Predicting GDP, A Volume 3 Data Project

A Cherpeski, S McIntyre, A Melo, M Stafford

March 2024

Abstract

We use classical decomposition, ARMA models, the Kalman filter, and structural estimation to understand and forecast Gross Domestic Product (GDP) trends in the United States. We use potential explanatory variables such as dependency ratio, unemployment, and measures for education and conclude that they do not substantially improve modeling with basic functional forms.

1 Problem Statement and Motivation

Economic prosperity has long been a topic of interest for researchers; after all, an understanding of how to increase it could lead to an improvement in the lives of billions. One valuable proxy for such prosperity is the Gross Domestic Product (GDP), which is the measure of the final value of all finished goods produced in an economy in a single year. In this project, we observe how these factors apply to real-life empirical results. Thus, we set out to statistically model and predict GDP for the United States growth with investment, dependency ratios, and other factors.

The Solow Growth model serves as the theoretical motivation for our work. Developed by Robert Solow in 1956, this model relates capital to income. We thus focus our predictions using investment (how much money in a given year is spent on machinery, buildings, equipment, etc) and dependency ratios (the number of working-age people divided by the total population) [1]. Our last model is the neoclassical growth model, which is the successor to the Solow Growth model and is more successful at explaining GDP growth, but in its simple form is fairly similar

2 Data

The data from this project came from two sources: The Federal Reserve [2] and an education dataset from Statista [3], which we merged together. Our combined dataset contains quarterly GDP data for the United States, unemployment rates, and percentage of adults who have graduated from high school and college. The data ranged from 1947-2023, although some values were missing. When appropriate, we used linear interpolation to fill in missing quarterly values. There were several instances of missing data prior to 1960 and after 2021, so the models which required them use only data from 1960-2021. Since both of our datasets come from United States government agencies, we consider this data to be trustworthy.

Models were trained from 1947 to 2000 unless stated otherwise. The test data is from 2000 onward.

3 Methods

3.1 Regression Techniques

Figure 7 (Appendix) shows the GDP of the United States every quarter from 1947-2023. GDP tends to grow exponentially, with occasional hiccups due to recessions. Figure 8 in the Appendix shows the results of fitting an exponential curve to this data and the resulting residuals. While the exponential curve fits neatly, it is clear that the resulting residuals are not covariance-stationary. Therefore, a simple regression model does not adequately explain the variation in the data and we will proceed with using other models.

3.2 Kalman Filter

Next, we used the Kalman filter to model GDP, using these equations:

$$x_{t+1} = Fx_t + w_k, w_k \sim N(0, Q)$$
$$z_t = Hx_t + v_k, v_k \sim N(0, R)$$

We used the Kalman filter in two ways: with x_t and z_t being one-dimensional, and with x_t and z_t being three dimensional.

3.2.1 One-Dimensional Kalman Filter

In our one-dimensional model, z_t represents the GDP observations and x_t are the hidden states which can be interpreted as the "true" GDP values. w_k and v_k are noise variables that are normally distributed with mean zero and covariance Q and R respectively. To simplify our calculations, we set $Q = 1$ and $R = 1$. We assumed that $H = 1$, because it it reasonable to assume that the observations of GDP are not consistently biased from the true GDP values in any particular direction. F represents the growth in GDP between each quarter. So a good initial estimate for F would be the average quarterly GDP growth for the United States during our training data years (1960-2000). According to our data, that value is 0.782815% which we use as an initial guess. We then constructed a likelihood function based on our data and used BFGS to maximize this likelihood function with respect to F. We found that the optimal value of F is the same as the initial value: $\tilde{F} = 1.00782815$. Other tests for different initial values of F yielded the same results. Therefore, we proceed with our Kalman Filter using this value for F.

3.2.2 Three-Dimensional Kalman Filter

In our three-dimensional model, z_t represents the GDP, dependency rate, and investments. We experimented having our x_t be just GDP and on a second case, GDP, dependency rate, and investments. We set Q and R to be the identity matrix.

In order to successfully estimate F and H for this Kalman filter, initial values must be given and the model is very sensitive to this choice. We initialized F with $\tilde{F} = 1.00782815 \cdot I$ and H as the identity matrix.

3.3 ARMA Model

Because the ARMA model is an additive model, we work with log values. Although many factors possibly influence GDP, we will focus here on a vector of values that involves the logarithm of GDP, the logarithm of Investment, and the logarithm of the dependency ratio that serves as a proxy for population structure. Our model takes the following form:

$$\Delta Z_t = \mathbf{c} + \left(\sum_{i=1}^{p} \Phi_i \mathbf{y}_{t-i} \right) + \left(\sum_{j=0}^{q} \Theta_j \epsilon_{t-j} \right)$$

Each $y_t = Z_t - Z_{t-1}$, where $Z_{1,t}$ is the log of GDP, $Z_{2,t}$ is the log of total investment, and $Z_{3,t}$ is the log of the current dependency ratio–all at time

t. After performing a grid search we decided upon $p = 6$ and $q = 6$ because it minimized the AIC for the model on the training data. For this model to strongly predict GDP outcomes, we need the change in real log GDP to be covariance stationary. Naturally, this is likely too strong of an assumption to actually hold, but it can inform the analysis we make from the ARMA results.

3.4 Structural Estimation

We also attempted to estimate the relationship using a structural model. In essence, this entails specifying a functional relationship between many different variables and then minimizing a loss function of how well predictions line up with the truth. In this case, we specify a linear relationship between each variable, including a constant, and use a quadratic loss function

The benefit of this model is that it takes into account the fact that in economic time series, every variable is plausibly related to every other variable. For instance, the unemployment rate may influence investment which could influence consumption and so on. It also gives actual coefficients on the linear model, much like OLS, which allows us to do some interpretation of the results and ensure that they look reasonable.

The process is as follows: on the training data, estimate a linear relationship between all parameters. This allows for every variable to be both its own independent variable, and a dependent variable in every other relationship. The exception is GDP, which is only ever a dependent variable. Then we generate predicted values using those relationships, and then make predictions again using exclusively the predicted values. The loss function is the sum of squares between predicted values and true values for every time period in the training data, for every variable. We also include an L_2 regularization term.

This forces everything in the objective function to be in terms of predicted values which adhere to a certain functional form. Once we have those coefficients, we can follow the same process on the testing set. It is important to recognize that the data for GDP on the testing set is never used, and the inputs to predicting it are predicted values, not the true data.

Finally, we also consider moving on from a linear model to a more general form called the Neoclassical Growth Model. Without going into every detail of the model, it boils down to an optimization problem over investment. Essentially, the economy spends all of its output on consumption or invests it for the future, which leads to increased productive capacity in the next time period. We specify standard functional forms for all variables and use

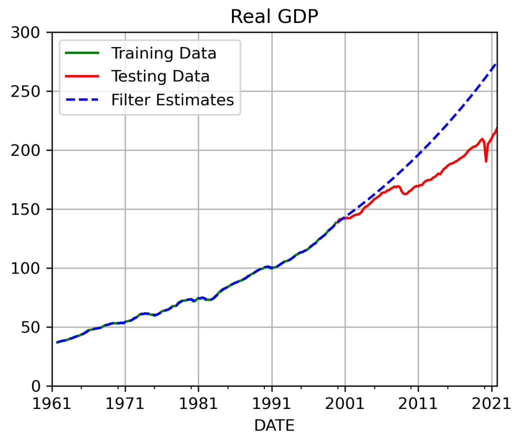

Figure 1: One-dimensional Kalman Filter on GDP

as inputs real GDP, Consumption, and Investment numbers from the United States.

We try to estimate several parameters of the model, which are savings rate s, the fraction of production put towards investment, capital depreciation rate δ, capital productivity share α, and a rate of technological growth A. For a more detailed description of the equations involved, see Acemoglu [4].

4 Results

4.1 Kalman Filter

The results of our Kalman filter are as follows:

4.1.1 One-Dimensional Kalman Filter

We trained the one-dimensional Kalman filter on the training data (years 1960-2000) and used it to predict GDP from 2001-2021. The results, compared with the actual data, is shown in Figure 2 .

This model fits the training data well, but overestimates the test data. The predictions are fairly close to the test data initially, but recessions such as the dot-com burst (in 2000-2002), the 2008 recession, and the 2020 COVID-19 pandemic were obviously not captured by the model.

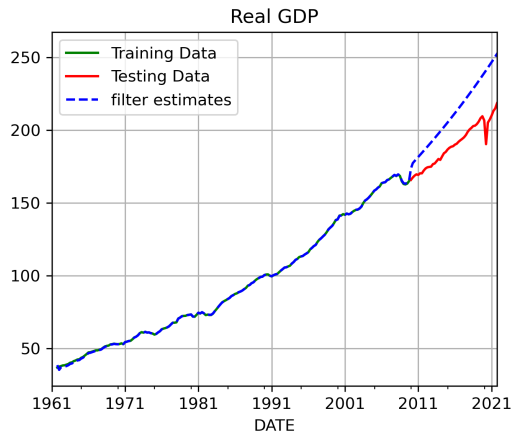

Figure 2: Three-dimensional Kalman Filter on GDP

4.1.2 Three-Dimensional Kalman Filter

For our three-dimensional Kalman Filter model we performed two experiments. Figure 2 shows the Kalman filter with only GDP as the observation variable with three hidden states. Figure 3 shows the Kalman filter for three dimensions of observation: GDP, dependency rate, and investments with three corresponding states, representing the "true" states of each variable.

In Figure 3, we can see similar results to the one-dimensional model. An interesting thing to note is how this model dealt with the first recession. It captured the recession, but then it assumed the prediction would go back to normal right after, following the same trend as before. When attempting to predict more than just GDP, the model performed fairly well. Again the GDP results were very similar to the one-dimensional case. One thing to note about the three dimensional model is the time to train the model. On average, the three-dimensional took 4 to 5 times as long to train as the one-dimensional.

4.2 ARMA

The results for the ARMA model are shown in Figure 5. The ARMA model proved fairly effective at predicting GDP performance in the short term. The main errors in our prediction came from the Great Recession and the Covid Recession. Recessions are notoriously difficult to predict, so it comes as no surprise that they could mess up our prediction's validity. We also

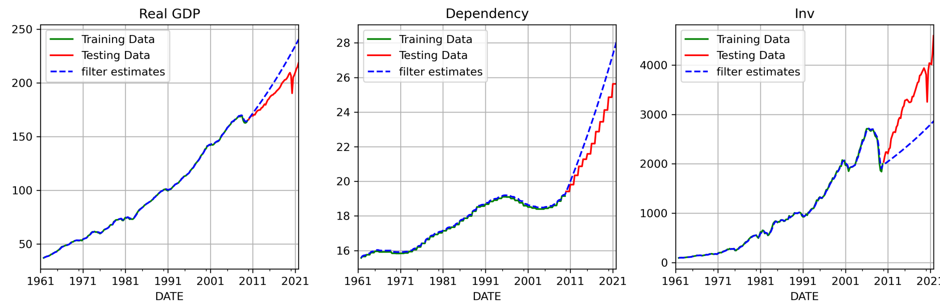

Figure 3: Three-dimensional Kalman Filter on GDP, Dependency, and Investments

encountered some issues with choosing the endpoint for the training set for the data. Because ARMA is primarily based on recent values, predicting from the end of a period of expansion or the beginning of a recession proved to make predictions fail almost immediately. Thus, these results are encouraging but limited.

4.3 Structural Estimation

Our structural estimation starts with a baseline of a simple linear regression of GDP on a constant and 12 other economic indicators. In order for a linear form to make sense, we first take the log of all monetary variables. As expected, simple linear regression is fantastic, 9 (Appendix) shows the fit in this case. Note that this is not true forecasting, since we incorporate data from the years 2000 - 2024, it is only the regression coefficients that are trained on the earlier period and then mapped to the second period. This indicates that, in a linear model, the coefficients seem to be quite stable over time.

To form a more realistic model, we detrend each variable individually, since the trend itself is capable of explaining a good deal of the variance (as in linear regression), even if the independent variable is unrelated to the dependent variable. Our linear structural model performs quite well when only using 6 out of 13 parameters if given a moderately sized regularization constant, somewhere around 1 to 10. If the constant is too small the model clearly overfits and gives nonsensical values. A constant in the realm of 50 or higher is too restrictive, and the model is not flexible enough to explain all the variation. Figure 5 shows the fit with lambda = 10 and 6 variables.

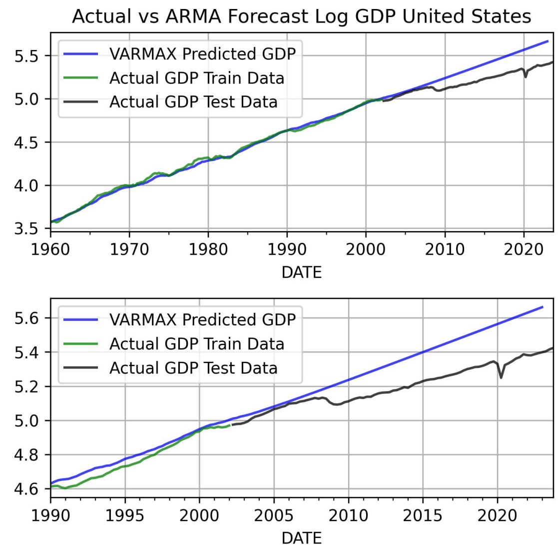

Figure 4: Forecasting GDP with ARMA

If we use all 13 variables the relationship between some of them is too weak to make good predictions, which actually leads to a worse fit because of the model's need to use only predicted data. As such, the relationships are much more noisy, so restricting analysis to only the most relevant variables is helpful. In addition, training a model with 13 variables involves estimating 169 interrelated parameters, which is very computationally expensive.

Our neoclassical growth structural model predicts $\alpha = 0.99$, $\delta = 0.17$, $s = 0.37$, and $A = 0.0087$. Predicted values by more complicated models range considerably, but these values appear to be in the correct arena with the exception of α, which is typically around 0.5. When used to predict future values of investment and consumption they actually perform quite well, as seen in figure 6. Although they are smoother than the true values of consumption and investment, they follow the correct trend and seem to capture the important details. We do not plot GDP because this model has no scope for government spending, imports, or exports, so our estimates would be systematically biased downwards.

The neoclassical growth model is not the industry standard today, that would be Dynamic Stochastic General Equilibrium (DSGM) model. The DSGM carries the same functional form as the neoclassical growth model but incorporates random shocks as well as more room for inflation, unemployment, interest rates, heterogeneity, government expenditure, etc. To create that model is beyond the scope of the project and this class, but the linear and neoclassical growth models give some of the flavor of the full structural model that is used in modern macroeconomics.

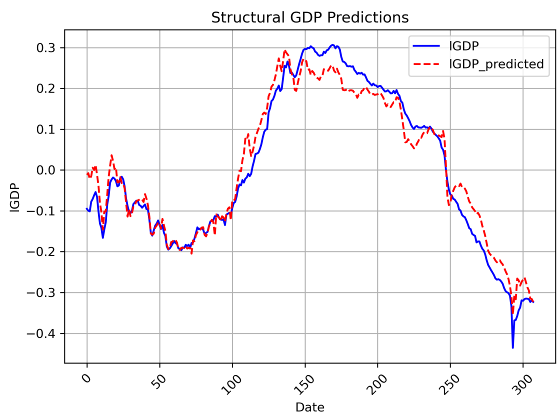

Figure 5: Linear Structural Estimation Predicting GDP 2000-2024, lambda=10

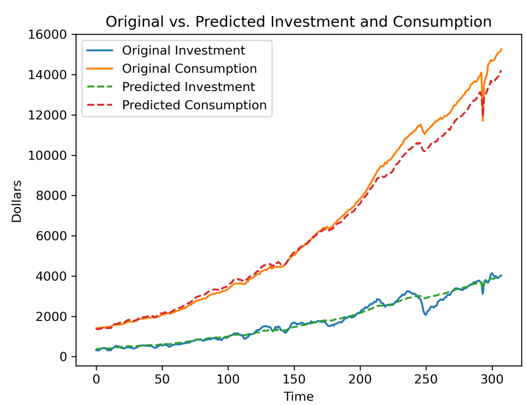

Figure 6: NGM Structural Estimation Used to Predict GDP 2000 - 2024

5 Ethical Considerations

Since our data primarily came from public sources such as the world bank and the Federal Reserve, privacy is not an ethical challenge for our research. However, much more could be said about the potential for miscommunication that comes when discussing complicated economic issues. For example, GDP should never be confused with general societal well-being. While GDP may be a valuable proxy, it remains distinct from well-being itself. Therefore, attempts to optimize GDP growth would not necessarily optimize the well-being of society. Moreover, a simple look at GDP fails to acknowledge the challenges that can result from income inequality and other cultural challenges that exist on the boundary of society and the economy.

Moreover, these models by necessity represent a vastly simplified picture of the world. In many cases the relationships expressed by these models do not fit neatly into a causal framework where one causes another. Indeed, economic data often features reverse causality and hidden correlations that make interpreting the results notoriously difficult. Because of that, readers should be careful to recognize that any estimated parameters or relationships only represents best attempts at fitting and forecasting, and do not necessarily imply a causal relationship of any particular sort between the data in question.

6 Conclusion

GDP is fundamentally difficult to predict with standard time-series tools and none of these models can adequately predict recessions or other changes in productivity which cause GDP to deviate from an exponential growth path. While a reasonable approximation may be made with either the Kalman Filter or ARMA models, neither one performs much better than guessing a growth rate and following an exponential curve. Because the growth rate over the last several decades has been lower than the preceding ones, these models tend to overfit the data. Structural models can explain what would lead to such a change, but they cannot truly predict when those changes might happen.

Our models also show that including many explanatory variables in these simple models is not particularly helpful unless the relationships between these variables and GDP are carefully established, as is done in industry. Most of our predictions are not substantially improved beyond using a handful of the most necessary variables.

References

[1] Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics*, 70(1), 65-94.

[2] Federal Reserve Bank of St. Louis.(Federal Reserve Economic Data) website: fred.stlouisfed.org/

[3] Statista, Educational attainment distribution in the United States from 1960 to 2022 www.statista.com

[4] Acemoglu, Daron. Economic Growth: Lectures 5 and 6, Neoclassical Growth, economics.mit.edu , Accessed April 2024

7 Appendix: Additional Figures

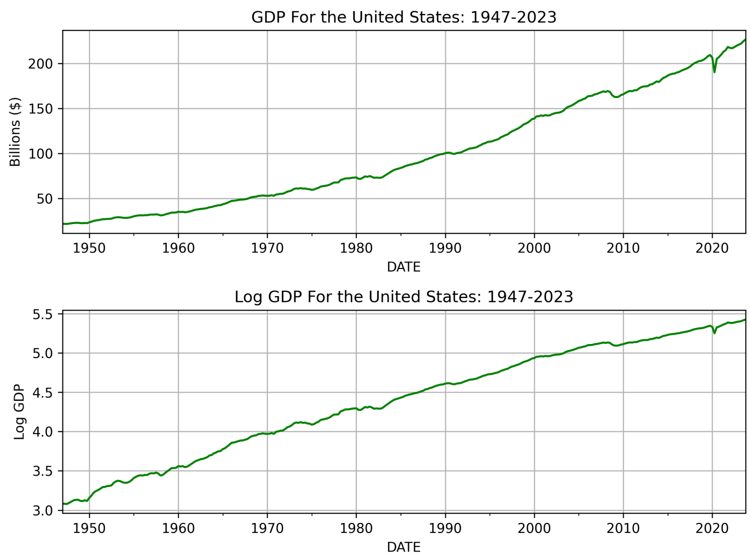

Figure 7: Decomposition of GDP Time Series

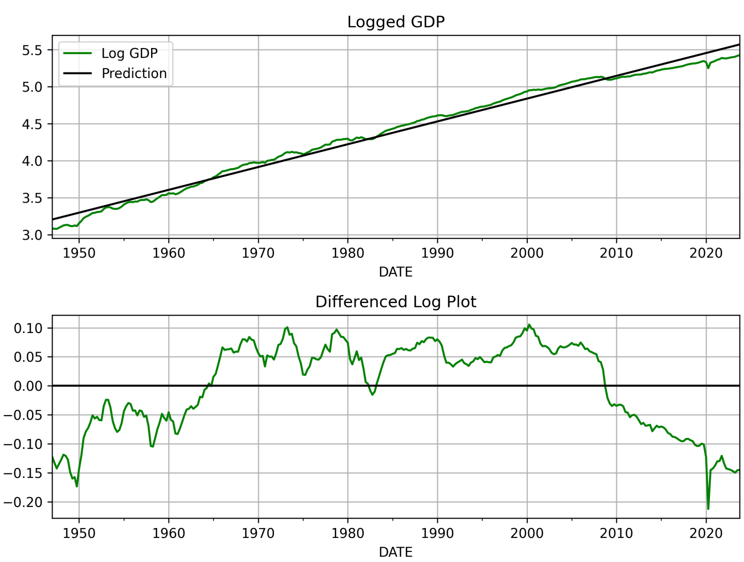

Figure 8: Log GDP with OLS and Differenced Plot, $\hat{\beta}_0 = 3.2056, \hat{\beta}_1 = 0.0077$

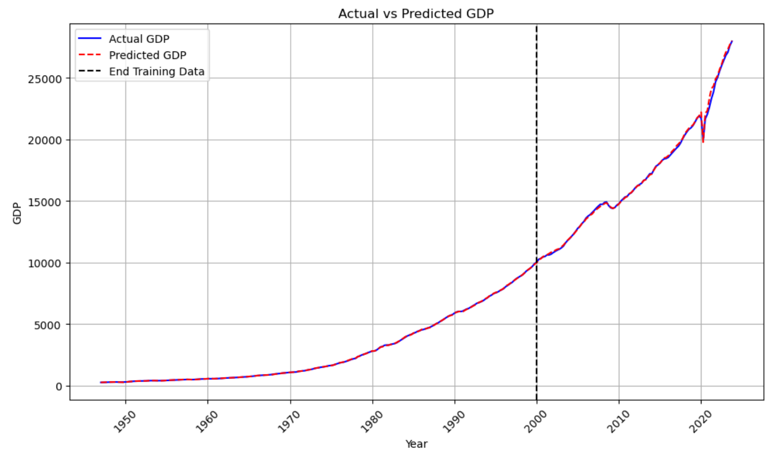

Figure 9: Linear Regression Used to Predict GDP 2000 - 2024

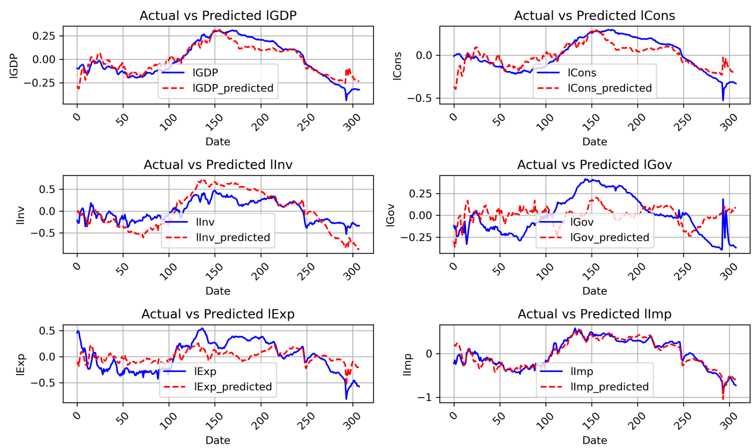

Figure 10: Structural Estimation, lambda = .1, 6 Variables

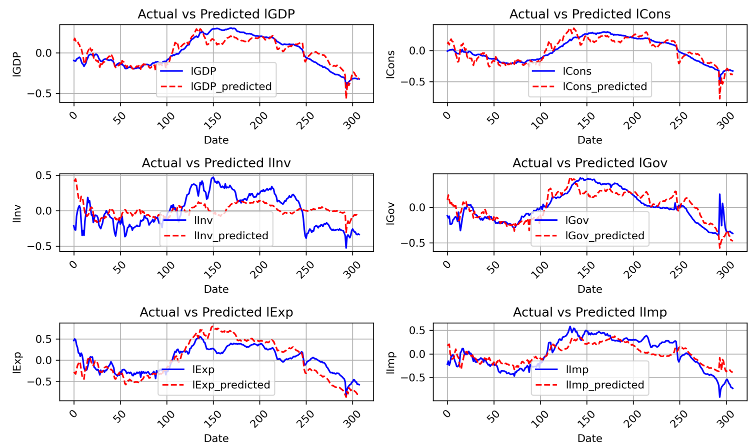

Figure 11: Structural Estimation, lambda = 1, 6 Variables

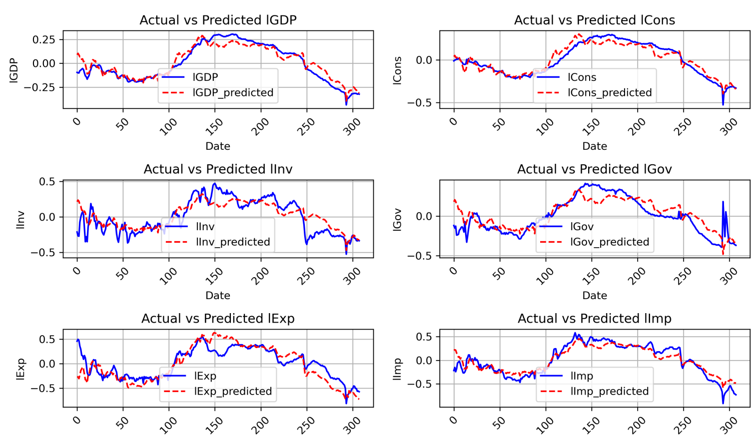

Figure 12: Structural Estimation, lambda = 10, 6 Variables

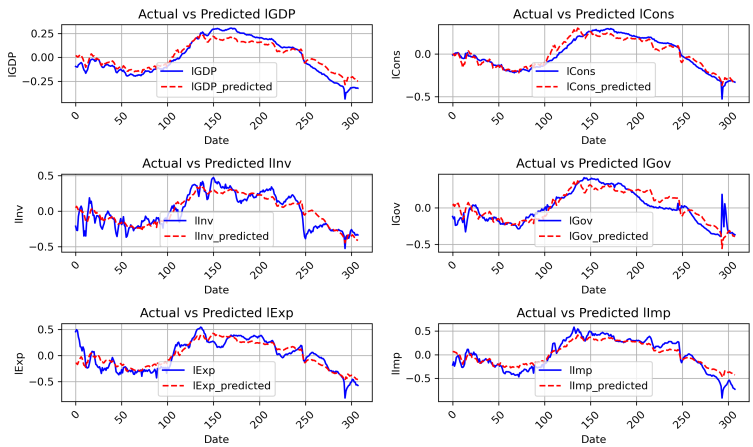

Figure 13: Structural Estimation, lambda = 50, 6 Variables

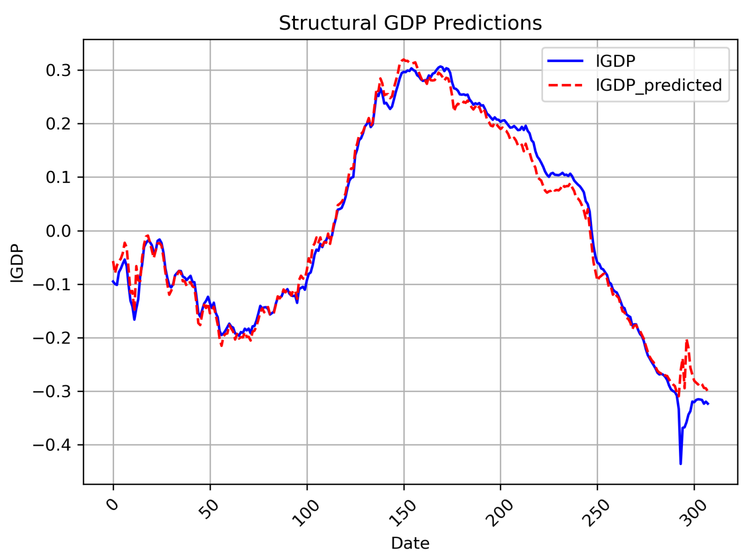

Figure 14: Structural Estimation, lambda = .1, 6 Variables

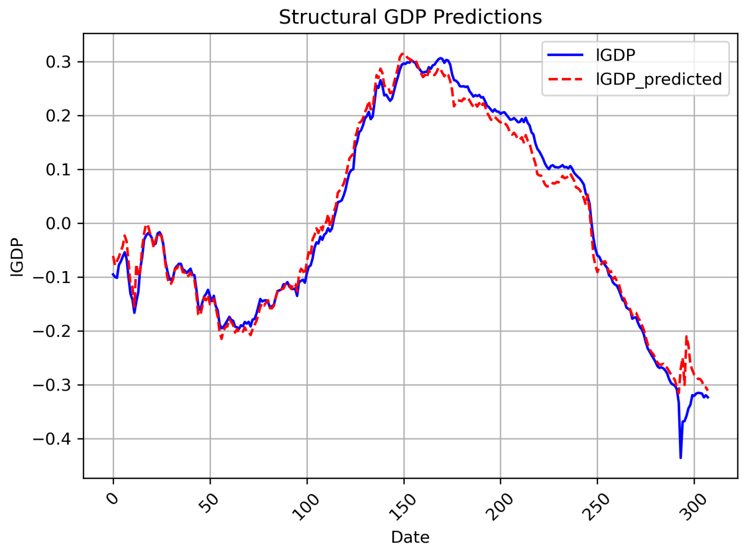

Figure 15: Structural Estimation, lambda = 1, 6 Variables

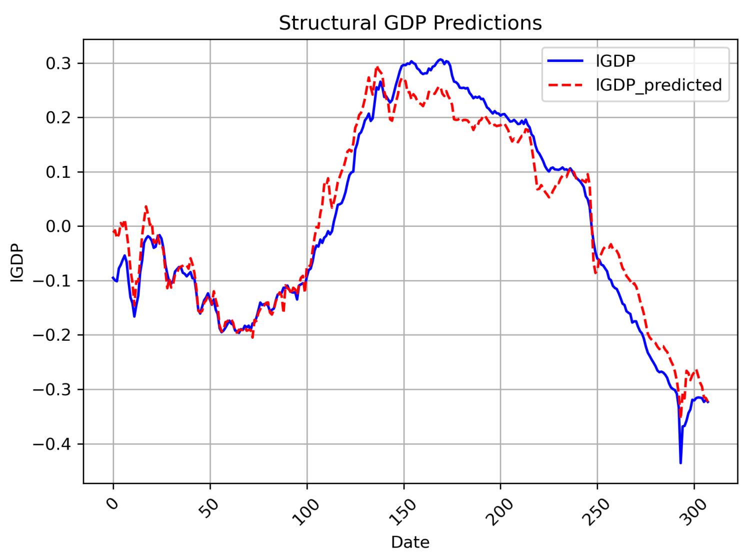

Figure 16: Structural Estimation, lambda = 10, 6 Variables

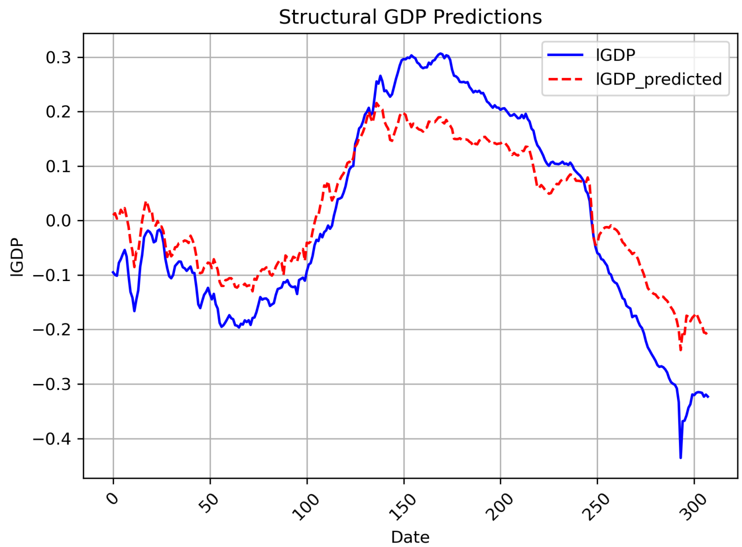

Figure 17: Structural Estimation, lambda = 50, 6 Variables

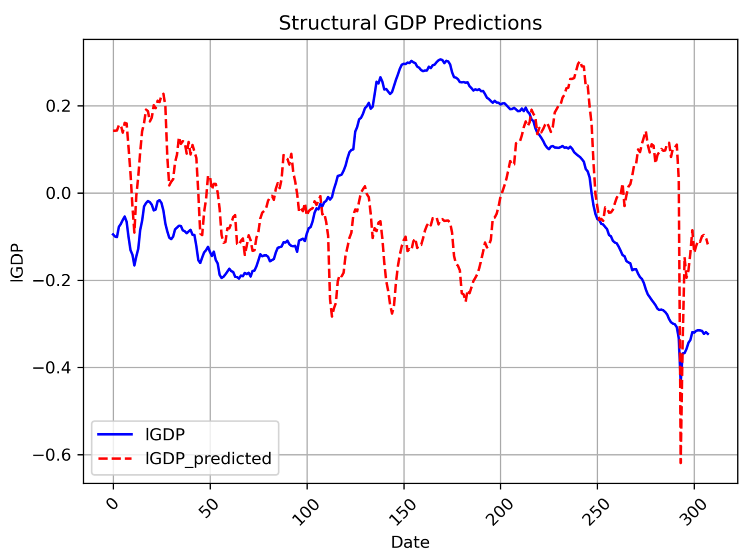

Figure 18: Structural Estimation, lambda = 10, 13 Variables

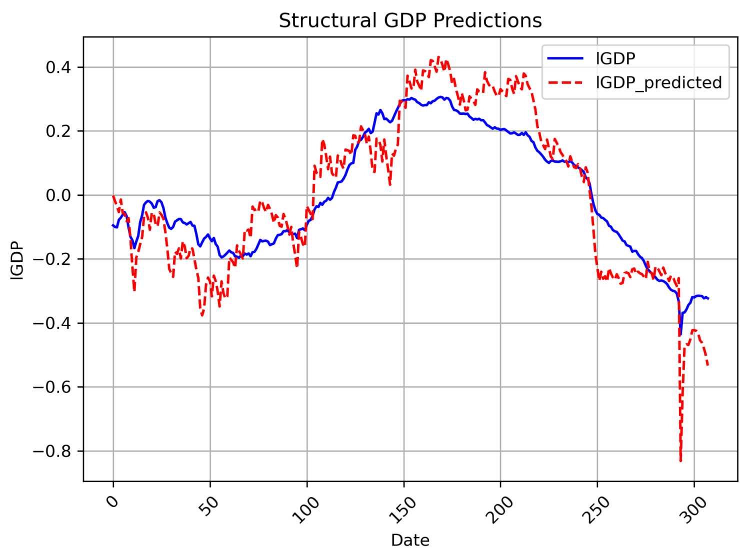

Figure 19: Structural Estimation, lambda = 50, 13 Variables