The Influence of Government Shutdown on Trump's Popularity

Linjing Li, Linda Meng, William Yang

March 4, 2019

1 Introduction

The US government was shut down from Dec 22, 2018 to Jan 25, 2019 for over one month. The incident provokes US citizens to have great concern over Trump's administration. We believe that such concern will further lead to a decrease in rate of support for Trump. Thus, we carried out this project to evaluate the correlation between Trump's approval rate and government shut down by using multivariable regression and econometric analysis.

We collected data from Nov 1, 2018 to Feb 25, 2019, so that three period of time are well represented, namely before, during and after the shutdown. The data sets used in this project are Trump's approval rate, government shutdown status whis is modeled with three variables, and SP 500 Dow Jones index. The software we used to perform regression and analysis is STATA 14.1.

The key conclusion we reached is that there exists a negative correlation between Trump's approval rate and government shutdown.

2 Data Description

Two separate types of data were acquired for this study:

(1) A daily calculation of the president's approval rating from November 1st 2018 to February 25th 2019;

This approval-rate data was taken from FiveThirtyEight. We used data from November 1st 2018 to February 25th 2019 (117 days in total, including 77 working days and 40 non-working days), because the government shutdown took place from December 22nd 2018 to January 25th 2019 (35 days in total, including 22 working days and 13 non-working days). By comparing the president's approval rate one month before the shutdown, during shutdown, and one month after the shutdown, we were able to gain a comprehensive understanding of how government shutdown affects the president's popularity.

In order to better track the trivial change in the approval rate due to government shutdown, we also made some adjustments to the raw data using Excel:

(a) In order to compare the approval rate with and without shutdown, we add a binary variable called "shutdown" to each date: we assign "1" to the date from December 22nd 2018 to January 25th 2019 (inclusive), and "0" otherwise.

(b) In order to study the trend of approval rate during the shutdown, we add another variable called "shutdown_time" to each date: we assign a number x ($x \in [1, 35]$) to each date from December 22nd 2018 to January 25th 2019 respectively, representing the xth day of shutdown, and "0" otherwise.

(c) In order to study the trend of approval rate after the shutdown, we add a third variable called "days_after_sd" to each date: we assign a number y ($y \in [1, 31]$) to each date from January 26th 2019 to February 25th 2019 respectively, representing the yth day after shutdown, and "0" otherwise.

(2) The daily S&P 500 index of each trading day from November 1st 2018 to February 25th 2019.

This S&P data was taken from FRED Economic data. The S&P 500 index includes 500 leading companies in leading industries of the U.S. economy. Apart from the political factors such as government shutdown, we also want to investigate the influence of non-political factors, e.g. financial indicators including stock market indexes, on the president's approval rate, in order to avoid omitted variable bias.

Using Excel, we made some adjustments to this data as well. In order to precisely match the stock market index with the condition of government shutdown on a daily basis, we assign the S&P index of day x to the S&P index of day $(x-1)$, since the approval rate of day x is determined by the stock market condition of day $(x-1)$.
The SP data set we used have a mean of 2657.26 and a standard deviation of 101.00.

3 Econometric Analysis

Since we want to take SP 500 variable into account, and SP 500 index is only calculated in each trading days, we ignore all observations on non-trading days in all models in order to make comparisons possible among models. We perform four linear regression models.

3.1 Model 1

For the first model, we perform a simple linear regression, where the only independent variable is the binary variable "shutdown", which equals 1 if the government is at the shutdown stage and 0 otherwise, and the dependent variable is the approval rate.

Dependent Variable	**Independent Variable**
Approval Rate	Shutdown

By running the regression function in Stata, we get the following result:

	Coefficient	**Standard Error**	**t**	**P> |t|**
Shutdown	-.8947138	.2249768	-3.98	0.000
Intercept	41.73254	.1450278	287.76	0.000
F(1, 75) =15.82, Prob > F= 0.0002, R-squared=0.1421, Root MSE=1.0061				

(1) When the government operates as usual, the estimated approval rate of Donald Trump is 41.73254%; When shutting down, the estimated approval rate falls by 0.8947138%.

(2) The predictor variable, "shutdown", is highly significant because its p-value is 0.000, which means that we reject the hypothesis that beta1(coefficient correspond to shutdown) equals zero. Thus, we should keep the variable shutdown in our model.

(3) The R-squared value is only 0.1421, meaning that this model only explains a small proportion of variability of the response data around its mean, and we should seek more independent variables to fit the data.

3.2 Model 2

For the second model, we keep the dependent variable the same, and add two independent variables: "shutdown_time" and "days_after_sd". We add this because we suspect that as the time of shutdown becomes longer, approval rate would continue to decrease, and when the government reopens, the approval rate would gradually rise.

Dependent Variable	**Independent Variable**
Approval Rate	Shutdown
	Shutdown Time
	Days after Shutdown

By running the regression function in Stata, we get the following result:

	Coefficient	Standard Error	t	P> |t|
Shutdown	.3799431	.1801001	2.11	0.038
Shutdown Time	-.0759958	.0053607	-14.18	0.000
Days After Shutdown	-.0285604	.0168361	-1.70	0.094
Intercept	41.89144	.1400164	299.19	0.000
F(3, 73) =82.41, Prob > F= 0.0000, R-squared=0.3243, Root MSE=.90505				

(1) Before the shutdown, the estimated approval rate of Donald Trump is 41.89144%. When shutting down(keeping all other independent variables unchanged), the estimated approval rate rises by 0.3799431%, which is, unexpectedly, a positive number. The reason for this is that the binary shutdown variable is related to the shutdown time and days after shutdown. When the shutdown time(day) rises by 1, the estimated approval rate falls by 0.0759958. When the days after shutdown rises by 1, the estimated approval rate falls by 0.0285604. The reason for this abnormality is the existence of the shutdown variable.

(2) Two predictor variables, shutdown time and days after shutdown, are highly significant because their p-values are 0.000 and 0.040, which means that we reject the hypothesis that beta1(coefficient correspond to shutdown) equals zero. However, the p-value of the binary variable shutdown is 0.389, implying that we should remove it from the regression when the other two variables exists.

(3) The R-squared value is only 0.3243, which is an improvement comparing to the first model. In other words, this model explains more variability of the response data around its mean.

3.3 Model 3

Keeping all variables the same as in model 2, we add another independent variable, SP 500 index, into our regression. SP 500, as an index that measures how the stock of the 500 companies with the top market capitalization is performing, is a relatively representative quantitative measurement of the financial condition. We add it because we expect a positive correlation between Trump's approval rate and financial conditions.

Dependent Variable	Independent Variable
Approval Rate	Shutdown
	Shutdown Time
	Days after Shutdown
	S&P 500

By running the regression function in Stata, we get the following result:

	Coefficient	Standard Error	t	P> |t|
Shutdown	.7305909	.4356779	1.68	0.098
Shutdown Time	-.085781	.0106193	-8.08	0.000
Days After Shutdown	-.0328966	.0165505	-1.99	0.051
S&P 500 Adjusted	.0012874	.0011711	1.10	0.275
Intercept	41.89144	.1400164	299.19	0.000
F(4, 72) =69.45, Prob > F= 0.0000, R-squared=0.3291, Root MSE=.90806				

(1) Before the shutdown, the estimated approval rate of Donald Trump is 38.44029%. When shutting down(keeping all other independent variables unchanged), the estimated approval rate rises by 0.7305909%. When the shutdown time(day) rises by 1, the estimated approval rate falls by 0.085781. When the days after shutdown rises by 1, the estimated approval rate falls by 0.0328966. When the SP 500 index rises by 1, the approval rate is estimated to rise by 0.0012874

(2) Only the shutdown time variable is considered significant under the 95% significance level as its p-values is 0.000. However, the p-value of the other three independent variables are all above the threshold, implying that they are not significant enough in this model.

(3) The R-squared value is only 0.3291, which makes an small improvement comparing to the second model. In other words, this model explains slightly more variability of the response data around its mean.

3.4 Model 4

Based on the feedback of previous models, we notice that S&P 500 Adjusted variable is not significant enough when the other variables are presented. When "Shutdown Time" and "Days After Shutdown" variables exists, variable "Shutdown" is also predetermined, as it will be 1 when "Shutdown Time" is non-zero, and will be 0 when "Days After Shutdown" is non-zero or both variables are zero. So we want to form a model of only two independent variables, which are "Shutdown Time" and "Days After Shutdown."

Dependent Variable	Independent Variable
Approval Rate	Shutdown Time
	Days after Shutdown

By running the regression function in Stata, we get the following result:

	Coefficient	Standard Error	t	P> |t|
Shutdown Time	-.061927	.0061867	-10.0108	0.000
Days After Shutdown	-.030617	.0166449	-1.84	0.070
Intercept	41.93235	.1234718	399.61	0.000
F(2, 74) =50.15, Prob > F= 0.0000, R-squared=0.3174, Root MSE=.90353				

(1) Before the shutdown, the estimated approval rate of Donald Trump is 41.93235%. When the shutdown time(day) rises by 1, the estimated approval rate falls by 0.061927. When the days after shutdown rises by 1, the estimated approval rate falls by 0.030617.

(2) Only the shutdown time variable is considered significant under the 95% significance level as its p-values is 0.000. However, the p-value of "days after shutdown" variable is 0.070, implying that it is slightly less significant.

(3) The R-squared value is only 0.3174, which means that this model explains a small proportion of variability of the response data around its mean.

3.5 Conclusions

According to the regression result, all four models are not perfect. For model 1, although the variable is highly significant, R-squared is much lower than other models. For model 2, although R-squared is more satisfying, it contains a less significant variable, and so does model 3 and 4. Also, the R-squared value for all four models haven't exceed 0.4, which suggests that there might be other underlying variables beyond the shutdown and stock conditions that are also related to the approval rate.

In order to see how the shutdown have influenced Trump's approval rate, regardless of all other factors, we need to look at model 1, which shows that there is a negative correlation between shutdown(change from 0 to 1 when shutdown) and Trump's approval rate. Looking at how the approval rate changes as time passes, we need to look at model 4, which shows a negative correlation between shutdown time and the approval rate, and a non-significant negative correlation between days after shutdown and the approval rate. While S&P 500, according to model 3, has a non-significant positive correlation with the approval rate.

We also want to choose the best model. Our answer will variance based on different needs. If we want to maximize the variability explained by the model, model 3 is the best choice, as it takes the most independent variables into account. But the returns of adding a new variable might not be satisfying. For example, as we move from model 2 to model 3, R-squared only increases by a tiny amount. Another approach is to maximize adjusted R, which puts a penalty on every variable added, so it will only increase when the new term improves the model more than expected. The four adjusted R-squared values are 0.1307, 0.2966, 0.2919, and

0.2989. So we will choose model 4 under this criterion.

4 Discussion of the Key Assumptions

First of all, we need to check upon the four least square assumption in order for our regression to be solid.

First assumption says that the conditional distribution of u_i given X_{1i}, X_{2i},...,X_{ki} has a mean of zero. We have calculated mean of the residual term, the residual is -0.033, which is small enough for us to assume that our coefficients are unbiased. The reason why it is not exactly 0 is probably due to systematic errors, but overall the assumption can be confirmed to be satisfied.

Second assumption says that $(X_{1i}, X_{2i}, ..., X_{ki}, Y_i)$, i = 1, 2, ... , n, are independently and identically distributed. Each set of data we used, i.e. Trump's approval rate on a particular day and the matching data of shutdown time, SP 500 etc. are independent of another set of data describing another day. The data set are identically distributed by the way the data are collected at the original source.

Third assumption says that large outliers are unlikely, in other words, $X_{1i}, X_{2i}, ..., X_{ki}$, and Y_i have finite fourth moment. We checked all our data and confirmed that all data have finite fourth moment. Shutdown status variable is binary with 0 and 1 only. Shutdown time variable falls in the range of 1-35. SP 500 data are all within 3 standard deviation above and below mean.

Fourth assumption says that there is no perfect multicollinearity. This assumption is satisfied as none of our regressors is a perfect linear function of the other regressors. It is also proved by STATA's regression report because there is no variable assigned with coefficient 0 and a note of 'omitted'.

Our regression might be subject to omitted variable bias. There may be a variable that is both correlated with one of our independent variables and also determining part of the dependent variable. For instance, as government shutdown prolongs, US/Euro exchange also seem to fall because of weaker confidence in US currency, which in turns contribute to decreasing approval rate of Trump. As such, US/Euro exchange rate might be an omitted variable.

https://github.com/williamyangcl/Econometrics

5 Reference

Stock, J.H. , Watson, M.W. "Introduction to Econometrics" third edition, 2002