Censored Models in EViews

Ass.Prof. Andriy Stavytskyy

Data

- File: Censor.wf1

Description

- Hrs – wife working hours per year
- Edu – years of education of wife
- Age – age of person in years
- Kid1 – number of children aged 0 to 5

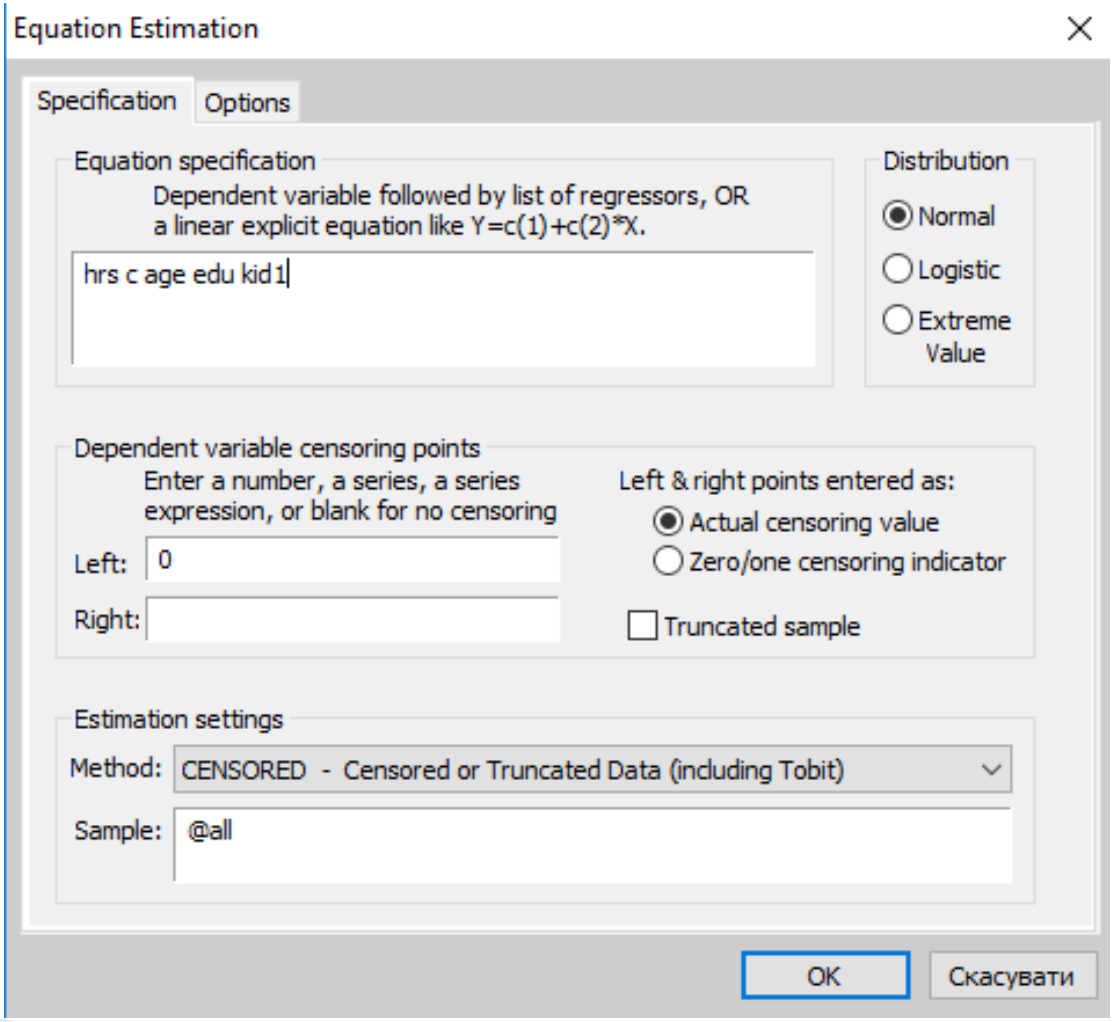
Model

- Suppose that we wish to estimate the model:

$$Hrs_t = \beta_0 + \beta_1 Age_t + \beta_2 Edu_t + \beta_3 Kid1_t + \varepsilon_t$$

where hours worked (HRS) is left censored at zero

Estimating Censored Models



Limit Points Known

- Left edit field: 0
- Right edit field: [blank]

- Left edit field: [blank]
- Right edit field: 20000

- Left edit field: 10000
- Right edit field: 20000

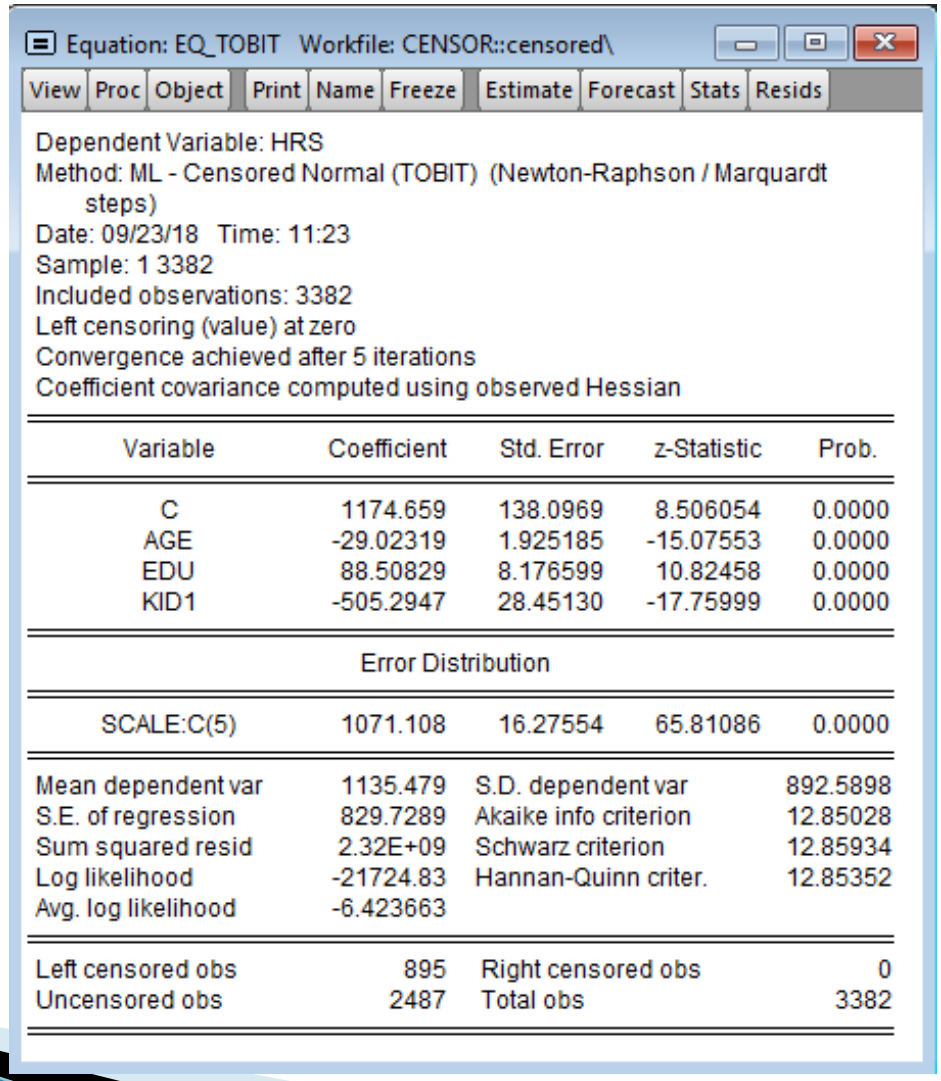
- Left edit field: lowinc
- Right edit field: vcens1+10

Limit Points Not Known

- This situation often occurs with data where censoring is indicated with a zero-one dummy variable, but no additional information is provided about potential censoring points.
- Simply select the Field is zero/one indicator of censoring option in the estimation dialog, and enter the series expression for the censoring indicator(s) in the appropriate edit field(s).

- Left edit field: [blank]
- Right edit field: rcens

Output



Equation: EQ_TOBIT Workfile: CENSOR::censored\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: HRS
Method: ML - Censored Normal (TOBIT) (Newton-Raphson / Marquardt steps)
Date: 09/23/18 Time: 11:23
Sample: 1 3382
Included observations: 3382
Left censoring (value) at zero
Convergence achieved after 5 iterations
Coefficient covariance computed using observed Hessian

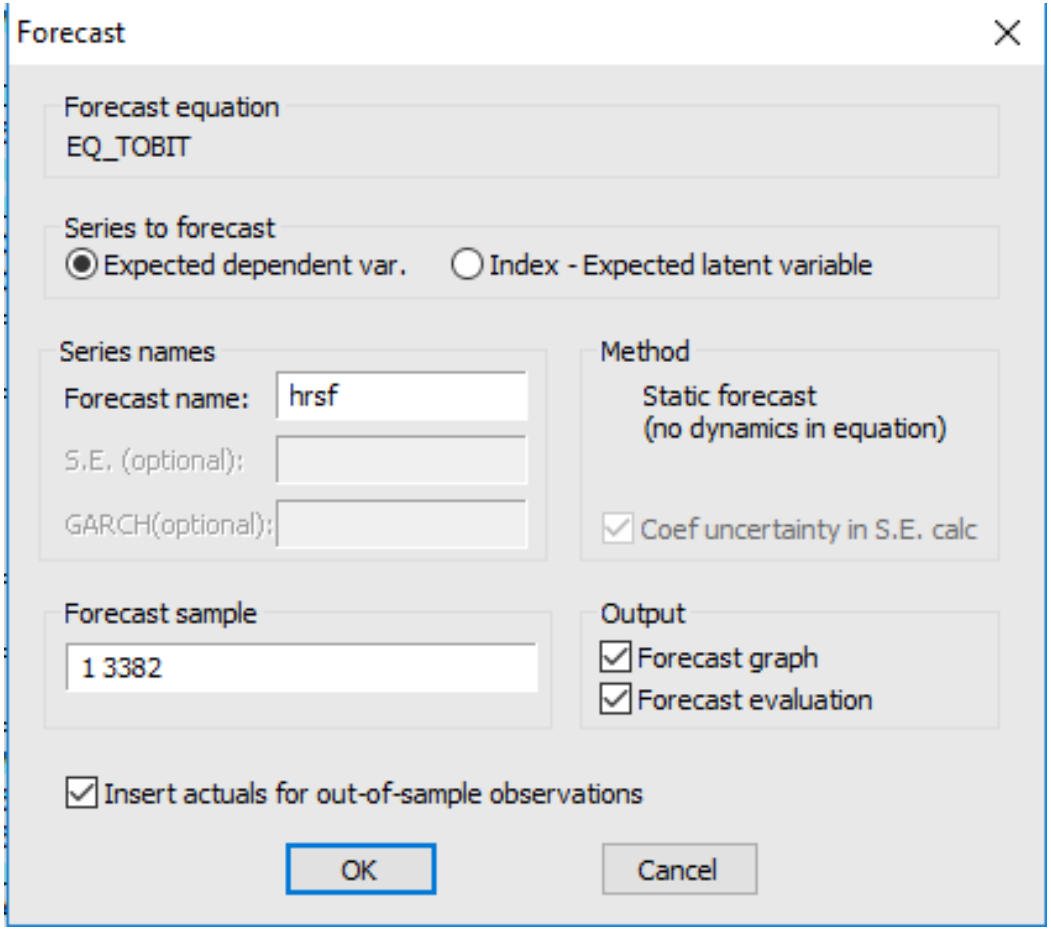
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	1174.659	138.0969	8.506054	0.0000
AGE	-29.02319	1.925185	-15.07553	0.0000
EDU	88.50829	8.176599	10.82458	0.0000
KID1	-505.2947	28.45130	-17.75999	0.0000
Error Distribution				
SCALE:C(5)	1071.108	16.27554	65.81086	0.0000

Mean dependent var	1135.479	S.D. dependent var	892.5898
S.E. of regression	829.7289	Akaike info criterion	12.85028
Sum squared resid	2.32E+09	Schwarz criterion	12.85934
Log likelihood	-21724.83	Hannan-Quinn criter.	12.85352
Avg. log likelihood	-6.423663		
Left censored obs	895	Right censored obs	0
Uncensored obs	2487	Total obs	3382

SCALE

- EViews reports an additional coefficient named SCALE, which is the estimated scale factor σ. This scale factor may be used to estimate the standard deviation of the residual, using the known variance of the assumed distribution.

Forecasting



To forecast the expected latent variable, click on **Index – Expected latent variable.**

To forecast the expected observed dependent variable, you should select **Expected dependent variable.**

Tests of Significance – 1

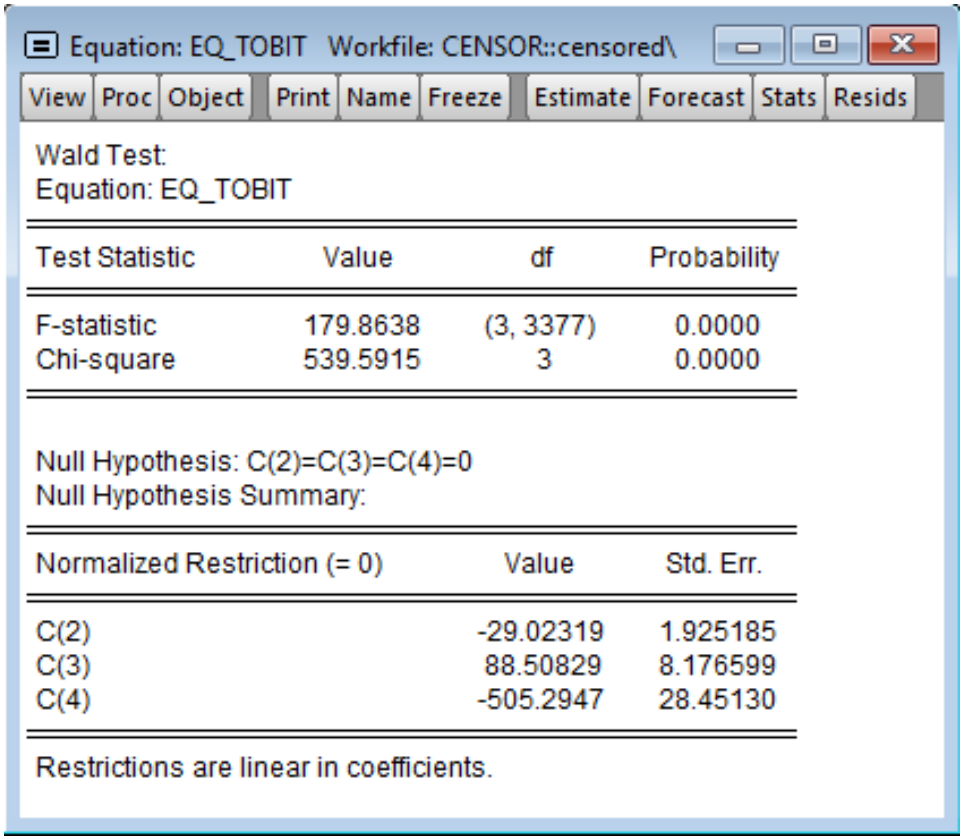
- EViews does not, by default, provide you with the usual likelihood ratio test of the overall significance for the tobit and other censored regression models.
- We can use the built-in coefficient testing procedures to test the exclusion of all of the explanatory variables. Select the redundant variables test and enter the names of all of the explanatory variables you wish to exclude. EViews will compute the appropriate likelihood ratio test statistic and the p-value associated with the statistic.

Tests of Significance – 2

- To take an example, suppose we wish to test whether the variables contribute to the fit of the model. Select **View/Coefficient Diagnostics/Redundant Variables – Likelihood Ratio…** and enter all of the explanatory variables:

$$c(2)=c(3)=c(4)=0$$

Tests of Significance – 3



Equation: EQ_TOBIT Workfile: CENSOR::censored\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Wald Test:
Equation: EQ_TOBIT

Test Statistic	Value	df	Probability
F-statistic	179.8638	(3, 3377)	0.0000
Chi-square	539.5915	3	0.0000

Null Hypothesis: C(2)=C(3)=C(4)=0
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2)	-29.02319	1.925185
C(3)	88.50829	8.176599
C(4)	-505.2947	28.45130

Restrictions are linear in coefficients.

The reported statistic is 179.86, with a p-value of less than 0.00001.

Estimating a Truncated Model

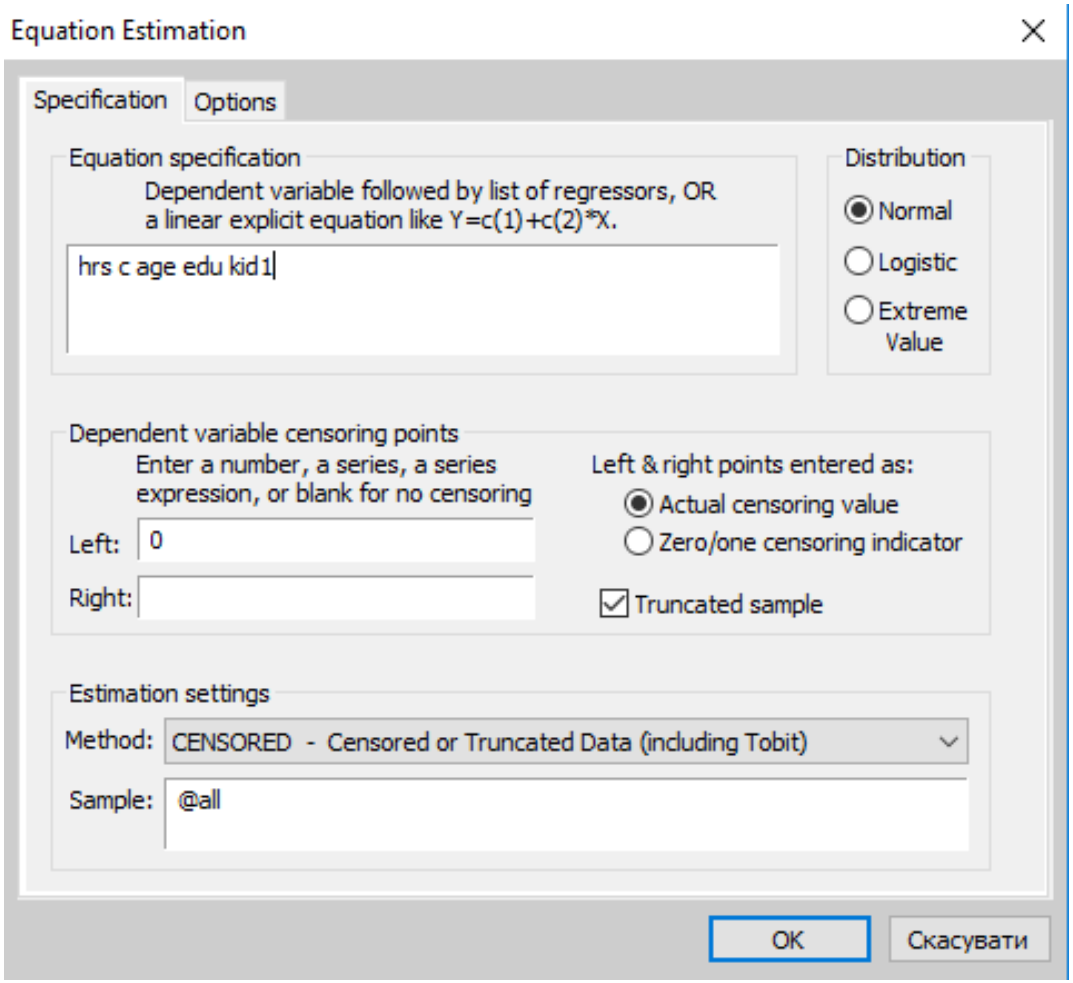
Estimation of a truncated regression model follows the same steps as estimating a censored regression:

- Select **Quick/Estimate Equation…** from the main menu, and in the Equation Specification dialog, select the **CENSORED estimation method**. The censored and truncated regression dialog will appear.
- Enter the name of the truncated dependent variable and the list of the regressors or provide explicit expression for the equation in the Equation Specification field, and select one of the three distributions for the error term.
- Indicate that you wish to estimate the truncated model by checking the **Truncated sample option**.
- Specify the truncation points of the dependent variable by entering the appropriate expressions in the two edit fields. If you leave an edit field blank, EViews will assume that there is no truncation along that dimension.

Keep in mind

- Truncated estimation is only available for models where the truncation points are known, since the likelihood function is not otherwise defined. If you attempt to specify your truncation points by index, EViews will issue an error message indicating that this selection is not available.
- EViews will issue an error message if any values of the dependent variable are outside the truncation points. Furthermore, EViews will automatically exclude any observations that are exactly equal to a truncation point. Thus, if you specify zero as the lower truncation limit, EViews will issue an error message if any observations are less than zero, and will exclude any observations where the dependent variable exactly equals zero.

Estimation


Equation Estimation
Specification
Options
Equation specification
Dependent variable followed by list of regressors, OR
a linear explicit equation like Y=c(1)+c(2)*X.
hrs c age edu kid1
Distribution
Normal
Logistic
Extreme Value
Dependent variable censoring points
Enter a number, a series, a series expression, or blank for no censoring
Left: 0
Right:
Left & right points entered as:
Actual censoring value
Zero/one censoring indicator
Truncated sample
Estimation settings
Method: CENSORED - Censored or Truncated Data (including Tobit)
Sample: @all
OK
Скасувати

Output

Equation: EQ_TOBIT Workfile: CENSOR::censored\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: HRS
Method: ML - Censored Normal (TOBIT) (Newton-Raphson / Marquardt steps)
Date: 09/23/18 Time: 12:05
Sample: 1 3382
Included observations: 2487
Truncated sample
Left censoring (value) at zero
Convergence achieved after 7 iterations
Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	1534.915	104.3856	14.70427	0.0000
AGE	-4.994707	1.507413	-3.313429	0.0009
EDU	21.67178	6.264716	3.459340	0.0005
KID1	-265.0019	22.96683	-11.53846	0.0000
Error Distribution				
SCALE:C(5)	681.9777	11.20692	60.85325	0.0000

Mean dependent var	1544.106	S.D. dependent var	672.5555
S.E. of regression	653.5106	Akaike info criterion	15.77823
Sum squared resid	1.06E+09	Schwarz criterion	15.78993
Log likelihood	-19615.23	Hannan-Quinn criter.	15.78248
Avg. log likelihood	-7.887103		

Left censored obs	0	Right censored obs	0
Uncensored obs	2487	Total obs	2487

Comparisson

Equation: EQ_TOBIT Workfile: CENSOR::censored\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: HRS
Method: ML - Censored Normal (TOBIT) (Newton-Raphson / Marquardt steps)
Date: 09/23/18 Time: 11:23
Sample: 1 3382
Included observations: 3382
Left censoring (value) at zero
Convergence achieved after 5 iterations
Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	1174.659	138.0969	8.506054	0.0000
AGE	-29.02319	1.925185	-15.07553	0.0000
EDU	88.50829	8.176599	10.82458	0.0000
KID1	-505.2947	28.45130	-17.75999	0.0000
Error Distribution				
SCALE:C(5)	1071.108	16.27554	65.81086	0.0000

Mean dependent var	1135.479	S.D. dependent var	892.5898
S.E. of regression	829.7289	Akaike info criterion	12.85028
Sum squared resid	2.32E+09	Schwarz criterion	12.85934
Log likelihood	-21724.83	Hannan-Quinn criter.	12.85352
Avg. log likelihood	-6.423663		

Left censored obs	895	Right censored obs	0
Uncensored obs	2487	Total obs	3382

Equation: EQ_TOBIT Workfile: CENSOR::censored\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: HRS
Method: ML - Censored Normal (TOBIT) (Newton-Raphson / Marquardt steps)
Date: 09/23/18 Time: 12:05
Sample: 1 3382
Included observations: 2487
Truncated sample
Left censoring (value) at zero
Convergence achieved after 7 iterations
Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	1534.915	104.3856	14.70427	0.0000
AGE	-4.994707	1.507413	-3.313429	0.0009
EDU	21.67178	6.264716	3.459340	0.0005
KID1	-265.0019	22.96683	-11.53846	0.0000
Error Distribution				
SCALE:C(5)	681.9777	11.20692	60.85325	0.0000

Mean dependent var	1544.106	S.D. dependent var	672.5555
S.E. of regression	653.5106	Akaike info criterion	15.77823
Sum squared resid	1.06E+09	Schwarz criterion	15.78993
Log likelihood	-19615.23	Hannan-Quinn criter.	15.78248
Avg. log likelihood	-7.887103		

Left censored obs	0	Right censored obs	0
Uncensored obs	2487	Total obs	2487

Heckman Selection model

- The wage equation is given by

$$wage_t = \beta_0 + \beta_1 Exper_t + \beta_2 Exper_t^2 + \beta_3 Educ_t + \beta_4 City_t + \varepsilon_t$$

where EXPER is a measure of each woman's experience, EDUC is her level of education, and CITY is a dummy variable for whether she lives in a city or not.

- The selection equation is given by:

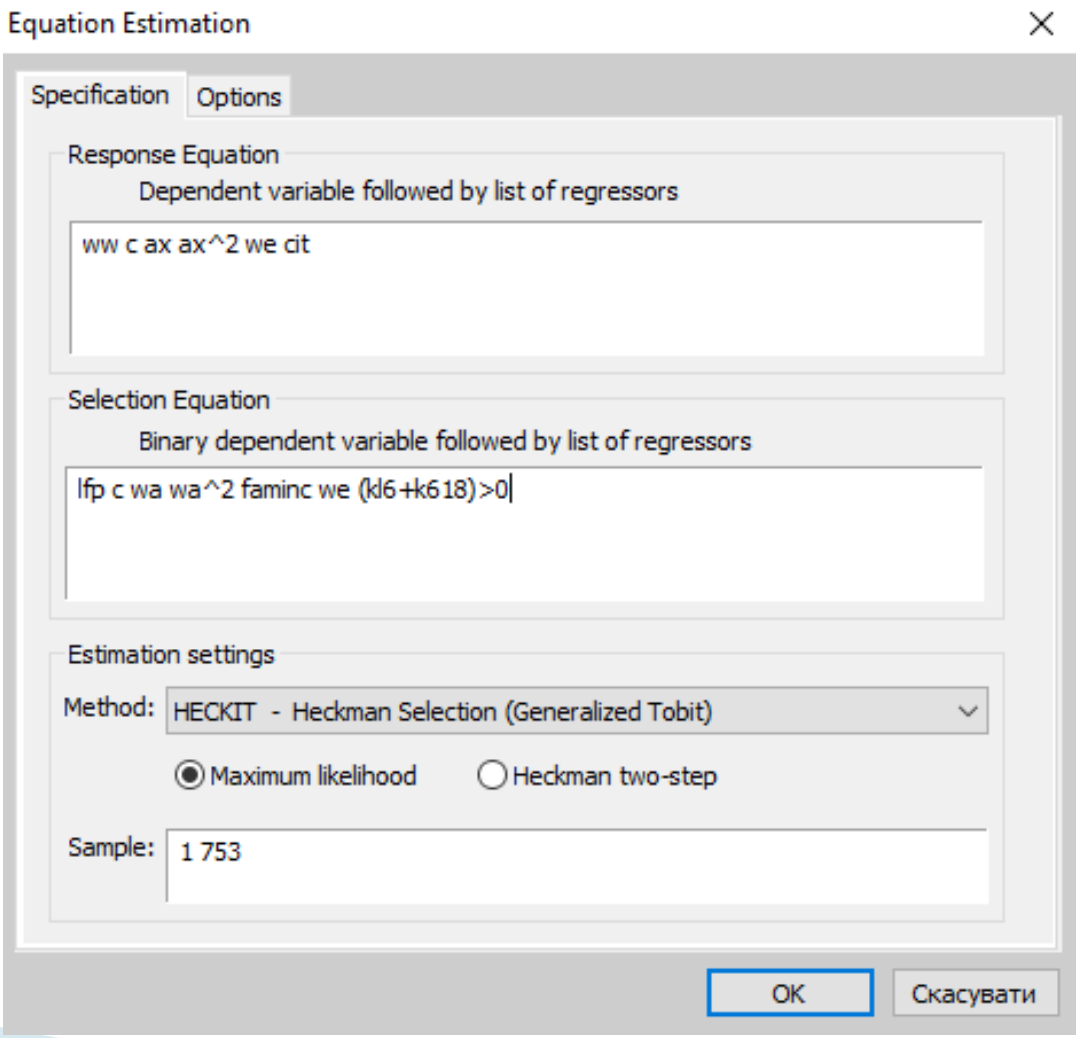
$$LFP_t = \gamma_0 + \gamma_1 Age_t + \gamma_2 Age_t^2 + \gamma_3 Faminic_t + \gamma_4 Educ_t + \gamma_5 Kids_t + u_t$$

where LFP is a binary variable taking a value of 1 if the woman is in the labor force, and 0 otherwise, AGE is her age, FAMINC is the level of household income not earned by the woman, and KIDS is a dummy variable for whether she has children.

Data

- File: heckit.wf1
- In this data, the wage data are in the series WW, experience is AX, education is in WE, the city dummy is CIT, labor force participation is LFP, age is WA, and family income is FAMINC. There is no kids dummy variable, but there are two variables containing the number of children below K6 education (KL6), and the number of kids between K6 education and 18 (K618). We can create the dummy variable simply by testing whether the sum of those two variables is greater than 0.

Estimation


Equation Estimation
Specification
Options
Response Equation
Dependent variable followed by list of regressors
ww c ax ax^2 we cit
Selection Equation
Binary dependent variable followed by list of regressors
lfp c wa wa^2 faminc we (kl6+k618)>0
Estimation settings
Method:
HECKIT - Heckman Selection (Generalized Tobit)
Maximum likelihood
Heckman two-step
Sample:
1 753
OK
Скасувати

Comparisson

Equation: UNTITLED Workfile: HECKIT::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: WW
Method: ML Heckman Selection (Newton-Raphson / Marquardt steps)
Date: 09/23/18 Time: 13:06
Sample: 1 753
Included observations: 753
Selection Variable: LFP
Convergence achieved after 5 iterations
Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Response Equation - WW				
C	-1.963024	1.198221	-1.638282	0.1018
AX	0.027868	0.061551	0.452764	0.6509
AX^2	-0.000104	0.001839	-0.056483	0.9550
WE	0.457005	0.073230	6.240688	0.0000
CIT	0.446529	0.315921	1.413420	0.1580
Selection Equation - LFP				
C	-4.119692	1.400516	-2.941552	0.0034
WA	0.184015	0.065867	2.793729	0.0053
WA^2	-0.002409	0.000772	-3.118875	0.0019
FAMINC	5.68E-06	4.42E-06	1.286180	0.1988
WE	0.095281	0.023153	4.115194	0.0000
(KL6+K618)>0	-0.450615	0.130185	-3.461331	0.0006
Interaction terms				
@LOG(SIGMA)	1.134100	0.036621	30.96833	0.0000
TFORM(RHO)	-0.210301	0.270853	-0.776440	0.4377
SIGMA	3.108376	0.113833	27.30651	0.0000
RHO	-0.131959	0.165127	-0.799133	0.4245

Mean dependent var	4.177682	S.D. dependent var	3.310282
S.E. of regression	2.361759	Akaike info criterion	4.234416
Sum squared resid	4127.650	Schwarz criterion	4.314247
Log likelihood	-1581.258	Hannan-Quinn criter.	4.265171

Equation: UNTITLED Workfile: HECKIT::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: WW
Method: Two-Step Heckman Selection
Date: 09/23/18 Time: 13:07
Sample: 1 753
Included observations: 753
Selection Variable: LFP
Coefficient covariance computed using two-step Heckman method

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Response Equation - WW				
C	-0.971200	2.132849	-0.455353	0.6490
AX	0.021061	0.062532	0.336804	0.7364
AX^2	0.000137	0.001882	0.072842	0.9420
WE	0.417017	0.104157	4.003746	0.0001
CIT	0.443838	0.316531	1.402194	0.1613
Selection Equation - LFP				
C	-4.156807	1.402086	-2.964730	0.0031
WA	0.185395	0.065967	2.810436	0.0051
WA^2	-0.002426	0.000774	-3.136096	0.0018
FAMINC	4.58E-06	4.21E-06	1.088918	0.2765
WE	0.098182	0.022984	4.271744	0.0000
(KL6+K618)>0	-0.448987	0.130911	-3.429697	0.0006

Mean dependent var	4.177682	S.D. dependent var	3.310282
S.E. of regression	2.418304	Akaike info criterion	6.017314
Sum squared resid	4327.663	Schwarz criterion	6.084863
Log likelihood	-2254.519	Hannan-Quinn criter.	6.043337

Questions?

Self study