Task 1

Name /ID

Date

# PART A

## Question 1

Considering the information provided on the system of equations of the IV regression, from time, , we begin with the following equations to derive the reduce their form:

Since in the function of we have the single right-hand-side endogenous regressor, we can substitute the function into as:

Putting like terms together and simplifying, we obtain:

However, we know the following conditions hold:

And

Therefore:

Where and have a normal distribution. We verify this distribution of the errors from the above equation as normally distributed (thus, expectation of 0, ). By letting we can derive the variance of the errors as:

Further, for and , the following relationships can be established:

## Question 2

From the general OLS covariance matrix, we have , , and as matrices in the following forms:

With the variance-covariance matrix of and

The reduced-form of the OLS estimation with the above parameters on Y is therefore in a linear form as:

However, we know that is matrix containing from the matrix and Y since it is derived that the OLS estimator for in terms of and matrices is given as in the form:

But

Therefore:

## Question 3

To derive the confidence set for in:

We use the regression on and , where is the predicted value from:

From , we see that limits the values of , and thus, the confidence set for are:

Assuming that is the 2-step OLS estimate of where is the percentile from the standard normal distribution, with the standard threshold of 5%.

## Question 4

We know that:

and the variance-covariance matrix of be denoted as . The null is tested upon and

or

The -statistic is therefore:

where .  
Based on Fieller method, The confidence set for is:

## Question 5

Anderson-Rubin confidence sets for in

is derived by solving for :

Where:

.

The roots are set as:

.

Hence,

## Question 6

The methods above will not give the same results. When there is a lack of a strong relationship between the instruments and the endogenous regressors, traditional methods of estimating and drawing conclusions for instrumental variables become incorrect. The vast majority of assumptions are often called into question by evidence collected in reality; nevertheless, this observation is only valid for data that is independent and homoscedastic. These results suggest that weak instruments continue to be a worry for empirical practise and that researchers may make simple steps to handle weak instruments more effectively in applications.

## Question 7

From the general knowledge of statistics, we know that under the that:

.

The statistic is closely approximated by a random variable, which has a conventional normal distribution, and its square root is a chi-square distribution with one unpredictability degree. For testing hypotheses at the 5% threshold level, this approximation supports the use of the usual normal critical values of 1.96. The crucial values , which represent the (quantile of the chi-squared distribution with one degree of freedom, are often used in the modified t-test tests.

The econometric literature has acknowledged, even for large samples, that t has a known non-normal distribution that, although sometimes "near" to the normal distribution, may occasionally diverge dramatically. This allows us to develop a formula for using 2SLS-based Wald test statistics. The results will demonstrate that under can be modelled using two jointly normal random variables and a single endogenous regressor. The two normal variables can be represented as and , where has a mean of:

is the asymptotic variance of and is a standard normal distribution with . The connection between and is identical to the connection between and .

Thus, the formula precisely quantifies the degree of inference bias caused by rejecting the null hypothesis with the criterion. This relationship is depicted with the argument that we have rejection probabilities, such as the probability that is greater than 1.96 if the null hypothesis is true.

# PART B

The data used has the following features as shown in table 1 below.

Table 1: Descriptive statistics for the entire data used

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Mean | Std. Dev. | Min | Max |
|  | - | - | - | - |
|  | .799 | .577 | -.431 | 2.899 |
|  | 4.686 | .042 | 4.588 | 4.768 |
|  | .576 | .005 | .564 | .586 |
|  | -.963 | 2.41 | -11.39 | 5.595 |
|  | .725 | 1.12 | -9.362 | 7.276 |
|  | .005 | .287 | -.888 | 1.339 |

We have seven variables in consideration, one being the date of observation for each parameter. To answer how has changed over time, we consider 3 sub-settings of the data; before 1984 (), between 1984 and 2007 (), and between 2008 and 2021.

Summary statistics for each subset are shown in table 2, with the period between 1984 to 2007 having the most observations. **Summary statistics: For each subset   
sub-set: 1**

Table : Summary statistics for the 3 subsets

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | N | mean | Standard deviation | min | max |
|  | 95 | 4382.589 | 2517.276 | 91 | 8674 |
|  | 95 | 1.198 | 0.699 | .132 | 2.899 |
|  | 95 | 4.717 | 0.018 | 4.687 | 4.768 |
|  | 95 | .58 | 0.002 | .576 | .586 |
|  | 95 | -.295 | 2.997 | -8.158 | 5.595 |
|  | 95 | .849 | 1.068 | -2.082 | 3.791 |
|  | 94 | .005 | 0.338 | -.888 | .927 |
|  | 95 | 1 | 0.000 | 1 | 1 |

**Sub-set: 2**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Observation date | 96 | 13102.938 | 2543.629 | 8766 | 17440 |
|  | 96 | .616 | 0.219 | .146 | 1.188 |
|  | 96 | 4.693 | 0.021 | 4.648 | 4.737 |
|  | 96 | .577 | 0.003 | .572 | .583 |
|  | 96 | -.733 | 1.384 | -3.647 | 2.168 |
|  | 96 | .796 | 0.506 | -.915 | 1.936 |
|  | 96 | -.003 | 0.165 | -.392 | .58 |
|  | 96 | 2 | 0.000 | 2 | 2 |

**sub-set: 3**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 54 | 19951.407 | 1436.524 | 17532 | 22371 |
|  | 54 | .425 | 0.306 | -.431 | 1.483 |
|  | 54 | 4.619 | 0.021 | 4.588 | 4.676 |
|  | 54 | .568 | 0.003 | .564 | .575 |
|  | 54 | -2.547 | 1.982 | -11.39 | .255 |
|  | 54 | .38 | 1.770 | -9.362 | 7.276 |
|  | 54 | .02 | 0.357 | -.756 | 1.339 |
|  | 54 | 3 | 0.000 | 3 | 3 |

Before deriving all the confidence sets, we first fit the New Keynesian Phillips curve equations for the 3 subsets. The results obtained are as shown in tables that follow.

Table : Subset 1 (Before 1984)

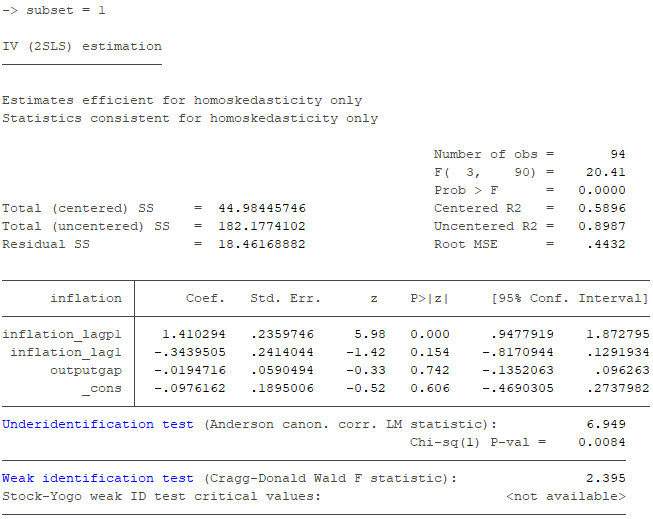


Table : Subset 2 (Between 1984 and 2007)

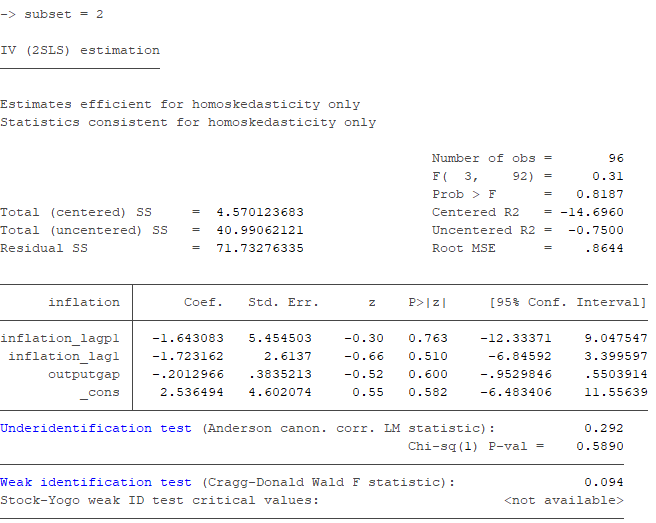
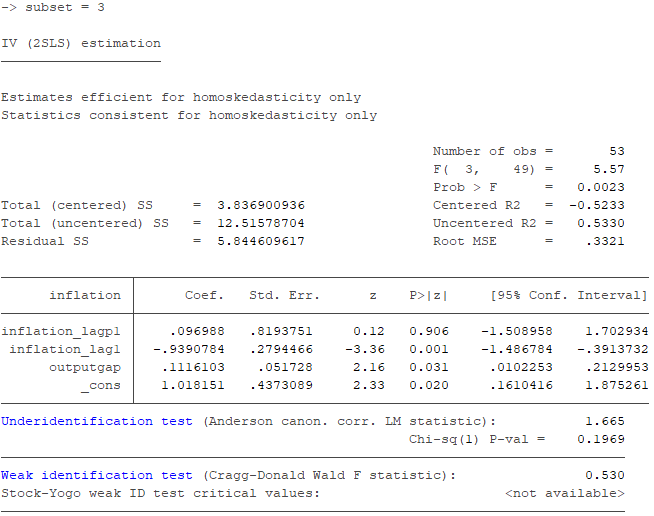


Table : Subset 3 (After 2008 until 2021)



For these 3 models, we derive the confidence interval sets (**Table 6**) for the coefficient of (that is ) using the command, to obtain the CLR; The null hypothesis, . It is presumable that the exogeneity constraints E(Zu)=0 are satisfied in this scenario. The range of B values over which the alternative hypothesis, beta=b0, cannot be rejected with a significance level of between 95% and 100% is referred to as the 95% confidence interval.

|  |  |  |
| --- | --- | --- |
| Sample | Confidence Interval Sets | |
| Subset 1 | -0.0789389 | 0.0567098 |
| Subset 2 | -0.0886202 | 0.0354146 |
| Subset 3 | 0.1081852 | 0.2896157 |

# PART C