# 파이낸스 어낼리틱스 789교시 4조 최종보고서

바이오메디컬공학부 201800308 강민구 통계학과 201902732 이윤지

## 1. 제목

선형회귀를 통한 물가 영향 지표 분석

## 2. 연구주제 및 배경

- 최근 물가수준의 급등세로 인해 일자리 사정도, 생활에도 서민 고통은 더 커지고 있다. 이에 외식물가가 8.9% 증가하여 짜장면 한 그릇과, 떡볶이 1인분도 먹기 겁나는 수준에 이르렀다. 또한 서민의 체감 경제적 어려움을 수치화한 서민경제고통지수가 집계 이후 역대 가장 높게 나타났고, 실업률과 소비자물가를 반영한 경제고통지수도 2019년 이래로 계속 증가하는 추세이다. 이에 더해 올해 상반기 무역적자가 103억 달러를 기록하면서 역대 최고 기록이었던 91억 달러를 넘어섰다. 또한 경상수지 흑자는 지난해 동월 대비 66억 달러 감소한 10억 달러를 기록하였다. 우리나라 경제를 책임지는 두 수지가 악화됨과 동시에 환율이 상승하면서 물가가 오르는 악순환을 겪고 있다.
- 따라서 물가에 영향을 주는 지표를 분석하여 이를 조정할 수 있는 정책을 고안해 서민들의 고통을 완화하고자 한다.

## 3. 자료(변수들) 설명

## 가. 유가

: 대표적인 원자재 지수이기에 우선적으로 고려하였다. 현대경제연구원에 의하면 2020년 기준으로 한국은 OECD 회원국 중 1위의 원유의존도와 4위의 1인당 원유 소비량을 기록하는 등 석유 의존성이 굉장히 높았다. 특히 철 강, 화학, 운수업, 서비스업 등 산업이 국제유가 상승에 영향을 받아 높은 원가상승률을 보일 것으로 예상된다.

#### 나. 금리와 환율

: 원화 가치의 상승으로 인해 수입물가의 하락이 국내 물가를 하락시키는 직접적 요인으로 작용한다. 반면 환율이 상승하면 원화가치가 하락하면서 물가가 상승하고, 이에 원화가치를 올리기 위해 금리를 인상하여 물가를 다시 낮추는 식의 경제 흐름이 이어진다.

# 다. 코스피

: 환율이 낮아지면 국내 시장으로의 투자가 활발해져 코스피지수가 오르는 등 영향을 줄 것이라 생각하였다.

# 라. COFIX 금리

: 기준금리보다 가산금리가 높아지는 한계를 보완하고자 COFIX 금리를 사용하기로 결정하였다.

# 마. 상품수지, 무역수지, 경상수지

: 경상수지와 무역수지가 급격히 하락해 쌍둥이 적자에 대한 공포가 커지고 있다. 상품수지와 경상수지, 무역수지의 흑/적자가 물가에 영향을 줄 것이다. 다만, 계절조정을 하지 않은 수치이므로 후처리가 필요함을 고려하였다.

### 바. 코로나19 확진자 수

: 코로나 확산에 따른 사회적 거리두기 및 대면접촉 기피는 총 수요를 감소시키면서 물가상승률을 낮추는 용도로 작용하였다. 대면 서비스를 중심으로 한 수요가 위축되고 코로나19 정부 대응 정책도 물가하방압력으로 작용하면서 물가 상승률을 둔화시킨 것으로 보인다.

#### 사. 소비자물가지수

: 주로 사용되는 물가지수에는 생산자물가지수와 수출입물가지수, GDP 디플레이터등이 있다. 이 중 소비자 입장에서 물가를 가장 잘 나타낼 것이라 생각되는 소비자물가지수를 종속변수로 선택하였다. 다만 올해의 물가지수를 제외한 2018년부터 2021년의 월별 물가지수를 사용하여 분석하고자 한다. 올해는 전쟁, 미국의 금리 정책 등과 같이 독립변수로 통제할 수 없는 노이즈가 생겨 모델을 구축하는데 방해가 될 것으로 판단하였다. 따라서 지난 4개년간의 물가만을 분석하여 주요 지표들을 선정하고자 한다.

# 4. 종속변수(Y)

: 소비자물가지수(priceindex)

#### 5. 독립변수(X)

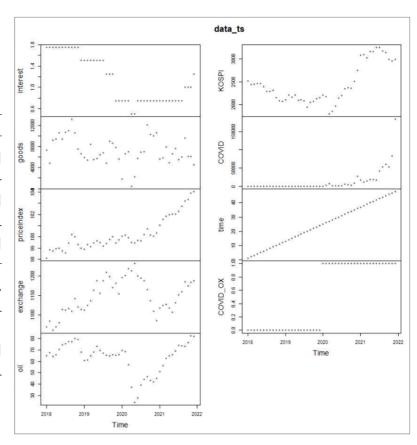
: 기준금리(interest), 상품수지(goods), 환율(exchange), 유가(oil), 코스피(KOSPI), 코로나19 확진자 수(COVID), 코로나 발생여부(COVID\_OX), 추세가변수(time)

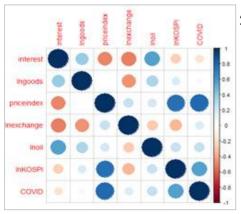
# 6. 자료 분석과정

1) 시간에 따른 데이터 추세 확인

이 자료는 시계열 자료이므로, 시간에 따른 추세를 확인해야 한다. 분산이 점차 커지는 변수가 있으므로 로그 변환하기로 결정하였다. 종속변수인 물가지수는 전년비, 전년동월비로 단위가 비율(%)이기에 로그변환을 취하지 않는다. 금리 또한 단위가 비율(%)이므로 로그변환 하지않는다. COVID 변수는 0을 포함하는 자료이므로 로그변환하지 않으며, COVID\_OX 변수는 범주형 자료임으로로그변환을 진행하지 않는다.

따라서 로그변환한 자료는 상품수지 (goods), 환율(exchange), 유가(oil), 코 스피(KOSPI)이다.





#### 2) 로그변환 후 상관계수 확인

: 범주형 자료인 COVID\_OX를 제외하고 상관계수 분석을 진행하였다.

독립변수 간 관련이 높은 변수가 존재한다면 다중공선성 문제가 발생할 것이라 예상하였다. 변수들 간의 상관관계를 확인해 본 결과, 강한 음의 상관관계를 보인 쌍은 없었지만, 강한 양의 상관관계를 보이는 쌍이 존재하였다. 무역수지, 상품수지, 경상수지 간에 높은 상관관계를 보였기에 이 중 음수 값이 없는 상품수지만 변수로 선택하였다. COFIX금리와 기준금리도 높은 상관성을 보이기에 보편적인 기준금리만을 이용하기로 결정하였다.

## 3) 모델 생성 과정

- 모든 변수를 포함한 다중회귀모형에서부터 시작하였다. 유의하지 않은 변수를 하나씩 제거하며 각 독립변수의 p-value가 유의할 때까지 진행하였다. 필요에 따라 절편을 제거한 경우 R-square 값이 1이 되는 결과를 초래하였다(진행된 모든 모형에서 절편을 제거한 경우 R-square값은 1이 되었다.). overfitting 문제가 예상되었고, ACF와 DW-test 결과도 유의하지 않으므로 해당 모형을 채택하지 않았다.
- 생성하는 모든 모델의 잔차의 자기상관이 매우 높아 어려움을 겪었다. 자기상관 문제를 해결하기 위해 독립변수로 추세 가변수를 추가하기도 하고 차분도 진행했으나 좋은 모델을 선정하기 힘들었다.
- 여러 모델 생성의 실패를 겪고 오차의 자기상관 문제를 해결하기 위한 방법으로 독립변수의 과거값을 새로운 독립변수로 추가하기로 결정하였다. 종속변수가 동월의 독립변수에 영향을 받은 것이 아닌, 과거의 값에 영향을 받았을 것이라 추측하고 모델을 생성하였다. 과거 값을 포함한 변수를 바탕으로 여러 모델의 구축을 진행하였다. R code에는 세 가지 모델만 정리하였지만 실제로는 100여 개 이상의 모델을 테스트했다.
- 최종 모형 생성을 위해서 각 변수의 유의성, Adjusted R-squared 값, ACF, DW-test, VIF가 모두 유의함을 확인하였으며, 이를 바탕으로 최종 모형을 채택하였다.

#### 4) 모형가정체크

- 가정1 : 변수 Y와 X의 관계는 선형(Linear)이다.
- 가정2 : X는 확률변수가 아닌 주어진 상수값이다.
- 가정3 : X값이 주어져 있을 때, 오차항의 평균은 0이다.
- 가정4: X값이 주어져 있을 때, 오차항의 분산은 시그마^2 로 모든 개체 i에 대해 동일하다.
- 가정5 : 서로 다른 개체간 오차항들은 상관되어있지 않다.
- 가정6 : X변수들이 여러 개 있을 때, X변수들 사이에는 선형관계가 없다
- 가정7 : 모형 설정 오류가 없음
- 가정8 : 오차항은 정규분포를 따름을 가정
- \* 본 모형은 가정2, 가정7을 만족한다는 전제하에 분석하였다.
- \* 가정1, 3, 4, 5, 6, 8은 최종모형 설정 후 진단하였고 7. 최종모형에서 확인가능하다.
- \* 가정1은 6-(1)에서 로그변환한 변수들이 시간에 따라 약한 양의 선형관계를 가짐을 확인하여 검증하였다.
- \* 가정3은 최종모형의 residual plot 확인을 통해 검증하였다.
- \* 가정4는 오차항의 등분산성 가정을 만족함을 확인하였다.
- \* 가정5를 만족시키기 위해 독립변수로 과거값을 사용하였다.
- \* 가정6은 최종모형의 VIF 값을 통해 만족 여부를 파악할 수 있다.
- \* 가정8은 오차항의 정규성 가정 만족 여부를 통해 확인하였다.

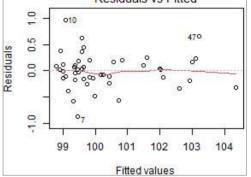
### 7. 최종모형

```
call:
lm(formula = final_target ~ final_interest:(final_COVID) + final_lnexchang
e:(COVID_OX) +
    final_lnoil + final_lnexchange + final_lnKOSPI, data = final)
Residuals:
     Min
               10
                     Median
-0.87142 -0.14990 0.01518 0.17462 0.96495
Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
                             -7.058e+01 1.860e+01 -3.796 0.000478 ***
(Intercept)
final_lnoil
                             1.395e+00 2.652e-01
                                                      5.258 4.88e-06 ***
final_Inexchange
                                                     7.705 1.69e-09 ***
                                        2.133e+00
                              1.643e+01
final_lnKOSPI 6.300e+00
final_interest:final_coVID 8.536e-06
                                         5.913e-01
                                                     10.656 2.20e-13 ***
                                                      3.639 0.000759 ***
                                         2.346e-06
final_lnexchange:COVID_OX 7.268e-02 2.822e-02
                                                     2.576 0.013705 *
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '1
Residual standard error: 0.3451 on 41 degrees of freedom
  (결측으로 인하여 1개의 관측치가 삭제되었습니다.)
Multiple R-squared: 0.9494, Adjusted R-squared: 0
F-statistic: 153.9 on 5 and 41 DF, p-value: < 2.2e-16
                                 Adjusted R-squared: 0.9432
```

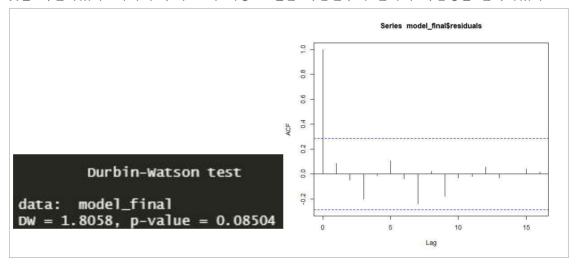
금리는 코로나 확진자의 수에 따라 영향을 받을 수 있을 것이라 생각해 교호작용을 고려하였으며, 환율의 경우 코로나 유무에 따른 영향을 파악하기 위해 교호작용을 확인하였다. 이를 통해 각각의 변수에 대한 조합을 검증하였다. 최종 모델의 summary 결과는 다음과 같다.

- 1. R-square를 통해 모델이 종속변수를 94% 가량 설명하는 것을 파악하였다. 또한 각 변수의 p-value가 0.05보다 작으므로 유의하다고 판단하였다. 각 변수가 모델에 미치는 영향을 확인하기 위해 estimate을 살펴보았다.
- 2. 교호작용이 없는 변수부터 해석하면, 코스피지수가 100% 증가하는 경우 물가는 6%p 증가한다. 또한, 원유의 가격이 100% 증가하면 물가지수는 1.4%p 가량 증가한다.

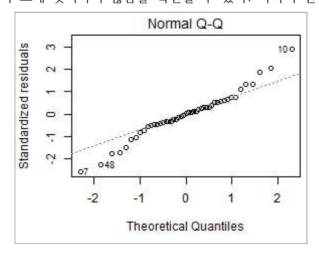
- 3. 교호작용이 있는 변수들은 다음과 같이 해석할 수 있다. 우선 환율과 코로나 발생여부의 교호작용을 분석 결과 코로나가 없던 시기 환율이 100% 증가하면 물가지수는 16%p 증가하였지만, 코로나의 유무를 고려하면 16.073%p 증가하여 약 0.073 기울기 차이를 보인다. 금리와 확진자 수의 교호작용을 보면, 확진자 수가 증가할수록 그 수에 비례하여 금리의 영향이 커짐을 알 수 있다. 확진자 한 명이 증가하면 금리는 평균적으로 물가지수를 8.536e-6 만큼 높아진다. 확진자 한 명의 영향은 굉장히 미미해 보이지만, 최대 확진자 수가 18만 명까지 발생했기에 이를 고려하면 그 영향력은 유의하다. 이렇게 COVID19와 금리, 환율의 각 교호작용을 바탕으로 COVID19가 물가지수 상승에 유의한 영향을 주었음을 확인하였다.
- 4. 이어서 잔차 plot을 통해 모델의 성능을 평가하였다. residual-fitted plot에서 0 주변에서 관측되는, 기울기가 0인 빨간 선을 기준으로 이상치는 7번, 10번, 47번 데이터로 보였으나 0을 중심으로 특정한 패턴을 가지고 있지 않음을 확인하였다. 또한 잔차가 0을 중심으로 크게 벗어나지않았기 때문에. 잔차의 등분산성을 만족한다고 평가하였다.



5. 더빈 왓슨 검정과 ACF를 통해 잔차의 독립성을 검정하였다. 더빈 왓슨 검정에서 p-value가 0.05보다 크므로 1차 자기상관이 없다는 귀무가설을 기각하지 못하였다. 또한 ACF 그래프에서 시차 0을 제외하면 모두 파란 구간 안에 포함되는 것을 확인하였다. 따라서 우리 조의 최종 모델은 독립변수의 잔차의 독립성을 만족하였다.



6. QQplot을 보면 잔차가 직선에서 크게 벗어나지 않음을 확인할 수 있다. 따라서 잔차의 정규성도 만족하였다.



7. VIF를 확인하면 10이하의 값을 취하므로, 다중공산성 문제가 없다고 판단할 수 있다.

#### 8. 결론

지난 4년간의 물가지수를 분석한 결과, 국내 물가를 잡기 위해 환율의 안정화가 필요하다고 판단하였다. 특히 우리나라는 경제적 대외의존도가 높기 때문에 환율상승이 물가상승에 가장 큰 요인으로 이어짐을 검증할 수 있었다. 따라서국내 물가 통제를 위해서는 원화가치가 떨어지지 않고 유지될 수 있도록 힘써야 할 것이다. 이에 더하여, 원유에 대한의존도 역시 높기에 유가 상승에 의한 경제 파급력이 크다는 것을 검증할 수 있었다. 다만, 기름 한 방울 나지 않는국가에서 여러 서비스업과 운송업, 철강업과 중화학 등의 산업을 지탱하기 위해서는 어쩔 수 없는 일이라 생각한다.

우리나라는 이를 조정하기 위해서 여러 가지 정책들을 시행중이다. 중앙은행은 통화정책의 수단과 환율의 변화에 대응하고자 공개시장을 운영해 BOK채권을 매도/매수하며, 여수신금리를 인상/인하시킴으로서 시장금리를 조절한다. 또한, 금리와 환율 조정을 위해 지준율을 조정하는 등 여러 가지 정책을 시행중이다. 이에 더하여, 2003년부터 2020년에 걸쳐 시행했던 승용차 요일제가 있었다. 비록 몇 년 전 폐지되었지만 교통체증과 온실가스 감소, 그리고 원유 사용량을 줄이기 위해 시행한 제도였다. 이러한 다양한 정책을 바탕으로 우리 조는 우리나라의 물가 안정에 기여할 수 있는 정책에 대해 고민하였다.

결론적으로 환율안정화가 물가안정화에 가장 유의한 영향을 준다고 판단하여, 우리 조는 물가와 환율 안정을 위해 통화 스와프 체결을 제안한다. 실제로 외환스왑은 서울외환시장의 최대거래 외환상품으로써, 통화를 단기간에 싸게 빌릴수 있다는 측면에서 그 유용함을 나타내고 있다. 이러한 장점을 살려 현재 RCEP(역내 포괄적 경제동반자협정)에서 아세안+3 통화스와프의 원활한 가동을 위한 협력이 추진되는 중이다. 이는 통화스와프 뿐 아니라 RCEP 회원국 간 공급망 안정화와 중소기업 경쟁력 강화 등을 통해 금융시장의 변동성을 대비하는 목적이다. 결과적으로. 이러한 협정이 우리나라의 물가 안정에 기여할 것이라 예상한다.

# 9. 한계점 및 소감

프로젝트를 진행하며 느낀 점은 선형회귀를 통해 물가의 움직임이나 분산을 파악하기에 어려움이 있었다는 것이다. 이는 물가지수가 우리가 선정한 지표만이 아닌 다양한 경제 지표의 영향을 받고, 더 나아가 사회 분위기에 의해서도 영향을 받는 종합 경제지수이기 때문일 것이라 생각한다. 따라서 보다 많은 범주형 변수들을 사용해 수치로 표현하기 힘든 부분을 짚었더라면 이상치 분석에 도움이 되지 않았을까 생각한다. 또한 원한 만큼의 데이터를 모으지 못한 점이다소 아쉽다. 데이터 중에 최근 5개년의 자료만 존재하는 경우가 있어 처음 목표로 했던 10개년 데이터를 추출할 수 없었기 때문이다. 만약 더욱 많은 데이터를 이용했더라면 더 유의한 결과를 얻지 않았을까 하는 아쉬움이 남는다.

그럼에도 프로젝트를 진행하면서 많은 것을 배웠다. 특히 모델을 생성하고 가정을 검증하면서 잔차의 분산과 자기상 관을 해결하기 위해 겪은 실패들이 기억에 남는다. 이 실패를 통해 최적의 모델은 단순히 함수의 사용으로 구축하는 것이 아닌 교호작용과 추세를 잡기 위한 변수, 절편 제거 등 많은 고민을 바탕으로 이뤄짐을 느꼈다. 각 단계별로 유의한 모델을 만들기 위해 노력하며 수치를 기반으로 결론을 내리는 법을 배운, 우리에게 많은 것이 남은 프로젝트였다.

## 10. 자료 출처

우리 조가 프로젝트를 위해 참고한 레퍼런스이다. 주제 선정을 위해 다양한 물가 관련 기사들과 한국은행, 경제연구원의 리포트를 참고하였다. 또한 데이터는 대한민국 통계청 KOSIS, 질병관리청과 몇몇 통계 사이트로부터 추출하였다.

# 1) 프로젝트 배경 선정 출처

[고공행진 물가에…5월 경제고통지수 21년 만에 최대 - 뉴시스]

https://mobile.newsis.com/view.html?ar\_id=NISX20220620\_0001912877

[일자리 줄고, 물가 치솟고…'서민경제고통지수' 역대 최대\_중앙일보]

https://www.joongang.co.kr/article/25040445#home

[김밥 한 줄·떡볶이 1인분도 "먹기 겁나네"…외식물가 껑충\_한경]

https://www.hankyung.com/economy/article/202211021066i

[무역수지악화→환율상승→물가급등..'악순환의 고리'에 갇힌 한국경제[양철민의 경알못]\_서울경제]

https://www.sedaily.com/NewsView/268D66XTP4

[7월 상품수지 10년만 적자…소비자물가 상승 '도미노 효과' 우려\_소비자경제]

http://www.dailycnc.com/news/articleView.html?idxno=213507

- 코로나 19가 물가에 미치는 영향 분석 (한국은행 BOK 이슈노트, 2020.11.10.)

https://www.bok.or.kr/portal/bbs/P0002353/view.do?nttId=10061158&menuNo=200433

- OECD 1위의 경제 원유의존도, 그 개선이 시급하다 - 국제유가 상승이 산업경쟁력에 미치는 영향과 시사점 (현대 경제연구원, 2022.02.09.)

https://hri.co.kr/kor/report/report-view.html?mode=1&uid=30296

- 통화정책 효과의 파급 (한국은행 통화정책국 정책분석팀)

https://www.bok.or.kr/portal/main/contents.do?menuNo=200289

- 코스피, 美 물가 충격에 2.5%대 급락…하루 만에 2,400선 하회(상보) (연합인포맥스, 2022.09.14.)

https://news.einfomax.co.kr/news/articleView.html?idxno=4233144

- 코픽스 (KDI 경제정보센터 click 경제교육, 2010.08.31.)

https://eiec.kdi.re.kr/material/clickView.do?click\_yymm=201512&cidx=1309

## 2) 데이터 출처

- 소비자물가지수(종속변수 ): [KOSIS 소비자물가지수, 20221102]

https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT\_1J20003&vw\_cd=MT\_ZTITLE&list\_id=P2\_6&seqNo=&lang\_mode=ko&language=kor&obj\_var\_id=&itm\_id=&conn\_path=MT\_ZTITLE

- 국제수지 (정량): [KOSIS 경상수지, 20221103]

- 국제유가 도입가 (정량) : [KOSIS 국제유가, 20221018]

https://kosis.kr/statHtml/statHtml.do?orgId=392&tblId=DT\_AA123&vw\_cd=MT\_ZTITLE&list\_id=T\_21&seqNo=&lang\_mode=ko&language=kor&obj\_var\_id=&itm\_id=&conn\_path=MT\_ZTITLE

- 경상수지 (정량): [KOSIS 경상수지, 20221111]

 $\frac{\text{https://kosis.kr/statHtml/statHtml.do?orgId=301\&tblId=DT\_301Y017\&vw\_cd=MT\_ZTITLE\&list\_id=S2\_301008\_00}{1\&\text{seqNo=\&lang\_mode=ko\&language=kor\&obj\_var\_id=\&itm\_id=\&conn\_path=MT\_ZTITLE}}$ 

- 상품수지 (정량): [KOSIS 상품수지, 20221111]

https://kosis.kr/statHtml/statHtml.do?orgId=301&tblId=DT\_301Y017&vw\_cd=MT\_ZTITLE&list\_id=S2\_301008\_00 
1&seqNo=&lang\_mode=ko&language=kor&obj\_var\_id=&itm\_id=&conn\_path=MT\_ZTITLE

-무역수지 (정량): [KOSIS 수출입총괄, 20221021]

https://kosis.kr/statHtml/statHtml.do?orgId=134&tblId=DT\_134001\_001&vw\_cd=MT\_ZTITLE&list\_id=&scrId=&seqNo=&lang\_mode=ko&obj\_var\_id=&itm\_id=&conn\_path=E1&docId=0388621915&markType=S&itmNm=%EC%A0%84%EA%B5%AD

- 기준금리 (정량): [한국은행 기준금리 추이]

https://www.bok.or.kr/portal/singl/baseRate/list.do?dataSeCd=01&menuNo=200643

- 환율 (정량): [환율 (USD/KRW) ,20221103]

https://kr.investing.com/currencies/usd-krw-historical-data

- 코스피지수 (정량): [코스피지수 ,20221103]

https://kr.investing.com/indices/kospi-historical-data

- COFIX 금리 (정량): [은행연합회 소비자포털 COFIX 금리, 20221111]

https://portal.kfb.or.kr/fingoods/cofix.php

- 코로나19 확진자 수 : [질병관리청 코로나바이러스19 누적 확진자 통계]

https://ncov.kdca.go.kr/