THE EFFECT OF ECONOMIC INDICATORS ON US PRESIDENTIAL APPROVAL RATINGS

BERJER DING

1. INTRODUCTION

In the 1992 presidential election, advisor to candidate Bill Clinton, James Carville, coined the phrase "It's the economy, stupid" as a retort against the incumbent's (George H.W. Bush) prevailing economic recession. The understanding was that the economy was such a universal force for the electorate that its handling is the fundamental measure that determined the presidency. (Goddard, 2017) Carville's retort was effective in his era as news media relied on economic statistics as a popular metric to evaluate federal action. However in the present mass media age the dynamic and actors are different. It is vital to establish whether this sentiment is still held in voters as its determination is essential to political action and electoralism, e.g. the coming 2024 presidential election. We therefore ask if there exists a relationship between the key statistics of the United States economy and positive presidential approval polls between 1999 and 2024.

The presidential job approval is a measure of the U.S. citizenry's opinion on the actions of the incumbent U.S. president. The economic indicators we look at are real GDP growth (RGDP), real consumer expenditure, real wage growth, change in business and manufacturing labour productivity, consumer inflation rate (the inflation rate), producer inflation rate, unemployment rate, and change in import and export price index. These set of indicators provide a high-level overview of the state of the U.S. economy. (World Bank Group, 2024).

Choi et al. (2016) observed a nonlinear relationship between approval ratings and unemployment, inflation, GDP growth, and consumer sentiment to test the contemporary "Threshold Hypothesis": that unemployment is the most significant variable when it is high. Berlemann and Enkelmann (2014) conducted an aggregate survey of literature on this very topic to find some consensus of opinions and an analysis of methodologies. They found that conclusions are divided with "about half of all studies for the United States find a significant effect of unemployment and inflation on presidential popularity, the others do not." (Berlemann & Enkelmann, 2016) With respect to historical analysis we also seek an explanatory model, but we broaden our scope to include other macroeconomic variables beyond the classic inflation, unemployment, and RGDP. In so doing we seek to establish a more complete understanding of the U.S. economy and therefore a more insightful analysis on its relation to presidential approval.

2. METHODS

We use the classic Gallup presidential approval poll as the dataset for approval ratings. The poll has been gathered by Gallup using periodic multiday interviews with a weight based on demography. Every series is conducted with at least 1,000 adults and asks the same question: "Do you approve or disapprove of the way [President Name] has been handling his job as president?" (Gallup, 2024) For economic variables we source all of our data from the United States Bureau of Labour Statistics

(BLS). All data is publicly available and is synthesized from government statistics, e.g. from government agencies and economists, and on-the-ground correspondence. (U.S. Bureau of Labour Statistics, 2024) All data we use are those collected monthly or quarterly.

We analyze our dataset using R. We first normalize our dataset using min-max scaling. We compute the VIFs between the predictors to find any multicollinearities. After evaluating the VIF we utilize a $60/40$ split to obtain a training and a testing dataset. (We set a sampling seed for reproducibility) We now perform a naïve multiple linear regression on the training dataset and obtain its diagnostic plots. After analysis of the diagnostic plots, i.e. determining normality of noise, trends in residuals, leverages and outliers, we consider a transformation on the data. We use two methods. First we try transforming the predictors to multinormality and then using an inverse response plot to transform the response. Second we try a multivariate Box-Cox transformation on the predictors and the response. We now perform a multiple linear regression on the two datasets and obtain their diagnostic plots. We evaluate the diagnostic plots, in particular utilizing the residuals-leverage plot to find outliers to remove from our dataset. After renormalizing we graph the added-variables plot of the better model. Our analysis after is twofold. We reduce redundancies in our model by performing using best-subset computation to determine the best models for a given criterion and evaluate it. Second, using prior information on the regressions we perform an F-test on the predictors to determine and verify statistically-insignificant predictors, and then use a partial F-test to allow us to conclude if the partial model without those predictors is of statistical worth. We now consider all criterions and again consider all diagnostics. We lastly evaluate our models on the test dataset, conclude which is most robust and obtain our final model, and consider any discrepancies.

3. Results

We split our dataset into 60% ($n = 177$) training $\mathbf{X}_0$ and 40% ($n = 119$) test, with a seed $s = 100$. For brevity our variables are denoted as follows: Y is presidential approval, X_1 is real GDP growth, X_2 is real consumer expenditure, X_3 is real wage growth, X_4 is change in business productivity, X_5 is change in manufacturing productivity, X_6 is consumer inflation rate, X_7 is producer inflation rate, X_8 is unemployment rate, X_9 is change in import price index, and X_{10} is change in export price index.

We compute the VIF of our variables to determine if there is any multicollinearity,

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
VIF	1.140	1.461	1.472	1.682	1.393	1.691	5.735	1.694	3.418	3.493

TABLE 3.1. VIFs of the predictors of $\mathbf{X_0}$

We see that only X_7 has a VIF above the threshold of five. All other variables however have a VIF markedly below this threshold. Therefore we will not face issues of multicollinearity. We confirm this conjecture with a naïve multiple linear regression $\mathcal{M}_0$ on the training dataset $\mathbf{X}_0$ with the unaltered formula $Y = \sum_{i=1}^{10} X_i$. This yields TABLE 3.2 and FIGURE 3.1.

We begin with the positives: analysis of the normal Q-Q plot gives evidence that our standardized residuals do exhibit behaviors of normality; there is little skew in both directions. Moreover our residual standard error is quite low and we conclude that with a p-value of 7×10^{-12} that our model is statistically significant at the $\alpha = 0.05$ level.

$\mathcal{M}_0$	*Estimate*	*Std. Error*	$\mathbb{P}(> \lvert \boldsymbol{t} \rvert)$
β_0	0.253	0.183	0.168
β_1	0.312	0.160	0.052
β_2	−0.044	0.242	0.857
β_3	−0.016	0.103	0.554
β_4	0.122	0.118	0.301
β_5	0.513	0.068	0.000
β_6	−0.111	0.077	0.153
β_7	0.018	0.211	0.932
β_8	−0.101	0.083	0.228
β_9	−0.312	0.177	0.079
β_{10}	0.098	0.177	0.581

Res. Std. Error	0.149
df	167
R^2_{adj}	0.314
$F_{10,167}$-*Statistic*	9.082
p-value	7×10^{-12}

TABLE 3.2. Parameters of $\mathcal{M}_0$

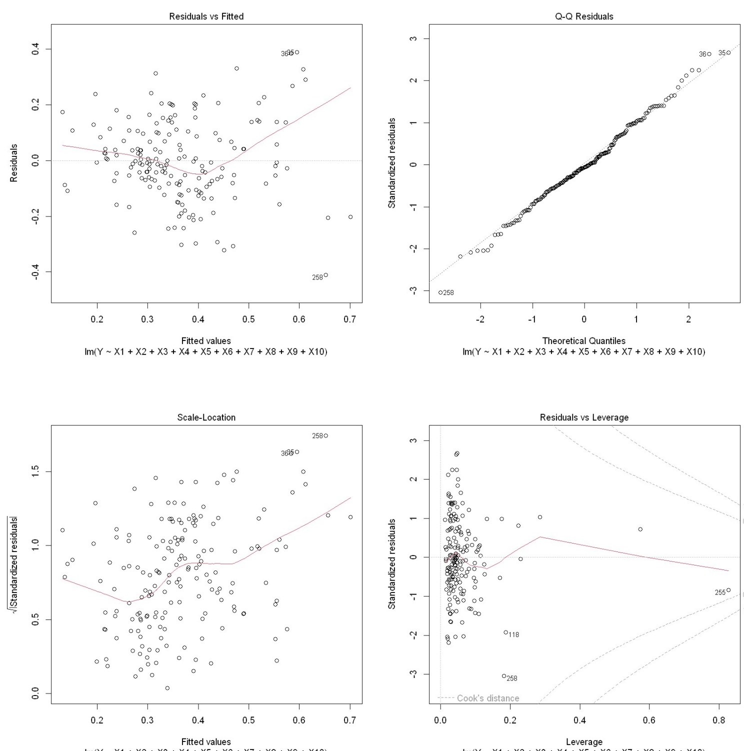

FIGURE 3.1. Diagnostic matrix of $\mathcal{M}_0$

Note that our leverage cutoff is $2(11)/177 = 0.124$. We now proceed with the negatives. First we observe that only one of our variables are significant at the α level, namely X_5 corresponding to change in manufacturing productivity. Second there is significant standard error amongst almost all regressors with values in $[0.1, 0.2]$, which under normalization is 10-20% of the range. Thirdly the R^2_{adj} of our dataset is markedly low, indicating a weak fit and high variability in the data respective to the regression plane. From the regression-fitted and scale-location plots we cannot conclude that there is constant variance in the noise in the naïve regression. Namely there is a strong upward trend in both graphs as the fitted value increases, indicating the necessity for a log-like transformations. Lastly we observe a substantial number of outliers that must be considered.

We perform an inverse response plot $\mathcal{M}_1$ and multivariate Box-Cox $\mathcal{M}_2$ on the data. From the diagnostic plots of the resultant models we observe that they appear more Gaussian, affirming the necessity for a transformation. Secondly their Q-Q normal plot and leverages plots, same as that of the diagnostic plots of $\mathcal{M}_0$, directly point to the existence of extremely influential outliers in the dataset, with multiple points with $D_i \geq 0.5$.

In inspecting the outliers we determine two prominent domains with which they arise. $\mathbf{y}_{255}$ and surrounding outliers correspond to May 2020 and the COVID-19 pandemic in the U.S. During this period almost all economic activity was halted by federal epidemiological action which resulted in record-breaking economic figures (e.g. 15.7% unemployment in May 2020). (Parker, Minkin, & Bennett, 2020) In assessing our research question these circumstances are extraordinary and contrary to standard analysis. We therefore remove this period between the start of the pandemic, January 2020, and its end, April 2021, corresponding to $\mathbf{y}_{251}$ to $\mathbf{y}_{266}$. The second domain of note is $\mathbf{y}_{32}$ and surrounding outliers. These datapoints correspond to September 2001 and the subsequent U.S. War on Terror. Following the 9/11 terrorist attack president George W. Bush saw incredibly high approval rating as the U.S. rallied after the tragedy, resulting in the highest recorded approval rating of 90% in the immediate aftermath. (Moore, 2001) We remove these dates, September 2001 to December 2002, $\mathbf{y}_{32}$ to $\mathbf{y}_{47}$ as they correspond to significant sociopolitical events that are significantly biased from the political norm. We renormalize to yield our modified training dataset ($n = 158$) $\mathbf{X}_2$ and test dataset ($n = 111$) $\mathbf{T}$.

We perform our analysis again. Namely we redetermine the VIFs and a naïve regression $\mathcal{M}_3$.

	$\boldsymbol{X_1}$	$\boldsymbol{X_2}$	$\boldsymbol{X_3}$	$\boldsymbol{X_4}$	$\boldsymbol{X_5}$	$\boldsymbol{X_6}$	$\boldsymbol{X_7}$	$\boldsymbol{X_8}$	$\boldsymbol{X_9}$	$\boldsymbol{X_{10}}$
VIF	1.250	1.147	1.500	1.713	1.858	1.700	5.280	1.710	3.858	3.535

$\mathcal{M}_3$	*Estimate*	*Std. Error*	$\mathbb{P}(>\lvert\boldsymbol{t}\rvert)$
β_0	−0.039	0.151	0.799
β_1	0.398	0.109	0.000
β_2	0.099	0.103	0.338
β_3	0.064	0.095	0.497
β_4	0.207	0.100	0.039
β_5	0.140	0.106	0.189
β_6	−0.094	0.090	0.292
β_7	0.067	0.253	0.790
β_8	0.091	0.064	0.160
β_9	−0.105	0.219	0.632
β_{10}	−0.050	0.244	0.838

Res. Std. Error	0.177
df	148
R^2_{adj}	0.228
$F_{10,167}$-*Statistic*	5.314
p-value	8×10^{-7}

TABLE 3.3. VIF and parameters of $\mathcal{M}_3$

We again observe no issues with multicollinearity. With the removal of aberrant points our diagnostic plots imply a much stronger adherence to our assumptions. Notably both the residuals-fitted and scale-location plots are now constant in slope with stochastic spread of data. The normal Q-Q plot has small skew and there are no outlier leverages in the leverage plot any longer. However we still face low statistical significance and fit in the parameters of $\mathcal{M}_3$. We thus seek transformations.

The table for our lambda values are given below,

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
Est.	1.651	1.011	0.968	0.782	0.768	0.176	1.140	0.234	1.409	1.333

TABLE 3.4. Estimated λ for power transformation of predictors of $\mathbf{X}_2$

We use the rounded to nearest 0.5 transformations to obtain the transformed dataset $\mathbf{X}_3$. We again proceed with the two approaches given in Sheather (2009). We first use an inverse response plot to estimate the transformation to be $\lambda_Y \approx 1$. We second use a multivariate Box-Cox method to estimate $\lambda_Y \approx 1$. In the case $\lambda_Y = 1$ the inverse response plot and multivariate Box-Cox agree so we have confidence and obtain one transformed model $\mathcal{M}_4$. The diagnostic matrix is given in FIGURE 6.1.

After transformation we still face low fit, but we do have a clearer understanding of which variables are statistically significant. In particular based on the parameters we have extremely insignificant variables in X_6 and X_{10}. The added-variable plot FIGURE 6.2 of $\mathcal{M}_4$ reinforces our conjecture.

We seek to determine the redundant variables and obtain the most precise model. To do so we utilize variable-selection and F-tests. A best-subset computation finds that predictors $\mathcal{M}_5 = \{X_1, X_4, X_5, X_8\}$ yields simultaneously the highest R^2_{adj}, AIC, and BIC. The parameters for $\mathcal{M}_5$ are given in TABLE 3.6. In particular we find our R^2_{adj} increased with substantially lower standard deviation in all variables, signalling a more robust fit. Meanwhile an F-test on $\mathcal{M}_4$ finds TABLE 3.5.

	$\mathbb{P}(>\lvert t\rvert)$
X_1	5.529×10^{-9}
X_2	1.617×10^{-1}
X_3	5.623×10^{-2}
X_4	1.545×10^{-4}
X_5	3.607×10^{-2}
X_6	3.511×10^{-2}
X_7	9.149×10^{-1}
X_8	6.772×10^{-2}
X_9	2.347×10^{-1}
X_{10}	9.996×10^{-1}

TABLE 3.5. $\mathcal{M}_4$ F-test

In particular we see that $X_2, X_3, X_7, X_8, X_9, X_{10}$ are all not significant at the α level. Consequently we perform a partial F-test on the nested model $\mathcal{M}_6$ with formula $Y \sim X_1 + X_4 + X_5 + X_6$. The resultant p-value is $0.488 > 0.05$ and thus we fail to reject the nested model. In particular all its variables are statistically significant at level α.

Lastly to determine which model is superior between $\mathcal{M}_5$ and $\mathcal{M}_6$ we apply our models on the test dataset $\mathbf{T}$. The aggregated results are given in TABLE 3.7. We see in that $\mathcal{M}_5$ is only minorly superior to $\mathcal{M}_6$ based on the marginally smaller residual standard error and lower p-value, though realistically this difference is completely negligible. Hence $\mathcal{M}_5 = \mathcal{M}_{\text{F}}$

However on a higher level we must recognize that overall $R^2_{\text{adj}} = 0.318$ on $\mathbf{T}$ and $R^2_{\text{adj}} = 0.243$ on $\mathbf{X}_3$ is quite low even after all of our optimization. Still our final model is not entirely worthless. First the p-value is of statistical significance in all models. An inspection on the diagnostic plots of $\mathcal{M}_{\text{F}}$ on $\mathbf{T}$ in FIGURE 6.3 displays a familiar adherence to normality in the normal Q-Q plot. Moreover we observe constant variance of noise in the residuals-fitted and scale-location plots, which may only better after considering the outliers in the leverage plot. Therefore the model does transform data to the correct assumptions for regression. Lastly we must emphasize that the model does generalize nicely to the test dataset with a within-an-order-of-magnitude R^2_{adj} and like std. error. Thus, though the model has a weak fit, its robustness to new data exhibits a strikingly low variance and not a significantly high bias.

$\mathcal{M}_{\mathrm{F}}$	*Estimate*	*Std. Error*	$\mathbb{P}(>\lvert \boldsymbol{t} \rvert)$
β_0	-0.497	0.063	0.000
β_1	0.554	0.136	0.000
β_4	0.235	0.090	0.010
β_5	0.165	0.097	0.090
β_8	0.033	0.012	0.001

Res. Std. Error	0.174
df	154
R^2_{adj}	0.243
$F_{10,148}$-*Statistic*	13.64
p-value	2×10^{-9}

TABLE 3.6. Parameters of $\mathcal{M}_{\mathrm{F}}$

	Res. Std. Error	R^2_{adj}	$F_{4,102}$-*statistic*	*p-value*
$\mathcal{M}_{\mathrm{F}}$	0.1639	0.318	13.37	0.8×10^{-8}
$\mathcal{M}_6$	0.1644	0.315	13.17	1.1×10^{-8}

TABLE 3.7. Model $\mathcal{M}_5 = \mathcal{M}_{\mathrm{F}}$ and $\mathcal{M}_6$ fit on test dataset $\mathbf{T}$

4. DISCUSSION

Our final model is,

$$\begin{aligned}\text{Approval R.} = -0.497 &+ 0.554(\Delta\text{RGDP})^{1.5} + 0.235(\Delta\text{Business Productivity})\\ &+ 0.165(\Delta\text{Manufacturing Productivity})\\ &+ 0.033\ln(\text{Unemployment Rate})\end{aligned}$$

We observe, that like the literature, RGDP Growth is a significant component of the relationship boasting the strongest trend in the normalized relationship. We also observe that unemployment rate is also a key indicator, but its contribution to the relationship is much weaker than its contemporaries being to smallest slope and to the logarithm. An insight we found is that unlike the literature, instead of the inflation rate it is change in business and manufacturing productivity that is a stronger predictor of approval rate. This extends the famous amorphism about the importance of productivity: "...in the long run, it's (productivity) almost everything", to the political sphere as well.

We interpret the problem of low R^2_{adj}. To frame this we consider the added-variables plot of FIGURE 6.9. In particular there are two consistent outliers in all plots – datapoints 86 and 77. An inspection into their dates reveal that they are December 2000 and January 2001 respectively, the date of the end of the controversial 2000s presidential election. Namely they are outliers in that even though the U.S. economy was booming due to the dotcom bubble, the populace was still significantly biased against the president due to sociopolitical turmoils. Such phenomena were especially significant in the points we removed too. Indeed our low R^2_{adj} supports the hypothesis that there are other significant forces, such as events of the popular sphere to presidential approval beyond economic variables. Still as we saw at the end of the **§RESULTS**, economic variables are still a stable and robust predictor.

Our results corroborate that of Choi et al and the negative of Berlemann and Enkelmann: there exists a statistically significant relationship, but a linear model of simple macroeconomic variables is not sufficient to completely capture it. Therefore, in the present age of mass media we must retort to Carville, "it's not just the economy, stupid".

5. References

Berlemann, M., & Enkelmann, S. (2016). The economic determinants of U.S. presidential approval: A survey. *European Journal of Political Economy*, 41-54.

Choi, S.-W., James, P., Li, Y., & Olson, E. (2016). Presidential approval and macroeconomic conditions: evidence from a nonlinear model. *Applied Economics*, 4558-4572.

Gallup. (2024). *How Does the Gallup Poll Social Series Work?* Retrieved from Gallup: https://www.gallup.com/175307/gallup-poll-social-series-methodology.aspx

Goddard, T. (2017, January). *It's the Economy Stupid*. Retrieved from Political Dictionary: https://politicaldictionary.com/words/its-the-economy-stupid/

Mariani, D. (2013, May 1). Predicting the US Presidential Approval and Applying the Model to Foreign Counties.

Moore, D. W. (2001). *Bush Job Approval Highest in Gallup History.* Gallup.

Parker, K., Minkin, R., & Bennett, J. (2020). *Economic Fallout From COVID-19 Continues To Hit Lower-Income Americans the Hardest.* Pew Research Center.

U.S. Bureau of Labour Statistics. (2024, January 29). *Occupational Requirements Survey: Data sources*. Retrieved from U.S. Bureau of Labor Statistics: https://www.bls.gov/opub/hom/ors/data.htm

World Bank Group. (2024). *Economy*. Retrieved from The World Bank: https://datatopics.worldbank.org/world-development-indicators/themes/economy.html#:~:text=Economic%20indicators%20include%20measures%20of,and%20the%20balance%20of%20payments).

6. APPENDIX

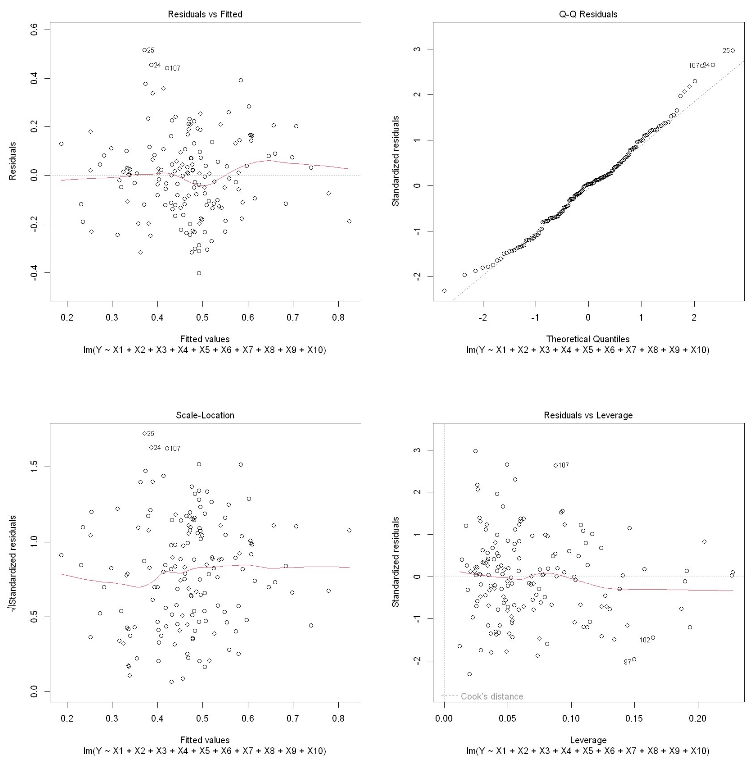

FIGURE 6.1. Diagnostic matrix of $\mathcal{M}_4$

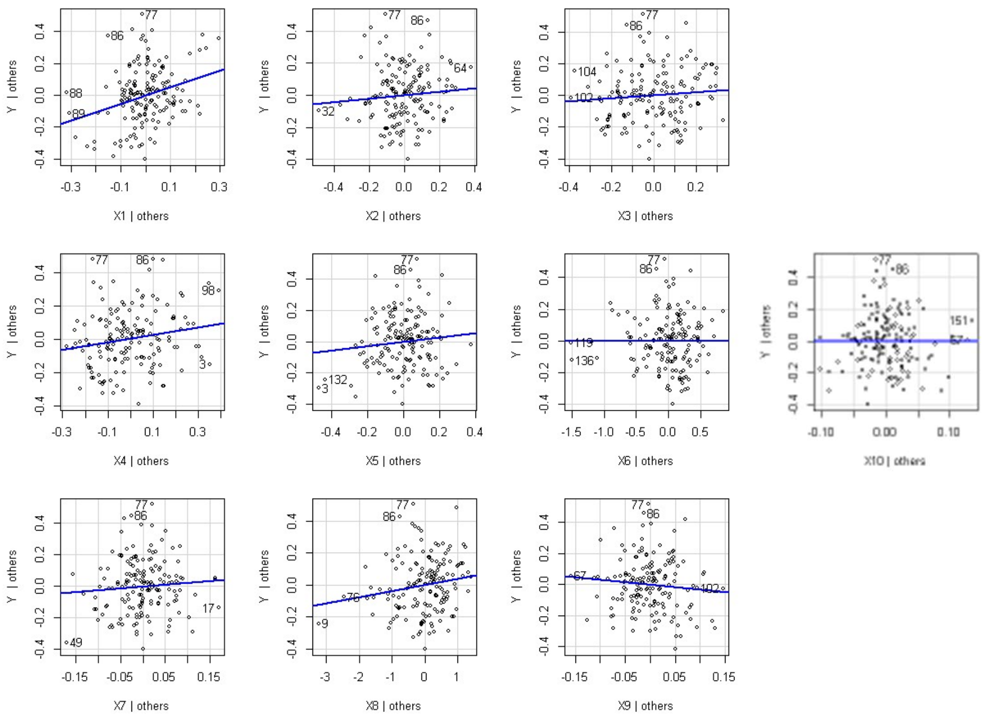

FIGURE 6.2. Added-variables plot of $\mathcal{M}_4$

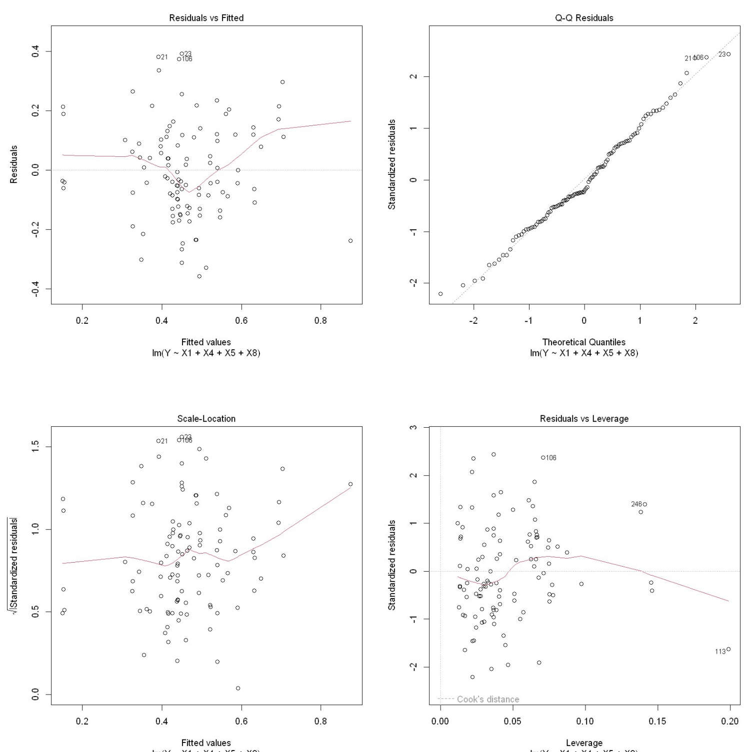

FIGURE 6.3. Diagnostic matrix of $\mathcal{M}_F$ on test dataset $\mathbf{T}$