FACULTY OF ACTUARIAL SCIENCE & ECONOMETRICS

A & E REPORT

REPORT AE 23/85

THE IMPORTANCE AND PERFORMANCE OF TESTS FOR
THE SELECTION OF INSTRUMENTAL VARIABLES

J.F. Kiviet

University of Amsterdam

Title: *THE IMPORTANCE AND PERFORMANCE OF TESTS FOR THE SELECTION OF INSTRUMENTAL VARIABLES*

Author : Jan F. Kiviet

Address: Faculty of Actuarial Science and Econometrics
University of Amsterdam
Jodenbreestraat 23
1011 NH Amsterdam

Date: November 1985

Series and Number: Report AE 23/85

Pages: 22

Price: No charge

JEL Subject Classification: 210

Keywords: specification testing, instrumental variables, exogeneity, interdependence of tests

Abstract:
In this paper it is demonstrated that testing models for misspecification by OLS-based variable addition may lead to the acceptance of reduced form equations or other hybrid relationships derived from the structural form. The awkward result is obtained that in order to acquire consistent estimates of structural parameters some regressors that appear as significant are sometimes better removed from the specification. This illustrates the need for tests on the validity of exogeneity assumptions. From application of the variable addition approach to a single linear simultaneous regression model three basic types of tests for instrument adequacy are developped, viz. tests for exogeneity of the maintained instruments, for exogeneity of excluded regressors, and for exogeneity of included regressors. Finally it is shown which orderings of these tests and tests on the adequacy of the specification will lead to asymptotic independence of the test statistics. Such orderings enable to exercise control over the overall type I error probability in a comprehensive model selection procedure.

1. INTRODUCTION AND SUMMARY

In specification testing of econometric relationships most attention is usually paid to the investigation of the adequacy of the set of explanatory variables and to the validity of the assumptions on the second moment matrix of the disturbance vector. In this paper it is demonstrated in which aspects the interpretation of many specification test statistics may be blurred through the adoption of incorrect exogeneity assuptions. First it is shown that in situations where little a priori information on the specification of the structural relationship is available the use of (mis)specification tests according to the variable addition principle may induce revisions of a correct specification, and then may lead to the acceptance of hybrid forms of structural and reduced form equations or of other derived relationships as model specification. By a simple example it is illustrated that the habitual judgement of OLS regressions by t ratio's is not void of pitfalls: The awkward result is obtained that in order to acquire consistent estimates of structural parameters, some regressors that appear as significant are sometimes better removed from the specification.

Having pointed out the importance of the validity of exogeneity assumptions, we review tests on joint dependence of regressors and disturbances, on instrumental variable adequacy and on overidentifying restrictions. Three basic types of tests on exogeneity emerge from application of the variable addition principle to the single linear simultaneous regression model, viz. tests for exogeneity of the set of maintained instruments, and two types of tests for exogeneity of variables that are not (yet) included in the set of instruments, viz. tests for variables that are excluded from the regression and tests for exogeneity of regressors of the structural model. Finally it is investigated how these three types of tests can be combined with the test on coefficient restrictions to solve empirically the comprehensive problem of both finding an adequate specification of the structural form and composing a set of admissible instrumental variables. Examination of the interdependence of the various tests, when applied to particular sequences of hypotheses, leads to some suggestions for ordenings in the employment of the different test procedures. These ordenings permit control of the overall type I error of the complete model selection procedure, as for all tests particular sequences are found that entail asymptotic independence of the test statistics involved.

2. VARIABLE ADDITION AND THE IMPORTANCE OF EXOGENEITY

Many checks on the validity of the specification of a regression model can be formulated as tests of the significance of extra regressors. Pagan (1984) presents a number of examples. For a single linear model estimated by ordinary least-squares (OLS) the framework in which such tests are used as tools for model validation and possibly for model (re)specification may be described as follows.

Let the present model specification for the Tx1 vector of observations on the dependent variable y be

$$y = X\beta + \varepsilon \tag{2.1}$$

where X is a TxK matrix of regressors and ε a Tx1 stochastic disturbance vector. Then a modelbuilder usually wants to supplement the OLS estimate $b=(X'X)^{-1}X'y$ of the coefficient vector β not only with an estimate of its dispersion, but also with one or more statistics that support the (often implicit) assumption that OLS is an appropriate technique for the supposed relationship (1.1). We shall clarify here, as has been done before inter alia by Pagan and Hall (1983) and Davidson and MacKinnon (1985), that insignificant values of statistics on the significance of extra regressors may provide such evidence, and that significant values of such tests may disqualify the original model, and perhaps may induce a revision of this specification. However, it will be demonstrated also that this type of inference is heavily dependent on the validity of a number of assumptions. We will focus on the effects of improper exogeneity assumptions which may entail that incorrect conclusions with respect to the adequacy of the model specification are drawn due to the application of an incorrect or suboptimal estimation technique.

Evidence obtained from variable addition on the likely consistency of b for β may be based on the following reasoning. Suppose that a consistent estimator of β could in theory be obtained from the application of OLS to the regression

$$y = X\beta + \bar{X}\bar{\beta} + \bar{\varepsilon} \ , \tag{2.2}$$

where $\bar{X}$ is a Tx$\bar{K}$ matrix of (usually unknown) extra regressors, and let the evidence on the consistency of b in (2.1) take the form of a test on the significance of the coefficients $\tilde{\beta}$ in the regression

$$y = X\beta + \tilde{X}\tilde{\beta} + \tilde{\varepsilon} \ , \tag{2.3}$$

where $\tilde{X}$ is a Tx$\tilde{K}$ matrix of observable regressors. The matrix $\tilde{X}$ may contain powers or other functions of variables from X, other variables, lagged residuals from (2.1), dummies etc. Of course we have

$$\varepsilon = \bar{X}\bar{\beta} + \bar{\varepsilon} \quad \text{and} \quad \tilde{\varepsilon} = \bar{X}\bar{\beta} - \tilde{X}\tilde{\beta} + \bar{\varepsilon} \ . \tag{2.4}$$

As X may contain redundant regressors (giving zero elements in β), the only restrictions so far are the linearity of the models, and the necessity to have $\tilde{K}<T-K$. As consistent estimates of β may be obtained from various specifications, (2.2) is not necessarily unique.

Now let us further assume that

$$\begin{aligned} &\text{plim } T^{-1}X'X = \Pi_{xx} \; ; \; \text{plim } T^{-1}\bar{X}'\bar{X} = \Pi_{\bar{x}\bar{x}} \\ &\text{plim } T^{-1}X'\bar{X} = \Pi_{x\bar{x}} \; ; \; \text{plim } T^{-1}\tilde{X}'\tilde{X} = \Pi_{\tilde{x}\tilde{x}} \\ &\text{plim } T^{-1}X'\tilde{X} = \Pi_{x\tilde{x}} \; ; \; \text{plim } T^{-1}\bar{X}'\tilde{X} = \Pi_{\bar{x}\tilde{x}} \end{aligned} \tag{2.5}$$

and that all these matrices have finite elements, and further that Π_{xx}, $\Pi_{\bar{x}\bar{x}}$ and $\Pi_{\tilde{x}\tilde{x}}$ are of full rank, while the matrices $[X:\bar{X}]$ and $[X:\tilde{X}]$ have full column rank. Finally we suppose that

$$T^{-\frac{1}{2}}X'\bar{\varepsilon} \xrightarrow{L} N(0,\sigma^2\Pi_{xx}^{-1}) \tag{2.6}$$

and

$$T^{-\frac{1}{2}}X'M_{\bar{x}}\bar{\varepsilon} \xrightarrow{L} N(0,\sigma^2[\Pi_{xx}-\Pi_{x\bar{x}}\Pi_{\bar{x}\bar{x}}^{-1}\Pi'_{x\bar{x}}]) \tag{2.7}$$

where M_A denotes $I-A(A'A)^{-1}A'$ for full column rank matrices A. From these assumptions it follows that $\text{plim } T^{-1}X'\bar{\varepsilon}=0$ and $\text{plim } T^{-1}\bar{X}'\bar{\varepsilon}=0$, hence the regressors are predetermined in (2.2) and - at least in theory - β could be estimated consistently by OLS from regression (2.2).

2.1. <u>The variable addition test under full exogeneity</u>

We now derive the asymptotic distribution of the test statistic on the significance of the extra regressors in (2.3) under some extra suppositions. The test statistic is given by

$$S(\tilde{\beta}) = y'M_x\tilde{X}(\tilde{X}'M_x\tilde{X})^{-1}\tilde{X}'M_xy/\hat{\sigma}^2 \ , \tag{2.8}$$

where $\hat{\sigma}^2$ is some estimator of σ^2. We will consider separately $\hat{\sigma}_0^2=T^{-1}y'M_xy$ and $\hat{\sigma}_1^2=T^{-1}y'M_{x:\tilde{x}}y$, which give rise to $S_0(\tilde{\beta})$ and $S_1(\tilde{\beta})$ respectively. From substitution of (2.2) in (2.8) we find for i=0,1

$$\begin{aligned} S_i(\tilde{\beta}) = [\,&\bar{\beta}'\bar{X}'M_x\tilde{X}(\tilde{X}'M_x\tilde{X})^{-1}\tilde{X}'M_x\bar{X}\bar{\beta} \\ &+ 2\,\bar{\beta}'\bar{X}'M_x\tilde{X}(\tilde{X}'M_x\tilde{X})^{-1}\tilde{X}'M_x\bar{\varepsilon} + \bar{\varepsilon}'M_x\tilde{X}(\tilde{X}'M_x\tilde{X})^{-1}\tilde{X}'M_x\bar{\varepsilon}\,]\,/\hat{\sigma}_i^2 \end{aligned} \tag{2.9}$$

and we use this to investigate the distribution of $S_i(\tilde{\beta})$ in various situations.

If we suppose

$$T^{-\frac{1}{2}}\tilde{X}'M_x\bar{\varepsilon} \xrightarrow{L} N(0,\sigma^2[\Pi_{\tilde{x}\tilde{x}}-\Pi'_{x\tilde{x}}\Pi_{xx}^{-1}\Pi_{x\tilde{x}}]) \tag{2.10}$$

then using (2.2) through (2.7) this implies $\text{plim}\ T^{-1}\tilde{X}'\bar{\varepsilon}=0$ and we find

$$\text{plim}\ \hat{\sigma}_0^2 = \text{plim}\ T^{-1}y'M_x y \tag{2.11}$$

$$= \sigma^2 + \bar{\beta}'[\Pi_{\bar{x}\bar{x}}-\Pi'_{x\bar{x}}\Pi_{xx}^{-1}\Pi_{x\bar{x}}]\bar{\beta} \geq \sigma^2$$

and

$$\text{plim}\ \hat{\sigma}_1^2 = \text{plim}\ T^{-1}y'M_{x:\tilde{x}}y \tag{2.12}$$

$$= \text{plim}\ T^{-1}(\bar{X}\bar{\beta}+\bar{\varepsilon})'[M_x-M_x\tilde{X}(\tilde{X}'M_x\tilde{X})^{-1}\tilde{X}'M_x](\bar{X}\bar{\beta}+\bar{\varepsilon}) \geq \sigma^2 .$$

Both probability limits exist (i.e. have finite constant values); besides we find

$$\text{plim}\ \hat{\sigma}_0^2 \geq \text{plim}\ \hat{\sigma}_1^2 \geq \sigma^2 \tag{2.13}$$

and it can also be derived that the equality signs apply under the extra assumption

$$\bar{\beta} = 0 . \tag{2.14}$$

If we adopt (2.14) then we have $\bar{\varepsilon}=\varepsilon$ and hence $\text{plim}\ T^{-1}X'\varepsilon=0$ which gives $\text{plim}\ b=\beta$.

From the above the following general results follow. Given (2.1) through (2.7) and (2.10), hence under exogeneity of X, $\bar{X}$ and $\tilde{X}$, we find that the statistics $S_0(\tilde{\beta})$ and $S_1(\tilde{\beta})$ tend to a $\chi^2(\tilde{K})$ variate for $T\to\infty$ if $\bar{\beta}=0$. If however $\bar{\beta}\neq 0$ then the denominators of $S_0(\tilde{\beta})$ and $S_1(\tilde{\beta})$ tend to a finite value whereas it follows from (2.9) that the numerator tends to an infinite value in general. Except when $\text{plim}\ T^{-\frac{1}{2}}\bar{X}'M_x\tilde{X}$ is finite, which implies

$$\text{plim}\ T^{-1}\bar{X}'M_x\tilde{X} = 0, \tag{2.15}$$

the numerator of the test statistics will assume finite values asymptotically. So we may conclude that under (2.1) through (2.7) and (2.10) we will obtain a value of the test statistic exceeding the α-level critical value of the $\chi^2(\tilde{K})$ distribution whith probability one asymptotically, provided that $\bar{\beta}\neq 0$ and $\text{plim}\ T^{-1}\bar{X}'M_x\tilde{X}\neq 0$. If $\bar{\beta}\neq 0$ but (2.15) applies, hence at any rate when $M_x\bar{X}$ and $M_x\tilde{X}$ span orthogonal vector spaces, then the inequality signs apply in (2.13) and we have

$$\frac{\operatorname{plim} \hat{\sigma}_i^2}{\sigma^2} \; S_i(\tilde{\beta}) \xrightarrow{L} \chi^2(\tilde{K}) \; . \qquad (2.16)$$

So then the asymptotic rejection probability of $S_0(\tilde{\beta})$ will be smaller than that of $S_1(\tilde{\beta})$, and this will be even smaller than α.

Thus under the maintained assumptions (2.5), (2.6), (2.7) and (2.10) the tests $S_i(\tilde{\beta})$ constitute consistent tests for the consistency of b, except for the case (2.15), where the search for misspecification goes in a completely wrong direction. A significant result of $S_i(\tilde{\beta})$ may be used to extend the regression model in such a way that it better suits the requirements to produce consistent OLS estimates.

2.2. The variable addition tests when regressors are jointly-dependent

The above results depend heavily upon the validity of the assumptions (2.6), (2.7) and (2.10). Let we extend the analysis now and let (2.2) represent the structural form of the relationship while

$$\operatorname{plim} T^{-1}X'\bar{\varepsilon} = \pi_{x\bar{\varepsilon}} \; , \; \operatorname{plim} T^{-1}\bar{X}'\bar{\varepsilon} = \pi_{\bar{x}\bar{\varepsilon}} \text{ and } \operatorname{plim} T^{-1}\tilde{X}'\bar{\varepsilon} = \pi_{\tilde{X}\bar{\varepsilon}}, \qquad (2.17)$$

where $\pi_{x\bar{\varepsilon}}$, $\pi_{\bar{x}\bar{\varepsilon}}$ and $\Pi_{\tilde{X}\bar{\varepsilon}}$ consist of finite elements not all necessarily equal to zero. Under the assumptions (2.1) through (2.5) and (2.17) the probability limits of $\hat{\sigma}_0^2$ and $\hat{\sigma}_1^2$ are still finite. It is easily seen from (2.9) that now $S_i(\tilde{\beta})$ will tend to infinity for $T\to\infty$ if

$$\operatorname{plim} T^{-1}\tilde{X}'M_x\bar{\varepsilon} = \pi_{\tilde{X}\bar{\varepsilon}} - \Pi'_{x\tilde{X}}\Pi_{xx}^{-1}\pi_{x\bar{\varepsilon}} \neq 0 \; , \qquad (2.18)$$

irrespective of the validity of (2.15) and the value of $\bar{\beta}$ and of $\pi_{\bar{x}\bar{\varepsilon}}$. Hence, even if model (2.1) is a 'correct specification' (X contains all relevant regressors thus $\bar{K}$=0) then significant values of $S_i(\tilde{\beta})$ will be found asymptotically if (2.18) applies. If we have $\bar{\beta}$=0, (2.18) may be written as

$$\operatorname{plim} \frac{1}{T}\tilde{X}'M_x\varepsilon = \pi_{\tilde{X}\varepsilon} - \Pi'_{x\tilde{X}}\Pi_{xx}^{-1}\pi_{x\varepsilon} \neq 0 \; , \qquad (2.19)$$

as then $\bar{\varepsilon}=\varepsilon$, and two special cases of (2.19) deserve further considerations:

(i) $\pi_{\tilde{X}\varepsilon}=0$ and $\pi_{x\varepsilon}\neq 0$ and hence $\Pi_{x\tilde{X}}\neq 0$.

Here (some of) the initial regressors are jointly dependent with y.

In regression (2.1) the exogenous regressors $\tilde{X}$ do not occur in the explanatory part of this (structural) relationship. However, they appear as significant due to $\pi_{x\varepsilon}\neq 0$ (the wrong application of OLS) and $\Pi_{x\tilde{x}}\neq 0$. Acceptance of the regressors $\tilde{X}$ in the specification would involve the conversion of the relationship in a hybrid form of both the structural and the reduced form of this equation. Here the better course of action in a modelling exercise would not be inclusion of the (seemingly significant) extra regressors $\tilde{X}$, but the replacement of OLS by an instrumental variable technique. The variables $\tilde{X}$ are feasible instruments here as $\pi_{\tilde{x}\varepsilon}=0$ and $\Pi_{x\tilde{x}}\neq 0$.

(ii) $\pi_{\tilde{x}\varepsilon}\neq 0$ and $\pi_{x\varepsilon}=0$.

Here the regressors in X are predetermined in (2.1) and hence b is a consistent estimator. However, due to correlation between $\tilde{X}$ and ε doubts are raised with respect to the adequacy of the specification because $\tilde{X}$ appears to contribute significantly to the explanation of y. This may occur when a single equation of a seemingly unrelated regression (SUR) system is estimated by OLS and dependent variables of some of the other SUR-equations are used in the matrix $\tilde{X}$. Then a (seemingly) significant decrease in the residual variance will result, although inclusion of the extra regressors may impede the assesment of the actual structural parameters of the relationship, if $\Pi_{x\tilde{x}}\neq 0$. This is illustrated by a simple example in the appendix.

The two cases discussed above constitute serious pitfalls the modelbuilder faces when he validates estimation results implicitly or explicitly by use of variable addition. These examples illustrate that it may be wise to remove regressors from a specification because they appear as significant! This analysis can easily be extended to cover the case of a single equation estimated by use of instrumental variables (IV), where again it can be shown that an estimated coefficient may assume a significant value not only because it concerns a genuine structural coefficient, but also when the exogeneity assumptions of the estimation technique employed do not hold.

3. TESTS FOR SPECIFICATION AND IV ADMISSIBILITY

We extend our framework here in such a way that we can indicate separately both the problems of specifying the structural relationship (testing for omitted regressors and for coefficient restrictions) and of composing an adequate set of instrumental variables (testing for exogeneity or for instrumental variable validity) for the same particular structural relationship. Therefore we use a notation for the single linear simultaneous equation here in which alternative assumptions on both issues can easily be expressed. Let the model now be

$$y = Y_1\alpha + Z_1\gamma + W_1\delta + \varepsilon = X\beta + \varepsilon \tag{3.1}$$

and

$$Y = [Y_0 : Y_1],\ Z = [Z_0 : Z_1],\ W = [W_0, W_1]\ ,$$

$$X = [Y_1 : Z_1 : W_1] \quad \text{and} \quad \beta' = (\alpha' : \gamma' : \delta')\ ,$$

where Y_i, Z_i and W_i are TxG_i, TxL_i and TxM_i matrices respectively and X is a TxK matrix with $K=G_1+L_1+M_1$ the number of coefficients in β, and where ε is a stochastic disturbance term of T elements. Hence, T is again the sample size, and the regressors $[Y_0 : Z_0 : W_0]$ are excluded a priori from the relationship. We suppose that the following probability limits exist:

$$\begin{aligned} &\text{plim}\ T^{-1}Z'Z = \Pi_{zz}\ ;\ \text{plim}\ T^{-1}W'W = \Pi_{ww}\ ;\ \text{plim}\ T^{-1}Z'W = \Pi_{zw}\ ; \\ &\text{plim}\ T^{-1}Z'X = \Pi_{zx}\ ;\ \text{plim}\ T^{-1}W'X = \Pi_{wx}\ , \end{aligned} \tag{3.2}$$

where Π_{zz}, Π_{ww} and Π_{zx} all have full column rank, hence

$$L = L_0 + L_1 \geq G_1 + L_1 + M_1 = K\ . \tag{3.3}$$

Further we write

$$\text{plim}\ T^{-1}Y'\varepsilon = \pi_{Y\varepsilon}\ ;\ \text{plim}\ T^{-1}Z'\varepsilon = \pi_{z\varepsilon}\ ;\ \text{plim}\ T^{-1}W'\varepsilon = \pi_{w\varepsilon}$$

and assume

$$T^{-\frac{1}{2}}Z'\varepsilon \xrightarrow{L} N(0,\sigma^2\Pi_{zz})\ , \tag{3.4}$$

hence a priori we have $\pi_{z\varepsilon} \equiv 0$, thus Z contains admissible instrumental variables, and because of (3.3) β is consistently estimable by means of the IV technique. We admit $\pi_{Y\varepsilon} \neq 0$ so the matrix Y contains jointly dependent variables, while W contains variables of which the status has yet to be assessed empirically. We shall consider tests now for (zero) restrictions on elements of β, for checks on $\pi_{z\varepsilon}=0$ and also for zero restrictions on $\pi_{w\varepsilon}$. In the following we will not always state explicitly the various (usually trivial) requirements with respect to the rank of (partitions of) Π_{zw} and Π_{wx}.

The test of H_0: $R\beta=r$, where R is a HxK matrix of rank H, is easily derived according to the Wald principle, see for instance Bowden and Turkington (1984). The statistic

$$S(R\beta-r) = (R\hat{\beta}-r)'[R(X'P_zX)^{-1}R']^{-1}(R\hat{\beta}-r)/\hat{\sigma}^2 \tag{3.5}$$

tends under H_0 in distribution to a χ^2 variate with H degrees of freedom. Here

$$\hat{\beta} = (X'P_zX)^{-1}X'P_z\,y \tag{3.6}$$

with

$$P_z = Z(Z'Z)^{-1}Z \quad \text{and} \quad K \leq L < T$$

is the IV estimator of β and $\hat{\sigma}^2$ is a consistent estimator of σ^2, for instance

$$T^{-1}(y-X\hat{\beta})'(y-X\hat{\beta}) = T^{-1}.RSS(X|Z) \;.$$

RSS(X|Z) indicates the sum of squares of the residuals obtained from the regressors X and the IV estimator based on the instrumental variable set Z.

The test statistic (3.5) can also be expressed as

$$S(R\beta-r) = T.[RRSS(\hat{X}|Z) - RSS(\hat{X}|Z)]/RSS(X|Z) \;, \tag{3.7}$$

where RSS($\hat{X}$|Z) indicates the residual sum of squares obtained when OLS is applied in the regression of y on $\hat{X}=P_zX$. We may call this the second stage RSS. This regression produces the IV estimator $\hat{\beta}$ which enables the calculation of RSS(X|Z). By RRSS($\hat{X}$|Z) we indicate the RSS that results from regressing y on $\hat{X}$ under the restrictions $R\beta=r$.

Statistic (3.7) and asymptotic equivalents thereof enable to test coefficient restrictions in model (3.1) under the assumptions (3.2), (3.3) and (3.4). The same test procedure can also be employed to test the significance of regressors added to specification (3.1). Pagan (1984) presents some (mis)-specification tests in this form such as for functional form and serial correlation. We consider here the possibility to generate tests for exogeneity and instrumental variable admissibility along these lines for the single simultaneous equation model given above.

3.1. Tests for the status of excluded regressors

We first consider the problem of assessing the status of the variables W_0 by a test procedure. An asymptotically more efficient estimate of β than $\hat{\beta}$ of (3.6) can be found if the variables in W_0 can be used as extra instruments. For the present purpose we examine the test of $H_0:\delta_0=0$ in the augmented model

$$y = X\beta + W_0\delta_0 + \tilde{\varepsilon} \tag{3.8}$$

where now $\tilde{\varepsilon}=\varepsilon-W_0\delta_0$ and where we use $[Z:W_0]$ as instrumental variables. According to formula (3.7) the test statistic, which we indicate by $S(\delta_0)$, equals

$$S(\delta_0) = [RSS(\hat{X}|Z:W_0) - RSS(\hat{X}:W_0|Z:W_0)]/\hat{\sigma}^2 \ . \tag{3.9}$$

In these second-stage residual sums of squares the degree of overidentification is $L+M_0-K$ and $L-K$ respectively. This Wald test is derived from the IV estimator of δ_0 in (3.8) which is

$$\hat{\delta}_0 = (W_0'M_{\hat{x}}W_0)^{-1}W_0'M_{\hat{x}}\, y \tag{3.10}$$

where

$$\hat{X} = P_{z:w_0}X$$

and P_A denotes the projection matrix $A(A'A)^{-1}A'$ for full column matrices A. As

$$W_0'M_{\hat{x}}X = W_0'[X-\hat{X}] = W_0'[I-P_{z:w_0}]X = 0$$

substitution of $y=X\beta+\varepsilon$ in (3.10) leads to

$$\hat{\delta}_0 = (W_0'M_{\hat{x}}W_0)^{-1}W_0'M_{\hat{x}}\varepsilon \quad , \tag{3.11}$$

and it is found that $S(\delta_0) \xrightarrow{L} \chi^2(M_0)$ and that $\text{plim}\ \hat{\delta}_0=0$ if we have (3.4) and

$$T^{-\frac{1}{2}}W_0'\varepsilon \xrightarrow{L} N(0,\sigma^2\Pi_{w_0w_0}) \quad , \tag{3.12}$$

which implies $\pi_{w_0\varepsilon} = \text{plim}\ T^{-1}W_0'\varepsilon=0$. Hence we find that the combination of $\pi_{z\varepsilon}=0$ and $\pi_{w_0\varepsilon}=0$ belongs to the implied null-hypothesis of test statistic $S(\delta_0)$. Note however that $S(\delta_0)$ may assume significant values not only due to $\pi_{w_0\varepsilon}\neq 0$, but also in case $\pi_{z\varepsilon}\neq 0$. Hence, if one lacks firm information on $\pi_{z\varepsilon}$, the test is better interpreted as a general misspecification test with a non-specific alternative hypothesis.

The test under consideration here also appears in the literature as a test for overidentifying restrictions. Usually it is presented then for models like (3.1) for the case where $M_0=0=M_1$ and then L-K elements of Z_0 are added to the specification and their significance is tested. Let $Z_0=[Z_{00}:Z_{01}]$ with Z_{01} a Tx(L-K) matrix, then the statistic may be written

$$S(\gamma_{01}) = [RSS(\hat{X}|Z) - RSS(\hat{X}:Z_{01}|Z)]/\hat{\sigma}^2 \ . \tag{3.13}$$

As now the extended regression is just identified we have for the second-stage residuals $RSS(\hat{X}:Z_{01}|Z)=RSS(Z|Z)$. Hence if Z contains all the predetermined variables of a complete simultaneous system then $RSS(\hat{X}:Z_{01}|Z)$ corresponds with the RSS of the reduced form equation for y. Therefore we can rewrite the numerator of (3.13) as follows

$$RSS(\hat{X}|Z) - RSS(\hat{X}:Z_{01}|Z) =$$

$$(y-\hat{X}\hat{\beta})'(y-\hat{X}\hat{\beta}) - y'M_Z y = (y-\hat{X}\hat{\beta})'(y-\hat{X}\hat{\beta}) - (y-\hat{X}\hat{\beta})M_Z(y-\hat{X}\hat{\beta})$$

$$= (y-\hat{X}\hat{\beta})P_Z(y-\hat{X}\hat{\beta}) = (y-X\hat{\beta})P_Z(y-X\hat{\beta}) \ .$$

This is a quadratic form in the IV residuals and it appears that the actual partition of the matrix Z_0 is irrelevant. The statistic (3.13), written as

$$S(\gamma_{01}) = T.(y-X\hat{\beta})'P_Z(y-X\hat{\beta}) \ / \ (y-X\hat{\beta})'(y-X\hat{\beta}), \tag{3.14}$$

is presented without much of a formal derivation in Sargan (1964, p29).

In Basmann (1960, p651) and in Harvey (1981, p338), a comparable test is presented in disguise. Basmann suggests to take as the numerator of the test statistic the minimum over α of

$$(y-Y_1\alpha)'[M_{Z_1}-M_Z](y-Y_1\alpha) \ . \tag{3.15}$$

Taking the derivative with respect to α and equating this to zero leads to the minimand

$$\alpha^* = (Y_1'[M_{Z_1}-M_Z]Y_1)^{-1}Y_1'[M_{Z_1}-M_Z]y$$

and as $M_{Z_1}-M_Z=P_Z-P_{Z_1}=P_Z M_{Z_1} P_Z$ we find

$$\alpha^* = (\hat{Y}_1' M_{Z_1}\hat{Y})^{-1}\hat{Y}_1' M_{Z_1} y = \hat{\alpha} \ ,$$

which equals the IV estimator of α in (3.1) for M=0. Hence the numerator of the test is, as in Harvey's statistic,

$$(y-Y_1\hat{\alpha})'[M_{Z_1}-M_Z](y-Y_1\hat{\alpha}) =$$

$$(y-Y_1\hat{\alpha}-Z_1\hat{\gamma})'[P_Z-P_{Z_1}](y-Y_1\hat{\alpha}-Z_1\hat{\gamma}) =$$

$$(y-X\hat{\beta})P_Z(y-X\hat{\beta}) \quad ,$$

which equals Sargan's, and where we made use of $M_{Z_1}Z_1\hat{\gamma}=0$, $M_Z Z_1\hat{\gamma}=0$ and of $\hat{X}'(y-X\hat{\beta})=\hat{X}'y-\hat{X}'X(\hat{X}'X)^{-1}\hat{X}'y=0$ which includes $Z_1'(y-X\hat{\beta}) = 0$ and gives $P_{Z_1}(y-X\hat{\beta}) = 0$. Basmann and Harvey take as denominator the estimator of σ^2

$$(y-Y_1\hat{\alpha})M_Z(y-Y_1\hat{\alpha})/T = (y-X\hat{\beta})M_Z(y-X\hat{\beta})/T \ .$$

If Z contains valid instruments this estimator is consistent and therefore the tests of Sargan and Basmann are asymptotically equivalent. Basmann also presents a straightforward modification of the test in an F statistic, viz.

$$\frac{T-K}{L-K} \; \frac{(y-X\hat{\beta})'P_Z(y-X\hat{\beta})}{(y-X\hat{\beta})'M_Z(y-X\hat{\beta})} \tag{3.16}$$

which, as we proved, is a simpler expression for the test given in Harvey.

The tests on exogeneity of excluded regressors $S(\delta_0)$ and $S(\gamma_{01})$ both follow from the rather heuristic variable addition approach. In Smith (1983) comparable tests are derived according to the Wald, Lagrange Multiplier and likelihood ratio principles. He supposes the disturbance in the single structural equation to be normally distributed and Z to be the full set of explanatory variables in the reduced form of the system which also has normally distributed disturbances. Then it appears that according to the LM principle again the statistic $S(\gamma_{01})$ of (3.14) is obtained. Therefore we may hope that the test has favourable characteristics also in the less regular cases considered here.

3.2. A test for the status of included regressors

We now consider the case where the exogeneity of the variables in W_1 is to be assessed. This problem is also addressed in Smith (1983) and again for the more specific case, where ε is normal and the reduced form of Y_1 has normal disturbances as well, he derives a statistic according to the LM principle. We will denote this test statistic as

$$S(\delta_1) = [RSS(\hat{X}|Z:W_1) - RSS(\hat{X}:M_Z W_1|Z:W_1)]/\hat{\sigma}^2 . \tag{3.17}$$

It corresponds to the test statistic that is obtained by using the instrumental variable set $[Z:W_1]$ to test $H_0:\delta_1=0$ in the augmented model.

$$y = X\beta + M_Z W_1 \delta_1 + \tilde{\varepsilon} . \tag{3.18}$$

Note that in (3.17) we have $\hat{X}=P_{Z:W_1}X$ and that we made use of $P_{Z:W_1}M_Z W_1 = W_1 - P_Z W_1 = M_Z W_1$. The statistic $S(\delta_1)$ can also be applied in our less regular circumstances; it is simply based on the IV estimator of δ_1 which equals

$$\hat{\delta}_1 = [W_1'M_Z M_{\hat{X}} M_Z W_1]^{-1} W_1'M_Z M_{\hat{X}} y \tag{3.19}$$

and hence its implied null-hypothesis can be investigated by analyzing the distribution of $T^{-\frac{1}{2}}W_1'M_Z M_{\hat{X}}y$. Because we may write $W_1'M_Z = W_1'M_Z P_{Z:W_1}$ we find

$$\begin{aligned} W_1'M_Z M_{\hat{X}}(X\beta+\varepsilon) &= W_1'M_Z P_{Z:W_1}[I-\hat{X}(\hat{X}'\hat{X})^{-1}\hat{X}'](X\beta+\varepsilon) \\ &= W_1'M_Z M_{\hat{X}}\varepsilon = -W_1'P_Z M_{\hat{X}}\varepsilon. \end{aligned} \tag{3.20}$$

Here the latter equality follows from $\hat{X}=P_{Z:W_1}X=[\hat{Y}_1:Z_1:W_1]$ which gives $W_1'M_{\hat{X}}=0$. As it also gives $Z_1'M_{\hat{X}}=0$ we can reduce (3.20) further by partitioning P_Z and we then obtain

$$T^{-\frac{1}{2}}W_1'M_Z M_{\hat{X}}y = -T^{-1}W_1'M_{Z_1}Z_0(T^{-1}Z_0'M_{Z_1}Z_0)^{-1}T^{-\frac{1}{2}}Z_0'M_{\hat{X}}\varepsilon . \tag{3.21}$$

It follows from (3.2) that plim $T^{-1}Z_0'M_{Z_1}Z_0$ is non-singular and also that plim $T^{-1}W_1'M_{Z_1}Z_0$ has finite elements. That the rank of the latter matrix expression will equal M_1 corresponds with the feasibility requirement of regression (3.18): estimator (3.19) only exists if $[X:M_Z W_1]$ or $[Y_1:Z_1:W_1:P_Z W_1]$ has full column rank. This implies that $Z_0'M_{Z_1}W_1$ has to have full column rank. It is obvious now from (3.21) that $S(\delta_1)$ is distributed asymptotically as a χ^2 variate with M_1 degrees of freedom if Z_0 and $\hat{X}=P_{Z:W_1}X=[P_{Z:W_1}Y_1:Z_1:W_1]$ are independent of ε. Hence $\pi_{Z\varepsilon}=0$ and $\pi_{W\varepsilon}=0$

are part of the implied null-hypothesis. If ε and W_1 are dependent we will generally have $S(\delta_1) \xrightarrow{L} \infty$ and $\text{plim}\ \hat{\delta}_1 \neq 0$. The test $S(\delta_1)$ is a general form of what is usually called the Wu-Hausman procedure for testing dependence of regressors and disturbances.

3.3. Another variable addition test for IV validation

In our initial design in (3.8) and (3.9) of the variable addition test in the context of IV estimation we tested the significance of the excluded regressors W_0 and, while using $[Z:W_0]$ as the set of instruments, a test for IV validity of W_0 emerged. In the test for exogeneity of W_1 we added the regressors $M_Z W_1$ to the model, and these variables were also added to the set of instrumental variables, as $[Z:M_Z W_1]$ and $[Z:W_1]$ span the same subspace. Therefore these two tests are particular examples of a straightforward extension of the $S(\tilde{\beta})$ test in the OLS regression of (2.3). Now we have

$$y = X\beta + \tilde{X}\tilde{\beta} + \tilde{\varepsilon} \tag{3.22}$$

and the statistic

$$S(\tilde{\beta}) = [RSS(\hat{X}|Z:\tilde{X}) - RSS(\hat{X}:\hat{\tilde{X}}\,|\,Z:\tilde{X})]/\hat{\sigma}^2 \tag{3.23}$$

tests the significance of $\tilde{\beta}$ using the instrumental variable set $[Z:\tilde{X}]$. In (2.3) we have the special case where $Z \equiv X$. Apart from the tests $S(\delta_0)$ and $S(\delta_1)$ the test for predictive failure given in Kiviet (1985) is another example of such a variable addition test.

Next to this we could use the variable addition principle without adding the extra variables to the set of instruments. Of course the number of extra regressors should not exceed L-K then. We distinguish this approach from the $S(\tilde{\beta})$ test by writing the extended model as

$$y = X\beta + \dot{X}\dot{\beta} + \dot{\varepsilon} \tag{3.24}$$

with $\dot{\varepsilon} = \varepsilon - \dot{X}\dot{\beta}$ and $\dot{X}$ a $T \times \dot{K}$ matrix of extra regressors with $\dot{K} \leq L-K$. The test statistic is

$$S(\dot{\beta}) = [RSS(\hat{X}|Z) - RSS(\hat{X}:\hat{\dot{X}}|Z)]/\hat{\sigma}^2 \tag{3.25}$$

and it is based on the IV estimator

$$\hat{\dot{\beta}} = (\hat{\dot{X}}'M_{\hat{X}}\hat{\dot{X}})^{-1}\hat{\dot{X}}'M_{\hat{X}}y \ , \tag{3.26}$$

where $\hat{X}=P_Z X$ and $\hat{\dot{X}}=P_Z\dot{X}$. As $\hat{\dot{X}}'M_{\hat{X}}X=\dot{X}'P_Z[X-\hat{X}]=0$ it follows from substitution of $y=X\beta+\varepsilon$ that the requirement for plim $\hat{\dot{\beta}}=0$ is plim $T^{-1}\hat{\dot{X}}'M_{\hat{X}}\varepsilon=0$. This obtains if either $\pi_{z\varepsilon}=0$ or plim $T^{-1}Z'\dot{X}=0$. However, the latter case may be excluded because then IV estimation is not feasible. Hence the test $S(\dot{\beta})$ of (3.23) allows (asymptotically) the following inference. If $S(\dot{\beta})$ is significant this indicates that Z is an inappropriate set of instruments for the model $y=X\beta+\varepsilon$ and hence $\hat{\beta}$ is inconsistent, while an insignificant value corroborates the adequacy of the instruments Z in this specification. However, a significant value gives no clue whether it is the specification of the explanatory part and hence of the disturbance term ε, or whether it is the matrix Z that has to be adapted (or both). Hence, test $S(\dot{\beta})$ provides no information on the status of $\dot{X}$ at all, but it uses the added regressors only to provide a diagnostic on the validity of $\pi_{z\varepsilon}=0$ and thus on the consistency of $\hat{\beta}$. From Kiviet (1985) it follows that particular tests for serial correlation in single simultaneous equations are particular examples of the $S(\dot{\beta})$ test.

4. CONSIDERATIONS ON A GENERAL TESTING STRATEGY FOR IV EQUATIONS

In the foregoing section we presented four types of tests for checks on the adequacy of the specification of the structural form and the composition of a set of instrumental variables for a single linear regression equation. All four tests have as a part of their implied null hypothesis $\pi_{z\varepsilon}=0$. For test $S(\dot{\beta})$ of (3.26) this is the main element of the null hypothesis. Test $S(\delta_1)$ of (3.19) tests also for $\pi_{w_1\varepsilon}=0$; test $S(\delta_0)$ of (3.9) tests also for $\pi_{w_0\varepsilon}=0$ [the test $S(\gamma_{01})$ of (3.14) is a special form of test $S(\delta_0)$]; and finally we have test $S(R\beta-r)$ of (3.7) which tests the validity of the restrictions on the coefficient vector $R\beta=r$.

From the extended regressions that correspond to the exogeneity tests for excluded regressors $S(\delta_0)$ and included regressors $S(\delta_1)$ we obtain a test for the exogeneity of both. This is based on the extended regression

$$y = X\beta + W_0\delta_0 + M_{z:w_0}W_1\delta_1 + \tilde{\varepsilon}\ , \tag{4.1}$$

where the joint significance of δ_0 and δ_1 is tested by using [Z:W] as instruments. In practice, however, a sequential test for the exogeneity of single variables is much more helpful. Such a procedure avoids the problem that remains when a joint test leads to a rejection, because then it is still possible that some of the tested variables are acceptable as instrument individually.

For a sequential test procedure, see Mizon (1977) and Kiviet and Phillips (1985), in which the same set of data is used repeatedly to test various hypotheses, it is desirable to have a unique ordering of successive hypotheses, and also that this ordering leads to test statistics that are independent under the most restrictive null hypothesis. We shall investigate these phenomena here for a sequential testing strategy in which the comprehensive problem is tackled of composing the explanatory part of a linear regression equation and assessing the validity of the instruments. For the moment we make the simplifying (but unrealistic) assumption that the modelbuilder has already established correctly that the structural relationship for the dependent variable y must contain the regressors Y_1, Z_1 and W_1, and further that the matrix Z contains $L \geq K$ admissible instruments. For this particular situation we now try to devise a sequential test procedure to verify the legitimacy of the instruments Z, to establish the (non-)exogeneity of all the regressors in W (one by one) and to validate particular restrictions on the coefficients β.

4.1. The interdependence of the test statistics

To begin with we consider pairs of statistics from the four types of tests of section 3. We start by investigating the sequential procedure where first the exogeneity of W_0 and next the exogeneity of W_1 is tested, and where in the second test W_0 is added to the set of instrumental variables Z if the first test statistic assumes an insignificant value. Then, according to (3.11), the asymptotic distribution of the first test statistic depends on $T^{-\frac{1}{2}}W_0'M_{\hat{X}}\varepsilon$, with $\hat{X}=P_{Z:W_0}X$, and according to (3.20) the distribution of the second test statistic depends on $T^{-\frac{1}{2}}W_1'P_{Z:W_0}M_{\hat{X}}\varepsilon$, with here $\hat{X}=P_{Z:W}X$. Each of these two stochastic vectors is under the overall null hypothesis of the exogeneity of W asymptotically normally distributed with mean zero. Whether the two test statistics are asymptotically independent depends on the value of

$$W_0'[I-P_{Z:W_0}X(X'P_{Z:W_0}X)^{-1}X'P_{Z:W_0}][I-P_{Z:W}X(X'P_{Z:W}X)^{-1}X'P_{Z:W}]P_{Z:W_0}W_1 \tag{4.2}$$

If this expression equals zero the two vectors are uncorrelated asymptotically, and hence, the two test statistics are asymptotically independent; in that case the asymptotic overall type I error of the procedure in which these tests are employed successively can be controlled. Making use of $P_{Z:W_0}P_{Z:W}X= P_{Z:W}P_{Z:W_0}X=P_{Z:W_0}X$ expression (4.2) can be reduced to

$$W_0'[I-P_{Z:W_0}X(X'P_{Z:W_0}X)^{-1}X'P_{Z:W_0}]P_{Z:W_0}W_1$$

and as $P_{Z:W_0}W_1$ constitutes one of the partitions of $P_{Z:W_0}X$ the expression equals zero indeed.

Notice that independence is not obtained if in the second stage W_0 is not added to the set of instruments. One can also verify that reversing the order of the two tests, and hence testing the exogeneity of the included regressors W_1 first, and next the exogeneity of the excluded regressors W_0, will lead to dependent test statistics under exogeneity of W, irrespective of the use of W_1 as instruments in the second stage. Upon addition of W_1 to the set of instruments in the second stage the relevant expression is

$$W_0'[I-P_{Z:W_1}X(X'P_{Z:W_1}X)^{-1}X'P_{Z:W_1}]P_ZW_1 \ .$$

This generally differs from zero since neither P_ZW_1 nor W_0 should be completely spanned by the columns of $[P_{Z:W_1}Y_1:Z_1:W_1]$; the case $Z_0'W_1=0$ has to be excluded, because then the first test cannot be performed due to extreme multicollinearity in the extended regression.

We now examine whether partitions of the matrix W_0 can be tested sequentially in such a way that the statistics are independent under exogeneity of W_0. If $W_0=[W_{00}:W_{01}]$, the expression to be investigated with respect to the asymptotic independence of the successive tests for the exogeneity of two mutually exclusive sets of excluded regressors is

$$W_{00}'[I-P_{Z:W_{00}}X(X'P_{Z:W_{00}}X)^{-1}X'P_{Z:W_{00}}] \times$$

$$[I-P_{Z:W_0}X(X'P_{Z:W_0}X)^{-1}X'P_{Z:W_0}]W_{01}$$

$$= W_{00}'[I-P_{Z:W_{00}}X(X'P_{Z:W_{00}}X)^{-1}X'P_{Z:W_{00}}]W_{01}$$

and this again differs from zero in general. When we happen to have $W_{01}'[Z:W_{00}]=0$ independence is obtained.

If we test sequentially partitions of the included regressors $W_1=[W_{10}:W_{11}]$ and add after testing W_{10} this matrix to the set of instruments in the second stage, then the relevant expression is

$$W_{10}'P_Z[I-P_{Z:W_{10}}X(X'P_{Z:W_{10}}X)^{-1}X'P_{Z:W_{10}}] \times$$

$$[I-P_{Z:W_1}X(X'P_{Z:W_1}X)^{-1}X'P_{Z:W_1}]P_{Z:W_{10}}W_{11}$$

$$= W_{10}'P_Z[I-P_{Z:W_{10}}X(X'P_{Z:W_{10}}X)^{-1}X'P_{Z:W_{10}}]P_{Z:W_{10}}W_{11} = 0 \ ,$$

as $P_{Z:W_{10}}W_{11}$ just contains some of the columns of $P_{Z:W_{10}}X$. Hence, independence is acquired here irrespective of orthogonality of the partitions of W_1.

The results obtained so far indicate that in a sequence of tests on the exogeneity of the variables W independence can only be obtained by first employing the $S(\delta_0)$ test once, and then performing a series of $S(\delta_{1i})$ tests for $i=0,1,\ldots$, on appropriate partitions or individual columns of W_1, under sequential addition of the not rejected exogeneity hypotheses.

We now examine how the $S(\delta_0)$ and $S(\delta_1)$ tests relate to the test $S(\dot{\beta})$. This has a less restricted null hypothesis than the tests on the status of W (when it is used with instrumental variable set Z). If test $S(\dot{\beta})$ is employed first, followed by test $S(\delta_0)$, the relevant expression concerning (in)dependence reduces to

$$\dot{X}P_Z[I-P_ZX(X'P_ZX)^{-1}X'P_Z]W_0 \ ,$$

and if the order of the tests is reversed it is

$$W_0'[I-P_{z:w_0}X(X'P_{z:w_0}X)^{-1}X'P_{z:w_0}]P_{z:w_0}\dot{X} \ .$$

Neither of these equals zero in general. If test $S(\dot{\beta})$ is employed first, and then the test $S(\delta_1)$ on exogeneity of the included regressors W_1, the relevant expression reduces to

$$\dot{X}P_z[I-P_zX(X'P_zX)^{-1}X'P_z]P_zW_1 \ ,$$

and this does equal zero as P_zW_1 is a partition of P_zX. If the order of the tests is reversed the expression is

$$W_1'P_z[I-P_{z:w_1}X(X'P_{z:w_1}X)^{-1}X'P_{z:w_1}]P_{z:w_1}\dot{X}$$

and in general this will differ from zero.

We will not consider explicitly the effects on the interdependence of these tests if $\dot{X}$ is included in the regression when the tests $S(\delta_0)$ or $S(\delta_1)$ are employed, and next the significance of $\dot{X}$ is tested. The results for the $S(\dot{\beta})$ test when used in this way will follow directly from our findings for the $S(R\beta-r)$ test, as this test and the $S(\dot{\beta})$ test are of exactly the same nature; the only difference is that model (3.1) constitutes the null hypothesis for test $S(\dot{\beta})$ and that it is the maintained hypothesis for test $S(R\beta-r)$.

From (3.5) it follows that under validity of $R\beta=r$ the distribution of $S(R\beta-r)$ depends on the distribution of $R(X'P_zX)^{-1}X'P_z\varepsilon$. Hence, employment of $S(\delta_0)$ first, followed by addition of W_0 to the set of instruments, and application of test $S(R\beta-r)$ next leads to independent test statistics since

$$W_0'[I-P_{z:w_0}X(X'P_{z:w_0}X)^{-1}X'P_{z:w_0}]P_{z:w_0}X = 0 \ .$$

If instead of $S(\delta_0)$ the test $S(\delta_1)$ is employed first, we obtain

$$W_1'P_z[I-P_{z:w_1}X(X'P_{z:w_1}X)^{-1}X'P_{z:w_1}]P_{z:w_1}X = 0$$

and, if test $S(\dot{\beta})$ precedes the $S(R\beta-r)$ test the relevant expression is

$$\dot{X}P_z[I-P_zX(X'P_zX)^{-1}X'P_z]P_zX = 0 \ .$$

The three test sequences above all constitute examples of what Kiviet and Phillips (1985) call the testing of 'superposed' alternatives: the null-hypothesis of the first test conforms with the maintained hypothesis of the second test. Here it is again shown that in such sequences the individual test statistics are asymptotically independent under validity of the most restrictive null-hypothesis.

We now consider the reverse order, in which test $S(R\beta-r)$ is employed first, followed by application of one of the other tests to the model in which the restrictions $R\beta=r$ are imposed. Such a sequence will in general lead to dependent test statistics as the alternative hypotheses of the two test statistics are non-nested, while both tests have the same null-hypothesis. Only in particular circumstances, such as for test $S(\dot{\beta})$ followed by test $S(\delta_1)$, the testing of these 'juxtaposed' alternatives will lead to independent test statistics.

The results obtained sofar may be summarized in the following scheme:

Second test / First test	$S(\dot{\beta})$	$S(\delta_{00})$	$S(\delta_{01})$	$S(\delta_{10})$	$S(\delta_{11})$	$S(R\beta-r)$
$S(\dot{\beta})$	-	d	d	i	i	i
$S(\delta_{00})$	d	-	d	i	i	i
$S(\delta_{01})$	d	d	-	i	i	i
$S(\delta_{10})$	d	d	d	-	i	i
$S(\delta_{11})$	d	d	d	i	-	i
$S(R\beta-r)$	d	d	d	d	d	-

Indicated is the asymptotic (in)dependence obtained in the sequential application of two tests out of the four types of tests considered here, where the first test is applied to model (3.1) and the second test is applied to model (3.1) after the imposition of the restrictions tested by the first statistic. The letter 'i' indicates asymptotic independence under the overall null-hypothesis and the letter 'd' denotes that the tests are dependent in general but can be independent in incidental situations which usually cannot be forced by the modelbuilder.

4.2. A strategic ordering of the test statistic

In a modelling exercise usually many tests are applied to the same set of data and if one wants to exert some control over the overall type I error of such an operation then the individual tests should be linked preferably in such a way that independence under the most restrictive null hypothesis is acquired. From the scheme in the foregoing subsection we see that we therefore should order the tests as follows. If one starts with one (or a sequence of superposed) $S(\dot{\beta})$ test(s) then one should avoid $S(\delta_0)$ tests altogether. Next, a (sequence of superposed) $S(\delta_1)$ test(s) may be applied, followed eventually by a (sequence of superposed) $S(R\beta-r)$ test(s). A $S(\delta_0)$ test should be applied only once, and, if applied at all, it should be used at the very outset of the test sequence; any $S(\dot{\beta})$ tests should be avoided then. This only $S(\delta_0)$ test should precede any sequence of $S(\delta_1)$ tests which again has to precede any sequence of $S(R\beta-r)$ tests. Note, that the order of testing meant above does not necessarily mean the actual order in which the computation of the tests takes place. This ordening of the tests merely corresponds with the degree of generallity of the model specifications to which the tests are applied.

The approach outlined above is only applicable if the initial number of instruments L in the matrix Z is sufficient. Whether the approach will turn out to be profitable depends amongst others on the admissibility of the instruments in Z, and hence on the adequacy of the initial model specification: in fact one should have available an overparameterization of the data generating process from the very outset. Further, some modified version of the various asymptotic tests should be used that behaves satisfactory in small samples with respect to the size, the power and the interdependence. Finally, the success of the complete exercise will be heavily dependent on the adequacy of the decisions taken by the modelbuilder in case hypotheses are rejected.

Appendix

That variable addition may contaminate consistent estimators is illustrated by the following simple example. Let the equation under study be

$$y_1 = \beta_1 x_1 + \beta_2 x_2 + \varepsilon_1 \tag{A.1}$$

and let a seemingly unrelated second equation be

$$y_2 = \beta_3 x_3 + \beta_4 x_1 + \varepsilon_2 \tag{A.2}$$

where

$$\varepsilon_1 = \rho\varepsilon_2 + \eta \text{ , with } \rho \neq 0 \text{ and } E\varepsilon_2\eta = 0 \text{ ,} \tag{A.3}$$

and let x_1, x_2 and x_3 be non-stochastic and ε_1 and ε_2 such that OLS estimation of the equation leads to (inefficient but) consistent estimators. Now using y_2 as an extra regressor in (A.1) to check the consistency of OLS leads to the equation

$$y_1 = \beta_1 x_1 + \beta_2 x_2 + \tilde{\beta} y_2 + \tilde{\varepsilon}_1 \text{ .} \tag{A.4}$$

It is simply seen that the test statistic $S(\tilde{\beta})$ will in general assume significant values asymptotically. From (A.1) through (A.3) it follows that we have

$$y_1 = (\beta_1 - \rho\beta_4) x_1 + \beta_2 x_2 + \rho y_2 - \rho\beta_3 x_3 + \eta \text{ ,} \tag{A.5}$$

from which the four parameters can be estimated consistently and more efficiently. In (A.5) the regressors are exogenous. In (A.4) the OLS estimates of β_1 and β_2 will be biased and inconsisyent due to the omission of x_3 in that regression, which leads to joint dependence of (some of) the regressors and the error term $\tilde{\varepsilon}_1$.

REFERENCES

Basmann, R.L. (1960), "On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics", JASA, 650-659.

Bowden, R.J. and D.A. Turkington (1984), "Instrumental Variables", Econometric Society Monographs in Quantative Economics, Cambridge University Press.

Davidson, R. and J.G. MacKinnon (1984), "The Interpretation of Test Statistics", Queen's University Institute of Economic Research, mimeo.

Harvey, A.C. (1981), "The Econometric Analysis of Time Series", Philip Allan.

Kiviet, J.F. (1985), "Model Selection Test Procedures in a Single Linear Equation of a Dynamic Simultaneous System and their Defects in Small Samples", forthcoming in Journal of Econometrics.

Kiviet, J.F. and G.D.A. Phillips (1985), "Testing Stategies for Model Selection", University of Amsterdam/University of Leeds, mimeo.

Mizon, G.E. (1977), "Inferential procedures in nonlinear models: an application in a UK industrial cross section study of factor substitution and returns to scale", Econometrica, 45, 1221-1242.

Pagan, A.R. (1984), "Model Evaluation by Variable Addition", in: Econometrics and Quantitative Economics", editors Hendry, D.F. and K.F. Wallis, Basil Blackwell, Oxford, 103-133.

Pagan, A.R. and A.D. Hall (1983), "Diagnostic tests as residual analysis", Econometric Reviews, 2, 159-218.

Sargan, J.D. (1964), "Wages and Prices in the UK: A Study in Econometric Methodology", in: Econometric Analysis for National Economic Planning, editors: Hart, Mills and Whitaker. London: Butterworths.

Smith, R.J. (1983), "Limited Information Classical Tests for the Independence of Stochastic Variables and Disturbances of a Single Lineair Stochastic Simultaneous Equation", Discussion Paper ES 142, University of Manchester.