GEOG788P Final Project Paper

Yuhao Wang

12/15/2021

How Gas Consumption Pattern Changed and Mental Health Condition Changed in Maryland Before and During the Pandemic

**Introduction:**

The coronavirus (COVID-19) quickly swept the world in December 2019 and was declared a global pandemic on Mar 11, 2020, by the World Health Organization. To this day, the pandemic has not ended yet; more than 160 Million cases of COVID-19 have been reported worldwide, most countries are restricting the population movement through social distancing interventions to stop the spread of COVID-19, such as work-from-home and close the public places. These policies have successfully decreased the spread of disease. The continuing spread of COVID-19 in the world has led to both public health as well as an economic crisis in the United States. With more than 32.8 million confirmed cases and 582K deaths reported inside the US, the COVID-19 has now more than the American soldiers lost in World War II. The economic crisis is unprecedented. From demand to supply sectors as well as the financial market, they are all suffered from it.

Due to the social distancing policy and fear of getting infected in public places, people choose to avoid unnecessary travel, leading to a vast decrease in the economy. The demand shock in the US is resulting from quarantine, unemployment, and business closures that dealt a blow to consumer services (Bauer et al., 2020). The economy’s capacity to produce goods and services has been primarily hit by lockdown measures and social distancing policies (Bauer et al., 2020). What is more, the National Bureau of Economic Research (NBER) determined that a peak in monthly economic activity occurred in the US economy in February 2020, marking the end of the longest recorded US expansion, which began in June 2009 (Bauer et al., 2020.). One obvious indicator of economy downturn is the unemployment rate. According to the US Bureau of Labor Statistics, the Employment–population ratio has dropped from the 61.1% in Jan 2020 to the lowest 51.3% in April 2020, which is the lowest in the nearest 20 years, until April 2021, the Employment–population ratio is 57.9%, which are still lower than the lowest point during the economic crisis back to 2008.

The US's gross domestic product (GDP) was dropped 3.5% compared with 2009, with a dramatic swing in household spending in the early development of pandemic in the US. Retail sales declined 8.7% from February to March 2020, the most significant month-to-month decrease since the Census Bureau started tracking the data. However, some retail sectors have increased, such as grocery, pharmacy and non-store retailers, due to the panic demand while other retail sectors such as auto, clothing are declined. As some states lifted social distancing restrictions in early May, sales began to recover in most goods sectors (Bauer et al., 2020). Overall, U.S. retail sales increased 17.7% from April to May, with the most significant monthly jump on record. It regrouped 63% of March and April’s losses (Bauer et al., 2020). Total retail sales increased 6.9% to $4.04 trillion in 2020 from $3.78 trillion in 201. According to Digital Commerce 360’s analysis from the latest US Department of Commerce figures, this is the highest growth since 1999 (this excludes sales of items not normally purchased online, such as spending at restaurants, bars, automobile dealers, gas stations and fuel dealers).

Although throughout the year of 2021, with the widespread of COVID-19 vaccines, the new confirmed cases of COVID-19 keeps decreasing. But new COVID-19 variants are being discovered along with potential to lower the vaccines’ effectiveness, the It seems like the pandemic would end in a not far future. In addition to that, as the spread of COVID-19 is successfully limited inside the US, policymakers now pushing the U.S market to recovery stage. With the help from government funding, the economy will no doubt increase. While people’s life are slowly going back to pre-pandemic era. However, it is still unknown that how people’s lifestyle has been permanently changed due to the pandemic.

Therefore, it is important to understand how the people’s lifestyle has any difference during the pandemic compared with pre-pandemic era. In this project, I aim to detect how gas consumption pattern has changed and mental health condition has changed in Maryland before and during the pandemic, from 2016 to 2020 among different census tracts. I choose because the gas consumption as an indication on to what degrees people choose to change their lifestyles based on the pandemic, since people would choose to stay at home to avoid the potential infection during the pandemic (before the large distribution of vaccines), people will drive their car less often, therefore, the amount to gasoline towards vehicle usage would decrease. However, it is unknown what could be the main driver for the decrease usage of gasoline, the reasons for decreased gasoline usage could be but not limited with: people who lose their job during pandemic; people who choose to work from home instead of work at office; people who choose to use online shopping to purchase grocery instead; the household income; education attainment. One of this project’s objective aims to find out the major driver of gas usage among different regions in Maryland before and during pandemic.

Another objective of this project is to find the drivers of mental health issues in among different regions in Maryland before and during pandemic. The pandemic does not only affect people’s physical health by its virus, it also affects people’s mental health because people has to stay at home to avoid affection, people may potentially lose their jobs or have less income due to the economic shock caused by pandemic, additionally, people may lose their family or friends because of pandemic. Therefore, similar to gas usage decrease during pandemic, the mental health issues become more severe, it is important to find out the drivers behind this and analyze whether the drivers were changes compared with pre-pandemic era. This pandemic does not simply affect people who are infected with COVID-19, but the affect general society who may even have infect COVID-19 with mental health issues.

This project aims to use SimplyAnalytics as a web application to download data variables; and use python to detect the drivers of gas usage and mental health from 2016 to 2020 in different census tracts in Maryland. One potential outcome of this project is to find the drivers, another exception of this project is to test the dataset from SimplyAnalytics. The methods this project aims to use is to conduct linear regression, spatial lagged regression, spatial error regression, geographically weighted regression and multiscale geographically weighted regression to find driver for different years and make a comparison between different years at the end.

**Literature Review:**

Several studies were discussing about the gasoline consumption during the pandemic in different places of the world with different focuses. Ou et al. published a study in 2020 to project the impact of COVID-19 on US motor gasoline demand. By studying the relationship between the new confirmed cases/death of COVID-19 and mobility data, they create a machine learning method to predict the future mobility trends depending on COVID-19 projections. Since the distribution of COVID-19 vaccines becomes more available to the public starting from late 2020, the newly confirmed cases decreased in the USGüngör et al. discussed how the COVID-19 outbreak impact on Turkish Gasoline consumption in 2020. During this paper, them investigate the effects of Covid-19 outbreak on Turkish gasoline consumption by employing a unique data set of daily gasoline consumption data covering the 2014-2020 period. An ARIMA (Autoregressive Integrated Moving Average) model was used in this paper to forecast the gasoline usage for both before and after the outbreak. The result shown that the model predicts only 0.8% difference between actual and forecasted gasoline consumption data. However, there is approximately a 30% difference after the outbreak.

Another Research Published by Smith et al. discussed the impact of COVID-19 on global fossil fuel consumption and CO2 emissions. In this paper, they assess the effect of the COVID-19 pandemic on global fossil fuel consumption and CO2 emissions over the two-year horizon from the first quarter of 2020 to the last quarter of 2021. They apply a global vector autoregressive (GVAR) model, which can be used to capture complex spatial-temporal interdependencies across countries associated with the international propagation of economic impact due to the pandemic. They produce forecasts of coal, natural gas and oil consumption, conditional on GDP growth scenarios based on alternative IMF World Economic Outlook forecasts that were made before and after the outbreak, and their results suggests that the pandemic’s influence on energy consumption will only have a significant drop on the early stages of pandemic, but during the recovery stage the boost energy consumption will significantly raise, hence that the CO2 emissions caused by burning fossil fuels could be substantially lower than in the case of no COVID-19 shock for the advanced economies, but the impact is more limited for the emerging economies. Thus COVID-19 would not provide countries with a reason to delay climate-change mitigation efforts.

Finally, Nyga-Łukaszewska & Aruga in 2020 discussed how the pandemic has influenced oil and gas prices, using energy market reactions in the United States and Japan, they applied the Auto-Regressive Distributive Lag (ARDL) approach to understand the impact of the COVID-19 cases on the crude oil and natural gas markets from Jan 2020 to Jun 2020, which is so called the “first pandemic wave”. They Conclude that in the US, the pandemic had a statistically negative impact on the crude oil price while it positively affected the gas price. In Japan, this negative impact was only apparent in the crude oil market with at wo-day lag. It could possibly be due to the different development during the first wave of pandemic and the diverse roles both countries have in energy markets.

Several Studies have also discussed on mental issues during the pandemic affect among different groups of people. Holingue et al. researched on how Mental distress conditions US adults without a pre-existing mental health condition in 2020. In study they want to assess the frequency and risk and protective factors of psychological distress, during the beginning of the pandemic. They focused at individuals with no prior history of a mental health condition, using survey data from the Pew Research Center's American Trends Panel (ATP). From the survey, they found that 15% of the sample group experienced 2 psychological distress symptoms for at least 3 days over the past week; 13% had three or more symptoms. Similar study conducted by Wang et al. in 2020 investigates the mental health of us college students during the pandemic also using survey data. They looking for the mental health status and severity of depression and anxiety of college students in a large university system in the United States during the pandemic. From their survey results, 48.14%showed a moderate-to-severe level of depression, 38.48% showed a moderate-to-severe level of anxiety, and 18.04% had suicidal thoughts. 71.26% of survey participants indicated that their stress/anxiety levels had increased during the pandemic.

In addition to the two papers mentioned above, another research conducted by Liu et al., 2020 are looking for the factors associated with depression, anxiety, and PTSD symptomatology during the COVID-19 pandemic among young adults in US. The researchers conducted a cross-sectional online study from April 2020 to May 2020. They find their study respondents reported high levels of depression (43%), high anxiety scores (45.4%), and high levels of PTSD symptoms (31.8%). High levels of loneliness, high levels of COVID-19-specific worry, and low distress tolerance were significantly associated with clinical levels of depression, anxiety, and PTSD symptoms. They further compared to Whites to Asian Americans were less likely to report high levels across mental health symptoms, and Hispanic/Latinos were less likely to report high levels of anxiety.

During the GEOG788P class, we studied models and methods for spatial data science. The class provided readings as a summarization of models and data science terms I learned during this class, also since most of this project’s methods are coming from the class reading, It is important to conduct a short literature review based on the class materials.

David Donoho in 2017 published a paper discussed the general idea of how data science evolved during past 50 years. Data science as a relatively new term compare with other science subjects were only becoming popular in the recent decades. But it actually is academic statistics to expand its boundaries beyond the classical domain of theoretical statistics. With the development of technology, data science are now dealing with extremely large datasets. In his paper, Donoho presents a vision of data science based on the researchers who are “learning from data,” and describes an academic field dedicated to improving how to conduct data science research.

Robert Haining in the book “Handbook of Regional Science” published in 2013 has a chapter review about spatial data and statistical methods. In this chapter, he discussed the spatial statistics from the beginning of different types of spatial data, and how different types of location data may need special consideration based on the shape of the data regions. He overviews of some of the major developments in spatial statistics with relevance to geography and regional science. In addition to that, he also mentioned about the emergence of spatial econometrics and its links (and overlaps) with spatial statistics. The third section of his paper considers the emergence of exploratory spatial data analysis and the development of “local statistics” for analyzing spatial heterogeneity. The fourth section is an overview spatial data mining. Followed by Sergio J. Rey, in the same book in 2013, he writes a chapter on spatial dynamics and space-time data analysis, where he provides an overview of spatial dynamics in the field of regional science. In his chapter, he focuses on the exploratory methods for space-time data, the development of spatial patterns as well as the identification of temporal dynamics that cluster in space.

Finally, this project used GWR (geographically weighted regression) and MGWR (multiscale geographically weighted regression) as a method to detect the spatial autocorrelation. Where there are several class readings discussed about GWR and MGWR. Fotheringham et al. published the MGWR in 2017. Since the classical GWR assumes that all of the processes being modeled operate at the same spatial scale, however it could not be truth in the actual scenario, where they bring up the MGWR as by deriving an optimal bandwidth vector in which each element indicates the spatial scale at which a particular process takes place. In MGWR, the model calibration and bandwidth vector selection were conducted using a back-fitting algorithm. They also compared the MGWR’s performance with GWR, and conclude that MGWR out performed GWR by replicating parameter surfaces with different levels of spatial heterogeneity and it also can provide valuable information on the scale at which different processes operate. Oshan et al. published a paper in 2018 which discussed the implementation process of MGWR to Python, in order to operate multiscale analyze of spatial heterogeneity. This paper also provides a literate programming style of how this model would work in Python, as well as a comparison of his MGWR model to two other software implementations to compare computational efficiency.

**Method/Data:**

One expectation of this project is to explore the SimplyAnalytics’ data quality and quantity, SimplyAnalytics is a web-based mapping, analytics, and data visualization application that makes it easy for anyone to create interactive maps, charts. SimplyAnalytics contains extensive data including demographic, historic census, business, health, real estate, housing, employment, consumer spending, and marketing (over 70,000 variables total). Users can create customized maps and reports based on the web platform. The finest spatial resolution of this dataset could goes down to the Block Group level. While most of their data is available at the State, County, City, ZIP Code, Census Tract and Block Group level, custom trade area, and the entire United States. However, data with spatial resolution with Block Group level are mainly comes from census dataset. One advantage of using SimplyAnalytics application is it provides science researchers with premium paid datasets such as SimplyAnalytics’ own premium datasets, which are Community Demographics and Consumer Expenditure Estimate, it also contains EASI’s health, light stage clusters and MRI consumer survey. Figure below provides a list of available datasets from SimplyAnalytics. SimplyAnalytics can directly export data as shapefiles or CSV depending on the data variables as well.

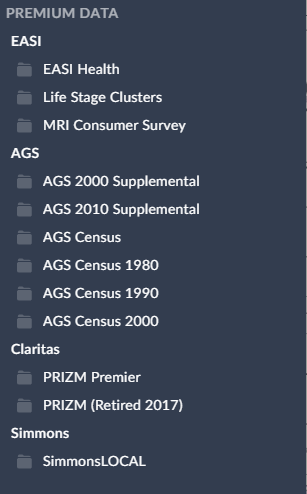
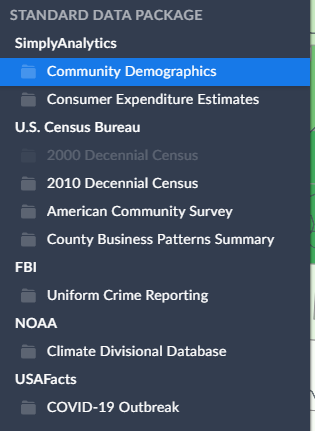


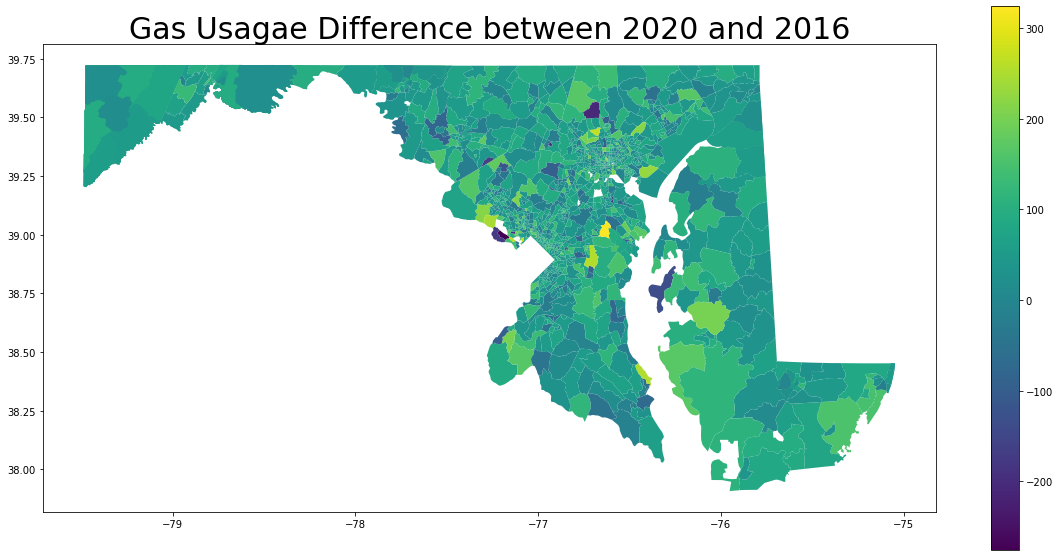
Figure1. Lists of data folders in SimplyAnalytics

The study area for this project is Maryland, and the spatial resolution for this project is census tract. There are 1942 number of census tracts in Maryland. It is because county-level dataset is not detailed enough to distinguish spatial difference within the state of Maryland, while the block groups does not have enough data variables to conduct the research, as well as the zip codes does not really work very well in terms of spatial regressions. The temporal coverage of this study is from 2016-2020, which has four yearly datasets. The extracted variables are showed in the table below, there are 24 variables in total, while the population density is not available before 2018. The independent variables could be categorized as following categories: Transportation methods; Vehicles availability for each housing units; Internet purchasing behavior; food expenditures; housing expenditures; and demographic data. While the dependent variables are the house average expenditures on gasoline, and the mental health survey of how many percent of general public have experienced sadness. Because the house average expenditures on gasoline does not reflect to the actual gas usage of each year due to the gas price changes over time, therefore, a yearly averaged gas price of Maryland from 2016 to 2020 is retrieved to convert the household average expenditure of gas to household average of used gas. The mental health variable of sadness are converted from the aggregation of Sadness - Some of the time and Sadness- All or most of the time since both of these variable are a percentage to the total population.

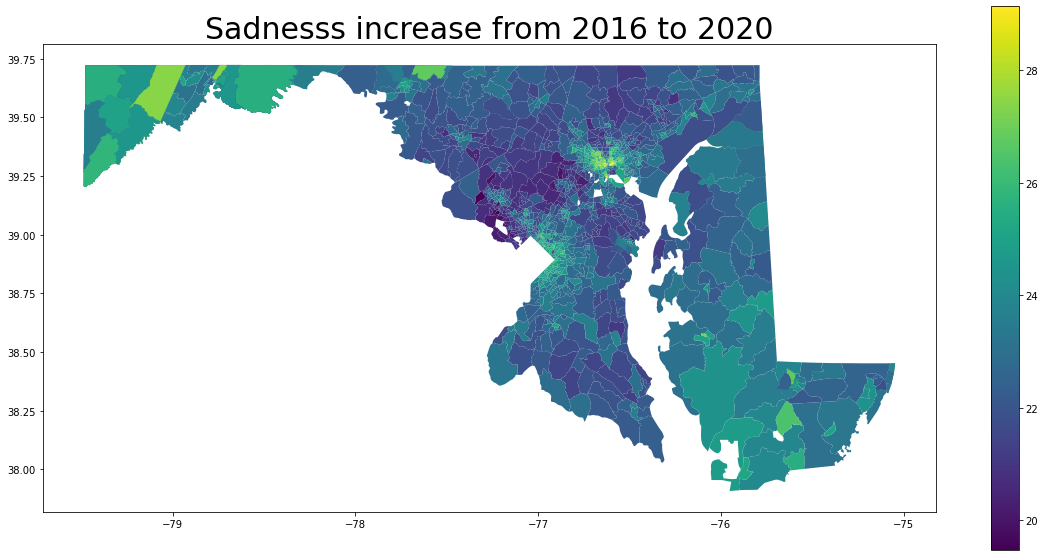
Table1. Variables Retrieved from SimplyAnalytics

|  |  |
| --- | --- |
| Variable Name | Dataset Downloaded from |
| # Educational Attainment | High school graduate (includes equivalency) | SimplyAnalytics Community Demographics |
| # Employment Status | In labor force | SimplyAnalytics Community Demographics |
| # Total Population | SimplyAnalytics Community Demographics |
| # Vehicles Available | No vehicle available | SimplyAnalytics Community Demographics |
| # Vehicles Available | Occupied housing units | SimplyAnalytics Community Demographics |
| % Educational Attainment | High school graduate (includes equivalency) | SimplyAnalytics Community Demographics |
| % Employment Status | In labor force | SimplyAnalytics Community Demographics |
| % Poverty Status by Age | In Poverty | SimplyAnalytics Community Demographics |
| % Purchasing: By Internet: In last 12 months: Food (Groceries) | MRI Consumer Survey |
| % Purchasing: Catalog And Internet, Phone & Mail Shopping - Amount Spent In Total: $800+ | MRI Consumer Survey |
| % Purchasing: Catalog And Internet, Phone & Mail Shopping - Amount Spent In Total: Less than $50 | MRI Consumer Survey |
| % Sadness - Some of the time | EASI Health Care |
| % Sadness- All or most of the time | EASI Health Care |
| % Transportation to Work | Public transportation | SimplyAnalytics Community Demographics |
| % Transportation to Work | Car, truck, or van | SimplyAnalytics Community Demographics |
| % Transportation to Work | Worked at home | SimplyAnalytics Community Demographics |
| % Vehicles Available | No vehicle available | SimplyAnalytics Community Demographics |
| Average annual expenditures (Household average) | SimplyAnalytics Consumer Expenditure Estimates |
| Food | Food at home (Household average) | SimplyAnalytics Consumer Expenditure Estimates |
| Food | Food away from home (Household average) | SimplyAnalytics Consumer Expenditure Estimates |
| Housing | Utilities, fuels, and public services (Household average) | SimplyAnalytics Consumer Expenditure Estimates |
| Median Household Income | SimplyAnalytics Community Demographics |
| Population Density (per square mile) | SimplyAnalytics Community Demographics |
| Transportation | Gasoline, other fuels, and motor oil (Household average) | SimplyAnalytics Community Demographics |

The methodology of this projects fellows the general structure of the GEOG788P’s lab2 methods, which is starting from using ordinary linear regression (OLS) to detect the main driver of gas usage and mental health issues from 2020 to 2016, then apply spatial error and spatial lagged model to find the main drivers of driver of gas usage and mental health issues; finally geographically weighted regression and multiscale geographically weighted regression will applied to the yearly datasets.

地图

描述已自动生成

地图

描述已自动生成

Figure 2. Gas Usage, Sadness, Number of People with Employment, Income increase from 2016 to 2020

**Result & Discussion:**

The project is down by Python at Google Colab platform. At the method/data section in this paper, I mentioned this project contains 24 variables. However, not all variables are used into the models. It is important to detect the multicollinearity issue among explanatory variables using VIF score. Since in this project, there are two dependent variables that I need to find the independent variables correspond to them, therefore, there are two series of VIF scores need to check. Generally speaking, the VIF scores for the explanatory variables should be lower than 5 while the ideal situation should be lower than 3. After series of variable exploring to eliminate the multicollinearity among variables, the select variables for the regression to check drivers of gas usage from 2016 to 2020 are: '% Transportation to Work | Worked at home', '% Transportation to Work | Public transportation', 'Median Household Income', '% Educational Attainment | High school graduate (includes equivalency)', '% Vehicles Available | No vehicle available, 2020', '% Poverty Status by Age | In Poverty'. Their VIF scores from 2016 to 2020 are shown below.

Table 2. VIF Scores for Gas Usage Regression from 2016-2020

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| VIF feature | 2020 | 2019 | 2018 | 2017 | 2016 |
| Transportation to Work | Worked at home | 2.979576 | 3.320314 | 3.920313 | 2.718244 | 3.698483 |
| Transportation to Work | Public transportation. | 2.829596 | 2.676291 | 2.925743 | 2.211910 | 2.855828 |
| Median Household Income | 3.966830 | 4.427177 | 5.146285 | 5.685329 | 4.963443 |
| Educational Attainment | High school graduate | 3.714985 | 2.969737 | 3.041029 | 3.222603 | 3.075490 |
| Vehicles Available | No vehicle available | 4.936511 | 3.199693 | 3.425987 | 6.358287 | 3.581328 |
| Poverty Status by Age | In Poverty | 4.652689 | 3.099841 | 3.342227 | 3.462155 | 3.421208 |

The select variables for the regression to check drivers of mental health- sadness issues from 2016 to 2020 are: ' '%Transportation to Work | Public transportation', '% Transportation to Work | Worked at home', 'Population Density (per square mile)' (not available after 2018), 'Educational Attainment | High school graduate (includes equivalency)', 'Vehicles Available | No vehicle available', '% Employment Status | In labor force', '% Poverty Status by Age | In Poverty',. Their VIF scores from 2016 to 2020 are shown below.

Table 3. VIF Scores for Mental Health- Sadness Regression from 2016-2020

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| VIF feature | 2020 | 2019 | 2018 | 2017 | 2016 |
| Transportation to Work | Public transportation' | 2.927120 | 3.169669 | 2.718244 | 3.052942 | 3.137602 |
| Transportation to Work | Worked at home' | 2.394444 | 2.602651 | 2.211910 | 2.701968 | 2.592157 |
| 'Population Density (per square mile) | 3.146627 | 3.273070 | NA | NA | NA |
| Educational Attainment | High school graduate | 4.062867 | 4.433431 | 5.685329 | 4.263047 | 4.418373 |
| Vehicles Available | No vehicle available | 3.040242 | 3.390436 | 3.222603 | 3.518269 | 3.589040 |
| Employment Status | In labor force | 7.299001 | 8.221266 | 6.358287 | 7.826083 | 7.734239 |
| Poverty Status by Age | In Poverty | 2.804152 | 3.129929 | 3.462155 | 3.406053 | 3.414812 |

After the VIF score has been confirmed, the ordinary linear regressions, spatial error models and spatial lagged models are calculated for different dependent variables from 2016-2020, the results are shown below:

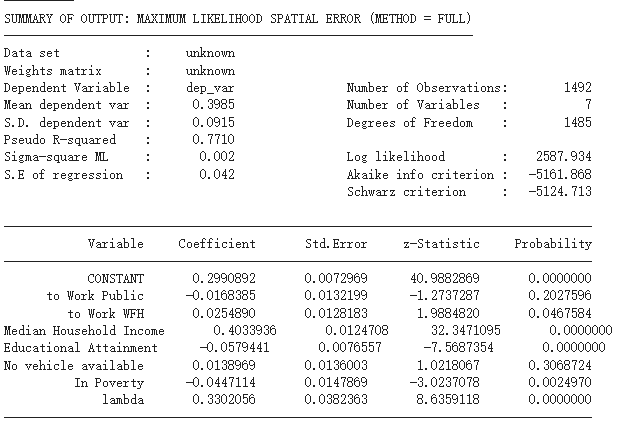
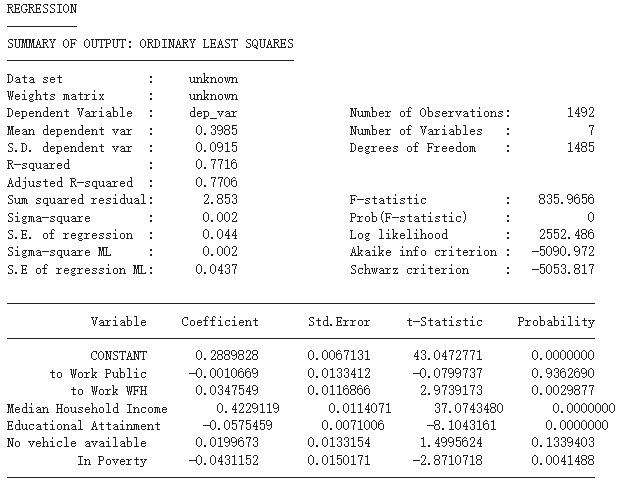


Figure 3. The OLS, Spatial Error, Spatial Lag model results of 2020’s gas usage

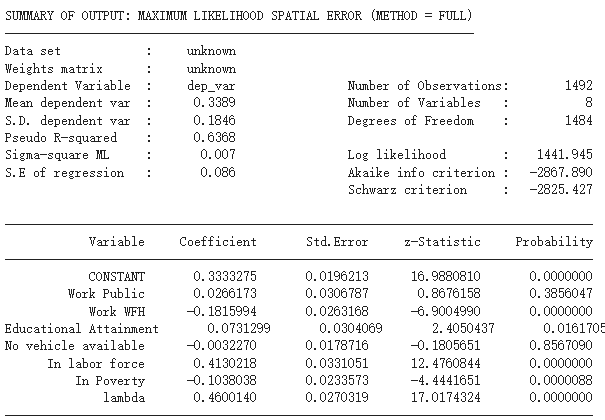
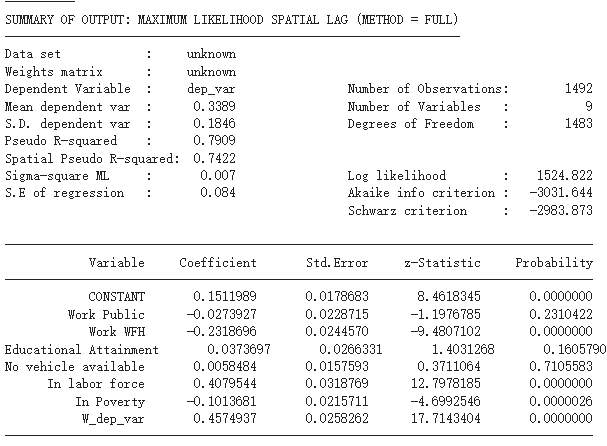
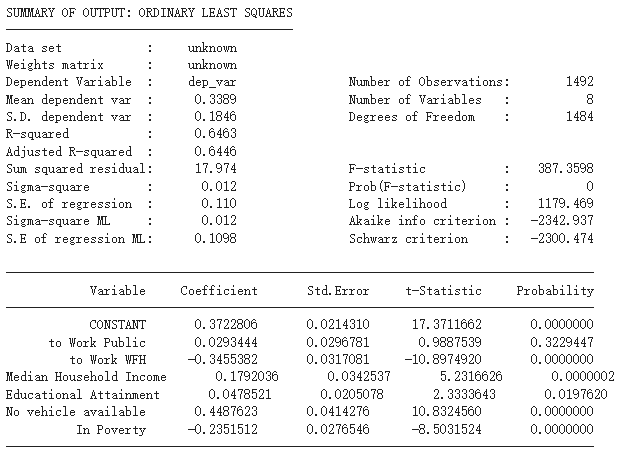


Figure 4. The OLS, Spatial Error, Spatial Lag model results of 2020’s mental health

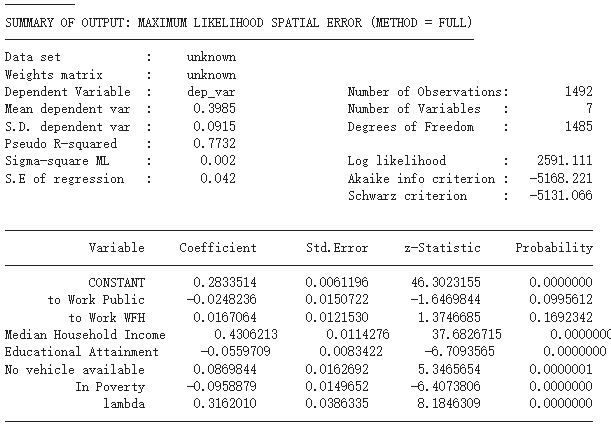
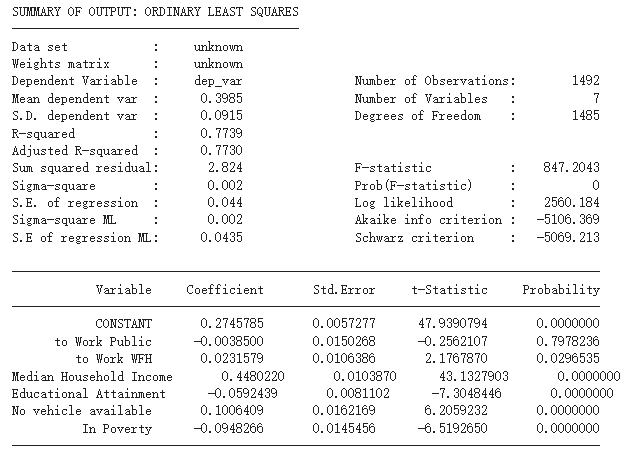


Figure 5. The OLS, Spatial Error, Spatial Lag model results of 2019’s gas usage

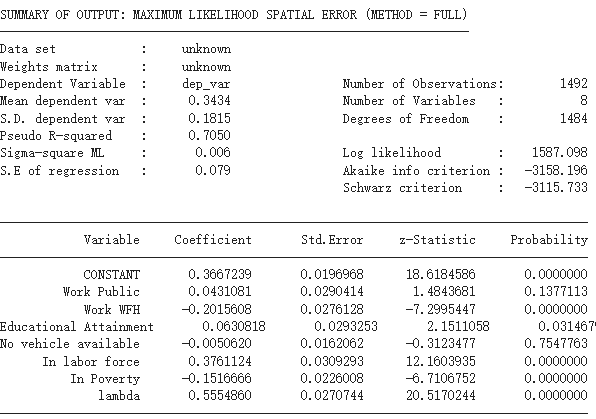
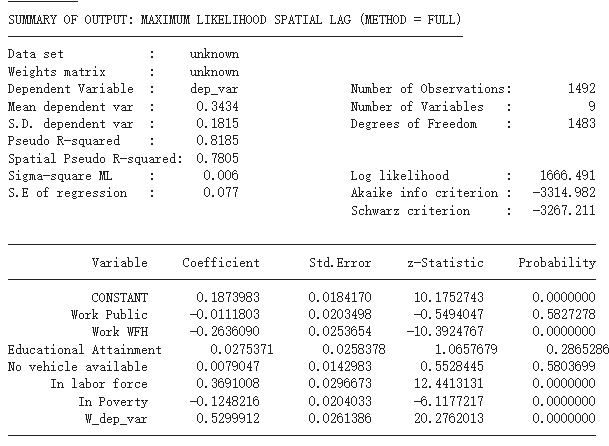
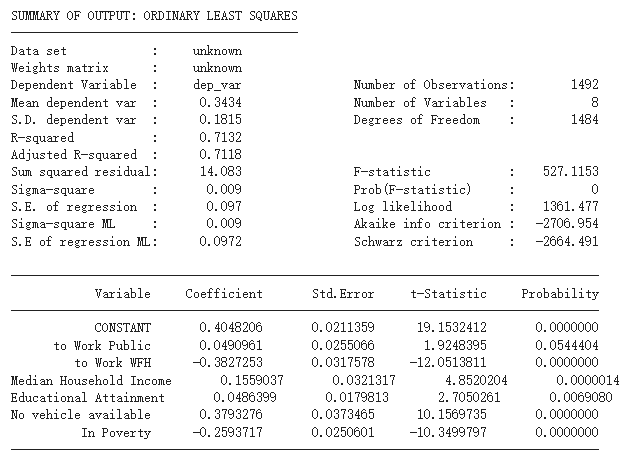


Figure 6. The OLS, Spatial Error, Spatial Lag model results of 2019’s mental health

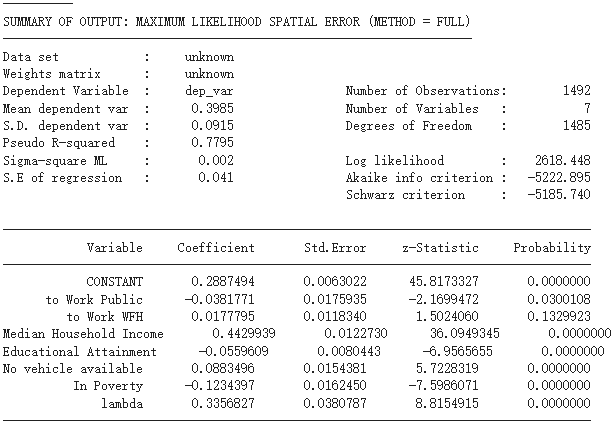
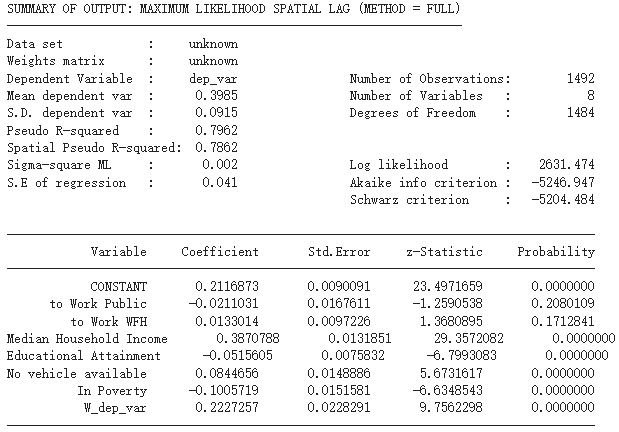
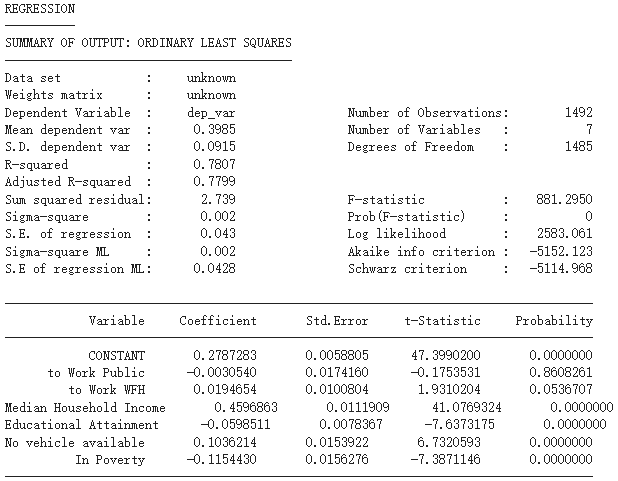


Figure 7. The OLS, Spatial Error, Spatial Lag model results of 2018’s gas usage

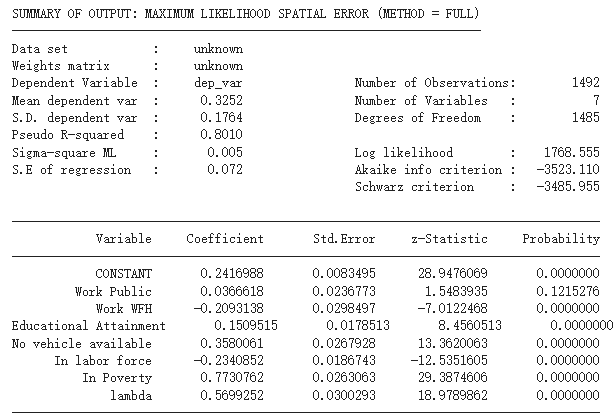
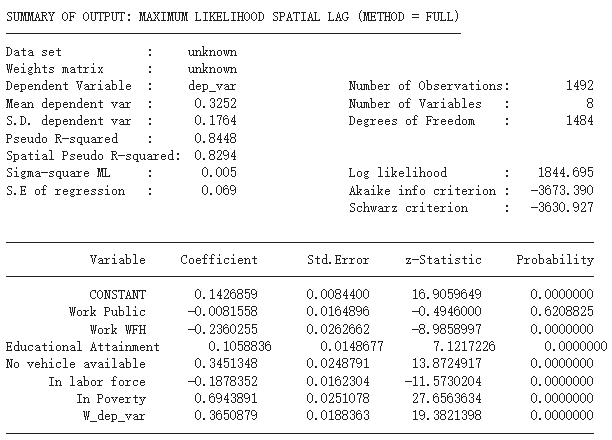
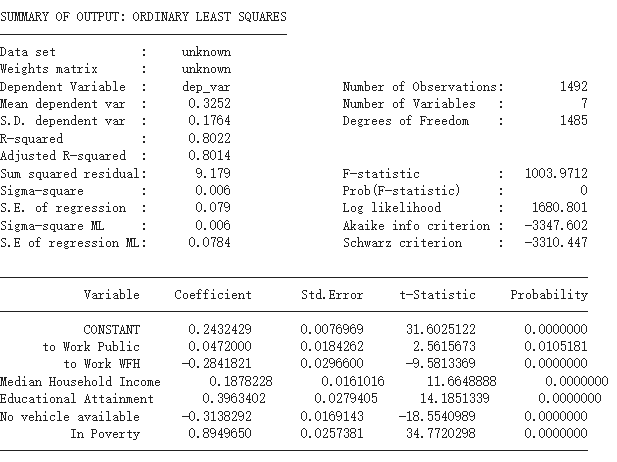


Figure 8. The OLS, Spatial Error, Spatial Lag model results of 2018’s mental health

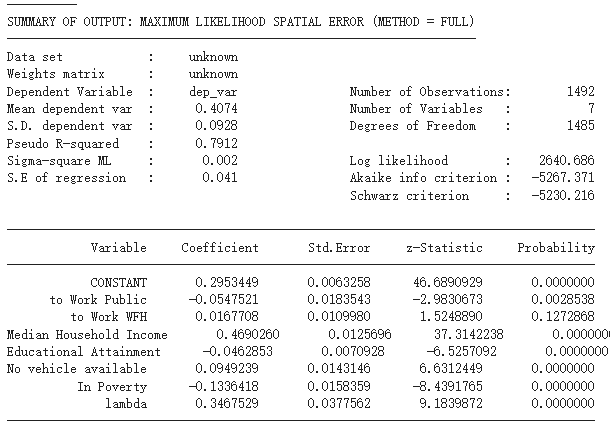
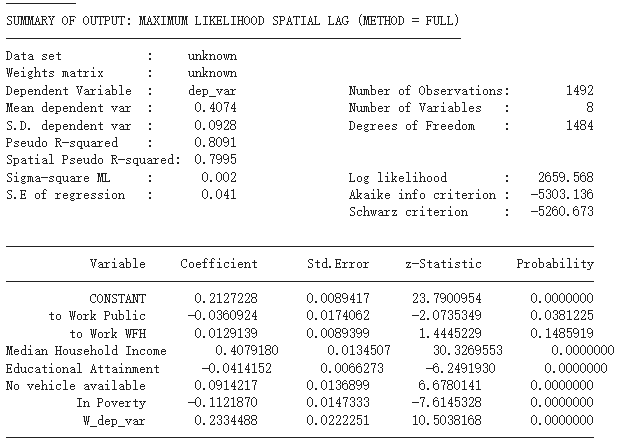
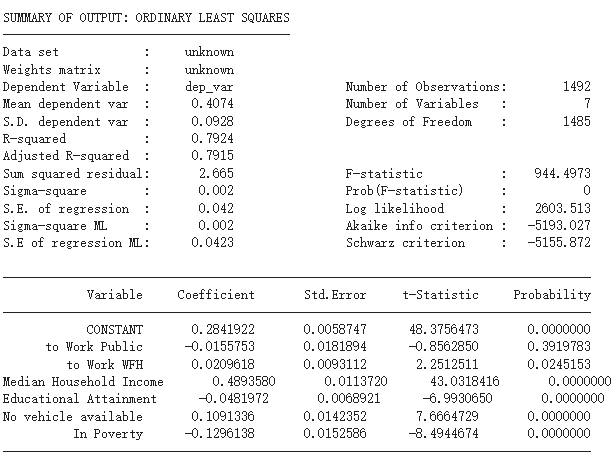


Figure 9. The OLS, Spatial Error, Spatial Lag model results of 2017’s gas usage

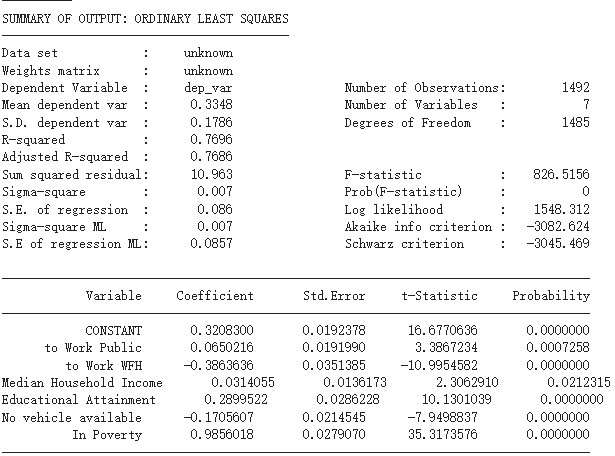
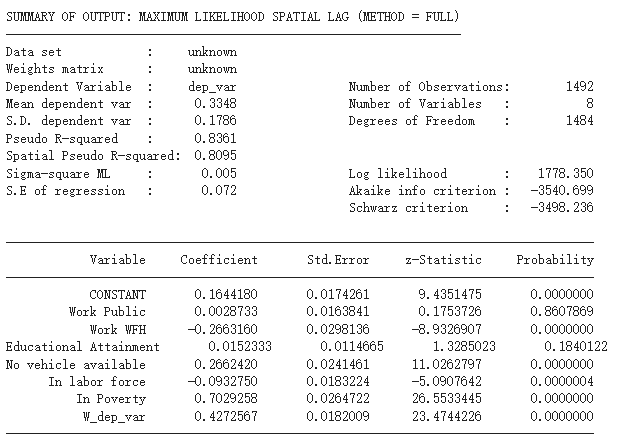
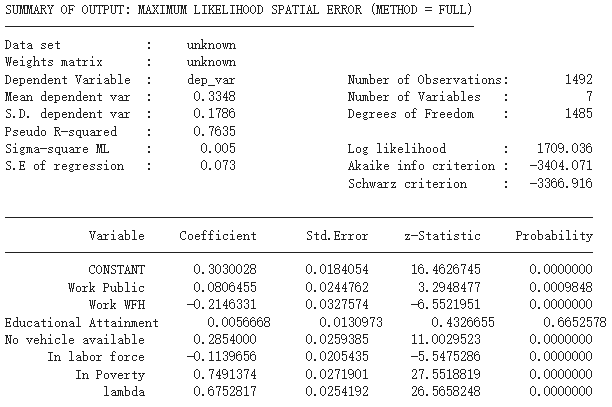


Figure 10. The OLS, Spatial Error, Spatial Lag model results of 2017’s mental health

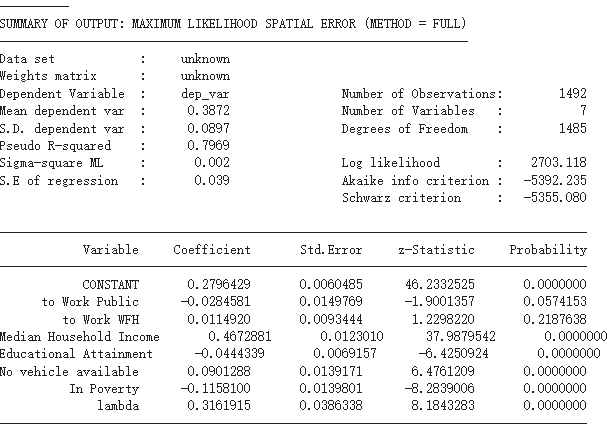
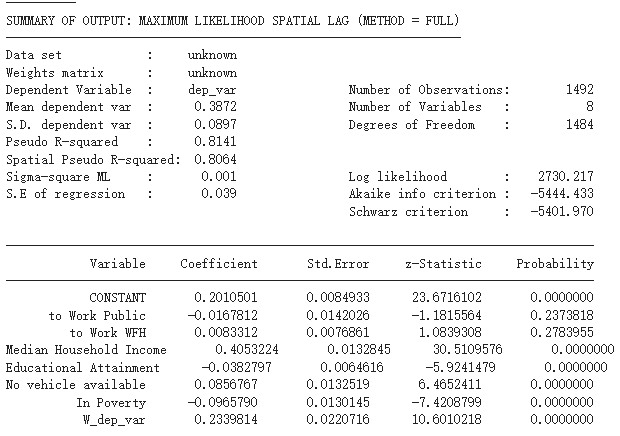
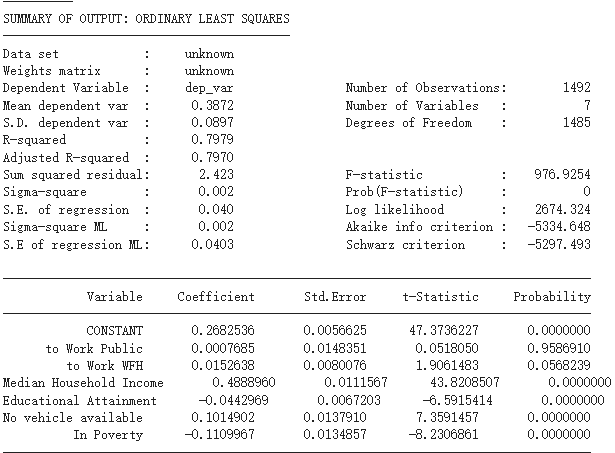


Figure 11. The OLS, Spatial Error, Spatial Lag model results of 2016’s gas usage

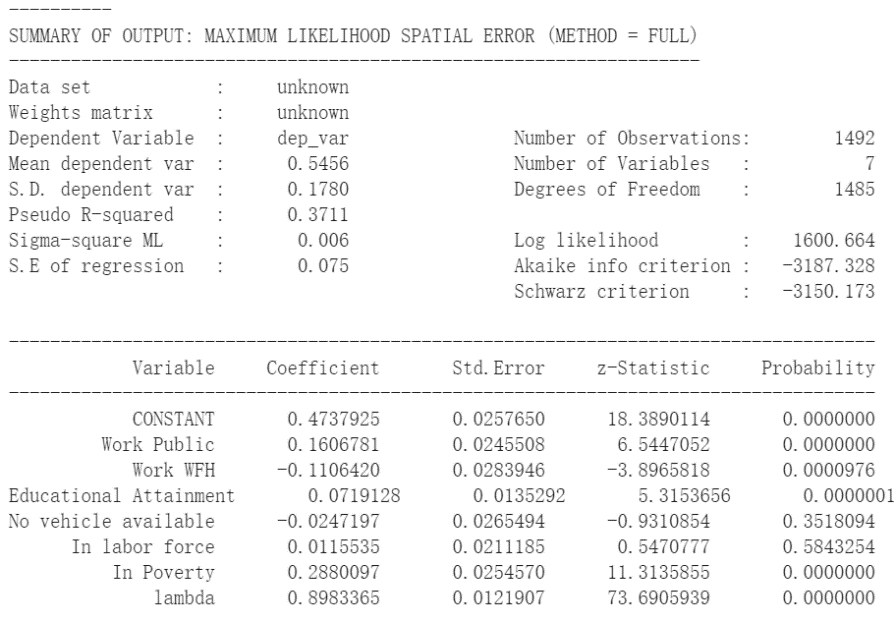
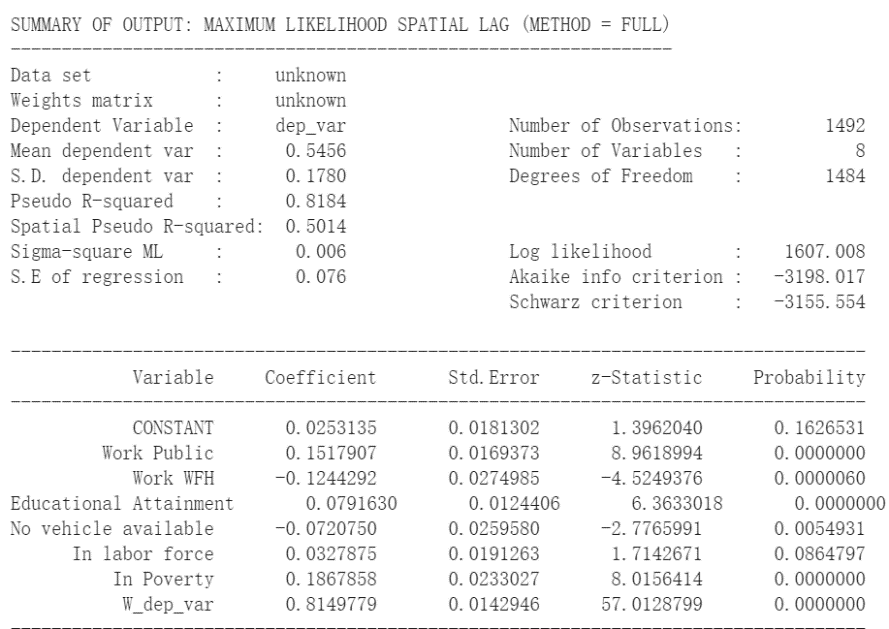
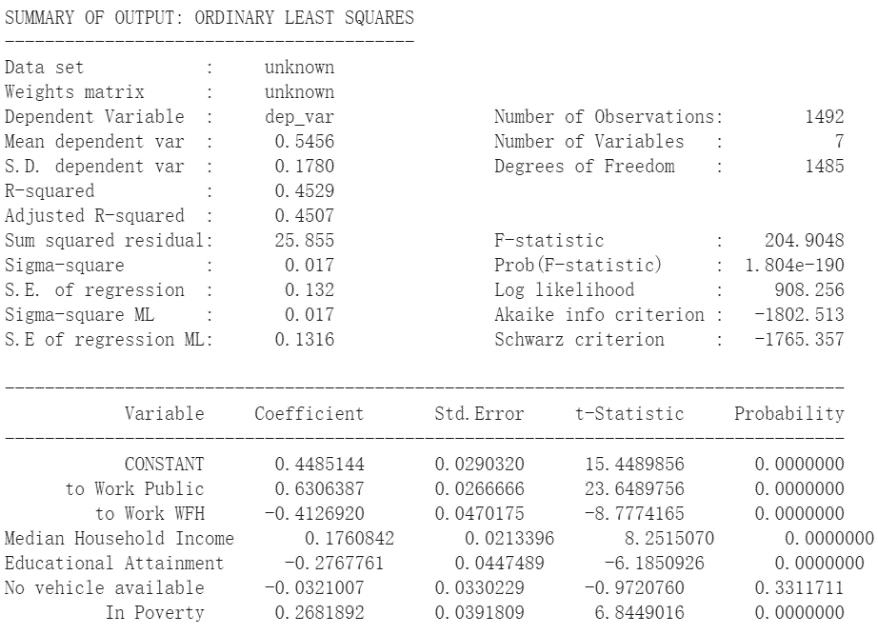


Figure 12. The OLS, Spatial Error, Spatial Lag model results of 2016’s mental health

From the result shown above, it is clear to notice that the major driver caused the increase of gas from 2016 to 2020 is the median household income, the estimates are from 0.44~0.40. While the major driver for increase on sadness from 2016 to 2020 in Maryland is poverty, the estimates are from 0.75~0.98. Since the household income is deeply connected with the poverty status, it is safe to say that the higher household income, less mental sadness while higher usage on Gas. The result from GWR and MGWR are the same in terms of the major driver to gas usage and mental health conditions is the average household income. From the Lagrange multiplier test result, the spatial lag model has better performance that the spatial error model. The results are shown above. The GWR and MGWR results are shown below, generally speaking, the MGWR results overperformed the GWR result, in terms of find the suitable BW(band width ) for each variables and the higher R2 among all GWR and MGWR results.

However, since the above conclusion exists through out the entire study years, it means the Pandemic does not majorly transfer the people’s lifestyle. But the other drivers could affect the gas usage and mental health conditions differently before the pandemic and during the pandemic. Using 2020 and 2016’s regression result as an example. The Worked at home, in 2020 has negative influence of gas usage while in 2016 it has positive influence on gas usage, similarly, the public transportation has high positive influence on gas usage in 2020 compare with 2016. During this study, one of the original goals is to looking into the relationship between people who drive to work and gas usage. However, during the VIF tests, variable represent people who drive to work failed the test. But the regression result shows that not only people who driver to work could have positive relationship with gas usage, but also people who use public transportation also leads to higher gas usage. One possible explanation of this scenario could be people who use public transportation tools would also need to driver to the public transportation stations by cars or motorcycles. Another interest thing to point out in this regression is the families with no vehicles available also has positive relationship with the gas usage, it could being even the familiar does not have cars, they cold also spend on taxis, carpooling, or other methods that could eventually involve with car’s gasoline usage.

The sadness conditions also has similar trends as gas usage, where people who has jobs (in labor force) would have higher negative influence on not get sadness during the pandemic compare with 2016. Similarly, people who choose to work from home does not have such higher positive influence on get worse mental health conditions compare with 2016, which it could be everyone were forced to stay at home and it increased the general population of people who choose to work at home. One interesting finding at the mental health regression is the high school graduates would have higher chance of have mental health issues, through out 2016-2020’s data. In addition to that, the population density also affect the mental health condition positively, meanly higher population density areas would have higher chance of having sadness some of time or all time. It could because that the high density population areas such as cities could lead to mental health issues or people lives in the city would have higher chance to report their mental health conditions.

However, this project also has limits. First of all, selected variables are having multicollinearity issues, especially driving to work, online purchasing, as well as food expenditures. The original plan of this project is to find out how the driving to work could effect the gas usage before and during the pandemic. Same issue for the online purchase, studies have shown that the online purchasing are becoming more often and even a dominant ways for household to get their grocery during the pandemic, there is a huge increase from 2016 to 2020 of the amount of people who use online shopping. But it cannot be put into the regression and measure the contribution of it. Another issues of his project due to the short temporal coverage of this dataset, it cannot conduct a time series regression but only spatial regression.

Figure 12. The GWR and MGWR results of 2020’s gas usage

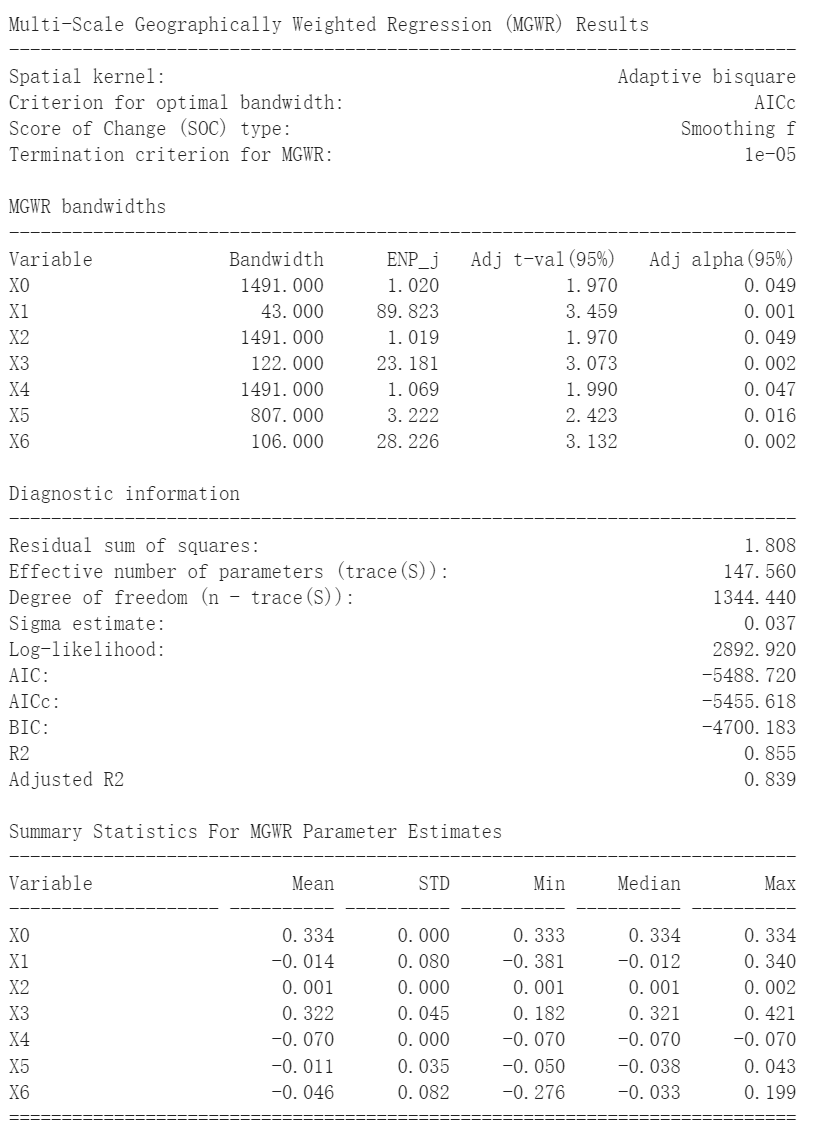
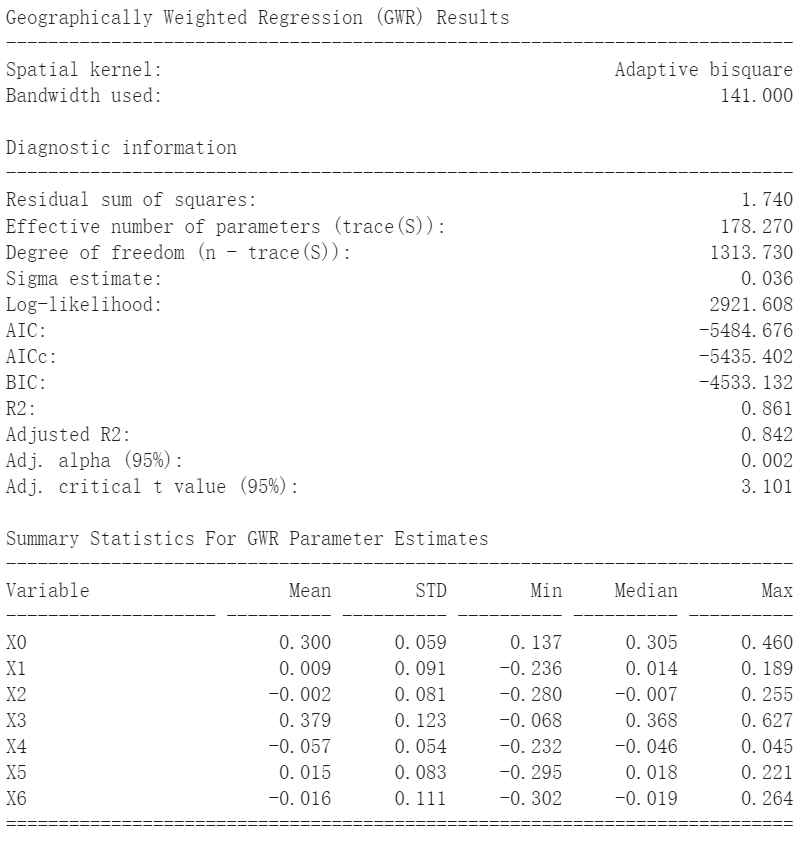


Figure 13. The GWR and MGWR results of 2019’s gas usage

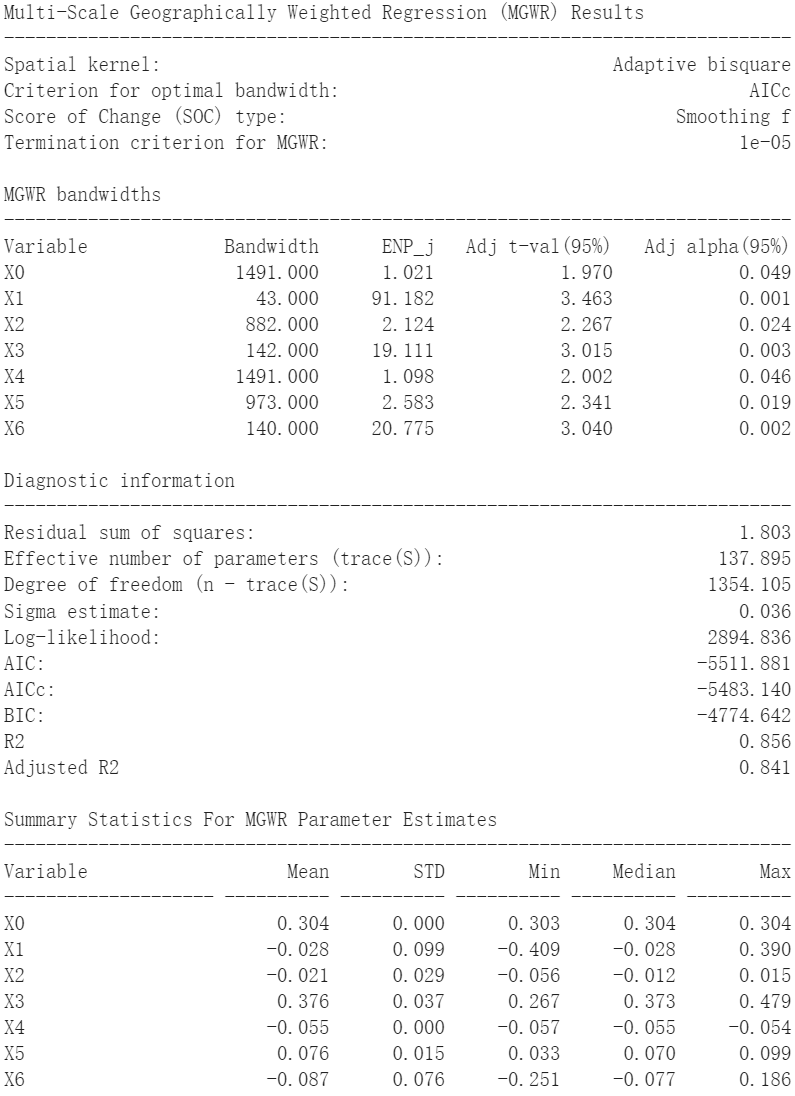
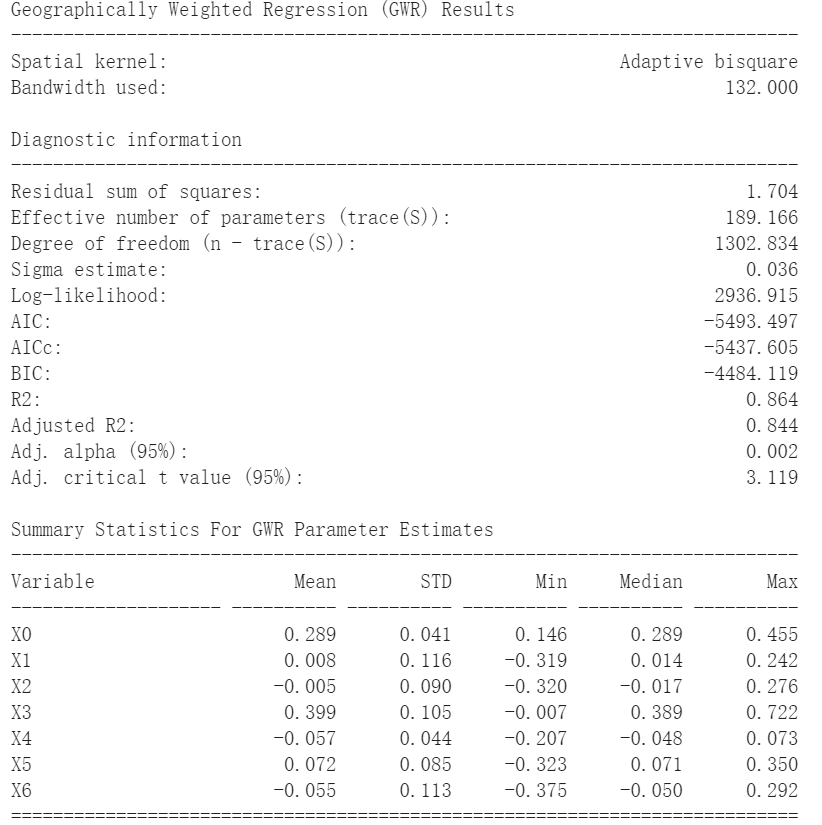


Figure 14. The GWR and MGWR results of 2018’s gas usage

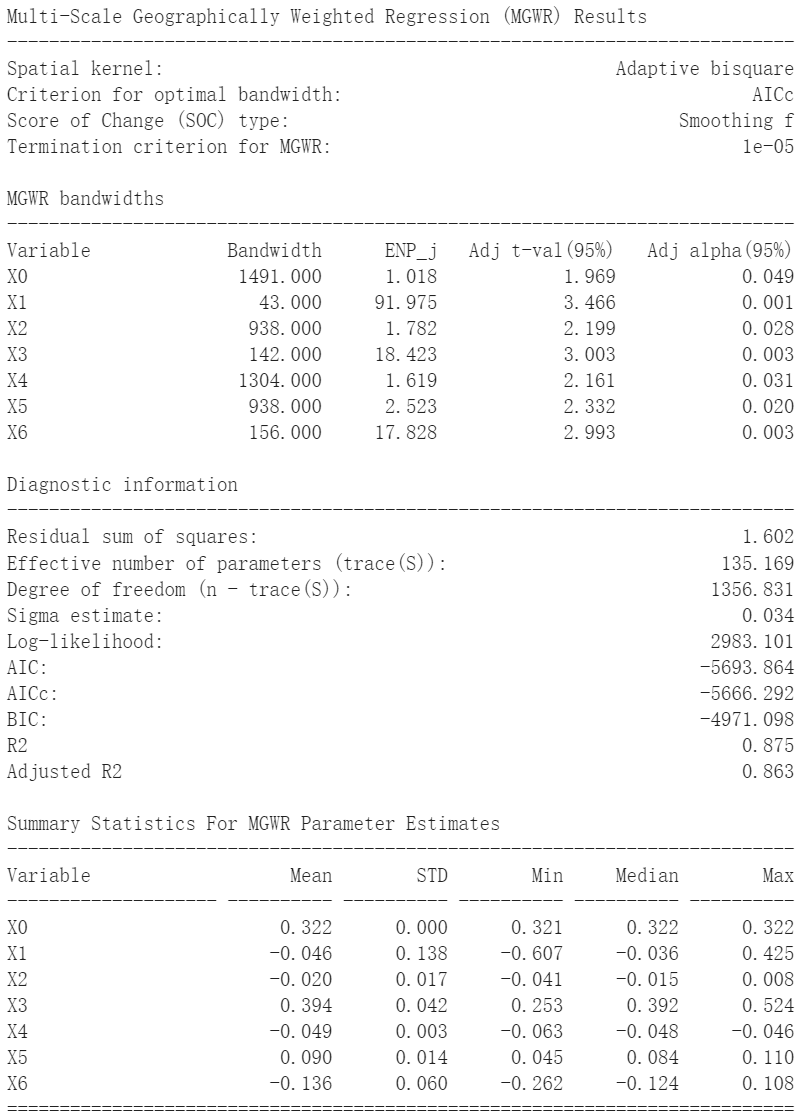
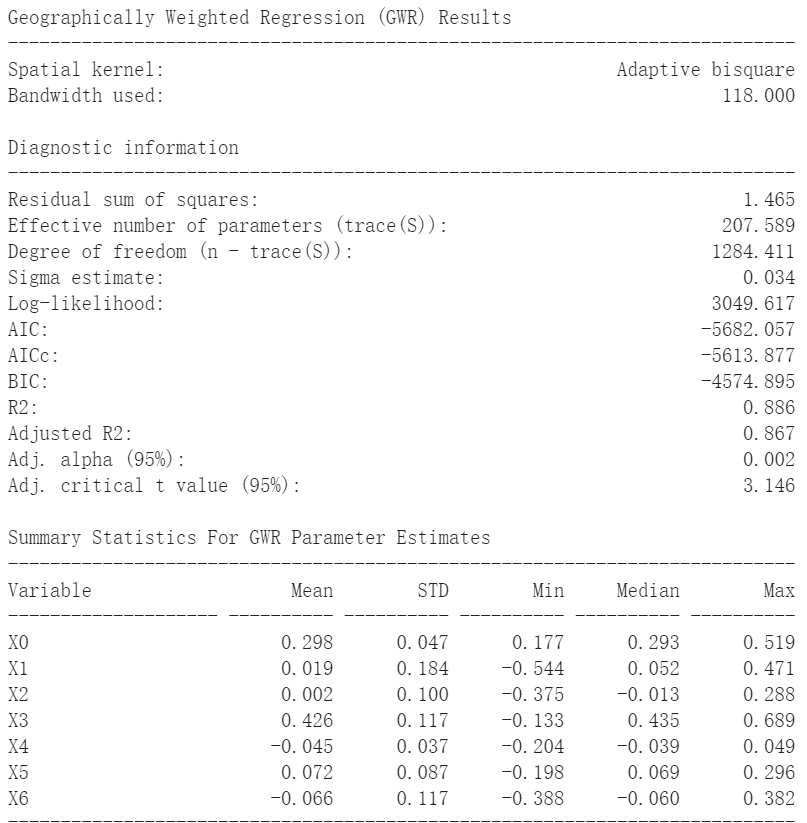
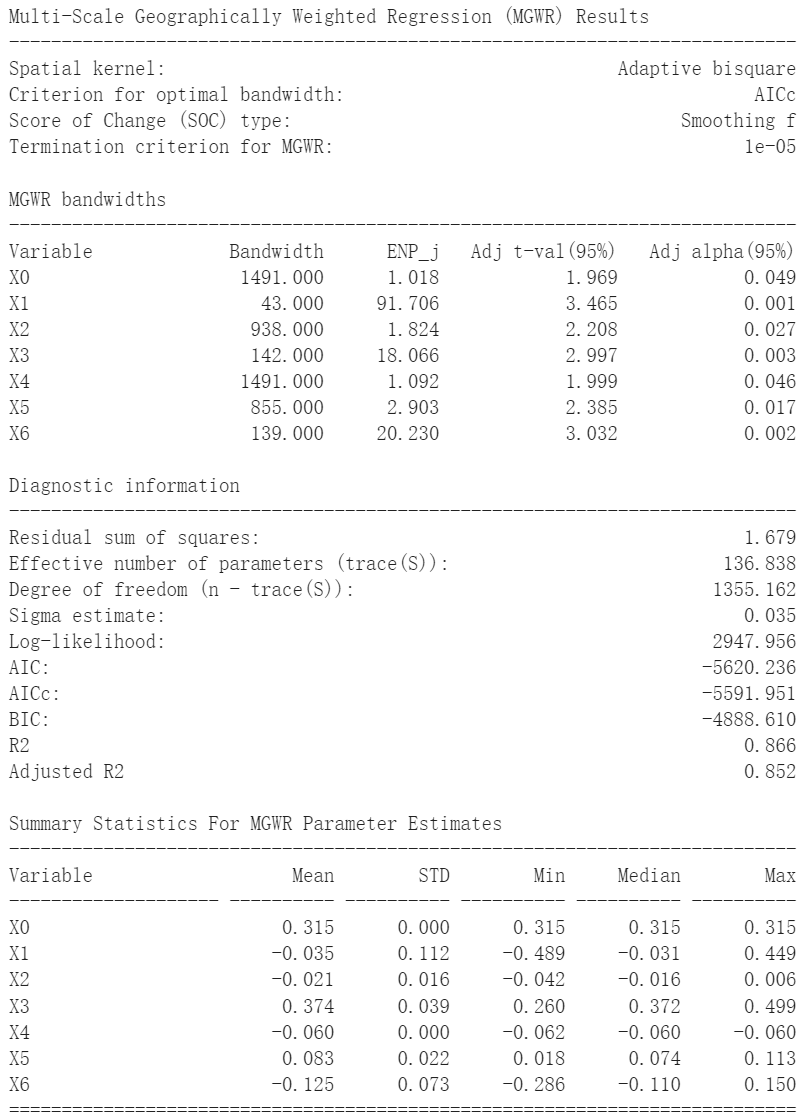
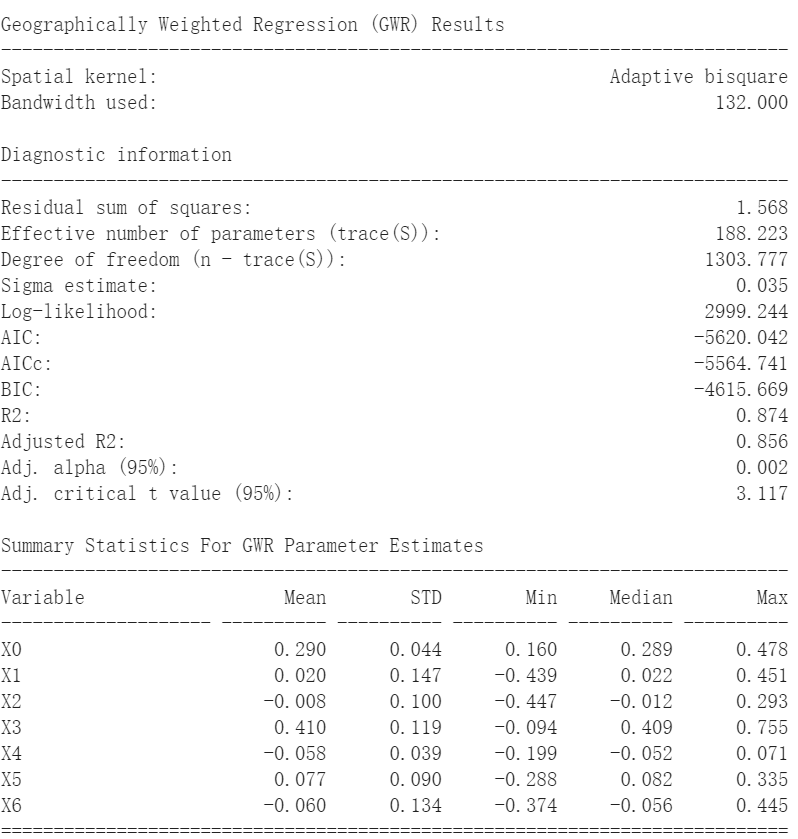


Figure 15. The GWR and MGWR results of 2017’s gas usage

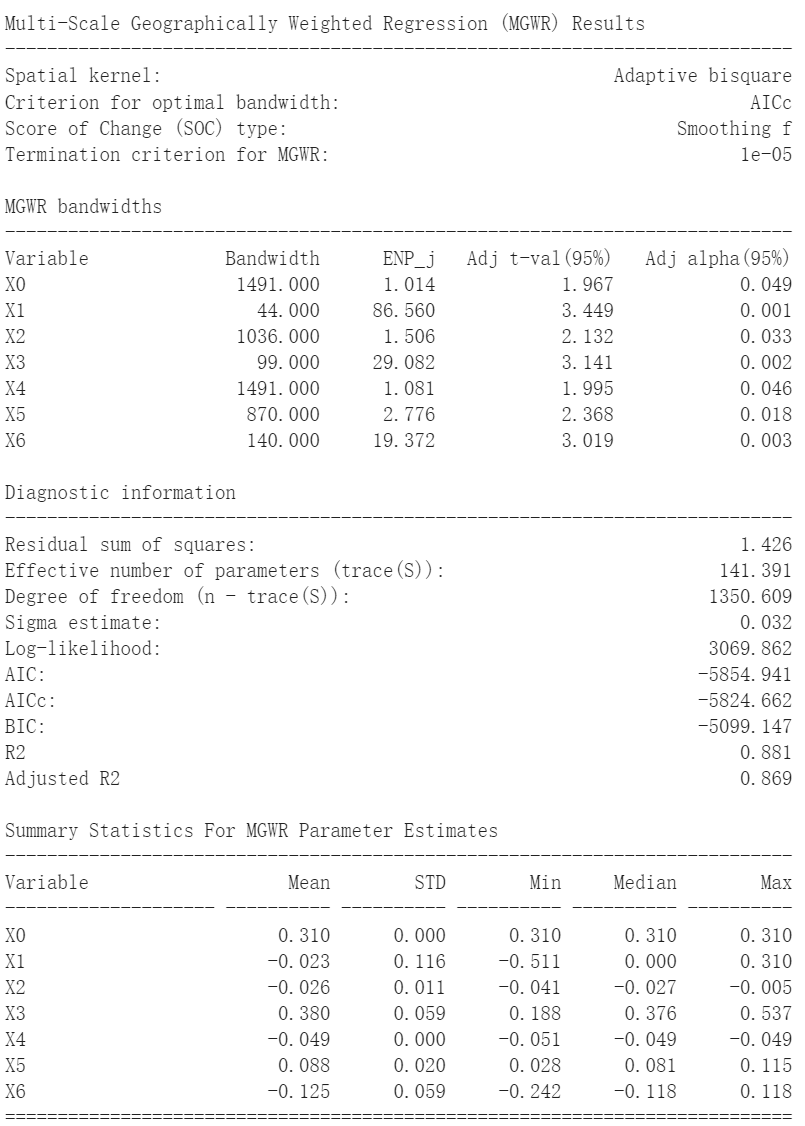
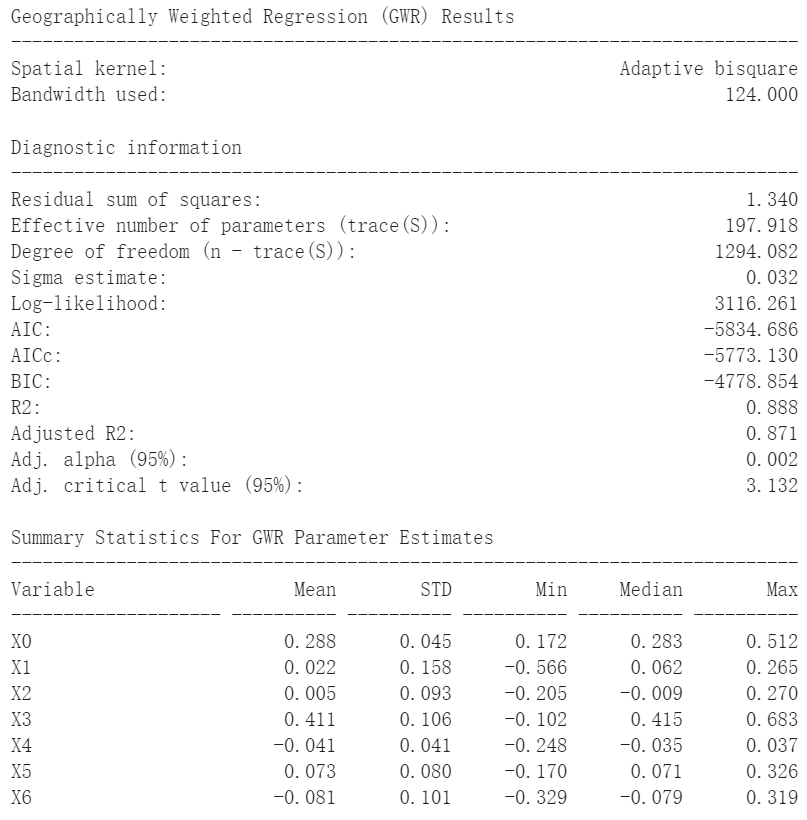


Figure 16. The GWR and MGWR results of 2016’s gas usage

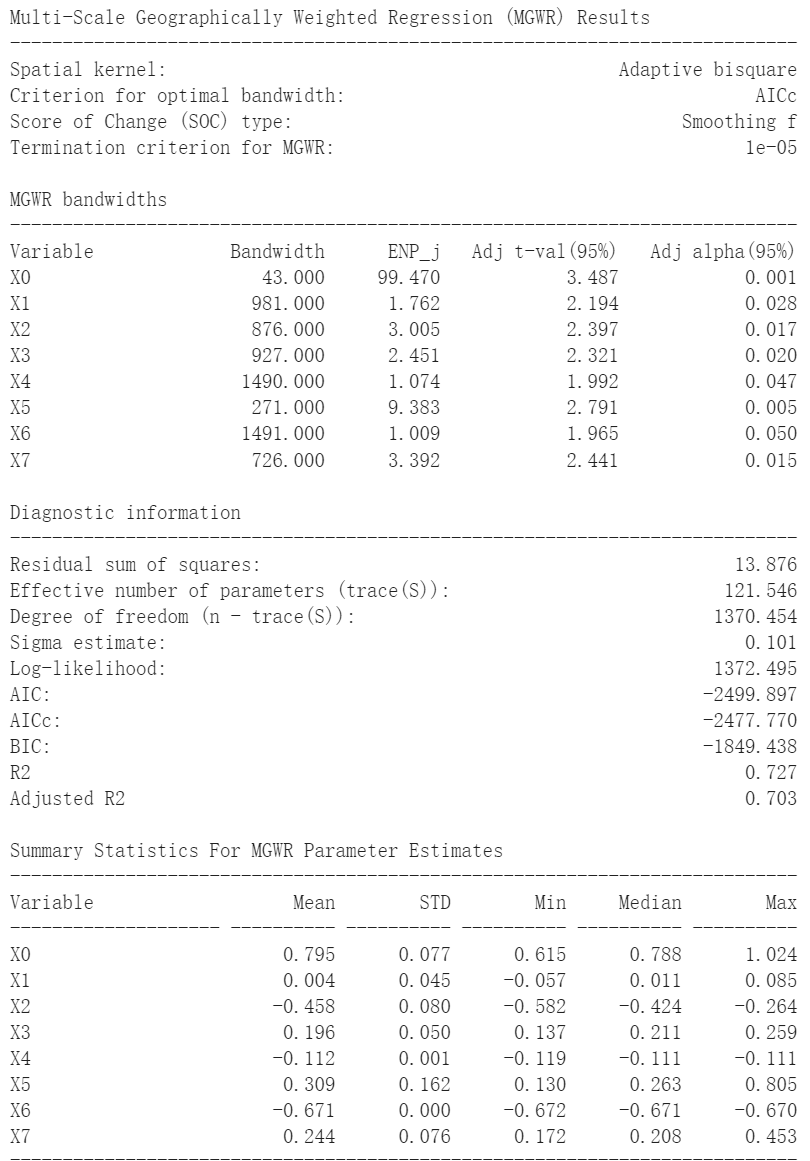
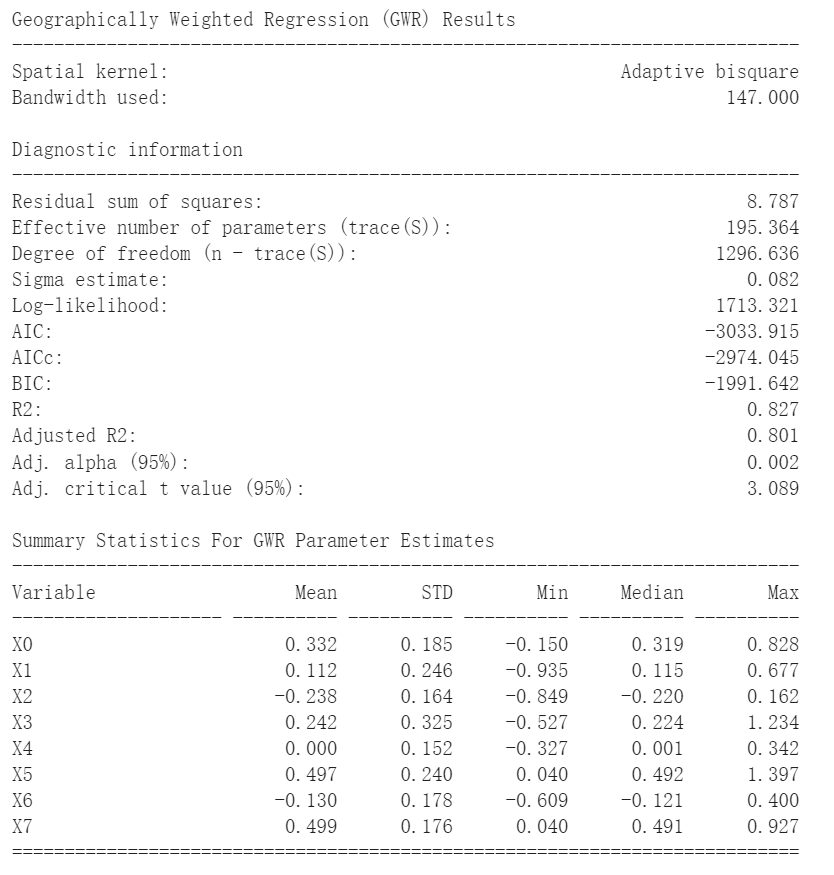


Figure 17. The GWR and MGWR results of 2020’s mental health

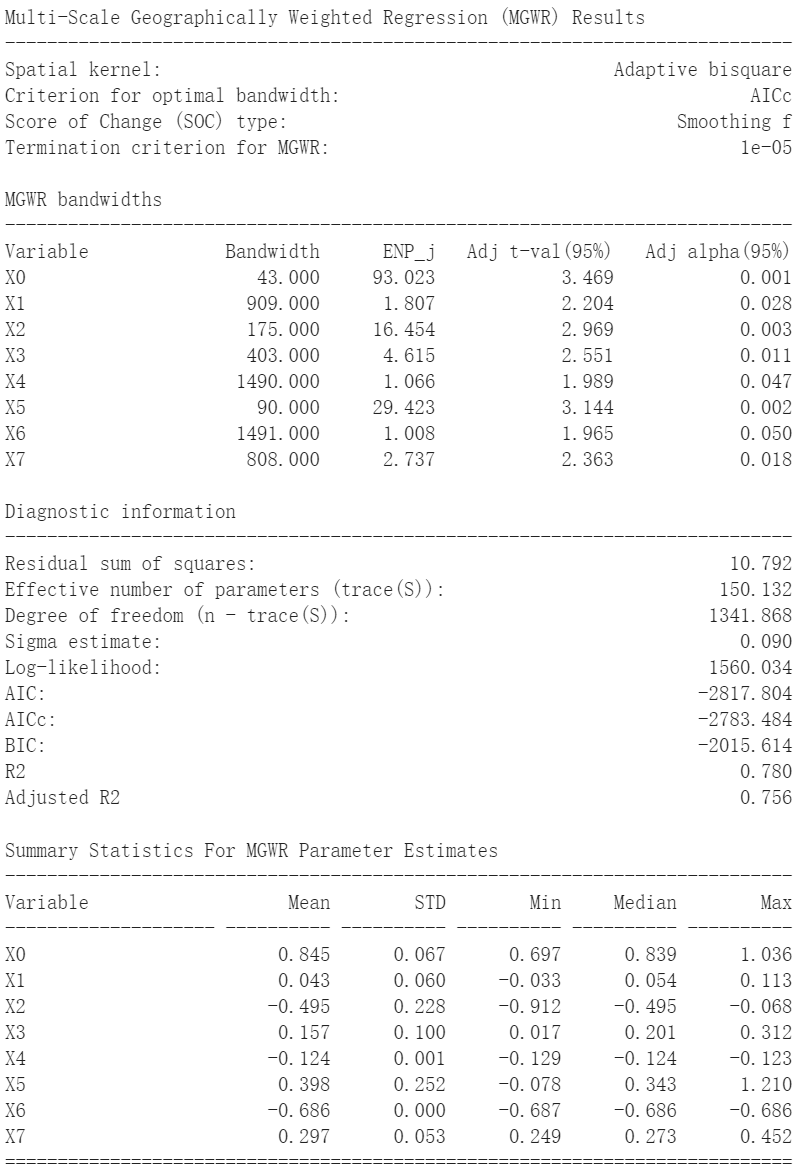
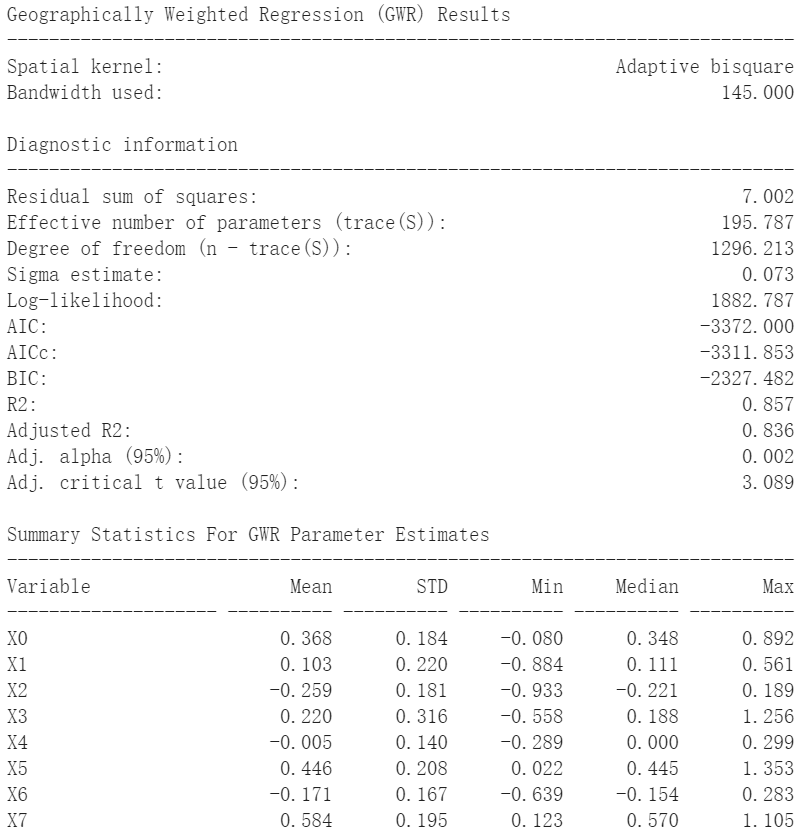


Figure 18. The GWR and MGWR results of 2019’s mental health

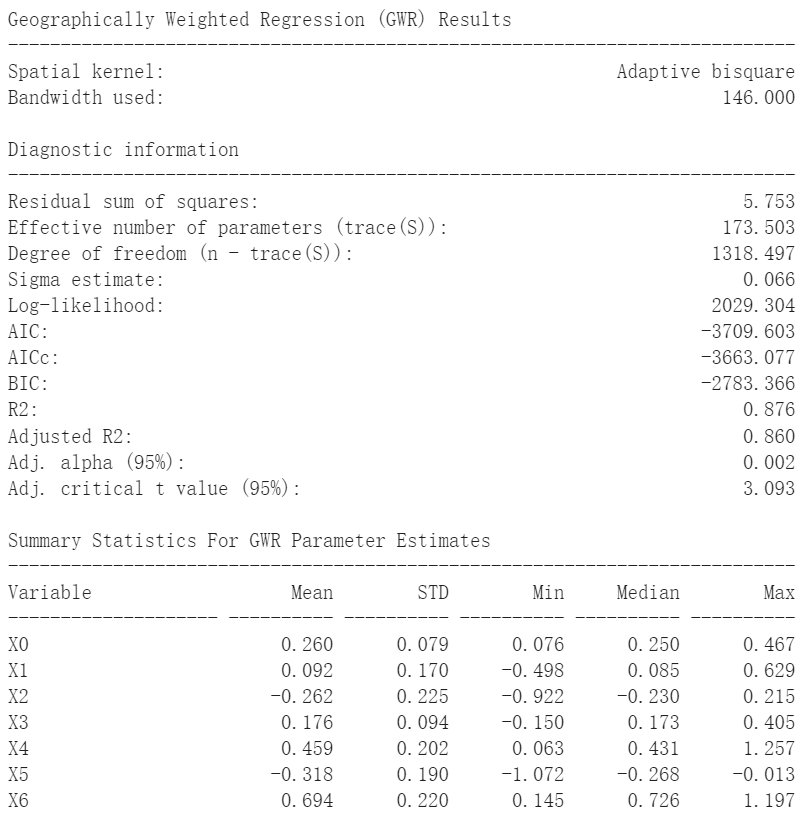
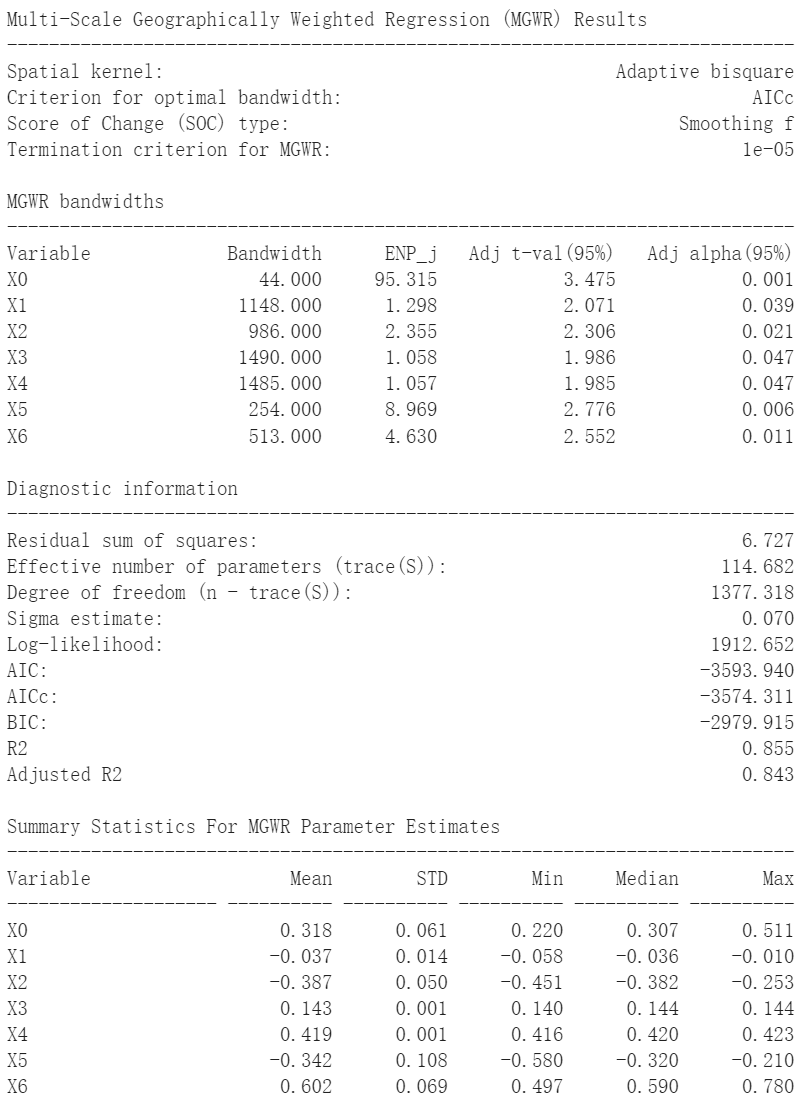


Figure 19. The GWR and MGWR results of 2018’s mental health

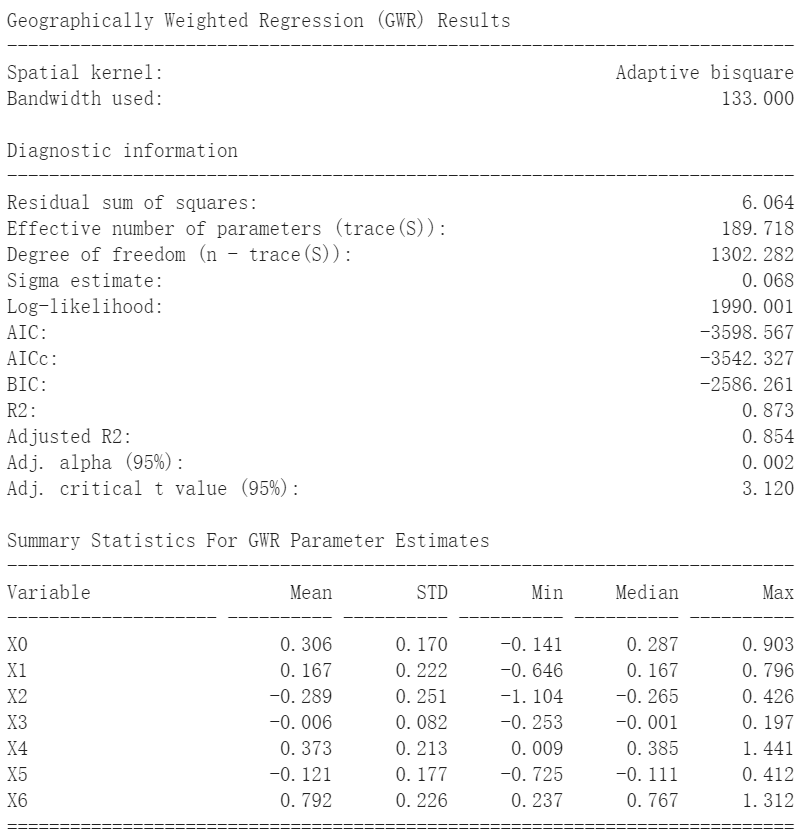
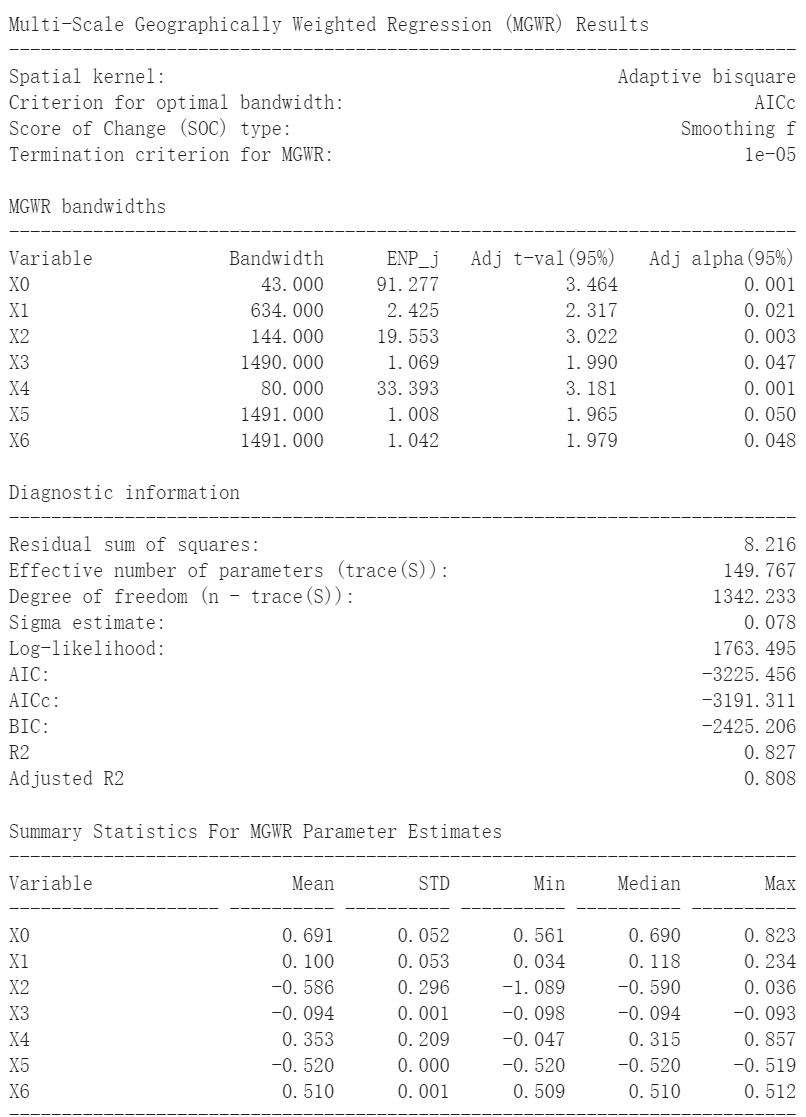


Figure 20. The GWR and MGWR results of 2017’s mental health

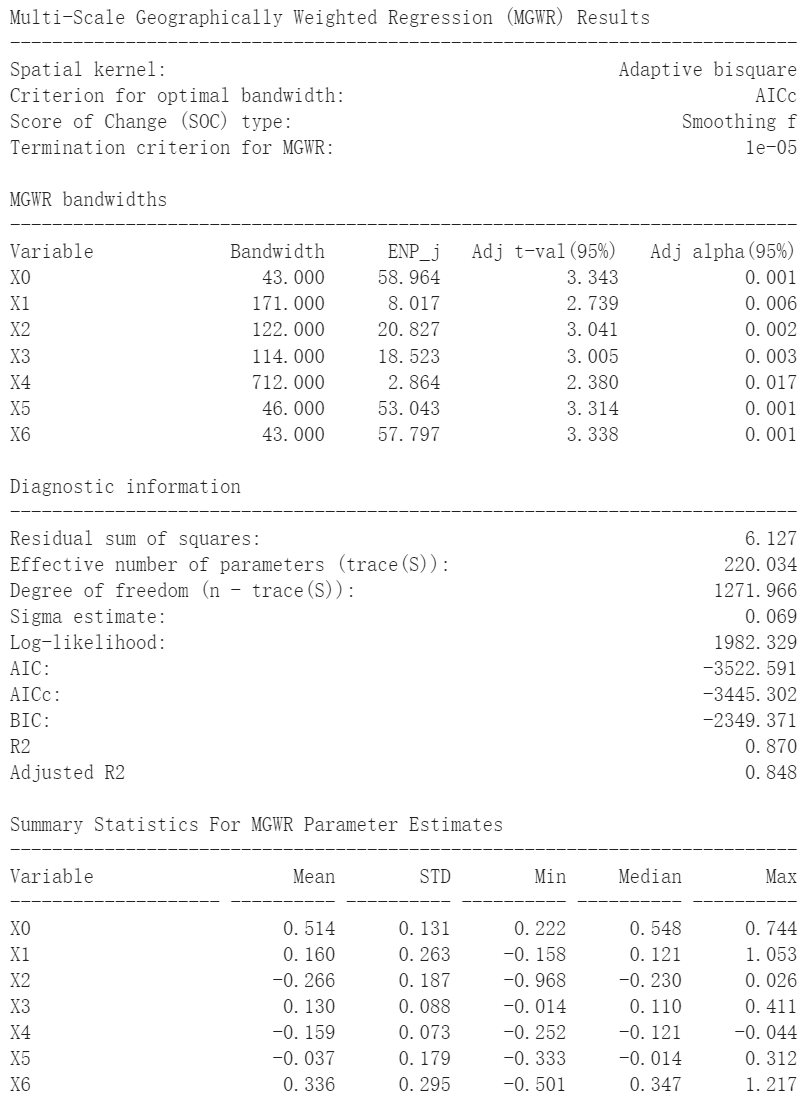
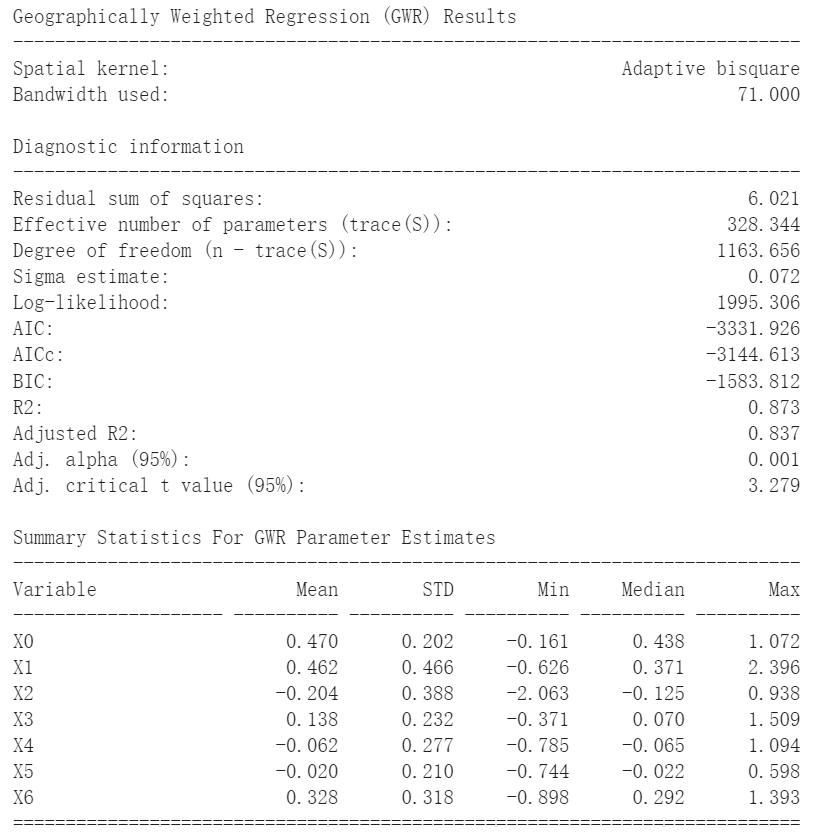


Figure 21. The GWR and MGWR results of 2016’s mental health

**Bibliography:**

Abouk, R., & Heydari, B. (2020). The immediate effect of COVID-19 policies on social distancing behavior in the United States. <https://doi.org/10.1101/2020.04.07.20057356>

Bauer, L., Broady, K. E., Edelberg, W., & O’Donnell, J. (2020, Sept 18). Ten facts about COVID-19 and the US economy. Brookings. <https://www.brookings.edu/research/ten-facts-about-covid-19-and-the-u-s-economy/>

Digital Commerce 360. (2021, Jan 29). US ecommerce grows 44.0% in 2020. <https://www.digitalcommerce360.com/article/us-ecommerce-sales/>

Employment–population ratio 54.6 percent in June 2020, down from 61.1 percent in February 2020 : The economics daily: US Bureau of Labor Statistics. (2020, Jul 8). US Bureau of Labor Statistics. <https://www.bls.gov/opub/ted/2020/employment-population-ratio-54-point-6-percent-in-june-2020-down-from-61-point-1-percent-in-february-2020.htm>

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers*, *107*(6), 1247-1265. <https://doi.org/10.1080/24694452.2017.1352480>

Güngör, O., Ertugrul, H. M., & Soytas, U. (2020). The impact of COVID-19 outbreak on Turkish gasoline consumption. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3611788>

Haining, R. (2013). Spatial data and statistical methods: A chronological overview. *Handbook of Regional Science*, 1277-1294. <https://doi.org/10.1007/978-3-642-23430-9_71>

Holingue, C., Badillo-Goicoechea, E., Riehm, K. E., Veldhuis, C. B., Thrul, J., Johnson, R. M., Fallin, M. D., Kreuter, F., Stuart, E. A., & Kalb, L. G. (2020). Mental distress during the COVID-19 pandemic among US adults without a pre-existing mental health condition: Findings from American trend panel survey. *Preventive Medicine*, *139*, 106231. <https://doi.org/10.1016/j.ypmed.2020.106231>

Liu, C. H., Zhang, E., Wong, G. T., Hyun, S., & Hahm, H. “. (2020). Factors associated with depression, anxiety, and PTSD symptomatology during the COVID-19 pandemic: Clinical implications for U.S. young adult mental health. *Psychiatry Research*, *290*, 113172. <https://doi.org/10.1016/j.psychres.2020.113172>

Nyga-Łukaszewska, H., & Aruga, K. (2020). Energy prices and COVID-immunity: The case of crude oil and natural gas prices in the US and Japan. *Energies*, *13*(23), 6300. <https://doi.org/10.3390/en13236300>

Oshan, T., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, A. S. (2018). Mgwr: A python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. <https://doi.org/10.31219/osf.io/bphw9>

Smith, L. V., Tarui, N., & Yamagata, T. (2021). Assessing the impact of COVID-19 on global fossil fuel consumption and CO2 emissions. *Energy Economics*, *97*, 105170. <https://doi.org/10.1016/j.eneco.2021.105170>

Wang, X., Hegde, S., Son, C., Keller, B., Smith, A., & Sasangohar, F. (2020). Investigating mental health of US college students during the COVID-19 pandemic: Cross-sectional survey study. *Journal of Medical Internet Research*, *22*(9), e22817. <https://doi.org/10.2196/22817>