"The Impact of Trade Dynamics, Inflation, Tax and Unemployment on Economic Growth, An Empirical Analysis"

by Aman Dongre

Abstract :

In this empirical paper, I would like to discuss and investigate the effects of tax, effects of Imports/exports, Inflation, and unemployment and how these variables would affect the economic growth that is GDP growth rate for the country.

We would discuss how these variables will cause the growth rate to either increase or decrease in the subsequent years. We'll see changes in these variables according to the economic changes which would also cause the growth rate of the economy to change. Also, we will notice if some other factors could affect the GDP growth rate and whether they should be included in the model. This paper will provide valuable insights into understanding the dynamics of economic growth at the national level.

Introduction :

Economic research has long been interested in the relationship between different macroeconomic parameters and economic growth. Through an empirical analysis, we examine in this paper how trade dynamics, inflation, tax policy, population, and unemployment affect economic growth. We seek to determine the strength and direction of these associations using econometric approaches to further our understanding of the factors that influence economic growth. The results of this study have significance for practitioners and policymakers who are trying to come up with creative ways to support sustainable economic development in a variety of settings.

Literature Review :

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10030258/ - discusses **The impact of the COVID–19 pandemic on global GDP growth** by Joseph E. Gagnon, Steven B. Kamin and John Kearns. It was one of the first attempts to gauge the effect of the COVID–19 pandemic on real GDP.

Similarities :

- The dependent variable real GDP in this research paper and the GDP growth rate in our study are the same except the GDP growth rate is the Percentage change in real GDP.
- Both papers have an analysis of how global trade(import/export) affects the GDP although the sample size is different as this one is limited to the COVID -19 dynamics while ours includes all the 3 major events from 1990 – 2022.

While both the papers discuss the dynamics of trade, inflation on GDP we are focusing more on discussing more about how these variables affect GDP growth rate and what variables would be more significant than the others in our model.

Econometric Models and Methodology Used :

In our study for the GDP growth rate, the econometric models in EViews would be **ARMA** models which are used to analyze time series data. They also help to capture the autocorrelation of the data and OLS models with Operational Least square methodology along with **forecasting**, **Serial correlations Testing**, **Wald test** for Hypothesis Testing, **Heteroskedasticity** Tests, Stationary or non – stationary variables, ARMA maximum Likelihood tests.

Empirical Analysis :

In our econometric model for GDP growth rate, we will use the sample size from 1990q1 to 2022q4 where q represents the quarters in a year from 1 to 4 which is a **time-series** data model. The data is obtained from the FRED economic data | St. Louis FED. (https://fred.stlouisfed.org/) , For GDP growth rate - https://fred.stlouisfed.org/series/USAGDPRQPSMEI

The dependent variable for this model would be the GDP growth rate and the independent variables would be the **tax in receipts, import/export, real net exports(excluding imports), Unemployment rate, Federal Funds effective rate, Nasdaq, inflation,** and **population** with every variable being quarterly adjusted.

The main tests are Wald tests to check whether the variables are individually significant or jointly significant in the model. Also, the model could **suffer** from multicollinearity, heteroskedasticity, and autocorrelation hence we are doing those tests to make sure the model is adjusted before moving forward to **forecasting**.

For this model the estimated equation would be which can be used in Eviews -

GDP growth rate = B_0 + B_1 imp_gr + B_2 exp_gr + B_3 Fed er+ B_4 inf + B_5 Nasdaq_gr+ B_6 rnet exp + B_7 Tax +B_8 Unemp + u (Growth rate same period previous year)

(We will use a 5% significance level for this model)

GDP growth rate = B_0 +B_1*Population + B_2*Tax + B_3*imp_gr + B_4*exp_gr + B_5*Fed_er+ B_6*inf + B_7*Nasdaq_gr+ B_8*rnet exp + B_9*Tax +B_{10}*Unemp + u

The Final equation is after removing the insignificant variable –

B_1*imp + B_2*exports + B_3*Fed_er+ B_4*inf + B_5*Nasdaq+ B_6*rnet exp + B_7*Tax +B_8*Unemp + u .

1.1

We can estimate the model by least squares and the estimated output table can be seen below:–

Dependent Variable: GDP_GROWTH_RATE
Method: Least Squares
Date: 04/10/24 Time: 19:23
Sample: 1990Q1 2022Q4
Included observations: 132

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	29.08662	8.956601	3.247506	0.0015
POPULATION	-6.54E-05	3.58E-05	-1.827059	0.0701
TAX	-0.002148	0.001615	-1.330399	0.1859
IMP_GR	0.005157	0.003280	1.572542	0.1184
EXP_GR	-0.004271	0.004365	-0.978461	0.3298
FED_ER	-0.507360	0.141231	-3.592408	0.0005
INF	0.424531	0.267852	1.584947	0.1156
NASDAQ_GR	0.000330	0.000153	2.158061	0.0329
RNET_EXP	0.005081	0.003665	1.386313	0.1682
UNEMP	-0.993893	0.162262	-6.125224	0.0000

R-squared	0.378390	Mean dependent var	2.454822
Adjusted R-squared	0.332533	S.D. dependent var	2.071563
S.E. of regression	1.692439	Akaike info criterion	3.962953
Sum squared resid	349.4507	Schwarz criterion	4.181347
Log-likelihood	-251.5549	Hannan-Quinn criteria.	4.051698
F-statistic	8.251605	Durbin-Watson stat	0.715411
Prob(F-statistic)	0.000000		

- The model has an **r-squared** value of 0.36 which means that **36%** of the values in this model of GDP growth rate can be explained by these variables.
- The **negative coefficients** in the output of the variables represent that a 1 unit decrease in those variables would **increase the GDP** growth rate by the same number.
- If we remove the population and tax variables one after the other, all the other variables except the **population** and **Tax** variable have a p-value of less than 0.05 which means they are significant in the model.

1.2

Therefore we could try to run the output again excluding the population variable first then estimating again excluding the Tax variable to see if it makes any difference.

Dependent Variable: GDP_GROWTH_RATE
Method: Least Squares
Date: 04/10/24 Time: 19:25
Sample: 1990Q1 2022Q4
Included observations: 132

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12.35717	1.886418	6.550602	0.0000
IMP_GR	0.006800	0.003207	2.120351	0.0360
EXP_GR	-0.008950	0.003829	-2.337252	0.0210
FED_ER	-0.480495	0.136366	-3.523576	0.0006
INF	0.626239	0.251438	2.490629	0.0141
NASDAQ_GR	0.000253	0.000123	2.064998	0.0410
RNET_EXP	0.008988	0.003211	2.798644	0.0060
UNEMP	-1.009843	0.155433	-6.496970	0.0000

R-squared	0.354281	Mean dependent var	2.454822
Adjusted R-squared	0.317829	S.D. dependent var	2.071563
S.E. of regression	1.710980	Akaike info criterion	3.970701
Sum squared resid	363.0039	Schwarz criterion	4.145416
Log likelihood	-254.0663	Hannan-Quinn criter.	4.041697
F-statistic	9.719137	Durbin-Watson stat	0.726481
Prob(F-statistic)	0.000000		

- It can be observed now that every variable in this model is significant.

1.3

Also by Wald test, we can confirm this as shown below:-

Wald Test:
Equation: EQ01

Test Statistic	Value	df	Probability
F-statistic	2.365854	(2, 122)	0.0982
Chi-square	4.731707	2	0.0939

Null Hypothesis: C(2)=0,C(3)=0
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2)	-6.54E-05	3.58E-05
C(3)	-0.002148	0.001615

Restrictions are linear in coefficients.

Hence the result is the same that these variables are jointly insignificant too due to their p-values being greater than 0.05.

2

Hypothesis Testing

Now we will do hypothesis testing for these variables which can be done through the Wald Test to determine whether the import and export variables affect the dependent variable GDP growth rate equally or not :

- For that the Null Hypothesis will be H_0 – The import and export variables affect the dependent variable equally.
- Alternative hypothesis H_1 – The variables do not affect the GDP growth rate equally.

Wald Test:
Equation: EQ01

Test Statistic	Value	df	Probability
t-statistic	-3.765193	124	0.0003
F-statistic	14.17668	(1, 124)	0.0003
Chi-square	14.17668	1	0.0002

Null Hypothesis: C(4)=C(5)
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(4) - C(5)	-1.106734	0.293938

Restrictions are linear in coefficients.

Here the p-value(0.0003) is less than our alpha (0.05) hence we can reject the null hypothesis and say that the import/export variables do not affect the GDP growth rate equally.

2.1

Testing for heteroskedasticity in the error term of the model to describe the unequal spread or error:-

Heteroskedasticity Test: Breusch-Pagan-Godfrey
Null hypothesis: Homoskedasticity

F-statistic	3.647776	Prob. F(7,124)	0.0013
Obs*R-squared	22.54026	Prob. Chi-Square(7)	0.0020
Scaled explained SS	65.56584	Prob. Chi-Square(7)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 04/10/24 Time: 19:32
Sample: 1990Q1 2022Q4
Included observations: 132

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.353278	7.314311	-0.458454	0.6474
IMP_GR	0.010325	0.012434	0.830404	0.4079
EXP_GR	-0.014622	0.014848	-0.984795	0.3266
FED_ER	-0.158752	0.528738	-0.300247	0.7645
INF	0.990848	0.974915	1.016343	0.3114
NASDAQ_GR	0.001077	0.000476	2.262651	0.0254
RNET_EXP	0.006905	0.012452	0.554550	0.5802
UNEMP	1.203396	0.602669	1.996778	0.0480

R-squared	0.170760	Mean dependent var	2.750030
Adjusted R-squared	0.123948	S.D. dependent var	7.087863
S.E. of regression	6.634074	Akaike info criterion	6.681007
Sum squared resid	5457.357	Schwarz criterion	6.855722
Log likelihood	-432.9465	Hannan-Quinn criter.	6.752003
F-statistic	3.647776	Durbin-Watson stat	2.055708
Prob(F-statistic)	0.001299		

As the P-value is less than alpha (0.05) we can reject the null hypothesis which states that the model has homoskedasticity and conclude that the model has **heteroskedasticity** in the error term in the regression model and the usual testing statistics would not be valid.

2.2

Hence we will re-run the regression model with adjusted standard errors and the p-values which are observed below:-

Dependent Variable: GDP_GROWTH_RATE
Method: Least Squares
Date: 04/10/24 Time: 19:35
Sample: 1990Q1 2022Q4
Included observations: 132
Huber-White-Hinkley (HC1) heteroskedasticity consistent standard errors and covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12.35717	2.541012	4.863092	0.0000
IMP_GR	0.006800	0.004854	1.400888	0.1637
EXP_GR	-0.008950	0.005740	-1.559359	0.1215
FED_ER	-0.480495	0.147595	-3.255487	0.0015
INF	0.626239	0.355978	1.759210	0.0810
NASDAQ_GR	0.000253	0.000189	1.340845	0.1824
RNET_EXP	0.008988	0.004776	1.881752	0.0622
UNEMP	-1.009843	0.253256	-3.987438	0.0001

R-squared	0.354281	Mean dependent var	2.454822
Adjusted R-squared	0.317829	S.D. dependent var	2.071563
S.E. of regression	1.710980	Akaike info criterion	3.970701
Sum squared resid	363.0039	Schwarz criterion	4.145416
Log likelihood	-254.0663	Hannan-Quinn criter.	4.041697
F-statistic	9.719137	Durbin-Watson stat	0.726481
Prob(F-statistic)	0.000000	Wald F-statistic	4.510880
Prob(Wald F-statistic)	0.000165		

Here the standard errors and the p-values are adjusted now and are valid because we have adjusted those values with the Huber – White Hinkley heteroskedasticity test which contains consistent standard errors and covariance.

2.3

We know that serial correlation can cause unreliable hypothesis test results in a regression model, So there’s a need to check if the model has serial correlation or not which is shown below:-

Breusch-Godfrey Serial Correlation LM Test:
Null Hypothesis: No serial correlation at up to 16 lags

F-statistic	6.082488	Prob. F(16,108)	0.0000
Obs*R-squared	62.56685	Prob. Chi-Square(16)	0.0000

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 04/10/24 Time: 19:37
Sample: 1990Q1 2022Q4
Included observations: 132
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.232790	1.591399	-0.146280	0.8840
IMP_GR	-0.001007	0.002682	-0.375546	0.7080
EXP_GR	0.001178	0.003182	0.370157	0.7120
FED_ER	0.029575	0.114057	0.259297	0.7959
INF	-0.228161	0.232692	-0.980527	0.3290
NASDAQ_GR	-2.59E-05	0.000100	-0.258464	0.7965
RNET_EXP	-0.001067	0.002684	-0.397516	0.6918
UNEMP	0.024792	0.139033	0.178320	0.8588
RESID(-1)	0.634602	0.096834	6.553476	0.0000
RESID(-2)	0.036360	0.114850	0.316590	0.7522
RESID(-3)	0.071087	0.114839	0.619016	0.5372
RESID(-4)	-0.282361	0.120780	-2.337817	0.0212
RESID(-5)	0.072702	0.126701	0.573803	0.5673
RESID(-6)	-0.065710	0.144366	-0.455164	0.6499
RESID(-7)	0.279639	0.216422	1.292101	0.1991
RESID(-8)	-0.377933	0.254009	-1.487870	0.1397
RESID(-9)	0.227047	0.249874	0.908642	0.3656
RESID(-10)	-0.309196	0.247009	-1.251757	0.2134
RESID(-11)	0.204370	0.242713	0.842024	0.4016
RESID(-12)	0.005160	0.254497	0.020275	0.9839
RESID(-13)	-0.077320	0.249770	-0.309564	0.7575
RESID(-14)	-0.121491	0.245624	-0.494622	0.6219
RESID(-15)	0.047594	0.237240	0.200617	0.8414
RESID(-16)	-0.035675	0.170253	-0.209543	0.8344

R-squared	0.473991	Mean dependent var	-3.57E-16
Adjusted R-squared	0.361971	S.D. dependent var	1.664639
S.E. of regression	1.329659	Akaike info criterion	3.570688
Sum squared resid	190.9432	Schwarz criterion	4.094833
Log likelihood	-211.6654	Hannan-Quinn criter.	3.783676
F-statistic	4.231296	Durbin-Watson stat	1.985956
Prob(F-statistic)	0.000000		

Here the null hypothesis for testing for serial correlation is that H_0 = No serial correlation; H_1 = There is a serial correlation.

In the output table from the Breusch Godfrey Serial correlation LM test, it is observed that the p-value is less than alpha (p-value<0.05) for **resid(-1) and resid(-4) terms** and we can conclude that there is no serial correlation in this model.

2.4

Now to construct an AR model that helps to adjust a model to account for the dependence of a variable on its own past values we will use a correlogram to check which AR terms could be added to the model.

Date: 04/10/24 Time: 19:51
Sample: 1990Q1 2022Q4
Included observations: 132

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.637	0.637	54.723	0.000
		2	0.392	-0.022	75.622	0.000
		3	0.205	-0.060	81.403	0.000
		4	-0.004	-0.179	81.405	0.000
		5	-0.053	0.064	81.800	0.000
		6	-0.060	0.016	82.298	0.000
		7	-0.075	-0.037	83.093	0.000
		8	-0.107	-0.102	84.712	0.000
		9	-0.120	-0.024	86.769	0.000
		10	-0.156	-0.072	90.313	0.000
		11	-0.145	0.010	93.384	0.000
		12	-0.125	-0.036	95.697	0.000
		13	-0.160	-0.116	99.496	0.000
		14	-0.168	-0.063	103.71	0.000
		15	-0.151	-0.008	107.17	0.000
		16	-0.119	0.010	109.34	0.000

- It is observed that AR(1) is more prominent and AR (4) is more prominent than others we will add these terms to the model.

Therefore with the ARMA Maximum Likelihood (OPG – BHHH) method, this can be done and the output Table is given below:-

As the AR(4) term has a p-value of 0.0507 which is **greater than** 0.05, the term is insignificant and we will keep only the AR(1) in the model.

Dependent Variable: GDP_GROWTH_RATE
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 04/23/24 Time: 12:21
Sample: 1990Q1 2022Q4
Included observations: 132
Convergence achieved after 63 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.779148	3.458510	2.538419	0.0124
IMP_GR	-0.000879	0.004639	-0.189509	0.8500
EXP_GR	0.000880	0.005879	0.149653	0.8813
FED_ER	-0.203629	0.330875	-0.615426	0.5394
INF	0.322365	0.153701	2.097359	0.0380
NASDAQ_GR	-4.93E-05	0.000210	-0.235073	0.8145
RNET_EXP	0.001843	0.004746	0.388271	0.6985
UNEMP	-0.799737	0.150371	-5.318431	0.0000
AR(1)	0.732643	0.057864	12.66140	0.0000
AR(4)	-0.227104	0.115087	-1.973334	0.0507
SIGMASQ	1.470952	0.122470	12.01068	0.0000

R-squared	0.654614	Mean dependent var	2.454822
Adjusted R-squared	0.626070	S.D. dependent var	2.071563
S.E. of regression	1.266758	Akaike info criterion	3.398153
Sum squared resid	194.1656	Schwarz criterion	3.638387
Log likelihood	-213.2781	Hannan-Quinn criter.	3.495773
F-statistic	22.93327	Durbin-Watson stat	2.076475
Prob(F-statistic)	0.000000		

Inverted AR Roots	.72+.40i	.72-.40i	-.35-.46i	-.35+.46i

The addition of the AR(1) and AR(4) terms causes the R-squared value to go up to 65.4% and is a good fit for the model.

3

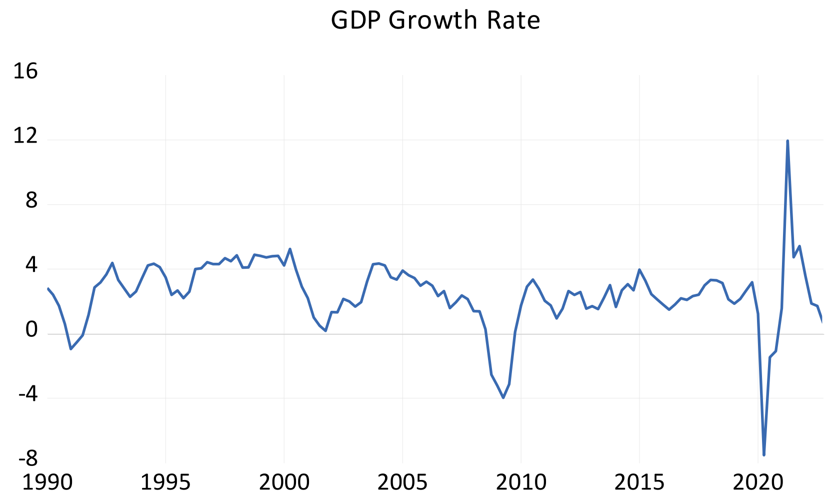
Testing stationary on non-stationary variables:-

Whether the dependent variable is stationary or not is important to know in a model as non – stationary variables can lead to misleading conclusions in the model and improper forecasting. Hence we will observe whether the dependent variable in our model which is the GDP growth rate is stationary or not and how to adjust it.

3.1

A graph of GDP growth is given below –

The GDP growth rate ratio appears to be stationary in this case as it often revolves around the mean after a few years till now.



It is important to find the best suitable ARMA model in an economic model as It makes sure that all the data is effectively captured in the model. Sample Size (1990q1 2020q4)

Tests for autocorrelation also need to be done to find the best fit which can be done with the help of a correlogram which is shown below. Hence from here onwards we will find the best-fitting ARMA model for our analysis.

Date: 04/23/24 Time: 12:22
Sample: 1990Q1 2022Q4
Included observations: 132
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.038	-0.038	0.1997	
		2	0.063	0.061	0.7326	
		3	0.110	0.116	2.4065	0.121
		4	-0.129	-0.127	4.7215	0.094
		5	0.036	0.013	4.9000	0.179
		6	0.055	0.063	5.3175	0.256
		7	0.049	0.080	5.6570	0.341
		8	-0.021	-0.050	5.7191	0.455
		9	-0.034	-0.055	5.8896	0.553
		10	-0.081	-0.081	6.8400	0.554
		11	-0.057	-0.035	7.3133	0.605
		12	-0.064	-0.065	7.9261	0.636
		13	-0.077	-0.079	8.8078	0.640
		14	-0.066	-0.077	9.4570	0.663
		15	-0.082	-0.067	10.462	0.656
		16	-0.029	-0.012	10.592	0.718

*Probabilities may not be valid for this equation specification.

Here we can observe that all of the p-values of the lags are highly insignificant and there is no serial correlation present in the model.

3.2

Re-estimating the model again with AR terms one by one for AR(1) First –

Dependent Variable: GDP_GROWTH_RATE
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 04/23/24 Time: 12:23
Sample: 1990Q1 2022Q4
Included observations: 132
Convergence achieved after 50 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.03308	2.938391	3.754803	0.0003
IMP_GR	0.003138	0.004424	0.709361	0.4795
EXP_GR	-0.004265	0.005361	-0.795580	0.4278
FED_ER	-0.338577	0.307164	-1.102267	0.2725
INF	0.398730	0.210959	1.890089	0.0611
NASDAQ_GR	0.000165	0.000165	0.995670	0.3214
RNET_EXP	0.005851	0.004311	1.357119	0.1772
UNEMP	-0.988983	0.159382	-6.205098	0.0000
AR(1)	0.665201	0.061907	10.74518	0.0000
SIGMASQ	1.587215	0.139448	11.38215	0.0000

R-squared	0.627315	Mean dependent var	2.454822
Adjusted R-squared	0.599822	S.D. dependent var	2.071563
S.E. of regression	1.310464	Akaike info criterion	3.455800
Sum squared resid	209.5124	Schwarz criterion	3.674194
Log likelihood	-218.0828	Hannan-Quinn criter.	3.544545
F-statistic	22.81712	Durbin-Watson stat	1.982718
Prob(F-statistic)	0.000000		

Inverted AR Roots	.67

Here AR(1) is highly significant due to its p -p-value less than 0.05 Schwarz Criterion is 3.67 and the AIC value is 3.45.

Now we will compare the Schwarz criterion and Akaike info criterion for all the estimation output tables to see which will fit the model the best among these.

The Schwarz Criterion and Akaike info criterion are the smallest for the AR(1), and AR(4) model which is 3.39 and 3.63 respectively. The smallest value of these criteria suggests that the AR(1), and AR(4) models would be the best fit among the other ones.

3.3

Now we can check whether the model has serial correlation or not including the AR(1) term,

The output table is given below :

Date: 04/23/24 Time: 12:29
Sample: 1990Q1 2022Q4
Included observations: 132
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.038	-0.038	0.1997	
		2	0.063	0.061	0.7326	
		3	0.110	0.116	2.4065	0.121
		4	-0.129	-0.127	4.7215	0.094
		5	0.036	0.013	4.9000	0.179
		6	0.055	0.063	5.3175	0.256
		7	0.049	0.080	5.6570	0.341
		8	-0.021	-0.050	5.7191	0.455
		9	-0.034	-0.055	5.8896	0.553
		10	-0.081	-0.081	6.8400	0.554
		11	-0.057	-0.035	7.3133	0.605
		12	-0.064	-0.065	7.9261	0.636

*Probabilities may not be valid for this equation specification.

It is observed that all the p-values are highly insignificant (all the p-values are greater than 0.05) up to 1 lag hence the null hypothesis can be rejected and we can conclude that there is no serial correlation in the model.

3.4

Unit root tests help determine whether a time series is stationary or non-stationary. Stationarity is an important assumption in many time series models because it ensures that the statistical properties of the series, such as mean and variance, remain constant over time.

Hence we will now check whether the GDP growth rate contains a unit root or not-

Null Hypothesis: GDP GROWTH_RATE has a unit root
Exogenous: Constant
Lag Length: 4 (Automatic - based on SIC, maxlag=4)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.564433	0.0078
Test critical values:	1% level	-3.482453	
	5% level	-2.884291	
	10% level	-2.578981	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GDP_GROWTH_RATE)
Method: Least Squares
Date: 04/23/24 Time: 12:50
Sample (adjusted): 1991Q2 2022Q4
Included observations: 127 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP_GROWTH_RATE(-1)	-0.302796	0.084949	-3.564433	0.0005
D(GDP_GROWTH_RATE(-1))	0.029334	0.086035	0.340949	0.7337
D(GDP_GROWTH_RATE(-2))	0.163010	0.084894	1.920176	0.0572
D(GDP_GROWTH_RATE(-3))	0.138223	0.085834	1.610359	0.1099
D(GDP_GROWTH_RATE(-4))	-0.391652	0.083734	-4.677336	0.0000
C	0.771923	0.242247	3.186508	0.0018

R-squared	0.373440	Mean dependent var	0.012613
Adjusted R-squared	0.347549	S.D. dependent var	1.694469
S.E. of regression	1.368698	Akaike info criterion	3.511688
Sum squared resid	226.6733	Schwarz criterion	3.646059
Log likelihood	-216.9922	Hannan-Quinn criter.	3.566281
F-statistic	14.42360	Durbin-Watson stat	2.001789
Prob(F-statistic)	0.000000		

Dependent Variable: D(GDP_GROWTH_RATE)
Method: Least Squares
Date: 04/23/24 Time: 12:52
Sample (adjusted): 1991Q2 2022Q4
Included observations: 127 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.189018	0.187381	-6.345446	0.0000
GDP_GROWTH_RATE	0.487077	0.062496	7.793724	0.0000
D(GDP_GROWTH_RATE(-1))	-0.328716	0.067511	-4.869045	0.0000
D(GDP_GROWTH_RATE(-2))	-0.234711	0.070769	-3.316557	0.0012
D(GDP_GROWTH_RATE(-3))	-0.257020	0.070642	-3.638333	0.0004
D(GDP_GROWTH_RATE(-4))	-0.502751	0.063260	-7.947327	0.0000

R-squared	0.539048	Mean dependent var	0.012613
Adjusted R-squared	0.520001	S.D. dependent var	1.694469
S.E. of regression	1.173961	Akaike info criterion	3.204736
Sum squared resid	166.7604	Schwarz criterion	3.339108
Log likelihood	-197.5008	Hannan-Quinn criter.	3.259330
F-statistic	28.30010	Durbin-Watson stat	1.307002
Prob(F-statistic)	0.000000		

Since the p- value is significant, we can conclude that the null hypothesis can be rejected which is that the GDP_growth_rate contains a unit root, and say that it does not have a unit root and go ahead with the forecasting.

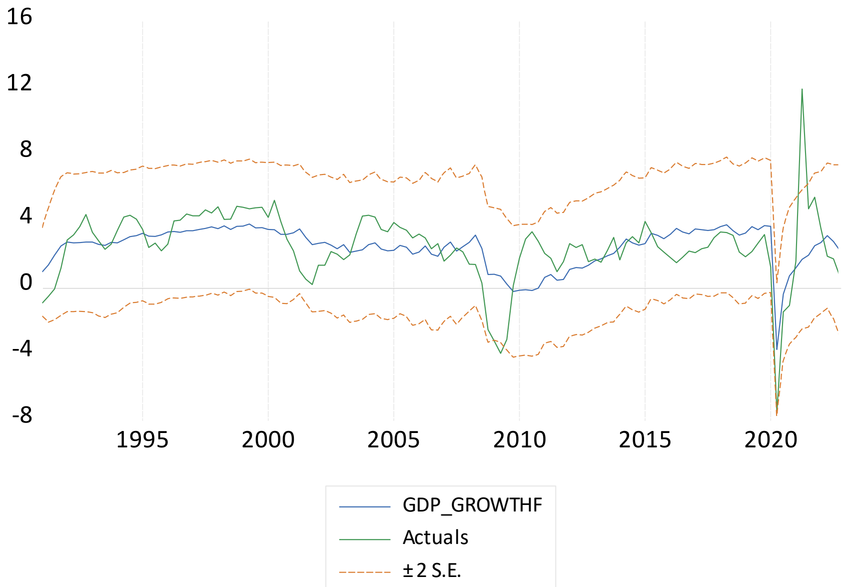
4.0

As the model is appropriate now we can forecast the data first for in-sample data as

Forecasting is typically associated with predicting future values based on historical data.

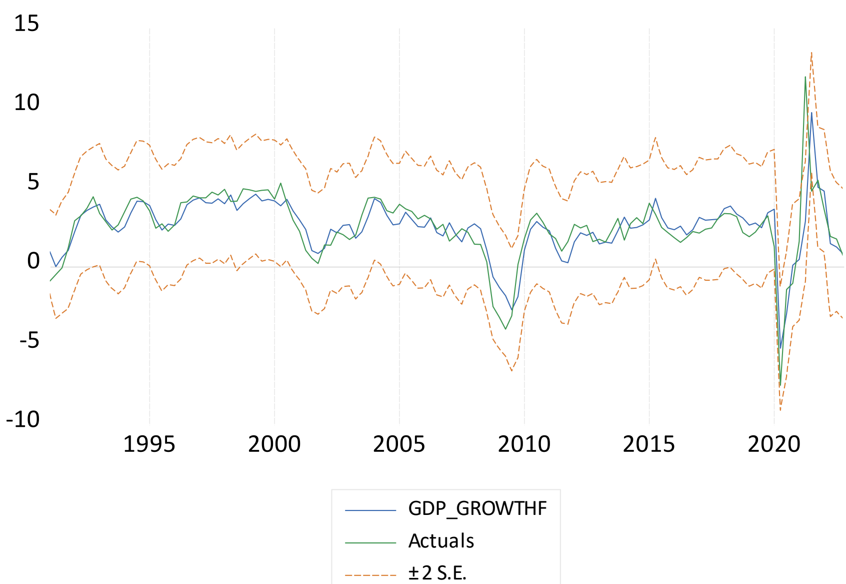
Forecasting on in-sample data can serve as a means to evaluate the performance of the model. By comparing the model's forecasts to the actual observed values within the sample period, we can assess how well the model captures the underlying patterns in the data and compare the dynamic and static forecasts which could be a better one for the model.

Below is the dynamic and static forecast of all of the sample data –



Forecast: GDP_GROWTHF	
Actual: GDP_GROWTH_RATE	
Forecast sample: 1990Q1 2022Q4	
Adjusted sample: 1991Q1 2022Q4	
Included observations: 128	
Root Mean Squared Error	1.727551
Mean Absolute Error	1.258239
Mean Abs. Percent Error	88.53677
Theil Inequality Coef.	0.290922
Bias Proportion	0.000501
Variance Proportion	0.280014
Covariance Proportion	0.719485
Theil U2 Coefficient	1.127797
Symmetric MAPE	57.86853

Dynamic Forecast



Forecast: GDP_GROWTHF	
Actual: GDP_GROWTH_RATE	
Forecast sample: 1990Q1 2022Q4	
Adjusted sample: 1991Q1 2022Q4	
Included observations: 128	
Root Mean Squared Error	1.224737
Mean Absolute Error	0.742669
Mean Abs. Percent Error	54.77958
Theil Inequality Coef.	0.196220
Bias Proportion	0.000160
Variance Proportion	0.084949
Covariance Proportion	0.914892
Theil U2 Coefficient	0.485764
Symmetric MAPE	35.50742

Static Forecast

4.1

Now we'll change the sample size to check if the model can perform out-of-sample forecasting as well. First, we changed the sample size to 1990q1 – 2021q4 for in sample and we selected the AR(1) model by looking at the correlogram and checking the serial correlation as well as the AR(1) model doesn't have a serial correlation in it which can be seen

below.

Dependent Variable: GDP_GROWTH_RATE
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 04/23/24 Time: 13:08
Sample: 1990Q1 2021Q4
Included observations: 128
Convergence achieved after 63 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.611682	3.615690	2.381753	0.0188
IMP_GR	7.57E-05	0.005194	0.014571	0.9884
EXP_GR	-0.000132	0.006537	-0.020183	0.9839
FED_ER	-0.181845	0.352940	-0.515230	0.6074
INF	0.468909	0.214373	2.187351	0.0307
NASDAQ_GR	4.21E-05	0.000204	0.206437	0.8368
RNET_EXP	0.001887	0.005266	0.358272	0.7208
UNEMP	-0.893315	0.184594	-4.839341	0.0000
AR(1)	0.690965	0.065493	10.55024	0.0000
SIGMASQ	1.585451	0.150276	10.55026	0.0000

R-squared	0.635504	Mean dependent var	2.470593
Adjusted R-squared	0.607704	S.D. dependent var	2.093791
S.E. of regression	1.311416	Akaike info criterion	3.460066
Sum squared resid	202.9377	Schwarz criterion	3.682881
Log likelihood	-211.4442	Hannan-Quinn criter.	3.550597
F-statistic	22.85944	Durbin-Watson stat	2.015035
Prob(F-statistic)	0.000000		

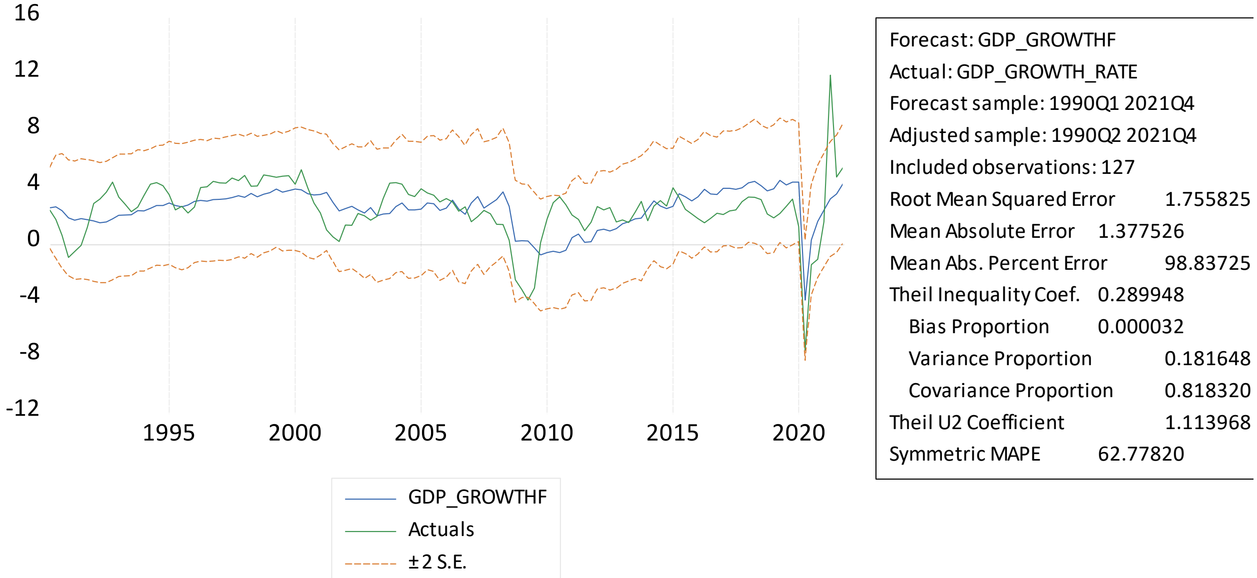
Inverted AR Roots	.69

Date: 04/23/24 Time: 13:12
Sample: 1990Q1 2021Q4
Included observations: 128
Q-statistic probabilities adjusted for 1 ARMA term

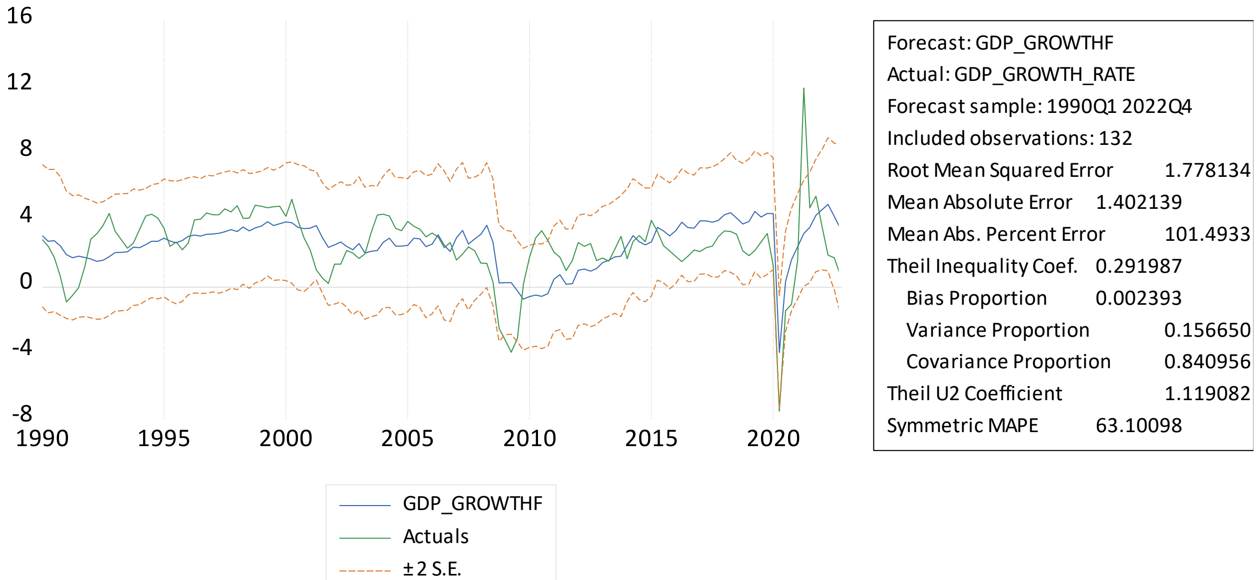
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.008	-0.008	0.0081	
		2	0.013	0.013	0.0308	0.861
		3	0.145	0.146	2.8427	0.241
		4	-0.112	-0.112	4.5265	0.210
		5	-0.093	-0.100	5.6874	0.224
		6	0.018	0.000	5.7315	0.333
		7	0.009	0.048	5.7434	0.453
		8	-0.035	-0.021	5.9173	0.549
		9	0.006	-0.022	5.9216	0.656
		10	-0.047	-0.063	6.2302	0.717
		11	-0.033	-0.017	6.3885	0.782
		12	-0.020	-0.016	6.4459	0.842

*Probabilities may not be valid for this equation specification.

The in-sample forecasting is shown below –



Dynamic Forecast



Static Forecast

When we compare the Root Mean Squared Error, Mean Absolute error, and Mean Abs. Percentage error of both the forecasts which are all measures of forecast accuracy.

The RMSE focuses on large errors, MAE gives equal weight to all errors, and MAPE provides a percentage measure of error relative to the actual values.

We notice that in the Dynamic forecast, all the values of the RMSE, MAE, and MAPE are smaller than those of the Static forecast.

Therefore it can be said that the **Dynamic forecast** is better in this context.

We can notice in the forecasts that the GDP growth rate fell sharp around 2000 – 2002 due to the market crash. The 2000 stock market crash was a direct result of the bursting of the dot-com bubble. NASDAQ subsequently lost nearly 80% and the S&P 500 lost 50% to reach the October 2002 lows.

In 2008 it fell sharp again much more than the last time when the 2008 financial crisis happened.

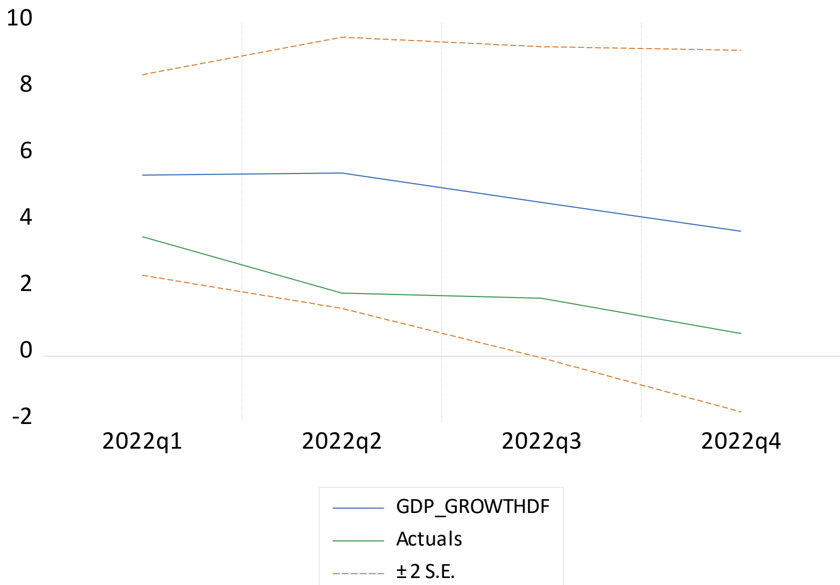
In 2020 COVID – 19 hit the whole world abruptly. It disrupted the supply chain industry all around the whole which made the trade dynamics much worse. A rise in inflation, tax, and unemployment were all the result of the COVID – 19 which caused the GDP growth rate to fall sharp around 2019-2020.

The FED at that time implemented expansionary monetary policies to counter the inflation and implemented interest rate cuts and the disruption in the supply chain industry which slowly recovered around 2021 as the economy recovered.

4.2

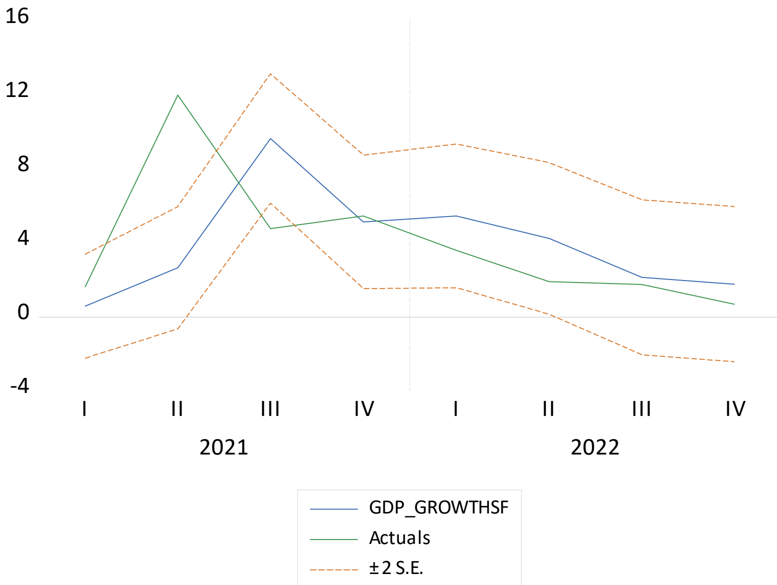
Now we will try to do out-of-sample data forecasting for sample sizes 2022q1 to 2022q4 as out-of-sample forecasting allows for the evaluation of a model's performance in predicting unseen data. By comparing the model's forecasts with the actual observed values in the out-of-sample period, analysts can assess the model's accuracy and reliability.

The forecasts are shown below :



Forecast: GDP_GROWTHDF	
Actual: GDP_GROWTH_RATE	
Forecast sample: 2022Q1 2022Q4	
Included observations: 4	
Root Mean Squared Error	2.927625
Mean Absolute Error	2.857363
Mean Abs. Percent Error	221.4236
Theil Inequality Coef.	0.413968
Bias Proportion	0.952577
Variance Proportion	0.012656
Covariance Proportion	0.034768
Theil U2 Coefficient	3.278443
Symmetric MAPE	92.86205

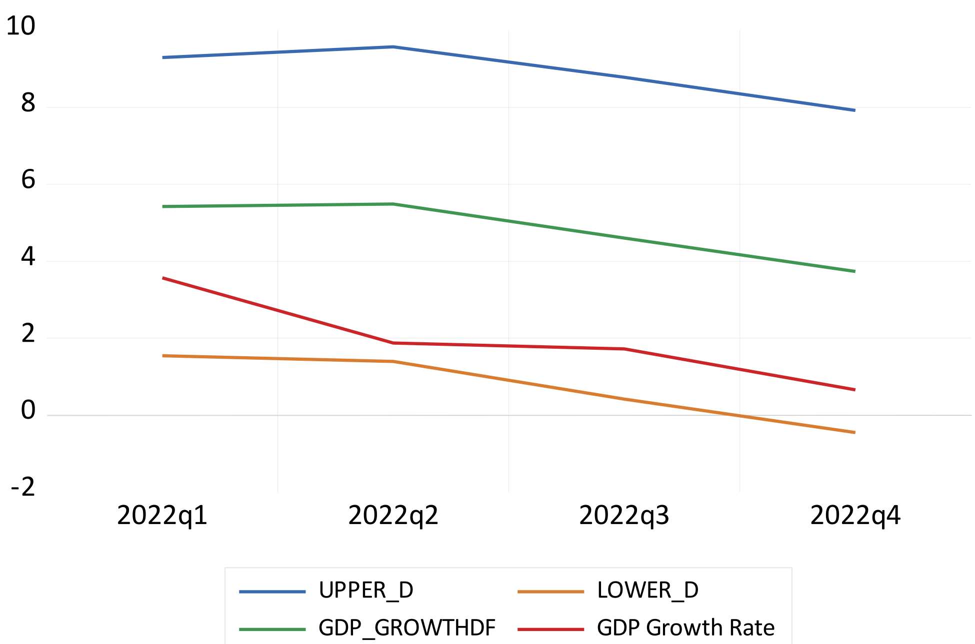
Dynamic Forecast

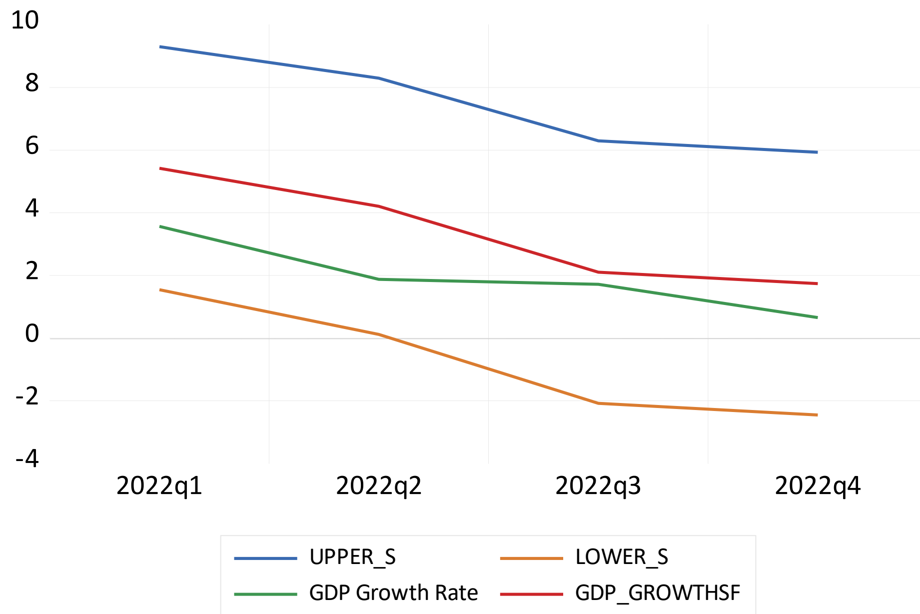


Forecast: GDP_GROWTHSF	
Actual: GDP_GROWTH_RATE	
Forecast sample: 2021Q1 2022Q4	
Included observations: 8	
Root Mean Squared Error	3.907125
Mean Absolute Error	2.651528
Mean Abs. Percent Error	77.25253
Theil Inequality Coef.	0.392908
Bias Proportion	0.000028
Variance Proportion	0.035149
Covariance Proportion	0.964823
Theil U2 Coefficient	0.902364
Symmetric MAPE	66.11389

Static Forecast

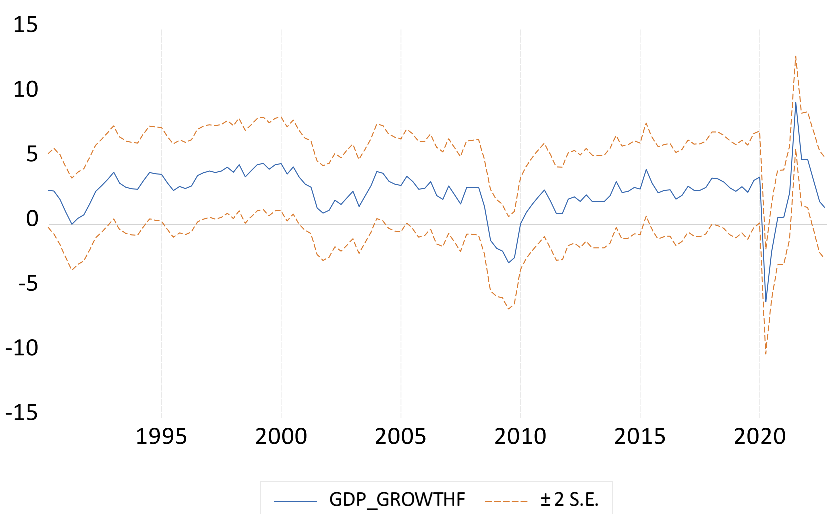
Now for the 95% Confidence Interval,





Hence the forecasted values for both the dynamic and the static forecasts fall inside the 95% confidence intervals.

When we compare the same way we compared in-sample data, In this case also as the values of except RMSE, MAE, and MAPE. The static forecast is better for out-of-sample data too. If we forecast all the values in the model to compare with the out-of-sample forecast of 2021 – 2022, then it would show that the out-of-sample forecast shown



Forecast: GDP_GROWTHF	
Actual: GDP_GROWTH_RATE	
Forecast sample: 1990Q1 2022Q4	
Adjusted sample: 1990Q2 2022Q4	
Included observations: 131	
Root Mean Squared Error	1.264554
Mean Absolute Error	0.779929
Mean Abs. Percent Error	62.31022
Theil Inequality Coef.	0.203025
Bias Proportion	0.000006
Variance Proportion	0.058691
Covariance Proportion	0.941303
Theil U2 Coefficient	0.764122
Symmetric MAPE	37.01820

above is the same as the forecast with all sample ranges till end included in it as seen below:-

Findings and Conclusion :

- In our study the Tax and population variables made the other variables highly insignificant and were removed from our model.
- The model suffered from heteroskedasticity and was adjusted accordingly.
- For our study, the static forecast was suitable for the model for both in-sample and out-of-sample data.
- The forecasts were accurate and worked for both out-of-sample data as well.
- The model explains the historical events accordingly when the GDP growth rate came way down and how the variables like trade and unemployment affected it according to their p-values with these values the variables go down and hence the GDP growth rate goes down.

In all the model has been adjusted and is able to explain the variables included in it with the forecasts being accurate.

References –

- (https://fred.stlouisfed.org/)
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10030258/