Revised draft
Please do not cite without authors' permission

# A New Specification Test for the Validity of Instrumental Variables*


Jinyong Hahn
University of Michigan

And

Jerry Hausman
MIT


**March, 2000**


Excellent research assistance has been provided by Derek Ho Cheung and Karen Hull. Whitney Newey provided some helpful suggestions. We received helpful comments from numerous seminar participants and from the referees. 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 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.

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 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 $\pi_2$. 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 will consider the non-normal case later in the paper. We use the following notation:

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

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

A common finding in empirical research is that when 2SLS is used the coefficient estimate increases in magnitude from the OLS estimate. However, in finite samples under certain situations even when 2SLS is used on equation (2.1), bias remains because an estimate of $\pi_2$ 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

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

Observe that

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

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

We acknowledge that the denominator of equation (3.2) is random so we have only an approximation, but we justify the expression subsequently on the basis of the asymptotic approximations that we carry out. 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 $\sigma_{\varepsilon v_2}$, $K$, $R_f^2$.

### 3.2 Forward and Reverse Regressions

Let

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

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

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

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

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

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

Below, we examine whether his asymptotics captures our motivation.

It can be shown that[1]

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

It can also be shown that[2]

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

[1] See Bekker (1994, p.663).
[2] See Bekker (1994).

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

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

which coincides with equation (3.2) and justifies our approximation result with the addition of the parameter $\alpha$.

## 5. 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 $\alpha$ 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, $\beta$, 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 $\beta$ 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 $\beta$ is equal to

$$(5.1) \qquad B = -\alpha \frac{\Theta\sigma_\varepsilon^2 + \alpha \det(\Omega)}{(\Theta + \alpha\omega_{22})(\beta\Theta + \alpha\omega_{12})}.$$

Bekker (1994, eq. (4.7)) shows that 2SLS is asymptotically normal. Therefore, $\sqrt{n}\left(b_{2SLS} - 1/c_{2SLS} - B\right)$ is also asymptotically normal. Because we do not know $B$ in general, we would like to deal with an asymptotic result of the form

$$(5.2) \qquad \sqrt{n}\left(b_{2SLS} - \frac{1}{c_{2SLS}} - \hat{B}\right) \to N(0,V)$$

where $\hat{B}$ is a $\sqrt{n}$ -consistent estimator of $B$.

In terms of a formal null hypothesis we test

$$(5.3) \qquad H_0 : \text{plim}\sqrt{n}\left(b_{2SLS} - \frac{1}{c_{2SLS}} - B\right) = 0$$

where we are taking the plim relative to the usual (higher order) asymptotic approximation. If we reject $H_0$ we decide that the usual first order inference based on the asymptotic normal approximation to the 2SLS estimator is not sufficiently accurate to be used. Two primary reasons exist for a rejection. First, the orthogonality assumptions of the instruments may be false. The traditional Sargan test of overidentifying restrictions also tests this assumption, but it is well known to have poor size properties. Our Monte Carlo results demonstrate that the new specification test has considerably better size properties than the Sargan test. Alternatively, a rejection may occur because the finite sample properties of the first order asymptotic approximation are not sufficiently accurate (weak instruments) in the current situation to be used.

$\hat{B}$ will be a $\sqrt{n}$ 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

$$\text{(5.4)} \qquad \Theta + \alpha\omega_{22} = \text{plim}\frac{1}{n}y_2' P_z y_2, \quad \text{and} \quad \beta\Theta + \alpha\omega_{12} = \text{plim}\frac{1}{n}y_2' P_z y_1 .$$

Further, it can be shown by using Lemma 1 in Appendix A.1 that

$$\text{(5.5)} \qquad \text{plim}\hat{\Xi} = \Theta\sigma_\varepsilon^2 + \alpha\det(\Omega),$$

where

$$\hat{\Xi} \equiv \left(\frac{1}{n}y_1' P_z y_1 - \frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_1' M_z y_1\right)\frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2' M_z y_2 - 2\left(\frac{1}{n}y_2' P_z y_1 - \frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2' M_z y_1\right)\frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2' M_z y_1$$
$$+\left(\frac{1}{n}y_2' P_z y_2 - \frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2' M_z y_2\right)\frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_1' M_z y_1$$
$$+\left(\frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_1' M_z y_1\right)\left(\frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2' M_z y_2\right) - \left(\frac{\hat{\alpha}}{1-\hat{\alpha}}\frac{1}{n}y_2' M_z y_1\right)^2$$

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

$$\text{(5.6)} \qquad \hat{B} \equiv -\frac{\hat{\Xi}}{\frac{1}{n}y_2 P_z y_2 \cdot \frac{1}{n}y_2' P_z y_1} .$$

By the delta method based on Lemma 1 in Appendix A.1, it can be shown that

**Theorem 5.1:** $\sqrt{n}\left(b_{2SLS} - \frac{1}{c_{2SLS}} - \hat{B}\right) \to N\left(0, \frac{2\alpha}{1-\alpha}\frac{(\sigma_\varepsilon^2)^2\Theta^2}{(\Theta+\alpha\omega_{22})^2(\beta\Theta+\alpha\omega_{12})^2}\right).$

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. Thus, the specification test takes the form of an asymptotic $t$ statistic:

$$m = \frac{\hat{d}}{\hat{w}^{0.5}}, \tag{5.7}$$

where $\hat{d}$ is the LHS of Theorem 5.1, and $\hat{w}$ is a consistent estimate of the variance in Theorem 5.1. We discuss later how to estimate this variance term.

## 6. 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 $\beta$ is:

$$b \equiv \frac{y_2' P_z y_1 - \lambda y_2' M_z y_1}{y_2' P_z y_2 - \lambda y_2' M_z y_2}, \qquad \text{where} \qquad \lambda = \frac{\frac{K-2}{n}}{1 - \frac{K-2}{n}}. \tag{6.1}$$

We can also estimate $\beta$ by the reverse IV specification:

$$\frac{1}{c} \equiv \frac{y_1' P_z y_1 - \lambda y_1' M_z y_1}{y_2' P_z y_1 - \lambda y_2' M_z y_1}. \tag{6.2}$$

By the delta method based on Lemma 1 in Appendix A.1, we can show that

**Theorem 6.1:**

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

.

We will use Theorem 6.2 below to compare the forward and reverse bias adjusted estimators of $\beta$ to form a test of the model specification.

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

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

which coincides with the asymptotic variance of LIML as derived by Bekker (1994, eq. (4.7)). 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 $\lambda$ 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 needs be made as in Theorem 5.1 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 6.1, we obtain

**Theorem 6.2:** $\sqrt{n}\left(b-\frac{1}{c}\right)\to N\left(0,\frac{2\alpha}{1-\alpha}\frac{(\sigma_\varepsilon^2)^2}{\beta^2\Theta^2}\right)$.

**Proof**: Follows from

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

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

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

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

Our second specification test has the form of an asymptotic $t$ statistic:

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

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 6.2. We subsequently discuss how to consistently estimate the variance term.

We now compare the two specification tests, which are based on the 2SLS estimator and the Nagar-type estimator. We find that the two tests have unitary correlation asymptotically using Bekker asymptotics.

**Theorem 6.3:**

$$(\beta\Theta+\alpha\omega_{12})(\Theta+\alpha\omega_{22})\bullet\sqrt{n}\left(b_{2SLS}-\frac{1}{c_{2SLS}}-\hat{B}\right)-\beta\Theta^2\bullet\left(b-\frac{1}{c}\right)=o_p(1).$$

**Proof:** See Appendix A.2.

Having established the asymptotic equivalence of the two tests, we now examine the robustness of our result to departures from normality under conditions adopted by Donald and Newey (1998). We impose a symmetry assumption:

**Condition (6.1):** (1) $E\left[v^{j}{}_{2i}\varepsilon^{k}{}_{1i}\right]=0$ for $j+k$ a positive odd integer such that $j+k\leq 5$;
(ii) $E\left[\left|v_{2i}\right|^{j}\left|\varepsilon_{1i}\right|^{k}\right]<\infty$ for *j* and *k* any positive integers such that $j+k\leq 5$.

It can be shown that:

**Theorem 6.4:** Suppose that equation (4.1) and Conditions (7.1) and (7.2) hold true. The approximate variance of $\sqrt{n}(b-1/c)$ equals

$$\frac{2\alpha}{1-\alpha}\frac{\left(\sigma_{\varepsilon}^{2}\right)^{2}}{\beta^{2}\Theta^{2}}+\left(\left(\frac{1}{1-\alpha}\right)^{2}\frac{\sum_{i=1}^{n}d_{i}^{2}}{n}-\left(\frac{\alpha}{1-\alpha}\right)^{2}\right)\frac{E\left[\varepsilon_{1i}^{4}\right]-3\left(\sigma_{\varepsilon}^{2}\right)^{2}}{\beta^{2}\Theta^{2}},$$

where $d_i$ for $i=1,\ldots,n$ denote diagonal elements of $P_z$.
**Proof:** See Appendix A.3.

In typical applications, the second term of this expression is expected to be small[3] compared to the first term unless the kurtosis of the distribution is extremely large. Thus, the approximation should work well in the symmetric case, expect for extreme departures, as our subsequent Monte Carlo experiments demonstrate where we use a *t*-distribution to allow for the non-normal situation.

---

[3] Because $\sum d_i=K$, we can without loss of generality write $n^{-1}\sum_{i=1}^{n}d_i^2-\alpha^2=\mathrm{Var}(d_i)$, where $\mathrm{Var}(d_i)\equiv n^{-1}\sum\left(d_i-K/n\right)^2$ denotes the sample variance of $d_i$. Therefore, the magnitude of $\left(\left(\frac{1}{1-\alpha}\right)^{2}\frac{\sum_{i=1}^{n}d_{i}^{2}}{n}-\left(\frac{\alpha}{1-\alpha}\right)^{2}\right)$ relative to $\frac{\alpha}{1-\alpha}$ can be readily computed.

## 7. 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 12.

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

$$\frac{(K-2)\sigma_{\varepsilon v_2}}{\Theta\sqrt{n}}, \tag{7.1}$$

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

$$\beta + \alpha\frac{\sigma_{\varepsilon v_2}}{\Theta}. \tag{7.2}$$

Observe that equation (7.2) 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

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

and (approximate) variance

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

which is similar to the Bekker-based result we derived for LIML in equation (6.1), except the approximating factor $\alpha/(1-\alpha)$ in equation (6.3) has changed to $\alpha$ in equation (7.4). 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

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

Notice that equation (7.5) again agrees with a Bekker-based asymptotic variance of the Donald-Newey estimator in Theorem 5.1 except that, again, Rothenberg's Edgeworth correction terms are of order $\alpha$, 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

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

Observe that equation (7.6) is smaller than equations (7.4) or (7.5), which suggests that the variance of 2SLS is smaller than that of a Nagar-type estimator or LIML.[4] We could not tell whether Bekker's asymptotics predicts the same pattern of variances. There is good reason to believe that equation (7.6) 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 (7.6) is negative, especially when the first stage $R_f^2$, and hence $\Theta$, is extremely small which can correspond to the "weak instrument" situation. Because 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.[5] 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.[6] 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 11 where we perform Monte Carlo experiments.

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

[5] 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 11, we investigate such a potential outcome by Monte Carlo simulation.

[6] See Mariano and Sawa (1972). As for the forward Nagar estimator, it does not even possess first moments as established by Sawa (1972). It can be demonstrated that the Donald-Newey estimators that we consider here have similar moment properties to the Nagar estimator.

## 8. Estimation of Asymptotic Variance Terms and Included Predetermined Variables

In this section we consider some practical consideration for application of the new specification test. For the first specification test, we need to estimate the asymptotic variance

$$(8.1) \qquad w_1 = \frac{2\alpha}{1-\alpha} \frac{\left(\sigma_\varepsilon^2\right)^2 \Theta^2}{\left(\Theta + \alpha\omega_{22}\right)^2 \left(\beta\Theta + \alpha\omega_{12}\right)^2} .$$

We use consistent estimates for the unknown parameters and follow Bekker and use $\hat{\alpha} = (K-1)/(n-1)$. Using Lemma 1 in Appendix A.1, we can show that a consistent estimator for asymptotic variance is given by

$$(8.2) \qquad \hat{w}_1 = 2\frac{K-1}{n-K} \frac{\left(\sum_{i=1}^n \left(y_i - \beta_{LIML} y_{2i}\right)^2\right)^2 \left(y_2' P_z y_2 - \frac{K-1}{n-K} y_2' M_z y_2\right)^2}{\left(y_2' P_z y_2\right)^2 \left(y_2' P_z y_1\right)^2} .$$

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

$$(8.3) \qquad w_2 = \frac{2\alpha}{1-\alpha} \frac{\sigma_\varepsilon^2}{\beta^2 \Theta^2} .$$

By the same calculation, a consistent estimator is given by

$$(8.4) \qquad \hat{w}_2 = 2\frac{K-1}{n-K} \frac{\left(\sum_{i=1}^n \left(y_i - \beta_{LIML} y_{2i}\right)^2\right)^2}{\beta_{LIML}^2 \left(y_2' P_z y_2 - \frac{K-1}{n-K} y_2' M_z y_2\right)^2} .$$

In either of the variance estimates of equations (8.2) and (8.4), a different consistent estimator other than LIML can be used, with no change in the distribution of the estimated test statistic.

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). 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

$$
\begin{aligned}
Y_{1i} &= \beta Y_{2i} + Z_{1i}'\gamma + \mathrm{E}_i \\
Y_{2i} &= Z_{1i}'\phi + Z_{2i}'\pi_2 + V_{2i}
\end{aligned} \tag{8.5}
$$

where $Z_{1i}$ is a $k_1$ dimensional vector of included predetermined variables in equation (8.5) 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 (8.5), 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 $Z_1$, $Z_2$, $\mathrm{E}$, and $V_2$ similarly. With

$$
y_1 = M_{Z_1}Y_1, \quad y_2 = M_{Z_1}Y_2, \quad z = M_{Z_1}Z_2, \quad \varepsilon = M_{Z_1}\mathrm{E}, \quad v_2 = M_{Z_1}V_2, \tag{8.6}
$$

we obtain equations (2.1) and (2.2) premultiplying equation (8.5) by $M_{Z_1}$.

As usual with partialling out with projection matrices a convenient computational procedure follows where we

- 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 $y_2'P_z y_2 = \hat{y}_2'\hat{y}_2$, $y_2'P_z y_1 = \hat{y}_2'\hat{y}_1$, $y_1'P_z y_1 = \hat{y}_1'\hat{y}_1$, $y_2'M_z y_2 = \widetilde{y}_2'\widetilde{y}_2$, $y_2'M_z y_1 = \widetilde{y}_1'\widetilde{y}_1$, $y_1'M_z y_1 = \widetilde{y}_1'\widetilde{y}_1$, and plug into equations (6.1), (8.2), and (8.4).

As for $K$ in equations (6.6), (8.2), and (8.4), 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.

## 9. 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 first 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 then consider the first specification test in a similar situation.

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

$$y_1 = \beta_2 y_2 + \beta_3 y_3 + \varepsilon_1 \qquad (9.1)$$

$$\begin{aligned} y_2 &= z\pi_2 + v_2 \\ y_3 &= z\pi_3 + v_3 \end{aligned} \qquad (9.2)$$

where we use the same matrix and vector notation as before. We consider estimation of $\beta_2$ and $\beta_3$ in equation (9.1) 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 $(\beta_2, \beta_3)$. 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 Appendix B.2, 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. To be specific, the asymptotic variance matrix has rank equal to one. 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:

$$(9.3) \qquad y_2 = \frac{1}{\beta_2} y_1 + \left( \frac{-\beta_3}{\beta_2} \right) y_2 + \varepsilon_2 .$$

In Appendix B.2, 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

$$(9.4) \quad 2\frac{K-1}{n-K} \frac{\left( \sum_{i=1}^{n} \left( y_{1i} - \beta_{2,LIML} y_{2i} - \beta_{3,LIML} y_{3i} \right)^2 \right)^2}{\beta_{2,LIML}^2 \left( y_2' P_z y_2 - \frac{K-1}{n-K} y_2' M_z y_2 - \frac{\left( y_2' P_z y_3 - \frac{K-1}{n-K} y_2' M_z y_3 \right)^2}{\left( y_3' P_z y_3 - \frac{K-1}{n-K} y_3' M_z y_3 \right)} \right)^2} .$$

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

$$(9.5) \qquad m_3 = \frac{\hat{d}_3}{\hat{w}_3^{\ 0.5}} ,$$

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

Having developed a specification test based on Nagar type estimators for the case of additional RHS jointly endogenous variables, we develop another specification test based on 2SLS estimators for additional RHS jointly endogenous variables. Let $b_{2SLS,1}$ and $b_{2SLS,2}$ denote the 2SLS versions of $b_1$ and $b_2$. Again we need to subtract off a bias term which is the $\sqrt{n}$-consistent estimate $\hat{C}$ of plim$b_{2SLS,1}$ - plim$b_{2SLS,2}$. In Appendix B.3, we demonstrate that

$$(9.6) \qquad \hat{C} = -\frac{-n_1 d_2 + n_2 d_1 + N_1 D_2 - N_2 D_1}{d_1 d_2}$$

where

$$D_1 = A_{22}A_{33} - \lambda A_{22}a_{33} - \lambda a_{22}A_{33} + \lambda^2 a_{22}a_{33} - A^2{}_{23} + 2\lambda A_{23}a_{23} - \lambda^2 a^2{}_{23}$$
$$D_2 = -A_{33}A_{12} + \lambda A_{33}a_{12} + \lambda a_{33}A_{12} - \lambda^2 a_{33}a_{12} + A_{23}A_{13} - \lambda A_{23}a_{13} - \lambda a_{23}A_{13} + \lambda^2 a_{23}a_{13}$$
$$d_1 = A_{22}A_{33} - A^2{}_{23}$$
$$d_2 = -A_{12}A_{33} + A_{23}A_{13} \quad \text{and}$$
$$N_1 = A_{12}A_{33} - \lambda A_{33}a_{12} - \lambda a_{33}A_{12} + \lambda^2 a_{33}a_{12} - A_{23}A_{13} + \lambda A_{23}a_{13} + \lambda a_{23}A_{13} - \lambda^2 a_{23}a_{13}$$
$$N_2 = -A_{11}A_{33} + \lambda A_{11}a_{33} + \lambda a_{11}A_{33} - \lambda^2 a_{11}a_{33} + A^2{}_{13} - 2\lambda A_{13}a_{13} + \lambda^2 a^2{}_{13}$$
$$n_1 = A_{12}A_{33} - A_{23}A_{13}$$
$$n_2 = -A_{11}A_{33} + A^2{}_{13}$$

Here we use the notation $\lambda = \hat{\alpha}/(1-\hat{\alpha})$, $y_j' P_z y_k = A_{jk}$ and $y_j' M_z y_k = a_{jk}$ for $j,k$ = 1,2,3.

**Theorem 9.1:**

$$\sqrt{n}(b_{2SLS,1} - b_{2SLS,2} - \hat{C}) \to N\left(0, \frac{2\alpha Var(\varepsilon_{1i})^2}{1-\alpha} \bullet \frac{\Theta^2{}_{33}(\Theta_{22}\Theta_{33} - \Theta^2{}_{23})^2}{Q^2{}_1 Q^2{}_2}\right)$$

where

$$Q_1 = \Theta_{22}\Theta_{33} + \alpha\Theta_{22}\omega_{33} + \alpha\omega_{22}\Theta_{33} + \alpha^2\omega_{22}\omega_{33} - \Theta^2{}_{23} - 2\alpha\Theta_{23}\omega_{23} - \alpha^2\omega^2{}_{23},$$

$$Q_2 = -\beta_2\Theta_{22}\Theta_{33} - \alpha\Theta_{33}\omega_{12} - \alpha\beta_2\omega_{33}\Theta_{22} - \alpha\beta_3\omega_{33}\Theta_{23} - \alpha^2\omega_{12}\omega_{33} + \beta_2\Theta^2{}_{23} + \alpha\Theta_{23}\omega_{13}$$
$$+ \alpha\beta_2\omega_{23}\Theta_{23} + \alpha\beta_3\omega_{23}\Theta_{33} + \alpha^2\omega_{23}\omega_{13}$$

**Proof:** See Appendix B.3.

Also, in Appendix B.3 we demonstrate that the two test statistics based on the Nagar estimator and the 2SLS estimator have unitary correlation under Bekker asymptotics. Thus, a similar result holds that all tests based on a single difference of the 2SLS estimators will have the same operating characteristics. We also derive a consistent estimate of the asymptotic variance for the test statistic derived in the theorem. Thus, we have generalized the specification tests to additional RHS endogenous variables and find the result that any single difference of the estimators provides a test that cannot be improved upon.

Inclusion of exogenous and predetermined variables in the specification as in equation (8.5) in Section 8 raises no new complications. The partialling-out methodology we used in Section 8 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.

## 10. 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.[7] 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 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

[7] 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).

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 (8.1) to calculate the variance and estimate our specification test statistic to be:

$$m = \frac{\hat{d}}{\hat{w}^{0.5}} = \frac{10.03}{5.43} = 1.85 \,. \tag{10.1}$$

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

$$(10.2) \qquad m = \frac{\hat{d}}{\hat{w}^{0.5}} = \frac{247.3}{17.4} = 14.21.$$

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.

## 11. 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

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

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

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

$$n = 100,\quad 250,\quad 1000,\quad 10000$$

$$\sigma_{\varepsilon v_2} = -.9,\quad -.5,\quad .5,\quad .9$$

$$\tilde{R}_f^2 = .001,\quad .01,\quad .1,\quad .3$$

$$K = 5,\quad 10,\quad 30$$

We examined performance of our tests by 5000 Monte Carlo replications. We used 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.[8] Because of space limitations we only report some or the results here with other results reported on a website.[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 Tables 1 and 2 the conventional test of overidentification often has actual size of above 0.3, when the nominal size is smaller

[8] We set the values of $\beta$ such that Var(ε)=1.

[9] http://web.mit.edu/jhausman/www/

[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.

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. 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 7. 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.

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. Instead of calibrating

[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.

$\sigma_{\varepsilon v_2}$ keeping $\sigma_{\varepsilon} = \sigma_{v_2} = 1$ and varying $\beta$, we calibrated $\omega_{12} = \mathrm{Cov}(v_1, v_2)$ as well as $\beta$ keeping $\sigma_{v_1} = \sigma_{v_2} = 1$, thereby varying $\sigma_{\varepsilon v_2} = \sigma^2_{v_1} - \beta\sigma_{v_1,v_2} = 1 - \beta \cdot \omega_{12}$ to a much greater extent than in previous Tables 1 - 4. 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 7, 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 $\widetilde{R}^2_f$ 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 $\beta$ 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 $\widetilde{R}^2_f$ 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.

In Table 7 we repeat Table 1 for the non-normal case. We use a log normal distribution for the RHS variable standardized with mean zero and variance one, and the stochastic disturbance has a *t* distribution with 12 degree of freedom, again standardized to have mean zero and variance one. A comparison of Table 7 with Table 1 shows that the new specification test continues to perform well with the actual sizes of the tests in

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

columns (a)-(b) and columns (h)-(i) quite close to the nominal sizes. The $nR^2$ of the test of overidentification in columns (f)-(g) performs more poorly than in Table 1 and we again find some evidence of a "moments problem" for the Nagar type estimator. These results carry over to the other Monte Carlo experiments for the non-normal situation, and we report the results on the website.

## 12. 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 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 (5.7) 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 (5.6) 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 (6.4). 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

[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.

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. We expect to extend our results to 3 or more RHS jointly endogenous variables in the near future.

# 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<br>(.147) | -2.03<br>(.465) | -2.31<br>(.500) | -2.03<br>(.465) | -2.31<br>(.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<br>(.147) | -1.24<br>(.194) | -8.01<br>(.789) | -1.21<br>(.194) | -4.64<br>(.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 One Endogenous Regressor

### A.1 A Technical Lemma

**Lemma 1** *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* $\Theta$*. Let* $\overline{S} \equiv U' P_z U$ *and* $S^{\perp} \equiv U' M_z U$*. We then have*

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

*where* $\overline{\Lambda}$ *and* $\Lambda^{\perp}$ *denote symmetric* $3 \times 3$ *matrices such that*

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

*and*

$$\begin{aligned}
\Lambda_{1,1}^{\perp} &= 2(1-\alpha)\,\omega_{11}^2 \\
\Lambda_{1,2}^{\perp} &= 2(1-\alpha)\,\omega_{11}\omega_{12} \\
\Lambda_{1,3}^{\perp} &= 2(1-\alpha)\,\omega_{12}^2 \\
\Lambda_{2,2}^{\perp} &= (1-\alpha)\,\omega_{11}\omega_{22} + (1-\alpha)\,\omega_{12}^2 \\
\Lambda_{2,3}^{\perp} &= 2(1-\alpha)\,\omega_{22}\omega_{12} \\
\Lambda_{3,3}^{\perp} &= 2(1-\alpha)\,\omega_{22}^2
\end{aligned}$$

**Proof.** Let $U \equiv \begin{bmatrix} y_1 & y_2 \end{bmatrix}$, $M \equiv [\beta \cdot z\pi_2, z\pi_2] = z\pi_2(\beta, 1)$, and $V \equiv U - M$. Note that the rows of $V$ are i.i.d. normal with zero mean and variance $\Omega$. Using Bekker (1994, Lemma 2), we obtain

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

The conclusion follows by combining this observation with Bekker (1994, Lemma 1) along with the fact $\overline{S}$ and $S^{\perp}$ are independent of each other due to normality. ■

## A.2 Proof of Theorem 6.3: Asymptotic Equivalence of Two Tests

Because $\frac{\widehat{\alpha}}{1-\widehat{\alpha}} = \frac{\alpha}{1-\alpha} + o_p\left(\frac{1}{\sqrt{n}}\right)$, and $\lambda = \frac{K-2}{n} / \left(1 - \frac{K-2}{n}\right) = \frac{\alpha}{1-\alpha} + o_p\left(\frac{1}{\sqrt{n}}\right)$, we may without loss of generality assume that $\frac{\widehat{\alpha}}{1-\widehat{\alpha}} = \lambda$ for the asymptotic argument. For simplicity, we will rewrite $A_1 = \frac{1}{n} y_1' P y_1$, $A_2 = \frac{1}{n} y_1' P y_2$, $A_3 = \frac{1}{n} y_2' P y_2$, $a_1 = \frac{1}{n} y_1' M y_1$, $a_2 = \frac{1}{n} y_1' M y_2$, and $a_3 = \frac{1}{n} y_2' M y_2$. Observe that

$$A_2 A_3 = (\beta\Theta + \alpha\omega_{12})(\Theta + \alpha\omega_{22}) + o_p(1), \quad \text{and} \quad (A_3 - \lambda a_3)(A_2 - \lambda a_2) = \beta\Theta^2 + o_p(1). \tag{A.1}$$

We may then write

$$\widehat{B} = -\frac{\widehat{\Xi}}{A_1 \cdot A_2},$$

where

$$\widehat{\Xi} = (A_1 - \lambda a_1)\lambda a_3 - 2(A_2 - \lambda a_2)\lambda a_2 + (A_3 - \lambda a_3)\lambda a_1 + (\lambda a_1)(\lambda a_3) - (\lambda a_2)^2.$$

Therefore, we have

$$S \equiv \sqrt{n}\left(b_{2SLS} - \frac{1}{c_{2SLS}} - \widehat{B}\right) = \sqrt{n}\left(\frac{A_2}{A_3} - \frac{A_1}{A_2} + \frac{\widehat{\Xi}}{A_2 A_3}\right) = \sqrt{n}\frac{A_2^2 - A_1 A_3 + \widehat{\Xi}}{A_2 A_3}, \tag{A.2}$$

$$T \equiv \sqrt{n}\left(b - \frac{1}{c}\right) = \sqrt{n}\left(\frac{A_2 - \lambda a_2}{A_3 - \lambda a_3} - \frac{A_1 - \lambda a_1}{A_2 - \lambda a_2}\right) = \sqrt{n}\frac{(A_2 - \lambda a_2)^2 - (A_1 - \lambda a_1)(A_3 - \lambda a_3)}{(A_3 - \lambda a_3)(A_2 - \lambda a_2)}. \tag{A.3}$$

The numerators on the far RHS in (A.2) and (A.3), i.e., $A_2^2 - A_1 A_3 + \widehat{\Xi}$ and $(A_2 - \lambda a_2)^2 - (A_1 - \lambda a_1)(A_3 - \lambda a_3)$ are identical. Call it $\digamma$. Using Theorem 6.2 and (A.1), we can show that $\sqrt{n}\digamma \to N\left(0, 2\frac{\alpha}{1-\alpha}\Theta^2\left(\sigma_\varepsilon^2\right)^2\right)$, i.e.,

$$\sqrt{n}\digamma = O_p(1). \tag{A.4}$$

Combining (A.4) with (A.1), we obtain

$$\begin{aligned}(\beta\Theta + \alpha\omega_{12})(\Theta + \alpha\omega_{22}) S - \beta\Theta^2 T &= \left(\frac{(\beta\Theta + \alpha\omega_{12})(\Theta + \alpha\omega_{22})}{A_2 A_3} - \frac{\beta\Theta^2}{(A_3 - \lambda a_3)(A_2 - \lambda a_2)}\right)\sqrt{n}\digamma \\ &= o_p(1) \cdot O_p(1) = o_p(1).\end{aligned}$$

## A.3 Proof of Theorem 6.4: Higher Order Approximation with Nonnormal Error

We adopt the same structure of argument as in Donald and Newey (1998). Our proof is much simpler due to explicit use of Bekker's assumption that $\frac{1}{n}\pi_2' z' z \pi_2$ is fixed at $\Theta$ and $\frac{K}{n} \to \alpha + o\left(\frac{1}{\sqrt{n}}\right)$. Consider the normalized difference of the two Nagar type estimators

$$\sqrt{n}\left(\frac{y_2' P_z y_1 - \lambda y_2' M_z y_1}{y_2' P_z y_2 - \lambda y_2' M_z y_2} - \frac{y_1' P_z y_1 - \lambda y_1' M_z y_1}{y_2' P_z y_1 - \lambda y_2' M_z y_1}\right) = \sqrt{n}\left(\frac{y_2' P_z y_1 - \lambda v_2' M_z v_1}{y_2' P_z y_2 - \lambda v_2' M_z v_2} - \frac{y_1' P_z y_1 - \lambda v_1' M_z v_1}{y_2' P_z y_1 - \lambda v_2' M_z v_1}\right). \tag{A.5}$$

Because

$$\begin{aligned} y_1' P_z y_1 &= n\beta^2\Theta + 2\beta(z\pi_2)' v_1 + v_1' P_z v_1, \\ y_2' P_z y_1 &= n\beta\Theta + (z\pi_2)' v_1 + \beta(z\pi_2)' v_2 + v_2' P_z v_1, \\ y_2' P_z y_2 &= n\Theta + 2(z\pi_2)' v_2 + v_2' P_z v_2, \end{aligned}$$

we have

$$\frac{y_2' P_z y_1 - \lambda v_2' M_z v_1}{y_2' P_z y_2 - \lambda v_2' M_z v_2} - \beta = \frac{\frac{1}{n}(z\pi_2)' \varepsilon_1 + \frac{1}{n} v_2' P_z \varepsilon_1 - \frac{1}{n}\lambda v_2' M_z \varepsilon_1}{\Theta + \frac{1}{n} 2 (z\pi_2)' v_2 + \frac{1}{n} v_2' P_z v_2 - \frac{1}{n}\lambda v_2' M_z v_2}$$
$$= \frac{\frac{1}{n}(z\pi_2)' \varepsilon_1 + \frac{1}{n}(1+\lambda) v_2' P_z \varepsilon_1 - \frac{1}{n}\lambda v_2' \varepsilon_1}{\Theta + \frac{1}{n} 2 (z\pi_2)' v_2 + \frac{1}{n}(1+\lambda) v_2' P_z v_2 - \frac{1}{n}\lambda v_2' v_2}, \quad \text{(A.6)}$$

and

$$\frac{y_1' P_z y_1 - \lambda v_1' M_z v_1}{y_2' P_z y_1 - \lambda v_2' M_z v_1} - \beta = \frac{\frac{1}{n}\beta (z\pi_2)' \varepsilon_1 + \frac{1}{n} v_1' P_z \varepsilon_1 - \frac{1}{n}\lambda v_1' M_z \varepsilon_1}{\beta\Theta + \frac{1}{n}(z\pi_2)' v_1 + \frac{1}{n}\beta (z\pi_2)' v_2 + \frac{1}{n} v_2' P_z v_1 - \frac{1}{n}\lambda v_2' M_z v_1}$$
$$= \frac{\frac{1}{n}\beta (z\pi_2)' \varepsilon_1 + \frac{1}{n}(1+\lambda) v_1' P_z \varepsilon_1 - \frac{1}{n}\lambda v_1' \varepsilon_1}{\beta\Theta + \frac{1}{n}(z\pi_2)' v_1 + \frac{1}{n}\beta (z\pi_2)' v_2 + \frac{1}{n}(1+\lambda) v_2' P_z v_1 - \frac{1}{n}\lambda v_2' v_1}. \quad \text{(A.7)}$$

We first take care of (A.6). Let

$$H_1 \equiv \Theta, \quad \widetilde{H}_1 \equiv -\left(\frac{1}{n} 2 (z\pi_2)' v_2 + \frac{1}{n}(1+\lambda) v_2' P_z v_2 - \frac{1}{n}\lambda v_2' v_2\right),$$

and

$$h_1 \equiv \frac{1}{\sqrt{n}} (z\pi_2)' \varepsilon_1, \quad \widetilde{h}_1 \equiv \frac{1}{\sqrt{n}}(1+\lambda) v_2' P_z \varepsilon_1 - \frac{1}{\sqrt{n}}\lambda v_2' \varepsilon_1.$$

Recall that $\lambda = O(1)$ under Bekker's asymptotics. Therefore, we may write

$$\widetilde{H}_1 = \widetilde{H}_{1,1} + \widetilde{H}_{1,2} + \widetilde{H}_{1,3} + O_p\left(\frac{1}{n}\right),$$

where

$$\widetilde{H}_{1,1} = -\frac{1}{n} 2 (z\pi_2)' v_2 = O_p\left(\frac{1}{\sqrt{n}}\right),$$
$$\widetilde{H}_{1,2} = -\frac{1}{n}(1+\lambda)\left(v_2' P_z v_2 - K\sigma_{v_2}^2\right) = O_p\left(\frac{1}{\sqrt{n}}\right), \quad \text{(A.8)}$$
$$\widetilde{H}_{1,3} = \frac{1}{n}\lambda\left(v_2' v_2 - n\sigma_{v_2}^2\right) = O_p\left(\frac{1}{\sqrt{n}}\right). \quad \text{(A.9)}$$

Derivation of the second equality in (A.8) uses Bekker (1994, Lemma 1), which implies

$$\frac{1}{n}\operatorname{Var}\left(v_2' P_z v_2 - K\sigma_{v_2}^2\right) = 2\frac{K}{n}\left(\sigma_{v_2}^2\right)^2 + \frac{\sum d_i^2}{n}\left(E\left[v_{2i}^4\right] - 3\left(\sigma_{v_2}^2\right)^2\right),$$

and the fact that $\frac{\sum d_i^2}{n} \leq \frac{K}{n} = O(1)$.[1] Derivation of the second equality in (A.9) used the usual Central Limit Theorem. Similarly, we have

$$\widetilde{h}_1 = \widetilde{h}_{1,1} + \widetilde{h}_{1,2} + O_p\left(\frac{1}{\sqrt{n}}\right),$$

where

$$\widetilde{h}_{1,1} = \frac{1}{\sqrt{n}}(1+\lambda)\left(v_2' P_z \varepsilon_1 - K\sigma_{\varepsilon v_2}\right) = O_p(1), \quad \text{(A.10)}$$
$$\widetilde{h}_{1,2} = -\frac{\lambda}{\sqrt{n}}\left(v_2' \varepsilon_1 - n\sigma_{\varepsilon v_2}\right) = O_p(1). \quad \text{(A.11)}$$

[1] This is because $\sum d_i = K$ and $0 \leq d_i \leq 1$.

The second equalities in (A.10) and in (A.10) are derived by the same argument as in the derivation of those in (A.8) and (A.9). Using the expansion

$$\frac{1}{H_1 - \widetilde{H}_1} = \frac{1}{H_1} + \frac{1}{H_1^2}\widetilde{H}_1 + \frac{1}{H_1^3}\widetilde{H}_1^2 + \cdots,$$

we may write $\sqrt{n}$ times the far right side of (A.6) as

$$\begin{aligned}\frac{h_1 + \widetilde{h}_1}{H_1 + \widetilde{H}_1} &= \frac{h_1 + \widetilde{h}_{1,1} + \widetilde{h}_{1,2}}{H_1} + \frac{h_1 + \widetilde{h}_{1,1} + \widetilde{h}_{1,2}}{H_1^2}\left(\widetilde{H}_{1,1} + \widetilde{H}_{1,2} + \widetilde{H}_{1,3}\right) \\ &+ \frac{h_1 + \widetilde{h}_{1,1} + \widetilde{h}_{1,2}}{H_1^3}\left(\widetilde{H}_{1,1} + \widetilde{H}_{1,2} + \widetilde{H}_{1,3}\right)^2 + o_p(1) \\ &= \frac{h_1 + \widetilde{h}_{1,1} + \widetilde{h}_{1,2}}{H_1} + o_p(1). \qquad \text{(A.12)}\end{aligned}$$

We now take care of (A.7). Using the same argument as before, we may write $\sqrt{n}$ times the far right side of (A.7) as

$$\begin{aligned}\frac{h_2 + \widetilde{h}_2}{H_2 + \widetilde{H}_2} &= \frac{h_2 + \widetilde{h}_{2,1} + \widetilde{h}_{2,2}}{H_2} + \frac{h_2 + \widetilde{h}_{2,1} + \widetilde{h}_{2,2}}{H_2^2}\left(\widetilde{H}_{2,1} + \widetilde{H}_{2,2} + \widetilde{H}_{2,3}\right) \\ &+ \frac{h_2 + \widetilde{h}_{2,1} + \widetilde{h}_{2,2}}{H_2^3}\left(\widetilde{H}_{2,1} + \widetilde{H}_{2,2} + \widetilde{H}_{2,3}\right)^2 + o_p(1) \\ &= \frac{h_2 + \widetilde{h}_{2,1} + \widetilde{h}_{2,2}}{H_2} + o_p(1), \qquad \text{(A.13)}\end{aligned}$$

where

$$H_2 = \beta\Theta, \quad \widetilde{H}_2 = \frac{1}{n}(z\pi_2)' v_1 + \frac{1}{n}\beta(z\pi_2)' v_2 + \frac{1}{n}(1+\lambda) v_2' P_z v_1 - \frac{1}{n}\lambda v_2' v_1,$$

$$h_2 = \frac{1}{\sqrt{n}}\beta(z\pi_2)'\varepsilon_1, \quad \widetilde{h}_2 = \frac{1}{\sqrt{n}}(1+\lambda) v_1' P_z \varepsilon_1 - \frac{1}{\sqrt{n}}\lambda v_1'\varepsilon_1,$$

$$\begin{aligned}\widetilde{H}_{2,1} &= \frac{1}{n}(z\pi_2)' v_1 + \frac{1}{n}\beta(z\pi_2)' v_2 = O_p\left(\frac{1}{\sqrt{n}}\right), \\ \widetilde{H}_{2,2} &= \frac{1}{n}(1+\lambda)\left(v_2' P_z v_1 - K\sigma_{v_1 v_2}\right) = O_p\left(\frac{1}{\sqrt{n}}\right), \\ \widetilde{H}_{2,3} &= -\frac{1}{n}\lambda\left(v_2' v_1 - n\sigma_{v_1 v_2}\right) = O_p\left(\frac{1}{\sqrt{n}}\right),\end{aligned}$$

and

$$\widetilde{h}_{2,1} = \frac{1}{\sqrt{n}}(1+\lambda)\left(v_1' P_z \varepsilon_1 - K\sigma_{\varepsilon v_1}\right) = O_p(1), \qquad \widetilde{h}_{2,2} = -\frac{1}{\sqrt{n}}\lambda\left(v_1'\varepsilon_1 - n\sigma_{\varepsilon v_1}\right) = O_p(1).$$

Combining (A.12) and (A.13), we obtain an approximation to (A.5):

$$\begin{aligned}&\frac{h_1 + \widetilde{h}_{1,1} + \widetilde{h}_{1,2}}{H_1} - \frac{h_2 + \widetilde{h}_{2,1} + \widetilde{h}_{2,2}}{H_2} + o_p(1) \\ &\quad = -\frac{\frac{1}{\sqrt{n}}(1+\lambda)\left(\varepsilon_1' P_z \varepsilon_1 - K\sigma_\varepsilon^2\right) - \frac{1}{\sqrt{n}}\lambda\left(\varepsilon_1'\varepsilon_1 - n\sigma_\varepsilon^2\right)}{\beta\Theta} + o_p(1) \\ &\quad = -\frac{\frac{1}{\sqrt{n}}\left(\varepsilon_1' P_z \varepsilon_1 - K\sigma_\varepsilon^2\right) - \frac{1}{\sqrt{n}}\lambda\left(\varepsilon_1' M_z \varepsilon_1 - (n-K)\sigma_\varepsilon^2\right)}{\beta\Theta} + o_p(1)\end{aligned}$$

Using Bekker (1994, Lemma 1) and Condition (6.1), it can be shown that

$$\mathrm{Var}\left(\varepsilon_1' P_z \varepsilon_1 - K\sigma_\varepsilon^2\right) = 2K\left(\sigma_\varepsilon^2\right)^2 + \sum_{i=1}^n d_i^2 \cdot \left(E\left[\varepsilon_{1i}^4\right] - 3\left(\sigma_\varepsilon^2\right)^2\right),$$

$$\mathrm{Var}\left(\varepsilon_1' M_z \varepsilon_1 - (n-K)\sigma_\varepsilon^2\right) = 2(n-K)\left(\sigma_\varepsilon^2\right)^2 + \left(n - 2K + \sum_{i=1}^n d_i^2\right) \cdot \left(E\left[\varepsilon_{1i}^4\right] - 3\left(\sigma_\varepsilon^2\right)^2\right),$$

$$\mathrm{Cov}\left(\varepsilon_1' P_z \varepsilon_1 - K\sigma_\varepsilon^2, \varepsilon_1' M_z \varepsilon_1 - (n-K)\sigma_\varepsilon^2\right) = \left(K - \sum_{i=1}^n d_i^2\right) \cdot \left(E\left[\varepsilon_{1i}^4\right] - 3\left(\sigma_\varepsilon^2\right)^2\right).$$

We therefore obtain

$$\begin{aligned}
&\mathrm{Var}\left(\frac{1}{\sqrt{n}}\left(\varepsilon_1' P_z \varepsilon_1 - K\sigma_\varepsilon^2\right) - \frac{1}{\sqrt{n}}\lambda\left(\varepsilon_1' M_z \varepsilon_1 - (n-K)\sigma_\varepsilon^2\right)\right) \\
&= 2\left(\frac{K}{n} + \frac{\lambda^2}{n}(n-K)\right)\left(\sigma_\varepsilon^2\right)^2 \\
&+ \left(\frac{1}{n}\sum d_i^2 + \frac{\lambda^2}{n}\left(n - 2K + \sum d_i^2\right) - 2\frac{\lambda}{n}\left(K - \sum d_i^2\right)\right)\left(E\left[\varepsilon_{1i}^4\right] - 3\left(\sigma_\varepsilon^2\right)^2\right) \\
&= \frac{2\alpha}{1-\alpha}\left(\sigma_\varepsilon^2\right)^2 + \left(\left(\frac{1}{1-\alpha}\right)^2\left(\frac{\sum d_i^2}{n}\right) - \left(\frac{\alpha}{1-\alpha}\right)^2\right)\left(E\left[\varepsilon_{1i}^4\right] - 3\left(\sigma_\varepsilon^2\right)^2\right) + o(1),
\end{aligned}$$

from which we obtain the desired conclusion.

# B Two Endogenous Regressors

It can be seen that $(\beta_2, \beta_3)$ 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

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

We will call such estimators $(b_2, c_2)$, and $(b_3, c_3)$. Note that we have three estimators for $(\beta_2, \beta_3)$: $(b_1, c_1)$, $\left(\frac{1}{b_2}, -\frac{c_2}{b_2}\right)$, $\left(-\frac{c_3}{b_3}, \frac{1}{b_3}\right)$.

## B.1 A Technical Lemma

Utilizing Bekker (1994, Lemma 2) again, we can establish that

**Lemma 2**

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

*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* $\Lambda^{\perp}$*, which are summarized below in Tables A.1 and A.2.*

## B.2 Nagar Based Specification Test

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 $\frac{2\alpha\,\mathrm{Var}(\varepsilon_{1i})^2}{1-\alpha}$ 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 $\frac{2\alpha\,\mathrm{Var}(\varepsilon_{1i})^2}{1-\alpha}\frac{1}{{\beta_2}^2\left(\Theta_{22}-\frac{{\Theta_{23}}^2}{\Theta_{33}}\right)^2}$. Observe that $\mathrm{Var}(\varepsilon_{1i})$ and $\beta_2$ can be estimated consistently utilizing the consistency of LIML. Also note that

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

for any consistent estimator $\widehat{\alpha}$ of $\alpha$. 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 A.1: $\overline{\Lambda}$

| | |
|---|---|
| $\overline{\Lambda}_{11}$ | $4\omega_{11}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+2\alpha\omega_{11}^2$ |
| $\overline{\Lambda}_{12}$ | $\omega_{11}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+\omega_{12}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)$ $+\omega_{11}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+\omega_{12}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+\alpha\left(\omega_{11}\omega_{12}+\omega_{11}\omega_{12}\right)$ |
| $\overline{\Lambda}_{13}$ | $\omega_{11}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+\omega_{13}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)$ $+\omega_{11}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+\omega_{13}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+\alpha\left(\omega_{11}\omega_{13}+\omega_{11}\omega_{13}\right)$ |
| $\overline{\Lambda}_{14}$ | $4\omega_{12}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+2\alpha\omega_{12}^2$ |
| $\overline{\Lambda}_{15}$ | $2\omega_{13}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+2\omega_{12}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+2\alpha\omega_{12}\omega_{13}$ |
| $\overline{\Lambda}_{16}$ | $4\omega_{13}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+2\alpha\omega_{13}^2$ |
| $\overline{\Lambda}_{22}$ | $\omega_{11}\Theta_{22}+\omega_{22}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)$ $+2\omega_{12}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+\alpha\left(\omega_{11}\omega_{22}+\omega_{12}^2\right)$ |
| $\overline{\Lambda}_{23}$ | $\omega_{11}\Theta_{23}+\omega_{23}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)+\omega_{12}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)$ $+\omega_{13}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+\alpha\left(\omega_{11}\omega_{23}+\omega_{12}\omega_{13}\right)$ |
| $\overline{\Lambda}_{24}$ | $\omega_{22}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+\omega_{12}\Theta_{22}+\omega_{22}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)$ $+\omega_{12}\Theta_{22}+\alpha\left(\omega_{22}\omega_{12}+\omega_{22}\omega_{12}\right)$ |
| $\overline{\Lambda}_{25}$ | $\omega_{22}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+\omega_{13}\Theta_{22}+\omega_{12}\Theta_{23}$ $+\omega_{23}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+\alpha\left(\omega_{22}\omega_{13}+\omega_{12}\omega_{23}\right)$ |
| $\overline{\Lambda}_{26}$ | $2\omega_{23}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+2\omega_{13}\Theta_{23}+2\alpha\omega_{13}\omega_{23}$ |
| $\overline{\Lambda}_{33}$ | $\omega_{11}\Theta_{33}+\omega_{33}\left(\beta_2^2\Theta_{22}+2\beta_2\beta_3\Theta_{23}+\beta_3^2\Theta_{33}\right)$ $+2\omega_{13}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+\alpha\left(\omega_{11}\omega_{33}+\omega_{13}^2\right)$ |
| $\overline{\Lambda}_{34}$ | $2\omega_{23}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+2\omega_{12}\Theta_{23}+2\alpha\omega_{12}\omega_{23}$ |
| $\overline{\Lambda}_{35}$ | $\omega_{33}\left(\beta_2\Theta_{22}+\beta_3\Theta_{23}\right)+\omega_{12}\Theta_{33}+\omega_{13}\Theta_{23}$ $+\omega_{23}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+\alpha\left(\omega_{33}\omega_{12}+\omega_{13}\omega_{23}\right)$ |
| $\overline{\Lambda}_{36}$ | $\omega_{33}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)+\omega_{13}\Theta_{33}+\omega_{33}\left(\beta_2\Theta_{23}+\beta_3\Theta_{33}\right)$ $+\omega_{13}\Theta_{33}+\alpha\left(\omega_{33}\omega_{13}+\omega_{33}\omega_{13}\right)$ |
| $\overline{\Lambda}_{44}$ | $4\omega_{22}\Theta_{22}+2\alpha\omega_{22}^2$ |
| $\overline{\Lambda}_{45}$ | $\omega_{22}\Theta_{23}+\omega_{23}\Theta_{22}+\omega_{22}\Theta_{23}+\omega_{23}\Theta_{22}+\alpha\left(\omega_{22}\omega_{23}+\omega_{22}\omega_{23}\right)$ |
| $\overline{\Lambda}_{46}$ | $4\omega_{23}\Theta_{23}+2\alpha\omega_{23}^2$ |
| $\overline{\Lambda}_{55}$ | $\omega_{22}\Theta_{33}+\omega_{33}\Theta_{22}+2\omega_{23}\Theta_{23}+\alpha\left(\omega_{22}\omega_{33}+\omega_{23}^2\right)$ |
| $\overline{\Lambda}_{56}$ | $\omega_{33}\Theta_{23}+\omega_{23}\Theta_{33}+\omega_{33}\Theta_{23}+\omega_{23}\Theta_{33}+\alpha\left(\omega_{33}\omega_{23}+\omega_{33}\omega_{23}\right)$ |
| $\overline{\Lambda}_{66}$ | $4\omega_{33}\Theta_{33}+2\alpha\omega_{33}^2$ |

Table A.2: $\Lambda^{\perp}$

| | | | |
|---|---|---|---|
| $\Lambda^{\perp}_{11}$ | $2(1-\alpha)\,\omega_{11}^2$ | $\Lambda^{\perp}_{33}$ | $(1-\alpha)\left(\omega_{11}\omega_{33}+\omega_{13}^2\right)$ |
| $\Lambda^{\perp}_{12}$ | $(1-\alpha)\,(\omega_{11}\omega_{12}+\omega_{11}\omega_{12})$ | $\Lambda^{\perp}_{34}$ | $2(1-\alpha)\,\omega_{12}\omega_{23}$ |
| $\Lambda^{\perp}_{13}$ | $(1-\alpha)\,(\omega_{11}\omega_{13}+\omega_{11}\omega_{13})$ | $\Lambda^{\perp}_{35}$ | $(1-\alpha)\,(\omega_{33}\omega_{12}+\omega_{13}\omega_{23})$ |
| $\Lambda^{\perp}_{14}$ | $2(1-\alpha)\,\omega_{12}^2$ | $\Lambda^{\perp}_{36}$ | $(1-\alpha)\,(\omega_{33}\omega_{13}+\omega_{33}\omega_{13})$ |
| $\Lambda^{\perp}_{15}$ | $2(1-\alpha)\,\omega_{12}\omega_{13}$ | $\Lambda^{\perp}_{44}$ | $2(1-\alpha)\,\omega_{22}^2$ |
| $\Lambda^{\perp}_{16}$ | $2(1-\alpha)\,\omega_{13}^2$ | $\Lambda^{\perp}_{45}$ | $(1-\alpha)\,(\omega_{22}\omega_{23}+\omega_{22}\omega_{23})$ |
| $\Lambda^{\perp}_{22}$ | $(1-\alpha)\left(\omega_{11}\omega_{22}+\omega_{12}^2\right)$ | $\Lambda^{\perp}_{46}$ | $2(1-\alpha)\,\omega_{23}^2$ |
| $\Lambda^{\perp}_{23}$ | $(1-\alpha)\,(\omega_{11}\omega_{23}+\omega_{12}\omega_{13})$ | $\Lambda^{\perp}_{55}$ | $(1-\alpha)\left(\omega_{22}\omega_{33}+\omega_{23}^2\right)$ |
| $\Lambda^{\perp}_{24}$ | $(1-\alpha)\,(\omega_{22}\omega_{12}+\omega_{22}\omega_{12})$ | $\Lambda^{\perp}_{56}$ | $(1-\alpha)\,(\omega_{33}\omega_{23}+\omega_{33}\omega_{23})$ |
| $\Lambda^{\perp}_{25}$ | $(1-\alpha)\,(\omega_{22}\omega_{13}+\omega_{12}\omega_{23})$ | $\Lambda^{\perp}_{66}$ | $2(1-\alpha)\,\omega_{33}^2$ |
| $\Lambda^{\perp}_{26}$ | $2(1-\alpha)\,\omega_{13}\omega_{23}$ | | |

### B.3 2SLS Based Specification Test

It can be shown that

$$b_1 - \frac{1}{b_2} = \frac{N_1}{D_1} - \frac{N_2}{D_2}, \qquad b_{2SLS,1} - \frac{1}{b_{2SLS,2}} = \frac{n_1}{d_1} - \frac{n_2}{d_2},$$

and

$$d_1 d_2 \left(\frac{n_1}{d_1} - \frac{n_2}{d_2} - \widehat{C}\right) = D_1 D_2 \left(\frac{N_1}{D_1} - \frac{N_2}{D_2}\right)$$

Note that, by design, the two test statistics would have unitary (asymptotic) correlation. Furthermore, the asymptotic variance of 2SLS based test would be equal to $\left(\operatorname{plim}\frac{1}{n}D_1 \cdot \operatorname{plim}\frac{1}{n}D_2\right) / \left(\operatorname{plim}\frac{1}{n}d_1 \cdot \operatorname{plim}\frac{1}{n}d_2\right)$ times the asymptotic variance of the corresponding Nagar based test. Because $\operatorname{plim}\frac{1}{n}D_1 = \Theta_{22}\Theta_{33} - \Theta_{23}^2$, $\operatorname{plim}\frac{1}{n}D_2 = -\Theta_{33}\beta_2\Theta_{22} + \beta_2\Theta_{23}^2$, and because the asymptotic variance of the Nagar based test is equal to $\frac{2\alpha \operatorname{Var}(\varepsilon_{1i})^2}{1-\alpha}\frac{1}{{\beta_2}^2\left(\Theta_{22}-\frac{{\Theta_{23}}^2}{\Theta_{33}}\right)^2}$, we can conclude that the asymptotic variance of the 2SLS based test is equal to

$$\frac{2\alpha \operatorname{Var}(\varepsilon_{1i})^2}{1-\alpha} \cdot \frac{\Theta_{33}^2\left(\Theta_{22}\Theta_{33} - \Theta_{23}^2\right)^2}{\left(\operatorname{plim}\frac{1}{n}d_1\right)^2 \left(\operatorname{plim}\frac{1}{n}d_2\right)^2},$$

where

$$\begin{aligned}
\operatorname{plim}\frac{1}{n}d_1 &= \Theta_{22}\Theta_{33} + \Theta_{22}\alpha\omega_{33} + \alpha\omega_{22}\Theta_{33} + \alpha^2\omega_{33}\omega_{22} - \Theta_{23}^2 - 2\Theta_{23}\alpha\omega_{23} - \alpha^2\omega_{23}^2, \\
\operatorname{plim}\frac{1}{n}d_2 &= -\Theta_{33}\beta_2\Theta_{22} - \Theta_{33}\alpha\omega_{12} - \alpha\omega_{33}\beta_2\Theta_{22} - \alpha\omega_{33}\beta_3\Theta_{23} - \alpha^2\omega_{33}\omega_{12} \\
&\quad + \beta_2\Theta_{23}^2 + \Theta_{23}\alpha\omega_{13} + \alpha\omega_{23}\beta_2\Theta_{23} + \alpha\omega_{23}\beta_3\Theta_{33} + \alpha^2\omega_{23}\omega_{13}.
\end{aligned}$$

As for the estimator of the asymptotic variance, we can simply use $\frac{D_1^2 D_2^2}{d_1^2 d_2^2}$ times a consistent estimator for the asymptotic variance of the Nagar based test.

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

| $n$ | $K$ | Cov(ε,$v_2$) | (a) | (b) | (c) | (d) | (e) | (f) | (g) | (h) | (i) | (j) | (k) | (l) | (m) | (n) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 |
| 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 |

(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) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 |
| 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 |

(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) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 |
| 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 |

(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) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 |
| 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 |

(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

| $n$ | $K$ | $\omega_{12}$ | $\widetilde{R}_f^2$ | $\beta$ | (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: $\omega_{12}$ 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 $\widetilde{R}_f^2 = .001$.

Table 7: Nonnormal Error with $\widetilde{R}_f^2 = .1$

| n | K | Cov(ε,$v_2$) | (a) | (b) | (c) | (d) | (e) | (f) | (g) | (h) | (i) | (j) | (k) | (l) | (m) | (n) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 250 | 5 | -0.9 | 0.093 | 0.059 | -0.142 | -0.025 | -0.103 | 0.194 | 0.130 | 0.083 | 0.053 | -0.188 | 0.053 | 0.116 | 0.136 | 0.086 |
| 250 | 5 | -0.5 | 0.098 | 0.067 | -0.075 | 0.130 | -0.162 | 0.110 | 0.059 | 0.088 | 0.060 | 0.012 | 0.075 | 0.073 | 0.098 | 0.052 |
| 250 | 5 | 0.5 | 0.103 | 0.070 | 0.080 | -0.153 | 0.187 | 0.115 | 0.061 | 0.088 | 0.062 | -50.38 | -0.106 | -0.029 | 0.097 | 0.052 |
| 250 | 5 | 0.9 | 0.088 | 0.055 | 0.140 | 0.024 | 0.103 | 0.184 | 0.123 | 0.075 | 0.045 | -0.006 | -0.050 | -0.072 | 0.130 | 0.082 |
| 250 | 10 | -0.9 | 0.088 | 0.047 | -0.283 | -0.099 | -0.185 | 0.300 | 0.217 | 0.061 | 0.036 | -0.046 | -0.013 | 0.072 | 0.144 | 0.093 |
| 250 | 10 | -0.5 | 0.079 | 0.046 | -0.159 | 0.255 | -0.392 | 0.128 | 0.073 | 0.077 | 0.049 | 0.047 | 0.066 | 0.058 | 0.099 | 0.054 |
| 250 | 10 | 0.5 | 0.081 | 0.041 | 0.158 | -0.253 | 0.390 | 0.122 | 0.064 | 0.074 | 0.045 | -0.076 | -0.080 | -0.035 | 0.093 | 0.049 |
| 250 | 10 | 0.9 | 0.086 | 0.046 | 0.287 | 0.104 | 0.180 | 0.310 | 0.226 | 0.060 | 0.036 | -0.077 | -0.061 | -0.069 | 0.147 | 0.093 |
| 250 | 30 | -0.9 | 0.132 | 0.098 | -0.545 | -0.287 | -0.263 | 0.621 | 0.514 | 0.073 | 0.046 | -2.191 | 0.274 | 0.095 | 0.155 | 0.107 |
| 250 | 30 | -0.5 | 0.106 | 0.065 | -0.305 | 0.492 | -0.798 | 0.145 | 0.074 | 0.066 | 0.041 | 0.433 | 0.366 | 0.052 | 0.088 | 0.041 |
| 250 | 30 | 0.5 | 0.111 | 0.067 | 0.304 | -0.488 | 0.796 | 0.140 | 0.073 | 0.065 | 0.034 | 0.749 | -0.136 | -0.048 | 0.078 | 0.038 |
| 250 | 30 | 0.9 | 0.132 | 0.098 | 0.544 | 0.285 | 0.264 | 0.620 | 0.515 | 0.076 | 0.046 | 0.100 | -0.058 | -0.106 | 0.159 | 0.107 |
| 1000 | 5 | -0.9 | 0.126 | 0.092 | -0.037 | -0.008 | -0.022 | 0.127 | 0.070 | 0.102 | 0.068 | -0.001 | 0.006 | 0.011 | 0.112 | 0.057 |
| 1000 | 5 | -0.5 | 0.130 | 0.094 | -0.021 | 0.033 | -0.041 | 0.107 | 0.058 | 0.106 | 0.077 | -0.001 | 0.013 | 0.006 | 0.103 | 0.055 |
| 1000 | 5 | 0.5 | 0.128 | 0.098 | 0.019 | -0.035 | 0.041 | 0.107 | 0.059 | 0.105 | 0.079 | -0.001 | -0.015 | -0.008 | 0.103 | 0.058 |
| 1000 | 5 | 0.9 | 0.112 | 0.080 | 0.035 | 0.007 | 0.022 | 0.114 | 0.062 | 0.087 | 0.061 | -0.001 | -0.007 | -0.011 | 0.101 | 0.053 |
| 1000 | 10 | -0.9 | 0.093 | 0.057 | -0.087 | -0.028 | -0.054 | 0.167 | 0.101 | 0.082 | 0.054 | 0.003 | 0.009 | 0.013 | 0.116 | 0.063 |
| 1000 | 10 | -0.5 | 0.093 | 0.061 | -0.048 | 0.064 | -0.101 | 0.108 | 0.058 | 0.085 | 0.059 | 0.002 | 0.014 | 0.008 | 0.094 | 0.050 |
| 1000 | 10 | 0.5 | 0.099 | 0.063 | 0.048 | -0.064 | 0.102 | 0.113 | 0.059 | 0.091 | 0.061 | -0.002 | -0.014 | -0.008 | 0.100 | 0.051 |
| 1000 | 10 | 0.9 | 0.092 | 0.052 | 0.089 | 0.030 | 0.054 | 0.163 | 0.099 | 0.082 | 0.049 | 0.000 | -0.006 | -0.010 | 0.110 | 0.061 |
| 1000 | 30 | -0.9 | 0.095 | 0.043 | -0.243 | -0.102 | -0.140 | 0.373 | 0.269 | 0.085 | 0.040 | 0.010 | 0.010 | 0.013 | 0.112 | 0.064 |
| 1000 | 30 | -0.5 | 0.096 | 0.047 | -0.135 | 0.171 | -0.300 | 0.148 | 0.080 | 0.090 | 0.051 | 0.006 | 0.016 | 0.009 | 0.097 | 0.051 |
| 1000 | 30 | 0.5 | 0.093 | 0.049 | 0.137 | -0.169 | 0.299 | 0.147 | 0.080 | 0.090 | 0.050 | -0.004 | -0.013 | -0.007 | 0.098 | 0.049 |
| 1000 | 30 | 0.9 | 0.098 | 0.047 | 0.245 | 0.102 | 0.141 | 0.385 | 0.289 | 0.090 | 0.044 | -0.009 | -0.010 | -0.014 | 0.120 | 0.068 |

(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.**