Correlations of Gross Domestic Product with Crime Rate and Opening Stock Price

Research Statistics 3

Period 5

6/5/2022

Thomas Jefferson High School for Science and Technology

Rationale

Gross Domestic Product (GDP) is the monetary measure of all the goods and services in a country during a certain period. It is commonly used as a measure of a country's wealth as it encompasses all the monetary transactions that take place in a country, however, it can also be used as an indirect measurement of the quality of life that the citizens of a country have. Higher GDPs tend to correlate to higher economic growth and a higher quality of life for most cases. However, a GDP is only a rough indicator because it does not directly account for levels of health and education, leisure, changes in income, and cultural values. Similarly, higher stock prices are generally considered a sign of economic prosperity, which usually relates to a higher quality of life. Violent crime rates are one of the factors that inhibit the quality of life for many citizens. The purpose of this study is to analyze how the factors of economic growth, such as stock prices, or the factors relating to a quality of life, such as violent crime rates, affect GDP.

The positive correlation between stock prices and GDP can be seen in Tursoy &Faisal's research (2016). Their study investigated the relationship between stock prices and GDP in Turkey using the quarterly data from 1989's second quarter to 2014's second quarter of Turkey. Using an autoregressive distributed lag framework, which is a version of a least squares regression line that utilizes the lags of both the dependent and the explanatory variables. They found there was strong evidence that both the stock prices and the GDP positively cointegrated in the long run. The adverse relationship was analyzed in Shiblee's research (2009). For an investor stock price is very important, and it is influenced by a variety of factors. Shiblee used data from the New York Exchange and the Federal Reserve website; during the time period, 1994-2007. The findings from this research were that each of the different variables had a different effect on the stock price. The most influential factor was the money supply which was

used as a replacement for GDP. These findings are echoed by a similar study, Reddy (2012). Reddy used the stock prices of various companies from 1997 to 2009 and India's GDP over the years. The results from this study showed that 95.6% of the variation in stock prices was explained through Real Gross Domestic Product (RGDP), Interest Rate (INT), and Inflation Rate and that RGDP showed a positive correlation with stock prices. On the other hand, as we see from Rosenfeld and Fornango's study (2008) that economic growth had significant effects on robbery and property crime rates when controlling for factors thought to influence variation in crime rates. They had used data from the Federal Reserve and the U.S. Department of Justice to reach this conclusion.

Methods

The unit of observations would be the individual years from 1990-2018. Although these years were not randomly chosen due to the limitations of the dataset from which we obtained the crime rate, the index fund was randomly chosen. We did this by compiling a list of profitable index funds. The chosen index fund was S&P 500. The crime rate data were obtained from the World Bank. Their data was calculated using the number of reported crimes in the U.S. We used the opening historical stock prices provided by Yahoo! The data provided by them came in formats of months. However, the World Bankonly provided data by year. To fix this, we found the average opening price of each year and used that so the data would stay consistent. We decided to do a multiple regression to check if there is a statistically significant relationship between the crime rate and the average opening stock price on GDP per capita.

Table 1

Compiled Data

Year	Avg. Opening Stock Price	Crime Rate	GDP per capita
2018	2755.66	4.96	$63,064
2017	2434.68	5.32	$60,110
2016	2088.91	5.39	$58,021
2015	2054.11	4.95	$56,863
2014	1926.67	4.44	$55,050
2013	1618.35	4.53	$53,107
2012	1372.59	4.73	$51,603
2011	1281.40	4.71	$49,883
2010	1120.59	4.76	$48,467
2009	930.48	5.03	$47,100
2008	1262.03	5.43	$48,383
2007	1473.59	5.70	$47,976
2006	1304.93	5.81	$46,299
2005	1204.74	5.67	$44,115
2004	1125.63	5.52	$41,713
2003	948.59	5.70	$39,496
2002	1010.94	5.65	$38,023
2001	1200.10	6.69	$37,133
2000	1432.14	5.53	$36,335
1999	1310.58	5.57	$34,514
1998	1066.29	6.17	$32,854
1997	856.72	6.70	$31,459
1996	664.45	7.32	$29,968
1995	533.82	8.15	$28,691
1994	461.24	8.89	$27,695
1993	450.91	9.45	$26,387
1992	415.56	9.25	$25,419
1991	374.28	9.71	$24,342
1990	334.62	9.30	$23,889

Results

Our null hypothesis was that $\beta_i=0$ and our alternative hypothesis was $\beta_i\neq0$ where i either equals the coefficient of Crime Rate or Avg. Opening stock price. First, we ran regressions with just "Crime Rate" or "Avg. Opening stock price". These both had significant p-values for the predictors and R^2 values of 70.62% and 82.2%, respectively. The conditions of a multiple regression include lack of multicollinearity, normally distributed errors, a linear relationship, homoscedasticity along with a large enough sample size for each independent variable. We had a sample size of 29 which is about enough. To check multicollinearity, we ran a regression between the independent variables to get the correlation coefficient. This turned out to be 0.5697 which is low enough to indicate the absence of multicollinearity. By itself, the p-values of the regression turned out to be significant with them all being less than 0.01 and there was a pretty strong R^2 value of 88.75%.

Table 2

Regression of GDP per Capita on Crime Rate and Average Opening Stock Price

Coefficients	Estimate	Std. Error	T-value	P-value	Fit
(Intercept)	43629.32	6450.04	6.764	$3.53*10^{-7}$***	
Crime Rate	-2650.86	716.34	-3.701	$1.02*10^{-3}$**	
Avg. Opening stock price	12.07	1.92	6.284	$1.19*10^{-6}$***	
					R^2=88.75%
					R^2(adj)=88.75%

***p<0. **p<0.01

However, the residuals were then graphed. As seen from the two figures below, there is both a lack of homoscedasticity and an indication that the data is not linear and is left-skewed. To solve this, we attempted to do regression by raising the GDP per Capita to a positive power.

Figure 1

Normal Probability Plot of Regression with Original Data

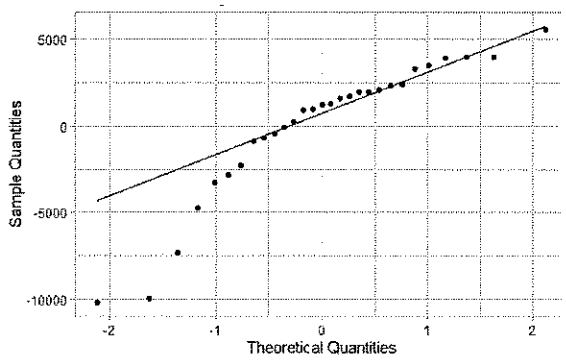
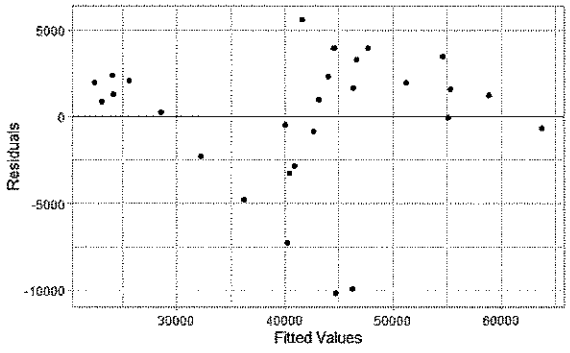


Figure 2

Residual Plot of Regression with Original Data



We tried to continue raising the power but the residual plots did not become normal. We found raising GDP per Capita to a power of 11 provided the best results. Table 3 shows the results of this transformation. Since this was so high we tried an inverse transformation but that actually turned the data from left-skewed to right-skewed indicating that it was too strong. A log transformation of the independent variables also did not fix anything. The p-values shown in table 3 even more strongly suggest a statistically significant relationship between crime rate and average opening stock and the transformed GDP per capita. However, the R^2 value did decrease from 88.75% to 83.29%. The residual plot still shows a lack of homoscedasticity but there also

seems to be an outlier present. The right-most point in Figures 3 and 4 is what we identified as a potential outlier with a standardized residual of almost 4 along with having a Cook's distance value of greater than one, marking it as influential.

Table 3

Transformed Regression of GDP per Capita on Crime Rate and Average Opening Stock Price

Coefficients	Estimate	Std. Error	T-value	P-value	Fit
(Intercept)	$-5.900*10^{52}$	$9.280*10^{51}$	-6.358	$9.83*10^{-7}$***	
Crime Rate	$4.982*10^{51}$	$1.031*10^{51}$	4.835	$5.20*10^{-5}$***	
Avg. Opening stock price	$2.876*10^{49}$	$2.763*10^{48}$	10.408	$9.15*10^{-11}$***	
					R^2=83.29%
					R^2(adj)=82.00%

***p<0. **p<0.01

Figure 3

Normal Probability Plot of Regression with Transformed Data

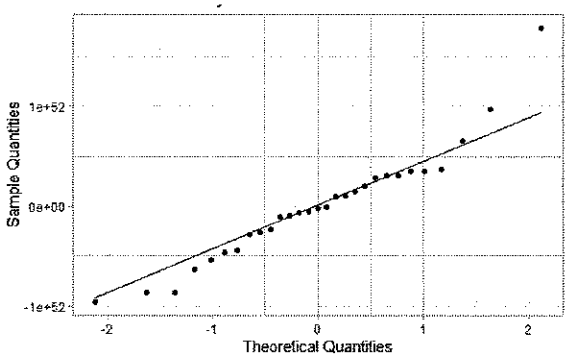
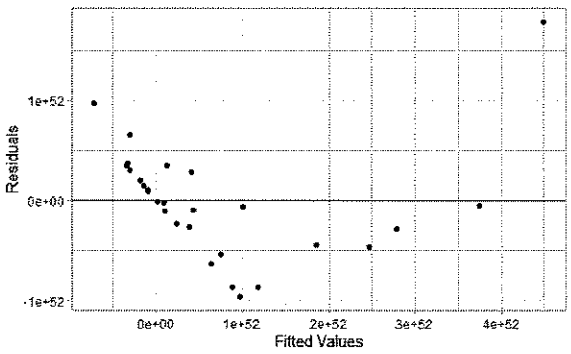


Figure 4

Residual Plot of Regression with Transformed Data



Figures 5 and 6 display plots for the regression of the original data without the outlier identified earlier. Judging from them, the plots appear to be not normal so inferences cannot be drawn from the regression. They appear to be left-skewed so once again we raised the response variables to powers greater than one.

Figure 5

Normal Probability Plot of Regression with Original Data without Outlier

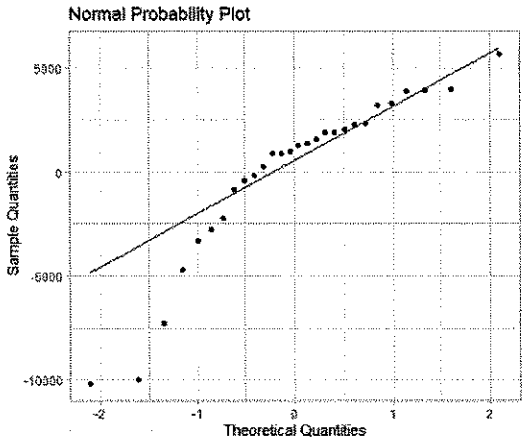
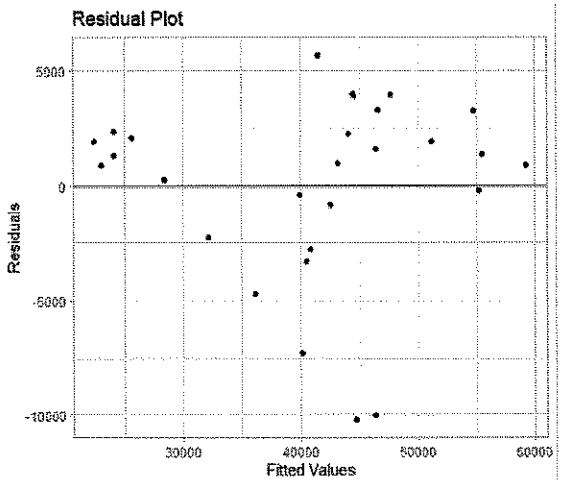


Figure 6

Residual Plot of Regression with Original Data Without Outlier



We were able to transform the regression into one with relatively normal residuals by raising GDP to the seventh power. As seen in table 4, the p-values were all still significant when considering an alpha value of 0.05. The residuals that resulted from this regression turned out to be the most normal. However, the R^2 value of 81.13% is lower than previous regressions but still relatively high.

Table 4

Transformed Regression of GDP per Capita on Crime Rate and Average Opening Stock Price Without Outlier

Coefficients	Estimate	Std. Error	T-value	P-value	Fit
(Intercept)	$-2.418*10^{33}$	$5.988*10^{32}$	-4.038	$4.5*10^{-4}$***	
Crime Rate	$1.764*10^{32}$	$6.406*10^{31}$	2.753	$1.083*10^{-2}$*	
Avg. Opening stock price	$1.614*10^{30}$	$1.942*10^{29}$	8.315	$1.15*10^{-08}$***	
					R^2=81.13%
					R^2(adj)=79.62%

***p<0. **p<0.01. *p<0.05

Figure 7

Normal Probability Plot of Regression with Transformed Data without Outlier

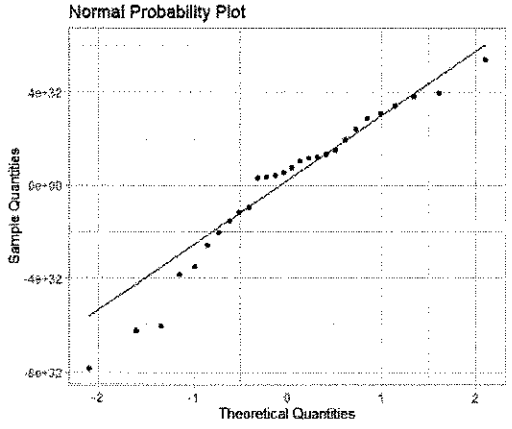
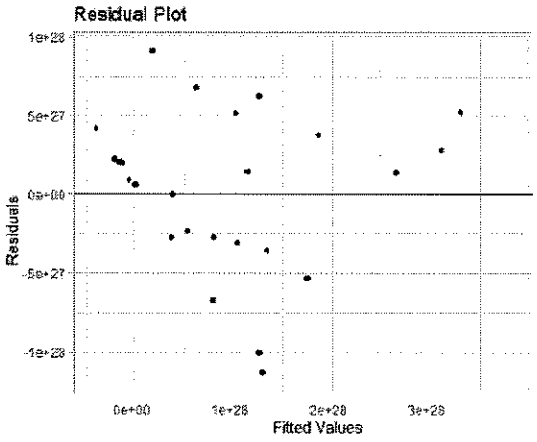


Figure 8

Residual Plot of Regression with Transformed Data without Outlier



Although the original regression model provided significant p-values for predictors and a high R^2 value of 88.75%, it did not satisfy the multiple regression requirements. Therefore we cannot make any conclusions about the cause and effect relationship between each of the variables, "Crime Rate" and " Avg. Opening stock price", have on Gross Domestic Product (GDP). We attempted to do a transformation by raising the response variable to the eleventh power to fix the left skew. This did not result in a normal residual graph but it was better. However, for the sake of this investigation, we proceeded as if the data we had fulfilled all of the requirements to create a multiple regression line. Both predictors remained significant with the p-values of the predictors decreasing after the transformation. When the outlier was removed and the data then transformed, thus providing a regression that practically met the conditions, the regression still produced significant p-values. Given this, both of the variables proved to be significant in determining GDP. As such we could conclude that yearly GDP is influenced by both the crime rate in the US and the opening stock price of the S&P 500. However, it does not actually provide a significant improvement in determining the variability explained by the model when using both variables versus just "Avg. Opening stock price" with an increase of 1.09% when using both when compared to the transformed regression. In fact, there is actually a decrease when comparing just "Avg. Opening stock price" with the transformed regression without the outlier. However, there was a much greater increase between both variables and just "Crime Rate" with an 11.58% increase. Due to the significant p-values of both the original and transformed regressions — with and without the outlier — we have enough evidence to reject the null hypothesis.

Discussion

Through our research, we found results similar to the different studies we reviewed. The positive correlation between stock prices and GDP that we observed can be seen in Tursoy and Faisal's research (2016). While our t value was even higher than theirs, the results accumulated to the same conclusion for the opening stock price. Even though they used quarterly data from only Turkey, both of us arrived at the same conclusion. Furthermore, Shiblee's research (2009) only furthers our sentiment when compared with ours. Finally, as we see from Rosenfeld and Fornango's study (2008) our results closely resemble theirs. Crime Rate has a significant effect on GDP as Rosenfeld and Fornago concluded. There are many implications of our study as we find a correlation between the US's GDP, the S&P 500 opening stock price, and the crime rate in the US. It shows us that we should focus on reducing crime if we want the general wealth of a country to increase. Therefore some possible applications could be government-sponsored programs that would aim to reduce the crime rate in order to increase the general welfare and the general economic affluence of the country. Before any action takes place, however, a more thorough investigation into the different types of crime and their implications for the GDP of the country. Future research that would need to be done would be finding the different types of crime and their individual impacts on GDP. Along with this, by looking into the different types of stocks' opening prices we could get a more in-depth understanding of what types of crime affect the GDP. Furthermore, our research only had implications for the United States. If we had the opportunity to pursue further research on this subject, we would try to find correlations internationally so we could make generalizations about a larger community. These generalizations could then be used to implement programs targeted to improve the GDP of every

country, which would let the citizens of the entire world reap the benefits of an improved economy.

Reflection

Through this research project, we gained experience in taking large datasets and running different significance tests. We also made our own more streamlined version of a dataset with different variables, taken from Yahoo!, that was used for our analysis. Furthermore, we went through many different transformations to try and fit the data to a linear model, which will serve as important knowledge in the future. We also learned other statisticians' approaches to ideas of wealth through our own background research. Finally, the future implications of our study were really eye-opening as while we could have predicted the outcome of our study, we learned of the many confounding variables present with crime and opening stock prices.

References

Reddy, L. (2012). *Impact of inflation and GDP on stock market returns in India - graph.* International Journal of Advanced Research in Management and Social Sciences . Retrieved June 8, 2022, from https://garph.co.uk/IJARMSS/Dec2012/9.pdf

ROSENFELD, R. I. C. H. A. R. D., & FORNANGO, R. O. B. E. R. T. (2008). *The impact of economic conditions on robbery and property crime: The ...* Wiley Online Library. Retrieved June 8, 2022, from https://onlinelibrary.wiley.com/doi/10.1111/j.1745-9125.2007.00096.x

Shiblee, L. S. (2009, December 30). *The impact of inflation, GDP, unemployment, and money supply on stock prices*. SSRN. Retrieved June 7, 2022, from https://deliverypdf.ssrn.com/delivery.php?ID=750066083101069103066079090114101076007056010023061049023084013108094106006127005111041119107107108043037095091115118103112089060086008008061121024081127087024087004064001071084067015089083024001082121121070010025029074064106086004099004028106001067114&EXT=pdf&INDEX=TRUE

Tursoy, T., & Faisal, F. (2016). *Bun Venit La Institutul Național de statistică | institutul național de ...* Romanian Statistical Review. Retrieved June 8, 2022, from https://insse.ro/cms/

U.S. Crime Rate & Statistics 1990-2022. MacroTrends. (n.d.). Retrieved June 7, 2022, from https://www.macrotrends.net/countries/USA/united-states/crime-rate-statistics

Yahoo! (2022, June 8). *S&P 500 (^GSPC) historical data*. Yahoo! Finance. Retrieved June 7, 2022, from https://finance.yahoo.com/quote/%5EGSPC/history?p=%5EGSPC