A simple procedure to estimate k structural parameters on conditionally endogenous variables with one conditionally mean independent instrument in linear models

Süß, Philipp

Goethe University Frankfurt

10 February 2015

A simple procedure to estimate k structural parameters on conditionally endogenous variables with one conditionally mean independent instrument in linear models

Philipp Süß*

February 10, 2015

Abstract

The following note proposes a simple procedure to estimate k parameters of interest in a linear model with potentially k conditionally endogenous variables of interest and m endogenous control variables in the presence of at least one instrumental variable under the assumption of conditional mean independence.

Keywords: Instrumental variables; Conditional independence assumption; Underidentified model
JEL Classification: C26

* Goethe University Frankfurt, Department of Applied Econometrics and International Economic Policy, Grüneburgplatz 1 60323 Frankfurt am Main, Germany Tel.: +49-69-798-34841; E-mail: Philipp.Suess@wiwi.uni-frankfurt.de

1 Introduction

In an attempt to uncover causal effects applied empirical researchers frequently resort to instrumental variables estimation using the 2SLS estimator to solve endogeneity issues. The well known order condition for identification states that the number of instruments has to be greater than or equal to the number of endogenous variables. This note proposes a procedure to estimate k parameters of interest in a linear model with potentially k conditionally endogenous variables of interest and m endogenous control variables with at least one conditionally independent instrumental variable z and describes a sufficient condition for consistency. In discussing the conditional mean independence assumption Stock (2010) states, that "focus shifted from developing a full model for y to estimating a single effect well". Following this reasoning, suppose a researcher is interested in the parameters β on k variables denoted by x and obtained m control variables w as well as at least one instrumental variable z. Assume that the researcher has a random sample of $\{(y_i, x_i, w_i, z_i)\}_{i=1}^{N}$, the necessary moments exist and the true model is linear in parameters.

$$y = x'\beta + w'\gamma + u \tag{1}$$

As is well known, if u is mean independent of x given w and the conditional mean of u is linear in w ($E(u|x,w) = w'\delta$), then the OLS estimator ($\hat{\beta}$) for the coefficient vector β is consistent, assuming the rank condition holds.[1] If the error term (u) is not mean independent of the variables of interest (x) conditional on the control variables (w), then OLS is generally inconsistent. The failure of conditional mean independence leads many researchers to consider instrumental variables strategies. It is easy to show that the the error (u) need not be mean independent of the instrumental variables (z). In the just- and overidentified case, it suffices for the consistency of the 2SLS estimator that u is mean independent of z given w and the conditional mean of u is linear in w ($E(u|z,w) = w'\delta$). The remainder of the note is structured as follows. In the next section I provide a proof for the consistency of the 2SLS estimator in the just- and overidentified case. In the third section I adapt the proof for the "underidentified" case, propose the main procedure and describe a (stronger) conditional independence assumption, which is sufficient for consistency. The main method is briefly outlined as follows. Assume there are k variables of interest and one valid instrumental variable for all k variables. To obtain consistent estimates for all k parameters perform k separate 2SLS regressions for each x_j $j \in \{1, ..., k\}$ using z as an instrument for x_j and x_{-j}, w serving as their own instruments. The last section concludes.

2 Proof just- and overidentified case

Proof of $plim(\hat{\beta}_{2SLS}) = \beta$ if $E(u|w,z) = E(u|w) = w'\delta$, in the case $dim(z) \geq dim(x)$:
Stacking all N observations, the model is given by:

$$Y \quad = \quad X\beta + W\gamma + U \tag{2}$$

[1] Frölich (2008) emphasizes that the linearity of the conditional mean of u in w ($E(u|x,w) = w'\delta$) is vital for consistency. Wooldridge (2005) shows that the inclusion of certain control variables can lead to a failure of the conditional mean independence assumption.

The 2SLS estimator can be written as an OLS estimator of a regression of y on $\hat{x}$ and w, where $\hat{x}$ denotes the predicted value from the linear projection from x on z and w ($\hat{x} = z'\hat{\theta}_1 + w'\hat{\theta}_2$). Denoting $\hat{e}$ as the corresponding residual, one can write $x = \hat{x} + \hat{e}$. Adding and subtracting $\hat{X}\beta$ to the equation above yields:

$$Y = \hat{X}\beta + W\gamma + (U + (X - \hat{X})\beta) \tag{3}$$

Using the Frisch Waugh Lovell theorem the coefficients $\hat{\beta}_{2SLS}$ can be obtained by partialling out W from $\hat{X}$ and Y. Let M_W denote the residual maker matrix of W given by $M_W = I - W(W'W)^{-1}W'$.

The 2SLS estimator is then:

$$\begin{aligned}
\hat{\beta}_{2SLS} &= ((M_W\hat{X})'(M_W\hat{X}))^{-1}(M_W\hat{X})'(M_W Y) \\
&= (\hat{X}'M_W\hat{X})^{-1}\hat{X}'M_W(X\beta + W\gamma + U) \\
&= (\hat{X}'M_W\hat{X})^{-1}\hat{X}'M_W X\beta + 0 + (\hat{X}M_W\hat{X})^{-1}\hat{X}'M_W U
\end{aligned}$$

The third line is due to $M_W W = 0$. Next I obtain the probability limit of the last term in the equation above:

$$\begin{aligned}
plim((\hat{X}'M_W\hat{X})^{-1}\hat{X}'M_W U) &= plim(N^{-1}\hat{X}'M_W\hat{X})^{-1}plim(N^{-1}\hat{X}'M_W U) \\
&= A \cdot E(\hat{X}M_W E(U|W,Z)) \\
&= A \cdot E(\hat{X}M_W W\delta) \\
&= 0
\end{aligned}$$

The third line is due to the key assumption $E(u|w,z) = E(u|w) = w'\delta$ and random sampling. The fourth line is due to $M_W W = 0$.

It remains to show that

$$\begin{aligned}
plim((\hat{X}'M_W\hat{X})^{-1}\hat{X}'M_W X\beta) &= \beta \\
plim(N^{-1}\hat{X}'M_W\hat{X})^{-1}plim(N^{-1}\hat{X}'M_W X)\cdot\beta &= \beta
\end{aligned}$$

If the first two terms on the LHS are equal, then $plim(\hat{\beta}) = \beta$. Use $\hat{X} = X - \hat{E}$.

$$plim(N^{-1}\hat{X}'M_W\hat{X}) = plim(N^{-1}\hat{X}'M_W X) - plim(N^{-1}\hat{X}'M_W\hat{E})$$

$\hat{E}$ is the vector of residuals from a regression of x on z and w. Hence the correlation of the predicted values and the residuals is zero and any regressor is uncorrelated with the residual (see FOC´s of OLS).

$$\begin{aligned}
N^{-1}\hat{X}'M_W\hat{E} &= N^{-1}\hat{X}'(I - W(W'W)^{-1}W')\hat{E} \\
&= N^{-1}\hat{X}'\hat{E} + \hat{X}'W(W'W)^{-1}N^{-1}W'\hat{E} \\
&= 0 + \hat{X}'W(W'W)^{-1}0 \\
&= 0
\end{aligned}$$

This completes the proof in the just- and overidentified case.

3 Proof "underidentified" case, the main procedure and example

The main procedure used to obtain consistent estimates for $\hat{\beta}_j$ $j \in \{1, ..., k\}$ is the following: Run k separate 2SLS regressions with the following set of instruments. For the estimation of the j'th coefficient ($\hat{\beta}_j$) use z as an instrument for x_j and x_{-j}, w serving as their own instruments. x_{-j} denotes all $x's$ without x_j. The following condition is sufficient for the consistency of each $\hat{\beta}_j$ ($plim(\hat{\beta}_j) = \beta_j \ \forall j$):

$$E(u|z, x_{-j}, w) = x'_{-j}\delta_{-j,1} + w'\delta_{-j,2} \ \forall j. \tag{4}$$

The proof of consistency $plim(\hat{\beta}_j) = \beta_j$ is just a matter of redefinition. Let $W^{new}_{-j} = [W, X_{-j}]$ be the new "W" and the proof above applies. The question if the condition $E(u|z, x_{-j}, w) = x'_{-j}\delta_{-j,1} + w'\delta_{-j,2}$ $\forall j$ holds, depends on the application at hand. An example where the assumption seems plausible is the following. Consider the goal of estimating the "causal" effect of various types of employment (for example industrial and service sectors employment) on violent crimes in a linear setting. Assume the error contains criminogenic traits, which lower employment prospects and have a stimulating effect on violent crimes. Consider firm bankruptcies as an instrument. In this example, the condition would hold if the number of bankruptcies is unrelated to criminogenic traits conditional on the number of employees in the service sector and that the number of bankruptcies are unrelated to criminogenic traits conditional on the number of employees in the industrial sector which could be seen as plausible.

4 Conclusion

The note showed that there are cases where one instrumental variable is sufficient for the estimation of $k \geq 1$ structural parameters in a linear model. A sufficient condition for consistency was spelled out which may help applied researchers assesing the plausibility of there results when estimating the k effects separately.

References

[1] Frölich, M., (2008): *Parametric and Nonparametric Regression in the Presence of Endogenous Control Variables, International Statistical Review 2008 (2)*, 214-227

[2] Stock, J. H., (2010): *The Other Transformation in Econometric Practice: Robust Tools for Inference, Journal of Economic Perspectives, 2010 (2)*, 83-94

[3] Wooldridge, J. M., (2005): *Violating Ignorability of treatment by controlling for too many factors, 2005 Econometric Theory 2005 (21),* 1026-1028