



# BOK 경제연구



빅데이터를 이용한  
실시간 민간소비 예측

신승준, 서범석



2022. 9





# 빅데이터를 이용한 실시간 민간소비 예측

신승준\*, 서범석\*\*

이 연구내용은 집필자 개인의 견이며 한국은행의 공식 견해와는 무관합니다. 따라서 본 논문의 내용을 보도하거나 인용할 경우에는 집필자명을 반드시 명시하여 주시기 바랍니다.

\* 고려대학교 통계학과 부교수 (전화: 02-3290-2247, E-mail: sjshin@korea.ac.kr)

\*\* 한국은행 경제통계국 통계연구반 과장 (전화: 02-759-5253, E-mail: bsseo@bok.or.kr)

본 논고의 분석 및 작성에 커다란 도움을 주신 심정은, 김희준 고려대학교 연구원께 깊은 감사의 뜻을 표합니다. 또한 본 논고에 유익한 논평을 주신 이인규 경제통계국 지출국민소득팀 팀장, 김남주 조사국 동향분석팀 차장께 감사의 말씀을 드립니다. 본 논고에 혹시 남아있을 수 있는 오류는 저자의 책임임을 밝힙니다.

## 차 례

I. 서론 .....	1
II. 배경 및 관련 연구 리뷰 .....	2
III. 방법론 .....	6
IV. 실증 분석 .....	11
V. 결론 및 제언 .....	19

## 빅데이터를 이용한 실시간 민간소비 예측

최근 코로나19 등으로 경제 불확실성이 확대됨에 따라 민간 경제주체의 경제상황을 직접적으로 반영하는 민간소비 동향을 신속히 파악할 필요성이 높아지고 있다. 이에 본 연구는 기존 거시경제지표 뿐만 아니라 빅데이터를 종합적으로 활용하여 민간소비를 실시간으로 추정(Nowcasting)하는 방법을 제안하였다. 특히 초고차원 빅데이터의 적합을 위해 활용 가능한 다양한 기계학습 방법론을 비교분석하여 민간소비 추정의 정확도를 향상시키고자 하였다. 실증 분석 결과, 빅데이터를 비롯한 가용 공변량의 수가 많은 경우에는 변수를 미리 선별하여 모형적합에 활용하는 것이 민간소비 예측 성능을 향상시킬 수 있음을 확인하였다. 또한 코로나19 이후 빅데이터의 반영이 민간소비 예측 성능을 더욱 크게 향상시킴에 따라 경제 불확실성이 높은 상황일수록 새로운 정보를 적시에 반영할 수 있는 고빈도 빅데이터의 활용가치가 높은 것으로 판단된다.

**핵심 주제어:** 커널분위수회귀(Kernel Quantile Regression), MIDAS, 빅데이터, 변수선별

**JEL Classification:** D12, C53, C55

## I. 서론

국민소득통계(이하 GDP)는 국가의 경제상황을 종합적으로 보여주는 경제지표로서 경제정책의 방향성을 설정하는 중요한 근거로 사용되고 있다. 국내 GDP 통계는 한국은행에서 분기 및 연간 통계를 속보, 잠정, 확정으로 구분하여 발표하고 있다. 대부분의 국가통계와 마찬가지로 GDP 통계는 조사기준 시점과 공표시점 간의 차이가 발생하는 공표지연 문제가 발생한다. 한국은행에서는 이를 줄이기 위한 많은 노력을 기울이고 있으나, 통계생산을 위한 기초자료의 최초 가용시점, 정확한 통계를 생산하기 위해 소요되는 물리적인 시간 등을 고려할 때, 공표시점을 앞당기는 것은 현실적으로 한계가 있다. 공표시차의 발생은 비단 우리나라만의 문제는 아니며, 미국 등을 포함한 선진국 등에서도 GDP 통계의 공표시차 제약을 보완하기 위해 많은 노력을 기울이고 있다. 이러한 노력의 일환으로 GDP 통계와 상관관계가 높으면서 보다 빠르게 수집이 가능한 자료들을 바탕으로 계량 모형을 구축하여 당분기 GDP를 실시간으로 추정하는 방법(이하 Nowcasting)이 오래전부터 활발하게 연구되어 왔다 (Bańbura et al., 2013; 김치호 and 김현정, 2016). 특히, 최근과 같이 경제 불확실성이 높은 상황에서는 시의성이 있는 정책의 실행과 평가가 무엇보다 중요하므로, 정확한 Nowcasting의 필요성은 어느 때보다 크다 하겠다.

GDP를 구성하는 다양한 항목 중 민간소비는 국민들이 체감하는 경제상황을 직접적으로 보여준다는 측면에서 정부의 정책 결정과 평가에 주요한 지표로 활용된다. 하지만 GDP 총액과는 달리 민간소비를 실시간으로 추정하는 방법에 대한 연구는 상대적으로 많이 이루어지지 않았다. GDP Nowcasting의 경우 GDP 통계보다 공표주기가 빠른 다른 거시경제지표 등을 활용하는 것이 가능하지만, GDP 구성항목의 하나인 민간소비는 활용 가능한 기초자료가 적어 추정이 더 어려운 측면이 있음을 예상할 수 있다. 그러나 최근 4차산업의 발전과 더불어 민간소비와 밀접한 관련이 있는 빅데이터에 대한 접근성이 크게 증대되었다. 대표적인 예로, 신용카드 소비액 등을 들 수 있는데, 이러한 빅데이터를 적극 활용하면 민간소비 Nowcasting의 예측력을 보다 향상시킬 수 있을 것으로 예상된다.

본 연구에서는 보다 정확한 민간소비 Nowcasting을 위해 빅데이터를 활용할 수 있는 기계학습 방법의 사용을 제안하고, 이를 실증적으로 비교 분석하였다. 본 연구의 의미는 크게 세 가지로 나눌 수 있다. 첫째, GDP 총액이 아닌 민간소비에

대한 Nowcasting을 다루었다는 점이다. 전술하였다시피 민간소비의 경우 실시간으로 활용가능한 카드 빅데이터, 신문기사 텍스트 데이터 등과 더욱 밀접한 연관을 지니고 있어, 이러한 빅데이터를 민간소비 예측에 적극 활용하고자 하였다. 둘째, 본 연구에서는 다양한 기계학습 방법을 활용한 Nowcasting 방법을 제안하고 그 성능을 실증적으로 비교 분석하였다. Nowcasting에서 널리 쓰이는 MIDAS-회귀 모형 (Ghysels et al., 2004)을 다양한 기계학습 알고리즘을 적용할 수 있는 형태로 확장하였다. 마지막으로, 본 연구에서는 다양한 종류의 빅데이터뿐만 아니라 기존에 널리 쓰이는 거시경제지표에 대한 자료도 Nowcasting에 모두 활용한 고차원 모형의 활용을 제안하였다. 차원의 저주 (curse of dimensionality)를 최소화하고 모형의 성능을 최적화하기 위해 변수선별 (feature screening)의 활용을 제안하고 그 효용을 확인해 보았다. 그 결과 코로나19 팬데믹으로 인해 경제 불확실성이 높은 최근의 상황에서는 변수선별이 Nowcasting의 성능에 큰 영향을 미치는 것을 확인하였다.

본 논고의 구성은 다음과 같다. 2절에서는 연구 배경과 Nowcasting에 관련된 선행연구들에 대해 살펴보았다. 3절에서는 본 연구에서 제안하는 민간소비 Nowcasting 방법에 대해 자세히 소개하였다. MIDAS-회귀모형을 바탕으로 다양한 기계학습 방법론을 제안하고 추가적인 변수선별에 대한 내용도 상세히 기술하였다. 4절에서는 제안된 방법을 실제자료에 적용하고 그 결과를 심층적으로 분석하였다. 마지막으로 본 연구의 결론과 시사점을 5절에서 제시하였다.

## II. 배경 및 관련 연구 리뷰

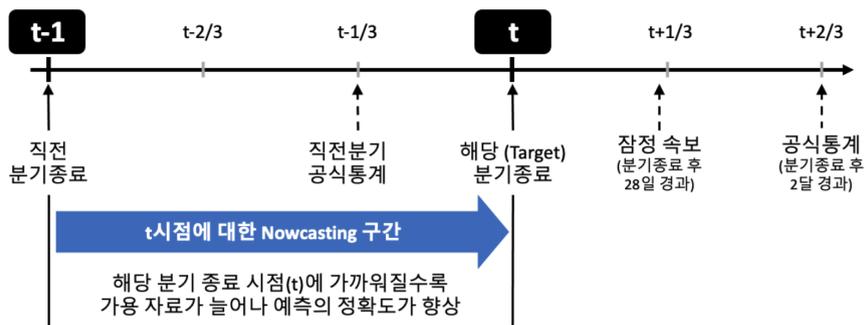
### 1. 연구 배경

Nowcasting은 장기전망을 의미하는 Forecasting과 대비되는 개념으로, GDP와 같이 공표시차가 발생하는 경우 다양한 연관 정보를 활용하여 공표시점 이전에 빠르게 추정치를 계산해 내는 것을 의미한다. 국내 GDP의 경우 해당 분기 종료 이후 약 28일 이내에 속보치를 발표하고 분기 잠정치 발표까지는 약 2달의 시차가 발생한다. (그림 1 참조).

GDP Nowcasting은 GDP보다 공표주기가 빠르거나 일찍 발표되는 다양한 관련 자료를 공변량으로 활용하여 GDP를 실시간으로 예측한다. 따라서 GDP에 활용가능한 공변량 계열은 통상적으로 GDP보다 빠른 주기를 가지는 시계열 자료

이다. 이러한 관점에서 GDP Nowcasting의 통계/계량적 방법론의 핵심은 자료별 관측주기가 다르기 때문에 발생하는 혼합주기 (mixed frequency)와 자료의 가용시점이 다르기 때문에 발생하는 문제 (ragged-edge problem)를 어떻게 해결하는가에 있다. 혼합주기 문제를 해결하는 가장 간단한 방법은 고주기 자료에서 나오는 다수의 값들을 하나의 값으로 결합하는 방법이다. 2절에서 소개한 교량방정식(Bridge Equation)이나 MIDAS-회귀모형 (Mixed Data Sampling Regression Model) 등에서 활용하는 방법으로 빠르고 효율적이지만 정보의 손실이 발생하는 방법이다. 다른 대안으로 혼합주기 문제를 결측으로 인식하여 풀어내는 방법이 있다. 이 경우, 자료의 분포함수를 직접 추정하는 방식으로 접근하면 자연스럽게 혼합주기 문제를 해결할 수 있다. 뿐만 아니라 자료의 서로 다른 가용시점(ragged-edge) 문제 역시 결측의 한 형태로 볼 수 있기 때문에 유사한 방법으로 처리가 가능하다는 장점이 있다. GDP Nowcasting에 가장 널리 쓰이는 동적인자모형(Dynamic Factor Model) 등이 이에 해당한다.

한편, Nowcasting에서 빅데이터의 활용이 점점 늘어나고 있는 것이 최근의 추세이다. 자료 수집주기가 매우 짧고 공표시차가 거의 없다고 할 수 있는 빅데이터의 특성상, 이들 자료를 잘 활용한다면 Nowcasting의 성능을 크게 향상시킬 수 있다. 이러한 빅데이터의 활용을 위해 알고리즘 효율성이 높은 기계학습 방법론이 많이 이용되고 있다. (Richardson et al., 2018; Chen et al., 2019; Babii et al.,



〈그림 1〉 GDP Nowcasting: 해당 분기 종료( $t$ )시점의 GDP Nowcasting은 다양한 시점에서 가능하다. 해당 분기 종료시점( $t$ )에 가까울수록 Nowcasting에 활용가능한 가용 자료의 수가 증가하므로,  $t$ 시점에서의 예측값이 가장 정확하다. 예를 들어, 해당 분기 종료시점 두달 전( $t - 2/3$ )에 직전분기( $t - 1$ )의 속도치를 활용할 수 있으며 한달전 ( $t - 1/3$ )에 공식통계를 활용할 수 있다.

2021; 이현창 et al., 2022). 빅데이터는 사용할 수 있는 정보의 양 (quantity)이 많고 수집 속도가 빠른 장점이 있는 반면, 관찰자료이기 때문에 정보의 질이 통상적인 조사자료에 비해 떨어진다. 이러한 경우, 변수선택은 필수적이다. 모형에 사용되는 변수의 개수가 늘어날수록 추정해야 하는 모수의 개수가 증가하기 때문에 해당 변수가 Nowcasting에 필요한 (양과 질 모든 측면에서) 충분한 수준의 정보를 보유하고 있지 않다면 오히려 예측성능이 저하되는 역효과가 발생하게 된다.

본 연구에서는 기존 거시경제지표 자료와 빅데이터를 모두 활용한 기계학습 기반 고차원 Nowcasting 방법론에 대해 분석하였다. 이를 위해 고차원 변수를 효율적으로 다룰 수 있는 다양한 알고리즘을 제안하고 그 성능을 실증적으로 비교 평가하였다.

## 2. 선행연구

GDP Nowcasting을 위한 방법은 크게 결합모형 (Joint Model)과 부분모형 (Partial Model) 두가지로 나눌 수 있다. 결합모형은 종속변수 시계열  $Y_t$ 와 공변량 시계열  $X_t$ 의 결합분포  $(X_t, Y_t) \sim \mathcal{P}_{X,Y}$ 를 추정하는 방법으로, 동적인자모형 (Dynamic Factor Model, Giannone et al., 2008)이 그 대표적인 예이다. Nowcasting을 위한 결합모형은 상태-공간 표현형식 (state-space Representation, Harvey, 1990)을 활용하여 각 시계열을 공통성분 (common component)과 고유성분 (idiosyncratic component)으로 직교분해한다. 이 때 공통성분은 모든 시계열이 공유하는 공통의 부분을 의미하며, 이를 바탕으로 종속변수 시계열의 값을 예측한다. 결합모형은 분포가정에 의존한 방법으로 통상적인 최대가능도추정이나 베이지안 방법으로 모형을 적합할 수 있다. 특히 Kalman-Filter나 EM 알고리즘 등이 널리 활용된다. 결합모형은 종속변수와 공변량의 결합분포를 추정하므로 혼합주기와 서로 다른 가용시점 문제 (ragged-edge problem)를 결측(missing)의 형태로 인식하여 자연스럽게 해결할 수 있다. 뿐만 아니라, 결합모형은 추정된 결합분포로부터 각 공변량별로 예측치를 계산해 낼 수 있어 이를 바탕으로 종속변수에 미치는 영향을 정량적으로 평가할 수 있다. 이러한 장점 때문에 결합모형의 대표적인 예인 동적인자모형이 GDP Nowcasting에서 가장 널리 쓰이고 있다 (미국, Giannone et al. (2008); 뉴질랜드 Matheson (2010); 노르웨이, Luciani and Ricci (2013); 프랑스, Barhoumi et al. (2010); 캐나다, Chernis and Sekkel (2017) 등).

이와는 달리 부분모형은 공변량이 주어졌을 때, 종속변수의 조건부 분포  $Y_t | X_t \sim \mathcal{P}_{Y_t|X_t}$ 를 추정하는 방법으로 교량방정식 (Bridge Equation)과 MIDAS-회귀 (MIDAS-Regression, Ghysels et al., 2004)가 그 대표적인 예이다. 종속변수  $Y_t$ 와  $X_t$ 보다  $m$ 배 빠른 고주기 공변량  $X_t$ 간의 교량방정식은 다음과 같다.

$$Y_t = \alpha + \beta \bar{X}_t + \varepsilon_t$$

여기서  $\bar{X}_t = m^{-1} \sum_{j=1}^m X_{t-(j-1)/m}$ 이다. 즉, 교량방정식은 고주기 공변량의 구간별 평균값을 활용하여 종속변수의 주기와 일치시킴으로써 혼합주기의 문제를 해결한다. MIDAS-회귀는 교량방정식을 확장한 형태로서, 평균값 계산에 활용되는 가중치를 자료로부터 학습시켜 구하는 방법이다<sup>1)</sup>. 부분모형은 ragged-edge 문제를 해결하기 위해 결측값 대체 (missing imputation)를 해주어야 한다. 만약 그렇지 않으면 정보의 손실이 발생하여 모형의 성능이 떨어질 수 있다는 단점이 있다. 또한 결합모형과 비교하여 종속변수와 공변량간의 연관성을 정량화하는데 어려움이 발생할 수 있다. 하지만 부분모형은 결합모형에 비해 단순한 가정을 활용하기 때문에 데이터의 크기가 증가하더라도 효율적인 계산이 가능하다.

최근 GDP Nowcasting에도 새로운 방법이 많이 활용되고 있는 추세이다. 특히 빅데이터의 활용이 주목 받고 있다. 신용카드나 POS (point of sales) 데이터, 혹은 스마트폰에서 수집된 GPS 데이터 등을 효과적으로 활용하면 GDP Nowcasting의 성능을 크게 향상시킬 수 있다 (서진용, 2017; Moriwaki, 2019). 빅데이터는 자료의 생성 직후 신속하게 수집이 가능하여 현재시점의 경제상황을 지체없이 확인할 수 있다. 이와 같은 이유로 다양한 국가에서 빅데이터에 기반한 Nowcasting 방법론을 개발하고자 노력하고 있다 (Barnett et al., 2016; Galbraith and Tkacz, 2018; Kim and Swanson, 2018; Raju and Balakrishnan, 2019; Aastveit et al., 2020, 등). Nowcasting에서 빅데이터를 활용하는 데 가장 큰 어려움은 계산의 효율성 확보이다. 동적인자모형을 비롯한 결합모형의 경우, 대용량 빅데이터를 활용하는 데 계산상의 어려움이 있으며, 이러한 단점을 보완하기 위한 연구가 이루어지고 있지만 그 성과가 제한적이다 (Chan and Jeliazkov, 2009; Delle Monache and Petrella, 2019). 이러한 관점에서 빅데이터에 기반한 Nowcasting 문제에서 부분모형의 활

1) 자세한 내용은 3.1절 참조

용가치는 매우 크다 하겠다 (Babii et al., 2021).

### Ⅲ. 방법론

#### 1. MIDAS-회귀모형

본 연구에서는 빅데이터 활용에 용이한 부분모형을 고려하였다. 특히 최근 Babii et al. (2021)이 제안한 고차원 MIDAS-회귀모형을 활용하였다. 종속변수 시계열  $y_t$ 와  $K$ 개의 공변량 시계열  $x_{t,k}, k = 1, 2, \dots, K$ 에 대한 혼합주기 회귀모형 (Mixed Frequency Regression)은 다음과 같이 정의된다.

$$\phi(L; J)y_t = \rho_0 + \sum_{k=1}^K \psi(L^{1/m_k}; \theta_k)x_{t,k} + \varepsilon_t, \quad t = 1, 2, \dots, T. \quad (1)$$

여기서  $\varepsilon_t$ 는 랜덤오차,  $\phi(L; J)y_t = (I - \rho_1 L - \rho_2 L^2 - \dots - \rho_J L^J)y_t$ 는 저주기 종속변수  $y_t$ 에 대한 시차-다항식 (Lag-polynomial)을 나타낸다.

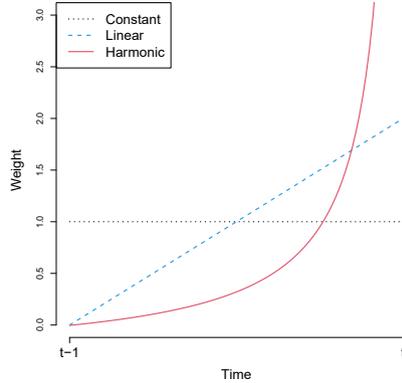
우변에 있는  $\psi(L^{1/m_k}; \theta_k)x_{t,k} = m_k^{-1} \sum_{j=1}^{m_k} \theta_{j,k} x_{t-(j-1)/m_k, k}$ 는 종속변수보다 주기가  $m_k$ 배 빠른 공변량  $x_{t,k}$ 에 대한 시차-다항식을 의미하며,  $\theta_k = (\theta_{1,k}, \dots, \theta_{m_k, k})^T, k = 1, \dots, K$ 는 회귀계수벡터를 나타낸다.

혼합주기 문제를 해결하기 위한 MIDAS (Ghysels et al., 2004) 아이디어를 (1)에 적용하면 회귀계수  $\theta_k$ 를 다음과 같이  $d$ 개의 모수  $\beta_k = (\beta_{1,k}, \dots, \beta_{d,k})^T$ 에 의해 조절되는 가중치함수  $w(\cdot)$ 를 통해 표현할 수 있다.

$$\theta_{j,k} = w\left(\frac{j-1}{m_k}; \beta_k\right) = \sum_{d=1}^D \beta_{k,d} v_d\left(\frac{j-1}{m_k}\right) = \beta_k^T \mathbf{v}_{j,k}, \quad j = 1, \dots, m_k. \quad (2)$$

여기서  $\{v_d(\cdot), d = 1, \dots, D\}$ 는 사용자가 특정할 수 있는 가중치함수의 기저함수 셋을 의미하며  $\mathbf{v}_{j,k} = \{v_d((j-1)/m_k), d = 1, \dots, D\}^T$ 이다. 즉,  $D$ 가 증가할수록 표현할 수 있는 가중치의 형태가 복잡해지므로 모형의 복잡도가 증가한다. 본 연구에서는 Bajari et al. (2015)가 제안한 3차 다항식 (Legendre Polynomial) 대신 아래와 같은 세 가지 형태의 서로 다른 기저함수를 고려하였다. 본 연구에서는 고려하는 공변량의 개수가 매우 많아 가중치함수의 복잡도를 높이는 것이 모형의 분산을 크게 만들기 때문에 오히려 Nowcasting의 성능을 저하시키는 것을 확인하

였다. 한가지 대안으로, 최적의 가중치함수를 찾기 위해 교차-타당성 검증 (Cross Validation)과 같은 방법을 고려할 수 있으며, 이는 후속연구를 통해 살펴볼 가치가 있는 것으로 판단한다.



〈그림 2〉 가중치 함수: 세가지 함수 모두  $(t-1, t]$  구간에서 관찰된 고주기 관찰치들의 가중치를 결정한다. 상수(constant) 함수는 단순평균값을 제공하며, 나머지 두 함수는  $t$ 시점에 가까울수록 높은 가중치를 주어 평균을 계산하여 주기를 맞춰준다.

이제, (2)를 (1)에 대입하면 다음과 같은 선형모형의 형식으로 (1)을 재표현할 수 있다.

$$y_t = \mathbf{b}^T \mathbf{z}_t + \varepsilon_t, \quad t = 1, 2, \dots, T \quad (3)$$

여기서  $\mathbf{b} = \{\rho_1, \dots, \rho_L, \beta_1^T, \dots, \beta_K^T\}$  이고  $\mathbf{z}_t = (1, y_{t-1}, \dots, y_{t-L}, m_1^{-1} \sum_{j=1}^{m_1} \mathbf{v}_{j,1}^T \cdot x_{j,1}, \dots, m_k^{-1} \sum_{j=1}^{m_k} \mathbf{v}_{j,k}^T \cdot x_{j,k})^T$  이다. Babii et al. (2021)에서는 (3)의 모수를 추정하기 위해 벌점화 최소제곱법 (Penalized Least Square Estimation)을 활용할 것을 제안하였다.

## 2. 모형 확장과 기계학습의 활용

한편, 기계학습에서는  $y_t$ 와  $\mathbf{z}_t$ 간의 관계가 반드시 단순한 선형함수일 필요가 없다. 따라서 다음과 같은 모형의 확장을 고려할 수 있으며 이는 Nowcasting 문제를

다양한 기계학습 알고리즘을 통해 해결할 수 있는 연결고리를 제공한다.

$$y_t = f(\mathbf{z}_t) + \varepsilon_t, \quad t = 1, 2, \dots, T \quad (4)$$

이제, (4) 하에서 다양한 기계학습 방법을 적용하여 비선형 함수를 포함한 다양한 회귀함수  $f$ 를 추정할 수 있다. 가장 간단한 방법으로는 Babii et al. (2021)이 처음 제안한 바와 같이  $f$ 에 대한 선형성을 가정한 뒤 벌점화를 통해 고차원 모수를 추정하는 방법이다. 이 때 사용할 수 있는 대표적인 벌점함수로는 LASSO (Tibshirani, 1996), Elastic Net (Zou and Hastie, 2005), 그리고 SCAD(Smoothly Clipped Absolute Deviance) (Fan and Li, 2001) 등이 있다. 특히 SCAD 벌점함수는 LASSO에서 발생하는 편향을 줄임으로써 근사적으로 더 좋은 추정량을 제공한다는 것이 알려져 있다. 다만, SCAD 벌점함수는 비볼록함수이므로 LASSO에 비해 계산이 좀더 어렵다고 알려져 있으나, 좌표별 강하법 (Coordinate Descent Algorithm) 등을 활용하면 대용량 자료에 대해서도 효율적 계산이 가능하다 (Breheny and Huang, 2011). 본 연구에서는  $f$ 를 선형함수로 가정한 뒤, LASSO와 SCAD 벌점함수를 적용한 모형을 활용하였다.

선형 모형을 비선형으로 확장하는 데 가장 중요한 것은 차원의 저주를 해결하는 것인데, 차원의 저주를 피할 수 있는 대표적인 방법으로 다음과 같은 단일지표 모형 (Single Index Model)을 고려할 수 있다.

$$y_t = g(\boldsymbol{\beta}^T \mathbf{z}_t) + \varepsilon_t, \quad t = 1, 2, \dots, T \quad (5)$$

즉, 종속변수와 다차원 공변량의 연관성에 대한 정보를 모두 포함하고 있는 단일 지표 모수  $\boldsymbol{\beta}$ 를 우선 추정한 뒤 1차원으로 축약된 공변량  $\boldsymbol{\beta}^T \mathbf{z}_t$ 와  $y_t$ 의 관계를 나타내는 회귀함수  $g$ 를 비모수적으로 추정하는 것이다. 본 연구에서는 단일지표모형 (5)에서  $\boldsymbol{\beta}$ 를 추정하기 위해 충분차원축약 (Sufficient Dimension Reduction, SDR; Li, 2018)을 활용하였으며, 차원축약 후 1차원 회귀함수를 적합하기 위해서는 스플라인 (spline) 방법을 활용하였다. 기존의 연구에서는 차원축소의 방법으로 주성분분석(Principal Component Analysis)을 많이 활용하였지만(Marcellino and Schumacher, 2010; Kim and Swanson, 2018), Nowcasting과 같은 지도학습 문제에서는 충분차원축약이 주성분분석보다 더 좋은 성능을 나타내는 것으로 알려져

있다.

또한 비선형 모형 (4)를 직접 추정하는 방법도 다양하게 존재한다. 예를 들어 나무 모형이나 나무모형을 기반으로 앙상블 방법을 활용할 수 있으며, 본 연구에서는 가장 대표적인 앙상블 방법 중 하나인 랜덤 포레스트를 고려하였다. 이 외에도 평활-스플라인 모형이나 인공신경모형도 대안으로 고려할 수 있다.

비선형 모형을 직접 추정하는 방법으로, 본 연구에서는 커널분위수회귀 (Kernel Quantile Regression, KQR)도 고려하였다. 커널을 활용하면 공변량이 많은 고차원 상황에서 회귀함수를 효율적으로 추정할 수 있다는 것이 잘 알려져 있다 (Zhang, 2002). 이러한 점에 착안하여 커널분위수회귀 (Li et al., 2007)를 이용하여 (4)를 추정하였다. 커널분위수회귀는 다음과 같은 최적화 문제를 통해 (4)의 회귀함수  $f$ 를 추정한다:

$$\hat{f} = \operatorname{argmin}_{f \in \mathcal{H}_K} \sum_{i=1}^n \rho_{\tau}(y_i - f(\mathbf{z}_i)) + \frac{\lambda}{2} \|f\|_{\mathcal{H}_K}^2 \quad (6)$$

여기서  $\rho_{\tau}(u) = u(\tau - \mathbb{1}\{u \geq 0\})$ 는  $100 \times \tau\%$ 번째 분위수 함수를 추정하기 위한 체크 손실함수이며,  $\mathcal{H}_K$ 는 주어진 커널함수  $K(\mathbf{z}, \mathbf{z}')$ 에 의해 생성되어진 재생커널힐버트 공간 (Reproducing Kernel Hilbert Space)을 나타낸다. 위 모형은 조율모수  $\lambda > 0$ 를 통해 모형복잡도와 적합도 사이에 균형점을 찾아 최적의 성능을 구현할 수 있다. 주어진 최적화 문제는 함수공간상에 정의된 무한차원의 문제로 보이지만, 대표자 정리 (Representer Theorem) 등에 의해 유한차원의 최적화 문제로 재표현할 수 있어 표본으로부터 계산이 매우 용이하다. 특히 커널을 이용할 경우 모수의 개수가 공변량의 개수가 아닌 표본의 크기 (종속변수의 크기)에 의존하게 되어 공변량의 수가 많은 고차원 모형에 특히 강점을 가진다. 본 연구에서는  $\tau = 0.5$ 로 설정하고 중위수 회귀를 고려하였다. 중위수 회귀 대신 통상적인 제곱 손실함수를 이용할 수 있으나, 본 연구에서는 추정의 강건성을 위해 분위수회귀를 고려하였다<sup>2)</sup>.

### 3. 변수선별

고빈도 데이터 등 다양한 종류의 빅데이터를 활용한 GDP Nowcasting은 고차원 문제이다. 즉, 주어진 데이터의 크기에 비해 공변량의 개수가 많다. 이 경우

2) 본 연구에서 활용된 기계학습 방법들에 관한 보다 자세한 설명은 부록 참조

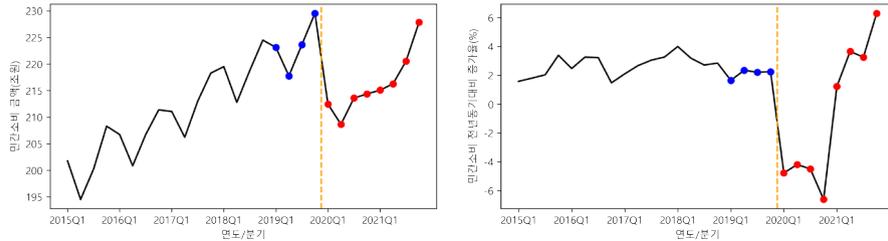
차원의 저주가 발생하는데, 이를 해결하기 위해서 기계학습에서는 희박성 가정 (sparsity assumption)을 활용한다. 즉, 많은 공변량 중 소수의 공변량만이 종속변수에 영향을 준다는 것이다. 희박성 가정하에서 LASSO나 SCAD 등의 벌점함수를 이용하면 변수선택을 통해 추정량의 성능이 향상된다는 사실이 잘 알려져 있다 (Hastie et al., 2015). 그런데 공변량의 개수가 월등히 많은 초고차원 (ultrahigh-dimensional) 상황이 되면 추정해야 할 모수가 주어진 정보에 비해 지나치게 많아져 벌점화를 활용하여 변수선택을 하더라도 모형의 성능이 현저히 떨어진다는 것도 알려져 있다 (Fan and Lv, 2008). 이를 해결하기 위한 방법으로 Fan et al. (2009)는 종속변수와 개별 공변량 간의 주변 연관성 (marginal association)을 정량화하고 이를 기반으로 한 변수선별 (feature screening)의 활용을 제안하였다.

변수선별의 핵심은 공변량 간의 상호작용은 배제하고, 종속변수와의 주변 연관성만을 바탕으로 빠르게 종속변수에 영향을 주는 신호변수를 선별하는 것이다. 변수선별 후 남은 변수들 중에는 종속변수에 영향을 주는 변수도 있고 그렇지 않은 변수도 있다. Fan and Lv (2008)는 주변 연관성을 이용한 변수선별이 신호변수를 제거할 확률이 0으로 수렴함을 보였으며, 이를 확신허선별성질 (sure screening property)이라고 명하였다. 확신허선별성질은 주변 연관성을 기반으로 한 변수선별의 이론적 정당성을 제공하여 이후 다양한 형태의 변수선별 방법이 제안될 수 있는 기반을 마련하였다 (Fan et al., 2012; Zhu et al., 2011; Li et al., 2012; He et al., 2013; Chang et al., 2013, 등).

Nowcasting을 위한 변수선별을 위해 종속변수  $y_t$ 와 주어진  $k$ 번째 공변량  $\mathbf{x}_k$  간에 다음과 같은 주변 모형 (marginal model)을 고려할 수 있다.

$$y_t = \mu_k + m_k^{-1} \sum_{j=1}^{m_k} x_{j,k} \mathbf{v}_{j,k}^T \boldsymbol{\beta}_k + u_t, \quad t = 1, \dots, T \quad (7)$$

여기서,  $\mu_k$ 는 주변모형의 상수항을 나타낸다. 위의 주변모형 (7)에 근거하여 가장 최근의 종속변수 값을 가장 잘 추정하는 변수를 선별할 수 있다. 즉, 각 공변량별로 (7)에 근거한 검증오차를 계산한 뒤 검증오차가 가장 작은 값 순으로 변수를 선별하는 것이다. 이때 (7)의 모형적합을 위해 다양한 방법을 활용할 수 있다. 본 연구에서는 변수선별의 안정성을 높이기 위하여 가중평균회귀, 가중중위수회귀, 그리고 간편화된 형태의 직교회귀를 이용하여 세 가지 검증오차를 계산한 뒤, 그 평균에 근거하여 변수를 선별하였다. 가중치는 최근 관찰치에 더 많은 중요도를



〈그림 3〉 2015년도 1분기부터 2021년도 3분기까지의 민간소비 금액과 전년동기대비 증가율

부여하기 위해 검증 시점과의 차이에 반비례하도록 설정하였다. 그리고 간편화된 직교회귀라 함은 데이터의 기울기 ( $s.d(y_t)/s.d(x_{k,t})$ )를 바탕으로 회귀계수를 추정하는 방법으로 측정오차 (measurement error)가 있는 경우에 활용하는 직교회귀 (orthogonal regression)의 아이디어를 응용한 것이다.

## IV. 실증 분석

### 1. 데이터

예측의 목표가 되는 종속변수는 GDP 구성항목 중 민간소비로 설정하고, 실질 민간소비 전년동기대비 증가율을  $y_t$ 로 고려하였다. 분석을 위해서 2000년 1분기부터 21년 3분기까지 약 20년간의 분기 데이터를 활용하였다. 여기서 민간소비는 연구 시점에서 가장 완전한 정보를 바탕으로 추계된 수치를 Ground-truth로 가정하고자 2000년부터 2020년까지는 확정치를 그리고 2021년은 잠정치를 이용하였다. 그림 3은 2015년 1분기부터 21년 3분기 실질 민간소비 금액과 전년동기대비 증가율을 나타낸다. 본 연구에서는 마지막 11개 분기의 전년동기대비 증가율을 Nowcasting하여 실제값과 비교해봄으로써 그 성능을 평가하였다. 11개 분기의 예측을 위해서는 해당 분기 종료시점까지 가용할 수 있는 모든 데이터를 활용하였다. 특히 2020년 초에 발생한 코로나19 팬데믹을 기점으로 이전 4개 분기와 이후 7개 분기의 Nowcasting 성능을 비교해 봄으로써 경제 불확실성이 증가한 상황에서 빅데이터의 활용이 Nowcasting에 어떠한 영향을 미치는지 확인해 보았다.

민간소비에 영향을 미치는 공변량은 매우 광범위하다. 통상적인 GDP Nowcasting에 사용되어 온 거시경제지표 및 통계청에서 주기적으로 발표하는 경제

관련 조사자료부터 최근 민간 소비지표 추정에 흔히 활용되는 신용카드 빅데이터까지 여러가지 공변량들을 활용할 수 있다. 본 연구에서는 한국은행과 통계청 데이터베이스에서 제공하는 다양한 공변량을 최대한 고려하여 총 281개의 후보 변수를 수집하였다 (전체 공변량 변수 목록은 부록 참조). 이 중 각 변수들이 제공하는 정보들간의 중복성, 공표주기 및 공표지연 등으로 인한 가용성 등을 고려하여 최종적으로 208개의 변수를 이용한 모형을 고려하였다. 총 208개의 변수를 그 특성에 따라 11개의 범주로 나누었고, 크게는 통상적인 거시경제데이터와 빅데이터로 구분하여 정리하였다 (표 1 참조). 빅데이터는 자료의 수집빈도가 높고 공표지연이 없어 Nowcasting에 유용한 정보를 제공할 것으로 기대하였다.

중속변수를 기준으로 모형적합에 활용할 수 있는 표본의 크기는 최대 90개를 넘지 못한다. 더욱이 공변량의 경우 이용가능 시점이 더 최근인 경우도 많기 때문에 경우에 따라 30개 미만의 데이터를 바탕으로 200여개의 공변량을 포함하는 고차원 모형을 적합시켜야 하므로, 이는 전형적인 고차원 혹은 초고차원 문제가 된다. 특히 MIADS 모형의 가중치 함수의 복잡도를 높이기 위해 여러개의 기저함수를 활용하게 되면 추정해야 하는 모수가 크게 늘어날 뿐만 아니라, 추가적인 표본이 수집되는 속도에 비해 활용가능한 변수의 개수가 더욱 빠르게 증가할 것으로 예상된다. 따라서 의미있는 공변량을 신속 정확하게 선별하는 것이 매우 중요할 것으로 예상할 수 있다.

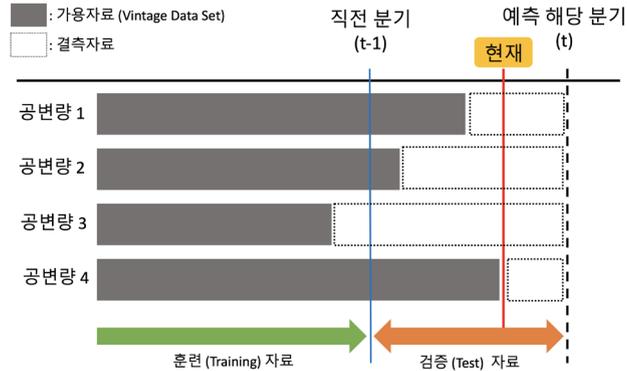
대분류	중분류	주기	설명
거시경제 데이터	금융 (10)	월별	주가, 수신금리, 대출금리 등
	물가 (7)	월별	생산자물가지수, 유가, 기대인플레이션, 소비자물가지수 등
	소비 (41)	월별	소비자 동향조사, 소매업체별 판매액지수, 경제심리지수 등
	생산(5)	월별	산업생산지수
	부동산(12)	월별	월세가격, 민간부문 미분양 건수
	고용 (6)	월별	규모별 고용현황
	경기종합 (3)	월별	경기종합지수 (선행, 동행, 후행)
	경제활동 (24)	월별	연령대 별 경제활동인구, 비경제활동인구, 실업자, 취업자
빅데이터	카드 (91)	월별	소비 유형별 신용카드 소비액 / 업종별 카드 사용금액 및 건수
	텍스트 (4)	일별	뉴스심리지수, 경제불확실성지수, 경제뉴스 텍스트에서 추출한 물가 지표
	고빈도 (5)	일별	영화매출액, 수출입, 고속도로통행량, 전력사용량 등

〈표 1〉 민간소비 Nowcasting에 활용된 공변량 요약

## 2. 빈티지 데이터셋 구축과 결측값 예측

Nowcasting의 핵심은 예측 대상분기와 관련된 정보를 최대한 많이 활용하는 것이다. 하지만 공변량으로 사용되는 많은 경제지표들도 공표지연의 문제가 발

생하고 지연 기간도 다르기 때문에 현재 시점에서 실제로 활용가능한 데이터, 즉 빈티지 데이터셋 (vintage data set)을 구성하는 것이 필요하다 (그림 4 참조).



〈그림 4〉 빈티지 데이터셋의 구성

MIDAS를 활용하는 경우, 결측값을 대체하지 않더라도 모형을 적합시키는 것이 불가능한 것은 아니지만, 본 연구에서는 Nowcasting의 성능을 최적화하기 위해 추가적인 결측치 대체를 실시하였다. 이는 코로나19 팬데믹 상황과 같이 경제의 불확실성이 큰 경우에는 공표지연 등으로 결측이 발생한 단기간 동안 경제상황이 급변할 수 있고, 결측치를 무시하는 경우 예측의 정확도가 현저히 떨어질 수 있기 때문이다. 본 연구에서는 가중회귀대체법 (weighted regression imputation)을 활용하였다. 즉 결측자료의 시점에 가까운 값들일수록 상대적으로 많은 가중치를 주고 회귀모형을 적합시켜 그 예측치로 대체하는 방법이다. 이때, 대체의 정확도와 안정성을 높이기 위해 1차부터 3차 다항회귀까지 세 가지 모형을 적합시켜 각 모형의 예측치 평균으로 대체하는 방법을 활용하였다. 기존 연구 등에서는 자기회귀모형을 결측치 대체 모형으로 많이 사용하였다. 하지만, 경제의 불확실성이 높은 경우에 자기회귀모형은 직전 관찰치의 영향을 지나치게 많이 받아 그 활용이 제한적일 것으로 판단하였다.

### 3. 변수선별

빈티지 데이터셋을 구축하고 결측치를 대체한 뒤, 3절에서 제안된 방법을 바탕으로 매분기 30개의 변수를 선별하여 모형을 적합하였다. 선별하는 변수의 개수는 통상적으로 표본의 크기를 넘지 않게 설정하는 것이 좋다고 알려져 있다 (Fan and

Lv, 2008). 표 2는 각 예측 시기별로 선별된 변수들의 구성을 나타낸 표이다. 신용카드 빅데이터와 소비자 동향조사나 소매업태별 판매액지수 등을 포함하는 소비 관련 지표들이 민간소비 Nowcasting에 유용한 정보를 제공한다는 사실을 확인하였다. 한 가지 주목할 점은 두 종류의 변수들이 차지하는 상대적 비중이 코로나19 경과와 함께 경제 불확실성 정도에 따라 변화하였다는 점이다. 경제 불확실성이 높은 시기(20년 1분기 및 21년 1분기 전후)에는 신용카드 빅데이터가 Nowcasting에 더 유용한 것으로 나타난 반면, 동 시기를 제외한 다른 시기에서는 전통적인 소비 관련 지표들이 더 중요한 역할을 하는 것을 확인하였다. 전통적인 거시경제지표의 경우 공표주기가 빠르지 않고 공표지연 역시 불가피하므로 경제상황이 급변하는 상황에서 민간소비 상황을 반영하는 데 어려움이 있기 때문으로 판단된다.

	빅데이터			거시경제 데이터							
	카드	텍스트	고빈도	금융	물가	소비	생산	부동산	고용	경기 종합	경제 활동
2019/1	8	1	0	0	1	12	0	0	1	2	5
2019/2	9	1	1	0	0	13	0	1	3	0	2
2019/3	8	0	1	1	1	12	1	1	3	1	1
2019/4	13	0	1	1	1	10	1	0	1	0	2
2020/1	10	2	0	3	0	7	4	0	0	1	3
2020/2	4	0	1	2	0	15	0	3	2	0	3
2020/3	5	0	1	3	1	11	0	1	1	1	6
2020/4	19	0	1	0	0	3	2	0	2	0	3
2021/1	13	0	1	0	1	7	0	0	2	0	6
2021/2	5	0	0	4	2	14	0	0	1	1	3
2021/3	7	1	2	1	1	11	0	0	0	2	5

〈표 2〉 Nowcasting 시점별 선별된 변수 요약

#### 4. Nowcasting 성능 비교

이제 3절에서 소개한 다양한 방법을 적용하여 민간소비 Nowcasting을 실시하고 그 성능을 평가해 보았다. 우선 가장 기본적인 모형으로 종속변수만을 활용한 1차 자기상관회귀모형(AR(1))을 고려하였다. 기계학습 방법으로는 LASSO와 SCAD 별점화를 이용한 선형회귀모형(LASSO, SCAD), 커널분위수회귀(KQR), 충분차원축약에 기반한 단일지표모형(SDR), 랜덤 포레스트(RF) 그리고 이 모든 추정량을 결합<sup>3)</sup>한 앙상블 방법(결합)을 고려하였다. 실제 연구과정에서는 Elastic Net

3) 각 방법의 추정치 단순 평균

(Zou and Hastie, 2005) 등과 같은 다른 벌점함수도 적용해 보았으나, 예측성능의 측면에서 LASSO와 SCAD가 가장 안정적인 성능을 보이는 것을 확인하였다. 단일 지표모형 추정을 위한 충분차원축약은 최소제곱추정량을 활용하였으며, 차원축약 이후 함수추정에는 자연 3차 스플라인 회귀 (cubic natural spline regression)를 이용하였다. 커널분위수회귀에서는  $\tau = 0.5$ 로 설정하고 선형커널을 이용하였다. 가우시안 커널 (Gaussian kernel)도 적용하였으나 선형커널에 비해 예측성능이 좋지 않음을 확인하였다 (5절 참조). 이 외에도 LASSO의 비선형 확장인 COSSO (Lin and Zhang, 2006), 인공신경망모형,  $L_2$ -부스팅 등도 적용해 보았으나 해당 방법들은 모형의 복잡도가 지나치게 높아서 과적합이 일어나거나 초기치에 지나치게 민감하게 반응하여 예측성능이 크게 떨어지는 것을 확인하였다. 마지막으로, 각 기계학습모형의 적합에 필요한 조율모수는 교차검증 (cross-validation)을 통해 추정하였다<sup>4)</sup>.

본 실증분석에서는  $t$ 시점 (혹은 분기)의 민간소비를  $t$  시점에 Nowcasting하는 상황을 설정<sup>5)</sup>하였다. 2절에서 설명한 바와 같이 Nowcasting 시점은  $t$  시점 이전에도 가능하지만  $t$  시점에 가까울수록 공변량의 가용정보가 많기 때문에 보다 정확한 예측이 가능해진다. 모형의 성능을 평가하기 위해서 평균절대오차 (Mean Absolute Deviance)을 고려하였다. 예측하고자 하는 분기의 관찰값을  $y_t$ ,  $t$ 시점에서 사용가능한 공변량 정보를 최대한 이용하여 Nowcasting한 값을  $\hat{y}_t$ 라 하면, 평균절대오차는 다음과 같이 절대오차의 평균으로 계산된다.

$$MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|$$

본 연구에서 검증의 대상이 되는 분기는 총 11개 이므로  $T = 11$ 로 하여 전체 MAE를 계산하였다. 나아가 코로나19 발생 이전의 4개 분기와 이후의 7개 분기를 구분하여 각각의 MAE를 계산하여 코로나19 전후의 Nowcasting 성능도 비교해 보았다. 표 3은 본 연구에서 제안한 다양한 방법의 예측 성능을 기본모형인 1차 자기상관회귀모형과의 MAE값에 대한 비로 나타낸 것이다. 즉, 해당 값이 0에 가까울수록 좋은 성능의 의미하며, 1이 넘는 경우 기본모형에 비해 성능이 떨어짐을

4) 시계열자료의 특성을 반영하여 자료를 분할하도록 수정

5) 'End of Quarter' 시나리오

의미한다. 기본모형인 1차 자기상관회귀모형(AR(1))의 경우, 코로나19 이전의 안정적인 경제상황에서는 어느 정도의 예측력을 보여주고 있지만, 코로나19 발생 이후에는 예측력이 현저하게 떨어진다. 이는 이전 분기의 관찰치에 의존하는 자기상관회귀모형의 특성상 자연스러운 결과로 해석된다.

본 연구에서 제안한 MIADS-회귀모형에는 1차 자기상관 항을 넣은 경우( $L=1$ )와 그렇지 않은 경우( $L=0$ )를 모두 고려하였다. 변수선별을 하지 않은 경우, 동 모형은 기본모형인 1차 자기상관모형에 비해서는 향상된 성능을 보이지만 코로나19 이후 분기의 민간소비 증가율 Nowcasting에는 어려움을 보인다. 동 모형은 SCAD와 같은 변수선택을 반영하는 경우에도 예측성능이 좋지 않은 것으로 나타난다. 통상적인 회귀모형의 경우, 종속변수와 공변량 간의 전체적인 상관관계를 통해 예측치를 추정하게 되는데, Nowcasting의 경우 데이터 전체구간의 상관성 보다는 예측 대상과 가장 가까운 가장자리 값들의 연관성이 상대적으로 훨씬 중요하다. 이러한 점을 고려하지 않고 다수의 변수를 포함하면, 종속변수의 전반적인 트렌드는 잘 예측하지만, Nowcasting의 대상이 되는 바깥쪽 끝부분의 값은 제대로 예측하지 못하는 문제가 발생한다. 표 3에서 변수선별을 한 경우의 전체 기간 예측성능을 살펴보면, 선형모형인 SCAD나 선형커널분위수회귀(KQR)가 비선형 모형인 랜덤포레스트(RF)나 충분차원축소 기반 방법(SDR)보다 좋은 성능을 보이는데, 이 역시 주어진 데이터의 바깥쪽 부분을 예측하는 Nowcasting의 특성상 모형의 복잡도가 지나치게 높은 경우 과적합 등의 문제로 인하여 예측성능이 저하되기 때문이다. 특히 코로나19 이후의 상황처럼 데이터의 변동성이 커지는 경우, 비선형 모형을 이용한 Nowcasting의 예측오차가 높아지는 것을 확인할 수 있다. 한편, 변수선별을 실시한 뒤 선별된 변수만으로 모형적합을 한 경우 변수선별을 실시하지 않은 경우에 비해 Nowcasting의 성능이 크게 향상됨을 확인하였다. 특히 모형의 복잡도가 높지 않은 SCAD나 커널분위수회귀가 안정적인 예측성능을 보여주었으며, 커널분위수회귀가 가장 좋은 성능을 보여주었다. 커널분위수회귀의 경우, SCAD와는 달리 추가적인 변수선택을 하지 않지만, 절대 손실함수를 사용하기 때문에 데이터의 바깥쪽 끝부분을 예측하는 Nowcasting 문제에서 좀 더 안정적인 성능을 보여주는 것으로 판단된다. 실제 분위수회귀는 중위수 뿐만 아니라 다양한 수준의 분위수를 예측치로 활용할 수 있다는 장점이 있어, 다양한 계량경제 모형에 널리 활용되고 있다(Koenker and Hallock, 2001). 충분차원축소 기반의 단일지표 모형은 변수선별을 하지 않았을 때는 성능이 좋지 못했지만, 변수선별 후에는 그

● 변수선별을 하지 않은 경우

	LASSO	SCAD	KQR	SDR	RF	결합
$L=0$ 전체	0.844	0.964	0.622	0.908	1.077	0.713
$L=0$ 코로나이전	0.235	0.321	0.120	0.211	0.187	0.199
$L=0$ 코로나이후	1.196	1.336	0.912	1.311	1.591	1.009
$L=1$ 전체	0.779	1.014	0.574	0.797	1.065	0.714
$L=1$ 코로나이전	0.222	0.207	0.107	0.114	0.165	0.134
$L=1$ 코로나이후	1.101	1.481	0.844	1.192	1.585	1.050

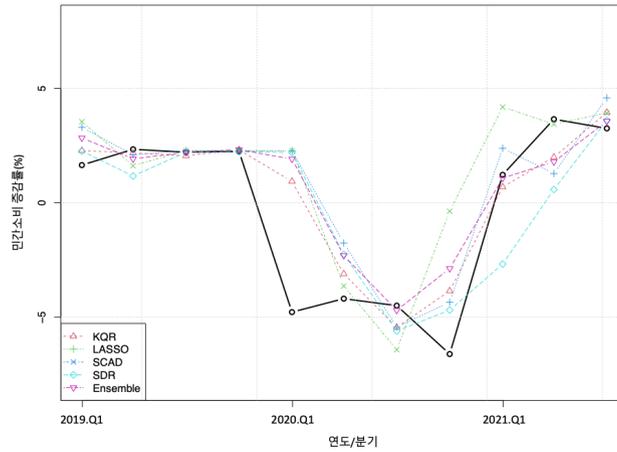
● 변수선별을 한 경우

	LASSO	SCAD	KQR	SDR	RF	결합
$L=0$ 전체	0.431	0.395	0.396	0.426	0.693	0.413
$L=0$ 코로나이전	0.131	0.152	0.129	0.185	0.042	0.109
$L=0$ 코로나이후	0.605	0.535	0.549	0.565	1.068	0.588
$L=1$ 전체	0.523	0.656	0.431	0.402	0.671	0.447
$L=1$ 코로나이전	0.102	0.102	0.056	0.016	0.116	0.056
$L=1$ 코로나이후	0.766	0.976	0.648	0.624	0.992	0.673

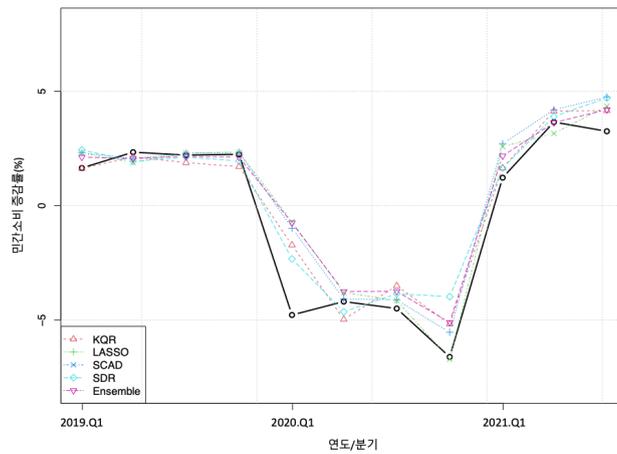
〈표 3〉 제안된 모형들간의 MAE 비교 (AR(1)모형 대비 상대 효율): 변수선별을 할 경우 적용된 학습방법에 상관없이 Nowcasting의 성능이 크게 향상됨을 알 수 있다.

성능이 크게 개선된 것을 확인할 수 있다. 이는 변수선별을 통해 모형의 복잡도를 크게 줄여 비모수적 추정방법의 안정성을 향상시킨 것이 예측성능을 개선하였기 때문으로 보인다. 실증분석에 활용한 모든 모형을 결합하여 예측하는 것도 매우 우수한 성능을 보이고 있는 사실에도 주목할 필요가 있다. 결합 모형의 경우, 양상불 효과로 인해 개별모형이 가진 편향의 효과가 감소하고 예측값의 분산이 줄어들기 때문이다. 예측의 정확성의 관점에서 본다면 nowcasting에서 결합모형의 활용 가능성을 좀 더 면밀히 검토해볼 가치 있다고 판단된다.

마지막으로, 그림 5는 변수선별에 따른 예측 결과를 AR(1) 항이 없는 ( $L=0$ ) 모형 하에서 비교한 것이다. (AR(1)항이 있는 모형에 대한 결과는 부록 참조). 코로나19 이전의 4분기 동안에는 모든 방법이 안정적인 성능을 보이고 있지만, 코로나19 이후의 시기에는 변수선별에 기반한 방법이 월등히 우수한 성능을 보이고 있다.



(a) 변수선별을 하지 않은 경우



(b) 변수선별을 한 경우

〈그림 5〉 AR(1)항이 없는 ( $L=0$ ) 모형간의 민간소비 예측치 비교: 변수선별의 효과를 확인할 수 있다.

## V. 결론 및 제언

본 연구에서는 빅데이터 기반 민간소비의 Nowcasting 문제를 다루었다. Nowcasting은 2000년대 이후 매우 활발히 연구되어 온 분야지만, 최근 빅데이터의 출현과 기계학습 방법론의 비약적인 발전으로 더욱 다양하고 새로운 시도들이 이루어지고 있다. 민간소비의 경우, 국민들의 경제활동 결과를 직접적으로 반영하므로 다양한 종류의 빅데이터를 활용하여 Nowcasting 성능을 높일 필요가 있다. 본 연구에서는 대용량 빅데이터 활용이 용이한 부분모형에 다양한 기계학습 방법론을 적용하여 Nowcasting을 실시하고 그 활용 가능성을 확인하였다.

본 연구 결과가 시사하는 내용을 요약하면 다음과 같이 세 가지로 정리할 수 있다. 첫째, Nowcasting을 위한 변수선별의 중요성이다. 데이터 관련 기술의 발전으로 Nowcasting에 활용할 수 있는 변수의 개수는 점차 증가하는 추세이다. 특히 빅데이터의 활용이 더 대중화되면 분석에 포함될 가능성이 있는 후보변수들이 크게 증가할 것이다. 하지만 변수, 혹은 그에 해당하는 정보의 증가가 반드시 모형의 성능을 향상시키는 방향으로 작용하는 것은 아니다. 특히 GDP Nowcasting과 같이 모형적합에 활용할 수 있는 표본의 수가 충분하지 않은 경우에는 예측력 향상에 도움이 되는 변수를 상황에 맞게 적절히 선별하는 것이 매우 중요하다. 본 연구에서는 초고차원 기계학습에서 널리 활용되는 방법인 변수선별방법의 활용을 제안하고 그 효과를 확인하였다. 하지만 변수선별이 반드시 데이터 기반 알고리즘에 의해서만 가능한 것은 아니다. GDP Nowcasting과 같이 오랜 시간 연구되어 온 주제의 경우, 해당 분야의 전문가 의견과 경제이론 등을 바탕으로 1차적으로 변수를 선별하는 것이 가능하다. 나아가 이러한 경험적 변수선별과 데이터 기반 변수선별을 적절한 방법으로 결합하면 Nowcasting의 성능을 추가적으로 개선할 수 있을 것으로 예상된다.

둘째, 본 연구를 통해 민간소비 Nowcasting을 위한 빅데이터의 활용 가능성을 확인하였다. 통상적인 거시경제지표 혹은 조사자료 등은 대부분 국가에서 관리하는 공식통계이므로 공표주기가 월별인 경우가 많으며 공표지연이 발생한다. 따라서 국민들이 체감하는 경제상황을 적시에 파악하기에는 한계가 있다. 따라서 민간 기업으로부터 수집이 가능한 신용카드 사용액 등의 자료와 뉴스 기사나 포털 검색 정보 등을 바탕으로 추출한 텍스트 자료 등 빅데이터를 활용하면 민간소비 Nowcasting의 성능을 향상시킬 수 있을 것으로 예상된다(이금희 and 황상필,

2016). 본 연구에서는 신용카드 데이터, 경제 뉴스기사 기반의 텍스트 자료 등을 포함한 빅데이터를 민간소비 Nowcasting에 활용해 보았다. 그 결과, 코로나19 팬데믹 사태와 같이 경제 불확실성이 높은 상황에서는 빅데이터 활용가치가 상대적으로 높아진다는 것을 확인하였다.

마지막으로, 빅데이터를 활용한 Nowcasting에서 교량방정식이나 MIDAS 모형과 같은 부분모형의 활용가능성을 제고하였다. 현재 GDP Nowcasting에서 가장 널리 쓰이는 방법은 동적인자모형 같은 결합모형에 근거한 방법이다. 결합모형을 활용하면 Nowcasting에서 발생하는 혼합주기나 ragged-edge 문제를 손쉽게 해결할 수 있을 뿐만 아니라, 모형의 해석이나 공변량 예측과 같은 보조적인 정보의 획득이 용이하기 때문이다. 하지만 대용량 빅데이터를 활용하기 위해서는 알고리즘의 계산효율성 (computational efficiency)과 확장성 (scalability)이 매우 중요하다. 이러한 측면에서 본다면, 빅데이터를 활용함에 있어 부분모형을 활용하는 것이 동적인자모형과 같은 결합모형에 비해 더 나은 선택이 될 수 있을 것이다.

한편, 본 연구의 한계점으로 다음과 같은 부분을 지적할 수 있다. 본 연구에서는 Nowcasting을 위해 기계학습 방법론을 적용하였는데, 이는 시계열 방법론에 근거한 전통적인 계량경제모형과는 차이가 있다. 특히 시계열자료가 가지는 특성을 제대로 모형에 반영하지 않았다는 비판을 받을 수 있으며, 이는 타당한 지적이라 하겠다. 그럼에도 불구하고, 본 연구에서는 빅데이터를 보다 효율적으로 활용하기 위해 통상적인 시계열 방법론에서 활용하는 정상성 등의 모형가정을 고려하지 않았다. 그렇기 때문에 본 연구를 통해 제안된 방법에 시계열자료의 특성을 세밀하게 고려하여 반영할 수 있도록 개선한다면 Nowcasting 성능을 더욱 향상시킬 수 있을 것으로 기대한다. 또한, 본 연구는 실증적 분석에 주안점을 두었으며, 제안된 방법론들과 관련된 이론적 내용에 대해서는 자세히 다루지 못하였다. 이후, 추가적인 연구를 통해 실증적 결과를 뒷받침 할 수 있는 이론적 결과를 유도할 수 있을 것으로 기대한다.

## References

- Aastveit, K. A., Fastbø, T. M., Granziera, E., Paulsen, K. S. and Torstensen, K. N. (2020). Nowcasting norwegian household consumption with debit card transaction data.
- Babii, A., Ghysels, E. and Striaukas, J. (2021). Machine learning time series regressions with an application to nowcasting, *Journal of Business & Economic Statistics* pp. 1–23.
- Bajari, P., Nekipelov, D., Ryan, S. P. and Yang, M. (2015). Machine learning methods for demand estimation, *American Economic Review* 105(5): 481–85.
- Bañbura, M., Giannone, D., Modugno, M. and Reichlin, L. (2013). Now-casting and the real-time data flow, *Handbook of economic forecasting*, Vol. 2, Elsevier, pp. 195–237.
- Barhoumi, K., Darné, O. and Ferrara, L. (2010). Are disaggregate data useful for factor analysis in forecasting french gdp?, *Journal of Forecasting* 29 (1-2): 132–144.
- Barnett, W., Chauvet, M., Leiva-Leon, D. and Su, L. (2016). Nowcasting nominal gdp with the credit-card augmented divisia monetary aggregates.
- Breheny, P. and Huang, J. (2011). Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection, *The Annals of Applied Statistics* 5(1): 232.
- Chan, J. C. and Jeliazkov, I. (2009). Efficient simulation and integrated likelihood estimation in state space models, *International Journal of Mathematical Modelling and Numerical Optimisation* 1(1-2): 101–120.
- Chang, J., Tang, C. Y. and Wu, Y. (2013). Marginal empirical likelihood and sure independence feature screening, *Annals of Statistics* 41(4).

- Chen, J. C., Dunn, A., Hood, K., Driessen, A. and Batch, A. (2019). Off to the races: A comparison of machine learning and alternative data for predicting economic indicators, *Big Data for 21st Century Economic Statistics*, University of Chicago Press.
- Chernis, T. and Sekkel, R. (2017). A dynamic factor model for nowcasting canadian gdp growth, *Empirical Economics* 53(1): 217–234.
- Delle Monache, D. and Petrella, I. (2019). Efficient matrix approach for classical inference in state space models, *Economics Letters* 181: 22–27.
- Fan, J., Feng, Y. and Song, R. (2012). Nonparametric independence screening in sparse ultra-high-dimensional additive models, *journal of the American Statistical Association* .
- Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties, *journal of the American statistical Association* 96 (456): 1348–1360.
- Fan, J. and Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space, *journal of the Royal Statistical Society: Series B* 70(5): 849–911.
- Fan, J., Samworth, R. and Wu, Y. (2009). Ultrahigh dimensional feature selection: beyond the linear model, *The journal of Machine Learning Research* 10: 2013–2038.
- Galbraith, J. W. and Tkacz, G. (2018). Nowcasting with payments system data, *International journal of Forecasting* 34(2): 366–376.
- Ghysels, E., Santa-Clara, P. and Valkanov, R. (2004). The MIDAS touch: Mixed data sampling regression models.
- Giannone, D., Reichlin, L. and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data, *journal of Monetary Economics* 55(4): 665–676.

- Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*, Cambridge university press.
- Hastie, T., Tibshirani, R. and Wainwright, M. (2015). *Statistical learning with sparsity*, Monographs on statistics and applied probability 143: 143.
- He, X., Wang, L. and Hong, H. G. (2013). Quantile-adaptive model-free variable screening for high-dimensional heterogeneous data, *The Annals of Statistics* 41(1): 342–369.
- Kim, H. H. and Swanson, N. R. (2018). Methods for backcasting, nowcasting and forecasting using factor-midas: With an application to korean gdp, *Journal of Forecasting* 37(3): 281–302.
- Koenker, R. and Hallock, K. F. (2001). Quantile regression, *journal of economic perspectives* 15(4): 143–156.
- Li, B. (2018). *Sufficient dimension reduction: Methods and applications with R*, Chapman and Hall/CRC.
- Li, R., Zhong, W. and Zhu, L. (2012). Feature screening via distance correlation learning, *journal of the American Statistical Association* 107(499): 1129–1139.
- Li, Y., Liu, Y. and Zhu, J. (2007). Quantile regression in reproducing kernel Hilbert spaces, *journal of the American Statistical Association* 102(477): 255–268.
- Lin, Y. and Zhang, H. H. (2006). Component selection and smoothing in multivariate nonparametric regression, *The Annals of Statistics* 34(5): 2272–2297.
- Luciani, M. and Ricci, L. (2013). *Nowcasting norway*, Available at SSRN 2211647 .
- Marcellino, M. and Schumacher, C. (2010). Factor midas for nowcasting and forecasting with ragged-edge data: A model comparison for german gdp, *Oxford Bulletin of Economics and Statistics* 72(4): 518–550.

- Matheson, T. D. (2010). An analysis of the informational content of new zealand data releases: The importance of business opinion surveys, *Economic Modelling* 27(1): 304–314.
- Moriwaki, D. (2019). Nowcasting unemployment rates with smartphone gps data, *International Workshop on Multiple-Aspect Analysis of Semantic Trajectories*, Springer, pp. 21–33.
- Raju, S. and Balakrishnan, M. (2019). Nowcasting economic activity in india using payment systems data, *journal of Payments Strategy & Systems* 13 (1): 72–81.
- Richardson, A., Mulder, T. et al. (2018). Nowcasting New Zealand GDP using machine learning algorithms.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso, *journal of the Royal Statistical Society: Series B (Methodological)* 58(1): 267–288.
- Zhang, T. (2002). Covering number bounds of certain regularized linear function classes, *journal of Machine Learning Research* 2(Mar): 527–550.
- Zhu, L.-P., Li, L., Li, R. and Zhu, L.-X. (2011). Model-free feature screening for ultrahigh-dimensional data, *journal of the American Statistical Association* 106(496): 1464–1475.
- Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net, *journal of the royal statistical society: series B (statistical methodology)* 67(2): 301–320.
- 김치호 and 김현정 (2016). GDP 성장률의 nowcasting에 관한 연구, [BOK] 국민계정 2016(2): 1–23.
- 서진용 (2017). 국내 가계수입과 소비지출의 추정을 위한 카드사용 정보의 활용 가능성 연구, *journal of The Korean Data Analysis Society* 19(1): 403–412.
- 이궁희 and 황상필 (2016). 빅데이터를 이용한 경기판단지표 개발: 네이버 검색 경기지수 작성과 유용성 검토, 한국은행 경제분석.

이현창, 최동규, 김용건 and 허정 (2022). 디지털 신기술을 이용한 실시간 당분기 경제전망 (GDP nowcasting) 시스템 개발, BOK 이슈노트 2022(7).

## [부록 1] 연구에 활용된 기계학습 방법론 추가 설명

### - SCAD 벌점화 선형회귀

주어진 자료  $(y_i, \mathbf{x}_i) \in \mathbb{R} \times \mathbb{R}^p$ 에 대해 다음과 같은 선형모형을 고려하자.

$$y_i = \beta_0 + \mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2),$$

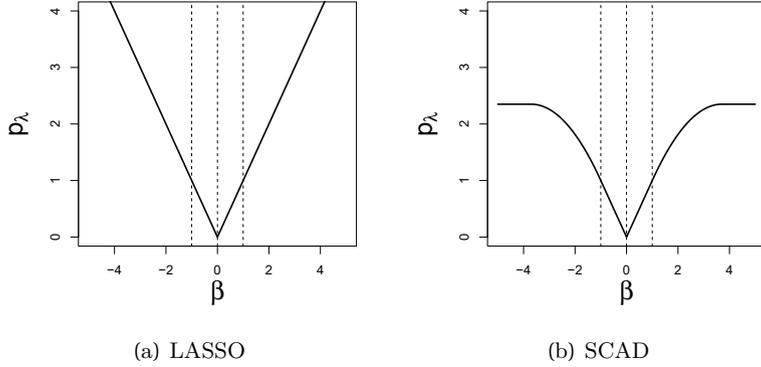
이 때 공변량의 개수  $p$ 가 많은 경우 다음과 같은 벌점화 회귀를 사용하여 회귀계수  $\boldsymbol{\beta}$ 를 추정할 수 있다.

$$\min_{\beta_0, \boldsymbol{\beta}} \sum_{i=1}^n (y_i - \beta_0 - \boldsymbol{\beta}^T \mathbf{x}_i)^2 + \sum_{j=1}^p p_\lambda(|\beta_j|)$$

여기서  $p_\lambda(\theta)$ 는 벌점함수를 나타내며, 대표적인 예로는 LASSO  $p_\lambda(\theta) = \lambda\theta$ 와 능형 벌점함수  $p_\lambda(\theta) = \frac{\lambda}{2}\theta^2$  등이 있다. 특히 LASSO 벌점함은 추정과 동시에 변수 선택도 가능하여 다양한 분야에 널리 활용된다. 하지만 LASSO 추정량은 편향이 크다는 사실이 알려져 있으며 이로 인해 추정의 효율이 떨어지게 된다. 이러한 단점을 개선하기 위한 방법으로 다음의 SCAD 벌점함수가 개발되었다 (Fan and Li, 2001). SCAD 벌점함수는 다음과 같이 도함수의 형태로 정의되며, LASSO와 달리 회귀계수의 값이 0에서 멀어지면 추가적인 벌점을 부여하지 않아 편향을 줄일 수 있다 (그림 A1 참조).

$$p'_\lambda(\theta) = \lambda \left\{ \mathbb{1}(\theta \leq \lambda) + \frac{(a\lambda - \theta)_+}{(a-1)\lambda} \mathbb{1}(\theta > \lambda) \right\}, \theta > 0$$

여기서  $a$ 는 2보다 큰 상수로 통상적으로 2.7을 사용한다. 다음은 LASSO와 SCAD 벌점함수를 비교한 그림이다. SCAD 벌점함은 더 좋은 추정량을 제공하지만 목적함수를 비볼록(non-convex)하게 만들어 그 계산이 더 어려워진다. 하지만 다양한 계산 알고리즘이 개발되었으며, 적어도 제곱손실함수를 이용한 회귀문제에서는 LASSO만큼이나 그 계산이 간단하다 (Breheny and Huang, 2011).



〈그림 A1〉 LASSO와 SCAD 벌점함의 비교: LASSO는 0에서 멀리 떨어질수록 벌점함의 증가하지만, SCAD는 0으로부터 떨어진 거리가 일정수준 이상이면 추가적인 벌점을 부여하지 않으므로써 편향을 줄여준다.

#### – 충분차원축소를 이용한 단일지표 모형의 추정

단일지표모형 (5)에서  $\beta$ 를 추정하기 위한 충분차원축소는 다음과 같은 가정에서 출발한다:

$$E(Y | \mathbf{X}) = E(Y | \beta^T \mathbf{X}) \quad (\text{A.1})$$

즉, (A.1)의 가정하에  $\beta$ 를 찾게 되면  $p$ 차원 공변량  $\mathbf{X}$ 를 1차원 공변량  $\beta^T \mathbf{X}$ 로 정보의 손실없이 축약할 수 있게 된다. 여기서 (A.1)의 가정을 만족하는  $\beta$ 는 유일하지 않지만,  $\beta$ 로 생성되는 공간은 유일하기 때문에 충분차원축약에서는  $\beta$ 가 아닌  $\beta$ 로 생성되는 공간, 즉  $\text{span}(\beta)$ 를 추정하는 것을 목표로 한다.  $\text{span}(\beta)$ 를 추정하는 방법은 매우 다양하지만, 가장 간단한 방법은 최소제곱추정량을 이용하는 것이다. 즉,  $\hat{\beta}$ 를 주어진 자료  $(y_i, \mathbf{x}_i) \in \mathbb{R} \times \mathbb{R}^p$ 로 부터 구한 최소제곱추정량이라 하면 다음이 성립한다.

$$\hat{\beta} \in \text{span}(\beta)$$

따라서 최소제곱추정법을 통해 우선  $\hat{\beta}$ 를 계산한 뒤  $\mathbf{z}_i = \mathbf{x}_i^T \hat{\beta}$ 와  $y_i$ 사이의 관계를 비모수적 방법(예, 스플라인 회귀 등)으로 추정하면 단일지표모형의 추정을 완료할 수 있다. 단, 본 연구에서는 공변량의 개수가 많기 때문에 최소제곱추정량을 바로

구할 수가 없으며, 대안으로 SCAD 벌점화 추정량을 구하여 변수선택을 진행한 뒤, 선택된 변수들만을 바탕으로 최소제곱 추정량을 계산하는 방법을 활용하였다.

### - 커널 분위수회귀

커널분위수회귀 (6)를 적합하기 위해서 대표자 정리 (Representer Theorem)를 이용할 수 있다. 대표자 정리는 (6)의 해가 다음과 같은 형태로 주어진다라는 것을 보장한다는 내용이다.

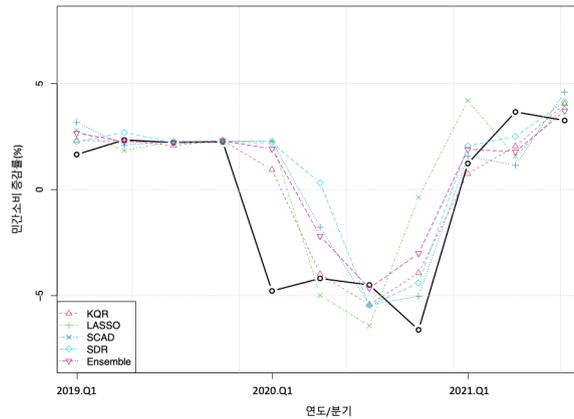
$$f(\mathbf{z}) = \alpha_0 + \sum_{i=1}^n y_i \alpha_i K(\mathbf{z}, \mathbf{z}_i) \quad (8)$$

이제, (8)를 (6)에 대입하면 다음과 같은 최적화 문제로 표현할 수 있다.

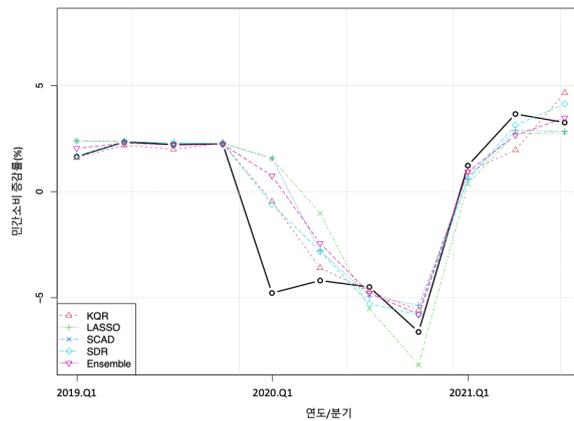
$$\min_{\alpha_0, \alpha} \sum_{i=1}^n \rho_{\tau} \left( y_i - \alpha_0 - \sum_{j=1}^n y_j \alpha_j K(\mathbf{z}_i, \mathbf{z}_j) \right) + \frac{\lambda}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j K(\mathbf{z}_i, \mathbf{z}_j)$$

위의 문제는 통상적인 이차계획법(quadratic programming)의 형태로 표현할 수 있음이 잘 알려져 있다(Li et al., 2007). 또한 추정해야하는 모수의 개수가  $p$ 가 아닌 표본의 크기  $n$ 에 의존하는 것을 알 수 있으며, 공변량  $\mathbf{z}$ 에 대한 정보가 커널함수  $K$ 를 통해서만 전달되기 때문에  $\mathbf{z}$ 가 고차원인 경우에도 계산량을 증가시키지 않는다는 장점이 있다.

## [부록 2] AR(1)항을 포함한 ( $L = 0$ ) 모형 예측 결과



(a) 변수선별을 하지 않은 경우



(b) 변수선별을 한 경우

〈그림 A2〉 AR(1)항을 포함한 ( $L = 0$ ) 모형간의 민간소비 예측치 비교:  
변수선별의 효과를 확인할 수 있다.

## [부록 3] 공변량 전체 목록

설명	원_지수	주기	시작시점	공표지연	분석포함
<b>환율</b>					
원/미국달러(매매기준율)	원자료	일	2000-01-04	1	○
원/일본엔	원자료	일	2000-01-04	1	○
원/유로	원자료	일	2000-01-04	1	○
<b>고빈도데이터</b>					
동행종합지수	지수	월	2012-01-31	0	X
GDP(성장률)	지수	분기	2012-03-31	2	X
최대전력사용량	원자료	일	2012-01-01	1	○
개인신용카드사용액	원자료	일	2016-01-01	1	○
수출액	원자료	일	2012-01-01	1	○
수입액	원자료	일	2012-01-01	1	○
영화매출액	원자료	일	2012-01-01	1	○
고속도로통행량	원자료	일	2012-01-01	1	○
<b>통화금융통계</b>					
M2(광의통화,평잔)(십억원)	원자료	월	2000-01-31	2	X
LF(평잔)(십억원)	원자료	월	2000-01-31	2	X
<b>카드데이터</b>					
식료품 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
식료품 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
의류 및 신발 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
의류 및 신발 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
주택 및 연료 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
주택 및 연료 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
가구, 가사 비품 및 기타 생활용품 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
가구, 가사 비품 및 기타 생활용품 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
보건 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
보건 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
개인운송장비 구매 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
개인운송장비 구매 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
개인운송장비 운영 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
개인운송장비 운영 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
여객운송서비스 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
여객운송서비스 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
상품운송서비스 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
상품운송서비스 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
정보통신 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
정보통신 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
오락용품 및 애완동물 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
오락용품 및 애완동물 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
오락문화서비스 및 여행 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
오락문화서비스 및 여행 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
교육서비스 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
교육서비스 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
일반음식점 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
일반음식점 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
주점 등 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
주점 등 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
숙박서비스 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
숙박서비스 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
보험 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
보험 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
개인미용 및 기타 서비스 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
개인미용 및 기타 서비스 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
백화점 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
백화점 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
할인점 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
할인점 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
편의점 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○

설명	원_지수	주기	시작시점	공표지연	분석포함
편의점 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
홈쇼핑 및 인터넷판매 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
홈쇼핑 및 인터넷판매 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
면세점 및 기타 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
면세점 및 기타 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
상품권+세금공과금 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
상품권+세금공과금 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
기타 기간중 카드 사용 금액	원자료	월	2015-11-30	0	○
기타 기간중 카드 사용 건수	원자료	월	2015-11-30	0	○
<b>경기종합지수</b>					
선행종합지수	지수	월	2000-01-31	0	○
동행종합지수	지수	월	2000-01-31	0	○
후행종합지수	지수	월	2000-01-31	0	○
<b>경제심리지수</b>					
경제심리지수	지수	월	2003-01-31	0	○
<b>경제활동인구조사-연령별</b>					
15 - 19세__경제활동인구 (천명)	원자료	월	2000-01-31	1	○
15 - 19세__비경제활동인구 (천명)	원자료	월	2000-01-31	1	○
15 - 19세__실업자 (천명)	원자료	월	2000-01-31	1	○
15 - 19세__취업자 (천명)	원자료	월	2000-01-31	1	○
20 - 29세__경제활동인구 (천명)	원자료	월	2000-01-31	1	○
20 - 29세__비경제활동인구 (천명)	원자료	월	2000-01-31	1	○
20 - 29세__실업자 (천명)	원자료	월	2000-01-31	1	○
20 - 29세__취업자 (천명)	원자료	월	2000-01-31	1	○
30 - 39세__경제활동인구 (천명)	원자료	월	2000-01-31	1	○
30 - 39세__비경제활동인구 (천명)	원자료	월	2000-01-31	1	○
30 - 39세__실업자 (천명)	원자료	월	2000-01-31	1	○
30 - 39세__취업자 (천명)	원자료	월	2000-01-31	1	○
40 - 49세__경제활동인구 (천명)	원자료	월	2000-01-31	1	○
40 - 49세__비경제활동인구 (천명)	원자료	월	2000-01-31	1	○
40 - 49세__실업자 (천명)	원자료	월	2000-01-31	1	○
40 - 49세__취업자 (천명)	원자료	월	2000-01-31	1	○
50 - 59세__경제활동인구 (천명)	원자료	월	2000-01-31	1	○
50 - 59세__비경제활동인구 (천명)	원자료	월	2000-01-31	1	○
50 - 59세__실업자 (천명)	원자료	월	2000-01-31	1	○
50 - 59세__취업자 (천명)	원자료	월	2000-01-31	1	○
60세이상__경제활동인구 (천명)	원자료	월	2000-01-31	1	○
60세이상__비경제활동인구 (천명)	원자료	월	2000-01-31	1	○
60세이상__실업자 (천명)	원자료	월	2000-01-31	1	○
60세이상__취업자 (천명)	원자료	월	2000-01-31	1	○
<b>미분양주택현황보고-규모별미분양</b>					
공공부문미분양	원자료	월	2007-01-31	1	○
민간부문미분양	원자료	월	2007-01-31	1	○
<b>전국주택가격동향조사-규모별월세가격지수</b>					
아파트전국__규모40㎡이하	지수	월	2017-11-30	1	○
아파트전국__규모40㎡ 60㎡	지수	월	2015-06-30	1	○
아파트전국__규모60㎡ 85㎡	지수	월	2015-06-30	1	○
아파트전국__규모85㎡ 102㎡	지수	월	2015-06-30	1	○
아파트전국__규모102㎡ 135㎡	지수	월	2015-06-30	1	○
아파트전국__규모135㎡초과	지수	월	2015-06-30	1	○
연립전국__규모40㎡이하	지수	월	2017-11-30	1	○
연립전국__규모40㎡ 60㎡	지수	월	2015-06-30	1	○
연립전국__규모60㎡ 85㎡	지수	월	2015-06-30	1	○
연립전국__규모85㎡초과	지수	월	2015-06-30	1	○
<b>사업체노동력조사-고용부문</b>					
300인미만__근로자__상용 (명)	원자료	월	2009-06-30	1	○
300인미만__근로자__임시일용 (명)	원자료	월	2009-06-30	1	○
300인미만__종사자__기타 (명)	원자료	월	2009-06-30	1	○
300인이상__근로자__상용 (명)	원자료	월	2009-06-30	1	○
300인이상__근로자__임시일용 (명)	원자료	월	2009-06-30	1	○
300인이상__종사자__기타 (명)	원자료	월	2009-06-30	1	○
<b>통화금융통계-대출금리</b>					
기업대출(연리%)	원자료	월	2000-01-31	0	○

설명	원_지수	주기	시작시점	공표지연	분석포함
가계대출(연리%)	원자료	월	2000-01-31	0	○
공공및기타부문대출(연리%)	원자료	월	2000-01-31	0	○
<b>생산자물가지수</b>					
생산자물가지수	지수	월	2000-01-31	1	○
<b>소비자동향조사</b>					
현재생활형편	지수	월	2009-12-31	0	○
현재경기관단	지수	월	2009-12-31	0	○
생활형편전망	지수	월	2009-12-31	0	○
향후경기전망	지수	월	2009-12-31	0	○
취업기회전망	지수	월	2009-12-31	0	○
금리수준전망	지수	월	2009-12-31	0	○
가계수입전망	지수	월	2009-12-31	0	○
소비지출전망	지수	월	2009-12-31	0	○
내구재지출전망	지수	월	2009-12-31	0	○
의류비지출전망	지수	월	2009-12-31	0	○
의식비지출전망	지수	월	2009-12-31	0	○
여행비지출전망	지수	월	2009-12-31	0	○
교육비지출전망	지수	월	2009-12-31	0	○
의료,보건비지출전망	지수	월	2009-12-31	0	○
교양,오락,문화생활비지출전망	지수	월	2009-12-31	0	○
교통비및통신비지출전망	지수	월	2009-12-31	0	○
주거비지출전망	지수	월	2013-01-31	0	○
현재가계지축	지수	월	2009-12-31	0	○
가계지축전망	지수	월	2009-12-31	0	○
현재가계부채	지수	월	2009-12-31	0	○
가계부채전망	지수	월	2009-12-31	0	○
물가수준전망(1년후)	지수	월	2009-12-31	0	○
주택가격전망	지수	월	2013-01-31	0	○
임금수준전망	지수	월	2013-01-31	0	○
소비자심리지수	지수	월	2009-12-31	0	○
<b>소비자동향조사-기대인플레이션</b>					
물가인식	지수	월	2013-01-31	0	○
기대인플레이션율	지수	월	2002-02-28	0	○
<b>소비자물가지수</b>					
소비자물가지수	지수	월	2000-01-31	1	○
<b>통화금융통계-수신금리</b>					
순수저축성예금(연리%)	원자료	월	2000-01-31	0	○
시장형금융상품(연리%)	원자료	월	2000-01-31	0	○
<b>외국인국제이동</b>					
입국자	원자료	월	2013-01-31	6	X
출국자	원자료	월	2013-01-31	6	X
<b>전산업생산지수</b>					
전산업생산지수(농림어업 제외)	지수	월	2000-01-31	1	○
광공업생산지수	지수	월	2000-01-31	1	○
건설업생산지수	지수	월	2000-01-31	1	○
서비스업생산지수	지수	월	2000-01-31	1	○
공공행정생산지수	지수	월	2000-01-31	1	○
<b>주민등록총인구</b>					
주민등록총인구	원자료	월	2011-01-31	1	X
<b>식유제품가격통계-주요소 평균 판매가격</b>					
보통휘발유	원자료	월	2000-01-31	1	○
실내등유	원자료	월	2000-01-31	1	○
자동차용 경유	원자료	월	2000-01-31	1	○
<b>소비 유형별 전국 신용카드 소비</b>					
가구/가전-가구	원자료	월	2009-12-31	0	○
가구/가전-가전제품/정보통신기기	원자료	월	2009-12-31	0	○
공과금/개인 및 전문 서비스	원자료	월	2009-12-31	0	○
교육	원자료	월	2009-12-31	0	○
금융/보험	원자료	월	2009-12-31	0	○
기타	원자료	월	2009-12-31	0	○
숙박/음식-숙박	원자료	월	2009-12-31	0	○
숙박/음식-음식점	원자료	월	2009-12-31	0	○
식료품-건강보조식품	원자료	월	2009-12-31	0	○

설명	원_지수	주기	시작시점	공표지연	분석포함
식료품-일반식료품	원자료	월	2009-12-31	0	○
여행/교통-대중교통	원자료	월	2009-12-31	0	○
여행/교통-여행사/자동차임대	원자료	월	2009-12-31	0	○
여행/교통-항공사	원자료	월	2009-12-31	0	○
연료	원자료	월	2009-12-31	0	○
오락/문화-서적/문구	원자료	월	2009-12-31	0	○
오락/문화-스포츠/오락/여가	원자료	월	2009-12-31	0	○
의료/보건-일반병원/기타의료기관	원자료	월	2009-12-31	0	○
의료/보건-종합병원	원자료	월	2009-12-31	0	○
의류/잡화-복식잡화	원자료	월	2009-12-31	0	○
의류/잡화-시계/귀금속/안경	원자료	월	2009-12-31	0	○
의류/잡화-의복/직물	원자료	월	2009-12-31	0	○
의류/잡화-화장품	원자료	월	2009-12-31	0	○
자동차-국산자동차신품	원자료	월	2009-12-31	0	○
자동차-기타운송수단	원자료	월	2009-12-31	0	○
자동차-자동차 부품 및 정비	원자료	월	2009-12-31	0	○
전자상거래/통신판매	원자료	월	2009-12-31	0	○
종합소매-대형마트/유통전문점	원자료	월	2009-12-31	0	○
종합소매-면세점	원자료	월	2009-12-31	0	○
종합소매-백화점	원자료	월	2009-12-31	0	○
종합소매-슈퍼마켓	원자료	월	2009-12-31	0	○
종합소매-편의점	원자료	월	2009-12-31	0	○
	추가지수				
코스피	원자료	월	2000-02-28	1	○
코스닥	원자료	월	2003-12-31	1	○
	소매업태별 판매액지수				
백화점	지수	월	2010-01-31	1	○
대형마트	지수	월	2010-01-31	1	○
면세점	지수	월	2010-01-31	1	○
체인 슈퍼마켓	지수	월	2010-01-31	1	○
일반 슈퍼마켓 및 잡화점	지수	월	2010-01-31	1	○
편의점	지수	월	2010-01-31	1	○
승용차 소매점	지수	월	2010-01-31	1	○
연료 소매점	지수	월	2010-01-31	1	○
가전, 컴퓨터, 통신기기 소매점	지수	월	2010-01-31	1	○
의복, 신발, 가방 소매점	지수	월	2015-01-31	1	○
음식, 가정, 문화상품 소매점	지수	월	2015-01-31	1	○
의약품, 화장품, 기타상품 소매점	지수	월	2015-01-31	1	○
인터넷 쇼핑	지수	월	2010-01-31	1	○
홈쇼핑	지수	월	2010-01-31	1	○
방문 및 배달 소매점	지수	월	2010-01-31	1	○
	텍스트				
경제불확실성지수	지수	일	2005-01-15	0	○
뉴스심리지수	지수	일	2005-01-03	0	○
경제뉴스텍스트에서 추출한 소비지표	지수	일	2005-01-01	0	○
경제뉴스텍스트에서 추출한 물가지표	지수	일	2005-01-08	0	○

## BOK 경제연구 발간목록

한국은행 경제연구원에서는 Working Paper인 『BOK 경제연구』를 수시로 발간하고 있습니다. 『BOK 경제연구』는 주요 경제 현상 및 정책 효과에 대한 직관적 설명 뿐 아니라 깊이 있는 이론 또는 실증 분석을 제공함으로써 엄밀한 논증에 초점을 두는 학술논문 형태의 연구이며 한국은행 직원 및 한국은행 연구용역사업의 연구 결과물이 수록되고 있습니다. 『BOK 경제연구』는 한국은행 경제연구원 홈페이지(<http://imer.bok.or.kr>)에서 다운로드하여 보실 수 있습니다.

제2019-1	Deciphering Monetary Policy Board Minutes through Text Mining Approach: The Case of Korea	Ki Young Park · Youngjoon Lee · Soohyon Kim
2	The Impacts of Macroeconomic News Announcements on Intraday Implied Volatility	Jieun Lee · Doojin Ryu
3	Taking a Bigger Slice of the Global Value Chain Pie: An Industry-level Analysis	Chong-Sup Kim · Seungho Lee · Jihyun Eum
4	Trend Growth Shocks and Asset Prices	Nam Gang Lee
5	Uncertainty, Attention Allocation and Monetary Policy Asymmetry	Kwangyong Park
6	Central Bank Digital Currency and Financial Stability	Young Sik Kim · Ohik Kwon
7	은행의 수익 및 자산구조를 반영한 통화정책 위험선호경로	김익진 · 정호성
8	혁신기업에 대한 산업금융 지원: 이론모형 분석	강경훈 · 양준구
9	가계부채 제약하의 통화정책: 2주체 거시모형(TANK)에서의 정량적 분석	정용승 · 송승주
10	Alchemy of Financial Innovation: Securitization, Liquidity and Optimal Monetary Policy	Jungu Yang
11	Measuring Monetary Policy Surprises Using Text Mining: The Case of Korea	Youngjoon Lee · Soohyon Kim · Ki Young Park
12	Tracking Uncertainty through the Relative Sentiment Shift Series	Seohyun Lee · Rickard Nyman
13	Intra-firm and Arm's Length Trade during the Global Financial Crisis: Evidence from Korean Manufacturing Firms	Moon Jung Choi · Ji Hyun Eum

---

14	특허자료를 이용한 우리나라 지식전파의 지역화 분석	이지홍 · 남윤미
15	Overhead Labour and Skill-Biased Technological Change: The Role of Product Diversification	Choong Hyun Nam
16	Does the Number of Countries in an International Business Cycle Model Matter?	Myunghyun Kim
17	High-Frequency Credit Spread Information and Macroeconomic Forecast Revision	Bruno Deschamps · Christos Ioannidis · Kook Ka
18	경제 분석을 위한 텍스트 마이닝	김수현 · 이영준 · 신진영 · 박기영
19	Takeover, Distress, and Equity Issuance: Evidence from Korea	Euna Cho
20	The Cash-Flow Channel of Monetary Policy: Evidence from Mortgage Borrowers	Sang-yoon Song
21	부의 효과의 분위 추정: 분위 정준 공적분 회귀를 중심으로	김기호
22	Identifying Government Spending Shocks and Multipliers in Korea	Kwangyong Park · Eun Kyung Lee
23	Systemic Risk of the Consumer Credit Network across Financial Institutions	Hyun Hak Kim · Hosung Jung
24	Impact of Chinese Renminbi on Korean Exports: Does Quality Matter?	Jihyun Eum
25	Uncertainty, Credit and Investment: Evidence from Firm-Bank Matched Data	Youngju Kim · Seohyun Lee · Hyunjoon Lim
26	A Structural Change in the Trend and Cycle in Korea	Nam Gang Lee · Byoung Hoon Seok

---

---

제2020 -1	인구 고령화가 실질 금리에 미치는 영향	권오익 · 김명현
2	달러라이제이션이 확산된 북한경제에서 보 유외화 감소가 물가 · 환율에 미치는 영향	문성민 · 김병기
3	상태공간 벡터오차수정모형을 이용한 월별 GDP 추정: 깃스표본추출 접근	김기호
4	우리나라 외환시장 오퍼레이션의 행태 및 환율변동성 완화 효과	박준서 · 최경욱
5	Common Factor Augmented Forecasting Models for the US Dollar–Korean Won Exchange Rate	Hyeongwoo Kim · Soohyon Kim
6	북한 「경제연구」로 분석한 경제정책 변화: 텍스트 마이닝 접근법	김수현 · 손 옥
7	북한의 광물 수출과 품목별 수입: 대중무역을 중심으로	김병연 · 김민정 · 김다울
8	Network–Based Measures of Systemic Risk in Korea	Jaewon Choi · Jieun Lee
9	Aggregate Productivity Growth and Firm Dynamics in Korean Manufacturing 2007–2017	Kyoo il Kim · Jin Ho Park
10	2001년 이후 한국의 노동생산성 성장과 인적자본: 교육의 질적 개선 효과를 중심으로	유혜미
11	House Prices and Household Consumption in Korea	Seungyoon Lee
12	글로벌 가치사슬 변화가 경제성장에 미치는 영향: 2008년 금융위기 전후 전 · 후방참여 효과의 국제비교를 중심으로	김세완 · 최문정
13	산업구조조정이 고용 및 성장에 미치는 영향	서병선 · 김태경
14	Cross–border Trade Credit and Trade Flows During the Global Financial Crisis	Moon Jung Choi · Sangyeon Hwang · Hyejoon Im

---

- 
- |    |  |  |
|----|--|--|
| 15 | International Co-movements and Determinants of Public Debt                               | Hasan Isomitdinov · Vladimir Arčabić · Junsoo Lee · Youngjin Yun |
| 16 | 북한 비공식금융 실태조사 및 분석 · 평가  | 이주영 · 문성민  |
| 17 | 북한의 장기 경제성장률 추정: 1956~1989년  | 조태형 · 김민정  |
| 18 | Macroeconomic and Financial Market Analyses and Predictions through Deep Learning        | Soohyon Kim  |
| 19 | 제조업의 수출과 생산성 간 관계 분석: 사업체 자료 이용  | 이윤수 · 김원혁 · 박진호  |
| 20 | 우리나라 제조업 수출기업의 내수전환 결정요인 분석  | 남윤미 · 최문정  |
| 21 | A Model of Satisficing Behaviour   | Rajiv Sarin · Hyun Chang Yi                                      |
| 22 | Vulnerable Growth: A Revisit   | Nam Gang Lee   |
| 23 | Credit Market Frictions and Coessentiality of Money and Credit                           | Ohik Kwon · Manjong Lee  |
| 24 | 북한의 자본스톡 추정 및 시사점  | 표학길 · 조태형 · 김민정  |
| 25 | The Economic Costs of Diplomatic Conflict  | Hyejin Kim · Jungmin Lee   |
| 26 | Central Bank Digital Currency, Tax Evasion, Inflation Tax, and Central Bank Independence | Ohik Kwon · Seungduck Lee · Jaevin Park                          |
| 27 | Consumption Dynamics and a Home Purchase   | Dongjae Jung   |
| 28 | 자본유입과 물가상승률 간의 동태적 상관관계 분석: 아시아의 8개국 소규모 개방경제를 중심으로                                      | 최영준 · 손종철  |
-

---

29	The Excess Sensitivity of Long-term Interest rates and Central Bank Credibility	Kwangyong Park
30	Wage and Employment Effects of Immigration: Evidence from Korea	Hyejin Kim
제2021-1	외국인력 생산성 제고 방안—직업훈련 프로그램의 노동시장 성과 분석을 중심으로	김혜진 · 이철희
2	한국경제의 추세 성장을 하락과 원인	석병훈 · 이남강
3	Financial Globalization: Effects on Banks' Information Acquisition and Credit Risk	Christopher Paik
4	The Effects of Monetary Policy on Consumption: Workers vs. Retirees	Myunghyun Kim · Sang-yeon Song
5	북한지역 토지자산 추정에 관한 연구: 프레임워크 개발 및 탐색적 적용	임송
6	김정은 시대 북한의 금융제도 변화 - 북한 문헌 분석을 중심으로 -	김민정 · 문성민
7	Chaebols and Firm Dynamics in Korea	Philippe Aghion · Sergei Guriev · Kangchul Jo
8	한국의 화폐환상에 관한 연구	권오익 · 김규식 · 황인도
9	재원조달 방법을 고려한 재정지출 효과 분석 : 미국의 사례를 중심으로	김소영 · 김용건
10	The Impact of Geopolitical Risk on Stock Returns: Evidence from Inter-Korea Geopolitics	Seungho Jung · Jongmin Lee · Seohyun Lee
11	Real Business Cycles in Emerging Countries: Are Asian Business Cycles Different from Latin American Business Cycles?	Seolwoong Hwang · Soyoung Kim
12	우리 수출의 글로벌 소득탄력성 하락 요인 분석	김경근
13	북한의 경제체제에 관한 연구: 실태와 평가	양문수 · 임송

---

---

14	Distribution–Dependent Value of Money: A Coalition–Proof Approach to Monetary Equilibrium	Byoung–Ki Kim · Ohik Kwon · Suk Won Lee
15	A Parametric Estimation of the Policy Stance from the Central Bank Minutes	Dong Jae Jung
16	The Immigrant Wage Gap and Assimilation in Korea	Hyejin Kim · Chulhee Lee
17	Monetary Non–Neutrality in a Multisector Economy: The Role of Risk–Sharing	Jae Won Lee · Seunghyeon Lee
18	International Transmission of Chinese Monetary Policy Shocks to Asian Countries	Yujeong Cho · Soyoung Kim
19	The Impact of Robots on Labor Demand: Evidence from Job Vacancy Data for South Korea	Hyejin Kim
20	전공 불일치가 불황기 대졸 취업자의 임금에 미치는 장기 효과 분석	최영준
21	Upstream Propagation of the U.S.–China Trade War	Minkyu Son
제2022 –1	Immigration and Natives’ Task Specialization: Evidence from Korea	Hyejin Kim · Jongkwan Lee
2	Transmission of Global Financial Shocks: Which Capital Flows Matter?	Bada Han
3	Measuring the Effects of LTV and DTI Limits: A Heterogeneous Panel VAR Approach with Sign Restrictions	Soyoung Kim · Seri Shim
4	A Counterfactual Method for Demographic Changes in Overlapping Generations Models	Byongju Lee
5	Housing Wealth, Labor Supply, and Retirement Behavior: Evidence from Korea	Jongwoo Chung

---

- 
- |    |   |   |
|----|---|---|
| 6  | Demand Shocks vs. Supply Shocks: Which Shocks Matter More in Income and Price Inequality?                 | Seolwoong Hwang · Kwangwon Lee · Geunhyung Yim          |
| 7  | Financial Literacy and Mutual Fund Retail Investing: Evidence from Korea During the 2008 Financial Crisis | Jongwoo Chung · Booyuel Kim                             |
| 8  | Exchange Rate Regime and Optimal Policy: The Case of China  | Yujeong Cho · Yiping Huang · Changhua Yu                |
| 9  | 북한 수출입단가지수 추정: 북중무역 데이터를 중심으로   | 이종민 · 김민정   |
| 10 | 탄소배출을 감안한 국가별 녹색 총요소생산성 분석  | 안상기   |
| 11 | 북한 소비자 지급수단 조사 및 분석   | 이주영   |
| 12 | Selection into Outsourcing versus Integration Strategies for Heterogeneous Multinationals                 | Sangho Shin   |
| 13 | Central Bank Digital Currency and Privacy: A Randomized Survey Experiment                                 | Syngjoo Choi · Bongseop Kim · Young Sik Kim · Ohik Kwon |
| 14 | Technological Change, Job Characteristics, and Employment of Elderly Workers: Evidence from Korea         | Jongwoo Chung · Chulhee Lee                             |
| 15 | Machine-Learning-Based News Sentiment Index (NSI) of Korea  | Beomseok Seo · Younghwan Lee · Hyungbae Cho             |
| 16 | 빅데이터를 이용한 실시간 민간소비 예측   | 신승준 · 서범석   |
-



---

## BOK 경제연구 제 2022-16

---

2022년 9월 26일 인쇄

2022년 9월 26일 발행

발행인 이 창 용

편집인 박 양 수

발행처 한국은행

경제연구원

서울시 중구 남대문로 55 (소공동 41번지)

인쇄처 (사)한국장애인문화콘텐츠협회

서울시 중구 수표로 10길 10 (저동 2가)

---

© 한국은행, 2022

■ 본 자료는 한국은행 홈페이지에서 무료로 다운로드 받으실 수 있습니다. <http://www.bok.or.kr>

