ISSN 0819-2642
ISBN 978 0 7340 3730 5

THE UNIVERSITY OF MELBOURNE

DEPARTMENT OF ECONOMICS

RESEARCH PAPER NUMBER 1020

December 2007

Conceptual Frameworks and Experimental Design in Simultaneous Equations

by

D.S. Poskitt & C.L. Skeels

Department of Economics
The University of Melbourne
Melbourne Victoria 3010
Australia.

Conceptual Frameworks and Experimental Design in Simultaneous Equations ⋆

D.S. Poskitt [a] C.L. Skeels [b,∗]

[a] *Department of Econometrics and Business Statistics, Monash University, Vic 3800, Australia*

[b] *Department of Economics, The University of Melbourne, Vic 3010, Australia*

Abstract

Using examples drawn from two important papers in the recent literature on weak instruments, we demonstrate how observed experimental outcomes can be profoundly influenced by the different conceptual frameworks underlying two experimental designs commonly employed when simulating simultaneous equations.

Key words: Simultaneous equations, Experimental design, Simulation experiment
JEL classification codes: C12, C15, C30

1 Simultaneous Equations and Experimental Design

The classical linear simultaneous equations model has enjoyed renewed interest of late as a consequence of the problems associated with inference in weakly identified models; see, for example, the papers discussed by Chesher, Dhaene, and van Dijk (2007). Several authors have addressed the specific problem of inference on the coefficient of an endogenous regressor in a structural equation and various suggestions have been made about how to proceed in such circumstances. Using invariance principles and similar regions, Andrews, Moreira, and Stock (2006) restrict attention to a class of tests from which they extract members with desirable optimality properties. However, in the absence of a

⋆ We thank both Adrian Pagan and an anonymous referee for useful feedback on an earlier version of this note, and acknowledge the support of the Australian Research Council Grant DP0771445.

∗ Corresponding author. Tel.: +61 3 8344 3783; fax: +61 3 8344 6899.
Email address: Chris.Skeels@unimelb.edu.au (C.L. Skeels).

uniformly most powerful test, comparison of the absolute and relative merits of different tests continues to be guided by simulation experiments.

One aspect of this analysis that has received little attention is the observation that there are two formulations of the underlying model in common use. These formulations are probabilistically equivalent, in that one can move from one to the other via non-singular linear transformations. However, these transformations involve parameters that may be of interest in certain types of simulation experiments and, in such cases, the competing formulations of the models can be conceptually quite different. In this note we demonstrate that the lessons one might hope to learn through the use of simulation studies can vary enormously depending upon which conceptual framework is chosen to guide the experiments. We illustrate this point by exploring the behaviour of three statistics that have been important in the analysis of weakly-identified simultaneous equations models, namely the AR test (Anderson and Rubin, 1949), the K test (Kleibergen, 2002), and the conditional likelihood ratio (CLR) test (Moreira, 2003).

To begin, consider the classical structural equation model

$$y = Y\beta + X\gamma + u, \tag{1}$$

where the endogenous matrix variables y and Y are $N \times 1$ and $N \times n$, respectively, the matrix of exogenous variables X is $N \times k$, and u denotes a $N \times 1$ vector of uncorrelated stochastic disturbances with zero mean and variance σ_u^2. The vectors of structural coefficients β and γ are $n \times 1$ and $k \times 1$, respectively.

There are two commonly encountered ways of completing the specification of this model. The first involves augmenting (1) by reduced form equations for Y; namely,

$$Y = X\Pi_1 + Z\Pi_2 + V, \tag{2}$$

with Π_1 and Π_2 of dimension $k \times n$ and $\nu \times n$, respectively. The stochastic specification is then completed by assumptions about the conditional distribution of $[u\ V]$ given X and Z, e.g. $[u\ V]|[X\ Z] \sim N(0, \Phi \otimes I_N)$, where

$$\Phi = \begin{bmatrix} \sigma_u^2 & \sigma_{uV} \\ \sigma_{Vu} & \Omega \end{bmatrix}.$$

That is, the rows of the $N \times (n+1)$ matrix $[u\ V]$ are uncorrelated random vectors with zero mean and common $(n+1) \times (n+1)$ covariance matrix Φ. Henceforth, the couplet of equations (1) and (2), together with the accompanying distributional assumption, will be referred to as the structural equation specification (SES).

The second specifies a reduced form for all of the endogenous variables in the system; namely

$$[y\ Y] = [X\ Z] \begin{bmatrix} \pi_1 & \Pi_1 \\ \pi_2 & \Pi_2 \end{bmatrix} + [v\ V], \tag{3}$$

where $[v\ V]|[X\ Z] \sim N(0, \Sigma \otimes I_N)$ with

$$\Sigma = \begin{bmatrix} \sigma_v^2 & \sigma_{vV} \\ \sigma_{Vv} & \Omega \end{bmatrix}, \tag{4}$$

and the coefficient vectors π_1 and π_2 are $k \times 1$ and $\nu \times 1$, respectively. Hereafter equations (1) and (3), together with their distributional assumption, will be referred to as the reduced form specification (RFS). This model is comprised of more equations than there are endogenous variables and so compatibility of equations (1) and (3) requires the parameter restrictions

$$\pi_1 - \Pi_1\beta = \gamma, \qquad \pi_2 = \Pi_2\beta, \qquad \sigma_u^2 = [1, -\beta']\Sigma[1, -\beta']', \tag{5}$$

which, together with (2), imply that

$$\Sigma = \begin{bmatrix} 1 & \beta' \\ 0 & I_n \end{bmatrix} \Phi \begin{bmatrix} 1 & 0' \\ \beta & I_n \end{bmatrix}. \tag{6}$$

Equations (5) and (6) show how the parameters of the two formulations of the model are related to each other. They also make clear that, in simulation experiments where values of β are varied, it is impossible to simultaneously keep fixed the remaining parameters of both formulations. Hence, contingent upon which formulation of the model you prefer, the other becomes something of a moving feast as β is varied, making comparison of simulation experiments across such paradigms very difficult. When investigating power, for example, the difficulty lies in presenting the power as a univariate function of β when in fact the power curve sits on a multidimensional manifold. Making a choice between the two formulations of the model implies that one is traversing this manifold and passing through observationally equivalent parameter points in very different ways.

In order to illustrate the point, let us consider the problem of testing H_0 : $\beta = 0$ against the two sided alternative $H_1 : \beta \neq 0$ under the following two experimental designs based upon SES and RFS.

$$
\begin{aligned}
y &= Y\beta + u \\
Y &= Z\Pi_2 + V \\
[u\ V] &\sim N\left(0, \Phi \otimes I_N\right) \\
\Phi &= \begin{bmatrix} 1 & \rho_{uV} \\ \rho_{uV} & 1 \end{bmatrix}
\end{aligned}
$$

Experimental Design 1 (ED1)

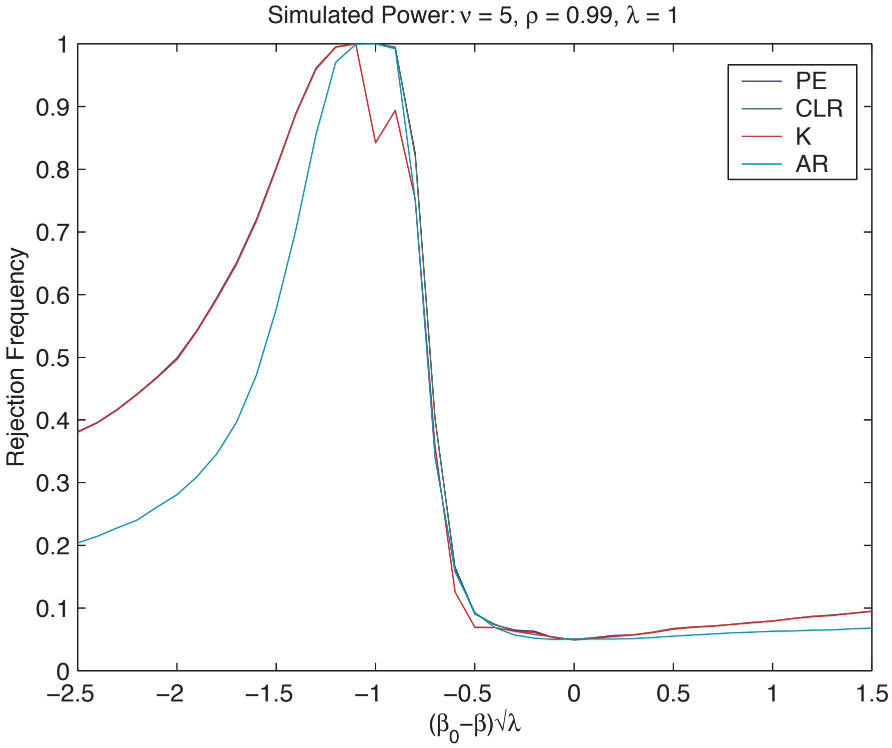
$$
\begin{aligned}
y &= Z\Pi_2\beta + v \\
Y &= Z\Pi_2 + V \\
[v\ V] &\sim N\left(0, \Sigma \otimes I_N\right) \\
\Sigma &= \begin{bmatrix} 1 & \rho_{vV} \\ \rho_{vV} & 1 \end{bmatrix}
\end{aligned}
$$

Experimental Design 2 (ED2)

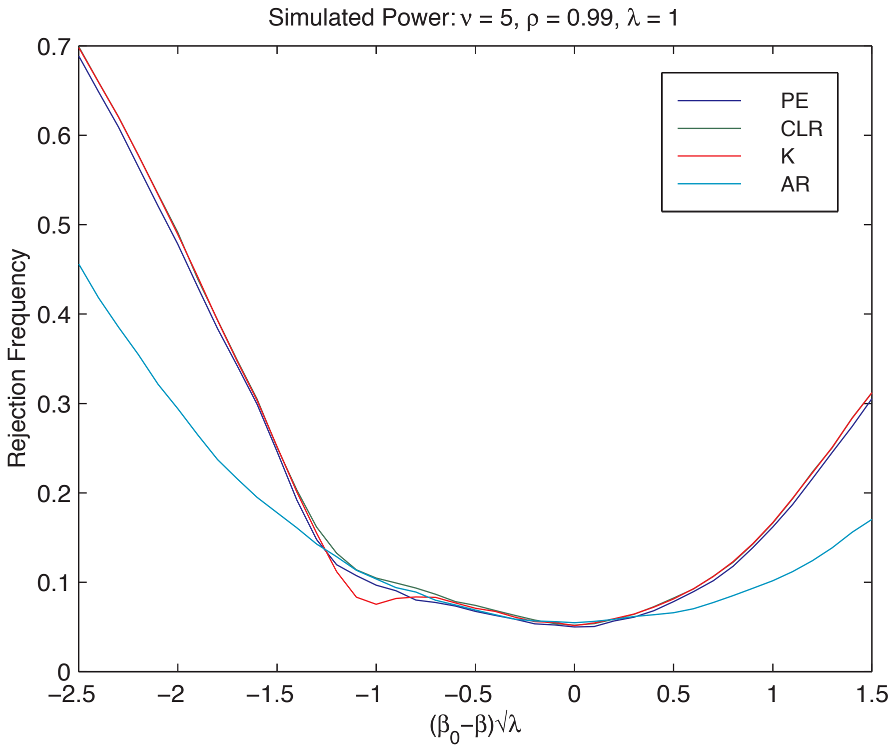
Note from (6) that, under H_0, $\Sigma = \Phi$ and the two data generating mechanisms are observationally equivalent for any fixed values of the nuisance parameters. But the latter is not true in general. From equation (6) it is apparent that, as β varies under H_1, one can choose to hold fixed either Σ or Φ, but not both. Hence it will be immaterial whether simulation experiments designed to investigate size properties are based on ED1 or ED2, but if one allows β to vary under H_1, as one does when considering the power of certain tests, then the implications of using ED1 or ED2 can be very different.

Figures 1 and 2 present power functions for the AR, K and CLR tests based on experimental designs ED1 and ED2, together with PE, the asymptotically efficient two-sided power envelope for invariant, similar tests, as described in Andrews et al. (2006). For both designs we have $n = 1$, $N = 100$ and $\nu = 5$. To obtain Figure 1 we have set $\lambda = \Pi_2' Z' Z \Pi_2 = 1.0$, so the instruments are weak. Figure 1(a) is based on ED1 with $\rho_{uV} = 0.99$, following Kleibergen (2002, Figure 4). Figure 1(b) examines the effect of such weak instruments in ED2 where, instead of fixing ρ_{uV}, we set $\rho_{vV} = 0.99$. In Figure 2 we have set $\lambda = 5.0$, so the instruments are stronger than for Figure 1. Figure 2(b), as in Andrews et al. (2006, Figure 1(a)), is based on ED2 with $\rho_{vV} = 0.95$, so that the degree of endogeneity is weaker than in Figure 1. Figure 2(a) examines the corresponding power curves under ED1, where now $\rho_{uV} = 0.95$.

The figures clearly demonstrate that the choice of experimental design has a profound effect upon the observed power characteristics of the different tests and hence a substantial influence on any conclusions that are likely to be drawn from the associated simulation experiments. That said, one result that is common across the figures is that the CLR test dominates the K test and has a performance that is remarkably close to that of the power envelope. At least in our examples there is little between them although the K test appears to have points in the parameter space where it displays erratic behaviour, a property not shared by CLR. However, it is only with ED2 that either K or CLR have properties that might be considered desirable. Clearly, the attractiveness of these procedures appears to be contingent upon the conceptual framework

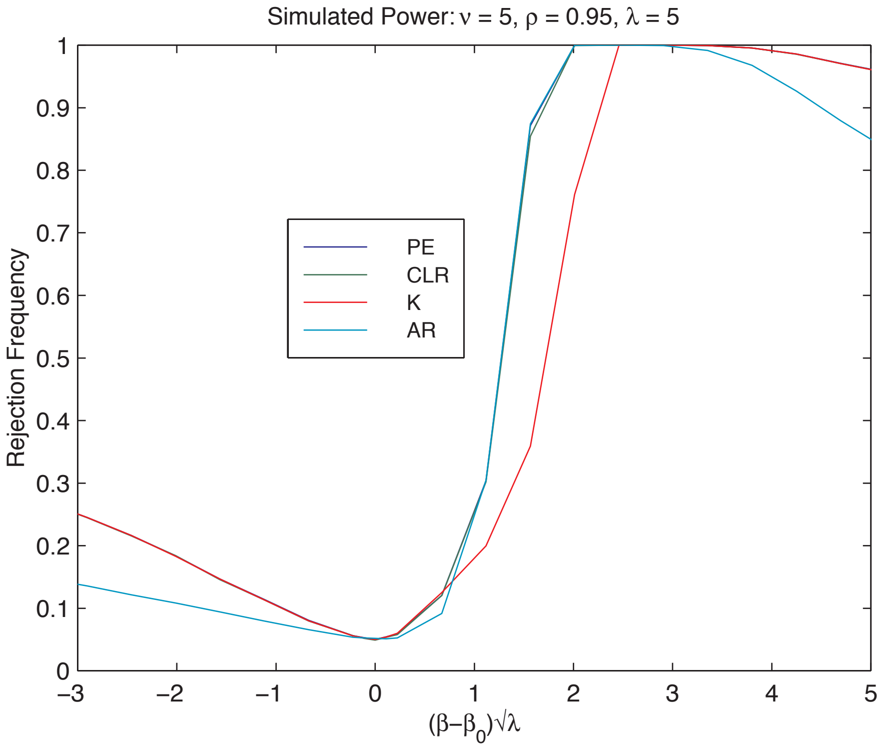


(a) Power Functions in ED1 (Kleibergen, 2002, Figure 4)

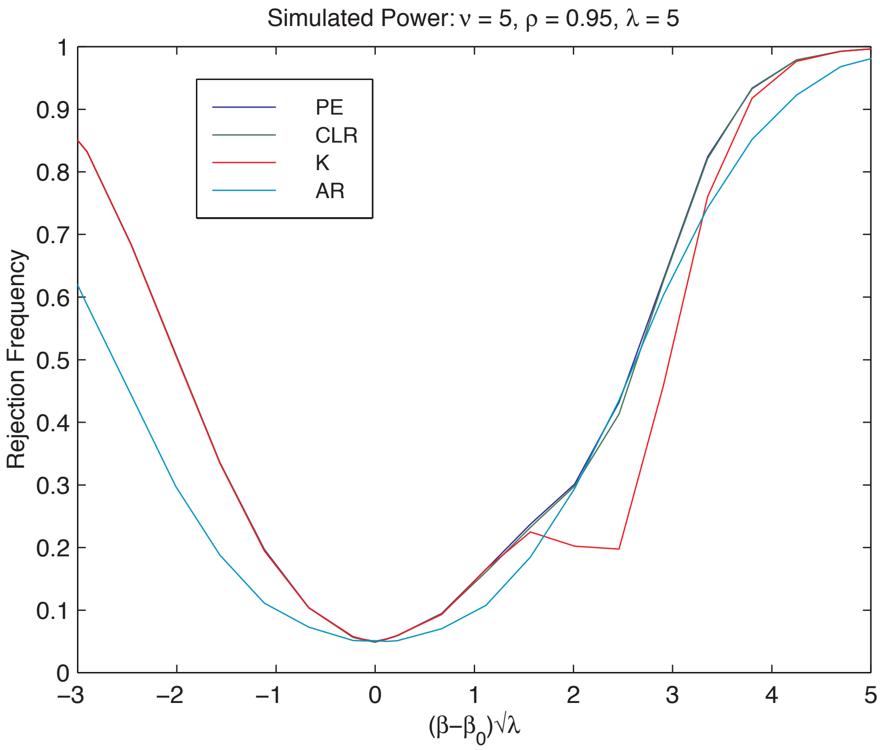


(b) Power Functions in ED2

Fig. 1. Power Functions: Weak Instruments and Strong Endogeneity



(a) Power Functions in ED1



(b) Power Functions in ED2 (Andrews et al., 2006, Figure 1(a))

Fig. 2. Power Functions: Stronger Instruments But Weaker Endogeneity

within which their behaviour is analysed.

2 Discussion

Let us state at the outset that we are not suggesting that either ED1 or ED2 is incorrect. Nevertheless, when faced with alternative specifications one is required to make choices. A preference for the specification that is comprised of both (i) a structural equation and (ii) the complete reduced form rather than just a subset of the reduced form, might be justified on two grounds. First, the overall reduced form is an unrestricted regression and so the accompanying conditional distribution of $[v\ V]$ given X and Z is, to our minds, somewhat easier to interpret than is that of $[u\ V]$ given X and Z. Although structural equation (1) is of primary economic interest, its probabilistic standing is less clear without reference to the corresponding complete reduced form model (3) and the compatibility conditions (5). Second, if one wishes to appeal to asymptotic arguments, such as the local-to-zero asymptotics of Staiger and Stock (1997), then a pervasive result is that Σ can be consistently estimated. [1] This suggests that a model in which Σ is fixed *a priori*, as in RFS, is the appropriate statistical model and, consequently, an experimental design where Σ is held constant, as in ED2, is of intrinsic appeal. Conversely, given that it is the structural equation that is of primary interest, it can be argued that σ_u^2 and σ_{uV} are natural parameters of the model that determine the variation and the degree of endogeneity in the SES, and therefore holding Φ constant *a priori*, as in ED1, is a sensible feature of an experimental design.

We recognize that a choice between specifications SES and RFS remains a matter of taste but, at the same time, stress that this choice carries with it strong implications for the observed operational characteristics of test procedures. Although it has not been completely ignored (see, for example, Forchini and Hillier, 2003), we believe that the literature has dramatically underestimated the practical consequences of this observation. In the absence of compelling arguments in favour of one conceptual framework, and its associated experimental design, over the other our results suggest that it would be prudent to explore both designs when analysing the behaviour of tests such as those considered here. Ultimately, of course, what is needed is an explanation of the observed operational characteristics that is valid whatever conceptual framework is adopted. Preliminary investigations using the ideas and results presented in Poskitt and Skeels (2007) suggest that such an explanation might be possible, but a detailed examination is beyond the scope of this paper.

[1] We have benefited from a discussion with James Stock on this point.

References

Anderson, T. W., Rubin, H., 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations. Annals of Mathematical Statistics 20 (1), 46–63.

Andrews, D. W. K., Moreira, M. J., Stock, J. H., 2006. Optimal two-sided invariant similar tests for instrumental variables regression. Econometrica 74 (3), 715–752.

Chesher, A., Dhaene, G., van Dijk, H., 2007. Endogeneity, instruments and identification. Journal of Econometrics 139 (1), 1–3.

Forchini, G., Hillier, G. H., October 2003. Conditional inference for possibly unidentified structural equations. Econometric Theory 19 (5), 707–743.

Kleibergen, F., 2002. Pivotal statistics for testing structural parameters in instrumental variables regression. Econometrica 70 (5), 1781–1803.

Moreira, M. J., 2003. A conditional likelihood ratio test for structural models. Econometrica 71 (4), 1027–1048.

Poskitt, D. S., Skeels, C. L., 2007. Approximating the distribution of the instrumental variables estimator when the concentration parameter is small. Journal of Econometrics 139 (1), 217–236.

Staiger, D., Stock, J. H., 1997. Instrumental variables regression with weak instruments. Econometrica 65 (3), 557–586.