Econometrics 2

Lecture 7: The Just-Identified Case ($q = p$)

IV Estimator, Constrained Spaces & Introduction to LAD

AUEB | Spring Semester 2026

Taught by: Prof. S. Arvanitis

Notes transcribed & digitized by: T. Kourtalis

1. Recap of the IV Estimator and Objective Function

The linear model with instrumental variables supports the system of moment conditions:

$$\mathbb{E}\left[\frac{1}{n}Z_n'(Y_n - X_n\beta)\right] = 0_{q\times 1} \quad (*)$$

Provided the relevant identification conditions hold, $(*)$ is satisfied if and only if $\beta = \beta_0$.

Choosing to work with Mahalanobis-type squared distances and independently of the weighting matrix W, we arrived at the usual version of the Instrumental Variables Estimator (IVE):

$$\beta_n \in \arg\min_{\beta\in\Theta}(Y_n - X_n\beta)'Z_nWZ_n'(Y_n - X_n\beta)$$

We previously questioned how the properties of β_n depend on the choice of W. By examining the case where $\Theta = \mathbb{R}^p$, the optimization is solved analytically:

$$\beta_n = (X_n'Z_nWZ_n'X_n)^{-1}X_n'Z_nWZ_n'Y_n$$

Since $Y_n = X_n\beta_0 + \epsilon_n$, we substituted this to find:

$$\beta_n = \beta_0 + (X_n'Z_nWZ_n'X_n)^{-1}X_n'Z_nWZ_n'\epsilon_n$$

This led to the conclusion that β_n is not generally unbiased.

2. The Just-Identified Case ($q = p$)

Let us investigate what happens to our estimator when $q = p$. This means the number of instrumental variables (q) perfectly coincides with the number of regressors (p).

Consider the matrix $X_n'Z_n$:

- Generally, it has dimensions $p \times q$.
- When $q = p$, it becomes a square $p \times p$ matrix.

By our general rank condition assumption, $rank(X_n' Z_n) = p$. Therefore, when $q = p$, the square matrix $X_n' Z_n$ is **invertible**. Consequently, its transpose $Z_n' X_n$ is also invertible.

Furthermore, the weighting matrix W is now $p \times p$. Since it is positive definite, it is also invertible.

These properties allow us to expand the inverse of the matrix product. Using the matrix property $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$:

$$\Big((X_n' Z_n) W (Z_n' X_n)\Big)^{-1} = (Z_n' X_n)^{-1} W^{-1} (X_n' Z_n)^{-1}$$

(Note: Each of these individual separated matrices is $p \times p$ and invertible).

3. Simplification of the Estimator

We can now substitute this expanded inverse back into the analytical formula for β_n:

$$\beta_n = \underbrace{(Z_n' X_n)^{-1} W^{-1} (X_n' Z_n)^{-1}}_{\text{Inverse part}} (X_n' Z_n) W Z_n' Y_n$$

Notice the immediate cancellations via identity matrices ($I_{p \times p}$):

- $(X_n' Z_n)^{-1}(X_n' Z_n) = I_{p \times p}$
- $W^{-1} W = I_{p \times p}$

This perfectly collapses the estimator down to:

$$\beta_n = (Z_n' X_n)^{-1} Z_n' Y_n$$

Conclusion for $q = p$: When $q = p$ and $\Theta = \mathbb{R}^p$, we find that $\beta_n = (Z_n' X_n)^{-1} Z_n' Y_n$. Therefore, our estimator—and consequently the properties of the IVE—become **completely independent** of the choice of the weighting matrix W.

4. OLS as a Special Case

What if, additionally, the instrumental variables simply coincide with the regressors themselves? (i.e., $Z_n = X_n$). *(Note: This relies on the core exogeneity assumption $\mathbb{E}(\epsilon_n | X_n) = 0_{n \times 1}$).*

Substituting $Z_n = X_n$ into our simplified just-identified estimator yields:

$$\beta_n = (X_n' X_n)^{-1} X_n' Y_n$$

This is exactly the Ordinary Least Squares Estimator (OLSE). Therefore, for $\Theta = \mathbb{R}^p$, we see that the Least Squares estimator becomes a special case of the instrumental variables estimator.

Caveat: If we restrict the parameter space such that $\Theta \subset \mathbb{R}^p$ (i.e., we operate in some smaller, constrained space), we can no longer guarantee this exact analytical outcome, as we do not know what happens with the inverse at the boundaries.

5. The Just-Identified Case in a Constrained Space

When $q = p$ but the parameter space is constrained ($\Theta \subset \mathbb{R}^p$), does the independence of the IV estimator from the choice of W still hold?

Let's define the *unconstrained* estimator as:

$$\beta_n^{(un)} = (Z_n' X_n)^{-1} Z_n' Y_n$$

The residuals based on this unconstrained estimator are:

$$e_n = Y_n - X_n \beta_n^{(un)} = Y_n - X_n (Z_n' X_n)^{-1} Z_n' Y_n$$
$$e_n = \Big(I_{n \times n} - X_n (Z_n' X_n)^{-1} Z_n' \Big) Y_n$$

Now, let's examine the sample moment conditions evaluated at these residuals:

$$\frac{1}{n} Z_n' e_n = \frac{1}{n} \Big(Z_n' - Z_n' X_n (Z_n' X_n)^{-1} Z_n' \Big) Y_n$$

Since $Z_n' X_n (Z_n' X_n)^{-1} = I_{p \times p}$, this simplifies to:

$$\frac{1}{n} Z_n' e_n = \frac{1}{n} (Z_n' - Z_n') Y_n = \frac{1}{n} O_{p \times n} Y_n = O_{p \times 1}$$

(Note: The sample property $\frac{1}{n} Z_n' e_n = 0$ perfectly mirrors our theoretical moment condition $\frac{1}{n}\mathbb{E}(Z_n' \epsilon_n) = 0_{p \times 1}$. Exercise: Does this exact zero hold if $q > p$?)

6. Reformulating the Objective Function

In the case we are examining, we have $\beta_n \in \arg\min_{\beta \in \Theta} \mu_n(\beta)$. Let's substitute $Y_n - X_n \beta$ in the objective function using our unconstrained estimator:

$$\begin{aligned} Z_n'(Y_n - X_n \beta) &= Z_n' \Big(Y_n - X_n \beta_n^{(un)} + X_n \beta_n^{(un)} - X_n \beta \Big) \\ &= Z_n' \Big(e_n + X_n (\beta_n^{(un)} - \beta) \Big) \\ &= Z_n' e_n + Z_n' X_n (\beta_n^{(un)} - \beta) \end{aligned}$$

Since $Z_n' e_n = 0$, we are left with:

$$Z_n'(Y_n - X_n \beta) = Z_n' X_n (\beta_n^{(un)} - \beta) \quad (*)$$

Consequently, because of $(*)$, our objective function becomes:

$$\begin{aligned} \mu_n(\beta) &= \frac{1}{n^2} (\beta_n^{(un)} - \beta)' X_n' Z_n W Z_n' X_n (\beta_n^{(un)} - \beta) \\ &= (\beta_n^{(un)} - \beta)' \left(\frac{X_n' Z_n}{n} W \frac{Z_n' X_n}{n} \right) (\beta_n^{(un)} - \beta) \end{aligned}$$

Let $A_n = \frac{X_n' Z_n}{n} W \frac{Z_n' X_n}{n}$. Since W is positive definite, we can use the Cholesky decomposition ($W = LL'$), which helps us prove that A_n is also a positive definite matrix.

Therefore, for $q = p$, the estimator β_n can equivalently be expressed as minimizing a quadratic form of the Mahalanobis distance with respect to A_n over the set Θ:

$$\beta_n \in \arg\min_{\beta \in \Theta} (\beta_n^{(un)} - \beta)' A_n (\beta_n^{(un)} - \beta) \quad (**)$$

Dependence on W

From $(**)$, we observe that the objective function strictly depends on W (through A_n). Consequently, the estimator is generally **expected to depend on** W.

It will only be independent of W when the unconstrained minimum happens to fall inside our parameter space: $\beta_n^{(un)} \in \Theta$. (For example, when $\Theta = \mathbb{R}^p$, which we already proved).

(Exercise: Show if we can form the Mahalanobis distance if $p > q$).

7. The Linear Model and the LAD Estimator

LAD = Least Absolute Deviations

Recall that in the usual version of the Least Squares estimator, the objective function is:

$$\mu_n(\beta) = \frac{1}{n}(Y_n - X_n\beta)'(Y_n - X_n\beta) = \frac{1}{n}\sum_{i=1}^{n}(Y_{(i)} - X_{(i)}\beta)^2$$

Let our matrices be defined as:

$$X_n = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{np} \end{pmatrix}$$

Where $X_{(i)} = (X_{i1}, X_{i2}, \dots, X_{ip})$ is the i-th row of X_n, and $Y_{(i)}$ is the i-th element of $Y_n = (Y_1, \dots, Y_n)'$.

Questions to ponder:

1. What would result if, as our objective function, we chose:

$$\frac{1}{n}\sum_{i=1}^{n}\left|Y_{(i)} - X_{(i)}\beta\right|$$

2. From what statistical model structure could such a choice be justified?