# 36-402 Homework 2

*Eu Jing Chua*
*eujingc*

*January 29, 2019*

## Question 1

Table 1: Coefficients & Std. Error of linear model

| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | -0.03525 | 0.00665 | -5.30037 | 0.00000 |
| underval | 0.00476 | 0.00218 | 2.18614 | 0.02898 |
| log(gdp) | 0.00630 | 0.00079 | 7.96591 | 0.00000 |

Since the coeficient of $log(gdp)$ is positive, this model does not seem to support the idea of "catching-up" as countries with higher GDP have a higher economic growth rate. However, it does support the idea that under-valuing a currency boosts economic growth as the coefficient of *underval* is positive, indicating a positive *underval* index, which represents undervaluing, leads to higher economic growth.

## Question 2

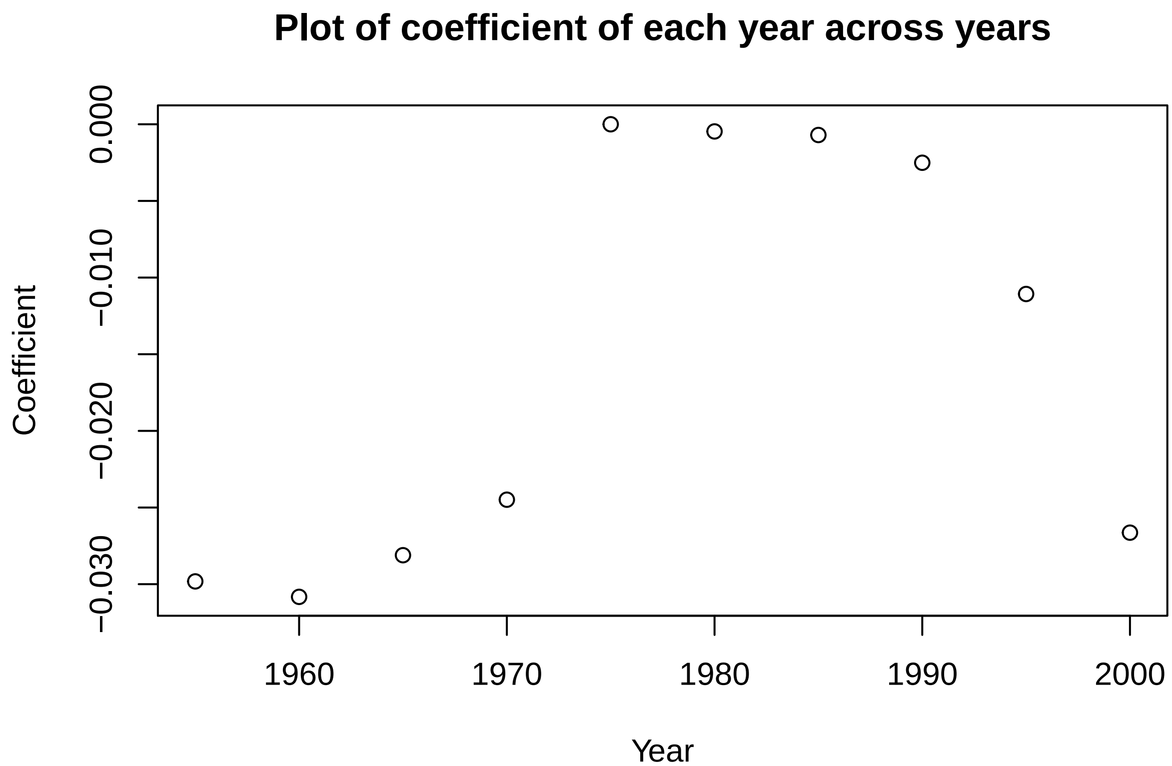
**Q2 a)**

Table 2: Coefficients & Std. Error of linear model

| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| underval | 0.01361 | 0.00290 | 4.69667 | 0 |
| log(gdp) | 0.02892 | 0.00317 | 9.13254 | 0 |

**Q2 b)** It is more appropriate to use `factor(year)` as there are only unique years that are 5 years apart. As such, modelling this way we will have a slope for each 5-year interval rather than a single slope for each increment of year.

**Q2 c)**




**Q2 d)**

Since the coeficient of $log(gdp)$ is positive, this model does not seem to support the idea of "catching-up" as countries with higher GDP have a higher economic growth rate. However, it does support the idea that under-valuing a currency boosts economic growth as the coefficient of *underval* is positive, indicating a positive *underval* index, which represents undervaluing, leads to higher economic growth.

## Question 3

**Q3 a)**

Table 3: $R^2$ values for each linear model

| | Model 1 | Model 2 |
|---|---|---|
| $R^2$ | 0.04855 | 0.42924 |
| Adj. $R^2$ | 0.04709 | 0.33214 |

**Q3 b)**

Table 4: $M\hat{S}E$ of linear models by LOOCV

| | x |
|---|---|
| Model 1 | 0.0010303 |

| | x |
|---|---|
| Model 2 | 0.0009528 |

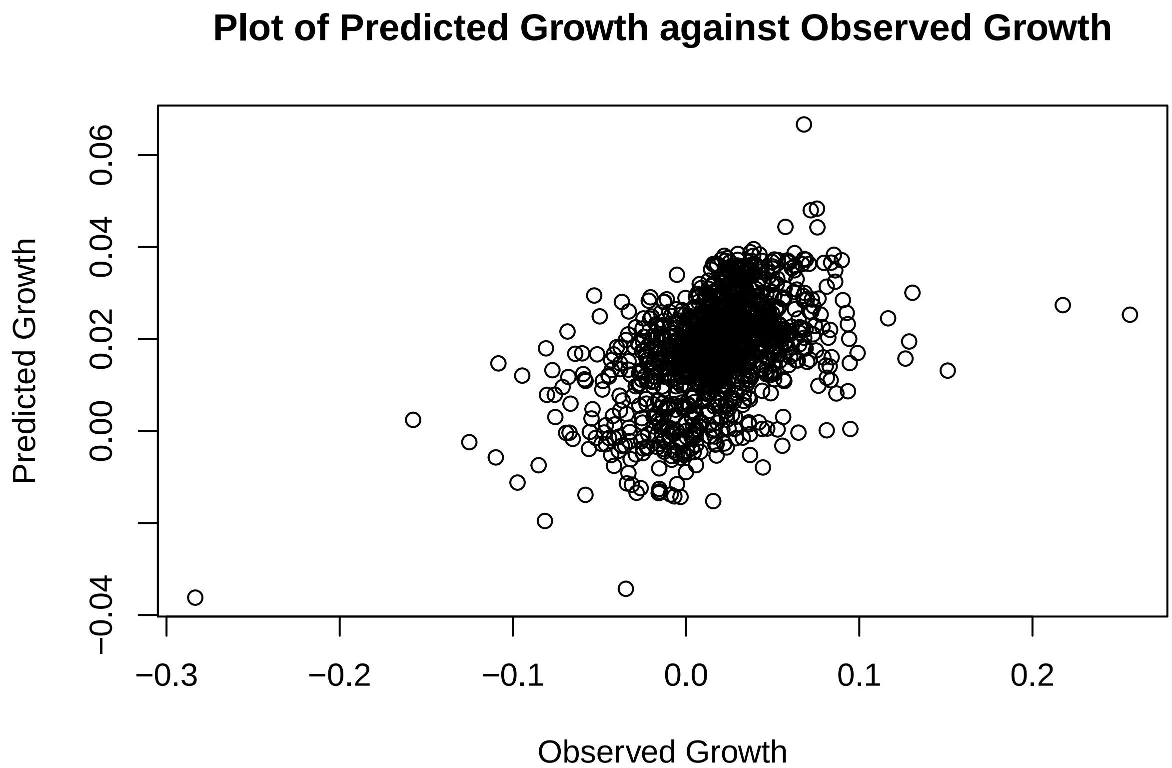
# Question 4

**Q4 a)**

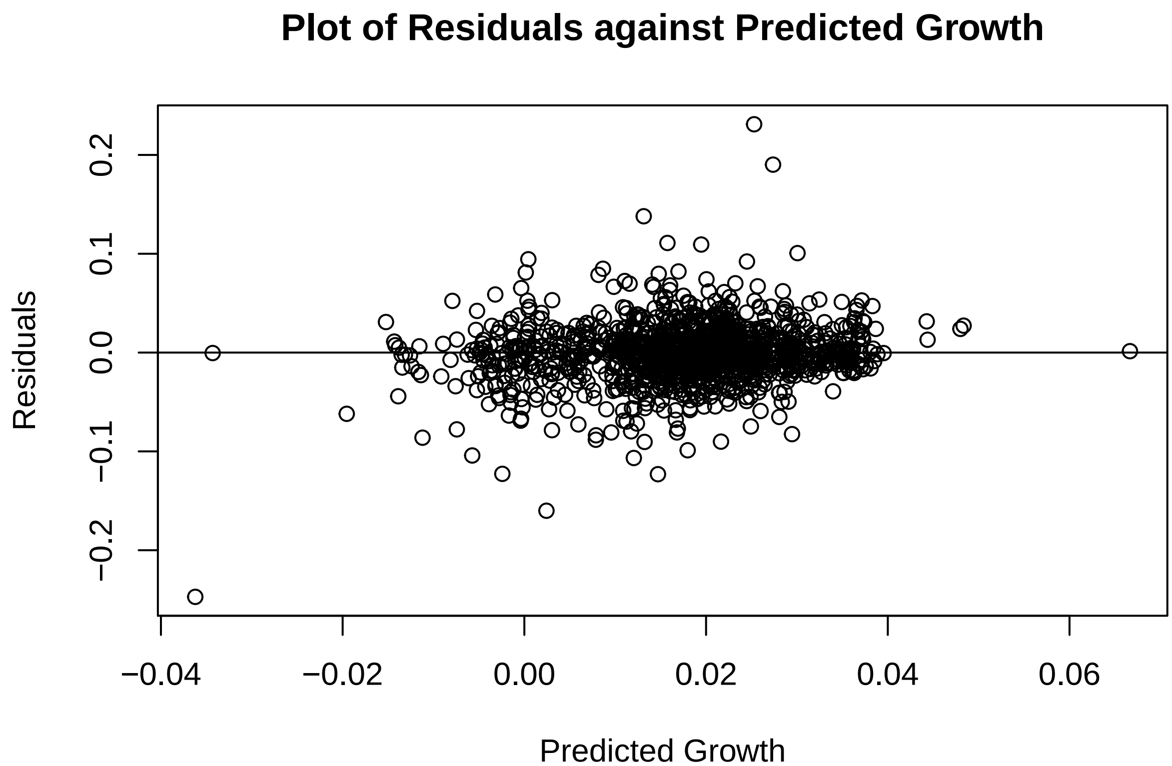
```
##
Multistart 1 of 3 |
Multistart 1 of 3 |
Multistart 1 of 3 |
Multistart 1 of 3 /
Multistart 1 of 3 -
Multistart 1 of 3 |
Multistart 1 of 3 |
Multistart 1 of 3 /
Multistart 1 of 3 -
Multistart 2 of 3 |
Multistart 2 of 3 |
Multistart 2 of 3 /
Multistart 2 of 3 -
Multistart 2 of 3 |
Multistart 2 of 3 |
Multistart 2 of 3 /
Multistart 2 of 3 -
Multistart 3 of 3 |
Multistart 3 of 3 |
Multistart 3 of 3 /
Multistart 3 of 3 -
Multistart 3 of 3 |
Multistart 3 of 3 |
Multistart 3 of 3 /
Multistart 3 of 3 -
```

There are no coefficients to report for kernel regression, as it is a non-parametric smoothing method that smooths the data with a kernel. The smoothing is only controlled by the choice of the kernel and the related bandwidth for the kernel, without any coefficients on the predictors.

**Q4 b)**


Plot of Predicted Growth against Observed Growth
Predicted Growth
0.06
0.04
0.02
0.00
−0.04
−0.3
−0.2
−0.1
0.0
0.1
0.2
Observed Growth

**Q4 c)**




The points should be scattered around a flat line at 0 if the model was right, as we assume $\mathbb{E}[\epsilon] = 0$. In this case, it would seem that the residuals indeed are roughly scattered around a flat line at 0.

**Q4 d)**

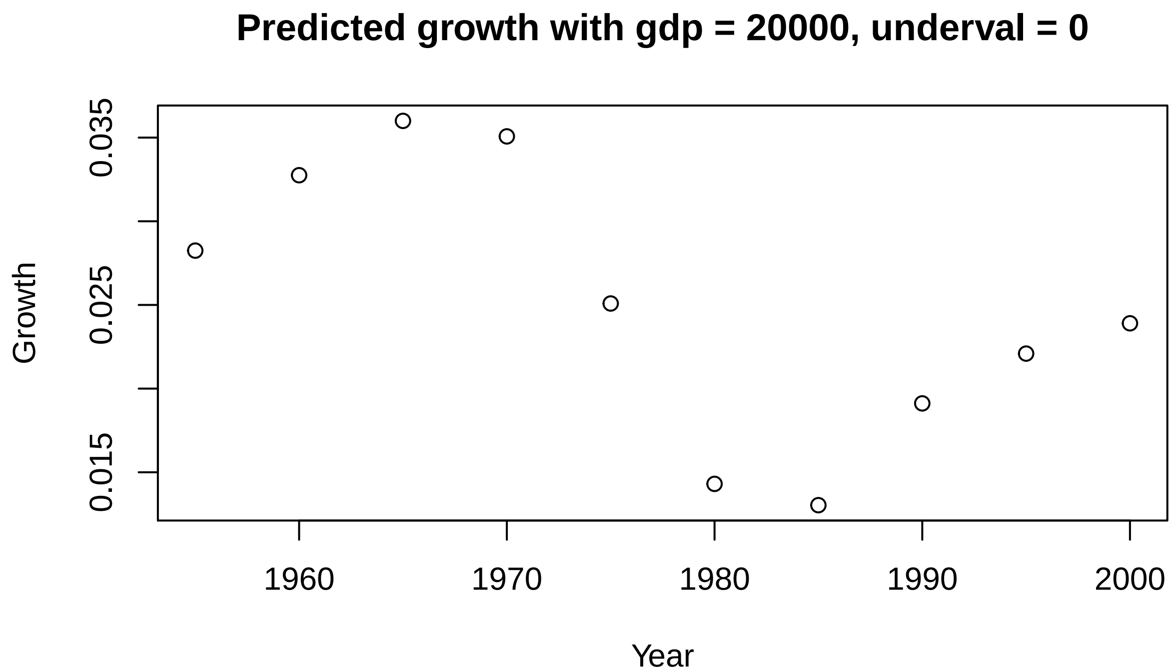
Table 5: $\hat{MSE}$ of the regressions via cross validation

| | x |
|---|---|
| Linear Regression | 0.0009688 |
| Kernel Regression | 0.0009481 |

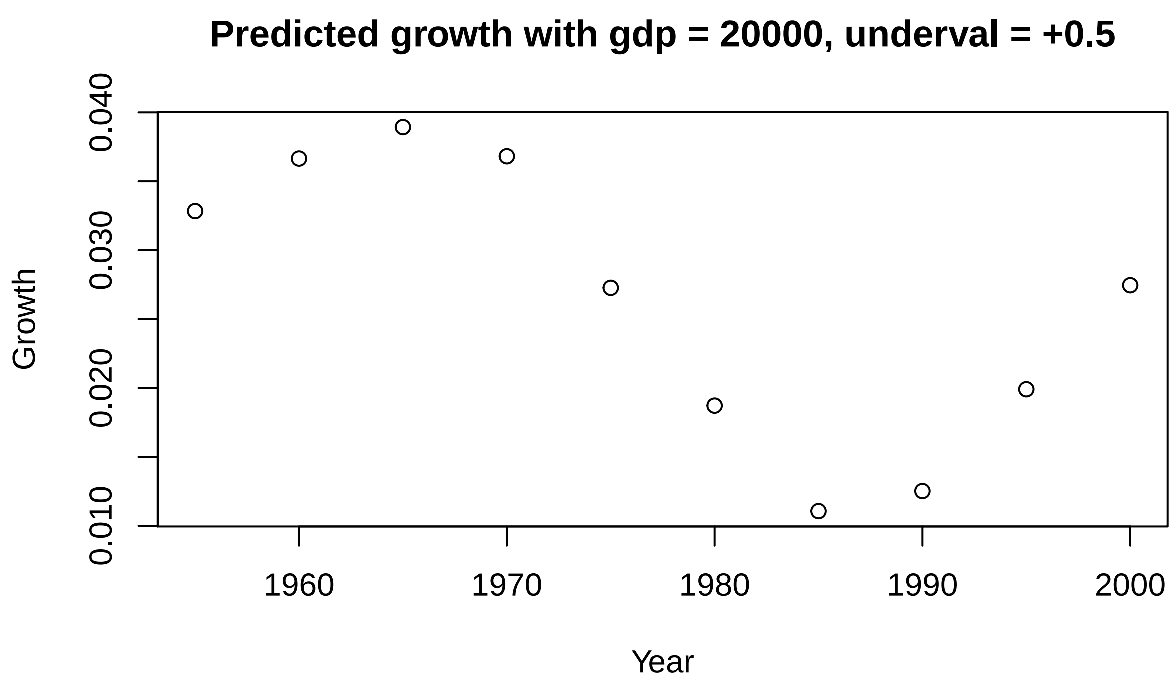
As seen from above, kernel regression has a lower estimated MSE than a linear model with the same covariates. Hence, the kernel regression is better in generalizing and predicting better than the linear model.

## Question 5

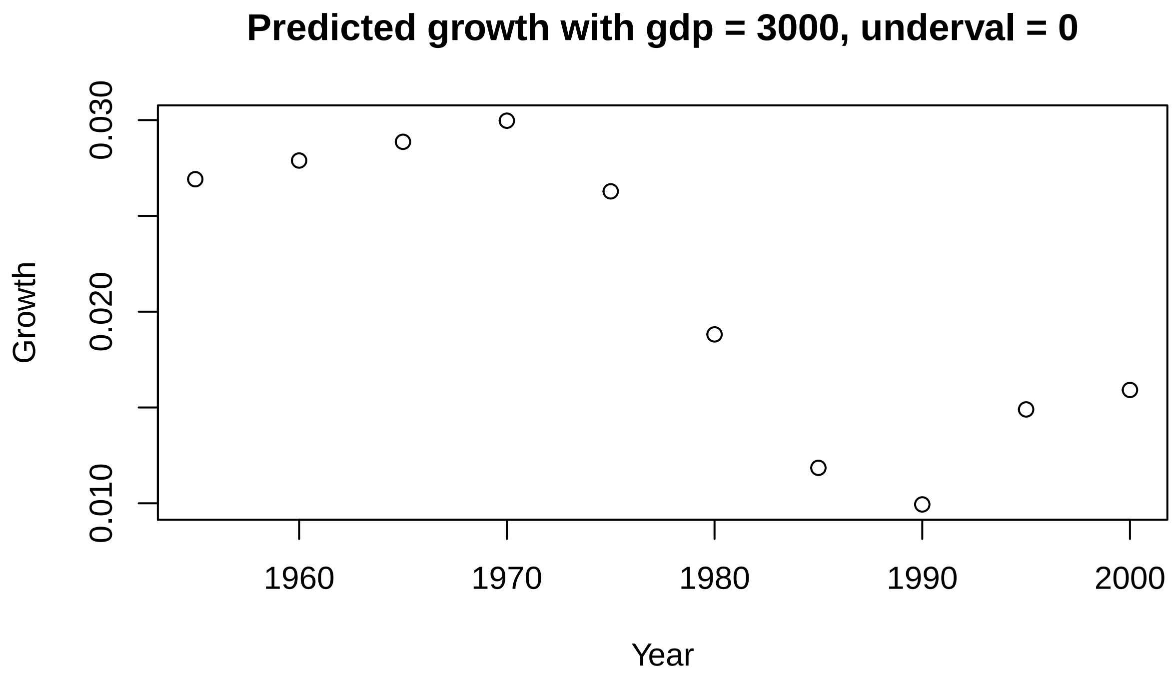
**Q5 a)**


Predicted growth with gdp = 20000, underval = 0
Growth
0.015
0.025
0.035
1960
1970
1980
1990
2000
Year

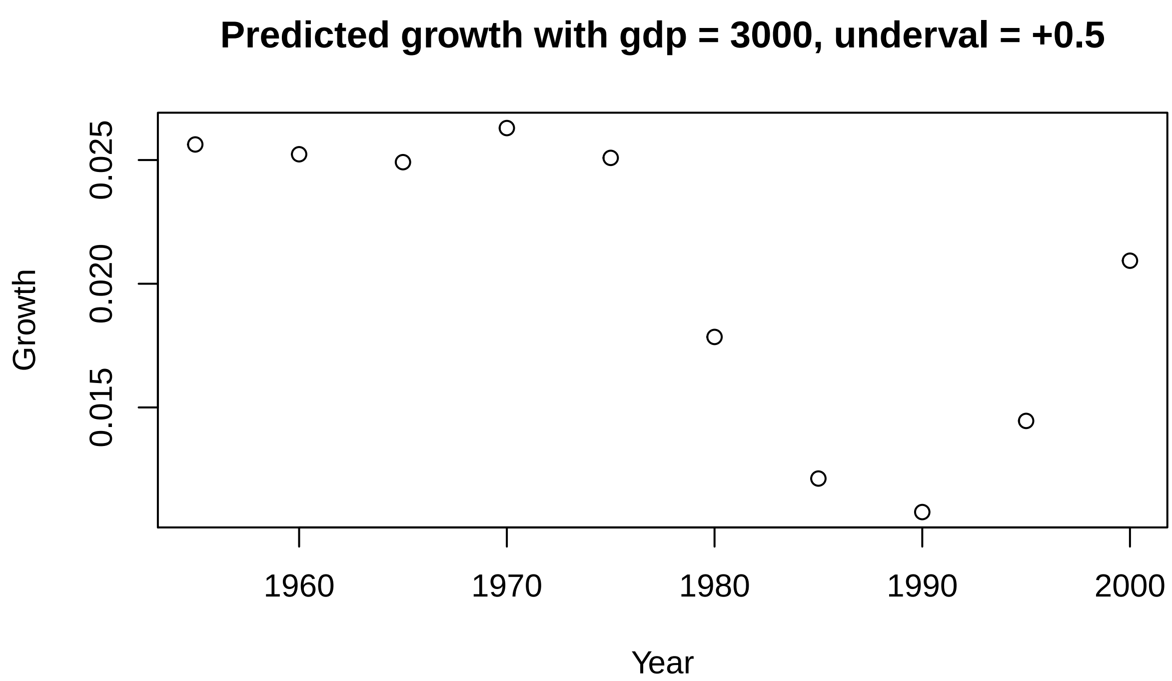
**Q5 b)**


Predicted growth with gdp = 20000, underval = +0.5
Growth
0.010
0.020
0.030
0.040
1960
1970
1980
1990
2000
Year

**Q5 c)**



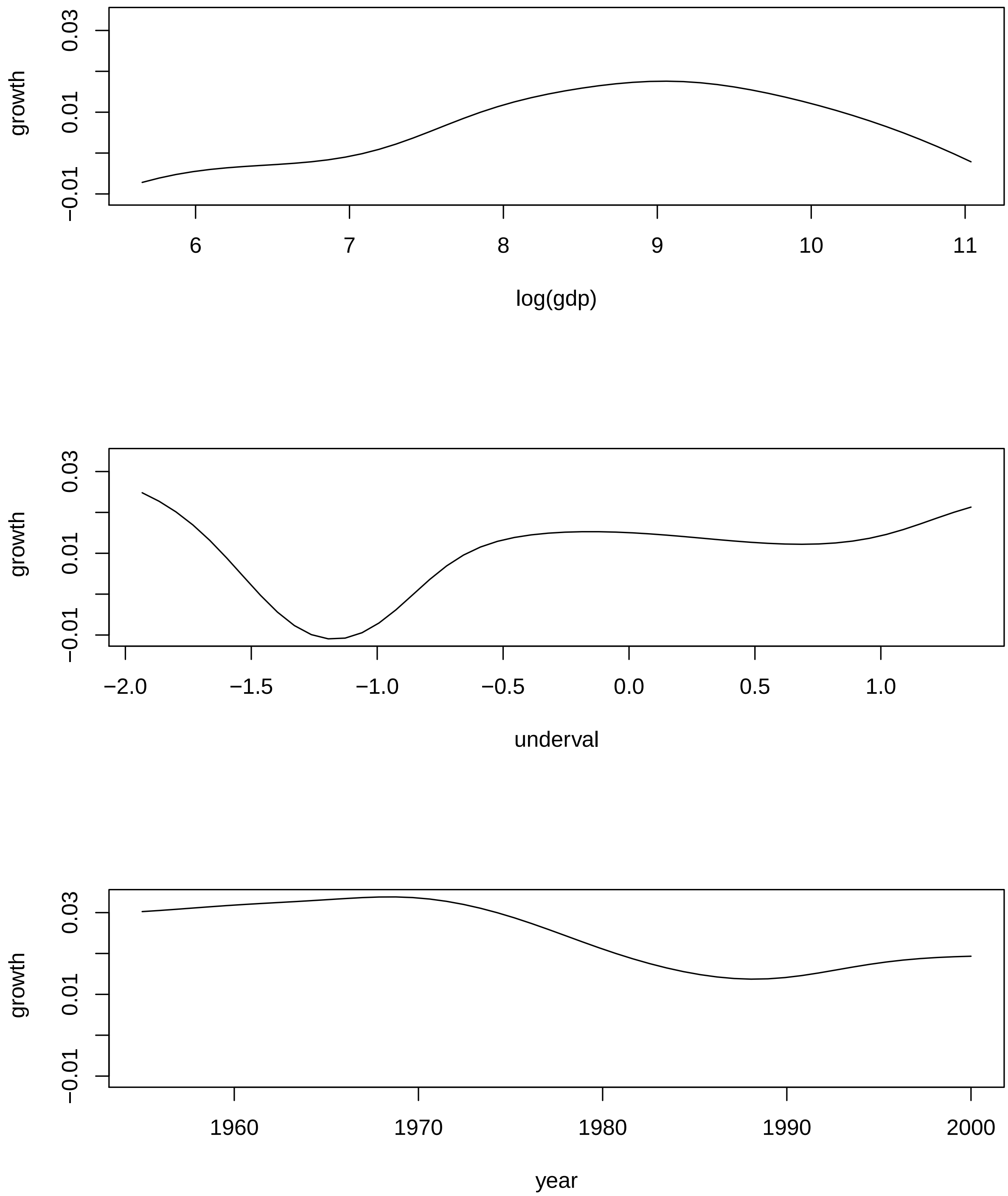

**Q5 d)**




**Q5 e)**

By comparing the graphs with initial $GDP = 20000$, we can see that by increasing under-valuation from 0 to 0.5, the predicted values increased from 1955 to 1970 but for $GDP = 3000$, increasing under-valuation from 0 to 0.5 decreased the predicted values from 1955 to 1970. The graphs are not parallel across and the effects not similar when we change intial GDP and under-valuation, hence there should be some interaction between these two variables.

**Q5 f)**



From the plots of growth against each variable, we can see that both GDP and under-valuation have strong relationshipswith growth. For the plot of growth against $log(GDP)$, there seems to be significant variation from varying $log(GDP)$ from 7 to 11. For the plot of growth against under-valuation, there also seems to be significant variation from varying under-valuation from -2.0 to -0.5. Finally, there is also some relationship, a weaker one than the previous two, between growth and year. For the plot of growth against year, there seems to be a slightly significant variation from avarying year from 1970 to 1985.