Basic Idea of Instrumental Variables

Consider $A2linear$ with two sets of regressors, X^G and X^B:

$$y = X^G \beta^G + X^B \beta^B + \epsilon^{true}$$

or, in observation-by-observation form:

$$y_s = x_s^{G'} \beta^G + x_s^{B'} \beta^B + x_s^{B'} \beta^B + \epsilon_s^{true} = x_s' \beta + \epsilon_s^{true}$$

Given $A1$ and $A2$, the Sampling Error Vector of OLS is:

$$SEV(\hat{\beta}_{ols}) = (B_s)^{-1} \cdot \sum_s a_s^{ols} = (x_s x_s')^{-1} \cdot \sum_s x_s \epsilon_s$$

But:

$$Ea_s^{ols} = E \begin{bmatrix} x_s^G \epsilon_s \\ x_s^B \epsilon_s \end{bmatrix} = \begin{bmatrix} Ex_s^G \epsilon_s \\ Ex_s^B \epsilon_s \end{bmatrix} = \begin{bmatrix} = 0 \\ \neq 0 \end{bmatrix}$$

because we are told that the 'good' variables satisfy the weak exogeneity condition $Ex^G \epsilon_s = 0$, while the 'bad' variables do not (since they are *endogenous* w.r.t. to the error).
Therefore, OLS will be inconsistent for all the βs since in general X^G and X^B are correlated.

Reverse Engineering the Instrumental Variables Estimator

Suppose we can find a data matrix W of the same dimension as the original X and of full rank k. We then define:

$$W'\epsilon^{true} = \sum_s w_s \epsilon_s^{true} = \sum_s a_s^{ive}$$

such that:

$$E\sum_s a_s^{ive} = 0$$

at the true parameter values.
We now use the GMM idea and rely on the true *population* orthogonality conditions implied by the true model: $A2 : y = X\beta + \epsilon^{true}$ and

$$A3Rsru : W, X^G E w_s \epsilon^{ture} = 0, E x_s^{G'} = 0$$

Therefore, we define the GMM=IVE by using the *sample* orthogonality conditions:

$$W'\hat{\epsilon}^{ive} = W'(y - X\hat{\beta}_{ive}) = 0$$

to mimic the population OCs. Finally, solving for $\hat{\beta}^{ive}$ we obtain:

$$\hat{\beta}_{ive}) = (W'X)^{-1}W'y$$

because $W'X$ is square and invertible given that $rank(X) = rank(W) = k$.
In conclusion, the IVE will be consistent provided every column used to construct W is a weakly exogenous variable w.r.t. the true error (i.e., satisfies A3Rsru).