Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/49/20230534

Analysis and Prediction of Poverty Rate along with Other Economic Factors in the United States

Xiaodong Zhao[1,a,*], Vivian Tao[1,b]

[1]*Canadian International School of Beijing, Beijing, China*
a. edisonzxd@proton.me, b. taovivian1204@163.com
**corresponding author*

Abstract: Poverty is an everlasting issue that affects everyone on the globe, and it is of crucial importance to lift people out of poverty. It is part of the UN SDGs and a macroeconomic goal of most countries. Thus, it is important to understand patterns and make predictions of poverty rate. In this paper, data from the US Census Bureau and The Balance are used to form timeseries analysis. In addition to poverty rate alone, unemployment rate and inflation rate are used to generate dynamic regression model for more complete and accurate predictions. The result of modeling and prediction shows that poverty rate in the United States will most likely increase in the following years to around 15 to 16 precent and then gradually decrease. It is also likely that it will fluctuate around 12%. Overall, the result is a sign that potential measures should be taken to mitigate the possible effects of a predicted increasing poverty rate.

Keywords: prediction, poverty rate, timeseries, economic indicators

1. Introduction

Using statistical measures to find patterns and make predictions for economic indicators are of crucial importance in understanding and addressing many of today's persistent challenges. One such challenge is poverty. Poverty is a prevalent issue that affects every nation on the globe and is causing detrimental impact to citizens. Poverty limits access to necessities, such as food, water, education, and so on. It also causes many other problems like housing instability, psychological issues, and crime.

In 2021, according to the US Census Bureau, the official poverty rate of the United States was 11.6 percent, equivalent of 37.9 million citizens in poverty [1]. Speaking globally, using poverty measurements of lower-middle-income countries, the global estimated poverty rate concluded by the World Bank is around 23.6 percent, representing around 28 million people. While with measurements for upper-middle-income countries, the global estimated poverty rate is around 46.9 percent, representing around 44 million people [2].

Due to the global impact and detrimental effects of poverty on people's well-being, the United Nation set lowering poverty rate as one of the 2030 Sustainable Development Goals, "No Poverty", aiming to reduce the number of people living in poverty by at least half [3]. Thus, evaluating and predicting poverty rate is crucial for effective policy interventions, ensuring efficient resources allocation, socioeconomic stability, and general happiness levels.

In this paper, the poverty status by percentage of the entire population is taken from the United States Census Bureau, a governmental data base that collects economic and demographic data in US

DOI: 10.54254/2754-1169/49/20230534

[1]. The data of unemployment rate and inflation rate are taken from the finance company "The Balance", a credible source considering its 20 years' worth of operation in the financial field, offering a complete summary of economic indicators in the past 60 years or so [4,5]. After data collection, timeseries analysis will be performed upon the poverty rate of United States, while including unemployment rate and inflation rate for dynamic regression models.

2. Method

2.1. Data collection

According to the US Census Bureau, the poverty rate is the ratio of the number of people below the poverty line and the entire population. Usually, the poverty line is taken as half of the median household income [6].

The unemployment rate refers to the number of unemployed citizens who are currently looking for a job divided by the entire labor force, which consists of people within the working age of around 15 to 64 [7].

Inflation rate is calculated using the ratio of the Consumer Price Index (CPI) for two consecutive years. This is a measure of change in value for the currency of nations, in this case, the US dollar [8].

In this paper, the US poverty rate (PovRate), unemployment rate (UnemRate), and inflation rate (InflRate) from 1959 to 2021 are examined and used for modeling and forecasting, mainly to predict the poverty rate of US and compare the result of different dynamic regression models.

2.2. Data pre-processing

Data before the time of 1959 was removed from the data set because the effect of the Great Depression and the Second World War would interfere with fitting a model with greater usability. As shown in table 1, the Box-Pierce test of the poverty rate data set shows a strong auto-correlation within the data set, thus, taking lag for the data set may be used later during the modeling process. Also, Augmented Dickey-Fuller Test shows that the data set is potentially stationarity, which should also be considered when fitting the model.

Table 1: Data pre-processing and examination.

Box-Pierce test		Ljung-Box test		ADF test	
X^2	48.119	Q*	121.04	Dickey-Fuller	-4.1813
df	1	df	10	Lag order	3
p-value	4.011 e-12	p-value	< 2.2 e-16	p-value	0.01

2.3. Modeling

2.3.1. Direct ARIMA modeling

The ARIMA (1, 1, 0) model is used to fit the poverty rate data directly. By considering autocorrelation of the dataset, a lag equals to 1 is used to produce the following result shown in figure 1 and table 2.

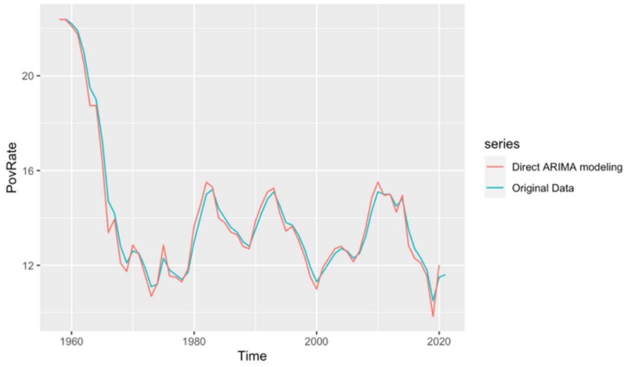

Figure 1: Fitting poverty rate with ARIMA (1, 1, 0) model.

Table 2: Ljung-Box Test of the ARIMA (1, 1, 0) model.

Q*	0.091327
df	0
p-value	< 2.2 e-16

As shown in table 2, using the auto.arima function in R, the ARIMA model (1, 1, 0) is determined to be the best model fitting the poverty rate data, with a p-value much smaller than 0.05, meaning the result is quite significant.

2.3.2. Dynamic regression models with different regressors

In order to better fit the data of poverty rate, the data of unemployment rate, annual GDP, and rate of inflation is used to generate dynamic regression models.

Firstly, the data of the three other economic indicators are used to fit ARIMA models, which will then be used as regressors for the dynamic regression model. For unemployment rate, an ARIMA (1, 0, 0) model, as shown in figure 2, is determined to be the best model and is thus used to fit the original data. The p-value from the Ljung-Box test, as shown in table 3, is significantly smaller than 0.05, which means that the model fits well.

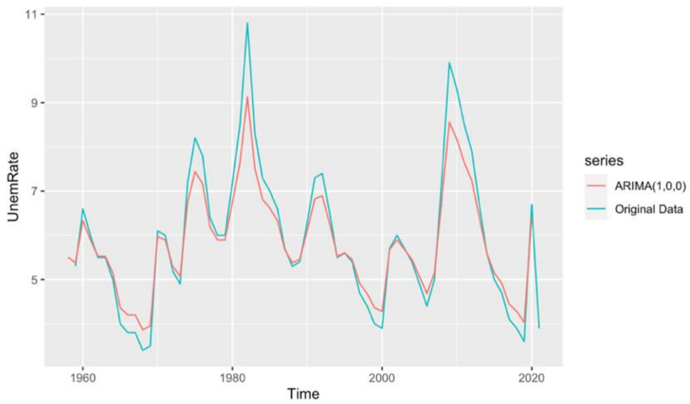

Figure 2: Fitting unemployment rate with ARIMA (1, 1, 0) model.

Table 3: Ljung-Box Test of the ARIMA (1, 1, 0) model.

Q*	1.1055
df	0
p-value	< 2.2 e-16

Similarly, ARIMA (0, 1, 1) model is used to fit the inflation rate data, as shown in figure 3, which also yields a significant result, with the p-value smaller than 0.05, shown in table 4. However, although the p-value indicates that the degree of fitting is quite high for the ARIMA model, the model did not capture the final sharp increase in inflation rate in 2021. This will probably affect the model prediction later.

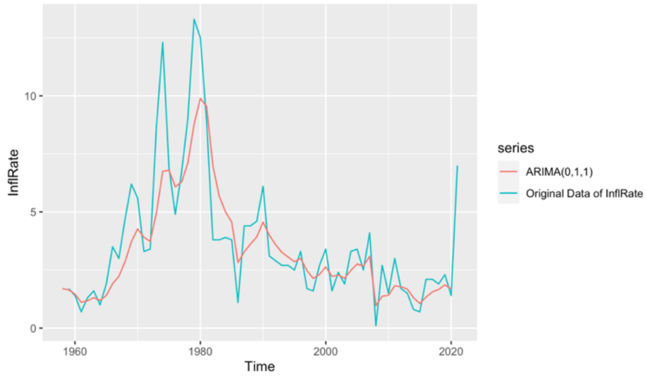

Figure 3: Fitting inflation rate with ARIMA (0, 1, 1).

Table 4: Ljung-Box Test of the ARIMA (0, 1, 1) model.

Q*	0.045143
df	0
p-value	< 2.2 e-16

Then, the ARIMA models fitting unemployment rate and inflation rate is used as the regressor to create dynamic regression models, prediction the poverty rate of the United States, as shown in figure 4 and figure 5.

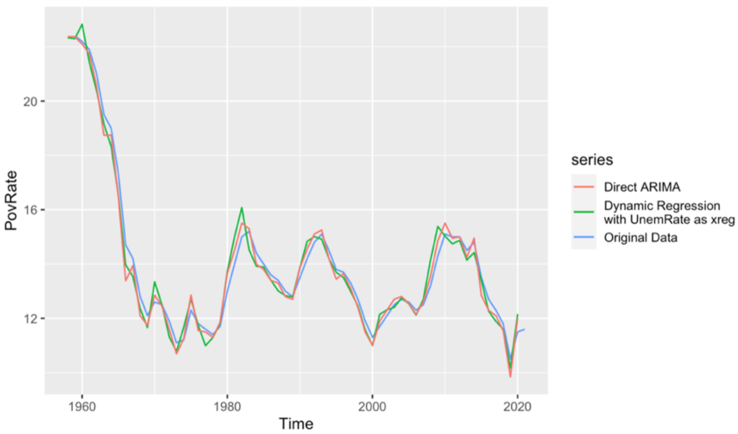

Figure 4: Dynamic regression model of poverty rate with unemployment rate as regressor.

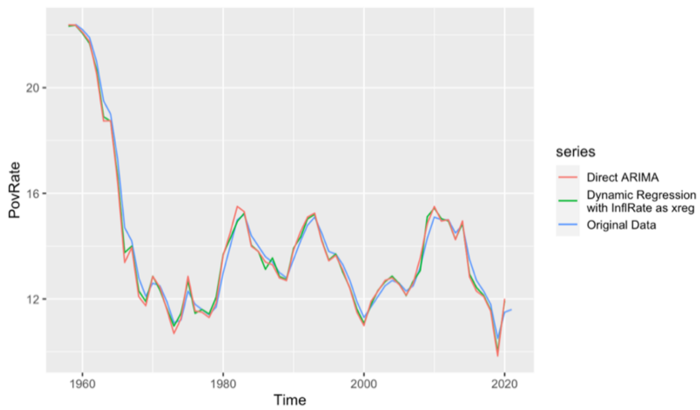

Figure 5: Dynamic regression model of poverty rate with inflation rate as regressor.

Table 5: Ljung-Box Test of the two dynamic regression models for poverty rate.

Unemployment rate		Inflation rate	
Q*	121.99	Q*	122.39
df	10	df	10
p-value	< 2.2 e-16	p-value	< 2.2 e-16

According to table 5, it can be seen that both models make pretty good fittings, for their p-value is below the significant level of 0.05.

So far, there are three different kinds of modeling for the US poverty rate, ARIMA (1, 1, 0), dynamic regression with unemployment rate, and dynamic regression with inflation rate. The three models are expected to yield different results in forecasting.

3. Result

3.1. Result of ARIMA (1, 1, 0) model

The ARIMA (1, 1, 0) model produces a constant point forecast of 12.0, as shown in figure 6. Without any additional changes, the ARIMA model works similarly to a naïve method or taking the mean of

DOI: 10.54254/2754-1169/49/20230534

the dataset; similar to analyzing for stationarity. The upper and lower 95% range expands in a rather consistent rate, reaching 18.3 and 5.7 in 2030 respectively.

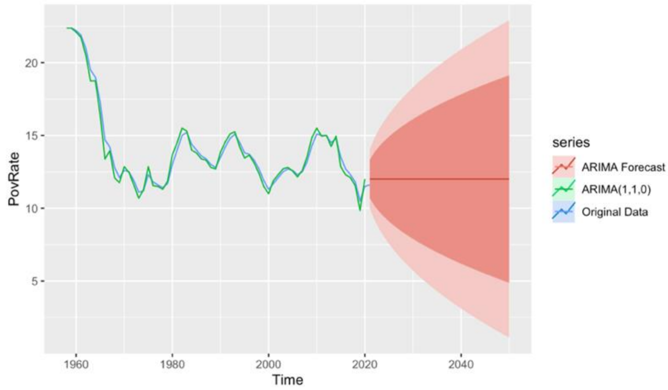

Figure 6: Forecast of ARIMA (1, 1, 0).

3.2. Result of Dynamic Regression using Unemployment Rate as regressor

The dynamic regression model using unemployment rate generates a forecast with more changes, as shown in figure 7. The forecasting poverty rate first increases to a maximum value in the year 2038, with a point forecast value of 15.9, an upper 95% boundary of 22.6 and a lower 95% boundary of 11.2. Then it comes into a local minimum in the year 2062, with a point forecast value of 12.6, an upper 95% boundary of 19.2 and a lower 95% boundary of 8.3. Finally, it increases again to a local maximum in the year 2072, with a point forecast value of 15.3, an upper 95% boundary of 23.4 and a lower 95% boundary of 10.0. The values afterwards are no longer within the accuracy of the model and is not valuable for predictions.

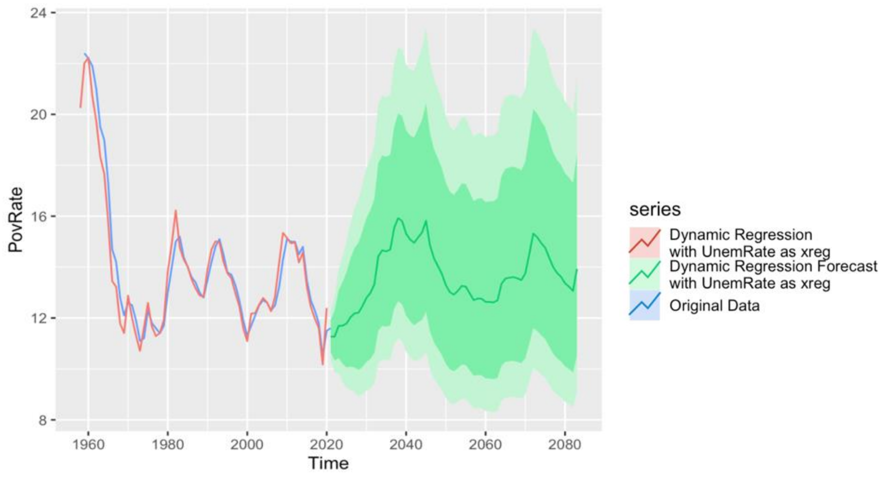

Figure 7: Forecast of Dynamic Regression with Unemployment rate as regressor.

3.3. Result of Dynamic Regression using Inflation Rate as Regressor

The dynamic regression model using inflation rate generates a forecast that is simpler, as shown in figure 8. The forecasting poverty rate increases as well, until a maximum point in the year 2037, with a point forecasting value of 15.2, an upper 95% boundary of 22.3 and a lower 95% boundary of 10.4. Then, it decreases to a local minimum in the year 2049, with a point forecasting value of 13.2, an

DOI: 10.54254/2754-1169/49/20230534

upper 95% boundary of 20.0 and a lower 95% boundary of 8.8. Finally, the forecasting value slowly decreases to around 13.3 in the next 50 years.

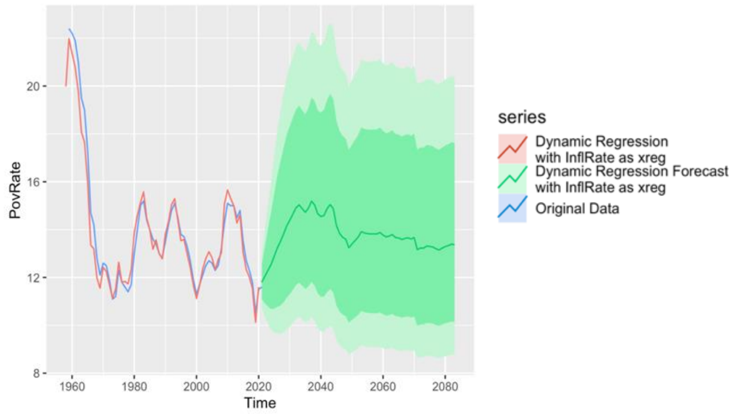

Figure 8: Forecast of Dynamic Regression with Inflation Rate as regressor.

4. Discussion

4.1. Model-to-model comparison

The fitting and modeling of the poverty rate data for the three methods are practically the same, as shown by the comparison in figure 9. It is hard to distinguish any differences from the data or the graph. For example, as shown above, the four lines, original data and the three models, are almost indistinguishable. This might suggest a possibility for over fitting or is a sign that all three methods are consistent with each other, therefore generating a similar result in modeling.

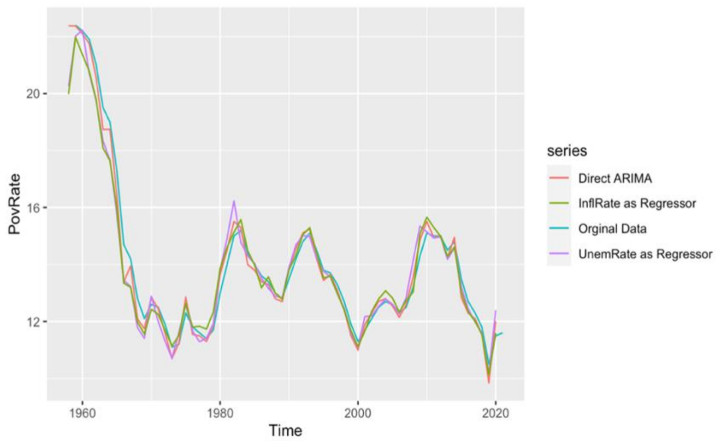

Figure 9: Modeling of poverty rate in the US using three different methods.

In terms of forecasting, the three models do have many differences, as shown in figure 10. The ARIMA (1, 1, 0) forecast is fairly simple, presenting a horizontal line and a confidence range that is ever expanding. The two dynamic regression models both predicted an immediate increase in poverty rate, which will reach a maximum level around the year 2045. The two dynamic regression models then display a difference starting around the year 2050, where the forecasting value using inflation rate increases slightly and the forecasting value using unemployment rate keeps decreasing. Afterwards, the dynamic regression model using inflation rate keeps decreasing with varying rate, while the regression model using unemployment rate increases to another local maximum in the year

2072. Finally, both of the poverty rate predicted by the dynamic regression models decrease to a similar level, which is about 1.2 higher than the ARIMA (1, 1, 0) model.

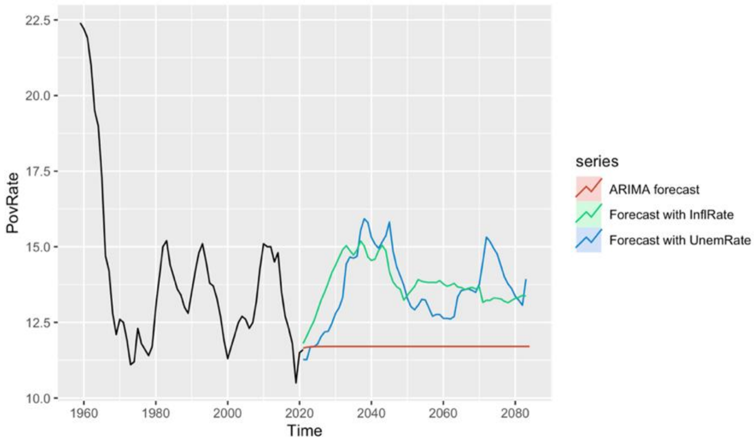

Figure 10: Forecast of poverty rate in the US using three different methods.

4.2. Introducing further model

Apart from unemployment rate and inflation rate, other economic factors can also be considered. However, some of them might not produce results that seems reliable. For example, in this part of the discussion, the dynamic regression model using annual GDP is considered and presented in figure 11.

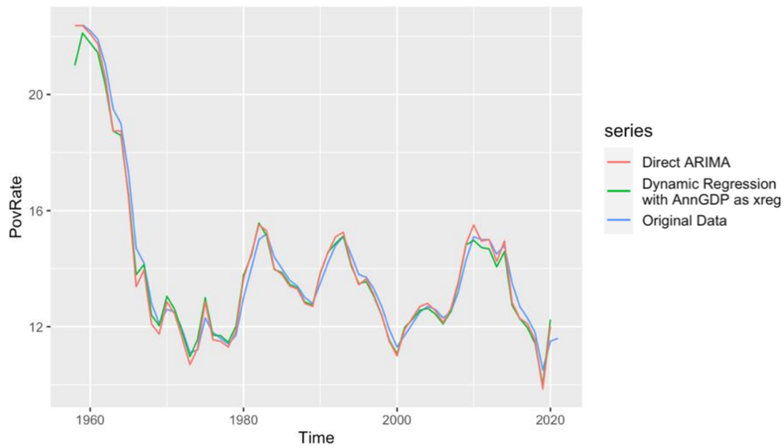

Figure 11: Fitting of dynamic regression model of poverty rate with annual GDP.

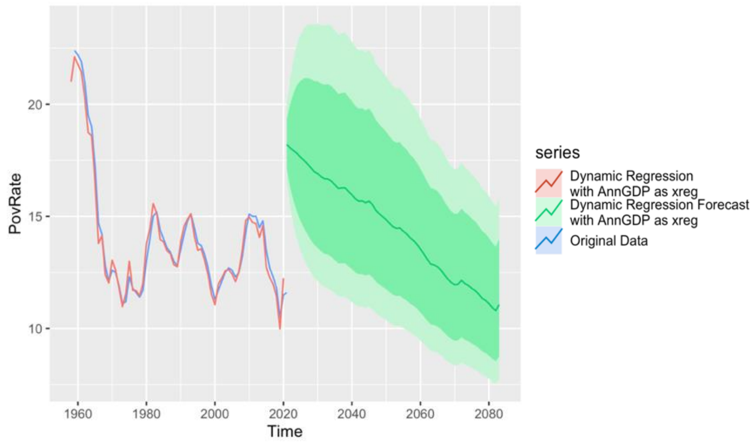

Figure 12: Forecast of dynamic regression model of poverty rate with annual GDP.

Still, as presented in figure 12, the poverty rate predicted using annual GDP is quite absurd, this would be an example of overfitting while still applying the model for forecasts. This could be because different economic factors relate to each other differently.

4.3. Overview of all models and their statistical linkage to the real world

In order to explain the different results obtained in the forecast for poverty rate when different regressors are used, further research needs to be done.

Based on researches conducted by the National Bureau of Economic Research, poverty rate is related to many other economic indicators For example, according to the authors, on average, an increase in the unemployment rate by 1% would lead to an 0.4% to 0.7% increase in poverty rate [9].

Also, according to a group of researchers who examined the relationship between poverty rate and other economic factors in Indonesia, inflation rate does not have any significant effect on unemployment rate, and thus poverty rate. Although the economic situation in Indonesia could be completely different from that of the United States, the overall relationship between inflation rate and unemployment rate is transferable [10].

4.4. Potential improvements and future outlooks

Apart from the economic factors discussed in this paper, there are many other interesting statistical relationships waiting to be discovered, such as the level of government subsidies, family structure, and reserve rate. All of these indicators can be used to generate dynamic regression models and forecasts, thus examining the change of poverty rate.

Still, the method used in this paper does not indicate which economic factor is the best for forecasting poverty rate, that is, which indicator relates to poverty rate most directly. Therefore, it can be proposed that, the data of poverty rate can be split into a training set and a test set and used to test whether the poverty rate predicted by the economic indicator fits the actual data.

Finally, one potential application is using the relationship between certain economic factors to statistically verify the validity of certain measures to lower the poverty rate. For example, if the government aims to reduce poverty rate, their measure probably should include lowering the unemployment rate as well, since it has been statistically proven to have a strong relationship with poverty rate.

DOI: 10.54254/2754-1169/49/20230534

5. Conclusion

To conclude, upon modeling and forecasting, the poverty rate of the United States is predicted to go through an increase that would reach around 15 to 16 percent and then gradually decrease or stabilizes around 12% for the nest 30 years. This predicted increase in poverty rate might be caused by the cyclicity of the poverty rate at seems and its sudden increase in 2019 due to the global pandemic. This is probably a sign that immediate action upon decreasing poverty rate should be taken so that situations do not worsen to an irreparable level. The US government might also want to consider offering measures that might reduce the impact of an increasing poverty level, such as providing subsidies and loans to citizens.

The potential cyclicity of the poverty rate could be caused by the actions of the government. It is possible that after the increase in the poverty rate and the implemented governmental measures, the government was able to lower the poverty rate, after which the government reduces the strength of their measures, thus poverty rate increases naturally.

In the future, more economic indicators can be taken into account and different models can be compared using fitting and test data sets, thus providing a more precise outcome in predicting the poverty rate of the United States. It is only through more accurate and complete data analysis can we use statistical measures to explain changes in economic indicators and find solutions to today’s challenges.

References

[1] *Census.gov. Poverty in the United States: 2021.https://www.census.gov/library/publications/2022/demo/p60-277.html, last accessed 2023/8/2.*

[2] *World Bank Blogs. March 2023 global poverty update from the World Bank: The Challenge of Estimating Poverty in the pandemic. https://blogs.worldbank.org/opendata/march-2023-global-poverty-update-world-bank-challenge-estimating-poverty-pandemic#:~:text=At%20%246.85%2C%20the%20global%20poverty,Asia%20and%20Sub%2DSaharan%20Africa. , last accessed 2023/8/2.*

[3] *UNDP. Sustainable development goals: United Nations Development Programme. https://www.undp.org/sustainable-development-goals/no-poverty?gclid=Cj0KCQjw2qKmBhCfARIsAFy8buJ_aTlE3v2q3SSjWuj0yx4ZDUTNrrwDYjqM9AlJl5n2c-kZItqyOa4aAmSgEALw_wcB, last accessed 2023/8/2.*

[4] *Lebergott, S. (1957). Annual estimates of unemployment in the United States, 1900-1954. In The measurement and behavior of unemployment (pp. 211-242). NBER.*

[5] *The Balance. An annual review of the U.S. economy since 1929. https://www.thebalancemoney.com/us-gdp-by-year-3305543, last accessed 2023/8/2.*

[6] *The United States Census Bureau. How the Census Bureau Measures Income and Poverty, https://www.census.gov/newsroom/blogs/random-samplings/2016/09/how-the-census-bureau-measures-income-and-poverty-4.html, last accessed 2023/8/2.*

[7] *U.S. Bureau of Labor Statistics. How the government measures unemployment. https://www.bls.gov/cps/cps_htgm.htm, last accessed 2023/8/2.*

[8] *Patel, N., & Villar, A. (2016). Measuring inflation. BIS Paper, (89b).*

[9] *Gunning, & Willem, J. . (2013). Why poverty persists: poverty dynamics in asia and africa edited by bob baulch. Agricultural Economics, 44(2), 253-254..*

[10] *Hasiholan, D., Hamzah, M. Z., Sofilda, E., & Hariyanti, D. (2022). Relationship Analysis between Unemployment and Poverty in 33 Provinces in Indonesia. OIDA International Journal of Sustainable Development, 15(02), 27-36.*