ALTERNATIVE APPROXIMATIONS OF THE BIAS AND MSE
OF THE IV ESTIMATOR UNDER WEAK IDENTIFICATION
WITH AN APPLICATION TO BIAS CORRECTION

By

John C. Chao and Norman R. Swanson

May 2003

Alternative Approximations of the Bias and MSE of the IV Estimator Under Weak Identification With an Application to Bias Correction*

John Chao
University of Maryland

Norman R. Swanson
Rutgers University

First Draft: January 2000
Current Version: May 2003

Abstract

We provide analytical formulae for the asymptotic bias (ABIAS) and mean squared error (AMSE) of the IV estimator, and obtain approximations thereof based on an asymptotic scheme which essentially requires the expectation of the first stage F-statistic to converge to a finite (possibly small) positive limit as the number of instruments approaches infinity. The approximations so obtained are shown, via regression analysis, to yield good approximations for ABIAS and AMSE functions, and the AMSE approximation is shown to perform well relative to the approximation of Donald and Newey (2001). Additionally, the manner in which our framework generalizes that of Richardson and Wu (1971) is discussed. One consequence of the asymptotic framework adopted here is that consistent estimators for the ABIAS and AMSE can be obtained. As a result, we are able to construct a number of bias corrected OLS and IV estimators, which we show to be consistent under a sequential asymptotic scheme. These bias-corrected estimators are also robust, in the sense that they remain consistent in a conventional asymptotic setup, where the model is fully identified. A small Monte Carlo experiment documents the relative performance of our bias adjusted estimators versus standard IV, OLS, LIML estimators, and it is shown that our estimators have lower bias than LIML for various levels of endogeneity and instrument relevance.

---

1 Introduction

Over the last decade there have been a great number of papers written on the subject of instrumental variables (IV) regression with instruments that are only weakly correlated with the endogenous explanatory variables. A very few of the important recent contributions include Nelson and Startz (1990a), Dufour (1997), Shea (1997), Staiger and Stock (1997), Wang and Zivot (1998), Hahn and Inoue (2000), Hahn and Kuersteiner (2002), Stock, Wright and Yogo (2002), Stock and Yogo (2002), and the references contained therein.[1] Important related contributions include Bekker (1994), Hall, Rudebusch and Wilcox (1996), and Hall and Peixe (2000). Much of this literature focuses on the impact that the use of weak instruments has on interval estimation and on hypothesis testing, although there are also notable results on the properties of point estimators. Both of these areas are of interest to applied researchers who first documented the weak instrument problem in empirical work (see e.g. Nelson and Startz (1990b), Bound, Jaeger, and Baker (1995), and Angrist and Krueger (1995)).

This paper focuses on point estimation properties. In particular, we focus on the IV estimator, and give new asymptotic bias (ABIAS) and asymptotic mean-squared error (AMSE) approximations based on an examination of explicit analytical formulae for the ABIAS and AMSE, under the local-to-zero framework.[2] These approximations can be viewed as having been derived from a sequential limit procedure whereby the sample size, T, is first allowed to grow to infinity followed by the passage to infinity of the number of instruments, k_{21}. The AMSE approximation so derived provides an alternative to the MSE approximation of the 2SLS estimator obtained in Donald and Newey (2001) under a different asymptotic scheme. Numerical calculations provided in this paper suggest that our first order AMSE approximation outperforms the approximation of Donald and Newey (2001), particularly in cases with small values of the concentration parameter (i.e., cases with weak instruments)[3]. Additionally, when the approximation method is applied to the bias, the lead term of the expansion (when appropriately standardized by the ABIAS of the OLS estimator) is exactly the relative bias measure given in Staiger and Stock (1997) in the case where there is only one endogenous regressor. Furthermore, the lead term of the MSE expansion is the square of the lead term of the bias expansion, implying that the variance component of the MSE is of a lower order vis-a-vis the bias component in a scenario where the number of instruments used is large relative to the value of the population analogue of the first stage F-statistic. In order to tie our findings in with the IV literature, we note also that our formulae for the asymptotic bias and MSE, correspond to the exact bias and MSE functions of the 2SLS estimator, as derived by Richardson and Wu (1971), when a fixed instrument/Gaussian model is assumed, and in this sense our findings generalize their results to the more general setting with possibly non-normal errors and stochastic instruments.

A consequence of the sequential limit approach which we adopt here is that consistent estimators for

---

[1] A related literature which examines the implications for statistical inference when the underlying simultaneous equations model is partially identified includes papers by Phillips (1989), Choi and Phillips (1992), and Kitamura (1994).

[2] Another approach used to examine the behavior of the IV/2SLS estimator under weak identification, which is based on conditional inference, is discussed in Forchini and Hillier (1999) and Moreira (2001, 2002).

[3] The performance of the MSE approximation of Donald and Newey (2001) has also been examined in the Monte Carlo study reported in Hahn, Hausman, and Kuersteiner (2002). Although the latter paper does not study our approximation, the results they obtained on the Donald-Newey approximation are in rough agreement with the results reported in section 3 of this paper.

the ABIAS and AMSE can be obtained.[4] This, in turn, enables us to construct bias-corrected OLS and IV estimators, which consistently estimate the structural coefficient of the IV regression even when the instruments are weak in the local-to-zero sense. In addition, we show that in the conventional setup where the model is fully identified, all but one of our proposed bias-corrected estimators remain consistent. We include a small Monte Carlo experiment in order to document the relative performance of our bias adjusted estimators as compared with unadjusted IV and OLS estimators and also with the LIML estimator.

This rest of the paper is organized as follows. Section 2 contains preliminaries, and Section 3 presents formulae for the ABIAS and AMSE, as well as approximations thereof. In Section 4 we discuss consistent estimation of the ABIAS and AMSE, and suggest a number of bias corrected OLS and IV estimators. Monte Carlo results are discussed in Section 5, and Section 6 concludes. All proofs are contained in two appendices. Before proceeding, we briefly introduce some notation. In the sequel, the symbols "$\Longrightarrow$" and "$\equiv$" denote convergence and equivalence in distribution, respectively. Also, $P_X = X(X'X)^{-1}X'$ is the matrix which projects orthogonally onto the range space of X and $M_X = I - P_X$.

2 Setup

Consider the simultaneous equations model (SEM):

$$
\begin{align}
y_1 &= y_2\beta + X\gamma + u, \tag{1}\\
y_2 &= Z\Pi + X\Phi + v, \tag{2}
\end{align}
$$

where y_1 and y_2 are $T \times 1$ vectors of observations on the two endogenous variables, X is an $T \times k_1$ matrix of observations on k_1 exogenous variables included in the structural equation (1), Z is a $T \times k_2$ matrix of observations on k_2 exogenous variables excluded from the structural equation, and u and v are $T \times 1$ vectors of random disturbances[5]. Let u_t and v_t denote the tth component of the random vectors u and v, respectively; and let Z_t' and X_t' denote the tth row of the matrices Z and X, respectively. Additionally, let $w_t = (u_t, v_t)'$ (or $w = (u, v)$) and let $\overline{Z}_t = (X_t', Z_t')'$ (or $\overline{Z} = (X, Z)$); assume that $E(w_t) = \mathbf{0}$, $E(w_t w_t') = \Sigma = \begin{pmatrix} \sigma_{uu} & \sigma_{uv} \\ \sigma_{uv} & \sigma_{vv} \end{pmatrix}$, and $E\overline{Z}_t w_t' = \mathbf{0}$ for all t and assume that $E(w_t w_s') = \mathbf{0}$ for all $t \neq s$, where $t, s = 1, ..., T$. Following Staiger and Stock (1997), we make the following assumptions.

Assumption 1: $\Pi = \Pi_T = C/\sqrt{T}$, where C is a fixed $k_2 \times 1$ vector.

Assumption 2: *The following limits hold jointly:* (i) $(u'u/T, u'v/T, v'v/T) \xrightarrow{p} (\sigma_{uu}, \sigma_{uv}, \sigma_{vv})$, (ii) $\overline{Z}'\overline{Z}/T \xrightarrow{p} Q$, and (iii) $(T^{-1/2}u'X, T^{-1/2}u'Z, T^{-1/2}v'X, T^{-1/2}v'Z)' \Longrightarrow (\psi_{Xu}', \psi_{Zu}', \psi_{Xv}', \psi_{Zv}')'$, where $Q = E(\overline{Z}_t \overline{Z}_t')$ and where $\psi \equiv (\psi_{Xu}', \psi_{Zu}', \psi_{Xv}', \psi_{Zv}')'$ is distributed $N(\mathbf{0}, (\Sigma \otimes Q))$.

We consider IV estimation of β, where the IV estimator may not make use of all available instruments. Define $\widehat{\beta}_{IV} = (y_2'(P_H - P_X)y_2)^{-1}(y_2'(P_H - P_X)y_1)$, where $H = (Z_1, X)$ is an $T \times (k_{21} + k_1)$ matrix of instruments, and Z_1 is an $T \times k_{21}$ submatrix of Z formed by column selection. It will prove convenient to partition Z as $Z = (Z_1, Z_2)$, where Z_2 is an $T \times k_{22}$ matrix of observations of the excluded exogenous variables

---

[4] Chao and Swanson (2001) discuss related consistency properties under a pathwise asymptotic scheme whereby the number of instruments approaches infinity as a function of the sample size.

[5] Although for notational simplicity we only study the case with one endogenous explanatory variable in this paper, we do not see any reason why many of the qualitative conclusions reached here will not continue to hold in more general settings.

not used as instruments in estimation. Note that when $Z_1 = Z$ and $H = [Z, X]$ (i.e. when all available instruments are used), the IV estimator defined above is equivalent to the 2SLS estimator. Additionally, partition Π_T, $T^{-\frac{1}{2}} Z'u$, $T^{-\frac{1}{2}} Z'v$, ψ_{Zu}, and ψ_{Zv} conformably with $Z = (Z_1, Z_2)$ by writing $\Pi_T = (\Pi'_{1,T},$ $\Pi'_{2,T})' = (C'_1/\sqrt{T}, C'_2/\sqrt{T})'$, $T^{-\frac{1}{2}} Z'u = (T^{-\frac{1}{2}} u' Z_1, T^{-\frac{1}{2}} u' Z_2)'$, $T^{-\frac{1}{2}} Z'v = (T^{-\frac{1}{2}} v' Z_1, T^{-\frac{1}{2}} v' Z_2)'$, $\psi_{Zu} = (\psi'_{Z_1 u}, \psi'_{Z_2 u})'$, and $\psi_{Zv} = (\psi'_{Z_1 v}, \psi'_{Z_2 v})'$, where from part (iii) of Assumption 2 we have that $(T^{-\frac{1}{2}} u' Z_1,$ $T^{-\frac{1}{2}} u' Z_2, T^{-\frac{1}{2}} v' Z_1, T^{-\frac{1}{2}} v' Z_2)' \Rightarrow (\psi'_{Z_1 u}, \psi'_{Z_2 u}, \psi'_{Z_1 v}, \psi'_{Z_2 v})'$. Furthermore, partition Q conformably with $\overline{Z} = (X, Z_1, Z_2)$ as

$$Q = \begin{pmatrix} Q_{XX} & Q_{XZ_1} & Q_{XZ_2} \\ Q_{Z_1 X} & Q_{Z_1 Z_1} & Q_{Z_1 Z_2} \\ Q_{Z_2 X} & Q_{Z_2 Z_1} & Q_{Z_2 Z_2} \end{pmatrix}. \tag{3}$$

Finally, define

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega'_{12} & \Omega_{22} \end{pmatrix} = \begin{pmatrix} Q_{Z_1 Z_1} - Q_{Z_1 X} Q_{XX}^{-1} Q_{XZ_1} & Q_{Z_1 Z_2} - Q_{Z_1 X} Q_{XX}^{-1} Q_{XZ_2} \\ Q_{Z_2 Z_1} - Q_{Z_2 X} Q_{XX}^{-1} Q_{XZ_1} & Q_{Z_2 Z_2} - Q_{Z_2 X} Q_{XX}^{-1} Q_{XZ_2} \end{pmatrix} \tag{4}$$

and $\Omega_{1*} = (\Omega_{11}, \Omega_{12})$. To ensure that the ABIAS and AMSE of the IV estimator are well-behaved assume that:

Assumption 3: *There exists finite positive integer T_0 such that $\sup_{T \geq T_0} E(|U_T|^{2+\delta}) < \infty$, for some $\delta > 0$, where $U_T = \widehat{\beta}_{IV,T} - \beta_0$, $\widehat{\beta}_{IV,T}$ denotes the IV estimator of β for a sample of size T , and β_0 is the true value of β.*

Assumption 3 is sufficient for the uniform integrability of $(\widehat{\beta}_{IV,T} - \beta_0)^2$ (see Billingsley (1968), pp. 32). Under Assumption 3, $\lim_{T \to \infty} E(\widehat{\beta}_{IV,T} - \beta_0) = E(U)$ and $\lim_{T \to \infty} E(\widehat{\beta}_{IV,T} - \beta_0)^2 = E(U^2)$, where U is the limiting random variable of the sequence $\{U_T\}$, whose explicit form is given in Lemma A1 in Appendix A. Hence, the ABIAS and AMSE correspond to the bias and MSE implied by the limiting distribution of $\widehat{\beta}_{IV,T}$. Note also that for the special case where $(u_t, v_t)' \sim i.i.d. N(0, \Sigma)$, $k_{21} \geq 4$ implies Assumption 3 (see e.g. Sawa (1969)). Throughout this paper, we shall assume $k_{21} \geq 4$ so as to ensure that our results also apply in the Gaussian case. In addition, note that Assumption 3 rules out the limited information maximum likelihood (LIML) estimator in the Gaussian case since no positive integer moment exists for the finite sample distribution of LIML in this case (see e.g. Mariano and McDonald (1979) and Phillips (1984, 1985)).

3   Bias and MSE Formulae and Their Approximations

We begin with a proposition stating the ABIAS and AMSE in our context.

Proposition 3.1[6]**:** *Given the SEM described above, and under Assumptions 1, 2, and 3, the following results hold for $k_{21} \geq 4$:*

(Bias)

(a)

$$b_{\widehat{\beta}_{IV}} (\mu'\mu, k_{21}) = \sigma_{uu}^{1/2} \sigma_{vv}^{-1/2} \rho e^{-\frac{\mu'\mu}{2}} \, {}_1F_1 \left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\mu'\mu}{2} \right), \tag{5}$$

---

[6] The proof of this proposition follows directly from the exact results of Richardson and Wu (1971) given the equivalence of the local-to-zero asymptotic distribution of the IV estimator and its exact distribution under Gaussian errors, as stated in Staiger and Stock (1997). An alternative proof of this proposition is also given in Chao and Swanson (2000).

where $b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = \lim_{T\to\infty} E(\widehat{\beta}_{IV,T} - \beta_0)$ is the asymptotic bias function of the IV estimator which we write as a function of $\mu'\mu = \sigma_{vv}^{-1}C'\Omega_{1*}'\Omega_{11}^{-1}\Omega_{1*}C$ and k_{21}, and where $\rho = \sigma_{uv}\sigma_{uu}^{-\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}$, $\Gamma(\cdot)$ denotes the gamma function, and $_1F_1(\cdot;\cdot;\cdot)$ denotes the confluent hypergeometric function;

(b) For k_{21} fixed, as $\mu'\mu \to \infty$, $b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) \to 0$;

(c) For $\mu'\mu$ fixed, as $k_{21} \to \infty$, $b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) \to \sigma_{uv}/\sigma_{vv} = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho$;

(d) The absolute value of the asymptotic bias function (i.e. $|b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21})|$) is a monotonically decreasing function of $\mu'\mu$ for k_{21} fixed and $\sigma_{uv} \neq 0$;

(e) The absolute value of the bias function (i.e. $|b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21})|$) is a monotonically increasing function of k_{21} for $\mu'\mu$ fixed and $\sigma_{uv} \neq 0$;

(MSE)

(f)

$$m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\left[\frac{1}{\rho^2}\left(\frac{1}{k_{21}-2}\right){}_1F_1\left(\frac{k_{21}}{2}-1;\frac{k_{21}}{2};\frac{\mu'\mu}{2}\right)\right.$$
$$\left.+\left(\frac{k_{21}-3}{k_{21}-2}\right){}_1F_1\left(\frac{k_{21}}{2}-2;\frac{k_{21}}{2};\frac{\mu'\mu}{2}\right)\right], \tag{6}$$

where $m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = \lim_{T\to\infty} E\left(\widehat{\beta}_{IV,T} - \beta_0\right)^2$ is the asymptotic mean squared error function of the IV estimator;

(g) For k_{21} fixed, as $\mu'\mu \to \infty$, $m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) \to 0$;

(h) For $\mu'\mu$ fixed, as $k_{21} \to \infty$, $m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) \to \sigma_{uv}^2/\sigma_{vv}^2 = \sigma_{uu}\sigma_{vv}^{-1}\rho^2$;

(i) The asymptotic mean squared error function $m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21})$ is a monotonically decreasing function of $\mu'\mu$ for k_{21} fixed and $\sigma_{uv} \neq 0$.

It is well known that confluent hypergeometric functions have infinite series representations, so that we can also write:

$$b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho e^{-\frac{\mu'\mu}{2}}\left[\sum_{j=0}^{\infty}\frac{(\frac{k_{21}}{2}-1)_j}{(\frac{k_{21}}{2})_j}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\right], \tag{7}$$

$$m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu}{2}}\left[\frac{1}{\rho^2}\left(\frac{1}{k_{21}-2}\right)\sum_{j=0}^{\infty}\frac{(\frac{k_{21}}{2}-1)_j}{(\frac{k_{21}}{2})_j}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\right.$$
$$\left.+\left(\frac{k_{21}-3}{k_{21}-2}\right)\sum_{j=0}^{\infty}\frac{(\frac{k_{21}}{2}-2)_j}{(\frac{k_{21}}{2})_j}\frac{\left(\frac{\mu'\mu}{2}\right)^j}{j!}\right]. \tag{8}$$

One merit of such infinite series representations is that they provide explicit formulae for the ABIAS and AMSE of $\widehat{\beta}_{IV}$ under weak identification, which can be used in numerical calculations, as is done below.

Note that $|b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21})| < |\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho|$. Hence, even when the instruments are weak in the sense of Staiger and Stock (1997), the ABIAS of the IV estimator is less in absolute magnitude than that of the OLS estimator for $\mu'\mu \neq 0$, and the former only tends to the OLS bias as $k_{21} \to \infty$. Furthermore, the asymptotic biases of the two estimators are exactly equal only when $\mu'\mu = 0$, for finite values of k_{21}. Also note from

4

inspection of equation (7) that the bias of the IV estimator has the same sign as the OLS bias. Thus, our results are in agreement with results in Bound, Jaeger, and Baker (1995) and Angrist and Krueger (1995), who suggest that with weak instruments, the IV estimator is biased in the direction of the OLS estimator, and the magnitude of the bias approaches that of the OLS estimator as the R^2 between the instruments and the endogenous explanatory variable approaches zero (i.e. as $\mu'\mu \to 0$). Our results also generalize characterizations of the IV bias given in Nelson and Startz (1990a&b) for a simple Gaussian model with a single fixed instrument and a single endogenous regressor to the more general case of an SEM with an arbitrary number of possibly stochastic instruments and with possible non-normal errors.

The above expressions for ABIAS and AMSE can also be compared with those of Richardson and Wu (1971), who obtained the exact bias and MSE of the $2SLS$ estimator for a fixed instrument/Gaussian model[7]. To facilitate this comparison, write the above SEM in its reduced form as: $y_1 = Z\Gamma_1 + X\Gamma_2 + \varepsilon_1$ and $y_2 = Z\Pi + X\Phi + \varepsilon_2$, where $\Gamma_1 = \Pi\beta$, $\Gamma_2 = \Phi\beta + \gamma$, $\varepsilon_2 = v$, and $\varepsilon_1 = u + v\beta = u + \varepsilon_2\beta$. Richardson and Wu (1971) assume that $(\varepsilon_{1t}, \varepsilon_{2t})' \equiv i.i.d.N(\mathbf{0}, G)$, where ε_{1t} and ε_{2t} denote the tth component of the $T \times 1$ random vectors ε_1 and ε_2, respectively and where $G = (g_{ij})$, $i, j = 1, 2$, is the 2×2 covariance matrix of $(\varepsilon_{1t}, \varepsilon_{2t})'$. Now, consider the case where all available instruments are used (i.e. the IV estimator is simply the $2SLS$ estimator). In this case, it follows that $\mu'\mu = \sigma_{vv}^{-1}C'\Omega C = \sigma_{vv}^{-1}C'(Q_{ZZ} - Q_{ZX}Q_{XX}^{-1}Q_{XZ})C$. Moreover, note that in terms of the elements of the reduced form error covariance matrix G the elements of the structural error covariance matrix Σ given in Section 2 can be written as: $\sigma_{uu} = g_{11} - 2g_{12}\beta + g_{22}\beta^2$, $\sigma_{uv} = g_{12} - g_{22}\beta$, and $\sigma_{vv} = g_{22}$. Substituting these expressions into equations (5) and (6) yields:

$$b_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = -\frac{g_{22}\beta - g_{12}}{g_{22}}e^{-\frac{\mu'\mu}{2}} \, {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right), \text{ and} \tag{9}$$

$$m_{\widehat{\beta}_{IV}}(\mu'\mu, k_{21}) = \frac{g_{11}g_{22} - g_{12}^2}{g_{22}}\left(\frac{1}{k_{21} - 2}\right)(1 + \overline{\beta}^2)e^{-\frac{\mu'\mu}{2}} \, {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right)$$
$$+ \left(\frac{k_{21} - 3}{k_{21} - 2}\right)\overline{\beta}^2 e^{-\frac{\mu'\mu}{2}} \, {}_1F_1\left(\frac{k_{21}}{2} - 2; \frac{k_{21}}{2}; \frac{\mu'\mu}{2}\right), \tag{10}$$

where $\overline{\beta} = (g_{22}\beta - g_{12})(g_{11}g_{22} - g_{12}^2)^{-\frac{1}{2}}$. Comparing equations (9) and (10) with equations (3.1) and (4.1) of Richardson and Wu (1971), we see that in this case the formulae for the ABIAS and AMSE are virtually identical to the exact bias and MSE derived under their assumption of a fixed instrument/Gaussian model - the only minor difference being that the (population) concentration parameter $\mu'\mu$ enters into the asymptotic formulae given in expressions (9) and (10) above, whereas the expression $\sigma_{vv}^{-1}\Pi'Z'M_XZ\Pi$ appears in the exact formulae reported in Richardson and Wu (1971). In this sense, the result in Proposition 1 generalizes Richardson and Wu (1971) to the case of non Gaussian errors. Additionally, our results are also consistent with the result from Staiger and Stock (1997) that the limiting distribution of the $2SLS$ estimator under the local-to-zero assumption is the same as the exact distribution of the estimator under the more restrictive assumptions of fixed instruments and Gaussian errors.

Next, we construct approximations for the bias and MSE. The approximate formulae we provide below have the virtue that they greatly simplify the more complicated expressions given in (7) and (8). To proceed, assume that

Assumption 4: $\quad \frac{\mu'\mu}{k_{21}} = \tau^2 + O\left(k_{21}^{-2}\right)$ *for some constant* $\tau^2 \in (0, \infty)$, *as* $\mu'\mu$, $k_{21} \to \infty$.

---

[7]Other papers which have studied the bias and/or MSE of the IV estimator, but for a fully identified model, include Richardson (1968), Hillier, Kinal, and Srivastava (1984), and Buse (1992).

The next result gives a formal statement of our approximations based on Assumption 4.

Theorem 3.2 (Approximations): *Suppose that Assumption 4 holds. Write* $\mu'\mu = \tau^2 k_{21} + O\left(k_{21}^{-1}\right) = \mu'\mu(\tau^2, k_{21})$, *say, and reparameterize the bias and MSE functions given in equations (5) and (6) in terms of* τ^2 *and* k_{21} *so that:*

$$
b_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\, e^{-\frac{\mu'\mu(\tau^2, k_{21})}{2}}\, {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\mu'\mu(\tau^2, k_{21})}{2}\right), \tag{11}
$$

$$
m_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) = \sigma_{uu}\sigma_{vv}^{-1}\rho^2 e^{-\frac{\mu'\mu(\tau^2, k_{21})}{2}}\left[\frac{1}{\rho^2}\left(\frac{1}{k_{21} - 2}\right) {}_1F_1\left(\frac{k_{21}}{2} - 1; \frac{k_{21}}{2}; \frac{\mu'\mu(\tau^2, k_{21})}{2}\right)\right.
$$
$$
\left. + \left(\frac{k_{21} - 3}{k_{21} - 2}\right) {}_1F_1\left(\frac{k_{21}}{2} - 2; \frac{k_{21}}{2}; \frac{\mu'\mu(\tau^2, k_{21})}{2}\right)\right) \tag{12}
$$

Then, as $k_{21} \to \infty$, *the following results hold :*

(i)

$$
b_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left\{\left(\frac{1}{1+\tau^2}\right) - \frac{2}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right)^2\right\} + O(k_{21}^{-2}) \tag{13}
$$

(ii)

$$
m_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) = \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left\{\left(\frac{1}{1+\tau^2}\right)^2 + \left(\frac{1-\rho^2}{\rho^2}\right)\left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right) + \left(\frac{1}{k_{21}}\right)\left(\frac{1}{1+\tau^2}\right)\right.
$$
$$
\left.\left[1 - 7\left(\frac{1}{1+\tau^2}\right) + 12\left(\frac{1}{1+\tau^2}\right)^2 - 6\left(\frac{1}{1+\tau^2}\right)^3\right]\right\} + O(k_{21}^{-2}) \tag{14}
$$

A few quick remarks about Theorem 3.2 are in order. First, set $\widehat{b}_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho$ $\left[\left(\frac{1}{1+\tau^2}\right) - \frac{2}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right)^2\right]$. Since the ABIAS of $\widehat{\beta}_{OLS}$ takes the form $b_{\widehat{\beta}_{OLS}} = \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho$, it follows that the relative bias of IV to that of OLS is given by the ratio

$$
\frac{\widehat{b}_{\widehat{\beta}_{IV}}(\tau^2, k_{21})}{b_{\widehat{\beta}_{OLS}}} = \left(\frac{1}{1+\tau^2}\right) - \frac{2}{k_{21}}\left(\frac{1}{1+\tau^2}\right)\left(\frac{\tau^2}{1+\tau^2}\right)^2. \tag{15}
$$

Observe that the lead term of equation (15) is $(1+\tau^2)^{-1} = (1 + \mu'\mu/k_{21})^{-1}$. Note also that when all available instruments are used so that $IV = 2SLS$, $(1 + \mu'\mu/k_{21})^{-1}$ is the relative bias measure given in Staiger and Stock (1997), in the case where there is only a single endogenous explanatory variable. Staiger and Stock point out that this measure of relative bias is given by an approximation which holds for large k_{21} and/or large $\mu'\mu/k_{21}$. Our analysis shows that their relative bias measure can also be obtained, from an approximation that requires $\mu'\mu/k_{21}$ to approach a finite limit as $\mu'\mu, k_{21} \to \infty$. [8]

---

[8] Note also that even though we take $\mu'\mu$ to infinity in making our approximations, our framework is still one which is appropriate for the case of weak instruments since we require $\frac{\mu'\mu}{k_{21}}$ to converge to a finite limit as $T \to \infty$ and $k_{21} \to \infty$, in sequence. This is in contrast to the usual case of full identification and good instruments where the first stage F-statistic diverges in probability as $T \to \infty$. Note, in particular, that even within a weak instrument setup, $\mu'\mu$ increases as $k_{21} \to \infty$ if the added instruments are not completely uncorrelated with the endogenous regressor.

Note also that the lead term of the $\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$ approximation is given by $\sigma_{uu}\sigma_{vv}^{-1}\rho^2(1+\tau^2)^{-2}$, which is simply the square of the lead term of $\widehat{b}_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$. It follows that the variance component of the AMSE is of a lower order in k_{21}, relative to the bias component, so that the variance can be thought of as being negligible relative to the bias component when the number of instruments is large relative to the value of τ^2. Note further that a desirable feature of the approximation formula $\widehat{m}_{\widehat{\beta}_{IV}}(\tau^2, k_{21})$ is that it is non-negative for $k_{21} \geq 4$, as one expects a mean-square error formula to be (see Chao and Swanson (2000) for further details).

As outlined in the proofs of Lemma A2 and Theorem 3.2, the approximate ABIAS and AMSE given above are derived using a Laplace approximation of the confluent hypergeometric function which holds as $\mu'\mu$, $k_{21} \to \infty$ such that $\frac{\mu'\mu}{k_{21}} = \tau^2 + O\left(k_{21}^{-2}\right)$. Our technical result extends approximations of the confluent hypergeometric function given in Slater (1960) in an important way and, thus, may be of independent interest. Note, in particular, that the approximation given in equation (4.3.8) of Slater (1960) is derived using the binomial expansion; and hence, strictly speaking, the result in Slater (1960) does not apply to the case $\tau^2 > 1$, whereas our results, being based on the Laplace approximation, do apply in this case (as is needed if the results are to provide useful approximations for the ABIAS and AMSE).

In order to assess the potential usefulness of our approximations, we carried out some numerical calculations using a canonical SEM, where the reduced form error covariance matrix is taken to be the identity matrix (i.e., $G = I$). In particular, we performed a simple regression analysis (see Table 1), where we calculated values of the dependent variable using the analytical formulae for the bias and MSE given by expressions (7) and (8), but specialized to the case of a canonical model. The independent variables in the regressions were the first order approximations given in Theorem 3.2 (called *CS Approximations*), again specialized to the canonical case. In addition, regressions were run using as independent variable the MSE approximation first derived in Donald and Newey (2001) and further examined in Hahn, Hausman and Kuersteiner (2002). This approximation, called the *DN Approximation* in the table, takes the form $\left(1 + \beta^2\right)/\mu'\mu + k_{21}^2\beta^2/(\mu'\mu)^2$ when the underlying model is a canonical SEM[9]. Values for both the dependent and the independent variables were calculated for $\beta = \{-0.5, -1.0, -1.5, ..., -10\}$, $\mu'\mu = \{2, 4, 6, 8, ..., 100\}$, and $k_{21} = \{3, 5, 7, 9, 11, ..., 101\}$, so that in total 50000 observations were generated by taking all possible combinations.[10] A total of 50 regression were run for each approximation, with each regression including 1000 observations for a given value of $\mu'\mu$. Of note is that these regressions thus all include observations for low values of k_{21}. Interestingly, both our bias and MSE approximation fare very well, with R^2 values very near to unity for all cases considered. This contrasts with the *DN Approximation,* which has low R^2 values when the instruments are weak, and plateaus with an R^2 value of 0.961 for the boundary case of $\mu'\mu = 100$. Our approximations thus appear quite robust, and our MSE approximation fairs very well compared with the approximation examined by Hahn, Hausman and Kuersteiner (2002). Furthermore, while numerical calculations given in the finite sample literature on single-equation estimators have tended to parameterize the bias and MSE functions in terms of k_{21} and $\mu'\mu$ (see, for example, Richardson and Wu (1971)), our results suggest that, in agreement with the findings of

---

[9] The approximate bias formula presented here has also been discussed in Hahn and Hausman (2002), although, in that paper, this bias approximation is not given a rigorous justification based on Laplace's method, as is done here.

[10] Note that, for our regressions analysis, we have chosen only negative values of β. This is because the bias function is perfectly symmetrical with respect to positive and negative values of β.

Staiger and Stock (1997), both the bias and the MSE might be better viewed instead as functions of k_{21} and the ratio $\tau^2 = \mu'\mu/k_{21}$.

4 Estimation of Bias and MSE and Bias Correction

In this section, we obtain consistent estimators for the lead terms of the bias and MSE expansions given in Theorem 3.2. Let $M_1 = M_{(Z,X)}$ and $M_2 = M_X$ and define the following statistics: $\widehat{\sigma}_{vv,i} = \frac{y_2' M_i y_2}{T}$, for $i = 1, 2$; $s_{uv,i} = \frac{(y_1 - y_2 \widehat{\beta}_{IV})' M_i y_2}{T}$, for $i = 1, 2$; $s_{uu} = \frac{(y_1 - y_2 \widehat{\beta}_{IV})' M_2 (y_1 - y_2 \widehat{\beta}_{IV})}{T}$; $\widehat{g}_{ij} = \frac{y_i' M_1 y_j}{T}$, for $i = 1, 2$ and $j = 1, 2$; $W_{k_{21},T} = \left[\frac{y_2'(P_H - P_X) y_2}{\widehat{\sigma}_{vv,1}} \right] k_{21}^{-1} = \frac{W_{k_{21},T}^*}{k_{21}}$; $\widehat{\sigma}_{uv,i} = s_{uv,i} \left(\frac{W_{k_{21},T}}{W_{k_{21},T} - 1} \right) = s_{uv,i} \left(\frac{1}{1 - \frac{1}{W_{k_{21},T}}} \right)$, for $i = 1, 2$;

and $\widehat{\sigma}_{uu,i} = s_{uu} + 2 \frac{\widehat{\sigma}_{uv,i}^2}{\widehat{\sigma}_{vv,i}} \left(\frac{1}{W_{k_{21},T}} \right) - \frac{\widehat{\sigma}_{uv,i}^2}{\widehat{\sigma}_{vv,i}} \left(\frac{1}{W_{k_{21},T}} \right)^2$, for $i = 1, 2$. The following Lemma shows that we can consistently estimate the quantities σ_{vv}, σ_{uv}, σ_{uu}, and $(1 + \tau^2)$ under the sequential limit approach.

Lemma 4.1: *Suppose that Assumptions 1 and 2 hold. Let $T \to \infty$, and then let $k_{21}, \mu'\mu \to \infty$ such that Assumption 4 holds. Then: (i) $\widehat{\sigma}_{vv,i} \xrightarrow{p} \sigma_{vv}$, for $i = 1, 2$; (ii) $W_{k_{21},T} \xrightarrow{p} 1 + \tau^2$; (iii) $\widehat{\sigma}_{uv,i} \xrightarrow{p} \sigma_{uv}$, for $i = 1, 2$; and (iv) $\widehat{\sigma}_{uu,i} \xrightarrow{p} \sigma_{uu}$, for $i = 1, 2$.*

Based on this result, several estimators for the ABIAS and the AMSE can be constructed, including:

$$\widehat{BIAS}_i = \frac{\widehat{\sigma}_{uv,i}}{\widehat{\sigma}_{vv,i}} \left(\frac{1}{W_{k_{21},T}} \right), \text{ for } i = 1, 2; \tag{16}$$

$$\widetilde{BIAS}_i = \frac{\widehat{\sigma}_{uv,i}}{\widehat{\sigma}_{vv,i}} \left[\left(\frac{1}{W_{k_{21},T}} \right) - \frac{2}{k_{21}} \left(\frac{1}{W_{k_{21},T}} \right) \left(\frac{W_{k_{21},T} - 1}{W_{k_{21},T}} \right)^2 \right], \text{ for } i = 1, 2; \tag{17}$$

$$\widehat{MSE}_i = \frac{\widehat{\sigma}_{uv,i}^2}{\widehat{\sigma}_{vv,i}^2} \left(\frac{1}{W_{k_{21},T}} \right)^2, \text{ for } i = 1, 2; \tag{18}$$

$$
\begin{aligned}
\widetilde{MSE}_i &= \frac{\widehat{\sigma}_{uv,i}^2}{\widehat{\sigma}_{vv,i}^2} \left[\left(\frac{1}{W_{k_{21},T}} \right)^2 + \frac{1}{k_{21}} \left(\frac{\widehat{\sigma}_{uu,i} \widehat{\sigma}_{vv,i} - \widehat{\sigma}_{uv,i}^2}{\widehat{\sigma}_{uv,i}^2} \right) \left(\frac{1}{W_{k_{21},T}} \right) \right. \\
&\quad \left. + \frac{1}{k_{21}} \left(\frac{1}{W_{k_{21},T}} \right) \left(1 - \frac{7}{W_{k_{21},T}} + \frac{12}{W_{k_{21},T}^2} - \frac{6}{W_{k_{21},T}^3} \right) \right], \text{ for } i = 1, 2;
\end{aligned}
\tag{19}
$$

and

$$
\begin{aligned}
\overline{MSE}_i &= \frac{\widehat{\sigma}_{uv,i}^2}{\widehat{\sigma}_{vv,i}^2} \left[\left(\frac{1}{W_{k_{21},T}} \right)^2 + \frac{1}{k_{21}} \left(\frac{\widehat{g}_{11} \widehat{g}_{22} - \widehat{g}_{12}^2}{\widehat{\sigma}_{uv,i}^2} \right) \left(\frac{1}{W_{k_{21},T}} \right) \right. \\
&\quad \left. + \frac{1}{k_{21}} \left(\frac{1}{W_{k_{21},T}} \right) \left(1 - \frac{7}{W_{k_{21},T}} + \frac{12}{W_{k_{21},T}^2} - \frac{6}{W_{k_{21},T}^3} \right) \right], \text{ for } i = 1, 2.
\end{aligned}
\tag{20}
$$

Note that the above estimators differ in that some are constructed based only on the lead term of the expansions given in Theorem 3.2 while others make use of both the lead and the second order terms. Also, two different estimators of $\sigma_{uu}\sigma_{vv} - \sigma_{uv}^2$ are used, yielding different sets of estimators.

8

Theorem 4.2: *Suppose that Assumptions 1 and 2 hold. Let $T \to \infty$, and then let $k_{21}, \mu'\mu \to \infty$ such that Assumption 4 holds. Then, for $i = 1, 2$: (i) $\widehat{BIAS}_i \overset{p}{\to} \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left(\frac{1}{1+\tau^2}\right)$; (ii) $\widetilde{BIAS}_i \overset{p}{\to} \sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho\left(\frac{1}{1+\tau^2}\right)$; (iii) $\widehat{MSE}_i \overset{p}{\to} \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left(\frac{1}{1+\tau^2}\right)$; (iv) $\widetilde{MSE}_i \overset{p}{\to} \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left(\frac{1}{1+\tau^2}\right)^2$; and (v) $\overline{MSE}_i \overset{p}{\to} \sigma_{uu}\sigma_{vv}^{-1}\rho^2\left(\frac{1}{1+\tau^2}\right)$.*
These weak consistency results suggest that there is information which can be exploited when a large number of weakly correlated instruments are available, as consistent estimation may be achieved when the number of instruments is allowed to approach infinity.

The consistency results presented in Theorem 4.2 can be applied to construct bias-corrected estimators[11]. In particular, we propose the following bias-adjusted OLS and IV estimators:

$$
\begin{aligned}
\widetilde{\beta}_{OLS,i} &= \widehat{\beta}_{OLS} - \frac{\widehat{\sigma}_{uv,i}}{\widehat{\sigma}_{vv,i}}, \text{ for } i = 1, 2; \\
\widetilde{\beta}_{IV} &= \widehat{\beta}_{IV} - \widehat{BIAS}_1; \\
\widetilde{\widetilde{\beta}}_{IV,i} &= \widehat{\beta}_{IV} - \widetilde{BIAS}_i, for\, i = 1, 2.
\end{aligned}
$$

The following consistency results can be obtained for these estimators.

Theorem 4.3: *Suppose that Assumptions 1 and 2 hold. Let $T \to \infty$, and then let $k_{21}, \mu'\mu \to \infty$, such that Assumption 4 holds. Then: (i) $\widetilde{\beta}_{OLS,i} \overset{p}{\to} \beta_0$, for $i = 1, 2$; (ii) $\widetilde{\beta}_{IV} \overset{p}{\to} \beta_0$; and (iii) $\widetilde{\beta}_{IV,i} \overset{p}{\to} \beta_0$, for $i = 1, 2$.*

Note that, under the local-to-zero framework with many instruments, the bias-corrected estimators are consistent. This is in contrast to the uncorrected OLS and IV estimators which are still not consistent in this case. It should also be noted that if we fix k_{21} and only allow $T \to \infty$, then none of the bias-adjusted estimators are consistent either. In fact, in this latter case, both the uncorrected and the bias-corrected IV estimators converge weakly to random variables. Moreover, if $\tau^2 = 0$, then the bias-adjusted estimators would not consistently estimate β. Our results, thus, suggest that if one is faced with a situation where a great many weak instruments are available; then, the appropriate estimation strategy may be to make use of all of these poor quality instruments in constructing bias-corrected estimators so long as the instruments are not completely uncorrelated with the endogenous explanatory variable (i.e. $\tau^2 \neq 0$). Finally, note that while Theorem 4.3 establishes the consistency of the bias-corrected estimators on the basis of a sequential asymptotic scheme; under some stronger but more primitive conditions than those stipulated in this paper, the bias-corrected estimators proposed here have also been found to be consistent under a pathwise asymptotic scheme whereby the number of instruments is taken to approach infinity as a function of the sample size (see Chao and Swanson (2001)).

It is also of interest to examine the above estimators under the conventional framework, where the instruments are not assumed to be weak in the local-to-zero sense but rather the usual identification condition is assumed to hold, even asymptotically. More explicitly, the conventional full identification assumption can be stated in our context as

Assumption 1*: Let Π be a fixed $k_2 \times 1$ vector such that $\Pi \neq \mathbf{0}$.

Under assumption 1[*], we have the following asymptotic results for our bias-adjusted estimators

Theorem 4.4 *Suppose that Assumptions 1* and 2 hold. Then, as $T \to \infty$, the following limit results hold: (i) $\widetilde{\beta}_{OLS,1} \overset{p}{\to} \beta_0 - \frac{\sigma_{uv}}{\sigma_{vv}}\left(\frac{\Pi'\Omega\Pi}{\Pi'\Omega\Pi+\sigma_{vv}}\right)$; (ii) $\widetilde{\beta}_{OLS,2} \overset{p}{\to} \beta_0$; (iii) $\widetilde{\beta}_{IV} \overset{p}{\to} \beta_0$; and (iv) $\widetilde{\beta}_{IV,i} \overset{p}{\to} \beta_0$ for $i = 1, 2$.*

---

[11]Other recent work in the area of bias correction is discussed in Hausman, Hahn, and Kuersteiner (2001).

Thus, Theorem 4.4 shows that, in the conventional setup where the instruments are fully relevant, all bias-corrected estimators, with the lone exception of $\widetilde{\beta}_{OLS,1}$ are still consistent[12].

5   Monte Carlo Results

In this section, we report the results of a small Monte Carlo study of the sampling behavior of the bias adjusted estimators introduced above. The DGP's used in the study are based on equations (1) and (2), with $\gamma = 0$ and $\Phi = 0$, i.e., they are based on the two-equation system

$$
\begin{aligned}
y_{1t} &= y_{2t}\beta + u_t, \\
y_{2t} &= Z_t'\Pi + v_t,
\end{aligned}
$$

where $t = 1, ..., T$. In this context, the reduced form of the first equation above can be written as

$$ y_{1t} = Z_t'\Gamma_1 + \varepsilon_{1t}, $$

$t = 1, ..., T$. Data were generated with: $\varepsilon_{2t} = v_t$, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})' \equiv i.i.d.N(0, I_2)$, $k_{21} = 20$, $T = 2000$, $\Pi = (\pi_1, \pi_2, ..., \pi_{20})' = (\overline{\pi}, \overline{\pi}, ..., \overline{\pi})$, with $\overline{\pi} = \{0.0225, 0.0513, 0.0745, 0.1120, 0.1290\}$ and $\beta = \{-0.1111, -0.5, -0.65, -0.81, -1.0\}$. In addition, the exogenous instruments, Z_t, were generated by assuming that $Z_t \equiv i.i.d.N(0, I_{k_{21}})$, and that all k_{21} instruments are used in estimator construction. Finally, define $w_t = (u_t, v_t)'$ with $E(w_t w_t') = \Sigma = \begin{pmatrix} \sigma_{uu} & \sigma_{vu} \\ \sigma_{vu} & \sigma_{vv} \end{pmatrix}$, and note that our canonical model specification implies that $\sigma_{uu} = 1 + \beta^2$, $\sigma_{uv} = -\beta$, and $\sigma_{vv} = 1$. It follows that the degree of endogeneity is determined by the value of the parameter β (see footnote to Table 2 for further discussion).

Table 2 reports the sample bias and MSE averaged across 5000 Monte Carlo simulations for the OLS, IV, LIML, and our 5 bias corrected estimators (for LIML, median bias is given, although mean bias results are similar). Of note is that $\widetilde{\widetilde{\beta}}_{IV,1}$ yields lower bias than OLS, IV and LIML for 19 of 25 different levels of endogeneity and instrument relevance examined. In addition, all of our other bias corrected estimators, with the exception of $\widetilde{\beta}_{OLS,1}$ yield less biased estimates in numerous cases. However, there is a trade-off. The bias corrected estimators only MSE dominate OLS, IV, and LIML for low levels of endogeneity (i.e. R^2_{endog} = 0.1 - see the table for further details), although in most cases the MSE is very close to the "best" MSE (i.e. that of the IV estimator).[13]

6   Concluding Remarks

In this paper, we constructed approximations for the ABIAS and AMSE of the IV estimator when the available instruments are weak in the local-to-zero sense. These approximations were shown, via a series of numerical computations, to be quite accurate. Additionally, we are able to obtain consistent estimators of the bias and MSE, and construct a variety of consistent bias-corrected OLS and IV estimators. Finally, we

---

[12] Note also that since the correction terms in the bias-corrected IV estimators $\widetilde{\beta}_{IV}$, $\widetilde{\widetilde{\beta}}_{IV,1}$, and $\widetilde{\widetilde{\beta}}_{IV,1}$ are all of order $O_p(T^{-1})$ under Assumption 1* and 2, these estimators are also asymptotically normal in the usual sense under these assumptions.

[13] Further exploration of these and other bias corrected estimators is the subject of ongoing research.

show that in the more conventional case where the simultaneous equations model is fully identified, all but one of our proposed bias corrected estimators are still consistent.

Appendix A

This appendix outlines a number of lemmas which are used to establish the main results of the paper (see Appendix B). Before presenting the lemmas, however, we first introduce some useful notation. Define $Z_{u,1} = \Omega_{11}^{-\frac{1}{2}\prime}(\psi_{Z_1 u} - Q_{Z_1 X} Q_{XX}^{-1} \psi_{Xu})\sigma_{uu}^{-\frac{1}{2}}$ and $Z_{v,1} = \Omega_{11}^{-\frac{1}{2}\prime}(\psi_{Z_1 v} - Q_{Z_1 X} Q_{XX}^{-1} \psi_{Xv})\sigma_{vv}^{-\frac{1}{2}}$. Note that

$$\begin{pmatrix} Z_{u,1} \\ Z_{v,1} \end{pmatrix} \sim N\left(0, \left(\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \otimes I_{k_{21}}\right)\right). \tag{21}$$

In addition, define $v_1(\mu'\mu, k_{21}) = (\mu + Z_{v,1})'(\mu + Z_{v,1}) = \sum_{i=1}^{k_{21}} \left(\mu_i + Z_{v,1}^i\right)^2$ and $v_2(\mu'\mu, k_{21}) = (\mu + Z_{v,1})' Z_{u,1} = \sum_{i=1}^{k_{21}} \left(\mu_i + Z_{v,1}^i\right) Z_{u,1}^i$, where μ_i, $Z_{u,1}^i$, and $Z_{v,1}^i$ are the $i-th$ components of μ, $Z_{u,1}$, and $Z_{v,1}$, respectively. Note that we have written $v_1(\cdot, \cdot)$ as a function of $\mu'\mu$ and not μ because v_1 is a non-central χ^2 random variable which depends on μ only through the noncentrality parameter $\mu'\mu$. In addition, since $\mu' Z_{u,1} \equiv N(0, \mu'\mu)$, $v_2(\mu'\mu, k_{21}) = \mu' Z_{u,1} + Z_{v,1}' Z_{u,1}$ also depends on μ only through $\mu'\mu$. To simplify notation, we will often write v_1 and v_2 instead of $v_1(\mu'\mu, k_{21})$ and $v_2(\mu'\mu, k_{21})$ in places where no confusion is caused by not making explicit the dependence of v_1 and v_2 on $\mu'\mu$ and k_{21}.

Lemma A1: Let $\widehat{\beta}_{IV,T}$ be the IV estimator defined in Section 2 and suppose that (1), (2) and Assumptions 1 and 2 hold. Then, as $T \to \infty$

$$\widehat{\beta}_{IV,T} - \beta_0 \Longrightarrow \sigma_{uu}^{\frac{1}{2}} \sigma_{vv}^{-\frac{1}{2}} v_1^{-1} v_2. \tag{22}$$

Proof: The proof follows from slight modification of the proof of Theorem 1, part (a) of Staiger and Stock (1997) and is, thus, omitted.

Lemma A2: Suppose that Assumption 4 holds. Write $\mu'\mu = \tau^2 k_{21} + R^*(k_{21}) = \mu'\mu(\tau^2, k_{21})$ (say), where $R^*(k_{21}) = O\left(k_{21}^{-1}\right)$. Then, for a given value of τ^2, as $k_{21} \to \infty$, the following results hold:

(a)

$$
\begin{aligned}
& {}_1F_1\left(k_{21}/2 - 1;\ k_{21}/2;\ \mu'\mu/2\right) \exp\left\{-\left(\mu'\mu/2\right)\right\} \\
= {} & {}_1F_1\left(k_{21}/2 - 1;\ k_{21}/2;\ \mu'\mu(\tau^2, k_{21})/2\right) \exp\left\{-\left(\mu'\mu(\tau^2, k_{21})/2\right)\right\} \\
= {} & \left(1 + \tau^2\right)^{-1} - k_{21}^{-1}\left(1 + \tau^2\right)^{-1}\left[2 - 4\left(1 + \tau^2\right)^{-1} + 2\left(1 + \tau^2\right)^{-2}\right] \\
& -k_{21}^{-2}\left(1 + \tau^2\right)^{-2}\left[8 - 28\left(1 + \tau^2\right)^{-1} + 32\left(1 + \tau^2\right)^{-2} - 12\left(1 + \tau^2\right)^{-3}\right] \\
& -R^*(k_{21}) k_{21}^{-1}\left(1 + \tau^2\right)^{-2} + O\left(k_{21}^{-3}\right),
\end{aligned} \tag{23}
$$

(b)

$$
\begin{aligned}
& {}_1F_1\left(k_{21}/2 - 2;\ k_{21}/2 - 1;\ \mu'\mu/2\right) \exp\left\{-\left(\mu'\mu/2\right)\right\} \\
= {} & {}_1F_1\left(k_{21}/2 - 2;\ k_{21}/2 - 1;\ \mu'\mu(\tau^2, k_{21})/2\right) \exp\left\{-\left(\mu'\mu(\tau^2, k_{21})/2\right)\right\} \\
= {} & \left(1 + \tau^2\right)^{-1} - k_{21}^{-1}\left(1 + \tau^2\right)^{-1}\left[4 - 6\left(1 + \tau^2\right)^{-1} + 2\left(1 + \tau^2\right)^{-2}\right] \\
& -k_{21}^{-2}\left(1 + \tau^2\right)^{-2}\left[24 - 56\left(1 + \tau^2\right)^{-1} + 44\left(1 + \tau^2\right)^{-2} - 12\left(1 + \tau^2\right)^{-3}\right] \\
& -R^*(k_{21}) k_{21}^{-1}\left(1 + \tau^2\right)^{-2} + O\left(k_{21}^{-3}\right).
\end{aligned} \tag{24}
$$

Proof: We shall only prove part (a) since the proof for part (b) follows in an analogous manner. To show (a), we make use of a well-known integral representation of the confluent hypergeometric function (see Lebedev (1972) pp. 266) to write

$$
\begin{aligned}
& {}_1F_1\left(k_{21}/2 - 1;\ k_{21}/2;\ \mu'\mu(\tau^2, k_{21})/2\right)\exp\left\{-\left(\mu'\mu(\tau^2, k_{21})/2\right)\right\} \\
= & \ [0.5\,(k_{21} - 2)]\int_0^1 \exp\left\{k_{21}h_1(t)\right\}\exp\left\{0.5R^*(k_{21})(t-1)\right\}dt,
\end{aligned}
\tag{25}
$$

where $h_1(t) = 0.5\left[\tau^2(t-1) + \log t\right] - (2/k_{21})\log t$. Given the integral representation (25), we can obtain the expansion given by the right-hand side of expression (23) by applying a Laplace approximation to this integral representation. We note that the maximum of the integrand of (25) in the interval $[0,1]$ occurs at the boundary point $t = 1$, and as $k_{21} \to \infty$ the mass of the integral becomes increasingly concentrated in some neighborhood of $t = 1$. Hence, we can obtain an accurate approximation for this integral by approximating the integrand with its Taylor expansion in some shrinking neighborhood of $t = 1$ and by showing that integration over the domain outside of this shrinking neighborhood becomes negligible as k_{21} becomes large. To proceed, notice that the RHS of equation (25) can be written as:

$$
\begin{aligned}
& [0.5\,(k_{21} - 2)]\int_{1-1/\sqrt{k_{21}}}^1 \exp\left\{k_{21}h_1(t)\right\}\exp\left\{0.5R^*(k_{21})(t-1)\right\}dt \\
& + [0.5\,(k_{21} - 2)]\int_0^{1-1/\sqrt{k_{21}}} \exp\left\{k_{21}h_1(t)\right\}\exp\left\{0.5R^*(k_{21})(t-1)\right\}dt = I_1 + I_2 \quad (say),
\end{aligned}
\tag{26}
$$

Now, note that

$$
\begin{aligned}
I_2 & \leq [0.5\,(k_{21} - 2)]\exp\left\{-\left(0.5\tau^2\sqrt{k_{21}}\right)\right\}\left(1 - k_{21}^{-\frac{1}{2}}\right)^{(k_{21}-2)/2}\exp\left\{-0.5k_{21}^{-\frac{1}{2}}R^*(k_{21})\right\} \\
& = O\left(k_{21}\exp\left\{-\left(0.5\tau^2\sqrt{k_{21}}\right)\right\}\left(1 - k_{21}^{-\frac{1}{2}}\right)^{(k_{21}-2)/2}\right),
\end{aligned}
\tag{27}
$$

where the inequality holds for $k_{21} \geq 4$. Now, turning our attention to I_1, we first make the change of variable $r = t - 1$ and rewrite $I_1 = [0.5\,(k_{21} - 2)]\int_{-1/\sqrt{k_{21}}}^0 \exp\left\{k_{21}h_2(r)\right\}\exp\left\{0.5R^*(k_{21})r\right\}dr$ where $h_2(r) = 0.5\left[\tau^2 r + \log(1 + r)\right] - (2/k_{21})\log(1 + r)$. With this change of variable, we note that the maximum of the integrand of I_1 in the interval $[-1/\sqrt{k_{21}}, 0]$ now occurs at the boundary point $r = 0$. To apply the Laplace approximation to I_1, note first that the derivatives of $h_2(r)$ evaluated at $r = 0$ have the explicit forms: $h_2'(0) = 0.5\left(1 + \tau^2\right) - 2k_{21}^{-1}$ and $h_2^{(i)}(0) = (-1)^{i-1}(i-1)!\left[0.5 - 2k_{21}^{-1}\right]$ for integer $i \geq 2$. By Taylor's formula, we can expand $h_2(r)$ about the point $r = 0$ as follows

$$
h_2(r) = h_2(0) + h_2'(0)r + \left(h_2^{(2)}(0)/2!\right)r^2 + \left(h_2^{(3)}(0)/3!\right)r^3 + \left(h_2^{(4)}(r^*)/4!\right)r^4,
\tag{28}
$$

where r^* lies on the line segment between r and 0 and $h_2(0) = 0$. Moreover, for $-1/\sqrt{k_{21}} \leq r \leq 0$, $|h_2^{(4)}(r^*)| = |3 - 12k_{21}^{-1}|(1 + r)^{-4} \leq |3 - 12k_{21}^{-1}|k_{21}^2\left(\sqrt{k_{21}} - 1\right)^{-4} = M(k_{21})$ (say), and note that $M(k_{21}) \to 3$ as $k_{21} \to \infty$. Hence, for $-1/\sqrt{k_{21}} \leq r \leq 0$,

$$
\left|h_2(r) - \sum_{i=1}^3 \frac{h_2^{(i)}(0)}{i!}r^i\right| = \left|\frac{h_2^{(4)}(r^*)}{4!}r^4\right| \leq \left[M(k_{21})r^4\right]/4!.
\tag{29}
$$

It follows that $\sum_{i=1}^{3} \frac{h_2^{(i)}(0)}{i!} r^i - \frac{M(k_{21})r^4}{4!} \le h_2(r) \le \sum_{i=1}^{3} \frac{h_2^{(i)}(0)}{i!} r^i + \frac{M(k_{21})r^4}{4!}$, so that

$$\left(\frac{k_{21}-2}{2}\right) \int_{-\frac{1}{\sqrt{k_{21}}}}^{0} \exp\left\{ k_{21}\left(\sum_{i=1}^{3} \frac{h_2^{(i)}(0)}{i!} r^i - \frac{M(k_{21})r^4}{4!} \right) \right\} \exp\left\{ \frac{r}{2} R^*(k_{21}) \right\} dr$$

$$\le \quad \left(\frac{k_{21}-2}{2}\right) \int_{-1/\sqrt{k_{21}}}^{0} \exp\left\{ k_{21} h_2(r) \right\} \exp\left\{ 0.5 r R^*(k_{21}) \right\} dr$$

$$\le \quad \left(\frac{k_{21}-2}{2}\right) \int_{-\frac{1}{\sqrt{k_{21}}}}^{0} \exp\left\{ k_{21}\left(\sum_{i=1}^{3} \frac{h_2^{(i)}(0)}{i!} r^i + \frac{M(k_{21})r^4}{4!} \right) \right\} \exp\left\{ \frac{r}{2} R^*(k_{21}) \right\} dr. \qquad (30)$$

Let I_3 denote the upper bound integral in expression (30). To evaluate I_3, we rewrite it as

$$I_3 = \left(\frac{k_{21}-2}{2}\right) \int_{-\frac{1}{\sqrt{k_{21}}}}^{0} \exp\left\{ k_{21} h_2'(0) r \right\} \exp\left\{ k_{21}\left(\sum_{i=2}^{3} \frac{h_2^{(i)}(0)}{i!} r^i + \frac{M(k_{21})r^4}{4!} \right) \right\} \exp\left\{ \frac{r}{2} R^*(k_{21}) \right\} dr.$$
$$(31)$$

Expanding the latter two exponentials in the integrand above in power series and integrating term-by-term while noting the absolute and uniform convergence of the series involved in the interval $r \in [-1/\sqrt{k_{21}}, 0]$ for $k_{21} \ge 4$; we obtain, after some tedious but straightforward calculations,

$$I_3 = (0.5(k_{21}-2)) \left[\int_{-1/\sqrt{k_{21}}}^{0} \exp\left\{ k_{21} h_2'(0) r \right\} \left(1 + \left[k_{21} h_2^{(2)}(0)/2! \right] r^2 \right. \right.$$

$$\left. + \left[k_{21} h_2^{(3)}(0)/3! \right] r^3 + \left[k_{21} \left(h_2^{(2)}(0) \right)^2 (2!)^{-3} \right] r^4 + 0.5 R^*(k_{21}) r \right) dr + O\left(k_{21}^{-4} \right) \right] \qquad (32)$$

$$= (1+\tau^2)^{-1} - k_{21}^{-1} (1+\tau^2)^{-1} \left[2 - 4(1+\tau^2)^{-1} + 2(1+\tau^2)^{-2} \right] - k_{21}^{-2} (1+\tau^2)^{-2} [8-$$

$$28(1+\tau^2)^{-1} + 32(1+\tau^2)^{-2} - 12(1+\tau^2)^{-3}] - R^*(k_{21}) k_{21}^{-1} (1+\tau^2)^{-3} + O\left(k_{21}^{-3} \right). \qquad (33)$$

By a similar argument, it can be shown that the lower bound integral in expression (30) can also be approximated by the right-hand side of expression (33). It, thus, follows that

$$(0.5(k_{21}-2)) \int_{-1/\sqrt{k_{21}}}^{0} \exp\left\{ k_{21} h_2(r) \right\} \exp\left\{ 0.5 R^*(k_{21}) r \right\} dr$$

$$= (1+\tau^2)^{-1} - k_{21}^{-1} (1+\tau^2)^{-1} \left[2 - 4(1+\tau^2)^{-1} + 2(1+\tau^2)^{-2} \right] - k_{21}^{-2} (1+\tau^2)^{-2} [8-$$

$$28(1+\tau^2)^{-1} + 32(1+\tau^2)^{-2} - 12(1+\tau^2)^{-3}] - R^*(k_{21}) k_{21}^{-1} (1+\tau^2)^{-3} + O\left(k_{21}^{-3} \right) \qquad (34)$$

Finally, the result given in part (a) follows immediately from expressions (27) and (34).

Lemma A3: Suppose that (1), (2) and Assumptions 1 and 2 hold. Then, the following convergence results hold jointly as $T \to \infty$:

(a) $(u'M_X u/T, \, y_2'M_X u/T, \, y_2'M_X y_2/T) \xrightarrow{p} (\sigma_{uu}, \, \sigma_{uv}, \, \sigma_{vv})$.

(b) $Z_1'M_X Z_1/T \xrightarrow{p} \Omega_{11}$, where $\Omega_{11} = Q_{Z_1 Z_1} - Q_{Z_1 X} Q_{XX}^{-1} Q_{X Z_1}$.

(c) $\left\{ (Z_1'M_X Z_1)^{-\frac{1}{2}} Z_1'M_X u, \, (Z_1'M_X Z_1)^{-\frac{1}{2}} Z_1'M_X v \right\} \Longrightarrow \left\{ Z_{u,1} \sigma_{uu}^{\frac{1}{2}}, \, Z_{v,1} \sigma_{vv}^{\frac{1}{2}} \right\}$,

where $(Z_{u,1}', \, Z_{v,1}')'$ has joint normal distribution given by (21).

(d) $(Z_1'M_X Z_1/T)^{-\frac{1}{2}} \left(Z_1'M_X y_2/\sqrt{T} \right) \Longrightarrow (\mu + Z_{v,1}) \sigma_{vv}^{\frac{1}{2}}$.

14

(e) $\left(y_2' M_X Z_1 (Z_1' M_X Z_1)^{-1} Z_1' M_X u, \; y_2' M_X Z_1 (Z_1' M_X Z_1)^{-1} Z_1' M_X y_2,\right.$
$\left. u' M_X Z_1 (Z_1' M_X Z_1)^{-1} Z_1' M_X u\right) \Longrightarrow \left(\sigma_{vv}^{\frac{1}{2}} v_2 \sigma_{uu}^{\frac{1}{2}}, \; \sigma_{vv} v_1, \; \sigma_{uu} Z_{u,1}' Z_{u,1}\right).$

(f) $(u' M_{(Z, \; X)} u/T, \; y_2' M_{(Z, \; X)} u/T, \; y_2' M_{(Z, \; X)} y_2/T) \xrightarrow{p} (\sigma_{uu}, \sigma_{uv}, \sigma_{vv}).$

(g) $(y_1' M_{(Z, \; X)} y_1/T, \; y_1' M_{(Z, \; X)} y_2/T) \xrightarrow{p} (g_{11}, g_{12})$, where g_{11} and g_{12} are elements of the reduced form error covariance matrix G.

Proof: These results follow from Staiger and Stock (1997), and the reader is referred to Chao and Swanson (2000) for further details.

Lemma A4: Let Assumption 4 hold, so that $\mu'\mu/k_{21} = \tau^2 + O(k_{21}^{-2})$ for a fixed constant $\tau^2 \in (0, \infty)$; and write $\mu'\mu = \tau^2 k_{21} + O(k_{21}^{-1}) = \mu'\mu(\tau^2, k_{21})$. Then, as $k_{21} \to \infty$, (a) $\frac{v_1(\mu'\mu(\tau^2, \; k_{21}), k_{21})}{k_{21}} \xrightarrow{p} (1 + \tau^2)$ and (b) $\frac{v_2(\mu'\mu(\tau^2, \; k_{21}), k_{21})}{k_{21}} \xrightarrow{p} \rho.$

Proof: To prove (a), write $\frac{v_1(\mu'\mu(\tau^2, \; k_{21}), k_{21})}{k_{21}} = \frac{Z_{v,1}' Z_{v,1}}{k_{21}} + 2\frac{\mu' Z_{v,1}}{k_{21}} + \frac{\mu'\mu}{k_{21}}$. Next, note that $\frac{\mu' Z_{v,1}}{k_{21}} \equiv N\left(0, \frac{\mu'\mu}{k_{21}^2}\right)$ so that $E\left(\frac{2\mu' Z_{v,1}}{k_{21}}\right)^2 = \frac{4\mu'\mu}{k_{21}^2} = \frac{4\tau^2}{k_{21}} + O(k_{21}^{-3})$, and, thus, $2\frac{\mu' Z_{v,1}}{k_{21}} \xrightarrow{p} 0$ as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4. Moreover, note that, as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4, $E\left(\frac{Z_{v,1}' Z_{v,1}}{k_{21}} - 1\right)^2 = \frac{2}{k_{21}} \to 0$, so that $\frac{Z_{v,1}' Z_{v,1}}{k_{21}} \xrightarrow{p} 1$, and note also that $\frac{\mu'\mu}{k_{21}} \to \tau^2$. It follows by the Slutsky's Theorem that $\frac{Z_{v,1}' Z_{v,1}}{k_{21}} + 2\frac{\mu' Z_{v,1}}{k_{21}} + \frac{\mu'\mu}{k_{21}} \xrightarrow{p} 1 + \tau^2$, as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4. To show (b), write $\frac{v_2(\mu'\mu(\tau^2, \; k_{21}), k_{21})}{k_{21}} = \frac{\mu' Z_{u,1}}{k_{21}} + \frac{Z_{v,1}' Z_{u,1}}{k_{21}}$. First, from expression (21), we see that $Z_{u,1} \equiv N(0, I_{k_{21}})$, $Z_{v,1} \equiv N(0, I_{k_{21}})$, and $E(Z_{u,1} Z_{v,1}') = \rho I_{k_{21}}$. It follows from Khinchine's weak law of large numbers that, as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4, $\frac{Z_{v,1}' Z_{u,1}}{k_{21}} = (1/k_{21}) \sum_{i=1}^{k_{21}} Z_{v,1}^i Z_{u,1}^i \xrightarrow{p} \rho$, where $Z_{v,1}^i$ and $Z_{u,1}^i$ denote the $i - th$ component of $Z_{v,1}$ and $Z_{u,1}$, respectively. In addition, note that $\frac{\mu' Z_{u,1}}{k_{21}} \equiv N\left(0, \frac{\mu'\mu}{k_{21}^2}\right)$ so that $E\left(\frac{\mu' Z_{u,1}}{k_{21}}\right)^2 = \frac{\mu'\mu}{k_{21}^2} = \frac{\tau^2}{k_{21}} + O(k_{21}^{-3})$, and, thus, $\frac{\mu' Z_{u,1}}{k_{21}} \xrightarrow{p} 0$ as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4. The desired result, thus, follows by the Slutsky's Theorem.

Lemma A5: Suppose that (1), (2) and Assumptions 1* and 2 hold. Then, the following convergence results hold as $T \to \infty$.

(a) $(u' M_X u/T, \; y_2' M_X u/T, \; y_2' M_X y_2/T) \xrightarrow{p} (\sigma_{uu}, \sigma_{uv}, \Pi' \Omega \Pi + \sigma_{vv}).$

(b) $(u' M_{(Z, \; X)} u/T, \; y_2' M_{(Z, \; X)} u/T, \; y_2' M_{(Z, \; X)} y_2/T) \xrightarrow{p} (\sigma_{uu}, \sigma_{uv}, \sigma_{vv}).$

(c) $(Z_1' M_X Z_1/T, \; Z_1' M_X y_2/T) \xrightarrow{p} (\Omega_{11}, \Omega_{1*} \Pi).$

Proof: Each part of this lemma follows directly from Assumptions 1* and 2 and the Slutsky's Theorem. The arguments are standard and well-known, so we omit the details.

Appendix B

Proof of Theorem 3.2: To show part (a), note that direct application of part (a) of Lemma A2 to equation (5) yields

$$b_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) = \sigma_{uu}^{1/2} \sigma_{vv}^{-1/2} \rho \left\{\left(1 + \tau^2\right)^{-1} - 2k_{21}^{-1} \left(1 + \tau^2\right)^{-1} \left(\tau^2/\left(1 + \tau^2\right)\right)^2\right\} + O\left(k_{21}^{-2}\right). \qquad (35)$$

15

To show part (b), first rewrite equation (6) as follows:

$$
\begin{aligned}
m_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) \;=\;& \sigma_{uu}\sigma_{vv}^{-1}\rho^2 \left[\rho^{-2}(k_{21}-2)^{-1} \,{}_1F_1\left(k_{21}/2-1;\; k_{21}/2;\; \mu'\mu(\tau^2,k_{21})/2\right) e^{-\mu'\mu(\tau^2,k_{21})/2} \right.\\
&+ \left(\frac{k_{21}-3}{k_{21}-2}\right)\left(\frac{k_{21}-2}{2}\right) {}_1F_1\left(\frac{k_{21}}{2}-2;\frac{k_{21}}{2}-1;\frac{\mu'\mu(\tau^2,k_{21})}{2}\right) e^{-\frac{\mu'\mu(\tau^2,k_{21})}{2}}\\
&\left. - \left(\frac{k_{21}-3}{k_{21}-2}\right)\left(\frac{k_{21}-4}{2}\right) {}_1F_1\left(\frac{k_{21}}{2}-1;\frac{k_{21}}{2};\frac{\mu'\mu(\tau^2,k_{21})}{2}\right) e^{-\frac{\mu'\mu(\tau^2,k_{21})}{2}}\right],
\end{aligned}
\tag{36}
$$

where we have made use of the identity $(\gamma-\alpha-1)\,{}_1F_1(\alpha;\gamma;z) = (\gamma-1)\,{}_1F_1(\alpha;\gamma-1;z) - \alpha\,{}_1F_1(\alpha+1;\gamma;z)$. (See Lebedev (1972), pp. 262, for more details on this and identities involving confluent hypergeometric functions.) Applying the results of Lemma A2 to the confluent hypergeometric functions in an expansion of equation (6) discussed in Chao and Swanson (2000) yields, after some tedious algebra, the following expression

$$
\begin{aligned}
m_{\widehat{\beta}_{IV}}(\tau^2, k_{21}) \;=\;& \sigma_{uu}\sigma_{vv}^{-1}\rho^2 \left\{ (1+\tau^2)^{-2} + \rho^{-2}k_{21}^{-1}(1+\tau^2)^{-1} - k_{21}^{-1}(1+\tau^2)^{-2}\,[7\right.\\
&\left. -12(1+\tau^2)^{-1}+6(1+\tau^2)^{-2}\right] + O(k_{21}^{-2})\}\\
=\;& \sigma_{uu}\sigma_{vv}^{-1}\rho^2 \left\{ (1+\tau^2)^{-2} + \left((1-\rho^2)/\rho^2\right)k_{21}^{-1}(1+\tau^2)^{-1} + k_{21}^{-1}(1+\tau^2)^{-1}\times\right.\\
&\left. \left[1-7(1+\tau^2)^{-1}+12(1+\tau^2)^{-2}-6(1+\tau^2)^{-3}\right]\right\} + O(k_{21}^{-2}).
\end{aligned}
\tag{37}
$$

Proof of Lemma 4.1: To show part (i), note that since $\widehat{\sigma}_{vv,1} = \frac{y_2' M_{(Z,X)} y_2}{T}$ and $\widehat{\sigma}_{vv,2} = \frac{y_2' M_X y_2}{T}$, it follows directly from part (a) and (f) of Lemma A3 that, as $T\to\infty$, $\widehat{\sigma}_{vv,1} \xrightarrow{p} \sigma_{vv}$ and $\widehat{\sigma}_{vv,2} \xrightarrow{p} \sigma_{vv}$. Note, of course that these limits do not depend on either k_{21} or $\mu'\mu$, and the results of part (i) follow as an immediate consequence. To show part (ii), note that it follows from part (d) of Lemma A3, part (i) of this Lemma, and the continuous mapping theorem that as $T\to\infty$, $W_{k_{21},T} = \dfrac{\left(\frac{y_2' M_X Z_1}{\sqrt{T}}\right)\left(\frac{Z_1' M_X Z_1}{T}\right)^{-1}\left(\frac{Z_1' M_X y_2}{\sqrt{T}}\right)/k_{21}}{\widehat{\sigma}_{vv,1}} \Longrightarrow$ $\frac{(\mu+Z_{v,1})'(\mu+Z_{v,1})}{k_{21}} = \frac{v_1(\mu'\mu,k_{21})}{k_{21}}$. It then follows directly from part (a) of Lemma A4 that, as $k_{21}\to\infty$ and $\mu'\mu\to\infty$ under Assumption 4, $\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \xrightarrow{p} 1+\tau^2$ as desired. To prove part (iii), write $\widehat{\sigma}_{uv,1} = \frac{(y_1-y_2\widehat{\beta}_{IV})' M_{(Z,X)} y_2}{T}\left(\frac{W_{k_{21},T}}{W_{k_{21},T}-1}\right) = \left[\frac{u' M_{(Z,X)} y_2}{T} - \left(\widehat{\beta}_{IV}-\beta_0\right)\right.$ $\left.\frac{y_2' M_{(Z,X)} y_2}{T}\right]\left(\frac{W_{k_{21},T}}{W_{k_{21},T}-1}\right)$ and $\widehat{\sigma}_{uv,2} = \left[\frac{u' M_X y_2}{T} - \left(\widehat{\beta}_{IV}-\beta_0\right)\frac{y_2' M_X y_2}{T}\right]\left(\frac{W_{k_{21},T}}{W_{k_{21},T}-1}\right)$. Applying Lemma A1, parts (a) and (f) of Lemma A3, and the continuous mapping theorem; we see immediately that $\widehat{\sigma}_{uv,1} \Longrightarrow$ $\left[\sigma_{uv} - \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{\frac{1}{2}}\left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}}\right)^{-1}\left(\frac{v_2(\mu'\mu,k_{21})}{k_{21}}\right)\right]\left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}}-1\right)^{-1}\left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}}\right) = \mathcal{A}_{k_{21},\mu'\mu}$ (say) and also that $\widehat{\sigma}_{uv,2} \Longrightarrow \mathcal{A}_{k_{21},\mu'\mu}$ as $T\to\infty$, so that both estimators approach the same random limit as the same size approaches infinity. Moreover, applying Lemma A4, we deduce that, as $k_{21}\to\infty$ and $\mu'\mu\to\infty$ under Assumption 4, $\mathcal{A}_{k_{21},\mu'\mu} \xrightarrow{p} \left[\sigma_{uv} - \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{\frac{1}{2}}\rho\left(1+\tau^2\right)^{-1}\right]$ $\left(\frac{1+\tau^2}{\tau^2}\right) = \sigma_{uv}\left[1-\frac{1}{1+\tau^2}\right]\left(\frac{1+\tau^2}{\tau^2}\right) = \sigma_{uv}$, thus, establishing the desired results. Finally, to show part (iv), write $\widehat{\sigma}_{uu,1} = s_{uu} + 2\frac{\widehat{\sigma}_{uv,1}^2}{\widehat{\sigma}_{vv,1}}\left(\frac{1}{W_{k_{21},T}}\right) - \frac{\widehat{\sigma}_{uv,1}^2}{\widehat{\sigma}_{vv,1}}\left(\frac{1}{W_{k_{21},T}}\right)^2$ and $\widehat{\sigma}_{uu,2} = s_{uu} + 2\frac{\widehat{\sigma}_{uv,2}^2}{\widehat{\sigma}_{vv,2}}\left(\frac{1}{W_{k_{21},T}}\right) - \frac{\widehat{\sigma}_{uv,2}^2}{\widehat{\sigma}_{vv,2}}\left(\frac{1}{W_{k_{21},T}}\right)^2$. Note first that $s_{uu} = \frac{(y_1-y_2\widehat{\beta}_{IV})' M_X (y_1-y_2\widehat{\beta}_{IV})}{T} = \frac{u' M_X u}{T} - 2\left(\widehat{\beta}_{IV}-\beta_0\right)\frac{y_2' M_X u}{T} + \left(\widehat{\beta}_{IV}-\beta_0\right)^2\frac{y_2' M_X y_2}{T}$. Hence, it follows from Lemma A1; part (a) of Lemma A3; the proofs of parts (i), (ii), and (iii) of this Lemma; and the continuous mapping theorem that as $T\to\infty$, $\widehat{\sigma}_{uu,1} \Longrightarrow \sigma_{uu} - 2\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\sigma_{uv}\left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}}\right)^{-1}\left(\frac{v_2(\mu'\mu,k_{21})}{k_{21}}\right) +$

$\sigma_{uu} \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-2} \left(\frac{v_2(\mu'\mu,k_{21})}{k_{21}} \right)^2 + 2 \frac{\mathcal{A}^2_{k_{21},\mu'\mu}}{\sigma_{vv}} \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-1} - \frac{\mathcal{A}^2_{k_{21},\mu'\mu}}{\sigma_{vv}} \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-2} = \mathcal{B}_{k_{21},\mu'\mu}$ (say). Similarly, $\widehat{\sigma}_{uu,2} \Longrightarrow \mathcal{B}_{k_{21},\mu'\mu}$ as $T \to \infty$. Moreover, applying Lemma A4 and part (iii) of this Lemma, we easily deduce that $\mathcal{B}_{k_{21},\mu'\mu} \xrightarrow{p} \sigma_{uu} - 2\sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\sigma_{uv} \left(\frac{\rho}{1+\tau^2} \right) + \sigma_{uu} \left(\frac{\rho}{1+\tau^2} \right)^2 + 2 \frac{\sigma_{uv}^2}{\sigma_{vv}} \left(\frac{1}{1+\tau^2} \right) - \frac{\sigma_{uv}^2}{\sigma_{vv}} \left(\frac{1}{1+\tau^2} \right)^2 = \sigma_{uu}$ as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4, thus, establishing the desired result.

Proof of Theorem 4.2: We will only prove the convergence result for the estimator with subscript $i = 1$, since the proofs for $i = 2$ follow directly. First, to show (i) write $\widehat{BIAS}_1 = \frac{\widehat{\sigma}_{uv,1}}{\widehat{\sigma}_{vv,1}} \left(\frac{1}{W_{k_{21},T}} \right)$, and note that given Lemma 4.1 and the continuous mapping theorem, it is apparent that as $T \to \infty$, $\widehat{BIAS}_1 \Longrightarrow \frac{\mathcal{A}_{k_{21},\mu'\mu}}{\sigma_{vv}} \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-1} = \mathcal{C}_{k_{21},\mu'\mu}$ (say). Applying Lemma A4 and part (iii) of Lemma 4.1, note that as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4, $\mathcal{C}_{k_{21},\mu'\mu} \xrightarrow{p} \frac{1}{\sigma_{vv}} \left(\sigma_{uv} - \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{\frac{1}{2}} \left(\frac{\rho}{1+\tau^2} \right) \right) \left(\frac{1+\tau^2}{\tau^2} \right) \left(\frac{1}{1+\tau^2} \right) = \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho \left(\frac{1}{1+\tau^2} \right)$, as required. To show part (ii), write $\widetilde{BIAS}_1 = \widehat{BIAS}_1 - \left(\frac{2}{k_{21}} \right) \left[\left(\frac{\widehat{\sigma}_{uv,1}}{\widehat{\sigma}_{vv,1}} \right) \left(\frac{1}{W_{k_{21},T}} \right) \left(\frac{W_{k_{21},T}-1}{W_{k_{21},T}} \right)^2 \right]$. As above, it follows that as $T \to \infty$, $\widetilde{BIAS}_1 \Longrightarrow \mathcal{C}_{k_{21},\mu'\mu} - \frac{2}{k_{21}} \left[\frac{\mathcal{A}_{k_{21},\mu'\mu}}{\sigma_{vv}} \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-1} \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-2} \right.$ $\left. \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} - 1 \right)^2 \right] = \mathcal{E}_{k_{21},\mu'\mu}$ (say). Moreover, note that $\mathcal{E}_{k_{21},\mu'\mu} = \mathcal{C}_{k_{21},\mu'\mu} + Op \left(\frac{1}{k_{21}} \right)$, so that applying Lemma A4, part (iii) of Lemma 4.1, and part (i) of this theorem; it follows that as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4, $\mathcal{E}_{k_{21},\mu'\mu} \xrightarrow{p} \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho \left(\frac{1}{1+\tau^2} \right)$, as required. To show part (iii), write $\widehat{MSE}_1 = \left(\widehat{BIAS}_1 \right)^2$. It follows immediately from the proof of part (i) above and the continuous mapping theorem that as $T \to \infty$, $\widehat{MSE}_1 \Longrightarrow \mathcal{C}^2_{k_{21},\mu'\mu}$. Moreover, given part (i) of this theorem, as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4, we deduce easily that $\mathcal{C}^2_{k_{21},\mu'\mu} \xrightarrow{p} \sigma_{uu}\sigma_{vv}\rho^2 \left(\frac{1}{1+\tau^2} \right)^2$. To show part (iv), write $\widetilde{MSE}_1 = \widehat{MSE}_1 + \frac{1}{k_{21}} \left(\frac{\widehat{\sigma}_{uv,1}}{\widehat{\sigma}_{vv,1}^2} \right) \left(\frac{1}{W_{k_{21},T}} \right)$ $\left[\left(\frac{\widehat{\sigma}_{uu,1}\widehat{\sigma}_{vv,1}-\widehat{\sigma}_{uv,1}^2}{\widehat{\sigma}_{uv,1}^2} \right) + \left(1 - \frac{7}{W_{k_{21},T}} + \frac{12}{W_{k_{21},T}^2} - \frac{6}{W_{k_{21},T}^3} \right) \right]$. Hence, from the proof of Lemma 4.1, the proof of part (iii) of this theorem, and the continuous mapping theorem; it is apparent that as $T \to \infty$, $\widetilde{MSE}_1 \Longrightarrow$ $\mathcal{C}^2_{k_{21},\mu'\mu} + \frac{1}{k_{21}} \left(\frac{\mathcal{A}^2_{k_{21},\mu'\mu}}{\sigma_{vv}^2} \right) \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-1}$ $\left[\left(\frac{\sigma_{vv}\mathcal{B}_{k_{21},\mu'\mu}-\mathcal{A}^2_{k_{21},\mu'\mu}}{\mathcal{A}^2_{k_{21},\mu'\mu}} \right) + \left(1 - 7 \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-1} + 12 \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-2} - 6 \left(\frac{v_1(\mu'\mu,k_{21})}{k_{21}} \right)^{-3} \right) \right] = \mathcal{F}_{k_{21},\mu'\mu}$ (say). Furthermore, note that $\mathcal{F}_{k_{21},\mu'\mu} = \mathcal{C}^2_{k_{21},\mu'\mu} + Op \left(\frac{1}{k_{21}} \right)$, so that applying Lemma A4, parts (iii) and (iv) of Lemma 4.1, and part (iii) of this theorem; we readily deduce that, as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4, $\mathcal{F}_{k_{21},\mu'\mu} \xrightarrow{p} \sigma_{uu}\sigma_{vv}\rho^2 \left(\frac{1}{1+\tau^2} \right)^2$, as required. Finally, to show part (v), note that by comparing the expressions for $\widetilde{MSE}_1$ and $\overline{MSE}_1$, it is clear that the only difference between the two is that $\widetilde{MSE}_1$ estimates the quantity $\sigma_{uu}\sigma_{vv} - \sigma_{uv}$ using the consistent estimator $\widehat{\sigma}_{uu,1}\widehat{\sigma}_{vv,1} - \widehat{\sigma}_{uv,1}^2$ whereas $\overline{MSE}_1$ estimates the quantity $g_{11}g_{22} - g_{12}^2$ using the estimator $\widehat{g}_{11}\widehat{g}_{22} - \widehat{g}_{12}^2$. Since it is easy to verify that $g_{11}g_{22} - g_{12}^2 = \sigma_{uu}\sigma_{vv} - \sigma_{uv}$, all that is left to show is the consistency of the estimator $\widehat{g}_{11}\widehat{g}_{22} - \widehat{g}_{12}^2$. However, given that $\widehat{g}_{11} = \frac{y_1' M_{(Z,X)} y_1}{T}$, $\widehat{g}_{12} = \frac{y_1' M_{(Z,X)} y_2}{T}$, and $\widehat{g}_{22} = \frac{y_2' M_{(Z,X)} y_2}{T}$; we see immediately from parts (f) and (g) of Lemma A8 that, as $T \to \infty$, $\widehat{g}_{11} \xrightarrow{p} g_{11}$, $\widehat{g}_{12} \xrightarrow{p} g_{12}$, and $\widehat{g}_{11} \xrightarrow{p} \sigma_{vv} = g_{22}$; and, thus, by the Slutsky's theorem, $\widehat{g}_{11}\widehat{g}_{22} - \widehat{g}_{12}^2 \xrightarrow{p} g_{11}g_{22} - g_{12}^2 = \sigma_{uu}\sigma_{vv} - \sigma_{uv}$. Since these limits do not depend on k_{21} and $\mu'\mu$, the desired result follows as a direct consequence.

Proof of Theorem 4.3: We will only prove consistency results for $\widetilde{\beta}_{OLS,1}$, $\widetilde{\beta}_{IV}$, and $\widetilde{\beta}_{IV,1}$ since the results for $\widetilde{\beta}_{OLS,2}$ and $\widetilde{\beta}_{IV,2}$ can be shown in a manner similar to those for $\widetilde{\beta}_{OLS,1}$ and $\widetilde{\beta}_{IV,1}$, respectively. To prove part (i), write $\widetilde{\beta}_{OLS,1} = \widehat{\beta}_{OLS} - \frac{\widehat{\sigma}_{uv,1}}{\widehat{\sigma}_{vv,1}} = \beta_0 + (y_2' M_X y_2)^{-1}(y_2' M_X u) - \frac{\widehat{\sigma}_{uv,1}}{\widehat{\sigma}_{vv,1}}$. Making use of part (a) of Lemma A3, the proof of parts (i) and (iii) of Lemma 4.1, and the continuous mapping theorem; we see that as $T \to \infty$, $\widetilde{\beta}_{OLS,1} \Longrightarrow \beta_0 + \frac{\sigma_{uv}}{\sigma_{vv}} - \frac{\mathcal{A}_{k_{21},\mu'\mu}}{\sigma_{vv}} = \mathcal{L}_{k_{21},\mu'\mu}$ (say). It follows immediately that as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4, $\mathcal{L}_{k_{21},\mu'\mu} \xrightarrow{p} \beta_0$ since $\mathcal{A}_{k_{21},\mu'\mu} \xrightarrow{p} \sigma_{uv}$, as shown in the proof of part (iii) of Lemma 4.1. To show part (ii), write $\widetilde{\beta}_{IV} = \widehat{\beta}_{IV} - \widehat{BIAS}_1$. Making use of Lemma A1, the proof of part (i) of Theorem 4.2, and the continuous mapping theorem; we see that, as $T \to \infty$, $\widetilde{\beta}_{IV} \Longrightarrow \beta_0 + \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} \left[\frac{v_2(\mu'\mu,k_{21})}{v_1(\mu'\mu,k_{21})}\right] - \mathcal{C}_{k_{21},\mu'\mu} = \mathcal{M}_{k_{21},\mu'\mu}$ (say). Moreover, applying Lemma A4 and the fact that $\mathcal{C}_{k_{21},\mu'\mu} \xrightarrow{p} \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left(\frac{1}{1+\tau^2}\right)$, as shown in the proof of Theorem 4.2 part (i); we deduce that $\mathcal{M}_{k_{21},\mu'\mu} \xrightarrow{p} \beta_0 + \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left(\frac{1}{1+\tau^2}\right) - \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left(\frac{1}{1+\tau^2}\right) = \beta_0$ as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4. To show part (iii), write $\widetilde{\beta}_{IV,1} = \widehat{\beta}_{IV} - \widehat{BIAS}_1$. Now, given Lemma A1, the proof of part (ii) of Theorem 4.2, and the continuous mapping theorem; it is apparent that, as $T \to \infty$, $\widetilde{\beta}_{IV,1} \Longrightarrow \beta_0 + \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}} \left[\frac{v_2(\mu'\mu,k_{21})}{v_1(\mu'\mu,k_{21})}\right] - \mathcal{E}_{k_{21},\mu'\mu} = \mathcal{N}_{k_{21},\mu'\mu}$ (say). Moreover, applying Lemma A4 and the fact that $\mathcal{E}_{k_{21},\mu'\mu} \xrightarrow{p} \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left(\frac{1}{1+\tau^2}\right)$, as shown in the proof of Theorem 4.2 part (ii), we readily deduce that $\mathcal{N}_{k_{21},\mu'\mu} \xrightarrow{p} \beta_0 + \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left(\frac{1}{1+\tau^2}\right) - \sigma_{uu}^{\frac{1}{2}}\sigma_{vv}^{-\frac{1}{2}}\rho\left(\frac{1}{1+\tau^2}\right) = \beta_0$ as $k_{21} \to \infty$ and $\mu'\mu \to \infty$ under Assumption 4.

Proof of Theorem 4.4: Note, that under Assumption 1*, the SEM described in Section 2 is fully identified in the usual sense. The results follows from standard arguments, which are omitted here. See Chao and Swanson (2000) for details.

References

Angrist, J. D. and A. B. Krueger (1995). "Split-Sample Instrumental Variables Estimates of the Return to Schooling," *Journal of Business and Economic Statistics*, 13, 225-235.

Bekker, P.A. (1994). "Alternative Approximations to the Distributions of Instrumental Variable Estimators," *Econometrica*, 62, 657-681.

Bound, J., D.A. Jaeger and R. Baker (1995). "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variables is Weak," *Journal of the American Statistical Association*, 90, 443-450.

Buse, A. (1992). "The Bias of Instrumental Variable Estimators," *Econometrica*, 60, 173-180.

Choi, I. and P.C.B.Phillips (1992). "Asymptotic and Finite Sample Distribution Theory for IV Estimators and Tests in Partially Identified Structural Equations," *Journal of Econometrics*, 51, 113-150.

Chao, J.C. and N.R.Swanson (2000). "Bias and MSE Analysis of the IV Estimator Under Weak Identification with Application to Bias Correction," University of Maryland and Rutgers University Working Paper.

Chao, J.C. and N.R.Swanson (2001). "Pathwise Asymptotics for the Bias-Corrected Instrumental Variables Estimator When the Number of Weak Instruments Is Large," University of Maryland and Rutgers University Working Paper.

Donald, S.G. and W.K. Newey (2001). "Choosing the Number of Instruments", *Econometrica*, 69, 1161-1191.

Dufour, J.- M. (1997). "Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models," *Econometrica*, 68, 1365-1388.

Forchini, G. and G. Hillier (1999). "Conditional Inference for Possibly Unidentified Structural Equations," University of Southampton Working Paper.

Hahn, J. and A. Inoue (2000). "A Monte Carlo Comparison of Various Asymptotic Approximations to the Distribution of Instrumental Variables Estimators," University of Brown Working Paper.

Hall, A.R., G.D. Rudebusch, and D.W.Wilcox (1996). "Judging Instrument Relevance in Instrumental Variables Estimation," *International Economic Review*, 37, 283-298.

Hall, A.R. and F. P. M. Peixe (2000). "A Consistent Method for the Selection of Relevant Instruments," North Carolina State University Working Paper.

Hahn, J. and J. Hausman (2002). "Notes on Bias in Estimators for Simultaneous Equation Model", *Economic Letters, 75, 237-241.*

Hahn, J., J. Hausman, and G. Kuersteiner (2002). "Estimation With Weak Instruments: Accuracy of Higher Order Bias and MSE Approximations", M.I.T. Working Paper.

Hahn, J. and G. Kuersteiner (2002). "Discontinuities of Weak Instrument Limiting Distributions", *Economic Letters, 75, 325-331.*

Hausman J., J. Hahn, and G. Kuersteiner (2001). "Bias Corrected Instrumental Variables Estimation for Dynamic Panel Models with Fixed Effects," M.I.T. Working Paper.

Hillier, G.H., T.W. Kinal and V.K. Srivastava (1984). "On the Moments of Ordinary Least Squares Instrumental Variables Estimators in a General Structural Equation," *Econometrica*, 52, 185-202.

Kitamura, Y. (1994). "Specification Tests with Instrumental Variables and Rank Deficiency," University of Minnesota Working Paper.

Lebedev, N. N. (1972). *Special Functions and Their Applications* (Dover Publications, Inc.; New York).

Mariano, R. and J.B.McDonald (1979). "A Note on the Distribution Functions of LIML and 2SLS Structural Coefficient in the Exactly Identified Case," *Journal of the American Statistical Association,* 74, 847-848.

Moreira, M.J. (2001). "Tests with Correct Size when Instruments Can Be Arbitrarily Weak," U.C. Berkeley Working Paper.

Moreira, M.J. (2002). "A Conditional Likelihood Ratio Test for Structural Models," U.C. Berkeley Working Paper.

Nelson, C. and R. Startz (1990a). "Some Further Results on the Exact Small Sample Properties of the Instrumental Variables Estimator," *Econometrica*, 58, 967-976.

Nelson, C. and R. Startz (1990b). "The Distribution of the Instrumental Variables Estimator and Its t-ratio when the Instrument is a Poor One," *Journal of Business*, 63, S125-S140.

Phillips, P.C.B. (1984). "The Exact Distribution of LIML: I," *International Economic Review,* 25, 249-261.

Phillips, P.C.B. (1985). "The Exact Distribution of LIML: II," *International Economic Review,* 26, 21-36.

Phillips, P.C.B. (1989). "Partially Identified Econometric Models," *Econometric Theory*, 5, 181-240.

Richardson, D. H. (1968). "The Exact Distribution of a Structural Coefficient Estimator," *Journal of the American Statistical Association*, 63, 1214-1226.

Richardson, D. H. and D.-M. Wu (1971). "A Note on the Comparison of Ordinary and Two-Stage Least Squares Estimators," *Econometrica*, 39, 973-982.

Sawa, T. (1969). "The Exact Finite Sampling Distribution of OLS and 2SLS Estimators," *Journal of the American Statistical Association,* 64, 923-937.

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

Slater, L.J. (1960). *Confluent Hypergeometric Functions* (Cambridge University Press, London).

Staiger, D. and J.H. Stock (1997). "Instrumental Variables Regression with Weak Instruments." *Econometrica*, 65, 557-586.

Stock, J.D., J. Wright, and M. Yogo (2002). "GMM, Weak Instruments, and Weak Identification", *Journal of Business and Economic Statistics*, 20, 518-529.

Stock, J.D. and M. Yogo (2002). "Testing for Weak Instruments in Linear IV Regression", Harvard University Working Paper.

Wang, J. and E. Zivot (1998). "Inference on Structural Parameters in Instrumental Variables Regressions with Weak Instruments," *Econometrica*, 66, 1389-1404.

Table 1: Approximation Accuracy for Various Values of $\mu'\mu$ *

$\mu'\mu$	Bias		MSE			
	CS Approximation		CS Approximation		DN Approximation	
	$\hat{\beta}(t_{\hat{\beta}})$	R^2	$\hat{\beta}(t_{\hat{\beta}})$	R^2	$\hat{\beta}(t_{\hat{\beta}})$	R^2
2	1.0004(1563.2)	0.9996	0.9993(701.76)	0.9980	0.0004(29.555)	0.4665
4	1.0030(836.01)	0.9986	0.9980(871.11)	0.9987	0.0017(34.233)	0.5398
6	1.0053(693.30)	0.9979	0.9982(1130.2)	0.9992	0.0037(38.209)	0.5937
8	1.0069(657.54)	0.9977	0.9985(1373.2)	0.9995	0.0064(41.777)	0.6360
10	1.0079(656.18)	0.9977	0.9986(1563.0)	0.9996	0.0097(45.084)	0.6705
12	1.0085(668.18)	0.9978	0.9986(1697.9)	0.9997	0.0135(48.213)	0.6994
14	1.0088(685.61)	0.9979	0.9985(1789.8)	0.9997	0.0179(51.212)	0.7242
16	1.0089(705.04)	0.9980	0.9983(1851.5)	0.9997	0.0226(54.110)	0.7456
18	1.0090(724.95)	0.9981	0.9981(1892.8)	0.9997	0.0276(56.929)	0.7644
20	1.0090(744.56)	0.9982	0.9978(1920.2)	0.9997	0.0330(59.684)	0.7810
22	1.0089(763.59)	0.9983	0.9976(1938.0)	0.9997	0.0386(62.385)	0.7957
24	1.0087(781.85)	0.9984	0.9973(1949.0)	0.9997	0.0444(65.043)	0.8090
26	1.0086(799.33)	0.9984	0.9970(1955.0)	0.9997	0.0504(67.663)	0.8209
28	1.0084(816.00)	0.9985	0.9967(1957.1)	0.9997	0.0566(70.250)	0.8317
30	1.0082(831.92)	0.9986	0.9964(1956.4)	0.9997	0.0629(72.810)	0.8414
32	1.0080(847.12)	0.9986	0.9961(1953.3)	0.9997	0.0692(75.346)	0.8504
34	1.0078(861.63)	0.9987	0.9958(1948.3)	0.9997	0.0757(77.861)	0.8585
36	1.0076(875.48)	0.9987	0.9954(1941.8)	0.9997	0.0822(80.357)	0.8660
38	1.0074(888.69)	0.9987	0.9951(1934.2)	0.9997	0.0888(82.838)	0.8729
40	1.0072(901.37)	0.9988	0.9948(1925.4)	0.9997	0.0954(85.304)	0.8793
42	1.0070(913.51)	0.9988	0.9945(1915.8)	0.9997	0.1020(87.758)	0.8852
44	1.0068(925.16)	0.9988	0.9942(1905.5)	0.9997	0.1086(90.200)	0.8906
46	1.0066(936.35)	0.9989	0.9940(1894.6)	0.9997	0.1152(92.633)	0.8957
48	1.0063(947.08)	0.9989	0.9937(1883.2)	0.9997	0.1218(95.057)	0.9004
50	1.0061(957.40)	0.9989	0.9934(1871.3)	0.9997	0.1283(97.473)	0.9049
52	1.0059(967.32)	0.9989	0.9931(1859.0)	0.9997	0.1349(99.882)	0.9090
54	1.0057(976.95)	0.9990	0.9928(1846.4)	0.9997	0.1414(102.28)	0.9128
56	1.0055(986.17)	0.9990	0.9926(1833.5)	0.9997	0.1479(104.68)	0.9165
58	1.0053(995.08)	0.9990	0.9923(1820.3)	0.9997	0.1543(107.07)	0.9198
60	1.0051(1003.7)	0.9990	0.9921(1806.9)	0.9997	0.1606(109.46)	0.9230
62	1.0049(1012.0)	0.9990	0.9918(1793.4)	0.9997	0.1670(111.84)	0.9260
64	1.0047(1020.0)	0.9990	0.9916(1779.7)	0.9997	0.1732(114.22)	0.9289
66	1.0045(1027.9)	0.9991	0.9913(1765.8)	0.9997	0.1795(116.60)	0.9316
68	1.0043(1035.4)	0.9991	0.9911(1751.8)	0.9997	0.1856(118.98)	0.9341
70	1.0041(1042.8)	0.9991	0.9909(1737.8)	0.9997	0.1917(121.35)	0.9365
72	1.0039(1049.9)	0.9991	0.9906(1723.7)	0.9997	0.1977(123.72)	0.9387
74	1.0037(1056.8)	0.9991	0.9904(1709.6)	0.9997	0.2037(126.09)	0.9409
76	1.0036(1063.5)	0.9991	0.9902(1695.4)	0.9997	0.2096(128.46)	0.9429
78	1.0034(1069.9)	0.9991	0.9900(1681.1)	0.9996	0.2154(130.83)	0.9449
80	1.0032(1076.3)	0.9991	0.9898(1666.9)	0.9996	0.2212(133.19)	0.9467
82	1.0030(1082.5)	0.9991	0.9896(1652.7)	0.9996	0.2269(135.56)	0.9484
84	1.0029(1088.5)	0.9992	0.9894(1638.5)	0.9996	0.2326(137.92)	0.9501
86	1.0027(1094.3)	0.9992	0.9892(1624.3)	0.9996	0.2381(140.29)	0.9517
88	1.0025(1100.0)	0.9992	0.9891(1610.1)	0.9996	0.2436(142.65)	0.9532
90	1.0024(1105.7)	0.9992	0.9889(1596.0)	0.9996	0.2491(145.01)	0.9547
92	1.0022(1111.0)	0.9992	0.9887(1581.9)	0.9996	0.2544(147.38)	0.9560
94	1.0020(1116.3)	0.9992	0.9886(1568.0)	0.9996	0.2597(149.74)	0.9574
96	1.0019(1121.5)	0.9992	0.9884(1554.1)	0.9996	0.2650(152.11)	0.9586
98	1.0017(1126.6)	0.9992	0.9882(1540.2)	0.9996	0.2702(154.47)	0.9598
100	1.0016(1131.4)	0.9992	0.9881(1526.4)	0.9996	0.2753(156.84)	0.9610

*Notes: 50000 actual bias and MSE values were generated using the analytical formulae given in Section 2 for various values of $\mu'\mu$, k_{21} and β (i.e. $\sigma_{uu}^{1/2}\sigma_{vv}^{-1/2}\rho$), as discussed above. For each value of $\mu'\mu$, a pseudo regression (with 1000 observations) was then run with the actual bias (MSE) regressed on an intercept and an approximate bias (MSE). Slope coefficients (with t-statistics in brackets) are reported, as well as regression R^2 values.

Table 2: Bias and MSE of OLS, IV, LIML and Bias Adjusted Estimators[*]

ρ_{uv}	R^2_{relev}	$\widehat{\beta}_{OLS}$	$\widehat{\beta}_{IV}$	$\overline{\beta}_{LIML}$	$\overline{\beta}_{OLS,1}$	$\overline{\beta}_{OLS,2}$	$\overline{\beta}_{IV}$	$\overline{\beta}_{IV,1}$	$\overline{\beta}_{IV,2}$
					Bias				
0.10	0.01	0.110056	0.055485	0.018593	0.000708	0.002863	0.001791	0.002979	0.004026
0.10	0.05	0.105647	0.017693	0.003799	-0.006580	-0.000031	-0.001088	0.000194	0.001178
0.10	0.10	0.100122	0.009350	0.002005	-0.011004	0.001004	0.000013	0.000794	0.001702
0.10	0.20	0.088960	0.004525	0.002084	-0.021911	0.001147	0.000269	0.000662	0.001459
0.10	0.25	0.083505	0.003512	0.001892	-0.027351	0.001107	0.000283	0.000587	0.001334
0.20	0.01	0.495437	0.243785	0.039562	-0.207689	-0.195773	-0.202685	-0.197053	-0.190255
0.20	0.05	0.475447	0.075283	0.006690	-0.034721	-0.004923	-0.009726	-0.003904	0.000566
0.20	0.10	0.450478	0.038763	0.002751	-0.053507	0.000951	-0.003544	-0.000006	0.004112
0.20	0.20	0.400122	0.018076	0.002988	-0.101504	0.002789	-0.001184	0.000596	0.004202
0.20	0.25	0.375531	0.013814	0.002428	-0.125700	0.002936	-0.000790	0.000586	0.003961
0.30	0.01	0.644019	0.316396	0.036167	-0.059195	-0.045779	-0.052677	-0.045312	-0.038563
0.30	0.05	0.618020	0.097534	0.006689	-0.045459	-0.006694	-0.012941	-0.005371	0.000442
0.30	0.10	0.585556	0.050132	0.003070	-0.069860	0.000967	-0.004878	-0.000278	0.005077
0.30	0.20	0.520090	0.023319	0.003080	-0.132171	0.003443	-0.001724	0.000590	0.005279
0.30	0.25	0.488123	0.017803	0.002292	-0.163600	0.003659	-0.001186	0.000603	0.004991
0.40	0.01	0.802475	0.393875	0.032691	-0.260180	-0.241607	-0.252047	-0.242810	-0.232556
0.40	0.05	0.770068	0.121286	0.003910	-0.056859	-0.008533	-0.016319	-0.006884	0.000360
0.40	0.10	0.729610	0.062272	0.002785	-0.087282	0.001001	-0.006285	-0.000551	0.006123
0.40	0.20	0.648034	0.028921	0.002399	-0.164863	0.004149	-0.002290	0.000594	0.006438
0.40	0.25	0.608201	0.022065	0.002058	-0.204007	0.004437	-0.001600	0.000629	0.006098
0.50	0.01	0.990607	0.485889	0.026662	-0.813962	-0.786146	-0.803914	-0.792443	-0.774906
0.50	0.05	0.950593	0.149506	0.002772	-0.070357	-0.010682	-0.020295	-0.008646	0.000299
0.50	0.10	0.900647	0.076699	0.002844	-0.107948	0.001057	-0.007939	-0.000861	0.007381
0.50	0.20	0.799944	0.035582	0.002363	-0.203664	0.004997	-0.002952	0.000608	0.007823
0.50	0.25	0.750773	0.027134	0.001982	-0.251967	0.005370	-0.002084	0.000669	0.007420
					Mean Square Error				
0.10	0.01	0.012631	0.029674	2.232538	3.344585	3.276442	3.342322	3.339166	3.273371
0.10	0.05	0.011658	0.008821	0.012163	0.013834	0.012351	0.012591	0.012293	0.012075
0.10	0.10	0.010495	0.004457	0.005174	0.006443	0.005153	0.005246	0.005171	0.005087
0.10	0.20	0.008331	0.002038	0.002173	0.003647	0.002153	0.002187	0.002172	0.002142
0.10	0.25	0.007364	0.001546	0.001622	0.003325	0.001607	0.001630	0.001621	0.001600
0.20	0.01	0.245980	0.087718	25.38972	6.183739	6.056343	6.180073	6.175192	6.051633
0.20	0.05	0.226561	0.015405	0.015113	0.018464	0.015991	0.016374	0.015845	0.015539
0.20	0.10	0.203425	0.006695	0.006366	0.010397	0.006500	0.006635	0.006500	0.006405
0.20	0.20	0.160551	0.002778	0.002678	0.013922	0.002693	0.002734	0.002707	0.002681
0.20	0.25	0.141454	0.002063	0.002000	0.018714	0.002007	0.002033	0.002017	0.002003
0.30	0.01	0.415271	0.129580	4.192306	20.86674	20.45195	20.86363	20.85753	20.44608
0.30	0.05	0.382455	0.020142	0.017161	0.021728	0.018523	0.019011	0.018315	0.017946
0.30	0.10	0.343372	0.008297	0.007230	0.013280	0.007468	0.007635	0.007456	0.007350
0.30	0.20	0.270960	0.003308	0.003046	0.021409	0.003082	0.003129	0.003092	0.003070
0.30	0.25	0.238712	0.002433	0.002275	0.029917	0.002296	0.002324	0.002303	0.002292
0.40	0.01	0.644463	0.186309	1.508342	4.737517	4.635030	4.732172	4.724438	4.627611
0.40	0.05	0.593505	0.026557	0.019876	0.026131	0.021933	0.022562	0.021639	0.021186
0.40	0.10	0.532830	0.010464	0.008403	0.017190	0.008779	0.008989	0.008750	0.008630
0.40	0.20	0.420430	0.004023	0.003545	0.031556	0.003608	0.003663	0.003614	0.003595
0.40	0.25	0.370378	0.002933	0.002648	0.045095	0.002686	0.002717	0.002689	0.002683
0.50	0.01	0.981786	0.269921	3.761505	332.5419	325.9601	332.5278	332.5176	325.9504
0.50	0.05	0.904124	0.036005	0.023845	0.032628	0.026965	0.027801	0.026545	0.025966
0.50	0.10	0.811670	0.013650	0.010137	0.022952	0.010709	0.010984	0.010656	0.010514
0.50	0.20	0.640417	0.005075	0.004281	0.046498	0.004382	0.004450	0.004382	0.004367
0.50	0.25	0.564161	0.003668	0.003199	0.067443	0.003260	0.003296	0.003258	0.003258

[*] Notes: The 1st and 2nd columns report values of the correlation (ρ_{uv}) between the errors in the canonical model (the degree of endogeneity) and the correlation (R^2_{relev}) between the instruments and the endogenous explanatory variable (instrument relevance). Average Bias (except for the LIML estimator, for which median bias is reported) and MSE are reported for the OLS, IV, LIML, and 5 bias corrected estimators in the 3rd through 11th columns, respectively. In the 1st column, the correlations correspond to β = -0.1111, -0.5, -0.65, -0.81, and -1.0, respectively, in the canonical model. In the 2nd column, the correlations correspond to $\overline{\pi}$ = 0.0225, 0.0513, 0.0745, 0.1120, and 0.1290, respectively, in the canonical model. All entries are based on 5000 Monte Carlo trials (see above for further details).