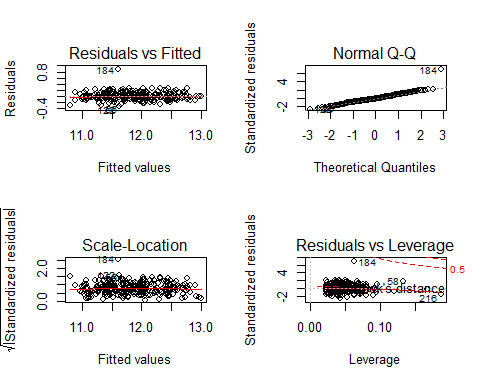
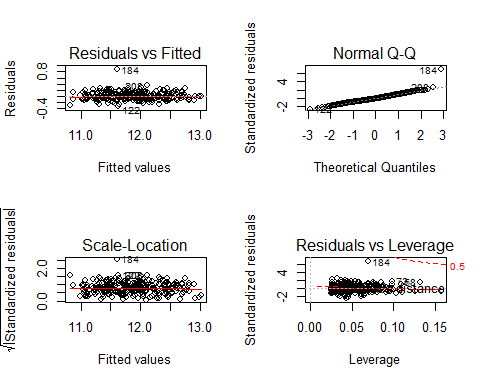
**Graph 4. Diagnostic plots for model 1**



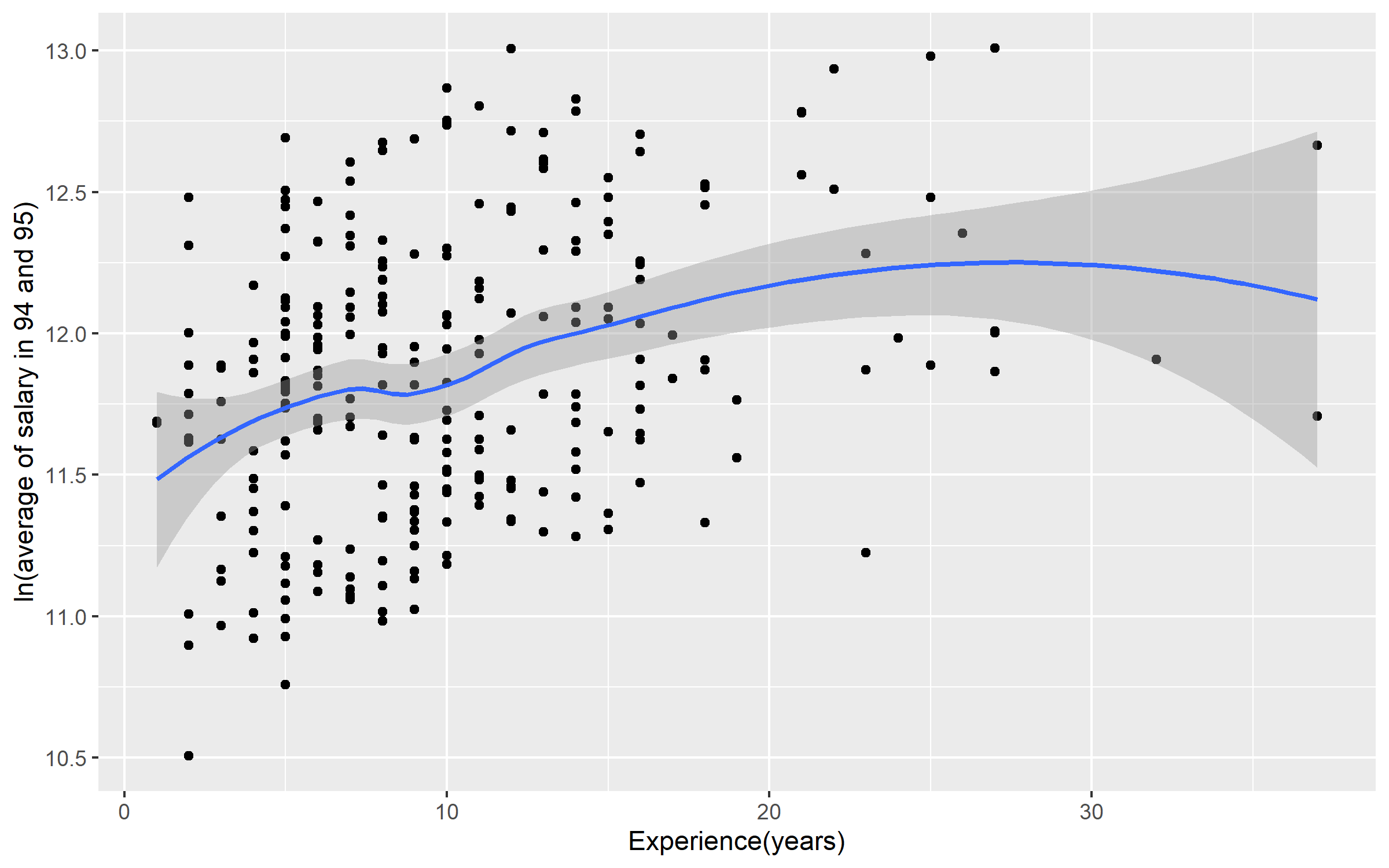
\* The curves in residuals vs fitted plot are horizontal and bounce around 0 which indicates that constant variance is satisfied in our model. Scale-location plot shows the same thing. No.184 observations deviates from the normal line on the qq plot means it might be an outlier. Except from this observation, other observations satisfy normality. No. 184 observation is close to the 0.5 cook's distance in residuals vs leverage plot, which means No.184 observation might be influential.

**Graph 5. Diagnostic plots for model 2**



\*The curves in residuals vs fitted plot are horizontal and bounce around 0 which indicates that constant variance is satisfied in our model. Scale-location plot shows the same thing. No.184 observations deviates from the normal line on the qq plot means it might be an outlier. Except from this observation, other observations satisfy normality. No. 184 observation is close to the 0.5 cook's distance in residuals vs leverage plot, which means No.184 observation might be influential.

**Graph 6. Scatter plot for experience against ln(average of salary in 94 and 95)**



\*In the scatter plot, experience is the X-axis variable, ln\_salavg (log transformation for mean salary) is the y-axis variable. This plot shows the potential non-linear relationship between them.

**Table 9. VIF of predictors after stratification by experience in model 1**

|  |  |  |  |
| --- | --- | --- | --- |
| **VIF** | **exper<10** | **10<exper<20** | **exper<20** |
| **gendermale** | 1.207 | 1.302 | 1.454 |
| **dept2** | 1.894 | 1.419 | 1.689 |
| **dept3** | 1.775 | 1.298 | 1.287 |
| **dept4** | 2.749 | 1.400 |  |
| **dept5** | 3.930 | 2.130 | 2.418 |
| **dept6** | 3.038 | 2.058 | 2.591 |
| **clin1** | 1.983 | 1.564 | 2.146 |
| **cert1** | 1.396 | 1.347 | 1.199 |
| **rank2** | 1.104 | 2.616 |  |
| **rank3** | 1.148 | 2.729 | 1.258 |

\* VIF > 5 was the criteria for deciding multicollinearity in the model; dept stands for department,

clin stands for primarily emphasis, cert stands for certified status; rank stands for career rank.

**Table 10. VIF value of predictors after stratification by rank in model 2**

|  |  |  |  |
| --- | --- | --- | --- |
| **VIF** | **rank = 1** | **rank = 2** | **rank = 3** |
| **gendermale** | 1.389 | 1.192 | 1.212 |
| **dept2** | 1.788 | 1.494 | 1.503 |
| **dept3** | 1.647 | 1.326 | 1.436 |
| **dept4** | 2.506 | 2.125 | 1.236 |
| **dept5** | 3.759 | 2.805 | 1.901 |
| **dept6** | 3.218 | 2.316 | 1.898 |
| **clin1** | 1.822 | 2.112 | 1.662 |
| **cert1** | 1.387 | 1.413 | 1.284 |
| **exper** | 1.259 | 1.167 | 1.115 |

\* VIF > 5 was the criteria for deciding multicollinearity in the model; dept stands for department,

clin stands for primarily emphasis, cert stands for certified status; exper stands for experience.

**Table 11. Potential influential observation in models**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model 1** | | **Model 2** | |
| **case** | **cook.d** | **case** | **cook.d** |
| 56 | 0.022 | 56 | 0.022 |
| 58 | 0.040 | 58 | 0.025 |
| 59 | 0.016 | 59 | 0.016 |
| 73 | 0.017 | 73 | 0.026 |
| 82 | 0.022 | 82 | 0.020 |
| 122 | 0.021 | 122 | 0.020 |
| 135 | 0.024 | 184 | 0.246 |
| 182 | 0.016 | 208 | 0.016 |
| 184 | 0.242 |  |  |
| 216 | 0.048 |  |  |
| 220 | 0.023 |  |  |

\* cook.d > 4 / n (n = 261 in our study) was the criteria for selecting potential influential observations in the models.