**Jawaharlal Nehru University**
**Centre for International Trade and Development**

**"Factors Affecting Inflation Rates of Different Countries of the World in Fiscal Year 2020-21"**

# Advanced Econometrics

Submitted To: -

**Prof Amit S Ray**

Group Members: -

**Ritika Priya**

**Aayush Neelkanth**

# Introduction

In our project, we have modeled the inflation rates of different countries of the world in the year 2020-21, i.e., the covid year. We regress inflation rates on unemployment, real interest rates, deaths, fuel import, and exports. We have used Stata for our econometric analysis.

## Data Source:

- We have collected Data of Inflation, Unemployment, Real Interest Rate, Death Rate per 1000, Exports of Goods and Services (% of GDP), Fuel Imports, Money Supply, and Political Stability using the World Development Indicator of the World Bank.
- We have collected data of crime rate from Numbeo. Numbeo is a crowd-sourced global database of quality-of-life data: housing indicators, perceived crime rates, healthcare quality, transport quality, and other statistics.

## Methods used:

1. For Multicollinearity: -
    a. Testing
        i. Pairwise Correlation
        ii. Variance Inflation Factor (VIF)
2. For Heteroscedasticity: -
    a. Testing
        i. White Heteroscedasticity
        ii. Breusch Pagan Test
    b. Correction Method:
        i. Robust Linear Regression
3. For Omitted Variable Bias: -
    a. Testing
        i. Ramset RESET Test
4. For Endogeneity / Reverse Causality -
    a. Method
        i. 2SLS

## Our regression equation: -

$$Inflation = \beta_0 + \beta_1 * Unemployment + \beta_2 * RealInterestRate + \beta_3 * DeathRateper1000 + \beta_4 * FuelImports + \beta_5 * Exports$$

According to the theory,

Inflation: Annual percentage of change in the cost to the average consumer of acquiring a basket of goods and services.

Unemployment: Percentage of People Unemployed out of the total population of the country. We expect unemployment to be negatively related to inflation, as it is shown by the popular Phillips curve.

Deathrateper1000: Number of Deaths Per Thousand. This number constitutes the casualty from covid as well. We expect this number to be positive.

RealInterestRate: The rate which shows the real costs to a borrower and real yield to the lender. According to the fisher equation, Real interest rate = nominal interest rate – rate of inflation i.e., the real interest rate and inflation are negatively related.

FuelImports: It is the fuel imports as a percentage of merchandise imports.
We expect the relationship between fuel imports and inflation to be positive. This is because as rise in imports imply the rise in demand which leads to increase in price.

Exports: Total Exports as a percentage of GDP.
It impacts inflation through the demand side. It is a case of demand-pull inflation. Rise in exports signifies rise in demand which shifts the AD curve to the right, hence, increasing the inflation. Hence, we expect this number to be positive as well.

CrimeRate: Crime Rate is our Instrumental Variable.
Crime Index is an estimation of overall level of crime in a given city or a country. We consider crime levels lower than 20 as very low, crime levels between 20 and 40 as being low, crime levels between 40 and 60 as being moderate, crime levels between 60 and 80 as being high and finally crime levels higher than 80 as being very high.

Political: Crime Rate is our Instrumental Variable.
Political Stability and Absence of Violence/Terrorism measures perceptions of the likelihood of political instability and politically-motivated violence, including terrorism. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e., ranging from approximately -2.5 to 2.5.

*Note: - We are doing our analysis at 10% level of significance.*

## Ordinary Least Square Regression

Firstly, performing the OLS regression, we get the following result.

```
. reg Inflation Exports RealInterestRate DeathRateper1000 FuelImports Unemployement
```

| Source | SS | df | MS | Number of obs | = | 203 |
|---|---|---|---|---|---|---|
| | | | | F(5, 197) | = | 25.50 |
| Model | 132941.925 | 5 | 26588.385 | Prob > F | = | 0.0000 |
| Residual | 205394.492 | 197 | 1042.61163 | R-squared | = | 0.3929 |
| | | | | Adj R-squared | = | 0.3775 |
| Total | 338336.416 | 202 | 1674.93275 | Root MSE | = | 32.289 |

| Inflation | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| Exports | -.0685742 | .0712064 | -0.96 | 0.337 | -.2089989 | .0718505 |
| RealInterestRate | -2.668264 | .2394183 | -11.14 | 0.000 | -3.140416 | -2.196112 |
| DeathRateper1000 | -1.293288 | .7257311 | -1.78 | 0.076 | -2.724487 | .1379112 |
| FuelImports | .6075692 | .3089789 | 1.97 | 0.051 | -.0017616 | 1.2169 |
| Unemployement | .1938946 | .3710442 | 0.52 | 0.602 | -.5378339 | .9256231 |
| _cons | 22.53776 | 6.732319 | 3.35 | 0.001 | 9.261091 | 35.81442 |

From the Table we can see that,

- **Unemployment** is coming out to be positive which should be negative by the theory of Phillips curve. But several studies have shown that sometimes positive relationships also exist between the two.

- **Deaths per 1000** is coming out to be negative which is also opposite to our understanding. This can be due to the time effect also. As the number of deaths peaked in the latter half of the year 2020, the inflation also began to fall gradually with time. This can probably explain the negative sign.
- **Real interest rates** should be negative which is coming out to be negative only, justifying our hypothesis.
- **Fuel imports** are also coming out to be positive.
- **Exports as a percentage of GDP** should come out to be positive but it is negative here but it should also be noted that export is an insignificant variable here as its p-value (0.337) > 0.10.

## Multicollinearity: -

This means the existence of a perfect or exact linear relationship among two or more explanatory variables. In numeric terms, multicollinearity is said to exist if the following condition holds,

$\lambda_1 X_1 + \lambda_2 X_2 + \lambda_3 X_3 + \ldots\ldots\lambda_k X_k + v_i = 0$
Where $v_i$ is a stochastic term.

Consequences of ignoring multicollinearity-

1. OLS estimates are still unbiased, consistent and efficient.
2. Standard Errors are higher.
3. t statistics are insignificant.

How do we detect multicollinearity-

1. Through **pairwise correlation** coefficients.

```
. pwcorr Unemployement FuelImports DeathRateper1000 RealInterestRate Exports
```

| | Unempl~t | FuelIm~s | Dea~1000 | RealIn~e | Exports |
|---|---|---|---|---|---|
| Unemployem~t | 1.0000 | | | | |
| FuelImports | 0.1149 | 1.0000 | | | |
| DeathRa~1000 | 0.1439 | 0.0661 | 1.0000 | | |
| RealIntere~e | 0.0066 | 0.0179 | -0.1590 | 1.0000 | |
| Exports | -0.0743 | 0.0574 | 0.1459 | -0.0525 | 1.0000 |

   Here, we can see that the pairwise correlation coefficients between any of the explanatory variables is **not** significantly higher.

2. Through **VIF**, which is a measure of the extent to which variance of OLS estimates are inflated due to the presence of multicollinearity.

$$VIF_i = \frac{1}{1 - Ri^2} = \frac{1}{Tolerance}$$

```
. vif
```

| Variable | VIF | 1/VIF |
|---|---|---|
| DeathRa~1000 | 1.08 | 0.929734 |
| Unemployem~t | 1.04 | 0.957259 |
| Exports | 1.04 | 0.965213 |
| RealIntere~e | 1.03 | 0.972382 |
| FuelImports | 1.02 | 0.980077 |
| Mean VIF | 1.04 | |

Usually, the VIF > 10 is considered to be a sure shot sign of multicollinearity. Here, the VIF which we are getting is very low, so we conclude that there does not exist a problem of multicollinearity in our model.

## Heteroscedasticity-

One of the most important assumptions of CLRM is the assumption of homoscedasticity which states that the variance of each disturbance term $u_i$, conditional on chosen values of the explanatory variable, is some constant $\sigma^2$.

$E(u_i^2 | x_i) = \sigma^2$

Consequences of ignoring heteroscedasticity-

1. Estimates are unbiased, consistent and inefficient.
2. OLS estimates are not BLUE.

Detecting Heteroscedasticity: -

a. White's heteroskedasticity test-

```
. imtest, white

White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(20)      =    195.03
         Prob > chi2   =    0.0000

Cameron & Trivedi's decomposition of IM-test
```

| Source | chi2 | df | p |
|---|---|---|---|
| Heteroskedasticity | 195.03 | 20 | 0.0000 |
| Skewness | 79.17 | 5 | 0.0000 |
| Kurtosis | 1.13 | 1 | 0.2884 |
| Total | 275.33 | 26 | 0.0000 |

From the p value being less than 0.1, we can conclude that we reject the null hypothesis and there exists heteroscedasticity in the model.

b. Breusch-Pagan test-

```
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of Inflation

         chi2(1)      =  1598.87
         Prob > chi2  =   0.0000
```

We can see that the p value < level of significance (0.1) and hence we can conclude that we reject the null hypothesis and heteroscedasticity is present in the model.

## Solution for Heteroscedasticity: -

We use robust standard error method. This would correct our result of heteroscedasticity and give us robust standard errors.

$$se(\hat{\beta}_j) = \sqrt{\frac{\sum_{i=1}^{N} e_{ji}^2 e_i^2}{RSS_j^2} * \frac{N}{N-K-1}}$$

$e_j$ = It is the residual when we regress $j^{th}$ independent variable on all other independent variables.
$e_i$= Residual from the normal regression
$RSS_j$= It is the residual from regressing variable j on all explanatory variables.
N= sample size
N-K-1 = degrees of freedom

```
. reg Inflation Unemployement FuelImports RealInterestRate DeathRateper1000 Exports, robust

Linear regression                               Number of obs     =        203
                                                F(5, 197)         =       1.02
                                                Prob > F          =     0.4064
                                                R-squared         =     0.3929
                                                Root MSE          =     32.289

------------------------------------------------------------------------------
             |               Robust
   Inflation |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 Unemployement |   .1938946   .3165466     0.61   0.541    -.4303602    .8181495
   FuelImports |   .6075692   .4021804     1.51   0.132    -.1855624    1.400701
RealInterestRate |  -2.668264   1.56014    -1.71   0.089    -5.744983    .4084547
DeathRateper1000 |  -1.293288   .7938703   -1.63   0.105    -2.858863    .2722872
       Exports |  -.0685742   .0419267    -1.64   0.104     -.151257    .0141086
         _cons |   22.53776    11.887     1.90   0.059    -.9043503    45.97986
------------------------------------------------------------------------------
```

## Omitted Variable Bias-

We perform the Ramsey RESET test to check for omitted variable bias and find that there are omitted variables present in our model. To rectify that, we tried to include many variables which have impacted inflation significantly like money supply, public debt but couldn't remove the omitted variable bias.

```
. ovtest

Ramsey RESET test using powers of the fitted values of Inflation
       Ho:  model has no omitted variables
                 F(3, 194) =     371.65
                  Prob > F =     0.0000
```

## Issue of reverse causality-

Issue of reverse causality arises when there is a two-way dependency between dependent and independent variables. To put it simply, Y is dependent on X and X is also dependent on Y. Here inflation and unemployment are dependent on each other. So, we will be using political stability and crime rate as an IV for unemployment.
Now, there are two conditions required for a variable to qualify as an IV: -

1. It should be uncorrelated with the error term. Political stability and Crime rate are uncorrelated with the error term.
2. The instrument variables should be correlated with endogenous variable. Political stability and Crime rate are correlated with Unemployment.

We first start off with performing the 2sls regression,

```
. ivregress 2sls Inflation Exports RealInterestRate DeathRateper1000 FuelImports (Unemployement = Political CrimeRate)

Instrumental variables (2SLS) regression          Number of obs   =        203
                                                  Wald chi2(5)    =     112.96
                                                  Prob > chi2     =     0.0000
                                                  R-squared       =     0.2781
                                                  Root MSE        =     34.687

------------------------------------------------------------------------------
   Inflation |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Unemployement |   2.458851   1.495542     1.64   0.100    -.4723569    5.39006
      Exports |  -.0240378   .0815754    -0.29   0.768    -.1839227    .1358471
RealInterestRate |  -2.703623   .2581763   -10.47   0.000     -3.20964   -2.197607
DeathRateper1000 |  -1.971421   .8910952    -2.21   0.027    -3.717936   -.2249066
  FuelImports |   .3975013   .3578311     1.11   0.267    -.3038348    1.098837
        _cons |   10.53857   10.51757     1.00   0.316    -10.07548    31.15262
------------------------------------------------------------------------------
Instrumented:  Unemployement
Instruments:   Exports RealInterestRate DeathRateper1000 FuelImports Political
               CrimeRate
```

We see that our endogenous variable **Unemployment** becomes significant in this case once we use political stability and crime rate as the IVs in this case.

## Test of endogeneity:

We then perform the test of endogeneity to see whether the variable is indeed endogenous.

```
. estat endog

  Tests of endogeneity
  Ho: variables are exogenous

  Durbin (score) chi2(1)          =  2.93602   (p = 0.0866)
  Wu-Hausman F(1,196)             =  2.87638   (p = 0.0915)
```

We see here that the variable is endogenous as the p value< 0.10 (level of significance) and there exists the problem of endogeneity in our model

## To check whether the instruments are weak or strong.

We use First Stage Regression: -

**Unemployment = $\beta_0$ + $\beta_1$*real interest rate + $\beta_2$*death per 1000 + $\beta_3$*fuel price + $\beta_4$*Exports + $\beta_4$*Political + $\beta_4$*CrimeRate + $v_i$**

Here, the R-sq denotes the coefficient of determination from running the first stage regression.

```
estat firststage

First-stage regression summary statistics
```

| Variable | R-sq. | Adjusted R-sq. | Partial R-sq. | F(2,196) | Prob > F |
|---|---|---|---|---|---|
| Unemployem~t | **0.1107** | **0.0835** | **0.0710** | **7.49352** | **0.0007** |

```
Minimum eigenvalue statistic = 7.49352

Critical Values                          # of endogenous regressors:    1
Ho: Instruments are weak                 # of excluded instruments:     2
```

| | 5% | 10% | 20% | 30% |
|---|---|---|---|---|
| 2SLS relative bias | | (not available) | | |

| | 10% | 15% | 20% | 25% |
|---|---|---|---|---|
| 2SLS Size of nominal 5% Wald test | **19.93** | **11.59** | **8.75** | **7.25** |
| LIML Size of nominal 5% Wald test | **8.68** | **5.33** | **4.42** | **3.92** |

Here, the R-sq comes out to be approximately 11 percent. The column "F (1, 196)" is an F statistic for the joint significance of $\beta_1$, $\beta_2$, $\beta_3$, and $\beta_4$, the coefficients on the additional instruments. The row

marked 2SLS Size of nominal 5% Wald Test contains critical values, on which we report our instrument as weak or strong. Here, our instrument variable comes out to be strong at 25% critical value. We can see that p-value < 0.10, which signifies that the model is jointly significant.

## Testing whether the instruments are good: -

We perform the **Sargan test** to see whether the instruments are good. The null hypothesis is that the instruments are valid.

```
estat overid

Tests of overidentifying restrictions:

Sargan (score) chi2(1) =  .120808  (p = 0.7282)
Basmann chi2(1)        =  .116712  (p = 0.7326)
```

We see that the p-value > 0.1 which implies that we cannot reject the null hypothesis therefore the instruments are-in fact valid.

## Hausman test: -

Ho: β_hat OLS = β_hat IV
H1: β_hat OLS ≠ β_hat IV
Let us say two independent variables $X_1$ and $X_2$ Suppose $X_2$ is our potentially endogenous regressor.

1. Regress $X_2$ on the other explanatory variable and the IV which gives us $X_2$_hat.
2. Then we calculate the endogenous component u_hat = $X_2$- $X_2$_hat.
3. Now we regress the dependent variable Y on all the explanatory variables and u_hat.
4. Our regression equation has an endogenous component with estimator ρ.
   $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \rho u\_hat + u$
   If $\rho = 0$
   - Endogenous part of X2 has no effect on Y.
   - Orthogonality problem did not exist i.e., β_hat OLS = β_hat IV

   This test is not robust, it assumes homoscedastic error which is not guaranteed.

We have performed the test step by step firstly, following the above-given steps.

```
. reg Unemployement Exports RealInterestRate DeathRateper1000 FuelImports Political CrimeRate
```

| Source | SS | df | MS | Number of obs | = | 203 |
|---|---|---|---|---|---|---|
| | | | | F(6, 196) | = | 4.07 |
| Model | 876.066717 | 6 | 146.011119 | Prob > F | = | 0.0007 |
| Residual | 7035.12085 | 196 | 35.8934737 | R-squared | = | 0.1107 |
| | | | | Adj R-squared | = | 0.0835 |
| Total | 7911.18757 | 202 | 39.1642949 | Root MSE | = | 5.9911 |

| Unemployement | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| Exports | -.0034416 | .0139488 | -0.25 | 0.805 | -.0309507 | .0240675 |
| RealInterestRate | .0278721 | .0445264 | 0.63 | 0.532 | -.0599403 | .1156844 |
| DeathRateper1000 | .2637988 | .1334892 | 1.98 | 0.050 | .0005394 | .5270583 |
| FuelImports | .0749739 | .0574345 | 1.31 | 0.193 | -.038295 | .1882428 |
| Political | -1.487346 | .4561276 | -3.26 | 0.001 | -2.386894 | -.5877981 |
| CrimeRate | .0255873 | .0173253 | 1.48 | 0.141 | -.0085806 | .0597552 |
| _cons | 4.342753 | 1.251404 | 3.47 | 0.001 | 1.874808 | 6.810699 |

```
. predict uhat, residual

. reg Inflation Exports RealInterestRate DeathRateper1000 FuelImports Unemployement uhat
```

| Source | SS | df | MS | Number of obs | = | 203 |
|---|---|---|---|---|---|---|
| | | | | F(6, 196) | = | 21.93 |
| Model | 135912.572 | 6 | 22652.0953 | Prob > F | = | 0.0000 |
| Residual | 202423.844 | 196 | 1032.77472 | R-squared | = | 0.4017 |
| | | | | Adj R-squared | = | 0.3834 |
| Total | 338336.416 | 202 | 1674.93275 | Root MSE | = | 32.137 |

| Inflation | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| Exports | -.0240378 | .0755784 | -0.32 | 0.751 | -.1730891 | .1250135 |
| RealInterestRate | -2.703623 | .2391965 | -11.30 | 0.000 | -3.175353 | -2.231894 |
| DeathRateper1000 | -1.971421 | .8255864 | -2.39 | 0.018 | -3.599594 | -.3432481 |
| FuelImports | .3975013 | .3315252 | 1.20 | 0.232 | -.2563132 | 1.051316 |
| Unemployement | 2.458851 | 1.385597 | 1.77 | 0.078 | -.2737422 | 5.191445 |
| uhat | -2.438146 | 1.437596 | -1.70 | 0.091 | -5.273288 | .3969969 |
| _cons | 10.53857 | 9.744367 | 1.08 | 0.281 | -8.678698 | 29.75584 |

Here, the p-value of uhat is 0.091 which is < 0.10. Hence, we conclude that there is indeed endogeneity present in the model and the β_ols and β_iv are different from each other.

By performing the test through the stata command-

```
. hausman B_iv B_ols
```

| | Coefficients | | | |
|---|---|---|---|---|
| | (b) B_iv | (B) B_ols | (b-B) Difference | sqrt(diag(V_b-V_B)) S.E. |
| Unemployem~t | 2.458851 | .1938946 | 2.264957 | 1.448783 |
| Exports | -.0240378 | -.0685742 | .0445364 | .039802 |
| RealIntere~e | -2.703623 | -2.668264 | -.0353592 | .0966121 |
| DeathRa~1000 | -1.971421 | -1.293288 | -.6781333 | .5170735 |
| FuelImports | .3975013 | .6075692 | -.210068 | .1804859 |

```
b = consistent under Ho and Ha; obtained from ivregress
B = inconsistent under Ha, efficient under Ho; obtained from regress

Test:  Ho:  difference in coefficients not systematic

            chi2(5) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                    =        2.44
          Prob>chi2 =      0.7849
```

We see that here the p-value> 0.10 which implies that β_ols and β_iv are not significantly different from each other.

## Conclusion: -

- Our model contained heteroscedasticity, for which we corrected through robust standard error method.
- Our model also suffers from omitted variable bias but even after including some other relevant variables like money supply and public debt, we are not able to correct for it.
- Our model has a problem of reverse causality between inflation and unemployment. We have taken the IVs crime rate and political stability which are valid and strong as well.