



2021.12.10

국회예산정책처 | 경제현안분석 | 제104호

단기 GDP 예측모형 연구

A short-term forecasting model for GDP

오현희



국회예산정책처
NATIONAL ASSEMBLY BUDGET OFFICE

단기 GDP 예측모형 연구

경제현안분석 제104호

단기 GDP 예측모형 연구

총 괄 | 이 정 은 경제분석국장

기획·조정 | 김 윤 기 거시경제분석과장

작 성 | 오 현 희 거시경제분석과 경제분석관

지 원 | 엄 상 미 거시경제분석과 행정실무원

「경제현안분석」은 국회가 관심 있게 다룰 재정 현안이나 정책 이슈에 대해 객관성 있는 분석정보를 적시성 있게 제공함으로써, 국회의 예산 및 법안 심사와 의제 설정을 실효성 있게 지원하기 위한 것입니다.

문 의 : 경제분석국 거시경제분석과 | 02) 6788-3775 | macro@nabo.go.kr

이 책은 국회예산정책처 홈페이지(www.nabo.go.kr)를 통하여 보실 수 있습니다.

“본 보고서는 담당 분석관의 연구 결과를 바탕으로 작성된 것으로
국회예산정책처의 공식견해와 다를 수 있음을 알려드립니다”

단기 GDP 예측모형 연구

2021. 12

이 보고서는 「국회법」 제22조의2 및 「국회예산정책처법」 제3조에 따라
국회의원의 의정활동을 지원하기 위하여, 국회예산정책처 「보고서발간심의
위원회」의 심의(2021. 11. 04.)를 거쳐 발간되었습니다.

발간사

2020년 이후 대내외 경제 환경은 코로나19로 인해 큰 변화를 맞이하였습니다. 비대면 경제의 활성화, 노동시장의 변화, 4차 산업혁명의 급진전 등 경제환경이 빠르게 변화하고 있습니다. 또한, 코로나19 전개상황에 대한 예측이 여전히 어려운 상황에서 포스트코로나 시대 통화·재정정책 정상화, 자동화 및 ESG 경영 본격화, 탄소중립경제로의 전환 등 다방면에서의 구조 변화는 경제의 불확실성을 확대시키는 요인이 되고 있습니다.

이와 같이 대내외 경제환경이 크게 변화하는 시기에 효과적인 정책대응을 위해서는 경제상황을 보다 정확하게 판단하고 예측하는 것이 중요하다고 하겠습니다. 특히, 실질국내총생산(GDP)은 경제상황을 가늠할 수 있는 대표적인 지표라고 할 수 있습니다. 그러나, GDP는 매 분기 실적치가 시차를 두고 공표되고 있어 시의성이 떨어지는 한계가 있습니다. 이에 대내외 기관들에서는 상대적으로 경제환경의 변화를 신속하게 반영할 수 있는 월별지표를 이용하여 단기 GDP를 예측하는 모형을 개발하여 운용하고 있습니다.

본 보고서는 비교적 빠르게 입수할 수 있는 26개의 월별자료를 이용하여 주성분분석 및 동태요인분석 방법을 통해 단기 GDP성장률 예측 모형을 구축하였습니다. 분석결과 본 모형은 기존 분기모형에 비해 예측오차가 작은 것으로 나타나 거시경제 전망의 정확도를 제고하는 보완도구로 활용될 수 있을 것으로 기대됩니다.

앞으로도 국회예산정책처는 경제전망의 예측력을 높이기 위해 거시경제 전망모형의 고도화를 지속적으로 추진해 나가도록 하겠습니다. 끝으로 본 보고서는 담당 분석관의 연구 결과를 바탕으로 작성된 것으로 국회예산정책처의 공식견해와 다를 수 있음을 알려드립니다.

2021년 12월

국회예산정책처장 임익상

차 례

I. 연구개요 / 1

1. 연구의 배경 및 목적	1
2. 단기예측모형 방법론	5
3. 기존 연구 검토	7

II. 단기 실질GDP 예측 모형설정 / 15

1. 분석모형	15
가. 주성분분석모형	15
나. 동태요인모형	19
2. 분석방법	23
가. 결측치 처리방법	23
나. 예측력 비교 방법	26
다. 분석기간	30
라. 이용자료	35

III. 분석결과 / 41

1. 주성분분석	41
2. 동태요인분석	47
3. 코로나19 기간을 포함한 분석 결과	49

IV. 요약 및 시사점 / 57

참고문헌 / 61

[부 록] 국내외 기관의 단기 예측 모형 운용현황 / 65

경제현안분석 목록 / 71

표 차례

[표 1] 주성분분석을 활용한 기존 연구	10
[표 2] 동태요인모형을 활용한 기존 연구	13
[표 3] 우리나라의 기준순환일 및 국면 지속기간	31
[표 4] 분석에 포함된 변수 목록	35
[표 5] 분기성장률 지표의 비교	39
[표 6] 주성분분석 결과	43
[표 7] 각 주성분의 주요 정보변수	44
[표 8] 주성분분석모형의 예측력 평가	45
[표 9] 동태요인분석모형의 예측력 평가	48
[표 10] 주성분분석결과	51
[표 11] 각 주성분의 주요 정보변수	52
[표 12] 2020년을 포함한 모형의 예측력 평가	53

그림 차례

[그림 1] 한국은행의 GDP 성장률 공표시기	2
[그림 2] 한국은행 실질성장률 속보치와 잠정치 추이	3
[그림 3] Kalman filter 추론과정	22
[그림 4] 월별자료의 공표시점	23
[그림 5] 추차예측(recursive forecast) 방식	29
[그림 6] 롤링예측(rolling window forecast) 방식	29
[그림 7] 실질GDP성장률의 예측 및 평가기간	30
[그림 8] 실질GDP성장률의 예측 및 평가기간(2020년 포함)	32
[그림 9] IMF의 전세계 불확실성 지수 추이	33
[그림 10] GDP성장률과 이용변수간 상관관계	37
[그림 11] 실질GDP성장률 추이(%)	38
[그림 12] 주성분분석 Scree Plot	42
[그림 13] 주성분분석 Parallel 분석	42
[그림 14] 주성분분석 Scree Plot(2020년 포함)	49
[그림 15] 주성분분석 Parallel 분석(2020년 포함)	50
[그림 16] 당분기 예측의 모형별, 측정방법별 예측오차 추이	55

요 약

I. 연구개요

- 최근 코로나19 등으로 대내외 경제환경이 급변하고 거시경제의 불확실성이 높아지면서 효과적인 대응을 위한 다양한 경제정책 시행이 필요한 상황
 - 보다 효과적인 정책대응을 위해서는 적시에 적절한 경제정책을 시행해야 하며 이를 위해서는 관련 거시경제지표 등을 통해 경제상황을 정확하게 판단하고 예측할 필요가 있음
 - 단기의 경제상황을 파악하고 향후 성장경로를 예측하기 위해서는 가장 대표적인 경제지표 중 하나인 실질GDP성장률의 1~2분기 예측이 중요
- 그러나 국내총생산(GDP)은 매 분기 약 1개월 이상의 시차를 두고 공표되고 있어 시의성이 떨어지는 문제가 발생
 - 한국은행은 분기 종료 후 28일 이내에 속보치를 발표하며, 해당분기 종료 후 70일 이내 잠정치치를 발표하고 있어 경제상황의 판단이나 예측에 GDP성장률을 이용하기가 쉽지 않은 상황
- 이에 따라 본 연구에서는 빠르게 입수할 수 있는 월별자료를 이용하여 분기 GDP성장률을 미리 예측할 수 있는 단기예측모형을 개발하고자 함
 - 분기지표인 GDP에 비해 상대적으로 공표시점이 빠른 월별지표를 이용하는 경우 경제변수의 변화를 보다 정확하게 측정하고 예측할 수 있음
 - 월별지표를 이용한 모형은 매월 발표되는 새로운 정보를 추가하여 예측하는 만큼 단기예측에 있어 정확성을 높일 수 있을 뿐만 아니라 당분기 또는 익분기에 대한 정확한 예측이 가능하게 되면 중장기 예측의 정확성 또한 제고할 수 있음

- 이를 위해 월별 자료를 이용하여 분기 자료를 추정하는 주성분분석 및 동태 요인분석 모형을 기본으로 설정하여 분기 GDP성장률을 예측하는 모형을 구축
 - 분기자료를 이용한 ARIMA, VAR 등 시계열모형은 월별자료 등 속보성 있는 자료를 이용하지 못할 뿐만 아니라 경제변수간의 유기적 관계 등을 고려하기 어렵다는 단점이 있음
 - 이들 시계열모형은 단독으로 연구하기보다는 대부분의 연구에서 다른 예측 모형의 비교대상으로 분석되고 있음
 - 월별지표를 이용하여 분기지표를 예측하는 모형 중 고전적 브릿지모델(bridge model)은 한정된 변수를 이용하기 때문에 월별지표들의 정보를 충분히 활용하기 어려운 단점이 있음
 - 고전적 브릿지모델은 GDP와 동행 및 선행성이 높은 월별변수를 통해 GDP 구성항목을 예측하고 GDP 총량간의 연계식으로 모형을 구축
 - 이러한 한계를 극복하기 위해 최근에는 보다 광범위한 월별데이터를 이용하여 요인(factor)을 추출하고 이를 이용하여 GDP를 예측하는 요인모형들이 이용되고 있으며 본 연구에서는 주성분분석과 동태요인분석 모형을 이용하여 GDP를 예측
 - 주성분(Principal Component)분석은 경제변수들이 지닌 특성을 반영하면서 상호독립인 공통요인들을 추출하는 방법이며, 이와 같이 추정된 공통요인과 예측하려는 변수인 GDP성장률간의 선형관계를 추정하여 예측방정식을 도출
 - 동태요인모형은 주성분분석을 이용하여 공통요인을 추출한 후 상태공간모형에 칼만필터(Kalman filter)를 적용하여 관측되지 않은 공통요인 F_t 를 추정하고 이를 GDP성장률 예측에 사용하는 방법
 - ※ 칼만필터는 추정된 계수(parameter)에 업데이트된 정보를 반영하여 계수의 추정치를 업데이트하는 축차적(recursive) 과정

II. 단기 실질GDP 예측 모형설정

- 월별지표를 이용하여 분기 GDP를 예측하는 방법은 Giannone et al.(2004), Giannone et al.(2005) 등 기존연구들에서 분석되어온 연계 모형(bridge equation)을 차용
 - 월별 변수($X_t = (x_{1t}, \dots, x_{nt})'$, $t = 1, \dots, T$)들의 공통요인 F_t 를 추출한 후 월별로 추출된 공통요인을 분기별로 변환한 데이터($aggregate(f_t^Q)$)를 이용하여 GDP성장률과의 회귀분석을 수행
 - 본 연구에서는 Stock and Watson(2002)과 Barhoumi(2008)의 주성분분석모 형과 Doz et al.(2007)과 Giannone et al.(2008)의 동태요인분석 모형을 통해 GDP성장률을 추정
 - 주성분분석을 통한 공통요인(common factor) 추정은 다변량 자료에 포함된 많은 정보들의 손실을 최소화하여 낮은 차원의 데이터로 축약하는 방식으로 제한된 변수를 활용하여 신뢰도가 높은 결과를 얻을 수 있음
 - 동태요인모형은 월별지표가 보유한 광범위한 정보를 공통요인으로 축약하는 한편 공통요인의 동태적 특성을 이용함으로써 자료의 주기와 공표시점 차이 등을 내생화하여 자료의 효율적 처리가 가능
- t분기 중 m_1 (초월), m_2 (중월), m_3 (말월) 등으로 각각 월에 발표되는 지표의 시점 차이가 발생하기 때문에 전망과정에서 월간지표별 공표 시차가 달라 발생하는 결측치 및 표본외 예측에 필요한 예측치를 생성할 필요
 - 주성분분석 모형의 경우 Barhoumi(2008)의 방식을 따라 월간자료 중 결측치 및 예측치가 발생할 경우 AR(p) 모형을 이용하여 값을 생성
 - 동태요인분석 모형에서는 자기회귀를 통해 공통요인의 동태적 특성을 분석하고 칼만필터 방법을 통해 결측치 및 예측치를 생성

□ 예측력을 평가하기 위한 첫 번째 단계로 표본 내(in-sample) 자료를 이용하여 모형을 추정하고 추정된 모형으로 7표본 외(out-of sample) 예측을 실시하여 실제 GDP성장률과 이를 비교하는 방법을 사용

○ 표본 외 예측치와 실제치를 비교하기 위해 RMSE(Root Mean Squared Error)와 MAE(Mean Absolute Error)를 계산하여 이를 비교

- RMSE(Root Mean Squared Error)는 실제값과 예측값의 차이인 오차들의 제곱 평균인 MSE(Mean Squared Error)에 제곱근을 씌운 값

$$RMSE = \sqrt{[\sum (\hat{y}_t - y_t)^2 / H]}$$

- MAE(Mean Absolute Error)는 실제값과 예측값의 차이인 오차들의 절댓값의 평균

$$MAE = [\sum |\hat{y}_t - y_t|] / H$$

○ 표본 외 예측기간 설정 시 축차예측(recursive forecast)과 롤링예측(rolling forecast) 방법을 모두 추정하여 예측오차를 비교

- 축차예측은 처음 정해진 구간을 기본으로 하여 구간을 추가하면서 추정하여 예측하는 방식
- 롤링예측은 표본의 수를 일정하게 유지한 상태에서 시점을 이동하면서 예측치를 도출하는 방법

□ 분석기간은 2000년부터 2019년까지로 설정하여 2008년 금융위기 기간을 포함하였으며, 코로나19 발생 직전까지의 기간을 대상으로 함

○ 예측오차의 평가기간은 2015년 1/4분기부터 2019년 4/4분기까지 20개 분기로 설정

○ 코로나19의 충격으로 경제성장률의 변동성이 매우 확대된 2020년을 분석에 포함할 경우 도출된 결론이 동 기간에 크게 좌우될 가능성 등을 고려하여 2019년까지를 기본적인 분석기간으로 설정

- 다만, 2020년을 포함한 분석결과도 함께 제시하여 코로나19가 예측성과에 미치는 영향을 함께 고려

III. 분석결과

- 분석결과 월별자료를 이용해 GDP를 예측한 주성분분석 모형이 분기자료를 이용해 GDP를 예측한 벤치마크 모형인 AR 모형에 비해 예측력이 우수한 것으로 나타남
 - 주성분분석 결과를 토대로 추출된 요인들과 GDP성장률에 대한 회귀분석을 통해 예측모형을 설정하여 표본 외 예측을 실시하였으며, 표본 외 예측결과를 실제 GDP성장률과 비교하여 예측오차를 산출
 - 주성분분석모형을 이용한 당분기 GDP예측의 경우 RMSE가 0.35~0.40%p, MAE가 0.28~0.38%p로 나타나, AR모형의 예측오차(RMSE 0.58%p, MAE 0.78%p)에 비해 작게 나타남
 - 주성분분석을 통한 예측 결과 전체 월(초월, 중월, 말월의 결과를 평균)과 초월, 중월, 말월의 예측오차 차이는 크게 나타나지 않은 모습
 - 이는 초월에서 말월로 이동하면서 모형의 모수 추정치가 크게 변하지 않았고 예측에 사용된 월별 변수 결측치도 어느 정도 정확성을 확보하고 있기 때문으로 판단

[주성분분석모형의 예측력 평가]

		당분기		다음 분기	
		RMSE	MAE	RMSE	MAE
축차예측	전체	0.397	0.315	0.445	0.608
	초월	0.399	0.312	0.446	0.625
	중월	0.385	0.352	0.445	0.595
	말월	0.386	0.377	0.438	0.588
롤링예측	전체	0.370	0.284	0.417	0.586
	초월	0.387	0.295	0.421	0.589
	중월	0.391	0.301	0.400	0.579
	말월	0.352	0.322	0.412	0.577
AR모형 ¹⁾	-	0.581	0.777	0.596	0.856

- 추차예측과 롤링예측을 비교해보면 당분기와 다음 분기 모두 롤링예측이 추차예측에 비해 예측오차가 작은 것으로 나타남
 - 롤링방식의 경우 오래된 자료를 배제하고 최신의 자료를 사용하여 월별자료의 변화를 반영하였기 때문에 예측오차를 줄인 것으로 판단
 - 또한, 글로벌 금융위기 시기인 2008년이 분석기간에 포함되어 있어, 이 시기에 성장추세의 구조적 변화가 발생하였다면 시계열의 불안정성을 초래하여 추차예측 방식의 예측력을 낮추었을 것으로 판단
- 월별자료를 이용해 GDP를 예측한 동태요인모형은 주성분분석모형과 마찬가지로 벤치마크 모형인 AR모형에 비해 예측력이 우수한 것으로 나타났으며, 주성분분석에 비해서도 예측오차가 작게 나타남
 - 월별지표를 이용한 동태요인분석모형의 당분기 GDP예측의 경우 RMSE는 0.34~0.38%p, MAE는 0.25~0.28%p로 나타남
 - 동태요인분석을 통한 GDP예측 결과 MAE와 RMSE가 주성분분석모형에 비해 낮아진 것으로 나타나 상대적으로 예측력이 우수한 것으로 판단
 - 주성분분석모형과 예측력을 비교해 보면 동태요인분석을 통한 실질GDP 예측은 당분기와 다음 분기, 추차와 롤링방식 모두에서 예측오차가 작게 나타남
 - 이는 AR모형을 통해 결측치를 처리하는 주성분분석모형에 비해 공통요인의 동학을 추정하여 결측치를 처리하는 동태요인분석 방식이 보다 유용한 정보를 포함하고 있어 예측오차를 낮춘 것으로 판단
 - 다만, 주성분 분석과 다르게 MAE보다 RMSE를 통해 계산된 예측오차가 크게 나타나, 작은 오차가 지속적으로 나타나기 보다는 한번씩 큰 오차를 발생시키는 것으로 보임
 - 주성분분석을 통한 예측결과와 마찬가지로 전체월과 초월, 중월, 말월의 예측오차 차이가 크게 나타나지 않았으며, 당분기와 다음 분기 모두 롤링예측이 추차예측에 비해 예측오차가 작은 것으로 나타남

1) 각 시점별로 AIC, BIC 기준에 따라 최적시차를 결정

[동태요인분석모형의 예측력 평가]

		당분기		다음분기	
		RMSE	MAE	RMSE	MAE
축차예측	전체	0.365	0.279	0.439	0.472
	초월	0.375	0.280	0.437	0.475
	중월	0.360	0.245	0.440	0.477
	말월	0.359	0.275	0.435	0.468
롤링예측	전체	0.338	0.263	0.352	0.390
	초월	0.337	0.266	0.357	0.396
	중월	0.342	0.259	0.351	0.390
	말월	0.335	0.263	0.351	0.385
AR모형 ²⁾	-	0.581	0.777	0.596	0.856

□ 한편, 코로나19의 충격이 반영된 2020년 자료를 분석에 포함시킬 경우 예측오차가 커지는 결과를 보임

- 2020년 GDP 성장률 추정치에 대한 예측오차가 상대적으로 매우 크게 나타났기 때문에 RMSE가 큰 폭으로 상승
 - 주성분분석모형에서는 0.37~0.40%p 수준이던 RMSE가 0.76~0.83%p로 높아졌으며, 동태요인분석모형에서는 0.34~0.37%p에서 0.75~0.83%p로 높아짐
- 2020년은 과거의 GDP성장률과 추세적으로 괴리되어 있으며, 주성분분석 결과 주요 정보변수들이 차이를 나타내고 있어 월별 지표들의 움직임도 과거 데이터와는 다른 패턴을 보이는 것으로 판단
- 이와 같은 큰 변동성, 추세적 이탈로 인해 추정치가 실적치를 벗어나면서 예측오차가 커지고 RMSE가 높아진 것으로 분석

2) 각 시점별로 AIC, BIC 기준에 따라 최적시차를 결정

IV. 요약 및 시사점

- 본 연구에서는 월별자료를 이용하여 주성분분석과 동태요인분석모형을 토대로 GDP성장률을 전망하고 예측결과를 비교하였으며, 구축된 예측모형을 통해 한국은행 속보치보다 1~2개월 빠르게 실질GDP성장률 예측치 생성이 가능
 - 분석결과 본 연구에서 구성한 단기예측모형인 주성분 및 동태요인분석 모형은 벤치마크 모형인 AR모형에 비해 예측력이 우수한 것으로 나타남
 - 한국은행의 실적치 발표 전에 분기별 실질GDP성장률을 생성함으로써 경제전망의 정확도를 제고할 수 있을 것으로 예상

- 다만, 본 연구에서 이용한 예측모형은 몇 가지 점에서 개선할 여지가 있음
 - 향후 시계열이 충분히 확보되면 2020년 자료를 분석에 포함시켜 모형을 재구축하고 예측오차를 평가해 볼 필요
 - 가장 최근 자료이지만, 불확실성이 매우 높았던 2020년을 분석에 포함할 경우 예측오차가 커져 예측력이 낮아짐
 - 2020년은 코로나19의 영향으로 변동성이 매우 높았을 뿐만 아니라 이전 GDP성장률과도 추세적으로 괴리되는 모습을 보이고 있어 예측력을 저하시킴
 - 또한, 후속연구에서는 속보치, 잠정치, 확정치를 구분한 모형을 구성하는 작업도 의미 있는 시도가 될 것으로 보임
 - 본 연구에서는 GDP의 속보치, 잠정치, 확정치를 구분하지 않고 시계열이 구성되어 있는 확정치를 사용
 - 실제 예측과정에서는 초월의 경우 전분기 GDP성장률의 속보치만 가용할 수 있어 속보치와 확정치의 차이 등에 따른 오차가 발생할 가능성

I. 연구개요

1. 연구의 배경 및 목적

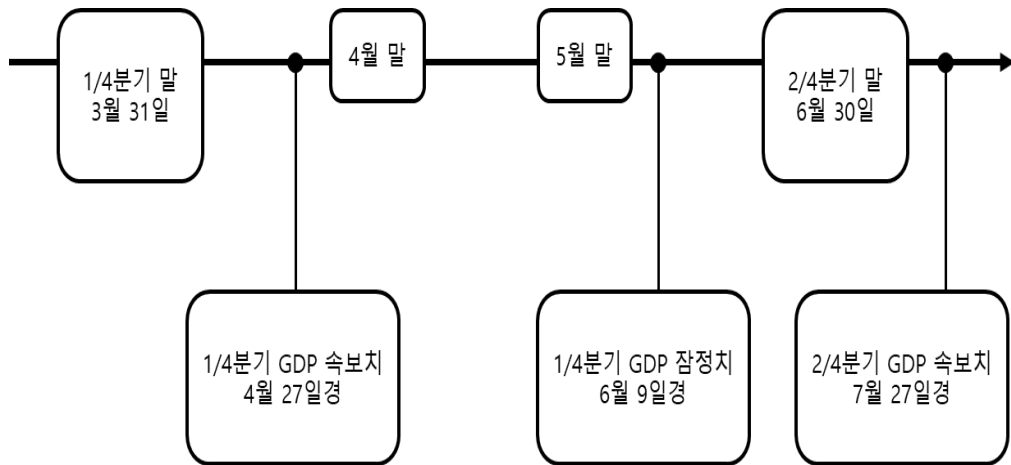
최근 국내외 경제환경이 급변하면서 거시경제 여건의 불확실성이 확대되고 있으며, 변화의 방향은 불투명한 상황이다. 특히 미·중간 무역분쟁과 코로나19를 계기로 세계경제는 새로운 변화의 시대에 직면하고 있다. 미·중간 무역분쟁과 보호무역주의 확산으로 세계교역신장이 위축되고 글로벌 가치사슬이 약화 되는 등 대외의존도가 높은 우리경제에 리스크를 높이고 있다. 우리나라의 1, 2위 교역대상국인 미·중간 갈등 장기화는 우리경제의 불확실성을 높이고 실물 및 금융부문에 부정적 영향을 미칠 가능성이 크다. 또한, 코로나19의 대유행은 국내외 경제에 수요와 공급측면에서 충격을 초래하는 전방위적인 영향을 미치고 있다. 세계경제 교역량이 위축되면서 생산 및 수출 부문에 충격을 미쳤으며, 각종 제한 조치들로 경제주체들의 소비심리가 위축되면서 내수 또한 부진한 모습을 보였다. 이에 따라 각국 정부 및 중앙은행들은 코로나19에 따른 경기둔화를 막기 위해 각종 완화정책을 시행하였다. 2021년 들어 백신공급 확대 및 접종의 진전으로 각종 봉쇄조치가 완화되면서 세계경제가 점차 회복세를 나타내고 있는 모습이나, 변이바이러스의 확산 등으로 인해 회복 경로의 불확실성은 아직 남아있는 상황이다.

거시경제의 불확실성이 높아지면서 추가적인 완화정책에 대한 사회적 요구가 높아졌으나, 최근 들어서는 주요국 물가상승 압력 증대 등으로 자산매입 축소와 금리인상 등 긴축정책 시행에 대한 논의가 증대되는 상황이다. 보다 효과적인 정책대응을 위해서는 적시에 적절한 경제정책을 시행할 필요성이 있으며 이를 위해서는 경제지표 등을 통해 경제상황을 정확하게 판단하고 예측할 필요가 있다.

경제상황을 가늠할 수 있는 가장 대표적인 지표 중 하나인 실질국내총생산(GDP)은 한 나라 경제가 이룩한 성과를 측정하는 척도³⁾이나 매 분기 실적치가

분기말과 약 1개월(속보치 기준) 이상의 시차를 두고 공표되고 있어 시의성이 떨어지는 문제가 발생한다. 한국은행은 분기 종료 후 28일 이내에 속보치를 발표하며, 해당분기 종료 후 70일 이내 잠정치를 발표하고 있어 경제상황의 판단이나 예측에 GDP성장률을 이용하기가 쉽지 않은 상황이다⁴⁾.

[그림 1] 한국은행의 GDP 성장률 공표시기



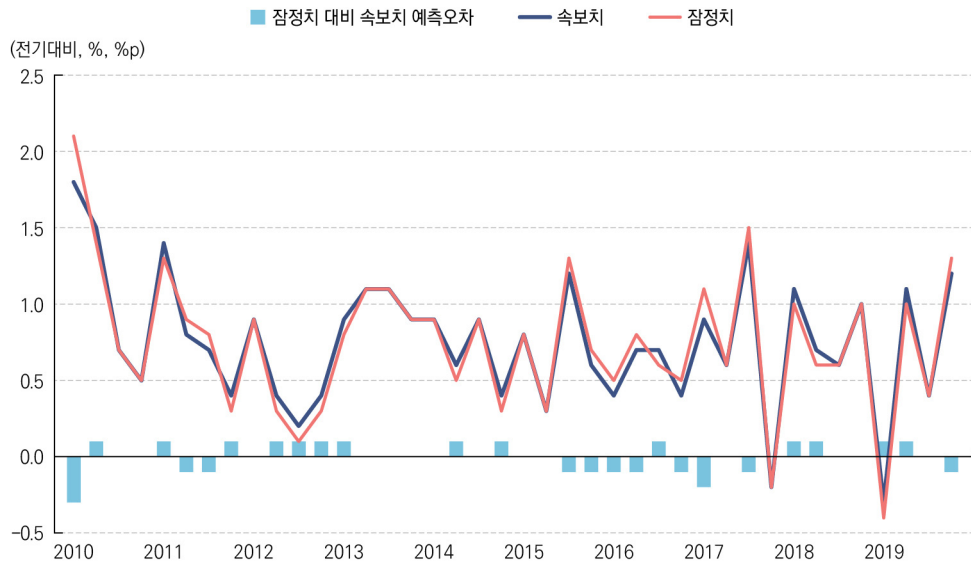
주: 2021년 기준

특히 속보치의 경우 각분기 말월에 대한 기초자료를 이용하지 못하고 편제한 수치로서 대부분의 경우 잠정치 및 확정치와 차이를 보이고 있다. 이에 따라 단기의 경제상황을 파악하고 향후 성장경로를 예측하기 위해 1~2분기 후 실질GDP 성장률을 예측할 필요가 있으며 예측오차를 최소화할 수 있는 효율적인 단기전망 모형을 개발할 필요가 있다.

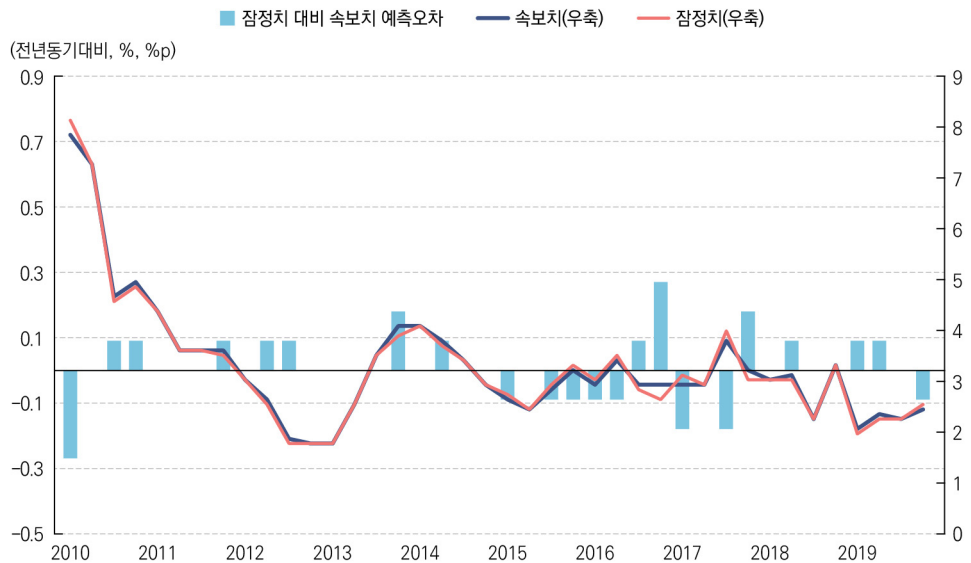
3)「우리나라의 국민계정체계」, 한국은행, 2015

4) 확정치의 경우에는 익년 6월에 공표

[그림 2] 한국은행 실질성장률 속보치와 잠정치 추이



자료: 한국은행



자료: 한국은행

통상적으로 0~2분기 정도의 시계에 대한 GDP를 예측하는 단기전망모형은 정교한 경제이론보다는 경제상황 등에 영향을 받는 경향이 있다. 따라서 상대적으로 국내외 경제환경의 변화를 즉각적으로 반영할 수 있는 월별지표를 이용하는

경우 경제변수의 변화를 보다 정확하게 측정하고 예측할 수 있다. 분기별 GDP가 발표되기 이전에 월별로 발표되는 자료들은 현재의 경제상황이나 여건 등을 보다 빠르게 반영할 수 있다는 장점을 가지고 있으며 이를 이용하여 분기 GDP를 선제적으로 예측하는 단기예측모형을 개발할 수 있다. 월별지표를 이용한 모형은 매월 발표되는 새로운 정보를 추가하여 예측하는 만큼 단기예측에 있어 정확성을 높일 수 있을 뿐만 아니라 당분기 또는 익분기에 대한 정확한 예측이 가능하게 되면 중장기 예측의 정확성 또한 제고할 수 있다. 당분기의 정확한 예측이 1년간의 예측의 정확성을, 정확한 당해연도 예측이 중장기 예측의 정확성을 향상(김양우·이금희·장동구(1997)) 시킬 수 있으므로 초단기 예측이 정확하게 이루어질 필요가 있다.

이러한 유용성에도 불구하고 월별지표를 이용한 단기 GDP예측모형은 월별 지표간 발표시점의 차이에 따른 문제, 월별 지표를 이용하여 분기지표를 예측하는 과정에서 나타날 수 있는 혼잡주기 자료 문제 등에 봉착할 수 있다.

본 보고서에는 이러한 점을 감안하여 그동안 이루어져 온 단기예측모형을 검토하고 효율적인 예측모형을 설정함으로써 예측력이 우수한 모형을 구축하고자 하였다. 1장에서는 기존에 연구되어 온 단기예측모형에 대해 살펴보고 관련 선행연구들을 검토하였다. 2장에서는 실질 GDP예측 분석모형인 주성분분석, 동태요인모형에 대해 정리하였으며, 결측치를 처리하는 방법과 예측력을 비교하는 방법에 대해 설명하였다. 3장에서는 앞선 논의를 바탕으로 단기 GDP 예측 모형을 추정하는 과정과 결과를 살펴보았다. 또한, 표본 외 예측(out-of-sample forecast)을 실시하여 각 모형 간 예측력을 비교하였다. 마지막으로 4장에서는 결과를 요약하고 단기예측모형의 개선방향과 시사점 등을 제시하였다.

2. 단기예측모형 방법론

GDP 성장률의 단기예측모형은 분기자료를 이용한 ARIMA, VAR 등 시계열모형, 월별자료를 분기자료와 연계시켜 예측하는 연계식모형(bridge equation model), 다수의 월별 자료가 지닌 공통정보를 소수의 주요인으로 요약하여 GDP를 예측하는 요인(factor)분석모형 등을 들 수 있다.

분기자료를 이용한 ARIMA, VAR 등의 시계열모형은 복잡한 경제이론에 근거한 회귀분석과 달리 손쉽게 모형을 추정하여 예측에 활용할 수 있다는 장점이 있다. 그러나 월별자료 등 속보성 있는 자료를 이용하지 못할 뿐만 아니라 경제 변수간의 유기적 관계 등을 고려하기 어렵다는 단점이 있어 이를 단독으로 연구하기보다는 대부분의 연구에서 다른 예측모형의 비교대상으로 분석되어 왔다. 국내에서는 자기회귀 및 벡터자기회귀모형을 통해 거시경제변수의 예측성과 자체를 비교한 이진희·김덕파(2014)의 연구가 대표적이다. 자기회귀 및 벡터자기회귀 뿐만 아니라 모형 80개를 결합하여 얻어진 14개 결합모형을 비교하여 경제성장률과 인플레이션, 이자율 등의 예측성과를 분석하였다.

분기자료를 이용한 시계열모형과 달리 월별지표를 이용하여 분기지표를 예측하는 경우 각 지표간 관측 주기가 달라서 발생하는 혼합주기(mixed frequency) 자료 문제가 나타나게 된다. 관측된 월별지표로 분기지표인 GDP를 예측하는 과정에서 생기는 문제는 월별지표와 분기지표를 연결하는 브릿지 모형(bridge model)을 이용하거나, 저빈도 자료를 고빈도 자료로 전환하여 월별 GDP를 추정하는 방법을 이용할 수 있다.

대표적인 연구로 Baffigi et al.(2003)은 연계식모형(bridge model)을 고안해 월별지표를 이용하여 GDP를 예측하였다. 월별자료에 포함된 정보와 분기 GDP성장률을 연결(bridge)하는 모델로 GDP와 주요 지출 부문간 연계식 모형을 구축하여 유로지역의 GDP를 예측하였다. Diron(2008)은 분기 GDP를 예측하기 위하여 다양한 월별 데이터를 이용한 8개의 bridge equation을 설정하였다. 산업생산, 건설액, 소매판매, 자동차 등록대수 등 경제지표를 이용한 모형,

ESI(Economic Sentiment Indicator), 소비자신뢰지수 등 서베이 자료를 이용한 모형, 실질실효환율, 금리 스프레드, 물가지수 등 금융지표를 이용한 모형, OECD 선행지수를 포함한 모형 등을 통해 GDP를 예측하였다. 월별로 결측치가 나타나는 경우에는 AR(6)를 이용하여 처리하였다. 국내 연구로 김양우·이공희·장동구(1997)는 GDP 구성항목 또는 GDP와 동행성이 높은 산업생산, 도소매판매, 비농업취업자수 등 주요변수에 대한 예측방정식을 먼저 도출하였다. 이어서 이들 구성항목의 합으로 GDP성장률을 잠정적으로 구한 후 이를 실제 GDP에 회귀시키는 연계식(bridge equation)을 이용하여 GDP성장률을 예측하였다.

이와 같이 지출부문과 GDP 총량 간의 연계식으로 모형을 구축하는 고전적인 연계식모형의 경우 한정된 변수를 이용하여 분석하기 때문에 이용 가능한 월별지표들의 정보를 충분하게 이용하기 어렵다는 단점을 가지고 있다. 즉, 다중공선성이나 자유도의 제약 등으로 인해 모형에 포함될 수 있는 변수가 한정적일 수 있다. 이러한 한계를 극복하기 위해 최근에는 보다 광범위한 월별데이터를 이용하여 요인을 추출하고 이를 이용하여 GDP를 예측하는 요인(factor)모형들이 이용되고 있다. 본 연구에서는 요인분석 방법인 주성분분석을 통한 공통요인 추정 방법과 공통요인의 동태적(dynamic) 특성을 이용한 동태요인분석을 통해 GDP 성장률을 예측하였다.

주성분분석을 통한 공통요인(common factor) 추정은 다변량 자료(multivariate data)에 포함된 많은 정보들의 손실을 최소화하여 낮은 차원의 데이터로 축약하는 방식이다. 변수들간의 분산-공분산 관계를 이용하여 모든 정보의 특성을 최대한 반영하면서 상호독립적인 주성분(principal component)⁵⁾을 찾는 분석 방법이다. 특히 단순 회귀분석을 이용할 경우 경제변수들 간에 발생할 수 있는 다중공선성과 내생성 문제를 해결⁶⁾하면서 변수들의 정보를 활용할 수 있다는 장점이 있을 뿐만 아니라 임의적인 설명변수 선택의 문제 등을 피할 수 있

5) 새롭게 추출된 주성분변수들은 서로 독립($Cov(Y_i, Y_k) = 0$)이며, 원래의 변수와는 선형 결합을 나타냄

6) 남준우·이한식(2005)에 따르면 변수들간에 다중공선성이 높을 때에는 회귀계수 추정량의 분산이 커지게 되고 분산이 커지면 회귀계수의 추정치가 실제 모수의 값과 큰 차이를 나타낼 가능성이 높을 뿐만 아니라 표본 선택에 따라 추정치의 값이 달라질 가능성이 커짐

다. 주성분분석은 소수지표만을 활용하여 신뢰도가 높은 결과를 얻을 수 있어 제한된 변수의 효율성 제고에 의의를 둘 수 있다(정갑영·강전은(2007)).

동태요인모형은 주성분분석을 이용하여 공통요인을 추정한 후 상태공간모형에서 칼만필터(Kalman filter)⁷⁾를 적용하여 결측치(F_t)를 추정하는 2단계 방법이다. 동태요인모형은 주성분분석과 마찬가지로 월별지표가 보유하고 있는 광범위한 정보를 공통요인으로 축약함으로써 분석의 효율성을 가져올 수 있다. Giannone et al.(2008)은 동태요인모형이 공통요인의 동태적 특성을 이용하여 결측치를 처리함으로써 단순 주성분분석보다 효율성이 향상될 수 있다고 주장하였다. 즉, 공통요인의 동태적 특성을 이용하여 자료주기, 공표시점 차이 등을 내생화 할 수 있는 특징이 있다(박강우·이중식(2009)). 다만 예측이 경제이론보다는 통계적 속성에 의존하고 있다는 한계에도 불구하고 적시에 효율적으로 자료를 축약하는 장점 등으로 인해 경제지표 구축(index construction) 및 단기예측(nowcasting & Forecasting)에 광범위하게 사용되고 있다.

3. 기존 연구 검토

주성분분석을 통해 경제변수를 예측한 국외 연구로는 Arnoštová et al.(2010), Mehrotra et al.(2011), Mingione et al.(2011), Hara et al.(2013) 등이 있으며 주성분분석을 사용하여 각각 체코, 중국, 일본의 GDP를 예측하였다.

Arnoštová et al.(2010)은 체코의 월간 지표를 활용한 분기 GDP 단기 예측모형의 예측력을 비교 및 평가하였다. 총 6개 모형의 표본 외 예측(Out of sample prediction)의 예측력을 비교한 결과 정적 주성분모형(Static Principal Components Model)을 통한 방식이 가장 우수한 예측력을 보였다. 구체적으로 해당 모형은 주성분분석을 통해 추출된 공통요인을 분기자료로 변환한 후 이를 GDP와 연계하여 연계 방정식의 선형회귀 추정을 통해 단기 GDP를 예측하였다.

7) 칼만필터는 추정된 계수(parameter)에 업데이트 된 정보를 반영하여 계수의 추정치를 업데이트 하는 축차적(recursive) 과정을 의미

Mehrotra et al.(2011)은 요인모형(Factor Model)을 통해 중국의 GDP성장률을 예측하였다. 먼저 주성분분석을 통하여 공통요인을 추출 후 GDP와의 선형회귀분석을 통해 유의 수준 10%를 만족하는 공통요인을 선별하였다. 선별된 공통요인을 바탕으로 예측된 GDP를 벤치마크 모형과 비교한 결과 유사한 예측력을 나타냈다. 반면, 아시아 및 글로벌 금융 위기와 같은 중국 경제의 충격이 발생한 기간 중 해당모형의 예측은 벤치마크 모형과 다소 차이가 있는 것으로 나타났다.

Hara et al.(2013)은 일본의 월별 지표를 이용하여 분기 GDP 예측을 시행하였다. 구체적으로 정적요인모형(Static Factor model)의 분석 절차에 따라 먼저 GDP와 연관성이 높은 다수의 경제지표들의 주성분분석을 통해 추출된 공통요인과 산업생산지수, 서비스업활동지수 등을 포함한 선형회귀모형으로 GDP를 추정하였다. 추정된 모형을 바탕으로 예측을 시행한 결과 공표시점 6주 전 예측에서 해당모형의 예측력이 벤치마크 모형인 ESP 모형⁸⁾의 예측력보다 우수한 것으로 나타났다.

국내에서 주성분분석을 사용하여 경제변수를 예측한 연구로는 김치호·김현정(2016), 김종화·이중식(2005)과 정갑영·강전은(2007)의 연구가 있으나, 주성분분석을 단독으로 사용한 연구보다는 동태요인모형 분석 과정에서 공통요인 추출을 위해 주성분분석을 사용한 연구가 대다수이다.

김치호·김현정(2016)은 당해분기 GDP성장률 예측을 위해 경제활동별 부가 가치를 합산하여 GDP 성장률을 추정하는 고전적 브릿지 모델뿐만 아니라 주성분분석과 주파수 영역 추출법을 적용하여 공통성분을 추출한 후 GDP성장률을 추정하는 동적인자모형(Dynamic Factor Model)⁹⁾을 통해 경제지표들 간 내재된 공통성분을 추출하여 GDP성장률을 추정하는 모형을 추정하였다. 분석결과 주파수 영역 추출법을 이용한 모형이 벤치마크 모형인 AR모형에 비해 추정오차가

8) 일본경제연구센터(JCER)가 전문가 40명의 의견을 토대로 작성한 예측조사

9) 김치호·김현정(2016)의 동적인자모형은 주성분분석과 주파수 영역 추출법을 이용하여 추출된 공통인자들의 시차변수들을 고려한 모형으로 구성지표 중 기준변수를 설정하고 나머지 구성변수들과 기준변수들 간의 시차상관계수 등을 구하는 방식으로 본고에서 다루는 동태요인분석과는 차이가 있음

작게 나타났다.

김종화·이중식(2005)은 주성분분석을 통하여 우리나라의 물가상승률을 예측하였다. 구체적으로 주성분분석을 통하여 실물, 해외 물가 및 금융시장부문 42개 개별 지표의 공통요인을 추출하였다. 추출된 공통요인을 바탕으로 자기회귀모형(Auto Regressive Model), 임의보행모형(Random Walk Model), 요인모형(Factor Model)을 비교한 결과 요인모형이 다른 모형들에 비하여 표본 외 예측에서 더 우수한 예측력을 보이는 것으로 나타났다. 반면, 임의보행모형의 경우 우리나라 물가상승률 예측 시 예측오차가 지나치게 큰 것으로 나타났다.

정갑영·강전은(2007)은 주성분분석을 통하여 북한의 국민총소득(GNI)을 추정하였다. 먼저 북한이 발표하는 20개의 실물지표와 과거의 추정치와의 주성분분석을 통해 공통요인을 추출하였다. 다음으로 추출된 공통요인과 GNI와의 관련성을 확인하기 위해 선형회귀분석을 시행하였다. 1990년부터 2005년까지 모형을 통해 추정된 결과를 한국은행에서 발표한 GNI와 비교한 결과 해당 모형의 설명력이 95.3% 이상으로 나타났다.

[표 1] 주성분분석을 활용한 기존 연구

저자(연도)	주요 분석 모형	주요 내용
Arnoštová et al. (2010)	정적주성분모형 (Static Principal Components Model)	- 체코의 월간지표를 활용한 분기 GDP 단기 예측 모형 평가 - 해당 모형이 표본 외 예측에서 가장 우수한 예측력을 보임
Mehrotra et al. (2011)	요인모형 (Factor Model)	- 추출된 공통요인의 선형회귀분석을 통해 공통요인을 선별 - 요인모형을 통해 중국의 GDP 예측 후 벤치마크 모형과 비교한 결과 유사한 예측력을 보임
Hara et al. (2013)	정적요인모형 (Static Factor Model)	- 일본의 월별 지표를 활용한 분기 GDP 예측을 시행 - 공표시점 6주 전 예측에서 벤치마크 모형(ESP)보다 우수한 예측력을 보임
김치호, 김현정 (2016)	동적인자모형 (Factor Model)	- 경제활동별 부가가치를 합산하여 GDP 성장률 추정 및 공통요인 추출 후 GDP 성장률 예측
김종화, 이중식 (2005)	요인모형 (Factor Model)	- 우리나라 물가상승률을 예측 및 비교 - 해당 모형이 표본 외 예측에서 벤치마크 모형(AR, RW)보다 우수한 예측력을 보임
정갑영, 강전은 (2007)	주성분분석 (Principal Components Analysis)	- 북한의 국민총소득(GNI)을 추정 - 선형회귀분석을 통해 한국은행에서 발표한 북한의 GNI와 비교한 결과 해당 모형의 설명력이 95.3%로 나타남

주: 해당 연구들은 모형의 공통요인 추출 과정에서 주성분분석을 활용함

공통요인 추정과 동태요인모형을 이용하여 GDP를 예측한 대표적인 연구로는 Barhoumi et al.(2008), Giannone et al.(2008), Doz et al.(2011) 등이

있으며 국내 연구로는 박강우·이중식(2009), 최병재·한진현(2014)의 연구가 있다.

Barhoumi et al.(2008)은 10개의 유럽 국가와 유로 지역 전체의 실질 GDP 성장에 대한 단기 예측을 위해 주성분분석, 동태요인분석, AR, VAR 등 다수의 모형을 비교 및 평가하였다. 분석 결과 월별 자료를 활용한 모형이 분기 자료만을 활용한 모형보다 우수한 예측력을 보이는 것으로 나타났으며 그 중 동태요인모형(Dynamic Factor Model)이 가장 우수한 예측력을 보이는 것으로 확인되었다.

Eickmeier and Ziegler(2008)는 메타분석을 통해 동태요인모형(Dynamic Factor Model)을 이용한 문헌들의 예측력을 비교 및 평가하였다. 분석 결과 동태요인모형의 예측성도가 다른 모형에 비해 우수하다는 것을 밝혀냈다. 구체적으로 동태요인모형의 경우 영국의 변수들보다 미국의 변수들에 대해, 유로 지역 생산량보다 미국 생산량에 대해, 미국 물가상승률보다 유로 지역 물가상승률에 대해 더 나은 예측결과를 나타내는 것으로 확인되었다. 더불어 공통요인 추출 시 가용되는 자료의 개수가 많을수록 상대적인 요인에 대한 예측성도에 긍정적인 영향을 미치는 것을 밝혀냈다.

Giannone et al.(2008)은 초단기예측(nowcasting) 모형을 통한 실질 GDP 예측 방법을 제시하였다. 구체적으로 미국의 1982년부터 2005년까지의 자료를 활용하여 주성분분석을 통한 공통요인 추정 후 칼만필터를 적용한 동태요인모형을 분석하였다. 분석결과 가용 자료가 많아지는 분기 말 시점에서 모형의 예측력이 개선되는 것으로 확인되었고 고용지표와 설문조사자료가 단기 GDP 예측력을 크게 개선시킨 것으로 확인되었다.

Barhoumi et al.(2010)은 프랑스의 월간지표를 이용한 분기 GDP 예측 모형들을 비교 분석하였다. 구체적으로 동태요인모형 활용 시 고려되는 공통요인 추정방법에 대하여 집계(aggregate) 자료 20개와 이를 개별로 분할(disaggregate)한 120개 자료의 예측력을 비교하였다. 분석 결과 Stock and Watson(2002a)의 모형과 나우캐스팅 모형을 사용할 경우 집계 자료를 사용할 때와 개별 분할

자료를 사용할 때의 예측력에 큰 차이가 없는 것으로 나타났다.

Doz et al.(2011)은 다변수 시계열 패널에서 동적근사요인모형(Dynamic approximate factor model)¹⁰⁾의 2단계 추정을 통해 일치성을 확인하였다. 구체적으로 동적근사요인모형의 첫 번째 단계에서 주성분분석을 통해 공통요인 추출 후 선형회귀식을 이용하여 모형의 매개 변수 추정을 시행하였다. 두 번째 단계에서는 칼만 필터를 이용하여 요인을 추정하는 방식으로 연구를 수행하였다. 일치성 확인을 위해 몬테카를로 시뮬레이션을 시행한 결과 표본의 크기가 작을수록 요인에 대한 일치성이 높아짐이 확인되었다.

국내연구로는 동태요인모형을 이용한 단기 GDP 예측을 시행한 박강우·이중식(2009)의 연구가 대표적이다. 먼저 주성분분석을 활용하여 51개 금융 및 실물 지표의 공통요인을 추출한 후 추정된 공통요인을 바탕으로 벡터자기회귀(VAR)모형을 구축하여 모수를 추정하였다. 다음으로 추정된 값들을 바탕으로 상태공간모형으로 전환 후 칼만필터를 통해 단기 GDP를 예측하였다. 해당 모형의 표본 외 예측력을 ARIMA, VAR, 연계식 예측모형과 비교한 결과 동태요인모형이 ARIMA 모형보다 우수하고 VAR모형과 연계식 모형과는 비슷한 예측력을 가지고 있는 것으로 나타났다.

최병재·한진현(2014)은 월별변수를 이용하여 시차변수 모형, 주성분분석, 동태요인분석 등 전망모형별 전망성과를 벤치마크 모형인 AR(4)모형과 비교하였다. 분석결과 월별자료를 이용한 당분기 및 단기 전망은 벤치마크 모형에 비해 예측오차가 더 작아 월별자료들이 전망에 유용한 정보를 포함하고 있음을 보였다.

요인분석 방법은 아니지만, 가장 최근에 이루어진 월별지표를 이용한 GDP 추정모형에 관한 연구로 김덕파·강성진(2020)의 연구를 들 수 있다. 외생변수를 포함한 자기회귀모형(ARX)을 기본모형으로 15개 월별자료와 자기회귀 차수, 추정법 등을 각각 다르게 적용하는 240개의 예측모형을 구축하였다. 연구결과 대부분의 예측모형에서 축차표본보다 순차이동표본이, 월별 자료 중 산업생산지수와 경기종합지수가 우수한 예측력을 나타내었다.

10) 1900년대 초 확률적 오차항을 허용하지 않는 Exact Factor Model과 구분하여 고유성분(idiosyncratic components)에 대해 약한 자기상관이 존재할 수 있음을 가정한 모델

[표 2] 동태요인모형을 활용한 기존 연구

저자(연도)	주요 분석 모형	주요 내용
Barhoumi et al. (2008)	동태요인모형 (Dynamic Factor Model)	- 10개 유럽 국가와 유로 지역 전체의 단기 실질 GDP 예측 - 월별 자료를 활용한 모형이 더 우수한 예측력을 보임
Eickmeier and Ziegler (2008)	동태요인모형 (Dynamic Factor Model)	- 동태요인모형을 이용한 경제 전망 문헌의 예측력 비교 - 해당 모형은 미국의 변수 예측 시 더 우수한 것으로 나타남
Giannone et al. (2008)	동태요인모형 (Dynamic Factor Model)	- 공통요인 추출 후 칼만 필터를 적용한 동태요인모형 활용 - 고용지표 및 설문조사자료가 예측력을 크게 개선시킴
Barhoumi et al. (2010)	동태요인모형 (Dynamic Factor Model)	- 프랑스의 월간지표를 이용한 분기 GDP 예측 모형을 비교 - 단기 예측에서 추차 및 롤링 예측을 결합한 모형의 예측력이 개선됨
Doz et al. (2011)	동태근사요인모형 (Dynamic Approximate Factor Model)	- 추출된 공통요인으로 선형회귀식을 통해 매개 변수를 추정 후 칼만필터로 예측 시행 - 표본의 크기가 작을수록 일치성이 높아짐을 확인
박강우, 이중식 (2009)	동태요인모형 (Dynamic Factor Model)	- 추출된 공통요인으로 벡터자기회귀(VAR)모형을 통해 모수를 추정 후 칼만필터로 예측 시행 - 해당 모형이 ARIMA 모형보다 우수한 것으로 나타남
최병재, 한진현 (2014)	동태요인모형 (Dynamic Factor Model)	- 월별지표를 이용한 모형과 AR(4)모형을 비교한 결과 월별 자료를 이용한 단기 예측의 예측 오차가 더 낮은 것으로 나타남

주: 해당 연구들은 모형의 공통요인 추출 과정에서 주성분분석을 활용함

II. 단기 실질GDP 예측 모형설정

1. 분석모형

월별지표를 이용하여 분기 GDP를 예측하는 방법은 Giannone et al.(2004), Giannone et al.(2005) 등 기존연구들에서 분석되어온 연계모형(bridge equation)을 차용하였다. 먼저 (식1)과 같이 월별 경제변수 $X_t = (x_{1t}, \dots, x_{nt})'$, $t = 1, \dots, T$ 들의 공통요인을 추출한 후(식1) 월별로 추출된 공통요인을 분기별로 변환한 데이터(aggregate(f_t^Q))를 이용하여 GDP성장률(y_t^Q)과의 회귀분석을 수행하는 방식이다(식2).

$$x_t = \Lambda f_t + \xi_t \quad (\text{식1})$$

$$y_t^Q = \mu + \beta' f_t^Q + \epsilon_t \quad (\text{식2})$$

본 연구에서는 두 가지 방법으로 GDP성장률을 추정하고자 한다. Stock and Watson(2002)과 Barhoumi(2008)의 주성분분석모형(Simple Principal Component)과 Doz et al.(2007)과 Giannone et al.(2008)의 동태요인모형이다.

가. 주성분분석모형

주성분분석(Principal Component) 방법을 이용하여 GDP를 예측하기 위해서 먼저 경제변수로부터 공통요인을 추출하였으며, 추정된 공통요인과 예측하려는 변수인 GDP성장률간의 선형관계를 추정하여 예측방정식을 도출하였다.

주성분분석은 변수들간의 분산-공분산 관계를 이용하여 상관관계가 높은 변수들의 선형결합을 추출하는 방법으로 본 보고서에서는 p개의 경제변수들을 k개

의 공통요인으로 축소시켜($k < p$) 다차원의 경제변수를 낮은 차원의 데이터로 축약하였다.

$p \times 1$ 인 경제변수 $X = (x_1, x_2, \dots, x_p)'$ 가 평균 μ , 공분산행렬 Σ 를 갖는 확률벡터일 때, 이들 p 개 설명변수들의 선형결합으로 이루어진 F 행렬은 다음과 같이 표현된다.

$$\begin{aligned} F_1 &= a_1'X = a_{11}x_1 + a_{21}x_2 + \dots + a_{p1}x_p \\ F_2 &= a_2'X = a_{12}x_1 + a_{22}x_2 + \dots + a_{p2}x_p \\ &\vdots \\ F_p &= a_p'X = a_{1p}x_1 + a_{2p}x_2 + \dots + a_{pp}x_p \end{aligned}$$

F_i 의 분산 공분산 행렬은 다음과 같이 표현될 수 있다.

$$Var(F_i) = a_i' \Sigma a_i, \quad i = 1, 2, \dots, p$$

$$Cov(F_i, F_k) = a_i' \Sigma a_k, \quad i, k = 1, 2, \dots, p$$

주성분분석은 이와 같이 주어진 경제변수인 $x_1, x_2, \dots, x_p$ 의 정보를 최대한 포함할 수 있도록 주성분요인 $F_1, F_2, \dots, F_p$ 의 분산을 극대화 하는 가중치 행렬을 찾는 문제이다. 각각의 주성분 요인은 서로 상관관계가 없는 선형결합(uncorrelated linear combination)이 되며 첫 번째 주성분요인은 가장 큰 분산을 가지는 선형결합이 된다. 즉, $F_i = a_i'X$ 의 선형결합 중 분산 $Var(a_i'X)$ 을

극대화 시키면서 F 가 X 의 가중합이라는 의미를 갖기 위한 제약조건인 $a_i' a_i = 1$ 을 만족시키는 주성분을 찾게 되며 (식3)과 같이 표현될 수 있다¹¹⁾.

$$\max \{ a' \Sigma a - \lambda (a' a - 1) \} \quad (\text{식3})$$

이는 $x_1, x_2, \dots, x_p$ 의 분산공분산행렬의 고유값(eigenvalue)과 고유벡터(eigenvector)를 구하여 도출하게 되며, (식3)의 일계조건으로부터 구한 p 개의 해인 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ 가 이들의 고유값(eigenvalue)과 같으며 $a_{1p}, a_{2p}, \dots, a_{pp}$ 가 고유벡터(eigenvector)가 된다. 각 고유값은 주성분의 분산과 일치하게 되며, 가장 큰 고유값을 가지는(분산이 가장 큰) F_t 가 제1 주성분이 된다. 이와 같이 새롭게 추출된 p 개의 주성분변수들은 서로 독립 ($Cov(a_i' X, a_k' X) = 0$)인 특성을 갖는다. 위와 같이 추출된 주성분요인들은 원래 경제변수들이 지니고 있는 특성을 반영하면서 상호독립적이기 때문에 다중공선성과 내생성의 문제를 해결할 수 있어 다수의 경제변수들을 분석에 포함시킬 수 있다.

이러한 조건에서 구한 p 개의 주성분 요인 중 k 번째 요인의 설명력은 다음과 같이 표현될 수 있으며, 이는 각 주성분이 원래 변수의 정보를 설명하는 정도를 나타내준다.

$$\text{proportion} = \frac{\lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad k=1, 2, \dots, p$$

한편, 주성분분석을 통해 p 개의 경제변수들의 공통요인을 추정하게 되면 p 개의 주성분 요인이 추출되게 된다. 그러나 추출된 p 개의 공통요인이 모두 의미

11) $\max Var(a_i' X)$ subject to $a_i' a_i = 1$

를 가지는 것은 아니기 때문에 추출된 적정한 k의 수를 결정하여 공통요인을 축약할 필요가 있다. k의 수를 결정하는 방법으로는 카이저 기준, 설명력 기준, Scree 도표, Parallel 분석방법 등이 이용가능하다.

카이저(Kaiser, 1960) 기준은 고유값이 1 이상인 주성분을 선택하는 방법이다. 분석을 통해 도출된 각 주성분들은 각 요인이 설명할 수 있는 설명력(분산)을 의미하며 분석대상 변수의 개수만큼 추출된다. 주성분분석 시 각 변수들은 평균이 0, 분산이 1이 되도록 표준화하기 때문에 각 주성분들의 고유값의 합은 분석에 포함된 변수의 개수(전체 변수의 분산)와 같게 되며, 고유값이 1보다 작은 주성분들은 한 개의 변수도 설명하기 어려운 것으로 해석된다. Stevens(1986)은 특히 분석대상 변수가 30개 미만으로 많지 않은 경우 이러한 방식이 적절하다고 분석하였다.

두 번째로는 Scree plot을 이용하는 방법이다. Scree plot은 고유값이 내림차순으로 정렬된 도표로서 감소폭이 급격하게 작아지는 지점(break 또는 elbow)까지의 요인을 선택하는 것으로 break 전에 나타나는 주성분까지를 의미 있는 것으로 고려하는 방법이다. Stevens(1986)은 표본이 200개 이상으로 큰 경우 Scree 도표를 이용한 방법이 합리적인 결과를 제공하는 것으로 분석하였다.

세 번째로는 설명력 기준을 사용하는 방법으로 전체 분산의 5% 이상 ($proportion = \frac{\lambda_k}{p} > 5\%$)을 차지하는 주성분들을 선택하는 분산비율 방법과 누적설명력이 70~80% 이상이 되는 주성분들을 선택하는 방법 등이 있다.

네 번째로는 자료에서 추정된 eigenvalue와 임의로 추출한 eigenvalue의 값을 비교하여 임의 추출한 값보다 큰 값을 가지는 요인의 개수로 결정하는 Parallel 분석방법이다. 즉, 위에서 설명한 Scree Plot과 임의로 추출한 고유값을 비교하는 방법이다.

공통요인이 추정된 후에는 Stock and Watson(2002), Bahoumi et al(2007)을 따라 추출된 공통요인 F_t 와 예측하고자 하는 변수인 GDP성장률 Y_t 와의 선형관계인 (식4)를 추정하고 공통요인 F_t 의 결측치 및 예측치를 이용하여

(식5)를 통해 h기 후의 GDP성장률 y_{t+h} 를 예측한다. 이때 추정된 공통요인들은

$$F_{j,t}^Q = \frac{1}{3} \sum_{s=0}^2 F_{j,t-s} \text{ 와 같은 방법으로 분기 변환하였다.}$$

$$Y_t^Q = \rho Y_{t-1}^Q + \beta \widehat{F}_t^Q + \epsilon_t \quad (\text{식4})$$

$$Y_{t+h}^Q = \rho Y_t^Q + \beta \widehat{F}_{t+h}^Q + \epsilon_{t+h} \quad (\text{식5})$$

나. 동태요인모형

동태요인모형은 주성분분석을 이용하여 공통요인을 추출¹²⁾한 후 칼만필터(Kalman filter)를 이용하여 공통요인 F_t 를 추정하는 2단계 방법이다. 즉, 추정된 공통요인으로 (식6)의 선형방정식과 (식7)의 VAR 모형을 구성할 수 있다.

먼저, 1단계로 관측 가능한 N개의 정보변수 X_t 와 r개의 공통요인 F_t 는 (식6)과 같이 표현된다. 즉, 공통요인에 의한 부분인 ΛF_t 와 개별요인(idiosyncratic component)인 ξ_t 로 구분될 수 있으며, (식6)을 이용하여 요인계수 Λ 와 잔차항(ξ_t)의 행렬 $R(\Sigma_\xi)$ 을 추정한다.

이어서 주성분분석을 통해 추출된 공통요인 F_t 의 동학을 고려한 VAR 모형을 추정하여 Λ 와 잔차항의 행렬 $Q(\Sigma_{u_t})$ 추정치를 도출한다.

$$X_t = \chi_t + \xi_t = \Lambda F_t + \xi_t, \quad \xi_t \sim iidN(0, R) \quad (\text{식6})$$

$$F_t = \Lambda F_{t-1} + u_t, \quad u_t \sim iidN(0, Q) \quad (\text{식7})$$

12) Bai(2003), Bai and Ng(2002), Forni et al.(2005), Stock and Watson(2002) 등의 연구에서 관찰된 변수들의 주성분분석을 통해 관측되지 않은 공통요인 F를 추정하고 있음

다음 단계로 관측방정식인 (식6)과 상태방정식인 (식7)에서 요인계수 Λ , 잔차항(ξ_t)의 행렬 R , 모수 A , Q 의 추정치가 도출되면 공통요인을 비관측요소로 하는 상태공간모형(state space representation)에 칼만필터를 적용하여 최종 $\hat{F}_t$ 를 추정하게 된다. 즉, 칼만필터 방법을 통해 t 기의 이용 가능한 정보에 기초하여 상태변수 F_t 에 대한 최적추정량을 축차적으로 계산하여 계수 추정치를 업데이트하게 된다. 즉, 칼만필터는 Λ , Σ_ξ , A , Σ_{u_t} 가 주어졌을 때 $\hat{F}_t$ 를 1기에서 T 기까지 추론해가는 과정이다. 칼만필터는 직접 관측할 수 없는 변수를 추출하는 방식으로 공통요인의 다이내믹스(dynamics)를 이용하고 과거의 정보와 새로운 측정값을 활용하여 결측치를 추정할 수 있는 효율적인 방법으로 단순 주성분분석에 비해 효율성을 향상시킬 수 있다는 장점이 있다.

(식6)~(식7)과 같은 확률계수모형의 계수 추정 시 비관측 요소인 F_t 를 도출하는 칼만 필터의 recursive algorithm은 다음과 같은 과정을 따른다. 주어진, 예측과정(prediction)과 수정과정(updating)을 통해 구해진 Λ , R , A , Q 하에서 F_t 를 추정하게 된다.

$$F_{t|t-1} = AF_{t-1|t-1} \quad (\text{식8})$$

$$P_{t|t-1} = AP_{t-1|t-1}A' + Q \quad (\text{식9})$$

$$X_{t|t-1} = \Lambda F_{t|t-1} \quad (\text{식10})$$

$$\eta_{t|t-1} = X_t - X_{t|t-1} \quad (\text{식11})$$

$$f_{t|t-1} = \Lambda P_{t|t-1}A' + R \quad (\text{식12})^{13)}$$

13) 예측오차($\eta_{t|t-1}$)의 분산($f_{t|t-1}$)은 다음과 같이 도출됨.

$Var(\eta_t|\psi_{t-1}) = E[(X_t - X_{t|t-1})(X_t - X_{t|t-1})'|\psi_{t-1}] = E[(\Lambda(F_t - F_{t|t-1}) + \xi_t)(\Lambda(F_t - F_{t|t-1}) + \xi_t)']$
 이때, $X_t = \Lambda F_t + \epsilon_t$, $X_{t|t-1} = \Lambda F_{t|t-1}$ 이며, F 에 대한 예측오차와 ξ_t 에 대한 예측오차는 서로 독립
 이므로 교차곱의 예측오차는 '0'이 되므로 다음과 같이 정리할 수 있음.

먼저, 예측과정에서는 t 기에 $t-1$ 기까지 입수 가능한 정보를 이용하여 F_t 의 최적추정량인 $F_{t|t-1}$ 을 계산한다(식 8~12). $P_{t|t-1}$ 은 (식8)에서 구한 $F_{t|t-1}$ 의 분산행렬로 $F_{t|t-1}$ 의 분산($P_{t|t-1}$)이 커진다는 것은 $F_{t|t-1}$ 의 정보의 불확실성이 높다는 의미이다. (식10)의 $X_{t|t-1}$ 은 $t-1$ 기의 정보를 바탕으로 예측한 계수 Λ 를 이용한 예측치이며, (식11)의 $\eta_{t|t-1}$ 은 $t-1$ 기의 자료를 바탕으로 예측한 $X_{t|t-1}$ 과 관측치인 X_t 와의 차이를 나타내는 예측오차이다. 식(12)의 $f_{t|t-1}$ 은 예측오차의 분산을 나타낸다.

수정과정은 예측과정에서 구한 예측오차 $\eta_{t|t-1}$ 과 분산 $f_{t|t-1}$ 을 이용하여 상태변수 $F_{t|t}$ 와 그 분산인 $P_{t|t}$ 에 대한 추정치를 업데이트하는 과정으로 (식13)과 (식14)로 나타낼 수 있다. 업데이트된 수치들을 이용하여 (식8)~(식12)를 이용하여 $t+1$ 기의 예측치를 얻고 수정과정을 통해 이를 업데이트하는 과정을 반복하여 F_t 를 1기에서 T 기까지 추정한다.

$$F_{t|t} = F_{t|t-1} + K_t \eta_{t|t-1} \quad (\text{식13})$$

$$\text{단, } K_t = P_{t|t-1} \Lambda' f_{t|t-1}^{-1}$$

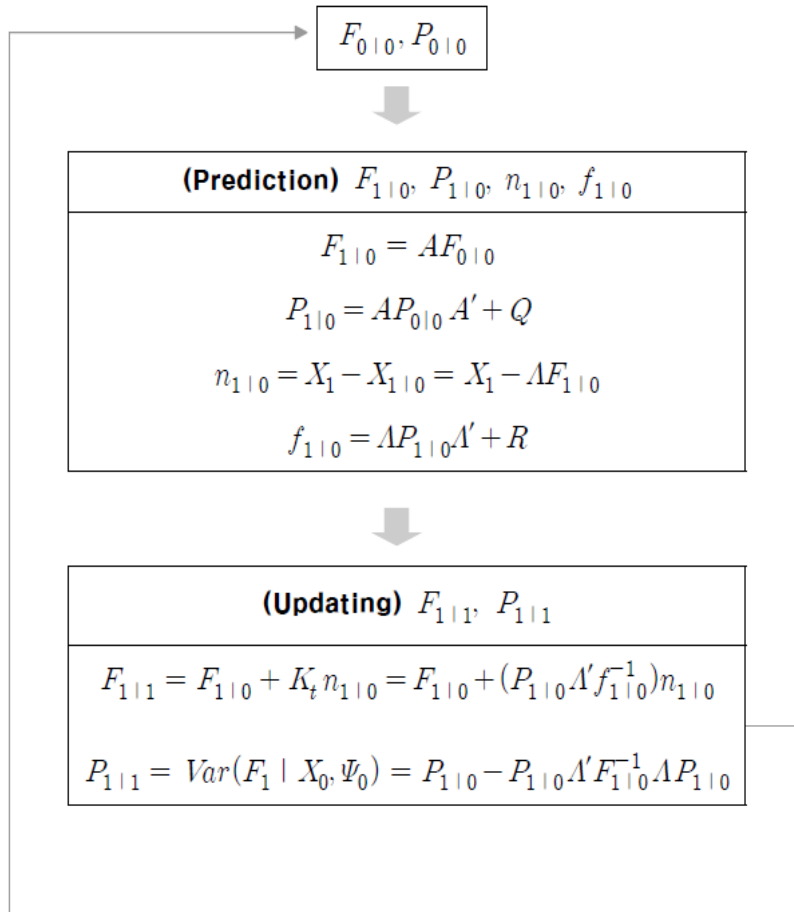
$$P_{t|t} = \text{Var}(F_t | X_{t-1}, \Psi_{t-1}) = P_{t|t-1} - P_{t|t-1} \Lambda' f_{t|t-1}^{-1} \Lambda P_{t|t-1} = P_{t|t-1} K_t \Lambda P_{t|t-1} \quad (\text{식14})$$

위와 같이 추정된 월별 공통요인을 분기 평균으로 전환하여 GDP성장률과의 회귀방정식 (식15)을 통해 예측치를 도출하게 된다.

$$Y_{t+h}^{q_0} = \rho Y_t^{q_0} + \beta \widehat{F}_t^{q_0} + \epsilon_{t+h} \quad (\text{식15})$$

$$\text{Var}(\eta_t | \psi_{t-1}) = E[\Lambda(F_t - F_{t|t-1})(F_t - F_{t|t-1})\Lambda'] + E(\xi_t \xi_t') = \Lambda P_{t|t-1} \Lambda' + R$$

[그림 3] Kalman filter 추론과정



2. 분석방법

가. 결측치 처리방법

분기데이터인 GDP를 예측하기 위해서는 t 분기에 포함된 m 월까지의 정보를 이용해야 한다. 즉, 월별 자료를 이용한 $t+h$ 기 GDP성장률을 전망하기 위해 t 분기 m 월의 정보를 이용하게 된다. 그러나 t 분기 중 m_1 (초월), m_2 (중월), m_3 (말월) 각각 월에 발표되는 지표의 차이가 발생하기 때문에 전망과정에서 월간지표별 공표 시차가 달라 발생하는 결측치를 생성할 필요가 있다. 예를들어, t 분기 m_1 말 시점에서 금리, 추가 등 금융변수는 m_1 까지의 자료를 모두 이용할 수 있으나, 물가와 노동부문 지표들은 $t-1$ 분기 m_3 의 자료가 이용가능하다. 경기, 산업활동, 통화량 등 2개월 후행하는 지표들의 경우 $t-1$ 분기 m_2 월의 자료까지 이용할 수 있다.

[그림 4] 월별자료의 공표시점



Barhoumi(2008)의 방식을 따라 월간자료 중 결측치가 발생할 경우 AR(p) 모형을 이용하여 값을 생성하였다. Barhoumi(2008)는 다변량 모형과 Stock and Watson(2003)이 제안한 EM 알고리즘 등을 이용하여 결측치를 생성한 결과 AR모형과의 차이가 거의 없었으며 AR모형을 통한 결측치 생성의 전망성보다 더 우수함을 확인하였다. t분기 m_3 말 시점에서 m_1 까지 발표되는 자료의 경우 AR(p)모형을 이용하여 m_2 와 m_3 를 생성하여 t분기 예측에 이용하였으며, t+1~t+h 분기까지 AR(p) 모형을 이용하여 월별지표의 결측치를 생성하여 전망의 정보변수로 활용하였다. 자기회귀 차수(p)는 아카이케 정보기준 (Akaike information criterion, AIC) 또는 베이지언 정보기준 (Bayesian information criterion, BIC)을 이용하여 설정하였으며, 허용 가능한 최대차수 $p_{\max}$ 를 설정하고 동 범위내에서 아래의 정보기준을 최소화하는 차수로 자기회귀 차수를 설정하였다.

$$AIC(p) = \log(\hat{\sigma}_\rho^2) + \frac{2(p+1)}{T-p_{\max}}$$

$$BIC(p) = \log(\hat{\sigma}_\rho^2) + \frac{\log(T)(p+1)}{T-p_{\max}}$$

단순 주성분모형의 경우 위와 같이 AR(p) 모형을 통해 결측치를 생성하였으나, 동태요인모형은 결측치를 처리하는 방식에 차이가 있다. 주성분분석 방법과 마찬가지로 동태요인분석 방법에서도 월별로 추정된 공통요인을 분기 평균으로 전환하는 과정에서 결측치를 처리하는 문제가 발생하게 된다. 그러나 단순 주성분분석 방법과 달리 동태요인분석에서는 칼만필터를 이용하여 공통요인의 예측치를 도출할 수 있다는 차이점이 있다. 즉, 자기회귀를 통해 공통요인의 동태적 특성을 분석하고 칼만스무딩(Kalman smoothing) 방법을 통해 결측치를 처리한다. 칼만필터를 이용하게 되면 공통요인의 지속성 및 매월 달라지는 정보의 이질

성(heterogeneity)을 고려한 가중치를 적용하게 되어 현재 및 과거 데이터들의 적절한 조합을 통한 결측치를 생성할 수 있다(Barhoumi(2008)).

공통요인에 칼만필터를 적용하기 위해 (식7)($F_t = \Lambda F_{t-1} + u_t, u_t \sim iidN(0, Q)$)과 같이 VAR 모형을 설정하여 주성분의 동태(dynamics)를 매개변수화(parameterize) 하였다. 관측방정식인 (식6)($X = \Lambda F_t + \xi_t$)은 공통요인에 의한 부분인 ΛF_t 와 개별요인(idiosyncratic component)인 ξ_t 로 구분될 수 있으며, 개별요인들의 측정오차는 다음과 같이 횡단면적으로 직교(모든 시차에서 상관관계가 없음)하며, 평균이 0, 공분산이 ψ_i 인 정규분포를 하는 특성이 있다.

$$E(\xi_{t|v_j} \xi_{t|v_j}) = \Psi_{t|v_j} = diag(\tilde{\psi}_{1,t|v_j}, \dots, \tilde{\psi}_{n,t|v_j}) \quad (\text{식16})$$

$$E(\xi_{t|v_j} \xi_{t-s|v_j}) = 0, s > 0 \text{ for all } v, j \quad (\text{식17})$$

$$\tilde{\psi}_{i,t|v_j} = \begin{cases} \psi_i & \text{if } x_{i,t|v_j} \text{ is available} \\ \infty & \text{if } x_{i,t|v_j} \text{ is not available} \end{cases}$$

먼저, 데이터 중 관측치가 가장 짧은 기간¹⁴⁾ 기준 이용 가능한 변수들을 대상으로 계수(parameter)들을 추정한다. 이와 같은 과정을 통해 추정된 계수(parameter)를 μ_θ 로, v_j 시점에서 수집된 정보집합을 Ω_{v_j} 라고 하면, 상태변수의 예측치와 분산은 다음과 같이 표현될 수 있다.

$$\hat{F}_{t|v_j} = E[F_t | \Omega_{v_j}; \mu_{\hat{\theta}}]$$

$$\hat{V}_{v_j} = E[(F_t - \hat{F}_t)(F_t - \hat{F}_t)'; \mu_{\hat{\theta}}]$$

14) $\min\{T_{1v_j}, \dots, T_{nv_j}\}$

v_j 시점에서 이용 가능한 정보를 이용하여 $\hat{F}_t$ 에 대한 최적예측치 및 이의 분산 행렬을 계산하며 예측방정식(prediction)과 수정방정식(updating)을 통해 F_t 에 대한 최적 추정량을 축차적으로 계산하게 된다.

한편 이러한 결측치 추정방법은 모형 설정 후 표본 외 예측을 위한 예측치 추정 과정에서도 동일하게 적용된다.

나. 예측력 비교 방법

예측력을 평가하기 위한 첫 번째 단계로 표본 내(in-sample) 자료를 이용하여 모형을 추정하고 추정된 모형으로 표본 외(out-of sample) 예측을 실시하여 실제 GDP성장률과 이를 비교하는 방법을 사용하였다.

표본 외 예측치와 실제치를 비교하기 위해 RMSE(Root Mean Squared Error)와 MAE(Mean Absolute Error)를 계산하여 이를 비교하였다. 또한, 표본 외 예측 전 구간에 대해 종합적으로 RMSE와 MAE를 산출하는 한편 각 분기 중 첫 번째 월, 두 번째 월, 세 번째 월 각각에 대해서도 예측력을 평가하였다.

RMSE(Root Mean Squared Error)는 실제값과 예측값의 차이인 오차들의 제곱 평균인 MSE(Mean Squared Error)에 제곱근을 씌운 값이며, MAE(Mean Absolute Error)는 실제값과 예측값의 차이인 오차들의 절댓값의 평균으로 계산된다. RMSE의 경우 제곱을 통해 절댓값이 1보다 큰 오차를 더욱 크게 만들어 특이치에 더 민감하게 반응하는 경향이 있다.

$\hat{y}_{T,1}, \hat{y}_{T,2}, \hat{y}_{T,3}$ 을 T분기 GDP성장률(y_T)에 대한 T분기 중 초월말, 중월말, 말월말의 예측치라 하고 H를 $T_e - h - T_s + 1$ 이라 하면, 전체 T_e 개의 자료 중 T_s 까지의 자료를 이용해 모형을 추정하여 h 기간 후의 예측을 수행하는 경우 RMSE와 MAE는 다음과 같이 정의된다.

$$\text{종합 RMSE} = \sqrt{[\sum (\hat{y}_{T,1} - y_T)^2 + (\hat{y}_{T,2} - y_T)^2 + (\hat{y}_{T,3} - y_T)^2] / 3H}$$

$$\text{초월 RMSE} = \sqrt{[\sum (\hat{y}_{T,1} - y_T)^2] / H}$$

$$\text{중월 RMSE} = \sqrt{[\sum (\hat{y}_{T,2} - y_T)^2] / H}$$

$$\text{말월 RMSE} = \sqrt{[\sum (\hat{y}_{T,3} - y_T)^2] / H}$$

$$\text{종합 MAE} = [\sum |\hat{y}_{T,1} - y_T| + |\hat{y}_{T,2} - y_T| + |\hat{y}_{T,3} - y_T|] / 3H$$

$$\text{초월 MAE} = [\sum |\hat{y}_{T,1} - y_T|] / H$$

$$\text{중월 MAE} = [\sum |\hat{y}_{T,2} - y_T|] / H$$

$$\text{말월 MAE} = [\sum |\hat{y}_{T,2} - y_T|] / H$$

분석에 사용된 자료들은 2000년 1월부터 2019년 12월까지이며 표본 외 예측 기간은 2015년 1/4분기부터 2019년 4/4분기까지의 20개 분기로 정하였다. 2015년 1월말에 2015년 1/4분기를 예측하는 경우의 주성분분석모형의 메커니즘은 다음과 같다. 첫째, 2000년 1월부터 2014년 11월¹⁵⁾의 데이터를 이용하여 주성분분석을 통해 공통요인 F_t 를 추정한다. 둘째, AR모형을 통해 2014년 12

15) 2015년 1월말 기준으로 시차가 2개월인 데이터로 인해 모든 데이터가 입수 가능한 시점이 2014년 11월

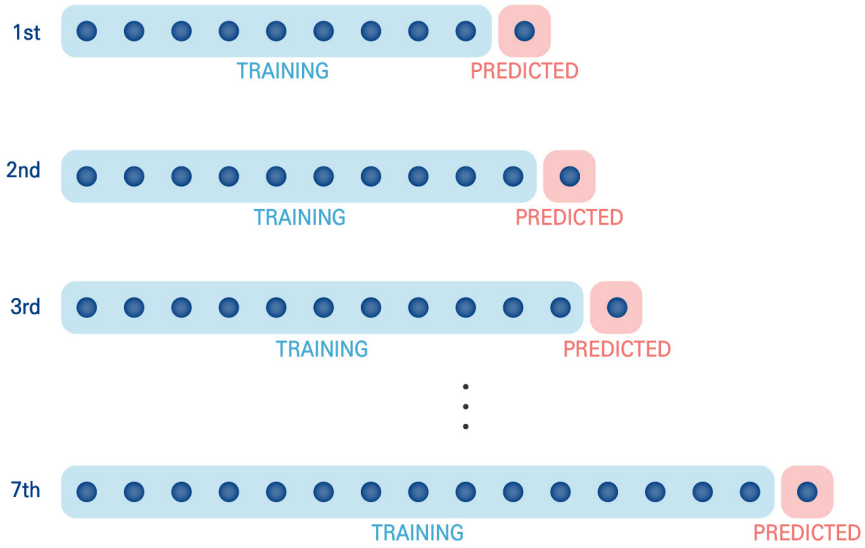
월~2015년 6월까지의 각 변수들의 결측치 및 예측치를 추정한다. 셋째, 추정된 변수로 공통요인을 구성하고, 이를 분기전환하여 2000년 1/4분기~2014년 4/4분기 GDP성장률과 회귀분석을 수행한다. 넷째, 추정된 회귀식에 공통요인의 예측치를 대입하여 GDP성장률을 추정하였으며 이를 실제 GDP성장률과 비교하였다.

2015년 1월말에 2015년 1/4분기를 예측하는 경우의 동태요인분석모형의 메커니즘은 다음과 같다. 첫째, 2000년 1월부터 2014년 11월의 데이터를 이용하여 주성분분석을 통해 공통요인 F_t 를 추정한다. 둘째, 칼만필터와 상태공간모형을 이용하여 2014년 12월~2015년 6월까지의 주성분의 예측치를 추정한다. 셋째, 추정된 공통요인을 분기전환하여 2000년 1/4분기~2014년 4/4분기 GDP 성장률과 회귀분석을 수행한다. 넷째, 추정된 회귀식에 공통요인의 예측치를 대입하여 GDP성장률을 추정하였으며 이를 실제 GDP성장률과 비교하였다.

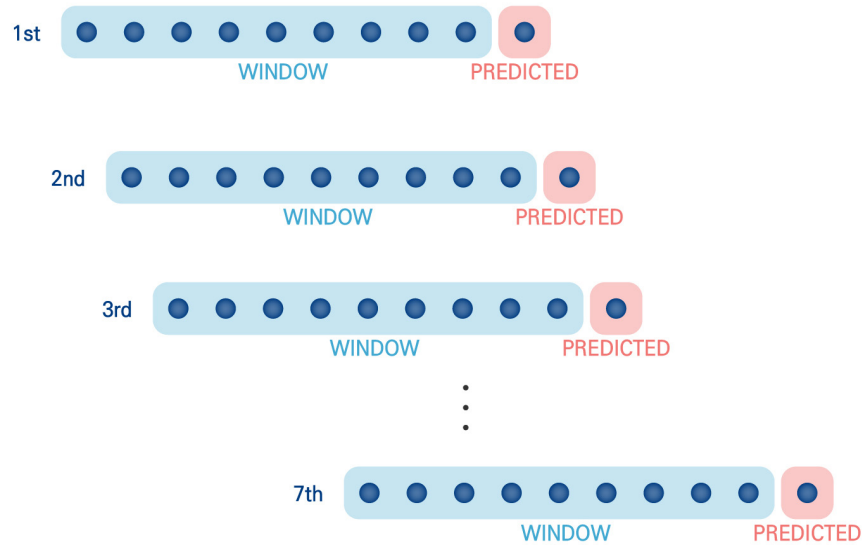
표본 외 예측기간을 설정할 때 일반적으로 사용되는 방법으로 축차예측(recursive forecast)과 롤링예측(rolling forecast) 방법이 있으며 본 연구에서는 두 가지 방법을 모두 추정하여 예측오차를 비교하였다. 롤링예측은 윈도우의 크기를 고정(표본의 수를 일정하게 유지)한 상태에서 시점을 이동하면서 예측치를 도출하는 방법이며, 축차예측은 처음 정해진 구간을 기본으로 하여 기간을 추가하면서 추정하여 예측하는 방식이다. Swanson and White(1997), Tashman(2000), 최병재·한진현(2014) 등 기존연구에서는 롤링예측 방식이 전망 성과를 보다 개선시키고 있음을 보였으며, 본 연구에서는 두 가지 방법을 모두 비교하였다.

이진희·김덕파(2014)에 따르면 축차예측방식의 경우 시계열이 안정적인 모습을 보일 때는 우수한 성과를 보일 것으로 예상되나, 구조변경 등 불안정성이 발생할 경우에는 그로부터 지속적인 영향을 받게 되는 단점이 있다. 따라서 분석대상 시계열이 안정적인 경우 축차예측방식이 유리하며, 구조변화가 포함된 시계열의 경우 롤링예측방식이 유리할 것으로 보인다.

[그림 5] 축차예측(recursive forecast) 방식



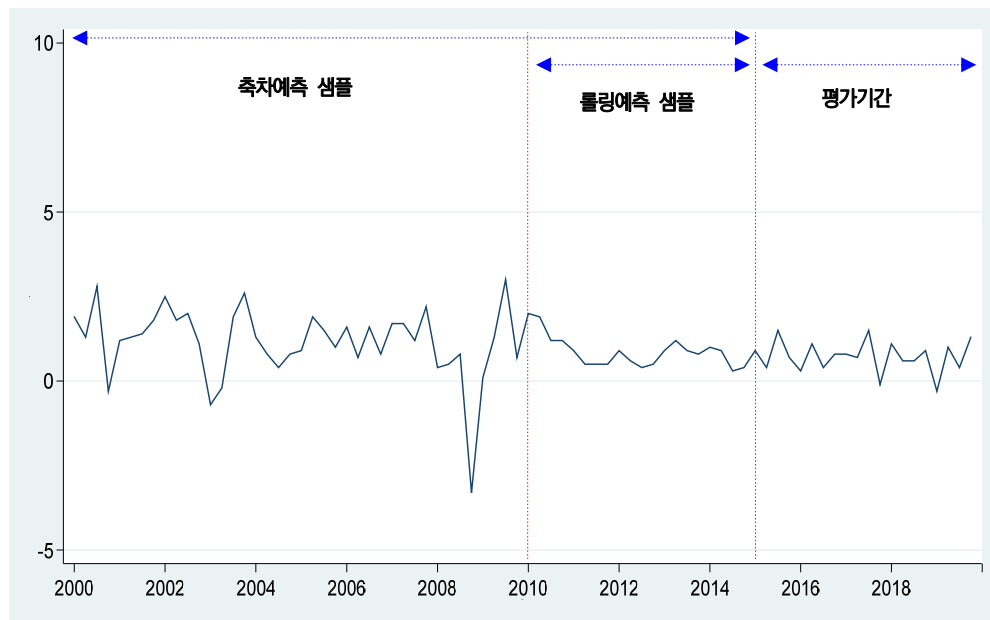
[그림 6] 롤링예측(rolling window forecast) 방식



다. 분석기간

분석기간은 2000년부터 2019년까지로 설정하여 2008년 금융위기 기간을 포함하였으며, 코로나19 발생 직전까지의 기간을 대상으로 하였다. 예측오차의 평가기간은 2015년 1/4분기부터 2019년 4/4분기까지 20개 분기로 정하였다. 추차예측의 경우 2000년부터 평가기간 전까지의 기간으로 모형을 구성하여 평가기간의 GDP성장률을 예측함에 따라 평가기간이 지나면서 분석기간 또한 늘어나게 된다. 롤링예측의 경우 5년을 기준으로 점차 분석기간을 변동하며 GDP성장률을 예측하였다. 즉, 예측시행 시점 직전 5년이 추정 표본으로 설정된다. 롤링방식의 예측력은 이동구간의 크기에 따라 달라지고 있어, 5~10년의 기간을 비교한 결과 가장 예측력이 우수하게 나타난 5년을 기준으로 측정하였다.

[그림 7] 실질GDP성장률의 예측 및 평가기간



주: 그림에 표시된 추차예측과 롤링예측 기간은 각각 2015년 1/4분기 예측을 위한 샘플 기준

또한, 우리나라의 평균 경기순환주기가 49개월(4년)로 롤링예측 샘플로 설정된 5년은 경기순환주기를 포함하고 있다는 점에서 적절한 추정기간으로 판단된

다. 경기순환은 경제의 장기추세를 중심으로 상승과 하강을 반복하는 현상으로, 경기저점에서 정점까지 경제활동이 활발한 확장국면, 경기정점에서 저점까지 경제활동이 위축된 수축국면으로 나눌 수 있다. 확장과 수축의 경기국면에서 저점에서 다음 저점까지, 또는 정점에서 다음 정점까지의 기간을 순환주기라고 하며 1972년 3월 저점부터 2017년 9월 정점까지 평균 확장기 33개월, 평균 수축기 18개월로 평균 순환주기는 49개월로 나타났다¹⁶⁾.

[표 3] 우리나라의 기준순환일 및 국면 지속기간

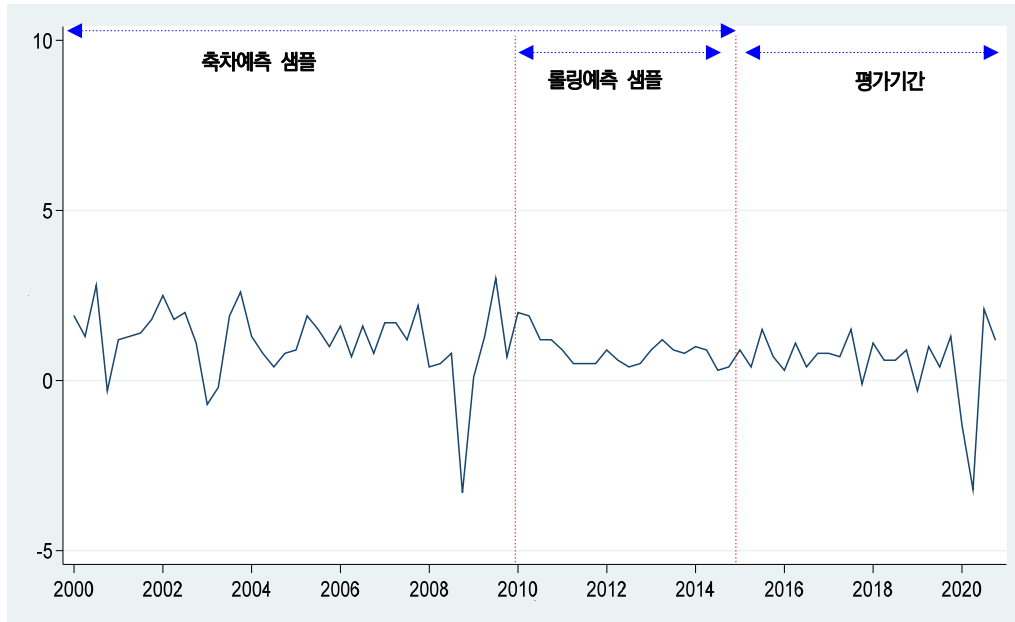
	기 준 순 환 일			지 속 기 간(개월)		
	저 점	정 점	저 점	확장기	수축기	순환기
제 1 순환기	1972. 3	1974. 2	1975. 6	23	16	39
제 2 순환기	1975. 6	1979. 2	1980. 9	44	19	63
제 3 순환기	1980. 9	1984. 2	1985. 9	41	19	60
제 4 순환기	1985. 9	1988. 1	1989. 7	28	18	46
제 5 순환기	1989. 7	1992. 1	1993. 1	30	12	42
제 6 순환기	1993. 1	1996. 3	1998. 8	38	29	67
제 7 순환기	1998. 8	2000. 8	2001. 7	24	11	35
제 8 순환기	2001. 7	2002.12	2005. 4	17	28	45
제 9 순환기	2005. 4	2008. 1	2009. 2	33	13	46
제10순환기	2009. 2	2011. 8	2013. 3	30	19	49
제11순환기	2013. 3	2017. 9	-	54	-	-
평 균	-	-	-	33	18	49

자료: 통계청

코로나19의 충격으로 경제성장률의 변동성이 매우 확대된 2020년을 분석에 포함할 경우 도출된 결론이 동 기간에 좌우될 가능성 등을 고려하여 2019년까지를 기본적인 분석기간으로 설정하되, 2020년을 포함한 분석결과도 함께 제시하여 코로나19가 예측성가에 미치는 영향도 살펴보았다.

16) 자료: 통계청

[그림 8] 실질GDP성장률의 예측 및 평가기간(2020년 포함)

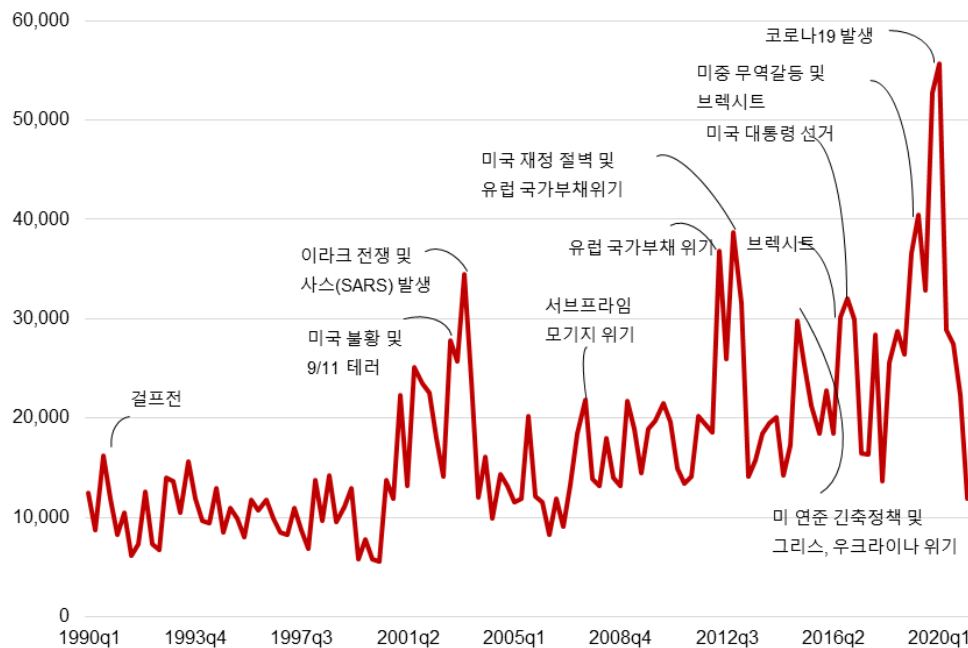


주: 그림에 표시된 추차예측과 롤링예측 기간은 각각 2015년 1/4분기 예측을 위한 샘플 기준

2020년 코로나19가 중국에서 발생한 이후 전세계로 확산되면서 세계보건기구(WHO)는 전염병 경보단계 중 최고위험등급인 팬데믹을 공식 선언하였다. 코로나19의 확산을 막기 위해 각국 정부는 국경을 봉쇄하고 이동 제한, 업무중지 조치 등을 시행하였으며, 이와 같은 대응조치는 국가 간 인적 및 물적 이동을 제한하고, 상품과 서비스 생산활동의 위축을 초래하였다. 세계경제는 코로나19의 충격으로 글로벌 공급망이 훼손되고 유효수요가 축소되는 등 공급과 수요 충격이 동시에 발생한 전례 없는 위기 상황을 경험하였다. 코로나19로 인해 산업활동이 위축되었으며 생산차질로 공급 및 유통망이 훼손되고 격리, 봉쇄, 사회적거리두기 등을 통해 노동공급이 감소하는 공급 충격이 나타났다. 또한, 전염병 확산으로 소비심리가 하락하고 대외활동이 제한되면서 소비가 둔화되었고 경제활동 위축에 따른 투자감소와 실업증가로 총수요 부문의 충격도 동반되었다. 과거에 나타난 위기들이 주로 수요 또는 공급측면의 문제에 기인하였던 것과 달리 코로나19는 공급 및 수요측면에서 동시에 충격을 발생시키며 기존에 경험해 보지 못한 높은

불확실성을 야기하였다. IMF에서 추정하고 있는 불확실성 지수(World Uncertainty Index)는 코로나19 발발 초기에 사상 최고치에 도달하는 등 이전의 위기들에 비해 불확실성이 매우 높은 모습을 보였다. 이와 같은 코로나19의 높은 불확실성과 변동성은 경제 예측 및 정책대응의 어려움을 가중시키고 있으며 실질GDP성장률 또한 정상적인 성장경로에서 벗어나 있을 가능성이 높다.

[그림 9] IMF의 전세계 불확실성 지수 추이

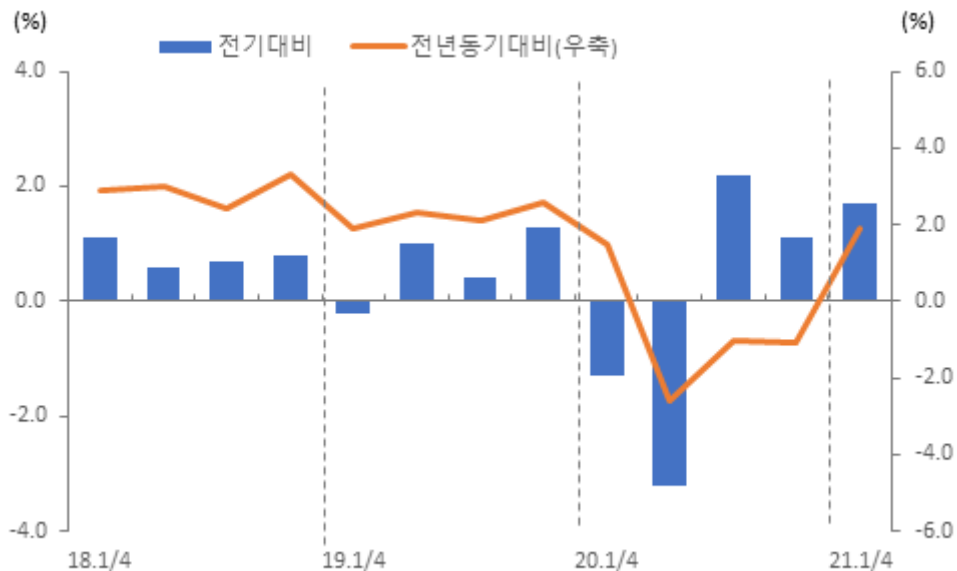


주: 세계 불확실성 지수(WUI)는 글로벌 경제정보서비스인 EIU(Economist Intelligence Unit)의 국가별 보고서에 '불확실성'과 관련된 단어가 언급된 횟수를 집계한 후 지수화 하여 IMF가 공표
 자료: IMF자료를 바탕으로 국회예산정책처 재작성

[참고] 2020년 경제성장률 추이

2020년 실질GDP는 코로나19 확산에 따른 경제성장 위축으로 전년대비 0.9% 감소하였다. 특히 상반기 중 코로나19 확산에 따라 경제활동이 위축되고 세계교역량이 감소하면서 민간소비와 수출을 중심으로 큰 폭으로 급감하였다. 1/4분기와 2/4분기 실질GDP성장률은 각각 전기대비 -1.3%, -3.2%를 기록하였으며, 전년동기대비로는 2/4분기 2.6% 감소하여 외환위기로 인해 성장률이 큰 폭으로 하락하였던 1998년 4/4분기(-3.8%) 이후 가장 낮은 성장률을 기록하였다. 하반기 중에는 코로나19 확산이 다소 진정되는 양상을 보이면서 세계경기 개선세와 경제활동 재개 등으로 인해 수출과 설비투자를 중심으로 회복되는 모습을 보였다. 3/4분기와 4/4분기 실질GDP성장률은 전기대비 2.1%, 1.2%를 기록하며 증가세로 전환되었다. 부문별로는 코로나19 확산에 따른 방역조치와 각국의 봉쇄조치로 민간소비와 수출이 급격하게 감소하는 모습을 나타냈으나, 대외수요가 점차 개선되면서 수출과 설비투자는 비교적 빠른 회복세를 보이고 있다. 반면, 민간소비는 11월 들어 코로나19 3차 확산과 사회적거리두기 강화 등으로 하반기에도 감소하는 모습을 보였다.

[GDP성장률 추이]



자료: 한국은행

라. 이용자료

본 분석모형을 추정하기 위해 실물부문과 금융부문 등을 포괄하는 26개의 변수를 사용하였다. [표 4]는 분석에 사용된 자료들을 생산부문, 소비부문, 투자부문, 가격지수, 노동부문, 금융부문으로 분류하고 있다.

[표 4] 분석에 포함된 변수 목록

분류	변수	발표시차
생산부문	산업생산지수(전산업)	2개월
	산업생산지수(제조업)	
	산업생산지수(서비스업)	
	제조업평균가동률	
경기부문	경기선행지수 순환변동치	2개월
	경기동행지수 순환변동치	
소비부문	소매판매지수	2개월
	소매판매지수(내구재)	
	소매판매지수(비내구재)	
투자부문	설비투자지수	2개월
	건설기성액	
가격지수	소비자물가지수	1개월
	생산자물가지수	
	수입물가지수	
노동부문	취업자수	1개월
	실업률	
	고용률	
	실질임금상승률	2개월
금융부문	M1, M2, LF	2개월
	국고채(3년) 금리	0개월
	KOSPI 지수	
대외부문	수출금액, 수입금액	1개월
	실질실효환율	

생산부문에는 산업생산지수와 제조업 가동률 등이 포함되었으며, 경기부문에는 현재와 미래의 경기동향을 살펴볼 수 있는 경기선행지수 순환변동치와 경기동행지수 순환변동치를 포함하였다. 산업생산지수는 국내에서 생산한 재화와 용역에 대한 생산활동 동향을 나타낸 지표로 작성주기가 월간단위로서 신속한 동향파악이 가능한 장점이 있다. 제조업 평균가동률은 제조업체의 생산설비 이용도를 보여주는 지표로 제조업 부문의 생산능력과 설비이용정도를 나타내는 지표이다. 경기선행지수 순환변동치와 경기동행지수 순환변동치는 각각 선행종합지수와 동행종합지수에서 추세변동분을 제거한 지표로 현재 경기국면 및 향후경기국면, 전환점 예측에 이용되는 지표이다.

소비부문에서는 통계청의 소매판매지수를 포함하였으며, 특성이 상이한 내구재와 비내구재 부문을 구분하였다. 투자부문은 설비투자와 건설투자를 각각 포괄하는 설비투자지수와 건설기성액을 포함시켰다. 소매판매지수는 소비동향을 파악하기 위해 작성되는 월별 지표로 백화점, 대형마트, 면세점, 슈퍼마켓, 전문소매점, 편의점 등 소매판매점의 매월 판매금액을 조사하여 작성하고 있다. 내구재는 1년 이상 반복적으로 사용이 가능하고 주로 고가의 상품이 해당되며, 비내구재는 주로 이용기간이 짧고 반복적으로 이용이 곤란한 상품이 해당된다. 설비투자지수는 신속한 설비투자 동향파악을 위해 매월 국내에 공급되는 설비투자재의 투자액을 추정하는 월별지표이며, 건설기성액은 건설업체의 시공실적을 금액으로 평가하는 월별지표이다.

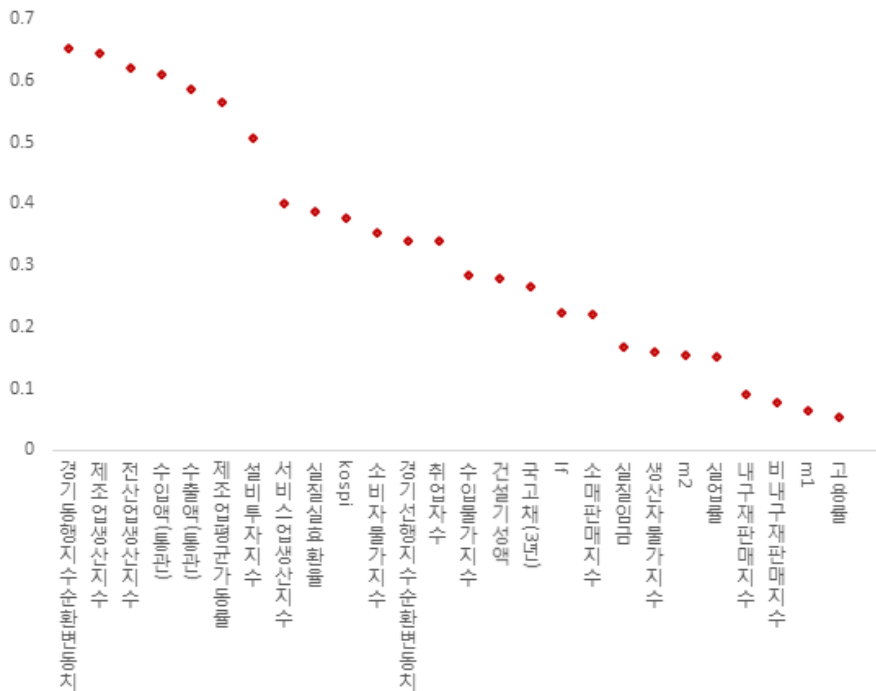
가격지수로는 소비자물가지수, 생산자물가지수, 수입물가지수를 포함시켰으며, 노동부문에는 취업자수, 실업률, 고용률 등 고용지표와 실질임금상승률을 포함시켰다. 금융부문에서는 M1(현금통화와 요구불예금, 수시입출식예금), M2(M1과 만기 2년 미만 금융상품), LF(금융기관 유동성) 등 통화량 지표, 금리(국고채 3년물)와 주가지수(코스피 지수) 등 금융시장 지표들이 포함되었다. 대외부문은 통관 수출입액과 실질실회환율이 포함되었다.

생산, 소비, 경기, 투자, 고용 등 실물지표와 유동성 지표의 경우 통계생성기관에서 발표하는 계절조정계열을 사용하였다. 계절조정계열이 공표되지 않는 코스피, 통관수출입, 임금상승률의 경우 X12-Arima를 통해 계절조정하였으며, 로그차분을

통해 안정적 시계열로 변환하였다.

GDP성장률과 이용변수간 상관관계를 산출한 결과 경기동행지수, 제조업생산지수, 전산업생산지수 등 생산부문 지표들과의 상관관계가 높은 것으로 나타났으며, 수입액 및 수출액도 상관관계가 높은 편으로 나타났다. 이어서 제조업가동률, 설비투자지수, 서비스업생산지수 등 생산관련 지표들의 순으로 상관관계가 높게 나타났다.

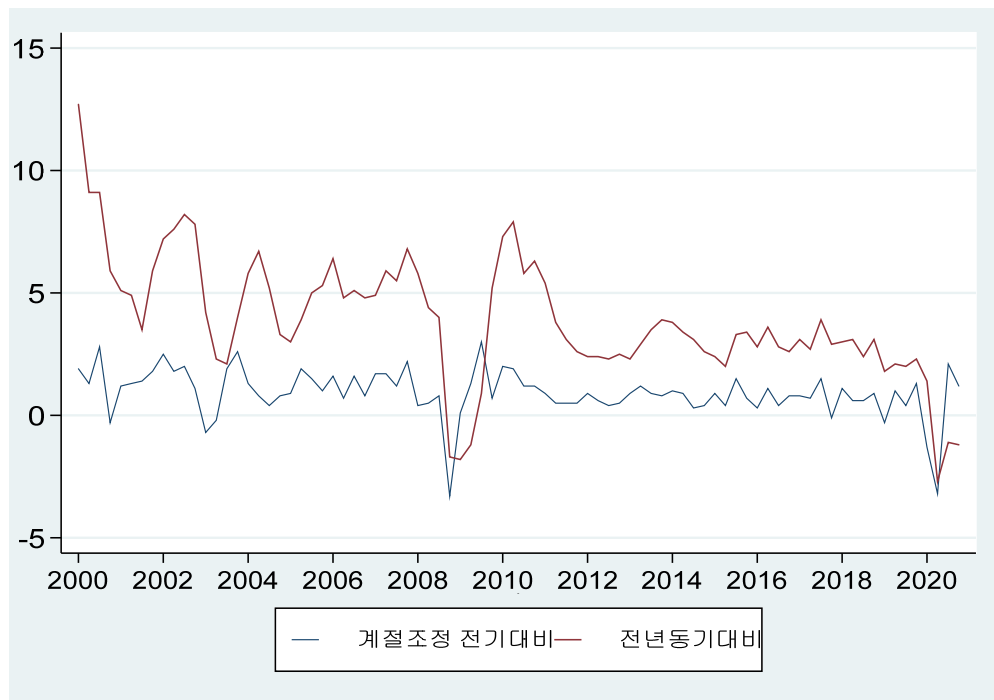
[그림 10] GDP성장률과 이용변수간 상관관계



예측의 대상이 되는 GDP성장률의 경우 계절조정된 전기대비 성장률을 사용하였다. 한국은행은 2006년부터 전기대비 성장률을 사용하고 있으며, 전년동기대비 성장률은 보조지표로 활용하고 있다. 전년동기대비 성장률은 비교대상이 1년 전 시점이기 때문에 단기적인 경기변동을 파악하기가 어려울 뿐만 아니라 경기전환점 포

착시기가 지연되는 문제점을 내포하고 있어 한국은행은 2006년 1/4분기부터 분기 성장률 주지표를 전년동기대비에서 전기대비로 전환하였다¹⁷⁾. 본 연구에서도 주지표인 전기대비 실질GDP성장률을 기준으로 예측모형을 설정하였다. 0~1분기를 전망하는 단기 GDP 예측모형은 비교적 단기간인 최근의 경제상황을 분석하고 예측하기 위한 모형인만큼 경기변동의 흐름을 신속하게 포착할 수 있는 장점이 있는 전기대비 지표를 사용하는 것이 바람직할 것으로 판단된다.

[그림 11] 실질GDP성장률 추이(%)



자료: 한국은행

17) 한국은행(2006)

[표 5] 분기성장률 지표의 비교

	전년동기대비	전기대비
정의	- 전년의 같은 분기에 비해 당해분기가 얼마나 성장했는지를 보여주는 지표 - 전년의 같은 분기와의 비교를 통해 분기의 계절적 영향을 제거	- 현재분기가 전분기보다 나아지고 있는지 나빠지고 있는지를 판단하는 지표 - 통계적 기법에 의해 계절조정된 계열에 기초하여 작성
산출식	$\left(\frac{Y_t}{Y_{t-4}} - 1 \right) \times 100$ Y_t: 당해분기 GDP Y_{t-4}: 4분기 전 분기 GDP	$\left(\frac{Y_t^{sa}}{Y_{t-1}^{sa}} - 1 \right) \times 100$ Y_t^{sa}: t분기의 계절조정 통계
장점	- 전기비에 비해 변동성이 상대적으로 작게 나타남 - 연간성장률과 안정적인 근사성을 보임에 따라 연간성장률을 사전적으로 가늠하는데 적합 - 작성이 간편하고 이해하기 쉬움	- 경기회복의 전환점을 신속하게 포착 - 경기변동의 방향과 속도를 보다 잘 보여줌 - 미국, 일본, 영국, 캐나다 등 주요 선진국들이 주지표로 채택
단점	- 경제구조 변화가 계절성분에 미치는 영향(이동계절성)이 반영되지 않음 - 명절, 공휴일 수, 영업일수 차이 등 달력효과 제거 불가능 - 전년동분기 파업 등과 같은 특이 현상 발생시 기저효과가 크게 발생 - 당해년 경기움직임이 동일하더라도 전년도에 변동패턴에 따라 전년동기대비 성장률은 상이하게 나타날 수 있음 - 경기전환 포착시점 지연으로 경기국면 판단을 흐리게 함	추세, 순환요인뿐만 아니라 파업, 자연재해 등 불규칙 요인도 포함하고 있어 전년동기대비에 비해 상대적으로 큰 변동 기록을 보임

자료: 한국은행(2006) 자료를 토대로 작성

Ⅲ. 분석결과

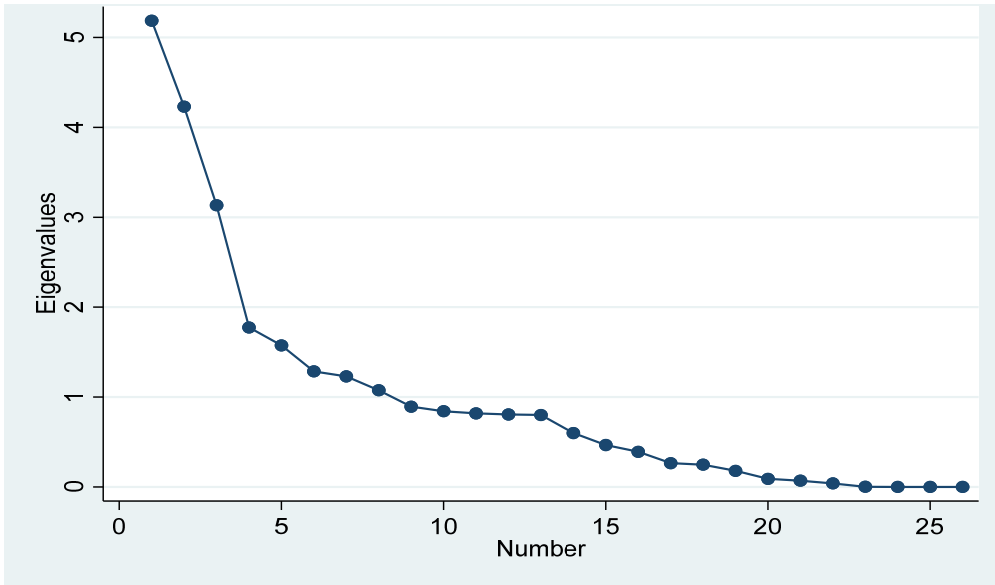
1. 주성분분석¹⁸⁾

실물부문과 금융부문을 포괄하는 거시경제의 26개 변수를 사용하여 주성분분석을 수행하였다. 변수들간의 분산-공분산 관계를 이용하여 모든 정보의 특성을 최대한 반영하면서 상호독립적인 주성분(principal component)을 추출하였다.

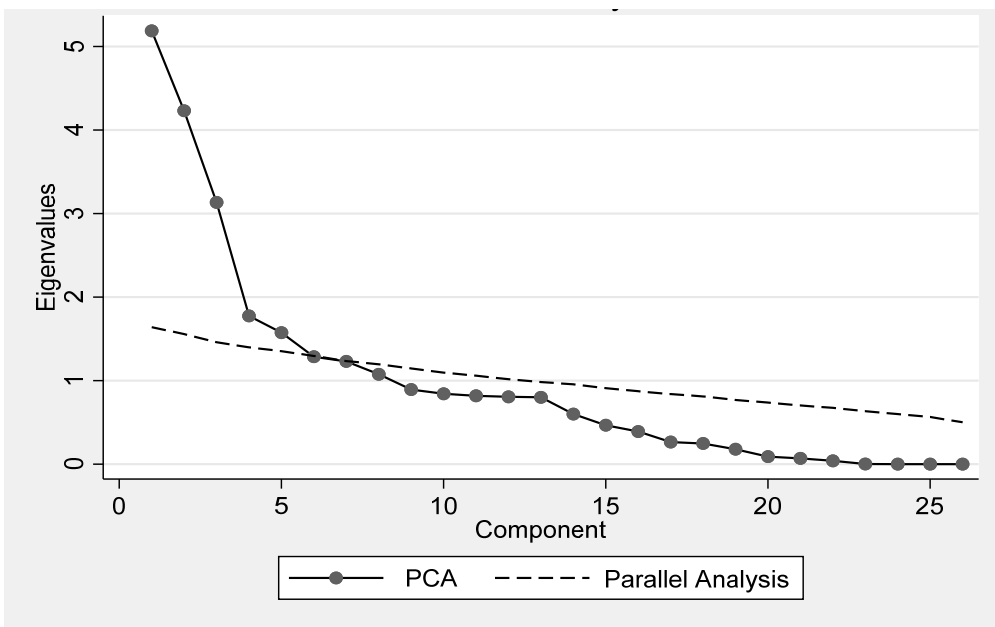
분석에 포함된 경제변수의 개수와 같은 26개의 주성분요인이 추출되나, 모두 의미를 가지는 것은 아니기 때문에 적절한 주성분 요인의 개수를 선택해야 한다. 주성분분석에서 주로 사용하는 카이저기준, 설명력기준, Scree Plot, Parallel 분석을 모두 고려하였다. 카이저기준에 따르면 고유값(eigenvalue)이 1보다 큰 제9주성분(고유값 1.1032) 까지 선택될 수 있다(표5). 또한 Scree Plot에 의해서는 고유값이 급격한 차이가 나타나는 지점 이상인 제4주성분 까지, Parallel 분석에 따라서는 임의 추출한 값보다 큰 6개의 주성분이 선택될 수 있다. 이에 따라 본 분석에서는 4~8개의 주성분을 모두 이용하여 분석을 수행하였으며, 분석결과 예측오차가 가장 적은 것으로 나타난 5개를 적정 주성분의 개수로 선택하였다.

18) 매기 주성분분석을 통해 공통요인을 추정하여 분석하였으나, 2000년 1월~2019년 12월 기준 주성분분석 결과를 토대로 살펴봄

[그림 12] 주성분분석 Scree Plot



[그림 13] 주성분분석 Parallel 분석



누적비중(cumulative)에 따르면 선택된 5개의 주성분은 전체 분산의 51.9%를 설명하고 있다. 각 주성분이 원래 변수의 정보를 설명하는 정도를 나타내는

비중(proportion)을 살펴보면 제1주성분은 전체 17.6%의 설명력을, 제2주성분은 12.6%의 설명력을 보이고 있다.

[표 6] 주성분분석 결과

	고유값(Eigenvalue)	비중(Proportion)	누적비중(Cumulative)
PRIN1	4.5779	0.1761	0.1761
PRIN2	3.2831	0.1263	0.3023
PRIN3	2.0605	0.0792	0.3816
PRIN4	1.9380	0.0745	0.4561
PRIN5	1.6443	0.0632	0.5194
PRIN6	1.4787	0.0569	0.5763
PRIN7	1.3397	0.0515	0.6278
PRIN8	1.2118	0.0466	0.6744
PRIN9	1.1032	0.0424	0.7168
PRIN10	0.9599	0.0369	0.7537

주성분분석을 통해 추정된 고유벡터들에서 각 변수들에 대응되는 벡터값을 통해 각 주성분이 어떤 요인들에 의해 주도되는지 살펴볼 수 있다¹⁹⁾. 각 주성분 별로 가중치가 크게 나타난 지표들을 살펴보면 제1주성분의 경우 생산지수, 소매판매, 수출입액 등 실물부문과 경기동행지수순환변동치와 경기선행지수순환변동치 등 경기

19) 김중호·이중식(2005)에 따르면 일반적으로 주성분분석에 의해 도출되는 공통요인들은 서로 직교관계를 형성하는 통계적 가공변수로 그 경제적 의미를 식별하기가 쉽지 않으나, 구성변수들로 이루어진 2차 적률행렬의 고유벡터에서 개별변수들에 대응되는 벡터값의 크기를 통해 주성분의 움직임이 어느 변수에 의해 주도되고 있는지 파악할 수 있음

요인이 주도하고 있는 것으로 나타났다. 특히 제1주성분의 경우 앞서 살펴본 변수들 중 GDP와의 상관계수가 높게 나타난 항목들이 많이 포함되어 있으며, GDP와 상관관계가 높은 상위 5개 변수가 모두 포함되어 있는 것으로 나타났다.

[표 7] 각 주성분의 주요 정보변수

주성분	해당 주성분의 주요 정보변수
PRIN1	전산업생산지수(0.3775), 제조업생산지수(0.3198), 서비스생산지수(0.2973), 경기동행지수순환변동치(0.3096), 소매판매지수(0.2523), 수입액(0.2348), 수출액(0.2307), 경기선행지수순환변동치(0.2159), 건설기성액(0.1992)
PRIN2	국고채(0.4025), 소비자물가지수(0.4442), 생산자물가지수(0.3683), 수입물가지수(0.3589), 제조업평균가동률(0.2411), 전산업생산지수(0.1591)
PRIN3	M2(0.4080), LF(0.3903), 소매판매지수(0.3148), M1(0.2805), 고용률(0.2566), 비내구재판매지수(0.2443), 내구재판매지수(0.1710)
PRIN4	LF(0.4602), M2(0.4363), 수입액(0.2370), 수출액(0.2267), 제조업평균가동률(0.1058)
PRIN5	실질실효환율(0.3609), 코스피(0.2915), 국고채(0.2848), 취업자수(0.2822), 실업률(0.2441), 경기선행지수순환변동치(0.2524), 제조업평균가동률(0.2301) 경기동행지수순환변동치(0.1820)

주: () 안은 분산공분산행렬의 고유벡터 값으로 해당요인과 변수간의 상관관계를 나타냄

제2주성분의 경우 국고채, 소비자물가, 생산자물가, 수입물가 등 금리와 물가의 영향력이 크게 나타났다. 제3주성분과 제4주성분의 경우 M2와 LF 등 통화량 지표의 영향력이 크게 나타났으며, 3주성분의 경우 소매판매 관련 지표가, 4주성분의 경우 수출입 지표가 포함되었다. 제5 주성분의 경우 환율, 주가, 금리 등 금융지

표가 주도하고 있는 것으로 나타났으며, 취업자 수와 실업률 등 고용부문의 영향력도 크게 나타났다.

[표 8] 주성분분석모형의 예측력 평가

		당분기		다음 분기	
		RMSE	MAE	RMSE	MAE
축차예측	전체	0.397	0.315	0.445	0.608
	초월	0.399	0.312	0.446	0.625
	중월	0.385	0.352	0.445	0.595
	말월	0.386	0.377	0.438	0.588
롤링예측	전체	0.370	0.284	0.417	0.586
	초월	0.387	0.295	0.421	0.589
	중월	0.391	0.301	0.400	0.579
	말월	0.352	0.322	0.412	0.577
AR모형 20)	-	0.581	0.777	0.596	0.856

주성분분석 결과를 토대로 추출된 요인들과 GDP성장률의 회귀분석을 통해 예측모형을 설정하여 표본 외 예측을 실시하였으며, 표본 외 예측결과를 실제 GDP성장률과 비교하여 예측오차를 산출하였다. 분석결과 월별자료를 이용해 GDP를 예측한 주성분분석모형이 분기자료를 이용해 GDP를 예측한 벤치마크 모형인 AR모형에 비해 예측력이 우수한 것으로 나타났다. 주성분분석모형을 이용한 당분기 GDP예측의 경우 RMSE 0.35~0.40%p, MAE 0.28~0.38%p로 나타났다. 다음 분기 예측의 경우 RMSE 0.40~0.50%p, MAE 0.58~0.63%p 수준을

20) 각 시점별로 AIC, BIC 기준에 따라 최적시차를 결정

나타냈다. 반면 AR모형의 당분기 예측오차는 RMSE 0.58%p, MAE 0.78%p로 월별지표를 이용한 주성분분석모형에 비해 크게 나타났다.

주성분분석을 통한 예측 결과 전제월(초월, 중월, 말월의 결과를 평균)과 초월, 중월, 말월의 예측오차 차이는 크게 나타나지 않는 모습이다. 이는 초월에서 말월로 이동하면서 모형의 모수 추정치가 크게 변하지 않았고 예측에 사용된 월별 변수 결측치에 대한 예측치도 어느 정도 정확성을 확보하고 있기 때문으로 판단된다²¹⁾. 초월은 가장 적은 정보를 바탕으로 예측하고 있으나 예측력이 크게 떨어지지 않는 모습이며, 가장 많은 정보를 바탕으로 하는 말월의 경우에도 예측력은 초월이나 중월과 비슷한 모습이다.

예측오차 측정방법 중에서는 당분기의 경우 MAE와 RMSE가 비슷하게 나타났으나, 다음분기 예측에는 MAE에 비해 RMSE가 작게 나타나 편차가 상대적으로 큰 예측오차들이 많지 않은 것을 알 수 있다. 일반적으로 예측오차가 작은 값을 갖고 있음에도 불구하고 편차가 상대적으로 큰 오차값이 존재하는 경우 RMSE가 오차정도를 과대평가할 수 있어 MAE에 비해 RMSE가 크게 나오는 경향이 있다.

축차예측과 롤링예측을 비교해보면 당분기와 다음 분기 모두 롤링예측이 축차예측에 비해 예측오차가 작은 것으로 나타났다. 롤링방식의 경우 오래된 자료를 배제하고 최신의 자료를 사용하여 상대적으로 월별자료의 변화를 반영하였기 때문에 예측오차를 줄인 것으로 보인다. 앞서 이진희·김덕파(2014)는 구조변화가 포함된 시계열의 경우 구조변화의 지속적인 영향을 받게 되는 축차예측방식보다 롤링예측방식이 유리할 수 있음을 지적하였다. 글로벌 금융위기 시기인 2008년 이 분석기간에 포함되었으며, 이 기간 중 한국경제의 성장추세에서 구조적 변화가 발생²²⁾되었다면 경제구조변화가 시계열의 불안정성을 가져와 축차예측방식의 예측력을 낮추었을 것으로 판단된다.

21) 김덕파, 강성진(2020)에 따르면 예측치에 직접 포함된 월별변수의 값은 초월말에서 말월말로 진행하면서 결측치에 대한 대체치가 실측치로 교체되기는 하나 대체치가 실측치에 대해 일정한 설명력을 지남에 따라 실질GDP 예측치에 미치는 영향이 제한적일 가능성

22) 진익(2020년)은 실질국내총생산의 자연대수값을 대상으로 Chow 검정을 수행한 결과 2008년 한국경제의 성장이 구조적으로 둔화되었음을 확인하였다.

2. 동태요인분석

주성분분석과 마찬가지로 동태요인분석을 통해 추출된 요인들과 GDP성장률의 회귀분석을 통해 모형을 설정하였다. 모형을 통해 예측한 결과를 토대로 표본 외 예측을 실시하였으며, 표본 외 예측결과를 실제 GDP성장률과 비교하여 예측오차를 산출하였다. 분석결과 월별자료를 이용해 GDP를 예측한 동태요인모형은 주성분분석모형과 마찬가지로 분기자료를 이용해 GDP를 예측한 벤치마크 모형인 AR 모형에 비해 예측력이 우수한 것으로 나타났으며, 주성분분석에 비해서도 예측력이 우수한 것으로 나타났다.

월별지표를 이용한 동태요인분석모형의 당분기 GDP예측의 경우 축차 및 롤링방법을 통해 추정된 경우 RMSE가 0.34~0.38%p, MAE가 0.25~0.28%p로 나타났다. 다음 분기의 경우 당분기보다 예측오차가 크게 나타나 RMSE 0.35~0.44%p, MAE 0.39~0.48%p 범위로 나타났다. 반면 AR모형의 예측오차는 RMSE 0.58%p, MAE 0.78%p로 동태요인분석모형에 비해 크게 나타났다. 또한, 주성분분석을 통한 예측결과와 마찬가지로 전체월과 초월, 중월, 말월의 예측오차 차이가 크게 나타나지 않았으며, 당분기와 다음 분기 모두 롤링예측이 축차예측에 비해 예측오차가 작은 것으로 나타났다.

주성분분석모형과 예측력을 비교해 보면 동태요인분석을 통한 실질GDP 예측은 당분기와 다음 분기, 축차와 롤링방식 모두에서 예측오차가 작게 나타났다. 이는 AR모형을 통해 결측치를 처리하는 주성분분석모형에 비해 공통요인의 동학을 추정하여 결측치를 처리하는 동태요인분석 방식이 보다 유용한 정보를 포함하고 있어 예측오차를 낮춘 것으로 판단된다. 즉, 공통요인의 동태적 특성을 이용하여 자료주기, 공표시점 차이 등을 내생화시켜 자료의 비정규성을 효율적으로 처리함²³⁾으로써 예측력을 높일 수 있었던 것으로 보인다.

동태요인분석을 통한 GDP예측 결과 MAE와 RMSE가 주성분분석모형에 비해 낮아진 것으로 나타나 상대적으로 예측력이 우수한 것으로 판단된다. 다만, 주

23) 박강우·이중식(2009)

성분 분석과 반대로 당분기의 경우 MAE 보다 RMSE를 통해 계산된 예측오차가 크게 나타났다. 즉, 작은 오차가 지속적으로 나타나기 보다는 한번씩 큰 오차를 범하는 것으로 나타났다. 앞서 언급한 바와 같이 예측오차가 작은 값을 갖고 있음에도 불구하고 편차가 상대적으로 큰 오차값이 존재하는 경우 RMSE가 오차정도를 과대평가할 수 있어 MAE에 비해 RMSE가 크게 나오는 경향이 있다. 즉, RMSE는 실적치에서 예측치를 뺀 절댓값이 1보다 큰 경우 제곱을 통해 오차를 더욱 크게 만들기 때문에 평가기간 동안 큰 예측오차가 나타나는 경우 MAE에 비해 크게 나타나게 된다.

[표 9] 동태요인분석모형의 예측력 평가

		당분기		다음분기	
		RMSE	MAE	RMSE	MAE
축차예측	전체	0.365	0.279	0.439	0.472
	초월	0.375	0.280	0.437	0.475
	중월	0.360	0.245	0.440	0.477
	말월	0.359	0.275	0.435	0.468
롤링예측	전체	0.338	0.263	0.352	0.390
	초월	0.337	0.266	0.357	0.396
	중월	0.342	0.259	0.351	0.390
	말월	0.335	0.263	0.351	0.385
AR모형 24)	-	0.581	0.777	0.596	0.856

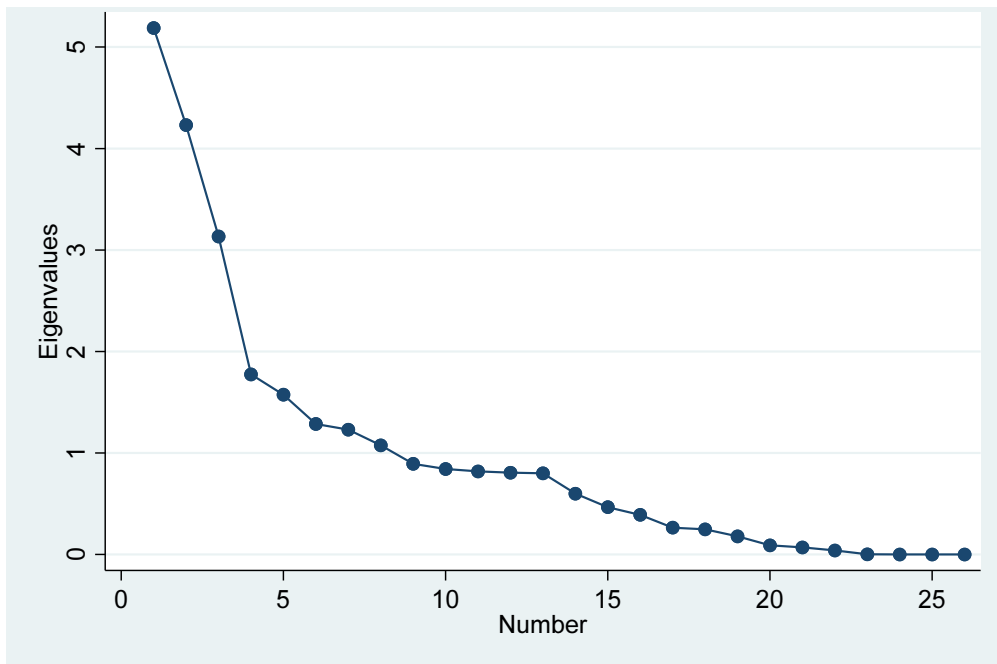
24) 각 시점별로 AIC, BIC 기준에 따라 최적시차를 결정

3. 코로나19 기간을 포함한 분석 결과

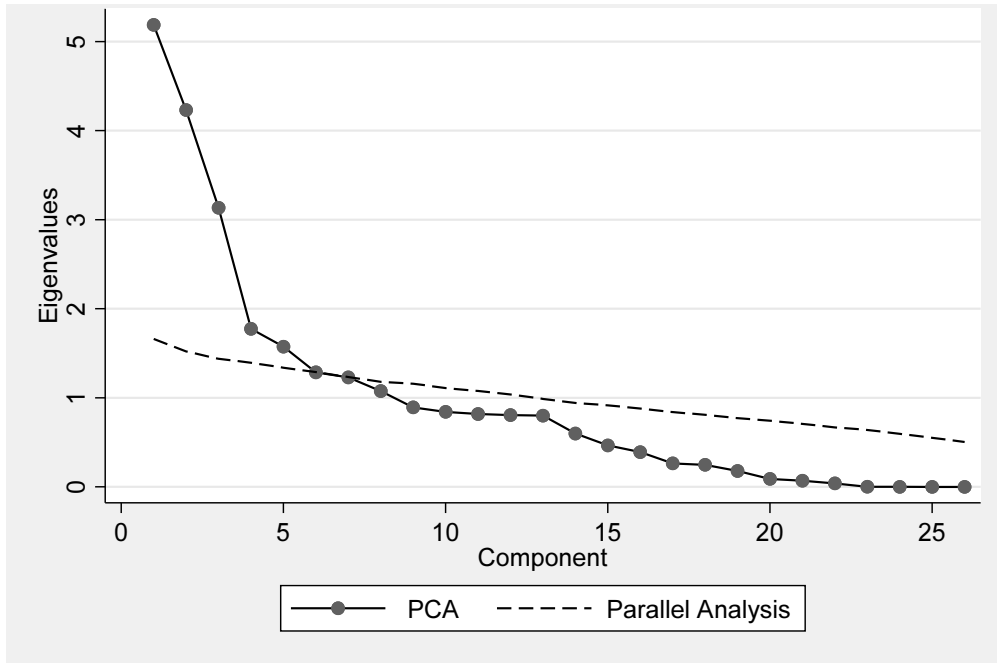
코로나19 확산으로 충격을 받은 2020년을 분석에 포함하여 주성분을 추출한 결과 고유값, 비중, 주성분의 주요 정보변수가 2019년까지 자료를 대상으로 한 분석과 다소 차이를 나타냈다.

먼저, 적절한 주성분 요인의 개수를 선택하기 위해 카이저기준, 설명력기준, Scree Plot, Parallel 분석 결과는 2019년까지 자료를 대상으로 한 분석과 비슷한 모습을 나타냈으며, 앞선 분석과 마찬가지로 분석결과 예측오차가 가장 적은 것으로 나타난 5개를 적정 주성분의 개수로 선택하였다.

[그림 14] 주성분분석 Scree Plot(2020년 포함)



[그림 15] 주성분분석 Parallel 분석(2020년 포함)



주성분분석 결과 누적비중(cumulative)에 따르면 선택된 5개의 주성분은 전체 분산의 61.2%를 차지하고 있어 2019년까지의 분석에서 나타난 51.9%에 비해 설명력은 높은 모습을 보였다. 각 주성분이 원래 변수의 정보를 설명하는 정도를 나타내는 비중(proportion)을 살펴보면 제1주성분은 전체 20.0%의 설명력을, 제2주성분은 16.3%의 설명력을 보이고 있다.

주성분분석을 통해 추정된 고유벡터들에서 각 변수들에 대응되는 벡터값은 2019년까지를 대상으로 한 분석과 차이를 나타냈다. 각 주성분 별로 가중치가 크게 나타난 지표들을 살펴보면 제1주성분의 경우 통화량과 임금요인에 의해 주도되고 있는 것으로 나타났으며, 제2주성분의 경우 생산지수, 경기지수 등 경기요인이 주도하고 있는 것으로 나타났다. 제3주성분의 경우 물가, 금리, 생산, 소비 등 다양한 부분의 움직임을 반영하고 있으며 특히 물가요인의 영향력이 큰 것으로 나타났다. 제4주성분은 소매판매지수, 비내구재 및 내구재 판매지수 등 소비요인이 주도하고 있는 것으로 나타났으며, 실업률이 유일하게 포함된 주성분으로 나타났다. 제5주성

분은 고용률이 가장 크게 나타났으며, 물가, 수출입, 소비, 투자, 생산 등 다양한 실물부문 지표들이 주도하고 있는 것으로 나타났다. 앞서 분석한 2019년까지의 결과에서는 제1주성분의 경우 실물지표 및 경기지표의 영향력이 높게 나타났으나, 코로나19의 영향을 받은 2020년을 포함할 경우 통화량 지표가 높게 나타난 점에서 차이가 있다. 또한, 환율, 주가 등 금융지표들이 포함되지 않았다는 점에서 차이가 있다.

[표 10] 주성분분석결과

	고유 값(Eigenvalue)	비중(Proportion)	누적비중(Cumulative)
PRIN1	5.1875	0.1995	0.1995
PRIN2	4.2310	0.1627	0.3622
PRIN3	3.1337	0.1205	0.4828
PRIN4	1.7744	0.0682	0.551
PRIN5	1.5745	0.0606	0.6116
PRIN6	1.2860	0.0495	0.6610
PRIN7	1.2298	0.0473	0.7083
PRIN8	1.0757	0.0414	0.7497
PRIN9	0.8930	0.0343	0.7841
PRIN10	0.8425	0.0324	0.8165

[표 11] 각 주성분의 주요 정보변수

주성분	해당 주성분의 주요 정보변수
PRIN1	M2(0.4329), LF(0.4329), M1(0.4327), 실질임금(0.4324)
PRIN2	전산업생산지수(0.4073), 제조업생산지수(0.3440), 서비스생산지수(0.2977), 경기동행지수순환변동치(0.3276), 수입액(0.2752), 수출액(0.2712), 소매판매지수(0.2272)
PRIN3	소비자물가지수(0.4557), 국고채금리(0.3983), 생산자물가지수(0.3903), 수입물가지수(0.3950), 제조업 평균가동률(0.1550), 전산업생산지수(0.1457), 소매판매지수(0.1271), 비내구재판매지수(0.1235), 서비스업생산지수(0.1213)
PRIN4	소매판매지수(0.5538), 비내구재판매지수(0.4492), 내구재판매지수(0.2749), 서비스업생산지수(0.1679), 실업률(0.1274)
PRIN5	고용률(0.3800), 생산자물가지수(0.2077), 내구재판매지수(0.2994), 소매판매지수(0.2370), 수출액(0.1990), 수입액(0.1744), 수입물가지수(0.1774)

주: () 안은 분산공분산행렬의 고유벡터 값으로 해당요인과 변수간의 상관관계를 나타냄

코로나19 확산에 따른 경제적 충격이 나타난 2020년을 포함하여 주성분분석과 동태요인분석모형으로 GDP성장률을 추정하였다. 추출된 요인들과 GDP성장률의 회귀분석을 통해 모형을 설정하고 예측한 결과를 토대로 표본 외 예측을 실시하였으며, 표본 외 예측결과를 실제 GDP성장률과 비교하여 예측오차를 산출하였다. 분석결과 코로나19의 충격이 반영된 2020년 자료를 분석에 포함시킬 경우 예측오차가 높아지는 것으로 나타났다.

[표 12] 2020년을 포함한 모형의 예측력 평가

분석 모형	오차측정 방법	예측기간	당분기		다음분기	
			RMSE	MAE	RMSE	MAE
주성분 분석	축차	2019년까지	0.397	0.315	0.445	0.608
		2020년까지	0.845	0.524	1.021	0.752
	롤링	2019년까지	0.370	0.284	0.417	0.586
		2020년까지	0.764	0.490	0.952	0.744
동태요인분석	축차	2019년까지	0.365	0.279	0.439	0.472
		2020년까지	0.825	0.480	1.122	0.705
	롤링	2019년까지	0.338	0.263	0.352	0.390
		2020년까지	0.753	0.489	0.985	0.658
AR모형 ²⁵⁾	-	2019년까지	0.581	0.777	0.596	0.856
		2020년까지	0.920	0.895	1.131	0.954

예측기간, 분석모형, 측정방법 모두 분석기간에 2020년을 포함할 경우 MAE와 RMSE가 높아진 것으로 나타났다. 특히, 2020년을 분석에 포함할 경우 RMSE의 증가폭이 매우 크게 나타났다. 주성분분석모형에서는 0.37~0.40%p 수준이던 RMSE가 0.76~0.85%p로 높아졌으며, 동태요인분석모형에서는 0.34~0.37%p에서 0.75~0.83%p로 높아졌다. RMSE 방식으로 산출한 예측오차는 편차가 상대적으로 큰 오차값이 존재하는 경우 커질 수 있다. 2020년의 경우 GDP성장률 추정치에 대한 예측오차가 상대적으로 매우 크게 나타났기 때문에

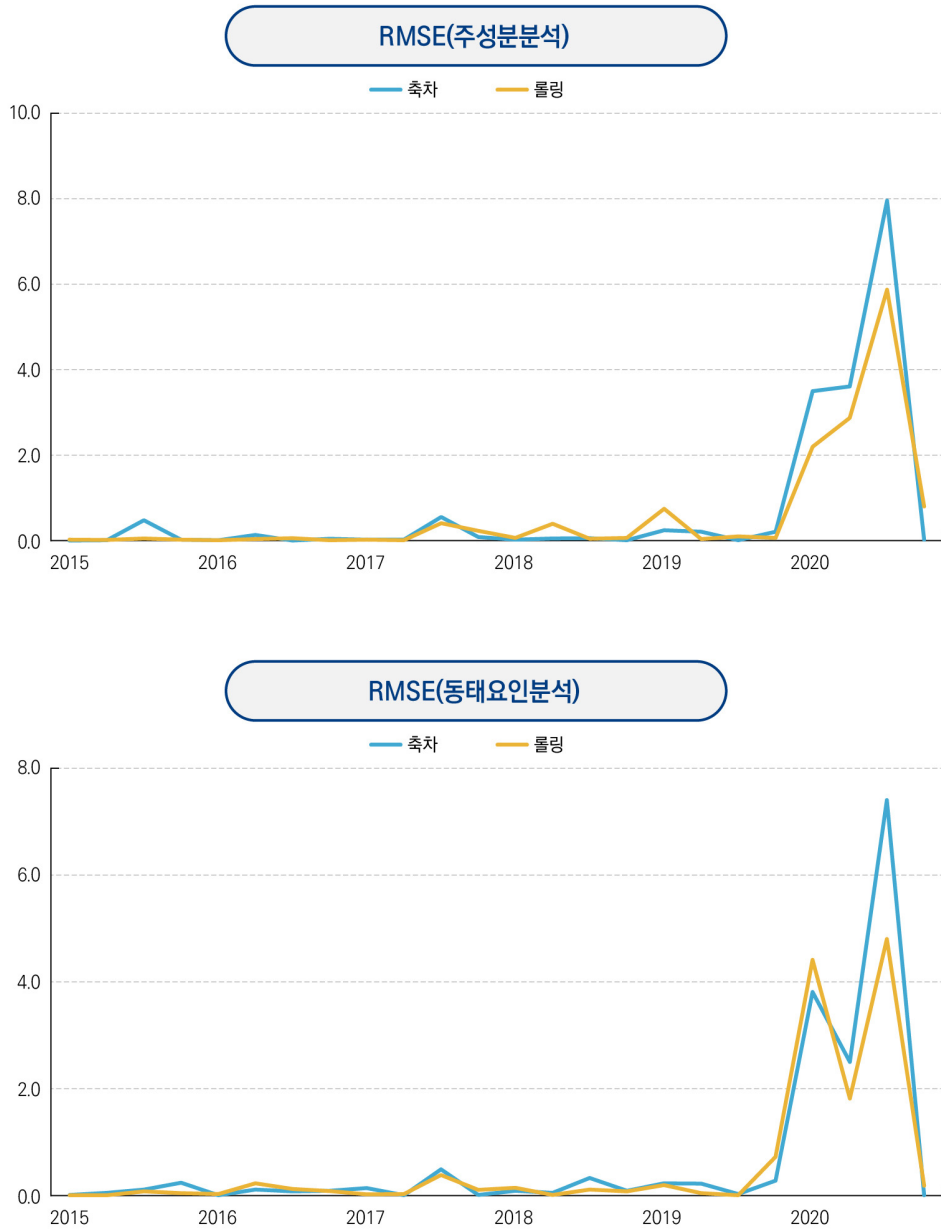
25) 각 시점별로 AIC, BIC 기준에 따라 최적시차를 결정

RMSE가 큰 폭으로 상승하였다.

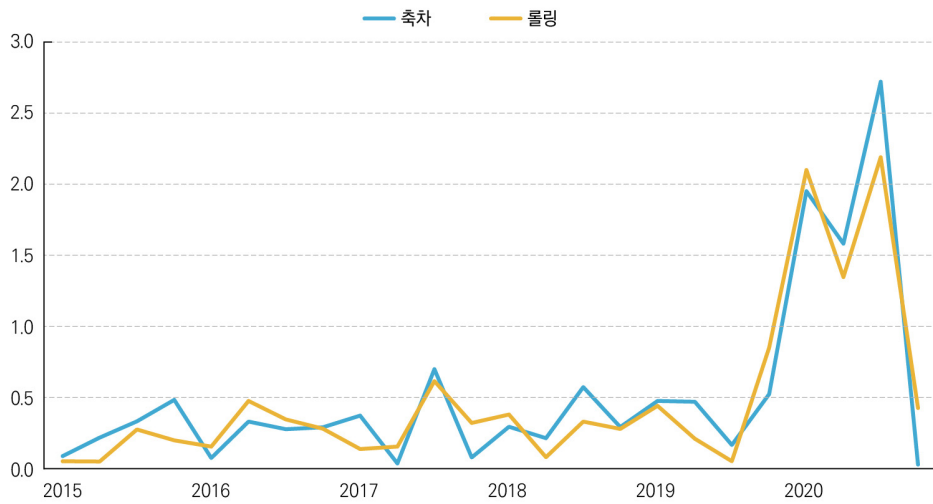
예측치를 추정하기 위해서는 기본적으로 과거의 자료가 미래를 예측하는데 유용한 정보를 지니고 있다고 가정하고 있다. 그러나 예측하고자 하는 변수의 구조변화(structural break) 등이 나타나 과거의 움직임과 다른 방향을 보인다면 예측오차가 크게 나타날 수 있다. 2020년은 코로나19의 충격으로 1/4분기 GDP 성장률이 -1.3%로 하락하였으며, 2/4분기에는 -3.2%로 하락폭이 확대되었다. 그러나 3/4분기 중에는 2.1%로 반등하며 GDP성장률이 큰 폭의 변동성을 보였다. 2020년은 과거의 GDP성장률과 추세적으로 괴리되어 있으며, 주성분분석 결과 주요 정보변수들이 차이를 나타내고 있어 월별 지표들의 움직임도 과거 데이터와는 다른 패턴을 보이고 있다고 추정된다. 이와 같은 큰 변동성, 추세적 이탈로 인해 추정치가 실적치를 벗어나면서 예측오차가 크게 나타났고 RMSE가 커진 것으로 분석된다. [그림 16]은 예측기간 중 모형별, 측정방법별 예측오차 추이를 나타내었다. 모형이나 측정방법과 상관없이 2020년의 경우 예측오차가 매우 커지는 것으로 나타나 이 기간을 포함할 경우의 RMSE 및 MAE가 높아지게 된다.

2020년 자료를 포함하는 경우 예측오차가 커지는 것으로 나타났으나, 벤치마크 모형인 분기데이터를 이용한 AR모형에 비해서는 2019년까지의 분석과 마찬가지로 예측력이 우수한 것으로 나타났다. 또한, 2019년까지 예측결과와 마찬가지로 축차방식 보다는 롤링방식이, 주성분분석 보다는 동태요인분석모형이 모형의 예측력을 개선시키는 것으로 나타났다.

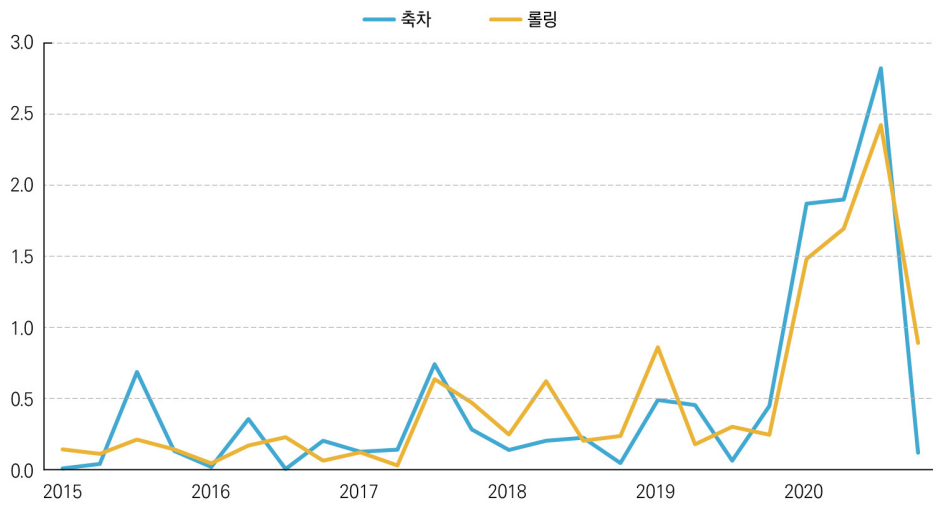
[그림 16] 당분기 예측의 모형별, 측정방법별 예측오차 추이



MAE(주성분분석)



MAE(동태요인분석)



IV. 요약 및 시사점

본 연구의 목적은 빠르게 입수할 수 있는 월별자료를 이용하여 분기 GDP성장률을 미리 예측할 수 있는 단기예측모형을 개발하는 것이다. 예측방식, 예측기간, 전망모형별로 예측성과를 비교하여 현재의 경제상황을 측정할 수 있는 가장 대표적인 변수라 할 수 있는 실질GDP성장률에 대한 전망성과를 개선시킬 수 있는 방안을 찾고자 하였다. 특히, 해당분기 중 초월, 중월, 말월로 갈수록 입수할 수 있는 정보가 많아지는 점을 감안하여 분기내에서도 첫 번째~세번째 달 각각에 대해 이용 가능한 정보를 통해 GDP성장률을 예측하는 과정을 거쳤다.

2000~2019년까지의 월간지표와 2019년 4/4분기까지의 실질GDP성장률 자료를 이용하여 주성분분석모형과 동태요인분석모형을 통해 성장률 예측치를 추정 한 결과 분기자료만을 이용한 AR모형에 비해 예측오차가 작게 나타나 월별자료를 이용한 전망모형이 예측력을 개선시키고 있는 것으로 나타났다.

전망모형별로는 월별자료의 공통정보를 이용하여 예측한다는 측면에서 주성분분석과 동태요인분석모형이 동일한 메커니즘을 따르고 있으나, 결측치를 처리하는 방식에서 AR모형을 이용하는 주성분분석과 자기회귀를 통해 공통요인의 동태적 특성을 분석하고 칼만스무딩(Kalman smoothing) 방법을 이용하는 동태요인분석모형 간에 차이가 있다. 분석결과 동태요인분석의 예측력이 보다 우수한 것으로 나타났으며, 이는 공통요인의 동학을 추정하여 결측치를 처리하는 동태요인 방식이 보다 유용한 정보를 포함하고 있기 때문으로 판단된다.

예측방식 중에서는 처음 정해진 구간을 기본으로 하여 기간을 추가하면서 추정하는 축차방식에 비해 표본의 수를 일정하게 유지한 채 데이터를 따라 이동하면서 예측치를 도출하는 롤링방식이 상대적으로 전망성과가 우수한 것으로 나타났다. 이는 분석기간에 경제의 성장추세에 구조적 변화를 가져올 수 있는 2008년 금융위기가 포함되어 있어 축차예측방식의 경우 구조변화의 지속적 영향을 받

게 되어 예측력이 상대적으로 낮아진 것으로 판단된다.

분기 중 초월, 중월, 말월 각각의 예측결과 전체월(초월, 중월, 말월의 결과를 평균)과 초월, 중월, 말월의 예측오차는 크게 달라지지 않는 모습이다. 이는 초월에서 말월로 이동하면서 모형의 모수 추정치가 크게 변하지 않았고 예측에 사용된 월별 변수 결측치에 대한 예측치도 어느 정도 정확성을 확보하고 있기 때문으로 판단된다.

코로나19의 충격으로 경제성장률의 변동성이 매우 확대된 2020년을 분석에 포함할 경우 도출된 결론이 동 기간에 좌우될 가능성 등을 고려하여 2019년까지를 기본 분석기간으로 설정하되 2020년을 포함한 분석도 수행하였다. 분석결과 당초 예상과 같이 GDP성장률이 큰 폭으로 변동하고 있는 2020년을 포함할 경우 예측오차가 매우 높아지는 것으로 나타났다. 그러나, 벤치마크 모형인 분기데이터를 이용한 AR(Autoregressive Mode)모형에 비해서는 2019년까지의 분석과 마찬가지로 예측력이 우수한 것으로 분석되었다. 또한, 2019년까지 예측결과와 마찬가지로 추차방식 보다는 롤링방식이, 주성분분석 보다는 동태요인분석모형이 모형의 예측력을 개선시키는 것으로 나타났다.

본 연구에서는 월별자료를 이용하여 주성분분석과 동태요인분석모형을 토대로 GDP성장률을 전망하고 예측결과를 비교하였으며, 구축된 예측모형을 통해 한국은행 속보치보다 1~2개월 빠르게 실질GDP성장률 예측치를 생성할 수 있었다. 한국은행 발표 전에 분기별 실질GDP성장률을 생성함으로써 경제전망의 정확도를 제고할 수 있을 것으로 예상되며, 신뢰성 있는 예측치를 제공함으로써 경제정책 방향을 결정하는 기초자료로 활용될 수 있을 것으로 기대된다.

다만, 가장 최근 자료이지만, 불확실성이 매우 높았던 2020년을 분석에 포함할 경우 예측오차가 커져 예측력이 낮아지는 모습을 보였다. 2020년은 코로나19의 영향으로 변동성이 매우 높을 뿐만 아니라 이전 GDP성장률과도 추세적으로 괴리되는 모습을 보이고 있어 예측에 어려움이 있다. 향후 시계열이 충분히 확보되면 2020년에 대해 structural break 등을 설정하여 모형을 재구축하고 예측오차를 평가해 볼 필요가 있겠다.

또한, 본 연구에서는 GDP의 속보치, 잠정치, 확정치를 구분하지 않고 시계열이 구성되어 있는 확정치를 사용하였다. 실제 예측과정에서는 초월의 경우 전분기 GDP성장률의 속보치만 가용할 수 있기 때문에 속보치와 확정치의 차이 등에 따른 오차가 발생할 가능성이 있으며, 이에 따라 초월의 예측오차를 과소평가하였을 가능성도 있기 때문에 이를 고려한 연구도 필요할 것으로 보인다.

참고문헌

- 김덕파, 강성진, 2020, “월별지표를 이용한 GDP추정을 위한 계량모형 구축”, 국회 예산정책처 연구용역보고서
- 김양우, 이금희, 장동구, 1997, “한국의 단기경제예측시스템”, 經濟分析, 3(3), 1-61.
- 김종화, 이중식, 2005, “인플레이션 전망에 있어서 정보변수의 역할”, 經濟分析, Vol. 11 No. 3, 34-58.
- 김치호, 김현정, 2016, “GDP 성장률의 Nowcasting에 관한 연구”, [Bok] 국민계 정리부, 2016년 제2호, 1-23
- 남준우, 이한식, 2005, 계량경제학, 2nd ed., 홍문사
- 박강우, 이중식, 2009, “GDP 예측력 제고를 위한 동태요인모형 구축”, 한국은행 조사통계월보, 제63권, 27-60
- 이진희, 김덕파, 2014, “한국 경제변수에 대한 자기회귀 및 벡터자기회귀 모형의 예측성과 비교”, 經濟分析, Vol. 20 No. 4, 115-151.
- 정갑영, 강전은, 2007, “구성분분석을 이용한 북한의 총국민소득(GNI) 추정”, 비교 경제연구, Vol. 14 No. 2, 1-33.
- 진익, 2020, 「경제구조 변화에 대응한 지속성장 과제」, 국회예산정책처
- 최광신, 2018, “K-SuperCast:빅데이터를 이용한 GDP성장률 적시예측 시스템”, FSS Working Paper
- 최병재, 한진현, 2014, “월별자료를 이용한 예측조합에 의한 GDP 전망”, 한국은행 「조사통계월보」, 제68권 제10호, 16-48.
- 통계청, 「『경기종합지수』통계정보 보고서」, 2020.6
- 한규숙, 형남원, 2019, “동태요인모형을 이용한 미국 주택투자 예측”, 주택연구, Vol. 27 No. 4, 35-68
- 한국은행, 2006, “분기 경제성장률 주지표 전환”, 「계간 국민계정」, 제1호, 106~140
- 한국은행, 2015, 「우리나라의 국민계정체계」

- Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L. & Rünstler, G., 2011, "Short-term forecasts of euro area GDP growth", *Econometrics Journal*, Vol. 14 No. 1, C25-C44.
- Arnoštová, K., Havrlant, D., Růžička, L. & Tóth, P., 2010, "Short-Term Forecasting of Czech Quarterly GDP Using Monthly Indicators", .
- Baffigi, A., Golinelli, R. & Parigi, G., 2004, "Bridge models to forecast the euro area GDP", *International Journal of Forecasting*, Vol. 20 No. 3, 447-460.
- Bai J., 2003, "Inferential Theory for Factor Models of Large Dimensions", *Econometrica*, Vol. 71 No. 1, 135-171.
- Barhoumi, K., Benk, S., Cristadoro, R., Den Reijer, A., Jakaitiene, A., Jelonek, P., Rua, A., Rünstler, G., Ruth, K. & Nieuwenhuyze, C. V., 2008. Short-term forecasting of GDP using large monthly datasets-A pseudo real-time forecast evaluation exercise. National Bank of Belgium Working Paper, (133).
- Barhoumi, K., Darné, O. & Ferrara, L., 2010, "Are disaggregate data useful for factor analysis in forecasting French GDP?", *Journal of Forecasting*, Vol. 29 No. 1, 132-144.
- Boivin J., Ng S., 2005, "UNDERSTANDING AND COMPARING FACTOR-BASED FORECASTS", Nber Working Paper Series, Vol. - No. 11285.
- D' Agostino A., Giannone D., 2012, "Comparing Alternative Predictors Based on Large-Panel Factor Models", *Oxford Bulletin of Economics and Statistics*, Vol. 74 No. 2, 306-326.
- Diron M., 2008, "Short-term forecasts of euro area real GDP growth: an assessment of real-time performance based on vintage data", *Journal of Forecasting*, Vol. 27 No. 5, 371-390.
- Doz, C., Giannone, D. & Reichlin, L., 2011, "A two-step estimator for large approximate dynamic factor models based on Kalman filtering", *Journal of Econometrics*, Vol. 164 No. 1, 188-205.
- Eickmeier S., Ziegler C., 2008, "How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach", *Journal of Forecasting*, Vol. 27 No. 3, 237-265.
- Forni, M., Hallin, M., Lippi, M. & Reichlin, L., 2005, "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting", *Journal of the American Statistical Association*, Vol. 100 No. 471, 830-840.

- Giannone, D., Reichlin, L. & Small, D, 2008, "Nowcasting: The real-time informational content of macroeconomic data", *Journal of Monetary Economics*, Vol. 55 No. 4, 665-676.
- Hara, N., Yamane, S., 2013. New monthly estimation approach for Nowcasting GDP growth: The case of Japan (No. 13-E-14). Bank of Japan.
- Stevens, J., 1986. *Applied multivariate statistics for the social sciences*. Lawrence Erlbaum Associates, Inc.
- Kaiser H. F., 1960, "The Application of Electronic Computers to Factor Analysis", *Educational and Psychological Measurement*, Vol. 20 No. 1, 141-151.
- Mehrotra A., Pääkkönen J., 2011, "Comparing China's GDP statistics with coincident indicators", *Journal of Comparative Economics*, Vol. 39 No. 3, 406-411.
- Mingione F., 2011, "Forecasting with Principal Components Analysis: An Application to Financial Stability Indices for Jamaica". Technical Report. Bank of Jamaica. Jamaica.
- Patrick Higgins, 2014, "GDPNow: A Model for GDP "Nowcasting", Federal Reserve Bank of Atlanta Working Paper 2014-7
- Stock J. H., Watson M. W., 2016, "Chapter 8 - Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics", *Handbook of Macroeconomics*, Vol. 2, 415-525.
- Stock J. H., Watson M. W., 2002, "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business & Economic Statistics*, Vol. 20 No. 2, 147-162.

[부 록] 국내외 기관의 단기 예측 모형 운용현황

1. 금융감독원 K-SuperCast²⁶⁾

금융감독원은 2018년 국내 경제 상황을 신속하게 파악하여 거시 건전성 감독을 강화하기 위해 빅데이터 기반의 GDP성장률 예측 모형인 K-SuperCast(Korea Supervisory Forecast)를 개발하였다. K-SuperCast는 금융감독원에서 자체적으로 실시하는 스트레스 테스트 시 국내 경제 변수에 대한 가정 등 향후 시나리오 마련에 활용하고자 개발되었다. 금융감독원은 K-SuperCast를 포함한 거시건전성 감독 3종세트(KOMPAS)²⁷⁾로 구성된 ‘거시건전성 감독 분석 체계’를 구축하여 2018년 7월부터 운용하고 있다. 다만, 2020년 코로나19 확산 이후 수집되는 기초통계 데이터의 급격한 변동으로 인해 운용을 잠시 중단한 상황이며, 지난해와 같은 예상치 못한 경제충격을 모형에 반영하기 위해 보완 중에 있다.

K-SuperCast 모형은 2008년 이후의 거시경제 시계열 자료 83개를 주성분 분석을 이용하여 3개의 팩터(factor)로 선별하고 각 거시경제 변수의 발표 주기가 달라 발생하는 데이터 공백을 없애고 연속적인 데이터 추정을 위해 칼만필터(Kalman Filtering) 기법을 적용하였다.

주성분분석과 칼만필터를 사용한다는 점에서 본고의 모형과 유사한 과정을 따르고 있으나, 월별과 분기 데이터를 함께 사용하고 있다는 점에서 차이가 있다. K-SuperCast 모형에서는 분기단위 데이터를 모형 구축 과정에서 월별 데이터로 변형한 후 다른 월단위 데이터와 함께 주성분 요인을 추출하게 되며, 이 과정에서 기본적으로 기하평균을 이용하고 있다. 예측하고자 하는 변수인 실질GDP 또한 Mariano & Murasawa(2003)의 방법을 이용하여 월단위 데이터로 변환하여 주성분과 OLS 분석을 통해 계수를 추정하였다.

26) 금융감독원 보도자료, “금융시스템의 불안 요인에 선제 대응을 목적으로 「거시건전성 감독 분석체계(KOMPAS)」를 확립”, 2018.12 참고

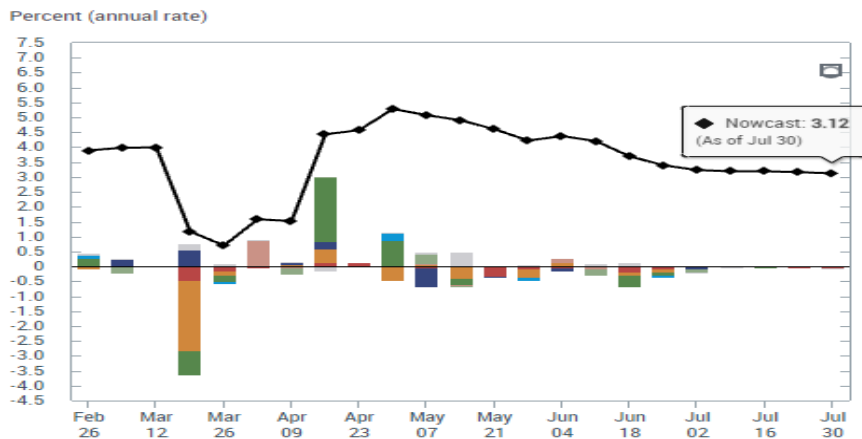
27) KOMPAS는 거시건전성 감독 스트레스 모형(K-STARS), 금융산업 조기경보 모형(K-SEEK), GDP 성장률 예측모형(K-SuperCast) 등 총 세 개의 모형으로 구성

예측결과 GDP성장률을 결정하는 주성분의 숫자는 3개로 팩터1은 수출 및 수입수지와 설비가동수지, 기업 경기지수 등이 포함되었으며, 팩터2는 매매와 전세시장을 보여주는 지표, 정부 재정지출과 임금 등이 포함되었다. 팩터3은 임금 및 노동관련 변수로 임금 및 노동시장을 반영하고 있어 팩터1, 2, 3을 각각 기업 부문, 건설 및 부동산, 임금 및 노동으로 분류하였다. 이와 같이 추출된 팩터를 이용하여 2015년 4/4분기~2018년 3/4분기까지 12분기에 대한 당분기 예측을 수행한 결과 평균오차는 0.198%p로 나타났다(최광신(2018)).

2. FRB NY Nowcast

뉴욕 연방준비은행(Federal Reserve Bank of New York)은 광범위한 거시데이터를 활용하 칼만 필터링과 동적요인모형(Dynamic Factor Model)을 통해 단기 GDP를 예측하고 있다. 198개의 거시변수들을 이용하여 10개의 요인을 추출하고 이를 GDP성장률과 회귀분석하여 성장률을 예측하였다. Nowcast의 GDP성장률은 매주 금요일(연방 공휴일 제외) 오전 10시까지 사용 가능한 데이터를 활용하여 오전 11시 15분에 업데이트된다.

[그림] Nowcast 모형을 이용한 2021년 2/4분기 시점별 GDP성장률 예측치



자료: FRB NY

Nowcast 모형을 통한 GDP성장률 예측치는 2014년 6월부터 매주 발표되었으나, 뉴욕 연방준비은행은 9월 3일 코로나19 팬더믹을 둘러싼 불확실성과 지

표의 변동성이 높아지고 있다고 판단하여 Nowcast 모델에 대한 방법론적 개선 작업을 계속하는 동안 발표를 잠정 중단하겠다고 밝혔다.

3. FRB Atlanta GDPNow

애틀랜타 연방준비은행(Federal Reserve Bank of Atlanta)는 2014년 7월부터 GDPNow라는 단기예측모형을 통해 월별로 발표되는 13개의 예측변수를 활용하여 이들의 합으로 해당분기의 GDP를 예측하고 있다. 즉, 예측력이 우수한 것으로 나타난 GDP의 하위 구성 요소들을 추적하여 미국 주요 경제지표가 발표되고 약 1.5~2시간 후 GDP 예측치를 업데이트 하고 있다. ISM 제조업 구매자지수, 승용차 판매, 무역, 고용지표, 소비자물가지수, 미 재무부 의회 제출 월간 보고서, 생산자물가지수, 수출입물가지수, 산업생산, 소매판매, 주택착공지수, 기존주택 판매, 신규주택판매, 내구재주문 등 지표 발표시마다 GDP 예측치를 발표하고 있다.

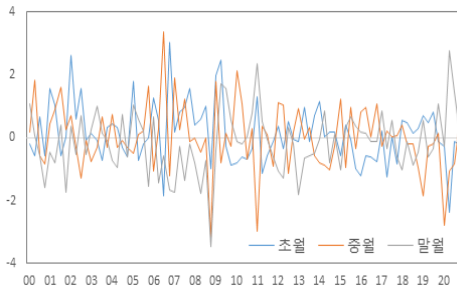
GDPNow 모델은 GDP 발표 기관인 미국 경제분석국(U.S. Bureau of Economic Analysis)의 연쇄 가중치 방식을 차용하여 GDP를 구성하는 13개의 하위요소²⁸⁾를 집계하여 성장률을 예측하고 있다. GDP성장률과 13개 하위요소에 대해 BVAR을 이용하여 예측하는 방식을 취하고 있으며, BEA의 연쇄가중법을 이용하여 구성요소별로 분기 예측을 하여 이를 GDP 예측과 결합하여 전망한다. 또한, 월별데이터를 분기데이터와 연결시키기 위해 Giannone et al.(2008)의 방식을 차용한 bridge 모형을 이용하였다(Patrick Higgins(2014)). 2011년 GDP성장률 추정 이후 실질GDP와의 예측오차는 MAE 0.81%p, RMSE 1.23%p로 나타났다.

FRB Atlanta는 코로나19 확산 이후에도 여전히 GDP성장률 예측치를 발표하고 있으나, 코로나19와 사회적 이동이 이미 발표된 경제 지표나 GDP 원천 자료에 미치는 영향을 포착하지 못한다고 밝혀 경제 충격에 따른 예측의 어려움을 인정하였다.

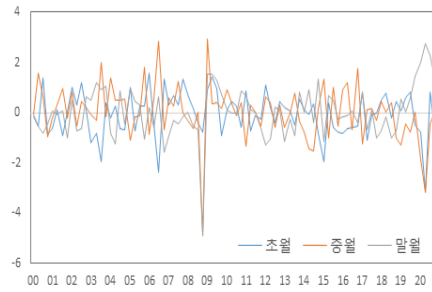
28) 개인 상품 소비지출, 개인 서비스 소비지출, 실질 민간 설비투자, 지적재산투자, 구조물 투자, 실질 민간 주거용 투자, 실질 상품 수출, 실질 서비스 수출, 실질 상품 수입, 실질 서비스 수입, 실질 정부 소비 및 투자, 실질 지방정부 소비 및 투자, 재고증감

[부 록] 월별자료 추이

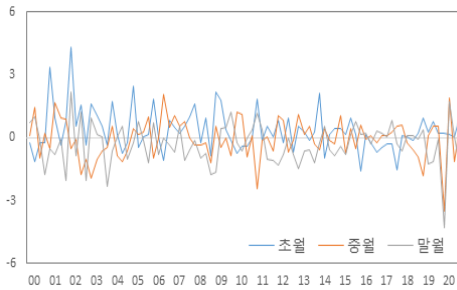
전산업생산지수



제조업생산지수



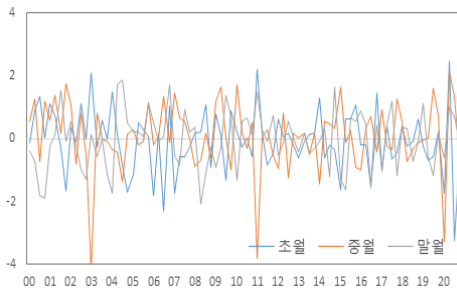
서비스업생산지수



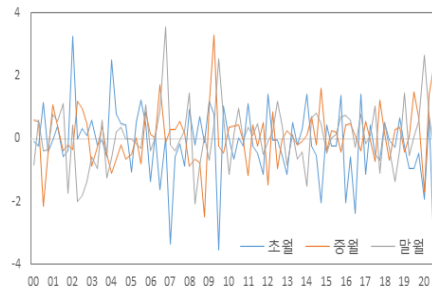
제조업가동률지수



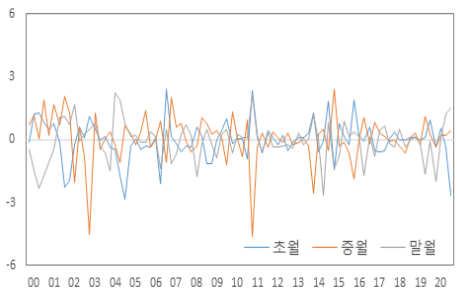
소매판매지수



내구재판매지수



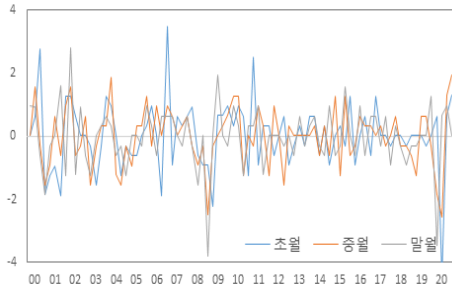
비내구재판매지수



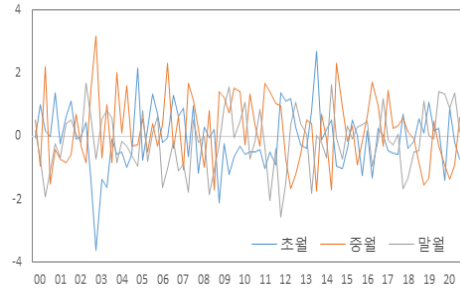
경기선행지수순환변동치



경기동행지수순환변동치



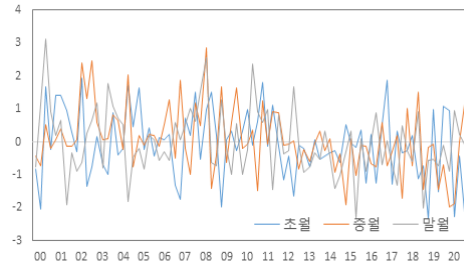
설비투자지수



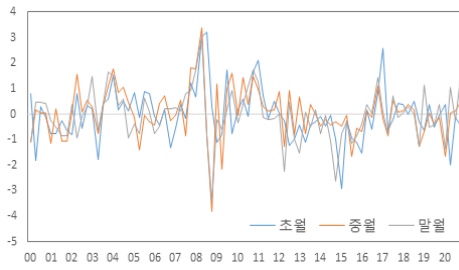
건설기성액



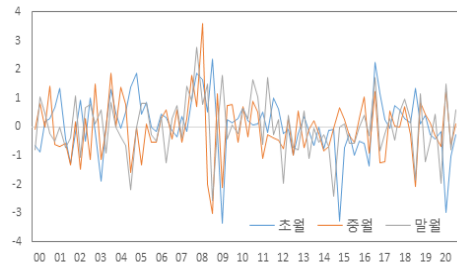
소비자물가지수



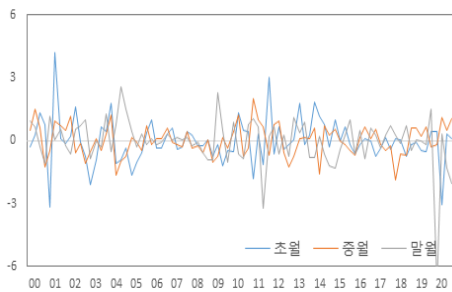
생산자물가지수



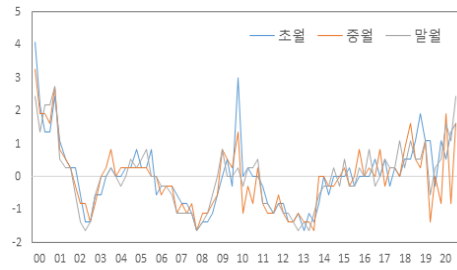
수입물가지수

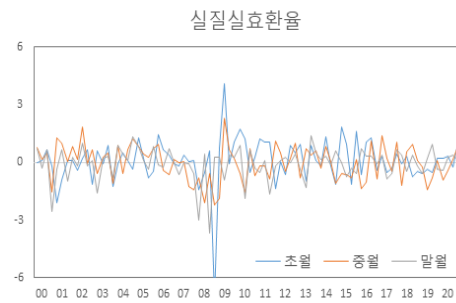
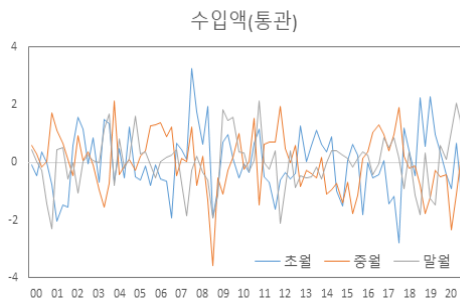
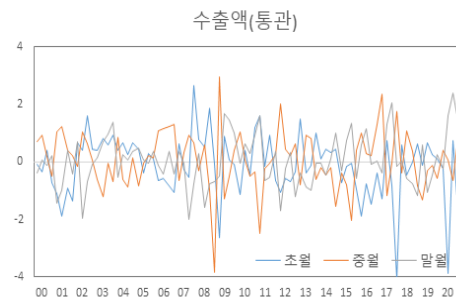
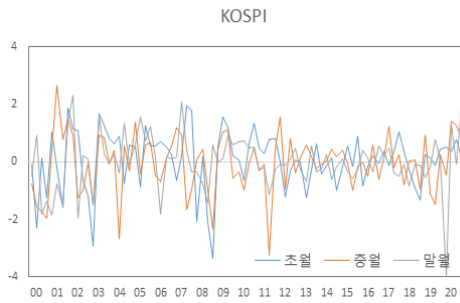
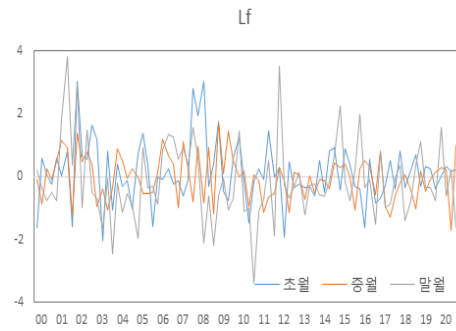
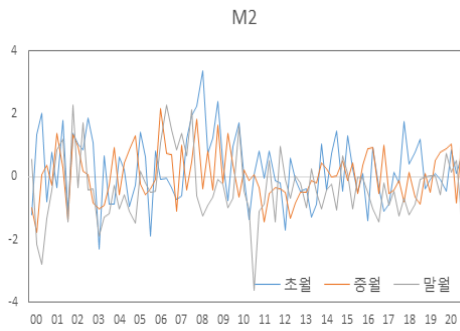
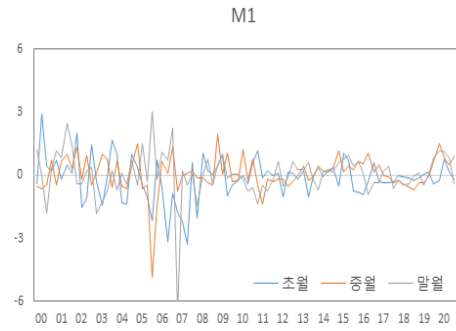
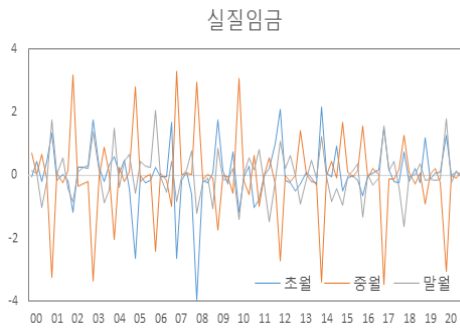


취업자수



실업률





경제현안분석 목록

	제 목	집 필	발 간
103	분위별 자산소득 분포 분석 및 국제비교	권일, 김미애	2021. 7
102	경기조정 재정수지와 잠재성장률 추정	황중률	2021. 6
101	NABO 내국인 인구 시범추계: 2020~2040년	김경수·김상미	2021. 3
100	주요 주력산업과 신산업의 동향 및 수출경쟁력분석	김상우·김미애· 허가형·권일· 최세중	2020. 12
99	혁신성장 전략투자의 현황 및 경제적 파급효과 분석	김상우·최세중	2020. 10
98	주요 산업별 수출의 경쟁력 및 경제적 기여도 분석	김상우·김미애· 최세중·신동진· 권 일	2019. 12
97	중국경제 현안 분석 -부채·부동산·그림자금융을 중심으로	김윤기·황중률· 오현희	2018. 12
96	5대 신산업 선도 프로젝트의 추진 현황과 정책효과 분석	김상우·신동진· 김미애·권 일· 장아련	2018. 12
95	북한 경제개발 재원조달을 위한 국제기구와의 협력방안	진 익·모주영· 박승호·조은영	2018. 12
94	우리나라 저출산의 원인과 경제적 영향	김경수·허가형· 김윤수·김상미	2018. 10
93	내수활성화 결정요인 분석	김윤희·진 익	2017. 12
92	주택가격 변화가 가계부채와 금융 안정성에 미치는 영향	현영진	2016. 11
91	아동 관련 복지분야의 조세지출과 재정지출 지원 현황 및 시사점	채은동	2016. 6
90	일본의 장기침체기 특성과 정책대응에 관한 연구	김윤기·유승선· 황중률·오현희	2016. 4
89	조세지출제도 국내외 동향 및 시사점	채은동·이영숙	2015. 9
88	글로벌 금융위기 이후 OECD 국가들의 세계개편 동향 연구	이영숙	2015. 2
87	취득세율 인하가 주택거래 및 지방재정에 미치는 영향	채은동·태정림	2015. 2

	제 목	집 필	발 간
86	북유럽 국가의 금융·재정위기 극복과 시사점	조은영	2014. 4
85	우리나라 투자지원 조세제도 현황과 주요국 제도와의 비교 연구	이영숙	2013.12
84	해외 주요국의 재정준칙 운용동향과 정책시사점	김정미·이강구	2013. 9
83	담배가격 인상에 따른 재정 영향 분석	신영임·서재만	2013. 7
82	가계부채의 현황 및 대응방안	신동진	2013. 7
81	소액주주 주식양도소득세 도입방안 및 세수효과분석	채은동	2013. 5
80	남부유럽재정위기가 국내외 경제에 미치는 영향	신후식·유승선	2012.10
79	고령화가 근속 및 연공임금체계에 미치는 영향과 정책 시사점	장인성	2012. 9
78	고령자 일자리 현황과 정책과제	서재만	2012. 9
77	파생금융상품에 대한 거래세 도입에 관한 연구	채은동	2012. 8
76	국민연금 장기 지속가능성 확보방안	김대철·심혜정	2012. 8
75	공적자금 상환대책의 현황 및 개선방안	신동진	2012. 8
74	세무조사 운영실태의 문제점과 개선방안	심혜정	2012. 7
73	2012 근로장려세제 확대시행의 소요재정과 분배효과	장윤정	2012. 7
72	출산·보육지원 재정소요 추계와 정책과제	조은영	2012. 6
71	발생주의 회계제도 도입이 세입 결산에 미치는 영향	신영임·장윤정	2012. 6
70	외국자본의 조세회피 방지를 위한 합리적 과세방안	최천규	2012. 5
69	신용카드 소득공제의 소득계층별 귀착 및 세수효과와 시사점	성명기	2011.12
68	경제성장률 단기예측 모형 - 베이지언 VAR 접근방식에 의한 예측 -	황종률	2011.12
67	재정통계 개편의 주요 쟁점과 과제	윤용중, 이강구, 윤준승, 서재만, 김정미	2011.11
66	소득계층별 물가지수의 차이가 체감물가에 미치는 영향	장인성	2011.11
65	자영업자 현황 및 정책 방향	서재만	2011. 9
64	통일비용에 대한 기존연구 검토	신동진	2011. 8
63	2010년 결산상 재정통의 문제점과 개선방안	윤준승	2011. 8

	제 목	집 필	발 간
62	재정의 경제안정화 효과 분석 - 자동안정화장치를 중심으로 -	박승준·이강구	2011. 6
61	한·중 신재생에너지 정책 비교와 시사점	원동아	2011. 2
60	고령화가 생산성 및 경제성장에 미치는 영향	장인성	2010.12
59	위안화 절상의 영향과 시사점	신후식·유승선	2010.12
58	재정정보 공개 현황 및 개선방안	서재만	2010.12
57	2010년 세법개정안의 세수효과	이영환·신영임	2010.10
56	조세법률주의 위반 사례 및 개선방향	윤준승, 정지은, 이남수	2010. 9
55	가계부채의 문제점과 정책개선방안	신동진	2010. 7
54	경기선행지수의 향후 경기에 관한 시사점	유승선	2010. 7
53	남유럽 재정위기와 정책시사점	김정미	2010. 7
52	외평기금 이자비용 처리문제로 본 통합재정통계의 문제점 및 개선방안	심혜정	2010. 6
51	경제위기와 각국의 조세정책 동향 및 시사점	신영임·이영환	2010. 5
50	외화예산의 환위험 관리방안	연훈수	2010. 2
49	국가재정운용계획의 평가 및 과제	나아정·박승준	2009.12
48	소득격차의 확대와 재분배 정책의 효과	장인성	2009.12
47	금융위기와 한국의 잠재성장률	황중률	2009.12
46	사회복지 기능 확대에 따른 지방재정 영향 분석	심혜정	2009.12
45	2009년 말 일몰도래 비과세·감면항목 운용현황	정지은	2009.11
44	비과세·감면제도 운용현황 및 개선과제	정지은	2009.11
43	지방소득세·지방소비세 도입과 향후 과제	이영환, 황진영,신영임	2009.11
42	금융안정화대책의 정책효과와 출구전략의 방향	신동진	2009.11
41	2008년 이후 세계개편의 세수효과	이영환·신영임	2009. 8
40	글로벌 금융위기 극복을 위한 금융정책 분석	신동진	2009. 7
39	재정확대의 거시경제적 효과분석	박승준	2009. 4
38	경제위기의 전개와 대응	신후식, 유승선, 연훈수	2009. 3
37	우리나라 외환금융시장 취약성 비교 분석	신후식, 유승선, 연훈수	2008.12
36	중국의 기업소득세법 제정에 따른 입법적 시사점 검토	황진영	2008.12

	제 목	집 필	발 간
35	지방정부 재정자주권의 국제비교와 정책적 시사점	심혜정	2008.12
34	2008년 세제개편안 분석 : 목적세 정비안을 중심으로	이영환·정지은	2008.11
33	2008년 세제개편안 분석 : 비과세·감면제도를 중심으로	정지은	2008.11
32	2008년 세제개편안 분석 : R&D지원 강화를 중심으로	이상훈	2008.11
31	2008년 세제개편안 분석 : 종합부동산세	이영환·신영임	2008.11
30	OECD 주요국가 초과세수 발생과 재정규율 사례	이남수·이성규	2008. 9
29	유가환급금 지급(안) 평가	정지은, 홍인기, 전승훈	2008. 9
28	청년층 고용현황과 시사점	정상훈·이충언	2008. 8
27	환율변동이 국내물가에 미치는 영향	연훈수	2008. 4
26	목적세와 특별회계의 문제점과 개편방향	이영환, 이성규	2008. 1
25	은행산업의 경쟁도 분석과 정책적 시사점	신동진	2007.12
24	「강제집행등과 체납처분의 절차조정법」의 입법 필요성 검토	황진영	2007.12
23	원화가치 변동이 수출가격에 미치는 영향	성명기	2007.11
22	한국의 실질 GDP 장기 예측 : 2007~2050년	장인성	2007.11
21	세법체계 개편작업의 동향분석	황진영	2007. 9
20	중소기업 신용보증제도의 운영성과와 개선방안	정상훈	2007. 9
19	과세정보 공개제도의 현황	문성환	2007. 8
18	구조조정 이후 은행산업의 효율성 분석	신동진	2007. 8
17	한국의 잠재성장률과 자연실업률 추정	황종률	2007. 7
16	유가 상승의 원인 및 유류세 인하를 둘러싼 쟁점 분석	이영환·전승훈	2007. 7
15	미국 기준선전망의 의의와 우리 예산과정에 대한 시사점	정문중	2007. 6
14	최근 일본의 재정개혁과 시사점	이남수·서세욱	2007. 6
13	물가상승에 의한 소득세 부담 증가 완화를 위한 정책대안 : 소득세 물가연동제에 대한 검토	전승훈	2007. 5
12	DDA 농업협상의 논의동향 및 영향에 대한 고찰	송원근	2006. 6
11	미국의 재정개혁 논의동향과 시사점	정문중	2006. 6

	제 목	집 필	발 간
10	2000~2005년 경제예측의 경험과 단기예측 방식의 개선방향	유승선	2006. 5
9	퇴직연금세제 관련 현안분석과 개선방향	문성환	2006. 5
8	자영업 진출 결정요인과 정책적 시사점	김기승	2006. 2
7	분야별 자원배분에 대한 국제비교 연구	전승훈	2006. 1
6	주택가격 안정을 위한 정책현황 및 과제	송원근	2005.12
5	국세행정예 대한 새로운 감독체제의 모색	문성환	2005. 7
4	재정 건전성 강화를 위한 재정규율의 확립- 지출상한선을 중심으로	정문종	2005. 6
3	일자리 창출정책의 현황과 과제	김기승	2005. 5
2	조세지출예산제도와 정책과제	전승훈	2004.12
1	재정지출 확대와 감세의 경제적 효과 분석	김기승, 임일섭, 전승훈	2004.10

단기 GDP 예측모형 연구

발간일	2021년 12월 10일
발행인	임익상 국회예산정책처장
편 집	경제분석국 거시경제분석과
발행처	국회예산정책처 서울특별시 영등포구 의사당대로 1 (tel 02·2070·3114)
인쇄처	경성문화사(tel 02·786·2999)

이 책은 국회예산정책처 홈페이지(www.nabo.go.kr)에서
보실 수 있습니다.

ISBN 979-11-6799-034-1 93350

© 국회예산정책처, 2021

내일을 여는 국민의 국회



(07233)서울특별시 영등포구 의사당대로 1
Tel. 02-2070-3114 www.nabo.go.kr

발 간 등 록 번 호

31-9700496-001921-01



국회에산정책처
NATIONAL ASSEMBLY BUDGET OFFICE