Documentation of STATA code for "Extremal Quantile Regression for Selection Models and the Black-White Wage Gap"

December 2016

Title

eqregsel - Estimation method using extremal quantile regression under endogenous selection

Syntax

eqregsel varlist, [options]

Options	Description
hom(#)	Estimate the coefficients of the first # independent variables; default is hom(1)
subs(#)	Subsampling size used for selecting optimal quantile index tau; default is based on the size of the dataset
grid(#)	Discretize the interval of tau into # subintervals; default is grid(40)
rep(#)	Perform # bootstrap replications; default is rep(150)

Description

eqregsel estimates and provides inference for the coefficient of variables of interest under endogenous selection. The estimation method mainly assumes that under endogenous selection, the effect of the outcome on selection dominates those of covariates for sufficiently large values of the outcome, which is proposed by d'Haultfoeuille, X., Maurel, A., & Zhang, Y. (2016), based on the paper Extremal Quantile Regression for Selection Models and the Black-White Wage Gap. The variables of interest are assumed to have homogeneous effects on the outcome across its distribution. The number of variables of interest can be more than one.

Options

The four options of this command have default values.

hom(#) specifies the number of variable of interests. The variables of interest are assumed to have homogeneous effects on the outcome across its distribution. Apart from specifying the number of variables of interests, users have to put the all the variables of interest right after the dependent variables, followed by other independent variables. The default value is 1.

subs(#) specifies the sample size used for selecting the optimal quantile index tau. To select tau, an information criteria considering tradeoff between biasedness and efficiency is computed using a selected subsample, whose size is specified by subs(). If no value is specified, the sample size is calculated by the following formula.

$0.6n - 0.2(n-500)^+ - 0.2(n-1000)^+ - 0.2[1 - \frac{ln2000}{lnn}](n-2000)^+$

Or, by the piecewise function

$$\begin{cases} 06n & n < 500 \\ 300 + 0.4(n-500) & 500 \leq n < 1000 \\ 500 + 0.2(n-1000) & 1000 \leq n < 2000 \\ 700 + 0.2 \times \frac{ln2000}{lnn} \times (n-2000) & n \geq 2000 \end{cases}$$

grid(#) specifies the number of subintervals to divide region of search when selecting optimal quantile index tau. The upper bound of region of search is 0.3, while the lower bound is $min(0.1, 0.8/b_n)$, where b_n is specified by subs(). The default value is 40.

rep(#) specifies the number of bootstrap replications when selecting optimal quantile index tau . The default value is 150.

Output

Name	Description
e(tau0)	Optimal quantile index
e(specificationtest)	Value of specification test
e(subs)	Subsampling size in the bootstrapping process
e(homvar)	Number of variable(s) with homogeneous effect(s) on the outcome
e(beta_hom)	Estimated coefficient(s) of interest
e(std_b)	Standard error of the estimator(s)

Remarks

1. Users should specify the number of variables of interest by using the option hom(). Apart from specifying number of variables whose coefficients are to be estimated, users should put names of those variables right after the dependent variable, followed by other independent variables. Users can specify more than one variables whose coefficients are to be estimated.

2. During computation, the command will display the estimated time and progress bar. After estimation, results will be displayed.

Examples

Here are some examples.

. eqregsel log_wage black hispanic age AFQT0 AFQT0_2
(black is the variable of interest)
. eqregsel log_wage black hispanic age AFQT0 AFQT0_2, hom(2)
(black and hispanic are variables of interest)
. eqregsel log_wage black hispanic age AFQT0 AFQT0_2, rep(300)

Version Requirements

This command requires Stata 12 or upper.

Methods and Formulas

See Xavier D'Haultfoeuille, Arnaud Maurel, Yichong Zhang (2016), Extremal Quantile Regression for Selection Models and the Black-White Wage Gap, Mimeo.

Supporting Package

This command requires -moremate- package. Before calling this command, the user needs to install -moremata- by typing:

. ssc install moremata