Solution to Exercise 12.28

⋆12.28 In the just-identified case of LIML estimation, for which, in the notation of (12.91), the number of excluded instruments in the matrix $\boldsymbol{W}_1$ is equal to the number of included endogenous variables in the matrix $\boldsymbol{Y}$, show that the minimized value of the ratio κ given by (12.92) is equal to the global minimum of 1. Show further that the vector of estimates $\hat{\boldsymbol{\beta}}_2$ that attains this minimum is the IV, or 2SLS, estimator of $\boldsymbol{\beta}_2$ for equation (12.90) with instruments $\boldsymbol{W}$.

In the overidentified case of LIML estimation, explicitly formulate a model containing the model consisting of (12.90) and (12.91) as a special case, with the overidentifying restrictions relaxed. Show that the maximized loglikelihood for this unconstrained model is the same function of the data as for the constrained model, but with $\hat{\kappa}$ replaced by 1.

As we saw in Section 12.5, the ratio (12.92) cannot be less than unity, because the numerator is the sum of squared residuals from a regression of $\boldsymbol{y} - \boldsymbol{Y\beta}_2$ on the matrix $\boldsymbol{Z}$, and the denominator is the sum of squared residuals from a regression of $\boldsymbol{y} - \boldsymbol{Y\beta}_2$ on the matrix $\boldsymbol{W}$. Since $\mathcal{S}(\boldsymbol{Z})$ is a subspace of $\mathcal{S}(\boldsymbol{W})$, the second regression cannot have less explanatory power than the first. Thus the global minimum cannot be less than 1.

If equation (12.90) is estimated by IV with instruments $\boldsymbol{W}$, the estimating equations are

$$\begin{bmatrix} \boldsymbol{Z}^\top \\ \boldsymbol{Y}^\top \boldsymbol{P_W} \end{bmatrix} (\boldsymbol{y} - \boldsymbol{Z\beta}_1 - \boldsymbol{Y\beta}_2) = \boldsymbol{0}. \tag{S12.42}$$

When equation (12.90) is exactly identified, $\mathcal{S}(\boldsymbol{Z}, \boldsymbol{P_W Y}) = \mathcal{S}(\boldsymbol{W})$. This follows from the facts that each column of the matrix $[\boldsymbol{Z} \quad \boldsymbol{P_W Y}]$ is simply a linear combination of the columns of $\boldsymbol{W}$, and that this matrix has as many columns as $\boldsymbol{W}$ itself. Since this matrix must have full rank if the equation is identified, it follows that $\mathcal{S}(\boldsymbol{Z}, \boldsymbol{P_W Y}) = \mathcal{S}(\boldsymbol{W})$. Consequently, we see from equation (S12.42) that

$$\boldsymbol{W}^\top (\boldsymbol{y} - \boldsymbol{Z}\hat{\boldsymbol{\beta}}_1^{\rm IV} - \boldsymbol{Y}\hat{\boldsymbol{\beta}}_2^{\rm IV}) = \boldsymbol{0}.$$

This equation states that the IV residuals are orthogonal to all of the instruments, including those in the matrix $\boldsymbol{Z}$. Therefore,

$$\begin{aligned} \boldsymbol{y} - \boldsymbol{Z}\hat{\boldsymbol{\beta}}_1^{\rm IV} - \boldsymbol{Y}\hat{\boldsymbol{\beta}}_2^{\rm IV} &= \boldsymbol{M_W}(\boldsymbol{y} - \boldsymbol{Z}\hat{\boldsymbol{\beta}}_1^{\rm IV} - \boldsymbol{Y}\hat{\boldsymbol{\beta}}_2^{\rm IV}) \\ &= \boldsymbol{M_Z}(\boldsymbol{y} - \boldsymbol{Z}\hat{\boldsymbol{\beta}}_1^{\rm IV} - \boldsymbol{Y}\hat{\boldsymbol{\beta}}_2^{\rm IV}). \end{aligned}$$

The second equation above implies that $\boldsymbol{M_W}(\boldsymbol{y} - \boldsymbol{Y}\hat{\boldsymbol{\beta}}_2^{\rm IV}) = \boldsymbol{M_Z}(\boldsymbol{y} - \boldsymbol{Y}\hat{\boldsymbol{\beta}}_2^{\rm IV})$, since $\boldsymbol{M_W Z} = \boldsymbol{M_Z Z} = \mathbf{O}$, which implies that the numerator and denominator of κ, as given by (12.92), are equal when κ is evaluated at $\hat{\boldsymbol{\beta}}_2^{\rm IV}$. This shows that, in the just-identified case, the minimized value of the ratio (12.92) is equal to 1, and this value is achieved when $\boldsymbol{\beta}_2$ is equal to the IV estimator using $\boldsymbol{P_W}$ as the matrix of instruments.

If, in the overidentified case, we relax the overidentifying restrictions by adding some additional columns of $\boldsymbol{W}$ as regressors, we obtain the new structural equation

$$\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\beta}_1 + \boldsymbol{Z}'\boldsymbol{\beta}_1' + \boldsymbol{Y}\boldsymbol{\beta}_2 + \boldsymbol{u}, \tag{S12.43}$$

where $\boldsymbol{Z}'$ denotes the additional columns of $\boldsymbol{W}$, and $\boldsymbol{\beta}_1'$ denotes the vector of coefficients on them. Now recall that the restricted value of the loglikelihood function is given by expression (12.93). The only part of this function that depends on what exogenous variables are included in the structural equation is the term involving $\hat{\kappa}$. For the unrestricted model, $\hat{\kappa} = 1$, and so the maximized loglikelihood function for the unconstrained model is that for the constrained model but with $\hat{\kappa}$ replaced by 1, as we were asked to show.