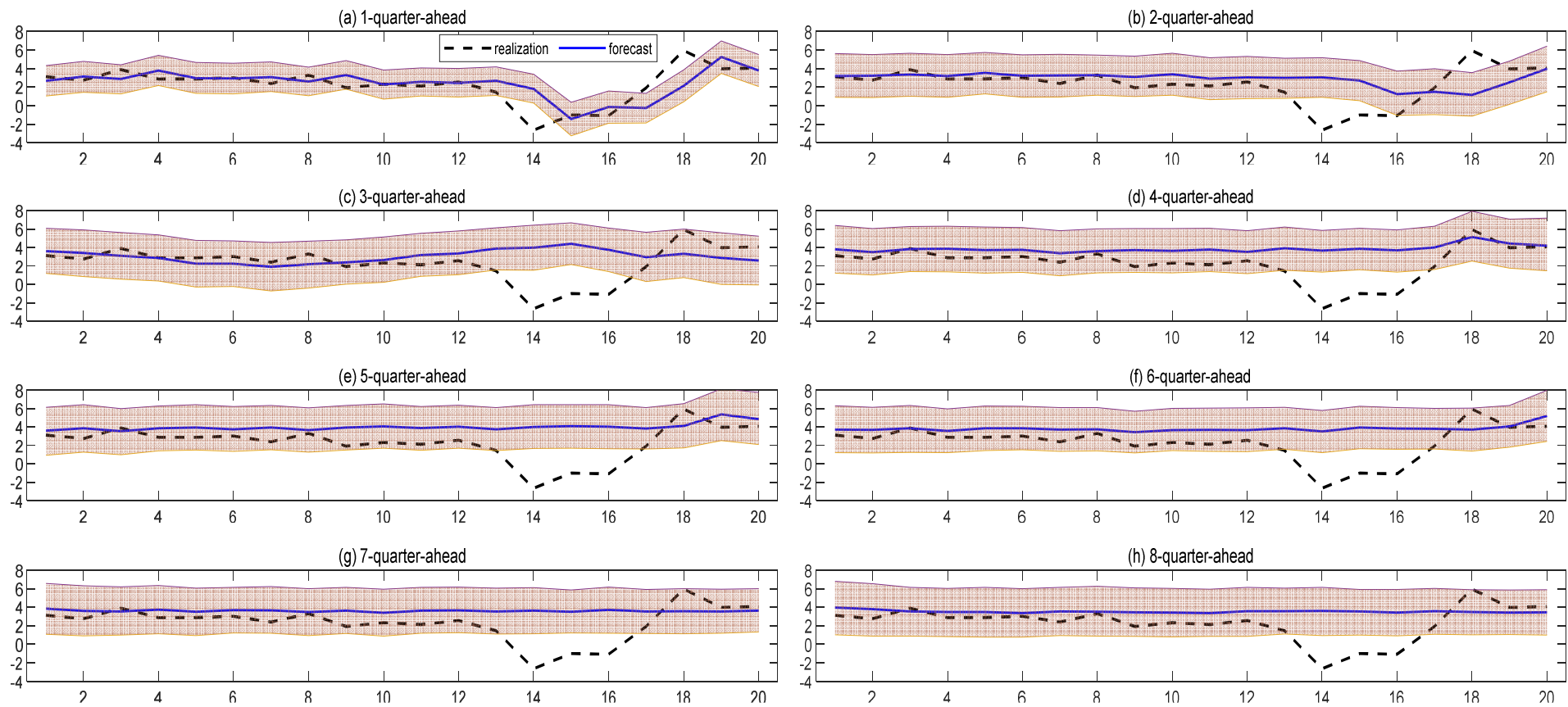
2. Uni-ADL(GDP)



3-1. Bi-ADL(CPI)



3-2. Bi-ADL(GDP)



Predictive Probability Estimation of Stagflation in Korea

Keunhyeong Park¹⁾, Kyu Ho Kang²⁾

The prediction of the short- and long-term stagflation probability has a profound influence on the Bank of Korea's preemptive monetary policy decision-making for economic and price stability. This study estimates the probability of stagflation in Korea over the next two years. To this end, a joint predictive distribution of the inflation rate and the real economic growth rate is generated using univariate and bivariate autoregressive distributed lag (ADL) models. In this process, the optimal prediction model is selected through Bayesian variable selection and precise tuning through out-of-sample prediction. As a result of long-term and short-term out-of-sample predictions for the last 5 years, the joint distribution predictive accuracy was maximized when each variable was independently predicted using the univariate ADL model in most forecasting horizons. As a result of the actual prediction using the optimal model, the probability of stagflation gradually increasing from the fourth quarter of 2021 to 38.4% in the first quarter of 2023. Thereafter, although it decreases slightly, it remains high at 32.3% until the third quarter of 2023, after which the probability drops to the 20% level, and the probability of an economic boom increases instead. Therefore, it is estimated that the probability of temporary stagflation for the next one to one and a half years is higher than that of the normal period.

JEL Classification Number: G12, C11, E43

Keywords: Out-of-sample joint prediction, monetary policy, business cycles, Bayesian variable selection

1) Human Resource Department, Bank of Korea, Seoul, 04514, South Korea (E-mail: keunhyeong.park@bok.or.kr, Tel: 02-759-4114)

2) Corresponding author. Department of Economics, Korea University, Seoul, 02841, South Korea (E-mail: kyuhu@korea.ac.kr, Tel: 02-3290-5132)