



# 단기 재정위험 진단 방법 연구

2024. 12.

오종현 · 배진수



# 단기 재정위험 진단 방법 연구

2024. 12.

오종현 · 배진수



## 서 언

우리나라 재정의 지속가능성이 위협받고 있다. 코로나바이러스감염증-19 확산 등의 위기에 대응하기 위해 정부가 적극적인 재정정책을 펼치면서 최근 국가채무가 빠르게 증가하였다. 2019년에 GDP의 35.4%였던 국가채무는 코로나바이러스감염증-19가 처음 발생한 2020년에 GDP의 40%를 넘어섰고, 2023년에는 GDP의 46.9%에 이르렀다. GDP 대비 국가채무 비율이 최근 4년간 11.5%p 상승한 것으로 그 상승 속도가 매우 가팔랐다. 우리나라 재정의 지속가능성 문제는 장기적으로도 우려를 낳고 있는 상황이다. 저출생·고령화로 인한 인구구조 변화는 재정을 심각하게 위협하는 구조적인 요인이다. 이뿐만 아니라 기후위기 대응이나 통일에 따른 비용 등도 미래 재정의 잠재적인 위협 요인이다.

재정의 지속가능성 악화로 재정의 체계적인 관리가 강조되고 있다. 재정을 체계적으로 관리하기 위해서는 재정준칙, 중기재정관리체계, 독립적인 재정기구 등 다양한 요소가 필요하다. 이러한 요소 중의 하나가 바로 재정위험을 체계적으로 진단하는 것이다. 정부가 국가채무나 재정수지와 같은 재정지표에 중장기적인 목표 수준을 설정하고, 이러한 목표에서 벗어났을 때 이 목표에 수렴하기 위한 경로 등을 제시할 필요가 있다. 이때 이러한 목표와 경로를 투명하고 신뢰성 있게 제시하는 데 가장 기본이 되는 것이 정량적인 근거를 제공할 수 있는 재정위험 진단이다.

본 연구는 기계학습(machine learning) 기법을 활용하여 단기 재정위험을 진단하는 방법론을 개발한다. 재정위험은 다양한 시계에서 다양한 방법으로 진단할 수 있다. 그간 우리나라에서 재정위험에 대한 논의는 주로 장기재정전망을 중심으로 이루어졌다. 장기재정전망을 통한 재정위험 진단은 인구구조 변화라는 구조적이고 장기적인 위험 요인을 식별하고 평가하는 것이 목적이다. 반면 우리나라에서 단기 재정위험 진단에 대한 논의는 그리 활발한



편은 아니었다. 그러나 최근 재정의 지속가능성이 급격히 악화되면서 단기적인 재정위험 진단에 대한 관심도 높아지고 있다. 동시에 해외 연구들은 최근 기계학습 기법이 단기 재정위험 진단에 유용하다는 결과들을 제시하고 있다. 이러한 상황에서 본 연구는 기존의 방법론이 아닌 기계학습이라는 새로운 방법론으로 재정위험을 진단하여 우리나라 재정위험 분석의 폭을 넓히는 데 의의가 있다.

본 보고서는 본원의 오종현 연구위원과 한국금융연구원의 배진수 부연구위원이 공동으로 집필하였다. 저자들은 본 연구에 도움을 주신 많은 분들께 감사의 인사를 전한다. 먼저 연구의 착수 단계부터 최종 출판 단계까지 많은 검토와 조언을 주신 이화여대 송호신 교수, 충남대학교 신상화 교수, 그리고 본원의 고창수 부연구위원, 김정환 부연구위원께 감사드린다. 착수, 중간, 최종보고회에 직접 참여하여 토론을 통해 소중한 의견을 주신 동료 및 선후배 박사들과 외부 전문가들께도 감사드린다. 그리고 최종 출판 단계에서 유익한 조언을 해 주신 두 분의 익명의 심사자께도 감사드린다. 그리고 본 연구를 진행하면서 자료 수집 및 정리 등에 도움을 주신 본원의 김인유, 장준희 선임연구원께도 감사드린다. 그 밖에 본 연구를 진행하는 데 있어 행정적으로 아낌없이 지원해 주신 많은 분들께 감사드린다.

마지막으로 본 보고서의 모든 내용은 본원의 공식적인 의견이나 입장이 아닌 저자들의 개인적인 견해를 밝힌다.

2024년 12월

한국조세재정연구원  
원장 이 영

# 요약 및 정책적 시사점

## 1. 연구 목적

본 연구는 기계학습(machine learning) 기법을 활용하여 단기 재정위험을 진단하는 방법론을 개발한다. 재정위험은 다양한 시계에서 다양한 방법으로 진단할 수 있다. 그간 우리나라에서 재정위험에 대한 논의는 주로 장기재정 전망을 중심으로 이루어졌다. 장기재정전망을 통한 재정위험 진단은 인구구조 변화라는 구조적이고 장기적인 위험 요인을 식별하고 평가하는 것이 목적이다. 반면 우리나라에서 단기 재정위험 진단에 대한 논의는 그리 활발한 편은 아니었다. 그러나 최근 재정의 지속가능성이 급격히 악화되면서 단기적인 재정위험 진단에 대한 관심도 높아지고 있다. 동시에 해외 연구들은 최근 기계학습 기법이 단기 재정위험 진단에 유용하다는 결과들을 제시하고 있다. 이러한 상황에서 본 연구는 기존의 방법론이 아닌 기계학습이라는 새로운 방법론으로 재정위험을 진단하여 우리나라 재정위험 분석의 폭을 넓히는 데 의의가 있다.

## 2. 선행연구와의 차별점

본 연구는 방법론 측면에서 기계학습 기법을 이용하여 재정위험을 진단하여 우리나라의 재정위험을 진단하는 선행연구들과 차이가 있다. 우리나라의 재정위험은 주로 장기재정전망을 통해 진단되는데, 이는 재정의 다양한 구성요소를 각각 추정한 뒤 통합하는 회계적 방법론을 사용한다. 최근에는 김명규·김민경(2023)과 한종석 외(2024)와 같이 장기재정전망이 아닌 다른 방법으로 우리나라 재정의 지속가능성을 평가하거나 진단한 연구도 있다. 그러나 두 연구 모두 유럽연합 집행위원회(European Commission)나 국제통화

기금(International Monetary Fund) 등 국제기구가 주기적으로 발표하는 재정위험 진단 보고서의 기본적인 방법론을 우리나라에 적용해 본 것으로, 본 연구의 방법론과는 차이가 있다.

본 연구는 재정위험의 정의에 있어 국내외 선행연구들과 차이가 있다. 선행 연구들은 재정위험(fiscal risk)을 채무불이행과 같은 심각한 재정위기(fiscal crisis)가 발생할 가능성으로 정의한다. 반면 본 연구는 재정위기를 향후 1~3년간 GDP 대비 정부부채 비율이 일정 수준 이상 증가하는 상황으로 정의하고, 이러한 재정위기가 발생할 가능성을 재정위험으로 정의하였다. 특히 기본 분석에서는 GDP 대비 정부부채 비율이 연평균 3%p 이상 증가하는 경우, 즉 1년간 3%p, 2년간 6%p, 3년간 9%p 이상 증가하는 경우를 재정위기가 발생한 것으로 정의하였다. 재정위험으로 채무불이행과 같은 심각한 재정위기보다 정부부채 비율이 급증할 가능성을 식별하는 것이 재정의 지속가능성을 관리하는 측면에서 우리나라의 현실에 더 부합하는 것으로 판단된다. 재정의 지속가능성을 가늠하는 핵심적인 지표는 GDP 대비 국가채무 또는 정부부채 비율이다. 이 비율이 안정적이지 않고 지속적으로 증가한다면 일반적으로 재정은 지속가능하지 않다고 말한다. 특히 이 비율이 너무 높아진다면 심각한 재정위기에 직면할 가능성이 커진다. 따라서 재정위기를 사전에 방지하기 위해서는 단기적으로 정부부채 비율의 증가 속도를 관리할 필요가 있다.

### 3. 연구 방법론

본 연구는 크게 두 가지의 분석을 수행한다. 먼저 우리나라에 적합한 재정위험을 정의하고 이를 정량적으로 측정하는 방법에 대해 논의한다. 재정위험을 진단하기 위해서는 우선 종속변수인 과거에 이미 발생한 재정위기를 식별하는 것이 중요하다. 재정위기의 측정 방법을 선택할 때 자료의 이용가능성과 적용의 용이성 등을 고려하여 자료에서 과거에 드문 빈도로 재정의 지속가능성을 악화시키는 방향으로 관찰되었던 사건들을 기준으로 재정위기를 정의한다. 그리고 이렇게 정의된 재정위기에 대해 통계적 기법으로 정형적

사실을 정리한다.

다음으로 단기 재정위험 진단에 현재 널리 활용되고 있거나 선행연구에서 최근에 예측력이 높은 것으로 분류된 몇 가지 방법론을 비교하여 앞에서 정의한 재정위기를 예측하는 방법을 개발한다. 구체적인 재정위험 진단 방법으로 신호(signaling), 로짓(logit), 라쏘(Least Absolute Shrinkage and Selection Operator, LASSO), 랜덤포레스트(random forest) 모형을 비교·검토한다. 신호, 로짓 모형은 전통적인 방법, 그리고 라쏘, 랜덤포레스트 모형은 기계학습 방법이다. 모형의 예측력을 비교하기 위해 정확도(accuracy), 정밀도(precision), 재현율(recall), F1 점수, F2 점수, AUC(Area Under the Curve) 등의 다양한 지표를 검토한다. 재정위험 진단을 위한 변수들은 국제통화기금(International Monetary Fund)이 주기적으로 발표하는 World Economic Outlook(IMF-WEO)의 자료와 경제협력개발기구(Organization for Economic Co-operation and Development)의 Economic Outlook(OECD-EO) 자료를 활용한다. 그리고 기본 분석의 대상 국가는 IMF-WEO에서 선진국으로 분류하는 국가 중 OECD-EO에도 자료가 존재하는 32개 국가이다.

#### 4. 주요 결과 및 정책시사점

재정위기 전후의 재정 및 거시경제 변수의 변화를 살펴본 결과는 다음과 같다. 재정위기가 발생한 기간에 기초재정수지와 정부부채 등 주요 재정지표가 크게 악화되고, 이러한 영향이 일시적인 것이 아닌 일정 기간 지속되는 것으로 나타났다. 특히 재정위기 기간이 끝난 이후에도 높은 수준의 재정지출이 지속되어 정부부채가 상당 기간 빠르게 증가하는 것으로 나타났는데, 이는 재정위기를 겪은 이후에도 재정관리가 중요함을 시사한다. 또한 주요 거시경제 변수의 변화를 살펴보면, 재정위기 기간에는 경제성장률이 급격히 낮아지는 것으로 분석되었다. 반면 이 기간에 실질이자율은 증가하거나 큰 변화가 없는 것으로 나타났다. 이에 이자율-성장률 격차가 크게 증가하며, 이 또한 상당 기간 증가한 상태로 머물러 있는 것으로 분석되었다. 즉 재정위기는 정부부채의 한도와 재정여력을 크게 축소시키는 것으로 해석할

수 있다.

재정위험 진단 모형을 비교·분석한 결과는 다음과 같이 요약할 수 있다. 첫째로 랜덤포레스트 모형이 다른 모형들에 비해 더 나은 예측력을 보여주었다. 이는 결정나무(decision tree)에 기반한 랜덤포레스트 모형이 변수들의 비선형적이고 복잡한 관계를 모형화하는 데 장점이 있으므로, 이러한 특성이 상대적으로 부족한 로짓이나 라쏘 모형에 비해 좋은 성능을 낸 것으로 판단된다. 만일 로짓이나 라쏘 모형에 변수들의 교차항을 추가하거나 비선형적인 관계를 추가할 수 있다면 모형의 성능을 향상할 수도 있을 것이다. 다만 모든 변수들의 교차항과 비선형적인 관계를 넣는 것은 현실적으로 어려우며 연구자의 자의적인 판단이 일부 필요할 수 있다.

둘째로 로짓 모형과 라쏘 모형의 예측 성능이 거의 유사한 것으로 나타났다. 대체로 훈련데이터에서는 로짓 모형이, 평가데이터에서는 라쏘 모형이 소폭 높은 예측력을 보였으나 그 차이가 크지는 않았다. 이는 라쏘 모형이 로짓 모형의 예측 성능을 추가적으로 개선하지 못했다는 것을 의미한다. 이 결과는 두 가지로 해석될 수 있다. 첫째, 로짓 모형이 이미 최적의 예측 성능을 보여주고 있어 개선의 여지가 거의 없었을 가능성이 있다. 둘째, 라쏘 모형의 변수 선택 및 정규화(regularization) 기능이 본 데이터에서 유효하게 작동하지 않았을 가능성이 있다. 본 연구의 분석 결과만으로는 이 두 가지 중 어느 경우인지는 확실하지 않지만, 로짓 모형의 예측 성능이 전반적으로 낮지 않다는 점을 고려할 때 첫 번째 가능성이 더 큰 것으로 보인다.

셋째로 신호 모형의 예측 성능이 다른 모형에 비해서 낮은 것으로 확인되었다. 다만 항상 그런 것은 아니고 일부 상황에서는 신호 모형이 로짓이나 라쏘 모형보다 더 나은 예측력을 보이는 경우도 있었다. 신호 모형은 재정위험 진단에 자주 사용되는 모형으로서 각 변수들마다 정해진 임계치가 있어서 이 임계치를 넘으면 신호를 보내는 것으로 간주하고 신호들의 합을 통해서 위기를 예측하는 것이다. 이 모형은 개별 변수들이 위험한 범위에 있는지, 아니면 안전한 범위에 있는지에 대한 명시적인 구분을 도와주기 때문에 개별 변수들을 관리하는 데 있어 시사점을 제공하기에는 큰 도움이 될 수 있다. 하지만 이 모형은 변수들을 모두 독립적인 것으로 판단하고 서로

의 상호작용은 고려하지 않아 예측력이 높지 않기 때문에 실효성 있는 모형으로 활용하기는 어려울 것으로 판단된다.

마지막으로 각 모형들이 재정위기 예측에 유용하다고 판단하는 변수들은 모형과 예측 시계에 따라 다른 것으로 나타났다. 예측력이 높은 랜덤포레스트 모형을 기준으로 살펴보면 1년 또는 2년간 정부부채 비율이 급증할 사건을 예측할 때는 재정수지와 정부부채 비율과 같은 재정변수가 예측에 중요한 영향을 미치는 것으로 나타났다. 하지만 3년간 정부부채 비율이 급증할 사건을 예측할 때는 재정변수도 중요하지만 경상수지, 이자율, 장단기 금리차 등 경제환경 변수의 역할이 강화되는 것으로 나타났다. 이는 비교적 짧은 시계에는 정부부채 비율을 직접적으로 결정하는 변수가 예측에 중요한 역할을 하지만, 비교적 긴 시계에는 재정변수뿐만 아니라 재정변수에 영향을 미칠 수 있는 더욱 근본적인 원인 변수의 역할이 중요해짐을 시사하는 것으로 해석할 수 있다.

본 연구는 민감도 분석도 수행하였다. 기본 분석에서 재정위기를 GDP 대비 정부부채 비율이 연평균 3%p 이상 증가한 상황으로 정의하였는데, 이는 자의적이라는 한계가 있다. 이에 정부부채 비율의 증가 속도에 대해 연평균 2%p와 4%p 이상과 같은 다른 기준을 적용해 보았는데, 모형의 예측력 간의 정성적인 특성은 기본 분석과 유사하였다. 정부부채 비율의 증가 속도를 통제하기 위한 목표는 재정이나 경제 환경에 따라 달라질 수 있다. 따라서 이를 특정한 하나의 수치로 정하여 진단하기보다는 재정관리 목표에 맞추어 분석하는 것이 적절할 것으로 판단된다.

한편 기본 분석에서는 선진국을 대상으로만 분석하였는데, 17개 신흥국을 포함한 민감도 분석도 수행하였다. 그 결과 재정위기를 예측하는 데 있어 신흥국의 포함 여부가 예측력을 뚜렷하게 높이거나 낮추지는 않았으며, 예측 시계에 따라 어떤 경우는 선진국만을 대상으로 분석할 때, 그리고 다른 경우는 신흥국을 포함하여 분석할 때 예측력이 소폭 높은 것으로 나타났다. 즉 실제로 재정위기를 예측할 때 신흥국을 포함한 모형과 그렇지 않은 모형을 모두 활용하여 비교해 보는 것이 적절할 것으로 판단된다.

본 연구에서 검토한 모형들은 재정위험 진단에 활용될 수 있지만, 재정위

기의 발생 원인을 직접적으로 설명하지는 않는다. 물론 일차적으로 모형에 활용된 변수들 중 재정위험 진단 결과에 중요한 영향을 미친 변수들을 식별하여 그 원인을 기늩할 수 있다. 다만 진단 결과 재정위기가 발생할 확률이 높다면 그 원인을 파악하고 대처하기 위해서는 당시의 경제 상황 등에 대한 별도의 분석이 필요하다.

## 목 차

I. 서론 .....	17
II. 재정위기의 정의 및 식별 .....	25
1. 분석 개요 .....	25
2. 재정위기의 정의 .....	25
가. 선행연구의 재정위기 정의 .....	25
나. 본 연구의 재정위기 정의 .....	30
3. 재정위기 전후의 특징 .....	40
가. 분석 모형 및 자료 .....	40
나. 기본 분석 결과 .....	42
다. 민감도 분석 결과 .....	50
III. 재정위험 진단 방법 .....	57
1. 연구 설계 .....	57
2. 예측 모형 설명 및 예측 결과 .....	63
가. 신호(signaling) 모형 .....	63
나. 로짓(logit) 모형 .....	68
다. 라쏘(LASSO) 모형 .....	72
라. 랜덤포레스트(random forest) 모형 .....	77
3. 분석 결과 비교 .....	81
4. 민감도 분석 .....	103
가. 정부부채 비율 급증 기준 하향 .....	103
나. 정부부채 비율 급증 기준 상향 .....	110
다. 신흥국 포함 .....	116

---



## CONTENTS

---

IV. 결론 및 정책시사점 .....	121
참고문헌 .....	126
부록: 분석에 활용된 독립변수 .....	130

---

---

## 표목차

〈표 II-1〉 Medas et al.(2018)의 재정위기 정의 .....	26
〈표 II-2〉 우리나라에 대한 국제통화기금(IMF) 차관협약 .....	27
〈표 II-3〉 주요 재정변수의 위기 전후 변동 .....	45
〈표 II-4〉 주요 거시경제 변수의 위기 전후 변동 .....	48
〈표 III-1〉 재정위험 예측에 관한 혼돈행렬 .....	59
〈표 III-2〉 모형별 예측력 비교(연평균 3%p 증가 기준) .....	86
〈표 III-3〉 1년간 정부부채 비율 급증에 영향을 미치는 주요 변수 .....	88
〈표 III-4〉 2년간 정부부채 비율 급증에 영향을 미치는 주요 변수 .....	95
〈표 III-5〉 3년간 정부부채 비율 급증에 영향을 미치는 주요 변수 .....	99
〈표 III-6〉 모형별 예측 성능 비교(연평균 2%p 증가 기준) .....	109
〈표 III-7〉 모형별 예측 성능 비교(연평균 4%p 증가 기준) .....	115
〈표 III-8〉 모형별 예측 성능 비교(신흥국 포함 모형) .....	120

---

## 그림목차

[그림 Ⅰ-1] GDP 대비 국가채무와 재정수지 비율 실적 및 정부 계획 .....	18
[그림 Ⅰ-2] 국회예산정책처의 국가채무 기준선 전망 .....	19
[그림 Ⅱ-1] 5년 만기 외평채 CDS 프리미엄 .....	29
[그림 Ⅱ-2] 우리나라의 국가 신용등급 .....	29
[그림 Ⅱ-3] 한국의 GDP 대비 정부부채 비율 .....	34
[그림 Ⅱ-4] 한국의 재정위기 식별 .....	34
[그림 Ⅱ-5] 선진국의 GDP 대비 정부부채 비율 .....	37
[그림 Ⅱ-6] 선진국의 재정위기 식별 .....	38
[그림 Ⅱ-7] 선진국과 신흥국의 재정위기 식별 .....	39
[그림 Ⅱ-8] 재정수지와 이자율-성장을 격차 변화에 따른 재정여력 변화 .....	41
[그림 Ⅱ-9] 주요 재정변수의 위기 전후 변동 .....	43
[그림 Ⅱ-10] 주요 거시경제 변수의 위기 전후 변동 .....	47
[그림 Ⅱ-11] 주요 재정변수의 위기 전후 변동(민감도 분석: 연평균 2%p) .....	51
[그림 Ⅱ-12] 주요 거시경제 변수의 위기 전후 변동(민감도 분석: 연평균 2%p) .....	52
[그림 Ⅱ-13] 주요 재정변수의 위기 전후 변동(민감도 분석: 연평균 4%p) .....	53
[그림 Ⅱ-14] 주요 거시경제 변수의 위기 전후 변동(민감도 분석: 연평균 4%p) .....	54
[그림 Ⅱ-15] 주요 재정변수의 위기 전후 변동(민감도 분석: 신흥국 포함) .....	55
[그림 Ⅱ-16] 주요 거시경제 변수의 위기 전후 변동(민감도 분석: 신흥국 포함) .....	56
[그림 Ⅲ-1] ROC 곡선 예시 .....	62
[그림 Ⅲ-2] 신호 모형으로 예측한 한국의 정부부채 비율 급증 확률 .....	66
[그림 Ⅲ-3] 신호 모형으로 예측한 선진국의 정부부채 비율 급증 확률 .....	67
[그림 Ⅲ-4] 로짓 모형으로 예측한 한국의 정부부채 비율 급증 확률 .....	70
[그림 Ⅲ-5] 로짓 모형으로 예측한 선진국의 정부부채 비율 급증 확률 .....	71

---

[그림 III-6] 라쏘 모형으로 예측한 한국의 정부부채 비율 급증 확률 .....	75
[그림 III-7] 라쏘 모형으로 예측한 선진국의 정부부채 비율 급증 확률 .....	76
[그림 III-8] 랜덤포레스트 모형으로 예측한 한국의 정부부채 비율 급증 확률 .....	79
[그림 III-9] 랜덤포레스트 모형으로 예측한 선진국의 정부부채 비율 급증 확률 .....	80
[그림 III-10] 모형별 ROC 곡선 비교(연평균 3%p 증가 기준) .....	83
[그림 III-11] 1년간 정부부채 비율 급증에 영향을 미치는 변수들의 부분의존도 ..	93
[그림 III-12] 2년간 정부부채 비율 급증에 영향을 미치는 변수들의 부분의존도 ..	98
[그림 III-13] 3년간 정부부채 비율 급증에 영향을 미치는 변수들의 부분의존도 ..	102
[그림 III-14] 모형별 한국의 정부부채 비율 급증 확률(연평균 2%p 증가 기준) ..	106
[그림 III-15] 모형별 선진국의 정부부채 비율 급증 확률(연평균 2%p 증가 기준) ..	107
[그림 III-16] 모형별 ROC 곡선 비교(연평균 2%p 증가 기준) .....	108
[그림 III-17] 모형별 한국의 정부부채 비율 급증 확률(연평균 4%p 증가 기준) ..	112
[그림 III-18] 모형별 선진국의 정부부채 비율 급증 확률(연평균 4%p 증가 기준) ..	113
[그림 III-19] 모형별 ROC 곡선 비교(연평균 4%p 증가 기준) .....	114
[그림 III-20] 모형별 한국의 정부부채 비율 급증 확률(신흥국 포함 모형) .....	117
[그림 III-21] 모형별 선진국의 정부부채 비율 급증 확률(신흥국 포함 모형) .....	118
[그림 III-22] 모형별 ROC 곡선 비교(신흥국 포함 모형) .....	119

---



---

# I. 서론

---

우리나라 정부가 코로나바이러스감염증-19(이하 코로나19) 확산 등의 위기에 대응하기 위해 적극적인 재정정책을 펼치면서 최근 재정의 지속가능성이 급격히 악화되었다. [그림 I-1-(가)]에 나타난 바와 같이 2019년에 35.4%이었던 우리나라의 GDP 대비 국가채무 비율은 코로나19가 처음 발생한 2020년에 41.1%로 5.7%p 상승하면서 40%를 넘어섰다. 그 이후에도 3년간 GDP 대비 국가채무 비율은 5.8%p 상승하여 2023년에는 46.9%에 이르렀다. 즉 지난 4년간 GDP 대비 국가채무 비율은 11.5%p 상승한 것이다. 그리고 대한민국정부(2024)가 발표한 『2024~2028년 국가재정운용계획』에 따르면 GDP 대비 국가채무 비율은 앞으로도 꾸준히 증가하여 2028년에는 50%를 넘어설 전망이다.<sup>1)</sup>

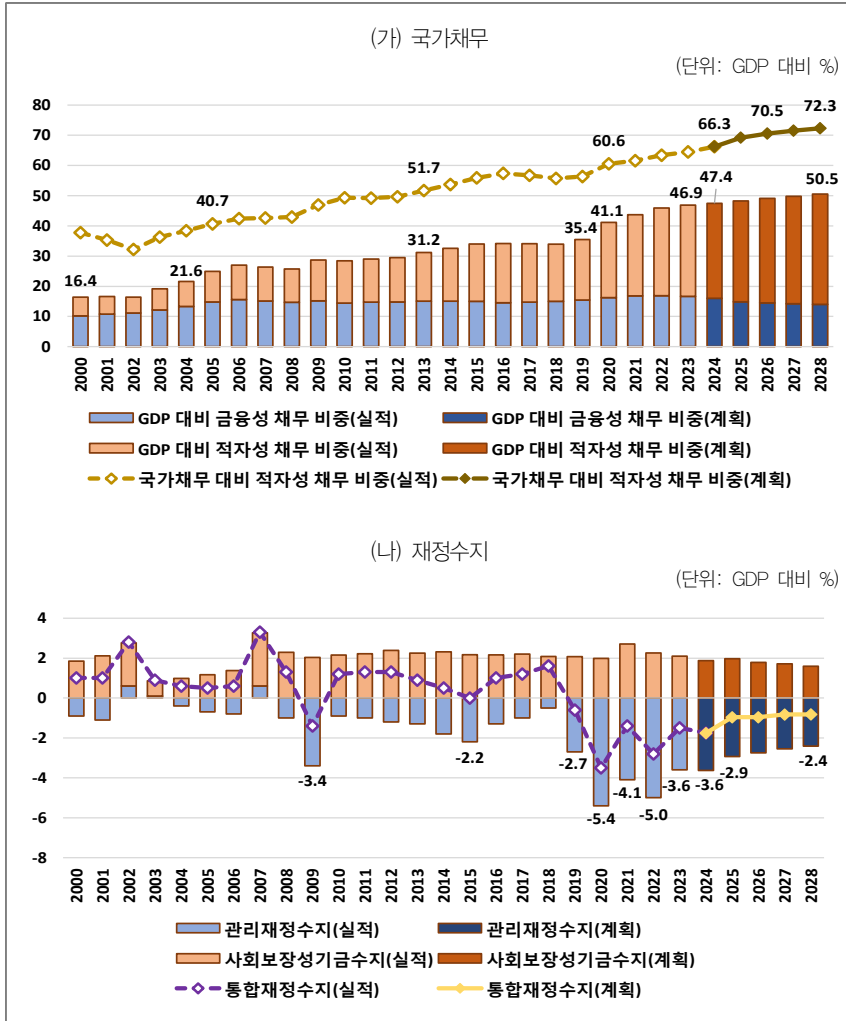
관리재정수지 적자 규모도 2020년 이후 커졌다. [그림 I-1-(나)]에 나타난 바와 같이 관리재정수지 적자는 2020년 GDP의 5.4%를 기록한 이후 2023년까지 매년 GDP의 3%를 넘어섰다. 현 정부가 건전재정을 강조하고 있음에도 불구하고 2024년 예산의 관리재정수지 적자 규모 또한 GDP의 3.6%로 예상되어 우리나라의 재정수입과 재정지출의 격차에 구조적인 문제가 있음을 드러냈다. GDP 대비 3.6%의 관리재정수지 적자 규모는 국제 금융위기 시기인 2009년보다도 높은 수준으로, 이는 이례적으로 큰 규모라 할 수 있다. 『2024~2028년 국가재정운용계획』에 따르면 정부는 향후 재정수지를 점차 개선해 나갈 계획이다. 그러나 2028년에도 관리재정수지 적자 규모는 GDP의 2.4%를 기록할 것으로 계획되어 강력한 지출구조조정이나 세입확충 등 재정수지를 개선하기 위한 특별한 조치가 없다면 관리재정수지 적자가 2019년 이전의 평상시 수준으로 돌아가는 데에는 오랜 기간이 필요할 것으로 보인다.<sup>2)</sup>

---

1) 대한민국정부, 『2024~2028년 국가재정운용계획』, 2024, p. 39.

2) 대한민국정부, 『2024~2028년 국가재정운용계획』, 2024, p. 39.

[그림 1-1] GDP 대비 국가채무와 재정수지 비율 실적 및 정부 계획



주: 사회보장성기금수지 = 통합재정수지 - 관리재정수지

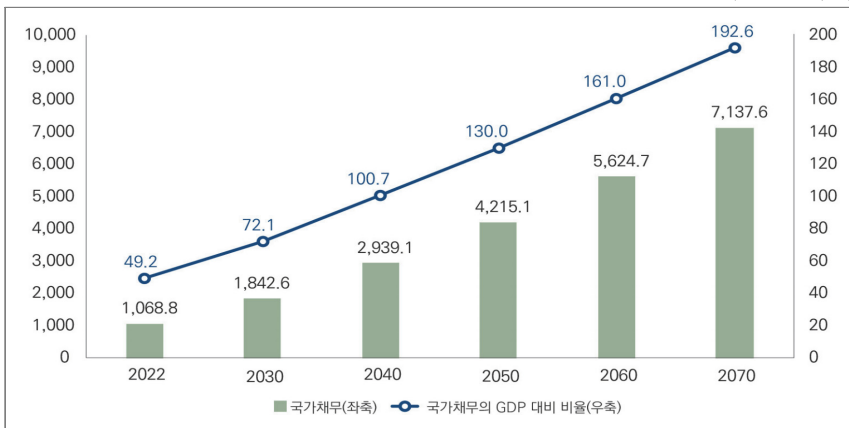
자료: 다음의 자료를 이용하여 저자 작성

1. 대한민국정부, 『2024~2028년 국가재정운용계획』, 2024. 9. p. 39.
2. 대한민국정부, 『2024~2028년 국가재정운용계획 첨부서류』, 2024. 9. p. 176.
3. 지표누리, e-나라지표, 국정모니터링시스템, 「국가채무추이」, [https://www.index.go.kr/unity/potal/main/EachDllPageDetail.do?idx\\_cd=1106](https://www.index.go.kr/unity/potal/main/EachDllPageDetail.do?idx_cd=1106), 검색일자: 2024. 11. 11.
4. 지표누리, e-나라지표, 국정모니터링시스템, 「통합재정수지」, [https://www.index.go.kr/unity/potal/main/EachDllPageDetail.do?idx\\_cd=1104](https://www.index.go.kr/unity/potal/main/EachDllPageDetail.do?idx_cd=1104), 검색일자: 2024. 10. 23.
5. 한국은행, 경제통계시스템(ECOS), 「국민계정」, “2.1.2.2.3. 국내총생산에 대한 지출(원계열, 명목, 분기 및 연간),” <https://ecos.bok.or.kr/#/SearchStat>, 검색일자: 2024. 11. 11.

우리나라 재정의 지속가능성 문제는 장기적으로도 우려를 낳고 있는 상황이다. 특히 저출생·고령화로 인한 인구구조 변화는 재정의 지속가능성을 심각하게 위협하는 요인이다. 국회예산정책처(2022)는 현행 제도에 변화가 없다면 인구구조 변화로 인하여 우리나라의 GDP 대비 국가채무 비율이 안정화되지 않고 지속적으로 상승하는 것으로 전망하였다. [그림 1-2]에 나타난 바와 같이 현행 제도하에서 우리나라의 국가채무 규모는 2040년에 GDP 규모를 넘어서고, 2070년에는 GDP의 192.6%에 이를 전망이다. 한편 이러한 전망에는 고령화가 진행될수록 재정에 심각한 문제가 발생할 것으로 예상되는 국민연금, 건강보험 등 사회보장제도의 재정 악화가 반영되지 않아, 실제로는 이 전망보다 재정의 지속가능성에 대한 위협이 더욱 심각한 상황이다. 인구구조의 변화뿐만 아니라 탄소중립 달성 등 기후위기 대응이나 통일에 따른 비용 등도 재정의 지속가능성을 위협하는 잠재적인 요인이다.

[그림 1-2] 국회예산정책처의 국가채무 기준선 전망

(단위: 조원, %)



주: 전망액은 2022년 불변가액 기준이며, 인구는 통계청 중위 기준

자료: 국회예산정책처, 『2022~2070년 NABO 장기 재정전망』, 2022. 8. p. 65, [그림 20].

코로나19 위기 대응으로 인한 급격한 재정건전성 악화와 저출생·고령화로 인한 장기적인 재정위험은 우리나라뿐만 아니라 정도의 차이는 있지만 다른 많은 국가도 공통적으로 직면한 위협 요인이다. 이에 최근 재정의 지속



가능성을 관리하기 위한 노력에 대해 많은 논의가 진행되고 있으며, 이러한 노력 중 하나가 재정위험을 진단하는 것이다. 재정의 지속가능성을 확보하고 유지하기 위해서는 투명하고 신뢰성이 있는 재정관리체계가 필요하다. 특히 이러한 재정관리체계에서 국가채무나 재정수지 등 재정총량에 대한 목표치나 한계치를 설정할 필요가 있는데, 이때 재정위험에 대한 분석이 필수적이다.

우리나라의 재정위험 분석은 주로 장기 시계에서 이루어지고 있다. 정부는 「국가재정법」에 의해 2015년부터 5년마다 40회계연도 이상의 기간에 대한 재정전망 결과를 발표하고 있다. 국회예산정책처 또한 주기적으로 [그림 I-2]와 같은 장기재정전망 결과를 다양한 민감도 분석과 함께 제시하고 있다. 이뿐만 아니라 김학수 외(2021)와 같이 개별 연구자들도 장기 시계에서의 재정위험을 진단하고 재정건전화 방안을 제시한다. 이러한 장기 시계에서의 재정위험 진단은 대체로 저출생·고령화로 인한 인구구조 변화에 따른 재정위험 진단에 초점이 맞춰져 있다. 우리나라의 경우 매우 빠르게 변화하고 있는 인구구조는 재정의 지속가능성을 크게 위협하는 구조적인 요인이다.

한편 중단기 시계에서의 재정위험 진단은 그간 활발한 편은 아니었다. 중기 시계에서 정부는 매년 국가재정운용계획을 통해 당해 연도를 포함한 향후 5년에 대한 총수입, 총지출, 재정수지, 국가채무 등 재정총량에 대한 정부의 관리계획을 수립하여 발표한다. 그리고 국회예산정책처는 정부가 발표한 예산안과 국가재정운용계획을 바탕으로 정부보다 5년의 시계를 연장한 향후 10년간의 재정을 전망하고 정부의 예산안과 국가재정운용계획을 평가한다. 그러나 정부의 계획과 국회예산정책처의 평가에는 재정총량에 대한 민감도 분석이나 위험 상황에 따른 스트레스 테스트 등의 재정위험 분석이 포함되지 않는다.

그러나 최근 우리나라의 국가채무 비율<sup>3)</sup>이 높아지고 있고 앞으로도 재정의 지속가능성이 계속 악화된다면 중단기적인 재정위험 진단의 필요성이 점

---

3) 본 연구에서 특별한 언급이 없는 한 국가채무 비율과 정부부채 비율은 모두 GDP 대비 비율을 의미한다.

차 커질 것으로 판단된다. 한중석 외(2024)에 따르면 정부는 과거 홍승현·윤성주(2013)를 바탕으로 한시적으로(2014~2016년) 재정수지, 부채, 공기업, 지방정부 등에 대한 다양한 재정지표를 설정하여 재정위험을 모니터링한 적이 있다. 그러나 이러한 재정위험 모니터링의 실효성이 크지 않아 현재는 운영되지 않는 것으로 보인다. 그러나 한중석 외(2024)는 정부가 재정건전성을 체계적으로 관리하기 위해서는 기존의 국가재정운용계획 외에 재정위험 진단을 포함한 재정의 지속가능성을 평가하는 보고서를 주기적으로 발간할 것을 제안하였다.

최근 김명규·김민경(2023)과 한중석 외(2024)와 같이 우리나라 재정의 지속가능성을 평가하거나 이를 진단하는 방법론에 대한 연구도 있다. 두 연구 모두 국제기구가 주기적으로 발표하는 재정위험 진단 보고서의 기본적인 방법론을 우리나라에 적용해 본 것이다. 김명규·김민경(2023)은 유럽연합 집행위원회(European Commission, EC)의 『재정 지속가능성 보고서(Fiscal Sustainability Report)』의 방법론을 우리나라의 상황에 맞춰 단기, 중기, 장기 재정위험 지표를 설정하고 분석하였다.<sup>4)</sup> 한중석 외(2024)는 국제통화기금(International Monetary Fund, IMF)의 『국가부채 위험 및 지속가능성 체계(Sovereign Risk and Debt Sustainability Framework, SRDSF)』 중 중기 분석(Medium Term Analysis) 방법론을 우리나라에 시험적으로 적용하였다.<sup>5)</sup>

그러나 아직 기계학습(machine learning) 방법론을 활용하여 재정위험을 진단하는 방법론을 연구한 국내 연구는 존재하지 않는다. 최근 기계학습(machine learning)을 활용한 분석 기법이 발전하여 국제기구를 중심으로 재정위험 진단에 이를 적용하려는 시도들을 하고 있다. 지금까지는 주로 단기 재정위험 진단을 위해 로짓(logit)·프로빗(probit) 모형과 같은 전통적인 계량경제학적 방법이나 신호접근법(signaling approach)과 같은 특정 지표들

---

4) 유럽연합 집행위원회(EC)의 재정위험 진단 방법론은 EC(2022, *Fiscal Sustainability Report 2021*, Volume 1)를 참고하기 바란다.

5) 국제통화기금(IMF)의 재정위험 진단 방법론은 IMF(2022, *Staff Guidance Note on the Sovereign Risk and Debt Sustainability Framework for Market Access Countries*)를 참고하기 바란다.

의 임계점을 이용한 방법에 의존하였다. Kaminsky et al.(1998)은 외환위기 예측을 위한 방법으로 신호접근법을 제안하였다. 국제통화기금(IMF)은 『국가부채 위험 및 지속가능성 체계(SRDSF)』에서 단기 재정위험을 다변수 로짓 모형으로 진단한다. 유럽연합 집행위원회(EC)는 『재정 지속가능성 보고서(Fiscal Sustainability Report)』에서 단기 재정위험에 대한 지표로 S0를 활용하는데, 이는 신호접근법에 해당한다. 그러나 최근에는 Jarmulaska(ECB<sup>6)</sup> Working Paper, 2020), IMF(IMF Technical Notes and Manuals, 2021), Hellwig(IMF Working Paper, 2021), Valencia et al.(IDB<sup>7)</sup> Working Paper, 2022) 등과 같이 국제기구를 중심으로 랜덤포레스트(random forest) 등 기계학습(machine learning) 기법을 활용한 재정위험 진단이 활발히 시도되고 있다.

본 연구는 기계학습(machine learning)을 활용한 재정위험 진단 방법을 개발하는 것을 목적으로 한다. 재정위험은 다양한 시계에서 다양한 방법으로 진단할 수 있는데, 본 연구는 기존의 방법론이 아닌 새로운 방법론을 통해 재정위험을 진단하여 분석의 폭을 넓히는 데 의의가 있다. 다만 기계학습의 특성상 중장기적으로 영향을 미치는 구조적인 위험을 진단하기보다는 선행 연구들처럼 다양한 경제 및 재정변수를 활용하여 단기 재정위험을 진단하는 것이 적합하다. 기계학습(machine learning)을 통한 재정위험 진단은 앞에서 설명한 장기재정전망이나 김명규·김민경(2023), 한종석 외(2024)의 방법론을 통한 우리나라의 재정위험 진단을 보완할 수 있을 것으로 판단된다.

단기 재정위험(fiscal risk)을 진단하기 위해서는 먼저 진단하고자 하는 재정위기(fiscal crisis)를 정의할 필요가 있다. 단기 재정위험이란 가까운 미래에 재정위기가 발생할 가능성이기 때문이다. 앞에서 언급한 기계학습을 이용하여 재정위험을 진단하는 선행연구들은 주로 Medas et al.(2018)의 정의를 이용하여 재정위기를 식별하였다. Medas et al.(2018)은 재정위기를 채무 불이행과 같은 상당히 극단적인 상황으로 정의하였다.

---

6) 유럽중앙은행(European Central Bank)

7) 미주개발은행(Inter-American Development Bank)

하지만 단기적으로 채무불이행의 가능성이 높지 않은 우리나라의 경우 재정위기를 Medas et al.(2018)과 같이 정의하는 것은 실효성이 크지 않을 것으로 판단된다. 우리나라의 재정은 당장의 채무불이행보다는 장기적인 지속가능성에 문제가 있어 재정의 지속가능성을 관리하기 위한 재정위험 진단이 더 유용할 수 있기 때문이다. 또한 단기간 내 채무불이행이 발생할 가능성이 높다면 이러한 위험은 금융시장 등 다양한 경로를 통해 빠르게 드러날 가능성이 높다. 그리고 이러한 위험을 진단하는 방법론을 개발하더라도 실제 위험에 대한 경고등이 켜졌을 때에는 이미 해당 위험이 실현되지 않도록 대처하기에 매우 어려운 상황일 것이다. 따라서 현재 우리나라의 재정 상황에서는 심각한 재정위기를 겪을 가능성을 진단하기보다는 이러한 위기가 발생하지 않도록 예방하기 위한 도구, 특히 국가채무 비율을 안정적으로 관리하기 위한 수단으로 재정위험을 진단하는 것이 더 유용할 것으로 판단된다.

본 연구는 단기 재정위험으로 국가채무 또는 정부부채 비율의 증가 속도에 초점을 두고 분석한다. 중장기적으로 재정의 지속가능성을 식별하는 데 중심이 되는 지표는 GDP 대비 국가채무 비율이다. 하지만 높은 수준의 국가채무는 하루아침에 쌓이는 것이 아니라 여러 해에 걸쳐 누적되어 나타난 결과이다. 따라서 단기적으로 국가채무 비율이 누적되어 높아지지 않도록 관리하는 것이 재정위기를 사전에 방지하는 가장 좋은 방법이다. 특히 단기에는 국가채무 비율의 증가 속도가 중요할 수 있다. 최근 코로나19 위기 과정에 나타난 바와 같이 위기 때 국가채무 비율이 급격히 상승하면서 재정의 지속가능성이 악화될 수 있다. 그리고 이러한 빠른 증가 속도는 위기가 지나간 이후에도 지속될 수 있다. 위기 때는 재정을 통해 유연하게 대처할 필요가 있지만 위기 이후에 이를 관리하는 모습이 필요하다. 그래야 장기적으로 재정이 지속될 수 있도록 관리할 수 있다.

본 연구는 크게 두 가지의 분석을 수행한다. 먼저 우리나라에 적합한 재정위험을 정의하고 이를 정량적으로 측정하는 방법에 대해 논의한다. 재정위험을 진단하기 위해서는 우선 종속변수인 과거에 이미 발생한 재정위험을 식별하는 것이 중요하다. 재정위험의 측정 방법을 선택할 때 자료의 이용

가능성과 적용의 용이성 등을 고려하여 자료에서 과거에 드문 빈도로 재정의 지속가능성을 악화시키는 방향으로 관찰되었던 사건들을 기준으로 재정위기를 정의하고자 한다. 그리고 이렇게 정의된 재정위기에 대해 정형적 사실을 정리하고자 한다. 다음으로 단기 재정위험 진단에 현재 널리 활용되고 있거나 선행연구에서 최근에 예측력이 높은 것으로 분류된 몇 가지 방법론을 비교하여 앞에서 정의한 재정위기를 예측하는 방법을 개발하고자 한다. 특히, 전통적인 방법(예: 신호접근법, 로짓 모형)과 기계학습 분석 기법(예: 라쏘, 랜덤포레스트)에 의한 재정위험의 진단 결과를 비교한다. 따라서 본 연구는 기계학습(machine learning)을 활용하여 재정위험을 진단한다는 측면에서 다른 국내 선행연구와 차이점이 있고, 진단하는 재정위험의 정의 측면에서도 다른 국내외 선행연구와 차별점이 있다.

본 연구는 다음과 같이 구성된다. 제Ⅱ장에서는 진단하고자 하는 재정위기를 정의하고 이를 식별하여 특성을 파악한다. 제Ⅲ장에서는 제Ⅱ장에서 정의한 재정위기를 예측하는 여러 방법론을 개발하여 비교한다. 그리고 제Ⅳ장은 결론 및 정책시사점을 제시한다.

---

## Ⅱ. 재정위기의 정의 및 식별

---

### 1. 분석 개요

본 연구는 우리나라의 재정위험을 진단하기 위한 방법론을 연구하는 것이 목적이다. 재정위험(fiscal risk)은 미래에 재정위기(fiscal crisis)가 발생할 가능성을 의미한다. 따라서 재정위험을 진단하기 위해서는 먼저 재정위기를 정의해야 한다. 이에 본 장에서는 재정위기를 정의하고자 한다. 그리고 과거 자료를 이용하여 본 연구가 정의한 재정위기 전후에 나타난 주요 재정 및 거시경제 변수의 움직임을 분석하여 정형적 사실을 살펴보고자 한다.

### 2. 재정위기의 정의

#### 가. 선행연구의 재정위기 정의

Hellwig(2021), Valencia et al.(2022) 등 재정위험을 진단하는 최근의 연구들은 Baldacci et al.(2011), Medas et al.(2018)과 유사한 방식으로 재정위기를 정의한다. Medas et al.(2018)은 Baldacci et al.(2011)이 정의한 재정위기를 조금 더 정교하게 다듬은 것인데, 다음의 네 가지 사건 중 하나 이상의 사건이 발생하면 해당 시기를 재정위기 기간으로 정의한다. 첫 번째는 국채와 관련된 신용 사건이 발생한 경우로 완전한 채무불이행이나 채무 구조조정이 여기에 해당된다. 두 번째는 국제통화기금(IMF)으로부터 대규모의 구제금융을 지원받는 경우이다. 세 번째는 초인플레이션(hyperinflation)이 발생하거나 국가채무의 연체가 급증하는 등 공공 채무불이행이 암묵적으로 발생한 경우이다. 마지막으로 네 번째는 국가에 대한 시장의 신뢰가 상실되어 시장에 접근할 수 없거나 국채 또는 CDS 스프레드가 급격히 상승하는 등의 사건이 발생한 경우이다.

## 〈표 II-1〉 Medas et al.(2018)의 재정위기 정의

Criterion		Thresholds			
		AMs	EMs	LIDCs	SDSs
(1) Credit event	Any operation that makes creditors incur material economic losses on the sovereign debt they hold (e.g. default, restructuring, or rescheduling) (i) of substantial size (in percent of GDP p.a.) (ii) if (i) holds <i>and</i> the defaulted nominal amount grows by a substantial amount (in percent p.a.)	> 0.2			
		≥ 10			
(2) Exceptionally large official financing	High-access IMF financial arrangement with fiscal adjustment objective in place (in percent of quota)	≥ 100			
(3) Implicit domestic public default	(a) High inflation rate (in percent of growth of annual average CPI p.a.)	≥ 35	≥ 100	≥ 35	
	(b) Steep increase in domestic arrears (in first difference of the ratio of 'other account payables (OAP)' to GDP in percentage points)	≥ 1			
(4) Loss of market confidence	(a) Loss of market access	when market access is lost (after maintaining market access for a 1/4 of the sample time and 2 consecutive years before the loss year)			
	(b) High price of market access (in basis points, sovereign spreads or CDS spreads)	≥ 1,000 bps			

주: Medas et al.(2017)은 Medas et al.(2018)의 Working Paper

자료: Medas et al.(2017), "Fiscal Crises," IMP Working Paper, p. 12, Table 2.1.

Medas et al.(2018)은 이러한 기준으로 1970~2015년에 188개 국가에서 발생한 재정위기를 식별하였다. 이에 따르면 우리나라의 재정위기 기간은 1980~1985년(6년간)과 1997~2000년(4년간)으로 나타났다. 두 차례 모두 국제통화기금(IMF)으로부터 대규모의 금융지원을 받은 시기로, Medas et al.(2018)의 두 번째 기준에 의해 재정위기 기간으로 식별되었다. 우리나라는 1980~1985년에 국제통화기금(IMF)으로부터 약 21억SDR의 대기성차관(stand-by arrangement)을 승인받아 약 16억SDR을 인출하였으며, 외환위기 기간인 1997년에는 155억SDR의 대기성차관을 승인받아 약 144억SDR을 인출하였다.<sup>8)</sup> 한편 1997~1998년은 시장의 신뢰도가 상실된 시기로, Medas et al.(2018)의 네 번째 기준에 의해서도 재정위기 기간으로 동시에 식별되었다.

Medas et al.(2018)이 식별한 우리나라의 재정위기 기간에는 우리나라의 경제가 역성장을 하는 등 경기가 매우 좋지 않은 시기도 포함된다. 우리나

8) SDR(Special Drawing Right, 특별인출권)은 국제통화기금(IMF)의 국제준비자산으로 1997년 12월 4일 기준 1SDR은 약 1,3572미국달러(IMF, "SDR Valuation," [www.imf.org/external/np/fin/data/rms\\_sdrv.aspx](http://www.imf.org/external/np/fin/data/rms_sdrv.aspx), 검색일자: 2024. 6. 27.)의 가치로 환산되었다.

라가 1954년 이후 역성장을 기록한 해는 1980년(-1.65%), 1998년(-5.13%), 2020년(-0.71%)으로 딱 세 번 있다.<sup>9)</sup> 이 중 Medas et al.(2018)의 자료 기간인 1970~2015년에는 두 번 있는데, 모두 Medas et al.(2018)이 식별한 재정위기 기간에 포함되었다. 1980년은 정치불안과 유가파동, 국제금리 상승이 겹쳤던 시기이며, 1998년은 외환위기 시기이다. 그리고 2020년은 Medas et al.(2018)의 자료 기간에는 포함되지 않지만 코로나19 위기 시기였다.

〈표 II-2〉 우리나라에 대한 국제통화기금(IMF) 차관협약

(단위: 천SDR)

유형	협약일	만기일	승인 규모	인출 규모
대기성차관협약	1997년 12월 4일	2000년 12월 31일	15,500,000	14,412,500
보충준비금융	1997년 12월 4일	1998년 12월 17일	9,950,000	9,950,000
대기성차관협약	1985년 7월 12일	1987년 3월 10일	280,000	160,000
대기성차관협약	1983년 7월 8일	1985년 3월 31일	575,775	575,775
대기성차관협약	1981년 2월 13일	1982년 2월 12일	576,000	576,000
대기성차관협약	1980년 3월 3일	1981년 2월 12일	640,000	320,000
대기성차관협약	1977년 5월 6일	1977년 12월 31일	20,000	0
대기성차관협약	1975년 10월 22일	1976년 6월 30일	20,000	9,000
대기성차관협약	1974년 5월 17일	1974년 12월 31일	20,000	20,000
대기성차관협약	1973년 4월 2일	1973년 12월 31일	20,000	0
대기성차관협약	1972년 1월 1일	1972년 12월 31일	30,000	0
대기성차관협약	1971년 1월 1일	1971년 12월 31일	25,000	0
대기성차관협약	1970년 3월 13일	1970년 12월 31일	25,000	0
대기성차관협약	1969년 4월 15일	1969년 12월 31일	25,000	0
대기성차관협약	1968년 4월 11일	1969년 4월 10일	25,000	12,500
대기성차관협약	1967년 3월 22일	1968년 3월 21일	18,000	0
대기성차관협약	1966년 3월 22일	1967년 3월 21일	12,000	0
대기성차관협약	1965년 3월 22일	1966년 3월 21일	9,300	0
합계	-	-	17,821,075	16,085,775

자료: IMF, "Korea: History of Lending Commitments as of April 30, 2020," <https://www.imf.org/external/np/fin/tad/extarr2.aspx?memberKey1=550&date1key=2020-04-30>, 검색일자: 2024. 6. 27.

9) 경제성장률은 2015년 기준년 국민계정 기준이다.

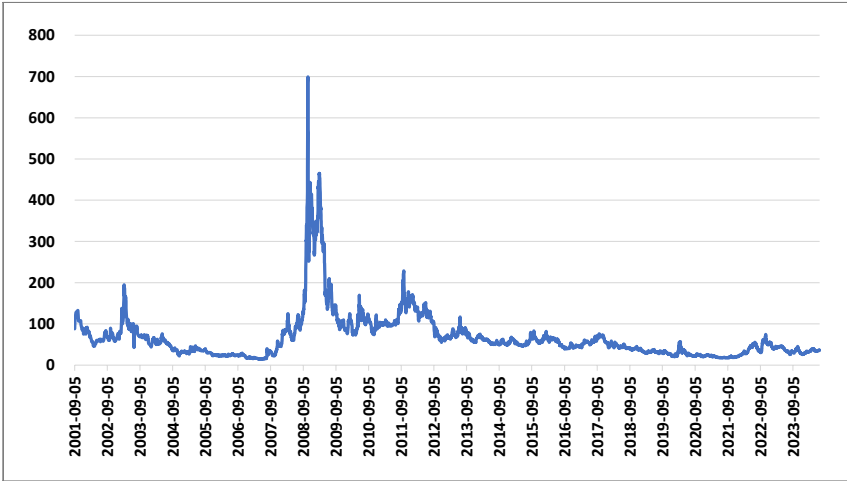


본 연구에서 재정위기를 Medas et al.(2018)과 유사하게 정의하여 식별하더라도 그 실효성이 크지 않은 것으로 판단된다. 이러한 정의는 매우 극단적인 상황만을 식별하기 때문이다. 재정위기를 Medas et al.(2018)과 같이 정의하고 이러한 재정위기가 발생할 가능성인 재정위험을 진단하는 연구들은 국제통화기금(IMF), 유럽연합 집행위원회(EC), 미주개발은행(IDB) 등 주로 국제기구를 중심으로 수행되고 있다. 국제기구의 경우 전 세계 많은 국가들의 재정위험을 진단하고 실제 채무불이행과 같은 위험을 측정하여 대응하는 것이 필요하기 때문에 이러한 재정위기의 정의가 적절할 수 있다. 그러나 우리나라의 재정위험을 진단하는 것이 목적인 본 연구에서 재정위기를 매우 극단적인 상황으로만 가정한다면 재정위험 진단의 유용성이 크지 않을 수 있다.

우리나라와 같이 현재 재정상태가 비교적 양호한 국가에서는 단기간 내에 채무불이행과 같은 심각한 재정위기가 발생할 가능성을 진단하기보다는 재정의 지속가능성이 악화될 위험을 사전에 진단하여 대처하는 것이 더 유용할 것으로 판단된다. 심각한 재정위기에 빠질 가능성이 높은 상황에서는 그 위기가 본 연구의 모형을 통해 감지되는 것보다 CDS 프리미엄이나 국제 신용평가사 등을 통해 시장에서 훨씬 더 빠르게 감지될 가능성이 높다. 실제로 국제 금융위기 기간에는 우리나라 5년 만기 외평채의 CDS 프리미엄이 급등하였으며, 외환위기 시기에도 국제 신용평가사가 평가한 우리나라의 국가 신용등급이 급락한 것을 관찰할 수 있다. 그러나 재정의 지속가능성을 악화시킬 위기는 비교적 작은 충격일 수도 있어 시장에서의 반응이 크지 않을 수 있다. 그럼에도 이러한 작은 위기들이 장기간 쌓이면 미래에는 재정의 지속가능성을 유지하기 어려워지는 심각한 상황에 놓일 수도 있다. 따라서 본 연구는 우리나라 재정의 지속가능성을 유지하고 관리하기에 적합한 재정 지표를 선정하여 재정위기를 진단함으로써 재정위험 진단의 유용성을 높이고자 한다.

[그림 II-1] 5년 만기 외평채 CDS 프리미엄

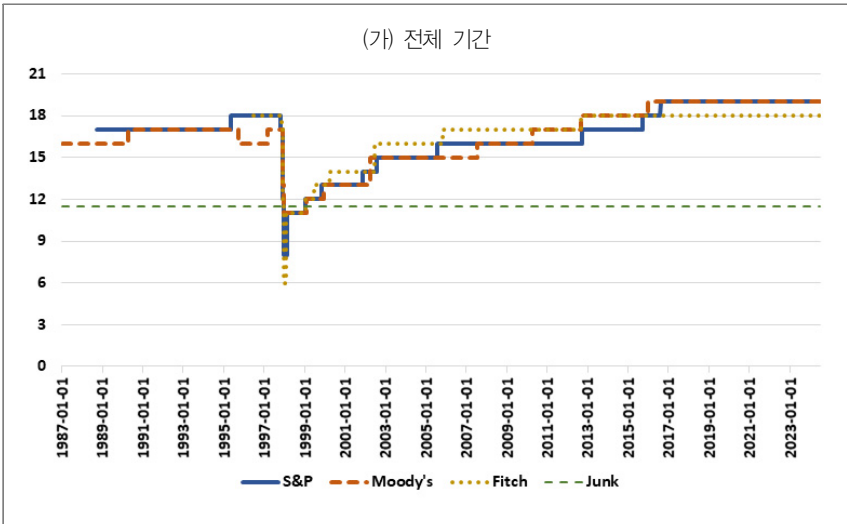
(단위: bp)



자료: 국제금융센터, 지표: 자료, 「CDS 프리미엄」, <https://www.kcof.or.kr/chart/intrList>, 검색일자: 2024. 6. 29.

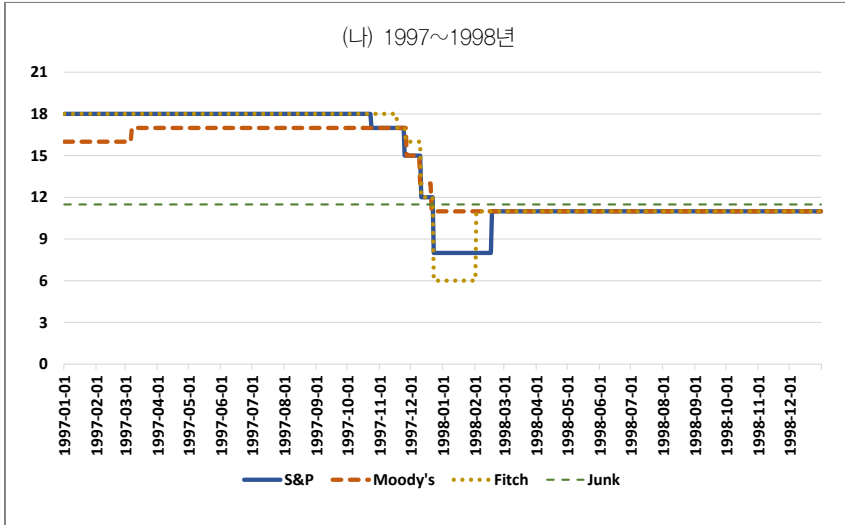
[그림 II-2] 우리나라의 국가 신용등급

(단위: 등급)



## [그림 II-2]의 계속

(단위: 등급)



주: 신용평가사별 최고등급은 21로 설정

자료: 행정안전부, 공공데이터포털, 「국가신용등급」, <https://www.data.go.kr/data/15117342/fileData.do>, 검색 일자: 2024. 6. 23.

### 나. 본 연구의 재정위기 정의

본 연구는 단기 재정위기를 정부부채 비율이 빠르게 증가하는 시기로 정의한다. 그리고 본 연구는 이러한 위기를 사전에 예측하기 위한 모형을 구축하는 것을 목적으로 한다. 재정위험에 대해 논의할 때 가장 많이 언급되는 지표가 GDP 대비 정부부채 비율이다. 정부부채 비율이 높을수록 재정여력이 축소되어 외생적인 경제적 충격에 더 취약해진다. 높은 정부부채 비율은 국가신용도를 낮추고 채무불이행의 가능성을 높이는 요인이기도 하다. 또한 장기적으로 재정의 지속가능성을 판단할 때 중심이 되는 지표가 바로 GDP 대비 정부부채 비율이다.

재정위기를 식별하기 위해서는 정부부채 비율의 증가 속도에 대한 임계점을 구체적으로 정의할 필요가 있다. 본 연구의 기본 분석에서는 재정위기를 정부부채 비율이 연평균 3%p 이상 증가한 시기로 정의한다. 이러한 정의에

기반하여 재정위험 진단을 위한 예측 시계는 1~3년이다. 예측 시계가 1년인 경우는 지난해( $t-1$ 년) 대비 올해( $t$ 년), 즉 1년간 정부부채 비율이 3%p 이상 증가하는 것을 의미하고, 예측 시계가 2년인 경우는 작년( $t-1$ 년) 대비 내년( $t+1$ 년), 즉 2년간 정부부채 비율이 6%p 이상 상승하는 것을 의미한다. 마찬가지로 예측 시계가 3년인 경우는 작년( $t-1$ 년) 대비 후년( $t+2$ 년), 즉 3년간 정부부채 비율이 9%p 이상 증가하는 것을 의미한다.

재정위기를 이렇게 정의한 이유는 크게 두 가지이다. 첫 번째는 이러한 정의는 매우 직관적이면서 재정위기를 쉽게 판별할 수 있기 때문이다. 재정위험을 진단할 때 그 위험의 정의가 복잡하고 실제 재정 측면에서 어떤 결과를 예측하는 것인지 분명하지 않다면 재정위기에 대한 경고 신호가 울리더라도 여기에 대처하는 것이 쉽지 않을 수 있다. 이러한 측면에서 정부부채 비율의 증가 속도는 쉽게 이해할 수 있으면서 정부가 예산을 편성하고 집행할 때 수입과 지출 등을 통해 직접적으로 통제하고 관리할 수 있다는 장점이 있다. 이뿐만 아니라 우리나라를 포함한 전 세계 많은 국가들이 정부부채 비율 통계를 작성하여 공표한다. 이에 재정위험의 실현 여부 및 관리에 대한 노력 정도를 투명하게 확인할 수 있다.

두 번째는 우리나라의 재정위험을 진단하는 목적에 부합하기 때문이다. 재정위험 진단의 주된 목적은 재정의 지속가능성을 확보하는 것이다. 지속가능한 재정의 핵심적인 부분은 정부부채 비율을 안정적으로 관리하는 것이다. 우리나라의 정부부채 비율은 다른 선진국들과 비교해 아직 양호한 수준이지만 고령화 등으로 꾸준히 증가할 전망이다. 따라서 중장기적으로는 정부부채 비율이 일정 수준을 넘지 않도록 관리하는 것이 재정의 지속가능성 확보에 중요한 부분이다. 그리고 단기적인 관점에서는 정부부채 비율의 증가 속도를 가능한 억제할 필요가 있다. 고령화 등 구조적인 요인에 의해 장기적으로 정부부채 비율이 증가할 것으로 전망되지만, 단기적인 경제적 충격 등이 이러한 정부부채 비율의 상승을 가속시킬 수 있기 때문이다. 정부부채 비율의 증가 속도를 관리하기 위해서는 정부부채 비율이 빠르게 증가할 가능성을 사전에 식별하여 이에 대처할 방안을 마련할 필요가 있다. 따라서

재정위기를 정부부채 비율의 증가 속도를 기준으로 정의하여 정부부채 비율의 증가 속도가 일정 수준을 넘어설 것으로 예상될 때 사전에 재정위기에 대한 경고 신호를 보내는 것에는 의의가 있다고 판단된다.

이러한 재정위기의 정의는 자의적이라는 문제점이 있다. 정부부채 비율이 어느 정도의 속도로 증가해야 위험한지에 대한 명확한 기준은 없다. 따라서 민감도 분석으로 정부부채 비율의 상승 속도를 연평균 2%p로 더 엄격하게 정의할 때와 연평균 4%p로 더 관대하게 정의할 때에 대해서도 분석한다. 다만 뒤에서 논의하듯이 정부부채 비율의 증가 속도를 연평균 3%p로 정의할 때 우리나라 과거의 재정위기 기간을 합리적으로 식별하는 것으로 판단된다. 그러나 이러한 기준은 정부부채 비율의 수준에 따라 다르게 정의할 필요가 있다. 정부부채 비율이 높을수록 정부부채의 증가 속도를 더욱 엄격하게 관리해야 하기 때문이다.

본 연구는 정부부채로 일반정부를 포괄하면서 국제 비교가 가능한 D2를 기준으로 분석한다. 정부부채에 대한 자료는 국제통화기금(IMF)의 World Economic Outlook(이하 IMF-WEO) 자료를 활용한다. IMF-WEO는 전 세계 약 200개 국가의 경제를 전망하는 자료이지만, 주요 경제 및 재정변수의 과거 실현된 수치도 함께 제공하여 본 연구에서 활용하기에 유용하다. 특히 IMF-WEO는 재정에 대한 정보로 일반정부의 정부부채뿐만 아니라 기초재정수지, 구조적 재정수지, 재정수입, 재정지출 등에 대한 정보를 제공한다.

분석 자료로는 IMF-WEO 자료에 경제협력개발기구(OECD)의 Economic Outlook(이하 OECD-EO) 자료를 보완하여 활용한다. OECD-EO 자료 또한 IMF-WEO와 유사하게 OECD 회원국과 그 외 신흥국의 경제를 전망하면서 주요 변수들에 대한 과거 자료를 제공한다. 특히 장·단기 금리와 같이 IMF-WEO에서 제공하지 않는 정보를 OECD-EO가 추가적으로 제공한다. 따라서 본 연구는 본 장에서만뿐만 아니라 제Ⅲ장의 재정위험을 진단할 때에도 IMF-WEO와 OECD-EO 자료를 이용하여 분석한다. 본 연구가 이용한 자료의 기간은 1981~2023년으로 모두 전망치를 제외한 실제 실현된 정보를 기준으로 분석한다. 다만 국가나 변수마다 이용할 수 있는 자료의 기간이 다르며 대

체로 1990년 이후의 정보를 제공한다. 우리나라의 실제 정부부채 비율 또한 1990~2022년까지 총 33년에 대한 자료가 존재한다.

본 연구의 분석 대상 국가는 IMF-WEO 자료에서 선진국으로 분류되면서 OECD-EO 자료에도 포함된 32개 국가로 한정한다. 분석 대상인 32개국 중 크로아티아를 제외한 31개국이 OECD 회원국이다. 국제기구에서 수행한 주요 선행연구들은 전 세계 국가 각각의 재정위험을 진단할 필요가 있어 상당히 많은 국가들을 대상으로 분석하였다. 그러나 선행연구와 달리 본 연구는 우리나라의 재정위험 진단이 중요한 문제이며, 이 경우에는 유의미한 표본 수를 확보하면서도 우리나라와 유사한 상황의 국가들에 한정하여 분석하는 것이 유리하다. 이에 선행연구와 달리 기본 분석에서는 선진국에 한정하여 분석하고, 민감도 분석으로 신흥국도 포함한 분석 결과를 제시한다. 신흥국을 포함하면 분석 대상 국가는 49개국으로 확대된다.<sup>10)</sup>

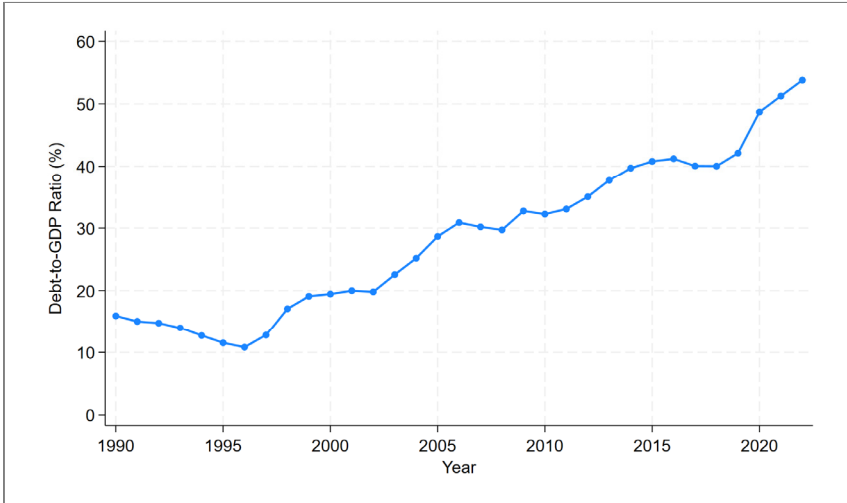
정부부채 비율의 증가 속도로 우리나라의 재정위기 시기를 식별해 보자. [그림 II-3]은 우리나라의 정부부채 비율에 대한 시계열 자료이다. 서론에서 국가채무(D1) 비율을 통해 살펴본 바와 같이 우리나라의 정부부채(D2) 비율도 대체로 상승하는 추세에 있다. 그러나 그 증가 속도는 시기마다 차이가 있다. [그림 II-3]의 자료를 차분한 자료는 [그림 II-4(가)]에 제시되어 있다. [그림 II-4(가)]에서 첫 번째 패널은 지난해 대비 올해의 정부부채 비율 증가폭을 시계열로 나타낸 것이다. 그리고 두 번째 패널은 작년 대비 내년의 정부부채 비율 증가폭이며, 세 번째 패널은 작년 대비 후년의 정부부채 비율 증가폭이다. [그림 II-4(나)]는 정부부채 비율의 증가폭에 대한 분포를 나타낸 것이다. 그리고 마지막으로 [그림 II-4(다)]는 정부부채 비율 증가폭에 일정한 기준을 적용하여 식별한 재정위기 시기를 나타낸다.

---

10) 본 연구에서 고려하는 신흥국은 아르헨티나, 브라질, 불가리아, 칠레, 중국, 콜롬비아, 코스타리카, 헝가리, 인도, 인도네시아, 멕시코, 페루, 폴란드, 루마니아, 러시아, 남아프리카공화국, 튀르키예 등 총 17개국이다.

[그림 II-3] 한국의 GDP 대비 정부부채 비율

(단위: GDP 대비 %)



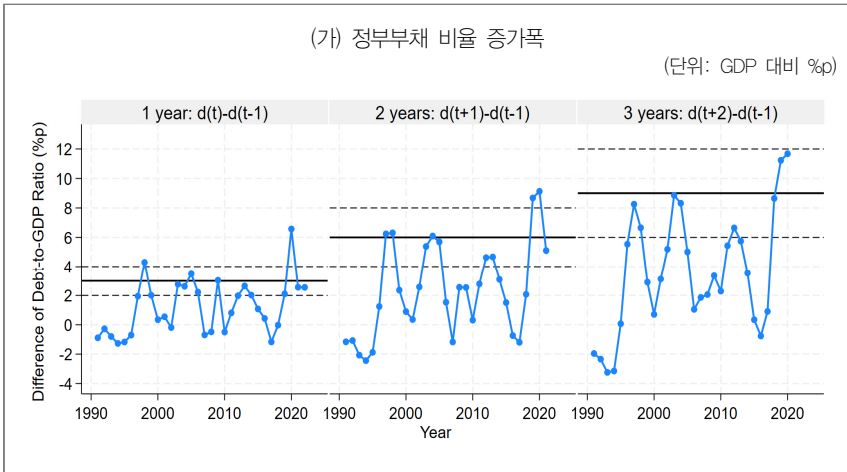
주: 원자료에서 2011년 이후 수치는 일반정부 부채(D2) 비율에 대한 통계이나, 2010년 이전의 수치는 국가채무(D1) 비율과 가까워 2010년 이전의 자료는 원자료에 2011년의 일반정부 부채(D2) 비율과 국가채무(D2) 비율의 차이인 2.761%p를 더한 값으로 조정

자료: IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.

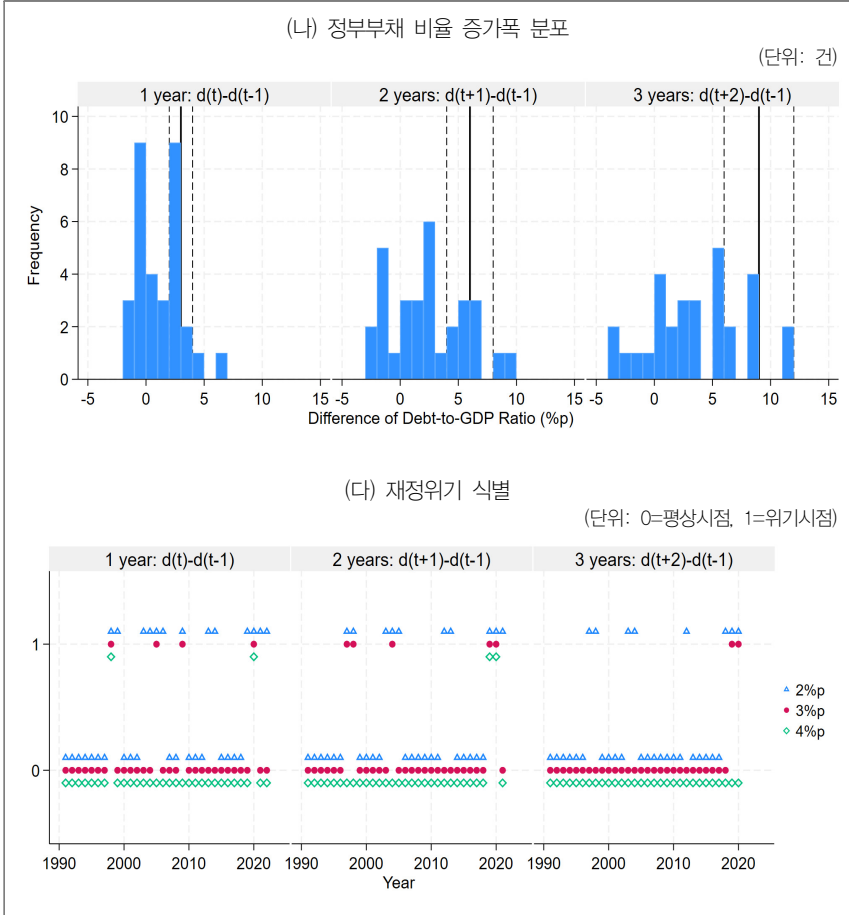
[그림 II-4] 한국의 재정위기 식별

(가) 정부부채 비율 증가폭

(단위: GDP 대비 %p)



[그림 II-4]의 계속



자료: IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24. 자료를 이용하여 저자 작성

기본 분석에서 재정위기를 겪은 해는 정부부채 비율이 1년간 3%p, 2년간 6%p, 3년간 9%p 이상 증가한 해를 의미한다. [그림 II-4(다)]에서 붉은 동그라미 표식이 1로 식별된 시점이 기본 분석의 재정위기 시기를 의미한다. 1년간 정부부채 비율이 3%p 이상 증가한 해는 1998년, 2005년, 2009년, 2020년으로 총 4개년이 식별된다. 1998년은 외환위기, 2009년은 국제 금융위기, 2020년은 코로나19 위기와 관련이 있다.<sup>11)</sup> 2년간 정부부채 비율이 6%p 이상 증가



한 시기에는 국제 금융위기 기간이 제외되지만 다른 시기는 포함이 된다. 그리고 3년간 정부부채 비율이 9%p 이상 증가한 시기는 코로나19 위기만 식별된다.

기본 분석과 같이 재정위기를 정의하기 위한 정부부채 비율 증가 속도의 기준이 꼭 연평균 3%p일 필요는 없다. 이는 정부부채 비율의 수준이나 향후 예상되는 구조적 위험 등을 고려하여 더 엄격한 기준을 적용하거나 아니면 더 완화된 기준을 적용할 수도 있다. 기준을 바꿔 재정위기 식별 기준을 정부부채 비율이 연평균 2%p 이상인 시기로 정의하면 다소 많은 시기가 재정위기 기간으로 식별된다. 1년간 정부부채 비율의 증가 속도를 기준으로 식별할 경우 2%p를 넘어선 해는 전체 32개년 중 13개년이다. 반면 재정위기를 정부부채 비율의 증가 속도가 연평균 4%p 이상인 시기로 정의하면 재정위기가 다소 적게 식별된다. 1991~2022년 중 1년간 정부부채 비율이 4%p 이상 증가한 해는 1998년과 2020년 단 2개년이다. 그리고 3년간 12%p 이상 증가한 해는 1991년 이후 한 차례도 관찰되지 않았다.

본 연구의 기본 분석에서는 정부부채 비율의 증가 속도가 연평균 3%p 이상인 해를 재정위기로 식별하고, 이를 사전에 예측하기 위한 모형을 구축하고자 한다. 재정위기를 정부부채 비율이 연간 3%p 이상 증가한 기간으로 정의하는 것은 다소 자의적이라는 단점이 있지만, 이 기준에 의해 과거 우리나라의 재정위기로 식별된 빈도 수가 본 연구의 기본 분석으로 진행하기에 적절한 것으로 판단된다.

본 연구의 분석 대상인 선진국과 신흥국에 대해 동일한 기준을 적용하면 상당히 많은 국가에서 재정위기가 식별된다. [그림 II-6]은 선진국을 대상으로, 그리고 [그림 II-7]은 선진국뿐만 아니라 신흥국도 포함하여 재정위기를 식별한 결과를 제시한다. 대부분의 시기에 재정위기를 겪은 국가가 존재하

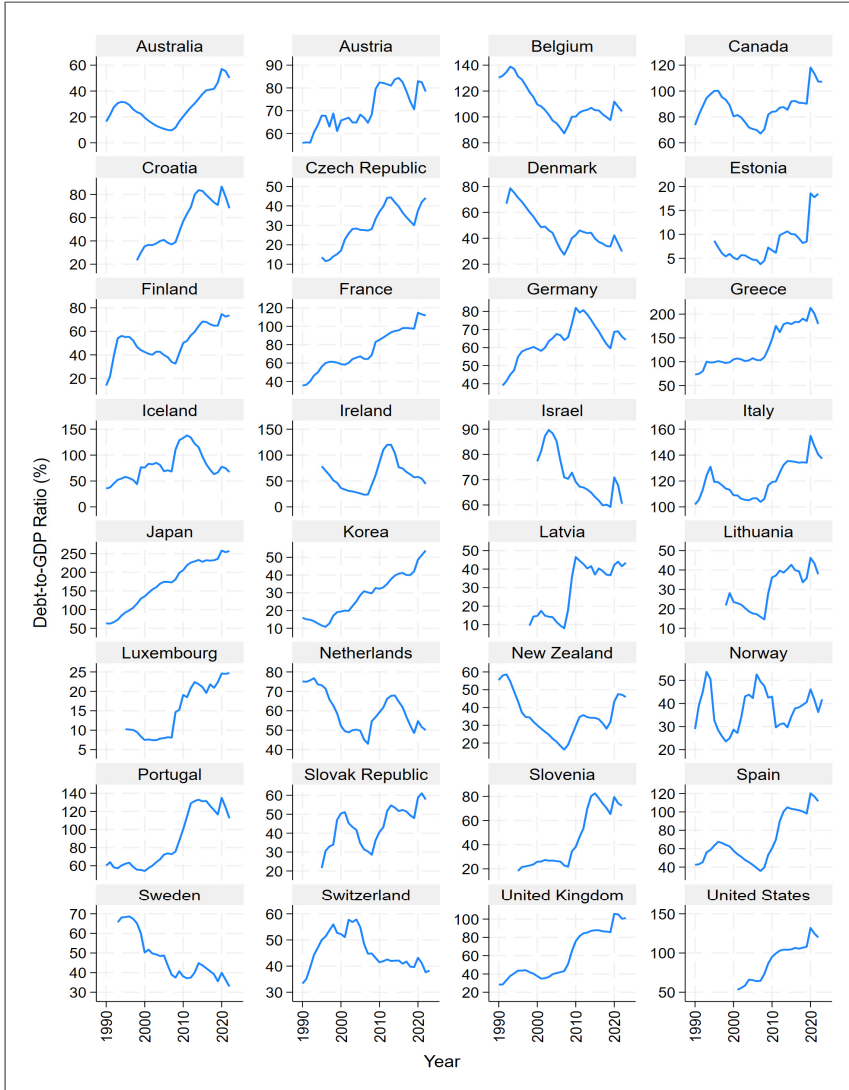
---

11) 2005년은 여기서 언급한 세 가지 경제위기처럼 큰 위기가 있었던 시기는 아니다. 그러나 2000년대 초반에는 외환위기의 여파와 신용카드 대란 등의 영향으로 경제에 어려움이 있었다. 한편, 정부부채 비율의 급증이 항상 경제위기와 관련된 것은 아니다. 사회복지 확대 등 정부의 정책적 변화나 인구구조 변화와 같은 사회구조적 영향 등 다양한 요인이 정부부채 비율에 영향을 미칠 수 있다.

며, 특히 국제 금융위기와 코로나19 위기에 정부부채 비율이 급증한 국가가 많이 관찰된다.

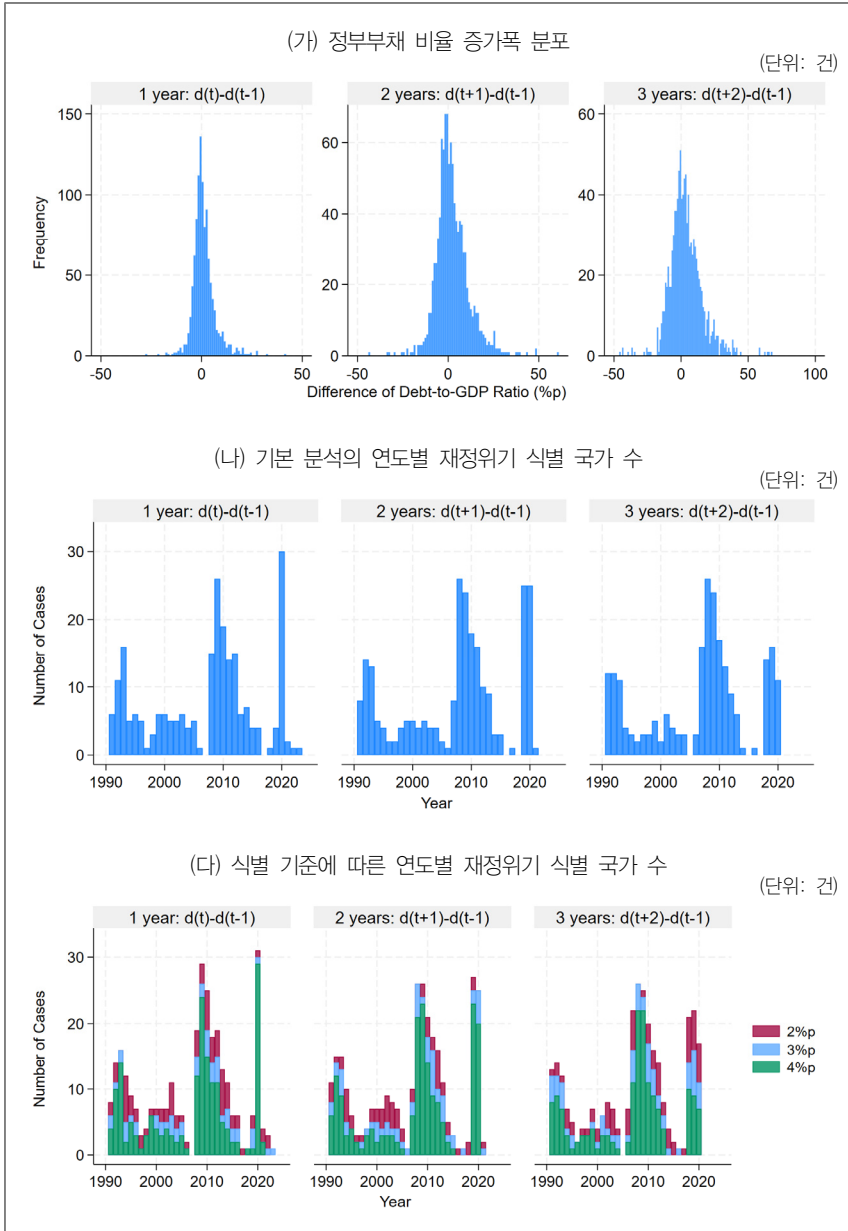
[그림 II-5] 선진국의 GDP 대비 정부부채 비율

(단위: GDP 대비 %)



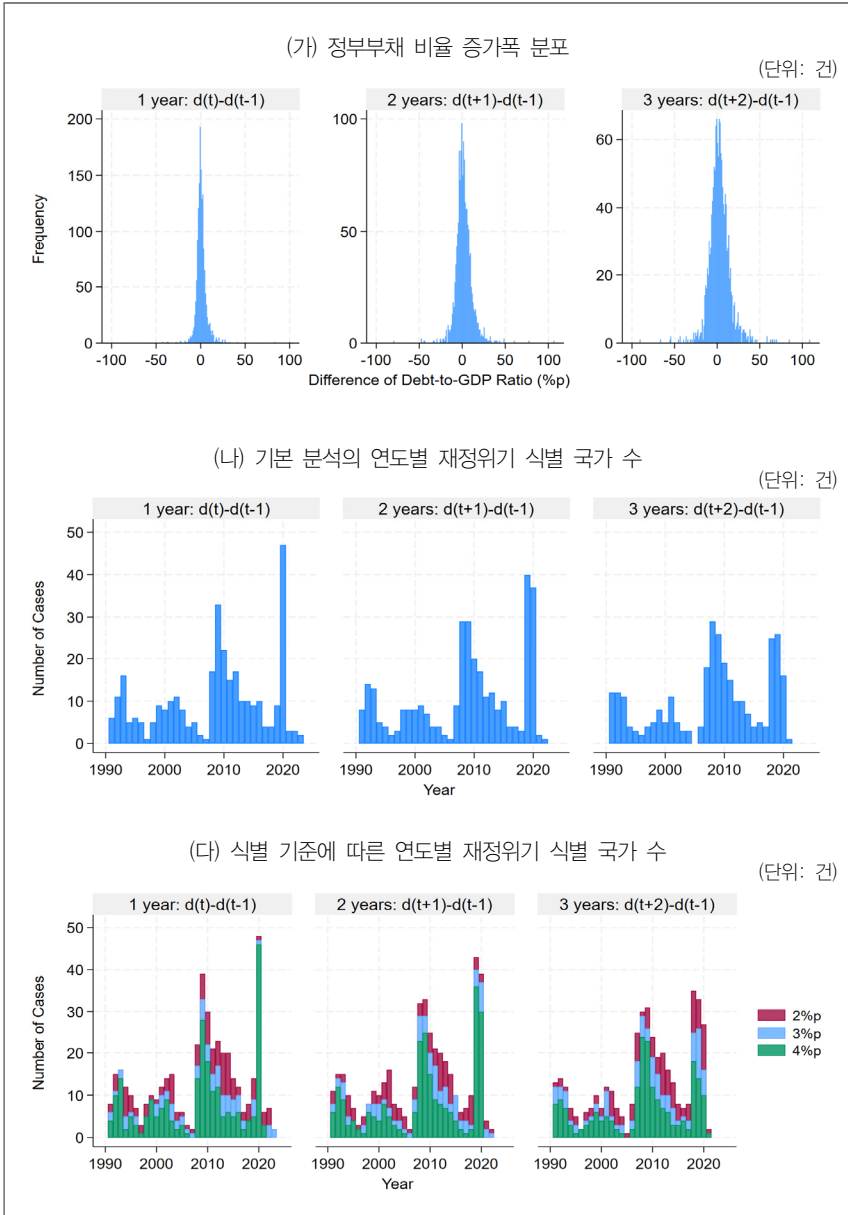
자료: IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.

[그림 II-6] 선진국의 재정위기 식별



자료: IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24. 자료를 이용하여 저자 작성

[그림 II-7] 선진국과 신흥국의 재정위기 식별



자료: IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24. 자료를 이용하여 저자 작성

### 3. 재정위기 전후의 특징

#### 가. 분석 모형 및 자료

재정위기로 식별된 시기의 전후 기간에 대해 주요 재정변수와 거시경제 변수의 움직임을 살펴보고자 한다. 이를 위해 Medas et al.(2018)과 같이 사건 분석(event study) 방법론을 활용한다. 구체적인 회귀방정식은 식 (II-1)과 같이 설정한다. 국가 고정효과를 가정한 패널회귀분석을 통해 재정위기 전후의 주요 변수의 움직임을 추정하는 것이다. 재정위기로 식별된 시기의 첫 번째 해를 0으로 설정하여 재정위기가 시작된 연도의 2년 전부터 5년 후까지 총 8년간의 변화를 살펴본다. 재정위기가 여러 해에 걸쳐 연속으로 식별되면 연속으로 식별된 전체 기간을 하나의 재정위기 기간으로 간주하고 첫 번째 해를 시점 0으로 정하여 분석한다.

$$y_{i,t} = \alpha_i + \sum_{j=-2}^5 \beta_j D_{i,t+j} + \varepsilon_{i,t} \quad \text{식 (II-1)}$$

- $y_{i,t}$ 는  $t$  연도의  $i$  국가에 대한 관심 변수
- $\alpha_i$ 는 국가 고정효과
- $D_{i,t+j}$ 는  $t+j$  연도에  $i$  국가의 재정위기가 시작했으면 1, 그렇지 않으면 0인 더미변수

한편 특정 시점에 둘 이상의 재정위기로 인한 영향이 미치는 해는 분석에서 제외한다. 달리 말하면 본 분석에서 재정위기 시작 시점 전후로 총 8년간의 재정 및 거시경제 변수의 움직임을 살펴보는데, 서로 다른 재정위기의 시작 시점이 최소한 8년 이상 떨어져 있는 경우만을 대상으로 분석을 수행한다. 그렇지 않으면 특정 시점의 반응이 둘 이상의 재정위기에 영향을 받을 수 있어 재정위기 전후의 주요 변수들의 반응을 해석하는 데 어려움이 있다. 또한 식 (II-1)에서 주요 관심 추정치는  $\beta_j$ 인데, 이는 재정위기 전후 8년의 기간에 속하지 않는 평상시 대비 재정위기 전후 기간의 해당 변수의 변화를 의미한다.

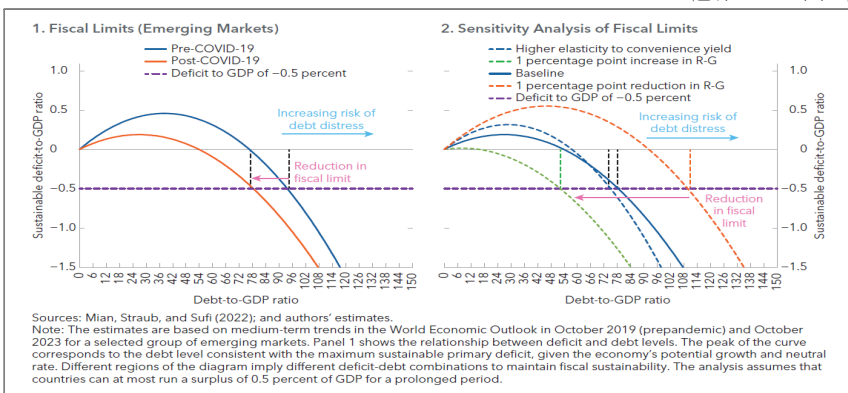
주요 재정변수로는 정부부채, 기초재정수지, 재정수입, 재정지출의 GDP 대비 비율을 살펴본다. 그리고 주요 거시경제 변수로는 경제성장률, 실질이자율, 이자율-성장률 격차를 살펴본다. 식 (II-2)는 채무 동학을 나타내는데, 정부부채 비율의 움직임을 결정하는 주요 변수는 기초재정수지, 경제성장률, 이자율-성장률 격차이다. 특히 [그림 II-8]에 나타난 바와 같이 정부부채 비율의 한도(fiscal limit)는 고정되어 있지 않고 이자율-성장률 격차에 민감하게 반응한다. 특히 아시아 신흥국의 경우 이자율-성장률 격차가 1%p 증가하면 자본조달 비용이 증가하여 정부부채 비율의 한도가 78%에서 약 50%로 급격히 낮아지고 재정여력이 축소되는 것으로 나타난다. 따라서 재정위기 전후 이러한 변수들의 움직임을 살펴보는 것은 의미가 있다.

$$\Delta d_{t+1} = \left( \frac{r_t - g_t}{1 + g_t} \right) d_t - pb_t + sfa_t \quad \text{식 (II-2)}$$

- $d$ 는 GDP 대비 정부부채 비율
- $r$ 는 이자율
- $g$ 는 경상성장률
- $pb$ 는 기초재정수지
- $sfa$ 는 유량-저량 조정

[그림 II-8] 재정수지와 이자율-성장률 격차 변화에 따른 재정여력 변화

(단위: GDP 대비 %)



자료: Flores et al., "Upgrading Fiscal Frameworks in Asia-Pacific," Departmental Paper, International Monetary Fund, 2024. 9., p. 23, Figure. 23.

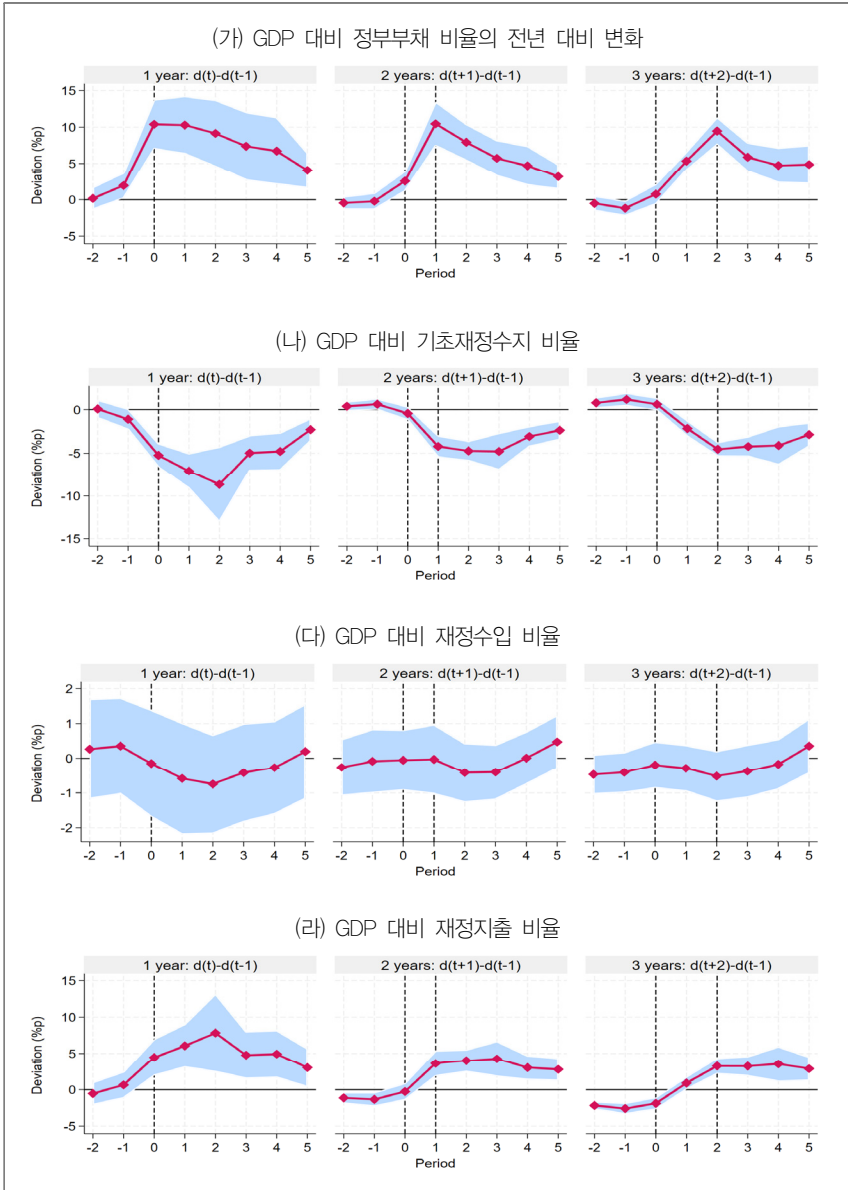
분석을 위한 재정변수와 경제성장률, 물가상승률 자료는 IMF-WEO 자료를 이용하였다. 그리고 이자율은 OECD-EO 자료 중 정부채권의 장기이자율 변수를 이용하였다. 실질이자율은 이러한 명목이자율에서 IMF-WEO의 해당 연도 평균 소비자물가상승률을 차감하여 계산한 변수이다. 이자율-성장률 격차는 실질이자율에서 경제성장률을 차감하여 계산하였다. 앞에서도 언급하였듯이 분석 기간은 1981~2023년이나 국가와 변수마다 이용할 수 있는 자료의 기간이 다르다.

#### 나. 기본 분석 결과

재정위기 시작 시점에 주요 재정변수들의 지표는 크게 악화되며, 이러한 영향은 일정 기간 지속되는 것으로 나타난다. [그림 II-9의 각 그림에 세로선이 있는데, 세로선이 하나인 그림에서는 해당 시점이, 그리고 세로선이 두 개인 그림에서는 세로선을 포함하여 두 세로선 사이의 기간이 재정위기 기간이다. 정부부채 비율은 재정위기 기간에 급증하는 것으로 나타나는데, 이는 재정위기의 정의에 따른 자연스러운 결과이다. 그러나 재정위기 기간이 지난 이후에도 정부부채 비율의 빠른 증가는 상당 기간 지속되는 것으로 나타난다. 정부부채 비율이 빠르게 증가할 때 기초재정수지는 악화되는데, 이는 재정수입의 감소보다는 재정지출의 증가로 인한 영향이 더 큰 것으로 보인다. 재정수입의 감소도 관찰되지만 이는 다른 재정변수들과 달리 일시적이며, 변동폭도 다른 재정변수들보다 작다. 그러나 재정지출의 증가는 상당 기간 지속되는 것으로 나타난다.

주요 거시경제 변수의 변화를 살펴보자. 재정위기 기간에는 경제성장률이 급격히 낮아지는 것으로 분석된다. 그리고 낮아진 경제성장률이 상당 기간 지속되는 것으로 나타난다. 하지만 재정위기 기간에 정부채권에 대한 장기 실질이자율은 증가하거나 최소한 통계적으로 유의미한 수준으로 감소하지는 않는 것으로 나타난다. 경기가 하강하면 이자율이 낮아질 가능성도 존재하지만 정부부채 비율의 급격한 증가는 정부채권의 공급을 증가시키는 동시에 위험도 높여, 정부채권의 장기 이자율을 상승시키는 효과가 반영된 것으로

[그림 II-9] 주요 재정변수의 위기 전후 변동



주: 푸른색 영역은 강건표준오차로 측정된 90% 신뢰구간

자료: IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24. 자료를 이용하여 저자 분석



해석된다. 이에 정부채권의 장기 실질이자율에서 경제성장률을 차감한 이자율-성장률 격차가 크게 증가하며 이 또한 상당 기간 증가한 상태에 머물러 있는 것으로 분석된다. 즉 재정위기는 정부부채 한도와 재정여력을 크게 축소시키는 것으로 해석할 수 있다.

〈표 II-3〉 주요 재정변수의 위기 전후 변동

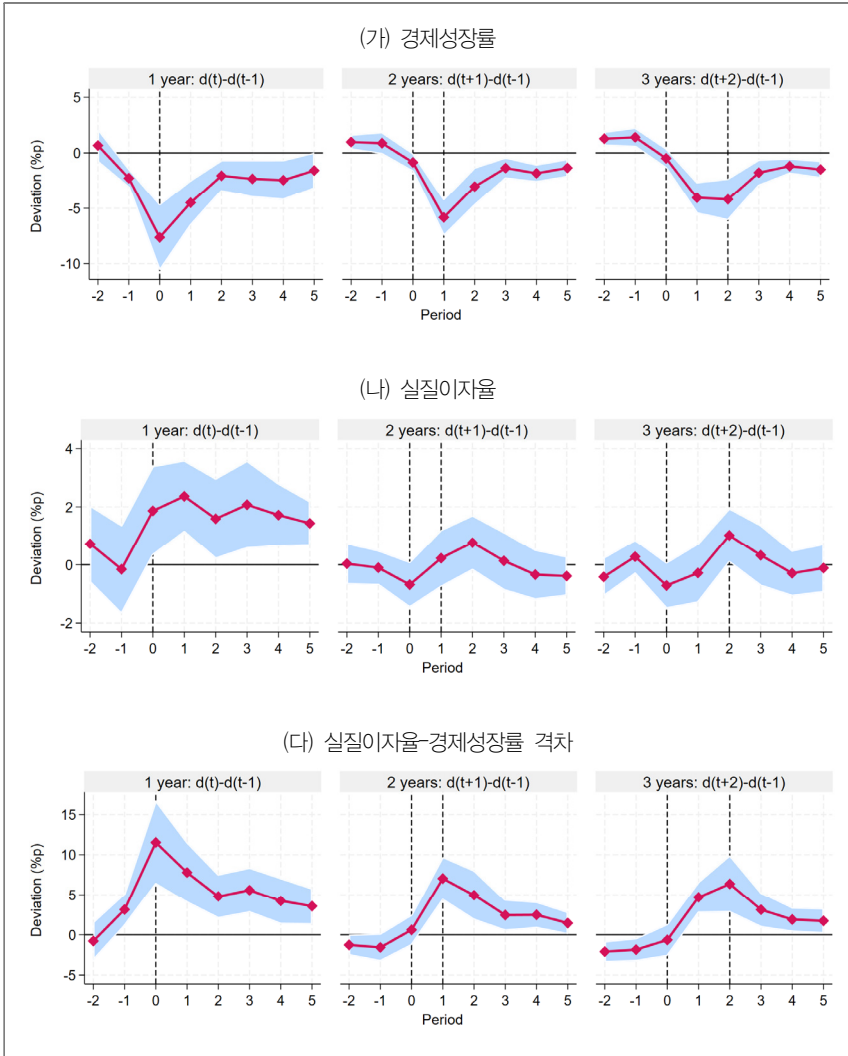
변수	GDP 대비 정부부채 비율의 전년 대비 변화			GDP 대비 기초재정수지			GDP 대비 재정수입			GDP 대비 재정지출		
	1년	2년	3년	1년	2년	3년	1년	2년	3년	1년	2년	3년
(t-2)기	0.175 (0.885)	-0.449 (0.511)	-0.514 (0.582)	0.0863 (0.604)	0.393 (0.315)	0.795** (0.347)	0.265 (0.843)	-0.265 (0.476)	-0.469 (0.333)	-0.533 (0.901)	-1.129** (0.437)	-2.165*** (0.310)
(t-1)기	1.937* (1.025)	-0.215 (0.656)	-1.173* (0.601)	-1.097 (0.693)	0.631 (0.373)	1.179** (0.435)	0.352 (0.813)	-0.0842 (0.540)	-0.410 (0.342)	0.652 (1.058)	-1.335** (0.555)	-2.580*** (0.435)
t기	10.39*** (1.990)	2.560*** (0.808)	0.763 (0.808)	-5.256*** (0.803)	-0.427 (0.451)	0.622 (0.423)	-0.151 (0.904)	-0.0538 (0.509)	-0.195 (0.393)	4.456*** (1.486)	-0.258 (0.676)	-1.879*** (0.478)
(t+1)기	10.29*** (2.323)	10.46*** (1.781)	5.346*** (0.750)	-7.078*** (1.202)	-4.223*** (0.729)	-2.161*** (0.536)	-0.587 (0.941)	-0.0298 (0.588)	-0.289 (0.389)	6.068*** (1.754)	3.622*** (1.040)	0.907* (0.507)
(t+2)기	9.147*** (2.683)	7.924*** (1.470)	9.471*** (1.111)	-8.661*** (2.575)	-4.748*** (0.667)	-4.553*** (0.478)	-0.749 (0.833)	-0.418 (0.500)	-0.519 (0.426)	7.826** (3.158)	4.005*** (0.898)	3.259*** (0.641)
(t+3)기	7.374** (2.753)	5.732*** (1.443)	5.871*** (1.169)	-4.994*** (1.192)	-4.801*** (1.239)	-4.238*** (0.676)	-0.419 (0.829)	-0.402 (0.462)	-0.375 (0.444)	4.778** (1.906)	4.257*** (1.430)	3.243*** (0.795)
(t+4)기	6.722** (2.707)	4.690*** (1.586)	4.738*** (1.403)	-4.806*** (1.266)	-3.077*** (0.684)	-4.123*** (1.297)	-0.271 (0.787)	0.00732 (0.446)	-0.173 (0.423)	4.926** (1.913)	3.033*** (0.981)	3.543** (1.419)
(t+5)기	4.056*** (1.421)	3.181*** (0.996)	4.860*** (1.544)	-2.296*** (0.764)	-2.363*** (0.630)	-2.867*** (0.793)	0.196 (0.801)	0.472 (0.452)	0.352 (0.462)	3.029* (1.564)	2.793*** (0.909)	2.885*** (0.964)
상수항	-1.523*** (0.391)	-0.947*** (0.307)	-0.512** (0.203)	1.317*** (0.207)	0.770*** (0.155)	0.653*** (0.127)	41.08*** (0.183)	41.58*** (0.148)	41.49*** (0.0966)	41.66*** (0.359)	42.92*** (0.255)	42.94*** (0.166)

〈표 II-3〉의 계속

변수	GDP 대비 정부부채 비율의 전년 대비 변화			GDP 대비 기초재정수지			GDP 대비 재정수입			GDP 대비 재정지출		
	1년	2년	3년	1년	2년	3년	1년	2년	3년	1년	2년	3년
관측치	371	614	678	369	617	681	378	627	692	378	627	692
국가 수	31	32	32	31	32	32	31	32	32	31	32	32
결정계수	0.391	0.321	0.268	0.405	0.294	0.250	0.00702	0.00495	0.00642	0.243	0.179	0.163

주: ( ) 안은 강건표준오차이며, \*\*\*, \*\*, \*는 각각 1%, 5%, 10% 수준에서 유의함을 의미  
 자료: IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24. 자료를 이용하여 저자 분석

[그림 II-10] 주요 거시경제 변수의 위기 전후 변동



주: 푸른색 영역은 강건표준준차로 측정된 90% 신뢰구간

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis/?t\[s0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&fc=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&d\[ag\]=OECD.ECO.MAD&df\[vs\]=1,0&dq=.GDPV\\_ANNPCT.A&iom=LASTNPERIODS&o=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis/?t[s0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&fc=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&d[ag]=OECD.ECO.MAD&df[vs]=1,0&dq=.GDPV_ANNPCT.A&iom=LASTNPERIODS&o=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

〈표 II-4〉 주요 거시경제 변수의 위기 전후 변동

변수	경제성장률			실질이자율			실질이자율-경제성장률 격차		
	1년	2년	3년	1년	2년	3년	1년	2년	3년
(t-2)기	0.673 (0.870)	0.987** (0.392)	1.284*** (0.369)	0.722 (0.779)	0.0334 (0.414)	-0.418 (0.387)	-0.785 (1.390)	-1.265 (0.749)	-2.100*** (0.755)
(t-1)기	-2.272*** (0.503)	0.879 (0.579)	1.414*** (0.504)	-0.156 (0.897)	-0.103 (0.348)	0.275 (0.345)	3.128** (1.181)	-1.578 (1.003)	-1.854** (0.832)
t기	-7.625*** (1.754)	-0.849* (0.490)	-0.480 (0.525)	1.864** (0.902)	-0.684 (0.458)	-0.714 (0.463)	11.53*** (3.053)	0.598 (1.113)	-0.650 (1.168)
(t+1)기	-4.510*** (1.201)	-5.835*** (0.980)	-4.063*** (0.836)	2.365*** (0.721)	0.223 (0.579)	-0.284 (0.595)	7.791*** (2.218)	7.039*** (1.593)	4.681*** (1.140)
(t+2)기	-2.056** (0.821)	-3.031*** (1.013)	-4.204*** (1.111)	1.591* (0.813)	0.765 (0.556)	1.019* (0.548)	4.797*** (1.615)	4.951** (1.826)	6.369*** (2.102)
(t+3)기	-2.338** (0.985)	-1.358** (0.556)	-1.777** (0.689)	2.075** (0.884)	0.132 (0.589)	0.324 (0.615)	5.586*** (1.646)	2.440** (1.135)	3.112** (1.273)
(t+4)기	-2.451** (1.056)	-1.830*** (0.466)	-1.186*** (0.399)	1.720** (0.636)	-0.340 (0.503)	-0.293 (0.458)	4.218** (1.688)	2.467** (0.955)	1.885** (0.890)
(t+5)기	-1.583 (0.955)	-1.349*** (0.473)	-1.488*** (0.455)	1.434*** (0.446)	-0.389 (0.394)	-0.115 (0.486)	3.553** (1.323)	1.445* (0.812)	1.723* (0.902)
상수항	3.698*** (0.173)	3.248*** (0.120)	2.966*** (0.0798)	2.549*** (0.146)	2.693*** (0.115)	2.549*** (0.0998)	-1.564*** (0.328)	-0.926*** (0.253)	-0.684*** (0.178)

〈표 II-4〉의 계속

변수	경제성장률			실질이자율			실질이자율-경제성장률 격차		
	1년	2년	3년	1년	2년	3년	1년	2년	3년
관측치	382	631	696	361	590	649	361	590	649
국가 수	31	32	32	29	31	31	29	31	31
결정계수	0,250	0,211	0,162	0,102	0,0126	0,0156	0,344	0,175	0,154

주: ( ) 안은 강건표준오차이며, \*\*\*, \*\*, \*는 각각 1%, 5%, 10% 수준에서 유의함을 의미

자료: 다음의 자료를 이용하여 저자 분석

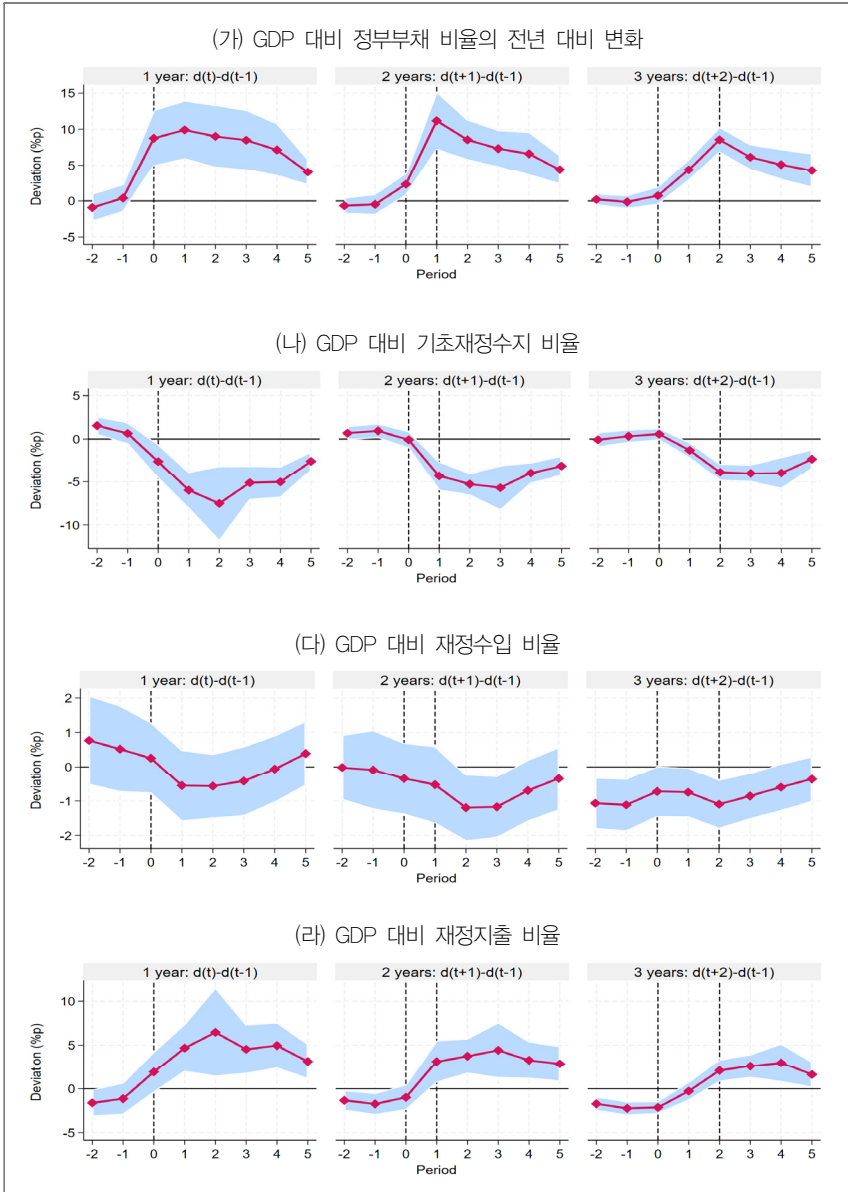
1. IIMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?ts\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&lom=LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?ts[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&lom=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

#### 다. 민감도 분석 결과

이하에서는 세 가지 민감도 분석을 실시한다. 기본 분석에서는 재정위기를 정부부채 비율이 연평균 3%p 이상 증가한 시기로 정의하여 분석하였다. 하지만 재정위기를 정의할 때 정부부채 비율의 증가 속도에 대한 임계점은 자의적일 수 있다. 이에 이 임계점을 2%p와 4%p로 가정한 두 가지 민감도 분석을 실시한다. 그리고 기본 분석은 선진국을 대상으로 분석하였다. 이에 정부부채 비율의 증가 속도에 대한 임계점은 기본 분석과 마찬가지로 3%p로 유지한 채 분석 대상 국가를 선진국뿐만 아니라 신흥국까지 확장한 민감도 분석을 실시한다.

민감도 분석 결과는 기본 분석 결과와 정성적 특징이 거의 같다. 모든 민감도 분석에서 정부부채, 기초재정수지 등의 재정변수는 재정위기 시기에 크게 악화되며, 그 영향이 상당 기간 지속되는 것으로 나타난다. 또한 재정수지의 악화는 재정수입의 감소보다 재정지출의 증가로 인한 영향이 더 큰 것으로 관찰된다. 거시경제 변수에 대한 분석 결과도 기본 분석과 민감도 분석이 유사하다. 재정위기 기간에 경제성장률이 크게 둔화되고, 이자율·성장률 격차도 커져 자본조달 비용도 상승하는 것으로 나타난다. 그리고 이러한 상황은 재정위기 이후에도 상당 기간 지속되는 것으로 분석된다.

[그림 II-11] 주요 재정변수의 위기 전후 변동(민감도 분석: 연평균 2%p)

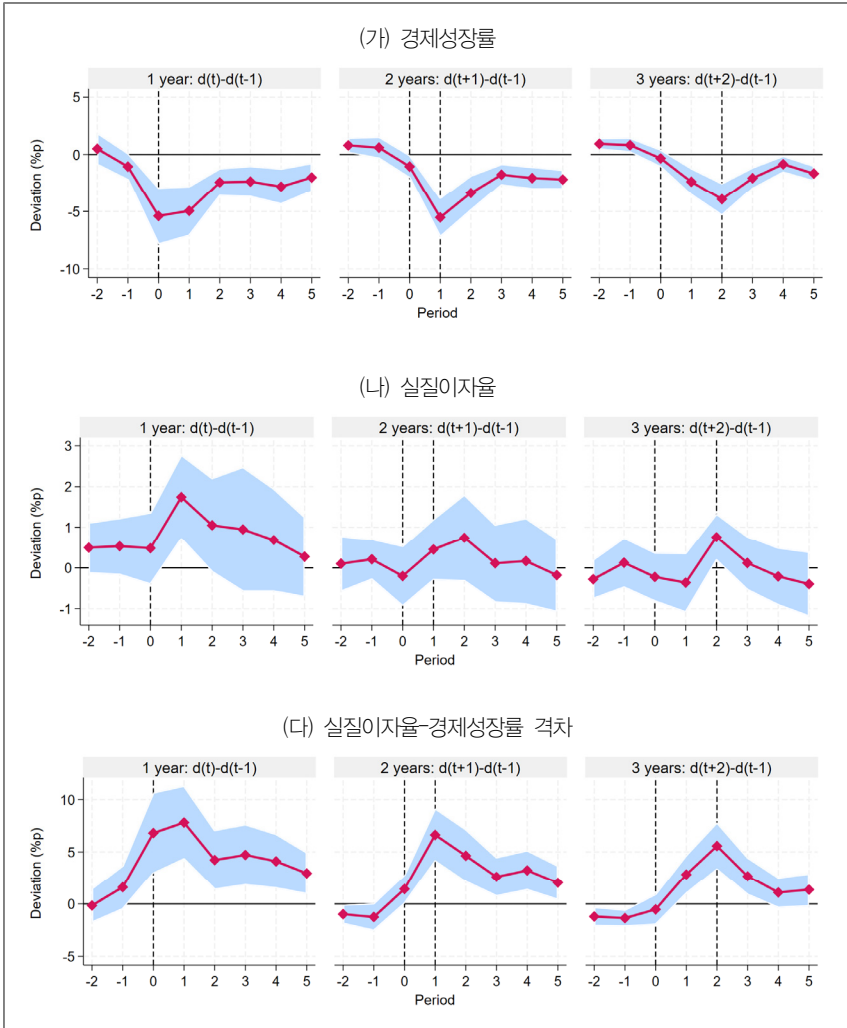


주: 푸른색 영역은 강건표준오차로 측정된 90% 신뢰구간

자료: IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24. 자료를 이용하여 저자 분석



[그림 II-12] 주요 거시경제 변수의 위기 전후 변동(민감도 분석: 연평균 2%p)

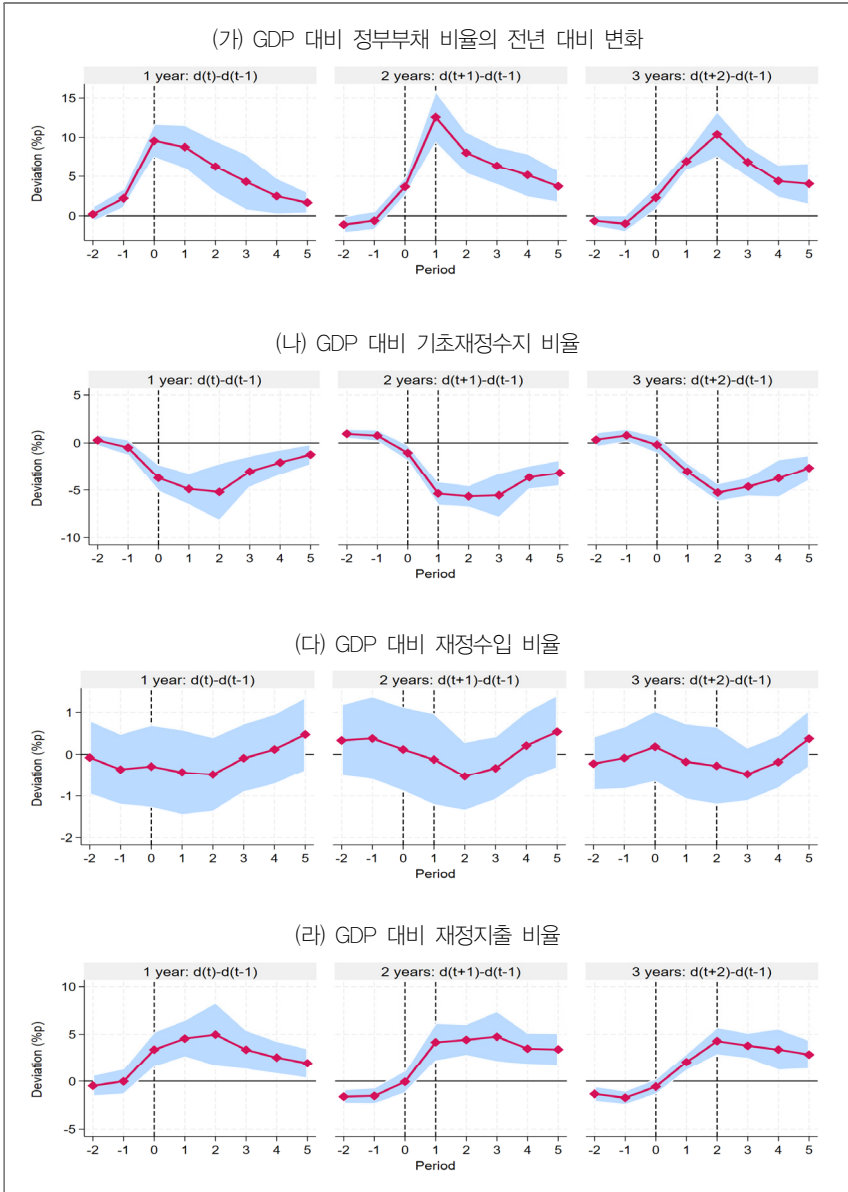


주: 푸른색 영역은 강건표준오차로 측정된 90% 신뢰구간

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?ts\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?ts[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

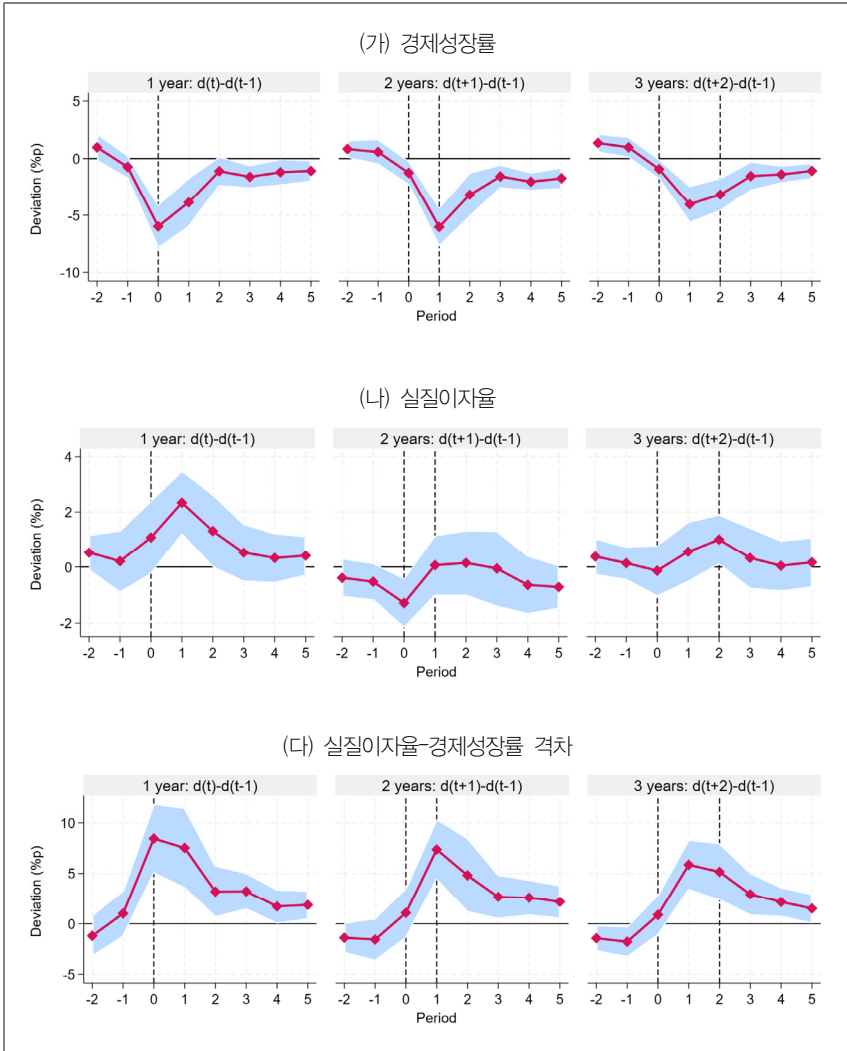
[그림 II-13] 주요 재정변수의 위기 전후 변동(민감도 분석: 연평균 4%p)



주: 푸른색 영역은 강건표준오차로 측정된 90% 신뢰구간

자료: IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24. 자료를 이용하여 저자 분석

[그림 II-14] 주요 거시경제 변수의 위기 전후 변동(민감도 분석: 연평균 4%p)

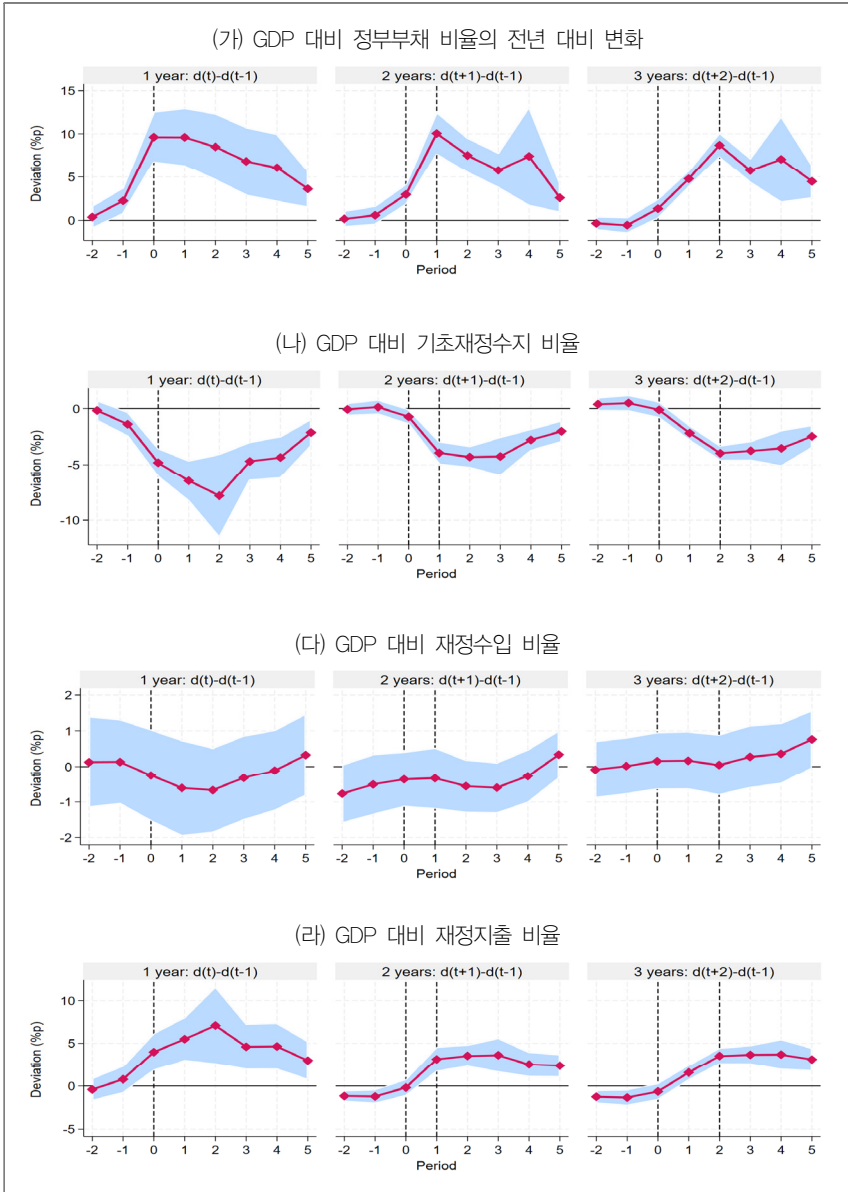


주: 푸른색 영역은 강건표준오차로 측정된 90% 신뢰구간

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&om=LASTNPERIODS&lo=5&io\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&om=LASTNPERIODS&lo=5&io[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

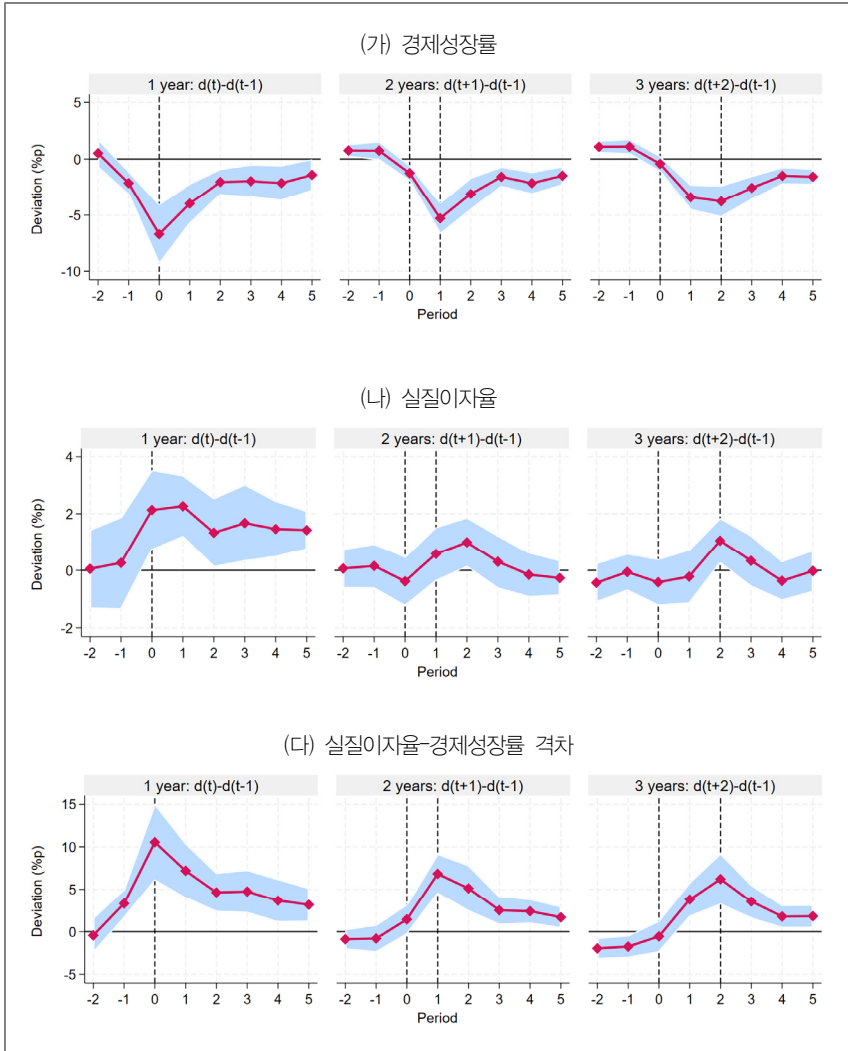
[그림 II-15] 주요 재정변수의 위기 전후 변동(민감도 분석: 신흥국 포함)



주: 푸른색 영역은 강건표준준차로 측정된 90% 신뢰구간

자료: IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24. 자료를 이용하여 저자 분석

[그림 II-16] 주요 거시경제 변수의 위기 전후 변동(민감도 분석: 신흥국 포함)



주: 푸른색 영역은 강건표준오차로 측정된 90% 신뢰구간

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&om=LASTNPERIODS&lo=5&io\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&om=LASTNPERIODS&lo=5&io[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

---

## Ⅲ. 재정위험 진단 방법

---

### 1. 연구 설계

본 장에서는 IMF-WEO와 OECD-EO에서 얻은 국가들의 경제 변수들을 이용하여 정부부채 비율이 급증할 사건을 예측하는 여러 모형을 구축하고 각 모형들의 예측력 및 특징들을 비교하고자 한다. 예측을 위한 분석에는 기존의 선행연구와 유사하게 GDP 관련 변수, 재정, 고용, 국제무역, 물가, 환율, 선진국 및 세계 경제 등과 관련된 변수 등이 활용되었다. 다만 IMF-WEO와 OECD-EO에 포함된 모든 자료를 이용하는 것이 아닌, 상관관계가 높은 변수들을 제외하고 정부부채 비율 예측과 관련이 있으면서 국제 비교가 가능한 변수를 중심으로 분석에 포함하였다. 변수들의 수준뿐만 아니라 증가율, 차분, 타 변수와의 격차 등 다양한 변수를 생성하여 분석에 포함하였다. 예측 시점에는 전년도 등 과거 시점에 실현된 변수만 관측할 수 있기 때문에 이를 고려하여 시차변수를 기준으로 분석하였다.<sup>12)</sup>

정부부채가 급증하는 사건은 제Ⅱ장에서와 같이 GDP 대비 정부부채 비율이 1년 동안 3%p, 2년 동안 6%p, 3년 동안 9%p 이상 증가한 시기로 정의하였다. 추가적으로 이 정의에 대해 민감도 분석을 실시하였다. 한편 본 장은 재정위험 진단을 위한 다양한 변수들을 고려하다 보니 제Ⅱ장보다 분석 기간이 축소되었다. 이에 모형을 통해 도출된 결과는 2000년 이후 기간에 대해 제시한다.

본 연구가 재정위험 진단을 위해 고려하는 모형은 크게 신호(signaling), 로짓(logit), 라쏘(Least Absolute Shrinkage and Selection Operator, LASSO), 랜덤포레스트(random forest) 모형이다. 로짓 모형은 종속변수가 0과 1로 표현

---

12) 분석에 활용된 구체적인 변수들은 부록을 참고하기 바란다.

되는 이진적 결과를 가질 때 예측 모형으로 자주 활용되며 전통적으로 재정 위험 진단 문헌에도 활용되었다. 신호 모형은 재정위험 진단 문헌에서 자주 활용되는 방법론으로, 특정 경제변수가 어떠한 임계점을 넘을 때 재정위기 신호가 발생했다고 간주하는 예측 모형을 만드는 것이다. 라쏘와 랜덤포레스트는 기계학습 방법론의 일종으로, 예측 문제에서 뛰어난 성능을 보여 자주 활용되고 있다. 각 모형에 대한 자세한 설명은 이어지는 절에서 다루고자 한다.

모형들의 재정위기 예측력을 비교하기 위해서는 동일한 조건하에서 예측 모형을 만들고 그 예측 결과를 측정할 수 있는 지표를 비교해야 한다. 이를 위해 본 연구는 다음과 같은 절차를 거쳤다. 우선 분석자료의 70%는 모형을 적합시키는 훈련데이터(training data)로 분류하고 나머지 30%는 모형의 적합도를 판별하는 평가데이터(test data)로 분류하였다.<sup>13)</sup> 이렇게 평가데이터를 따로 떼어 둔 이유는 재정위험 진단의 목적이 이미 관측된 표본 내 데이터(in-sample data)를 잘 설명하는 것보다는 관측되지 않은 표본 외 데이터(out-of-sample data)를 잘 예측하는 것이기 때문에 표본 외 데이터를 예측하는 성능을 비교·분석하기 위함이다.

비교 대상이 되는 모형들은 훈련데이터를 통해 최적의 모형을 적합한 후 평가데이터를 예측하는 성능 지표를 산출하였다. 여기서 사용된 지표들은 정확도(accuracy), 정밀도(precision), 재현율(recall), F1 점수, F2 점수, AUC(Area Under the Curve)이다. 이 중 AUC를 제외한 나머지 지표들은 종속변수가 이진변수일 때 활용할 수 있는 지표들로, 예측값과 실젯값을 비교해 보여주는 혼돈행렬(confusion matrix)의 정보들을 활용한 지표이다. 본 연구의 환경에서 혼돈행렬은 <표 III-1>과 같으며 예측 성능 지표는 혼돈행렬의 요소들로 계산될 수 있다.

13) 전체 데이터를 시간 순서와 관계없이 무작위로 70%의 훈련데이터와 30%의 평가데이터로 나누었다. 일반적인 예측 모형에서는 과거 데이터로 학습하고 이를 바탕으로 현재나 미래를 예측하는 것이 보편적이므로, 시간 순서대로 앞쪽 70%를 훈련데이터로, 뒤쪽 30%를 평가데이터로 사용하는 것이 더 자연스러울 수 있다. 하지만 본 연구는 예측 모형의 구조가 시간에 따라 크게 변하지 않는다고 가정하였으므로 데이터를 시간 순서대로 분할하든, 무작위로 분할하든 결과에 큰 차이가 없을 것으로 판단된다.

〈표 III-1〉 재정위험 예측에 관한 혼동행렬

구분		실제	
		위기(Crisis)	비위기(Non-Crisis)
예측	신호 발생 (Total Positive Signal, TPS)	참 양성 (True Positive, TP)	거짓 양성 (False Positive, FP)
	신호 미발생 (Total Negative Signal, TNS)	거짓 음성 (False Negative, FN)	참 음성 (True Negative, TN)
전체 (Total Observation, TO)		전체 위험 (Total Crises, TC)	전체 비위험 (Total Non-Crisis, TNC)

자료: 저자 작성

정확도는 전체 데이터에서 올바르게 예측된 비율을 나타내며 모형의 전반적인 성능을 평가하는 척도이다. 하지만 실재값이 한쪽으로 치우쳐 있는 불균형 데이터의 경우에는 정확도가 큰 의미를 가지지 않을 수 있다. 예를 들어, 실재값이 대부분 비위기 사례라고 한다면 모든 상황에서 비위험 진단을 내리는 모형이 매우 높은 정확도를 가질 수 있다. 특히 본 연구는 재정위험을 진단하는 것이 목적인데, 재정위기가 드물게 발생하고 어떤 모형이 이를 보수적으로 예측하여 대부분의 경우 위험 신호를 보내지 않는다면 정확도는 높지만 실제 재정위험을 진단하는 데에는 유용하지 않을 수 있다.

$$\text{정확도 (Accuracy)} = \frac{TP + TN}{\text{Total Observation} (= TC + TNC)}$$

정밀도는 위기라고 예측한 상황 중에서 실제 위기였던 비율을 나타내는 지표이다. 위기를 예측했는데 실제 위기가 아닌 거짓 양성(FP)을 많이 발생시키면 정밀도는 떨어지게 된다. 거짓 경고(false alarm)를 남발하지 않는 것이 중요한 상황에서 유용하게 적용되는 지표이다.<sup>14)</sup>

$$\text{정밀도 (Precision)} = \frac{TP}{TPS (= TP + FP)}$$

14) 스팸메일을 걸러 주는 필터 혹은 신용카드 오용을 예측하는 시스템을 대표적인 예시로 들 수 있다.



재현율<sup>15)</sup>은 실제 위기가 발생했을 때 이를 위기로 예측한 비율을 나타내는 지표이다. 실제 위기를 예측하지 못하는 거짓 음성(FN)이 많으면 재현율은 낮아지게 된다. 이는 위험한 상황을 예측하지 못하고 놓치는 것이 심각한 문제가 되는 상황에서 중요한 지표이다.<sup>16)</sup> 재정위기가 드물게 발생하지만 이를 사전에 잘 예측하기 위해서는 재현율이 높은 모형이 유용할 수 있다.

$$\text{재현율}(Recall) = \frac{TP}{TC(= TP + FN)}$$

F1 점수는 정밀도와 재현율의 조화평균으로 정밀도와 재현율 간의 균형을 평가하는 데 유용한 지표이다. 만일 정밀도 혹은 재현율 중 하나가 매우 낮으면 F1 점수도 낮아지기 때문에 두 지표 간의 균형을 유지하는 모형의 성능을 측정하는 데 적합하다. 재정위험을 진단하는 데 재현율이 높은 모형이 유용하지만 거짓 양성 신호를 남발하면서 재현율이 높은 모형은 좋은 모형이라 할 수 없다. 따라서 재정위험을 진단하는 데 있어 재현율과 정밀도 간의 균형이 중요하다.

$$F1\text{점수} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

F2 점수는 재현율이 더 중요한 경우에 활용되는 점수로 재현율에 가중치를 더 부여한 지표이다.<sup>17)</sup>

$$F2\text{점수} = 5 \times \frac{Precision \times Recall}{4 \times Precision + Recall}$$

15) 민감도(sensitivity)라고도 한다.

16) 예를 들면 암 진단에서 암 환자를 놓치는 것을 최소화하기 위해서 재현율이 중요하다.

17) 일반적으로 재현율(Recall)에 부여하는 가중치에 따라  $F_\beta$  점수를 계산할 수 있으며,  $\beta$ 에 1이 들어간 것이 F1 점수, 2가 들어간 것이 F2 점수이다.

$$F_\beta\text{점수} = (1 + \beta^2) \times \frac{Precision \times Recall}{\beta^2 \times Precision + Recall}$$

한편 하나의 예측 모형 내에서도 위험진단의 임계점을 어떻게 설정하느냐에 따라 예측 성능이 달라질 수 있다. 예를 들어 임계점을 매우 높게 설정하면 위험 신호를 거의 보내지 않기 때문에 거짓 양성(FP) 사례가 발생하는 경우는 거의 없겠지만 마찬가지로 참 양성(TP)으로 구분되는 사례도 거의 없을 것이다. 반면 임계점을 낮게 설정하면 참 양성(TP) 사례가 발생하는 경우도 많겠지만 거짓 양성(FP) 사례도 많이 발생하게 된다. 이처럼 예측의 임계점이 변함에 따라 발생하는 예측 성능의 상충관계를 직관적으로 보여주는 것이 ROC 곡선(Receiver Operating Characteristic Curve)이다.

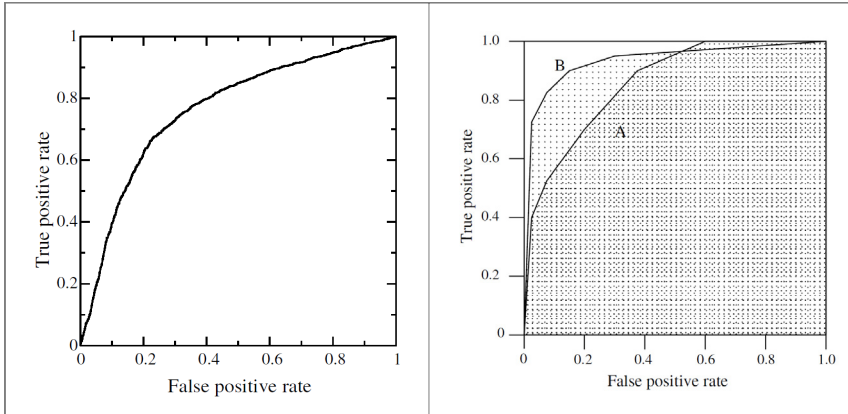
[그림 Ⅲ-1]의 왼쪽 그림은 ROC 곡선의 일반적인 예시를 보여준다. 수직축의 참 양성률(True positive rate = TP/TC)<sup>18)</sup>은 실제 위기 상황을 위험하다고 진단한 비율을 의미한다. 수평축의 거짓 양성률(False positive rate = FP/TNC)은 실제로는 위험 상황이 아닌데 위험하다고 잘못 진단한 비율을 의미한다. 좌측 하단 모서리인 (0, 0)은 어떠한 상황도 전혀 양성으로 분류하지 않는 경우를 의미하며, 우측 상단 모서리인 (1, 1)은 모든 상황을 항상 양성으로 분류하는 경우를 의미한다. 좌측 상단 모서리인 (0, 1)은 위험 상황과 비위험 상황을 완벽하게 분류하는 경우를 의미한다. 어떤 모형에서 임계점을 높게 설정하면 양성을 보수적으로 진단하므로 값이 ROC 곡선의 좌측 하단부에 나타난다. 그리고 임계점을 낮춰 가면 양성 진단이 많아짐과 동시에 참 양성률과 거짓 양성률이 높아지면서 그래프는 우상향하게 된다. 그리고 임계치가 매우 낮아지면 대부분의 사례를 양성으로 진단하므로 참 양성률과 거짓 양성률이 모두 높아져 ROC 곡선의 우측 상단으로 이동하게 된다.

모형을 통해 도출된 ROC 곡선의 전체적인 모습이 좌측 상단 모서리에 접근할수록 더 좋은 예측 성능을 가진 모형이라고 할 수 있다. ROC 곡선의 하단부의 면적을 계산하면 ROC 곡선이 의미하는 모형의 예측 성능을 하나의 숫자로 표현할 수 있으며, 이를 AUC(Area Under the Curve)라고 한다. 일반적으로 AUC가 높은 모형이 낮은 모형보다 예측력이 좋다고 할 수 있다. 하지만 [그림 Ⅲ-1]의 오른쪽 그림에서 보는 것과 같이 AUC가 높은 모형(B)

18) 재현율(recall)과 같으며 민감도(sensitivity)라고도 한다.

이 ROC 곡선의 특정 부분에서는 다른 모형(A)에 비해 예측 성능이 더 낮을 수도 있다. 한편 완전히 임의로 위험을 진단하여 예측력이 없는 모형의 ROC 곡선은 45도선과 같을 것이므로 AUC가 0.5보다 낮은 모형은 예측 모형으로서의 가치가 없다고 할 수 있다.

[그림 III-1] ROC 곡선 예시



자료: Fawcett, Tom, "An introduction to ROC analysis," *Pattern Recognition Letters*, 27(8), 2006, p. 868, Fig. 8. (a); p. 870, Fig. 9. (b).

ROC 곡선이 보여주는 바와 같이 위험진단을 위한 모형의 임계점을 어떻게 설정하느냐에 따라 예측 성능이 달라진다. 특히 앞에서 언급한 모형의 성능을 나타내는 지표 중 하나를 극대화하는 임계치를 선택할 수 있다. 본 연구는 F1 점수를 극대화하는 임계점을 기준으로 분석한다. F1 점수는 정밀도와 재현율을 동시에 고려한다. 즉 위험 신호를 보낼 때 실제 재정위기가 발생할 확률과 실제 위기가 발생할 때 위험 신호를 보낼 확률이 동시에 균형 있게 고려된다. 한편 F1 점수와 함께 ROC 곡선에서 참 양성률과 거짓 양성률의 차이를 극대화하는 값도 임계점 선택으로 많이 고려되는 방법인데, 본 연구에서 두 방법에 따른 임계점이 매우 유사한 것으로 나타났다. 이에 본 연구는 F1 점수를 극대화하는 임계점을 기준으로 분석한다.<sup>19)</sup>

19) F2 점수도 정밀도와 재현율을 동시에 고려하지만, 재현율에 더 가중치를 두는 지표이다.

## 2. 예측 모형 설명 및 예측 결과

### 가. 신호(signaling) 모형

재정위험 진단을 위해 관련 문헌에서 널리 쓰이는 전통적인 방법은 신호(signaling) 모형이다(Kaminsky et al., 1998; Hellwig, 2021; Jarmulska, 2022; Valencia et al., 2022). 이 방법론에서는 재정위험을 진단하는 여러 변수가 사용되며 각 변수는 특정 조건을 만족시키는 경우<sup>20)</sup> 재정위험 신호를 보내는 것으로 간주된다. 변수들의 신호를 합산하여 하나의 지표를 만들면 이는 재정위험을 예측하는 종합적인 지표로서 역할을 하게 된다.

신호 모형을 활용할 때 연구자가 설정해야 할 요소들은 다음과 같다. 우선 어떠한 상황을 재정위기라고 할지 정의해야 한다. 그리고 재정위험을 진단할 때 사용할 변수들을 설정해야 한다. 마지막으로 각 변수가 신호를 보내게 되는 임계점을 설정해야 한다.

변수들의 임계점을 설정할 때 중요하게 고려해야 할 점은 신호가 너무 불필요하게 많이 발생하지 않으면서도 필요한 경우에는 신호가 발생하도록 균형을 잡는 것이다. 이를 위해서 임계점이 신호 대 잡음비(signal to noise ratio)를 극대화하거나 총 오분류 오류(total mis-classification error)를 최소화하도록 설정하는 것이 일반적이다.

신호 대 잡음비는 식 (Ⅲ-1)과 같이 거짓 양성(잡음) 대비 참 양성(신호)의 비율로 정의되며, 이것을 극대화하는 것이 신호의 성능을 높게 된다. 신호가 유의미한 지표가 되기 위해서는 신호 대 잡음비는 1 이상이 되어야 한다. 총 오분류 오류는 식 (Ⅲ-2)와 같이 거짓 음성률과 거짓 양성률의 합으로 정의되며, 최적 신호를 얻기 위해서는 이를 최소화해야 한다. 신호 대 잡음비와 총 오분류 오류 중 어느 것이 선호된다고 단정 짓기는 어렵지만

---

검토 결과 F2 점수를 극대화하는 임계점은 대체로 낮아 재현율을 매우 높이지만 정밀도를 크게 낮추는 것으로 나타났다. 따라서 재정위험 진단을 위한 임계점 설정의 기준으로 F2 점수는 적절하지 않은 것으로 판단된다.

20) 변수가 통상적으로 가지는 값에서 상당히 벗어난 경우를 조건으로 하는 것이 일반적이다(Kaminsky et al., 1998).

본 연구에서는 총 오분류 오류를 사용하도록 한다.

$$SNR = \frac{TP}{TC} \times \left[ \frac{FP}{TNC} \right]^{-1} \quad \text{식 (III-1)}$$

$$TME = \frac{FN}{TC} + \frac{FP}{TNC} \quad \text{식 (III-2)}$$

재정위험은 이러한 각 변수들이 보내는 신호들을 묶어서 하나의 지표를 도출하여 진단한다. 식 (III-1) 또는 식 (III-2)의 기준을 활용하여 변수별로 임계점을 설정하면 각 변수별로 재정위험에 대한 신호를 보내는데, 이를 식 (III-3)과 같이 하나의 지표로 통합하여 최종적인 신호를 통해 재정위험을 진단하는 것이다. 식 (III-3)은 각 국가  $j$ 의 각 기간  $t$ 에 대하여 최종 지표  $L$ 을 산정하는 식이다. 여기서  $z_i$ 는 식 (III-4)에 의해 정의된 신호의 강도(Signal Power, SP)를 나타내며,  $h_{jt}^k$ 는 변수  $k$ 가 국가  $j$ 에서 시간  $t$ 에 관찰될 때 1의 값을 취하는 이진변수이다. 변수  $d_{jt}^i$ 는 변수  $i$ 가 임계점을 넘어서 신호를 보낼 때 1의 값을 가지는 이진변수이다.  $L_{jt}$ 는 재정위험을 나타내는 신호들의 가중된 합이라고 할 수 있다. 이 지표가 0이면 어떠한 변수도 위험 신호를 보내지 않는 저위험 상황을 의미하며, 이 지표가 1이면 모든 변수가 위험 신호를 보내는 고위험 상황이라고 할 수 있다.<sup>21)</sup> 따라서 재정위험을 진단하려면  $L_{jt}$ 에 적절한 임계점을 설정해야 하는데, 앞에서 설명한 것처럼 F1 점수를 극대화하는 값을  $L_{jt}$ 의 임계점으로 도출하여 분석한다.

$$L_{jt} = \sum_{i=1}^n \frac{z_i}{\sum_{k=1}^n h_{jt}^k z_k} d_{jt}^i \quad \text{식 (III-3)}$$

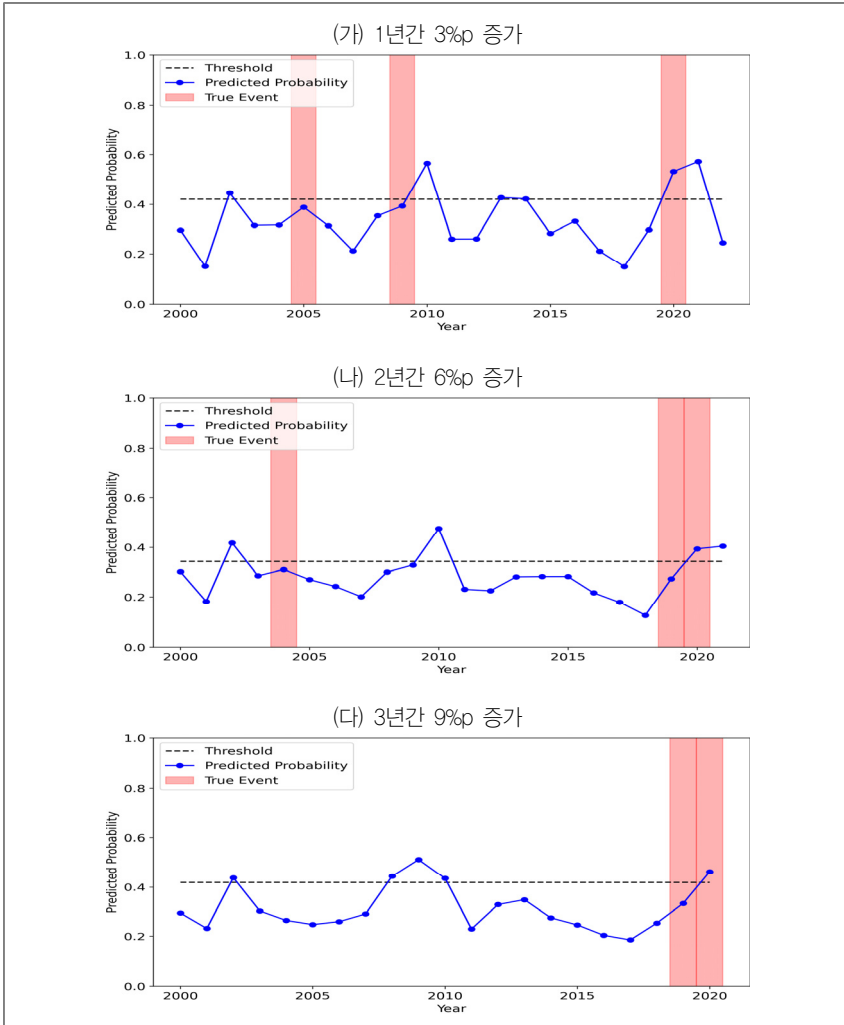
$$z_i = 1 - TME_i \quad \text{식 (III-4)}$$

21) 100을 곱하면 0에서 100의 범위를 가지게 된다.

[그림 Ⅲ-2]는 신호 모형으로 한국의 정부부채 비율이 급증할 확률을 산정하여 나타낸 것이다. [그림 Ⅲ-2]에서 푸른 실선은 재정위기가 발생할 확률을 의미하고, 검은 점선은 임계점이다. 푸른 실선이 검은 점선보다 높은 곳에 위치한 시기가 재정위기가 발생할 가능성이 높다는 경고 신호를 보내는 것으로 이해할 수 있다. 그리고 [그림 Ⅲ-2]에서 붉은색으로 표시한 시점은 본 연구에서 정의한 재정위기가 실제로 발생한 시기이다. [그림 Ⅲ-2(가)]를 보면 신호 모형이 재정위기를 정확히 예측하여 경보를 보낸 시점은 2020년 뿐이며 2005년과 2009년에 대한 경보는 보내지 못한 것으로 나타났다. 반면 2002년, 2010년, 2013년, 2014년, 2021년은 재정위기가 발생하지 않았음에도 재정위기를 겪을 것이라는 잘못된 경보를 보냈다. [그림 Ⅲ-2(나)]와 [그림 Ⅲ-2(다)]에서도 재정위기가 발생한 시기를 정확히 진단한 것은 모두 2020년이 유일하며, 재정위기가 없었던 시기에도 각각 세 차례와 네 차례의 오경보를 보냈다.

[그림 Ⅲ-3]은 신호 모형으로 예측한 선진국의 정부부채 비율 급증 확률의 평균과 실제 정부부채 비율이 급증한 국가들의 비율을 나타낸 것이다. 신호 모형이 보내는 경보를 높은 확률로 보내는 시기는 대체로 선진국들이 실제로 재정위기를 많이 겪은 시기를 후행하는 것으로 나타났다. [그림 Ⅲ-3(가)]에 나타난 바와 같이 선진국 중 많은 국가가 2009년과 2020년에 재정위기를 겪었는데, 신호 모형이 재정위기가 발생할 확률이 높다고 예측한 시기는 2010년과 2021년으로 재정위기 예측에는 한계가 있는 것으로 판단된다. 이러한 경향은 [그림 Ⅲ-3(나)]와 [그림 Ⅲ-3(다)]에도 나타났다.

[그림 III-2] 신호 모형으로 예측한 한국의 정부부채 비율 급증 확률

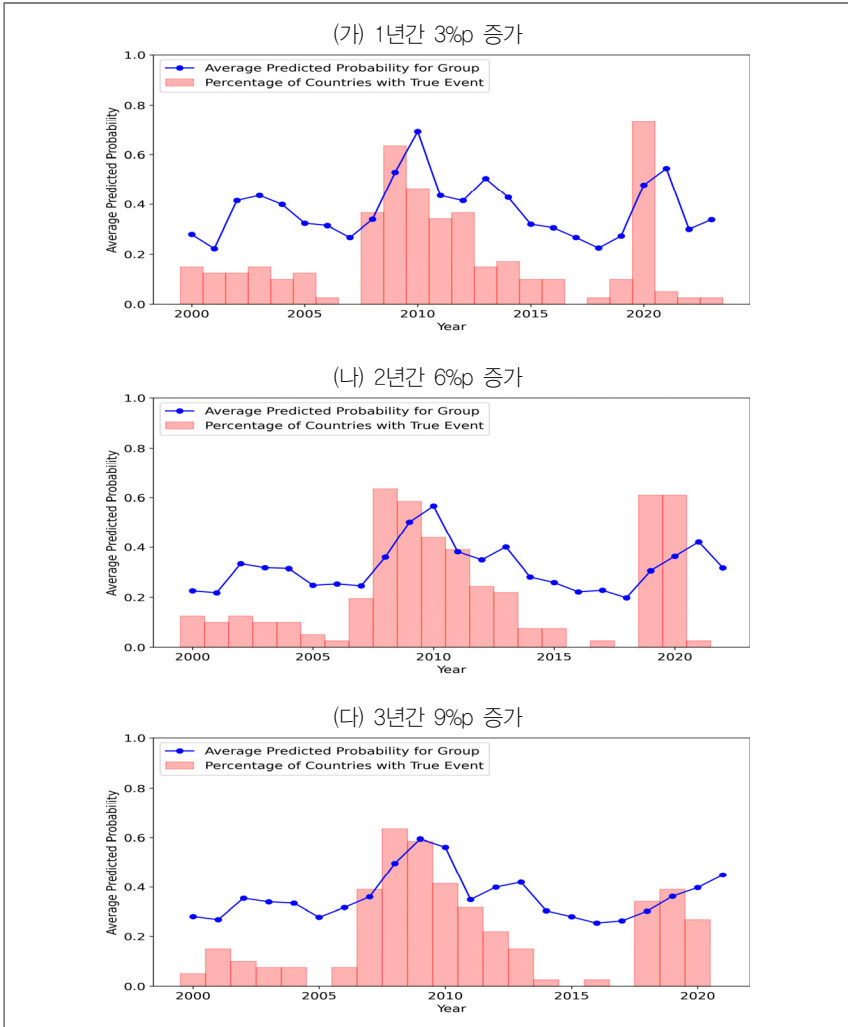


주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증하여 재정위기로 식별된 기간이며, 푸른 실선이 검은 점선보다 높은 시가 모형에 의해 재정위기로 예측된 기간

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&om=LASTNPERIODS&io=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&om=LASTNPERIODS&io=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

[그림 III-3] 신호 모형으로 예측한 선진국의 정부부채 비율 급증 확률



주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증한 국가의 비율이며, 푸른 실선은 모형이 예측한 재정 위기의 발생 확률의 국가 간 평균

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&it\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&it[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.



## 나. 로짓(logit) 모형

재정위험을 진단하는 또 다른 전통적인 모형은 다변량 로짓(logit) 혹은 프로빗(probit) 모형이다(Kaminsky et al., 1998). 이러한 모형은 위기 발생 여부가 이분법적일 때(예: 위기 대 비위기) 다양한 경제지표를 바탕으로 재정 위기 발생 가능성을 추정하는 데 활용될 수 있다.

본 연구에서는 로짓 모형을 기준으로 분석을 수행할 예정이다. 로짓 모형은 다음의 식 (Ⅲ-5)로 표현될 수 있다.  $y$ 는 사건이 발생할 확률을 의미한다. 그리고  $\beta_i X_i$ 는 각 설명변수  $X_i$ 와 계수  $\beta_i$ 의 곱을 의미한다. 이 모형을 기반으로 식 (Ⅲ-6)의 손실함수를 최소화하도록 계수를 추정한다.

$$y = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^N \beta_i X_i)}} \quad \text{식 (Ⅲ-5)}$$

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad \text{식 (Ⅲ-6)}$$

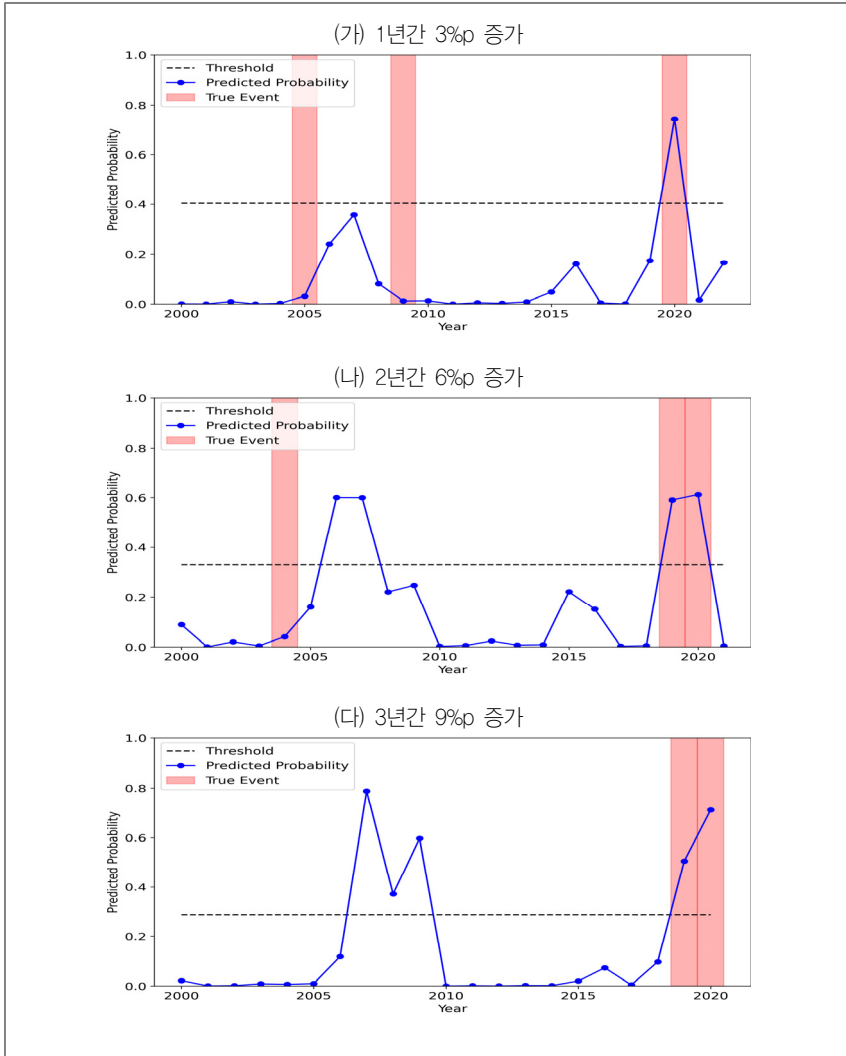
로짓 모형은 여러 설명변수를 동시에 고려하여 재정위기 확률을 추정할 수 있다는 점에서 유용하다. 하지만 이 모형은 재정위험 진단 모형으로서 한계가 있다(Kaminsky et al., 1998). 각 변수 중에서 어떠한 변수들이 참 양성 또는 참 음성을 예측하는 데 유용한지를 판단하기 어렵다는 특징이 있다. 또한 어떤 변수가 예측에 반드시 들어가야 하는지 등에 대해서도 판단을 제공해 주지 않는다는 한계가 있다.

[그림 Ⅲ-4]는 로짓 모형을 이용해 예측한 한국의 GDP 대비 정부부채 비율이 급증할 확률을 나타낸 것이다. [그림 Ⅲ-4(가)]에 나타난 바와 같이 로짓 모형은 1년간 정부부채 비율이 3%p 이상 급증할 위기에 대해 2020년에는 참 양성 신호를 보냈지만, 2005년과 2009년에는 거짓 음성 신호를 보낸 것으로 나타났다. [그림 Ⅲ-4(나)]를 살펴보면 2년 동안 정부부채 비율이 6%p 이상 급증할 위기에 대해서도 로짓 모형은 2019~2020년에는 참 양성

신호를 보냈지만, 2004년에는 거짓 음성 신호를, 그리고 2019~2020년에는 거짓 양성 신호를 보낸 것으로 분석되었다. [그림 Ⅲ-4(다)]를 보면 3년 동안 정부부채 비율이 9%p 이상 증가하는 위기에 대해서 로짓 모형은 2019~2020년에는 재정위기를 잘 진단하였지만, 2007~2009년에는 거짓 양성 신호를 보냈다.

[그림 Ⅲ-5]는 선진국 중 정부부채의 급증을 겪은 국가들의 비율과 로짓 모형으로 계산된 재정위기를 겪을 확률의 평균을 나타낸 것이다. 선진국 중 1~3년간 정부부채 비율이 급증한 국가들의 비율은 로짓 모형으로 예측한 재정위기를 겪을 확률의 평균과 대동소이하다. 다만 국제 금융위기 시기에는 로짓 모형을 통한 진단이 실제 위기보다 후행하는 것이 관찰되었다. [그림 Ⅲ-5-(나)]에서 2022년에 실제 관측된 위기는 없음에도 재정위험이 급증하는 것으로 나타났는데, 이는 2023년에 대한 관측치가 작기 때문인 것으로 판단된다.

[그림 III-4] 로짓 모형으로 예측한 한국의 정부부채 비율 급증 확률

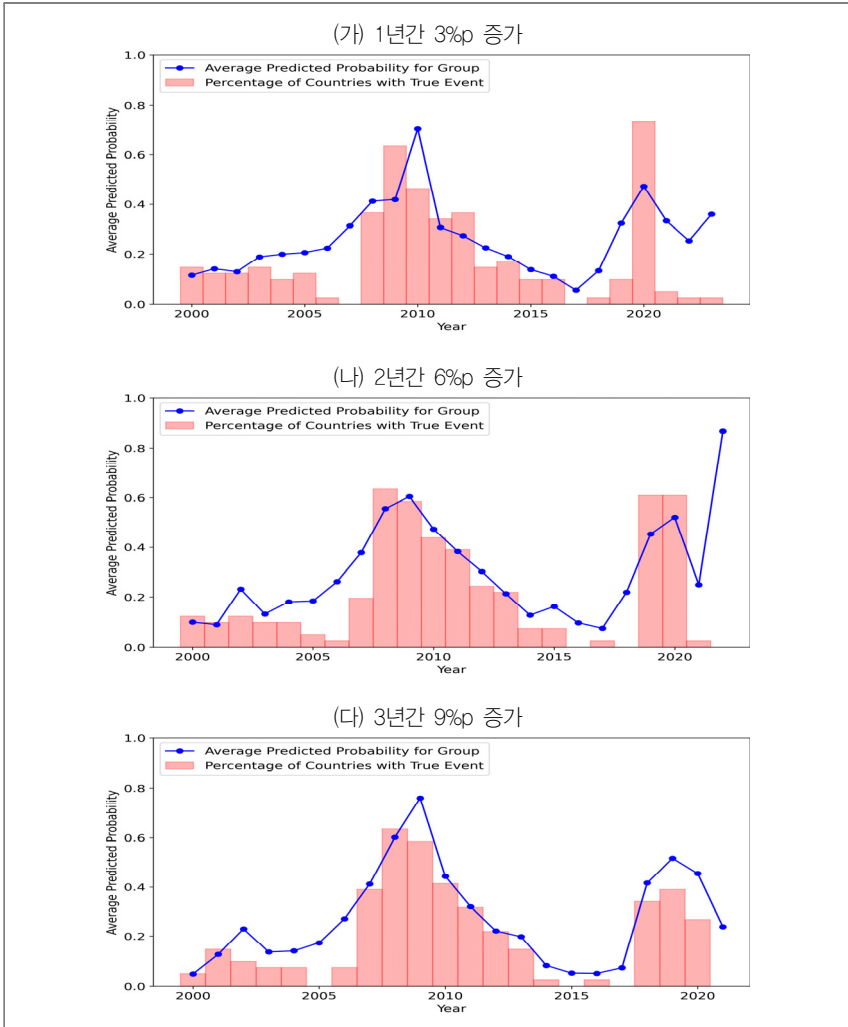


주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증하여 재정위기로 식별된 기간이며, 푸른 실선이 검은 점선보다 높은 시기가 모형에 의해 재정위기로 예측된 기간

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis/?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis/?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

[그림 III-5] 로짓 모형으로 예측한 선진국의 정부부채 비율 급증 확률



주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증한 국가의 비율이며, 푸른 실선은 모형이 예측한 재정 위기의 발생 확률의 국가 간 평균

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&fc=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&it\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&fc=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&it[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

#### 다. 라쏘(LASSO) 모형

재정위험 진단을 위해서 신호 모형이 널리 사용되고 있으나 이 방법론에는 크게 두 가지 취약점이 있다(Hellwig, 2021; Valencia et al., 2022). 첫째로 신호 모형은 각 변수들의 임계점을 정하는 방식이기 때문에 변수 간의 상호작용을 고려하기 어려워 결과의 편향을 가져올 수 있다는 점이다. 두 번째로 신호 모형은 표본 내 예측 성능을 높이는 데 중점을 두기 때문에 표본 외 예측 성능은 크게 다를 수 있다는 점이다.

최근 연구에서는 이러한 약점을 극복하기 위해 기계학습 방법론이 대두되고 있다. 이는 기계학습 방법론은 연구자들이 모두 고려하기 어려운 변수들의 복잡한 상호작용을 포착하는 데 용이하기 때문이며, 또한 기계학습 방법은 예측력을 높이는 방법으로 주로 활용되어 표본 외 예측 성능이 높을 것으로 기대할 수 있기 때문이다.

기계학습 방법론에는 다양한 방법이 있지만 본 연구에서는 라쏘(Last Absolute Shrinkage and Selection Operator, LASSO)와 랜덤포레스트(random forest) 방법론(다음 절에서 언급)을 활용하고자 한다. 라쏘는 선형 회귀와 로짓 모형 등의 회귀분석에서 사용되는 정규화 기법 중의 하나로, 주로 변수 선택과 모형의 복잡성 제어를 위해 사용된다. 본 연구에서는 로짓 모형에 정규화가 적용된 라쏘-로짓 모형을 중심으로 분석을 진행한다.

라쏘-로짓 모형은 기본적으로 앞에서 언급한 로짓 모형과 접근 방법이 유사하다. 다만 핵심적인 차이점은 식 (III-7)과 같이 손실함수에 마지막 항인 L1 정규화 항이 추가되어 있다는 것이다. L1 정규화 항은 회귀분석에서 얻어진 계수들( $\beta_j$ )의 절댓값의 합을 손실함수에 더하는 것으로, 계수의 값이 크면 그만큼 벌칙을 주는 것이다.<sup>22)</sup> 이렇게 정규화 항이 더해지면 어떠한 변수를 모형에 추가하여  $\beta_j$ 를 추정하는 것 자체가 손실함수를 크게 만들기 때문에 설명력이 부족한 변수는 그 계수를 0으로 만들어 모형에서 배제하게 된다. 이러한 특징이 있기에 라쏘를 사용하면 모형에 들어가는 변수를 모형

22) 손실함수에 회귀분석에서 얻어진 계수들  $\beta_j$ 의 제곱의 합인 L2 정규화 항을 추가한 Ridge 모형, L1, L2 정규화 항을 모두 추가한 Elastic net 모형도 활용된다.

자체가 선택하는 역할을 하게 되며, 로짓 모형에 비해 단순화된 모형을 얻을 수 있게 된다.  $\lambda$ 는 정규화 강도를 의미하는 변수로 값이 클수록 더 많은 변수가 0으로 추정되어 모형에서 탈락하게 된다. 본 연구의 라쏘 모형 추정은 식 (Ⅲ-7)과 같이 로짓 모형에 벌칙항을 추가하는 방법으로 수행하였다. 이 과정에서 변수의 절대적인 크기가 변수 선택에 큰 영향을 미치므로 라쏘 모형에서는 변수의 평균을 0으로, 표준편차를 1로 표준화하는 것이 일반적이다. 본 연구 또한 분석 대상 변수들을 이와 같이 표준화하였다.

$$L = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\tilde{y}_i) + (1-y_i) \log(1-\tilde{y}_i)] + \lambda \sum_{j=1}^p |\beta_j| \quad \text{식 (Ⅲ-7)}$$

라쏘 모형에서 최적  $\lambda$ 를 정하기 위해서 일반적으로 사용하는 방법은 교차검증 방법론이다. 이는 데이터의 일부분으로 모형을 훈련한 후 나머지 데이터로 모형의 성능을 검증하여 가장 높은 성능을 가져다주는  $\lambda$ 의 값을 최적  $\lambda$ 로 설정하는 것이다. 이때 모형의 성능을 검증하는 방법은 여러 가지가 있을 수 있으나 정밀도(precision), 재현율(recall), F1 점수, 평균 제곱 오차 (Mean Squared Error, MSE) 등이 활용될 수 있다.

교차검증 방법론 중에서도 가장 많이 활용되는 방법은 K겹 교차검증(K-fold cross-validation)이다. 이는 데이터를 K개로 분할(fold)하여 각 분할에 대해 모형을 훈련하고 검증하는 것이다. 예를 들어, 5겹 교차검증을 사용하면 데이터를 5개 부분으로 나누어 4개의 분할을 훈련데이터로, 1개의 분할을 검증데이터로 사용하며, 이를 모든 분할에 대해 반복한다.<sup>23)</sup> 이렇게 모형을 여러 번 훈련하면서 평균적으로 최고의 성능을 내는 모형을 선택한다. 본 연구에서는 라쏘 모형의  $\lambda$ 선택을 위해서 5겹 교차검증을 사용하였다.

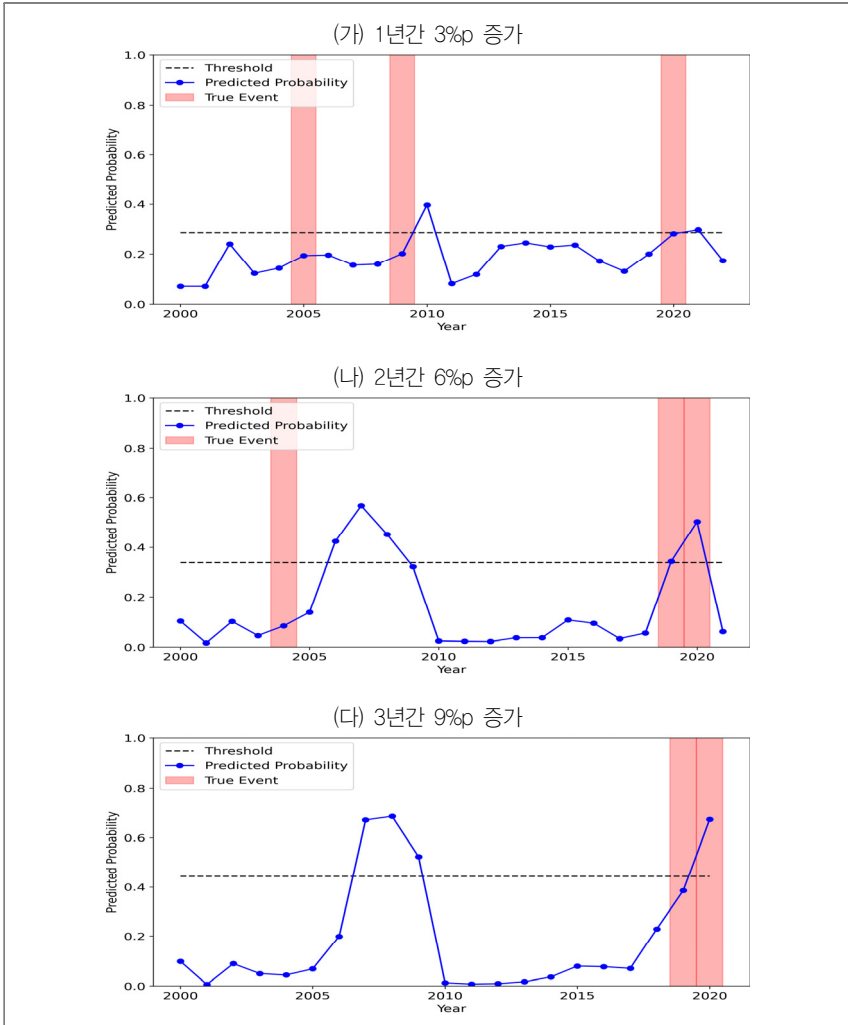
[그림 Ⅲ-6]은 라쏘 모형으로 예측한 한국의 GDP 대비 정부부채 비율이

23) 검증데이터는 모형을 학습시키는 데 직접적으로 활용하지는 않지만, 학습된 모형의 성능을 평가하여 학습 모형에 환류하는 역할을 하므로 학습에 간접적으로 기여한다. 반면 평가데이터는 학습에 전혀 관여하지 않으며 단지 최종 학습된 모형의 성능을 평가하는 데만 활용된다.

급증할 확률을 나타낸 것이다. [그림 Ⅲ-6-(가)]를 보면 1년간 정부부채 비율이 3%p 이상 급증할 확률에 대해서 라쏘 모형은 재정위기 시점을 정확히 진단하지 못하고, 2009년과 2020년 위기는 다음 해에 재정위기 경보를 보내는 것으로 나타났다. [그림 Ⅲ-6-(나)]에 나타난 바와 같이 2년간 정부부채 비율이 6%p 이상 증가하는 사건에 대해서 라쏘 모형은 2019~2020년에는 재정위기를 잘 예측하였다. 하지만 2004년의 위기는 예측하지 못했고, 국제 금융위기 기간과 가까운 2006~2008년에 대해서는 정부부채 비율이 크게 증가하지 않았음에도 재정위기에 대한 거짓 경보를 보냈다. [그림 Ⅲ-6-(다)]를 살펴보면 3년간 정부부채 비율이 9%p 이상 증가하는 사건에 대해서 라쏘 모형은 2020년은 잘 진단하였지만, 2019년에는 거짓 음성 신호를, 그리고 2007~2009년에는 거짓 양성 신호를 보냈다.

[그림 Ⅲ-7]은 라쏘 모형으로 예측한 선진국들의 정부부채 비율 급증 확률과 실제 정부부채 비율이 급증한 국가의 비율을 보여주고 있다. 라쏘 모형은 로짓 모형과 유사한 패턴을 보였다. 진단이 실제 위기를 1년 정도 후행하는 모습이 관찰되었다.

[그림 III-6] 라쏘 모형으로 예측한 한국의 정부부채 비율 급증 확률



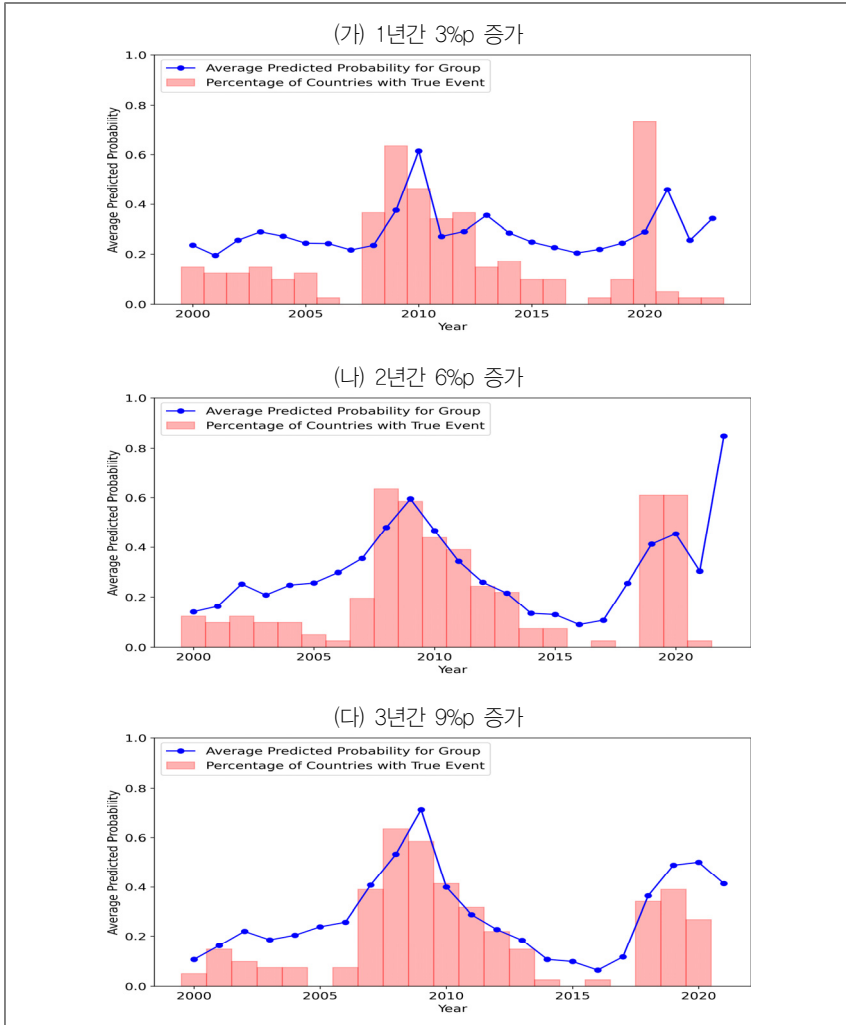
주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증하여 재정위기로 식별된 기간이며, 푸른 실선이 검은 점선보다 높은 시기가 모형에 의해 재정위기로 예측된 기간

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&fc=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&it\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&fc=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&it[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.



[그림 III-7] 라쏘 모형으로 예측한 선진국의 정부부채 비율 급증 확률



주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증한 국가의 비율이며, 푸른 실선은 모형이 예측한 재정 위기의 발생 확률의 국가 간 평균

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&om=LASTNPERIODS&io=5&it\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&om=LASTNPERIODS&io=5&it[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

## 라. 랜덤포레스트(random forest) 모형

랜덤포레스트(random forest)는 앙상블(ensemble) 학습 방법의 하나이다. 이는 분류(classification)와 회귀(regression) 문제에 널리 사용되는 결정나무(decision tree)를 다수 구축하여 그 결과를 종합해 안정적이고 높은 예측 결과를 얻기 위해서 활용된다. 결정나무는 데이터를 반복적으로 분할하면서 조건을 만족하는 하위 집단으로 나누는 알고리즘이다. 각각의 분할은 “예” 또는 “아니오”와 같은 이진 분류 방식으로 이루어진다.<sup>24)</sup> 재정위험을 진단하는 문제를 예로 들면, 첫 번째로 “GDP 대비 정부부채 비율이 90% 이상인가”라는 조건을 설정할 수 있다. 정부부채 비율이 90% 이상일 경우, 해당 국가는 재정위기를 겪을 가능성이 크기 때문에 높은 재정위험 집단으로 분류할 수 있다. 반면 정부부채 비율이 90% 미만일 경우 두 번째 조건으로 “GDP 대비 재정적자 비율이 3% 이상인가”라는 질문을 적용할 수 있다. 이러한 방식으로 각 조건에 따라 국가의 재정 상태를 단계적으로 분석하고, 여러 설명변수를 기반으로 최종적으로 재정위험을 진단한다.

랜덤포레스트는 원본 표본에서 일부의 표본을 추출하여 이를 가장 잘 설명할 수 있는 결정나무를 구축한다. 그리고 또다시 일부의 표본을 추출하여 이를 가장 잘 설명하는 또 다른 결정나무를 구축한다. 이를 부트스트랩 샘플링(bootstrap sampling) 방식이라고 하는데, 이러한 과정을 통해 각 결정나무가 다소 다른 표본에서 학습되도록 하여 다양한 결정나무를 만들어 낸다. 그리고 이처럼 다양하게 만들어진 결정나무를 평균하거나 혹은 결정나무들의 다수결 원칙을 통해 최종 예측 결과를 만들어 내는 것이 랜덤포레스트 방법이다.

이러한 랜덤포레스트는 설명변수와 피설명변수의 선형적 관계를 가정하지 않으므로 비선형 데이터 혹은 변수들의 복잡한 교차효과를 잘 다룰 수 있다는 장점이 있다. 하지만 이러한 장점은 모형을 직관적으로 해석하기 어렵게 한다는 단점으로도 작용한다. 특히 하나의 결정나무가 아닌 많은 수의

---

24) 자세한 설명은 정재현·이환웅(2020), 이환웅 외(2021)를 참고하기 바란다.

결정나무로 구성된 랜덤포레스트 모형을 직관적으로 해석하는 것은 매우 어려운 일이다.

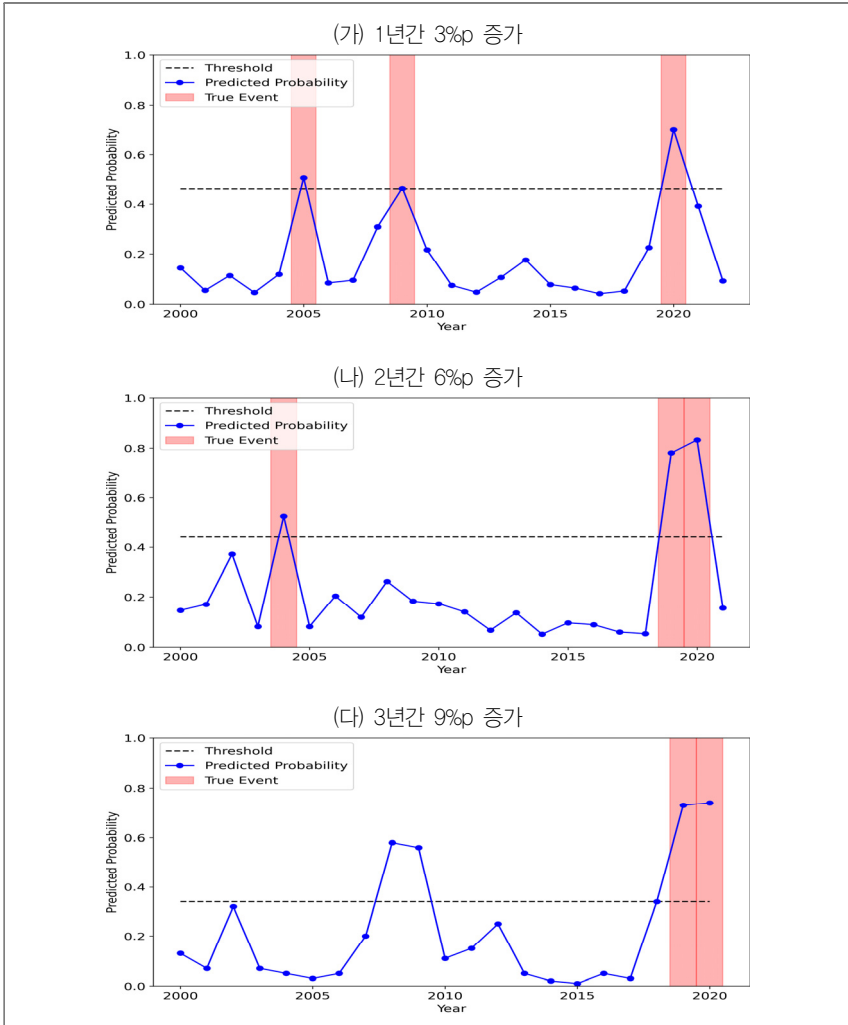
[그림 Ⅲ-8]은 랜덤포레스트 모형을 이용해 예측한 한국의 재정위기 확률을 나타낸 것이다.<sup>25)</sup> [그림 Ⅲ-8-(가)]에 나타난 바와 같이 1년간 정부부채 비율이 3%p 이상 증가하는 사건이 발생한 해에 랜덤포레스트 모형으로 진단한 재정위험도 높아지고 있어 좋은 예측력을 보여주고 있다. 특히 모든 시기에 참 양성 또는 참 음성 신호를 보내는 것으로 나타났다. [그림 Ⅲ-8-(나)]를 살펴보면 2년간 정부부채 비율이 6%p 이상 증가하는 사건에 대해서도 실제 사건이 발생한 기간에 맞춰서 예측 확률이 높아지는 것을 확인할 수 있는데, 이는 다른 모형에서는 발견되지 않는 좋은 예측 결과이다. 여기서도 모두 참 양성 또는 참 음성 신호였다. [그림 Ⅲ-8-(다)]에서 3년간 정부부채 비율이 9%p 이상 증가할 위기에 대해서도 2019~2020년에는 잘 진단하는 모습을 보여주고 있다. 다만 국제 금융위기 기간인 2008~2009년에는 거짓 양성 신호를 보낸 것으로 나타났다.

[그림 Ⅲ-9]은 랜덤포레스트 모형으로 예측한 선진국의 정부부채 비율 급증 확률과 실제 정부부채 비율이 급증한 국가들의 비율을 나타낸 것이다. 랜덤포레스트 모형은 선진국의 정부부채 비율 급증 확률을 잘 예측하고 있으며 로짓이나 라쏘 모형보다도 조금 더 실제 정부부채 비율이 급증한 국가의 비율을 잘 예측한다. 특히 앞에서 논의한 다른 세 개의 모형에서는 진단이 실제 위험을 후행하는 현상이 관찰되었는데, 랜덤포레스트 모형에서는 실제 많은 국가가 위기를 겪은 시기에 재정위험 진단 확률도 높아지는 것으로 나타났다.

---

25) 랜덤포레스트 모형의 조정 매개변수(tuning parameter)로는 부트스트랩 횟수(100, 200), 결정나무 최대 깊이(10, 20, 30), 분할 최소 표본 수(2, 3, 4, 5), 최종 노드 최소 표본 수(1, 2)를 설정하였으며, 이들 간 최적 조합은 scikit-learn 라이브러리의 GridSearchCV 함수를 통해 탐색하였다.

[그림 III-8] 랜덤포레스트 모형으로 예측한 한국의 정부부채 비율 급증 확률

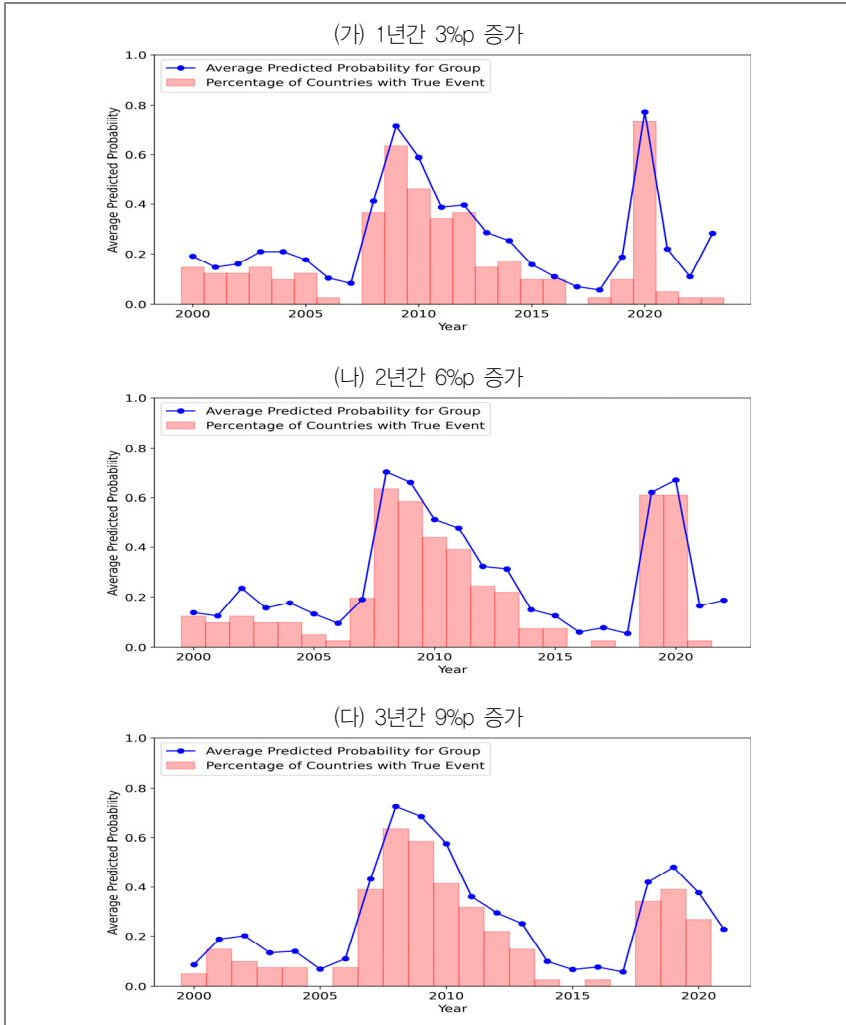


주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증하여 재정위기로 식별된 기간이며, 푸른 실선이 검은 점선보다 높은 시기가 모형에 의해 재정위기로 예측된 기간

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis/?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&it\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis/?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&it[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

[그림 III-9] 랜덤포레스트 모형으로 예측한 선진국의 정부부채 비율 급증 확률



주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증한 국가의 비율이며, 푸른 실선은 모형이 예측한 재정 위기의 발생 확률의 국가 간 평균

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&om=LASTNPERIODS&i=5&i\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&om=LASTNPERIODS&i=5&i[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

### 3. 분석 결과 비교

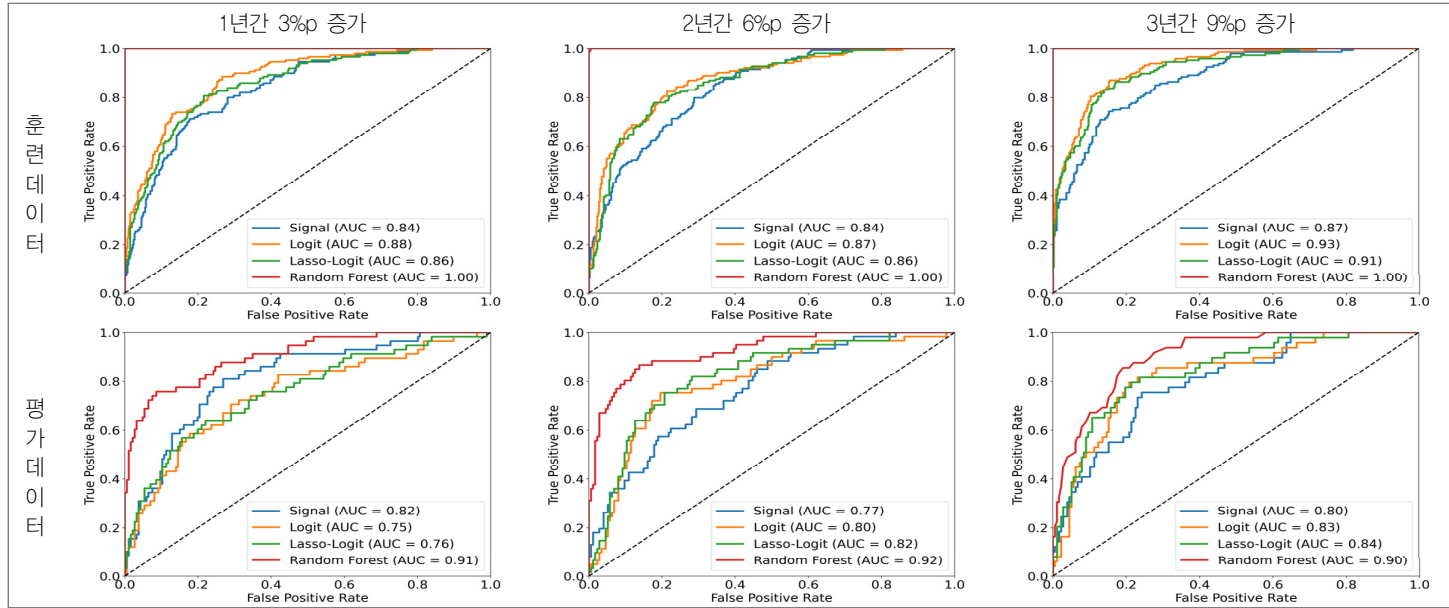
본 절에서는 각 모형의 예측력을 비교하고 각 모형이 어떠한 변수를 예측에 중요한 변수로 활용하고 있는지 확인하고자 한다. [그림 Ⅲ-10]은 GDP 대비 정부부채 비율이 급증할 사건을 예측할 때 각 모형의 예측력을 비교하기 위해 임계점을 다양하게 설정하여 도출한 ROC 곡선과 이 곡선의 아래쪽 면적을 의미하는 AUC를 비교한다. <표 Ⅲ-2>는 각 모형의 예측력을 AUC뿐만 아니라 정확도, 정밀도, 재현율, F1 점수, F2 점수 등의 다양한 지표를 통해 비교한다. 이러한 모형 간 예측력의 비교는 훈련데이터와 평가데이터를 구분하여 비교하는데, 훈련데이터의 예측력은 모형을 구축하는 데 사용된 데이터를 각 모형이 얼마나 잘 설명하는지를 의미하며, 평가데이터의 예측력은 모형을 구축하는 데 사용되지 않는 데이터를 얼마나 잘 설명하는지를 의미한다. 재정위험을 진단하는 모형의 핵심 목적은 아직 관측되지 않은 미래의 재정위기를 관측할 수 있는 변수를 통해 예측하는 것이므로, 모형의 예측력 비교는 훈련데이터보다 평가데이터를 기준으로 이루어지는 것이 적절하다고 판단된다.

[그림 Ⅲ-10]을 통해 각 모형의 ROC 곡선과 AUC를 비교해 보면, 훈련데이터와 평가데이터 모두에서 랜덤포레스트 모형의 예측력이 가장 뛰어난 것으로 나타난다. [그림 Ⅲ-10]의 모든 그림에서 모든 참 양성률에 대해 거짓 양성률을 최소화하는 모형이 랜덤포레스트 모형으로 나타나, ROC 곡선을 기준으로 예측력을 평가할 때 랜덤포레스트 모형이 다른 어떤 모형보다도 모든 상황에서 우월한 것으로 분석되었다. 따라서 각 모형의 AUC를 비교해 보면 [그림 Ⅲ-10]에서 살펴본 여섯 가지의 모든 상황에서 랜덤포레스트 모형의 AUC가 가장 높다. 특히 훈련데이터에서는 1~3년 시계에 대한 예측 모두 AUC가 1.0으로 나타나 재정위험을 완벽하게 진단하는 것으로 분석되었다. 평가데이터에서는 AUC가 0.90~0.92로 훈련데이터보다 낮아지지만, 이 수치 또한 다른 모형과 비교해 높은 수준이다.

랜덤포레스트 모형 다음으로는 대체로 로짓 모형과 라쏘 모형이 유사하게 높은 예측력을 보이고 신호 모형의 예측력이 가장 낮은 것으로 나타난다.

다만 세 모형 간 예측력의 순서가 항상 일정하게 유지되는 것은 아니다. 특히 평가데이터를 통해 1년간 정부부채 비율이 급증할 사건을 예측할 때 대체로 신호 모형이 로짓 모형과 라쏘 모형보다 우월한 성능을 보이는 것으로 분석되었다. 이 경우 신호 모형의 AUC가 0.82로 각각 0.75와 0.76인 로짓과 라쏘 모형의 AUC보다 높다. 또한 신호 모형의 AUC가 로짓이나 라쏘 모형보다 낮더라도 특정 참 양성률에 대해서는 신호 모형의 거짓 양성률이 로짓 모형과 라쏘 모형의 거짓 양성률보다 더 낮은 모습도 관찰된다. 이는 재정 위험 진단 모형으로 매우 높은 수준의 참 양성률을 요구할 때 신호 모형이 로짓이나 라쏘 모형보다 나은 선택일 수도 있음을 의미한다. 한편 훈련데이터에서는 로짓 모형이 라쏘 모형보다 미미하게나마 소폭 높은 예측력을 보이지만 평가데이터에서는 라쏘 모형이 로짓 모형보다 조금 더 높은 예측력을 보인다는 점도 눈에 띈다.

[그림 III-10] 모형별 ROC 곡선 비교(연평균 3%p 증가 기준)



자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?ts\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?ts[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.



재정위험 진단을 위한 모형으로 랜덤포레스트 모형이 다른 모형보다 우월하다는 점은 <표 Ⅲ-2>를 통해서도 확인된다. 랜덤포레스트 모형의 재정위기 예측력을 혼련데이터로 평가하면 모든 지표가 1에 가까워 다른 모형의 예측력 지표보다 월등히 높다. 이는 랜덤포레스트 모형이 혼련데이터에서 위기와 비위기 상황을 거의 완벽하게 분류했다는 것을 의미한다. 랜덤포레스트 모형은 다른 모형과 다르게 변수들의 복잡한 상호작용과 비선형 관계를 학습하는 데 강한 특성이 있으므로 혼련데이터를 잘 분류하는 것으로 이해된다.

평가데이터에 대한 예측력 지표를 살펴보다라도 랜덤포레스트 모형이 다른 모형보다 대체로 높은 예측력을 보인다. 특히 정밀도와 재현율의 조화평균으로 모형의 전반적인 예측력을 평가하는 F1 점수를 살펴보면 랜덤포레스트 모형의 예측력이 가장 우수하다고 평가할 수 있다. 이러한 점은 F1 점수와 유사하지만 재현율에 가중치를 더 많이 부여한 F2 점수에서도 동일하게 유지된다. 정확도와 정밀도, 재현율로 각 모형을 평가하더라도 대체로 랜덤포레스트 모형이 재정위기를 가장 잘 예측하는 것으로 나타난다.

한편 3년간 정부부채 비율이 급증할 사건에 대해 평가데이터로 예측력을 측정할 때 라쏘 모형의 정밀도가 가장 높지만 재현율은 라쏘 모형이 가장 낮은 것으로 나타난다. 높은 정밀도와 낮은 재현율은 재정위기에 대한 경보를 매우 엄격하고 제한적으로만 보낼 때 나타나는 현상이다. 이러한 경보 방식은 실제 재정위기가 발생하더라도 사전에 이를 알리기 위한 경보를 보내지 않는 경우가 빈번하게 발생할 수 있기 때문에 재정위험 진단 모형으로서 적절하다고는 말할 수 없다. 정밀도와 재현율을 종합적으로 평가하는 F1 점수와 F2 점수 등은 랜덤포레스트 모형보다 라쏘 모형에서 더 낮다. 또 다른 종합적인 예측력 평가 지표인 정확도 측면에서 라쏘 모형이 0.84로 가장 높지만 랜덤포레스트 모형의 정확도도 0.82로 라쏘 모형 못지않게 높다.

신호 모형은 라쏘 모형과는 반대로 1년간 정부부채 비율이 급증할 사건을 평가데이터로 예측력을 측정할 때 재현율이 가장 높으면서 동시에 정밀도는 가장 낮은 것으로 나타난다. 높은 재현율과 낮은 정밀도는 재정위기에 대한

경보를 너무 자주 보낼 때 발생할 수 있다. 앞의 라쏘 모형과는 반대로 이 경우에는 재정위기에 대한 경보를 지나치게 빈번하게 보내 신호 모형이 보내는 경보를 자칫 양치기 소년이 보내는 경보로 인식할 우려가 있어 이 또한 재정위험 진단 모형으로 적절하지는 않다. 이러한 맥락에서 높은 재현율에도 불구하고 정밀도와 함께 평가한 F1 점수, F2 점수 등은 신호 모형보다 랜덤포레스트 모형에서 더 높게 평가된 것으로 해석할 수 있다.

로짓 모형은 전반적으로 라쏘 모형과 유사한 재정위기 예측력을 보인다. 특히 1~2년간 정부부채 비율이 급증할 사건을 예측할 때는 로짓 모형과 라쏘 모형의 예측력 지표가 대동소이하다. 이는 본 연구의 라쏘 모형이 로짓 모형을 바탕에 두고 분석하여 예측하는 모형이기 때문으로 해석된다. 다만 3년간 정부부채 비율이 급증할 사건을 예측할 때는 정밀도와 재현율에서 두 모형 간에 차이가 발생한다.

예측 기간별 예측력을 비교해 보면 예측력에 있어 예측 기간 간에 차이가 크지 않은 것으로 나타난다. 특히 예측 시계가 3년으로 길어지더라도 예측력이 크게 떨어지지는 않는 것으로 분석되었다. 1년, 2년, 3년간 정부부채 비율이 급증할 사건에 대한 진단에 대하여 평가데이터로 측정한 랜덤포레스트 모형의 AUC는 각각 0.91, 0.92, 0.90으로 매우 유사하다. 동일한 상황에서의 랜덤포레스트 모형의 F1 점수도 각각 0.74, 0.77, 0.67로 3년간 정부부채 비율이 급증할 사건에 대한 예측력은 조금 낮아지는 것이 관찰되지만 평가데이터로 측정한 다른 모형의 어떤 시계에 대한 F1 점수보다도 높아 예측력이 급격히 낮아졌다고 할 수는 없다. 한편 로짓과 라쏘 모형에서는 오히려 1년에 대한 예측보다 2~3년에 대한 예측 성능이 더 높은 것으로 나타났다.

〈표 III-2〉 모형별 예측력 비교(연평균 3%p 증가 기준)

재정위험 구분	데이터 구분	성능측정 지표	신호	로짓	라쏘	랜덤 포레스트
1년간 3%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.74	0.83	0.81	1.00
		정밀도(precision)	0.51	0.68	0.67	1.00
		재현율(recall)	0.74	0.66	0.58	1.00
		F1 점수	0.60	0.67	0.62	1.00
		F2 점수	0.68	0.66	0.60	1.00
		AUC	0.84	0.88	0.86	1.00
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.76	0.76	0.77	0.88
		정밀도(precision)	0.49	0.50	0.52	0.77
		재현율(recall)	0.76	0.57	0.55	0.71
		F1 점수	0.60	0.53	0.54	0.74
		F2 점수	0.69	0.55	0.55	0.72
		AUC	0.82	0.75	0.76	0.91
2년간 6%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.73	0.80	0.81	1.00
		정밀도(precision)	0.52	0.63	0.64	1.00
		재현율(recall)	0.75	0.73	0.73	0.99
		F1 점수	0.61	0.68	0.68	0.99
		F2 점수	0.69	0.71	0.71	0.99
		AUC	0.84	0.87	0.86	1.00
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.70	0.79	0.78	0.88
		정밀도(precision)	0.45	0.57	0.56	0.78
		재현율(recall)	0.67	0.74	0.74	0.75
		F1 점수	0.54	0.64	0.63	0.77
		F2 점수	0.61	0.70	0.69	0.76
		AUC	0.77	0.80	0.82	0.92
3년간 9%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.81	0.85	0.84	1.00
		정밀도(precision)	0.65	0.68	0.73	0.99
		재현율(recall)	0.72	0.87	0.69	1.00
		F1 점수	0.69	0.76	0.71	1.00
		F2 점수	0.71	0.82	0.70	1.00
		AUC	0.87	0.93	0.91	1.00
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.75	0.79	0.84	0.82
		정밀도(precision)	0.46	0.51	0.62	0.56
		재현율(recall)	0.73	0.78	0.63	0.82
		F1 점수	0.56	0.61	0.63	0.67
		F2 점수	0.65	0.70	0.63	0.75
		AUC	0.80	0.83	0.84	0.90

자료: 다음의 자료를 이용하여 저차 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&io=5&ic\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&io=5&ic[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

〈표 Ⅲ-3〉은 1년 동안 정부부채 비율이 급증할 사건에 영향을 미치는 주요 변수들과 특성을 모아둔 것이다. 신호 모형에서는 총 오분류 비율(TME)이 가장 낮은 변수 10개를 선택하였고, 로짓 모형에서는 통계적 유의도가 가장 높은 변수 10개를, 라쏘 모형에서는 추정된 회귀계수의 절댓값이 가장 큰 변수 10개를, 그리고 랜덤포레스트 모형에서는 변수 중요도(variable importance)가 가장 높은 변수 10개를 모아 두었다.<sup>26)</sup>

예측력이 높은 랜덤포레스트 모형의 경우 정부부채 비율의 급증을 예측하는 데 가장 중요한 변수로 전년도의 잠재 GDP 대비 구조적 재정수지 비율을 꼽았다. 재정수지는 채무 동학을 나타내는 식에 들어가는 정부부채의 규모를 결정하는 직접적인 변수이기 때문에 이러한 결과는 직관에 부합한다. 다만 기초재정수지보다 구조적 재정수지의 중요성이 높다고 평가하는데, 이는 미래를 예측하는 데 있어 구조적 재정수지 적자는 빠르게 조정되기 어렵다는 점을 시사하는 것으로 해석할 수 있다. 한편 랜덤포레스트 모형은 전년도의 정부부채 비율의 증가분도 미래를 예측하는 데 중요한 변수로 꼽았는데, 이는 제Ⅱ장에서 살펴본 바와 같이 정부부채 비율이 급증했을 때 지속성이 있다는 점을 모형이 고려하는 것으로 이해된다.

다른 모형에서 중요하게 선정한 변수들도 살펴보자. 랜덤포레스트 모형과 마찬가지로 기계학습 기법을 이용하는 라쏘 모형 또한 전년도의 잠재 GDP 대비 구조적 재정수지 비율을 가장 중요한 변수로 꼽았다. 반면 기초재정수지와 정부부채 비율은 예측 시 고려하는 변수에서 탈락되었다. 신호 모형의 경우 전년도의 정부부채 비율 증가분을 가장 중요한 변수로 꼽았으며, 그다음으로 전년도의 잠재 GDP 대비 구조적 재정수지 비율과 전년도의 GDP 대비 기초재정수지 비율을 차례로 꼽았다. 즉 랜덤포레스트, 라쏘, 신호 모형은 미래 정부부채 비율의 급증을 예측하는 데 재정과 직접적으로 관련된

26) 랜덤포레스트 모형의 변수 중요도(variable importance)는 모형 전체적으로 특정 독립변수가 결괏값에 미치는 영향을 평가하는 중요한 지표이다. 그러나 총체적 변수 중요도는 모델 전반의 패턴을 파악하는 데 유용하지만, 독립변수가 특정 값을 가지는 경우 그 근방에서의 변수 중요도(local variable importance)는 다르게 나타날 수 있다. 특정 구간에서 변수 중요도는 이후에 나올 부분의존도(Partial Dependence Plot, PDP) 그래프를 통해 시각적으로 확인할 수 있도록 하였다.

변수들을 중요하게 고려하는 것으로 나타났다. 반면 로짓 모형은 선진국의 실업률, 선진국의 소비자물가지수 상승률, 명목 실효환율과 같이 대외환경과 밀접하게 관련된 변수들이 정부부채의 급증을 예측하는 데 높은 설명력이 있다고 분석하여 다른 모형의 결과와는 차이가 있다. 특히 로짓 모형에서 재정과 관련된 변수로 잠재 GDP 대비 구조적 재정수지 비율도 포함되었지만 추정된 회귀계수의 크기가 상대적으로 작다.

〈표 III-3〉 1년간 정부부채 비율 급증에 영향을 미치는 주요 변수

모형	변수명	변수 설명	총 오분류 비율(TME)	신호 모형 방향	로짓 계수	로짓계수 P-통계량	라쏘 계수	랜덤포레스트 변수 중요도
신호 모형	weo_ggx_wdg_ngdp_pd_l1	GDP 대비 정부부채 비율의 차분(전년도)	0.6462	양(+)	-0.0238	0.5152	-	0.0464
	weo_ggsb_npgdp_l1	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.6463	음(-)	-0.2059	0.0024	-0.6283	0.0663
	weo_ggx0_nlb_ngdp_l1	GDP 대비 기초재정수지 비율(전년도)	0.6778	음(-)	-0.0150	0.8294	-	0.0345
	weo_ngdp_r_gw_l1	경제성장률(전년도)	0.7094	음(-)	-0.2608	0.0203	-0.3053	0.0277
	weo_nid_ngdp_pd_l1	GDP 대비 총투자 비율의 차분(전년도)	0.7148	음(-)	-1.4624	0.0003	-0.5610	0.0289
	weo_ggx0_ngdp_pd_l1	GDP 대비 재정지출 비율의 차분(전년도)	0.7149	양(+)	-0.2192	0.1297	-0.1627	0.0199
	weo_ggx0_nlb_ngdp_pd_l1	GDP 대비 기초재정수지 비율의 차분(전년도)	0.7187	음(-)	-0.0582	0.7054	-	0.0184
	weo_tmng_rpch_l1	실질 상품수입 증가율(전년도)	0.7187	음(-)	-0.0618	0.1612	-0.2591	0.0287
	adv_lur_pd_l1	선진국의 실업률의 차분(전년도)	0.7240	양(+)	-2.4495	0.0091	-	0.0208
	adv_ngdp_rpch_l1	선진국의 경제성장률(전년도)	0.7347	음(-)	-0.8672	0.0260	-	0.0181

〈표 III-3〉의 계속

모형	변수명	변수 설명	총 오분류 비율(TME)	신호 모형 방향	로짓 계수	로짓계수 P-통계량	라쏘 계수	랜덤포레스트 변수 중요도
로짓 모형	adv_lur_l1	선진국의 실업률(전년도)	0.8386	음(-)	-1.1554	0.0000	-0.5628	0.0442
	adv_pcpi_pch_l2	선진국의 소비자물가지수 상승률(전전년도)	0.8036	양(+)	1.2643	0.0000	0.5114	0.0119
	oecd_exc_heb_l1	명목 실효환율(전년도)	0.8926	양(+)	0.0447	0.0000	0.3604	0.0162
	weo_nid_ngdp_pd_l1	GDP 대비 총투자 비율의 차분(전년도)	0.7148	음(-)	-1.4624	0.0003	-0.5610	0.0289
	adv_pcpi_pch_l1	선진국의 소비자물가지수 상승률(전년도)	0.8247	양(+)	-0.7898	0.0020	-0.0428	0.0166
	weo_ggsb_npgdp_l1	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.6463	음(-)	-0.2059	0.0024	-0.6283	0.0663
	weo_ngsd_ngdp_pd_l1	GDP 대비 총저축 비율의 차분(전년도)	0.7929	음(-)	1.0128	0.0062	0.0419	0.0117
	weo_bca_ngdp_pd_l1	GDP 대비 경상수지 비율의 차분(전년도)	0.8546	양(+)	-1.0174	0.0086	-	0.0135
	adv_lur_pd_l1	선진국의 실업률의 차분(전년도)	0.7240	양(+)	-2.4495	0.0091	-	0.0208
	oecd_gni_niq_pd_l2	GDP 대비 순이자지출의 차분(전전년도)	0.8382	양(+)	1.1199	0.0111	0.2598	0.0113
라쏘 모형	weo_ggsb_npgdp_l1	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.6463	음(-)	-0.2059	0.0024	-0.6283	0.0663
	adv_lur_l1	선진국의 실업률(전년도)	0.8386	음(-)	-1.1554	0.0000	-0.5628	0.0442
	weo_nid_ngdp_pd_l1	GDP 대비 총투자 비율의 차분(전년도)	0.7148	음(-)	-1.4624	0.0003	-0.5610	0.0289

〈표 III-3〉의 계속

모형	변수명	변수 설명	총 오분류 비율(TME)	신호 모형 방향	로짓 계수	로짓계수 P-통계량	라쏘 계수	랜덤포레스트 변수 중요도
라쏘 모형	adv_pcpi_pch_l2	선진국의 소비자물가지수 상승률(전전년도)	0.8036	양(+)	1.2643	0.0000	0.5114	0.0119
	weo_bca_ngdpd_l1	GDP 대비 경상수지 비율(전년도)	0.8291	음(-)	0.6180	0.0944	-0.3803	0.0204
	oecd_exc_heb_l1	명목 실효환율(전년도)	0.8926	양(+)	0.0447	0.0000	0.3604	0.0162
	weo_nid_ngdp_pd_l2	GDP 대비 총투자 비율의 차분(전전년도)	0.8600	음(-)	-0.1845	0.6275	-0.3471	0.0114
	weo_ngdp_r_gw_l1	경제성장률(전년도)	0.7094	음(-)	-0.2608	0.0203	-0.3053	0.0277
	weo_pcpi_gw_l2	소비자물가지수 상승률(전전년도)	0.8099	양(+)	-0.3442	0.0266	-0.2963	0.0145
	oecd_gni_ntq_pd_l2	GDP 대비 순이자지출의 차분(전전년도)	0.8382	양(+)	1.1199	0.0111	0.2598	0.0113
랜덤포레스트 모형	weo_ggsb_npgdp_l1	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.6463	음(-)	-0.2059	0.0024	-0.6283	0.0663
	weo_ggx_wdg_ngdp_pd_l1	GDP 대비 정부부채 비율의 차분(전년도)	0.6462	양(+)	-0.0238	0.5152	-	0.0464
	adv_lur_l1	선진국의 실업률(전년도)	0.8386	음(-)	-1.1554	0.0000	-0.5628	0.0442
	weo_ggxo_nlb_ngdp_l1	GDP 대비 기초재정수지 비율(전년도)	0.6778	음(-)	-0.0150	0.8294	-	0.0345
	weo_nid_ngdp_pd_l1	GDP 대비 총투자 비율의 차분(전년도)	0.7148	음(-)	-1.4624	0.0003	-0.5610	0.0289
	weo_tmgrpch_l1	실질 상품수입 증가율(전년도)	0.7187	음(-)	-0.0618	0.1612	-0.2591	0.0287
	weo_ngdp_r_gw_l1	경제성장률(전년도)	0.7094	음(-)	-0.2608	0.0203	-0.3053	0.0277

〈표 III-3〉의 계속

모형	변수명	변수 설명	총 오분류 비율(TME)	신호 모형 방향	로짓 계수	로짓계수 P-통계량	라쏘 계수	랜덤포레스트 변수 중요도
랜덤 포레스트 모형	oecd_irs_pd_l2	단기이자율의 차분(전전년도)	0.7541	양(+)	0.0464	0.7881	0.0747	0.0246
	weo_lur_pd_l1	실업률의 차분(전년도)	0.7676	양(+)	0.1695	0.4710	0.1368	0.0209
	adv_lur_pd_l1	선진국의 실업률의 차분(전년도)	0.7240	양(+)	-2.4495	0.0091	-	0.0208

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lo=LASTNPERIODS&lo=5&td\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lo=LASTNPERIODS&lo=5&td[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

[그림 III-11]은 랜덤포레스트 모형으로 1년간 정부부채 비율이 급증할 사건을 예측할 때 주요 변수들의 부분의존도를 보여준다. [그림 III-11]을 보면 이러한 변수들은 특정 수치까지는 큰 영향이 없다가 특정 수치보다 낮아지거나 높아지게 되면 급격하게 영향력이 커지는 것을 확인할 수 있는데, 이는 변수들의 비선형성을 잘 나타내는 것으로 해석할 수 있다.

더 나아가 부분의존도 그림은 특정한 임계점에 대한 정보를 제공하므로, 어떤 변수들을 특정 임계점 내에서 잘 관리해야 한다는 시사점도 제공할 수 있다. 가령 [그림 III-11]의 첫 번째 그림은 향후 GDP 대비 정부부채 비율이 급증하지 않도록 하기 위해서는 당장 올해의 정부부채 비율의 증가폭을 잘 관리할 필요가 있다는 점을 보여준다. 이는 GDP 대비 정부부채 비율이 급증하면 이러한 현상이 단기간에 그치는 것이 아니라 지속성을 갖고 다년간에 걸쳐 정부부채 비율의 증가폭이 높은 수준을 유지한다는 제II장의 분석 결과와도 일맥상통한다. 이뿐만 아니라 GDP 대비 구조적 재정수지 및 기초 재정수지 비율의 수준이나 전년 대비 변화폭, GDP 대비 정부지출 비율과



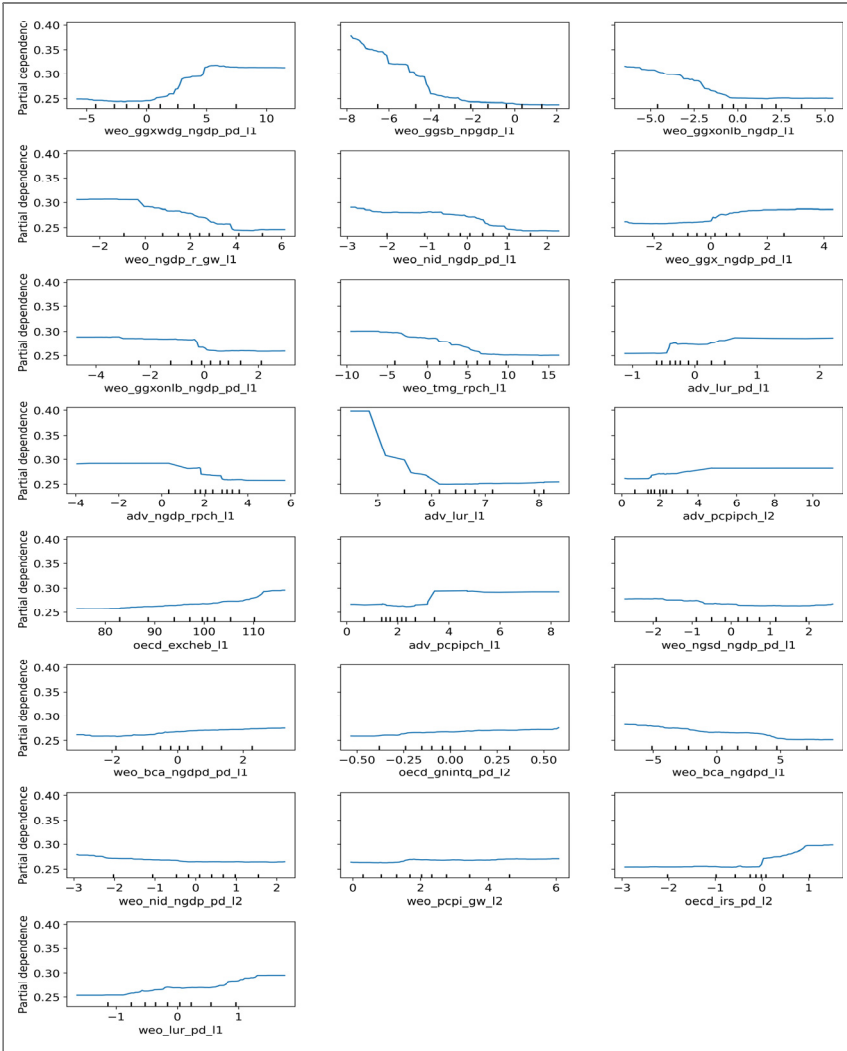
같은 재정변수들도 크게 악화되지 않도록 관리할 필요가 있음을 [그림 Ⅲ-11]에 제시된 부분의존도 그림을 통해 알 수 있다.

대외환경, 경제성장률, 이자율, 환율 등과 같이 정부가 직접적으로 통제하기 어려운 변수들도 있는데, 이러한 변수들의 경우 특정 임계점을 넘어서었을 때 정부부채 비율이 급증할 우려가 있어 주의해야 한다는 시사점도 얻을 수 있다. 즉 이러한 변수들이 정부부채 비율을 급증시킬 우려가 있는 지점에 있을 때 본 연구에서 정의한 재정위기의 관점에서만 보면 정부가 통제할 수 있는 재정변수들을 더 주의해서 관리해야 함을 의미한다.

다만 정부부채 비율의 급증을 촉발시킬 수 있는 각 변수의 임계점은 국가마다 다를 수 있는데, 본 연구의 분석은 이러한 점을 반영하지 못하는 한계가 있다. 따라서 부분의존도에 나타난 각 변수의 임계점이 우리나라의 임계점으로 해석하는 것은 곤란하고, 본 연구의 분석 대상인 32개 선진국의 평균 임계점으로 해석하는 것이 바람직하다.

한편 선진국의 전년도 실업률이 비교적 낮은 수준일 때 부분의존도가 급격히 커져 직관에는 부합하지 않는 현상이 관찰된다. 이는 국제 금융위기가 닥쳤을 때 GDP 대비 정부부채 비율과 선진국의 실업률이 동시에 급증하였는데, 당시 선진국의 전년도 실업률은 비교적 낮은 상황이었다는 점이 반영된 것으로 해석된다.

[그림 III-11] 1년간 정부부채 비율 급증에 영향을 미치는 변수들의 부분의존도



주: 변수명에 대한 설명은 <부록> 참고

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD,ECO,MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&lom=LASTNPERIODS&o=5&ic\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD,ECO,MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&lom=LASTNPERIODS&o=5&ic[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

〈표 Ⅲ-4〉는 2년간 정부부채 비율이 급증할 사건을 예측하는 데 영향을 주는 변수들을 보여주고 있다. 랜덤포레스트 모형은 전년도의 선진국의 실업률을 2년간 정부부채 비율이 급증할 사건을 예측하는 데 가장 중요한 변수로 뽑았다. 선진국의 실업률은 〈표 Ⅲ-3〉에서도 재정위험을 진단하는 데 세 번째로 중요하다고 뽑았던 변수이다. 또한 랜덤포레스트 모형뿐만 아니라 다른 모형도 이 변수를 중요한 변수로 뽑고 있다. 즉 정부부채 비율을 예측하는 데 대외환경도 중요함을 시사한다. 선진국의 실업률 다음으로 중요한 변수로는 재정변수인 구조적 재정수지, 정부부채 비율 증가분, 기초재정수지를 뽑았다. 신호, 로짓, 라쏘 모형도 대체로 재정 및 대외환경과 관련된 변수가 정부부채 비율의 급증을 예측하는 데 중요하다고 말하고 있다.

〈표 Ⅲ-5〉는 3년간 정부부채 비율이 급증할 사건을 예측하는 데 영향을 주는 변수들을 보여주고 있다. 이 경우에는 네 가지 모형 모두 전년도의 선진국의 실업률 변수가 재정위험 진단에 가장 중요한 변수로 뽑았다. 이뿐만 아니라 상대적으로 먼 시계를 예측할 때는 경상수지, 무역량, 환율 등 대외환경과 관련된 변수가 더욱 중요해지는 것으로 판단된다.

[그림 Ⅲ-12]와 [그림 Ⅲ-13]은 각각 2년과 3년간 정부부채 비율이 급증할 사건을 랜덤포레스트 모형으로 예측할 때 주요 변수들의 부분의존도를 보여 준다. 주요 시사점은 [그림 Ⅲ-11]에 나타난 바와 유사하다.

〈표 III-4〉 2년간 정부부채 비율 급증에 영향을 미치는 주요 변수

모형	변수명	변수 설명	총 오분류 비율(TME)	신호 모형 방향	로짓 계수	로짓계수 P-통계량	라쏘 계수	랜덤포레 스트 변수 중요도
신호 모형	weo_ggsb _npgdp_11	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.6934	음(-)	-0.2405	0.0003	-0.5095	0.0559
	adv_lur_11	선진국의 실업률 (전년도)	0.7044	음(-)	-1.3567	0.0000	-0.8353	0.0737
	weo_ggx wdg_ngdp _pd_11	GDP 대비 정부부채 비율의 차분(전년도)	0.7104	양(+)	0.0162	0.6852	0.0630	0.0380
	oecd_irs _pd_12	단기이자율의 차분(전전년도)	0.7195	양(+)	-0.0167	0.9270	0.0465	0.0198
	weo_bca_ ngdpd_11	GDP 대비 경상수지 비율 (전년도)	0.7438	음(-)	0.2098	0.5990	-0.3225	0.0274
	adv_pcpi pch_12	선진국의 소비자물가지수 상승률(전전년도)	0.7554	양(+)	0.4025	0.1654	0.1473	0.0196
	oecd_gni ntq_pd_11	GDP 대비 순이자지출의 차분(전년도)	0.7629	양(+)	0.3961	0.3216	0.0423	0.0190
	weo_ggx0 nlb_ngdp _pd_11	GDP 대비 기초재정수지 비율의 차분(전년도)	0.7709	음(-)	-0.1590	0.2471	-	0.0108
	weo_ggx0 nlb_ngdp _11	GDP 대비 기초재정수지 비율(전년도)	0.7710	음(-)	-0.0373	0.5800	-0.0507	0.0347
	weo_ngsd _ngdp_11	GDP 대비 총저축 비율(전년도)	0.7746	음(-)	-0.2762	0.4907	-	0.0162
로짓 모형	adv_lur_11	선진국의 실업률 (전년도)	0.7044	음(-)	-1.3567	0.0000	-0.8353	0.0737
	oecd_exc heb_11	명목 실효환율 (전년도)	0.7936	양(+)	0.0684	0.0000	0.6480	0.0283
	weo_ggsb _npgdp_11	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.6934	음(-)	-0.2405	0.0003	-0.5095	0.0559
	wld_trade pch_12	전 세계의 실질 무역량 증가율 (전전년도)	0.8024	양(+)	0.2949	0.0009	0.6734	0.0161

〈표 III-4〉의 계속

모형	변수명	변수 설명	총 오분류 비율(TME)	신호 모형 방향	로짓 계수	로짓계수 P-통계량	라쏘 계수	랜덤포레스트 변수 중요도
로짓 모형	adv_ngdp_rpch_l2	선진국의 경제성장률(전전년도)	0.8539	양(+)	-1.1836	0.0028	-0.9415	0.0077
	oecd_irs_l1	단기이자율(전년도)	0.8409	양(+)	0.2867	0.0088	0.2197	0.0156
	adv_lur_pd_l1	선진국의 실업률의 차분(전년도)	0.8145	양(+)	-1.6664	0.0558	-0.0282	0.0118
	weo_tmgrpch_l1	실질 상품수입 증가율(전년도)	0.8025	음(-)	-0.0687	0.0811	-0.2652	0.0168
	oecd_gnitq_l1	GDP 대비 순이자 지출(전년도)	0.9353	음(-)	-0.1770	0.1007	-0.2056	0.0127
	weo_ggxonlb_ngdp_pd_l2	GDP 대비 기초 재정수지 비율의 차분(전전년도)	0.8778	음(-)	-0.1917	0.1157	-0.0971	0.0127
라쏘 모형	adv_ngdp_rpch_l2	선진국의 경제성장률(전전년도)	0.8539	양(+)	-1.1836	0.0028	-0.9415	0.0077
	adv_lur_l1	선진국의 실업률(전년도)	0.7044	음(-)	-1.3567	0.0000	-0.8353	0.0737
	wld_trade_pch_l2	전 세계의 실질 무역량 증가율(전전년도)	0.8024	양(+)	0.2949	0.0009	0.6734	0.0161
	oecd_exc_heb_l1	명목 실효환율(전년도)	0.7936	양(+)	0.0684	0.0000	0.6480	0.0283
	weo_ggsb_npgdp_l1	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.6934	음(-)	-0.2405	0.0003	-0.5095	0.0559
	weo_nid_ngdp_pd_l2	GDP 대비 총투자 비율의 차분(전전년도)	0.8145	음(-)	-0.0570	0.8884	-0.4095	0.0132
	weo_nid_ngdp_pd_l1	GDP 대비 총투자 비율의 차분(전년도)	0.7928	음(-)	-0.5528	0.2156	-0.3395	0.0180
	weo_bcan_gdp_l1	GDP 대비 경상수지 비율(전년도)	0.7438	음(-)	0.2098	0.5990	-0.3225	0.0274

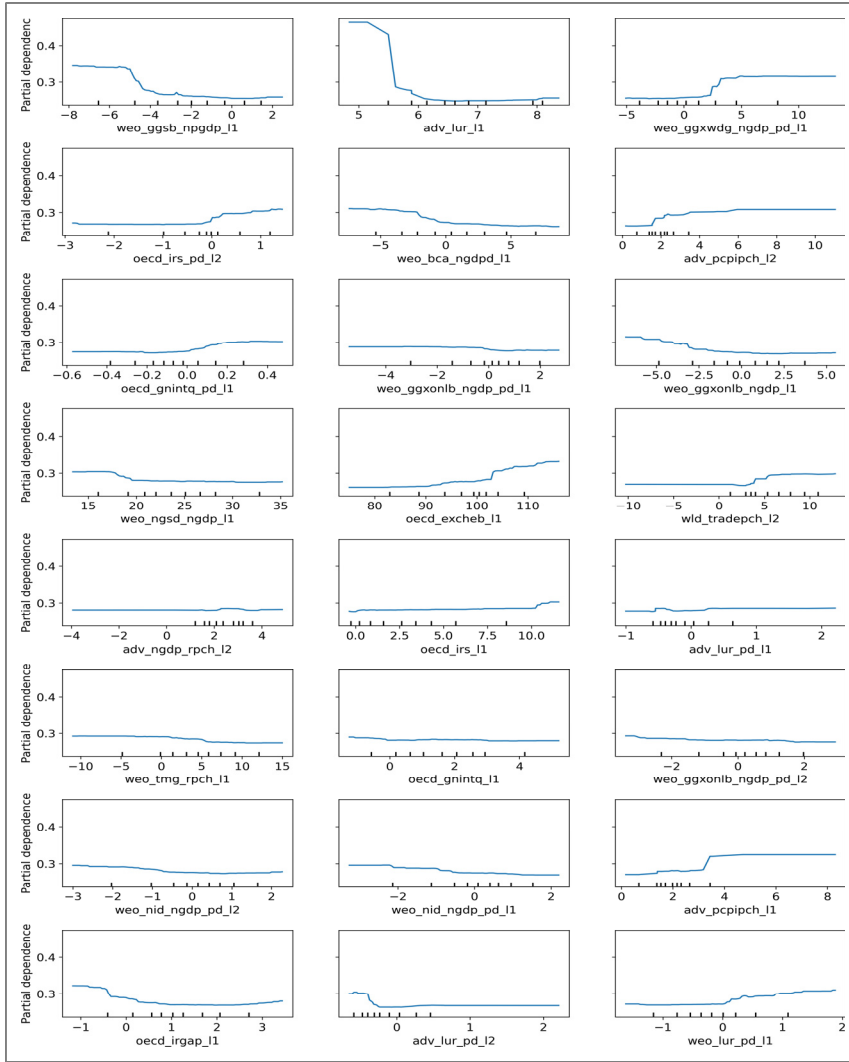
〈표 III-4〉의 계속

모형	변수명	변수 설명	총 오분류 비율(TME)	신호 모형 방향	로짓 계수	로짓계수 P-통계량	라쏘 계수	랜덤포레스트 변수 중요도
라쏘 모형	adv_pcpi_pch_11	선진국의 소비자물가지수 상승률(전년도)	0.8205	양(+)	0.0987	0.7309	0.3059	0.0243
	weo_tmgrpch_11	실질 상품수입 증가율(전년도)	0.8025	음(-)	-0.0687	0.0811	-0.2652	0.0168
랜덤포레스트 모형	adv_lur_11	선진국의 실업률(전년도)	0.7044	음(-)	-1.3567	0.0000	-0.8353	0.0737
	weo_ggsb_npgdp_11	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.6934	음(-)	-0.2405	0.0003	-0.5095	0.0559
	weo_ggxwdg_ngdp_pd_11	GDP 대비 정부부채 비율의 차분(전년도)	0.7104	양(+)	0.0162	0.6852	0.0630	0.0380
	weo_ggxonlb_ngdp_11	GDP 대비 기초재정수지 비율(전년도)	0.7710	음(-)	-0.0373	0.5800	-0.0507	0.0347
	oecd_exc_heb_11	명목 실효환율(전년도)	0.7936	양(+)	0.0684	0.0000	0.6480	0.0283
	weo_bcan_gdpdp_11	GDP 대비 경상수지 비율(전년도)	0.7438	음(-)	0.2098	0.5990	-0.3225	0.0274
	oecd_irgap_11	장단기 금리차(전년도)	0.7783	음(-)	0.2030	0.1953	-	0.0270
	adv_lur_pd_12	선진국의 실업률의 차분(전년도)	0.7884	음(-)	0.4892	0.5465	-	0.0251
	adv_pcpi_pch_11	선진국의 소비자물가지수 상승률(전년도)	0.8205	양(+)	0.0987	0.7309	0.3059	0.0243
	weo_lur_pd_11	실업률의 차분(전년도)	0.7917	양(+)	-0.0934	0.6713	0.0178	0.0243

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&fc=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&lom=LASTNPERIODS&io=5&ic\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&fc=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&lom=LASTNPERIODS&io=5&ic[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

[그림 III-12] 2년간 정부부채 비율 급증에 영향을 미치는 변수들의 부분의존도



주: 변수명에 대한 설명은 <부록> 참고  
 자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/viz/?ts\[0\]=Topic%2C1%7CEconomic%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=.GDPV\\_ANNPCT.A&lom=LASTNPERIODS&o=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/viz/?ts[0]=Topic%2C1%7CEconomic%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=.GDPV_ANNPCT.A&lom=LASTNPERIODS&o=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

〈표 III-5〉 3년간 정부부채 비율 급증에 영향을 미치는 주요 변수

모형	변수명	변수 설명	총 오분류 비율(TME)	신호 모형 방향	로짓 계수	로짓계수 P-통계량	라쏘 계수	랜덤포레스트 변수 중요도
신호 모형	adv_lur_l1	선진국의 실업률(전년도)	0.5855	음(-)	-1.7879	0.0000	-1.1306	0.0725
	oecd_irs_pd_l2	단기이자율의 차분(전전년도)	0.6830	양(+)	0.0021	0.9927	0.2578	0.0230
	weo_bca_ngdpd_l1	GDP 대비 경상수지 비율(전년도)	0.6886	음(-)	0.0342	0.9365	-0.6314	0.0412
	adv_pcpi_pch_l2	선진국의 소비자물가지수 상승률(전전년도)	0.7138	양(+)	-0.2546	0.5106	0.0808	0.0158
	adv_lur_pd_l2	선진국의 실업률의 차분(전년도)	0.7339	음(-)	-1.7875	0.0699	-	0.0292
	oecd_irs_pd_l1	단기이자율의 차분(전년도)	0.7385	양(+)	0.4947	0.0589	0.2566	0.0309
	oecd_irgap_l1	장단기 금리차(전년도)	0.7432	음(-)	-0.2621	0.1362	-0.1474	0.0313
	weo_usex_l1	대미국달러 환율(전년도)	0.7470	양(+)	1.2362	0.0092	0.3299	0.0291
	weo_ggxo_nlb_ngdp_l1	GDP 대비 기초재정수지 비율(전년도)	0.7543	음(-)	-0.0644	0.4757	-0.0847	0.0336
	weo_ngsd_ngdp_l1	GDP 대비 총저축 비율(전년도)	0.7545	음(-)	-0.2112	0.6241	-	0.0238
로짓 모형	adv_lur_l1	선진국의 실업률(전년도)	0.5855	음(-)	-1.7879	0.0000	-1.1306	0.0725
	oecd_exc_heb_l1	명목 실효환율(전년도)	0.8183	양(+)	0.0572	0.0000	0.4842	0.0246
	wld_trade_pch_l2	전 세계의 실질 무역량 증가율(전전년도)	0.8081	양(+)	0.5221	0.0001	0.5130	0.0099
	oecd_irgap_pd_l1	장단기 금리차의 차분(전년도)	0.9025	양(+)	1.0612	0.0002	0.6310	0.0164
	adv_ngdp_rpch_l2	선진국의 경제성장률(전전년도)	0.8231	양(+)	-2.1073	0.0003	-0.6101	0.0056
	adv_pcpi_pch_l1	선진국의 소비자물가지수 상승률(전년도)	0.7746	양(+)	1.3132	0.0026	0.6828	0.0191



〈표 III-5〉의 계속

모형	변수명	변수 설명	총 오분류 비율(TME)	신호 모형 방향	로짓 계수	로짓계수 P-통계량	라쏘 계수	랜덤포레스트 변수 중요도
로짓 모형	weo_use_x_1	대미국달러 환율 (전년도)	0.7470	양(+)	1,2362	0.0092	0.3299	0.0291
	weo_lur_pd_12	실업률의 차분 (전전년도)	0.9246	양(+)	0.8025	0.0095	0.2091	0.0127
	weo_ggx_ngdp_pd_1	GDP 대비 재정지출 비율의 차분(전년도)	0.7563	양(+)	0.4346	0.0142	0.2423	0.0163
	weo_ggsb_ngdp_1	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.7572	음(-)	-0.2218	0.0153	-0.4829	0.0276
라쏘 모형	adv_lur_1	선진국의 실업률 (전년도)	0.5855	음(-)	-1.7879	0.0000	-1.1306	0.0725
	adv_pcpi_pch_1	선진국의 소비자물가지수 상승률(전년도)	0.7746	양(+)	1.3132	0.0026	0.6828	0.0191
	weo_bca_ngdpd_1	GDP 대비 경상수지 비율(전년도)	0.6886	음(-)	0.0342	0.9365	-0.6314	0.0412
	oecd_irg_ap_pd_1	장단기 금리차의 차분(전년도)	0.9025	양(+)	1.0612	0.0002	0.6310	0.0164
	adv_ngdp_rpch_12	선진국의 경제성장률 (전전년도)	0.8231	양(+)	-2.1073	0.0003	-0.6101	0.0056
	wld_trad_epch_12	전 세계의 실질 무역량 증가율 (전전년도)	0.8081	양(+)	0.5221	0.0001	0.5130	0.0099
	oecd_ex_cheb_1	명목 실효환율 (전년도)	0.8183	양(+)	0.0572	0.0000	0.4842	0.0246
	weo_ggsb_ngdp_1	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.7572	음(-)	-0.2218	0.0153	-0.4829	0.0276
	weo_usex_1	대미국달러 환율 (전년도)	0.7470	양(+)	1,2362	0.0092	0.3299	0.0291
	weo_nid_ngdp_pd_12	GDP 대비 총투자 비율의 차분(전전년도)	0.9160	음(-)	-0.9933	0.0294	-0.3208	0.0134

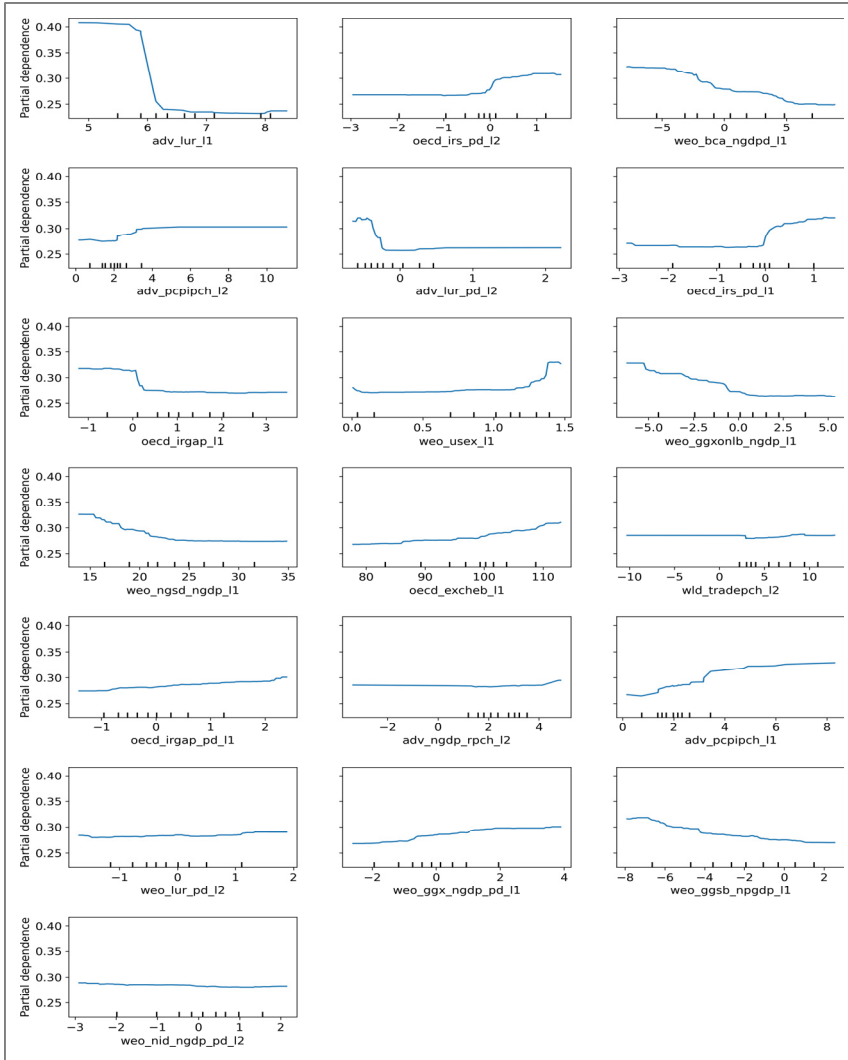
〈표 III-5〉의 계속

모형	변수명	변수 설명	총 오분류 비율(TME)	신호 모형 방향	로짓 계수	로짓계수 P-통계량	라쏘 계수	랜덤포레스트 변수 중요도
랜덤 포레스트 모형	adv_lur_l1	선진국의 실업률(전년도)	0.5855	음(-)	-1.7879	0.0000	-1.1306	0.0725
	weo_bca_ngdpd_l1	GDP 대비 경상수지 비율(전년도)	0.6886	음(-)	0.0342	0.9365	-0.6314	0.0412
	weo_ggxo_nlb_ngdp_l1	GDP 대비 기초재정수지 비율(전년도)	0.7543	음(-)	-0.0644	0.4757	-0.0847	0.0336
	oecd_irgap_l1	장단기 금리차(전년도)	0.7432	음(-)	-0.2621	0.1362	-0.1474	0.0313
	oecd_irs_pd_l1	단기이자율의 차분(전년도)	0.7385	양(+)	0.4947	0.0589	0.2566	0.0309
	adv_lur_pd_l2	선진국의 실업률의 차분(전년도)	0.7339	음(-)	-1.7875	0.0699	-	0.0292
	weo_usex_l1	대미국달러 환율(전년도)	0.7470	양(+)	1.2362	0.0092	0.3299	0.0291
	weo_ggsb_npgdp_l1	잠재 GDP 대비 구조적 재정수지 비율(전년도)	0.7572	음(-)	-0.2218	0.0153	-0.4829	0.0276
	oecd_exc_heb_l1	명목 실효환율(전년도)	0.8183	양(+)	0.0572	0.0000	0.4842	0.0246
	weo_ngsd_ngdp_l1	GDP 대비 총저축 비율(전년도)	0.7545	음(-)	-0.2112	0.6241	-	0.0238

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=TTopic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&Iom=LASTNPERIODS&io=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=TTopic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&Iom=LASTNPERIODS&io=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

[그림 III-13] 3년간 정부부채 비율 급증에 영향을 미치는 변수들의 부분의존도



주: 변수명에 대한 설명은 <부록> 참고

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/viz?ts\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=.GDPV\\_ANNPCT.A&om=LASTNPERIODS&o=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/viz?ts[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=.GDPV_ANNPCT.A&om=LASTNPERIODS&o=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

## 4. 민감도 분석

본 절에서는 제Ⅱ장에서와 같이 세 가지 민감도 분석을 수행한다. 기본 분석에서는 재정위기를 정부부채 비율이 연평균 3%p 이상 증가한 시기로 정의하여 모형들의 예측 결과를 비교하였다. 하지만 재정위기가 다른 방식으로 정의될 때도 모형들의 예측력 비교가 유사한 결과를 나타낼 것이라고 확신하기는 어렵다. 따라서 본 절에서는 GDP 대비 정부부채 비율이 연평균 2%p와 4%p 이상 증가하는 시기를 재정위기로 재정의하는 민감도 분석을 실시한다. 그리고 추가적으로 재정위기를 정의하기 위한 연평균 GDP 대비 정부부채 비율의 증가폭에 대한 임계점을 기본 분석과 같이 3%p로 유지하되 32개 선진국뿐만 아니라 17개 신흥국을 함께 포함한 모형을 분석하도록 한다.

### 가. 정부부채 비율 급증 기준 하향

[그림 Ⅲ-14]와 [그림 Ⅲ-15]는 재정위기를 정부부채 비율이 연평균 2%p 이상 증가한 시기로 정의했을 때 각 모형이 예측한 한국과 선진국의 재정위기 발생 확률과 실제 재정위기 기간을 나타낸 그림이다. 기본 모형에 비해서 재정위기를 정의하는 정부부채 비율의 증가폭이 작아졌기 때문에 실제 위기로 규정된 기간이 많은 것을 확인할 수 있으며 예측 확률도 그에 따라 대체로 더 높아진 것을 볼 수 있다.

기본 분석과 마찬가지로 [그림 Ⅲ-14]와 [그림 Ⅲ-15]에서 랜덤포레스트 모형이 다른 모형들보다 재정위기 시점을 가장 잘 예측하는 것으로 관찰된다. 먼저 [그림 Ⅲ-14]를 통해 우리나라의 재정위기 예측력을 살펴보자. 1년간 정부부채 비율이 급증할 사건에 대해서 랜덤포레스트 모형의 양성 신호는 모두 참이었다. 다만 위기 때 세 번의 거짓 음성 신호도 보내 재정위기를 정확히 예측하지 못하기도 한다는 점이 관찰되었다. 라쏘 모형의 경우 거짓 음성은 한 번으로 랜덤포레스트 모형보다 이 부분에서는 더 좋은 예측력을 보이나, 그만큼 거짓 양성 신호도 다섯 차례 보내 전체적인 위험 신호의 신

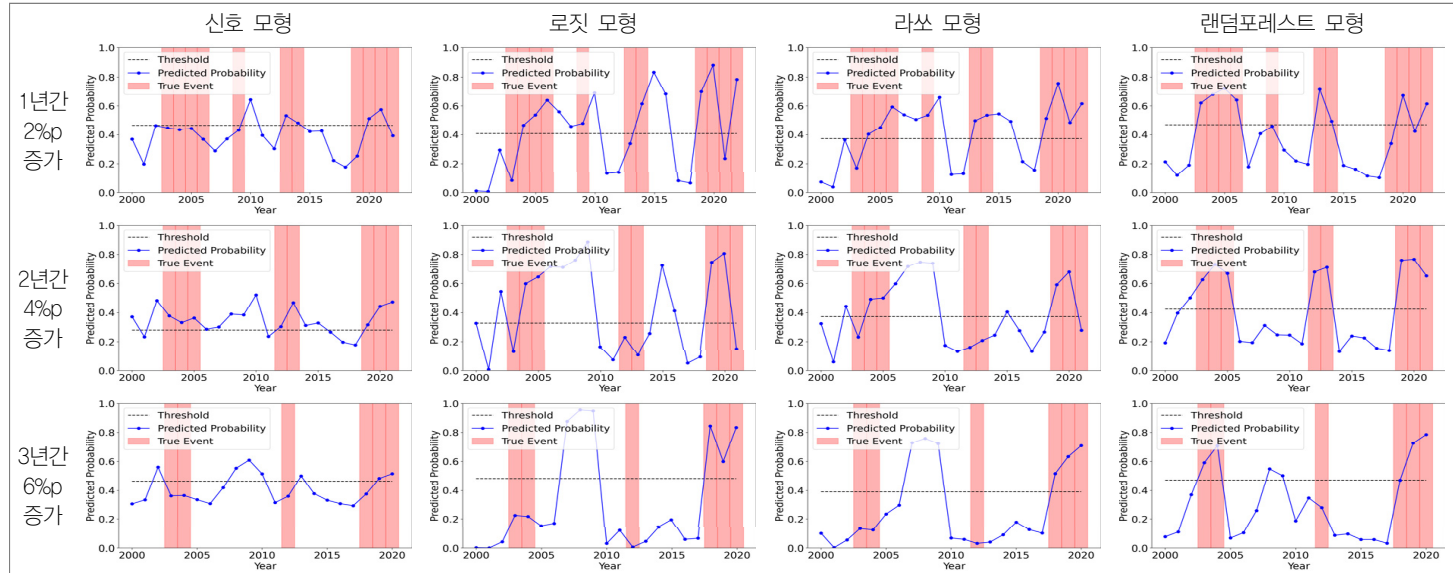
뢰성 측면에서는 랜덤포레스트 모형보다 낮은 것으로 판단된다. 로짓 모형은 세 차례의 위기를 예측하지 못하면서 다수의 거짓 양성 신호를 보냈고, 신호 모형은 다수의 위기를 예측하지 못하였다. 2년간 정부부채 비율이 급증할 사건에 대해서도 랜덤포레스트 모형은 단 한 차례의 거짓 양성 신호를 제외하면 모두 위기와 비위기를 정확히 예측하였는데, 다른 모형에서는 다수의 오신호가 관측되었다. 3년간 정부부채 비율이 급증할 사건에 대해서도 랜덤포레스트 모형은 2012년의 위기를 예측하지 못했고 2008~2009년을 위기가 아님에도 위기로 예측하였는데, 이는 다른 모형에서도 동일하게 관측되면서 다른 모형들은 그 외에도 더 많은 오신호를 보냈다. 다음으로 [그림 III-15]를 통해 선진국 전체의 재정위기 예측력을 비교해 보자. 이 경우에도 랜덤포레스트 모형으로 예측한 재정위기의 확률이 선진국 중 실제로 재정위기를 겪은 국가의 비중과 가장 유사한 모습을 볼 수 있다.

[그림 III-16]에 나타난 모습도 기본 분석과 유사하다. 훈련데이터와 평가데이터의 ROC 곡선 모두 랜덤포레스트 모형이 다른 모형보다 우월하다는 것을 보여준다. 한편 AUC를 살펴보면 훈련데이터에서는 라쏘 모형보다 로짓 모형의 예측력이 소폭 높지만 평가데이터에서는 그러한 현상이 관찰되지 않으며, 오히려 라쏘 모형이 소폭 높은 경우도 관찰된다. 또한 신호 모형의 AUC는 훈련데이터에서는 네 모형 중 가장 낮지만 평가데이터에서는 그렇지 않다는 점도 관찰된다. 특히 1년간 정부부채 비율이 급증할 사건에 대한 신호 모형의 AUC는 로짓과 라쏘 모형보다도 높다. 이러한 점은 모두 기본 분석에서도 유사하게 관측된 것이다.

〈표 III-6〉은 완화된 재정위기 기준에 대한 모형별 예측력을 보여주는데, 이 또한 기본 분석의 결과와 유사하게 전반적으로 랜덤포레스트 모형이 가장 높은 예측력을 보여주는 것으로 나타난다. 훈련데이터에서는 다른 모형과 달리 랜덤포레스트 모형의 예측력이 1에 가깝다. 평가데이터에서도 정확도와 정밀도, F1 점수, 그리고 AUC에서는 랜덤포레스트 모형이 가장 높다. 2년과 3년간 정부부채 비율이 급증할 사건에 대한 예측에서 신호 모형의 재현율이 가장 높은 것으로 나타나는데, 이는 정밀도에서 큰 손해를 보면서

달성된 수치이다. 즉 신호 모형의 경우 양성 신호를 남발하면서 재현율은 높였지만 거짓 양성도 많아 정밀도가 낮아진 것으로 다른 모형과 비교해 예측 모형으로 바람직한 것은 아니다. 이는 ROC 곡선에서 알 수 있듯이 랜덤 포레스트 모형도 신호 모형과 같은 재현율을 제약조건으로 걸면 더 높은 정밀도를 달성할 수 있기 때문이다.

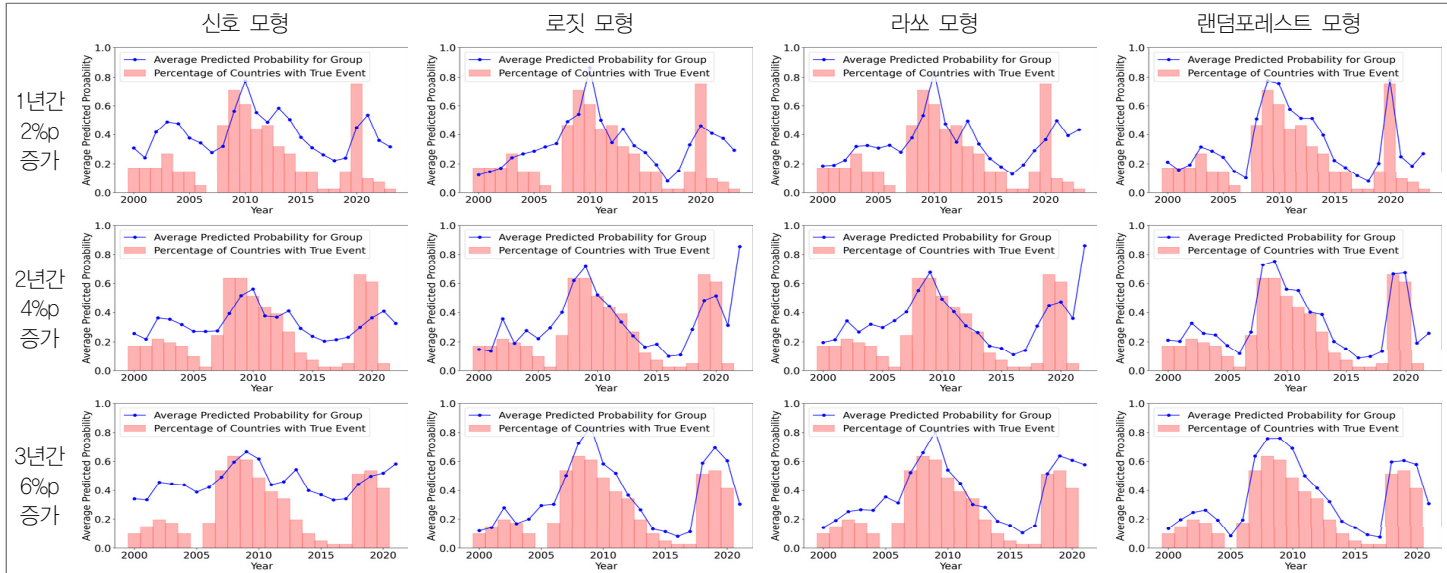
[그림 III-14] 모형별 한국의 정부부채 비율 급증 확률(연평균 2%p 증가 기준)



주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증하여 재정위기로 식별된 기간이며, 푸른 실선이 검은 점선보다 높은 시기가 모형에 의해 재정위기로 예측된 기간  
 자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?ts\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23EOO\\_OUT%23&pg=0&fc=Topic&bp=true&nb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&lom=LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?ts[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23EOO_OUT%23&pg=0&fc=Topic&bp=true&nb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&lom=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

[그림 III-15] 모형별 선진국의 정부부채 비율 급증 확률(연평균 2%p 증가 기준)

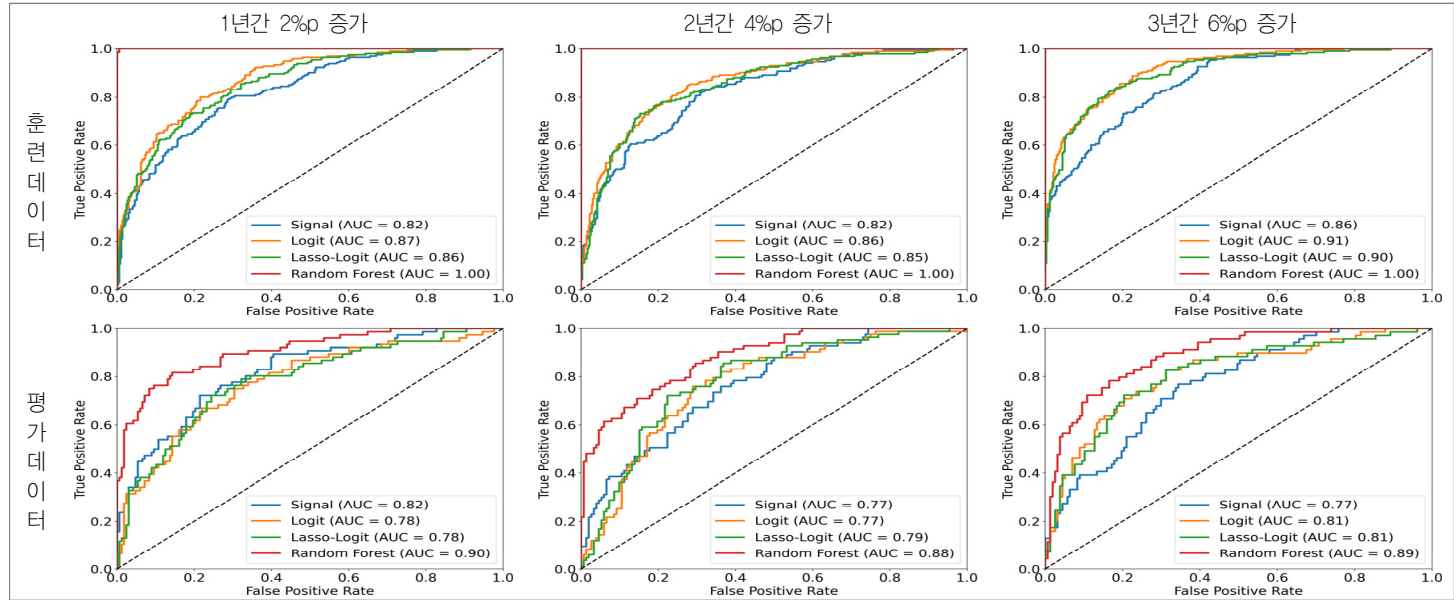


주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증한 국가의 비율이며, 푸른 실선은 모형이 예측한 재정위기의 발생 확률의 국가 간 평균  
 자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?ts\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lm=-LASTPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?ts[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lm=-LASTPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.



[그림 III-16] 모형별 ROC 곡선 비교(연평균 2%p 증가 기준)



자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&fc=Topic&bp=true&nb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&lm=-LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&fc=Topic&bp=true&nb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&lm=-LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

〈표 III-6〉 모형별 예측 성능 비교(연평균 2%p 증가 기준)

재정위험 구분	데이터 구분	성능측정 지표	신호	로짓	라쏘	랜덤 포레스트
1년간 2%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.75	0.79	0.78	0.99
		정밀도(precision)	0.62	0.70	0.67	1.00
		재현율(recall)	0.71	0.70	0.73	0.98
		F1 점수	0.66	0.70	0.70	0.99
		F2 점수	0.69	0.70	0.72	0.99
		AUC	0.82	0.87	0.86	1.00
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.76	0.74	0.74	0.86
		정밀도(precision)	0.60	0.57	0.57	0.80
		재현율(recall)	0.71	0.66	0.71	0.74
		F1 점수	0.65	0.61	0.63	0.77
		F2 점수	0.69	0.64	0.68	0.75
		AUC	0.82	0.78	0.78	0.90
2년간 4%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.63	0.78	0.80	1.00
		정밀도(precision)	0.48	0.65	0.69	1.00
		재현율(recall)	0.89	0.78	0.74	1.00
		F1 점수	0.63	0.71	0.71	1.00
		F2 점수	0.76	0.75	0.73	1.00
		AUC	0.82	0.86	0.85	1.00
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.63	0.72	0.75	0.80
		정밀도(precision)	0.49	0.58	0.63	0.72
		재현율(recall)	0.87	0.75	0.71	0.70
		F1 점수	0.62	0.66	0.67	0.71
		F2 점수	0.75	0.71	0.69	0.70
		AUC	0.77	0.77	0.79	0.88
3년간 6%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.75	0.84	0.82	1.00
		정밀도(precision)	0.62	0.79	0.72	1.00
		재현율(recall)	0.81	0.75	0.83	1.00
		F1 점수	0.70	0.77	0.77	1.00
		F2 점수	0.76	0.76	0.81	1.00
		AUC	0.86	0.91	0.90	1.00
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.69	0.77	0.77	0.84
		정밀도(precision)	0.49	0.60	0.60	0.74
		재현율(recall)	0.75	0.70	0.71	0.71
		F1 점수	0.59	0.64	0.65	0.73
		F2 점수	0.68	0.67	0.69	0.72
		AUC	0.77	0.81	0.81	0.89

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=.GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=.GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

## 나. 정부부채 비율 급증 기준 상황

[그림 Ⅲ-17]과 [그림 Ⅲ-18]은 재정위기를 GDP 대비 정부부채 비율이 연평균 4%p 이상 증가한 시기로 정의했을 때 각 모형이 예측한 재정위기 발생 확률과 실제 재정위기 기간을 나타낸 그림이다. 이러한 정의에 따르면 기본 분석과 비교해 재정위기를 정의하는 정부부채 비율의 증가폭이 커졌기 때문에 실제로 재정위기가 발생한 시기의 수가 확실히 줄어든다. 특히 한국에서는 3년간 정부부채 비율이 12%p 이상 증가한 시기를 찾아볼 수 없다.

한국의 사례를 보면 재정위기를 더 관대하게 정의한 경우에도 네 가지 모형 중 랜덤포레스트 모형의 예측력이 가장 높은 것으로 관찰된다. 1년과 2년간 정부부채 비율이 급증할 사건에 대해서 랜덤포레스트 모형은 단 한 차례의 거짓 양성 신호만 보냈고 나머지는 모두 참 양성 또는 참 음성 신호였다. 반면 다른 모형은 랜덤포레스트 모형보다 더 많은 오신호를 보냈으며, 특히 신호 모형과 라쏘 모형은 위기임에도 위기라고 예측하지 못한 사례도 관찰된다. 3년간 정부부채 비율이 급증할 사건에 대해서는 모든 모형이 공통적으로 2008~2009년에 거짓 양성 신호를 보냈다. 이는 국제 금융위기 기간 우리나라의 정부부채 비율이 본 분석이 정의한 재정위기의 기준만큼 급증하지는 않았으나 대외변수의 움직임으로 모든 모형에서 유사하게 위기가 올 것이라는 신호를 보낸 것으로 해석된다. 한편 랜덤포레스트 모형과 달리 다른 모형들은 2020년에 대해서도 거짓 양성 신호를 보냈다.

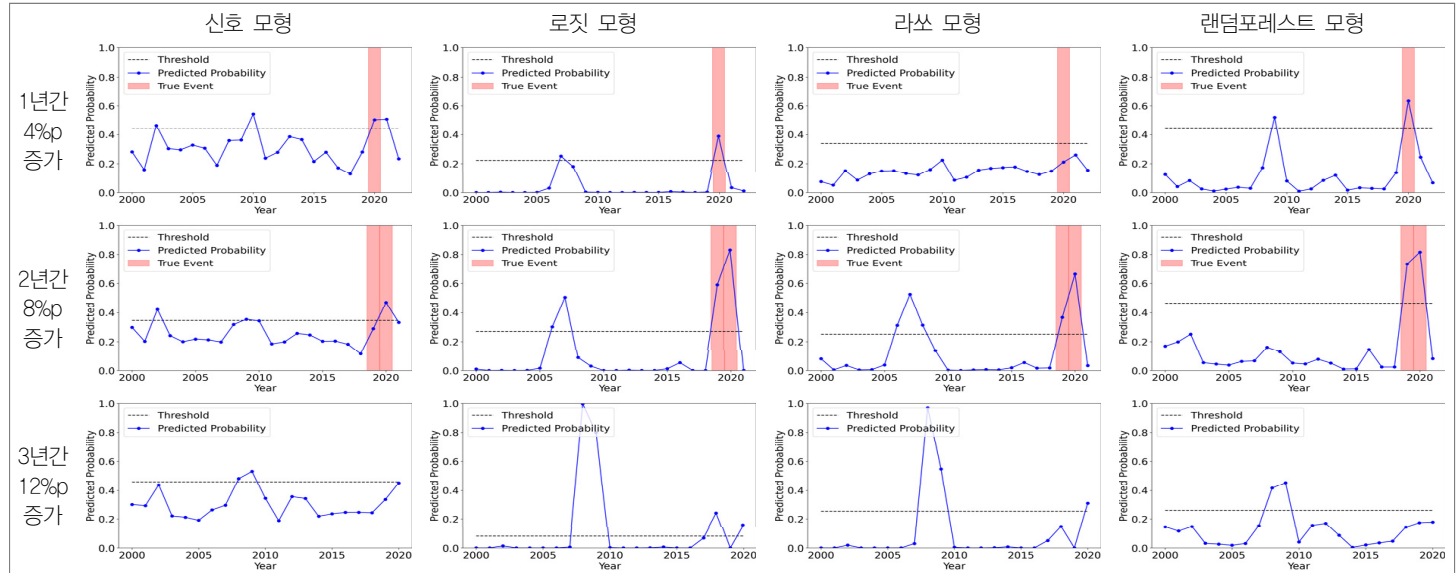
선진국의 평균적인 위기 예측을 살펴보다라도 랜덤포레스트 모형의 재정위기 예측 확률이 선진국 중 실제 위기를 겪은 국가의 비중과 가장 유사한 모습으로 움직이는 것을 확인할 수 있다. 신호 모형 등 다른 모형에서는 예측 확률의 최고점이 실제 위기를 겪은 선진국의 비중을 다소 후행하는 모습이지만 랜덤포레스트 모형에서는 이러한 모습이 훨씬 덜 관찰된다.

ROC 곡선과 AUC를 통해 살펴본 모형 간 예측력에 대해서도 기본 분석과 유사하게 랜덤포레스트 모형의 예측력이 가장 높게 나타난다. ROC 곡선에 나타나듯이 대부분의 경우에 랜덤포레스트 모형이 다른 모형보다 우월하다. 다만 평가데이터에서 3년간 정부부채 비율이 급증할 사건을 예측할 때

낮은 참 양성률에서 랜덤포레스트 모형과 신호 모형이 유사한 거짓 양성률을 보이는 구간이 존재하는데, 예측 모형의 특성상 일정 수준 이상의 높은 참 양성률을 보이도록 설정할 필요가 있어 이를 고려하면 이 구간에서의 신호 모형의 좋은 예측력은 모형 선택을 위한 고려 시 크게 중요하다고 판단하기는 어렵다.

〈표 Ⅲ-7〉은 강화된 정부부채 기준에 따른 모형별 재정위기의 예측력을 다양한 지표를 통해 보여주고 있다. 모형 간 예측력의 비교는 전반적으로 기본 분석과 정성적으로 유사하다. 훈련데이터에서 랜덤포레스트 모형은 모든 지표에서 가장 높은 예측력을 보여주고 있으며, 평가데이터에서도 대부분의 지표에서 랜덤포레스트 모형의 예측력이 가장 높다. 특히 평가데이터를 이용한 측정 지표로 1년과 2년간 정부부채 비율이 급증할 사건을 예측할 때의 재현율을 제외한 나머지 지표에서 모두 랜덤포레스트 모형의 예측력이 가장 높다. 랜덤포레스트 모형의 재현율이 낮은 경우에는 다른 모형과 비교해 랜덤포레스트 모형의 정밀도가 월등히 높는데, 이로 인해 이 두 지표를 동시에 고려하는 F1 점수 등에서 랜덤포레스트 모형이 높게 측정된 것으로 해석된다.

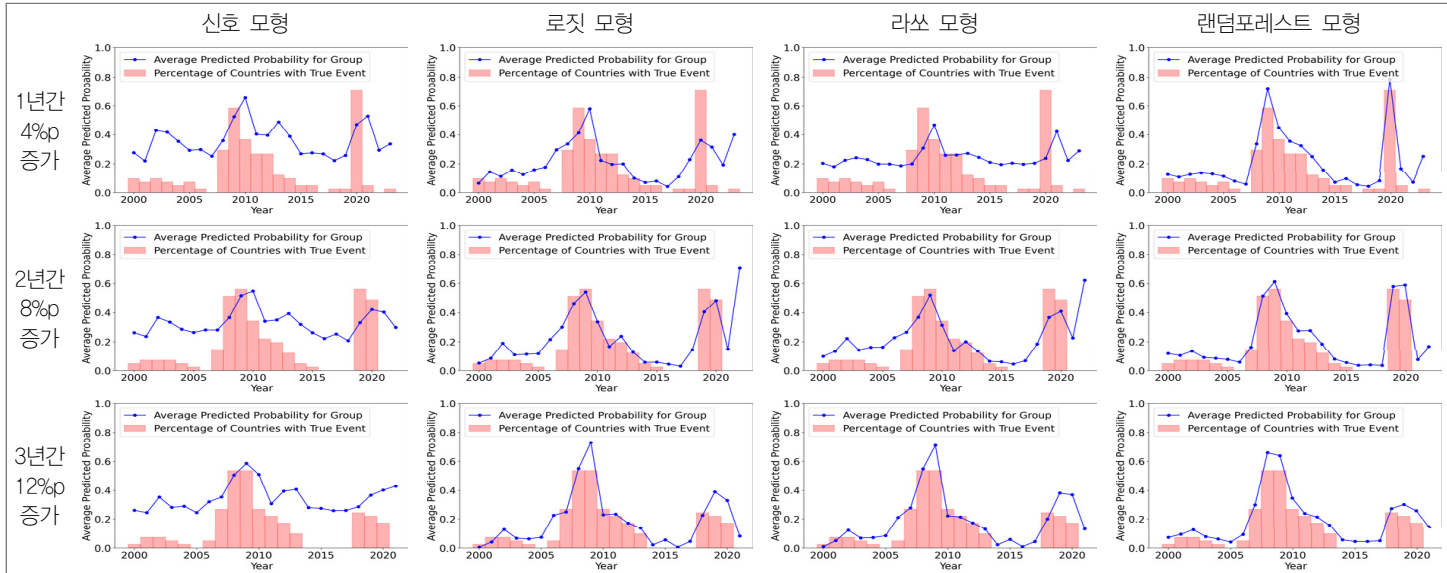
[그림 III-17] 모형별 한국의 정부부채 비율 급증 확률(연평균 4%p 증가 기준)



주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증하여 재정위기로 식별된 기간이며, 푸른 실선이 검은 점선보다 높은 시기가 모형에 의해 재정위기로 예측된 기간  
 자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?ts\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23EOO\\_OUT%23&pg=0&fc=Topic&bp=true&nb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&lo=LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?ts[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23EOO_OUT%23&pg=0&fc=Topic&bp=true&nb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&lo=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

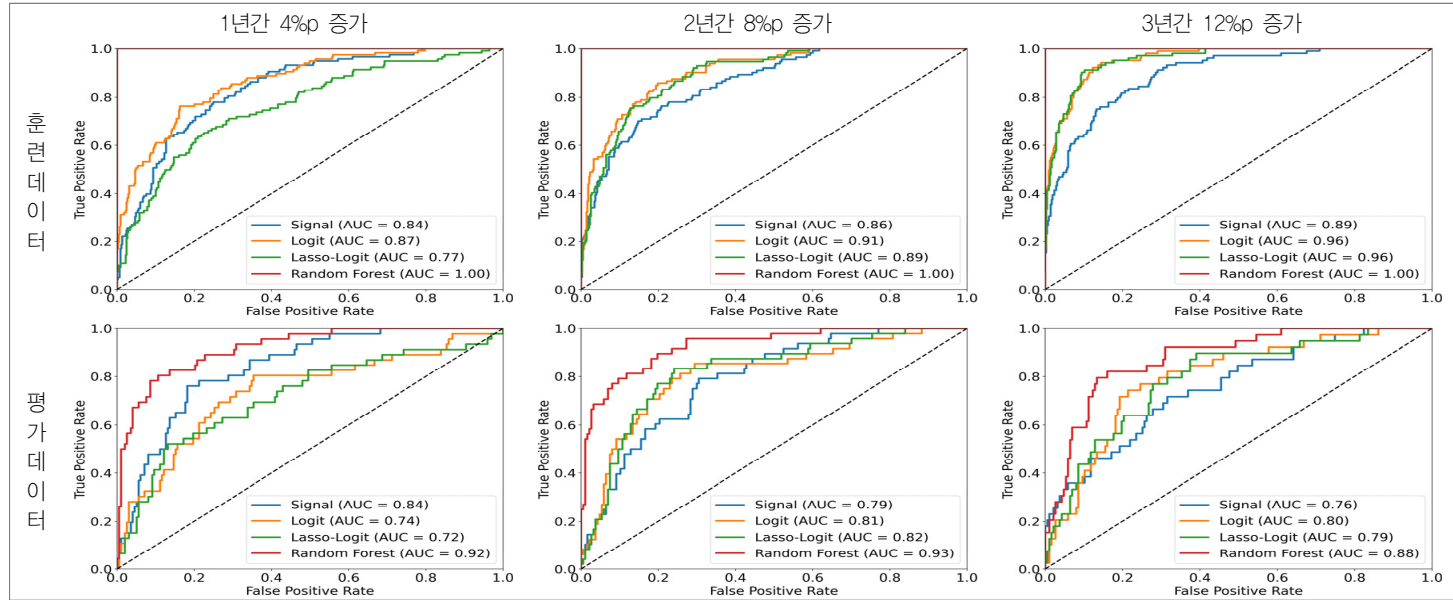
[그림 III-18] 모형별 선진국의 정부부채 비율 급증 확률(연평균 4%p 증가 기준)



주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증한 국가의 비율이며, 푸른 실선은 모형이 예측한 재정위기의 발생 확률의 국가 간 평균  
 자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?ts\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTPERIODS&lo=5&io\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?ts[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTPERIODS&lo=5&io[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

[그림 III-19] 모형별 ROC 곡선 비교(연평균 4%p 증가 기준)



자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&fc=Topic&bp=true&nb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&lm=-LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&fc=Topic&bp=true&nb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&lm=-LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

〈표 III-7〉 모형별 예측 성능 비교(연평균 4%p 증가 기준)

재정위험 구분	데이터 구분	성능측정 지표	신호	로짓	라쏘	랜덤 포레스트
1년간 4%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.78	0.79	0.80	0.99
		정밀도(precision)	0.49	0.50	0.51	1.00
		재현율(recall)	0.69	0.77	0.39	0.97
		F1 점수	0.57	0.61	0.44	0.98
		F2 점수	0.64	0.70	0.41	0.97
	AUC	0.84	0.87	0.77	1.00	
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.80	0.73	0.80	0.90
		정밀도(precision)	0.49	0.37	0.47	0.79
		재현율(recall)	0.74	0.67	0.50	0.65
		F1 점수	0.59	0.48	0.48	0.71
F2 점수		0.67	0.58	0.49	0.68	
AUC	0.84	0.74	0.72	0.92		
2년간 8%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.72	0.84	0.82	1.00
		정밀도(precision)	0.41	0.59	0.54	1.00
		재현율(recall)	0.82	0.77	0.80	1.00
		F1 점수	0.55	0.67	0.64	1.00
		F2 점수	0.68	0.73	0.73	1.00
	AUC	0.86	0.91	0.89	1.00	
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.71	0.80	0.79	0.91
		정밀도(precision)	0.39	0.51	0.49	0.84
		재현율(recall)	0.77	0.69	0.75	0.67
		F1 점수	0.52	0.58	0.60	0.74
F2 점수		0.65	0.64	0.68	0.70	
AUC	0.79	0.81	0.82	0.93		
3년간 12%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.86	0.80	0.91	0.98
		정밀도(precision)	0.66	0.49	0.70	0.91
		재현율(recall)	0.62	0.95	0.89	1.00
		F1 점수	0.64	0.65	0.79	0.95
		F2 점수	0.62	0.80	0.85	0.98
	AUC	0.89	0.96	0.96	1.00	
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.81	0.79	0.81	0.85
		정밀도(precision)	0.44	0.43	0.45	0.55
		재현율(recall)	0.44	0.69	0.51	0.77
		F1 점수	0.44	0.53	0.48	0.64
F2 점수		0.44	0.62	0.50	0.71	
AUC	0.76	0.80	0.79	0.88		

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&ic\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&ic[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

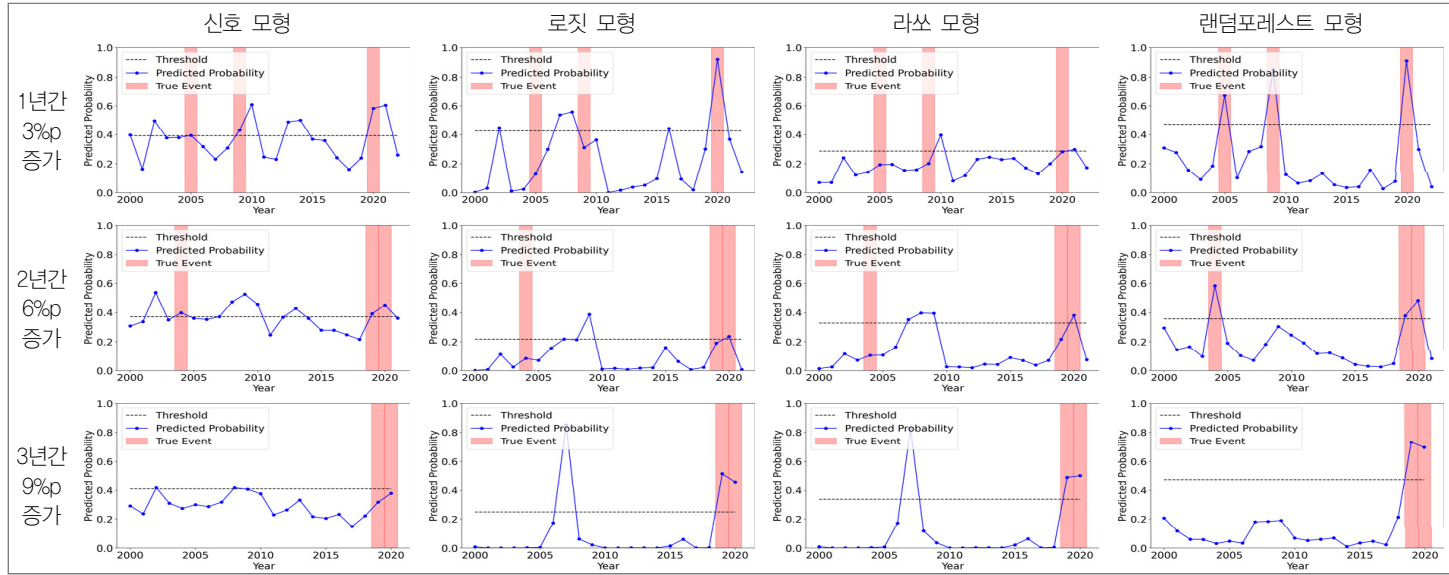


#### 다. 신흥국 포함

마지막 민감도 분석으로 기본 분석에 신흥국을 포함하여 관측치를 확대하면 재정위험 진단 모형의 예측력이 향상되는지 검토한다. [그림 Ⅲ-20]은 17개 신흥국을 포함하여 모형을 구축하였을 때 한국의 재정위기 발생에 대한 예측력을 보여주고 있다. 예측력이 가장 높은 랜덤포레스트 모형을 살펴보면, 신흥국을 포함하여 분석하였을 때 우리나라의 모든 재정위기를 정확히 예측한 것으로 나타난다. 특히 기본 분석에서 랜덤포레스트 모형은 3년간 정부부채 비율이 급증할 사건을 예측할 때 2008~2009년에 대해 거짓 양성 신호를 보냈지만, 신흥국을 포함한 분석에서는 이러한 오신호가 사라진 것이 눈에 띈다. 그 밖에 [그림 Ⅲ-20]~[그림 Ⅲ-22]에서 랜덤포레스트 모형이 다른 모형보다 재정위기를 예측하는 데 우월하다는 점은 선진국만을 대상으로 한 분석과 동일하게 나타난다.

신흥국 포함 여부에 따른 모형의 예측력을 비교하기 위해 <표 Ⅲ-8>을 <표 Ⅲ-2>와 비교해 보자. 마찬가지로 예측력이 가장 높은 랜덤포레스트 모형을 기준으로 비교해 보면 신흥국의 포함 여부가 재정위기에 대한 예측력을 획기적으로 개선시키는 것은 아니라는 점을 알 수 있다. 상황에 따라 신흥국을 포함하였을 때 재정위기에 대한 예측력이 소폭 높아지기도 하고 낮아지기도 한다. 1년과 2년간 정부부채 비율이 급증할 사건을 예측할 때는 선진국만을 대상으로 한 랜덤포레스트 모형의 예측력이 소폭 높지만, 3년간 정부부채 비율이 급증할 사건을 예측할 때는 신흥국을 포함한 모형의 예측력이 소폭 높다. 따라서 미래의 재정위기를 예측할 때 선진국을 대상으로 학습한 모형과 신흥국을 포함하여 학습한 모형을 동시에 활용하여 비교해 보는 것이 적절할 것으로 판단된다.

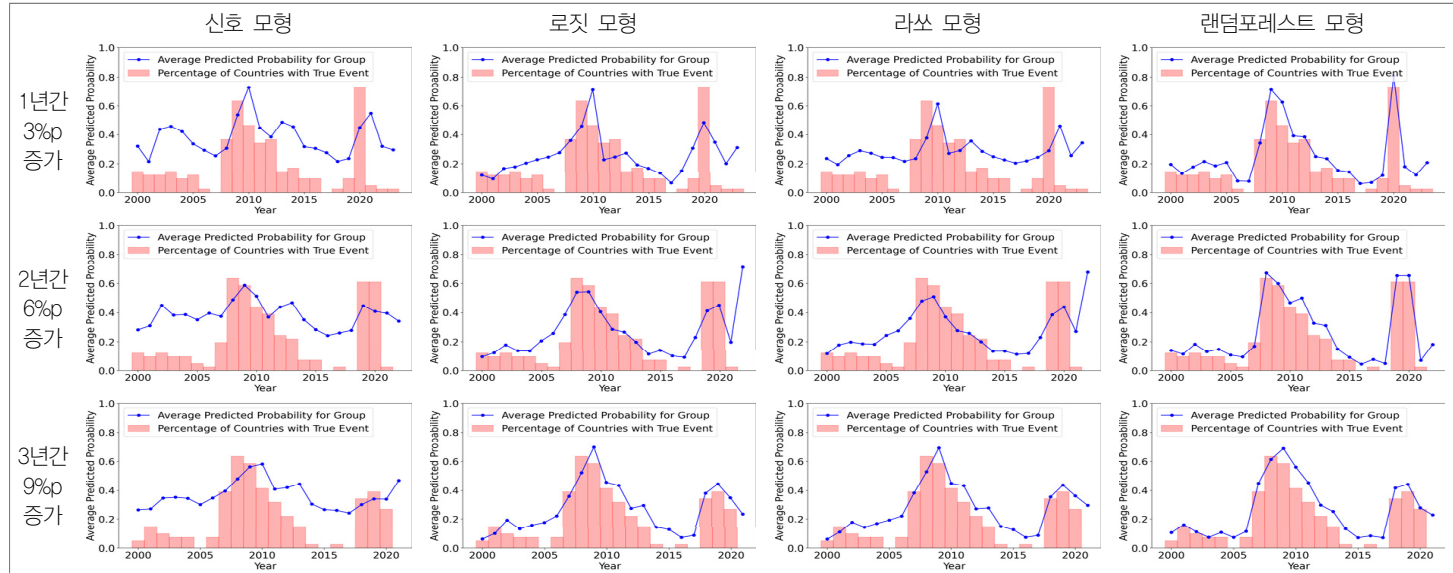
[그림 III-20] 모형별 한국의 정부부채 비율 급증 확률(신흥국 포함 모형)



주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증하여 재정위기로 식별된 기간이며, 푸른 실선이 검은 점선보다 높은 시기가 모형에 의해 재정위기로 예측된 기간  
 자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?is\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&nb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?is[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&nb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

[그림 III-21] 모형별 선진국의 정부부채 비율 급증 확률(신흥국 포함 모형)

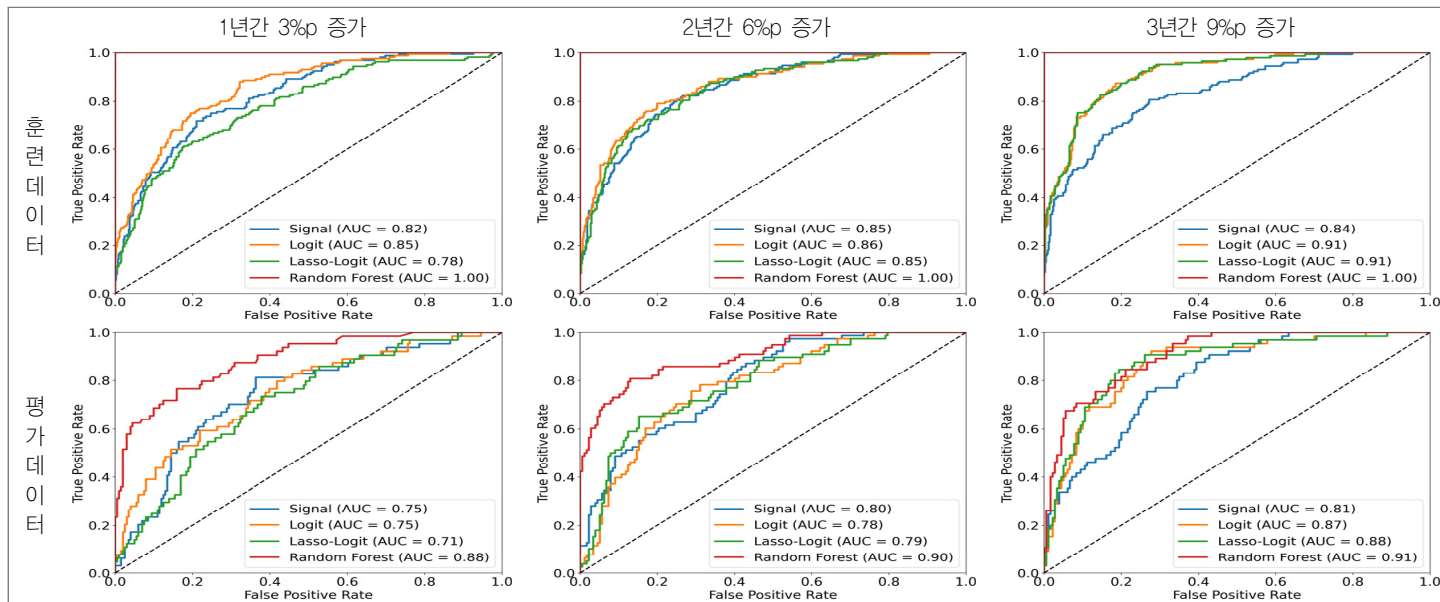


주: 붉은 막대는 실제 GDP 대비 정부부채 비율이 급증한 국가의 비율이며, 푸른 실선은 모형이 예측한 재정위기의 발생 확률의 국가 간 평균

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?ts\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23EOO\\_OUT%23&pg=0&fc=Topic&bp=true&nb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&lom=LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?ts[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23EOO_OUT%23&pg=0&fc=Topic&bp=true&nb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&lom=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

[그림 III-22] 모형별 ROC 곡선 비교(신흥국 포함 모형)



자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT,A&lom=LASTNPERIODS&l=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT,A&lom=LASTNPERIODS&l=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

〈표 III-8〉 모형별 예측 성능 비교(신흥국 포함 모형)

재정위험 구분	데이터 구분	성능측정 지표	신호	로짓	라쏘	랜덤 포레스트
1년간 3%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.69	0.80	0.70	1.00
		정밀도(precision)	0.45	0.64	0.45	1.00
		재현율(recall)	0.79	0.55	0.71	1.00
		F1 점수	0.57	0.59	0.55	1.00
		F2 점수	0.69	0.57	0.64	1.00
		AUC	0.82	0.85	0.78	1.00
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.68	0.77	0.65	0.87
		정밀도(precision)	0.41	0.52	0.38	0.81
		재현율(recall)	0.80	0.50	0.72	0.61
		F1 점수	0.54	0.51	0.49	0.70
		F2 점수	0.67	0.50	0.61	0.64
		AUC	0.75	0.75	0.71	0.88
2년간 6%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.66	0.76	0.80	1.00
		정밀도(precision)	0.43	0.52	0.60	1.00
		재현율(recall)	0.91	0.81	0.70	1.00
		F1 점수	0.58	0.64	0.64	1.00
		F2 점수	0.74	0.73	0.68	1.00
		AUC	0.85	0.86	0.85	1.00
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.66	0.71	0.78	0.85
		정밀도(precision)	0.47	0.52	0.65	0.73
		재현율(recall)	0.86	0.77	0.64	0.79
		F1 점수	0.61	0.62	0.65	0.76
		F2 점수	0.74	0.70	0.64	0.78
		AUC	0.80	0.78	0.79	0.90
3년간 9%p 증가	훈련 데이터 (in-sample)	정확도(accuracy)	0.77	0.84	0.85	1.00
		정밀도(precision)	0.53	0.64	0.68	1.00
		재현율(recall)	0.71	0.84	0.79	1.00
		F1 점수	0.61	0.72	0.73	1.00
		F2 점수	0.67	0.79	0.76	1.00
		AUC	0.84	0.91	0.91	1.00
	평가 데이터 (out-of-sample)	정확도(accuracy)	0.73	0.78	0.82	0.87
		정밀도(precision)	0.50	0.55	0.62	0.81
		재현율(recall)	0.74	0.89	0.82	0.66
		F1 점수	0.60	0.68	0.70	0.73
		F2 점수	0.67	0.79	0.77	0.69
		AUC	0.81	0.87	0.88	0.91

자료: 다음의 자료를 이용하여 저자 분석

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/wco-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\[0\]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

---

## IV. 결론 및 정책 시사점

---

재정의 지속가능성을 확보·유지하기 위해서는 투명하고 신뢰성 있는 재정 관리체계가 필요하다. 재정관리체계가 잘 작동하기 위해서는 다양한 요소가 필요한데 그중 하나가 재정위험(fiscal risk)을 정량적으로 진단하는 것이다. 재정관리체계는 재정위험에 대한 진단을 토대로 재정수지, 정부부채 등 재정총량에 대한 관리 목표를 구체적으로 설정하여 관리할 수 있다.

본 연구는 단기 재정위험 진단 방법으로 기계학습(machine learning)을 활용한 방법 등 다양한 방법을 검토하고 비교하였다. 재정위험은 다양한 세계에서 다양한 방법으로 진단될 수 있다. 우리나라의 경우 재정위험은 주로 장기적인 관점에서 인구구조의 변화가 재정의 지속가능성에 미치는 영향을 중심으로 논의되었다. 그러나 최근 코로나19 위기 대응 등으로 인하여 재정의 지속가능성이 악화되면서 김명규·김민경(2023), 한종석 외(2024) 등은 기존의 유럽연합 집행위원회(EC)와 국제통화기금(IMF)에서 활용하는 재정위험 진단 방법을 우리나라에 적용하는 연구를 수행하였다. 그러나 해외에서는 국제기구를 중심으로 기계학습(machine learning) 기법을 이용한 재정위험 진단을 시도하고 있는데, 국내에서는 이러한 시도가 아직 드문 것으로 파악된다. 기계학습 기법은 방법론의 특성상 장기보다는 단기적인 재정위험 진단에 적합하여 본 연구는 단기 재정위험 진단에 초점을 맞추었다. 재정위험은 다양한 각도에서 진단할 필요가 있는데, 본 연구는 우리나라의 재정위험 진단 방법론의 폭을 넓히는 데 의의가 있다.

본 연구는 재정위험의 정의에 있어 국내외 선행연구들과 차이가 있다. 본 연구는 재정위기(fiscal crisis)를 향후 1~3년간 GDP 대비 정부부채 비율이 일정 수준 이상 증가하는 사건으로 정의하고, 이러한 재정위기가 발생할 가능성을 재정위험(fiscal risk)으로 정의하였다. 특히 기본 분석에서는 GDP 대

비 정부부채 비율이 연평균 3%p 이상 증가하는 경우, 즉 1년간 3%p, 2년간 6%p, 3년간 9%p 이상 증가하는 경우를 재정위기가 발생한 것으로 정의하였다. 재정위험으로 채무불이행과 같은 심각한 재정위기보다 정부부채 비율이 급증할 가능성을 식별하는 것이 재정의 지속가능성 관리 측면에서 우리나라의 현실에 더 부합하는 것으로 판단된다. 재정의 지속가능성을 가늠하는 핵심적인 지표는 GDP 대비 국가채무 또는 정부부채 비율이다. 이 비율이 안정적이지 않고 지속적으로 증가한다면 일반적으로 재정은 지속가능하지 않다고 말한다. 특히 이 비율이 너무 높아진다면 심각한 재정위기에 직면할 가능성이 커진다. 따라서 재정위기를 사전에 방지하기 위해서는 단기적으로 정부부채 비율의 증가 속도를 관리할 필요가 있다.

이렇게 정의한 재정위기 전후의 재정 및 거시경제 변수의 변화를 살펴보았다. 그 결과 재정위기가 발생한 기간에 기초재정수지와 정부부채 등 주요 재정지표가 크게 악화되고, 이러한 영향이 일시적인 것이 아닌 일정 기간 지속되는 것으로 나타났다. 특히 재정위기 기간이 끝난 이후에도 높은 수준의 재정지출이 지속되어 정부부채가 상당 기간 빠르게 증가하는 것으로 나타났다는데, 이는 재정위기를 겪은 이후에도 재정관리가 중요함을 시사한다. 또한 주요 거시경제 변수의 변화를 살펴보면, 재정위기 기간에는 경제성장률이 급격히 낮아지는 것으로 분석되었다. 반면 이 기간에 실질이자율은 증가하거나 큰 변화가 없는 것으로 나타났다. 이에 이자율-성장률 격차가 크게 증가하며, 이 또한 상당 기간 증가한 상태로 머물러 있는 것으로 분석되었다. 즉 재정위기는 정부부채의 한도와 재정여력을 크게 축소시키는 것으로 해석할 수 있다.

구체적인 재정위험 진단 방법으로 신호(signaling), 로짓(logit), 라쏘(LASSO), 랜덤포레스트(random forest) 모형을 비교·검토하였다. 신호, 로짓 모형은 전통적인 방법, 그리고 라쏘, 랜덤포레스트 모형은 기계학습 방법이다. 모형의 예측력을 비교하기 위해 정확도(accuracy), 정밀도(precision), 재현율(recall), F1 점수, F2 점수, AUC(Area Under the Curve) 등의 다양한 지표를 검토하였다. 재정위험 진단을 위한 변수들은 국제통화기금(IMF)이 주기적으로 발

표하는 World Economic Outlook(IMF-WEO)의 자료와 경제협력개발기구(OECD)의 Economic Outlook(OECD-EO) 자료를 활용하였다. 그리고 기본 분석의 대상 국가는 IMF-WEO에서 선진국으로 분류하는 국가 중 OECD-EO에도 자료가 존재하는 32개 국가이다.

분석 결과는 다음과 같이 요약할 수 있다. 첫째로 랜덤포레스트 모형이 다른 모형들에 비해 더 나은 예측력을 보여주었다. 이는 결정나무(decision tree)에 기반한 랜덤포레스트 모형이 변수들의 비선형적이고 복잡한 관계를 모형화하는 데 장점이 있으므로, 이러한 특성이 상대적으로 부족한 로짓이나 라쏘 모형에 비해 좋은 성능을 낸 것으로 판단된다. 만일 로짓이나 라쏘 모형에 변수들의 교차항을 추가하거나 비선형적인 관계를 추가할 수 있다면 모형의 성능을 향상할 수도 있을 것이다. 다만 모든 변수들의 교차항과 비선형적인 관계를 넣는 것은 현실적으로 어려우며 연구자의 자의적인 판단이 일부 필요할 수 있다.<sup>27)</sup>

둘째로 로짓 모형과 라쏘 모형의 예측 성능이 거의 유사한 것으로 나타났다. 대체로 훈련데이터에서는 로짓 모형이, 평가데이터에서는 라쏘 모형이 소폭 높은 예측력을 보였으나 그 차이가 크지는 않았다. 이는 라쏘 모형이 로짓 모형의 예측 성능을 추가적으로 개선하지 못했다는 것을 의미한다. 이 결과는 두 가지로 해석될 수 있다. 첫째, 로짓 모형이 이미 최적의 예측 성능을 보여주고 있어 개선의 여지가 거의 없었을 가능성이 있다. 둘째, 라쏘 모형의 변수 선택 및 정규화(regularization) 기능이 본 데이터에서 유효하게 작동하지 않았을 가능성이 있다. 본 연구의 분석 결과만으로는 이 두 가지 중 어느 경우인지는 확실하지 않지만, 로짓 모형의 예측 성능이 전반적으로 낮지 않다는 점을 고려할 때 첫 번째 가능성이 더 큰 것으로 보인다.<sup>28)</sup>

셋째로 신호 모형의 예측 성능이 다른 모형에 비해서 낮은 것으로 확인되

27) 변수들이 많아질수록 조합 가능한 교차항의 수는 기하급수적으로 많아질 뿐만 아니라 각 변수들의  $n$ 차 항을 무한정 많이 넣는 것은 불가능하다.

28) 이를 실제로 확인하기 위해서는 라쏘 모형에서 최적 정규화 파라미터  $\lambda$  값이 얼마인지를 통해서 확인할 수 있는데, 만일  $\lambda$  값이 매우 작다면 벌점항이 거의 들어가지 않아 로짓 모형과 거의 같으므로 정규화가 예측 성능 개선에 거의 도움이 되지 않았다고 평가할 수 있을 것이다.



었다. 다만 항상 그런 것은 아니고 일부 상황에서는 신호 모형이 로짓이나 라쏘 모형보다 더 나은 예측력을 보이는 경우도 있었다. 신호 모형은 재정 위기 진단에 자주 사용된 모형으로서 각 변수들마다 정해진 임계치가 있어서 이 임계치를 넘으면 신호를 보내는 것으로 간주하고 신호들의 합을 통해서 위기를 예측하는 것이다. 이 모형은 개별 변수들이 위험한 범위에 있는지, 아니면 안전한 범위에 있는지에 대한 명시적인 구분을 도와주기 때문에 개별 변수들을 관리하는 데 있어 시사점을 제공하기에는 큰 도움이 될 수 있다. 하지만 이 모형은 변수들을 모두 독립적인 것으로 판단하고 서로의 상호작용은 고려하지 않아 예측력이 높지 않기 때문에 실효성 있는 모형으로 활용하기는 어려울 것으로 판단된다.

마지막으로 각 모형들이 재정위기 예측에 유용하다고 판단하는 변수들은 모형과 예측 시계에 따라 다른 것으로 나타났다. 예측력이 높은 랜덤포레스트 모형을 기준으로 살펴보면 1년 또는 2년간 정부부채 비율이 급증할 사건을 예측할 때는 재정수지와 정부부채 비율과 같은 재정변수가 예측에 중요한 영향을 미치는 것으로 나타났다. 하지만 3년간 정부부채 비율이 급증할 사건을 예측할 때는 재정변수도 중요하지만 경상수지, 이자율, 장단기 금리차 등 경제환경 변수의 역할이 강화되는 것으로 나타났다. 이는 비교적 짧은 시계에는 정부부채 비율을 직접적으로 결정하는 변수가 예측에 중요한 역할을 하지만 비교적 긴 시계에는 재정변수뿐만 아니라 재정변수에 영향을 미칠 수 있는 더욱 근본적인 원인 변수의 역할이 중요해짐을 시사하는 것으로 해석할 수 있다.

본 연구는 민감도 분석도 수행하였다. 기본 분석에서 재정위기를 GDP 대비 정부부채 비율이 연평균 3%p 이상 증가한 상황으로 정의하였는데, 이는 자의적이라는 한계가 있다. 이에 정부부채 비율의 증가 속도에 대해 연평균 2%p와 4%p 이상과 같은 다른 기준을 적용해 보았는데, 모형의 예측력 간의 정성적인 특성은 기본 분석과 유사하였다. 정부부채 비율의 증가 속도를 통제하기 위한 목표는 재정이나 경제 환경에 따라 달라질 수 있다. 따라서 이를 특정한 하나의 수치로 정하여 진단하기보다는 재정관리 목표에 맞추어

분석하는 것이 적절할 것으로 판단된다.

한편 기본 분석에서는 선진국을 대상으로만 분석하였는데, 17개 신흥국을 포함한 민감도 분석도 수행하였다. 그 결과 재정위기를 예측하는 데 있어 신흥국의 포함 여부가 예측력을 뚜렷하게 높이거나 낮추지는 않았으며, 예측 시계에 따라 어떤 경우는 선진국만을 대상으로 분석할 때, 그리고 다른 경우는 신흥국을 포함하여 분석할 때 예측력이 소폭 높은 것으로 나타났다. 즉 실제로 재정위기를 예측할 때 신흥국을 포함한 모형과 그렇지 않은 모형을 모두 활용하여 비교해 보는 것이 적절할 것으로 판단된다.

본 연구에서 검토한 모형들은 재정위험 진단에 활용될 수 있지만, 재정위기의 발생 원인을 직접적으로 설명하지는 않는다. 물론 일차적으로 모형에 활용된 변수들 중 재정위험 진단 결과에 중요한 영향을 미친 변수들을 식별하여 그 원인을 가늠할 수 있다. 다만 진단 결과 재정위기가 발생할 확률이 높다면 그 원인을 파악하고 대처하기 위해서는 당시의 경제 상황 등에 대한 별도의 분석이 필요하다.

---

## 참고문헌

---

### 〈국내 문헌〉

- 국회예산정책처, 『2022~2070년 NABO 장기 재정전망』, 2022. 8.
- 김명규·김민경, 『재정 지속가능성 복합지표 연구: EU 재정 지속가능성 지표를 중심으로』, 분석 23-02, 한국재정정보원, 2023. 5.
- 김학수·이태석·홍우형, 『코로나19 이후 조세·재정정책 방향의 재검토』, 연구보고서 2021-06, 한국개발연구원, 2021. 12.
- 대한민국정부, 『2024~2028년 국가재정운용계획』, 2024. 9.
- \_\_\_\_\_, 『2024~2028년 국가재정운용계획 첨부서류』, 2024. 9.
- 이환웅·고창수·배진수, 『머신러닝을 활용한 재정사업평가: 정책금융 사례를 중심으로』, 연구보고서 21-07, 한국조세재정연구원, 2021. 12.
- 정재현·이환웅, 『머신러닝(Machine Learning)을 활용한 조세·재정정책의 평가와 설계』, 연구보고서 20-08, 한국조세재정연구원, 2020. 12.
- 한중석·허준영·편주현·임태준, 『재정의 지속가능성 확보를 위한 지표 및 분석 방안 연구』, 국회예산정책처, 2024. 1.
- 홍승현·윤성주, 『재정분야 지속가능성 점검지표 설정』, 기획재정부·한국조세재정연구원, 2013. 9.

### 〈외국 문헌〉

- Baldacci, Emanuele, Iva Petrova, Nazim Belhocine, Gabriela Dobrescu, and Samah Mazraani, "Assessing Fiscal Stress," IMF Working Paper, WP/11/100, International Monetary Fund, 2011. 5.
- EC, *Fiscal Sustainability Report 2021*, Volume 1, European Commission, 2022. 4.

- Fawcett, Tom, "An introduction to ROC analysis," *Pattern Recognition Letters*, 27(8), 2006, pp. 861~874.
- Flores, Enrique, Pranav Gupta, Yinqiu Lu, Paulo Medas, Dinar Prihardini, Honda Selim, Weining Xin, and Masafumi Yabara, "Upgrading Fiscal Frameworks in Asia-Pacific," Departmental Paper, International Monetary Fund, 2024. 9.
- Hellwig, Klaus-Peter, "Predicting fiscal crises: A machine learning approach," IMF Working Paper, WP/21/150, International Monetary Fund, 2021.
- IMF, "How to Assess Country Risk: The Vulnerability Exercise Approach Using Machine Learning," Technical Notes and Manuals, TNM/21/03, International Monetary Fund, 2021.
- \_\_\_\_\_, *Staff Guidance Note on the Sovereign Risk and Debt Sustainability Framework for Market Access Countries*, International Monetary Fund, 2022. 8.
- Jarmulska, Barbara, "Random forest versus logit models: Which offers better early warning of fiscal stress?," Working Paper Series No. 2408, European Central Bank, 2020. 5.
- \_\_\_\_\_, "Random forest versus logit models: Which offers better early warning of fiscal stress?," *Journal of Forecasting* 41(3), 2022, pp. 455~490.
- Kaminsky, Graciela, Saul Lizondo, and Carmen M. Reinhart, "Leading indicators of currency crises," *IMF Staff Papers* 45(1), 1998, pp. 1~48.
- Medas, Paulo, Tigran Poghosyan, Yizhi Xu, Juan Farah-Yacoub, and Kerstin Gerling, "Fiscal Crises," IMP Working Paper, WP/17/86, 2017.
- \_\_\_\_\_, "Fiscal crises," *Journal of International Money and Finance* 88, 2018, pp. 191~207.
- Valencia, Oscar M., Juan Camilo Díaz, and Diego A. Parra, "Assessing

Macro-Fiscal Risk for Latin American and Caribbean Countries,” IDB Working Paper Series, IDB-WP-1346, Inter-American Development Bank, 2022. 10.

## 〈통계자료 및 누리집〉

국제금융센터, 지표·자료, 「CDS 프리미엄」, <https://www.kcif.or.kr/chart/intrList>,  
검색일자: 2024. 6. 29.

지표누리, e-나라지표, 국정모니터링시스템, 「국가채무추이」, [https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx\\_cd=1106](https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=1106), 검색일자:  
2024. 11. 11.

\_\_\_\_\_, e-나라지표, 국정모니터링시스템, 「통합재정수지」, [https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx\\_cd=1104](https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=1104), 검색일자:  
2024. 10. 23.

한국은행, 경제통계시스템(ECOS), 「국민계정」, “2.1.2.2.3. 국내총생산에 대한 지출(원계열, 명목, 분기 및 연간),” <https://ecos.bok.or.kr/#/SearchStat>,  
검색일자: 2024. 11. 11.

행정안전부, 공공데이터포털, 「국가신용등급」, <https://www.data.go.kr/data/15117342/fileData.do>, 검색일자: 2024. 6. 23.

IMF, “Korea: History of Lending Commitments as of April 30, 2020,” <https://www.imf.org/external/np/fin/tad/extarr2.aspx?memberKey1=550&date1key=2020-04-30>, 검색일자: 2024. 6. 27.

\_\_\_\_\_, “SDR Valuation,” [www.imf.org/external/np/fin/data/rms\\_sdrv.aspx](http://www.imf.org/external/np/fin/data/rms_sdrv.aspx),  
검색일자: 2024. 6. 27.

\_\_\_\_\_, “World Economic Outlook Database,” April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자:  
2024. 7. 24.

OECD Data Explorer, “Economic Outlook 115,” Edition 2024/1, <https://data>

-explorer.oecd.org/vis?fs[0]=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\_OUT%23&pg=0&fc=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD\_EO\_115%40DF\_EO\_115&df[ag]=OECD,ECO,MAD&df[vs]=1,0&dq=,GDPV\_ANNPCT,A&lom=LASTNPERIODS&lo=5&to[TIME\_PERIOD]=false, 검색일자: 2024. 7. 24.

## 부 록: 분석에 활용된 독립변수

〈부표 1〉 분석에 활용된 독립변수

출처	구분	변수명	변수 설명	단위
IMF-WEO	국민 계정	weo_ngdp_r_gw_l1	경제성장률(전년도)	%
		weo_ngdp_r_gw_l2	경제성장률(전전년도)	%
		weo_ngsd_ngdp_l1	GDP 대비 총저축 비율(전년도)	GDP 대비 %
		weo_ngsd_ngdp_pd_l1	GDP 대비 총저축 비율의 차분(전년도)	GDP 대비 %p
		weo_ngsd_ngdp_pd_l2	GDP 대비 총저축 비율의 차분(전전년도)	GDP 대비 %p
		weo_nid_ngdp_l1	GDP 대비 총투자 비율(전년도)	GDP 대비 %
		weo_nid_ngdp_pd_l1	GDP 대비 총투자 비율의 차분(전년도)	GDP 대비 %p
		weo_nid_ngdp_pd_l2	GDP 대비 총투자 비율의 차분(전전년도)	GDP 대비 %p
	재정	weo_ggxwdg_ngdp_l1	GDP 대비 정부부채 비율(전년도)	GDP 대비 %
		weo_ggxwdg_ngdp_pd_l1	GDP 대비 정부부채 비율의 차분(전년도)	GDP 대비 %p
		weo_ggxwdg_ngdp_pd_l2	GDP 대비 정부부채 비율의 차분(전전년도)	GDP 대비 %p
		weo_ggxonlb_ngdp_l1	GDP 대비 기초재정수지 비율(전년도)	GDP 대비 %
		weo_ggxonlb_ngdp_pd_l1	GDP 대비 기초재정수지 비율의 차분(전년도)	GDP 대비 %p
		weo_ggxonlb_ngdp_pd_l2	GDP 대비 기초재정수지 비율의 차분(전전년도)	GDP 대비 %p
		weo_ggsb_npgdp_l1	잠재 GDP 대비 구조적 재정수지 비율(전년도)	GDP 대비 %
		weo_ggsb_npgdp_pd_l1	잠재 GDP 대비 구조적 재정수지 비율의 차분(전년도)	GDP 대비 %p
		weo_ggsb_npgdp_pd_l2	잠재 GDP 대비 구조적 재정수지 비율의 차분(전전년도)	GDP 대비 %p

〈부표 1〉의 계속

출처	구분	변수명	변수 설명	단위
IMF-WEO	재정	weo_ggx_ngdp_l1	GDP 대비 재정지출 비율(전년도)	GDP 대비 %
		weo_ggx_ngdp_pd_l1	GDP 대비 재정지출 비율의 차분(전년도)	GDP 대비 %p
		weo_ggx_ngdp_pd_l2	GDP 대비 재정지출 비율의 차분(전전년도)	GDP 대비 %p
	고용	weo_lur_l1	실업률(전년도)	%
		weo_lur_pd_l1	실업률의 차분(전년도)	%p
		weo_lur_pd_l2	실업률의 차분(전전년도)	%p
	국제 무역 및 국제 거래	weo_tgx_rpch_l1	실질 상품수출 증가율(전년도)	%
		weo_tgx_rpch_l2	실질 상품수출 증가율(전전년도)	%
		weo_tmg_rpch_l1	실질 상품수입 증가율(전년도)	%
		weo_tmg_rpch_l2	실질 상품수입 증가율(전전년도)	%
		weo_bca_ngdpd_l1	GDP 대비 경상수지 비율(전년도)	GDP 대비 %
		weo_bca_ngdpd_pd_l1	GDP 대비 경상수지 비율의 차분(전년도)	GDP 대비 %p
		weo_bca_ngdpd_pd_l2	GDP 대비 경상수지 비율의 차분(전전년도)	GDP 대비 %p
	물가	weo_pcpi_gw_l1	소비자물가지수 상승률(전년도)	%
		weo_pcpi_gw_l2	소비자물가지수 상승률(전전년도)	%
	환율	weo_pppex_l1	구매력평가 환율(전년도)	국제달러 대비 자국화폐
		weo_pppex_pd_l1	구매력평가 환율의 차분(전년도)	국제달러 대비 자국화폐
		weo_pppex_pd_l2	구매력평가 환율의 차분(전전년도)	국제달러 대비 자국화폐
		weo_usex_l1	대미국달러 환율(전년도)	미국달러 대비 자국화폐
		weo_usex_pd_l1	대미국달러 환율의 차분(전년도)	미국달러 대비 자국화폐
		weo_usex_pd_l2	대미국달러 환율의 차분(전전년도)	미국달러 대비 자국화폐
	선진국 경제	adv_ngdp_rpch_l1	선진국의 경제성장률(전년도)	%
		adv_ngdp_rpch_l2	선진국의 경제성장률(전전년도)	%
		adv_lur_l1	선진국의 실업률(전년도)	%
		adv_lur_pd_l1	선진국의 실업률의 차분(전년도)	%p
		adv_lur_pd_l2	선진국의 실업률의 차분(전전년도)	%p



〈부표 1〉의 계속

출처	구분	변수명	변수 설명	단위
IMF-WEO	선진국 경제	adv_pccipch_11	선진국의 소비자물가지수 상승률(전년도)	%
		adv_pccipch_12	선진국의 소비자물가지수 상승률(전전년도)	%
	세계 경제	wid_tradepch_11	전 세계의 실질 무역량 증가율 (전년도)	%
		wid_tradepch_12	전 세계의 실질 무역량 증가율 (전전년도)	%
OECD-EO	재정	oecd_gnintq_11	GDP 대비 순이자지출(전년도)	%
		oecd_gnintq_pd_11	GDP 대비 순이자지출의 차분(전년도)	%p
		oecd_gnintq_pd_12	GDP 대비 순이자지출의 차분(전전년도)	%p
	이자율	oecd_irs_11	단기이자율(전년도)	%
		oecd_irs_pd_11	단기이자율의 차분(전년도)	%p
		oecd_irs_pd_12	단기이자율의 차분(전전년도)	%p
		oecd_irgap_11	장단기 금리차(전년도)	%p
		oecd_irgap_pd_11	장단기 금리차의 차분(전년도)	%p
		oecd_irgap_pd_12	장단기 금리차의 차분(전전년도)	%p
	환율	oecd_excheb_11	명목 실효환율(전년도)	지수
		oecd_excheb_pd_11	명목 실효환율의 차분(전년도)	지수
		oecd_excheb_pd_12	명목 실효환율의 차분(전전년도)	지수

자료: 다음의 자료를 이용하여 저자 작성

1. IMF, "World Economic Outlook," April 2024 Edition, <https://www.imf.org/en/Publications/WEO/weo-database/2024/April>, 검색일자: 2024. 7. 24.
2. OECD Data Explorer, "Economic Outlook 115," Edition 2024/1, [https://data-explorer.oecd.org/vis?fs\(0\)=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO\\_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df\[ds\]=dsDisseminateFinalDMZ&df\[id\]=DSD\\_EO\\_115%40DF\\_EO\\_115&df\[ag\]=OECD.ECO.MAD&df\[vs\]=1.0&dq=GDPV\\_ANNPCT.A&lom=LASTNPERIODS&lo=5&to\[TIME\\_PERIOD\]=false](https://data-explorer.oecd.org/vis?fs(0)=Topic%2C1%7CEconomy%23ECO%23%7CEconomic%20outlook%23ECO_OUT%23&pg=0&ic=Topic&bp=true&snb=3&df[ds]=dsDisseminateFinalDMZ&df[id]=DSD_EO_115%40DF_EO_115&df[ag]=OECD.ECO.MAD&df[vs]=1.0&dq=GDPV_ANNPCT.A&lom=LASTNPERIODS&lo=5&to[TIME_PERIOD]=false), 검색일자: 2024. 7. 24.

## 단기 재정위험 진단 방법 연구

---

오종현 · 배진수

본 연구는 우리나라의 단기 재정위험(fiscal risk)을 진단하기 위해 재정위기(fiscal crisis)를 정의하고 이를 진단할 수 있는 다양한 방법론을 비교·분석하였다. 본 연구는 재정위기를 GDP 대비 정부부채 비율이 1~3년간 연평균 일정 수준(예: 3%p) 이상 증가한 사건으로 정의하고, 이러한 재정위기가 발생할 가능성을 재정위험으로 정의하였다. 구체적인 재정위험 진단 방법으로 신호(signaling), 로짓(logit), 라쏘(Least Absolute Shrinkage and Selection Operator, LASSO), 랜덤포레스트(random forest) 모형을 비교·검토하였다. 재정위험 진단을 위한 변수들은 국제통화기금(IMF)이 주기적으로 발표하는 World Economic Outlook(IMF-WEO)의 자료와 경제협력개발기구(OECD)의 Economic Outlook(OECD-EO) 자료를 활용하였다. 분석 대상 국가는 IMF-WEO 자료에서 선진국으로 분류되며 OECD-EO 자료에도 존재하는 32개국이다. 모형의 예측력을 비교하기 위해 ROC 곡선(Receiver Operating Characteristic Curve) 및 AUC(Area Under the Curve)와 함께 정확도(accuracy), 정밀도(precision), 재현율(recall), F1 점수, F2 점수 등의 다양한 지표를 검토하였다. 분석 결과 기계학습(machine learning) 기법 중 하나인 랜덤포레스트 모형이 다른 모형보다 우수한 예측력을 보이는 것으로 나타났다.

## A Study on Methods for Diagnosing Short-Term Fiscal Risks

---

Jonghyeon Oh and Jinsoo Bae

This study defines fiscal crises suitable for diagnosing short-term fiscal risks in South Korea and compares several diagnostic methodologies. A fiscal crisis refers to an event where the government debt-to-GDP ratio increases by a certain annual average level (e.g., 3 percentage points) over one to three years, and the possibility of such a fiscal crisis occurring is defined as a fiscal risk. The study compares and examines various fiscal risk diagnostic methods, including the signaling approach, logit, LASSO(Least Absolute Shrinkage and Selection Operator), and random forest model. The variables used in the models are sourced from the IMF's World Economic Outlook(IMF-WEO) and the OECD's Economic Outlook (OECD-EO). The analysis covers 32 countries classified as advanced economies in the IMF-WEO, for which data are also available in the OECD-EO. To assess the predictive performance of each model, the study evaluates several indicators, including the Receiver Operating Characteristic(ROC) curve, Area Under the Curve(AUC), accuracy, precision, recall, F1 score, and F2 score. The results indicate that the random forest model, a machine learning technique, demonstrates superior predictive performance compared to the other models.

## ■ 저자약력

### 오종현

한양대학교 경제학 학사  
미국 The Ohio State University 경제학 박사  
현, 한국조세재정연구원 연구위원

### 배진수

연세대학교 경제학 학사  
미국 The Ohio State University 경제학 박사  
현, 한국금융연구원 부연구위원

### 자료 수집 및 정리

김인유 한국조세재정연구원 선임연구원  
장준희 한국조세재정연구원 선임연구원

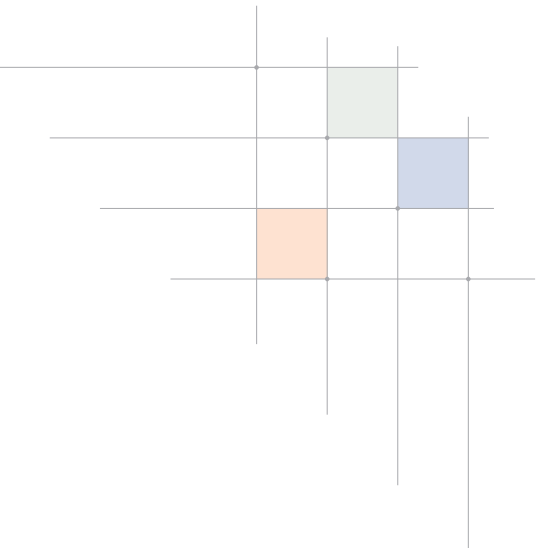
연구보고서 24-16

## 단기 재정위험 진단 방법 연구

발행	행	2024년 12월 31일
저자	자	오종현 · 배진수
발행인	이	영
발행처	처	한국조세재정연구원
주소	소	30147 세종특별자치시 시청대로 336
전화	화	(044)414-2114(대)
홈페이지	지	www.kjpt.re.kr
등록	록	1993. 7. 15. 제2014-24호
정가	가	14,000원
조판 및 인쇄	일	지사
I S B N		979-11-6655-325-7 93320

© 한국조세재정연구원 2024 \* 잘못 만들어진 책은 바꾸어 드립니다.

본 보고서는 친환경 용지를 사용하여 인쇄되었습니다.



KOREA INSTITUTE  
OF PUBLIC FINANCE

**kipf 한국조세재정연구원**

30147 세종특별자치시 시청대로 336  
TEL: (044)414-2114(대) www.kipf.re.kr



값 14000 원

93320

9 791166 553257  
ISBN 979-11-6655-325-7