**An empirical analysis of gas consumption to generate Electricity in the United States**

**Econometrics Project (Phase II)**

**Mohammad Saber**

**(MUID: 005852837)**

****

**Fall 2014**

**Table of Content**

|  |  |  |
| --- | --- | --- |
| **I** | **Introduction** | 6 |
| **II** | **Background and Literature review** | 8 |
| **III** | **Modeling Process and Data Issues** | 10 |
| **IV** | **Diagnostic tests and correction**  **Normality**  **Heteroscedasticity**  **Autocorrelation**  **Multicollinearity** | 14 |
| **V** | **Testing Hypotheses on the Regression Coefficients**  **Structural break test and correction**  **Hypotheses testing** | 16 |
| **VI** | **Comments and suggestions of Dr. Nourzad on phase-I and my reply** | 24 |
|  |  |  |
|  |  |  |
|  | **References** | 27 |
|  | **Appendix** | 29 |

**List of Figures**

|  |  |  |
| --- | --- | --- |
| **Figure 1** | Fossil fuel consumption for generation electricity in US | **7** |
| **Figure 2** | Recursive coefficients test | 29 |
| **Figure 3** | Recursive residual test | 30 |
| **Figure 4** | CUSUM test | 30 |
| **Figure 5** | CUSUM of squares test | 31 |
| **Figure 6** | Histogram Residuals | 31 |
| **Figure 7** | Residuals plot (Population vs Residual squared) | 32 |
| **Figure 8** | Residuals plot (Gross Domestic Product vs Residual squared) | 32 |
| **Figure 9** | Residuals plot (Unemployment rate vs Residual squared) | 33 |
| **Figure 10** | Residuals plot (Coal consumption vs Residual squared) | 33 |
| **Figure 11** | Residuals plot (Oil price vs Residual squared) | 34 |
| **Figure 12** | Residuals plot (Natural gas price vs Residual squared) | 34 |
| **Figure 13** | Residuals plot (Consumer Price Index vs Residual squared) | 35 |
| **Figure 14** | Residuals plot | 35 |

**List of Tables**

|  |  |  |
| --- | --- | --- |
| **Table 1** | Estimation of our linear regression model (before applying any test and correction) | 36 |
| **Table 2** | Estimation of linear regression model with Z1 (Mackinnon-White-Davidson test) | 37 |
| **Table 3** | Estimation of double logarithmic model with Z2 (Mackinnon-White-Davidson test) | 38 |
| **Table 4** | White’s test with cross terms | 38 |
| **Table 5** | White’s test without cross terms | 40 |
| **Table 6** | White’s HETCOV procedure to correct heteroscedasticity | 41 |
| **Table 7** | Dublin-Watson (DW) test | 42 |
| **Table 8** | Lagrange Multiplier (LM) test | 43 |
| **Table 9** | Newey-West HAC standard error | 44 |
| **Table 10** | Correlation Matrix | 45 |
| **Table 11** | Variance Inflation Factors (VIF) | 45 |
| **Table 12** | Structural break test using dummy variable for 1979 | 46 |
| **Table 13** | Wald test for DUM79 and DUM79\*POP | 47 |
| **Table 14** | Wald test for DUM79 and DUM79\*GDP | 47 |
| **Table 15** | Wald test for DUM79 and DUM79\*UNEMP | 48 |
| **Table 16** | Wald test for DUM79 and DUM79\*COAL\_CON | 48 |
| **Table 17** | Wald test for DUM79 and DUM79\*OIL\_PR | 49 |
| **Table 18** | Wald test for DUM79 and DUM79\*NG\_PR | 49 |
| **Table 19** | Wald test for DUM79 and DUM79\*CPI | 50 |
| **Table 20** | Corrected model for structural break at 1979 | 51 |
|  | **List of Tables (Continued)** |  |
| **Table 21** | Structural break test using dummy variable for 2008 | 52 |
| **Table 22** | Wald test for DUM08 and DUM08\*POP | 53 |
| **Table 23** | Wald test for DUM08 and DUM08\*GDP | 53 |
| **Table 24** | Wald test for DUM08 and DUM08\*UNEMP | 54 |
| **Table 25** | Wald test for DUM08 and DUM08\*COAL\_CON | 54 |
| **Table 26** | Wald test for DUM08 and DUM08\*OIL\_PR | 55 |
| **Table 27** | Wald test for DUM08 and DUM08\*NG\_PR | 55 |
| **Table 28** | Wald test for DUM08 and DUM08\*CPI | 56 |
| **Table 29** | OLS Model after fixing break points in 1979 and 2008 | 57 |
| **Table 30** | Wald test on POP, GDP, UNEMP and CPI (test 1) | 58 |
| **Table 31** | Wald test on POP, GDP, UNEMP and CPI (test 2) | 58 |
| **Table 32** | Wald test on COAL\_CON, OIL\_PR and NG\_PR (test 1) | 59 |
| **Table 34** | Wald test on COAL\_CON, OIL\_PR and NG\_PR (test 2) | 59 |

**I. Introduction:** A recent report of Federation of Indian Chambers of Commerce and Industry (FICCI) says "There is strong correlation between power consumption and the GDP of the country. Power shortages currently cost India a GDP loss of $68 billion (0.4 per cent of total GDP)" [1]. Thus in modern world generating sufficient amount of energy is very much important for sustainable development. Excess electricity generation is also not encouraged due to loss of time, man power and raw material cost for electricity production. Though electricity generation and distribution is a time consuming process, we need accurate forecast about how much electricity required for uninterrupted industrial production and household use. Electricity is mainly generated from fossil fuel, nuclear electric power and renewable energy. We should make sure about supplying enough raw materials timely to generate required amount of electricity. From the case study of Bangladesh electricity sector in [2], we can learn that inadequate raw material supply is one of the main reason behind electricity crisis in Bangladesh. Thus how much raw material are required to generate demanded electricity is also a very important factor to know for electricity producing companies.

According to US energy information administration report [3], 40.693 quadrillion Btu fossil fuel has been used between January to June 2014 which is more than 81% of total energy used in US. Coal, crude oil and natural gas are main sources of fossil fuel in United States. Earlier coal were mostly used for fossil fuel due to low price. But recently natural gas consumption to generate electricity increases significantly and in near future trend shows that natural gas consumption for electricity generation will surpass other fossil fuel alternatives (Figure 1). Thus accurate estimation and forecasting of natural gas is required for uninterrupted electricity supply which directly influence GDP growth of a country.

**Figure 1:** Fossil fuel consumption for generation electricity in US

According to the finding of this project population, real GDP and oil price have positive relation with natural gas consumption for electricity generation. On the other hand unemployment rate, coal consumption, natural gas price and consumer price index have negative relation with natural gas consumption for electricity generation. We have done all required diagnostic test and correction for our linear model to make our free from heteroscadasticity, autocorrelation and multicollinearity. Adjusted R-squared value of our estimated model is 0.97 after applying all correction methods. We have found lots of structural breaks in our model. Fixing all structural breaks is time consuming considering our huge dataset, thus we have fixed the break at 1979 and 2008 using dummy variable as an example. Similarly we can fix other break points.

Electricity generation and distribution is a huge industry in United States. In this project, we will do the empirical analysis on natural gas consumption for electricity generation in US. In section II, we will discuss about background study and literature review for this project. In section III, we will propose a mathematical model for total natural gas consumption in US for electricity generation and discuss about different data issues. In section IV, we will do all required diagnostic tests and if we find our model has normality, heteroscedasticity, autocorrelation or multicollinearity problem then we will fix it using existing formulas. In section V, we will check for any kind of structural break in our model and do hypotheses test to ensure that our model is not a sick model anymore after doing all required diagnostic test and correction in section IV. In last section, we will conclude the paper with Dr. Nourzad’s valuable comment/advice on first phase of this paper and my work/reply based on every comment.

**II. Background and Literature review:** In 2011, Himadri Shekhar Dey *et al.* [4], proposed a model to predict gas consumption in Bangladesh. According to the information given in this paper, more than 81% of the total power sector use natural gas raw material to generate electricity and trend shows it will increase future. They have tried to model natural gas demand as a function of price, gross domestic product and per capita income of the studied country, Bangladesh. They considered price and income as economic factor for their model. They used previous year’s gas demand as an input of current year’s gas demand which is called autoregressive model. They used lag length of three in their model to estimate gas demand for Bangladesh. Elasticity has been used in their model for smoothing and their final model is a double log model of Cobb-Douglas functional form. Finally they have shown economic forecast and time trend forecast till 2025 using their model compared the result with actual demand of gas consumption in Bangladesh. They have done Durbin’s alternative test, Breusch-Godfrey test and Ramsey RESET test to demonstrate that there is no serial auto correlation and no omitted variables in their model.

In 1970, US government took some initiative to reduce the electricity consumption through efficient use of it. Some incentives were declared at that time to motive people about efficient use of electricity [16]. In 2012, Richard Cebula [17], showed with mathematical model that energy efficient program worked perfectly. He modeled residential consumption of electricity as a function of cooling degree day, average real price of residential electricity, real per capita disposable income, real average price of natural gas, amount of natural gas used for heating, peak/maximum summer electricity generation capacity and cardinal measure (1, 2, 3) reflecting whether a given state is weekly, moderately or strongly involved with energy efficiency program. The author proposed three different models to analysis whether energy efficiency program performed perfectly or not. First model includes TREND term with double log model to estimate residential electricity consumption in United States. Finding of the paper shows that residential electricity consumption increases (decreases) with the increase (decrease) of cooling degree day, per capita personal income, natural gas price, and maximum summer electricity generating capacity. On the other hand, average real price of residential price, natural gas consumption for heating and energy efficiency program has negative relation with residential electricity consumption in United States. Then he replaced linear energy efficiency score (LEEP) with logarithm of LEEP2 and LEEP2 to generate two different models. Test result shows that coefficient of LEEP, logarithm of LEEP2 and LEEP2 in all three consecutive models are statistically significant which means that more aggressive energy efficiency program will reduce residential electricity consumption in United State keeping other economic factors constant.

In 2014, Richard Cebula and Nate Herder [5] proposed mathematical models to estimate the economic factor electricity consumption per commercial and industrial customer in United States as a function of cooling degree day, electricity price, per capita income, efficiency of electricity use by state and peak electricity generation capacity of each state. They have used same econometric factors stated above in different functional forms: linear model, double-log model and log-linear model. Their findings from the result shows that all prior expected sign holds for three different functional form. The total electricity consumption in United States increases with the increase of cooling degree days, per capita real disposable income and the peak summer electricity capacity, but consumption decreases due to increase in electricity price and more aggressive energy efficiency program.

In 2013, Sergio Ramos *et al.* [6] proposed four mathematical models to forecast electricity based on Halt-Winters exponential smoothing and an artificial neural network and then test these two models with known forecasting results. In two of their models they considered weekends and national holiday for forecasting electricity. Their result shows that considering national holidays and weekends for forecasting electricity gives more accurate result. Finally they compare both models performance with each other and find out neural network model is better than Halt-Winters exponential smoothing.

Sometime, we need data of electricity supply for smaller area to more accurate prediction of electricity consumption but we don’t have that data. In this case if a larger region electricity consumption data is available including that sub region then we can extract the data of our required small region from bigger region data using bus load distribution factor and decoupled extended Kalman filter neural network technique [7].

In fall 2013, Adriatik Hajdari and Paul Kaefer submitted a natural gas demand consumption model for Alaska as an econometrics project of this course. They used wind, local natural gas price, cooling degree day, local natural gas production, local oil production, heating degree day, unemployment rate, gross domestic product of Japan, real gross domestic product of US, natural gas production of US as a factor to forecast the natural gas consumption of Alaska. Their model combines Ronal H. Brown et al. [8] and [Jiří Vondráček](http://www.sciencedirect.com/science/article/pii/S0306261907001183) et al. [9] models.

In 2012, Ronald H. Brown *et.al.* [8], proposed an econometric model to predict residential natural gas consumption as function of cooling degree day, heating degree day, wind and day of the week. They used one period lag term, trend, two different reference temperature (55 and 65 degree Fahrenheit) to calculate cooling degree day and heating degree day in their baseline model. Wind and trend are used to create cross term with 65oF heating degree day. Then they extended their baseline model adding three econometric factors real gross domestic product, consumer price inflation and national unemployment rate. They forecasted residential natural gas consumption using both models for four major states of United States: California, Illinois, New York and Texas. Test result shows that extended model which considered economic factor does better for Illinois and New York.

**III. Modeling Process and Data Issues:** I have taken ideas from all of the papers and project discussed in section II to find an economic model for natural gas consumption in United States. At first I have tried to model [4], because of simplicity. They use only two economic factor to find out natural gas consumption in Bangladesh. Adriatik Hajdari and Paul Kaefer’s model uses too many factors (18 factors) which is very difficult to implement in a short time period. Richard Cebula and Nate Herder [5] proposed three models and one of them is linear regression model. But I have to collect 48 states data for all factors which is also not feasible within short time period. Thus I come up with my own model taking ideas from all of those models. I have used six factors (population, gross domestic product, unemployment rate, coal consumption, oil price and natural gas price of US) in my econometrics model.

Population is a major factor for natural gas consumption because normally when population increases we need more electricity for those extra people. Thus we need more natural gas to generate extra electricity.

Gross domestic product (GDP) is an indicator of economic growth for every country. If a country’s total production increases it requires more electricity to maintain its growth. Thus we include GDP as an economic factor to model gas consumption in US. There are two types of GDP data available: real and nominal GDP. Nominal GDP is the adjusted real GDP considering inflation. In first phase of the project I have used nominal GDP, but in phase II, I will use real GDP and CPI (Consumer Price Index) separately.

Though we have used real GDP in our model, thus we have to use inflation rate or consumer price index to normalize GDP. At first I have used inflation rate, but when I went for double log model testing, then I found that there are some negative data in inflation rate which creates problem to use log model. Thus I picked up consumer price index (CPI) as an alternative of inflation rate because CPI has no negative data.

Unemployment rate is directly related to any kind of consumption. From our prior knowledge we know when unemployment of a country increases it reduces any kind of consumption and vice versa. Thus I have used unemployment rate as a factor to design an economic model for natural gas consumption.

Coal is the main competitor of natural gas in United States due to its low price and availability. But recently we found lots of environmental issues about coal mining. In near past there were some protest in coal mine areas in Bangladesh for environmental issues which make the authority stop mining for several years. Currently coal is the cheapest fossil fuel available for electricity generation in US. Thus when coal consumption will drop for some reason (economical, environmental or political), we can expect to fill up the gap by natural gas. Because natural gas has the second lowest price among all fossil fuels in US market. Thus coal supply is an important economic factor to model natural gas consumption in US.

More than 81% of electricity in US generated by fossil fuel [3]. Two types of fossil fuels are mainly used other than natural gas. One is coal and other is crude oil. But crude oil price has been increased recently which makes other substitute more lucrative to generate electricity. Thus oil price is an important economic factor to model natural gas consumption.

Finally, the economic factor natural gas price is directly related to natural gas consumption. From the general theory of economics when price of some economic good goes up people consume that item less and vice versa considering other factors constant. Thus natural gas price in US is an important economic factor to model natural gas consumption in US.

Considering all economic factors our economic model for natural gas consumption to generate electricity is given below:

*NGC=β1+ β2\*POP+ β3\*GDP+ β4\*UMEMP+ β5\*COALC+ β6\*OILP+ β7\*NGP+ β8\*CPI*

Where (data sources are in parenthesis):

NGC = Total natural gas consumption for generating electricity in United States from 1949 to 2011, measured in trillion BTU [10]. The British thermal unit (BTU or Btu) is a traditional unit of energy equal to about 1055 joules. It is the amount of energy needed to cool or heat one pound of water by one degree Fahrenheit [14]. The unit trillion means we have to multiply with 1000,000,000,000 (12 zeros) for every entry of natural gas consumption in our table to get the original amount.

POP = Total number of people lived in United States from 1949 to 2011 [12], measured in million. The unit billion means we have to multiply each entry of population by 1000,000 (6 zeros) to get the original amount.

GDP = Gross Domestic Product of United States between 1949 and 2011 [12, 13], measured in billion. Nominal GDP is used to get better result. The unit billion means we have to multiply each entry of GDP by 1000,000,000 (9 zeros) to get the original amount.

UMEMP = Unemployment rate in United States between 1949 and 2011 [11], measured in percentage (%).

COALC = Total coal (one of the fossil fuel alternatives of natural gas) consumption in United States to generate electricity between 1949 and 2011 [10], measured in trillion BTU.

OILP = Oil price per million BTU in United States between 1949 and 2011 [10].

NGP = Natural gas price per million BTU in United States between 1949 and 2011 [10].

CPI = Consumer price index calculated in percentage (%) based on the year 1983 [18].

My data sources for this project are given below:

1. US Energy Information Administration (EIA)
2. Bureau of Labor Statistics (BLS)
3. Bureau of Economic Analysis (BEA)

**IV Diagnostic tests and correction:** Before doing any diagnostic test and correction our primary estimated model is given in Table 1 of appendix. Now we will do four kinds of diagnostic test (normality, heteroscedasticity, autocorrelation and multicollinearity) on our initial estimated model and if we find any kind of problem then we will correct the model using existing formula.

**Normality:** From histogram residual (see Figure 5 of appendix) we can observe skewness is -0.17 which very close to zero and kurtosis is 3.93 which is very close to 3. Thus our sample residuals pass in normality test.

Again, from the same figure, we can notice that Jarque-Bera (JB) value is 31.04 with p-value of 0, which means our sample residual is normal with any level of significance according to JB test. Thus we can conclude that there is no normality problem in our sample dataset.

**Heteroscedasticity:** From residual plot (see Figure 6-12 of appendix), we can observe that variance is not constant for any independent variable. Thus the model is suffering from heteroscedasticity problem.

Now, we will test the existence of heteroscedasticity using White’s test with cross terms and without cross terms. For both tests our hypotheses and alternative hypotheses is given below:

H0: σ2 is constant Vs. H1: σ2 is not constant

For White’s test with cross terms (see Table 4 of appendix), we can see that Obs\*R-squared is 425.64 with p-value 0 and without cross terms (see Table 5 of appendix) we can observe that Obs\*R-squared is 108.8 with p-value 0. In both cases, we can reject the null hypotheses in favor of alternative hypotheses with any level of significance. Thus we can conclude with high level of confidence that, our model is suffering from heteroscedasticity.

Though we have found heteroscedasticity problem in our model, thus we have corrected it using White’s HETCOV procedure (see Table 6 of appendix). The method corrects standard error as well as t-stats and p-values leaving all point estimates unchanged.

**Autocorrelation:** From the residual plot (see Figure 13 of appendix), we can observe that there exists positive autocorrelation in our estimated model. Now we will test the first order autocorrelation using Dublin-Watson and LM test.

Consider, for Dublin-Watson test, our null and alternative hypotheses are

H0: ρ ≤ 0 vs. H1: ρ > 0.

Using α = 0.05, k’ = 7 and n = 755 (though online table has 750 so I have used 750), we have found our critical values as dL = 1.864 and dU = 1.896 [19]. Since d= 1.243 < dL, we will reject the null hypotheses in favor of alternative hypotheses and conclude that with 5% level of significance there is evidence of positive first-order autocorrelation in the estimated model.

Now, we will use LM test, for verifying whether there exists any first order autocorrelation in our model. Though the LM test is a two-sided or two-tailed test so the null and alternatives are

H0: ρ = 0 vs. H1: ρ ≠ 0.

From test result, we can observe that nR2 has a Chi-squared value of 741.36 with a p-value of 0, which leads us to reject the null hypotheses with any level of significance (see Table 8 of appendix). Thus we can conclude with high level of confidence that there is evidence of first-order positive autocorrelation in the estimated model.

Though above tests confirm the existence of first order positive autocorrelation in our estimated model, we have fixed the problem using Newey-West HAC standard error (see Table 9 of appendix) which corrects all standard errors, t-stats and p-values leaving point estimates unchanged.

**Multicollinearity:** Though we are using time series data, there might be existence of multicollinearity. Thus we have tested the existence of this problem from correlation matrix and found that some independent variables are highly correlated to each other (see Table10 of appendix). For example Population is highly correlated with GDP, coal consumption and CPI.

From Variance Inflation Factor (VIF) test (see Table 11 of appendix), we can observe that only unemployment has VIF value less than 10 and all other variables have VIF value more than 10. That means in our model there is a big chance that population, GDP, coal consumption, oil price, natural gas price and CPI have multicollinearitty problem.

Though we have found that there exists high possibility of multicollinearity, still we will not do any correction because all of our coefficients are statistically significant and have correct sign in the estimated model. The rule says that if it’s not broke then we need not to fix multicollinearity.

**V Testing Hypotheses on the Regression Coefficients:** After all diagnostic tests and correction in previous section, we have our corrected model in Table 6 of appendix. In this section, we will interpret our final corrected model and then we will do some hypothesis testing to check our model’s performance.

According to the estimated coefficients of Table 6 in appendix, if population of United States increase (decrease) by 1 million then natural gas consumption for electricity generation increase (decrease) by 3.73 trillion BTU on an average holding other factors constant. This result satisfies our prior knowledge and expectation about population growth and consumption. When population will increase we will need more electricity for growing population and we need more natural gas to generate electricity.

According to Table 6 of appendix, on an average 1 billion dollar increase (decrease) in GDP cause natural gas consumption increase (decrease) by 0.17 trillion BTU holding other factors constant. This result also goes with our prior expectation. From common knowledge GDP increase of a country means more production and nowadays production is heavily depend on electricity. Thus we would need more natural gas for electricity generation. In phase I of this project we had unexpected sign for GDP. Thus we explain this unexpected result from energy efficiency program taken by the government in 1970 [15]. But in phase II, after including a new CPI and changing nominal GDP to real GDP, we get expected sign for the coefficient of GDP. We will check the significance of the sign in hypothesis testing part of this section.

According to Table 6 of appendix, one percentage point (100 basis point) increase (decrease) of unemployment rate in US is expected to decrease (increase) natural gas consumption for electricity generation by about 2.17 trillion BTU holding other factors constant. From common knowledge, we know when unemployment rate increase then people force to cut their consumption. Thus it makes sense.

According to Table 6 of appendix, on an average 1 trillion BTU coal consumption for electricity generation decrease (increase) has the effect of increasing (decreasing) natural gas consumption by about 0.18 trillion BTU holding other factors constant. Natural gas is the cheapest alternative of coal in US electricity section, thus reduction of coal supply positively impact on natural gas demand for electricity generation.

According to Table 6 of appendix, on an average the increase (decrease) of 1 dollar in oil price per million BTU cause natural gas consumption to increase (decrease) by about 7.8 trillion BTU for electricity generation in US holding other factors constant. When oil price became costly electricity companies searches for alternative sources like natural gas to reduce production cost. Thus our findings from the model goes with general concept of demand for any item and price of alternative.

According to Table 6 of appendix, when natural gas price increase (decrease) on an average by 1 dollar per million BTU then it’s consumption decrease (increase) by about 16.57 trillion BTU holding other factors constant which also goes with the common knowledge of price and demand of economics.

According to Table 6 of appendix, one percentage point (100 basis point) increase (decrease) of CPI is expected to decrease (increase) natural gas consumption for electricity generation by about 2.38 trillion BTU holding other factors constant. Increase of CPI means people’s purchasing power will decrease due to inflation. Thus people can consume less electricity using same amount of money over time. So, the coefficient sign is perfect considering our prior knowledge of CPI. The mean systematic effect of determinants of the natural gas consumption for electricity sector in US is about -544.12 trillion BTU.

**Structural break test and correction:** In phase 1 of this project, I have used chaw breakpoint test for testing structural breakpoint in our model. But chaw breakpoint have limitation when we try to check break point at the end or beginning of the time series data. It doesn’t work with small sample data. Now we will use another five methods (Recursive coefficient test, Recursive Residual test, CUSUM test, CUSUM of squares test, Structural test using dummy variable) to check break point in our model in this section.

According to recursive coefficient test (see Figure 1 of appendix), every independent variables are within ±2SE region. The independent variable population has structural breaks at 1954, 1958, 1966, 1972 and 1975. The unemployment rate has structural breaks at 1957, 1965, 1967 1972, 1980 and 2008. Coal consumption has structural breaks at 1951, 1959, 1966, 1973, 1975, 1984 and 1995. Natural gas price has structural breaks at 1951, 1959, 1974 and 1980. CPI has structural breaks at 1959, 1966, 1970 and 1979, GDP has structural breaks at 1959, 1963, 1966, 1974, 1981 and 1996. Oil price has structural breaks at 1950, 1974, 1975 and 1981.

According to recursive residual test (see Figure 2 of appendix), our model has structural breaks at 1959, 1964, 1965, 1968, 1970, 1972, 1974, 1975, 1977, 1979, 1990, 1993, 2004, 2006 and 2009.

According to CUSUM test (see Figure 3 of appendix), there is one large break point between 1979 and 2008.

According to CUSUM of squares test (see Figure 1 of appendix), there is one large break point between 1955 and 2006.

Now we will do dummy variable test for finding which variables are responsible for particular break point and then if we find any break point, we will fix the break point using dummy variables. Though there are lots of break point in our model and we have huge data set, thus we will test only test for break point at 1979 and 2008. Other points can be detected and fixed using same approach. At first, we will test the break point at 1979 using the following linear model:

*NGC=β1+β2\*DUM79+β3\*POP+β4\*GDP+β5\*UMEMP+β6\*COALC+β7\*OILP+β8\*NGP +β9\*CPI+β10\*DUM79\*POP+β11\*DUM79\*GDP+β12\*DUM79\*UNEMP +β13\*DUM79\*COAL\_CON+ β14\*DUM79\*OIL\_PR+β15\*DUM79\*NG\_PR+ β16\*DUM79\*CPI*

To find the responsible independent variable for structural break, we have to do the following two-tailed t-tests followed by some F-tests. For all required tests to check break point at 1979, our null hypotheses and alternative hypotheses are given below:

Test 1: H0: *β2* = 0 vs H1: *β2* ≠ 0 (t-test, see Table 12 of appendix)

Test 2: H0: *β10* = 0 vs H1: *β10* ≠ 0 (t-test, see Table 12 of appendix)

Test 3: H0: *β11* = 0 vs H1: *β11* ≠ 0 (t-test, see Table 12 of appendix)

Test 4: H0: *β12* = 0 vs H1: *β12* ≠ 0 (t-test, see Table 12 of appendix)

Test 5: H0: *β13* = 0 vs H1: *β13* ≠ 0 (t-test, see Table 12 of appendix)

Test 6: H0: *β14* = 0 vs H1: *β14* ≠ 0 (t-test, see Table 12 of appendix)

Test 7: H0: *β15* = 0 vs H1: *β15* ≠ 0 (t-test, see Table 12 of appendix)

Test 8: H0: *β16* = 0 vs H1: *β16* ≠ 0 (t-test, see Table 12 of appendix)

Test 9: H0: *β2* = 0 and *β10* = 0 vs H1: *β2* ≠ 0 and *β10* ≠ 0 (F-test, see Table 13 of appendix)

Test 10: H0: *β2* = 0 and *β11* = 0 vs H1: *β2* ≠ 0 and *β11* ≠ 0 (F-test, see Table 14 of appendix)

Test 11: H0: *β2* = 0 and *β12* = 0 vs H1: *β2* ≠ 0 and *β12* ≠ 0 (F-test, see Table 15 of appendix)

Test 12: H0: *β2* = 0 and *β13* = 0 vs H1: *β2* ≠ 0 and *β13* ≠ 0 (F-test, see Table 16 of appendix)

Test 13: H0: *β2* = 0 and *β14* = 0 vs H1: *β2* ≠ 0 and *β14* ≠ 0 (F-test, see Table 17 of appendix)

Test 14: H0: *β2* = 0 and *β15* = 0 vs H1: *β2* ≠ 0 and *β15* ≠ 0 (F-test, see Table 18 of appendix)

Test 15: H0: *β2* = 0 and *β16* = 0 vs H1: *β2* ≠ 0 and *β16* ≠ 0 (F-test, see Table 19 of appendix)

In test 1, 2, 3, 4, 6, 7 and 8, we have p-value of 0, thus we can reject the null hypotheses in favor of alternative hypotheses with any level of significance. But in test 5, p-value is 0.3245. Thus we can’t reject the null hypotheses in favor of alternative hypotheses with any 5% level of significance. So, we can conclude that only coal consumption is not responsible for break at 1979, but all other independent variables are responsible for the break at this point. Now from test 9, 10, 11, 12, 13, 14 and 15, we get p-values for every test result as 0, which means we can reject the null hypotheses in favor of alternative hypotheses with any level of significance. Thus we can conclude that at 1979, there is a break and all independent variables are jointly responsible for this break. Table 20 of appendix, contains the estimated model after fixing the break point at 1979.

Now, we will check another break point at 2008 using another dummy variable named DUM08. For this break point test, our linear model is given below:

*NGC=β1+β2\*DUM08+β3\*POP+β4\*GDP+β5\*UMEMP+β6\*COALC+β7\*OILP+β8\*NGP +β9\*CPI+β10\*DUM08\*POP+β11\*DUM08\*GDP+β12\*DUM08\*UNEMP +β13\*DUM08\*COAL\_CON+ β14\*DUM08\*OIL\_PR+β15\*DUM08\*NG\_PR+ β16\*DUM08\*CPI*

To find the responsible independent variable for structural break, we have to do the following two-tailed t-tests followed by some F-tests. For all required tests to check break point at 2008, our null hypotheses and alternative hypotheses are given below:

Test 1: H0: *β2* = 0 vs H1: *β2* ≠ 0 (t-test, see Table 21 of appendix)

Test 2: H0: *β10* = 0 vs H1: *β10* ≠ 0 (t-test, see Table 21 of appendix)

Test 3: H0: *β11* = 0 vs H1: *β11* ≠ 0 (t-test, see Table 21 of appendix)

Test 4: H0: *β12* = 0 vs H1: *β12* ≠ 0 (t-test, see Table 21 of appendix)

Test 5: H0: *β13* = 0 vs H1: *β13* ≠ 0 (t-test, see Table 21 of appendix)

Test 6: H0: *β14* = 0 vs H1: *β14* ≠ 0 (t-test, see Table 21 of appendix)

Test 7: H0: *β15* = 0 vs H1: *β15* ≠ 0 (t-test, see Table 21 of appendix)

Test 8: H0: *β16* = 0 vs H1: *β16* ≠ 0 (t-test, see Table 21 of appendix)

Test 9: H0: *β2* = 0 and *β10* = 0 vs H1: *β2* ≠ 0 and *β10* ≠ 0 (F-test, see Table 22 of appendix)

Test 10: H0: *β2* = 0 and *β11* = 0 vs H1: *β2* ≠ 0 and *β11* ≠ 0 (F-test, see Table 23 of appendix)

Test 11: H0: *β2* = 0 and *β12* = 0 vs H1: *β2* ≠ 0 and *β12* ≠ 0 (F-test, see Table 24 of appendix)

Test 12: H0: *β2* = 0 and *β13* = 0 vs H1: *β2* ≠ 0 and *β13* ≠ 0 (F-test, see Table 25 of appendix)

Test 13: H0: *β2* = 0 and *β14* = 0 vs H1: *β2* ≠ 0 and *β14* ≠ 0 (F-test, see Table 26 of appendix)

Test 14: H0: *β2* = 0 and *β15* = 0 vs H1: *β2* ≠ 0 and *β15* ≠ 0 (F-test, see Table 27 of appendix)

Test 15: H0: *β2* = 0 and *β16* = 0 vs H1: *β2* ≠ 0 and *β16* ≠ 0 (F-test, see Table 28 of appendix)

In test 1, 3, 5, 6 and 8, we have p-value of 0, thus we can reject the null hypotheses in favor of alternative hypotheses. But in test 2, 4 and 7 our p-values are 0.9402, 0.216 and 0.4594 respectively with any level of significance. Thus, for test 2, 4 and 7 we can’t reject the null hypotheses in favor of alternative hypotheses with 5% level of significance. So, we can conclude that population, unemployment rate and natural gas price are not responsible for structural break at 2008. Here GDP, coal consumption, oil price and CPI are responsible for the break at that point. Now from test 9, 10, 11, 12, 13, 14 and 15, we get p-values for every test result is 0, which means we can reject the null hypotheses in favor of alternative hypotheses with any level of significance. Thus we can conclude that at 2008, there is a break and all independent variables are jointly responsible for this break. Table 29 of appendix, contains the estimated model after fixing both break points at 1979 and 2008.

**Hypotheses testing:** Finally, we have the following linear model for hypotheses testing:

*NGC=β1+β2\*DUM79\_08+β3\*POP+β4\*GDP+β5\*UMEMP+β6\*COALC+β7\*OILP+β8\*NGP +β9\*CPI+β10\*DUM79\_08\*POP+β11\*DUM79\_08\*GDP+β12\*DUM79\_08\*UNEMP +β13\*DUM79\_08\*COAL\_CON+β14\*DUM79\_08\*OIL\_PR+β15\*DUM79\_08\*NG\_PR + β16\*DUM79\_08\*CPI*

From our econometric model of natural gas consumption for electricity section in US, we want to do the following hypotheses tests:

1. Testing the estimated model for overall significance: To perform the test our null and alternative hypotheses are given below:

H0: *β2* = 0 and *β3* = 0 and *β4* = 0 and *β5* = 0 and *β6* = 0 and *β7* = 0 and *β8* = 0 and *β9* = 0 and *β10* = 0 and *β11* = 0 and *β12* = 0 and *β13* = 0 and *β14* = 0 and *β15* = 0 and *β16* = 0 vs

H1: *β2* ≠ 0 or *β3* ≠ 0 or *β4* ≠ 0 or *β5* ≠ 0 or *β6* ≠ 0 or *β7* = 0 or *β8* = 0 or *β9* ≠ 0 or *β10* ≠ 0 or *β11* ≠ 0 or *β12* ≠ 0 or *β13* ≠ 0 or *β14* ≠ 0 or *β15* ≠ 0 or *β16* ≠ 0.

Alternatively, H0: R2 = 0 vs H1: R2 ≠ 0.

From the regression output (see Table 29 of appendix), we find the necessary F-statistic to be 2765.78 with a p-value of zero. This leads us to reject the null hypotheses in the favor of alternative hypotheses at nearly any level of significance and conclude that the estimated model is overall statistically significant.

1. Testing whether the estimated coefficients on population, GDP, unemployment rate and CPI are jointly statistically significant: Our null hypotheses and alternative hypotheses for the test are:

H0: *β3* = *β4* = *β5* = *β9* = 0 vs H1: *β3* ≠ 0 or *β4* ≠ 0 or *β5* ≠ or *β9* ≠ 0.

According to test result (see Table 30 of appendix), F-statistic is 426.7 with p-value of zero, meaning we can reject the null hypotheses in favor of alternative hypotheses at any level of significance. Thus we can conclude that US population, GDP, unemployment rate and CPI are jointly statistically significant with high level of confidence.

1. Testing whether the sum of estimated coefficients on population, GDP, and unemployment rate equal to zero: Our null hypotheses and alternative hypotheses for the test are: H0: *β3* + *β4* + *β5* + *β9* =0 vs H1: *β3* + *β4* + *β5* + *β9* ≠ 0.

The test result shows (see Table 32 of appendix), we can observe F-statistics is 11.39 with p-value of 0.0008 which leads us to reject the null hypotheses in favor of alternative hypotheses with any level of significance and conclude that the sum of the estimated effects of population, GDP, unemployment rate and CPI on natural gas consumption is statistically significantly different from zero with high level of confidence.

1. Testing whether the estimated coefficients on coal consumption, oil price, and natural gas price are jointly statistically significant: Our null hypotheses and alternative hypotheses for the test is given below:

H0: *β6* = *β7* = *β8* = 0 vs H1: *β6* ≠ 0 or *β7* ≠ 0 or *β8* ≠ 0.

From the test result (see Table 32 of appendix), we can see F-statistics is 70.55 with p-value of zero, which means we can reject the null hypotheses in favor of alternative hypotheses with any level of significance. Thus we can conclude that coal consumption, oil price, and natural gas price are jointly statistically significant with any high level of confidence.

1. Testing whether the sum of estimated coefficients on coal consumption, oil price, and natural gas price equal to zero: Our null hypotheses and alternative hypotheses for the test are: H0: *β6* + *β7* + *β8* = 0 vs H1: *β6* + *β7* + *β8* ≠ 0.

From the test result (see Table 33 of appendix), we can observe F-statistics is 19.07 with p-value of zero, which leads us to reject the null hypotheses in favor of alternative hypotheses with any level of significance and conclude that the sum of estimated effects of coal consumption, oil price, and natural gas price is statistically significantly different from zero with high level of confidence.

**VI Comments and suggestion of Dr. Nourzad on phase-I and my reply:**

**Comment and suggestion:** All in all, not a bad attempt at the first draft.

**Reply:** Thank you.

**Comment and suggestion:** Instead of nominal GDP, use real GDP

**Reply:** Done

**Comment and suggestion:** Include the inflation rate in your model as well.

**Reply:** Done. But later I have found that I can’t use log test using inflation rate because there are some negative numbers. Thus I asked you again and you suggested me to use consumer price index and I have worked accordingly.

**Comment and suggestion:** Test your model for double-log specification.

**Reply:** I have used Mackinnon White Davidson test for model selection between double-log and linear model. But unfortunately the test was inconclusive.

At first my null and alternative hypotheses was:

H0: purely functional form is correct vs H1: double-log functional form is correct

I have added the new independent variable  in our linear regression model (see Table 2 of appendix). We can observe in the table that, the t-stat of Z1 is -9 with p-value zero, which means we can reject the null hypotheses in favor of alternative hypotheses with any level of significance.

Now, for next step, our null and alternative hypotheses are given below:

H0: double-log functional form is correct vs H1: purely functional form is correct

I have included a new variable in Douglas-Cobb double log functional form (see Table 3 of appendix). We can observe from the table, the t-stat of Z2 is -4.48 with p-value zero, which means we can reject the null hypotheses in favor of alternative hypotheses with any level of significance.

Thus from above test results we can’t make a decision which model is better. Thus I have chosen linear model instead of double log model because of simplicity.

**Comment and suggestion:** You use annual data, why not monthly or at least quarterly data?

**Reply:** Phase II contains monthly data. I have tried to recollect all data for monthly. For some variables monthly data was not available. For those variables, I have converted quarterly and yearly data into monthly data using this tutorial: <http://www.eviews.com/Learning/freqconv_a.html>

**Problem in Phase I:** Introduction

**Work done:** In phase I, it was three paragraph. Now I have rewritten introduction section according to your suggestion (four paragraph- brief background on the subject from the literature, motivation, brief summary of findings, a paragraph describing layout of the rest of the paper)

**Problem in Phase I:** Page number missing

**Work done:** Included page number in Phase II

**Problem in Phase I:** Insufficient write-up for literature review section

**Work done:** Added two new papers in literature review section and also increased the writing on previously added paper to make it two and half pages long from earlier version of one page.

**Problem in Phase I:** Table number and heading is under the table

**Work done:** Now, in phase II, all table name and heading is top (beginning) of the table.

**Problem in Phase I:** Formatting problem. Reference section started from middle of the page.

**Work done:** Now I put some enter to make it start from next page in phase II.

**References:**

[1] Indian Power Sector.Com, [<http://indianpowersector.com/2013/09/power-shortages-cost-india-68-billion-in-gdp-ficci/> ] accessed October 12, 2014.

[2] Khairul Anam and Husnain-Al-Bustam, “Power Crisis & Its Solution through Renewable Energy in Bangladesh”, Journal of Selected Areas in Renewable and Sustainable Energy (JRSE), pp 13-18, 2011.

[3] US Energy Information Administration, Total energy consumption, [<http://www.eia.gov/totalenergy/>] accessed October 12, 2014.

[4] Himadri Shekhar Dey, Md. Ashfanoor Kabir, Zia Wadud, Shahidul Islam Khan and Md. Abul Kalam Azad, “Econometric Modeling and Forecasting of Natural Gas Demand for Power Sector in Bangladesh”, IEEE, TENCON, pp1383-1386, 2011.

[5] Richard Cebula and Nate Herder, “An Empirical Analysis of Determinants of Commercial and Industrial Electricity Consumption”, MPRA, paper no. 56763, 2014.

[6] Sergio Ramos, Joao Soares, Zita Vale and Sndra Ramos, “Short-term Load Forecasting Based on Load Profiling”, IEEE, 2013.

[7] Xiaorong Sun, Peter B. Luh and Kwok W. Cheung, “An Efficient Approach for Short-Term Substation Load Forecasting”, IEEE, 2013.

[8] Ronald H. Brown, David Clark, George F. Corliss, Farrokh Naurzad, Thomas Quinn, Catherine Twetten, “Forecasting Natural Gas Demand: The Role of Physical and Economic Factors”, Evanston,Midwest Economics Association Meetings, 2012.

[9] [Jiří Vondráček](http://www.sciencedirect.com/science/article/pii/S0306261907001183), [Emil Pelikán](http://www.sciencedirect.com/science/article/pii/S0306261907001183), [Ondřej Konár](http://www.sciencedirect.com/science/article/pii/S0306261907001183), “A Statistical Model for the Estimation of Natural Gas Consumption”, Applied Energy, pp 362-370, 2008.

[10] Fossil fuel Production Prices 1949-2011, US Energy Information Administration, [<http://www.eia.gov/totalenergy/data/annual/showtext.cfm?t=ptb0301>] accessed October 13, 2014.

[11] US Unemployment rate, Bureau of Labor Statistics,

[<http://data.bls.gov/pdq/SurveyOutputServlet>] accessed October 13, 2014.

[12] Population, US Gross Domestic Product and Implicit Price Deflator, 1949-2011, US Energy Information Administration, [<https://research.stlouisfed.org/fred2/data/GDP.txt>] accessed October 13, 2014.

[13] US Gross Domestic Product (GDP), US Department of Commerce: Bureau of Economic Analysis, [<https://research.stlouisfed.org/fred2/data/GDP.txt>] accessed October 13, 2014.

[14] BTU, Wikipedia, [<http://en.wikipedia.org/wiki/British_thermal_unit>] accessed October 14, 2014.

[15] Marvin J. [Horowitz](http://search.proquest.com/econlit/indexinglinkhandler/sng/au/Horowitz,+Marvin+J/$N?accountid=100), “Changes in Electricity Demand in the United States from the 1970s to 2003”, Energy journal, pp 93-119, 2007.

[16] Incentives for energy efficiency, [<http://www.dsireusa.org/>] Accessed November 4, 2014.

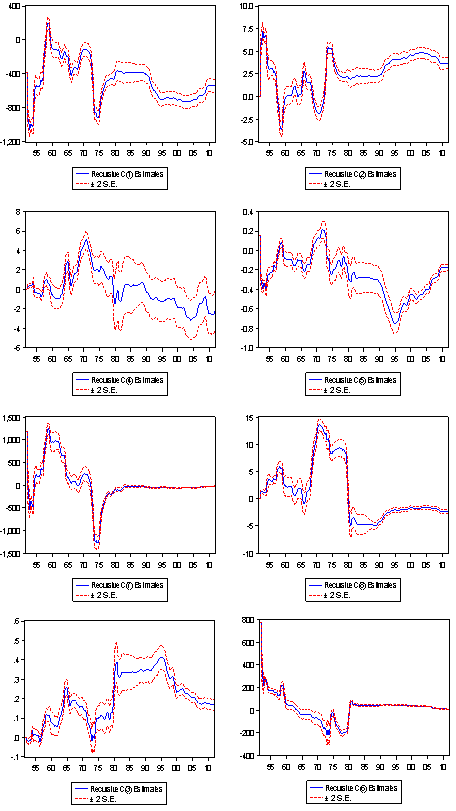
[17] Richar J. Cebula, “US residential electricity consumption: the effect of states’ pursuit of energy efficiency policies”, Applied Econometrics Letters, pp 1499-1503, 2012.

[18] Consumer Price Index, Bureau of labor statistics, [<http://www.usinflationcalculator.com/inflation/consumer-price-index-and-annual-percent-changes-from-1913-to-2008/>] Accessed November 4, 2014.

### [19] Critical Values for the Durbin-Watson Test, [<http://web.stanford.edu/~>[clint/bench/dwcrit.htm](http://web.stanford.edu/~clint/bench/dwcrit.htm)] Accessed November 13, 2014.

**Appendix:**

**Stability Tests:**



**Figure 2:** Recursive coefficients test



**Figure 3:** Recursive residual test



**Figure 4:** CUSUM test



**Figure 5:** CUSUM of squares test

**Normality test**



**Figure 6:** Histogram Residuals

**Heteroscedasticity tests**



**Figure 7:** Residuals plot (Population vs Residual squared)



**Figure 8:** Residuals plot (Gross Domestic Product vs Residual squared)



**Figure 9:** Residuals plot (Unemployment rate vs Residual squared)



**Figure 10:** Residuals plot (Coal consumption vs Residual squared)



**Figure 11:** Residuals plot (Oil price vs Residual squared)



**Figure 12:** Residuals plot (Natural gas price vs Residual squared)



**Figure 13:** Residuals plot (Consumer Price Index vs Residual squared)

**Autocorrelation test:**



**Figure 14:** Residuals plot

**Table 1:** Estimation of our linear regression model (before applying any test and correction)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: NG\_CON | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/06/14 Time: 22:06 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -544.1195 | 39.68684 | -13.71033 | 0.0000 |
| POP | 3.726407 | 0.322695 | 11.54777 | 0.0000 |
| GDP | 0.165990 | 0.013530 | 12.26793 | 0.0000 |
| UNEMP | -2.171742 | 1.032024 | -2.104353 | 0.0357 |
| COAL\_CON | -0.177541 | 0.016970 | -10.46187 | 0.0000 |
| OIL\_PR | 7.803850 | 0.893256 | 8.736412 | 0.0000 |
| NG\_PR | -16.57115 | 2.006785 | -8.257558 | 0.0000 |
| CPI | -2.380516 | 0.172633 | -13.78946 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.971455 | Mean dependent var | | 304.6971 |
| Adjusted R-squared | 0.971188 | S.D. dependent var | | 155.7872 |
| S.E. of regression | 26.44331 | Akaike info criterion | | 9.398409 |
| Sum squared resid | 523038.2 | Schwarz criterion | | 9.447383 |
| Log likelihood | -3544.599 | Hannan-Quinn criter. | | 9.417273 |
| F-statistic | 3636.665 | Durbin-Watson stat | | 0.019653 |
| Prob(F-statistic) | 0.000000 |  |  |  |

**Table 2:** Estimation of linear regression model with Z1 (Mackinnon-White-Davidson test)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: NG\_CON | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/06/14 Time: 22:05 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -716.6955 | 42.31231 | -16.93823 | 0.0000 |
| POP | 5.063917 | 0.340798 | 14.85902 | 0.0000 |
| GDP | 0.123034 | 0.013717 | 8.969680 | 0.0000 |
| UNEMP | -5.075738 | 1.032567 | -4.915651 | 0.0000 |
| COAL\_CON | -0.222156 | 0.016874 | -13.16588 | 0.0000 |
| OIL\_PR | 6.312166 | 0.865009 | 7.297222 | 0.0000 |
| NG\_PR | -12.64447 | 1.956580 | -6.462535 | 0.0000 |
| CPI | -2.068474 | 0.167700 | -12.33440 | 0.0000 |
| Z1 | -86.86803 | 9.649903 | -9.001959 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.974249 | Mean dependent var | | 304.6971 |
| Adjusted R-squared | 0.973973 | S.D. dependent var | | 155.7872 |
| S.E. of regression | 25.13288 | Akaike info criterion | | 9.298064 |
| Sum squared resid | 471851.3 | Schwarz criterion | | 9.353160 |
| Log likelihood | -3505.668 | Hannan-Quinn criter. | | 9.319286 |
| F-statistic | 3532.691 | Durbin-Watson stat | | 0.019989 |
| Prob(F-statistic) | 0.000000 |  |  |  |

**Table 3:** Estimation of double logarithmic model with Z2 (Mackinnon-White-Davidson test)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: LOG(NG\_CON) | | | |  |
| Method: Least Squares | | |  |  |
| Date: 11/06/14 Time: 22:04 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -22.82891 | 0.806101 | -28.32015 | 0.0000 |
| LOG(POP) | 5.559323 | 0.318566 | 17.45106 | 0.0000 |
| LOG(GDP) | -0.188612 | 0.141581 | -1.332185 | 0.1832 |
| LOG(UNEMP) | -0.085279 | 0.021101 | -4.041472 | 0.0001 |
| LOG(COAL\_CON) | 0.565527 | 0.033691 | 16.78568 | 0.0000 |
| LOG(OIL\_PR) | 0.261121 | 0.018664 | 13.99097 | 0.0000 |
| LOG(NG\_PR) | -0.256467 | 0.020540 | -12.48593 | 0.0000 |
| LOG(CPI) | -0.969579 | 0.031963 | -30.33482 | 0.0000 |
| Z2 | -0.001372 | 0.000306 | -4.480402 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.980142 | Mean dependent var | | 5.558491 |
| Adjusted R-squared | 0.979929 | S.D. dependent var | | 0.617410 |
| S.E. of regression | 0.087469 | Akaike info criterion | | -2.023227 |
| Sum squared resid | 5.715193 | Schwarz criterion | | -1.968131 |
| Log likelihood | 773.7797 | Hannan-Quinn criter. | | -2.002005 |
| F-statistic | 4608.754 | Durbin-Watson stat | | 0.015464 |
| Prob(F-statistic) | 0.000000 |  |  |  |

**Table 4:** White’s test with cross terms

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Heteroskedasticity Test: White | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| F-statistic | 26.50423 | Prob. F(35,720) | | 0.0000 |
| Obs\*R-squared | 425.6382 | Prob. Chi-Square(35) | | 0.0000 |
| Scaled explained SS | 610.5022 | Prob. Chi-Square(35) | | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Test Equation: | |  |  |  |
| Dependent Variable: RESID^2 | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/06/14 Time: 21:36 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 59548.27 | 55874.94 | 1.065742 | 0.2869 |
| POP^2 | 0.851785 | 3.594123 | 0.236994 | 0.8127 |
| POP\*GDP | -0.408016 | 0.358481 | -1.138177 | 0.2554 |
| POP\*UNEMP | 16.29539 | 17.25916 | 0.944159 | 0.3454 |
| POP\*COAL\_CON | 1.722823 | 0.421774 | 4.084709 | 0.0000 |
| POP\*OIL\_PR | 132.8890 | 45.20311 | 2.939821 | 0.0034 |
| POP\*NG\_PR | -405.6112 | 109.1043 | -3.717647 | 0.0002 |
| POP\*CPI | 0.935409 | 4.614308 | 0.202719 | 0.8394 |
| POP | -678.3441 | 891.0289 | -0.761304 | 0.4467 |
| GDP^2 | 0.006073 | 0.008036 | 0.755646 | 0.4501 |
| GDP\*UNEMP | -1.769794 | 0.802090 | -2.206477 | 0.0277 |
| GDP\*COAL\_CON | -0.119785 | 0.015046 | -7.961081 | 0.0000 |
| GDP\*OIL\_PR | -6.751727 | 1.062285 | -6.355853 | 0.0000 |
| GDP\*NG\_PR | 9.250779 | 2.678922 | 3.453173 | 0.0006 |
| GDP\*CPI | 1.053749 | 0.174951 | 6.023116 | 0.0000 |
| GDP | 90.60567 | 47.42473 | 1.910515 | 0.0565 |
| UNEMP^2 | -139.5407 | 32.41210 | -4.305202 | 0.0000 |
| UNEMP\*COAL\_CON | -1.518928 | 1.127550 | -1.347105 | 0.1784 |
| UNEMP\*OIL\_PR | 51.92190 | 70.79202 | 0.733443 | 0.4635 |
| UNEMP\*NG\_PR | 180.0201 | 156.2686 | 1.151991 | 0.2497 |
| UNEMP\*CPI | 29.09754 | 9.838917 | 2.957392 | 0.0032 |
| UNEMP | 130.7057 | 2042.244 | 0.064001 | 0.9490 |
| COAL\_CON^2 | 0.005336 | 0.012407 | 0.430090 | 0.6673 |
| COAL\_CON\*OIL\_PR | -0.133805 | 1.429320 | -0.093614 | 0.9254 |
| COAL\_CON\*NG\_PR | 21.54945 | 3.443085 | 6.258762 | 0.0000 |
| COAL\_CON\*CPI | 0.686459 | 0.202505 | 3.389845 | 0.0007 |
| COAL\_CON | -201.3520 | 54.88054 | -3.668915 | 0.0003 |
| OIL\_PR^2 | -24.49003 | 40.96413 | -0.597841 | 0.5501 |
| OIL\_PR\*NG\_PR | 598.1182 | 157.9104 | 3.787706 | 0.0002 |
| OIL\_PR\*CPI | 57.88341 | 15.51456 | 3.730909 | 0.0002 |
| OIL\_PR | -22411.22 | 7098.729 | -3.157075 | 0.0017 |
| NG\_PR^2 | -1261.891 | 187.9653 | -6.713424 | 0.0000 |
| NG\_PR\*CPI | 11.28754 | 38.09256 | 0.296319 | 0.7671 |
| NG\_PR | 44812.04 | 17508.17 | 2.559493 | 0.0107 |
| CPI^2 | -17.19918 | 1.459662 | -11.78299 | 0.0000 |
| CPI | -507.7582 | 629.9374 | -0.806046 | 0.4205 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.563013 | Mean dependent var | | 691.8494 |
| Adjusted R-squared | 0.541771 | S.D. dependent var | | 1185.107 |
| S.E. of regression | 802.2298 | Akaike info criterion | | 16.25912 |
| Sum squared resid | 4.63E+08 | Schwarz criterion | | 16.47950 |
| Log likelihood | -6109.946 | Hannan-Quinn criter. | | 16.34400 |
| F-statistic | 26.50423 | Durbin-Watson stat | | 0.134027 |
| Prob(F-statistic) | 0.000000 |  |  |  |

**Table 5:** White’s test without cross terms

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Heteroskedasticity Test: White | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| F-statistic | 17.96362 | Prob. F(7,748) | | 0.0000 |
| Obs\*R-squared | 108.8000 | Prob. Chi-Square(7) | | 0.0000 |
| Scaled explained SS | 156.0542 | Prob. Chi-Square(7) | | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Test Equation: | |  |  |  |
| Dependent Variable: RESID^2 | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/06/14 Time: 21:37 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -1265.508 | 384.8666 | -3.288173 | 0.0011 |
| POP^2 | 0.051016 | 0.015053 | 3.389097 | 0.0007 |
| GDP^2 | 0.000147 | 7.19E-05 | 2.046976 | 0.0410 |
| UNEMP^2 | -1.691761 | 2.874474 | -0.588546 | 0.5563 |
| COAL\_CON^2 | 0.000425 | 0.000262 | 1.621918 | 0.1052 |
| OIL\_PR^2 | 10.58453 | 2.105765 | 5.026454 | 0.0000 |
| NG\_PR^2 | -40.08083 | 9.622472 | -4.165336 | 0.0000 |
| CPI^2 | -0.183639 | 0.033904 | -5.416440 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.143915 | Mean dependent var | | 691.8494 |
| Adjusted R-squared | 0.135904 | S.D. dependent var | | 1185.107 |
| S.E. of regression | 1101.637 | Akaike info criterion | | 16.85751 |
| Sum squared resid | 9.08E+08 | Schwarz criterion | | 16.90648 |
| Log likelihood | -6364.138 | Hannan-Quinn criter. | | 16.87637 |
| F-statistic | 17.96362 | Durbin-Watson stat | | 0.048654 |
| Prob(F-statistic) | 0.000000 |  |  |  |

**Table 6:** White’s HETCOV procedure to correct heteroscedasticity

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: NG\_CON | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/06/14 Time: 22:09 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
| White heteroskedasticity-consistent standard errors & covariance | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -544.1195 | 32.72916 | -16.62492 | 0.0000 |
| POP | 3.726407 | 0.266662 | 13.97425 | 0.0000 |
| GDP | 0.165990 | 0.010986 | 15.10909 | 0.0000 |
| UNEMP | -2.171742 | 0.834307 | -2.603050 | 0.0094 |
| COAL\_CON | -0.177541 | 0.015993 | -11.10126 | 0.0000 |
| OIL\_PR | 7.803850 | 1.007753 | 7.743810 | 0.0000 |
| NG\_PR | -16.57115 | 2.099283 | -7.893719 | 0.0000 |
| CPI | -2.380516 | 0.146582 | -16.24015 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.971455 | Mean dependent var | | 304.6971 |
| Adjusted R-squared | 0.971188 | S.D. dependent var | | 155.7872 |
| S.E. of regression | 26.44331 | Akaike info criterion | | 9.398409 |
| Sum squared resid | 523038.2 | Schwarz criterion | | 9.447383 |
| Log likelihood | -3544.599 | Hannan-Quinn criter. | | 9.417273 |
| F-statistic | 3636.665 | Durbin-Watson stat | | 0.019653 |
| Prob(F-statistic) | 0.000000 | Wald F-statistic | | 9246.056 |
| Prob(Wald F-statistic) | 0.000000 |  |  |  |

**Table 7:** Dublin-Watson (DW) test

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: D(NG\_CON) | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/06/14 Time: 21:47 | | |  |  |
| Sample (adjusted): 1949M02 2011M12 | | | |  |
| Included observations: 755 after adjustments | | | |  |
| HAC standard errors & covariance (Bartlett kernel, Newey-West fixed | | | | |
| bandwidth = 7.0000) | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.784262 | 6.389514 | 0.122742 | 0.9023 |
| POP | 0.008458 | 0.052249 | 0.161882 | 0.8714 |
| GDP | 0.000904 | 0.002454 | 0.368485 | 0.7126 |
| UNEMP | -0.358700 | 0.169970 | -2.110376 | 0.0352 |
| COAL\_CON | -0.003197 | 0.003373 | -0.947659 | 0.3436 |
| OIL\_PR | -0.104030 | 0.200147 | -0.519767 | 0.6034 |
| NG\_PR | -0.258914 | 0.382721 | -0.676508 | 0.4989 |
| CPI | 0.018679 | 0.033766 | 0.553198 | 0.5803 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.063883 | Mean dependent var | | 0.826709 |
| Adjusted R-squared | 0.055111 | S.D. dependent var | | 3.185633 |
| S.E. of regression | 3.096607 | Akaike info criterion | | 5.109031 |
| Sum squared resid | 7162.965 | Schwarz criterion | | 5.158056 |
| Log likelihood | -1920.659 | Hannan-Quinn criter. | | 5.127915 |
| F-statistic | 7.282449 | Durbin-Watson stat | | 1.243115 |
| Prob(F-statistic) | 0.000000 | Wald F-statistic | | 3.453282 |
| Prob(Wald F-statistic) | 0.001201 |  |  |  |

**Table 8:** Lagrange Multiplier (LM) test

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Breusch-Godfrey Serial Correlation LM Test: | | | |  |
|  |  |  |  |  |
|  |  |  |  |  |
| F-statistic | 37815.78 | Prob. F(1,747) | | 0.0000 |
| Obs\*R-squared | 741.3555 | Prob. Chi-Square(1) | | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Test Equation: | |  |  |  |
| Dependent Variable: RESID | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/06/14 Time: 21:50 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
| Presample missing value lagged residuals set to zero. | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 3.153116 | 5.527325 | 0.570460 | 0.5685 |
| POP | -0.012674 | 0.044943 | -0.282002 | 0.7780 |
| GDP | -9.72E-05 | 0.001884 | -0.051563 | 0.9589 |
| UNEMP | -0.278809 | 0.143740 | -1.939676 | 0.0528 |
| COAL\_CON | 0.002744 | 0.002364 | 1.161010 | 0.2460 |
| OIL\_PR | -0.148471 | 0.124409 | -1.193418 | 0.2331 |
| NG\_PR | 0.307615 | 0.279495 | 1.100610 | 0.2714 |
| CPI | -0.012439 | 0.024043 | -0.517367 | 0.6051 |
| RESID(-1) | 0.997292 | 0.005128 | 194.4628 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.980629 | Mean dependent var | | 2.56E-13 |
| Adjusted R-squared | 0.980422 | S.D. dependent var | | 26.32044 |
| S.E. of regression | 3.682837 | Akaike info criterion | | 5.457077 |
| Sum squared resid | 10131.78 | Schwarz criterion | | 5.512173 |
| Log likelihood | -2053.775 | Hannan-Quinn criter. | | 5.478299 |
| F-statistic | 4726.972 | Durbin-Watson stat | | 1.238792 |
| Prob(F-statistic) | 0.000000 |  |  |  |

**Table 9:** Newey-West HAC standard error

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: NG\_CON | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/06/14 Time: 21:51 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
| HAC standard errors & covariance (Bartlett kernel, Newey-West fixed | | | | |
| bandwidth = 7.0000) | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -544.1195 | 81.99607 | -6.635922 | 0.0000 |
| POP | 3.726407 | 0.667989 | 5.578547 | 0.0000 |
| GDP | 0.165990 | 0.027411 | 6.055520 | 0.0000 |
| UNEMP | -2.171742 | 2.083651 | -1.042277 | 0.2976 |
| COAL\_CON | -0.177541 | 0.040565 | -4.376752 | 0.0000 |
| OIL\_PR | 7.803850 | 2.294130 | 3.401660 | 0.0007 |
| NG\_PR | -16.57115 | 4.615011 | -3.590706 | 0.0004 |
| CPI | -2.380516 | 0.370193 | -6.430476 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.971455 | Mean dependent var | | 304.6971 |
| Adjusted R-squared | 0.971188 | S.D. dependent var | | 155.7872 |
| S.E. of regression | 26.44331 | Akaike info criterion | | 9.398409 |
| Sum squared resid | 523038.2 | Schwarz criterion | | 9.447383 |
| Log likelihood | -3544.599 | Hannan-Quinn criter. | | 9.417273 |
| F-statistic | 3636.665 | Durbin-Watson stat | | 0.019653 |
| Prob(F-statistic) | 0.000000 | Wald F-statistic | | 1548.499 |
| Prob(Wald F-statistic) | 0.000000 |  |  |  |

**Table 10:** Correlation Matrix

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | POP | GDP | UNEMP | COAL\_CON | OIL\_PR | NG\_PR | CPI |
| POP | 1 | 0.99 | 0.33 | 0.98 | 0.77 | 0.85 | 0.97 |
| GDP | 0.99 | 1 | 0.25 | 0.96 | 0.8 | 0.89 | 0.99 |
| UNEMP | 0.33 | 0.25 | 1 | 0.27 | 0.39 | 0.21 | 0.29 |
| COAL\_CON | 0.98 | 0.96 | 0.27 | 1 | 0.68 | 0.83 | 0.96 |
| OIL\_PR | 0.77 | 0.8 | 0.39 | 0.68 | 1 | 0.87 | 0.81 |
| NG\_PR | 0.85 | 0.89 | 0.21 | 0.83 | 0.87 | 1 | 0.88 |
| CPI | 0.97 | 0.99 | 0.29 | 0.96 | 0.81 | 0.88 | 1 |

**Table 11:** Variance Inflation Factors (VIF)

|  |  |  |  |
| --- | --- | --- | --- |
| Variance Inflation Factors | | |  |
| Date: 11/06/14 Time: 22:00 | | |  |
| Sample: 1949M01 2011M12 | | |  |
| Included observations: 756 | | |  |
|  |  |  |  |
|  |  |  |  |
|  | Coefficient | Uncentered | Centered |
| Variable | Variance | VIF | VIF |
|  |  |  |  |
|  |  |  |  |
| C | 6723.356 | 2231.145 | NA |
| POP | 0.446209 | 6765.411 | 441.9940 |
| GDP | 0.000751 | 1563.202 | 584.6054 |
| UNEMP | 4.341602 | 42.74536 | 3.825066 |
| COAL\_CON | 0.001645 | 474.4775 | 196.0556 |
| OIL\_PR | 5.263033 | 41.08919 | 27.17057 |
| NG\_PR | 21.29832 | 43.78733 | 30.38705 |
| CPI | 0.137043 | 492.0945 | 227.9822 |

**Table 12:** Structural break test using dummy variable for 1979

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: NG\_CON | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/19/14 Time: 22:15 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
| White heteroskedasticity-consistent standard errors & covariance | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -2532.204 | 238.4692 | -10.61858 | 0.0000 |
| DUM79 | 2073.862 | 240.3944 | 8.626916 | 0.0000 |
| POP | 15.25134 | 1.423545 | 10.71363 | 0.0000 |
| GDP | -0.057172 | 0.030947 | -1.847400 | 0.0651 |
| UNEMP | -6.318344 | 1.903304 | -3.319671 | 0.0009 |
| COAL\_CON | -0.047157 | 0.034434 | -1.369498 | 0.1713 |
| OIL\_PR | 12.06416 | 1.196674 | 10.08141 | 0.0000 |
| NG\_PR | -11.08937 | 2.320964 | -4.777915 | 0.0000 |
| CPI | -5.680922 | 0.484006 | -11.73729 | 0.0000 |
| DUM79\*POP | -13.14316 | 1.441226 | -9.119434 | 0.0000 |
| DUM79\*GDP | 0.179062 | 0.046116 | 3.882840 | 0.0001 |
| DUM79\*UNEMP | 7.682834 | 2.021243 | 3.801044 | 0.0002 |
| DUM79\*COAL\_CON | -0.068527 | 0.069513 | -0.985820 | 0.3245 |
| DUM79\*OIL\_PR | -202.9971 | 16.96022 | -11.96901 | 0.0000 |
| DUM79\*NG\_PR | -183.9071 | 16.14802 | -11.38883 | 0.0000 |
| DUM79\*CPI | 14.33454 | 1.089201 | 13.16059 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.983260 | Mean dependent var | | 304.6971 |
| Adjusted R-squared | 0.982921 | S.D. dependent var | | 155.7872 |
| S.E. of regression | 20.35937 | Akaike info criterion | | 8.885897 |
| Sum squared resid | 306733.0 | Schwarz criterion | | 8.983845 |
| Log likelihood | -3342.869 | Hannan-Quinn criter. | | 8.923624 |
| F-statistic | 2897.735 | Durbin-Watson stat | | 0.056598 |
| Prob(F-statistic) | 0.000000 | Wald F-statistic | | 6265.999 |
| Prob(Wald F-statistic) | 0.000000 |  |  |  |

**Table 13:** Wald test for DUM79 and DUM79\*POP

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 45.26390 | (2, 740) | 0.0000 |
| Chi-square | 90.52780 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(10)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 2073.862 | 240.3944 |
| C(10) | | -13.14316 | 1.441226 |

**Table 14:** Wald test for DUM79 and DUM79\*GDP

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 38.43614 | (2, 740) | 0.0000 |
| Chi-square | 76.87228 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(11)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 2073.862 | 240.3944 |
| C(11) | | 0.179062 | 0.046116 |

**Table 15:** Wald test for DUM79 and DUM79\*UNEMP

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 37.22738 | (2, 740) | 0.0000 |
| Chi-square | 74.45476 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(12)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 2073.862 | 240.3944 |
| C(12) | | 7.682834 | 2.021243 |

**Table 16:** Wald test for DUM79 and DUM79\*COAL\_CON

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 37.23591 | (2, 740) | 0.0000 |
| Chi-square | 74.47182 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(13)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 2073.862 | 240.3944 |
| C(13) | | -0.068527 | 0.069513 |

**Table 17:** Wald test for DUM79 and DUM79\*OIL\_PR

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 104.0862 | (2, 740) | 0.0000 |
| Chi-square | 208.1724 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(14)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 2073.862 | 240.3944 |
| C(14) | | -202.9971 | 16.96022 |

**Table 18:** Wald test for DUM79 and DUM79\*NG\_PR

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 97.44944 | (2, 740) | 0.0000 |
| Chi-square | 194.8989 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(15)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 2073.862 | 240.3944 |
| C(15) | | -183.9071 | 16.14802 |

**Table 19:** Wald test for DUM79 and DUM79\*CPI

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 97.65894 | (2, 740) | 0.0000 |
| Chi-square | 195.3179 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(16)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 2073.862 | 240.3944 |
| C(16) | | 14.33454 | 1.089201 |

**Table 20:** Corrected model for structural break at 1979

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: NG\_CON | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/19/14 Time: 22:42 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
| White heteroskedasticity-consistent standard errors & covariance | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -2521.228 | 236.6225 | -10.65506 | 0.0000 |
| DUM79 | 2077.396 | 240.2852 | 8.645541 | 0.0000 |
| POP | 15.24079 | 1.419551 | 10.73635 | 0.0000 |
| GDP | -0.058524 | 0.030879 | -1.895295 | 0.0584 |
| UNEMP | -6.666656 | 1.835202 | -3.632655 | 0.0003 |
| COAL\_CON | -0.057295 | 0.030710 | -1.865689 | 0.0625 |
| OIL\_PR | 11.80325 | 1.130555 | 10.44022 | 0.0000 |
| NG\_PR | -10.66337 | 2.263453 | -4.711105 | 0.0000 |
| CPI | -5.590834 | 0.458406 | -12.19624 | 0.0000 |
| DUM79\*POP | -13.18243 | 1.439889 | -9.155173 | 0.0000 |
| DUM79\*GDP | 0.152189 | 0.038994 | 3.902906 | 0.0001 |
| DUM79\*UNEMP | 7.843378 | 1.992420 | 3.936609 | 0.0001 |
| DUM79\*OIL\_PR | -207.7430 | 15.14376 | -13.71806 | 0.0000 |
| DUM79\*NG\_PR | -187.9131 | 14.43071 | -13.02175 | 0.0000 |
| DUM79\*CPI | 14.49274 | 1.053088 | 13.76214 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.983246 | Mean dependent var | | 304.6971 |
| Adjusted R-squared | 0.982930 | S.D. dependent var | | 155.7872 |
| S.E. of regression | 20.35404 | Akaike info criterion | | 8.884077 |
| Sum squared resid | 306986.5 | Schwarz criterion | | 8.975904 |
| Log likelihood | -3343.181 | Hannan-Quinn criter. | | 8.919447 |
| F-statistic | 3106.301 | Durbin-Watson stat | | 0.057343 |
| Prob(F-statistic) | 0.000000 | Wald F-statistic | | 6639.740 |
| Prob(Wald F-statistic) | 0.000000 |  |  |  |

**Table 21:** Structural break test using dummy variable for 2008

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: NG\_CON | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/19/14 Time: 23:06 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
| White heteroskedasticity-consistent standard errors & covariance | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -1384.175 | 42.68809 | -32.42533 | 0.0000 |
| DUM08 | 735.2132 | 56.91341 | 12.91810 | 0.0000 |
| POP | 4.401030 | 0.149965 | 29.34703 | 0.0000 |
| GDP | -0.004805 | 0.000910 | -5.282615 | 0.0000 |
| UNEMP | -0.207570 | 0.125159 | -1.658450 | 0.0977 |
| COAL\_CON | 0.457789 | 0.002808 | 163.0423 | 0.0000 |
| OIL\_PR | 8.767933 | 0.142595 | 61.48830 | 0.0000 |
| NG\_PR | -38.53830 | 0.369519 | -104.2931 | 0.0000 |
| CPI | 0.059343 | 0.010629 | 5.582892 | 0.0000 |
| DUM08\*POP | -0.024808 | 0.330638 | -0.075030 | 0.9402 |
| DUM08\*GDP | 0.176106 | 0.011322 | 15.55371 | 0.0000 |
| DUM08\*UNEMP | -1.161274 | 0.937707 | -1.238419 | 0.2160 |
| DUM08\*COAL\_CON | -0.740555 | 0.023086 | -32.07773 | 0.0000 |
| DUM08\*OIL\_PR | 7.056572 | 1.368376 | 5.156896 | 0.0000 |
| DUM08\*NG\_PR | 1.983094 | 2.679043 | 0.740225 | 0.4594 |
| DUM08\*CPI | -1.879594 | 0.162374 | -11.57567 | 0.0000 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.975324 | Mean dependent var | | 304.6971 |
| Adjusted R-squared | 0.974824 | S.D. dependent var | | 155.7872 |
| S.E. of regression | 24.71863 | Akaike info criterion | | 9.273928 |
| Sum squared resid | 452147.8 | Schwarz criterion | | 9.371876 |
| Log likelihood | -3489.545 | Hannan-Quinn criter. | | 9.311656 |
| F-statistic | 1949.931 | Durbin-Watson stat | | 0.022801 |
| Prob(F-statistic) | 0.000000 | Wald F-statistic | | 373196.6 |
| Prob(Wald F-statistic) | 0.000000 |  |  |  |

**Table 22:** Wald test for DUM08 and DUM08\*POP

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 552.0954 | (2, 740) | 0.0000 |
| Chi-square | 1104.191 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(10)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 735.2132 | 56.91341 |
| C(10) | | -0.024808 | 0.330638 |

**Table 23:** Wald test for DUM08 and DUM08\*GDP

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 144.4412 | (2, 740) | 0.0000 |
| Chi-square | 288.8825 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(11)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 735.2132 | 56.91341 |
| C(11) | | 0.176106 | 0.011322 |

**Table 24:** Wald test for DUM08 and DUM08\*UNEMP

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 103.0151 | (2, 740) | 0.0000 |
| Chi-square | 206.0302 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(12)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 735.2132 | 56.91341 |
| C(12) | | -1.161274 | 0.937707 |

**Table 25:** Wald test for DUM08 and DUM08\*COAL\_CON

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 1489.925 | (2, 740) | 0.0000 |
| Chi-square | 2979.851 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(13)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 735.2132 | 56.91341 |
| C(13) | | -0.740555 | 0.023086 |

**Table 26:** Wald test for DUM08 and DUM08\*OIL\_PR

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 88.76903 | (2, 740) | 0.0000 |
| Chi-square | 177.5381 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(14)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 735.2132 | 56.91341 |
| C(14) | | 7.056572 | 1.368376 |

**Table 27:** Wald test for DUM08 and DUM08\*NG\_PR

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 88.15482 | (2, 740) | 0.0000 |
| Chi-square | 176.3096 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(15)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 735.2132 | 56.91341 |
| C(15) | | 1.983094 | 2.679043 |

**Table 28:** Wald test for DUM08 and DUM08\*CPI

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAKTEST08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 100.8026 | (2, 740) | 0.0000 |
| Chi-square | 201.6053 | 2 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis: C(2)=C(16)=0 | | |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(2) | | 735.2132 | 56.91341 |
| C(16) | | -1.879594 | 0.162374 |

**Table 29:** OLS Model after fixing break points in 1979 and 2008

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: NG\_CON | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 11/23/14 Time: 15:35 | | |  |  |
| Sample: 1949M01 2011M12 | | |  |  |
| Included observations: 756 | | |  |  |
| White Heteroskedasticity-Consistent Standard Errors & Covariance | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | -2870.633 | 234.8708 | -12.22218 | 0.0000 |
| DUM79\_08 | 2622.821 | 235.9842 | 11.11439 | 0.0000 |
| POP | 17.58079 | 1.381068 | 12.72985 | 0.0000 |
| GDP | -0.080603 | 0.031744 | -2.539151 | 0.0113 |
| UNEMP | -7.141672 | 2.180344 | -3.275479 | 0.0011 |
| COAL\_CON | -0.230269 | 0.051624 | -4.460507 | 0.0000 |
| OIL\_PR | 21.80490 | 2.332273 | 9.349203 | 0.0000 |
| NG\_PR | -32.01191 | 3.860464 | -8.292242 | 0.0000 |
| CPI | -5.065888 | 0.545606 | -9.284888 | 0.0000 |
| DUM79\_08\*POP | -16.85412 | 1.396701 | -12.06710 | 0.0000 |
| DUM79\_08\*GDP | 0.631625 | 0.043388 | 14.55773 | 0.0000 |
| DUM79\_08\*UNEMP | 14.80618 | 2.400236 | 6.168638 | 0.0000 |
| DUM79\_08\*COAL\_CON | -0.233476 | 0.060699 | -3.846476 | 0.0001 |
| DUM79\_08\*OIL\_PR | -22.00661 | 2.957071 | -7.442029 | 0.0000 |
| DUM79\_08\*NG\_PR | 35.52548 | 5.086138 | 6.984765 | 0.0000 |
| DUM79\_08\*CPI | -1.363652 | 0.662584 | -2.058080 | 0.0399 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.982476 | Mean dependent var | | 304.6971 |
| Adjusted R-squared | 0.982120 | S.D. dependent var | | 155.7872 |
| S.E. of regression | 20.83108 | Akaike info criterion | | 8.931706 |
| Sum squared resid | 321111.0 | Schwarz criterion | | 9.029654 |
| Log likelihood | -3360.185 | Hannan-Quinn criter. | | 8.969433 |
| F-statistic | 2765.778 | Durbin-Watson stat | | 0.058322 |
| Prob(F-statistic) | 0.000000 |  |  |  |

**Table 30:** Wald test on POP, GDP, UNEMP and CPI (test 1)

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAK\_79\_08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 426.7011 | (4, 740) | 0.0000 |
| Chi-square | 1706.804 | 4 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(3) | | 17.58079 | 1.381068 |
| C(4) | | -0.080603 | 0.031744 |
| C(5) | | -7.141672 | 2.180344 |
| C(9) | | -5.065888 | 0.545606 |
|  |  |  |  |
|  |  |  |  |
| Restrictions are linear in coefficients. | | | |

**Table 31:** Wald test on POP, GDP, UNEMP and CPI (test 2)

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAK\_79\_08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 11.39029 | (1, 740) | 0.0008 |
| Chi-square | 11.39029 | 1 | 0.0007 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(3) + C(4) + C(5) + C(9) | | 5.292625 | 1.568208 |
|  |  |  |  |
|  |  |  |  |
| Restrictions are linear in coefficients. | | | |

**Table 32:** Wald test on COAL\_CON, OIL\_PR and NG\_PR (test 1)

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAK\_79\_08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 70.54951 | (3, 740) | 0.0000 |
| Chi-square | 211.6485 | 3 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(6) | | -0.230269 | 0.051624 |
| C(7) | | 21.80490 | 2.332273 |
| C(8) | | -32.01191 | 3.860464 |
|  |  |  |  |
|  |  |  |  |
| Restrictions are linear in coefficients. | | | |

**Table 33:** Wald test on COAL\_CON, OIL\_PR and NG\_PR (test 2)

|  |  |  |  |
| --- | --- | --- | --- |
| Wald Test: | |  |  |
| Equation: LINEAR\_BREAK\_79\_08 | | |  |
|  |  |  |  |
|  |  |  |  |
| Test Statistic | Value | df | Probability |
|  |  |  |  |
|  |  |  |  |
| F-statistic | 19.06956 | (1, 740) | 0.0000 |
| Chi-square | 19.06956 | 1 | 0.0000 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| Null Hypothesis Summary: | | |  |
|  |  |  |  |
|  |  |  |  |
| Normalized Restriction (= 0) | | Value | Std. Err. |
|  |  |  |  |
|  |  |  |  |
| C(6) + C(7) + C(8) | | -10.43728 | 2.390104 |
|  |  |  |  |
|  |  |  |  |
| Restrictions are linear in coefficients. | | | |