Topic 16 Header Slide

- Topic 16. Extensions to $A2$ (non-linearities), $A4/A3$ (infinite moments), $A3$ (endogeneity)

Estimator	Notation	Under Complications	Extends to
1.	$\hat{\beta}_{OLS}$	$A2 Nonlinear.additive Error$	$\rightarrow \hat{\beta}_{NLLS}$
2.	$\hat{\beta}_{LAD}$		
3.	$\hat{\beta}_{Lstar}$		
4.	$\hat{\beta}_{GMM}$	$A3 Endogenous X/Exogenous W$	$\rightarrow \hat{\beta}_{IVE}$
$5a.$	$\hat{\beta}_{IGLS}$		
$5b.$	$\hat{\beta}_{FGLS}$		
6.	$\hat{\beta}_{MLE}$	$Multiple$	$\rightarrow \hat{\beta}_{FIMLE}$
7.	$\hat{\beta}_{IVE}$	$A2 Nonlinear.Additive Error$	$\rightarrow \hat{\beta}_{NLIVE}$

$A2$ **extensions:** $A2additively.nonlinear$

Suppose we have $A2linear$ along with $A3Rsru$.

- In this setting, we can avail the population orthogonality conditions (by $A3Rsru$).
- We know (due to the LLN) that we can estimate population moments consistently via sample moments. This motivates the use of sample orthogonality conditions for the purpose of GMM estimation of any unknown model parameters.
- The moment equations we obtain turn out to be equivalent to the first-order conditions for minimising the OLS objective function (up to scale).

Now suppose we have $A2.additively.nonlinear$ along with $A3Rsru$.

- In this setting, we can still avail the population orthogonality conditions (by $A3Rsru$).
- But now the moment equations we obtain turn out to be equivalent to the first-order conditions for minimising the NLLS objective function (up to scale).
- We have:

$$\hat{\beta}_{NLLS} = \arg\min_{b} \sum_{s=1}^{S} (y_s - g(x_s, b))^2 \,,$$

where $g(\cdot, \cdot)$ represents a non-linear link function between $\mathbb{E}(y_s|X)$ and the regressors.

$A2$ **and** $A3$ **extensions:** $A3Rsru.W$

Suppose we have $A2linear$ along with $A3Rsru.W$ whereby regressors may well be endogenous but there exists a weakly exogenous matrix of candidate instruments W.

- In this setting, we can avail the W matrix and undertake IVE (which applies the GMM principle to the endogenous regressor case).

Now suppose we have $A2additively.nonlinear$ along with $A3Rsru.W$

- In this setting, we can avail the W matrix and undertake NLIV (which applies the GMM principle to the endogenous regressor case with a nonlinear link function).

$A2$ and $A3$ extensions: joint consideration

Suppose we have:

* $A2linear$ or $A2additively.nonlinear$ or $A2nonadditively.nonlinear$
* $A3Rsru$ or $A3Rsru.W$
* and convenient $A4.(n)i.i.d.$

- We can then avail (full information) maximum likelihood — FIMLE. To achieve this, we employ all five classes of assumptions (including distributional family) and specify the full conditional density of all the endogenous variables, y and X^{bad}, given all the exogenous variables (regressors X^{good} and additional instrument variables Z) and the unknown parameter vector θ^{true}:

$$f(y, X^{bad}|X^{good}, Z; \theta^{true})$$

  Taking the (natural) log of the joint density as a function of any arbitrary θ defines the Loglikelihood Function and thus the FIMLE:

$$LLF(\theta; \text{all endogenous and exogenous data}) \equiv \ln f(y, X^{bad}|X^{good}, Z; \theta)$$

$$\theta_{FIMLE} = \arg\max_{\theta} LLF(\theta; \text{all data})$$

- The FIMLE in general requires the iterative solution of the FOCs, which will be a set of non-linear equations. If all five assumptions are correctly specified, the FIMLE will be the *best CUAN* estimator in that its asymptotic variance will achieve the Cramér-Rao lower bound, asymptotically with the sample size $S \to \infty$.

$A4$ extensions: infinite moments

Suppose we have $A1\text{-}A5iid.Cauchy$

- We can avail (full information) maximum likelihood given the following marginal PDF:

$$f(z) = \frac{1}{\pi b \left(1 + \left(\frac{z-a}{b}\right)^2\right)}$$

  where a and b are the parameters governing the Cauchy distribution.

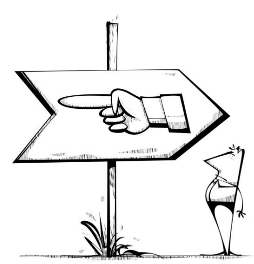
- In this case writing $\epsilon_s = z - a = y_s - x_s'\beta$, would allow us to define the likelihood contribution of data point s. Combining it with the independence assumption would give us the Likelihood function and the Loglikelihood function, thus allowing us to define the MLE for β and the scale parameter b.
- Despite the fact that all the even moments of the Cauchy distribution are infinite and all the odd moments are undefined, the first and second derivatives of the Loglikelihood function are well defined. Hence, the resulting MLE will be the best CUAN estimator for this case, asymptotically with the sample size $S \to \infty$.

Review Quiz for Topic 16

Question 1. [To be confirmed perhaps at a later date.]

Question 2. [To be confirmed perhaps at a later date.]

Question 3. [To be confirmed perhaps at a later date.]



Signpost 16

[To be confirmed perhaps at a later date.]