The Minimum Distance Estimator

James J. Heckman

THE UNIVERSITY OF CHICAGO

Econ 312, Spring 2019

- In order to test if there is a common structure in the parameter estimates for the choice of the alternative vertical information sources, a Minimum Distance Estimator (MDE) is used.

- A thorough discussion of the MDE and applications are presented in Kodde et al. (1990).

- Minimum Distance Estimation involves the estimation of the *R* reduced form parameter vectors in a first stage.

- In the present case, these reduced form parameters are the parameter estimates obtained from running two separate OLS regressions for the innovation intensity of cooperating and non–cooperating firms.

- In the second stage, the Minimum Distance Estimator is derived by minimizing the weighted difference between the auxiliary parameter vectors obtained in the first stage.

- Besides the practical advantage that the MDE can be easily implemented empirically, it has the further benefit that it provides the researcher with a formal test of common structures among the auxiliary parameter vectors.

- The MDE is derived by minimizing the distance between the auxiliary parameter vectors under the following set of restrictions:

$$f\left(\boldsymbol{\beta}, \widehat{\boldsymbol{\theta}}\right) = \boldsymbol{H}\,\boldsymbol{\beta} - \widehat{\boldsymbol{\theta}} = \mathbf{0}, \tag{1}$$

where the $R \cdot K \ \times K$ matrix $\boldsymbol{H}$ imposes $(R-1) \cdot K$ restrictions on $\boldsymbol{\theta}$.

- The $R \cdot K \times 1$ vector $\widehat{\boldsymbol{\theta}}$ contains the R stacked auxiliary parameter vectors.
- In the present case, $\boldsymbol{H}$ is defined by a $R \cdot K \times K$-dimensional stacked identity matrix.
- The MDE is given by the minimization of:

$$D(\beta) = f\left(\boldsymbol{\beta}, \widehat{\boldsymbol{\theta}}\right) \hat{V}\left[\widehat{\boldsymbol{\theta}}\right]^{-1} f\left(\boldsymbol{\beta}, \widehat{\boldsymbol{\theta}}\right), \quad (2)$$

where $\hat{V}[\widehat{\boldsymbol{\theta}}]$ denotes the common estimated variance–covariance matrix of the auxiliary parameter vectors.

- Minimization of D leads to

$$\widehat{\boldsymbol{\beta}} = \left(H' \, \hat{V}\left[\widehat{\boldsymbol{\theta}}\right]^{-1} H\right)^{-1} H' \, \hat{V}\left[\widehat{\boldsymbol{\theta}}\right]^{-1} \widehat{\boldsymbol{\theta}} \tag{3}$$

with variance–covariance matrix

$$\hat{V}\left[\hat{\beta}\right] = \left(H' \, \hat{V} \left[\hat{\theta}\right]^{-1} H\right)^{-1} \tag{4}$$

- In the present case, where the two equations were estimated using different samples, $V[\hat{\theta}]$ is a matrix carrying the estimated variance-covariance matrices of the first stage parameter vectors on its diagonal blocks.

- The off–diagonal blocks consist of zero–matrices.

- To test the null hypothesis that the R auxiliary parameter vectors coincide with one another, the following Wald-type test statistic can be applied:

$$W = f\left(\hat{\beta}, \hat{\theta}\right)' \, \hat{V}\left[\hat{\theta}\right]^{-1} f\left(\hat{\beta}, \hat{\theta}\right) \sim x^2_{(R-1)\cdot K}. \qquad (5)$$