Affecting Factors and Forecasting Trends of American GDP

Chenxiao Luo[1, †], Yuting Xue[2, *, †]

[1]Department of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China

[2]Department of Statistics and Applied Probability, Santa Barbara University, California, America

*Corresponding author: yutingxue@ucsb.edu

[†]These authors contributed equally.

Abstract. The predictors of Gross Domestic Product (GDP) have been a heated topic of research in recent years. Researchers have used diverse methods or models to find the factors that may affect GDP to some extent. However, there remain some gaps in the application of multiple regression models in GDP analyses. This paper aims to predict the trend of U.S. GDP by analyzing the trends of four variables and applying multiple regression models of these variables. As is found in the research, unemployment rate, total primary energy consumption and labor productivity of nonfarm business and nonfinancial corporations have an impact on GDP and can be utilized to predict GDP. Petroleum, hydroelectric power, solar energy and wind energy can also be used to forecast GDP. The findings of this paper may be conducive to more thorough and sophisticated analyses of GDP, the principal metric that is used to assess the performance of a nation's economy.

Keywords: GDP; unemployment rate; CPI; energy consumption; productivity.

1. Introduction

GDP is considered to be the best indicator of a country's economic activity. It not only reflects the economic situation of a country in a certain period of time but also shows the wealth and power of the country. Therefore, effective forecasting of GDP is important for the country to formulate policies: if GDP shows a downward trend, the country should propose solutions for the factors affecting GDP; if GDP shows an upward trend, the country needs to maintain the current situation and expand its advantages.

In this paper, four influencing factors are selected as variables to predict U.S. GDP from 2012 to 2022, including unemployment rate, Consumer Price Index (CPI), labor productivity and energy consumption, because the relationships between these four factors and GDP have been studied by many scholars. Firstly, according to Okun's Law, a 1% increase in unemployment is usually related to a 2% decrease in GDP [1]. Besides, Mehrnoosh and Feizolah's study verified that unemployment decreased economic growth from 1996 to 2012 in Iran by using the Autoregressive Distributed Lag Model [2]. However, in researches related to the analysis of the relationship between unemployment rate and GDP, there is no such close relationship in the Visegrád Four countries from 2008 to 2020 [3]. Hence, there are many factors, such as region or time, that can cause the relationship between unemployment rate and GDP to change and previous findings may not apply to the U.S. or a particular time period. In order to predict GDP more accurately, regression models are built in this paper to determine whether each variable is significantly related to GDP. Secondly, Mehrnoosh and Feizolah's study also presented that CPI had a negative effect on economic growth in Iran. In addition, in the study regarding income level, energy consumption and GDP in Sub-Saharan African countries, the panel estimation demonstrates that the Granger causation between CPI and GDP is unidirectional [4]. Therefore, in this paper, CPI is chosen as a variable to predict the U.S. GDP. Thirdly, according to the research on the relationship between labor productivity and GDP in seven OECD countries for the period between 2008 and 2014, the conclusion asserted that productivity had an impact on economic growth through decreasing input costs and efficient use of the production factor [5]. Besides, a previous study suggested that measuring total factor production is a necessary prerequisite for estimating a region's growth by using the data from the 20 Italian regions [6]. As a result, labor productivity can be considered as an influencing factor to predict the GDP. Finally, the study which

examined the association between energy consumption and GDP in Sub-Saharan Africa concluded that energy consumption and GDP are interdependent through using a panel cointegration method [4]. Moreover, another research divided 82 countries from 1972 to 2002 into four groups by income to examine the relationships between them and found that there was no evidence to prove that energy consumption could boost the GDP in any of the four income groups [7]. Unlike the classification in that study, in this paper, energy sources are classified by type and those that have a significant impact on GDP are selected by multiple regression models to predict GDP. For example, prior research verified that the nuclear energy affected GDP positively, and there is a bi-directional relationship in the long term between renewable energy consumption and economic growth [8, 9].

After discussing the four variables' possible relationships with GDP through various previous studies, this paper will use line graphs to further demonstrate their relationships with the U.S. GDP from 2012 to 2022 by comparing the changes in GDP and the four variables over time. Then multiple regression models can test whether these variables can be applied to predict the trend of GDP. In a previous study, the multiple linear regression model was used to test whether final consumption and total investments had an impact on the GDP of Romania [10]. As a result, this statistical method can be used to investigate the relationships between the four variables and the U.S. GDP in this paper. Then factors that have a statistically significant correlation with GDP can be applied to forecast GDP.

2. Methodology

2.1 Data

In order to predict the trend of U.S. GDP, four factors that are considered to have an impact on GDP are chosen as variables, including unemployment rate, CPI, labor productivity and energy consumption. The data on the unemployment rate, CPI and labor productivity comes from the latest news releases of the Bureau of Labor Statistics which serves as the Federal Government's primary fact-finding agency and provides a large amount of reliable data on U.S. economic and statistical fields. The data on energy consumption is from the U.S. Energy Information Administration which is a statistical government agency in the field of energy. Besides, the time of data includes each quarter between 2012 and 2022 to ensure a broad timespan to observe variables' trends. After confirming these four variables, the null hypothesis that unemployment rate, CPI, labor productivity and energy consumption do not affect GDP and cannot be used to predict it is formulated and then tested using line graphs and multiple regression model.

2.2 Line Graphs

To analyze the relationships between each variable and GDP, line graphs, which are drawn by Excel with time as the horizontal axis and each variable and GDP as the two vertical axes, are used to show the trends of four variables - unemployment rate, CPI, labor productivity and energy consumption - from the first quarter of 2019 to the last quarter of 2021. These graphs effectively show changes in different variables over time and compare these four variables with GDP. In addition, in order to predict GDP trends more accurately, labor productivity and energy consumption are broken down into smaller categories, which helps identify variables that are more closely related to GDP. Labor productivity is divided into labor productivity of manufacturing, nonfarm business and nonfinancial corporations. Energy consumption is categorized into fossil fuel consumption and renewable energy consumption. If a variable in the line graphs has the same trend as GDP, there may be a relationship between this variable and GDP, so these line graphs can be used to test the proposed hypothesis.

2.3 Multiple Regression Model

Since the line graphs do not directly prove the association between the variables and GDP, multiple regression models are able to provide more accurate statistical evidence to test the hypothesis. Multiple regression is a statistical method to analyze the relationships between more than one

independent variable and a dependent variable as well as help achieve forecasts. The equation of multiple regression is

$$y = b_1x_1 + b_2x_2 + \cdots + b_nx_n + c \quad (1)$$

In this equation, beta is a regression coefficient. If beta is positive, the independent variable and the dependent variables have a positive connection. Conversely, if beta is negative, the association between the independent variable and the dependent variable is negative. Besides, y is the dependent variable and x is the independent variable. In this paper, independent variables are unemployment rate, CPI, labor productivity and energy consumption. The outcome (dependent variable) is GDP. In addition, in the result of the regression analysis, p-value is applied to examine the null hypothesis. The null hypothesis is disproved if the p-value is less than 0.05, indicating that there is a statistical significance between the independent variable and the dependent variable. The null hypothesis is proved to be true if the p-value is less than 0.05, which means that the association between the independent variable and the dependent variable is statistically significant. Besides, R-squared is used to measure whether the regression model is fit. As a result, the multiple regression can be used to effectively test the relationships between the four variables and GDP.

3. Results and Discussion

3.1 Descriptive Analysis

U.S. GDP experienced a long-term steady growth from 2012 to 2019 except in the last quarter of 2013 when growth was slightly weighed down by a smaller contribution from exports and retail sales, as well as a sharp fall in federal government expenditures. However, the COVID-19 pandemic deeply affected U.S. economy, resulting in a recession of real GDP growth by 15 percent in the second quarter of 2020. It is also worth noticing that U.S. GDP recovered the fastest among G7 economies and even surpassed its pre-pandemic trend in 2021 despite elevated inflation.

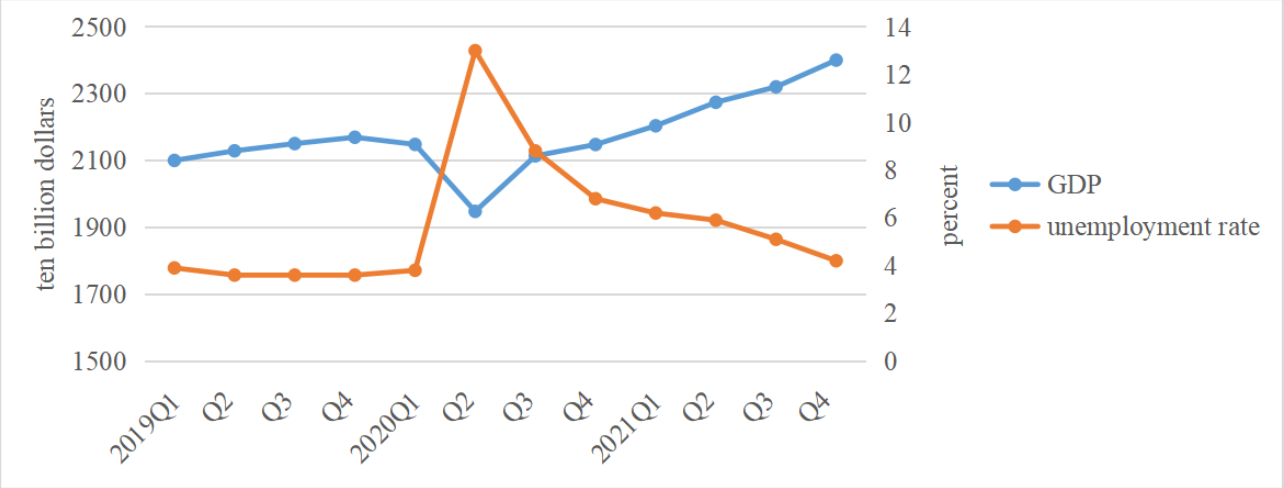

Fig. 1 Trend of GDP and unemployment rate over time

U.S. unemployment rate declined constantly from 2012 to 2019 but proliferated by a record amount during the pandemic. However, the labor market slowly recovered for the rest of 2020 and strongly rebounded in 2021, thanks to the stimulus bill of the Biden administration and the widespread availability of vaccines. Fig. 1 demonstrates a negative correlation between GDP and unemployment rate that as unemployment rate decreases, GDP increases. When unemployment rate is at the peak, GDP reaches a low point.

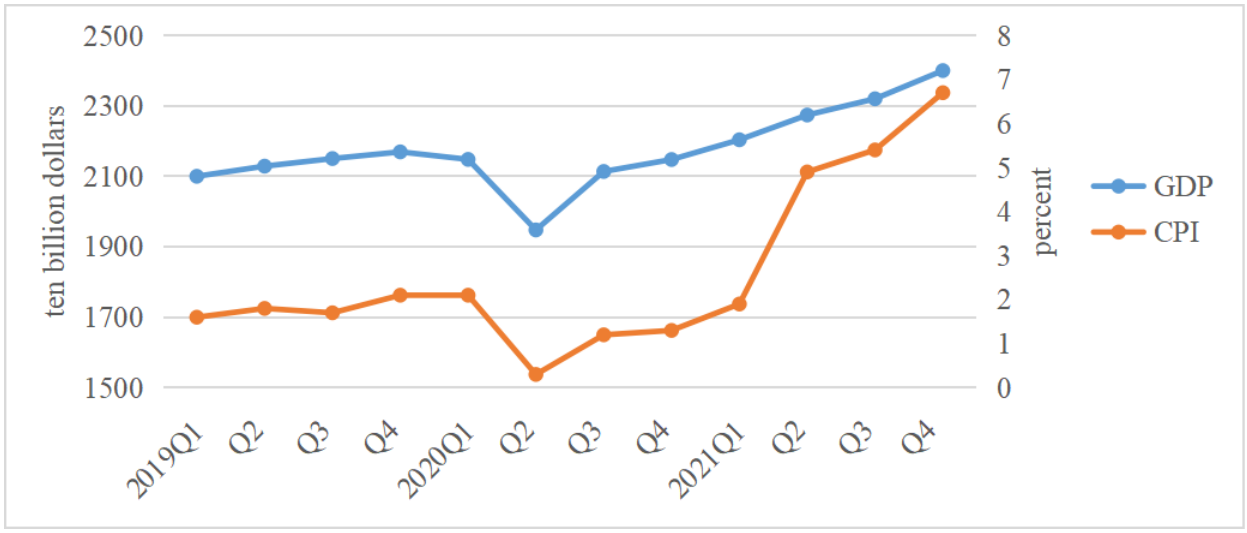

Fig. 2 Trend of GDP and CPI over time

According to the U.S. Bureau of Labor Statistics, CPI is used to evaluate the average change in household spendings on products and services over time. There are frequent fluctuations of U.S. CPI within the range between 1% and 3% with two exceptions. One exception is in 2015 when the energy prices were relatively low. The other is during the pandemic when CPI fell to almost 0 in the second quarter of 2020 and then surged above 6% in just one and a half years. The high inflation rate is partly attributed to the disruptions of the supply chain and the inflationary public spendings of the U.S. government and therefore weakens the purchase power of U.S. households. From Fig. 2, it can be considered that GDP and CPI roughly follow the same trend.

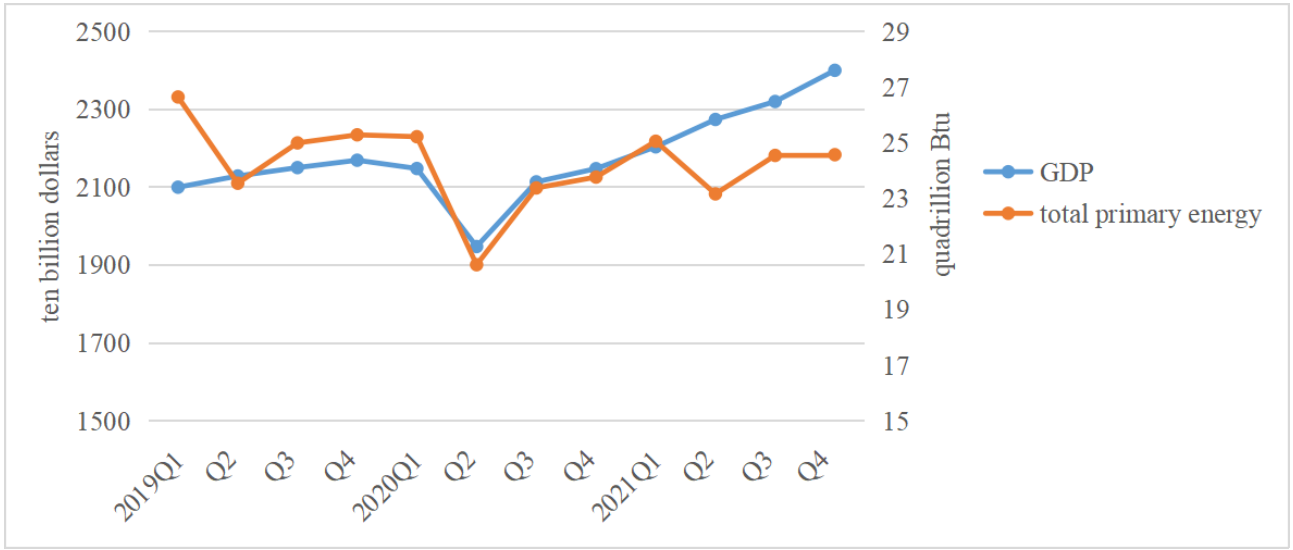

Fig. 3 Trend of GDP and total primary energy consumption over time

The U.S. overall energy consumption fluctuates around 25 quadrillion Btu with a valley in the second quarter of 2020. Fig. 3 seems to show some correlation between GDP and total primary energy consumption. But in a broader timespan, there is no relationship between them.

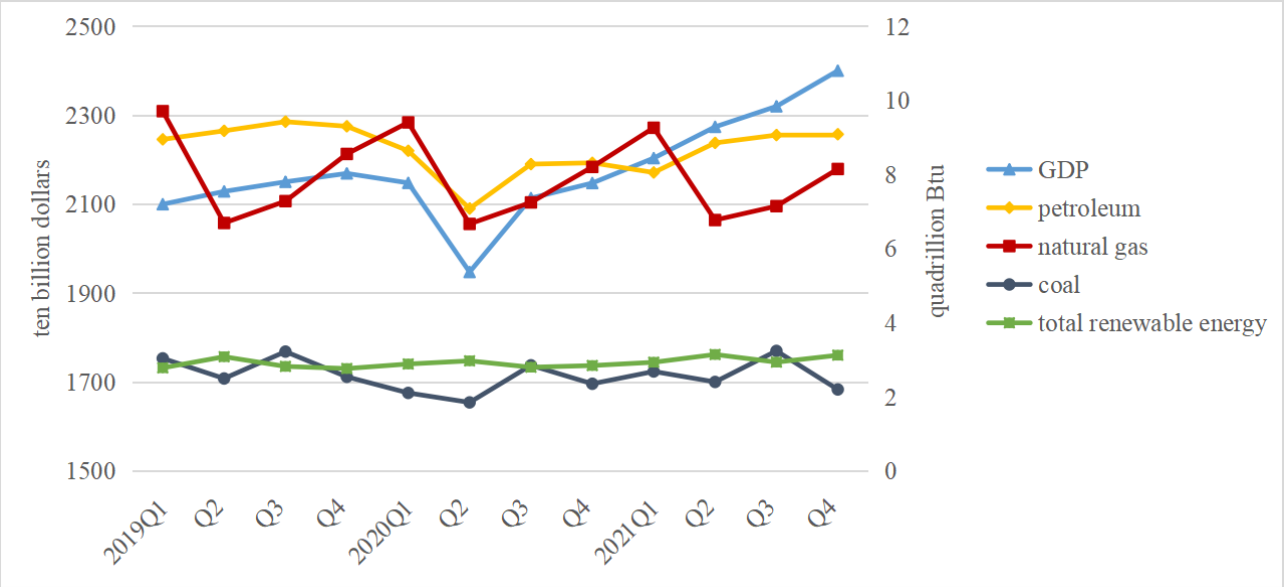

Fig. 4 Trends of GDP and primary energy over time

Similarly, the consumption of petroleum, which has been rather stable since 2012, also dropped during the pandemic. The consumption of natural gas has periodic variations that the peaks appear in winters and the valleys appear in summers. The consumption of coal shows a rapid decline, while the consumption of renewable energy gradually increases. As a result, coal has been exceeded by renewable energy in recent years. It is alleged that U.S. coal consumption will continue its decline and will not return to its pre-2010 levels even if some policies are conducted in order to encourage coal production. At the same time, renewable energy will play a more vital role in energy consumption. As is displayed in Fig. 4, there is relationship between GDP and petroleum consumption.

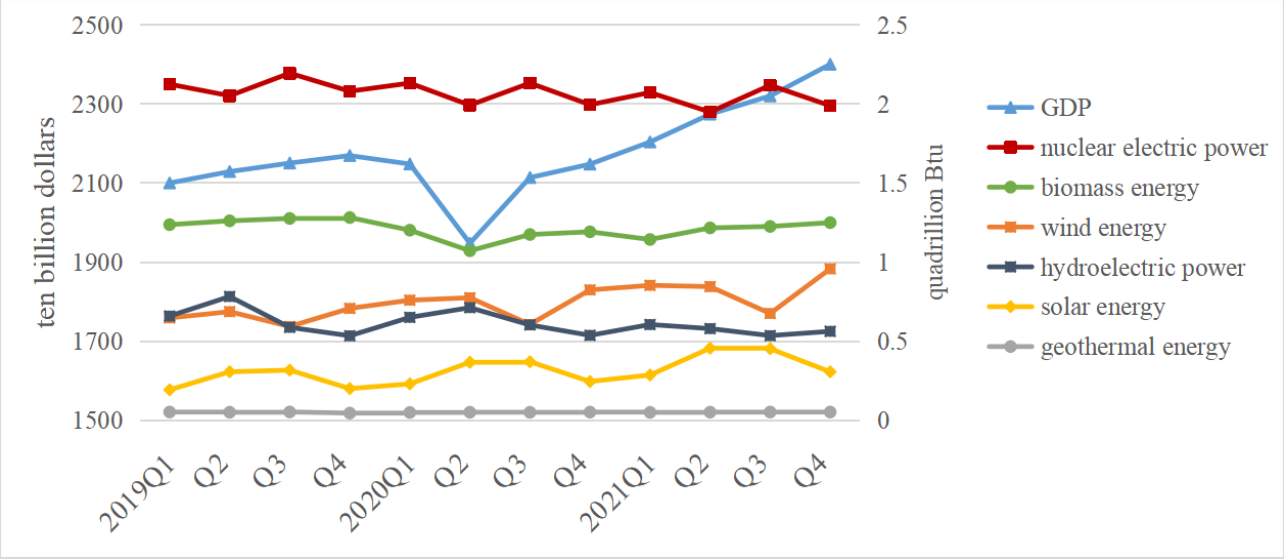

Fig. 5 Trends of GDP and renewable energy over time

When taking a closer look at the various classes of renewable energy, it is not difficult to distinguish which ones of them make the major contributions to the increase of renewable energy consumption. Nuclear power, biomass energy and hydroelectric power account for most of the proportion in renewable energy. Wind energy and solar energy benefit from the development of technology which makes them more efficient in energy conversion and less subject to the environmental conditions. Geothermal energy does not show obvious progress in development

compared with solar energy. Fig. 5 shows the steady growth of GDP and wind and solar energy consumption, which indicates their correlation.

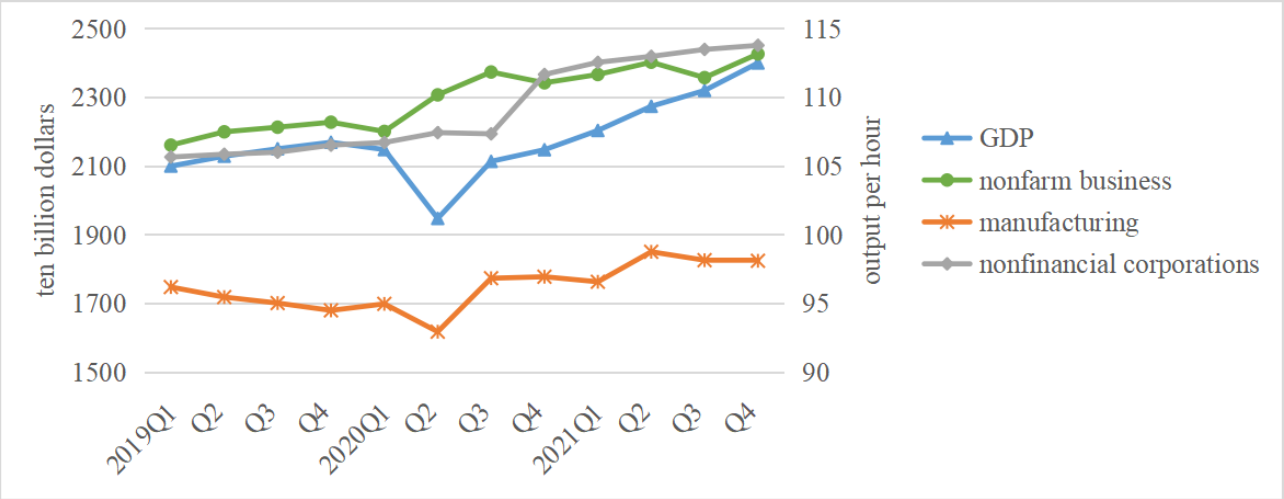

Fig. 6 Trends of GDP and labor productivity over time

Labor productivity is measured by output per hour of an individual labor. The labor productivity of nonfarm business and nonfinancial corporations have been steadily increasing since 2012 and were not affected by the pandemic. However, the labor productivity of manufacturing kept dropping even before the pandemic, which might be a result of a reduction in the rate of capital deepening in manufacturing. The productivity of manufacturing recovered slightly after the pandemic. In Fig. 6, GDP and labor productivity of nonfarm business and manufacturing follow similar trends.

3.2 Multiple Regression Analysis

There are two multiple regression models with the outcome, GDP. One includes the four variables: unemployment rate, CPI, labor productivity, which is divided into labor productivity of nonfarm business, manufacturing and nonfinancial corporations, and total primary energy consumption. The other model includes several primary types of energy consumption.

Table 1. The result of the multiple regression

Coefficients	Estimate	Standardized	Std. Error	t value	Pr(>|t|)
(Intercept)	-1.471×10^{3}	NA	3.534×10^{2}	-4.161	< 0.001
unemployment rate	-3.449×10^{1}	-3.160×10^{-1}	1.902	-18.128	< 0.001
CPI	1.324×10^{1}	8.205×10^{-2}	2.993	4.424	< 0.001
total primary energy	9.948×10^{1}	5.739×10^{-3}	2.752	0.361	0.720
nonfarm productivity	3.909×10^{1}	7.544×10^{-1}	3.218	12.149	< 0.001
manufacturing productivity	-8.180	-7.954×10^{-2}	2.511	-3.257	0.003
nonfinancial productivity	2.513	4.708×10^{-2}	2.876	0.874	0.389

In Table 1, p-values of unemployment rate, CPI and labor productivity of nonfarm business and manufacturing are well less than 0.05, which means that the null hypotheses are rejected and the relationships between these variables and GDP are statistically significant. Besides, the multiple R-squared is 0.995, so there is a substantial proportion of variance and the model fits the observations well. Furthermore, the value of beta constitutes a significant part in predicting the trend of GDP. For the unemployment rate and labor productivity of manufacturing, beta is negative so that their influence on GDP is negative. According to the regression equation, for each unit increase in unemployment rate or labor productivity of manufacturing, GDP decreases by the absolute values of their beta (3.449×101, 8.180). For CPI and labor productivity of nonfarm business, since the beta of them is positive, the relationships between them and GDP is positive. Based on the equation, for each

unit increase in the CPI or nonfarm business productivity, GDP increases by the absolute values of their beta (1.324×101, 3.909×101). In conclusion, unemployment rate, CPI and labor productivity of nonfarm business and manufacturing have an impact on the GDP and can be utilized to predict GDP.

Table 2. The result of the multiple regression of energy consumption

Coefficients	Estimate	Standardized	Std. Error	t value	Pr(>|t|)
(Intercept)	4.658×10^{2}	NA	3.451×10^{2}	1.350	0.187
coal	-3.148	-1.188×10^{-2}	1.803×10^{1}	-0.175	0.863
gas	1.318×10^{1}	7.122×10^{-2}	9.734	1.354	0.185
petroleum	1.481×10^{2}	2.857×10^{-1}	2.739×10^{1}	5.408	< 0.001
nuclear electric	1.541×10^{2}	5.713×10^{-2}	1.183×10^{2}	1.302	0.202
hydroelectric	-2.705×10^{2}	-1.041×10^{-1}	7.669×10^{1}	-3.528	0.001
geothermal	-3.203×10^{3}	-2.328×10^{-2}	4.909×10^{3}	-0.653	0.519
solar	8.133×10^{2}	4.330×10^{-1}	1.209×10^{2}	6.725	< 0.001
wind	6.258×10^{2}	5.120×10^{-1}	1.154×10^{2}	5.422	< 0.001
biomass	-3.243×10^{2}	-8.403×10^{-2}	2.275×10^{2}	-1.426	0.164

In Table 2, p-values of petroleum, hydroelectric power, solar energy and wind energy consumptions are well less than 0.05. Hence, the null hypothesis that these variables do not influence GDP is rejected and the relationships between these variables and GDP are statistically significant. Other energy consumptions whose p-values are larger than 0.05 have no effect on GDP. Besides, the multiple R-squared (R^2 = 0.979) is large enough to show that the regression model fits the observations well. In addition, the beta of hydroelectric power consumption is negative (b = -2.705×10^{2}) and presents that for each unit increase in hydroelectric consumption, GDP decreases by the absolute value of its beta. The betas of nuclear electric power, solar energy and wind energy consumptions are positive, which shows that for each unit increase in them, GDP decreases by the absolute values of their beta (1.481×10^{2}, 8.133×10^{2}, 6.258×10^{2}). In conclusion, petroleum, hydroelectric power, solar energy and wind energy consumptions have an effect on GDP and can be applied to forecast the trend of GDP.

4. Conclusion

It can be concluded from the regression models that unemployment rate, CPI and labor productivity of nonfarm business and manufacturing have impacts on GDP and can be utilized to predict GDP. Petroleum, hydroelectric power, solar energy and wind energy consumptions have impacts on GDP and can be applied to forecast the trend of GDP. Unemployment, CPI and labor productivity of nonfarm business and manufacturing are closely related to production and consumption, which determine the survival of most enterprises. Energy consumption and labor productivity of nonfinancial corporations tend to be steady or in recession in the U.S, therefore are not a key figure in assessing the trend of GDP. As for the energy sources that can be utilized to predict GDP, petroleum is the main source of energy and solar and wind energy are two promising renewable energy sources in the U.S. This research is beneficial to the deeper and more complex analyses of GDP which are expected to explore the profound driving forces of GDP. At the same time, this research can be an object of reference to individuals in terms of investment and other economic activities in the U.S. The variables of this research are selected based on the previous papers and more variables might be suitable for analysis as the U.S. economy evolves.

References

[1] Okun A. M. Potential GNP: Its Measurement and Significance, Proceedings of the Business and Economics Statistics Section of the American Statistical Association, 1962: 98-104.

[2] Mehrnoosh M. and Feizolah J. Examining the effects of inflation and unemployment on economic growth in Iran (1996-2012). Procedia Economics and Finance, 2016, 36: 381-389.

[3] Mura L., Zsigmond T., Kovacs A. and Baloghova E. Unemployment and GDP relationship analysis in the Visegrad Four Countries. Online Journal Modelling the New Europe, 2020, (34).

[4] Mulugeta S. K., Chali N., Peter V. S. and Tesfa G. G. Income level and the energy consumption–GDP nexus: Evidence from Sub-Saharan Africa. Energy Economics, 2012, 43(3): 739-746.

[5] Korkmaz S. and Korkmaz O. The relationship between labor productivity and economic growth in OECD Countries. International Journal of Economics and Finance, 2017, 9(5): 71-76.

[6] Napolitano O., Pietroluongo M. and Kounetas K. Stochastic convergence or divergence of total factor productivity and GDP of Italian Regions. Re-examing the evidence. Economics Bulletin, 2018, 38(4): 1857-1863.

[7] Bwo-Nung H., Hwang M.J. and Yang C.W. Causal relationship between energy consumption and GDP growth revisited: A dynamic panel data approach. Ecological Economics, 2008, 67(1): 41-54.

[8] Al-Mulali U. Investigating the impact of nuclear energy consumption on GDP growth and CO2 emission: A panel data analysis. Progress in Nuclear Energy, 2014, 73: 172-178.

[9] Al-Mulali U., Fereidouni, H. G., Lee, J. Y. and Sab, C. N. B. C. Examining the bi-directional long run relationship between renewable energy consumption and GDP growth. Renewable and sustainable energy reviews, 2013, 22: 209-222.

[10] Anghelache C., Manole A. and Anghel M. G. Analysis of final consumption and gross investment influence on GDP–multiple linear regression model. Theoretical and Applied Economics, 2015, 22(3):137-142.