working paper
department
of economics

A NEW SPECIFICATION TEST FOR THE VALIDITY OF INSTRUMENTAL VARIABLES

Jinyong Hahn
Jerry Hausman

No. 99-11 May, 1999

massachusetts
institute of
technology
50 memorial drive
cambridge, mass. 02139

WORKING PAPER
DEPARTMENT
OF ECONOMICS

A NEW SPECIFICATION TEST FOR THE VALIDITY OF INSTRUMENTAL VARIABLES

Jinyong Hahn
Jerry Hausman

No. 99-11 May, 1999

MASSACHUSETTS
INSTITUTE OF
TECHNOLOGY

50 MEMORIAL DRIVE
CAMBRIDGE, MASS. 02142

WP No. 99-11

Abstract for: A New Specification Test for the Validity of Instrumental Variables

By Hahn and Hausman

We develop a new specification test for the IV estimators adopting a particular second order approximation of Bekker (1994). The new specification test compares the difference of the forward (conventional) 2SLS estimator of the coefficient of the right hand side endogenous variable with the reserve 2SLS estimator of the same unknown parameter when the normalization is changed. Under the null hypothesis that conventional first order asymptotics provides a reliable guide, the two estimates should be very similar. Our test sees whether the resulting difference in the two estimates satisfies the results of second order asymptotic theory. Essentially the same idea is applied to develop another new specification test using second-order unbiased estimators of the type first proposed by Nagar (1959). If the forward and reverse Nagar-type estimators are not significantly different we recommend estimation by LIML, which we demonstrate is the optimal linear combination of the Nagar-type estimators (to second order). We also demonstrate the high degree of similarity for k-class estimators between the approach of Bekker (1994) and the Edgeworth expansion approach of Rothenberg (1983). Empirical example and Monte Carlo evidence are provided.

Preliminary—First draft
Please do not cite without authors' permission

A New Specification Test for the Validity of Instrumental Variables*

Jinyong Hahn
University of Pennsylvania

And

Jerry Hausman
MIT

May, 1999

Excellent research assistance has been provided by Derek Ho Cheung and Karen Hull. Whitney Newey provided some helpful suggestions. Please send email to jhausman@mit.edu.

1. Introduction

A significant understanding has emerged over the past few years that instrumental variable (IV) estimation of the simultaneous equation model can lead to problems of inference in the situation of "weak instruments," which can arise when the instruments do not have a high degree of explanatory power for the jointly endogenous variable(s) or when the number of instruments becomes large. The situation of limited information estimation of a single equation has been studied extensively in the presence of "weak instruments." These problems of inference in the weak instrument situation can arise when conventional (first order) asymptotic inference techniques are used. In particular, conventional first order asymptotics can lead to a lack of indication of a problem even though significant (large sample) bias is present because estimated standard errors are not very accurate.

A number of papers have recommended possible diagnostics for the presence of the problem, *e.g.* Shea (1997). The usual form of the recommended diagnostics is to examine the R^2 or the associated F statistic of the reduced form regression for the included endogenous variable(s). A more refined recommendation is to consider the partial R^2 (or its associated F statistic) after the predetermined variables have been partialled out of the equation being estimated. Another approach has been to consider the rank statistic originally put forward by Anderson and Rubin (1949). While both approaches yield valuable information, the R^2 approach lacks a distribution theory and the rank condition test, in some sense, does not answer the question at issue of how well conventional asymptotic theory does in forming statistics for inference.

In this paper, we take a new approach and use higher order asymptotic distribution theory to determine if the conventional first order IV asymptotics are reliable in a particular situation. We recommend a new specification test for the IV estimators, and we concentrate initially on the 2SLS estimator since it is by far the most commonly used estimator. Our new specification test takes the general approach as the specification test approach of Hausman (1978) and estimates the same parameter(s) in two different ways. In particular, we compare the difference of the forward (conventional) 2SLS estimator of the coefficient of the right hand side endogenous variable with the reverse 2SLS estimator of the same unknown parameter when the normalization is changed.

Under the null hypothesis that conventional first order asymptotics provides a reliable guide, the two estimates should be very similar. Indeed, they have unitary correlation according to first order asymptotic distribution theory. However, when second order asymptotic distribution theory is used, the two estimators will differ due to second order bias terms. Our test subtracts off these bias terms and then sees whether the resulting difference in the two estimates satisfies the results of second order asymptotic theory. If it does and the second order bias term is small, we do not reject the use of first order asymptotic theory. Furthermore, the second order asymptotic theory may provide a more reliable basis for inference. An added attraction of our approach is that it permits the econometrician to compare two estimates of a structural parameter, which will have a straightforward economic interpretation in many situations. Thus, the econometrician can use economic knowledge to determine if the two estimates are very different or are close together in terms of the economic problem under study.

If the new specification test rejects we then consider estimation of the equation by second-order unbiased estimators of the type first proposed by Nagar (1959). We again consider forward and reverse estimation by the Nagar-type estimators to determine if the estimates are significantly different according to the new specification test. If they are not significantly different we recommend estimation by LIML, which we demonstrate is the optimal linear combination of the Nagar-type estimators (to second order). If the second specification test rejects or the two Nagar-type estimators differ substantially based on economic considerations, we conclude that neither set of estimates, 2SLS or LIML, may provide reliable results for inference in the particular situation.

Our approach also provides some possible lessons about previous recommendations found in the literature. To second order, not only the R^2 of the reduced form affects the asymptotic bias of the 2SLS estimator, but also two other terms are important. These terms are the covariance between stochastic disturbances in the structural equation and in the reduced form equation(s), and in the number of instruments. These three factors interact in a nonlinear manner so it is unlikely that any single first order asymptotic theory based test statistic will suffice to indicate when 2SLS does not perform well. We point out similarities between this result and the well-known errors in variables model in econometrics to demonstrate how this outcome might be expected. We also develop conditions under which the true parameter will be lie in the interval of the forward and reverse estimates.

Lastly, we investigate the performance of Nagar-type second order bias corrected IV estimators. While these estimators and LIML can lead to improved performance, they may also not perform well in the weak instrument situation. Thus, we demonstrate that LIML need not be significantly better than 2SLS over a range of possible situations. In particular, inferences based on LIML may not do well in the "weak instruments" situation. While Rothenberg (1983) uses results of Pfanzagl and Wefelmeyer (1978, 1979) to demonstrate that, under certain conditions, LIML is second order efficient, our specification test should help determine when reliable inference can be based on the use of LIML. We also demonstrate the high degree of similarity for k-class estimators between the approach of Bekker (1994) and the Edgeworth expansion approach of Rothenberg (1983).

We analyze an empirical problem of a simultaneous equation specification of a demand equation. This type of model formed the original model consider by Haavelmo, who first demonstrated that least squares would lead to biased results. We consider the demand for railroad movements between different origin and destination pairs for a particular bulk commodity. The original specification has quantity as the left hand side variable and price along with other variables on the right hand side. As instruments we have short run marginal cost variables. We find that the 2SLS estimate of the demand elasticity is about 2 times larger than the least squares estimate. We then reverse the regression using price as the left hand side variable and quantity as the right hand side endogenous variable. The estimated elasticity increases, but the new specification test finds that the two estimates are close together enough so as not to reject the first order asymptotic results. We then include many more instruments by interacting the cost instruments with the indicator variables for each origin-destination pair. The estimated price elasticity decreases significantly in magnitude, back toward the least squares estimate. When we run the reverse 2SLS estimation, we find that the estimate is about 6 times higher than then forward estimate. Here our specification test easily rejects the use of the first order asymptotics. Also, LIML does not do well in this latter situation.

The previous literature on the presence of weak instruments begins with Nelson and Startz (1990 a and b) and Bound, Jaeger, and Baker (1995) who demonstrate the poor performance of IV estimators in the weak instruments situation. Analysis of conditions when the weak instruments problem may exist are given by Hall, Rudebusch, and Wilcox (1996), Shea (1997), and Staiger and Stock (1997). Improved inferential techniques are recommended by Startz, Nelson, and Zivot (1998), Wang and Zivot (1999), and Zivot, Startz, and Nelson (1999). All of these approaches are essentially first order asymptotic approximation approaches in terms of recognizing the weak instruments problem and offering alternative approaches to inference. The second order asymptotic approach to inference and to estimation that we use was initiated by Nagar (1959) and has been used by a number of researchers. We follow the particular second order approximation of Bekker (1994).

While many different conclusions can be drawn in the weak instruments situation, we tend to recommend that the IV estimates, or even the "improved" IV estimates not be used when the specification test rejects (unless the two estimates are close together). The reason for this conclusion is that the IV estimators typically have significant bias in these situations when the specification test rejects which recommends against their use. First order asymptotics assumes that no bias exists, but the second order approach can find significant bias depending on the underlying primitive conditions. When this bias is present as demonstrated by the specification test, we believe that use of the IV estimates may lead to misguided conclusions.

2. Model

We begin with the simplest model specification with one right hand side (RHS) jointly endogenous variable so that the left hand side variable (LHS) depends only on the single jointly endogenous RHS variable. In the class of models with only one RHS jointly endogenous variable, which is by far the most common specification used in econometrics, this model specification accounts for other RHS predetermined (or exogenous) variables, which have been "partialled out" of the specification. Thus, we do not lose any generality by not including predetermined variables in the initial specification. We demonstrate below how RHS predetermined variables may be included in the formulae and computations.

We will assume that

(2.1) $$y_1 = \beta y_2 + \varepsilon_1 = \beta z \pi_2 + v_1$$

(2.2) $$y_2 = z\pi_2 + v_2,$$

where $\dim(\pi_2) = K$. Thus, the matrix z is the matrix of all predetermined variables, and equation (2.2) is the reduced form equation for y_2 with coefficient vector π_2. We also assume homoscedastic normality:

(2.3) $$\begin{pmatrix} v_{1i} \\ v_{2i} \end{pmatrix} \sim N(0,\Omega) \sim N\left(0, \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{bmatrix}\right).$$

We use the following notation:

$$y \equiv \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad z \equiv \begin{pmatrix} z_1' \\ \vdots \\ z_n' \end{pmatrix}, \quad \sigma_\varepsilon^2 \equiv \mathrm{Var}(\varepsilon_{1i}), \quad \sigma_{\varepsilon v_2} \equiv \mathrm{Cov}(\varepsilon_{1i}, v_{2i}), \quad \sigma_{\varepsilon v_1} \equiv \mathrm{Cov}(\varepsilon_{1i}, v_{1i}).$$

The simultaneous equation problem, which causes least squares to be biased, arises when $\sigma_{\varepsilon v_2} \neq 0$. This situation is what specification tests of the type proposed by Hausman (1978) and others test.

3. Motivation

3.1 Errors in Variables

We first consider an analogy between the simultaneous equation model specification and the errors in variable (EIV) model specification. If y_2 in equation (2.1) were replaced with a mismeasured exogenous variable, under classical assumptions of uncorrelated measurement error, the least squares estimate of β would be biased (in magnitude) towards zero. Hausman, Newey and Powell (1995) call this result the "iron law" of econometrics – the magnitude of the coefficient estimates are less than expected.

While this result does not always occur when least squares is used on equation (2.1), since the direction of bias depends on the sign of ω_{12}, a common finding is that when 2SLS is used the coefficient estimate increases in magnitude. However, in finite samples under certain situations even when 2SLS is used on equation (2.1), bias remains because an estimate of π_2 from equation (2.2) is used, since the true parameters are unknown. We now demonstrate how this result occurs.

Suppose that $z\pi_2$ is measured without error. Then, OLS of y_1 on $z\pi_2$ would be unbiased. Instead, $z\pi_2$ must be estimated, *i.e.*, we have to rely on 2SLS. Let $\hat{\pi}_2$ denote the first stage OLS estimator. We have

$$(3.1) \quad b_{2SLS} - \beta = \frac{\sum_{i=1}^{n} \left(v_{1i} - \beta z_i' \cdot (\hat{\pi}_2 - \pi_2)\right) \cdot z_i' \pi_2}{\sum_{i=1}^{n} (z_i' \hat{\pi}_2)^2},$$

Observe that

$$\begin{aligned} E\left[\sum_{i=1}^{n} \left(v_{1i} - \beta z_i' \cdot (\hat{\pi}_2 - \pi_2)\right) \cdot z_i' \hat{\pi}_2\right] &= \sum_{i=1}^{n} E\left[v_{1i} \cdot z_i'(z'z)^{-1} z' v_2\right] - \beta \cdot E\left[(\hat{\pi}_2 - \pi_2)'(z'z)(\hat{\pi}_2 - \pi_2)\right] \\ &= \omega_{12} \sum_{i=1}^{n} z_i'(z'z)^{-1} z_i - \beta \omega_{22} \cdot K \\ &\equiv K \cdot \sigma_{\varepsilon v_2}. \end{aligned}$$

Also note that $\sum_{i=1}^{n} (z_i' \hat{\pi}_2)^2 = R_f^2 \cdot \sum_{i=1}^{n} y_{2i}^2$, where R_f^2 is the R^2 in the first stage regression to obtain $\hat{\pi}_2$. Therefore, we expect bias approximately equal to

$$(3.2) \quad \frac{K \cdot \sigma_{\varepsilon v_2}}{R_f^2} \frac{1}{\sum_{i=1}^{n} y_{2i}^2}.$$

We make some observations. Other things being equal,

- Bias is a monotonically increasing function of $\sigma_{\varepsilon v_2}$.
- Bias is a monotonically increasing function of K.
- Bias is a monotonically decreasing function of R_f^2.

Note that conventional asymptotics, which lets $n \to \infty$ keeping DGP fixed, ignores the influence of $\sigma_{\varepsilon v_2}$, K, R_f^2.

In terms of the analogy with the EIV model specification, note that the bias in the EIV model depends on the ratio of the variance of the observation error divided by the variance of the RHS variable, a result which occurs in equation (3.1) except that the covariance term $\sigma_{\varepsilon v_2}$ replaces the variance of the observation error. The bias in the EIV model specification also depends inversely on the R_f^2 of the EIV model specification, as we demonstrate below, a result we also find in equation (3.1). Thus, the finite sample bias in the simultaneous equation problem has similarities with the bias in the EIV model specification, with the major difference that equation (3.1) has a term in K, the number of instruments. No similar variable arises in the EIV model bias formula because no instruments are used when OLS estimation is undertaken.

3.2 Forward and Reverse Regressions

A well-known result in the EIV model specification is that the forward regression and the reverse regression, when the coefficient estimate is inverted, bound the true coefficient β, where by reverse regression we mean interchanging the RHS variable with the LHS variable in the regression specification. Perhaps a less well-known result is that the product of the forward and reverse estimates equal the R^2 of the regression model (which is the same for the forward or reverse regression). Thus, a high R^2 implies that the bounds for the true coefficient β are very tight, and vice versa.[1] We now explore a similar result in the context of the simultaneous equation model. Let

$$(3.3) \quad b_{OLS} \equiv \frac{\sum_i y_{2i} y_{1i}}{\sum_i y_{2i}^2}, \qquad \text{and} \qquad c_{OLS} \equiv \frac{\sum_i y_{1i} y_{2i}}{\sum_i y_{1i}^2}$$

denote the forward and reverse OLS estimates of the model (2.1). Note that

[1] Thus, the use of the "permanent income" consumption model specification cannot explain the finding of different estimates of the marginal and average propensity to consume, when estimates are done on aggregate time series data because the R^2 of the regression specification is extremely high.

(3.4) $$b_{OLS} \cdot c_{OLS} = R^2_{(y,x)},$$

where $R^2_{(y,x)}$ is the R^2 in the OLS regression of (2.1), a result that is the same as that of the EIV model specification although here the two estimates need not bound the true parameter β.

Now, let

(3.5) $$b_{2SLS} \equiv \frac{\sum_i \hat{y}_{2i}\hat{y}_{1i}}{\sum_i \hat{y}_{2i}^2}, \quad \text{and} \quad c_{2SLS} \equiv \frac{\sum_i \hat{y}_{1i}\hat{y}_{2i}}{\sum_i \hat{y}_{1i}^2}$$

denote forward and reverse 2SLS estimates, where $\hat{y}_{2i}$ and $\hat{y}_{1i}$ are the results of orthogonal projections onto the subspace spanned by z. They are based on moment restrictions

(3.6) $$E[z_i \cdot (y_{i1} - \beta \cdot y_{2i})] = 0, \quad \text{and} \quad E\left[z_i \cdot \left(y_{2i} - \frac{1}{\beta} y_{1i}\right)\right] = 0.$$

It can easily be shown that, under conventional (first order) asymptotics,

(3.7) $$\sqrt{n} \cdot \left(b_{2SLS} - \frac{1}{c_{2SLS}}\right) = o_p(1),$$

which implies that the forward and reverse estimates are perfectly correlated, *i.e.* the two estimates are exactly the same in a given sample up to first order asymptotics. Thus, using equation (3.2) amounts to the implicit assumption that

(3.8) $$R^2_{(\hat{y}_1,\hat{y}_2)} \approx 1,$$

asymptotically to first order. Empirically, the authors have observed that the forward and reverse 2SLS estimates can differ by large amounts numerically even with quite large samples, which by equation (3.2) implies that in these situations conventional first order asymptotics may not provide a particularly good guide to the actual sample situation in question. We use this observation and implication of equation (3.2) to provide an approach that attempts to determine when conventional first order asymptotics can be relied on, or when alternative approaches need to be employed.

4. Bekker's (1994) Asymptotics: Is It Sensible?

Since conventional first order asymptotics do not necessarily provide a reliable guide, we need to use a different approach to the asymptotics. We explore the approach of Bekker (1994) and see whether his approach to asymptotic expansion captures the main features of the bias in the estimators that concern us. We assume as in Bekker (1994) that

$$(4.1) \quad \frac{K}{n} \to \alpha \qquad \text{and} \qquad \frac{1}{n}\pi_2' z' z \pi_2 = \Theta .$$

Below, we examine whether his asymptotics captures our motivation.

4.1 Errors-in-Variables Motivation

It can be shown that[2]

$$(4.2) \qquad \operatorname{plim} b_{2SLS} = \beta + \alpha \frac{\omega_{12} - \beta\omega_{22}}{\Theta + \alpha\omega_{22}} .$$

It can also be shown that[3]

$$(4.3) \quad \operatorname{plim} \frac{1}{n}\sum_{i=1}^{n} y_{2i}^2 = \Theta + \omega_{22} \qquad \text{and} \qquad \operatorname{plim} \frac{1}{n}\sum_{i=1}^{n} \hat{y}_{2i}^2 = \Theta + \alpha\omega_{22} .$$

Using the fact that $\sigma_{\varepsilon v} = \omega_{12} - \beta\omega_{22}$, we may rewrite equation (4.2) as

$$(4.4) \qquad \operatorname{plim} b_{2SLS} = \beta + \frac{\alpha}{\operatorname{plim} n^{-1} \sum_{i=1}^{n} y_{2i}^2} \frac{\sigma_{\varepsilon v_2}}{\operatorname{plim} R_f^2} ,$$

---

[2] See Bekker (1994).
[3] See Bekker (1994).

which coincides with equation (3.1), and again points out the close similarity with the EIV model specification result with the addition of the parameter α.

By a similar argument, we can show that

(4.5) $$c_{2SLS} = \frac{1}{\beta} + \alpha \frac{\omega_{12} - \frac{1}{\beta}\omega_{11}}{\beta^2\Theta + \alpha\omega_{11}} + o_p(1),$$

and therefore,

(4.6) $$\frac{1}{c_{2SLS}} = \beta + \alpha \frac{\omega_{11} - \beta\omega_{12}}{\beta\Theta + \alpha\omega_{12}} + o_p(1).$$

Analogous to equation (4.4), we may also write

(4.7) $$\operatorname{plim} c_{2SLS} = \frac{1}{\beta} + \frac{\alpha}{\operatorname{plim} n^{-1}\sum_{i=1}^{n} y_{1i}^2} \frac{\sigma_{-\frac{1}{\beta}\varepsilon, v_1}}{\operatorname{plim} R_r^2}.$$

Here, R_r^2 is from the first stage regression of equation (2.1). Thus, we see from equations (4.4) and (4.7) that if the (asymptotic) R^2's of the reduced form equations were one, conventional first order asymptotics yields the correct results. However, in actual situations when this result does not hold the asymptotic approximation to the bias depends on the covariance term which creates the simultaneous equation problem in the first place, the size of the R^2 in the reduced form equation and the parameter α which approximates the dimension of the subspace spanned by the predetermined variables relative to the sample size. Thus, under this approach sample size alone does not indicate how well conventional first order asymptotics do, but instead the dimension of the subspace spanned by z must also be considered.

4.2 R^2 Motivation

By using Lemma 5 in Appendix, we can show that

$$(4.8)\quad \operatorname{plim} R^2_{(\hat{y}_1,\hat{y}_2)} = \frac{(\beta\Theta+\alpha\omega_{12})^2}{(\beta^2\Theta+\alpha\omega_{11})(\Theta+\alpha\omega_{22})} = 1-\frac{\alpha\sigma_\varepsilon^2+\alpha^2\det(\Omega)}{(\beta^2\Theta+\alpha\omega_{11})(\Theta+\alpha\omega_{22})},$$

which avoids the implicit and false R^2 assumption of equation (3.3). Thus, the use of the Bekker asymptotic approach does yield an implication consistent with the empirical result we attempt to capture and that affects the bias in the 2SLS estimates.

5. Biases of Forward and Reverse 2SLS

Because $R^2_{(\hat{y}_1,\hat{y}_2)} = b_{2SLS}\cdot c_{2SLS}$ by definition, equation (4.8) implies that

$$(5.1)\qquad 0\le \operatorname{plim} b_{2SLS}\cdot \operatorname{plim} c_{2SLS}\le 1,$$

which in turn implies that

$$(5.2)\qquad 0\le \operatorname{plim}\frac{b_{2SLS}}{\beta}\cdot \operatorname{plim}\frac{c_{2SLS}}{1/\beta}\le 1.$$

Inequality (5.2) suggests that the forward and reverse 2SLS estimates may bound the true parameter β as in the EIV case.

In order to understand the inequality from a different perspective, rewrite equations (4.4) and (4.8) as

$$(5.3)\qquad \frac{\operatorname{plim} c_{2SLS}-\dfrac{1}{\beta}}{\operatorname{plim} b_{2SLS}-\beta} = -\frac{1}{\beta}\frac{\sigma_{\varepsilon v_1}}{\sigma_{\varepsilon v_2}}\frac{\operatorname{plim} R_f^2}{\operatorname{plim} R_r^2}\frac{\operatorname{plim} n^{-1}\sum_{i=1}^n y_{2i}^2}{\operatorname{plim} n^{-1}\sum_{i=1}^n y_{1i}^2}$$

$$= -\frac{1}{(1-\sigma_{\varepsilon\varepsilon}/\sigma_{\varepsilon v_1})}\frac{\operatorname{plim} R_f^2}{\operatorname{plim} R_r^2}\frac{\operatorname{plim} n^{-1}\sum_{i=1}^n y_{2i}^2}{\operatorname{plim} n^{-1}\sum_{i=1}^n y_{1i}^2},$$

where the second equality is based on

$$\sigma_{\varepsilon v_2} = \frac{1}{\beta}\sigma_{\varepsilon v_1} - \frac{1}{\beta}\sigma_{\varepsilon\varepsilon}.$$

Assume that $\sigma_{\varepsilon\varepsilon}/\sigma_{\varepsilon v_1}\ge 1$. We would then have

(5.4) $\beta \in \left(\operatorname{plim} b_{2SLS}, \operatorname{plim} 1/c_{2SLS}\right)$ or $\beta \in \left(\operatorname{plim} 1/c_{2SLS}, \operatorname{plim} b_{2SLS}\right)$.

We can see that the ratio

$$(5.5) \quad \frac{\operatorname{plim} R_f^2}{\operatorname{plim} R_r^2} \frac{\operatorname{plim} n^{-1} \sum_{i=1}^{n} y_{2i}^2}{\operatorname{plim} n^{-1} \sum_{i=1}^{n} y_{1i}^2}$$

determines the relative magnitude of the bias of the two estimators. We can see that the bias of c_{2SLS} is small relative to that of b_{2SLS} if $R_f^2 << R_r^2$.

6. A Specification Test based on Forward and Reverse 2SLS

We now turn to the main contribution of the paper. We attempt to provide an answer to the question: When can you trust the conventional first order asymptotics given the well documented problems of the first order asymptotic approximation in certain cases? As our derivations demonstrate above, the 2SLS bias depends on 3 factors: the covariance of the stochastic terms in equations (2.1) and (2.2), the R^2 of the reduced form equations, and the parameter α which depends on both K and n. Thus, no simple single statistic, *e.g.* the R^2 of the reduced form equation (or the associated F statistic), seems likely to be sufficient to answer the question of how well the conventional asymptotic approximation is doing in a particular situation.

Instead we turn to one of the basic ideas of the specification test approach of Hausman (1978) and estimate the same parameter, β, in two different ways. If the difference between the estimates is small, one will not reject the underlying assumption of the model specification. If the difference is large, one will come to the opposite conclusion. Here a possible approach is to use the forward and reverse 2SLS estimates and see how far apart they are. Thus, the specification test will be used in model specifications with overidentification, but this situation holds in most instances. An "economic sense" of the difference of the two estimates can be gained because in many cases the econometrician will know how big a change in the true coefficient β is important, since the parameter will have a marginal interpretation.

To do a statistical test, we need to determine the variance of the difference of the two estimates. Here first order asymptotics will not suffice, since because the forward and reverse coefficient estimates have unit correlation, the variance of the difference of the two estimates will be zero when a first order asymptotic approximation is used. Thus, we turn to second order asymptotic approximations, which were pioneered by Nagar (1959) and have been used since by Kadane (1973), Sargan (1976), Rothenberg (1983), and numerous other authors.

Note that the probability limit of the difference between the two possible estimators of β is equal to

$$(6.1) \quad B = -\alpha \frac{\Theta\sigma_\varepsilon^2 + \alpha\det(\Omega)}{(\Theta + \alpha\omega_{22})(\beta\Theta + \alpha\omega_{12})}.$$

Bekker (1994) shows that 2SLS is asymptotically normal. Therefore, $\sqrt{n}(b_{2SLS} - 1/c_{2SLS} - B)$ is also asymptotically normal. Because we do not know B in general, we would like to deal with an asymptotic result of the form

$$(6.2) \quad \sqrt{n}\left(b_{2SLS} - \frac{1}{c_{2SLS}} - \hat{B}\right) \to N(0,V)$$

for our specification test.[4]

$\hat{B}$ will be a consistent estimator of the difference of the biases. Let P_z and M_z denote the projection matrices onto the column space spanned by z and its orthogonal complement. It can be shown that

$$(6.3) \quad \Theta + \alpha\omega_{22} = \text{plim}\frac{1}{n}y_2'P_z y_2, \quad \text{and} \quad \beta\Theta + \alpha\omega_{12} = \text{plim}\frac{1}{n}y_2'P_z y_1.$$

Further, it can be shown by using Lemma 5 in Appendix that

$$(6.4) \quad \text{plim}\,\hat{\Xi} = \Theta\sigma_\varepsilon^2 + \alpha\det(\Omega),$$

where

$$\hat{\Xi} \equiv \left(\frac{1}{n}y_1'P_zy_1 - \frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_1'M_zy_1\right)\frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2'M_zy_2 - 2\left(\frac{1}{n}y_2'P_zy_1 - \frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2'M_zy_1\right)\frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2'M_zy_1$$
$$+\left(\frac{1}{n}y_2'P_zy_2 - \frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2'M_zy_2\right)\frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2'M_zy_2$$
$$+\hat{\alpha}\left(\frac{1}{1-\hat{\alpha}}\frac{1}{n}y_1'M_zy_1\right)\left(\frac{1}{1-\hat{\alpha}}\frac{1}{n}y_2'M_zy_2\right) - \hat{\alpha}\left(\frac{1}{1-\hat{\alpha}}\frac{1}{n}y_2'M_zy_1\right)^2$$

and $\hat{\alpha}$ is any consistent estimator for α. We may therefore use

(6.6) $$\hat{B} \equiv -\hat{\alpha}\frac{\hat{\Xi}}{\frac{1}{n}y_2P_zy_2 \cdot \frac{1}{n}y_2'P_zy_1}.$$

By the delta method based on Lemma 5 in Appendix, it can be shown that

Theorem 1. $\sqrt{n}\left(b_{2SLS} - \frac{1}{c_{2SLS}} - \hat{B}\right) \to N\left(0, \frac{2\alpha}{1-\alpha}\frac{\left(\sigma_\varepsilon^2\right)^2\Theta^2}{\left(\Theta+\alpha\omega_{22}\right)^2\left(\beta\Theta+\alpha\omega_{12}\right)^2}\right).$

Thus, we compare the difference of the forward 2SLS estimator and the reverse 2SLS estimator after subtracting off the bias term which arises to the second order of approximation. Note that the order of the variance in Theorem 1 is $O(n^{-2})$ rather than $O(n^{-1})$ because of the second order approximation. Thus, the specification test takes the form of an asymptotic t statistic:

(6.7) $$m = \frac{\hat{d}}{\hat{w}^{0.5}},$$

where $\hat{d}$ is the LHS of Theorem 1, and $\hat{w}$ is a consistent estimate of the variance in Theorem 1. We discuss later how to estimate this variance term.

---

[4] A similar result was given in Hausman (1978) for the bias of the least squares coefficient when the RHS

7. A Specification Test based on Nagar-Type Estimators

We also explore an alternative approach, which is closely related to comparing the forward and reverse 2SLS estimators. Nagar (1959) calculated the second-order bias of the 2SLS (and other k class) estimators. He demonstrated how to bias-adjust these estimators to second order. Thus, we can estimate Nagar-type bias corrected IV estimators and then again compare forward and reverse bias-corrected estimators. The estimates should be very similar if the asymptotic approximations are sufficient for the particular simultaneous equation model specification. Thus, we follow a similar strategy as in the last section, but here we use bias-corrected forward and reverse regression estimators.

We use the B2SLS estimator of Donald and Newey (1998) to estimate the forward and reverse regressions. Note that this estimator is a k class estimator and is a member of the Nagar class of estimators. The forward IV estimator of β is:

$$(7.1)\quad b \equiv \frac{y_2'P_z y_1 - \lambda y_2'M_z y_1}{y_2'P_z y_2 - \lambda y_2'M_z y_2}, \qquad \text{where} \qquad \lambda = \frac{\frac{K-2}{n}}{1-\frac{K-2}{n}}.$$

We can also estimate β by the reverse IV specification:

$$(7.2)\quad \frac{1}{c} \equiv \frac{y_1'P_z y_1 - \lambda y_1'M_z y_1}{y_2'P_z y_1 - \lambda y_2'M_z y_1}.$$

By the delta method based on Lemma 5 in Appendix, we can show that

Theorem 2.

$$\sqrt{n}\left(\left(b, \frac{1}{c}\right)' - (\beta, \beta)'\right) \to N\left(0, \begin{bmatrix} \frac{\sigma_\varepsilon^2}{\Theta} + \frac{\alpha}{1-\alpha}\frac{\sigma_\varepsilon^2\sigma_{v_2}^2 + \sigma_{\varepsilon v_2}^2}{\Theta^2} & \frac{\sigma_\varepsilon^2}{\Theta} + \frac{\alpha}{1-\alpha}\frac{\sigma_\varepsilon^2\sigma_{v_1 v_2} + \sigma_{\varepsilon v_2}\sigma_{\varepsilon v_1}}{\beta\Theta^2} \\ & \frac{\sigma_\varepsilon^2}{\Theta} + \frac{\alpha}{1-\alpha}\frac{\sigma_\varepsilon^2\sigma_{v_1}^2 + \sigma_{\varepsilon v_1}^2}{\beta^2\Theta^2} \end{bmatrix}\right)$$

.

variable was correlated with the stochastic term in the equation.

Proof: See Appendix B.

As in Theorem 1, the terms of the variance matrix are of order $O(n^{-2})$ due to the second order nature of the asymptotic approximation. We will use Theorem 3 below to compare the forward and reverse bias adjusted estimators of β to form a test of the model specification.

We may want to consider linear combinations of b and $\frac{1}{c}$ for improved inference. It can easily be shown that the asymptotic variance, to second order, for the optimal linear combination is given by

$$(7.3) \qquad Var_a(b_{LIML}) = \frac{\sigma_\varepsilon^2}{\Theta} + \frac{\alpha}{1-\alpha}\frac{\det(\Omega)}{\Theta^2},$$

which coincides with the asymptotic variance of LIML as derived by Bekker (1994). Therefore, we may interpret LIML as an optimal linear combination of bias corrected forward 2SLS and reverse 2SLS. LIML is also known to be median unbiased for normal distributions of the stochastic disturbance of equation (2.1), as shown by Anderson (1977), and, more generally, for symmetric distributions of the stochastic disturbance of equation (2.1), by Rothenberg (1983). Thus, the optimality results of Pfanzagl and Wefelmeyer (1978, 1979) are applicable to claim that the resulting LIML estimator is admissible, while other k class estimators are inadmissible unless λ in the estimator definition above has a coefficient of unity.

We now calculate our second specification test by comparing the forward and reverse B2SLS estimators. Note that no bias correction need be made as in Theorem 1 and in the first specification test since our estimators here have no bias to second order. The variance of the difference of the estimators thus has a very simple form. As a consequence of Theorem 2, we obtain

Theorem 3: $\sqrt{n}\left(b - \frac{1}{c}\right) \to N\left(0, \frac{2\alpha}{1-\alpha}\frac{(\sigma_\varepsilon^2)^2}{\beta^2\Theta^2}\right).$

Proof: Follows from

$$\left(\frac{\sigma_\varepsilon^2}{\Theta}+\frac{\alpha}{1-\alpha}\frac{\sigma_\varepsilon^2\sigma_{v_2}^2+\sigma_{\varepsilon v_2}^2}{\Theta^2}\right)+\left(\frac{\sigma_\varepsilon^2}{\Theta}+\frac{\alpha}{1-\alpha}\frac{\sigma_\varepsilon^2\sigma_{v_1}^2+\sigma_{\varepsilon v_1}^2}{\beta^2\Theta^2}\right)-2\left(\frac{\sigma_\varepsilon^2}{\Theta}+\frac{\alpha}{1-\alpha}\frac{\sigma_\varepsilon^2\sigma_{v_1v_2}+\sigma_{\varepsilon v_2}\sigma_{\varepsilon v_1}}{\beta\Theta^2}\right)$$

$$=\frac{\alpha}{1-\alpha}\frac{1}{\beta^2\Theta^2}\left(\sigma_\varepsilon^2\cdot\left(\beta^2\sigma_{v_2}^2-2\beta\sigma_{v_1v_2}+\sigma_{v_1}^2\right)+\left(\beta^2\sigma_{\varepsilon v_2}^2-2\beta\sigma_{\varepsilon v_2}\sigma_{\varepsilon v_1}+\sigma_{\varepsilon v_1}^2\right)\right)$$

$$=\frac{\alpha}{1-\alpha}\frac{1}{\beta^2\Theta^2}\left(\sigma_\varepsilon^2\cdot\sigma_\varepsilon^2+\left(\beta\sigma_{\varepsilon v_2}-\sigma_{\varepsilon v_1}\right)^2\right)$$

$$=\frac{2\alpha}{1-\alpha}\frac{(\sigma_\varepsilon^2)^2}{\beta^2\Theta^2}.$$

Note that the denominator of the variance is again of order $O\left(n^{-2}\right)$ because of the second order approximation we use. Our second specification test has the form of an asymptotic t statistic:

$$(7.4)\quad m_2=\frac{\hat{d}_2}{\hat{w}_2^{0.5}},$$

where the numerator is the difference of the two estimators multiplied by $n^{1/2}$ and the denominator is the square root of the variance term in Theorem 3. We subsequently discuss how to consistently estimate the variance term.

8. Similarity of Bekker's (1994) Asymptotics to the Edgeworth Expansion for *k*-Class Estimators

In this section, we demonstrate that the relevance of Bekker's (1994) asymptotic approximation is not necessarily confined to the case where $\alpha = K/n$ is large. Given that Bekker's alternative limiting distribution is driven by the assumption that the number of instruments grows to infinity as a function of the sample size, his approximation may seem of limited applicability when the number of instruments is 'small.' We demonstrate that Bekker's approximation is in fact quite similar to the second order Edgeworth expansion with symmetrically distributed errors. Unlike the Edgeworth expansion based approximation, Bekker's approximation produces limiting normal distributions, which causes the resulting tests to be quite convenient. Normal approximations turn out to be quite reasonable approximations as supported by our Monte Carlo simulation discussed in Section 13.

Rothenberg (1983) computes higher order moments of k-class estimators. For symmetrically distributed errors, it can be shown by Rothenberg (1983, Theorem 2) that $\sqrt{n}(b_{2SLS} - \beta)$ has an (approximate) mean

$$\text{(8.1)} \qquad \frac{(K-2)\sigma_{\varepsilon v_2}}{\Theta\sqrt{n}},$$

which predicts that the mean of b_{2SLS} is approximately equal to

$$\text{(8.2)} \qquad \beta + \alpha \frac{\sigma_{\varepsilon v_2}}{\Theta}.$$

Observe that equation (8.2) is similar to the probability limit (4.2) of 2SLS under Bekker's asymptotics except that equation (4.2) uses $\Theta + \alpha\omega_{22}$ as the denominator of the bias. As for LIML, using an Edgeworth expansion we find that $\sqrt{n}(b_{LIML} - \beta)$ has an (approximate) mean

$$\text{(8.3)} \qquad -\frac{\sigma_\varepsilon^2}{\Theta\sqrt{n}} = o(1),$$

and (approximate) variance

$$\text{(8.4)} \qquad \frac{\sigma_\varepsilon^2}{\Theta} + \frac{K}{n}\frac{\sigma_\varepsilon^2\sigma_{v_2}^2 - \sigma_{\varepsilon v_2}^2}{\Theta^2} + o(1) \approx \frac{\sigma_\varepsilon^2}{\Theta} + \alpha\frac{\det(\Omega)}{\Theta^2},$$

which is similar to the Bekker-based result we derived for LIML in equation (7.1), except the approximating factor $\alpha/(1-\alpha)$ in equation (7.1) has changed to α in equation (8.4).

As for the (forward) k-class estimator b considered by Donald and Newey (1998), using an Edgeworth expansion it can be shown that $\sqrt{n}(b-\beta)$ has (approximate) mean 0, and (approximate) variance

(8.5) $$\frac{\sigma_\varepsilon^2}{\Theta}+\frac{K}{n}\frac{\sigma_\varepsilon^2\sigma_{v_2}^2+\sigma_{\varepsilon v_2}^2}{\Theta^2}+o(1)\approx\frac{\sigma_\varepsilon^2}{\Theta}+\alpha\frac{\sigma_\varepsilon^2\sigma_{v_2}^2+\sigma_{\varepsilon v_2}^2}{\Theta^2}.$$

Notice that equation (8.5) again agrees with a Bekker-based asymptotic variance of the Donald-Newey estimator in Theorem 2 except that, again, Rothenberg's Edgeworth correction terms are of order α, whereas Bekker's correction terms are of order $\alpha/(1-\alpha)$. These results suggest that Bekker's asymptotic approximation can be interpreted as a convenient method of Edgeworth expansion with wider applicability than might be thought considering Bekker-type asymptotics in isolation.

Bekker-type asymptotics or Edgeworth expansions do not always provide reasonable approximation to finite sample distribution of IV estimators. First of all, it should be noted that variance predicted by the Edgeworth expansion is not always guaranteed to be positive. It can be shown that the (approximate) variance of $\sqrt{n}(b_{2SLS}-\beta)$ calculated by Rothenberg (1983, Theorem 2) is equal to

(8.6) $$\frac{\sigma_\varepsilon^2}{\Theta}-\frac{K}{n}\frac{\sigma_\varepsilon^2\sigma_{v_2}^2+7\sigma_{\varepsilon v_2}^2}{\Theta^2}+o(1)\approx\frac{\sigma_\varepsilon^2}{\Theta}-\alpha\frac{\sigma_\varepsilon^2\sigma_{v_2}^2+7\sigma_{\varepsilon v_2}^2}{\Theta^2}.$$

Observe that equation (8.6) is smaller than equations (8.4) or (8.5), which suggests that the variance of 2SLS is smaller than that of a Nagar-type estimator or LIML.[5] We could not tell whether Bekker's asymptotics predicts the same pattern of variances. There is good reason to believe that equation (8.6) may be overly optimistic about the variance of 2SLS in certain situations: It is not difficult to come up with a parameter combination such that equation (8.6) is negative, especially when the first stage R_f^2, and hence Θ, is extremely small which can correspond to the "weak instrument" situation. Because

[5] The bias of 2SLS is larger, which leads to the optimality results of Rothenberg (1983).

Bekker's asymptotic variance of $\sqrt{n}(b_{2SLS} - \beta)$ is based on the delta method, it is guaranteed to be nonnegative. Therefore, Bekker's asymptotics may be interpreted as a way to fix such undesirable predictions of Edgeworth expansions in extreme situations.[6] However, a further caution should be recognized when using either Bekker's asymptotics or Edgeworth expansions for LIML or Nagar-type estimators. Neither LIML nor Nagar-type estimators possess finite sample second moments.[7] Thus, the performance of the asymptotic approximations may vary depending on sample size and whether a "weak instruments" situation is present. We explore this possibility in Section 13 where we perform Monte Carlo experiments.

9. Estimation of Asymptotic Variance Terms

9.1 2SLS

For the first specification test, we need to estimate the asymptotic variance

$$(9.1)\quad w_1 = \frac{2\alpha}{1-\alpha} \frac{\left(\sigma_\varepsilon^2\right)^2 \Theta^2}{\left(\Theta + \alpha\omega_{22}\right)^2 \left(\beta\Theta + \alpha\omega_{12}\right)^2}.$$

For this purpose we note that σ_ε^2 may be consistently estimated by

$$(9.2)\quad \hat{\sigma}_\varepsilon{}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(y_i - \hat{\beta} y_{2i}\right)^2$$

for some consistent estimator $\hat{\beta}$ for β. We also note that LIML is consistent for β. As for Θ, we note that

$$(9.3)\quad \hat{\Theta} \equiv \frac{1}{n} y_2' P_z y_2 - \frac{\hat{\alpha}}{1-\hat{\alpha}} \frac{1}{n} y_2' M_z y_2 = \Theta + o_p(1)$$

[6] The fact that Edgeworth expansion predicts a smaller variance for 2SLS suggests that if the bias of 2SLS is negligible 2SLS may dominate both Nagar-type estimators and LIML under reasonable loss functions. In Section 13, we investigate such a potential outcome by Monte Carlo simulation.
[7] See Mariano and Sawa (1972). As for the forward Nagar estimator, it does not even possess first moments as established by Sawa (1972).

by Lemma 5. As for $\hat{\alpha}$, we follow Bekker and use

$$(9.4) \quad \hat{\alpha} \equiv \frac{K-1}{n-1}.$$

Finally, we note that

$$(9.5) \quad \Theta + \alpha\omega_{22} = \operatorname{plim}\frac{1}{n}y_2'P_z y_2, \qquad \text{and} \qquad \beta\Theta + \alpha\omega_{12} = \operatorname{plim}\frac{1}{n}y_2'P_z y_1.$$

To summarize, our consistent estimator for asymptotic variance is given by

$$(9.6) \quad \hat{w}_1 = 2\frac{K-1}{n-K}\frac{\left(\sum_{i=1}^{n}(y_i - \beta_{LIML}y_{2i})^2\right)^2\left(y_2'P_z y_2 - \frac{K-1}{n-K}y_2'M_z y_2\right)^2}{(y_2'P_z y_2)^2(y_2'P_z y_1)^2}.$$

9.2 Bias Corrected 2SLS

For the second specification test based on the bias corrected 2SLS estimators, we need to estimate the asymptotic variance

$$(9.7) \quad w_2 = \frac{2\alpha}{1-\alpha}\frac{\sigma_\varepsilon^2}{\beta^2\Theta^2}.$$

By the same calculation as in the previous section, a consistent estimator is given by

$$(9.8) \quad \hat{w}_2 = 2\frac{K-1}{n-K}\frac{\left(\sum_{i=1}^{n}(y_i - \beta_{LIML}y_{2i})^2\right)^2}{\beta_{LIML}^2\left(y_2'P_z y_2 - \frac{K-1}{n-K}y_2'M_z y_2\right)^2}.$$

In either of the variance estimates of equations (9.6) and (9.8), a different consistent estimator other than LIML can be used, with no change in the distribution of the estimated test statistic.

10. Included Exogenous and Predetermined Variables

We have so far assumed that a single jointly endogenous RHS variable exhausts the list of explanatory variables. The results we have derived are fully general with respect to the inclusion of predetermined variables in equation (2.1). In this section, we demonstrate that our procedure would need to be modified if equations (2.1) and (2.2) are understood to be equations where included exogenous variables are partialled out.

Suppose that the full model is

$$
\begin{aligned}
Y_{1i} &= \beta Y_{2i} + Z_{1i}'\gamma + \mathrm{E}_i \\
Y_{2i} &= Z_{1i}'\phi + Z_{2i}'\pi_2 + V_{2i}
\end{aligned}
\tag{10.1}
$$

where Z_{1i} is a k_1 dimensional vector of included predetermined variables in equation (10.1) and Z_{2i} is a K dimensional vector containing all other predetermined variables. Let M_{Z_1} denote the projection operator partialling Z_{1i} out of equation (10.1), and let equations (2.1) and (2.2) be understood to be the resultant expression: Let Y_j denote a column vector consisting of Y_{ji}. Define Z_1, Z_2, E, and V_2 similarly. With

$$
y_1 = M_{Z_1}Y_1, \quad y_2 = M_{Z_1}Y_2, \quad z = M_{Z_1}Z_2, \quad \varepsilon = M_{Z_1}\mathrm{E}, \quad v_2 = M_{Z_1}V_2,
\tag{10.2}
$$

we obtain equations (2.1) and (2.2) premultiplying equation (10.1) by M_{Z_1}.

We ask if there is any simple way to compute $y_2'P_zy_2$, $y_2'M_zy_2$, etc. avoiding the projection of Z_2 on Z_1. Simple projection algebra based arguments show that they could be characterized quite easily. The following then provides a convenient computational procedure:

- Regress Y_1 and Y_2 on Z_1. Obtain residuals, and label them W_1 and W_2.
- Regress Y_1 and Y_2 on Z_1 and Z_2. Obtain residuals, and label them $\widetilde{W}_1$ and $\widetilde{W}_2$.
- Let $\hat{y}_1 \equiv W_1 - \widetilde{W}_1$ and $\hat{y}_2 \equiv W_2 - \widetilde{W}_2$.

- Compute $y_2'P_zy_2 = \hat{y}_2'\hat{y}_2$, $y_2'P_zy_1 = \hat{y}_2'\hat{y}_1$, $y_1'P_zy_1 = \hat{y}_1'\hat{y}_1$, $y_2'M_zy_2 = \tilde{y}_2'\tilde{y}_2$, $y_2'M_zy_1 = \tilde{y}_1'\tilde{y}_1$, $y_1'M_zy_1 = \tilde{y}_1'\tilde{y}_1$, and plug into equations (6.1), (9.1), and (9.2).
- As for σ_ε^2, we can estimate it by the average squared LIML residuals on the full regression: We can replace $\sum_{i=1}^n (y_i - \beta_{LIML} y_{2i})^2$ in equations (9.1) and (9.2) by $\sum_{i=1}^n (Y_{1i} - \beta_{LIML} Y_{2i} - Z_{1i}'\gamma_{LIML})^2$.
- As for K in equations (6.6), (9.6), and (9.8), we may conservatively use $K = \dim(Z_{1i}) + \dim(Z_{2i})$, although $K = \dim(z_i) = \dim(Z_{2i})$ may also be a reasonable choice.

Note also that one may want to adjust the sample size in the above equations to $n^* = n - k_1$ to take account of the loss of degrees of freedom from partialling out the Z_{1i} variables.

11. Additional RHS Jointly Endogenous Variables

To this point in the paper, we have only considered the situation of one RHS jointly endogenous variable, which is by far the most common situation encountered in empirical application of IV estimators (*e.g.* 2SLS). We now extend the model specifications to allow for additional RHS jointly endogenous variables. We derive the second specification test for 2 RHS jointly endogenous, which demonstrates how to generalize our results to $r_1 > 2$ RHS jointly endogenous variables. We leave the derivation of the first specification test in this situation to further research.

We extend our original simultaneous model specification of equations (2.1) and (2.2) to the situation of 2 RHS jointly endogenous variables:

(11.1) $$y_1 = \beta_2 y_2 + \beta_3 y_3 + \varepsilon_1$$

(11.2) $$\begin{aligned} y_2 &= z\pi_2 + v_2 \\ y_3 &= z\pi_3 + v_3 \end{aligned}$$

where we use the same matrix and vector notation as before. We consider estimation of β_2 and β_3 in equation (11.1) by use of the Donald and Newey (1998) B2SLS estimator. We will refer to the estimator as (b_1, c_1). Changing the normalization we could also estimate $(1/\beta_2, -\beta_3/\beta_2)$ or $(1/\beta_3, -\beta_2/\beta_3)$. Thus, we would have three potential estimators for (β_2, β_3). The question would naturally arise of how to combine these potential estimators to achieve the most powerful specification test of a given size.

However, as we demonstrate in Technical Appendices C and D, it turns out that we cannot stack the estimates to derive a more powerful test since the asymptotic variance matrix of the three tests is singular. Thus, all tests based on a single difference will have the same operating characteristics, and a more powerful test cannot be derived using additional differences (contrasts). Thus, we will use the estimator $b_1 - 1/b_2$, where $1/b_2$ is the estimator derived from application of B2SLS (or another Nagar-type estimator) to the equation:

$$(11.3) \qquad y_2 = \frac{1}{\beta_2} y_1 + \left(\frac{-\beta_3}{\beta_2} \right) y_2 + \varepsilon_2 .$$

In Appendix D we derive a consistent estimate of the asymptotic variance of the scaled difference of the two estimators $d_3 \equiv n^{1/2}(b_1 - 1/b_2)$ to be

$$(11.4) \quad 2\frac{K-1}{n-K} \frac{\left(\sum_{i=1}^{n} \left(y_{1i} - \beta_{2,LIML} y_{2i} - \beta_{3,LIML} y_{3i} \right)^2 \right)^2}{\beta_{2,LIML}^2 \left(y_2' P_z y_2 - \frac{K-1}{n-K} y_2' M_z y_2 - \frac{\left(y_2' P_z y_3 - \frac{K-1}{n-K} y_2' M_z y_3 \right)^2}{\left(y_3' P_z y_3 - \frac{K-1}{n-K} y_3' M_z y_3 \right)} \right)^2} .$$

As before, other Nagar-type estimators may replace LIML estimators in the above formula. The specification test will take the form:

$$(11.5) \qquad m_3 = \frac{\hat{d}_3}{\hat{w}_3^{0.5}},$$

where $\hat{w}_3$ is the estimated variance in the above equation.

Inclusion of exogenous and predetermined variables in the specification as in equation (10.1) in Section 10 raises no new complications. The partialling-out methodology we used in Section 9 is directly applicable to the current situation with 2 (or more) RHS jointly endogenous variables. The new jointly endogenous variable, Y_3, is partialled out by regressing Y_3 on Z_1. All other formulae follow as before, and the above variance formula can be used on the partialled out variables to form the second form of our specification test.

12. An Empirical Example

We analyze an empirical example of a simultaneous equation specification of a demand function. This type of specification is the original type of problem studied by Haavelmo, who demonstrated that least squares lead to bias results. The left hand side variable of the first specification represents movements of a homogenous bulk chemical commodity measured in log of ton miles. Data were collected on approximately 50 origin-destination (OD) pairs over a 33 month period. Each data point is an individual freight movement. As right hand side variables, we include the log of the price of the movement which is a jointly endogenous variable, a measure of economic activity, and OD indicator variables which change each year to allow for fixed effects for OD pairs. We also used a trucking price index variable, which was assumed to be predetermined. Altogether, we have 132 right hand side variables, one of which is jointly endogenous. As instruments for the jointly endogenous variable we use the log of a short run marginal cost variable for the appropriate movements of the bulk commodity, which is available for each shipment.[8] The other instrumental variable that we use is the monthly price index for diesel fuel.

In Table A in the first column we give the estimated price elasticity (and an estimate of the first order asymptotic standard error) along with the estimated standard

[8] While these data are accounting data that unlikely to be true measures of marginal cost, potential errors in variables in instruments do not create a problem in instrumental variable estimation under the usual assumptions. See Hausman (1977).

error and the R^2. Note that the price elasticity estimate is –1.36 (.147) and is estimated quite precisely. The R^2 is also quite high at .962. In column 2 we use the conventional 2SLS estimator. The estimated price elasticity increases in magnitude to –2.03 (.465) which is the expected outcome given the expected direction of the simultaneous equation bias of least squares. Again, we find a relatively small estimated standard error. When we consider possible diagnostics, we find that the R^2 of the reduced form is 0.941 with an F statistic of 154.5. The R^2 of the reduced form model after all of the predetermined RHS variables of the structural equation have been partialled out is .093 with an associated F statistic of 74.6. While the partialled out model has a lower R^2 and F statistic, as expected, they do not indicate a problem according to rules of thumb previously put forward in the literature.

We now interchange the jointly endogenous variable and put price on the LHS and quantity on the RHS. The results are given in Column 3 of Table A. We use the same instruments and find our estimate of $\frac{1}{c}$ to be -0.433 (.094) so the reverse estimate of the price elasticity is –2.31 (.500) so that the difference between the forward and reverse estimates of the price elasticity is 0.275. The question is whether these estimates, which should be exactly the same under first order asymptotics, are different enough to reject the conventional first order asymptotic approach.

Using the second order approximation of equation (6.6), we estimate the difference in the bias of the two estimators to be -.0012, which is quite small. We then use equation (9.1) to calculate the variance and estimate our specification test statistic to be:

$$m = \frac{\hat{d}}{\hat{w}^{0.5}} = \frac{10.03}{5.43} = 1.85. \tag{12.1}$$

Thus, up to a second order asymptotic approximation we do not reject the first order asymptotic approach or the associated estimators.

We now use the Nagar-type estimator of Donald and Newey in columns four and five. Here we find the same estimates in the forward and reverse direction because the degree of overidentification is 1. Using Theorem 2 to form the specification test, we find it to be 1.82, which is very similar to our previous estimate. Lastly, we find the LIML estimate to be –2.05 (.469), which does lie between the forward and reverse estimates, as expected, but note that it is quite close to the forward 2SLS estimate.

We now increase the number of instruments by 131 by interacting the cost instrumental variable with the corresponding OD indicator variables. This new variable allows for unobserved cost differences across the different OD pairs. The results are given in Table B. The first column has the forward 2SLS estimate of –1.24 (.194) which has decreased significantly in magnitude back towards the least squares estimate from Table B of –1.36. A situation of weak instruments may well be present. The R^2 of the reduced form is 0.972. The R^2 of the reduced form for the partialled out model is .219 with an associated F statistic of 2.79, which gives little indication of a weak instruments problem.

In the second column of Table B we present the reverse 2SLS estimate of –8.01 (.789), which is approximately 6.5 times higher than the forward 2SLS estimate. The difference between the two estimates of –6.77 would likely be considered significant, on economic terms, by most researchers. Here the R^2 of the reduced form of the partialled out model is .038 with an associated F statistic of .280, which could indicate that a "weak instruments" problem exists according to rules of thumb put forward in the past literature. The difference in second order bias terms is estimated to be -.041, much smaller than the actual difference in the forward and reverse estimates. The test statistic. is estimated to be

$$(12.2)\qquad m = \frac{\hat{d}}{\hat{w}^{0.5}} = \frac{247.3}{17.4} = 14.21.$$

Thus, the specification term rejects the conventional first order asymptotic approach, and we would recommend that the estimates not be used.

In columns 3 and 4 of Table B we present the Nagar-type bias corrected forward and reverse IV estimates recommended by Donald and Newey. The forward estimator is now –1.21 (.194), while the reverse estimator is –4.64 (.293). While some improvement has been made, the two estimates still differ by a large amount. The specification test is estimated to be 5.78, which again rejects. Lastly, the LIML estimate is –1.18 (.211), which, again, is quite close to the forward regression. Thus, we do not recommend the use of LIML in the weak instrument situation when the forward and reverse Nagar-type estimators differ significantly because it often has a significant asymptotic bias, as indicated in this example and in other empirical examples we have investigated.

We conclude that in a real world example that the IV estimators can perform poorly in the weak instrument situation. Using the forward and reverse estimate seems to give a convenient metric to analyze the performance of the estimators. The specification tests we have proposed also work as we would expect. We now turn to some Monte-Carlo results to explore further the performance of the tests.

13. Monte Carlo Experiments

We generated data from the model specification

$$y_{1i} = \beta z_i'\pi_2 + v_{1i}$$

$$y_{2i} = z_i'\pi_2 + v_{2i} \qquad i = 1,\ldots,n$$

such that

$$z_i \sim N(0, I_K), \qquad \pi_2 = (\phi,\ldots,\phi),$$

$$\Omega = \begin{bmatrix} 1 & \omega_{12} \\ \omega_{12} & 1 \end{bmatrix}, \qquad \widetilde{R}_f^2 \equiv \frac{\pi_2' E[z_i z_i']\pi_2}{\pi_2' E[z_i z_i']\pi_2 + \omega_{22}} = \frac{K\phi^2}{K\phi^2 + 1}.$$

Here, $\widetilde{R}_f^2$ denotes the theoretical R^2 in the first stage regression. We use following parameter combinations:

$$n = 100, \quad 250, \quad 1000, \quad 10000$$
$$\sigma_{\varepsilon v_2} = -.9, \quad -.5, \quad .5, \quad .9$$
$$\widetilde{R}_f^2 = .001, \quad .01, \quad .1, \quad .3$$
$$K = 5, \quad 10, \quad 30$$

We examined performance of our tests by 5000 Monte Carlo replications. Tables 1 – 4 report results using a range of instruments from 5-30, sample sizes of 100-10,000, and a range of covariances (correlations) where we vary the $\tilde{R}_f^2$ of the reduced form regression.[9] Columns (a) and (b) report the actual size of the test based on forward and reverse 2SLS with 10% and 5% nominal sizes. The actual sizes of the test are generally quite close to the nominal sizes, with only a small falling off above the nominal size when the number of instruments becomes large and the $\tilde{R}_f^2$ becomes quite low (.001). Columns (c) and (d) report actual biases of forward and reverse 2SLS estimators, and column (e) reports the expected value of $\hat{B}$. The estimates of the difference of second order bias terms are typically quite accurate, although when the expected difference of biases becomes quite large, the estimates can vary by quite a lot. However, in these situations, the test statistic should still work well because the presence of a large expected bias (even if not measured totally accurately) will alert the econometrician to the dangers of using 2SLS, or other IV estimators, in this situation. Importantly, the estimates of the expected value of $\hat{B}$ appear to do a good job of indicating the presence of "weak instruments," *e.g.* column (e) in Table 3 with "weak instruments" compared to column (e) of Table 4 where the instruments are better because the R^2 of the reduced form equation is much higher. Thus, the second order asymptotic approach seems to provide a useful tool to indicate when the "weak instrument" problem is present.

Columns (f) and (g) report the actual size of the traditional test of overidentification (based on forward 2SLS) with nominal sizes equal to 10%, and 5%.[10] The conventional test of overidentification, based on the forward 2SLS estimates, does not perform well in a large variety of situations, as has been noted numerous times in the previous literature. As shown in Table 1 the conventional test of overidentification often has actual size of above 0.3, when the nominal size is smaller than 0.1. Note that when the $\tilde{R}_f^2$ of the reduced form becomes high, the test of overidentification has approximately the correct size. This result is expected since the test of overidentification assumes

---

[9] In Tables 1 – 4, we set the values of β such that Var(ε)=1.

[10] We use $n \cdot R^2$ of the regression of the forward 2SLS residuals on instruments as the test statistic. Because forward and reverse 2SLS should be perfectly correlated under conventional asymptotics, tests of overidentification based on forward and reverse 2SLS should have the same operating characteristics if conventional asymptotics provides reasonable approximations to sampling distributions of various IV estimators.

implicitly that this $\widetilde{R}_f^2$ is unity, in the sense that it assumes that the reduced form coefficients are known with certainty. Also, for very large samples, *e.g.* 10,000, the size of the test becomes approximately correct. When the number of instruments begins to increase, the size performance of the test of overidentification falls off again. When the number of instruments becomes quite large (30) in Tables 1 – 4, the actual size of the conventional test of overidentification becomes abysmally large, sometimes exceeding 0.5 in the low $\widetilde{R}_f^2$ situation. Thus, we conclude that the second order asymptotic approximations work considerably better than the conventional first order asymptotic approximations when applied to the 2SLS estimator.

Columns (h) and (i) report results for cases where we consider the Nagar-type bias corrected estimator. We find that the actual size of the new specification test based on Donald and Newey's estimator with 10% and 5% nominal sizes again approximates the nominal size quite well with no tendency to be too large a size for the test. Columns (m) and (n) report the actual size of the traditional test of overidentification (based on Donald and Newey's forward estimator) with nominal sizes equal to10% and 5%.[11] While the use of the Nagar-type estimator improves the traditional test of overidentification, the conventional test of overidentification sometimes has an actual size of above 0.2, when the nominal size is smaller than 0.1. Note that when the $\widetilde{R}_f^2$ of the reduced form becomes high, the test of overidentification has approximately the correct size once again.

Also, note that in columns (j)-(l) where we report the means biases of the Nagar-type and LIML estimators, the mean bias of the Donald-Newey (Nagar-type) estimators and LIML estimators occasionally are found to be very large. This finding results from the non-existence of finite sample moments of Nagar-type and LIML estimators that we discussed in Section 8. These results should be a caution about using Nagar-type or LIML estimates even with the second order asymptotic approximations without further investigation or specification tests in a given empirical problem.

[11] We use $n \cdot R^2$ of the regression of the residuals from Donald and Newey's forward estimator on instruments as the test statistic.

Table 5 report Monte Carlo results in some "extreme cases" where the number of instruments is large, $K = 30$, and the $\widetilde{R}_f^2$ of the reduced form is low. The actual sizes of the new specification test in columns (a)-(b) and (h)-(i) are again close to the nominal sizes, although in a few cases the test based on the Nagar-type estimator does have too large size. However, these results should be compared to the traditional test of overidentification based on 2SLS in columns (f) and (g) where the actual sizes always exceed 0.85, even though the nominal size is 0.10! Similarly, the traditional tests of overidentification based on the Nagar-type estimators in columns (m) and (n) do better, but they still exceed the nominal size by factors of 2 to 5. These results, along with the second order bias estimates of column (e), which are again successful in indicating the presence of "weak instruments," demonstrate that tests based on the second order asymptotic approximations do considerably better than tests based on the conventional first order asymptotic approximations in these extreme situations.

As we discussed in Section 8, Edgeworth expansions predict smaller variances for 2SLS than for LIML. In Table 6, we compare 2SLS and LIML when the bias of 2SLS is negligible and $\widetilde{R}_f^2$ is small. In all cases, 2SLS dominates LIML under mean square error loss. This result is not surprising because LIML does not possess second moments. However, the dispersion of LIML around β measured in the interquartile range or interdecile range is much larger than that of 2SLS.[12] We conclude that Bekker's asymptotics may be a poor approximation when $\widetilde{R}_f^2$ is extremely small, which leads to the suggestion of using the specification test to help determine the usefulness of second order asymptotics in a given situation.

14. Conclusions

Using the forward and reverse 2SLS estimates to test for weak instruments to form a specification test seems to be a helpful approach. We use a second order asymptotic approximation to form a test statistic to see if the conventional first order asymptotic approach is accurate enough to provide reliable inferences. The first order

[12] The fact that 2SLS does better than LIML suggests that 2SLS should be used for Hausman tests of endogeneity of regressors.

asymptotics implies that the two estimates should be the same, while the second order asymptotic approach allows for different biases in the two estimators. The econometrician can also consider the estimates and see whether the difference in the estimates is large in economic terns relative to what would be expected. The test statistic is straightforward to compute using existing econometric software to calculate the 2SLS estimators, Nagar-type estimators, and LIML as well as the partialled out models.

While giving guidance to inference is often subjective based on the econometrician's beliefs, we suggest the following approach. We suggest that the new specification test of equation (6.7) based on forward and reverse 2SLS be done. If the 2SLS estimates are close and the estimate of the bias term $\hat{B}$ from equation (6.6) is small, the conventional first order asymptotics may be used, and the 2SLS estimates should be all right. If the test rejects or the estimated bias term is large, we then suggest using Nagar-type estimates to perform the second specification test based on equation (7.4). If the forward and reverse estimates are close and the specification test does not reject, we suggest using LIML, which is the optimal combination of the two estimators.[13] If the test rejects, we do not suggest using these estimates as either a failure of the orthogonality conditions or an extreme situation of "weak instruments" is likely to be present. If the Nagar-type forward and reverse estimates are not close but the specification test does not reject, a decision cannot typically be made based on the new specification test.

Our approach can be generalized when more than one jointly endogenous variable is on the right hand side of the model specification. Two variables can be interchanged as before to provide forward and reverse estimates. Second order asymptotic theory is again used to form the associated distributions of the second order distributions for the bias terms and for the specification tests. We derive the rather unexpected result that only one set of differences provides the optimal specification test. So far, we have limited the extension to 2 RHS jointly endogenous variables for the second specification test based on the Nagar-type estimator. We expect to extend our results to 3 or more RHS jointly endogenous variables and to the first specification test, which is based on the

2SLS estimator.

[13] If the LIML estimate differs markedly from the forward and reverse Nagar-type estimates, the LIML estimate should not be used because the problem of the absence of finite sample moments may well be present.

References

ANDERSON, T.W. (1977): "Asymptotic Expansions of the Distributions of Estimates in Simultaneous Equations for Alternative Parameter Sequences," *Econometrica*, 45, 509 – 518.

ANDERSON, T.W., AND H. RUBIN (1949): "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics,* 20, 46-63.

BEKKER, P.A. (1994): "Alternative Approximations to the Distributions of Instrumental Variable Estimators", *Econometrica,* 92, 657 – 681.

BOUND, J, JAEGER, D. A., AND BAKER, R. M. (1995): "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90, 443 – 450.

DONALD, S.G., AND NEWEY, W.K. (1998): "Choosing the Number of Instruments," *mimeo.*

HALL, A. R., RUDEBUSCH, G. D., AND WILCOX, D. W. (1996): "Judging Instrument Relevance in Instrumental Variables Estimation," *International Economic Review*, 37, 283 – 289.

HAUSMAN, J.A. (1977): "Simultaneous Equations with Errors in Variables, " *Journal of Econometrics*, 5, 389 – 401.

HAUSMAN, J.A. (1978): "Specification Tests in Econometrics," *Econometrica*, 76, 1251 – 1271.

HAUSMAN, J.A. (1983): "Specification and Estimation of Simultaneous Equation Models," in Z. Griliches and M. Intriligator, *eds., Handbook of Econometrics*, Vol. 1, Amsterdam: North Holland.

HAUSMAN, J.A., NEWEY, W.K., AND POWELL, J. (1995): "Nonlinear Errors in Variables: Estimation of Some Engel Curves," *Journal of Econometrics*, 65, 205-233.

KADANE, J. (1973): "Testing a Subset of the Overidentifying Restrictions," *Econometrica*, 39, p. 853 – 867.

MARIANO, R.S. AND T. SAWA (1972): "The Exact Finite-Sample Distribution of the Limited-Information Maximum Likelihood Estimator in the Case of Two Included Endogenous Variables," *Journal of the American Statistical Association,* 67, 159 – 163.

NAGAR, A.L. (1959): "The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations," *Econometrica*, 27, 575 – 595.

NELSON, C. R. AND STARTZ, R. (1990a): "Some Further Results on the Exact Small Sample Properties of the Instrumental Variables Estimator," *Econometrica*, 58, 967 – 976.

__________ AND ________ (1990b), "The Distribution of the Instrumental Variables Estimator and Its t-ratio when the Instrument is a Poor One," *Journal of Business*, 63, 5125 – 5140.

PFANZAGL, J., AND WEFELMEYER W. (1978): "A Third-Order Optimum Property of the Maximum Likelihood Estimator," *Journal of Multivariate Analysis*, 8, 1 – 29.

PFANZAGL, J., AND WEFELMEYER W. (1979): Addendum to: "A Third-Order Optimum Property of the Maximum Likelihood Estimator," *Journal of Multivariate Analysis*, 9, 179 – 182.

PHILLIPS, P.C.B. (1989): "Partially Identified Econometric Models," *Econometric Theory*, 5, 181 – 240.

ROTHENBERG, T.J. (1983): "Asymptotic Properties of Some Estimators In Structural Models," in *Studies in Econometrics, Time Series, and Multivariate Statistics*.

SAWA, T. (1972): "Finite-Sample Properties of the k-Class Estimators," *Econometrica*, 40, 653-680.

SARGAN, J.D. (1976): "Econometric Estimators and the Edgeworth Approximation," *Econometrica*, 44, p. 421 – 448.

SHEA, J. (1997), "Instrument Relevance in Multivariate Linear Models: A Simple Measure," *Review of Economics and Statistics*, 79, 348 – 352.

STAIGER, D., AND STOCK, J. H. (1997): "Instrumental Variables Regressions with Weak Instruments," *Econometrica*, 65, 557 – 586.

STARTZ, R., C.R. NELSON, AND E. ZIVOT (1998): "Improved Inference for the Instrumental Variable Estimator," *mimeo*.

WANG, J. AND ZIVOT, E. (1998): "Inference on Structural Parameters in Instrumental Variables Regression with Weak Instruments," *Econometrica*, 66,1389 – 1404.

ZIVOT, E., STARTZ, R., AND NELSON, C.R. (1998): "Valid Confidence Intervals and Inference in the Presence of Weak Instruments," *International Economic Review*, 39, 1119 – 1246.

Table A: Estimates with 1452 Observations, 134 Predetermined Variables and 136 Instruments. $\hat{\alpha} = .002$

	Least Squares	2SLS Forward	2SLS Reverse	Nagar Forward	Nagar Rev.
1. Price elasticity	-1.36 (.147)	-2.03 (.465)	-2.31 (.500)	-2.03 (.465)	-2.31 (.502)
2. Standard Error	.301	.303	.133	.303	.133
3. R^2	.962	-----	----	----	----
Reduced Form Regressions					
4. Standard Error		.053	.308		
5. R^2		.941	.960		
6. F statistic		154.5	234.1		
Partialled Out Reduced Form Regression					
7. Standard Error		.016	.037		
8. R^2		.093	.012		
9. F statistic		74.6	8.54		

Table B: Estimates with 1452 Observations, 134 Predetermined Variables and 266 Instruments $\hat{\alpha} = .092$

	Least Squares	2SLS Forward	2SLS Rev.	Nagar Forward	Nagar Rev.
1. Price elasticity	-1.36 (.147)	-1.24 (.194)	-8.01 (.789)	-1.21 (.194)	-4.64 (.293)
2. Standard Error	.301	.301	.060	.317	.080
3. R^2	.962	-----	----	----	----
Reduced Form Regressions					
4. Standard Error		.038	.295		
5. R^2		.972	.967		
6. F statistic		154.2	130.5		
Partialled Out Reduced Form Regression					
7. Standard Error		.016	.038		
8. R^2		.219	.027		
9. F statistic		2.79	.280		

Appendix

A Technical Lemmas

Let

$$U = \left[\begin{array}{cc} y_1 & y_2 \end{array}\right], \quad M \equiv [\beta \cdot z\pi_2, z\pi_2] = z\pi_2 (\beta, 1), \quad V = U - M.$$

Note that the rows of V are i.i.d. normal with zero mean and variance Ω. Also let

$$\overline{S} = U'P_zU \quad S^{\perp} = U'M_zU.$$

Lemma 1 *$\overline{S}$ and $S^{\perp}$ are independent of each other.*

Proof. Follows easily from normality of U. ∎

We need to establish asymptotic distributions of $\overline{S}$ and $S^{\perp}$. We first show that

Lemma 2 *Let $e_1 = (1,0)'$ and $e_2 = (0,1)'$. Then,*

$$\frac{1}{\sqrt{n}}\left(\left(\begin{array}{c} U'P_zUe_1 \\ U'P_zUe_2 \end{array}\right) - \mathrm{E}\left(\begin{array}{c} U'P_zUe_1 \\ U'P_zUe_2 \end{array}\right)\right)$$

and

$$\frac{1}{\sqrt{n}}\left(\left(\begin{array}{c} U'M_zUe_1 \\ U'M_zUe_2 \end{array}\right) - \mathrm{E}\left(\begin{array}{c} U'M_zUe_1 \\ U'M_zUe_2 \end{array}\right)\right)$$

converge in distribution to normal distributions.

Proof. The proof will only be given for $\left((U'P_zUe_1)', (U'P_zUe_2)'\right)'$. The proof for $\left((U'M_zUe_1)', (U'M_zUe_2)'\right)'$ is omitted. Observe that the conclusion would follow if $\left((U'P_zUa)', (U'P_zU\Delta)'\right)'$ is asymptotically normal for arbitrary a, Δ. By the Cramer-Wald device, the conclusion follows if $U'P_zUa$ is. Bekker (1994) shows that $U'P_zUa$ is asymptotically normal under appropriate conditions. Also note that P_z is symmetric idempotent with rank equal to K. Therefore,

$$\mathrm{E}\left[U'P_zUa\right] = M'P_zMa + K\Omega a$$

and

$$\begin{aligned}\mathrm{Var}\left[U'P_zUa\right] = {} & a'\Omega a M'P_zM + a'M'P_zMa\Omega + \Omega aa'M'P_zM \\ & + M'P_zMaa'\Omega + Ka'\Omega a\Omega + K\Omega aa'\Omega.\end{aligned}$$

If $\frac{1}{n}\operatorname{Var}[U'P_zUa]$ converges, say to W, then Bekker (1994, Lemma 2) notes that $\frac{1}{\sqrt{n}}(U'P_zUa - \mathrm{E}[U'P_zUa])$ is asymptotically normal with mean zero and variance W. In our case, because

$$\frac{1}{n}M'P_zM = \frac{1}{n}\begin{pmatrix}\beta\\1\end{pmatrix}\pi_2'z'P_zz\pi_2(\beta,1) = \left(\frac{1}{n}\pi_2'z'z\pi_2\right)\cdot\begin{pmatrix}\beta\\1\end{pmatrix}(\beta,1) \to \Theta\cdot\Delta,$$

and

$$\frac{K}{n} \to \alpha,$$

we have

$$\begin{aligned}\frac{1}{n}\operatorname{Var}[U'P_zUa] \to\ & a'\Omega a\cdot\Theta\cdot\Delta + \Theta\cdot a'\Delta a\cdot\Omega + \Theta\cdot\Omega aa'\Delta \\ & + \Theta\cdot\Delta aa'\Omega + \alpha\cdot a'\Omega a\cdot\Omega + \alpha\cdot\Omega aa'\Omega \qquad (1)\end{aligned}$$

convergent, where

$$\Delta = \begin{pmatrix}\beta\\1\end{pmatrix}(\beta,1).$$

The conclusion follows. ■

Lemma 3 *Let $\overline{\Lambda}$ denote a symmetric 3×3 matrix such that*

$$\begin{aligned}\overline{\Lambda}_{1,1} &= 4\omega_{11}\Theta\beta^2 + 2\alpha\omega_{11}^2\\ \overline{\Lambda}_{1,2} &= 2\omega_{11}\Theta\beta + 2\beta^2\Theta\omega_{12} + 2\alpha\omega_{11}\omega_{12}\\ \overline{\Lambda}_{1,3} &= 4\beta\Theta\omega_{12} + 2\alpha\omega_{12}^2\\ \overline{\Lambda}_{2,2} &= \omega_{11}\Theta + \beta^2\Theta\omega_{22} + 2\Theta\omega_{12}\beta + \alpha\omega_{11}\omega_{22} + \alpha\omega_{12}^2\\ \overline{\Lambda}_{2,3} &= 2\omega_{22}\Theta\beta + 2\Theta\omega_{12} + 2\alpha\omega_{22}\omega_{12}\\ \overline{\Lambda}_{3,3} &= 4\omega_{22}\Theta + 2\alpha\omega_{22}^2\end{aligned}$$

Then

$$\frac{1}{\sqrt{n}}\left(\begin{pmatrix}\overline{S}_{11}\\ \overline{S}_{12}\\ \overline{S}_{22}\end{pmatrix} - \mathrm{E}\begin{pmatrix}\overline{S}_{11}\\ \overline{S}_{12}\\ \overline{S}_{22}\end{pmatrix}\right) \Rightarrow \mathcal{N}(0,\overline{\Lambda}).$$

Proof. We will denote the asymptotic variance and covariance as Var_a and Cov_a.Note that

$$U'P_zUe_1 = \overline{S}e_1 = \begin{pmatrix}\overline{S}_{11}\\ \overline{S}_{12}\end{pmatrix},\quad U'P_zUe_2 = \overline{S}e_2 = \begin{pmatrix}\overline{S}_{12}\\ \overline{S}_{22}\end{pmatrix}.$$

Therefore, $\left(\overline{S}_{11}, \overline{S}_{12}, \overline{S}_{22}\right)'$ consists of elements of

$$\begin{pmatrix} U'P_zUe_1 \\ U'P_zUe_2 \end{pmatrix},$$

and the asymptotic normality follows easily from Lemma 2. Therefore, it remains to characterize the asymptotic variance. Using (1), we obtain

$$\mathrm{Var}_a \begin{bmatrix} \overline{S}_{11} \\ \overline{S}_{12} \end{bmatrix} = \mathrm{Var}_a \left[U'P_zUe_1\right] = \begin{pmatrix} \overline{\Lambda}_{1,1} & \overline{\Lambda}_{1,2} \\ \cdot & \overline{\Lambda}_{2,2} \end{pmatrix},$$

and

$$\mathrm{Var}_a \begin{bmatrix} \overline{S}_{12} \\ \overline{S}_{22} \end{bmatrix} = \mathrm{Var}_a \left[U'P_zUe_2\right] = \begin{pmatrix} \overline{\Lambda}_{2,2} & \overline{\Lambda}_{2,3} \\ \cdot & \overline{\Lambda}_{3,3} \end{pmatrix}.$$

So far, we have characterized $\mathrm{Var}_a\left(\overline{S}_{11}\right)$, $\mathrm{Var}_a\left(\overline{S}_{12}\right)$, $\mathrm{Var}_a\left(\overline{S}_{22}\right)$, $\mathrm{Cov}_a\left(\overline{S}_{11}, \overline{S}_{12}\right)$, and $\mathrm{Cov}_a\left(\overline{S}_{12}, \overline{S}_{22}\right)$. Therefore, it remains to characterize $\mathrm{Cov}_a\left(\overline{S}_{11}, \overline{S}_{22}\right)$. For this purpose, it is useful to note that

$$\begin{aligned}
&\mathrm{Cov}_a\left(U'P_zUe_1, \left(U'P_zUe_2\right)'\right) + \mathrm{Cov}_a\left(U'P_zUe_2, \left(U'P_zUe_1\right)'\right) \\
&= \begin{pmatrix} \mathrm{Cov}_a\left(\overline{S}_{11}, \overline{S}_{12}\right) & \mathrm{Cov}_a\left(\overline{S}_{11}, \overline{S}_{22}\right) \\ \mathrm{Cov}_a\left(\overline{S}_{12}, \overline{S}_{12}\right) & \mathrm{Cov}_a\left(\overline{S}_{12}, \overline{S}_{22}\right) \end{pmatrix} + \begin{pmatrix} \mathrm{Cov}_a\left(\overline{S}_{11}, \overline{S}_{12}\right) & \mathrm{Cov}_a\left(\overline{S}_{12}, \overline{S}_{12}\right) \\ \mathrm{Cov}_a\left(\overline{S}_{11}, \overline{S}_{22}\right) & \mathrm{Cov}_a\left(\overline{S}_{12}, \overline{S}_{22}\right) \end{pmatrix} \\
&= \overline{V}\left(e_1 + e_1\right) - \overline{V}\left(e_1\right) - \overline{V}\left(e_2\right),
\end{aligned}$$

where $\overline{V}(a)$ is the asymptotic variance of $U'P_zUa$ as discussed in (1). Therefore,

$$\mathrm{Cov}_a\left(\overline{S}_{11}, \overline{S}_{22}\right) + \mathrm{Cov}_a\left(\overline{S}_{12}, \overline{S}_{12}\right) = \left[\overline{V}\left(e_1 + e_2\right) - \overline{V}\left(e_1\right) - \overline{V}\left(e_2\right)\right]_{(1,2)}.$$

It can be shown that

$$\begin{aligned}
&\overline{V}\left(e_1 + e_2\right) - \overline{V}\left(e_1\right) - \overline{V}\left(e_2\right) = \\
&\begin{bmatrix} 4\Theta\omega_{12}\beta^2 + 4\Theta\beta\omega_{11} + 4\alpha\omega_{12}\omega_{11} & 6\Theta\beta\omega_{12} + \Theta\omega_{11} + \Theta\omega_{22}\beta^2 + 3\alpha\omega_{12}^2 + \alpha\omega_{22}\omega_{11} \\ 6\Theta\beta\omega_{12} + \Theta\omega_{11} + \Theta\omega_{22}\beta^2 + 3\alpha\omega_{12}^2 + \alpha\omega_{22}\omega_{11} & 4\Theta\omega_{12} + 4\Theta\beta\omega_{22} + 4\alpha\omega_{12}\omega_{22} \end{bmatrix},
\end{aligned}$$

from which we obtain

$$\begin{aligned}
\mathrm{Cov}_a\left(\overline{S}_{11}, \overline{S}_{22}\right) &= \left[\overline{V}\left(e_1 + e_2\right) - \overline{V}\left(e_1\right) - \overline{V}\left(e_2\right)\right]_{(1,2)} - \mathrm{Cov}_a\left(\overline{S}_{12}, \overline{S}_{12}\right) \\
&= 4\beta\Theta\omega_{12} + 2\alpha\omega_{12}^2.
\end{aligned}$$

The conclusion follows. ∎

Lemma 4 *Let $\Lambda^{\perp}$ denote a symmetric 3×3 matrix such that*

$$\Lambda^{\perp}_{1,1} = 2(1-\alpha)\,\omega_{11}^2$$
$$\Lambda^{\perp}_{1,2} = 2(1-\alpha)\,\omega_{11}\omega_{12}$$
$$\Lambda^{\perp}_{1,3} = 2(1-\alpha)\,\omega_{12}^2$$
$$\Lambda^{\perp}_{2,2} = (1-\alpha)\,\omega_{11}\omega_{22} + (1-\alpha)\,\omega_{12}^2$$
$$\Lambda^{\perp}_{2,3} = 2(1-\alpha)\,\omega_{22}\omega_{12}$$
$$\Lambda^{\perp}_{3,3} = 2(1-\alpha)\,\omega_{22}^2$$

Then,

$$\frac{1}{\sqrt{n}}\left(\begin{pmatrix} S_{11}^{\perp} \\ S_{12}^{\perp} \\ S_{22}^{\perp} \end{pmatrix} - \mathrm{E}\begin{pmatrix} S_{11}^{\perp} \\ S_{12}^{\perp} \\ S_{22}^{\perp} \end{pmatrix}\right) \Rightarrow N\left(0, \Lambda^{\perp}\right).$$

Proof. Similar to the proof of Lemma 3, and omitted. ■

Lemma 5 *Assume that*

$$\frac{K}{n} \to \alpha + o\left(n^{-1/2}\right),$$

and that $\pi_2' z' z \pi_2 / n$ is fixed at Θ. We then have

$$\sqrt{n}\left(\begin{pmatrix} n^{-1}\overline{S}_{11} \\ n^{-1}\overline{S}_{12} \\ n^{-1}\overline{S}_{22} \\ n^{-1}S_{11}^{\perp} \\ n^{-1}S_{12}^{\perp} \\ n^{-1}S_{22}^{\perp} \end{pmatrix} - \begin{pmatrix} \Theta\cdot\beta^2 + \alpha\cdot\omega_{11} \\ \Theta\cdot\beta + \alpha\cdot\omega_{12} \\ \Theta + \alpha\cdot\omega_{22} \\ (1-\alpha)\cdot\omega_{11} \\ (1-\alpha)\cdot\omega_{12} \\ (1-\alpha)\cdot\omega_{22} \end{pmatrix}\right) \Rightarrow \mathcal{N}\left(0, \begin{bmatrix} \overline{\Lambda} & 0 \\ 0 & \Lambda^{\perp} \end{bmatrix}\right).$$

Proof. Using Bekker (1994, Lemma 2), we obtain

$$\mathrm{E}\left[U' P_z U a\right] = (\beta, 1)'\,\pi_2' Z' Z \pi_2\,(\beta, 1)\,a + K\Omega a,$$
$$\mathrm{E}\left[U' M_z U a\right] = (n-K)\,\Omega a.$$

Combining this observation with Lemmas 1, 3, and 4, we obtain the desired conclusion. ■

B Proof of Theorem 2

We note that

$$\lambda = \frac{\alpha}{1-\alpha} + o\left(n^{-1/2}\right)$$

Therefore, without loss of generality, we may assume

$$b = \frac{n^{-1}\overline{S}_{12} - \frac{\alpha}{1-\alpha} n^{-1} S_{12}^{\perp}}{n^{-1}\overline{S}_{22} - \frac{\alpha}{1-\alpha} n^{-1} S_{22}^{\perp}} \quad \text{and} \quad \frac{1}{c} = \frac{n^{-1}\overline{S}_{11} - \frac{\alpha}{1-\alpha} n^{-1} S_{11}^{\perp}}{n^{-1}\overline{S}_{12} - \frac{\alpha}{1-\alpha} n^{-1} S_{12}^{\perp}}$$

and utilize the delta method. It can be easily seen that

$$\sqrt{n}\left(\left(b, \frac{1}{c}\right)' - (\beta, \beta)'\right)$$

is asymptotically normal with mean zero and variance

$$\Upsilon \equiv A \begin{bmatrix} \overline{\Lambda} & 0 \\ 0 & \Lambda^{\perp} \end{bmatrix} A',$$

where

$$A = \begin{bmatrix} 0 & \frac{1}{\Theta} & -\frac{\beta}{\Theta} & 0 & \frac{\alpha}{\Theta(-1+\alpha)} & -\alpha\frac{\beta}{\Theta(-1+\alpha)} \\ \frac{1}{\beta\Theta} & -\frac{1}{\Theta} & 0 & \frac{\alpha}{\beta\Theta(-1+\alpha)} & -\frac{\alpha}{\Theta(-1+\alpha)} & 0 \end{bmatrix}.$$

After some tedious algebra, it can be shown that the (1,1), (1,2), and (2,2)-elements are

$$\begin{aligned}
\Upsilon_{11} &= \frac{\beta^2\omega_{22} - 2\beta\omega_{12} + \omega_{11}}{\Theta} + \frac{\alpha}{1-\alpha}\frac{\omega_{11}\omega_{22} + \omega_{12}^2 - 4\beta\omega_{22}\omega_{12} + 2\beta^2\omega_{22}^2}{\Theta^2}, \\
\Upsilon_{12} &= \frac{\beta^2\omega_{22} - 2\beta\omega_{12} + \omega_{11}}{\Theta} + \frac{\alpha}{1-\alpha}\frac{-\omega_{11}\omega_{22}\beta + 2\omega_{11}\omega_{12} + 2\beta^2\omega_{22}\omega_{12} - 3\omega_{12}^2\beta}{\beta\Theta^2}, \\
\Upsilon_{22} &= \frac{\beta^2\omega_{22} - 2\beta\omega_{12} + \omega_{11}}{\Theta} + \frac{\alpha}{1-\alpha}\frac{\omega_{11}\omega_{22}\beta^2 - 4\omega_{11}\omega_{12}\beta + 2\omega_{11}^2 + \omega_{12}^2\beta^2}{\beta^2\Theta^2}.
\end{aligned}$$

We obtain the desired conclusion from the observation that

$$\begin{aligned}
\operatorname{Var}(\varepsilon) &= \beta^2\omega_{22} - 2\beta\omega_{12} + \omega_{11}, \\
\operatorname{Cov}(\varepsilon, v_1) &= \omega_{11} - \beta\omega_{12}, \\
\operatorname{Cov}(\varepsilon, v_2) &= \omega_{12} - \beta\omega_{22}.
\end{aligned}$$

C Two Endogenous Regressors

We will write

$$\begin{aligned}
y_{1i} &= \beta_2 y_{2i} + \beta_3 y_{3i} + \varepsilon_{1i} \\
y_{2i} &= z_i'\pi_2 + v_{2i} \\
y_{3i} &= z_i'\pi_3 + v_{3i}
\end{aligned}$$

It can be seen that (β_2, β_3) can be estimated by Donald and Newey's (1998) B2SLS applied to

$$y_{1i} = \beta_2 y_{2i} + \beta_3 y_{3i} + \varepsilon_{1i}.$$

We will call such estimator (b_1, c_1). Similarly, $\left(\frac{1}{\beta_2}, -\frac{\beta_3}{\beta_2}\right)$ and $\left(\frac{1}{\beta_3}, -\frac{\beta_2}{\beta_1}\right)$can be estimated by B2SLS applied to

$$y_{2i} = \frac{1}{\beta_2} y_{1i} + \left(-\frac{\beta_3}{\beta_2}\right) y_{3i} + \varepsilon_{2i},$$

and

$$y_{3i} = \frac{1}{\beta_3} y_{1i} + \left(-\frac{\beta_2}{\beta_1}\right) y_{2i} + \varepsilon_{3i}.$$

We will call such estimators (b_2, c_2), and (b_3, c_3).

Note that we have three estimators for (β_2, β_3):

$$(b_1, c_1), \quad \left(\frac{1}{b_2}, -\frac{c_2}{b_2}\right), \quad \left(-\frac{c_3}{b_3}, \frac{1}{b_3}\right).$$

C.1 Technical Lemmas

As before, let $U = M + V$, where M is a fixed, and n rows of V are i.i.d. normal with zero mean and covariance matrix Ω. We examine the first two moments of $U'PU$, where P is an *arbitrary* projection matrix of rank r onto the subspace spanned by the columns of z. Let y_j and m_j denote the jth columns of U and M. $(j = 1, 2, 3)$ We would like to characterize the expectation and variance of

$$S \equiv \begin{pmatrix} y_1' P y_1 \\ y_1' P y_2 \\ y_1' P y_3 \\ y_2' P y_2 \\ y_2' P y_3 \\ y_3' P y_3 \end{pmatrix}$$

Lemma 6 *By Bekker (1994, Lemma 1), we have*

$$E\left[y_i' P y_j\right] = m_i' P m_j + r\omega_{ij}$$

Therefore,

$$E \begin{pmatrix} y_1' P y_1 \\ y_1' P y_2 \\ y_1' P y_3 \\ y_2' P y_2 \\ y_2' P y_3 \\ y_3' P y_3 \end{pmatrix} = \begin{pmatrix} m_1' P m_1 + r\omega_{11} \\ m_1' P m_2 + r\omega_{12} \\ m_1' P m_3 + r\omega_{13} \\ m_2' P m_2 + r\omega_{22} \\ m_2' P m_3 + r\omega_{23} \\ m_3' P m_3 + r\omega_{33} \end{pmatrix}$$

Let Λ denote the variance matrix of S.

Lemma 7 *By Bekker (1994, Lemma 1), we have*

$$\operatorname{Var}\left(y_i' P y_j\right) = \omega_{ii} m_j' P m_j + \omega_{jj} m_i' P m_i + 2\omega_{ij} m_i' P m_j + r\left(\omega_{ii}\omega_{jj} + \omega_{ij}^2\right).$$

Therefore, we have

$$\begin{aligned}
\Lambda_{11} &= 4\omega_{11} m_1' P m_1 + 2r\omega_{11}^2 \\
\Lambda_{22} &= \omega_{11} m_2' P m_2 + \omega_{22} m_1' P m_1 + 2\omega_{12} m_1' P m_2 + r\left(\omega_{11}\omega_{22} + \omega_{12}^2\right) \\
\Lambda_{33} &= \omega_{11} m_3' P m_3 + \omega_{33} m_1' P m_1 + 2\omega_{13} m_1' P m_3 + r\left(\omega_{11}\omega_{33} + \omega_{13}^2\right) \\
\Lambda_{44} &= 4\omega_{22} m_2' P m_2 + 2r\omega_{22}^2 \\
\Lambda_{55} &= \omega_{22} m_3' P m_3 + \omega_{33} m_2' P m_2 + 2\omega_{23} m_2' P m_3 + r\left(\omega_{22}\omega_{33} + \omega_{23}^2\right) \\
\Lambda_{66} &= 4\omega_{33} m_3' P m_3 + 2r\omega_{33}^2
\end{aligned}$$

Lemma 8

$$\operatorname{Cov}\left(y_i' P y_j, y_i' P y_k\right) = \omega_{ii} m_j' P m_k + \omega_{jk} m_i' P m_i + \omega_{ij} m_i' P m_k + \omega_{ik} m_i' P m_j + r\left(\omega_{ii}\omega_{jk} + \omega_{ij}\omega_{ik}\right).$$

Therefore, we have

$$\begin{aligned}
\Lambda_{12} &= \omega_{11} m_1' P m_2 + \omega_{12} m_1' P m_1 + \omega_{11} m_1' P m_2 + \omega_{12} m_1' P m_1 + r\left(\omega_{11}\omega_{12} + \omega_{11}\omega_{12}\right) \\
\Lambda_{13} &= \omega_{11} m_1' P m_3 + \omega_{13} m_1' P m_1 + \omega_{11} m_1' P m_3 + \omega_{13} m_1' P m_1 + r\left(\omega_{11}\omega_{13} + \omega_{11}\omega_{13}\right) \\
\Lambda_{23} &= \omega_{11} m_2' P m_3 + \omega_{23} m_1' P m_1 + \omega_{12} m_1' P m_3 + \omega_{13} m_1' P m_2 + r\left(\omega_{11}\omega_{23} + \omega_{12}\omega_{13}\right) \\
\Lambda_{24} &= \omega_{22} m_2' P m_1 + \omega_{21} m_2' P m_2 + \omega_{22} m_2' P m_1 + \omega_{21} m_2' P m_2 + r\left(\omega_{22}\omega_{21} + \omega_{22}\omega_{21}\right) \\
\Lambda_{25} &= \omega_{22} m_1' P m_3 + \omega_{13} m_2' P m_2 + \omega_{21} m_2' P m_3 + \omega_{23} m_2' P m_1 + r\left(\omega_{22}\omega_{13} + \omega_{21}\omega_{23}\right) \\
\Lambda_{35} &= \omega_{33} m_1' P m_2 + \omega_{12} m_3' P m_3 + \omega_{31} m_3' P m_2 + \omega_{32} m_3' P m_1 + r\left(\omega_{33}\omega_{12} + \omega_{31}\omega_{32}\right) \\
\Lambda_{36} &= \omega_{33} m_3' P m_1 + \omega_{31} m_3' P m_3 + \omega_{33} m_3' P m_1 + \omega_{31} m_3' P m_3 + r\left(\omega_{33}\omega_{31} + \omega_{33}\omega_{31}\right) \\
\Lambda_{45} &= \omega_{22} m_2' P m_3 + \omega_{23} m_2' P m_2 + \omega_{22} m_2' P m_3 + \omega_{23} m_2' P m_2 + r\left(\omega_{22}\omega_{23} + \omega_{22}\omega_{23}\right) \\
\Lambda_{56} &= \omega_{33} m_3' P m_2 + \omega_{32} m_3' P m_3 + \omega_{33} m_3' P m_2 + \omega_{32} m_3' P m_3 + r\left(\omega_{33}\omega_{32} + \omega_{33}\omega_{32}\right)
\end{aligned}$$

Proof. It follows from

$$\begin{aligned}
2\operatorname{Cov}\left(y_i' P y_j, y_i' P y_k\right) &= \operatorname{Var}\left(y_i' P\left(y_j + y_k\right)\right) - \operatorname{Var}\left(y_i' P y_j\right) - \operatorname{Var}\left(y_i' P y_k\right) \\
&= \omega_{ii}\left(m_j + m_k\right)' P\left(m_j + m_k\right) + \left(\omega_{jj} + \omega_{kk} + 2\omega_{jk}\right) m_i' P m_i \\
&\quad + 2\left(\omega_{ij} + \omega_{ik}\right) m_i' P\left(m_j + m_k\right) + r\left(\omega_{ii}\left(\omega_{jj} + \omega_{kk} + 2\omega_{jk}\right) + \left(\omega_{ij} + \omega_{ik}\right)^2\right) \\
&\quad - \left(\omega_{ii} m_j' P m_j + \omega_{jj} m_i' P m_i + 2\omega_{ij} m_i' P m_j + r\left(\omega_{ii}\omega_{jj} + \omega_{ij}^2\right)\right) \\
&\quad - \left(\omega_{ii} m_k' P m_k + \omega_{kk} m_i' P m_i + 2\omega_{ik} m_i' P m_k + r\left(\omega_{ii}\omega_{kk} + \omega_{ik}^2\right)\right) \\
&= 2\omega_{ii} m_j' P m_k + 2\omega_{jk} m_i' P m_i + 2\omega_{ij} m_i' P m_k + 2\omega_{ik} m_i' P m_j + 2r\left(\omega_{ii}\omega_{jk} + \omega_{ij}\omega_{ik}\right)
\end{aligned}$$

■

Lemma 9 *Suppose that $j \neq i$. We then have*

$$\mathrm{Cov}\left(y_i' P y_i, y_j' P y_j\right) = 4\omega_{ij} m_i' P m_j + 2r\omega_{ij}^2$$

Therefore, we have

$$\Lambda_{14} = 4\omega_{12} m_1' P m_2 + 2r\omega_{12}^2$$

$$\Lambda_{16} = 4\omega_{13} m_1' P m_3 + 2r\omega_{13}^2$$

$$\Lambda_{46} = 4\omega_{23} m_2' P m_3 + 2r\omega_{23}^2$$

Proof. Observe that

$$\begin{aligned}\mathrm{Cov}\left(y_i' P y_i, y_j' P y_j\right) = {} & \mathrm{Cov}\left(y_i' P y_i + y_i' P y_j, y_i' P y_j + y_j' P y_j\right) - \mathrm{Cov}\left(y_i' P y_i, y_i' P y_j\right) \\ & - \mathrm{Cov}\left(y_j' P y_j, y_i' P y_j\right) - \mathrm{Var}\left(y_i' P y_j\right).\end{aligned}$$

Also observe that

$$\begin{aligned}& \mathrm{Cov}\left(y_i' P y_i + y_i' P y_j, y_i' P y_j + y_j' P y_j\right) \\ & = \mathrm{Cov}\left(\left(y_i + y_j\right)' P y_i, \left(y_i + y_j\right)' P y_j\right) \\ & = \left(\omega_{ii} + \omega_{jj} + 2\omega_{ij}\right) m_i' P m_j + \omega_{ij}\left(m_i + m_j\right)' P \left(m_i + m_j\right) \\ & + \left(\omega_{ii} + \omega_{ij}\right)\left(m_i + m_j\right)' P m_j + \left(\omega_{jj} + \omega_{ij}\right)\left(m_i + m_j\right)' P m_i \\ & + r\left(\left(\omega_{ii} + \omega_{jj} + 2\omega_{ij}\right)\omega_{ij} + \left(\omega_{ii} + \omega_{ij}\right)\left(\omega_{jj} + \omega_{ij}\right)\right),\end{aligned}$$

where the last line follows from Lemma 8. Because

$$\mathrm{Cov}\left(y_i' P y_i, y_i' P y_j\right) = \omega_{ii} m_i' P m_j + \omega_{ij} m_i' P m_i + \omega_{ii} m_i' P m_j + \omega_{ij} m_i' P m_i + r\left(\omega_{ii}\omega_{ij} + \omega_{ii}\omega_{ij}\right),$$

$$\mathrm{Cov}\left(y_j' P y_j, y_i' P y_j\right) = \omega_{jj} m_j' P m_i + \omega_{ij} m_j' P m_j + \omega_{jj} m_i' P m_j + \omega_{ij} m_j' P m_j + r\left(\omega_{jj}\omega_{ij} + \omega_{jj}\omega_{ij}\right),$$

and

$$\mathrm{Var}\left(y_i' P y_j\right) = \omega_{ii} m_j' P m_j + \omega_{jj} m_i' P m_i + 2\omega_{ij} m_i' P m_j + r\left(\omega_{ii}\omega_{jj} + \omega_{ij}^2\right),$$

we obtain

$$\begin{aligned}& \mathrm{Cov}\left(y_i' P y_i, y_j' P y_j\right) \\ & = \left(\omega_{ij} + \omega_{jj} + \omega_{ij} - \omega_{ij} - \omega_{ij} - \omega_{jj}\right) m_i' P m_i \\ & + \left(\omega_{ii} + \omega_{jj} + 2\omega_{ij} + 2\omega_{ij} + \omega_{ii} + \omega_{ij} + \omega_{jj} + \omega_{ij} - \omega_{ii} - \omega_{ii} - \omega_{jj} - \omega_{jj} - 2\omega_{ij}\right) m_i' P m_j \\ & + \left(\omega_{ij} + \omega_{ii} + \omega_{ij} - \omega_{ij} - \omega_{ij} - \omega_{ii}\right) m_j' P m_j \\ & + r\left(\left(\omega_{ii} + \omega_{jj} + 2\omega_{ij}\right)\omega_{ij} + \left(\omega_{ii} + \omega_{ij}\right)\left(\omega_{jj} + \omega_{ij}\right) - \omega_{ii}\omega_{ij} - \omega_{ii}\omega_{ij} - \omega_{jj}\omega_{ij} - \omega_{jj}\omega_{ij} - \omega_{ii}\omega_{jj} - \omega_{ij}^2\right) \\ & = 4\omega_{ij} m_i' P m_j + 2r\omega_{ij}^2\end{aligned}$$

■

Lemma 10 *Suppose that $j \neq i$, and $k \neq i$. We then have*

$$\mathrm{Cov}\left(y_i' P y_i, y_j' P y_k\right) = 2\omega_{ik} m_i' P m_j + 2\omega_{ij} m_i' P m_k + 2r\omega_{ij}\omega_{ik}.$$

Therefore, we have

$$\Lambda_{15} = 2\omega_{13} m_1' P m_2 + 2\omega_{12} m_1' P m_3 + 2r\omega_{12}\omega_{13},$$
$$\Lambda_{26} = 2\omega_{32} m_3' P m_1 + 2\omega_{31} m_3' P m_2 + 2r\omega_{31}\omega_{32},$$
$$\Lambda_{34} = 2\omega_{23} m_2' P m_1 + 2\omega_{21} m_2' P m_3 + 2r\omega_{21}\omega_{23}.$$

Proof. By normality, we may write

$$y_j = \frac{\omega_{ij}}{\omega_{ii}} y_i + e_j, \quad y_k = \frac{\omega_{ik}}{\omega_{ii}} y_i + e_k,$$

where $y_i \perp e_j, y_i \perp e_k$, and

$$E\left[e_j\right] = m_j - \frac{\omega_{ij}}{\omega_{ii}} m_i, \quad E\left[e_k\right] = m_k - \frac{\omega_{ik}}{\omega_{ii}} m_i.$$

We therefore have

$$\begin{aligned}\mathrm{Cov}\left(y_i' P y_i, y_j' P y_k\right) = \frac{\omega_{ij}}{\omega_{ii}}\frac{\omega_{ik}}{\omega_{ii}} \mathrm{Cov}\left(y_i' P y_i, y_i' P y_i\right) + \frac{\omega_{ik}}{\omega_{ii}} \mathrm{Cov}\left(y_i' P y_i, e_j' P y_i\right) \\ + \frac{\omega_{ij}}{\omega_{ii}} \mathrm{Cov}\left(y_i' P y_i, e_k' P y_i\right) + \mathrm{Cov}\left(y_i' P y_i, e_j' P e_k\right).\end{aligned}$$

Observe that the last term on RHS is zero by independence. Also observe that

$$\mathrm{Cov}\left(y_i' P y_i, y_i' P y_i\right) = 4\omega_{ii} m_i' P m_i + 2r\omega_{ii}^2,$$
$$\mathrm{Cov}\left(y_i' P y_i, y_i' P e_k\right) = 2\omega_{ii} m_i' P\left(m_k - \frac{\omega_{ik}}{\omega_{ii}} m_i\right),$$
$$\mathrm{Cov}\left(y_i' P y_i, y_i' P e_j\right) = 2\omega_{ii} m_i' P\left(m_j - \frac{\omega_{ij}}{\omega_{ii}} m_i\right),$$

where the last two equalities can be deduced by an argument similar to Lemma 8. We therefore have

$$\begin{aligned}\mathrm{Cov}\left(y_i' P y_i, y_j' P y_k\right) &= 4\frac{\omega_{ij}\omega_{ik}}{\omega_{ii}} m_i' P m_i + 2r\omega_{ij}\omega_{ik} \\ &\quad + 2\omega_{ik} m_i' P\left(m_j - \frac{\omega_{ij}}{\omega_{ii}} m_i\right) + 2\omega_{ij} m_i' P\left(m_k - \frac{\omega_{ik}}{\omega_{ii}} m_i\right) \\ &= 2\omega_{ik} m_i' P m_j + 2\omega_{ij} m_i' P m_k + 2r\omega_{ij}\omega_{ik}.\end{aligned}$$

■

Let

$$\begin{bmatrix} \Theta_{22} & \Theta_{23} \\ \Theta_{23} & \Theta_{33} \end{bmatrix} \equiv \mathrm{plim}\,\frac{1}{n}\begin{bmatrix} m_2' \\ m_3' \end{bmatrix} P_z\left[m_2, m_3\right] = \mathrm{plim}\,\frac{1}{n}\begin{bmatrix} \pi_2' z \\ \pi_3' z \end{bmatrix} P_z\left[z\pi_2, z\pi_3\right].$$

Utilizing Bekker (1994, Lemma 2), and the previous results, we can conclude that

$$\sqrt{n}\left(\frac{1}{n}\begin{pmatrix} y_1'P_zy_1 \\ y_1'P_zy_2 \\ y_1'P_zy_3 \\ y_2'P_zy_2 \\ y_2'P_zy_3 \\ y_3'P_zy_3 \end{pmatrix} - \begin{pmatrix} (\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33})+\alpha\omega_{11} \\ (\beta_2\Theta_{22}+\beta_3\Theta_{23})+\alpha\omega_{12} \\ (\beta_2\Theta_{23}+\beta_3\Theta_{33})+\alpha\omega_{13} \\ \Theta_{22}+\alpha\omega_{22} \\ \Theta_{23}+\alpha\omega_{23} \\ \Theta_{33}+\alpha\omega_{33} \end{pmatrix}\right)$$

and

$$\sqrt{n}\left(\frac{1}{n}\begin{pmatrix} y_1'M_zy_1 \\ y_1'M_zy_2 \\ y_1'M_zy_3 \\ y_2'M_zy_2 \\ y_2'M_zy_3 \\ y_3'M_zy_3 \end{pmatrix} - \begin{pmatrix} (1-\alpha)\,\omega_{11} \\ (1-\alpha)\,\omega_{12} \\ (1-\alpha)\,\omega_{13} \\ (1-\alpha)\,\omega_{22} \\ (1-\alpha)\,\omega_{23} \\ (1-\alpha)\,\omega_{33} \end{pmatrix}\right)$$

are independent of each other, and asymptotically normal with zero mean and variances equal to $\overline{\Lambda}$ and $\Lambda^{\perp}$, where

$$\overline{\Lambda}_{11} = 4\omega_{11}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+2\alpha\omega_{11}^2$$

$$\begin{aligned}\overline{\Lambda}_{12} &= \omega_{11}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+\omega_{12}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+\omega_{11}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)\\ &\quad+\omega_{12}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+\alpha\left(\omega_{11}\omega_{12}+\omega_{11}\omega_{12}\right)\end{aligned}$$

$$\begin{aligned}\overline{\Lambda}_{13} &= \omega_{11}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+\omega_{13}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+\omega_{11}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)\\ &\quad+\omega_{13}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+\alpha\left(\omega_{11}\omega_{13}+\omega_{11}\omega_{13}\right)\end{aligned}$$

$$\overline{\Lambda}_{14} = 4\omega_{12}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+2\alpha\omega_{12}^2$$

$$\overline{\Lambda}_{15} = 2\omega_{13}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+2\omega_{12}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+2\alpha\omega_{12}\omega_{13}$$

$$\overline{\Lambda}_{16} = 4\omega_{13}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+2\alpha\omega_{13}^2$$

$$\overline{\Lambda}_{22} = \omega_{11}\Theta_{22}+\omega_{22}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+2\omega_{12}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+\alpha\left(\omega_{11}\omega_{22}+\omega_{12}^2\right)$$

$$\begin{aligned}\overline{\Lambda}_{23} &= \omega_{11}\Theta_{23}+\omega_{23}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+\omega_{12}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+\omega_{13}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)\\ &\quad+\alpha\left(\omega_{11}\omega_{23}+\omega_{12}\omega_{13}\right)\end{aligned}$$

$$\overline{\Lambda}_{24} = \omega_{22}\left(\beta_2\Theta_{22} + \beta_3\Theta_{23}\right) + \omega_{12}\Theta_{22} + \omega_{22}\left(\beta_2\Theta_{22} + \beta_3\Theta_{23}\right) + \omega_{12}\Theta_{22} + \alpha\left(\omega_{22}\omega_{12} + \omega_{22}\omega_{12}\right)$$

$$\overline{\Lambda}_{25} = \omega_{22}\left(\beta_2\Theta_{23} + \beta_3\Theta_{33}\right) + \omega_{13}\Theta_{22} + \omega_{12}\Theta_{23} + \omega_{23}\left(\beta_2\Theta_{22} + \beta_3\Theta_{23}\right) + \alpha\left(\omega_{22}\omega_{13} + \omega_{12}\omega_{23}\right)$$

$$\overline{\Lambda}_{26} = 2\omega_{23}\left(\beta_2\Theta_{23} + \beta_3\Theta_{33}\right) + 2\omega_{13}\Theta_{23} + 2\alpha\omega_{13}\omega_{23}$$

$$\overline{\Lambda}_{33} = \omega_{11}\Theta_{33} + \omega_{33}\left(\beta_2^2\Theta_{22} + 2\beta_2\beta_3\Theta_{23} + \beta_3^2\Theta_{33}\right) + 2\omega_{13}\left(\beta_2\Theta_{23} + \beta_3\Theta_{33}\right) + \alpha\left(\omega_{11}\omega_{33} + \omega_{13}^2\right)$$

$$\overline{\Lambda}_{34} = 2\omega_{23}\left(\beta_2\Theta_{22} + \beta_3\Theta_{23}\right) + 2\omega_{12}\Theta_{23} + 2\alpha\omega_{12}\omega_{23}$$

$$\overline{\Lambda}_{35} = \omega_{33}\left(\beta_2\Theta_{22} + \beta_3\Theta_{23}\right) + \omega_{12}\Theta_{33} + \omega_{13}\Theta_{23} + \omega_{23}\left(\beta_2\Theta_{23} + \beta_3\Theta_{33}\right) + \alpha\left(\omega_{33}\omega_{12} + \omega_{13}\omega_{23}\right)$$

$$\overline{\Lambda}_{36} = \omega_{33}\left(\beta_2\Theta_{23} + \beta_3\Theta_{33}\right) + \omega_{13}\Theta_{33} + \omega_{33}\left(\beta_2\Theta_{23} + \beta_3\Theta_{33}\right) + \omega_{13}\Theta_{33} + \alpha\left(\omega_{33}\omega_{13} + \omega_{33}\omega_{13}\right)$$

$$\overline{\Lambda}_{44} = 4\omega_{22}\Theta_{22} + 2\alpha\omega_{22}^2$$

$$\overline{\Lambda}_{45} = \omega_{22}\Theta_{23} + \omega_{23}\Theta_{22} + \omega_{22}\Theta_{23} + \omega_{23}\Theta_{22} + \alpha\left(\omega_{22}\omega_{23} + \omega_{22}\omega_{23}\right)$$

$$\overline{\Lambda}_{46} = 4\omega_{23}\Theta_{23} + 2\alpha\omega_{23}^2$$

$$\overline{\Lambda}_{55} = \omega_{22}\Theta_{33} + \omega_{33}\Theta_{22} + 2\omega_{23}\Theta_{23} + \alpha\left(\omega_{22}\omega_{33} + \omega_{23}^2\right)$$

$$\overline{\Lambda}_{56} = \omega_{33}\Theta_{23} + \omega_{23}\Theta_{33} + \omega_{33}\Theta_{23} + \omega_{23}\Theta_{33} + \alpha\left(\omega_{33}\omega_{23} + \omega_{33}\omega_{23}\right)$$

$$\overline{\Lambda}_{66} = 4\omega_{33}\Theta_{33} + 2\alpha\omega_{33}^2$$

$$\Lambda_{11}^{\perp} = 2\left(1-\alpha\right)\omega_{11}^2$$

$$\Lambda_{12}^{\perp} = \left(1-\alpha\right)\left(\omega_{11}\omega_{12} + \omega_{11}\omega_{12}\right)$$

$$\Lambda_{13}^{\perp} = \left(1-\alpha\right)\left(\omega_{11}\omega_{13} + \omega_{11}\omega_{13}\right)$$

$$\Lambda_{14}^{\perp} = 2\left(1-\alpha\right)\omega_{12}^2$$

$$\Lambda_{15}^{\perp} = 2\left(1-\alpha\right)\omega_{12}\omega_{13}$$

$$\Lambda_{16}^{\perp} = 2\,(1-\alpha)\,\omega_{13}^2$$

$$\Lambda_{22}^{\perp} = (1-\alpha)\left(\omega_{11}\omega_{22} + \omega_{12}^2\right)$$

$$\Lambda_{23}^{\perp} = (1-\alpha)\,(\omega_{11}\omega_{23} + \omega_{12}\omega_{13})$$

$$\Lambda_{24}^{\perp} = (1-\alpha)\,(\omega_{22}\omega_{12} + \omega_{22}\omega_{12})$$

$$\Lambda_{25}^{\perp} = (1-\alpha)\,(\omega_{22}\omega_{13} + \omega_{12}\omega_{23})$$

$$\Lambda_{26}^{\perp} = 2\,(1-\alpha)\,\omega_{13}\omega_{23}$$

$$\Lambda_{33}^{\perp} = (1-\alpha)\left(\omega_{11}\omega_{33} + \omega_{13}^2\right)$$

$$\Lambda_{34}^{\perp} = 2\,(1-\alpha)\,\omega_{12}\omega_{23}$$

$$\Lambda_{35}^{\perp} = (1-\alpha)\,(\omega_{33}\omega_{12} + \omega_{13}\omega_{23})$$

$$\Lambda_{36}^{\perp} = (1-\alpha)\,(\omega_{33}\omega_{13} + \omega_{33}\omega_{13})$$

$$\Lambda_{44}^{\perp} = 2\,(1-\alpha)\,\omega_{22}^2$$

$$\Lambda_{45}^{\perp} = (1-\alpha)\,(\omega_{22}\omega_{23} + \omega_{22}\omega_{23})$$

$$\Lambda_{46}^{\perp} = 2\,(1-\alpha)\,\omega_{23}^2$$

$$\Lambda_{55}^{\perp} = (1-\alpha)\left(\omega_{22}\omega_{33} + \omega_{23}^2\right)$$

$$\Lambda_{56}^{\perp} = (1-\alpha)\,(\omega_{33}\omega_{23} + \omega_{33}\omega_{23})$$

$$\Lambda_{66}^{\perp} = 2\,(1-\alpha)\,\omega_{33}^2$$

D Application: Nagar-Type Estimator

Applying the delta method, we can find that

$$\sqrt{n}\left(b_1 - \frac{1}{b_2}, c_1 - -\frac{c_2}{b_2}, b_1 - -\frac{c_3}{b_3}, c_1 - \frac{1}{b_3}\right)'$$

is asymptotically normal with zero mean and variance equal to

$$\frac{2\alpha \operatorname{Var}(\varepsilon_{1i})^2}{1-\alpha}$$

times

$$\begin{bmatrix} \frac{\Theta_{33}{}^2}{\beta_2{}^2(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & -\frac{\Theta_{23}\Theta_{33}}{\beta_2{}^2(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & -\frac{\Theta_{23}\Theta_{33}}{\beta_2\beta_3(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & \frac{\Theta_{22}\Theta_{33}}{\beta_2\beta_3(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} \\ -\frac{\Theta_{23}\Theta_{33}}{\beta_2{}^2(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & \frac{\Theta_{23}{}^2}{\beta_2{}^2(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & \frac{\Theta_{23}{}^2}{\beta_2\beta_3(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & -\frac{\Theta_{22}\Theta_{23}}{\beta_2\beta_3(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} \\ -\frac{\Theta_{23}\Theta_{33}}{\beta_2\beta_3(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & \frac{\Theta_{23}{}^2}{\beta_2\beta_3(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & \frac{\Theta_{23}{}^2}{\beta_3{}^2(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & -\frac{\Theta_{22}\Theta_{23}}{\beta_3{}^2(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} \\ \frac{\Theta_{22}\Theta_{33}}{\beta_2\beta_3(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & -\frac{\Theta_{22}\Theta_{23}}{\beta_2\beta_3(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & -\frac{\Theta_{22}\Theta_{23}}{\beta_3{}^2(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} & \frac{\Theta_{22}{}^2}{\beta_3{}^2(\Theta_{22}\Theta_{33}-\Theta_{23}{}^2)^2} \end{bmatrix}.$$

With some tedious algebra, it can be shown that the above asymptotic variance matrix is singular: Postmultiplying the asymptotic variance by

$$\left(\frac{\Theta_{23}}{\Theta_{33}}, 1, 0, 0\right)', \quad \left(-\frac{\beta_2\Theta_{22}}{\beta_3\Theta_{33}}, 0, 0, 1\right)', \quad \text{or} \quad \left(\frac{\beta_2\Theta_{23}}{\beta_3\Theta_{33}}, 0, 1, 0\right)',$$

we obtain zero. Therefore, we cannot stack the estimates to derive a more efficient test since all tests based on a single difference will have the same operating characteristics. This implies that the test can be applied only to one component of

$$\left(b_1 - \frac{1}{b_2}, c_1 - -\frac{c_2}{b_2}, b_1 - -\frac{c_3}{b_3}, c_1 - \frac{1}{b_3}\right)',$$

say $b_1 - \frac{1}{b_2}$. It is to be noted that the asymptotic variance of such a test is given by

$$\frac{2\alpha \operatorname{Var}(\varepsilon_{1i})^2}{1-\alpha} \frac{1}{\beta_2{}^2\left(\Theta_{22} - \frac{\Theta_{23}{}^2}{\Theta_{33}}\right)^2}.$$

Observe that $\operatorname{Var}(\varepsilon_{1i})$ and β_2 can be estimated consistently utilizing the consistency of LIML. Also note that

$$\operatorname{plim}\begin{bmatrix} \widehat{\Theta}_{22} & \widehat{\Theta}_{23} \\ \widehat{\Theta}_{23} & \widehat{\Theta}_{33} \end{bmatrix} \equiv \operatorname{plim}\left(\frac{1}{n-2}\begin{bmatrix} y_2' \\ y_3' \end{bmatrix} P_z[y_2, y_3] - \frac{\widehat{\alpha}}{1-\widehat{\alpha}}\frac{1}{n-2}\begin{bmatrix} y_2' \\ y_3' \end{bmatrix} M_z[y_2, y_3]\right) = \begin{bmatrix} \Theta_{22} & \Theta_{23} \\ \Theta_{23} & \Theta_{33} \end{bmatrix}$$

for any consistent estimator $\widehat{\alpha}$ of α. We may therefore estimate the asymptotic variance consistently by

$$2\frac{K-1}{n-K}\frac{\left(\sum_{i=1}^{n}(y_{1i} - \beta_{2LIML}y_{2i} - \beta_{3LIML}y_{3i})^2\right)^2}{\beta_{2LIML}^2\left(y_2'P_zy_2 - \frac{K-1}{n-K}y_2'M_zy_2 - \frac{\left(y_2'P_zy_3 - \frac{K-1}{n-K}y_2'M_zy_3\right)^2}{y_3'P_zy_3 - \frac{K-1}{n-K}y_3'M_zy_3}\right)^2}.$$

Table 1: $\widetilde{R}_f^2 = .1$

n	K	$Cov(\varepsilon, v_2)$	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
100	5	-0.9	0.086	0.050	-0.223	-0.041	-0.178	0.235	0.166	0.074	0.045	0.027	0.081	0.096	0.154	0.104
100	5	-0.5	0.084	0.054	-0.123	0.312	-0.398	0.116	0.056	0.073	0.048	0.106	0.058	0.206	0.096	0.045
100	5	0.5	0.079	0.053	0.125	-0.303	0.396	0.114	0.058	0.072	0.050	-0.399	-0.220	0.021	0.096	0.048
100	5	0.9	0.078	0.045	0.219	0.027	0.502	0.233	0.161	0.067	0.041	0.080	-0.061	-0.158	0.153	0.100
100	10	-0.9	0.108	0.078	-0.405	-0.158	-0.258	0.369	0.275	0.067	0.049	-0.046	0.140	-0.094	0.172	0.116
100	10	-0.5	0.079	0.049	-0.228	0.417	-0.630	0.117	0.059	0.070	0.045	0.168	0.520	0.074	0.082	0.040
100	10	0.5	0.078	0.043	0.223	-0.430	0.637	0.124	0.064	0.066	0.042	0.184	1.217	-0.039	0.085	0.043
100	10	0.9	0.099	0.070	0.405	0.163	0.248	0.360	0.266	0.064	0.043	-0.097	-0.175	-0.107	0.164	0.115
100	30	-0.9	0.137	0.112	-0.656	-0.384	-0.277	0.548	0.382	0.111	0.095	-0.361	0.135	0.544	0.180	0.113
100	30	-0.5	0.110	0.081	-0.361	0.662	-1.035	0.098	0.036	0.083	0.061	0.870	0.116	-216.133	0.053	0.017
100	30	0.5	0.103	0.071	0.365	-1.241	1.691	0.101	0.039	0.076	0.058	0.038	0.133	-0.902	0.056	0.020
100	30	0.9	0.142	0.115	0.655	0.380	0.280	0.539	0.383	0.113	0.094	-1.737	-0.232	-0.344	0.176	0.107
250	5	-0.9	0.108	0.074	-0.095	-0.014	-0.065	0.169	0.108	0.090	0.057	0.011	0.028	0.043	0.130	0.075
250	5	-0.5	0.112	0.077	-0.055	0.094	-0.118	0.107	0.056	0.094	0.065	0.004	0.041	0.021	0.097	0.050
250	5	0.5	0.117	0.081	0.052	-0.098	0.117	0.115	0.061	0.098	0.068	-0.007	-0.046	-0.027	0.103	0.055
250	5	0.9	0.105	0.071	0.094	0.016	0.064	0.162	0.100	0.085	0.057	-0.011	-0.025	-0.038	0.120	0.072
250	10	-0.9	0.084	0.039	-0.212	-0.074	-0.132	0.249	0.172	0.068	0.039	0.040	0.027	0.037	0.132	0.079
250	10	-0.5	0.079	0.045	-0.112	0.177	-0.269	0.118	0.062	0.076	0.049	0.019	0.052	0.036	0.092	0.045
250	10	0.5	0.079	0.046	0.118	-0.177	0.275	0.121	0.063	0.075	0.052	-0.021	-0.051	-0.040	0.090	0.048
250	10	0.9	0.081	0.037	0.209	0.067	0.137	0.250	0.173	0.063	0.035	-0.043	-0.038	-0.048	0.127	0.070
250	30	-0.9	0.101	0.069	-0.460	-0.223	-0.239	0.570	0.463	0.065	0.030	-0.163	0.156	0.056	0.130	0.087
250	30	-0.5	0.090	0.051	-0.256	0.389	-0.640	0.145	0.068	0.063	0.032	0.067	0.095	0.040	0.080	0.038
250	30	0.5	0.087	0.048	0.255	-0.385	0.637	0.137	0.070	0.066	0.037	-0.249	-0.054	0.014	0.080	0.039
250	30	0.9	0.103	0.069	0.460	0.224	0.238	0.565	0.464	0.062	0.030	-0.207	-0.050	-0.052	0.141	0.084

(a), (b): Actual sizes of the new test based on 2SLS with nominal sizes = 10%, and 5%
(c), (d), (e): Mean biases of forward and reverse 2SLS, and mean of $\hat{B}$
(f), (g): Actual sizes of the tests based on nR^2 of the residual of forward 2SLS with nominal sizes = 10%, and 5%
(h), (i): Aetual size of the new test based on Nagar with nominal size = 10%, and 5%
(j), (k), (l): Mean biases of forward Nagar, reverse Nagar, and LIML
(m), (n): Actual sizes of the tests based on nR^2 of the residual of forward Nagar with nominal sizes = 10%, and 5%

The reported numbers are based on 5000 Monte Carlo replications.

Table 1 (Cont.): $\widetilde{R}_f^2 = .1$

n	K	Cov(ε, v_2)	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
1,000	5	-0.9	0.131	0.095	-0.026	-0.006	-0.015	0.118	0.061	0.102	0.068	-0.002	0.003	0.007	0.108	0.055
1,000	5	-0.5	0.131	0.096	-0.016	0.020	-0.027	0.106	0.057	0.105	0.078	-0.002	0.007	0.002	0.102	0.054
1,000	5	0.5	0.137	0.099	0.016	-0.021	0.027	0.107	0.053	0.108	0.073	0.002	-0.007	-0.002	0.104	0.050
1,000	5	0.9	0.126	0.094	0.026	0.006	0.015	0.115	0.067	0.100	0.074	0.002	-0.003	-0.007	0.105	0.059
1,000	10	-0.9	0.097	0.064	-0.062	-0.020	-0.038	0.138	0.083	0.085	0.058	0.001	0.005	0.008	0.106	0.062
1,000	10	-0.5	0.106	0.071	-0.035	0.043	-0.071	0.118	0.063	0.097	0.066	-0.001	0.008	0.004	0.105	0.056
1,000	10	0.5	0.103	0.070	0.035	-0.043	0.071	0.113	0.062	0.093	0.064	0.000	-0.008	-0.004	0.101	0.054
1,000	10	0.9	0.106	0.065	0.063	0.021	0.038	0.154	0.089	0.093	0.059	0.001	-0.004	-0.007	0.117	0.062
1,000	30	-0.9	0.094	0.046	-0.183	-0.073	-0.108	0.306	0.214	0.089	0.042	0.004	0.007	0.010	0.117	0.066
1,000	30	-0.5	0.097	0.052	-0.102	0.124	-0.219	0.145	0.074	0.090	0.052	0.003	0.012	0.007	0.098	0.052
1,000	30	0.5	0.097	0.049	0.101	-0.124	0.219	0.145	0.078	0.094	0.051	-0.003	-0.012	-0.007	0.101	0.051
1,000	30	0.9	0.098	0.048	0.183	0.073	0.108	0.309	0.215	0.091	0.047	-0.004	-0.006	-0.009	0.116	0.062
10,000	5	-0.9	0.135	0.096	-0.002	0.000	-0.002	0.099	0.053	0.102	0.075	0.000	0.001	0.001	0.098	0.052
10,000	5	-0.5	0.124	0.093	-0.001	0.003	-0.003	0.095	0.047	0.098	0.068	0.001	0.001	0.001	0.094	0.047
10,000	5	0.5	0.134	0.100	0.001	-0.003	0.003	0.102	0.049	0.106	0.076	0.000	-0.001	-0.001	0.102	0.049
10,000	5	0.9	0.125	0.092	0.002	0.000	0.002	0.095	0.045	0.097	0.066	-0.001	-0.001	-0.001	0.094	0.045
10,000	10	-0.9	0.101	0.065	-0.007	-0.003	-0.004	0.102	0.051	0.088	0.056	-0.001	0.000	0.000	0.096	0.048
10,000	10	-0.5	0.115	0.075	-0.003	0.005	-0.007	0.108	0.053	0.100	0.064	0.000	0.001	0.001	0.107	0.052
10,000	10	0.5	0.103	0.067	0.004	-0.004	0.007	0.098	0.050	0.089	0.058	0.001	0.000	0.000	0.097	0.049
10,000	10	0.9	0.111	0.073	0.006	0.002	0.004	0.109	0.056	0.097	0.064	0.000	-0.001	-0.001	0.107	0.053
10,000	30	-0.9	0.103	0.055	-0.022	-0.008	-0.014	0.122	0.066	0.100	0.051	0.000	0.000	0.001	0.102	0.052
10,000	30	-0.5	0.096	0.057	-0.012	0.013	-0.025	0.097	0.054	0.092	0.055	0.000	0.001	0.000	0.094	0.051
10,000	30	0.5	0.099	0.054	0.013	-0.013	0.025	0.109	0.053	0.096	0.053	0.000	-0.001	0.000	0.102	0.050
10,000	30	0.9	0.099	0.054	0.022	0.008	0.014	0.118	0.062	0.094	0.051	0.000	-0.001	-0.001	0.097	0.050

(a), (b): Actual sizes of the new test based on 2SLS with nominal sizes = 10%, and 5%
(c), (d), (e): Mean biases of forward and reverse 2SLS, and mean of $\hat{B}$
(f), (g): Actual sizes of the tests based on nR^2 of the residual of forward 2SLS with nominal sizes = 10%, and 5%
(h), (i): Actual size of the new test based on Nagar with nominal size = 10%, and 5%
(j), (k), (l): Mean biases of forward Nagar, reverse Nagar, and LIML
(m), (n): Actual sizes of the tests based on nR^2 of the residual of forward Nagar with nominal sizes = 10%, and 5%

The reported numbers are based on 5000 Monte Carlo replications.

Table 2: $\widetilde{R}_f^2 = .01$

n	K	Cov(ε,v₂)	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
100	5	-0.9	0.082	0.061	-0.745	-0.062	-1.748	0.213	0.134	0.119	0.108	4.579	-0.179	-0.551	0.151	0.092
100	5	-0.5	0.067	0.052	-0.407	-5.226	3.170	0.080	0.040	0.106	0.084	-0.504	0.785	1.807	0.057	0.026
100	5	0.5	0.068	0.051	0.410	-0.157	-1.143	0.079	0.034	0.110	0.086	1.073	-0.286	1.127	0.054	0.023
100	5	0.9	0.087	0.066	0.739	0.253	1.207	0.211	0.131	0.123	0.110	1.780	0.569	1.211	0.148	0.089
100	10	-0.9	0.089	0.070	-0.819	-0.526	-0.358	0.208	0.122	0.117	0.107	-0.681	-0.555	1.445	0.129	0.071
100	10	-0.5	0.085	0.062	-0.457	5.292	-7.591	0.077	0.035	0.112	0.089	-0.380	-1.332	-0.769	0.050	0.022
100	10	0.5	0.092	0.069	0.455	-1.742	0.131	0.086	0.038	0.107	0.088	0.337	-0.414	0.435	0.049	0.018
100	10	0.9	0.087	0.066	0.816	0.525	0.353	0.213	0.123	0.116	0.104	1.094	0.618	0.994	0.125	0.064
100	30	-0.9	0.110	0.089	-0.871	-0.637	-0.245	0.122	0.046	0.125	0.112	-1.848	-8.336	-9.463	0.060	0.023
100	30	-0.5	0.130	0.102	-0.480	1.124	-1.652	0.065	0.021	0.128	0.102	-1.012	3.453	0.667	0.026	0.009
100	30	0.5	0.119	0.094	0.485	-1.094	1.682	0.065	0.023	0.129	0.101	23.277	-12.186	0.799	0.035	0.008
100	30	0.9	0.107	0.086	0.869	0.633	0.247	0.126	0.052	0.130	0.116	0.935	0.059	3.031	0.060	0.022
250	5	-0.9	0.093	0.066	-0.581	0.720	-6.404	0.341	0.251	0.111	0.096	-0.476	-0.936	0.163	0.231	0.165
250	5	-0.5	0.068	0.049	-0.328	0.971	-1.427	0.091	0.046	0.084	0.062	-0.323	-2.610	0.065	0.067	0.034
250	5	0.5	0.057	0.042	0.327	-0.261	0.053	0.096	0.044	0.085	0.066	0.366	0.593	0.191	0.069	0.032
250	5	0.9	0.086	0.066	0.572	0.205	0.466	0.326	0.240	0.115	0.102	-13.996	-0.119	0.233	0.225	0.156
250	10	-0.9	0.112	0.086	-0.709	-0.401	-0.371	0.377	0.264	0.137	0.124	-0.504	0.169	-0.468	0.214	0.143
250	10	-0.5	0.088	0.066	-0.389	2.396	-3.260	0.091	0.044	0.096	0.073	0.187	0.319	-0.282	0.057	0.027
250	10	0.5	0.090	0.065	0.396	-2.142	3.312	0.095	0.049	0.106	0.083	0.431	0.141	2.101	0.058	0.030
250	10	0.9	0.116	0.095	0.716	0.401	0.379	0.372	0.267	0.144	0.131	0.129	0.220	0.784	0.210	0.139
250	30	-0.9	0.109	0.091	-0.830	-0.576	-0.266	0.333	0.207	0.145	0.130	-0.866	0.005	72.662	0.167	0.094
250	30	-0.5	0.123	0.094	-0.462	1.169	-1.729	0.089	0.037	0.116	0.095	1.454	2.040	-1.142	0.041	0.014
250	30	0.5	0.115	0.088	0.461	-1.126	1.661	0.080	0.036	0.105	0.082	-2.396	-0.128	0.354	0.043	0.017
250	30	0.9	0.113	0.096	0.830	0.577	0.264	0.322	0.208	0.146	0.134	0.806	1.358	1.300	0.159	0.094

(a), (b): Actual sizes of the new test based on 2SLS with nominal sizes = 10%, and 5%

(c), (d), (e): Mean biases of forward and reverse 2SLS, and mean of $\hat{B}$

(f), (g): Actual sizes of the tests based on nR^2 of the residual of forward 2SLS with nominal sizes = 10%, and 5%

(h), (i): Actual size of the new test based on Nagar with nominal size = 10%, and 5%

(j), (k), (l): Mean biases of forward Nagar, reverse Nagar, and LIML

(m), (n): Actual sizes of the tests based on nR^2 of the residual of forward Nagar with nominal sizes = 10%, and 5%

The reported numbers are based on 5000 Monte Carlo replications.

Table 2 (Cont.): $\widetilde{R}_f^2 = .01$

n	K	Cov(ε,v_2)	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
1,000	5	-0.9	0.075	0.043	-0.242	-0.043	-0.209	0.248	0.177	0.064	0.040	-0.096	0.064	0.115	0.156	0.104
1,000	5	-0.5	0.084	0.054	-0.141	0.301	-0.403	0.112	0.059	0.074	0.053	0.298	0.157	-0.409	0.095	0.047
1,000	5	0.5	0.087	0.055	0.139	-0.287	0.380	0.120	0.062	0.079	0.052	0.635	-0.150	-0.006	0.100	0.053
1,000	5	0.9	0.087	0.050	0.240	0.045	0.205	0.244	0.177	0.069	0.044	-0.067	-0.198	-0.347	0.154	0.103
1,000	10	-0.9	0.116	0.079	-0.424	-0.174	-0.267	0.388	0.304	0.073	0.052	0.266	0.222	0.309	0.165	0.120
1,000	10	-0.5	0.084	0.053	-0.237	0.514	-0.764	0.129	0.069	0.067	0.046	-0.103	-0.324	0.213	0.091	0.046
1,000	10	0.5	0.083	0.053	0.235	-0.384	0.584	0.127	0.075	0.069	0.046	6.848	-0.151	-0.180	0.094	0.051
1,000	10	0.9	0.107	0.072	0.426	0.176	0.267	0.395	0.311	0.072	0.049	0.789	-0.068	-0.261	0.181	0.128
1,000	30	-0.9	0.155	0.134	-0.671	-0.396	-0.285	0.678	0.575	0.116	0.096	-0.519	0.003	-2.470	0.203	0.153
1,000	30	-0.5	0.109	0.077	-0.372	0.702	-1.090	0.138	0.074	0.071	0.053	-0.135	-3.483	0.766	0.069	0.035
1,000	30	0.5	0.114	0.078	0.372	-0.724	1.119	0.136	0.072	0.069	0.049	-0.098	-0.254	-1.195	0.077	0.039
1,000	30	0.9	0.159	0.135	0.670	0.394	0.287	0.678	0.579	0.122	0.100	0.213	-0.055	-0.121	0.212	0.159
10,000	5	-0.9	0.127	0.089	-0.026	-0.004	-0.017	0.118	0.066	0.097	0.070	0.001	0.007	0.010	0.103	0.058
10,000	5	-0.5	0.119	0.088	-0.013	0.027	-0.030	0.096	0.048	0.096	0.066	0.002	0.012	0.007	0.093	0.046
10,000	5	0.5	0.128	0.096	0.013	-0.027	0.030	0.104	0.053	0.105	0.074	-0.002	-0.012	-0.007	0.101	0.050
10,000	5	0.9	0.119	0.085	0.024	0.002	0.017	0.111	0.057	0.092	0.062	-0.003	-0.008	-0.011	0.099	0.050
10,000	10	-0.9	0.095	0.056	-0.069	-0.023	-0.042	0.146	0.084	0.082	0.052	0.000	0.004	0.007	0.107	0.056
10,000	10	-0.5	0.108	0.069	-0.037	0.049	-0.077	0.114	0.061	0.095	0.064	0.002	0.011	0.006	0.105	0.054
10,000	10	0.5	0.096	0.065	0.039	-0.047	0.078	0.107	0.058	0.084	0.059	0.001	-0.009	-0.004	0.096	0.050
10,000	10	0.9	0.103	0.061	0.067	0.021	0.042	0.153	0.090	0.088	0.056	-0.002	-0.006	-0.010	0.112	0.062
10,000	30	-0.9	0.096	0.045	-0.199	-0.081	-0.116	0.337	0.239	0.092	0.044	0.004	0.006	0.009	0.119	0.066
10,000	30	-0.5	0.096	0.052	-0.110	0.132	-0.237	0.136	0.076	0.087	0.052	0.002	0.010	0.005	0.092	0.053
10,000	30	0.5	0.097	0.049	0.111	-0.133	0.238	0.147	0.081	0.093	0.052	-0.002	-0.011	-0.006	0.101	0.052
10,000	30	0.9	0.095	0.049	0.197	0.080	0.116	0.320	0.228	0.088	0.048	-0.006	-0.007	-0.010	0.114	0.063

(a), (b): Actual sizes of the new test based on 2SLS with nominal sizes = 10%, and 5%
(c), (d), (e): Mean biases of forward and reverse 2SLS, and mean of $\hat{B}$
(f), (g): Actual sizes of the tests based on nR^2 of the residual of forward 2SLS with nominal sizes = 10%, and 5%
(h), (i): Actual size of the new test based on Nagar with nominal size = 10%, and 5%
(j), (k), (l): Mean biases of forward Nagar, reverse Nagar, and LIML
(m), (n): Actual sizes of the tests based on nR^2 of the residual of forward Nagar with nominal sizes = 10%, and 5%

The reported numbers are based on 5000 Monte Carlo replications.

Table 3: $\widetilde{R}_f^2 = .001$

n	K	Cov(ε, v_2)	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
100	5	-0.9	0.079	0.062	-0.884	-0.359	-2.085	0.079	0.039	0.112	0.100	-3.437	-0.720	-2.111	0.062	0.029
100	5	-0.5	0.077	0.061	-0.484	-0.528	4.529	0.067	0.033	0.117	0.092	-0.757	1.651	-5.630	0.049	0.021
100	5	0.5	0.075	0.057	0.489	-0.579	2.218	0.063	0.027	0.113	0.089	0.851	1.464	6.667	0.045	0.019
100	5	0.9	0.084	0.068	0.877	0.603	0.249	0.085	0.041	0.113	0.102	1.262	-168.415	0.834	0.060	0.029
100	10	-0.9	0.097	0.077	-0.892	-0.640	-0.302	0.080	0.037	0.120	0.107	-1.314	-1.118	-1.157	0.046	0.020
100	10	-0.5	0.091	0.070	-0.497	3.495	-4.726	0.068	0.031	0.119	0.096	0.108	1.644	-0.008	0.046	0.019
100	10	0.5	0.097	0.076	0.498	-3.618	4.631	0.072	0.034	0.113	0.090	0.115	-3.987	-1.273	0.042	0.018
100	10	0.9	0.090	0.070	0.892	0.635	0.313	0.089	0.041	0.112	0.098	0.859	0.796	0.046	0.053	0.024
100	30	-0.9	0.127	0.103	-0.897	-0.677	-0.232	0.060	0.021	0.128	0.114	-0.269	-1.310	-0.918	0.031	0.008
100	30	-0.5	0.128	0.102	-0.495	1.667	-2.498	0.062	0.020	0.127	0.104	1.516	-0.945	0.209	0.027	0.009
100	30	0.5	0.116	0.089	0.499	-1.556	2.241	0.063	0.022	0.125	0.103	0.047	0.023	-1.076	0.031	0.007
100	30	0.9	0.121	0.098	0.895	0.672	0.234	0.067	0.025	0.131	0.111	1.589	2.212	0.968	0.035	0.012
250	5	-0.9	0.073	0.057	-0.860	-1.063	11.367	0.114	0.061	0.106	0.095	-1.207	-0.756	-58.292	0.081	0.040
250	5	-0.5	0.065	0.051	-0.488	-0.056	-0.417	0.065	0.032	0.106	0.083	0.203	-1.664	-0.927	0.046	0.022
250	5	0.5	0.074	0.060	0.488	-5.079	4.522	0.068	0.030	0.107	0.084	41.542	0.060	0.173	0.049	0.022
250	5	0.9	0.077	0.057	0.856	0.499	0.456	0.114	0.061	0.107	0.096	3.040	-0.011	-7.534	0.085	0.045
250	10	-0.9	0.092	0.074	-0.875	-0.614	-0.323	0.106	0.054	0.111	0.097	-1.061	5.569	-0.720	0.066	0.034
250	10	-0.5	0.099	0.078	-0.482	1.951	-2.368	0.071	0.032	0.123	0.098	-0.106	0.715	-0.206	0.044	0.018
250	10	0.5	0.096	0.076	0.485	-2.446	2.977	0.075	0.035	0.119	0.094	1.033	-0.877	20.133	0.044	0.018
250	10	0.9	0.102	0.079	0.878	0.621	0.307	0.105	0.049	0.114	0.105	0.252	0.051	0.802	0.068	0.032
250	30	-0.9	0.121	0.092	-0.893	-0.667	-0.236	0.092	0.044	0.135	0.116	-0.743	0.028	-4.268	0.044	0.019
250	30	-0.5	0.125	0.099	-0.497	1.132	-1.607	0.075	0.030	0.121	0.098	1.404	-1.071	-1.118	0.037	0.012
250	30	0.5	0.118	0.092	0.496	-1.201	1.756	0.066	0.030	0.114	0.089	-2.415	0.123	0.287	0.037	0.015
250	30	0.9	0.124	0.102	0.893	0.668	0.234	0.093	0.043	0.128	0.113	1.225	1.253	-2.604	0.047	0.022

(a), (b): Actual sizes of the new test based on 2SLS with nominal sizes = 10%, and 5%
(c), (d), (e): Mean biases of forward and reverse 2SLS, and mean of $\hat{B}$
(f), (g): Actual sizes of the tests based on nR^2 of the residual of forward 2SLS with nominal sizes = 10%, and 5%
(h), (i): Actual size of the new test based on Nagar with nominal size = 10%, and 5%
(j), (k), (l): Mean biases of forward Nagar, reverse Nagar, and LIML
(m), (n): Actual sizes of the tests based on nR^2 of the residual of forward Nagar with nominal sizes = 10%, and 5%

The reported numbers are based on 5000 Monte Carlo replications.

Table 3 (Cont.): $\widetilde{R}_f^2 = .001$

n	K	Cov(ε, v_2)	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
1,000	5	-0.9	0.078	0.062	-0.750	-0.175	-1.376	0.221	0.145	0.112	0.100	-0.552	2.395	-2.964	0.152	0.096
1,000	5	-0.5	0.071	0.054	-0.418	0.895	0.738	0.075	0.038	0.102	0.078	-5.998	0.097	-0.151	0.053	0.027
1,000	5	0.5	0.063	0.047	0.422	-0.917	1.382	0.088	0.043	0.102	0.076	0.290	-0.059	1.123	0.063	0.028
1,000	5	0.9	0.081	0.061	0.745	0.347	0.635	0.218	0.139	0.113	0.102	2.105	1.884	0.133	0.153	0.094
1,000	10	-0.9	0.100	0.077	-0.816	-0.528	-0.350	0.221	0.140	0.125	0.115	0.595	-0.346	-0.152	0.132	0.079
1,000	10	-0.5	0.091	0.070	-0.458	0.935	-1.196	0.080	0.038	0.112	0.087	-0.724	-0.016	7.981	0.050	0.020
1,000	10	0.5	0.091	0.072	0.450	-0.436	1.358	0.080	0.043	0.120	0.092	0.526	2.132	-2.410	0.054	0.025
1,000	10	0.9	0.098	0.078	0.817	0.530	0.363	0.214	0.130	0.128	0.116	1.532	1.284	4.417	0.132	0.078
1,000	30	-0.9	0.106	0.085	-0.871	-0.636	-0.245	0.186	0.100	0.134	0.119	-0.849	-0.940	-2.584	0.100	0.046
1,000	30	-0.5	0.119	0.088	-0.482	0.222	-0.448	0.091	0.042	0.105	0.084	0.845	-0.126	-15.053	0.045	0.019
1,000	30	0.5	0.127	0.095	0.484	-1.280	1.860	0.088	0.044	0.108	0.085	-0.144	0.427	-4.637	0.048	0.021
1,000	30	0.9	0.114	0.090	0.870	0.633	0.248	0.186	0.106	0.133	0.119	0.312	0.206	-0.016	0.102	0.052
10,000	5	-0.9	0.078	0.043	-0.236	-0.039	-0.206	0.256	0.179	0.067	0.039	0.466	0.041	0.172	0.157	0.105
10,000	5	-0.5	0.075	0.054	-0.128	0.330	-0.486	0.108	0.056	0.067	0.048	0.053	0.321	0.079	0.086	0.046
10,000	5	0.5	0.080	0.051	0.132	-0.605	0.738	0.115	0.064	0.074	0.051	0.061	-0.152	-0.437	0.097	0.053
10,000	5	0.9	0.072	0.046	0.232	0.028	0.227	0.241	0.171	0.062	0.042	-0.840	-0.164	-0.204	0.148	0.098
10,000	10	-0.9	0.112	0.085	-0.425	-0.176	-0.269	0.385	0.306	0.071	0.052	-0.024	0.346	0.165	0.171	0.118
10,000	10	-0.5	0.083	0.049	-0.234	0.497	-0.754	0.129	0.071	0.071	0.048	-0.005	0.300	-0.074	0.092	0.048
10,000	10	0.5	0.073	0.043	0.239	-0.418	0.642	0.117	0.060	0.063	0.043	-0.031	0.189	-0.047	0.083	0.042
10,000	10	0.9	0.107	0.074	0.425	0.174	0.272	0.391	0.312	0.073	0.053	-1.078	0.239	-0.202	0.176	0.126
10,000	30	-0.9	0.151	0.121	-0.673	-0.398	-0.285	0.692	0.598	0.116	0.098	0.100	-1.022	0.165	0.213	0.159
10,000	30	-0.5	0.115	0.080	-0.373	0.700	-1.094	0.131	0.070	0.071	0.049	-0.198	0.233	0.220	0.068	0.032
10,000	30	0.5	0.117	0.085	0.375	-0.740	1.143	0.132	0.070	0.068	0.049	0.721	0.293	-0.148	0.068	0.033
10,000	30	0.9	0.155	0.126	0.671	0.396	0.287	0.685	0.588	0.121	0.093	-0.080	-0.008	-0.571	0.209	0.157

(a), (b): Actual sizes of the new test based on 2SLS with nominal sizes = 10%, and 5%

(c), (d), (e): Mean biases of forward and reverse 2SLS, and mean of $\hat{B}$

(f), (g): Actual sizes of the tests based on nR^2 of the residual of forward 2SLS with nominal sizes = 10%, and 5%

(h), (i): Actual size of the new test based on Nagar with nominal size = 10%, and 5%

(j), (k), (l): Mean biases of forward Nagar, reverse Nagar, and LIML

(m), (n): Actual sizes of the tests based on nR^2 of the residual of forward Nagar with nominal sizes = 10%, and 5%

The reported numbers are based on 5000 Monte Carlo replications.

Table 4: $\widetilde{R}_f^2 = .3$

n	K	Cov(ε,v_2)	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
100	5	-0.9	0.123	0.087	-0.063	-0.012	-0.041	0.145	0.086	0.100	0.064	0.002	0.014	0.023	0.122	0.065
100	5	-0.5	0.124	0.086	-0.036	0.059	-0.073	0.106	0.053	0.099	0.071	0.001	0.025	0.013	0.100	0.051
100	5	0.5	0.123	0.091	0.036	-0.061	0.074	0.109	0.055	0.101	0.072	0.000	-0.027	-0.014	0.102	0.051
100	5	0.9	0.117	0.081	0.062	0.011	0.041	0.141	0.084	0.094	0.063	-0.004	-0.015	-0.024	0.116	0.066
100	10	-0.9	0.090	0.045	-0.147	-0.046	-0.093	0.197	0.124	0.078	0.045	0.011	0.020	0.029	0.120	0.066
100	10	-0.5	0.094	0.056	-0.084	0.117	-0.180	0.113	0.059	0.088	0.059	0.003	0.031	0.018	0.097	0.047
100	10	0.5	0.098	0.055	0.081	-0.118	0.178	0.118	0.056	0.090	0.056	-0.007	-0.033	-0.018	0.099	0.046
100	10	0.9	0.091	0.045	0.150	0.050	0.091	0.195	0.123	0.077	0.047	-0.005	-0.014	-0.024	0.118	0.067
100	30	-0.9	0.087	0.047	-0.368	-0.169	-0.193	0.400	0.261	0.071	0.030	0.077	0.016	0.027	0.114	0.063
100	30	-0.5	0.086	0.041	-0.201	0.279	-0.466	0.111	0.045	0.072	0.037	0.133	0.047	0.027	0.070	0.028
100	30	0.5	0.088	0.042	0.204	-0.285	0.474	0.115	0.048	0.076	0.039	-0.032	-0.048	-0.031	0.074	0.028
100	30	0.9	0.089	0.050	0.365	0.166	0.194	0.397	0.261	0.063	0.027	-0.038	-0.019	-0.030	0.111	0.057
250	5	-0.9	0.131	0.097	-0.024	-0.003	-0.016	0.121	0.067	0.105	0.074	0.001	0.007	0.011	0.111	0.058
250	5	-0.5	0.130	0.096	-0.015	0.023	-0.029	0.103	0.053	0.102	0.072	-0.001	0.009	0.004	0.100	0.050
250	5	0.5	0.138	0.098	0.013	-0.025	0.029	0.107	0.054	0.106	0.074	-0.001	-0.011	-0.006	0.103	0.052
250	5	0.9	0.129	0.092	0.025	0.005	0.016	0.117	0.062	0.099	0.069	0.000	-0.005	-0.009	0.107	0.054
250	10	-0.9	0.100	0.059	-0.065	-0.021	-0.039	0.145	0.083	0.086	0.055	0.000	0.005	0.008	0.106	0.057
250	10	-0.5	0.096	0.060	-0.032	0.049	-0.073	0.105	0.051	0.085	0.056	0.004	0.013	0.008	0.092	0.047
250	10	0.5	0.097	0.063	0.036	-0.047	0.074	0.102	0.053	0.085	0.057	-0.001	-0.010	-0.005	0.094	0.046
250	10	0.9	0.092	0.055	0.061	0.017	0.040	0.139	0.072	0.079	0.050	-0.005	-0.009	-0.012	0.100	0.052
250	30	-0.9	0.088	0.042	-0.188	-0.075	-0.111	0.279	0.176	0.083	0.041	0.006	0.009	0.012	0.100	0.051
250	30	-0.5	0.098	0.048	-0.104	0.130	-0.227	0.128	0.062	0.092	0.048	0.004	0.014	0.008	0.088	0.042
250	30	0.5	0.093	0.047	0.104	-0.129	0.227	0.120	0.060	0.087	0.047	-0.004	-0.013	-0.008	0.086	0.039
250	30	0.9	0.092	0.046	0.189	0.075	0.111	0.291	0.186	0.086	0.044	-0.005	-0.008	-0.011	0.102	0.054

(a), (b): Actual sizes of the new test based on 2SLS with nominal sizes = 10%, and 5%
(c), (d), (e): Mean biases of forward and reverse 2SLS, and mean of $\hat{B}$
(f), (g): Actual sizes of the tests based on nR^2 of the residual of forward 2SLS with nominal sizes = 10%, and 5%
(h), (i): Actual size of the new test based on Nagar with nominal size = 10%, and 5%
(j), (k), (l): Mean biases of forward Nagar, reverse Nagar, and LIML
(m), (n): Actual sizes of the tests based on nR^2 of the residual of forward Nagar with nominal sizes = 10%, and 5%

The reported numbers are based on 5000 Monte Carlo replications.

Table 4 (Cont.): $\widetilde{R}_f^2 = .3$

n	K	Cov(ε,v_2)	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
1,000	5	-0.9	0.138	0.098	-0.007	-0.002	-0.004	0.105	0.053	0.106	0.072	-0.001	0.000	0.001	0.103	0.051
1,000	5	-0.5	0.135	0.100	-0.005	0.005	-0.007	0.103	0.053	0.107	0.078	-0.001	0.001	0.000	0.102	0.052
1,000	5	0.5	0.136	0.101	0.005	-0.005	0.007	0.103	0.050	0.109	0.074	0.001	-0.001	0.000	0.102	0.049
1,000	5	0.9	0.132	0.097	0.007	0.002	0.004	0.105	0.056	0.105	0.076	0.001	0.000	-0.001	0.102	0.054
1,000	10	-0.9	0.102	0.069	-0.017	-0.005	-0.010	0.109	0.062	0.090	0.063	0.000	0.001	0.002	0.099	0.054
1,000	10	-0.5	0.114	0.076	-0.010	0.011	-0.019	0.110	0.057	0.100	0.067	-0.001	0.002	0.001	0.108	0.056
1,000	10	0.5	0.106	0.073	0.010	-0.011	0.019	0.104	0.054	0.094	0.064	0.000	-0.002	-0.001	0.101	0.052
1,000	10	0.9	0.111	0.074	0.017	0.006	0.010	0.118	0.062	0.096	0.064	0.001	-0.001	-0.001	0.108	0.057
1,000	30	-0.9	0.099	0.050	-0.055	-0.020	-0.034	0.160	0.087	0.094	0.049	0.000	0.002	0.003	0.106	0.052
1,000	30	-0.5	0.100	0.057	-0.030	0.035	-0.063	0.113	0.058	0.093	0.055	0.000	0.003	0.002	0.096	0.052
1,000	30	0.5	0.103	0.055	0.030	-0.035	0.063	0.115	0.059	0.096	0.052	-0.001	-0.003	-0.002	0.100	0.050
1,000	30	0.9	0.105	0.053	0.055	0.020	0.034	0.157	0.090	0.098	0.050	0.000	-0.002	-0.002	0.107	0.053
10,000	5	-0.9	0.136	0.096	0.000	0.000	0.000	0.098	0.052	0.102	0.075	0.000	0.000	0.000	0.098	0.052
10,000	5	-0.5	0.125	0.094	0.000	0.001	-0.001	0.094	0.047	0.099	0.067	0.000	0.001	0.000	0.094	0.047
10,000	5	0.5	0.134	0.101	0.000	-0.001	0.001	0.102	0.049	0.108	0.076	0.000	0.000	0.000	0.102	0.049
10,000	5	0.9	0.125	0.093	0.000	0.000	0.000	0.094	0.044	0.098	0.067	0.000	0.000	-0.001	0.094	0.044
10,000	10	-0.9	0.101	0.066	-0.002	-0.001	-0.001	0.096	0.048	0.088	0.057	0.000	0.000	0.000	0.095	0.047
10,000	10	-0.5	0.115	0.076	-0.001	0.001	-0.002	0.108	0.052	0.101	0.065	0.000	0.000	0.000	0.107	0.052
10,000	10	0.5	0.103	0.067	0.001	-0.001	0.002	0.098	0.049	0.089	0.059	0.000	0.000	0.000	0.098	0.049
10,000	10	0.9	0.111	0.074	0.002	0.000	0.001	0.107	0.052	0.097	0.063	0.000	0.000	0.000	0.106	0.051
10,000	30	-0.9	0.104	0.056	-0.006	-0.002	-0.004	0.106	0.053	0.100	0.052	0.000	0.000	0.000	0.100	0.050
10,000	30	-0.5	0.096	0.057	-0.003	0.003	-0.007	0.096	0.052	0.092	0.054	0.000	0.000	0.000	0.094	0.051
10,000	30	0.5	0.100	0.055	0.003	-0.003	0.007	0.104	0.050	0.096	0.053	0.000	0.000	0.000	0.103	0.049
10,000	30	0.9	0.100	0.055	0.006	0.002	0.004	0.102	0.052	0.095	0.051	0.000	0.000	0.000	0.095	0.049

(a), (b): Actual sizes of the new test based on 2SLS with nominal sizes = 10%, and 5%
(c), (d), (e): Mean biases of forward and reverse 2SLS, and mean of $\hat{B}$
(f), (g): Actual sizes of the tests based on nR^2 of the residual of forward 2SLS with nominal sizes = 10%, and 5%
(h), (i): Actual size of the new test based on Nagar with nominal size = 10%, and 5%
(j), (k), (l): Mean biases of forward Nagar, reverse Nagar, and LIML
(m), (n): Actual sizes of the tests based on nR^2 of the residual of forward Nagar with nominal sizes = 10%, and 5%

The reported numbers are based on 5000 Monte Carlo replications.

Table 5: Extreme Cases

n	K	ω_{12}	$\widetilde{R}_f^2$	β	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	(m)	(n)
1000	30	-0.9	0.001	5	0.056	0.047	-5.711	-7.739	2.281	0.996	0.994	0.173	0.150	-6.324	-21.806	-11.574	0.607	0.560
100	30	-0.9	0.010	5	0.043	0.034	-5.713	-7.774	2.281	0.993	0.984	0.172	0.146	-5.613	-16.686	-12.761	0.590	0.529
250	30	-0.9	0.010	5	0.086	0.075	-5.444	-18.456	15.504	0.993	0.989	0.208	0.182	-4.063	-8.329	30.184	0.492	0.437
1000	30	0.9	0.001	5	0.053	0.046	-3.968	-3.233	-0.792	0.989	0.982	0.168	0.147	-3.515	0.469	-1.909	0.577	0.528
250	30	0.9	0.010	5	0.090	0.076	-3.782	-2.538	-1.335	0.987	0.980	0.196	0.168	-7.067	0.373	-206.265	0.482	0.425
100	30	0.9	0.010	5	0.046	0.040	-3.967	-3.236	-0.775	0.973	0.939	0.161	0.143	-10.529	-1.393	-4.978	0.533	0.470
250	30	-0.5	0.010	5	0.093	0.081	-5.077	-16.519	16.499	0.969	0.953	0.185	0.156	-33.998	-14.437	-181.137	0.477	0.417
1000	30	-0.9	0.010	5	0.164	0.145	-4.399	3.468	-8.535	0.959	0.942	0.158	0.117	0.591	1.264	0.747	0.265	0.215
10000	30	-0.9	0.001	5	0.164	0.139	-4.401	-10.374	4.202	0.958	0.941	0.141	0.103	3.253	1.293	4.417	0.251	0.204
1000	30	0.9	0.010	5	0.161	0.143	-3.059	-1.184	-1.943	0.957	0.940	0.113	0.080	-3.988	1.095	0.943	0.252	0.207
10000	30	0.9	0.001	5	0.161	0.140	-3.064	-1.197	-1.945	0.956	0.939	0.118	0.081	0.377	0.229	0.698	0.254	0.206
250	30	-0.9	0.001	5	0.046	0.040	-5.852	-6.439	0.635	0.947	0.905	0.068	0.059	-14.582	-6.682	3.729	0.560	0.488
10000	30	-0.5	0.001	5	0.162	0.138	-4.102	1.893	-6.970	0.943	0.923	0.137	0.096	-5.156	1.373	1.047	0.243	0.193
1000	30	-0.5	0.010	5	0.160	0.134	-4.100	7.585	-10.363	0.942	0.921	0.143	0.097	-1.881	0.572	0.996	0.259	0.209
250	30	0.5	0.010	5	0.095	0.082	-4.154	-1.537	-2.652	0.940	0.905	0.171	0.130	-1.067	-8.462	4.416	0.449	0.383
10000	30	0.5	0.001	5	0.161	0.138	-3.362	-0.356	-3.100	0.936	0.908	0.116	0.073	-0.990	0.779	0.714	0.246	0.194
250	30	0.0	0.010	5	0.088	0.078	-4.615	2.629	-8.799	0.934	0.899	0.140	0.092	-0.906	0.078	2.272	0.443	0.381
10000	30	0.0	0.001	5	0.164	0.140	-3.732	0.965	-4.838	0.932	0.902	0.122	0.078	-1.398	4.593	0.963	0.240	0.190
1000	30	0.5	0.010	5	0.165	0.141	-3.355	-0.358	-3.070	0.929	0.901	0.124	0.079	-11.288	0.684	1.076	0.245	0.196
1000	30	0.0	0.010	5	0.159	0.136	-3.728	-0.064	-4.001	0.928	0.896	0.135	0.088	3.761	0.945	0.943	0.249	0.197
1000	30	-0.5	0.001	5	0.071	0.060	-5.323	-9.849	2.004	0.926	0.879	0.173	0.149	-12.138	-6.498	13.322	0.517	0.446
100	30	-0.9	0.100	5	0.142	0.124	-4.303	7.817	-12.307	0.920	0.877	0.167	0.125	9.616	1.320	1.048	0.242	0.180
1000	30	-0.9	0.010	1	0.163	0.140	-1.417	-3.825	2.785	0.918	0.889	0.264	0.238	0.369	-0.015	0.201	0.247	0.204
10000	30	-0.9	0.001	1	0.161	0.140	-1.417	-3.965	3.080	0.915	0.884	0.244	0.218	1.292	1.183	0.714	0.243	0.196
100	30	0.9	0.100	5	0.142	0.124	-2.990	-1.138	-1.863	0.910	0.858	0.114	0.085	-1.431	0.560	0.499	0.239	0.182
100	30	-0.5	0.100	5	0.143	0.123	-4.005	6.504	-10.517	0.897	0.836	0.155	0.106	-0.881	1.255	0.935	0.240	0.174
250	30	-0.9	0.010	1	0.100	0.086	-1.753	-2.364	0.662	0.886	0.824	0.047	0.039	-1.199	-13.265	-3.013	0.411	0.345
100	30	0.5	0.100	5	0.144	0.125	-3.278	-0.272	-3.026	0.873	0.804	0.121	0.087	10.961	0.546	1.531	0.225	0.161
100	30	0.0	0.100	5	0.149	0.130	-3.639	0.345	-4.030	0.871	0.805	0.140	0.090	-1.074	0.653	0.323	0.233	0.164
250	10	-0.9	0.001	5	0.058	0.049	-5.768	-7.567	2.579	0.863	0.823	0.115	0.100	-8.002	9.639	7.835	0.540	0.485

Note: ω_{12} denotes the covariance between v_1 and v_2. Columns (a) – (n) denote same objects as in previous tables.

Table 6: Comparison of 2SLS and LIML

$\left(\sigma_{\varepsilon v_2}, \beta, \sigma_\varepsilon^2\right)$		(Forward) 2SLS	LIML
(-0.1, 0.1, 1.01)	Mean Bias	-0.09633	-0.45529
	Median Bias	-0.09545	-0.0786
	Standard Deviation	0.188187	40.32329
	75 percentile – 25 percentile	0.254975	1.828275
	90 percentile – 10 percentile	0.47662	5.41373
(-0.01, 0.01, 1.0001)	Mean Bias	-0.00922	0.065113
	Median Bias	-0.00845	-0. 0063
	Standard Deviation	0.188241	20.9529
	75 percentile – 25 percentile	0.25485	1.8313
	90 percentile – 10 percentile	0.47805	5.39492
(0.01, -0.01, 1.0001)	Mean Bias	0.010141	1.618553
	Median Bias	0.0107	0.0088
	Standard Deviation	0.188259	103.4206
	75 percentile – 25 percentile	0.25465	1.831725
	90 percentile – 10 percentile	0.47745	5.37822
(0.1, -0.1, 1.01)	Mean Bias	0.097252	-0.18974
	Median Bias	0.0977	0.08535
	Standard Deviation	0.188368	37.93867
	75 percentile – 25 percentile	0.2552	1.8266
	90 percentile – 10 percentile	0.47512	5.43348

Note: The reported numbers are based on 5000 Monte Carlo repetitions. For every parameter combination, we set $n = 1000$, $K = 30$, and $\widetilde{R}_f^2 = .001$.