A Short Note on the Application of Chow Test of Structural Break in US GDP

Hari S. Luitel[1] & Gerry J. Mahar[1]

[1] Department of Business & Economics, Algoma University, Sault Ste. Marie, Ontario, Canada

Correspondence: Hari S. Luitel, Department of Business & Economics, Algoma University, Sault Ste. Marie, Ontario, P6A 2G4, Canada. E-mail: hari.luitel@algomau.ca

Abstract

In 1997, the North American Industry Classification System (NAICS)--the standard for use by Federal statistical agencies in classifying business establishments for the collection, tabulation, presentation, and analysis of statistical data describing the US economy--replaced the Standard Industrial Classification (SIC) system. There are substantial differences in the method of data reporting under the SIC system and the under NAICS. In this short note, using a Chow Test, we formally provide evidence of a structural break in the US GDP data due to the switch from the SIC reporting system to the NAICS reporting system.

Keywords: Chow test, structural break, US GDP, statistical methods, econometric methods, single equation model

1. Introduction

In 1997, The US Department of Commerce Bureau of Economic Analysis (BEA) switched reporting gross domestic product (GDP) and other national accounts from Standard Industrial Classification (SIC) System to North American Industry Classification System (NAICS). The NAICS was developed jointly by the US Economic Classification Policy Committee (ECPC), Statistics Canada, and Mexico's Instituto Nacional de Estadistica y Geografia (INEGI), to allow for a high level of comparability in business statistics among North American countries. The NAICS was constructed within a supply-based, or production-oriented, conceptual framework where establishments using similar production processes to produce goods and services are grouped to form industries. The NAICS allows for the identification of 1,170 industries compared to the 1,004 found in the SIC system. The increase in the number of categories was substantial. For a detail discussion of the differences between the SIC and the NAICS, see Issue Papers 1 through 6 of Economic Classification Policy Committee (1993a, 1993b, 1993c, 1993d, 1993e, and 1993f). Our objective in this short note was to apply the Chow Test of a structural break to determine if this test would be capable of identifying a break in the time series data of the US GDP due to the switch from the SIC reporting system to the NAICS reporting system in 1997.

A time series data set contains observations ordered and recorded in time for the same variable. In macroeconomics, a structural break occurs when there is an unexpected shift in the data of a time series. Knowledge of a structural break in the time series data of US GDP is important for a number of reasons: Firstly, a structural break may affect any or all of the underlying model parameters which have different implications. For example, this can lead to forecasting errors and to unreliability of the model in general. Researchers using the time series data of US GDP could easily reach quite opposite conclusions--hardly an example of sound scientific practice (Note 1). Secondly, notwithstanding the fact that the time series data of US GDP has been one of the most widely studied and reported macroeconomic variables, it continues to be closely monitored as a leading indicator and measure by investors, academics as well as government officials all around the world. Thirdly, many leading econometrics textbooks have used the data series of US GDP to illustrate examples of various time series analysis in the economics curriculum and many influential articles in economics journals such as those by Nelson and Plosser (1982), Engel and Granger (1987), Perron (1989), Zivot and Andrews (1992), to name a few, have used the data series of US GDP in their analysis. Lastly, the National Bureau of Economic Research (NBER) uses the data series of US GDP to declare whether and/or when the US economy enters and exits a recession (Note 2). In this note, we hope to shed some light on previously unexplored characteristics of the time series data of US GDP that might be helpful to others working with the GDP measure in academia and those employed in

industry or in government.

2. Empirical Analysis and Results

From 1973 to 2014, the US GDP data covered two different time spans (Note 3). Up to 1996, the BEA reported US GDP data according to the Standard Industrial Classification (SIC) System. In 1997, the SIC system was replaced by the North American Industry Classification System (NAICS). There were substantial differences in the method of data reporting under the SIC system and under the NAICS. Therefore, we suspected that the parametric values governing the data generating process of US GDP under these two regimes would likely be different. To investigate further, we obtained annual data of US GDP for the period from 1973 to 2014. Data was gathered from the official website of the United States Department of Commerce, Bureau of Economic Analysis (BEA) (Note 4).

Visual inspection of the data is usually a first step in an analysis of structural break. *As can be seen in Figure 1, a visual inspection of the actual data of US GDP did not show a break.* Thus, we need to use a Chow test (F-test) to determine if a structural break had occurred.
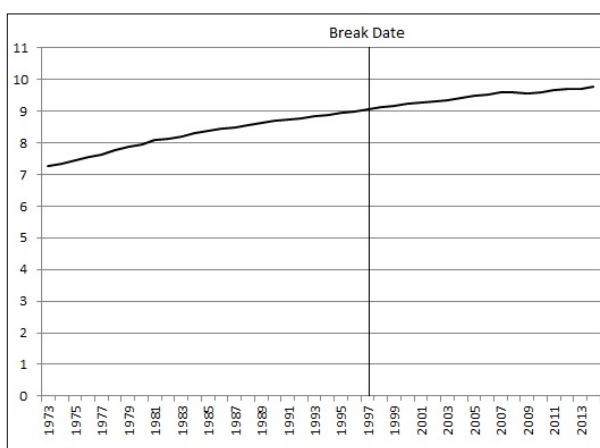

Figure 1. Log of US GDP (1973-2014)

Source: Bureau of Economic Analysis, http://www.bea.gov/.

Web access date: September 18, 2015.

A linear time-trend model is a simple regression model in which the independent variable is the raw index or any ascending sequence of equally spaced numbers (daily, weekly, monthly, quarterly, and annually etc.). A growth rate time-trend model includes the underlying rate of economic growth that can be maintained without inflationary pressures. The choice between a linear time-trend model and a growth rate time-trend model depends on whether one is interested in the absolute or relative change in GDP. For comparative purposes, the relative change is usually of greater interest to economists and more important than an absolute change in GDP. Therefore, we estimated a semi-log linear time-trend model of the form:

$$ln(Y_t) = \beta_0 + \beta_1 X_t + u_{1t} \tag{1}$$

where $ln(Y_t)$ represents natural log of US GDP, X_t is the time variable that varies from 1973 to 2014 and u_{1t} is the error term. Equation (1) is called a semi-log model because only the regressand appears in the logarithmic form.

We know the exact date when the BEA switched reporting the GDP and other national accounts from the SIC to the NAICS and we can use ordinary least squares (OLS) regression to estimate the parameters in equation (1). Our objective was to determine if the switch in the method of data reporting from the SIC to the NAICS had changed the model parameters. We posed the question: Did the regression coefficients--β_0 and β_1--remain stable over the entire time period? To address this question, we divided the data into two time periods, 1973-1996 and 1997-2014, and estimated the two separate regressions as follows:

$$\textbf{Period 1973-1996: } ln(Y_t) = \alpha_0 + \alpha_1 X_t + u_{2t} \tag{2}$$

$$\textbf{Period 1997-2014: } ln(Y_t) = \delta_0 + \delta_1 X_t + u_{3t} \tag{3}$$

where u_{2t} and u_{3t} are assumed iid.

If the data generating processes under the SIC and the NAICS systems were to be the same, we would expect that $\beta_0 = \alpha_0 = \delta_0$ (i.e. the intercept would not be statistically significantly different from each other) and $\beta_1 = \alpha_1 = \delta_1$ (i.e., the slope coefficients would not be statistically significantly different from each other). To examine this, we estimated three regressions, equations (1), (2), and (3) above, and obtained their respective residuals sums of squares. The results are reported in Table 1. Note that there were 42 observations: 24 in the 1973-1996 time period and 18 in 1997-2014 time period.

Table 1. OLS regression results

Regression under the assumption of parametric stability	Regressions under the assumption of parametric variability	
Time Period 1973-2014:	Time Period 1973-1996:	Time Period 1997-2014:
$ln(Y_t) = \beta_0 + \beta_1 X_t + u_{1t}$	$ln(Y_t) = \alpha_0 + \alpha_1 X_t + u_{2t}$	$ln(Y_t) = \delta_0 + \delta_1 X_t + u_{3t}$
$ln(\hat{Y}) = $ -109.834 + 0.059X_t	$ln(\hat{Y}) = $ -142.154 + 0.075X_t	$ln(\hat{Y}) = $ -71.239 + 0.040X_t
t = (33.93) (36.64)	t = (32.29) (34.16)	t = (21.31) (24.14)
$n = n_1 + n_2 = 42$	$n_1 = 24$	$n_2 = 18$
$R^2 = 0.9711$	$R^2 = 0.9815$	$R^2 = 0.9733$
$RSSR = 0.6506$	$SSR_1 = 0.1245$	$SSR_2 = 0.0215$
df = 40	df = 22	df = 16

Note. Figures in the parenthesis are absolute t statistics.

Source: Authors calculations.

Under the assumption that there would be stability of parameters (i.e., the regression coefficients did not change over the period), the residual sum of squares obtained from equation (1) is called the restricted residual sum of squares (RSSR), and will have (n-k) degrees of freedom (*df*), where $n = n_1 + n_2$. The assumption that the error terms (u_{2t} and u_{3t}) in equation (2) and equation (3) are iid implies that the two samples (i.e. sample periods 1973-1996 and 1997-2014) are independent. By adding the residual sum of squares from equation (2) and equation (3), we get the unrestricted residual sum of squares (USSR), which has ($n_1 + n_2 - 2k$) *df*, where n_1 and n_2 are the number of observations in the first and second periods and k is the number of parameters estimated in each model (two in our example). If parameters are stable, RSSR and USSR will not be statistically different. Alternatively, if there was no stability of parameters, the two residual sums of squares would differ. This can be tested as below:

$$F = \frac{(RSSR - SSR_1 - SSR_2)/k}{(SSR_1 + SSR_2)/(n-2k)} \sim F_{[k,(n-2k)]} \tag{4}$$

The F test above, or the variance ratio test, is often referred to in econometrics as the "Chow test," due to Chow (1960). This test statistic has k and (n-2k) degrees of freedom because the restricted regression model has k parameters whereas the unrestricted regression model has 2k parameters. It will be exactly distributed as F(k, n-2k) if the error terms were normal and independent of the fixed regressors X, and it will be asymptotically distributed as $x^2(k)$ under much weaker conditions (MacKinnon 1989, pages 78-79). If the computed F statistic is not statistically significant, say, at 1 percent, 5 percent or 10 percent level of significance, it would indicate that the parameters were stable. However, if the computed F statistics were statistically significant, we would reject the hypothesis that there was stability of parameters and it would mean that the data generating process of the US GDP had changed over time.

As reported in Table 1, the various sums of residuals squares are as follows: RSSR = 0.6506; SSR_1 = 0.1245; SSR_2 = 0.0215. Inserting these values in equation (4), we obtain:

$$\frac{(RSSR - SSR_1 - SSR_2)/k}{(SSR_1 + SSR_2)/(n-2k)} = \frac{(0.6506 - 0.1245 - 0.0215)/2}{(0.1245 + 0.0215)/(42 - 2 \times 2)} = 65.67 \tag{5}$$

In our example, this *F* statistics follows the F distribution with 2 and 38 degrees of freedom in the numerator and denominator. The 1 percent critical value is close to 5.18 as reported in Green (2000; Table B5, page 961). Because the computed F value far exceeds the tabulated critical value, we would reject the hypothesis of stability of parameters (i.e. reject H_0: $\beta_0 = \alpha_0 = \delta_0$ and $\beta_1 = \alpha_1 = \delta_1$). The results indicate that the models for the two periods are systematically different, beyond a simple switch from the SIC to the NAICS and we can conclude that the data generating process under the SIC and under the NAICS was not the same.

3. Summary and Conclusion

In 1997, the US Department of Commerce Bureau of Economic Analysis (BEA) switched reporting gross domestic product (GDP) and other national accounts from Standard Industrial Classification System (SIC) to North American Industry Classification System (NAICS). Because there are substantial differences in the method of data reporting under the SIC system and the NAICS, we were interested to determine if the switch in the method of data reporting from the SIC to the NAICS changed the model parameters of US GDP. In this short note, we applied a Chow Test of structural break. The test results showed a structural break in the US GDP time series data due to the switch from the SIC reporting system to the NAICS reporting system.

Our research findings have several implications. First, if the structural break in the US GDP of 1997 is not properly taken into account, researchers in public policy using US GDP time series data may easily arrive at misleading conclusions. Second, it is an interesting research question for researchers and practitioners working with the US GDP measure to know how any forecast errors due to a structural break in the US GDP of 1997 would compare with its actual observation either before and/or after 1997. This needs further investigation. Above all unit root property of the data is inconsistent with the structural break. A structural break implies that the parameter values governing the data generating process have changed. However, the current literature on unit root treats the unit root property of the data as not being affected by a structural break. The implications of the structural break findings in this paper in the context of the uniqueness of unit root and cointegration analysis further complement Luitel and Mahar (2015).

References

Chow, G. C. (1960). Test of equality between sets of coefficients in two linear regressions. *Econometrica, 28*(3), 591-605. http://dx.doi.org/10.2307/1910133

Economic Classification Policy Committee. (1993a). *Conceptual issues.* Issue Paper No. 1, Bureau of Economic Analysis, Department of Commerce, Washington D. C.

Economic Classification Policy Committee. (1993b). *Aggregation structure and hierarchies.* Issue Paper No. 2, Bureau of Economic Analysis, Department of Commerce, Washington D.C.

Economic Classification Policy Committee. (1993c). *Collectiblility of data.* Issue Paper No. 3, Bureau of Economic Analysis, Department of Commerce, Washington D.C., May 1993.

Economic Classification Policy Committee. (1993d). *Criteria for determining industries.* Issue Paper No. 4, Bureau of Economic Analysis, Department of Commerce, Washington D.C., October 1993.

Economic Classification Policy Committee. (1993e). *The impact of classification revisions on time series.* Issue Paper No. 5, Bureau of Economic Analysis, Department of Commerce, Washington D.C., July 1993.

Economic Classification Policy Committee. (1993f). *Services classifications.* Issue Paper No. 6, Bureau of Economic Analysis, Department of Commerce, Washington D.C., September 1993.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica, 55*(2), 251-276. http://dx.doi.org/10.2307/1913236

Green, W. H. (2000). *Econometric analysis* (4th ed.). New Jersey: Prentice Hall, Inc.

Leontief, W. (1971). Theoretical assumptions and nonobserved facts. *The American Economic Review, 61*(1), 1-7.

Luitel, H. S. (2014). *Is Tax Amnesty a Good Tax Policy? Evidence from state tax amnesty programs in the United States.* Lanham: Lexington Books.

Luitel, H. S., & Mahar, G. J. (2015). *Why most published results on unit root and cointegration are false.* SSRN Working Paper, July 2015. http://dx.doi.org/10.2139/ssrn.2628645

MacKinnon, J. G. (1989). Heteroskedasticity-robust tests for structural change. *Empirical Economics, 14*(2),

77-92. http://dx.doi.org/10.1007/BF01980588

Nelson, C. R., & Plosser, C. I. (1982). Trends and random walks in macroeconomic time series. *Journal of Monetary Economics, 10*(2), 139-162. http://dx.doi.org/10.1016/0304-3932(82)90012-5

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica, 57*(6), 1361-1401. http://dx.doi.org/10.2307/1913712

Zivot, E., & Andrews, D. W. K. (1992). Further evidence on the great crash, the oil price shock, and the unit root hypothesis. *Journal of Business and Economic Statistics, 10*(3), 251-270.

Notes

Note 1. For a criticism of modern mainstream economic modeling approach, which fails to properly account for the non-observed facts, see Leontief (1971). For a recent example of criticism of mainstream economic modeling approach, see Luitel (2014).

Note 2. The financial press often defines a recession in terms of two consecutive quarters of decline in real GDP. According to NBER, "a recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales."

Note 3. Perron (1989) shows a structural break in US GNP in 1973. For our data analysis, we, therefore, consider only the time period since 1973.

Note 4. The BEA reports the data in billions of dollars. We first took the natural log of the data as reported by the BEA without other manipulation and then analyzed it. The data used in the analysis is available from the authors upon request.

Copyrights