Same-sex marriage support rate and education level, marital status, age in Taiwan

POLI 706

Edwina Chih-Yu Chen

May 4, 2021

In May 2019, Taiwan became the first country in Asia to legalize same-sex marriage. Prior to the passage of the same-sex marriage bill, there are two critical events in the progress of advancing same-sex marriage in Taiwan. The first is the constitutional interpretation case in 2017, in which the Justices of Supreme Court considered that the current Civil Code is unconstitutional. The court requested the legislative branch to complete the enactment of the relevant law within two years. The outcome of the constitutional interpretation case has provoked opposition from conservative forces. As a result, a national referendum was held in 2018.

Hypothesis

This project focuses on this second event, the referendum in 2018. By analyzing the referendum results and the voter data of the region, the project aims to understand what personal factors may influence the support of same-sex marriage.

In this project, I hypothesized that people with higher levels of education, young people, would be more likely to support same-sex marriage. In Taiwan, the inadequate protection of property and medical-related issues is a critical reason for one to support the bill. Married people are more aware of the importance of marital rights from this perspective. Hence, I assume that married people sympathize with the LGBT community and are therefore more likely to support same-sex marriage. I also wanted to find out if there are more same-sex couples in an area, will the support rate for same-sex marriage be higher. This project aims to test the following Hypotheses:

Hypothesis 1: Higher education graduation rate and the percentage of same-sex marriage support is positively associated..
Hypothesis 2: Youth population percentage and the percentage of same-sex marriage support is positively associated.
Hypothesis 3: Marriage rate and the percentage of same-sex marriage support is positively associated.
Hypothesis 4: Same-sex marriage rate and the percentage of same-sex marriage support is positively associated.

Data

I am merging two datasets in this project. The first dataset, obtained from the g0v movement open source platform, has the referendum voting result in 2018 sorted by district. From the Ministry of the Interior open data platform, the second dataset included population data with basic information (sex, educational level, marriage status) in 2019. I am using the

2019 population data instead of 2018 because I am interested in the relationship between the same-sex marriage population and the support rate.[1] Using these two datasets, I create a dataset that contains information I need, which include following variables:

Dependent variable *Support rate* is the voter percentage that support same-sex marriage.[2] Explanatory variables including *Higher education rate*, the percentage of population with a higher education degree. *Youth Population rate*, the percentage of population between 20 to 34 years old. *Same-sex marriage rate*,the percentage of population who married same-sex partner. *Marriage rate*,the percentage of population who married different sex.
The level of analysis is district level, the basic administration unit in Taiwan. Some of them are town or city; to simplify, I use the term "district" to refer to a unit.

[1]The actual date of implementation after the passage of the same-sex marriage bill is May 24, 2019. Therefore this information does not become available until the 2019 population statistics.

[2]The agree vote count of Question 14th in 2018 referendum : Do you agree to the protection of same-sex marital rights with marriage as defined in the Civil Code?

Visualization

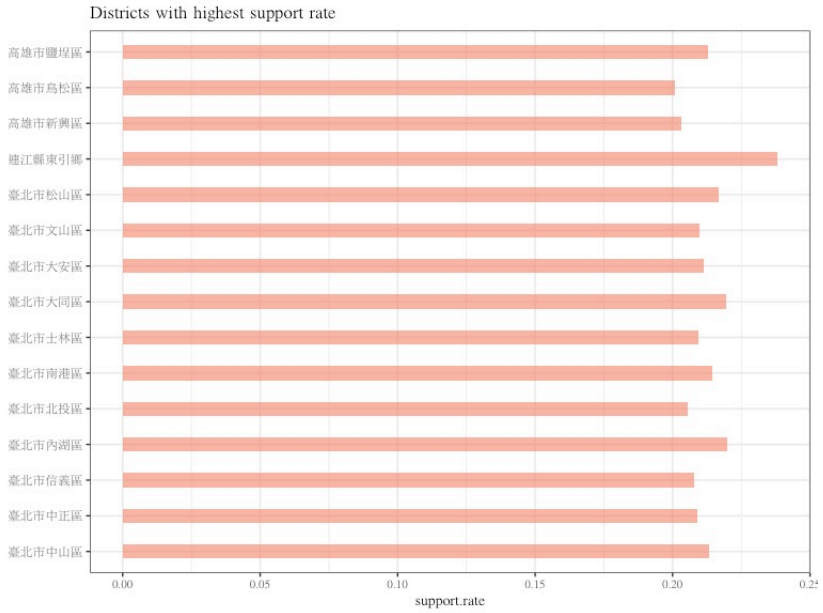

Figure 1: Districts with Highest support rate

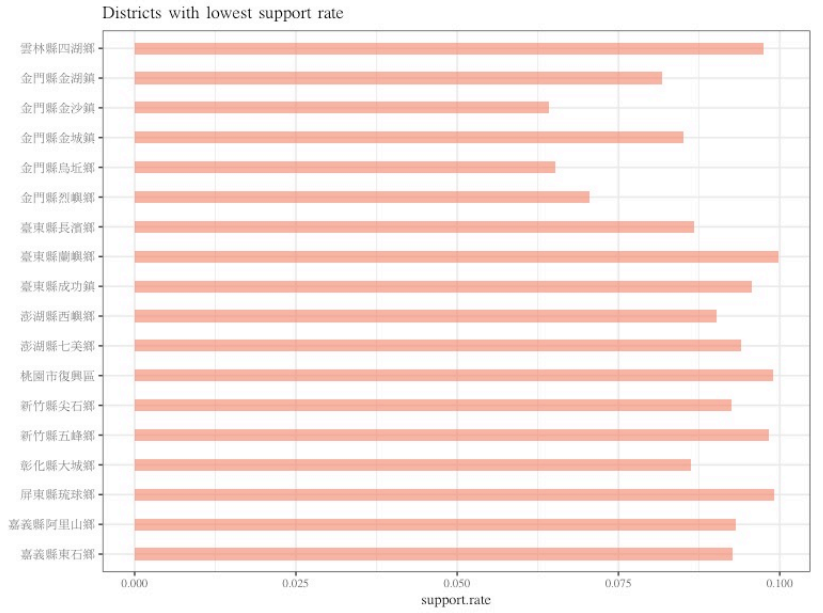

Figure 2: Districts with Lowest support rate

Figure 1 shows the districts with the highest support rate. Figure 2 shows the districts with the lowest support rate. The range of support rate is roughly between 0.08 to 0.24.

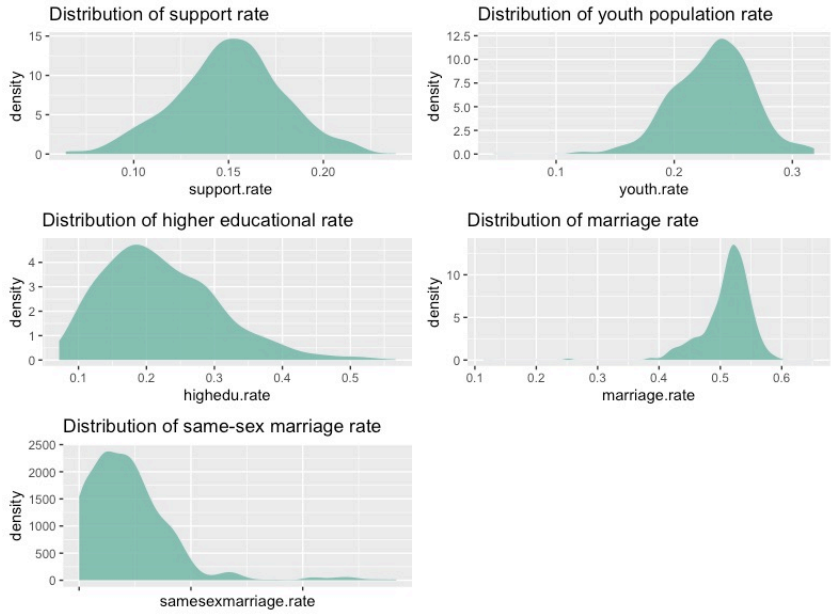

Figure 3: Distributions of variables

Figure 3 shows the distribution of all variables. The distribution of *Support rate*is normal. The distributions of *Youth Population rate* and *Marriage rate* are left-skewed. The distributions of *Higher education rate* and *Same-sex marriage rate* are right-skewed.

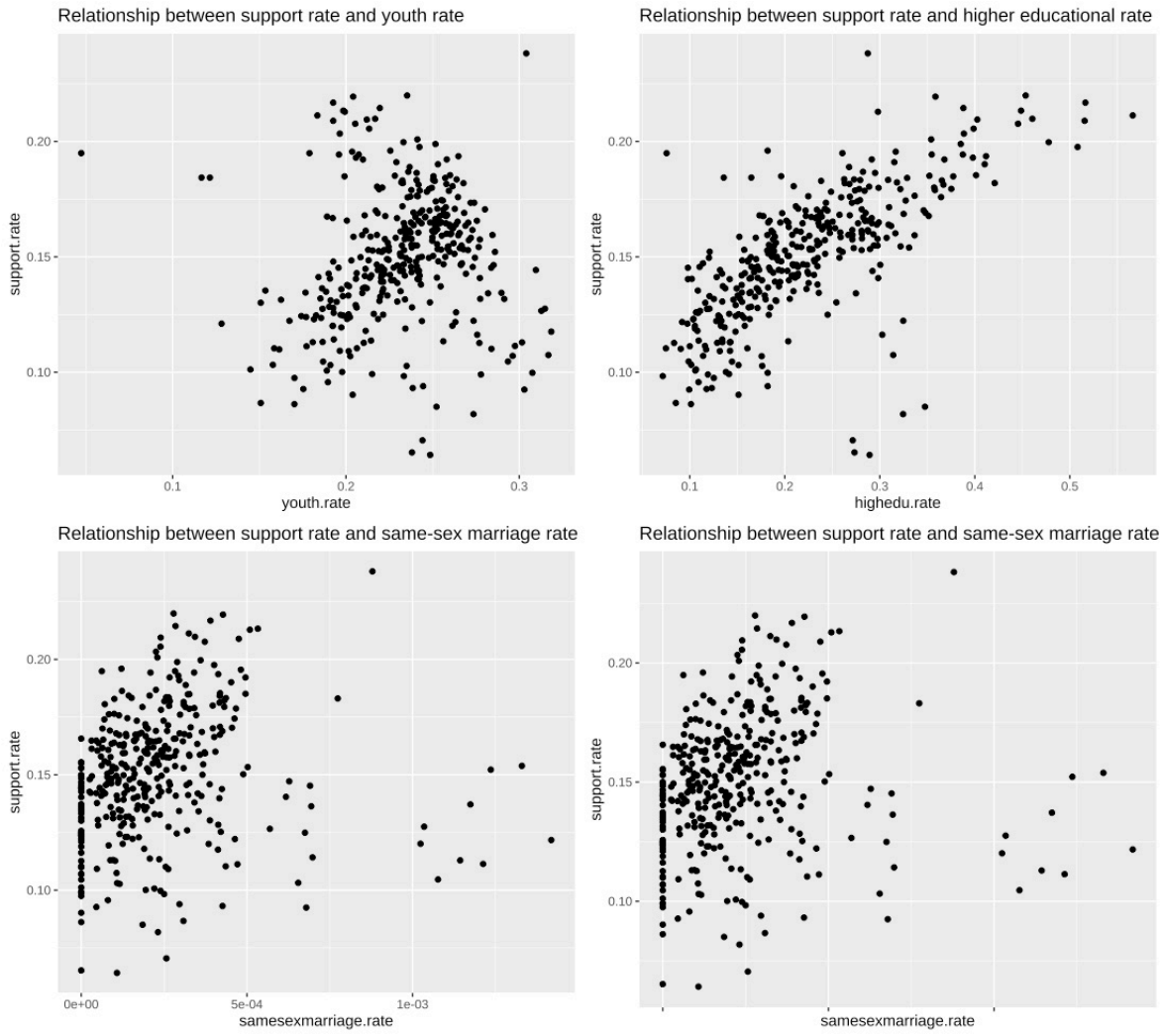

Figure 4: Relationship between Explanatory variables and Dependent variables

Figure 4 shows the relationship between explanatory variables and dependent variables. *Higher education rate* seems to have a strong linear relationship with *Support rate*.

Analysis

As shown in Figure 5, all variables are correlated with each other. We might have a multi-collinearity problem; we should check the VIF score of the model afterwards.

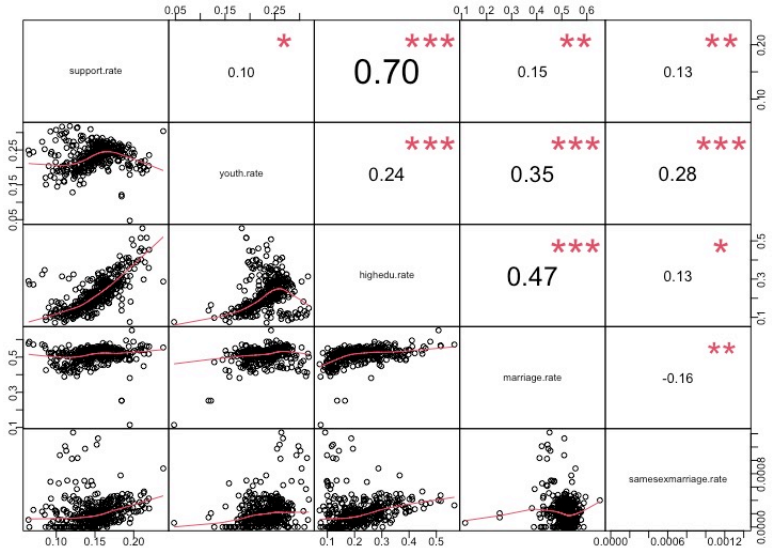

Figure 5: Correlation matrix of variables

I use the OLS method to measure the association between variables.In all possible combinations of explanatory variables, I identify the model with the highest adjusted R-squared. Here, Model 1 has the highest adjusted R- squared (with all four explanatory variables). Using backward selection (comparing p-value), model 2 excludes same-sex marriage rate, and model 3 excludes same-sex marriage rate and Youth population rate.

Table 1: Coefficients and 95 percent confidence bounds OLS

	DV = Same-sex marriage support rate		
	Model 1	Model 2	Model 3
Higher education rate	0.26***	0.26***	0.26***
	(0.24, 0.29)	(0.24, 0.29)	(0.24, 0.29)
Marriage rate	−0.14***	−0.14***	−0.14***
	(−0.19, −0.08)	(−0.19, −0.09)	(−0.19, −0.09)
Youth Population rate	−0.01	−0.01	
	(−0.08, 0.06)	(−0.08, 0.05)	
Same-sex marriage rate	−0.74		
	(−11.14, 9.66)		
Constant	0.16***	0.16***	0.16***
	(0.14, 0.19)	(0.14, 0.19)	(0.14, 0.19)
Observations	368	368	368
Adjusted R^2	0.52	0.52	0.52
Residual Std. Error	0.02	0.02	0.02
F Statistic	100.37***	134.18***	201.65***

*** p less than 0.001, ** p less than 0.01, * p less than 0.05

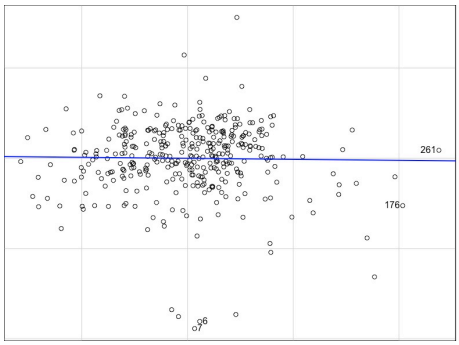

Figure 6: Added-Variable Plots: Youth rate

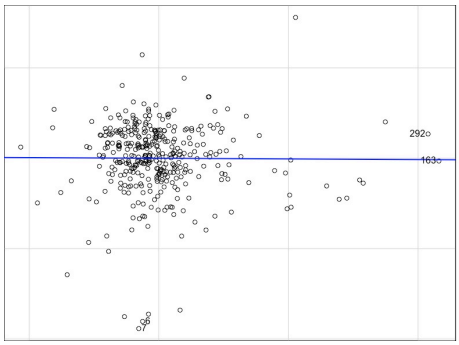

Figure 7: Added-Variable Plots: Same-sex marriage rate

We can see that model 3 has almost the same adjusted R-squared value compared to model 1 and model 2 (I round the number for table 1 so they look equal, but it does differ slightly.) As shown in figure 6 and 7, the partial effect plots of same-sex marriage rate and Youth population rate, the slopes are both low. These two variables are not very important. The best-fitting model is model 3.

For model 3, The intercept is 0.16, when both Explanatory Variables are 0, we expect the

support rate to be 0.16. The higher education rate is positively associated with the support rate and is statistically significant at the 0.001 level. Holding marriage rate constant, for every 1 unit increase in higher education rate, we expect the support rate to increase by 0.26. The marriage rate is negatively associated with the support rate and is statistically significant at the 0.001 level. Holding higher education rate constant, for every 1 unit increase in the marriage rate, we expect the support rate to decrease by 0.14. The adjusted R squared is 0.52. The explanatory variables in our model explain 52 percents of the variance. The F-statistic represents the overall significance of model 3. The probability of our coefficients are all zero is very small (p-value smaller than 0.05).

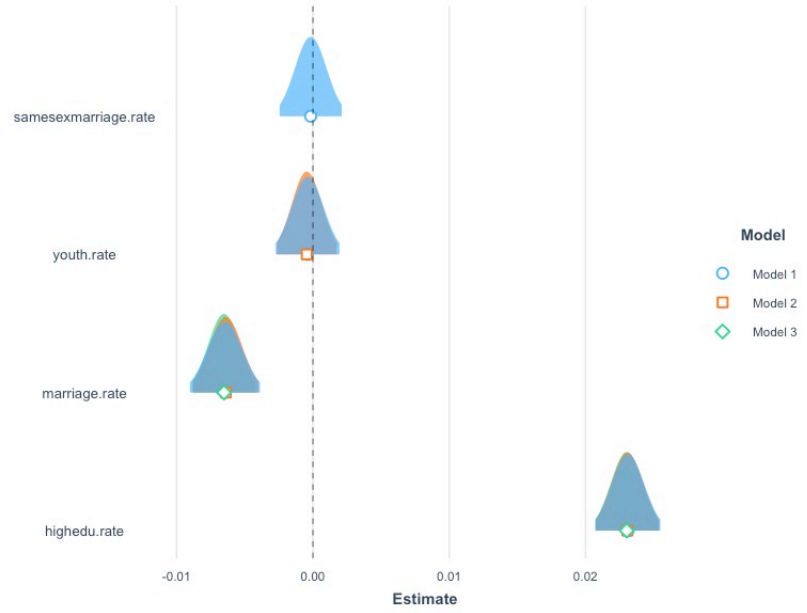

Figure 8: 95 percents interval width of distribution for each estimate in Model 1, 2 and 3

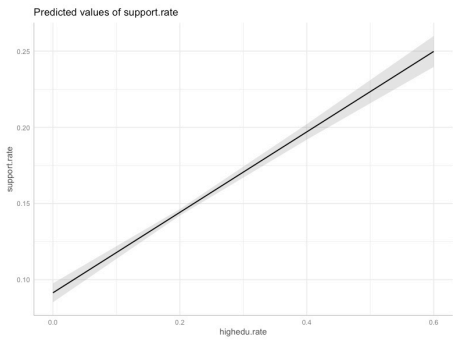

Figure 9: Predicted value across higher education rate

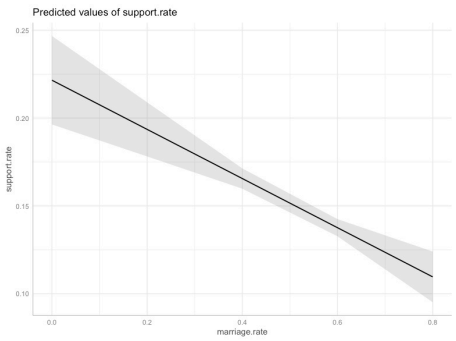

Figure 10: Predicted value across marriage rate

Figure 8 illustrates 95 percents interval width of distribution for each estimate for models. Figure 9 and 10 illustrate predicted values of support rate using higher education rate or marriage rate as predictor with 95 percents confidence intervals.

Robustness test for model 3

From Figure 11, we can check whether the model meets the assumptions of OLS. Firstly, the Residuals vs. Fitted plot has a horizontal line and no obvious pattern. Model 3 meets the linear relationship assumption.

Secondly, the Normal Q-Q plot can examine whether the residuals are normally distributed. Here the residuals points roughly follow with the line which means the model has normal residuals.

Thirdly, the Scale-Location plot can examine homoscedasticity, it seems to be close to a horizontal line, but the points don’ t spread equally. So, I also did a Breusch-Pagan test, which has a p-value of 0.095. We can’ t reject the null hypothesis that homoscedastic is present.

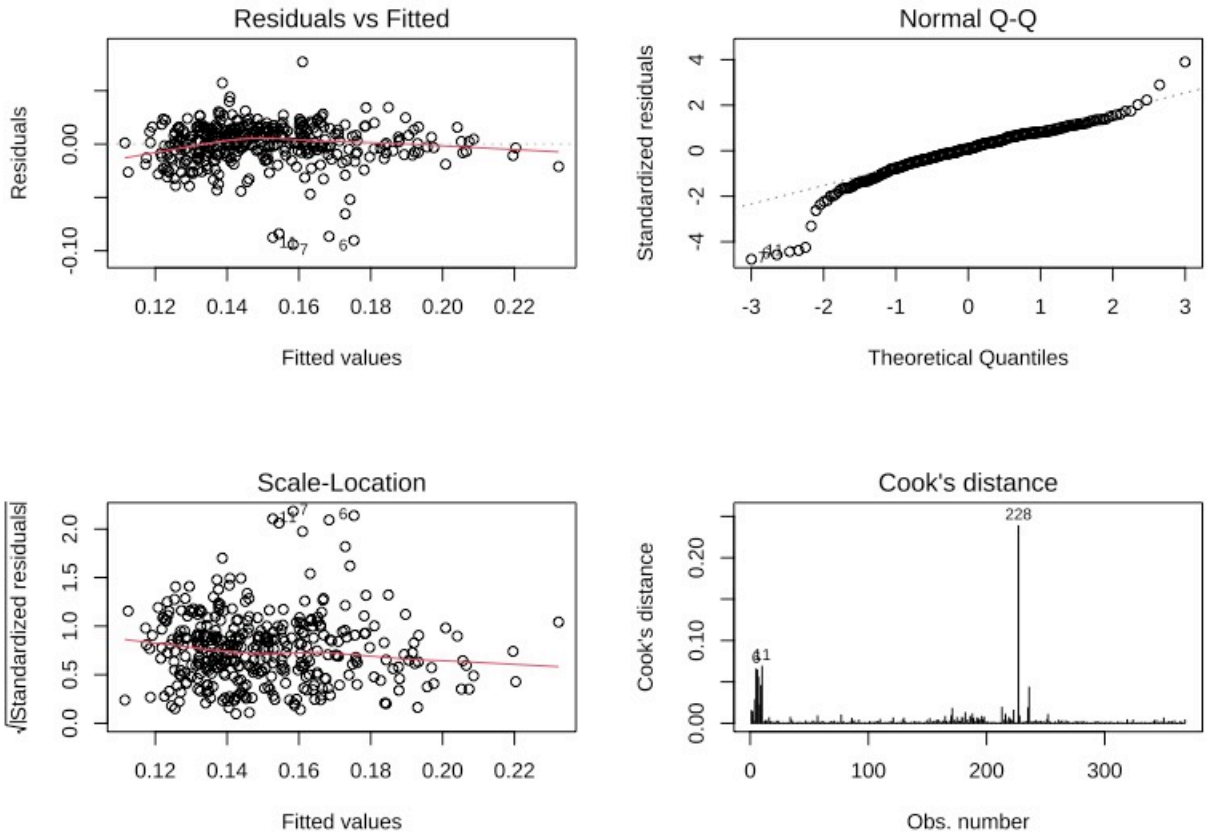

Figure 11: Diagnostic Plots

For the multicollinearity problem, the VIF score for both the higher education rate and marriage rate is 1.27. The scores indicate low multicollinearity between these two variables.

Lastly, I identified several observations that had a great impact on the coefficient and did not fit well in our model by removing obersvation one by one and recalculate coefficients (see Figure 12 and Figure 13). For *Higher education rate* coefficient, these observations include 連江縣東引鄉, 金門縣金湖鎮, 台北市信義區. The first two districts are underpopulated, so voting was polarized, with only solid supporters or opponents showing up to vote.As for

台北市信義區 (Xinyi District, Taipei City), it is probably the most educated district but does not have a very high approval rate. For *Marriage rate* coefficient, these observations include 連江縣莒光鄉, 彰化縣和美鎮, 台東縣東河鄉, 花蓮縣鳳林鎮.These towns do not have substantial marriage populations but show less support, mainly due to sparse and aging populations.

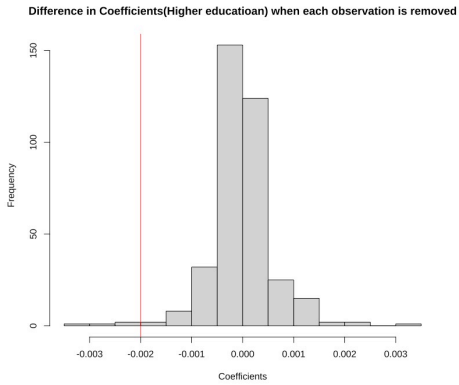

Figure 12: Distribution of Coefficients difference -Higher Education

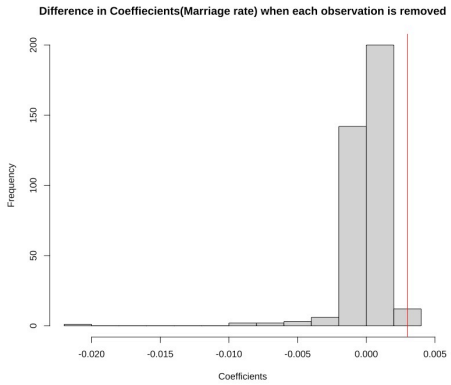

Figure 13: Distribution of Coefficients difference- Marriage rate

From Figure 11 (Cook's distance) and Figure 14 , we can see that the most influential observation is 高雄市楠梓區 (Nanzhi District, Kaohsiung City). Given the marriage rate and education level of the area (just a bit higher than average), the support rate should not be so high. The most likely explanation is that Nanzhi District has a technology industrial park and a high proportion of young people (0.257), so the residents are more progressive in human rights issues. We might better explain these outliers if population density and income level are included in the analysis.

For the multicollinearity problem, the VIF score for both the higher education rate and marriage rate is 1.27. The scores indicate low multicollinearity between these two variables.

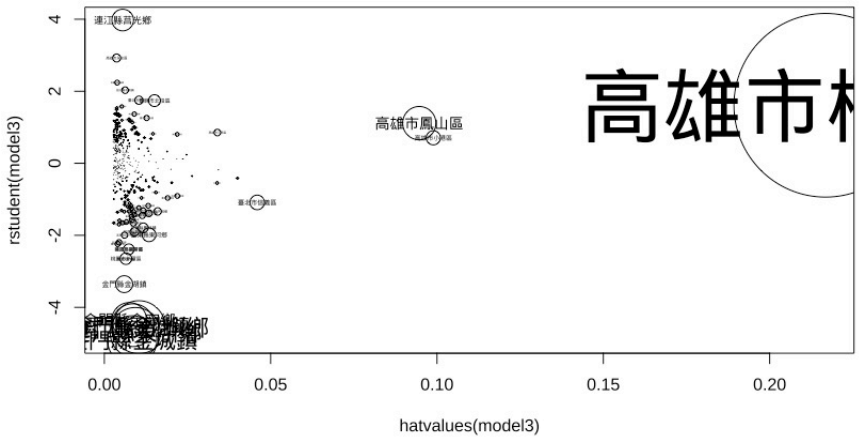

Figure 14: Influential observations in model 3

Conclusion

After quantitative analysis, I conclude that the percentage of higher education and the percentage of the married population are associated with the support rate of same-sex marriage in Taiwan. Contrary to the hypothesis, there is a negative association between the marriage population and the support rate. However, the higher education rate is positively associated with the support rate of same-sex marriage. There is no statistically significant relationship between the percentage of young people, the percentage of same-sex marriages, and the support rate.

References

Data.GOV.TW. 2021. *Population statistic with site id* 戶里人口資料. [online] Available at: <https://data.gov.tw/dataset/77132> [Accessed 26 April 2021].

Liu, Y., 2021. *g0v/referendum_report 2018*公投第十案. [online] GitHub. Available at: <https://github.com/g0v/referendum_report/blob/master/results/第10案：你是否同意民法婚姻規定應限定在一男一女的結合%3F.csv> [Accessed 26 April 2021].